Model Identification and Control Strategies for Batch Cooling Crystallizers

Stephen M. Miller and James B. Rawlings
Dept. of Chemical Engineering, The University of Texas at Austin, Austin, TX 78712

The open-loop optimal control strategy to regulate the crystal-size distribution of batch cooling crystallizers handles input, output, and final-time constraints, and is applicable to crystallization with size-dependent growth rate, growth dispersion, and fines dissolution. The objective function can be formulated to consider solid-liquid separation in subsequent processing steps.

A model-based control algorithm requires a model that accurately predicts system behavior. Uncertainty bounds on model parameter estimates are not reported in most crystallization model identification studies. This obscures the fact that resulting models are often based on experiments that do not provide sufficient information and are therefore unreliable. A method for assessing parameter uncertainty and its use in experimental design are presented. Measurements of solute concentration in the continuous phase and the transmittance of light through a slurry sample allow reliable parameter estimation. Uncertainty in the parameter estimates is decreased by data from experiments that achieve a wide range of supersaturation. The sensitivity of the control policy to parameter uncertainty, which connects the model identification and control problems, is assessed. The model identification and control strategies were experimentally verified on a bench-scale KNO_3-H_2O system. Compared to natural cooling, increases in the weight mean size of up to 48% were achieved through implementation of optimal cooling policies.

Introduction

Distributed parameter systems are processes with spatially varying states, controls, and parameters. The states, controls, and parameters of a crystallizer can be spatially distributed, but they are also distributed over a population of crystals, giving rise to the challenges of characterizing and controlling crystallizers. The population balance approach (Randolph and Larson, 1962; Hulburt and Katz, 1964) provides a modeling framework that enables representation of the distributed nature of dispersed-phase systems such as crystallizers.

The quality of a crystalline product is usually specified in terms of the crystal size, shape, and purity. Although the population balance approach allows consideration of the distribution of shape and purity of a population, this study deals with the modeling and control of crystal size and crystal-size distribution (CSD). Customer quality requirements of a product are often stated in terms of ability to flow, dissolution rate, aesthetic appeal—all primarily functions of crystal size and CSD. For products to be used in photographic materials, size uniformity is so critical that the CSD is the principal consideration of a customer. If acceptable CSD and purity standards are not met, the crystals must undergo further processing steps, such as milling or recrystallization. In addition to customer requirements, a concern of the manufacturer is the CSD-influenced efficiency of downstream processes such as thickening and filtration, often the time-limiting steps in crystallization operation.

As discussed in the review by Rawlings et al. (1993), there have been many attempts at continuous crystallizer control since the development of population balance models. Despite the activity in the area of continuous crystallizer control, there have been relatively few control algorithms developed for batch crystallizers.

As for any batch process, batch crystallizer control requires a dynamic operation policy. The control algorithms that have

Correspondence concerning this article should be addressed to J. B. Rawlings.

been developed for control of batch crystallizers are limited to specific crystallizer configurations and do not allow incorporation of constraints. The applicability of these methods has been restricted further by inaccurate process measurements and failures to properly identify models that characterize system behavior. As discussed by Rawlings et al. (1993) and Miller (1993), the limitations of measurement technology for crystallization systems has confounded parameter estimation, but this difficulty has been obscured by the fact that most of the previous studies of crystallization kinetic parameter estimation have not provided an assessment of the reliability of the parameter estimates.

This article presents the development of a flexible, model-based strategy for batch cooling crystallizers that handles input, output, and final-time constraints. The interplay between process measurements, model identification, and model-based control is significant. The determination of model parameter uncertainty allows analysis of the information content of data. The model identification technique discussed here includes an assessment of parameter uncertainty. This assessment is used as an aid for experimental design and to examine the sensitivity of the determined control policy to parameter uncertainty—the issue that connects the model identification and control problems.

Crystallization of potassium nitrate from water in a batch cooling crystallizer, the system used for verification of the model identification technique, was chosen for experimental testing of the control strategy.

Experimental Apparatus

Figure 1 shows the bench-scale, batch cooling crystallization system used to test the model identification and control strategies discussed in this article. The crystallizer is a 3-L jacketed vessel. Samples for the Malvern 3600Ec particle sizer and the Anton Paar densitometer are drawn continuously from the crystallizer using peristaltic pumps.

The sample stream for the densitometer is maintained free of solids by a 15 micron screen, and a heat exchanger is used to keep the inlet to the two-millimeter, flow-through measuring cell approximately isothermal. For two-component systems, the densitometer has proven to allow accurate determination of the solute concentration of the solids-free phase.

The Malvern 3600Ec particle sizer is an instrument that is based on laser light scattering. The theory behind light scattering methods for particle sizing is discussed by Boxman (1992) and Rawlings et al. (1993). The CSD measurement provided by the Malvern particle sizer is not reliable enough for the purposes of this study. The limitations of the particle sizer are due primarily to the facts that inference of the CSD from scattered light measurements requires the solution of an ill-conditioned inverse problem, and results are based on the assumption that the particles are spherical. As discussed by Witkowski et al. (1990), information about the CSD can be obtained from a measurement of the transmittance, the fraction of light that is transmitted through the crystal slurry. As explained by Rawlings et al. (1993), the transmittance measurement does not suffer from some of the problems encountered in CSD determination. Although the Malvern particle sizer is used to obtain a transmittance measurement, a relatively inexpensive spectrophotometer could be used to make this measurement.

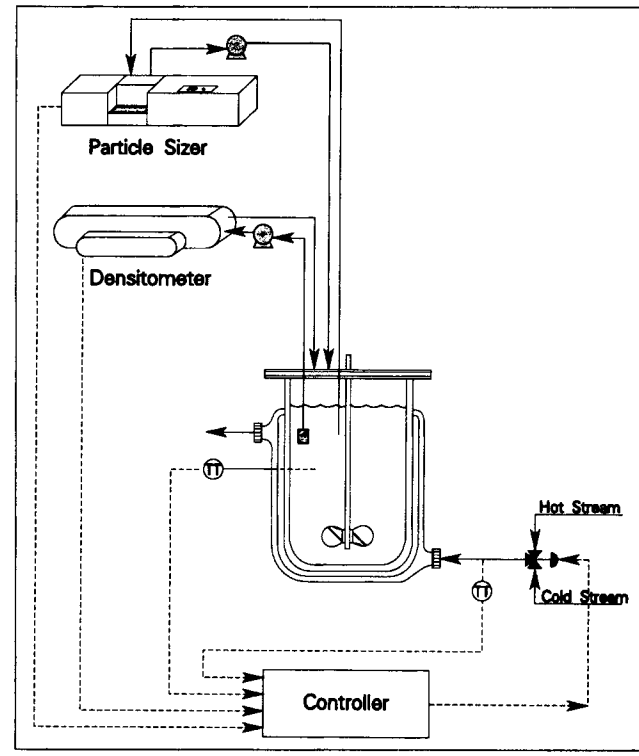

Figure 1. Batch crystallization system.

The Beer-Lambert law provides a simple model that relates the transmittance to the CSD:

$$\text{transmittance} \equiv \frac{I}{I_0} = \exp\left[-l \frac{k_a}{2} \int_0^\infty f(L) L^2 dL \right] \qquad (1)$$

where l is the instrument flow-cell width, L is the characteristic crystal size, f is the crystal distribution function ($\int_{L_1}^{L_2} f dL$ is the number of crystals per slurry density with size between L_1 and L_2), and k_a is the surface area shape factor. As can be seen, there is a one-to-one relationship between transmittance and the second moment of the CSD. The Beer-Lambert law is based on the assumption that there are no multiple scattering effects. Miller (1993) demonstrates that this expression is valid for $I/I_0 > 0.7$. For cases in which on-line measurements from light scattering instruments are required and the slurry density is very high, Jager et al. (1987) discuss the application of an automatic dilution unit. The challenge of this approach is maintaining a diluent of proper quality to assure crystal stability and a representative sample. Alternatively, an instrument designed such that the path length of light is very small would negate the limitation of the Beer-Lambert law.

