I&EC research
Industrial & Engineering Chemistry Research

Article

pubs.acs.org/IECR

Design of Dynamic Experiments: A Data-Driven Methodology for the Optimization of Time-Varying Processes

Christos Georgakis*

Department of Chemical and Biological Engineering and Systems Research Institute for Chemical and Biological Processes, Tufts University, Medford, Massachusetts 02155, United States

ABSTRACT: A new methodology for the design of experiments is presented that provides a way to optimize the operation of a variety of batch and semibatch or fed-batch processes without the use of a knowledge-driven or fundamental model describing the inner workings of the process, when one or more time-varying decision variables must be selected. The types of processes that can benefit from this approach include those producing specialty chemical, pharmaceutical, or food products whose production rate is not high enough to justify the development of a knowledge-driven model. The approach generalizes the classical and widely used Design of Experiments (DoE), which considers only decision variables, or factors, which are constant with time. The new approach, called the Design of Dynamic Experiments (DoDE), systematically designs experiments that explore a considerable number of dynamic signatures, called dynamic subfactors (DSFs), in the time dependency of the unknown decision variables or factors. Two example processes—a batch nonisothermal reactor and a semibatch penicillin fermentation process—amply demonstrate the utility of the method. In both cases, a small number of experiments lead to the quick and accurate optimization of the process. The calculated optimal operating conditions through the proposed DoDE approach are only slightly different from the optimum that would have been obtained if a knowledge-driven model were available.

1. INTRODUCTION

Over the last 50 years, there has been a very profound increase of the use of knowledge-driven models in the design, analysis, and operations of continuous chemical, petroleum, and petrochemicals processes. [Here, we are using the term "knowledge-driven" model instead of the possibly more widely used term of "knowledge based" or "fundamental" model for two reasons. The first is in juxtaposition to data-driven models. The second reason is because "knowledge-driven" models are also based on data that are needed to estimate the parameter of the model. Without such data, the model-driven model has limited utility.] In contrast, the use of such knowledge-driven models has not been as widely used for batch processes. Batch processes are less well understood and their production rates are, in most cases, at least 1 order of magnitude smaller than continuous processes. A variation of the knowledge-driven approach, using less-accurate models and called Tendency Modeling, has been also proposed.[1−4] Even though it has been successfully applied on several occasions,[5,6] its wider acceptability has not been as strong as was initially expected. This indicated that even approximate knowledge-driven tendency models might be perceived as not providing, at present, sufficient return for the needed developmental effort. In contrast, the use of data-driven models in the batch industry has been widely used especially, but not only, in the pharmaceutical industry. [A search in the "web of knowledge" database with topics "design of experiments" and "pharmaceuticals" yielded more than 5000 entries.]

The Design of Experiments (DoE) methodology is a classical data-driven approach that enables the correlation of the performance characteristics (output variables) of a given process on the values of the operating conditions (input variables or factors). It has found widespread applicability in many industrial applications over the past 50 plus years. The related literature is very extensive with numerous textbooks (for example, see refs 7 and 8 and a multitude of research articles (for example, refs 9−11). The power and utility of the method has been widely recognized. It designs a systematic plan of experimentation in which the decision variables, or factors, are varied over their expected ranges to estimate the overall behavior of the process in a statistically sound way. However, the DoE methodology is limited by the fact that all the factors examined are invariant with time. In a batch process, the time evolution of an operating condition, such as temperature, can have a substantial impact on the end result. In such cases, the classical DoE methodology does not offer an efficient and systematic way to explore the impact that a multitude of time dependencies of one or several operating conditions have in the process performance. The chances are high that time-variant conditions will be more optimal those of the best time-invariant one. In the present article, we remove this important limitation of the DoE methodology, proposing an important generalization that will be referred to as Design of Dynamic Experiments (DoDE).

We start by discussing the characteristics of the DoE methodology as well as its limitations. In section 2, we present the details of the new DoDE methodology. We start by formulating the general DoDE problem when there are only dynamic factors and also when we have both dynamic and static (classical) factors (section 2.1). We present an approximation of the problem with a finite number of experiments (section

Special Issue: John MacGregor Festschrift

Received: December 17, 2012
Revised: April 30, 2013
Accepted: May 3, 2013
Published: May 3, 2013

dx.doi.org/10.1021/ie3035114 | *Ind. Eng. Chem. Res.* 2013, 52, 12369−12382

2.2), and we detail the algorithmic steps of the proposed methodology (section 2.3). In section 3, we present some sets of time-varying experiments as examples of the power and versatility of the new DoDE approach. In section 4, we examine the case of a batch reactor, which we wish to optimize with respect to the operating temperature profile. A more complex process of penicillin fermentation is then examined through simulated experiments using a well-accepted model, in section 5. Here, the time-varying decision variable is the feeding flow rate of the substrate. The last section (section 6) presents the conclusions of this paper.

The Design of Experiment (DoE) approach enables the design of a set of experiments, so that the maximum information is derived from a given experimental effort or, conversely, the minimum number of experiments is performed to obtain the desired information. In most cases, a response surface model is built that summarizes and interpolates the obtained results and enables the optimization of the process. Based on the number of independent variables that one considers, and their related ranges of values, different types of designs are available.

When a time-varying decision variable, or dynamic factor, must be decided, one ad-hoc approach that the DoE approach can offer is to split the $(0, t_b)$ time interval into n subintervals within each of which the decision variable will be changing linearly with time. The decision variables are then the $n - 1$ times instants defining the subintervals and the $n + 1$ values of the temperature at the beginning and the end of each interval. Consequently, the total number of variables is $2n$. As n increases for higher accuracy, the number of DoE experiments needed increases substantially. For two-level experiments one needs 2^{2n} experiments and for three-lever designs, one needs 3^{2n}. In the case of $n = 5$ or 10, there is a need between 1024 (2^{10}) and 1048576 (2^{20}) for level 2 (high−low) experiments. For lever 3 (high−middle−low) experiments, the number of experiments is between 59049 (3^{10}) and 3.4868×10^9 (3^{20}), just for a single time-varying factor. In the case of a full center composite design (CCD), the number of experiments required is 1064 for $n = 5$. This is a very large number of experiments and a more-systematic approach is needed for time-varying decision variables or dynamic factors, as we will call them.

Before we define the mathematical framework of the new Design of Dynamic Experiments (DoDE) methodology, we revisit here the classical DoE approach. Assume that we have n time-invariant inputs u_i $(i = 1, 2, ..., n)$ that range between a minimum and a maximum value, $u_{i,\min}$ and $u_{i,\max}$, and a single process output y. It is customary to substitute each dimensional input variable by a dimensionless equivalent w_i, called a coded variable, which is defined as follows:

$$w_i = \frac{u_i - u_{io}}{\Delta u_i} \qquad i = 1, 2, ..., n \tag{1}$$

where

$$u_{io} = \frac{u_{i,\max} + u_{i,\min}}{2}$$

and

$$\Delta u_i = \frac{u_{i,\max} - u_{i,\min}}{2}$$

This definition implies that $-1 \le w_i \le 1$ $(i = 1, 2, ..., n)$. If enough experiments at different values of the coded variables are performed and the process performance, y, is recorded, then a response surface model, most likely quadratic in nature, can be estimated by linear regression.

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i w_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \beta_{ij} w_i w_j + \sum_{i=1}^{n} \beta_{ii} w_i^2 \tag{2}$$

We now rewrite this equation in a vector form to motivate its generalization.

$$y = \beta_0 + \mathbf{q}^T\mathbf{w} + \mathbf{w}^T\mathbf{Q}\mathbf{w} \quad \text{with}$$
$$q_i \equiv \beta_i,\ Q_{ii} \equiv \beta_{ii} (i = 1, 2, ..., n) \quad \text{and}$$
$$Q_{ij} = Q_{ji} \equiv \beta_{ij}/2\ (i, j = 1, 2, ..., n) \tag{3}$$

Here, the components of the vector $\mathbf{q}$ represents the local sensitivities at $\mathbf{w} = 0$ of the output y, with respect to each of the inputs w_i. The matrix $\mathbf{Q}$ determines whether y has an extremum stationary point (maximum or minimum) or a nonextremum stationary point. Such a stationary point is determined by first differentiating y in eq 3 with respect to the vector $\mathbf{w}$, resulting into the following equation and then solving this equation for $\mathbf{w}$.

$$\mathbf{0} = \mathbf{q} + 2\mathbf{Q}\mathbf{w} \tag{4}$$

If $\mathbf{Q}$ is full rank and has all its eigenvalues non-negative (nonpositive), then a minimum (maximum) exists for y. In the case that some of the eigenvalues of $\mathbf{Q}$ are positive and some are negative, a nonextremum stationary point exists. Whether such an extremum point is inside or outside the constraining region defined by $-1 \le w_i \le +1$ $(i = 1, 2, ..., n)$ is also of great interest. If the extremum stationary point is outside the constrained region, the minimum (maximum) will be at a point in the boundary of the constrained region.

We will now turn out attention to the description of the proposed novel approach of the Design of Dynamic Experiments (DoDE). Because the term previously has been used in the literature (e.g., see refs 12 and 13), there is a need for a clarification. The above-referenced citations aim to design dynamic experiments for the more-accurate estimation of a knowledge-driven model. While this is an important task, it presumes the existence of a model, a knowledge-driven model. This a different problem than the one we address here. We are assuming here that we do not have a knowledge-driven model at hand and we do not possess the complete process knowledge that is required to develop one. Here, we design experiments to develop a data-driven model, a task that does not require having a complete knowledge of the inner working of a process.

