ON THE DESIGN OF OPTIMALLY INFORMATIVE EXPERIMENTS FOR MODEL DISCRIMINATION AMONG DYNAMIC CRYSTALLIZATION PROCESS MODELS

Bing H. Chen and Steven P. Asprey
Centre for Process Systems Engineering, Imperial College
London SW7 2BY, United Kingdom

Sean K. Bermingham, Andreas M. Neumann, and Herman J. M. Kramer
Laboratory for Process Equipment, Delft University of Technology
Leeghwaterstraat 44, 2628 CA, Delft, the Netherlands

Abstract

A general, systematic procedure is presented to support the development and statistical verification of dynamic crystallization process models. Within this procedure, methods are presented to address the optimal design of *dynamic* experiments for model discrimination among several crystallization kinetics models. The problem of designing dynamic experiments is cast as an optimal control problem that enables the calculation of optimal discrete sampling points, experiment duration, fixed and variable external control profiles, and initial conditions of a dynamic experiment subject to general constraints on inputs and outputs. Within this framework, methods are presented to provide experiment design robustness, accounting for parametric and thus prediction uncertainty.

Keywords

Optimal experiment design, Dynamic modeling, Model discrimination, Dynamic optimization.

Introduction

In this work, we present a general, systematic procedure to support the optimal design of *dynamic* experiments for model discrimination. We present an extension to the criterion put forth by Buzzi-Ferraris and Forzatti (1984) to obtain an information measure for discrimination in the dynamic situation, in which the system can be modeled with a set of mixed differential and algebraic equations:

$$\mathbf{f}\left(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\varphi}, t\right) = \mathbf{0} \tag{1}$$

where $\mathbf{x}$ is a state vector, $\mathbf{y}$ a set of algebraic variables, $\boldsymbol{\theta}$ a set of parameters that must be estimated from collected experimental data, $\boldsymbol{\varphi}$ a set of experimental settings, and t is time. The problem is cast as an optimal control problem, in which one can calculate optimal discrete sampling points, final time, fixed and variable external controls with input constraints, and initial conditions of a dynamic experiment. To mathematically represent time-varying external controls to the process, we use the control vector parameterization (CVP) technique (see Asprey and Macchietto, 2002).

We show the application of these statistical techniques to the development of a dynamic predictive crystallization process model. Embedded within a first-principles-based model, three crystallization kinetic sub-models are proposed, for which we use model discrimination techniques applied to the entire crystallizer model to select the best model from the candidate set. The crystallization application is used to demonstrate the overall procedure of the proposed techniques in their ability to reduce the quantity of experimental work required, while increasing the quality of the results.

Crystallization Modeling

We use a compartmental modeling approach to model the crysCODE 1100-liter DTB crystallizer, an evaporative crystallizer with a fines destruction loop (Fig. 1). The compartmental modeling approach was selected to account for the spatial distribution of process conditions, such as the supersaturation, energy dissipation and crystal size distribution (CSD) throughout the crystallizer, thus allowing

separation of the *local* intrinsic kinetics and the *overall* hydrodynamics. In this work, the well-mixed compartment model is the only building block used for the construction of compartmental models (for a further discussion on why this is so and the compartmentation procedure, refer to Bermingham, 2002). All compartments within the compartmental model are described with the same *compartment* model, *i.e.*, the same equations of conservation, physical and thermodynamic property relations, kinetic rate expressions and parameters. Differences between compartments with respect to nucleation, growth, dissolution, attrition, breakage, and aggregation rates are therefore purely a result of varying process conditions.

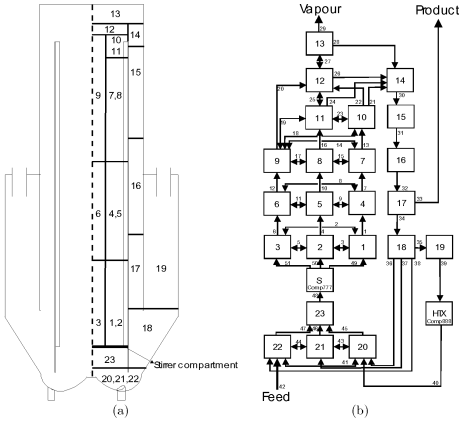

Figure 1. The compartment model network.