A cascade configuration is used to control the crystallizer temperature. Disturbances in the ambient temperature and in jacket supply water temperature are rejected by a secondary loop that uses a simple PI controller to maintain the jacket temperature at its setpoint by manipulating the ratio of the flow to the jacket of a hot stream and a cold stream. The primary controller is a model-predictive controller that uses the crystallizer temperature measurement and setpoint trajectory to compute a setpoint for the jacket temperature controller. A "natural cooling" profile is effected by disabling

the primary controller and entering a step change in jacket temperature setpoint.

All of the experiments discussed in this article were seeded crystallizer runs. The time at which the seeds were introduced corresponds to $t=0$. With the exception of a few specifically noted runs, the seed crystals were taken from the 180–212 μm sieve class, and the seed loading was low. While it has been demonstrated that seeding can take place at significantly higher supersaturation levels without affecting product quality (Karpinski et al., 1980), seeding was performed at or near the saturation temperature.

A detailed description of the construction and operation of this system, and the physical properties of the KNO_3-H_2O system are given by Miller (1993).

Model Identification

Models derived strictly from basic physical laws of nature do not require experimentation. Nevertheless, at least a portion of the model is usually empirical; in this case model identification is a twofold process that involves model formulation and parameter estimation.

The population balance approach provides a framework for determining the structure of models for crystallizers, but the kinetic parameters of the constitutive relations must be estimated for a particular chemical system. The model identification is not complete until the parameter uncertainty and the model reliability are assessed. Since the deficiency of a model can be due to inadequate information content of the data used in parameter estimation, experiments used for model identification should be carefully designed.

A more detailed discussion of the issues covered in this section is given by Miller (1993).

Model formulation and solution

As discussed by Rawlings et al. (1993), stochastic population balances are necessary for small populations in which the fluctuations about the mean number of particles become of the same order of magnitude as the mean itself; however, a deterministic population balance is appropriate for most cases of crystallizer modeling.

The population balance for a jacketed batch cooling crystallizer without fines dissolution is:

$$\frac{\partial f}{\partial t} = -\frac{\partial (Gf)}{\partial L} \tag{2}$$

where G is the crystal growth rate. The boundary condition for this partial differential equation is:

$$f(L_0, t) = \frac{B^o}{G|_{L=L_0}} \tag{3}$$

where B^o is the rate of formation of new crystals (nucleation rate), which have size L_0 upon formation.

Constitutive relations are required for growth and nucleation. These phenomena can occur only if the solute concentration exceeds its saturation value. The supersaturation (the degree to which the concentration exceeds the saturation concentration) is the driving force for the formation and growth of crystals. The following empirical power law expressions have become standard for growth and nucleation rates for seeded crystallizers:

$$G = k_g S^g \tag{4a}$$

$$B^o = k_b S^b \hat{\mu}_3 \tag{4b}$$

where $S = (C - C_{sat})/C_{sat}$ and $\hat{\mu}_3$ is the third moment of the CSD with units cm^3/g solvent. These expressions are appropriate for cases in which the solubility of the solvent is such a strong function of temperature that high yield can be obtained with small temperature variations. Otherwise, the temperature dependence can be incorporated using an Arrhenius-type expression.

Because of the dependence of growth and nucleation on concentration and temperature (S is a function of temperature), mass and energy balances are required to complete the description of the crystallization process. General forms of the solute mass balance in the continuous phase and the energy balance are given by Miller (1993). For a jacketed batch cooling crystaller without fines dissolution they can be reduced to:

$$\frac{dC}{dt} = -3\rho_c k_v h \int_0^\infty fL^2 G dL \tag{5}$$

$$\rho V c_p \frac{dT}{dt} = -3\Delta H_c \rho_c k_v V \int_0^\infty fL^2 G dL - UA_c(T - T_c) \tag{6}$$

where U is the jacket heat-transfer coefficient, A_c is the heat-transfer area, T_c is the cooling fluid temperature, ΔH_c is the heat of crystallization, k_v is the volume shape factor, ρ_c is the crystal density, C is the solute concentration on a per-mass-solvent basis, and h is a conversion factor equal to the volume of slurry per mass of solvent. The conversion of the basis for solute concentration is done to be consistent with convention for batch crystallizers. Equations 2, 5, and 6 are obtained by assuming that V, ρ, and the system pressure are constant and the cooling system is the only mechanism for external heat loss/gain.

The equations given above and the following initial conditions constitute the model used in this study:

$$f(L, t) = f_0(L), \quad t = 0 \tag{7a}$$

$$C(t) = C_0, \quad t = 0 \tag{7b}$$

$$T(t) = T_0, \quad t = 0 \tag{7c}$$

Analytical solutions of the coupled set of nonlinear integro-differential equations of crystallizer models do not exist. Nevertheless, a variety of numerical methods have been employed to solve specific models on a case-by-case basis; several of these methods are reviewed by Rawlings et al. (1993). For the case of size-independent growth rate and no fines dissolution, the method of moments provides a way to reduce the population balance to a set of ODEs, for which there are well-developed solution techniques (Hulburt and Katz, 1964). The recovery of the CSD from the quantities resulting from this transformation is difficult. However, the method of moments

is adequate for the model used in this study, because the mass and energy balances and transmittance equation can be expressed in terms of the second moment of the distribution:

$$\frac{dC}{dt} = -3\rho_c k_v h G \mu_2 \tag{8a}$$

$$\rho V c_p \frac{dT}{dt} = -3\Delta H_c \rho_c k_v V G \mu_2 - UA_c(T - T_c) \tag{8b}$$

$$\frac{I}{I_0} = \exp\left(-\frac{k_a}{2} l \mu_2\right) \tag{8c}$$

Therefore, predictions of the outputs of interest (concentration and transmittance) are provided by this method. Even predictions of sieve analysis data—the weight mean size and the coefficient of variation of the CSD—can be obtained from the first five moments.

The development of a solution method that is applicable to cases of size-dependent growth rate is given by Miller (1993). It should be noted that because of such solution techniques, the work presented in this article is not limited to crystallization with size-independent growth rate.

Parameter estimation

A wide variety of experimental configurations and techniques have been used for crystallization parameter estimation; Rawlings et al. (1993) provide a review of these studies. Parameter estimation can be viewed as the action of an operator on a set of data. With this abstraction of the relationship between experimental data and the parameters, the requirements for the unique solution of a parameter estimation problem can be stated succinctly (Miller, 1993). Nevertheless, the checks to determine if the problem is well-posed are not straightforward. Measurements are limited and corrupted with noise, the model structure (on which the parameter estimation operator depends) may be in error, and, as is the case for crystallizer kinetic parameters, the parameter estimation operator is often nonlinear. Also, there are challenges that are peculiar to parameter estimation for distributed parameter systems such as crystallizers (Seinfeld, 1969).

The general nonlinear parameter estimation problem can be cast as a nonlinear optimization problem in which the objective function is a measure of how well the model solution with a given set of parameter values predicts the experimental data. The weighted least-squares objective function is a common choice because it is intuitive and does not involve explicit consideration of the error structure of the experimental data. It also allows the scaling of measurements that have different dimensions and weighting so that parameter estimates will be most influenced by the measurements felt to be most reliable. Although the weights can be obtained from secondary experiments, they are usually set somewhat arbitrarily.

The maximum likelihood method enables the estimation of weights for the data along with the estimation of parameters. If one assumes that the model structure is correct, then the deviations between the data and the model predictions are due to random (and possibly biased) errors in the measurements. The maximum likelihood method can be thought of as determining the parameters that maximize the probability that a given set of data could be obtained, considering the error structure of the measurements.