2. THE DESIGN OF DYNAMIC EXPERIMENTS METHODOLOGY

Let us consider the case that we have m time-varying input profiles, $\mathbf{u}(t)$, as inputs to the process; we define, in a similar manner, the m-dimensional coded input functional variable $\mathbf{z}(\tau)$ as follows:

$$z_i(\tau) = \frac{u_i(\tau) - u_{i0}(\tau)}{\Delta U_i(\tau)} \qquad -1 \le z_i(\tau) \le 1 \tag{5}$$

where

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

$$u_{i0} = \frac{u_{i,\max}(\tau) + u_{i,\min}(\tau)}{2} \quad \text{and}$$
$$\Delta U_i(\tau) = \frac{u_{i,\max}(\tau) - u_{i,\min}(\tau)}{2}$$
$$\text{for } \tau \in (0, 1) \text{ and } i = 1, 2, ..., n$$

The functions $u_{i,\max}(\tau)$ and $u_{i,\min}(\tau)$ define the ranges of time variations of interest. Different input functions $z_i(\tau)$ for each value of i (i = 1, 2, ..., n) must be selected (designed) and the corresponding experiments performed and the output performance y of the system recorded. Here, we will assume that all functions of time belong to $\mathcal{H}_2(0,1)$, the Hilbert space of square integrable functions in the (0,1) interval.[14] This is the simplest compact functional space in which both the norm of a function and the inner product between two functions is defined and the norm is induced by the inner product.

$$\|f\| \equiv \sqrt{\int_0^1 |f(\tau)|^2 \, d\tau}, \quad (f, g) \equiv \int_0^1 f(\tau)g(\tau) \, d\tau,$$
$$\text{and} \quad \|f\| = \sqrt{(f, f)} \tag{6}$$

The quadratic response surface model, to be estimated from the experimental results, is of the following form:

$$y = \beta_0 + \int_0^1 \mathbf{p}^T(\tau)\mathbf{z}(\tau) \, d\tau + \int_0^1 \int_0^1 \mathbf{z}^T(\tau_1)\mathbf{P}(\tau_1, \tau_2)\mathbf{z}(\tau_2) \, d\tau_1 \, d\tau_2 \tag{7}$$

The sensitivity of y with respect to $\mathbf{z}(\tau)$ is

$$\frac{\partial y}{\partial \mathbf{z}^T(\tau)} = \mathbf{p}(\tau) + 2\int_0^1 \mathbf{P}(\tau, \tau_2)\mathbf{z}(\tau_2) \, d\tau_2 \tag{8}$$

[Here, we are using the convention that the differentiation of a scalar by a row vector produces a column vector and, correspondingly, the differentiation of a scalar by a column vector produces a row vector. This is justified by the fact that a scalar is the product of a row vector times a column vector.] Moreover, the calculation of the unconstrained extremum or stationary point is given by

$$0 = \mathbf{p}(\tau) + 2\int_0^1 \mathbf{P}(\tau, \tau_2)\mathbf{z}^*(\tau_2) \, d\tau_2 \tag{9}$$

Here, $\mathbf{p}(\tau)$ is an m-dimensional vector of sensitivity functions $p_i(\tau)$ in the interval (0, 1) and $\mathbf{P}(\tau_1,\tau_2)$ is an $m \times m$ Hessian matrix with entries following two symmetric relationships $P_{ij}(\tau_1,\tau_2) = P_{ji}(\tau_1,\tau_2)$ and $P_{ij}(\tau_1,\tau_2) = P_{ij}(\tau_2,\tau_1)$, for i, j = 1, 2, ..., m. The solution of eq 9 provides the optimal functional vector $\mathbf{z}^*(\tau)$ in the case that there are no constraints. The presence of constraints on $\mathbf{z}(\tau)$ (see eq 5) makes the optimization problem more challenging.

2.1. Static and Dynamic Factors. We now consider the most general case in which we have n time-invariant inputs $\mathbf{w}$ and m time-variant inputs $\mathbf{z}(\tau)$. Hereafter, we will call the former inputs *static factors* and the later *dynamic factors*. The generalization of eqs 3 and 7 is given by

$$y = \beta_0 + \mathbf{q}^T\mathbf{w} + \mathbf{w}^T\mathbf{Q}\mathbf{w} + \int_0^1 \mathbf{p}^T(\tau)\mathbf{z}(\tau) \, d\tau + \int_0^1 \int_0^1 \mathbf{z}(\tau_1)^T\mathbf{P}(\tau_1, \tau_2)\mathbf{z}(\tau_2) \, d\tau_1 \, d\tau_2 + \mathbf{w}^T \int_0^1 \mathbf{S}(\tau)\mathbf{z}(\tau) \, d\tau \tag{10}$$

Here, $\mathbf{q}$ and $\mathbf{w}$ are $n \times 1$ vectors, $\mathbf{p}(\tau)$ and $\mathbf{w}(\tau)$ are $m \times 1$ vectors of functions, $\mathbf{Q}$ is an $n \times n$ matrix, and $\mathbf{P}(\tau_1, \tau_2)$ and $\mathbf{S}(\tau)$ are matrices of functions with size $m \times m$ and $n \times m$, respectively. Differentiation with respect to $\mathbf{w}$ and/or $\mathbf{z}(\tau)$ yields

$$\frac{\partial y}{\partial \mathbf{w}^T} = \mathbf{q} + 2\mathbf{Q}\mathbf{w} + \int_0^1 \mathbf{S}(\tau_1)\mathbf{z}(\tau_1) \, d\tau_1 \tag{11}$$

and

$$\frac{\partial y}{\partial \mathbf{z}^T(\tau)} = \mathbf{p}(\tau) + 2\int_0^1 \mathbf{P}(\tau, \tau_1)\mathbf{z}(\tau_1) \, d\tau_1 + \mathbf{S}(\tau)^T\mathbf{w} \tag{12}$$

If we now take the second derivatives of y, with respect to all of the input variables, we obtain the generalized Hessian

$$\mathbf{H}(\tau_1, \tau_2) = \begin{pmatrix} 2\mathbf{Q} & \mathbf{S}(\tau_2) \\ \mathbf{S}(\tau_1)^T & 2\mathbf{P}(\tau_1, \tau_2) \end{pmatrix} \tag{13}$$

Here, $\mathbf{H}$ is an $(n + m) \times (n + m)$ symmetric matrix with the $\mathbf{Q}$ elements being constants while the elements of $\mathbf{P}$ are functions of two time variables and the elements of $\mathbf{S}$ are functions of one time variable. Even though we used an informal presentation, we remind ourselves that many of the above derivatives are Fréchet derivatives.[14]

2.2. Finite Number of Experiments. The challenge that we address here is the estimation of the parameters or parametric functions in eq 7 or eq 10. Here, $\mathbf{p}(\tau)$, $\mathbf{P}(\tau_1,\tau_2)$, and $\mathbf{S}(\tau)$ are functions of time; thus, infinite-dimensional vectors and theoretically an infinite number of experiments is needed for their most-accurate estimation. However, an approximate estimation with a reasonable number of experiments will be sufficient. To this end, we introduce an ortho-normal set of functions $\{\phi_1(\tau), \phi_2(\tau), ..., \phi_i(\tau), ...\}$ that is a basis in $\mathcal{H}_2(0,1)$. Consequently, any function in $\mathcal{H}_2(0,1)$ can be expanded as a linear combination of the basis functions. We start by letting $\mathbf{z}(\tau) = \sum_{i=1}^{\infty}\mathbf{x}_i\phi_i(\tau)$, but we limit ourselves to a finite sum of N terms, expecting that it will provide a sufficiently accurate approximation. Here, each of the $\mathbf{x}_j$ is an m-dimensional vector, in the general case that $\mathbf{z}(\tau)$ is an m-dimensional vector. In a similar fashion, we expand into the same basis the $\mathbf{p}(\tau)$, $\mathbf{S}(\tau)$, and $\mathbf{P}(\tau_1,\tau_2)$ functions that must be estimated.

$$\mathbf{z}(\tau) = \sum_{i=1}^{N}\mathbf{x}_i\phi_i(\tau), \quad \mathbf{p}(\tau) = \sum_{i=1}^{N}\boldsymbol{\beta}_i\phi_i(\tau),$$
$$\mathbf{S}(\tau) = \sum_{i=1}^{N}\boldsymbol{\Gamma}_i\phi_i(\tau), \quad \mathbf{P}(\tau_1, \tau_2) = \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Delta}_{ij}\phi_i(\tau_1)\phi_j(\tau_2) \tag{14}$$

Here, each $\boldsymbol{\beta}_i$ is an m-dimensional vector with entries $\beta_{i,k}$ with (k = 1, 2, ..., m), each $\boldsymbol{\Gamma}_i$ is an $m \times n$ matrix with entries $\gamma_{i,kl}$, (k = 1, 2, ..., m, l = 1, 2, ..., n), and each $\boldsymbol{\Delta}_{ij}$ is an $m \times m$ matrix with entries $\delta_{ij,kl}$, (k = 1, 2, ..., m, l = 1, 2, ..., m). Introducing these series expansions into eq 10, and utilizing the orthonormality

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

$$\int_0^1 \phi_i(\tau)\phi_j(\tau)\,\mathrm{d}\tau = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases} \tag{15}$$

of the basis functions $\{\phi_1(\tau), \phi_2(\tau), ..., \phi_i(\tau)\}$, we arrive at the following relationship:

$$y = \beta_0 + (\mathbf{q}_1^{\mathrm{T}} \quad \mathbf{q}_2^{\mathrm{T}})\begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{pmatrix} + (\mathbf{w}_1^{\mathrm{T}} \quad \mathbf{w}_2^{\mathrm{T}})\begin{pmatrix}\mathbf{T}_{11} & \mathbf{T}_{12}\\ \mathbf{T}_{21} & \mathbf{T}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{pmatrix} \tag{16}$$

This is very identical to the structure of eq 3. The relationship of the variables in eq 16 with the previously defined variables is as follows:

$$\mathbf{q}_1 = \mathbf{q} \in \mathbf{R}^n, \quad \mathbf{w}_1 = \mathbf{w} \in \mathbf{R}^n$$

$$\mathbf{q}_2 = (\boldsymbol{\beta}_1^{\mathrm{T}}, \boldsymbol{\beta}_2^{\mathrm{T}}, ..., \boldsymbol{\beta}_N^{\mathrm{T}})^{\mathrm{T}} \in \mathbf{R}^{mN},$$
$$\mathbf{w}_2 = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, ..., \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}} \in \mathbf{R}^{mN}$$

$$\mathbf{T}_{11} = \mathbf{Q} \in \mathbf{R}^{n\times n}, \quad \mathbf{T}_{12} = \mathbf{T}_{21}^{\mathrm{T}} = (\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, ..., \boldsymbol{\Gamma}_N) \in \mathbf{R}^{n\times mN}$$

$$\mathbf{T}_{22} = \begin{pmatrix} \boldsymbol{\Delta}_{11} & \boldsymbol{\Delta}_{12} & \cdots & \boldsymbol{\Delta}_{1N} \\ \boldsymbol{\Delta}_{21} & \boldsymbol{\Delta}_{22} & \cdots & \boldsymbol{\Delta}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Delta}_{N1} & \boldsymbol{\Delta}_{N2} & \cdots & \boldsymbol{\Delta}_{NN} \end{pmatrix} \in \mathbf{R}^{mN\times mN}$$

$$\text{with } \boldsymbol{\Delta}_{ij} = \boldsymbol{\Delta}_{ji}^{\mathrm{T}} \text{ for } i, j = 1, 2, ..., N \tag{17}$$