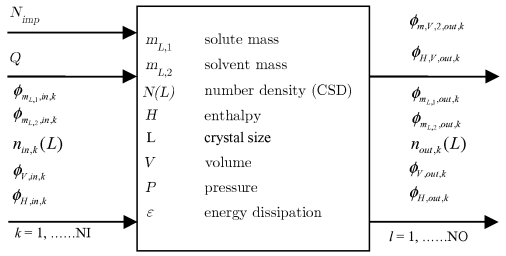

Figure 2. The compartment model variables.

A single compartment model is schematically depicted in Fig. 2, showing the states within the compartment. An arbitrary number of inlet streams and outlet streams, respectively, enter and exit the compartment. Each flow is characterized by an overall flow rate, component mass flow rates($\phi_{mL,i,in,k}$ or $\phi_{mL,i,out,k}$), crystal number densities ($n_{in,k}(L)$ or $n_{out,k}(L)$) and an enthalpy flow rate ($\phi_{H,in,k}$ or $\phi_{H,out,k}$).

In this study, we focus on the crystallization of ammonium sulphate from water. Hence, we have two liquid phase components (solute, A [component 1] and solvent, B [component 2]), one solid phase (crystals, C) and no liquid phase reactions. The crystallization reaction is thus:

$$\upsilon_{S,1}A \quad + \quad \upsilon_{S,2}B \quad \rightleftharpoons \quad C \qquad (2)$$

The stoichiometric coefficients $\upsilon_{S,1}$ and $\upsilon_{S,2}$ are equal to one and zero, respectively.

Mass balances for the solvent/solute and Energy Balance

We can write the liquid phase component mass balances for the solvent and solute generally as (see Figure 2):

$$\overbrace{\frac{dm_{L,i}(t)}{dt}}^{\text{mass rate of accumulation}} = \overbrace{\sum_{k=1}^{NI}\phi_{m_{L,i},in,k}(t)-\sum_{l=1}^{NO}\phi_{m_{L,i},out,l}(t)}^{\text{mass rates in and out via convective transport}} + \overbrace{\mathrm{M}_i\upsilon_{S,i}\left[\phi_{m,nu}(t)+\phi_{m,grow}(t)-\phi_{m,dis}(t)\right]}^{\text{interphase mass flux due to crystal growth and due to primary nucleation and dissolution at the critical nucleus size}} \qquad (3)$$

for i = 1,2, with initial condition $m_{L,i}(t=0)=m_{L,i,0}$. The subscripts *nu*, *grow* and *dis* represent nucleation, growth, and dissolution, respectively. M_i, i = 1,2 are the molecular weights for component 1 and 2, respectively.

The dynamics of the temperature are given by the energy balance:

$$\overbrace{\frac{dH(t)}{dt}}^{\text{rate of enthalpy accumulation}} = \overbrace{\sum_{k=1}^{NI}\phi_{H,in,k}(t)-\sum_{l=1}^{NO}\phi_{H,out,l}(t)-\phi_{H,V,out}(t)}^{\text{rate of total enthalpy in and out by convection}} + \overbrace{Q(t)}^{\text{net rate of heat addition}} \qquad (4)$$

With initial condition $H(t=0)=H_0$.

Population balance for crystals

The time evolution of a CSD is given by the population balance equation (PBE). The PBE for a uniformly mixed volume, with the amount and the size of particles expressed in terms of number density and particle length respectively, can be written as follows:

$$\overbrace{\frac{\partial(n(L,t)V(t))}{\partial t}}^{\text{number rate of accumulation}} = \overbrace{-V(t)\frac{\partial(n(L,t)\cdot G(L,t))}{\partial L}}^{\text{rate of number gain due to crystal growth and dissolution}} + \overbrace{\sum_{k=1}^{NI}\phi_{V,in,k}(t)n_{in,k}(L,t)}^{\text{number rate in}} - \overbrace{\sum_{l=1}^{NO}\phi_{V,out,l}(t)n_{out,l}(L,t)}^{\text{number rate out}} + \overbrace{\left[\phi_{n,nu}^{+}(L,t)-\phi_{n,dis}^{-}(L,t)\right]V}^{\text{number production rate due to primary nucleation and dissolution at the critical nucleus size}} + \overbrace{\left[\phi_{n,attr}^{\pm}(L,t)+\phi_{n,break}^{\pm}(L,t)+\phi_{n,aggl}^{\pm}(L,t)\right]V}^{\text{number production rate due to attrition, breakage and agglomeration}} \qquad (5)$$