Consider the following assumptions about the error structure of the measurements: independent variables are deterministic, errors at the jth sampling instant are normally distributed with zero mean and are homoscedastic, errors at different sampling instants are uncorrelated, and errors in the measured variables are independent (measured variable covariance matrix, V, is diagonal). With these assumptions, it can be shown (Bard, 1974) that the maximum likelihood method is equivalent to the minimization of:

$$\Phi(\theta) = \frac{N_d}{2} \sum_{i=1}^{N_m} \log\left[\sum_{j=1}^{N_d} e_{ij}^2(\theta)\right] \tag{9}$$

where N_m is the number of measured variables (that is, concentration, transmittance, and so on), N_d is the number of samples of each measured variable, $e_{ij} = y_{ij} - \tilde{y}_{ij}$, and y_{ij} and $\tilde{y}_{ij}$ denote the actual and model-predicted values of the ith measured variable at the jth sampling instant. In this study, the successive quadratic programming code NPSOL (Gill et al., 1986) was used to solve this optimization problem.

The method also yields an estimate for the elements of the diagonal measurement error covariance matrix:

$$V_{ii} \approx \tilde{\sigma}_i^2 = \frac{1}{N_d} \sum_{j=1}^{N_d} e_{ij}^2(\theta^*) \tag{10}$$

where θ^* is the set of optimal parameters.

Equations 9 and 10 correspond to the assumptions given above, but the maximum likelihood method is not restricted to these assumptions (Bard, 1974).

Interpretation of estimates

Although many crystallization parameter estimation studies do not provide an assessment of the reliability of the parameter estimates, the model identification is not complete until the parameter uncertainty is considered. As discussed above, the estimation process can be thought of as a nonlinear operator that transforms a data set with random errors into an estimate of the parameters. The parameter estimates are random variables with a probability distribution that depends on the nature of the estimator and on the distribution of the measurement errors. The distribution of the estimates, commonly referred to as the sampling distribution, can be used to establish a region of confidence for the parameter estimates.

If many replicates of the experiments used to estimate parameters were available, the sampling distribution could be rigorously characterized. However, replicate data from most crystallization experiments are difficult and time-consuming to obtain.

An approximation of the confidence region can be obtained by assuming that the model can be represented by linear functions in the vicinity of the estimate:

$$\tilde{y}_j(\theta) \approx \tilde{y}_j(\theta^*) + \boldsymbol{B}_j(\theta - \theta^*) \tag{11}$$

where $\tilde{y}_j^T(\theta) = [\tilde{y}_{1j}(\theta), \ldots, \tilde{y}_{N_m j}(\theta)]^T$, and $\boldsymbol{B}_j$ is the $N_m \times N_p$ matrix (where N_p is the number of parameters) given below:

$$B_j = \left.\frac{\partial \tilde{y}_j}{\partial \theta}\right|_{\theta=\theta^*}$$

With the linearized model and the assumption that the measurement errors are random and normally distributed, the estimates are normally distributed, and the quantity:

$$(\theta-\theta^*)^T\left(\sum_{j=1}^{N_d} B_j^T V^{-1} B_j\right)(\theta-\theta^*)$$

is distributed as χ^2 with N_p degrees of freedom (Bard, 1974). V is the measurement error covariance matrix. Therefore, letting $V_\theta^{-1} = \Sigma_{j=1}^{N_d} B_j^T V^{-1} B_j$ (V_θ is the parameter covariance matrix for the linearized problem),

$$\Pr\{(\theta-\theta^*)^T V_\theta^{-1}(\theta-\theta^*) \le \chi^2_{N_p}(\alpha)\} = 1-\alpha \qquad (12)$$

where $Pr\{E\}$ denotes the probability of some event E. In other words, the approximate $100(1-\alpha)\%$ confidence region from the linearization method is the ellipsoidal region defined by:

$$(\hat{\theta}-\theta^*)^T V_\theta^{-1}(\hat{\theta}-\theta^*) = \chi^2_{N_p}(\alpha) \qquad (13)$$

where $\hat{\theta}$ represents the true parameter set.

Note that for cases in which the sampling rates of different measurements are not equal:

$$V_\theta^{-1} = \sum_{i=1}^{N_m}\sum_{j=1}^{N_{d_i}} B_{ij}^T V^{-1} B_{ij}$$

where B_{ij} is the $1 \times N_p$ vector defined as:

$$B_{ij} = \left.\frac{\partial \tilde{y}_{ij}}{\partial \theta}\right|_{\theta=\theta^*}$$

The meaning of the report of a confidence region is usually not immediately clear for cases in which $N_p > 2$; therefore, confidence intervals—$\Pr\{\hat{\theta} \in [\theta^* \pm \delta\theta(\alpha)]\} = (1-\alpha)$—are often desired. Conservative estimates of these can be obtained by locating the box in the parameter space that circumscribes the ellipsoid.

Experimental design

Experiments should be designed such that the data contain information sufficient for reliable parameter estimates. As discussed by Rippin (1988), experimental design involves choices of experimental configuration, variables measured, and operating conditions, each of which can influence the information content of the data.

An assessment of parameter uncertainty allows a precise statement of the goal of experimental design: choose the experimental configuration, measurements, and operating conditions such that the confidence region is minimized.

The test system for the control strategy discussed subsequently was a batch KNO_3-H_2O crystallizer. This is advantageous for experimental design, because the dynamic nature of batch crystallization allows much information about a system to be obtained from a single run.

Table 1. Results of Parameter Estimation Based on Concentration Measurements

	$\ln(k_g)$	g	$\ln(k_b)$	b
Estimate	8.014	1.10	14.710	0.85
Interval	±0.423	±0.12	±6.268	±2.48

An exploration of the remaining choices of experimental design—measured variables and operating condition—is provided by Rawlings et al. (1993). It is shown that measurement of continuous-phase concentration, the sole data source in many parameter estimation studies, provides insufficient information for identifying a model. Although the model fit is good, there are many combinations of parameters that have roughly the same ability to predict the observations. This is illustrated by the large uncertainty shown in Table 1 for the parameter estimates resulting from concentration measurements; the parameters are those for the kinetic expressions given in Eq. 4.

As intuition would suggest, solid-phase information is necessary to obtain kinetic parameters for both crystal growth and nucleation. Although there is a limit to its range of applicability, the Beer-Lambert law provides the model given by Eq. 1 that relates the CSD and the transmittance measurement; the result is a one-to-one connection between transmittance

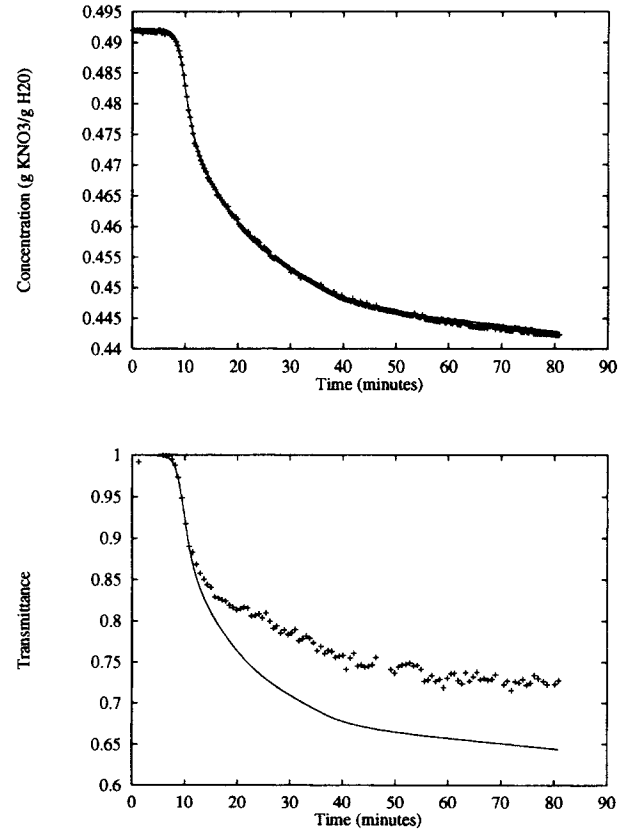

Figure 2. Concentration and transmittance data and predictions from model with parameters given in Table 1.