With the parametrization that was introduced in eq 14, the DoDE problem in eq 17 shares many commonalities with the classical DoE problem in eq 3. In the simplest possible DoDE case, we are examining no static factors $\mathbf{w}$ $(n = 0)$ and we have a single dynamic factor $z_1(\tau)$ $(m = 1)$. The response surface model that we wish to estimate from the experimental results then is given by

$$\begin{aligned} y &= \beta_0 + \mathbf{q}_2^{\mathrm{T}}\mathbf{w}_2 + \mathbf{w}_2^{\mathrm{T}}\mathbf{T}_{22}\mathbf{w}_2 \\ &= \beta_0 + (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_N)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + (x_1 \quad x_2 \quad \cdots \quad x_N) \\ &\quad \begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1N} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{N1} & \Delta_{N2} & \cdots & \Delta_{NN} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \end{aligned} \tag{18}$$

Here, the matrices Δ_{ij} are 1×1, because $m = 1$, and they can be substituted by $\delta_{ij,11}$. The orthonormality of the basis functions $\varphi_i(\tau)$ contributes greatly in simplifying eq 16. Experiments with select values of the coefficients x_i must be performed, the experimental output in y must be recorded, and the coefficients β_0, β_i, and Δ_{ij} in eq 18 must be estimated. The estimation problem in eq 16 or eq 18 is no different than a typical estimation problem in classical DoE. In eq 18, we have N dynamic subfactors instead of the n factors in a DoE design.

The proposed DoDE generalization of the DoE approach then consists of the following four essential steps:

(a) Select a functional basis $\varphi_i(\tau)$ to parametrize the time-varying input function $\mathbf{z}(\tau)$ and the $\mathbf{p}(\tau)$, $\mathbf{S}(\tau)$, and $\mathbf{P}(\tau_1, \tau_2)$ parametric functions of the model.
(b) Design a set of time-varied experiments characterized by a properly selected the constants in the vectors $\boldsymbol{x}_i$ in eq 14.
(c) Estimate the values of β_0, $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{T}_{ij}$ in eq 16 or β_0, $(\beta_1, \beta_2, ..., \beta_N)$ and Δ_{ij} in eq 18, using the values of y that correspond to each of the performed experiments.
(d) Calculate the values of vectors $\mathbf{x}_i^*$ that optimize y.

Concerning item (a) above, the choice of the functional basis depends on the problem at hand. A good choice should be motivated by our intuitive expectation of the form of the optimal profile. The two obvious choices of bases functions are (a) polynomial or (b) trignometric (sines and cosines). There are other choices of orthogonal bases, almost an infinite number. Choosing a less appropriate basis will not derail the solution of a problem. However, it will require a larger number of experiments (larger N) to achieve the same accuracy in the solution of the problem. In the problems examined below, we will use the shifted Legendre polynomials as the basis functions, because we expect that the optimal solution can be accurately approximated by a low-order polynomial.

2.3. The Algorithmic Steps. We provide here some additional details of the four steps described above. We will assume that we have the simplest case of no static factors $(n = 0)$ and a single dynamic factor $(m = 1)$.

(1) Select the reference batch time and define the dimensional time $\tau = t/t_b$.

(2) Define the region within which the experiments are to take place by selecting $u_{max}(\tau)$ and $u_{min}(\tau)$, from which $u_0(\tau)$ and $\Delta u(\tau)$ can be calculated using the following equations:

$$u_0(\tau) = \frac{u_{max}(\tau) + u_{min}(\tau)}{2}$$

$$\Delta u(\tau) = \frac{u_{max}(\tau) - u_{min}(\tau)}{2}$$

Conversely, one can first define $u_0(\tau)$ and $\Delta u(\tau)$ and then calculate $u_{max}(\tau)$ and $u_{min}(\tau)$.

(3) The dimensionless dynamic variable $z(\tau)$, which varies between -1 and $+1$, is defined by

$$z(\tau) \equiv \frac{u(\tau) - u_0(\tau)}{\Delta u(\tau)}$$

and is the coded time-dependent dynamic factor.

(4) Select an appropriate functional basis $\{\phi_i(\tau)|i = 1, 2, ...\}$ defined in the interval $(0, 1)$. These functions must be a linearly independent set that is complete. It is preferred if it is an orthogonal basis. A nonorthogonal basis can also be used, but the calculations are more complex, but certainly doable. Preferably, each member function should be normalized to a norm of unity. The selection of this basis should be influenced by the expected character of the problem's solution in order to reduce the number of needed expansion terms for a given accuracy and thus the number of experiments.

(5) The dynamic factors $z(\tau)$ is expanded as a finite linear combination of the basis functions $\varphi_i(\tau)$, given in eq 14.

(6) A finite number of experiments n_e is designed, motivated by the classical Design of Experiments (DoE) methodology. It selects an appropriate set of values of the subfactor x_i $(i = 1, 2, ..., N)$. Each set of values of the subfactors correspond to a

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

specific time-dependent function $z(\tau)$ and, for this reason, it is called Design of Dynamic Experiments (DoDE). The selection of the x_i (i = 1, 2, ..., N) values need satisfy a set of constraints to ensure that the constraints, $-1 \le z(\tau) \le +1$, on the $z(\tau)$ are satisfied. Depending on the nature of the orthonormal basis $\phi_i(\tau)$, constraints on the values of the coefficients x_i are imposed. This will be demonstrated in the examples to be treated in the following.

(7) Perform the n_e experiments and record the process performance in an n_e-dimensional vector $\mathbf{y}_e$. There might be cases that more than one performance characteristic of the process need to be recorded. In such a case, the experimental results are recorded in a series of r n_e-dimensional vectors $\mathbf{y}_e^1$, $\mathbf{y}_e^2$, ..., $\mathbf{y}_e^r$. The overall performance of the process J might be related to one or more of these r experimental outputs and the corresponding values of J are recorded in the vector $\mathbf{J}_e$.

(8) Develop the appropriate response surface model (RSM) relating $\mathbf{J}_e$ to the values of the x_i (i = 1, 2, ..., N) in the form of eq 18. The unknown parameters of the model β_0, β_i, and $\Delta_{ij} = \delta_{ij,11}$ are estimated by a linear regression algorithm. An analysis of variance (ANOVA) is performed to reveal which of the terms are significant based on the accuracy of the experimental measurements.

(9) Calculate the optimal values of the x_i^* coefficients that optimize J. This is a constrained optimization task since each of the coefficients x_i (i = 1, 2, ..., N) is individually constrained by an upper and lower value, usually $-1 \le x_i \le +1$, and the necessary constraints on x_i which ensure that $-1 \le z(\tau) \le +1$. Through the optimal values of the x_i^* coefficients, one obtains the approximation of the optimum profile $z^*(\tau)$.

The proposed methodology generalizes the classical design of experiments (DoE) methodology, with respect to dynamically varying decision variables. For this reason, the term Dynamic Design of Experiments (DoDE) has been coined to describe it.[15−17]

3. EXAMPLE DESIGNS OF DYNAMIC EXPERIMENTS

Here, we present some example designs of the DoDE experiments. We define the basis function $\varphi_i(\tau)$, in terms of the shifted Legendre polynomial as the basis in the Hilbert space $\mathcal{H}_2(0,1)$. This is the simplest set of basis functions that enables simple polynomial dependence of the experimental conditions with time.

$$\phi(\tau) = \eta_{i-1} P_{i-1}(\tau) \qquad i = 1, 2, ..., N, ... \tag{19}$$

Here, the values of the η_i constants are used to normalize the orthogonal set of shifted Legendre polynomial. The first three Legendre polynomials are

$$P_0(\tau) = 1, \quad P_1(\tau) = -1 + 2\tau, \quad P_2(\tau) = 1 - 6\tau + 6\tau^2 \tag{20}$$

and the corresponding values of η_i are

$$\eta_0 = 1, \quad \eta_1 = \sqrt{3}, \quad \eta_2 = \sqrt{5} \tag{21}$$

or, more generally, $\eta_N = (2N + 1)^{1/2}$. A transformation of the expansion of eq 14, in terms of the shifted Legendre polynomials, can now be written as

$$z(\tau) = \sum_{i=1}^{N} x_i \phi_i(\tau) = \sum_{i=1}^{N} \tilde{x}_i P_{i-1}(\tau) \quad \text{with} \quad \tilde{x}_i = \eta_{i-1} x_i, \quad i = 1, 2, ..., N \tag{22}$$

We will use these orthogonal polynomials to define the dynamic experiments in many of the examples to be discussed here. Because the values that each shifted Legendre polynomial can attain in the (0, 1) interval are between −1 and 1, the following inequalities on the coefficients $\tilde{x}_i$

$$-1 \le \tilde{x}_1 \pm \tilde{x}_2 \pm ... \pm \tilde{x}_N \le +1 \tag{23}$$

guarantees that the constraints on $z(\tau)$ in eq 5 will be satisfied. For the case of N = 3, eq 23 is a compact representation of the following eight inequalities:

$$\begin{aligned} -1 &\le \tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 \le +1 \\ -1 &\le \tilde{x}_1 - \tilde{x}_2 + \tilde{x}_3 \le +1 \\ -1 &\le \tilde{x}_1 + \tilde{x}_2 - \tilde{x}_3 \le +1 \\ -1 &\le \tilde{x}_1 - \tilde{x}_2 - \tilde{x}_3 \le +1 \end{aligned} \tag{24}$$

For convenience, from now on, we will assume that we are working with the shifted Legendre polynomials and we will simple write x_i instead of $\tilde{x}_i$.