with initial condition $n(L,t=0)=n_0(L)$. Subscripts *attr*, *break*, and *aggl* indicate attrition, breakage, and agglomeration, respectively. The initial condition can be used to express a trivial situation such as a clear liquid (no crystals) or a size distribution of seed crystals (often used in batch processes).

The classical boundary condition with respect to crystal size for the population balance equation is:

$$n(L=0,t)=\frac{B_0(t)}{G(L=0,t)} \qquad G(L)\geq 0\ \forall\ L \qquad (6)$$

B_0 denotes the birth rate of crystals with size zero, and G is growth rate. Depending on the employed kinetic model,

nucleation is either modeled using a birth rate at the boundary or as a distributed process with respect to crystal size. For a more detailed discussion of these boundary conditions, the reader is referred to Bermingham (2002).

Crystallization kinetic models

Crystallization kinetics are described using three competing models (see Table 1). The reader is referred to Bermingham (2002) for further details.

Table 1. Implemented kinetic models.

Kinetic Model	Basis	Predictive Value	Kinetic Parameters
Power Law	Purely empirical	Limited	5
Ó Meadhra (1996)	Partly first principles	Reasonable	9
Gahn (1999)	Largely first principles	Good	2

Designing Dynamic Experiments for Model Discrimination

The problem of model discrimination occurs when N_M rival models have been proposed to describe a system and it is not certain which of the models is "best". For any experiment, it is expected that the "best" model will provide the most accurate prediction of the observed measurements. Very little past work has been presented for discriminating between rival *dynamic* models (see, *e.g.*, Cooney and McDonald, 1995).

For steady-state models, Buzzi-Ferraris and Forzatti (1984) took an alternative approach to model discrimination in which they proposed a design criterion as a ratio of the average squared difference between predictions to the average variance in the predictions. As such, this criterion is similar to an *F*-statistic, and its value can be used to judge the ability to discriminate between the models. For instance, if it is possible to discriminate between models at a given experimental condition, the value of the criterion should be greater than *M*, the number of responses. If the criterion value is less than *M* at all possible experimental conditions, no discrimination can be gained, and the sequential design for model discrimination stopped.

In this work, we extend the Buzzi-Ferraris criterion for design of *dynamic* experiments, using the matrix:

$$\mathbf{V}_{l,sp}=\begin{bmatrix} \dfrac{\partial x_{l,1}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,1}} & \dfrac{\partial x_{l,1}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,2}} & \cdots & \dfrac{\partial x_{l,1}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,P_l}} \\ \dfrac{\partial x_{l,2}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,1}} & \dfrac{\partial x_{l,2}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,2}} & \cdots & \dfrac{\partial x_{l,2}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,P_l}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_{l,M}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,1}} & \dfrac{\partial x_{l,M}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,2}} & \cdots & \dfrac{\partial x_{l,M}\left(t_{sp_1},\hat{\boldsymbol{\theta}}_l\right)}{\partial\hat{\theta}_{l,P_l}} \end{bmatrix} \quad (7)$$

where $\mathbf{V}_{l,sp} \equiv$ the sensitivity matrix for model l responses at sampling time t_{sp1}. The parametric sensitivity information for models comprising a set of differential and algebraic equations (*cf.* Eqs. (2)-(6)) can be obtained by the solution of the first-order derivative of the model equations (1) with respect to the model parameters:

$$\mathbf{f}_{\dot{\mathbf{x}}}\frac{\partial\dot{\mathbf{x}}}{\partial\boldsymbol{\theta}}+\mathbf{f}_{\mathbf{x}}\frac{\partial\mathbf{x}}{\partial\boldsymbol{\theta}}+\frac{\partial\mathbf{f}}{\partial\boldsymbol{\theta}}=\mathbf{0} \quad (8)$$