Table 2. Parameter Estimates Resulting from Concentration and Transmittance Measurements from Two Batch Runs

	$\ln(k_g)$	g	$\ln(k_b)$	b
Estimate	8.918	1.32	17.207	1.78
Interval	±0.127	±0.03	±0.356	±0.09

and the second moment of the CSD. The fit of the model with the parameters obtained using concentration data from a natural cooling run is shown in Figure 2. The fit to the concentration data is good, but the fit to the transmittance data is not satisfactory.

As shown by Rawlings et al. (1993), the transmittance data used together with the concentration data provide sufficient information for reliable parameter estimation. The uncertainty in the parameter estimates can be reduced further through experiments with operating conditions chosen to achieve a wide range of supersaturation. The parameter estimates resulting from concentration and transmittance measurements from batch runs with two different cooling profiles is given in Table 2. The model fit is shown in Figure 3.

The seed load used in the experiments discussed above was quite light. The rationale behind this was that experiments with light seed loading would be expected to yield more information about nucleation than experiments in which the solute is primarily deposited on the seeds. This initial condition is another operating variable that could be used to affect the quality of parameter estimation.

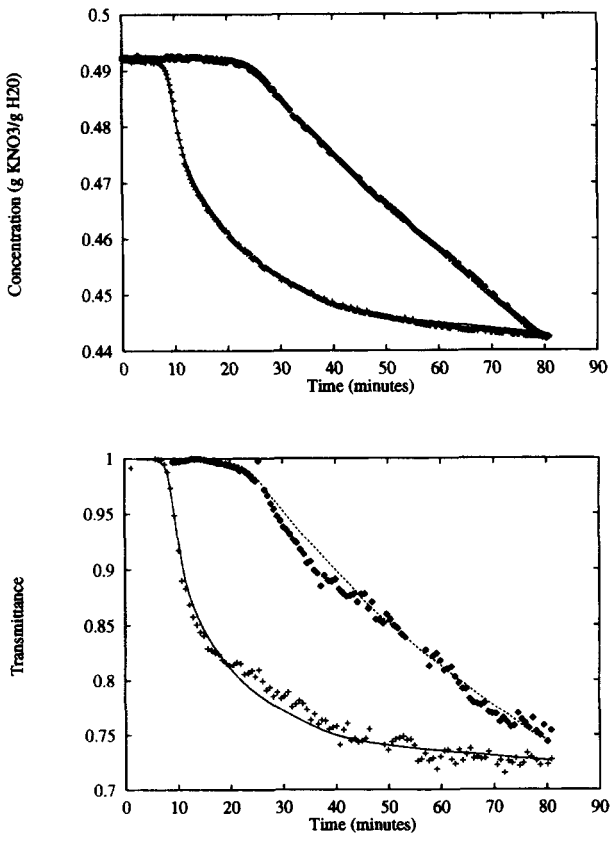

Figure 3. Fit of model with parameters given in Table 2 to concentration and transmittance data.

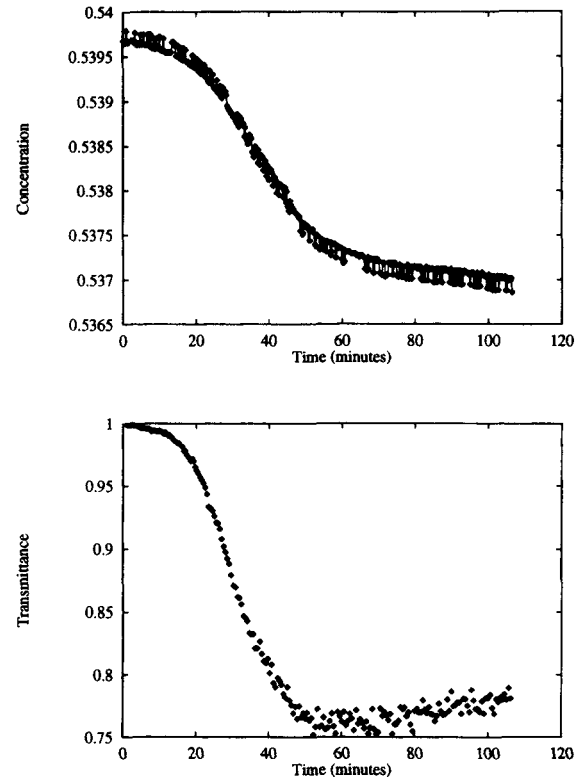

Figure 4. Data from seeded, isothermal batch crystallization of naphtalene from naphthalene-toluene solution.

Naphthalene-toluene system

The methodology described in this article is not limited to inorganic chemical systems with cubic crystal morphology such as potassium nitrate and water. A similar approach, but with a weighted least-squares objective function, has been applied to model identification for batch crystallization of naphthalene from a naphthalene-toluene solution; the details are given by Witkowski et al. (1990). Concentration and transmittance data obtained by Witkowski for an isothermal run is given in Figure 4. The naphthalene crystals of this organic system are plate-like, and the character of the system is much different from the aqueous KNO_3-H_2O system with nearly cubic crystals. Nevertheless, parameter estimation based on the concentration and transmittance measurements was accomplished. It should be pointed out that the convergence of the optimization problem and the quality of the resulting parameter estimates are inferior to those obtained for the KNO_3-H_2O system.

The ability to obtain reasonably reliable parameter estimates indicates that the method is applicable to the naphthalene-toluene system. The difficulties encountered probably indicate a deficiency in the experimental design. The data shown in Figure 4 is qualitatively similar to that obtained from the KNO_3-H_2O system. However, the data from the naphthalene-toluene system correspond to an isothermal run. Therefore, the noise-to-signal ratio is accentuated by the small dynamic range of the measurements. Also, the information content of the data is less than that obtained for the KNO_3-H_2O system, because the variation of the supersaturation over the run is much smaller for isothermal operation.

Verification of parameter estimation technique

The maximum likelihood method requires some assumptions about the measurement error structure. Parameter estimation should be followed by a review of these assumptions. For example, Eq. 9 is based on the assumption that the errors in the concentration and transmittance measurements are independent; this assumption was verified by using a check of the residuals, e_{ij}.

Another assumption was that there is no autocorrelation (serial correlation) in the errors of each of the measurement types. Tests of this assumption for the ith measurement type are also usually based on the residuals. A commonly-used indicator of autocorrelation is referred to as the 1st autocorrelation, which is the correlation between the series e_{ij} and the series of residuals shifted one step, $e_{i(j-1)}$. The residuals for the model fit shown in Figure 3 indicate significant autocorrelation. In fact, the 1st autocorrelations for the concentration and transmittance measurements are 0.78 and 0.72, respectively. As explained by Brook and Arnold (1985), high autocorrelations could suggest that the model neglects some physical effects, but the presence of high autocorrelations is not a basis for model rejection and does not mean that the estimated parameters are biased. Bard (1974) points out that for models fitted to good data, nonrandomness of residuals is the rule rather than the exception; this is because any neglected effects are prominent when the fit is very good, as in the case of these crystallization experiments.

For a particular model and experimental design, the decision to use linear statistics to obtain parameter confidence regions for the linearized model should also be reviewed. The validity of the linear statistics can be tested using a Monte Carlo study. To do so, the following procedure was followed:

- The data set shown in Figure 3 was used as the base data set.
- 500 "replicate runs" were constructed by adding error to the base data set according to pseudo-random numbers drawn from a normal distribution with zero mean and covariance V—the assumed error structure. Parameter estimation corresponding to the minimization of Eq. 9 is based on a diagonal measurement error covariance matrix with $V_{ii} = \sigma_i^2 \approx \tilde{\sigma}_i^2$. For this study, $i = 1$ corresponds to concentration measurements, and $i = 2$ corresponds to transmittance measurements. For the base-case parameter estimation, Eq. 10 gives $\tilde{\sigma}_1 = 0.0005$ and $\tilde{\sigma}_1 = 0.009$; these values were used to construct the Monte Carlo replicate data.
- For each Monte Carlo replicate data set, it was recorded whether or not the parameter estimates fall in the linear statistics confidence region defined by Eq. 13. This check was performed for various confidence levels (that is, various values of α).