3.1. A Simple Example of the DoDE Design Approach.

The simplest set of DoDE experiments is obtained by selecting the smallest value of N in eq 3, equal to 2. This implies that the dynamic profile of $u(\tau)$, or the coded variable $z(\tau)$, is a linear combination of the first two Legendre polynomials $P_0(\tau)$ and $P_1(\tau)$. This limits our consideration among constant or linear time dependencies. In deciding the values of the x_1 and x_2 subfactors, we can follow the classical DoE approach. If we perform two-level experiments for each subfactor, we will design the following four (2^2) experiments:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} +1 \\ +1 \end{pmatrix}, \begin{pmatrix} +1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ +1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \tag{25}$$

Translating, for example, the second experiment to the corresponding dynamic coded variable $z(\tau)$, we have $z(\tau) = P_0(\tau) - P_1(\tau) = 2\tau$. We realize that this profile does not meet its constraints $-1 \le z(\tau) \le +1$. This is easily remedied by imposing the following constraints on the x_1 and x_2 constants: $-1 \le x_1 \pm x_2 \le 1$ reducing the values of x_1 and x_2 without changing their relative magnitude, so the that the constraints on $w(\tau)$ are met. This leads to the following design of experiments:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} +0.5 \\ +0.5 \end{pmatrix}, \begin{pmatrix} +0.5 \\ -0.5 \end{pmatrix}, \begin{pmatrix} -0.5 \\ +0.5 \end{pmatrix}, \begin{pmatrix} -0.5 \\ -0.5 \end{pmatrix} \tag{26}$$

In this case, the DoDE design will be called Design D1. It consists of two dynamic subfactors (N = 2) and it is a two-level design in the classical DoE terminology. The four time variations in $z(\tau)$ are shown in Figure 1.

3.2. Other DoDE Examples. For N = 3, we are considering the first three Legendre polynomials, and if we consider only low and high values for the coefficients x_1, x_2, and x_3, we need to perform 8 (2^3) experiments. Here, we are considering the quadratic dependence on time, along with the constant and linear subfactors considered previously. The 8 different time variations of the $z(\tau)$ dynamic factor in these experiments are presented in Figure 2 and are characterized by the ($\pm^1/_3$, $\pm^1/_3$, $\pm^1/_3$) set of values for x_1, x_2, and x_3. The constraints that these three coefficients must satisfy are $-1 \le x_1 \pm x_2 \pm x_3 \le 1$. These are sufficient conditions for the time profiles to always stay within the boundary of $-1 \le z(\tau) \le 1$.

If we add cubic dependence by letting N = 4, we need 16 (2^4) experiments for a full factorial design. Here, we perform a

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

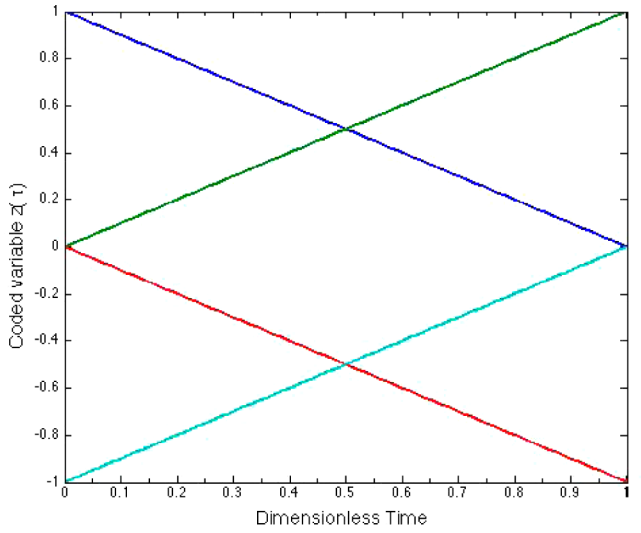

Figure 1. Two-level full factorial DoDE experiments with two dynamic subfactors (Design D1).

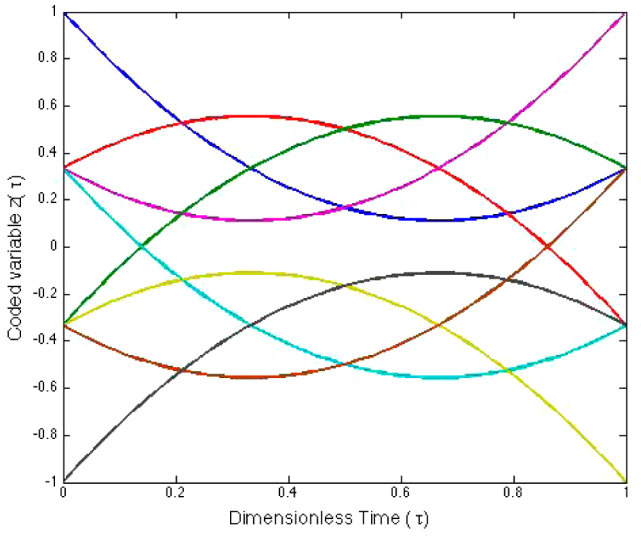

Figure 2. Two-level full factorial DoDE experiments with three dynamic subfactors (Design D2, 2^3 = 8 runs).

IV-optimal design for the estimation of a response surface model that includes the linear and the two-term interactions. The minimum number of experiments required is 11, and they are plotted in Figure 3. This design is called D3 and it uses the first four shifted Legendre polynomials. The fourth one, which previously was not used, is given by

$$P_3(\tau) = -1 + 12\tau - 30\tau^2 + 20\tau^3$$

Again, here, we are using the necessary constraints for the coefficients x_1, x_2, x_3, and x_4, as in the previous example.

For N = 5, we involve the first five shifted Legendre polynomials. This includes the four polynomials mentioned above, along with the fifth one:

$$P_4(\tau) = +1 - 20\tau + 90\tau^2 - 140\tau^3 + 70\tau^4$$

Here, the full factorial design requires 2^5= 32 experiments. An IV-optimal design for a quadratic, or a two-factor interaction response surface, model needs a minimum of 15 or 11 runs, respectively.

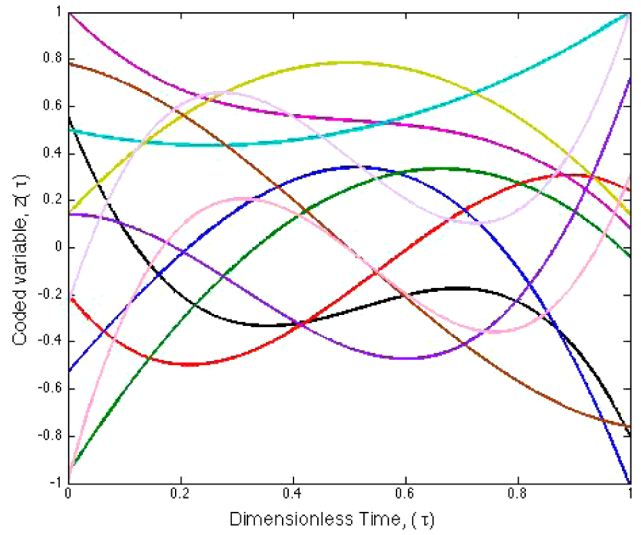

Figure 3. IV-optimal design for a two-factor interaction model with four dynamic subfactors (Design D3, 11 runs).

In this case, we design three-level full factorial experiments, still using the shifted Legendre polynomials as the basis functions; the following DoDE design is obtained for N = 2. We need 9 (3^2) experiments and the corresponding time variations of the $z(\tau)$ dynamic factor is presented in Figure 4. Here again, the two design variables x_1 and x_2 are constrained by the following two inequalities: $-1 \le x_1 \pm x_2 \le 1$

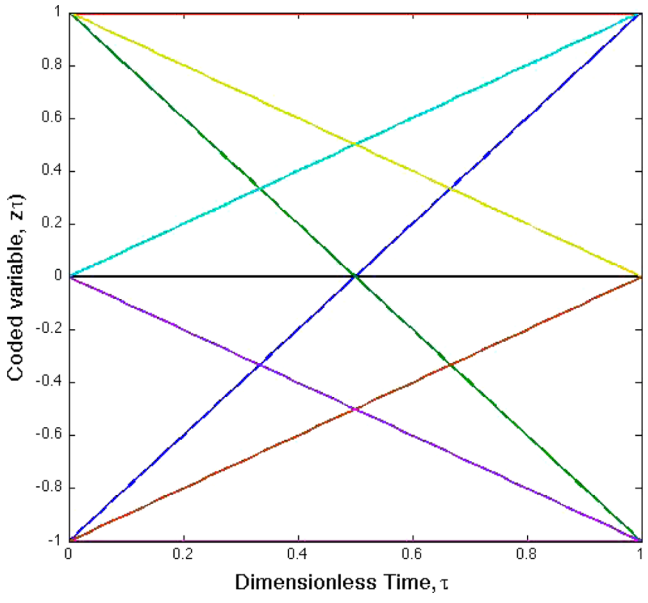

Figure 4. Three-level full factorial DoDE design with two dynamic subfactors (3^2 = 9 runs).

Because the three-level full factorial designs involve a rather large number of experiments, they are not used frequently. On the other hand, the 2-level experimental designs are not sufficient for the development of a quadratic Response Surface Model (RSM). Instead, the Center Composite Design (CCD) or an optimal design is used. In the case of three dynamic subfactors, a D-optimal (or a VI-optimal design), using the point exchange algorithm in the Design Expert software, defines the following 13 experiments instead of the 27 experiments in

the three-level full factorial. This set incudes the minimum of 10 experiments needed to estimate the 10 parameters of the quadratic model and 3 additional experiments to estimate the *lack-of-fit* (LoF) statistic. They are represented in Table 1 by the corresponding values of x_1, x_2, and x_3, which satisfy the following eight inequalities: $-1 \le x_1 \pm x_2 \pm x_3 \le 1$.