Subsequently, we can define:

$$\mathbf{W}_l=\mathbf{V}_{l,sp}\boldsymbol{\Sigma}_{\boldsymbol{\theta},l}^{-1}\mathbf{V}_{l,sp}^T \quad (9)$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta},l}^{-1}$ is the inverse parameter variance-covariance matrix (non-diagonal). As in Buzzi-Ferraris and Forzatti (1984), we still form the criterion as:

$$T_{l,l',sp}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_l,\hat{\boldsymbol{\theta}}_{l'}\right)=\left(\hat{\mathbf{x}}_l\left(\boldsymbol{\theta},t_{sp_k}\right)-\hat{\mathbf{x}}_{l'}\left(\boldsymbol{\theta},t_{sp_k}\right)\right)^T S_{l,l'}^{-1}\left(\hat{\mathbf{x}}_l\left(\boldsymbol{\theta},t_{sp_k}\right)-\hat{\mathbf{x}}_{l'}\left(\boldsymbol{\theta},t_{sp_k}\right)\right) \quad (10)$$

where $\mathbf{S}_{l,l'} = \mathbf{S} + \mathbf{W}_l + \mathbf{W}_{l'}$ and $\mathbf{S}$ is the (non-diagonal) covariance matrix of experimental error. In the case of dynamic design, we sum the criteria Eq. (10) over all discrete sampling times:

$$\max_{\boldsymbol{\varphi}}\sum_{sp=1}^{n_{sp}}T_{l,l',sp}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_l,\hat{\boldsymbol{\theta}}_{l'}\right) \quad (11)$$

With the new dynamic formulation, we extrapolate the Buzzi-Ferraris and Forzatti observation mentioned above, requiring maximization of (11) above ($n_{sp}\times M$) to ensure discriminatory power of the designed experiment.

We embed this criterion in the dynamic optimization framework for designing dynamic experiments as presented in Asprey and Macchietto (2002). For time-varying inputs to the process, we use the control vector parameterization technique, using a piecewise constant parameterization. As such, the vector of design variables $\boldsymbol{\varphi}$ is comprised as:

$$\boldsymbol{\varphi}=\left[t_{sp}^T,t_{sw_{1,1}},\ldots,t_{sw_{n_u,nsw-1}},z_{1,1},\ldots z_{n_u,nsw},y_0^T\right] \quad (12)$$

where $t_{sw_{i,j}}$ correspond to j^{th} times at which the i^{th} input switches levels, denoted $z_{i,j}$. In this formulation the sampling points are the same for all response variables, and we have a flexible final time for the end of the experiment, defined by the final designed sampling point. The resultant non-linear problem is solved using the NAG E04UFF SQP.

Results and Discussion

In the ammonium sulphate and water system, an accurate measurement of the supersaturation is extremely difficult; however, detailed CSD measurements are available at two-minute intervals for each experiment by use of light-scattering measurements. By using more characteristics of the CSD than the median size and/or using the observed

dynamics in the CSD, the absence of a supersaturation measurement can (in principle) be countervailed. For parameter estimation purposes, to characterize the CSD, the median size (L_{50}) and the 10%- and 90%-quantiles (L_{10} and L_{90}, respectively) were chosen. Three existing data sets were analyzed using our extended dynamic divergence measure, results of which are shown in Table 2 (500 sampling points were used in each case).

Table 2. Dynamic divergence measures (heat input of 120 kW/m^3 and impeller speed as indicated).

Data set	Div(G, O)	Div(O, P)	Div(G, P)
1 (640 RPM)	19038	6209	4077
2 (775 RPM)	1766	5943	5486
3 (910 RPM)	13663	5660	25613.7
Optimal	45134	51861	72560.0

As can be seen in Table 2, experiment #1 is relatively informative for discriminating between the Gahn (G) and Ó Meadhra (O) models, while experiment #3 is relatively informative for discriminating between the Gahn (G) and Power Law (P) models. All existing experiments are non-informative for discriminating between the Ó Meadhra (O) and Power Law (P) models.