The results of this study are summarized in Figure 5. This Monte Carlo study does not deal with the issue of the adequacy of the model structure or the validity of the error structure assumption; however, it does confirm that the elliptical confidence regions from linear statistics adequately characterize the parameter uncertainty for the system of this study.

Although it is not practical to perform enough experiments to allow construction of a confidence region, a replicate data set was obtained by repeating the two experiments used in the parameter estimation summarized in Table 2. The parameter estimation results for the replicate data are given in Table 3. A comparison of Tables 2 and 3 demonstrates agreement between the results from the two data sets.

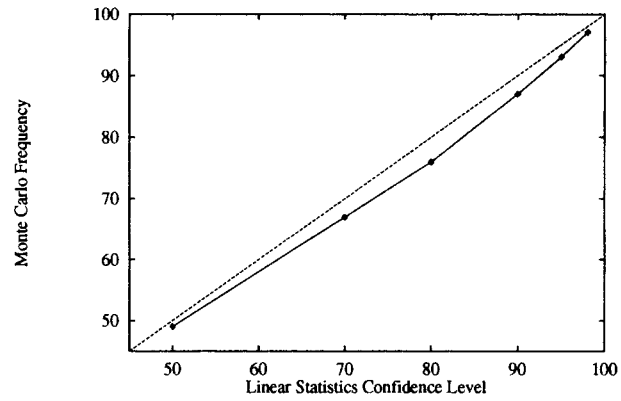

Figure 5. Comparison of linear statistics with Monte Carlo analysis of 500 replicate data sets.

Perfect agreement corresponds to the diagonal.

Table 4 gives the parameter estimation results obtained using data from all four runs. These parameters are used as the nominal parameters in subsequent discussion.

The addition of the solid-phase information provided by the transmittance measurement significantly reduces uncertainty in the parameter estimates for the KNO_3-H_2O system. The combination of data from experiments with various cooling profiles further reduces the parameter confidence region. The improvement is illustrated by a comparison of Tables 1 and 2. The parameter estimates given in Table 1, 2, and 4 are consistent (with only a few exceptions, the confidence intervals for each parameter overlap), and the confidence regions are consistently reduced as more information is added.

Further Model Validation

As discussed by Denn (1986), the objectives of the problem must be considered when determining the quality of a model. The model is of sufficient quality if it allows prediction of the variables of interest with desired accuracy. It has been established that, even over a wide range of operating conditions, the model identified for the KNO_3-H_2O system provides accurate and reproducible predictions of concentration and transmittance.

Sieve analysis yields qualitative information about a sample of crystals and, while subject to several sources of error, does provide measurements of the weight mean size, L_{wm}, and the coefficient of variation, $c.v.$, of the CSD. For further model validation, sieve analysis was performed on the products of several crystallization runs and the results compared with model predictions. A comparison of sieve analysis measurements and model predictions of L_{wm} and $c.v.$ for several experiments is

Table 3. Parameter Estimates Resulting from Replicate Runs of those Used to Obtain Table 2

	$\ln(k_g)$	g	$\ln(k_b)$	b
Estimate	8.817	1.32	16.766	1.69
Interval	±0.147	±0.04	±0.602	±0.15

Table 4. Parameter Estimates Resulting from Concentration and Transmittance Measurements from Four Batch Runs and Taken as Nominal Parameters of KNO_3-H_2O System

	$\ln(k_g)$	g	$\ln(k_b)$	b
Estimate	8.849	1.32	17.142	1.78
Interval	±0.112	±0.03	±0.360	±0.09

given subsequently, but Figures 6 and 7 show the results from two sample runs.

Figure 6 gives the model prediction and the sieve analysis results for the product of a crystallization experiment. There is reasonable agreement between the experimental and model-predicted maximum crystal size and location of the primary mode of the distribution. The weight mean size and coefficient of variation also agree fairly well—experiment: $L_{wm}=460$ μm, $c.v.=0.36$; model prediction: $L_{wm}=450$ μm, $c.v.=0.47$. Nevertheless, the experimental sieve analysis does show an unexpectedly low presence of crystals in the smaller size classes and an unexpectedly high peak at around 550 μm. This could be explained by agglomeration of small crystals during the drying process.

There is some discrepancy between the sieve analysis results and the model prediction, but they do demonstrate the same qualitative characteristics. It should also be noted that, while there are several sources for error in the sieve analysis, the analysis results are reproducible. The sieve analysis results for two replicate runs are given in Figure 7. The measured weight mean sizes for the two runs agree to within 16 μm, the coefficients of variation vary by only 0.03.

NLP Statement of Open-loop Optimal Control Problem

As discussed by Rawlings et al. (1993), there have been some studies of the determination of the temperature profile for a batch cooling crystallizer that optimizes some objective function that is based on the CSD. Unfortunately, the methods posed in these studies are restricted to very specific crystallizer configurations (for example, crystallizers with fines dissolution cannot be considered), and their applicability is limited by the difficulty of incorporating constraints. The method discussed in this section removes these limitations.

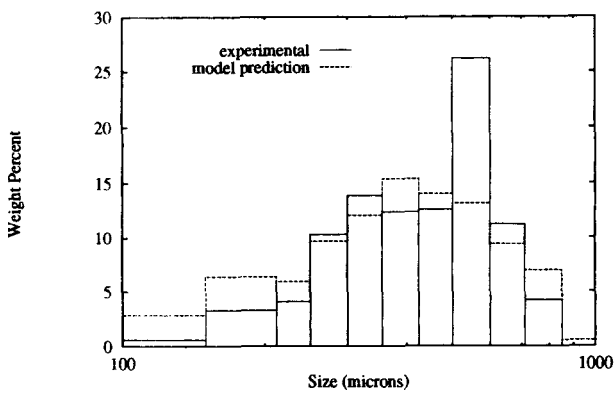

Figure 6. Comparison of sieve analysis results and the model prediction for a crystallizer product.

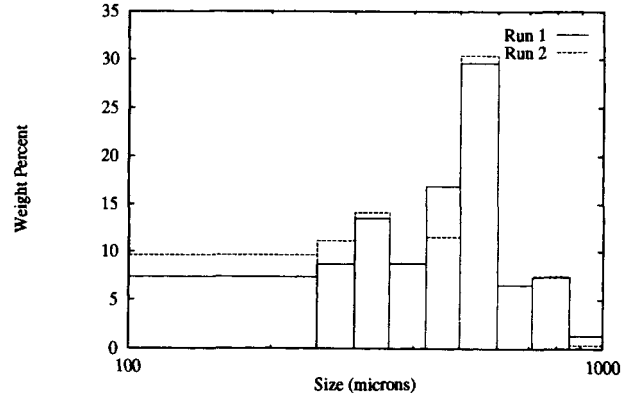

Figure 7. Sieve analysis results for two replicate experiments.