Table 1. The 13 DoDE D-Optimal Experiments Involving Three Dynamic Subfactors

run	x_1	x_2	x_3
1	0.00	0.00	−1.00
2	−0.49	−0.01	−0.50
3	0.51	0.00	−0.49
4	0.00	−0.54	−0.46
5	1.00	0.00	0.00
6	−0.47	−0.53	0.00
7	0.00	1.00	0.00
8	−0.98	0.02	0.00
9	0.00	−1.00	0.00
10	0.43	0.43	0.13
11	−0.36	0.38	0.24
12	0.40	−0.27	0.33
13	0.00	0.00	1.00

In most of the designs previously described, we talked of the absolute minimum of runs needed to estimate the number of parameters of the corresponding RSM. In such cases, the number of runs is equal to the number of parameters that must be estimated. These runs must be appended by 3−5 replicates to estimate the inherent variability of the process. Furthermore, one must add an additional set of 3−5 runs that are distinct from the previous runs in order to estimate the *lack-of-fit* (LoF) statistic, or, equivalently, the goodness-of fit (GoF) statistic (see page 115 in ref 8).

4. BATCH REACTOR WITH REVERSIBLE REACTION

Here, we consider the optimization of a batch reactor in which a reversible reaction between reactant A and product B takes place with the following characteristics:

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{B}, \qquad r = k_1 C_\mathrm{A} - k_2 C_\mathrm{B} \tag{27}$$

where

$$k_1 = k_{10} \exp\left(-\frac{E_1}{RT}\right)$$

$$k_2 = k_{20} \exp\left(-\frac{E_2}{RT}\right)$$

$$k_{10} = 1.32 \times 10^7\ \mathrm{h}^{-1}$$

$$k_{20} = 5.25 \times 10^{13}\ \mathrm{h}^{-1}$$

$$E_1 = 10\,000\ \mathrm{kcal}$$

$$E_2 = 20\,000\ \mathrm{kcal}$$

We select the activation energy of the reverse reaction to be larger than that for the forward reaction. This implies that the optimum temperature profile is a decreasing one.[18] This model will be used to perform *in silico* experiments, in order to demonstrate the DoDE methodology. Different DoDE designs will be examined, and the achieved DoDE-optimal operation of the batch reactor will be reported. Depending on the number and type of experiments performed the related DoDE-Optimal achieved will be different, because they are all suboptimal to the true optimal operation. The true optimal operation is the one that is calculated if a model of the process is at hand and the model is perfectly accurate, which is an idealized situation. The different DoDE-optimal operations will be compared to the true optimal to assess how close the DoDE optimal solutions are to the true optimal operation.

To properly simulate the experimental conditions a measurement error will be added to the results of the simulated experiments. The error will be proportional to the simulated value, as given in the following equation:

$$y_\mathrm{e} = y(1 + \sigma N(0, 1)) \tag{28}$$

Here, $N(0,1)$ is a normally distributed number with zero mean and standard deviation equal to 1. Unless otherwise specified, we will select the value of σ to be equal to 0.005 (defined as $\sigma = 0.01/1.96$). This implies that we can claim, with 95% confidence, that the measurements error is within the ±1.0% range. In the cases that the measurement error is larger, a larger value for σ should be selected. This does not alter the results that will be present below with the currently selected value of σ.

We will consider the operation of this batch reactor to be constrained between the lower and upper temperatures of 15 °C and 50 °C, respectively. We will also fix the batch time at 2.0 h. Different batch times can be considered, and this could be another factor in the design of dynamic experiments. This is examined in the second example. However, our interest here is confined in demonstrating with the simplest possible example how the optimal temperature profile is calculated using the DoDE methodology. We start with the calculation of the base case performance of the process when it is operated in the middle of the allowed temperature range (at 32.4 °C). The conversion at the end of 2 h is 70.46%. As a first attempt of optimizing the operation of the reactor, we are considering what the optimal time-invariant temperature might be. For this reason, we run five experiments: one at the maximum allowed temperature (50 °C), one at the lowest allowed temperature (15 °C) and three repeated experiments at the middle temperature (32.5 °C). The repeated experiments at the middle temperature aim to avail the Analysis of Variance (ANOVA) methodology, with an estimate of the inherent variability of the process. The conversions are given in Table 2.

Table 2. DoE Runs with Temperature Constant with Time

parameter	run 1	run 2	run 3	run 4	run 5
temperature, T	15 °C	32.5 °C	32.5 °C	32.5 °C	50 °C
coded variable, w	−1	0	0	0	1
simulated conversion, y	45.95%	70.46%	70.46%	70.46%	60.82%
measured conversion, y_e	46.20%	71.04%	70.54%	70.67%	60.56%

The resulting response surface model is quadratic and given as

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 {x_1}^2 \tag{29}$$

The values of the three coefficients, along with their 95% confidence intervals, are given by

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

$$\beta_0 = 70.75\% \pm 0.64\%$$
$$\beta_1 = 7.18\% \pm 0.79\%$$
$$\beta_2 = -17.37\% \pm 1.02\% \quad (30)$$

The maximum conversion is calculated to take place at $x_1 = 0.22$, corresponding to a constant temperature of 36.25 °C, and at which the estimated conversion is 71.49% ± 0.64%. Running the simulated experiments at this temperature yields a conversion of 71.16%.

4.1. Temperature Profiles Linear with Time. Here, we design the simplest of the DoDE designs, consisting of just two dynamic subfactors x_1 and x_2, representing constant and linear varying temperature profiles, respectively. Following the general relationships of eq 5 with $u(\tau) = T(\tau)$, we write

$$\begin{aligned} z(\tau) &= x_1 P_0(\tau) + x_2 P_1(\tau) \\ &= x_1 + x_2(-1 + 2\tau) \\ T(\tau) &= 32.5 + 17.5\{x_1 - x_2 + 2x_2\tau\} \end{aligned} \quad (31)$$

We now perform a set of experiments that are specified by the Center Composite Design (CCD) with $\alpha = 2$, for the purpose of estimating a quadratic RSM. The values of the above two degrees of freedom, x_1 and x_2, defining the time profiles, are given in the second and third columns of Table 3. The

Table 3. CCD DoDE Design with 12 Runs Defined with Two Dynamic Subfactors Characterizing Temperature Profiles That Are Linear in Time

	Coded Variable Coefficient				
run	x_1	x_2	temperature, T	simulated conversion, y	measured conversion, y_e
1	0	−1	50 → 15 °C	73.56%	73.88%
2	0	−1	50 → 15 °C	73.56%	73.16%
3	−0.5	−0.5	32.5 → 15 °C	62.84%	62.85%
4	0.5	−0.5	50 → 32.5 °C	72.71%	72.61%
5	−1	0	@ 15 °C	45.95%	45.68%
6	0	0	@ 32.5 °C	70.46%	70.53%
7	1	0	@ 50 °C	60.82%	60.09%
8	1	0	@ 50 °C	60.82%	60.50%
9	−0.5	0.5	15 → 32.5 °C	60.99%	61.09%
10	0.5	0.5	32.5 → 50 °C	63.00%	63.38%
11	0	1	15 → 50 °C	62.67%	62.64%
12	0	1	15 → 50 °C	62.67%	62.93%

characteristics of the temperature profiles are given in the fourth column of the same table. This set of 12 experiments consists of the 9 experiments depicted in Figure 4, with 3 duplicates that enable the calculation of the inherent experimental error. Because there are only 6 model parameters, there are 3 additional experiments than the absolute minimum, which are used to estimate the LoF statistic. The results of the simulated experiments are given in the fifth column of Table 3. In the sixth column of that table, the simulated experimental data are given, which are derived from the fifth column of that table by adding the previously defined experimental error.

The resulting RSM model, a quadratic one, is of the form

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} {x_1}^2 + \beta_{22} {x_2}^2 \quad (32)$$

with the following values for the coefficients, along with their 95% confidence intervals:

$$\beta_0 = 69.98\% \pm 1.19\%$$
$$\beta_1 = 6.98\% \pm 0.78\%$$
$$\beta_2 = -5.39\% \pm 0.67\%$$
$$\beta_{12} = -7.41\% \pm 6.17\%$$
$$\beta_{11} = -16.97\% \pm 1.54\%$$
$$\beta_{22} = -1.89\% \pm 1.42\% \quad (33)$$

The R^2-adj for the above regression is 0.99 and the LoF statistic is not significant ($p = 0.11$). The optimum reactor conversion is estimated to occur at a temperature profile characterized by the following values of the dynamic subfactors: $x_1^* = 0.28$, $x_2^* = 0.62$. The predicted conversion is 74.53% ± 1.76%. The simulated run with the optimal temperature profile of

$$\begin{aligned} T^*(t) &= 32.5 + 17.5[0.28P_0(\tau) - 0.62P_1(\tau)] \\ &= 50 - 22.05\tau \end{aligned} \quad (34)$$

gives a final conversion of 74.32%, which is within the confidence limit of the estimated value.

4.2. Temperature Profiles Quadratic in Time. An objection that might be raised about the DoDE design that we presented above is that it considers only linear temperature profiles and the optimal one might be a more complex one. For this reason, we are now considering a design of experiments that will involve three, instead of two, dynamic subfactors. Equation 31 is now written as

$$\begin{aligned} z(\tau) &= x_1 P_0(\tau) + x_2 P_1(\tau) + x_3 P_2(\tau) \\ &= x_1 + x_2(-1 + 2\tau) + x_3(+1 - 6\tau + 6\tau^2) \\ T(\tau) &= 32.5 + 17.5\{x_1 - x_2 + x_2 + 2(x_2 - 3x_3)\tau \\ &\quad + 6x_3\tau^2\} \end{aligned} \quad (35)$$

Here, there are three degrees of freedom defining each experiment: x_1, x_2, and x_3. A CCD design to estimate a quadratic model requires 14 experiments plus a minimum of three center points at the reference temperature profile to estimate the inherent variability of the process. Instead of the CCD design, we select a D-Optimal design that requires 10 experiments to estimate the 10 model parameters, three additional ones for the LoF statistic, and 3 replicates for estimating the inherent process variability. Table 4 details this set of experiments.