The final row of Table 2 shows results of optimal experiment designs obtained using our extended criterion (Eqn. (11)) and piece-wise constant control inputs, with a maximum experiment duration of 2 days. An *a priori* selection of the number of sampling points (500, based on the fixed two-minute intervals of the light-scattering measurements), along with the number of switching times/constant levels of the input profiles is required. We arbitrarily chose five time intervals for the control vector parameterization ($nsw_i = 5$, $i = 1,..,n_u$; where $n_u = 2$ corresponding to impeller speed and residence time, respectively). We impose the constraints:

$$t_{sw_{i,j}} - t_{sw_{i,j-1}} \geq 10.0 \quad i = 1 \ldots n_u; j = 1 \ldots nsw \qquad (13)$$

As can be seen in Table 2, these optimal designs are at least three times more informative than the existing data sets for the purpose of discrimination. For illustrative purposes, the model predictions for the optimal (G, O) model discrimination design are shown in Fig. 3, while the designed inputs are shown in Table 3.

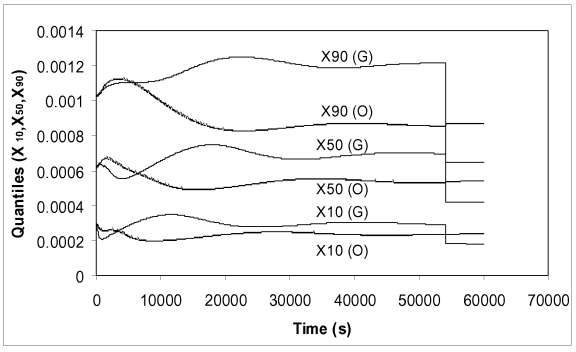

Fig. 3. The model predictions for optimal model discrimination between (G) and (O).

Table 3. The optimally designed inputs for model discrimination between (G) and (O).

$t_{sw,i}$; i=1,..5 (s)	0	600	1200	53400	54000
Residence time (s)	2000	2000	6500	6500	6500
Impeller Speed (RPM)	600	600	600	600	1080

Under the optimally designed conditions for discrimination between the Gahn (G) and Ó Meadhra (O) models, we subsequently simulated additional data using a "true" Gahn model under these experimental conditions, with a homoscedastic Normal pure error variance of 6.0% for all three responses, with 2.13% covariance between each response. Subsequent parameter estimation was carried out, giving the results in Table 4, in which we report a univariate χ^2 lack-of-fit (LOF) test for each model.

Table 4. Lack-of-fit analyses. Reference χ^2 =925.7

Quantile	Gahn	Ó Meadhra	Power Law
X_{10}	737.1	27517.5	25410.4
X_{50}	688.3	10224.9	11215.2
X_{90}	656.3	7025.2	6413.6

From inspection of Table 4, as expected, the Gahn model is the only model to pass the lack-of-fit test (calculated LOF < reference χ^2) among the three candidates, even after just one optimally designed dynamic experiment.

Concluding Remarks

In this paper, we have presented a general, systematic procedure to support the optimal design of *dynamic* experiments for model discrimination. The problem of designing dynamic experiments was cast as an optimal control problem, where, using our formulation, one can independently calculate optimal sampling points, experiment duration, fixed and variable external controls with state and input constraints, and initial conditions of a dynamic experiment. Within this framework, methods for accounting for parameter uncertainty were presented to provide experiment design robustness. Application of the tools was carried out using a compartmental crystallization model, demonstrating their effectiveness and power in aiding to build a dynamic process model.

References

Asprey, S. P. and S. Macchietto. (2002). Designing Robust Optimal Dynamic Experiments. *J. Proc. Control*, **12**, 545.

Bermingham, S. K. (2002). PhD Thesis, Delft University of Technology

Buzzi-Ferraris, G.; Forzatti, P. (1984). Sequential Experimental Design for Model Discrimination in the Case of Multiple Responses. *Chem. Eng. Sci*. **39**, 81.

Cooney, M. J. and K. A. McDonald. (1995). Optimal Dynamic Experiments for Bioreactor Model Discrimination. *Appl. Microbiol. Biotechnol*, **43**, 826.