The general, multivariable, open-loop optimal control problem can be stated as a nonlinear programming (NLP) problem, for which there are standard algorithms (Rawlings et al., 1993). Consider the case in which the crystallizer temperature is the manipulated variable. If the temperature is parameterized as piecewise linear, and the constraints consist of nonlinear state constraints and linear constraints on the temperature, then the optimal control problem can be restated as the following NLP problem:

$$\begin{aligned} &\min_{\boldsymbol{T}} && \Phi(\boldsymbol{T}, \boldsymbol{x}(t_f); \theta) \\ &\text{subject to:} && \text{Crystallizer Model} \\ & && \boldsymbol{b}_l \le \boldsymbol{A}\boldsymbol{T} \le \boldsymbol{b}_u \\ & && \boldsymbol{h}(\boldsymbol{x}(t_i)) = \boldsymbol{0} \quad i = 1, \ldots, m \\ & && \boldsymbol{g}(\boldsymbol{x}(t_i)) \ge \boldsymbol{0} \quad i = 1, \ldots, m \end{aligned} \qquad \text{(NLP1)}$$

where t_i is one of m points in time over the run, $\boldsymbol{A}$ is a constant matrix with n columns, and $\boldsymbol{T}$ is a vector of temperature values at n points in time between $t=0$ and $t=t_f$.

For the control problems discussed subsequently, the successive quadratic programming code NPSOL (Gill et al., 1986) was used to solve the NLP problem.

Final-time constraints

There are cases in which constraints on final-time states of the system are necessary. In the absence of final-time constraints, there are many objective functions for which the solution of the optimal control is trivial but undesirable. For example, to minimize the number of nucleated crystals (fines) in the product, optimal operation would simply be a constant temperature profile at the saturation temperature—no formation of new crystals. The objective function would be at a global minimum, but the yield would be zero.

Final time constraints allow the consideration of variables such as crystallizer yield. For a batch crystallizer, there is a one-to-one relationship between final solute concentration and yield. Therefore, if a desired yield is determined to correspond to a final concentration of, say C_y, then a minimum yield is assured by imposing the final-time constraint:

$$C_y - C(t_f) \ge 0$$

This constraint clearly fits into the framework of Eq. NLP1.

Input constraints

Due to the physical limitations of equipment, there are usually bounds on manipulated variables and constraints on the rate at which they can be changed. NLP1 allows the incorporation of such constraints.

Because the temperature profile is parameterized as piecewise linear, a rate-of-change constraint and bounds on the temperature can be expressed as a set of linear constraints:

$$\begin{bmatrix} T_{\min} \\ -R_{\max}\Delta t_i \\ \vdots \\ -R_{\max}\Delta t_{n-2} \\ -R_{\max}\Delta t_{n-1} \end{bmatrix} \leq \begin{bmatrix} I \\ -1 \quad 1 \quad 0 \quad \ldots \quad 0 \\ 0 \quad -1 \quad 1 \quad \quad \vdots \\ \vdots \quad \quad \ddots \quad \quad 0 \\ 0 \quad \ldots \quad \ldots \quad -1 \quad 1 \end{bmatrix} T \leq \begin{bmatrix} T_{\max} \\ R_{\max}\Delta t_i \\ \vdots \\ R_{\max}\Delta t_{n-2} \\ R_{\max}\Delta t_{n-1} \end{bmatrix}$$

where I is an $n \times n$ identify matrix, $R_{\max}$ is the maximum allowable cooling rate, $T_{\min} = [T_{\min}, \ldots, T_{\min}]^T$, $T_{\max} = [T_{\max}, \ldots, T_{\max}]^T$, $T = [T(t_1), \ldots, T(t_n)]^T$, and $\Delta t_i = t_{i+1} - t_i$. These constraints can be enforced in NLP1 by defining $\mathbf{A}$, $\mathbf{b}_u$, and $\mathbf{b}_l$ accordingly.

An illustration of the implementation of a cooling rate constraint is given below.

State constraints

There may be cases in which there are path constraints on the states (that is, constraints on the states of the crystallizer over the course of a batch run). For example, to attempt to avoid homogeneous nucleation, an upper bound on the supersaturation over the run may be desired. But the methods by which to deal with state path constraints are not as transparent as for final-time and input constraints.

There have been several methods suggested for handling path constraints. A path constraint can be converted to a set of nonlinear constraints by time discretization of the states of interest; this is only an approximation and it can lead to a very large number of nonlinear constraints. Bryson and Ho (1975) suggest augmenting the objective function with the integral of the path constraint violations; this, however, requires the choice of a suitable weighting parameter.

Sargent and Sullivan (1978) suggest converting the path constraints into a final-time constraint by defining a new state variable that is equal to the integral of the squares of the path constraint violations. Consider the case in which there are p path constraints of the form:

$$g_i(\boldsymbol{x}(t)) \geq 0$$

If a new state variable is defined as:

$$\dot{x}_{q+1} = \sum_{i=1}^{p} [\max(0, -g_i(\boldsymbol{x}(t)))]^2 \qquad (14)$$

$$x_{q+1}(t_0) = 0$$

where q is the number of system states, the path constraints are satisfied if and only if the terminal constraint $x_{q+1}(t_f) = 0$ is satisfied.

This method avoids the computational difficulties of the time discretization technique and the arbitrary nature of the penalty function approach. An example of the implementation of this method to impose a state path constraint is given below.

Control scheme example

Downstream processes such as thickening and filtration are often time consuming steps—sometimes the production-limiting steps. The formulation of the control problem given by NLP1 allows consideration of the efficiency of downstream processes.

Compared to unimodal distributions, there are indications that the precipitate permeability and filtration efficiency are lower for bimodal CSDs—an inherent property of products of seeded crystallizers. Thus, the minimization of the ratio of the final mass of nucleation-formed crystals, $m_n(t_f)$, to the final mass of seed crystals, $m_s(t_f)$, would be expected to favorably affect the rate of filtration. This optimal control problem corresponds to the solution of NLP1 with: A equal to the identity matrix, $b_l = [T_{\min} \ldots T_{\min}]^T$, $b_u = [T_{\max} \ldots T_{\max}]^T$, and the following objective function and constraint:

$$\Phi = \left.\frac{m_n}{m_s}\right|_{t_f}$$

$$C_y - C(t_f) \geq 0$$

where the constraint represents the requirement of a minimum acceptable yield. For this problem C_y was selected to be the final solute concentration resulting from cooling at a constant rate to $T = T_{\min}$.

For the kinetic parameters of the KNO_3-H_2O system given in Table 4, the temperature profile that is predicted to minimize $m_n(t_f)/m_s(t_f)$ for the KNO_3-H_2O system is given in Figure 8; the corresponding supersaturation profile is also given. The results for natural and linear cooling profiles are given for comparison. The optimal cooling policy requires a very rapid temperature decrease at the end of the batch run. Because of physical constraints or to avoid fouling of the cooling surface, this large cooling rate might not be feasible. As explained above, the control algorithm can handle a cooling rate constraint. Figure 9 shows the results of limiting the cooling rate to 0.1°C/min. To meet the yield constraint, the temperature must start its descent earlier for the case in which the cooling rate is constrained, resulting in an earlier and less pronounced supersaturation peak.

There have been attempts to characterize the metastable region of crystallization, the region between saturation and spontaneous homogeneous nucleation. The saturation concentration can usually be represented as a simple, often linear, function of temperature. The metastable limit is the concentration above which there is spontaneous homogeneous nucleation. While it is usually dependent on the cooling rate, it is often characterized as a curve that is approximately parallel to the saturation concentration. With this type of characterization of the metastable region, spontaneous nucleation would be avoided by imposing a simple upper bound on $C - C_{sat}$:

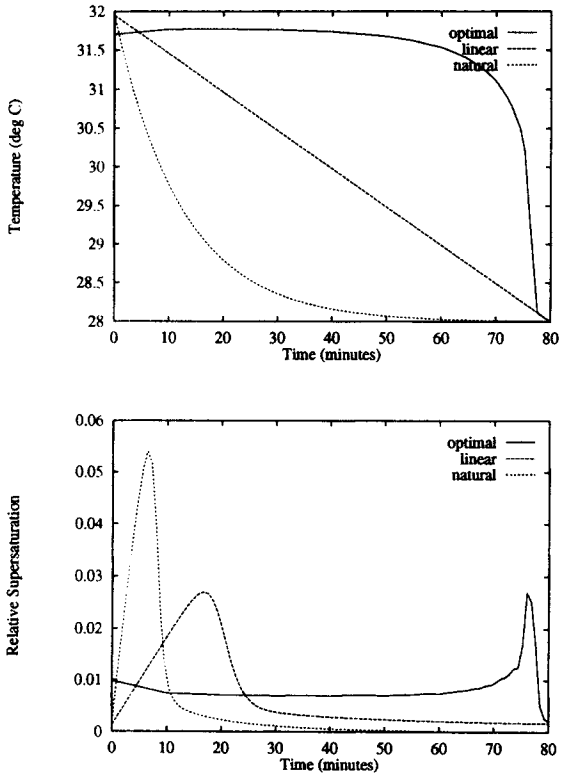

Figure 8. Temperature and supersaturation profiles for batch crystallizer.