The resulting quadratic RSM model is

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} {x_1}^2 \quad (36)$$

with the following values for the coefficients, along with their 95% confidence intervals:

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

Table 4. DoDE Runs with Three (3) Dynamic Subfactors Defining Temperature Profiles That Are Quadratic in Time

	Subfactor				
run	x_1	x_2	x_3	simulated conversion, y	measured conversion, y_e
1	0.00	0.00	−1.00	70.36%	70.67%
2	0.00	0.00	−1.00	70.36%	69.97%
3	0.00	−0.50	−0.50	71.84%	71.85%
4	−0.67	0.00	−0.33	56.71%	56.63%
5	0.00	−1.00	0.00	73.56%	73.13%
6	0.00	−1.00	0.00	73.56%	73.64%
7	0.50	−0.50	0.00	72.71%	71.84%
8	0.00	0.00	0.00	70.46%	70.09%
9	1.00	0.00	0.00	60.82%	60.92%
10	0.50	0.50	0.00	63.00%	63.38%
11	0.00	1.00	0.00	62.67%	62.64%
12	0.00	1.00	0.00	62.67%	62.93%
13	−0.67	0.00	0.33	56.64%	56.61%
14	0.50	0.00	0.50	66.23%	66.52%
15	0.00	0.50	0.50	64.52%	64.17%
16	0.00	0.00	1.00	67.02%	67.03%

$$\beta_0 = 68.29\% \pm 0.53\%$$
$$\beta_1 = 7.09\% \pm 1.19\%$$
$$\beta_2 = -5.36\% \pm 0.85\%$$
$$\beta_3 = -1.83\% \pm 0.91\%$$
$$\beta_{12} = -6.20\% \pm 5.35\%$$
$$\beta_{11} = -15.02\% \pm 1.78\% \quad (37)$$

The R^2-adj for the above regression is 0.98 and the LoF is 0.08. The optimum reactor conversion is estimated to take place at a temperature profile characterized by the following dynamic subfactor values: $x_1^* = 0.26$, $x_2^* = -0.67$, $x_3^* = 0.06$. The predicted conversion is 73.67% ± 2.09%. The simulated run with the optimal temperature profile

$$T(\tau) = 32.5 + 17.5\{0.26P_0(\tau) - 0.67P_1(\tau) + 0.06P_2(\tau)\} = 40 + 26.25\tau + 2.1\tau^2 \quad (38)$$

starting at 50 °C and ending at 15.85 °C at the end of the batch. Simulation of this optimal profile yields a final conversion of 74.31%, which is within the confidence limit of the estimated value.

By comparing the above two cases of DoDE design, we conclude that the quadratic temperature profiles do not provide any significant further optimization of the process, since the calculated optimal profiles yield a final conversion in each case that is indistinguishable from each other (74.32% vs 74.31%), especially if we take into account that the measurement error is 1%. If we assume that we know the exact model of the process and we perform a model-based optimization using the exact kinetic model in the simulated experiments, then the calculated optimal conversion is 74.6%, which is only slightly (0.38%) better than that calculated through the DoDE data-driven methodology. These calculations of the model-based optimum, were performed using the method of Randau collocations in finite elements in the GAMS environment.

In this section, we have demonstrated with a single example process that the DoDE approach has the potential of quickly optimizing a batch process, arriving at an operation that is not that far from—in this case, indeed very close to—the true optimal operation that can achieved if the exact model of the process was known. In the following section, we will consider a more-complex reaction example to offer further evidence of the DoDE promise. The example to be examined next relates to a penicillin fermentation process.

5. PENICILLIN FERMENTATION

Here, we simulate the penicillin fermentation model of Bajpai and Reuss,[19] which has been the center of attention in several model-based optimizations.[20] The model used to simulate the process consists of a volume balance (V), biomass (x) and substrate (s) balances, and a product (p) balance, as follows:

$$\frac{dV}{dt} = u$$
$$\frac{dx}{dt} = \mu x - \frac{xu}{V}$$
$$\frac{dV}{dt} = -\frac{\mu x}{Y_{X/S}} - \frac{\rho x}{Y_{P/S}} - \frac{m_S s x}{k_m + s} + \frac{(s_f - s)u}{V}$$
$$\frac{dp}{dt} = \rho x - k_d p - \frac{pu}{V} \quad (39)$$

with

$$\mu = \mu_{max}\left(\frac{s}{k_x x + s}\right) \quad \text{and} \quad \rho = \rho_{max}\left[\frac{s}{k_p + s + (s^2/k_{in})}\right] \quad (40)$$

The model parameters used here are those used by Riascos and Pinto[20] and are summarized in Table 5.

Table 5. Model Parameters for the Penicillin Fermentation Model

parameter	definition	value
μ_{max}	maximum specific biomass growth rate	0.1 h^{-1}
ρ_{max}	maximum specific production rate	5.5 × 10^{-3} g_P/g_X h
k_x	saturation parameter for biomass growth	6.0 × 10^{-3} g_S/g_X
k_p	saturation parameter for penicillin production	0.1 × 10^{-3} g_S/L
k_{in}	inhibition parameter for penicillin production	0.1 g_S/L
k_d	penicillin degradation	0.01 h^{-1}
k_m	saturation parameter for maintenance consumption	0.1 × 10^{-3} g_S/L
m_S	maintenance consumption rate	2.9 × 10^{-2} $g_S/(g_X$ h)
$Y_{X/S}$	yield factor, substrate (S) to biomass (X)	0.47 g_X/g_S
$Y_{P/S}$	yield factor, substrate (S) to product (P)	1.2 g_P/g_S
s_f	feed concentration of substrate (S)	500 g S/L

Here, we are considering a more-complex DoDE case of calculating the optimum time-varying profile, $u(\tau)$, but also considering the duration of the batch as an unknown. The batch time will be varied around the reference value of 130 h. We do this using the following definition of the first coded

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

variable, w_1. This is a traditional factor, similar to those considered in the classical DoE methodology.

$$t_b = 130 + 30w_1, \quad \text{with} \quad -1 \le w_1 \le +1 \tag{41}$$

In the set of experiments that we will design here, we include one additional traditional factor, as well as a dynamic factor. We set the initial volume of the biomass mixture in the bioreactor equal to 7.0 L and we will vary the initial biomass concentration between 1 and 2 g_X/L. Therefore, we set

$$x(0) = 1.5 + 0.5w_2, \quad \text{with} \quad -1 \le w_2 \le +1 \tag{42}$$

The initial concentration of the substrate in the bioreactor will be fixed and equal to 500 g_S/L. In a similar fashion, we fix the substrate concentration that is fed in semibatch mode to the same value. If one were to consider a more general case, a third factor (w_3) could have been defined to vary this substrate concentration over a range. We will not do this here, because we are focusing our attention onto the time-varying factors. The definition of the dynamic factor $z(\tau)$ is slightly more involved here, compared to the previous example. We are setting the maximum volume capacity of the bioreactor at 10 L, which is a value that was assumed in a previous publication[20] where the model-based optimum was calculated. The objective is for the total volume of the substrate fed during the duration of the batch fermentation reach this maximum volume. We know that, for such fermentations, the initial amount of the substrate fed should be high, so that it induces a rapid growth of the biomass. Furthermore, we expect that the incoming flow rate of the substrate should become equal to zero at the end of the fermentation. For the initial value of the reactor volume of 7.0 L, the reference value of the substrate inflow ($u_0(t)$) should satisfy the following total volume constraint:

$$V(0) + \int_0^{t_f} u_0(t)\, dt = V(t_f) \tag{43}$$

This simplifies to the following constraint:

$$\int_0^1 u_0(\tau)\, d\tau = \frac{3}{130} \triangleq \gamma(\text{L/h}) \tag{44}$$

We now assume a simple linear dependence of u_0 on the dimensionless time τ in the form of $u_0(\tau) = A + B\tau$ and impose the following final condition $u_0(1) = 0$. This condition is motivated by our knowledge that the feeding profile should be going to zero at the end of the batch. Along with the constraint of eq 44, we calculate that u_0 is given as

$$u_0(\tau) = A(1 - \tau), \quad A = 2\gamma = \frac{6}{130}\ \text{L/h} \tag{45}$$

We are also selecting $\Delta u(\tau) = A(1 - \tau)$, so that all possible feeding profiles between zero and $2u_0(\tau)$ are possible. This results in the following definition of the dynamic factor $z(\tau)$:

$$u(\tau)\ (\text{L/h}) = u_0(\tau) + \Delta u(\tau) z(\tau) \quad \text{with}$$
$$-1 \le z(\tau) \le +1 \quad \text{for} \quad \tau \in (0, 1) \tag{46}$$

This is not the only choice in defining the dynamic factor $z(\tau)$. There are a multitude of other choices for $u_0(\tau)$ and $\Delta u(\tau)$. The present choice is motivated by our desire to have the simplest possible definition that will incorporate our expectation that the feeding profile will be going to zero at the end of the batch. We now parametrize $z(\tau)$, in terms of the Legendre polynomials:

$$z(\tau) = x_1 P_0(\tau) + x_2 P_1(\tau) + x_3 P_2(\tau) \tag{47}$$

By imposing two constraints, the five degrees of freedom, consisting by the three dynamic subfactors (x_1, x_2, and x_3) and the two classical factors (w_1 and w_2), will be reduced to three. First, we will require that $z(1) = 0$. Since $\Delta u(1) = 0$, there is no incentive for $z(1)$ to have any other value. This implies that

$$x_3 = -(x_1 + x_2) \tag{48}$$

The objective of the second constraint is to fill the reactor but not overfill it. Consequently, an equation similar to eq 43 should be satisfied by $u(\tau)$:

$$V(0) + \int_0^{t_b} [u_0(t) + \Delta u(t) z(t)]\, dt = V(t_b) \tag{49}$$

This simplifies into the following relationship:

$$10\gamma w_1 + \frac{(1 + 10\gamma w_1)}{\gamma} \int_0^1 \Delta u(\tau) z(\tau)\, d\tau = 0 \tag{50}$$

The imposed two constraints of eqs 48 and 50 lead to the elimination of x_3 and w_1 as independent variables. They can be expressed in terms of x_1 and x_2 through eq 48 and the following expression:

$$w_1 = -4.33\left(\frac{3x_1 - x_2}{3x_1 - x_2 + 3}\right) \tag{51}$$

The resulting independence variables of the design are three: $A = w_2$, $B = x_1$, and $C = x_2$ while w_1 and x_3 are dependent factors through eqs 48 and 51. A D-optimal design of dynamic experiments is defined for the three factors A, B, and C, which stratify the following constraints:

$$-1 \le w_1 \le 1, \quad -1 \le A(= w_2) \le 1, \quad \text{and}$$
$$-1 \le z(\tau) \le 1 \quad \text{for} \quad \tau \in (0, 1) \tag{52}$$

The set of dynamic experiments is given in Table 6.