Cooling policies illustrated include natural cooling, constant-rate or linear cooling, and cooling controlled to minimize final-time value of m_n/m_s.

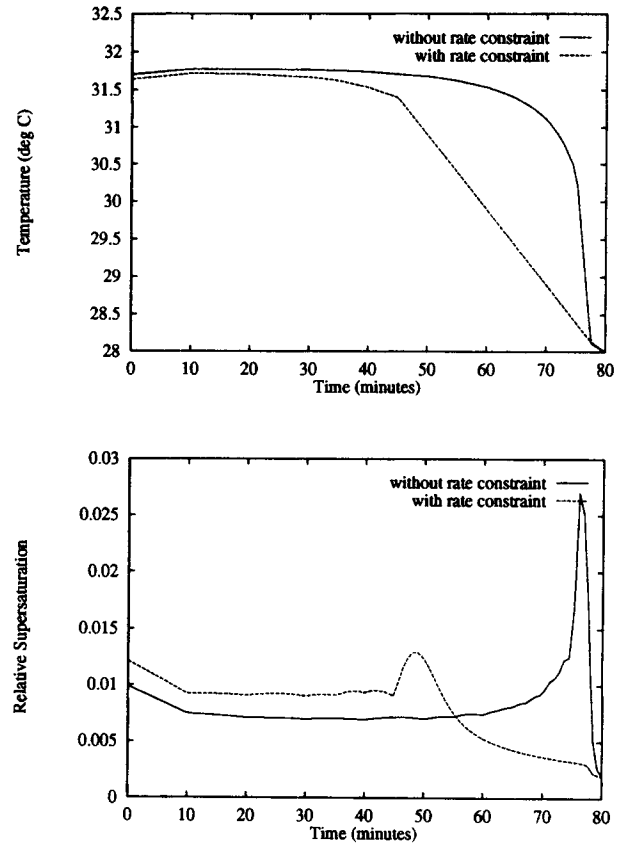

Figure 9. Demonstration of addition of cooling rate constraint of 0.1°C/min.

Profiles are those for minimization of the final-time value of m_n/m_s.

$$g(\boldsymbol{x}(t)) = (\text{metastable bound}) - (C - C_{sat}) \geq 0$$

As discussed above, such a constraint on the path of a state of the crystallizer can be incorporated into the solution of the optimal control problem. To do so, $g(\boldsymbol{x}(t))$ can be substituted into Eq. 14 to define an additional state variable that is constrained to a final-time value of zero. The result of placing a bound on the path of $C - C_{sat}$ is demonstrated in Figure 10.

The predicted effects of implementing temperature profiles to minimize the final-time value m_n/m_s, with and without constraints, are summarized in Table 5. Experimentally, the product of a batch crystallizer can be analyzed by sieve analysis. It is usually impossible to distinguish the difference between nucleated and seed crystals with sieve analysis, so this technique cannot provide a measurement of m_n/m_s. The weight mean size, L_{wm}, which can be assessed using sieve analysis, is included in Table 5 to provide a reference for the experimental study discussed in the section given below on experimental implementation. Notice that the optimal temperature profiles have a positive effect on the weight mean size.

Effectiveness of Method

In batch crystallization, nucleation and growth compete in the depletion of the solute concentration. Therefore, as Jones (1974) suggests, controlling the CSD of the crystallizer product may be regarded as optimally distributing a resource among a population.

Maximizing terminal seed size and minimizing nucleated crystal mass are similar problems in that the ideal for both is exclusive allocation of the solute to the seed crystal with no new crystal formation. While this ideal is not attainable, the difference in the order of the dependence of growth and of nucleation on supersaturation does allow one phenomenon to be favored over the other.

For the KNO_3-H_2O system,

$$G = \exp(8.849)(S)^{1.32} \quad B^o = \exp(17.142)(S)^{1.78}\hat{\mu}_3$$

The order of the dependence of nucleation rate on S is greater than that of growth rate. Therefore, the magnitude of both rates decrease as supersaturation is decreased, but the ratio of growth rate to nucleation rate increases. This explains the supersaturation profiles for the optimal control example given previously (see Figure 8). Optimal operation corresponds to keeping the supersaturation very small over most of the run. Only at the end of the runs is there a peak of supersaturation, leading to rapid growth of existing crystals and bursts of new crystals; the timing is such that deposition of solute to new crystals is minimized because of the limited opportunity for their growth.

As shown in Table 5, the benefit of implementing an optimal cooling profile for the KNO_3-H_2O system should be substantial.

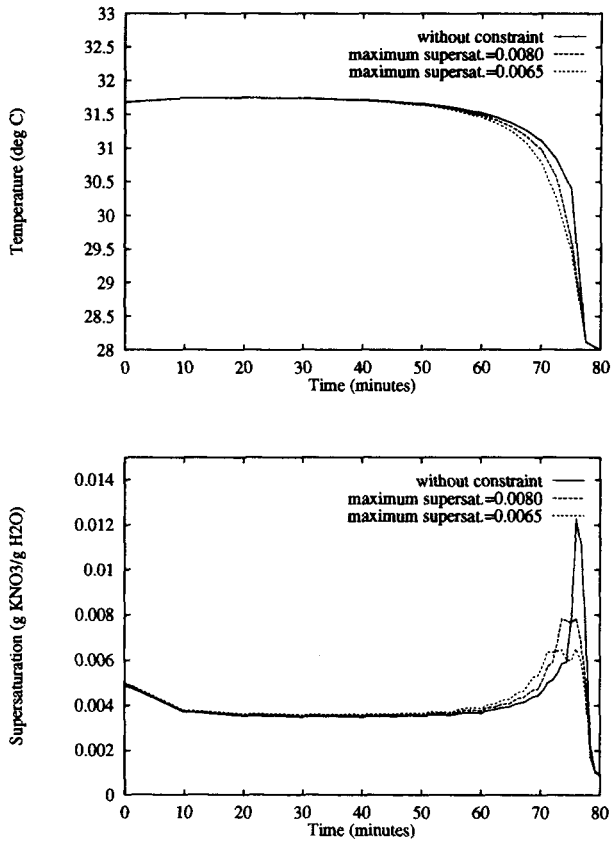

Figure 10. Implementation of a constraint on the maximum allowable value of C-C_{sat}.

Profiles are those for minimization of the final-time value of m_n/m_s.

The benefit is verified by the experimental implementation discussed subsequently. The attainable benefit could be more or less significant, depending on the difference in the dependence of the growth and nucleation rates on supersaturation. To illustrate, consider a hypothetical chemical system with the following kinetic expressions:

$$G = \exp(8.849)(S)^{1.32} \quad B^o = \exp(17.142)(S)^{2.78}\hat{\mu}_3$$

This is a system with the same growth kinetics as the KNO_3-H_2O system, but the order of dependence of nucleation rate on supersaturation is 2.78 instead of 1.78. As can be seen from Table 5, optimal cooling for the KNO_3-H_2O system leads to a predicted 33% decrease in the terminal value of m_n/m_s compared to that of natural cooling and a 20% decrease compared to that of linear cooling. As shown in Table 6, the relative benefit of optimal cooling for the hypothetical system is approximately twice as great.