In this table, we list 16 runs. The minimum number needed are 10, since there are 10 terms in a quadratic RSM involving the independent factors A, B, and C. Three more runs are

Table 6. D-Optimal Design of the DoDE Runs of the Penicillin Process

run	w_1	w_2[a]	x_1[b]	x_2[c]	x_3	y_s (g_P)[d]	y_e (g_P)[d]
1	0.99	0.33	−0.10	0.26	−0.16	73.46	72.94
2	−0.62	0.33	0.00	−0.50	0.50	34.16	33.30
3	1.04	−1.00	−0.27	−0.23	0.50	38.94	38.72
4	−0.99	−1.00	0.13	−0.50	0.37	28.40	28.42
5	−0.30	1.00	0.18	0.32	−0.50	78.20	79.85
6	−0.14	−1.00	0.03	−0.01	−0.02	53.71	53.91
7	−1.00	0.00	0.35	0.15	−0.50	56.51	57.46
8	0.99	−1.00	−0.02	0.50	−0.48	85.30	85.07
9	−0.30	1.00	0.18	0.32	−0.50	78.20	77.33
10	1.04	1.00	−0.27	−0.23	0.50	42.75	42.83
11	1.04	0.00	−0.27	−0.23	0.50	41.12	41.25
12	−0.99	−1.00	0.13	−0.50	0.37	28.40	28.56
13	0.03	−0.33	0.12	0.38	−0.50	78.62	79.98
14	1.04	1.00	−0.27	−0.23	0.50	42.75	41.42
15	−0.63	0.00	0.14	−0.09	−0.05	51.63	51.41
16	−1.00	1.00	0.24	−0.18	−0.06	48.26	49.15

[a] $w_2 = A$. [b] $x_1 = B$. [c] $x_2 = C$. [d] Units denote the mass of penicillin produced (in grams).

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

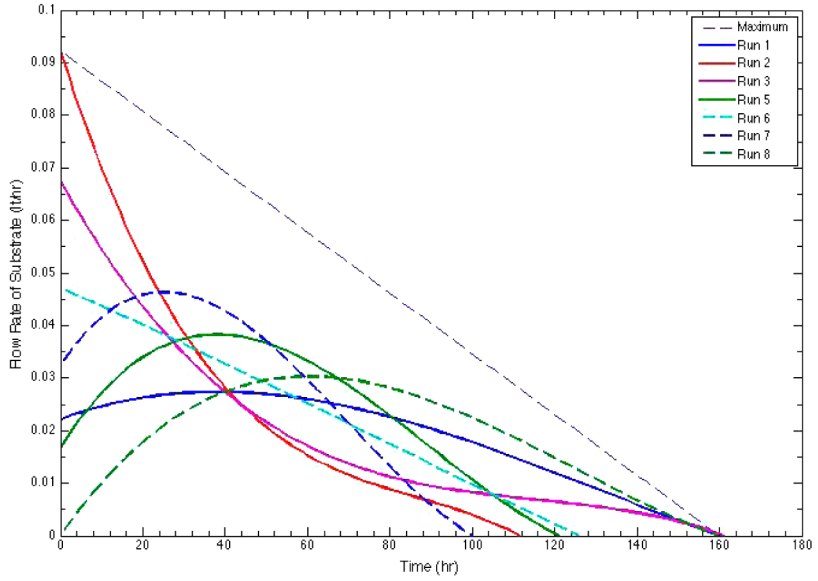

Figure 5. Some of the feeding profiles of the DoDE design for the penicillin process.

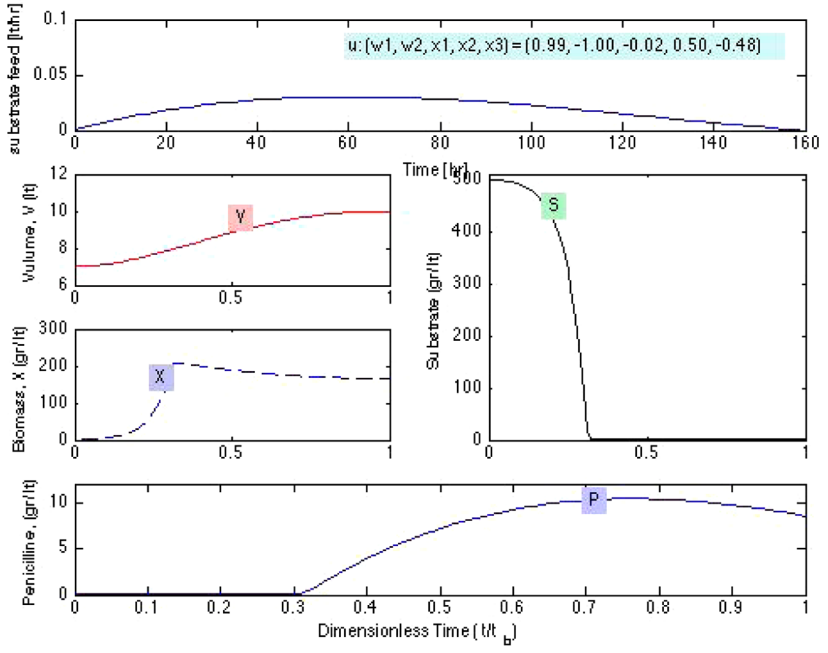

Figure 6. Time evolution of run 8, which produced the largest amount of penicillin among the 16 experiments of the DoDE design.

added to assess the LoF statistic. This statistic tells us if there is any significant variability in the data that has not been captured by the identified RSM. Three additional runs are also added (runs 9, 12, and 14) that are replicates of runs 5, 4, and 10, respectively. These replicate runs are used to assess the inherent variability of the process. Seven distinct feeding profiles of the substrate are presented in Figure 5. Some of the runs share the same substrate feeding profile but have different initial concentrations of the biomass. For example, see runs 10 and 11 in Table 6. The depicted profiles in Figure 5 do not fill the available region of flow between $u_0(\tau) - \Delta u(\tau)$ and $u_0(\tau) + \Delta u(\tau)$, because of the constraint that the total reactor volume should be increased from the initial value of 7 L to exactly 10 L.

The total amount of penicillin produced (in grams), calculated through the simulated model, is given in the seventh column of Table 6. The simulated results for runs 9, 12, and 14 are identical to runs 5, 4, and 10, respectively. To these numbers, we add a fractional experimental error using eq 28, with $\sigma = 0.02$. The measured values are given in the eighth column.

The maximum amount of penicillin produced is 85.07, obtained for run 8. Run 4 produces the minimum amount of 28.42 g of penicillin. The duration of runs 8 and 4 are 159.8 h and 100.3 h, respectively. The time evolution of run 8 is given in Figure 6.

Linear regression analysis of these results with only the three independent factors A, B, and C results in the following RSM:

$$\hat{y} = +58.00 + 3.06A - 1.77B + 27.70C + 1.39AB - 6.78B^2 + 2.56C^2 \quad (53)$$

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

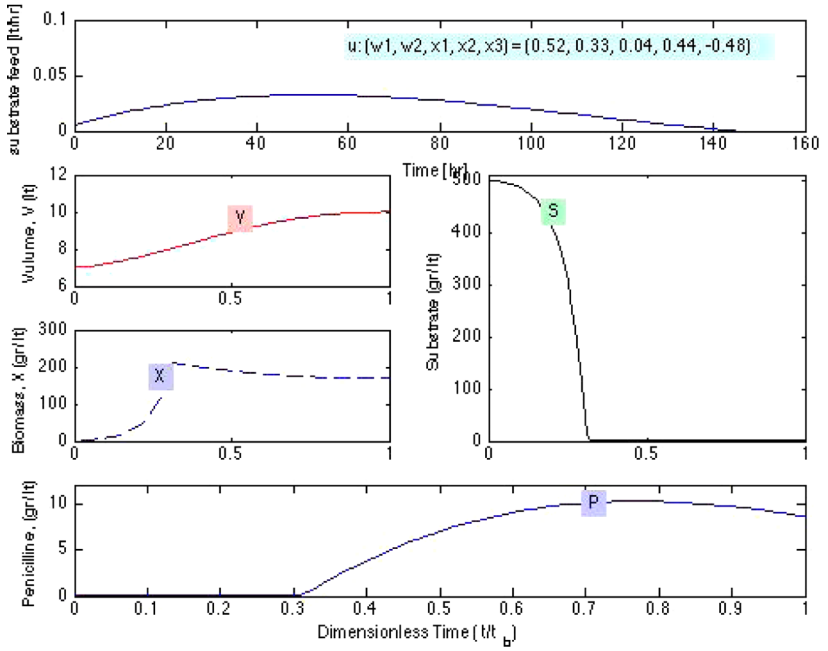

Figure 7. Time evolution of optimized run, which produced 85.89 g of penicillin.

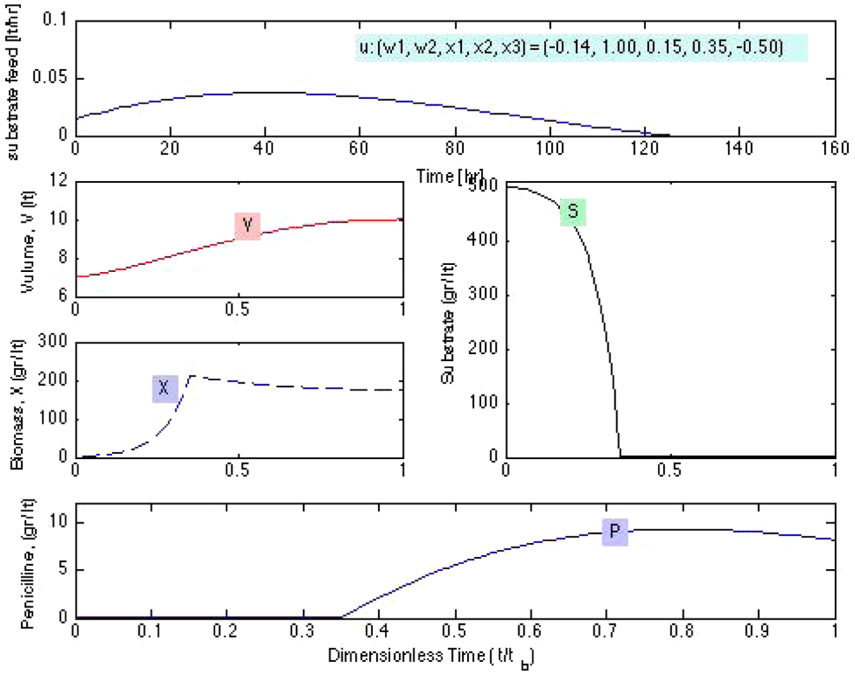

Figure 8. Time evolution of optimal productivity, which produced 80.81 g of penicillin and had a productivity of 0.643 g/h.

The related LoF statistic is $p = 0.83$, indicating that there is no variability that has not been accounted by the model. Maximization of the amount of penicillin produced in a single batch is obtained for the flowing values of independent and dependent factors.

$$\max_{A,B,C}\{\hat{y}\} \Rightarrow \begin{cases} w_1 = 0.52, \quad w_2 = A = 0.33, \quad \text{and} \\ x_1 = B = 0.04, \quad x_2 = C = 0.44, \\ x_3 = -x_1 - x_2 = -0.48 \end{cases} \tag{54}$$

The batch duration is 145.51 h and the estimated amount of penicillin that this run will yield is 82.36 ± 2.43 g. Simulation of this run yields 85.89 g of penicillin, which is not that different from the expected range. Its time evolution is depicted in Figure 7. This is only slightly better run than the best of the 16 initially designed, which was run 8.

We observe that the penicillin concentration reached a maximum before the end of the batch. If we were able to measure the penicillin concentration online, we would have observed, as revealed by Figure 7, that a penicillin concentration of >10.23 g/L would have been achieved at τ = 0.77 or t = 112.04 h. This would have occurred when the volume of the bioreactor reached 9.78, yielding a total of 100.05 g of penicillin.

dx.doi.org/10.1021/ie3035114 | Ind. Eng. Chem. Res. 2013, 52, 12369−12382

The optimum operation calculated through the use the exact model of the process has been discussed in ref 20. The calculation was restricted to a batch time of 130 h, and the maximum penicillin production was 87.9 g, which is only slightly better that that obtained using the proposed DoDE approach.

If we are interested in the productivity of the process, we should maximize the mass of penicillin (in grams) per hour. We divide the amount of penicillin produced (the last column of Table 6) by the duration of the batch increased by 24 h. This 24-h increase is used to approximately represent the time needed for emptying the completed batch and then cleaning and changing the next batch. The corresponding RSM model then is given as

$$\hat{y}_2 = \frac{\hat{y}}{130 + 30w_1 + 24} = 0.39 + 0.024A + 0.078B + 0.13C + 0.017AB + 0.051BC - 0.055B^2 \quad (56)$$

Here, the maximum productivity among the 16 runs is 0.55 g/h for run 5 and the minimum productivity is 0.21 g/h for run 3. Maximizing the penicillin productivity is achieved by a run that is characterized by the following values of the dependent and independent factors:

$$\max_{A,B,C}\left\{\hat{y}_2 = \frac{\hat{y}}{154 + 30w_1}\right\} \Rightarrow \begin{cases} w_1 = -0.14, \quad w_2 = A = 1.0, \quad \text{and} \\ x_1 = B = 0.15, \quad x_2 = C = 0.35, \\ x_3 = -x_1 - x_2 = -0.50 \end{cases}$$

The expected maximum penicillin productivity is 0.546 ± 0.025 g/h for a batch that lasts 125.8 h, and is expected to produce 80.81g of penicillin. Simulation of this run yields 80.89 g of penicillin, and its productivity is 0.643. Its time evolution is depicted in Figure 8.

6. CONCLUSIONS

Here, we have presented a new approach to experimentally optimize batch and semibatch (or fed-batch) processes, with respect to one or more time-varying decision variables, using data-driven models. The method, called the Design of Dynamic Experiments (DoDE), defines a set of experiments in which several systematically designed time-varying profiles of the decision variable is used. A response surface model, built from the performance values of each experiment, is used to optimize the process. This approach generalizes the classical Design of Experiments (DoE) methodology in which all the decision variables, or factors, are time-invariant. The proposed DoDE methodology enables us to design experiments in which the decision variables, or factors, are both time-varying and time-invariant. Two examples, a batch reaction and a penicillin fermentation, are used to demonstrate the powerful characteristics of the new methodology. The effect of measurement error (1%-4%) was investigated and it has been convincingly shown that its effect on the process optimization is not at all detrimental to the proposed approach. In the example examined, the data-driven optimum was very close to the optimum that would have been obtained if a very accurate knowledge-driven model was available.

In Figures 7 and 8, a limitation of the method was revealed, as a better performance, could be obtained at a shorter duration of the best batch. This is due to the restriction that only process measurements collected at the end of the batch were available and the characteristic of the present design that all experiments were expected to end with a final reactor volume of 10 lt. This limitation can be easily remedied by designing supplemental runs with the same characteristics as the best so far but with varying end times. Alternatively, the initial design could have been postulated with an additional time invariant factor, w_3, which was related to the final volume of the bioreactor:

$$V(t_f) = 9 + 0.5w_3, \text{ (L)}, \quad \text{with} \quad -1 \le w_3 \le +1 \quad (57)$$

With this extra factor, one must update some of the constraining equations, such as eq 49. This additional factor slightly increases the number of experiments. This is a classical dilemma of any design of experiments: the traditional experiment, where all the factors are time-invariant, or the experiment introduced here, where factors can vary with time. This dilemma is between the desire to include as many factors and the need to keep the number of experiments small.

One might also add that the proposed methodology can also be used in problems where one of the decision variables is varying with space rather with time, as well as when we have space- and time-varying factors.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +1-617-627-2573. E-mail: Christos.Georgakis@Tufts.edu.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Fotopoulos, J.; Georgakis, C.; Stenger, H. G. Use of tendency models and their uncertainty in the design of state estimators for batch reactors. *Chem. Eng. Process.* **1998**, *37* (6), 545−558.

(2) Fotopoulos, J.; Georgakis, C.; Stenger, H. G. Effect of process-model mismatch on the optimization of the catalytic epoxidation of oleic acid using tendency models. *Chem. Eng. Sci.* **1996**, *51* (10), 1899−1908.

(3) Fotopoulos, J.; Georgakis, C.; Stenger, H. G., Jr. Uncertainty issues in the modeling and optimization of batch reactors with tendency models. *Chem. Eng. Sci.* **1994**, *49* (24B), 5533−5547.

(4) Filippi, C.; Greffe, J. L.; Bordet, J.; Villermaux, J.; Barnay, J. L.; Bonte, P.; Georgaki., C. Tendency Modeling of Semibatch Reactiors for Optimization and Control. *Chem. Eng. J.* **1986**, *41* (4), 913−920.

(5) Cabassud, M.; Cognet, P.; Garcia, V.; Le Lann, M. V.; Casamatta, G.; Rigal, L. Modeling and optimization of lactic acid synthesis by the alkaline degradation of fructose in a batch reactor. *Chem. Eng. Commun.* **2005**, *192* (6), 758−786.

(6) Martinez, E. C. Model discrimination and selection in evolutionary optimization of batch processes with tendency models. *Eur. Symp. Comput.-Aided Process Eng.—15* **2005**, *20A−20B*, 463−468.

(7) Box, G. E. P.; Draper, N. R. *Response Surfaces, Mixtures, and Ridge Analysis*; Wiley: Hoboken, NJ, 2007.

(8) Montgomery, D. C. *Design and Analysis of Experiments*, 8th Edition; Wiley: New York, 2013; p 729.

(9) Atkinson, A. C.; Bailey, R. A. One hundred years of the design of experiments on and off the pages of Biometrika. *Biometrika* **2001**, *88* (1), 53−97.

(10) Draper, N. R.; Pukelsheim, F. An overview of design of experiments. *Stat. Pap.* **1996**, *37* (1), 1−32.

(11) Herzberg, A. M.; Cox, D. R. Recent Work on Design of Experiments—A Bibliography and a Review. *J. R. Stat. Soc. Ser. A* **1969**, *132*, 29−.
(12) Asprey, S. P.; Macchietto, S. Statistical tools for optimal dynamic model building. *Comput. Chem. Eng.* **2000**, *24*, 2−7.
(13) Asprey, S. P.; Macchietto, S. Designing robust optimal dynamic experiments. *J. Process Control* **2002**, *12* (4), 545−556.
(14) Hutson, V.; Pym, J. S.; Coud, M. J. *Applications of Functional Analysis and Operator Theory*, 2nd ed.; Elsevier Science & Technology Books: Amsterdam, 2005; Vol. *200*, p 426.
(15) Fiordalis, A.; Georgakis, C. Data-driven, using design of dynamic experiments, versus model-driven optimization of batch crystallization processes. *J. Process Control* **2013**, *23* (2), 179−188.
(16) Georgakis, C. A Model-Free Methodology for the Optimization of Batch Processes: Design of Dynamic Experiments. In *IFAC Symposium on Advanced Control of Chemical Processes (ADCHEM)*, Istanbul, Turkey, July 12−15, 2009; IFAC: Istanbul, Turkey, 2009.
(17) Troup, G. M.; Georgakis, C. Process systems engineering tools in the pharmaceutical industry. *Comput. Chem. Eng.* **2013**, *51*, 157−171.
(18) Rippin, D. W. T. Simulation of a Single- and Multiproduct Batch Chemical Plants for Optimal Design and Operation. *Comput. Chem. Eng.* **1983**, *7* (3), 137−156.
(19) Bajpai, R. K.; Reuss, M. A Mechanistic Model for Penicillin Production. *J. Chem. Technol. Biotechnol.* **1980**, *30* (6), 332−344.
(20) Riascos, C. A. M.; Pinto, J. M. Optimal control of bioreactors: a simultaneous approach for complex systems. *Chem. Eng. J.* **2004**, *99* (1), 23−34.