Table 5. Summary of Predicted Effects of Minimizing $m_n(t_f)/m_s(t_f)$

Cooling Policy	m_n/m_s	Max. Seed Size (μm)	L_{wm} (μm)
Natural	26.5	895	439
Linear	22.0	959	487
Optimal	17.5	1,042	524
w/Cooling rate constraint	18.5	1,023	521
w/0.0065 Bound on $C - C_{sat}$	17.6	1,042	527

Table 6. Summary of Predicted Effects of Minimizing $m_n(t_f)/m_s(t_f)$ for Hypothetical Chemical System

Cooling Policy	m_n/m_s	L_{wm} (μm)
Natural	6.18	876
Linear	3.64	1,139
Optimal	1.98	1,458

The impact of model identification on the control problem can be illustrated by considering the result of using the poorly determined parameters given in Table 1—those resulting from concentration data alone. The boundary condition of the population balance equation (the nuclei density) is:

$$f(L_0, t) = \frac{k_b}{k_g} S^{(b-g)}\hat{\mu}_3$$

For the parameters based on concentration data alone, the order of dependence of the nuclei density on supersaturation is of the opposite sign of that corresponding to parameters from more informative data. The use of the parameters given in Table 1 would have led to a significantly different (and erroneous) optimal cooling profile; for the cooling profiles shown in Figure 8, these parameters would suggest that the natural cooling profile would have been superior to the linear and "optimal" cooling profiles.

Sensitivity of Optimal Profile to Model Parameters

Model-based control policies are, of course, determined using a nominal set of parameters. While these parameters are not known with total certainty, the parameter estimation scheme provides an assessment of the uncertainty. An analysis of the sensitivity of control policies to parameter uncertainty and the impact of the uncertainty on the performance of the control strategy constitute a link between model identification and model-based control.

The optimal control problem can be stated as:

$$u = \mathfrak{N}(\theta)$$

where $\mathfrak{N}$ is a nonlinear operator, and u is a vector representing the optimal profile of the manipulated variable for a given set of parameters, θ.

Let δu be the difference between the optimal profile determined for parameters θ and for that determined for θ^*, the nominal set of parameters:

$$\delta u = \mathfrak{N}(\theta) - \mathfrak{N}(\theta^*)$$

One way to examine the relationship between the optimal control policy and the parameter uncertainty is to determine the θ in the confidence region that "maximizes" $\delta u = \mathfrak{N}(\theta) - \mathfrak{N}(\theta^*)$. In other words, one approach is to find the worst-case deviation of the optimal control policy from the nominal profile for a given parameter confidence region. Maximizing δu requires the definition of a norm of δu. While there are several reasonable choices, the 2 norm:

$$\|\delta u\|^2 = \delta u^T \delta u$$

provides a measure of the overall deviation of an optimal profile from the nominal optimal profile.

The vector that is amplified most (in the 2-norm sense) by the linear matrix operator L is the first column of the right singular matrix of L. The singular value decomposition can be performed efficiently by routines in LINPACK (Dongarra et al., 1979).

Linearizing $\mathfrak{N}$ about the nominal parameter set allows an approximate statement of the optimal control problem:

$$u \approx L\theta$$

$$L_{ij} = \left.\frac{\partial u_i}{\partial \theta_j}\right|_{\theta^*}$$

According to this linear approximation of the control problem,

$$\delta u = L(\theta - \theta^*) = L\delta\theta$$

and the $\delta\theta$ that maximizes $\delta u^T \delta u$ is a multiple of $\delta\theta_{\text{worst}}$, the first column of the right singular matrix of L. In other words, the vector $\delta\theta_{\text{worst}}$ is the direction in which θ^* can be perturbed that should lead to the largest change in u.

For the KNO_3-H_2O system:

$$\theta^* = [8.849 \quad 1.32 \quad 17.142 \quad 1.78]^T$$

For the optimal control problem of minimizing the final-time value of m_n/m_s, the singular value decomposition of the finite difference approximation of L gives:

$$\delta\theta_{\text{worst}} = [-0.161 \quad 0.930 \quad -0.065 \quad 0.324]^T$$

Figure 11 gives the results of solving the optimal control problem after perturbing the nominal parameter by $\pm 0.1(\delta\theta_{\text{worst}})$. The analysis presented above suggests that the effect of perturbation of θ^* by an equal or lesser amount in any other direction should be less than that illustrated in Figure 11.

As discussed, the $100(1-\alpha)\%$ approximate parameter confidence region is a hyperellipsoid that is bounded by:

$$\delta\theta^T V_\theta^{-1} \delta\theta = \chi^2_{N_p}(\alpha)$$

where $\chi^2_{N_p}$ is a random variable of a chi-square distribution with N_p degrees of freedom. The parameter sets in the confidence region that differ most from θ^* are the endpoints of the major axis of the hyperellipsoid. For $\Phi = m_n(t_f)/m_s(t_f)$,

major axis endpoints

$$\equiv \pm\delta\theta_{\max} = [-0.029 \quad -0.006 \quad 0.349 \quad 0.090]^T$$

Perturbing θ^* by $\pm\delta\theta_{\max}$ leads to an imperceptible change in u. As illustrated by Figure 11, perturbation of the same magnitude ($\|\delta\theta\| = 0.36$) in the direction of $\delta\theta_{\text{worst}}$ leads to a significant change in the optimal u.

While perturbing θ^* by a multiple of $\delta\theta_{\text{worst}}$ can lead to a markedly different predicted optimal temperature profile, $0.36(\delta\theta_{\text{worst}})$ is well outside of the 95% confidence region. The task of finding the parameter perturbation *within the confidence region* that has the greatest effect on δu can be stated as:

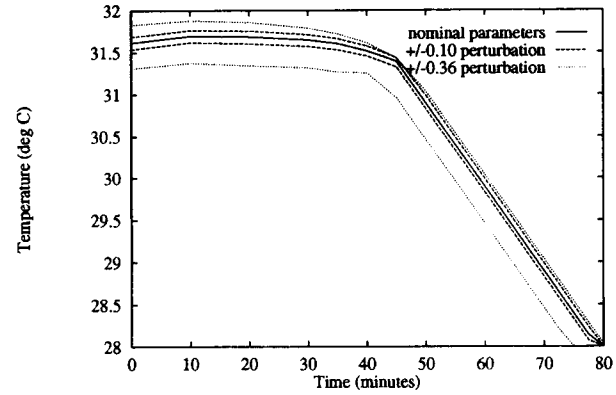

Figure 11. Sensitivity of control profile to parameter uncertainty.

Optimal control policy for nominal parameters and for parameter sets perturbed by ± 0.1 ($\delta\theta_{\text{worst}}$) and ± 0.36 ($\delta\theta_{\text{worst}}$).

$$\delta\theta = \theta^* \pm \gamma\delta\theta_{\text{worst}} \tag{15a}$$

$$\gamma = \left(\frac{\chi^2_{N_p}(\alpha)}{\delta\theta^T V_\theta^{-1} \delta\theta}\right)^{1/2} \tag{15b}$$

For the 95% confidence region of the KNO_3-H_2O system, $\gamma = 0.004$, which does not lead to a perceivable change in u from that determined for the nominal set of parameters. In other words, the sensitivity analysis suggests that for the tight confidence region of the KNO_3-H_2O kinetic parameters, the parameter uncertainty is not a concern in the determination of an optimal open-loop control policy.

Experimental Implementation of Open-loop Control

The bench-scale crystallizer with a KNO_3-H_2O chemical system was used to test the model-based control strategy. The example discussed above in which the objective was to minimize the terminal value of m_n/m_s was posed to possibly benefit efficiency of subsequent processing steps. This policy was chosen for implementation. As discussed, it is usually impossible to experimentally evaluate m_n/m_s, but the policy that minimizes m_n/m_s is expected to have measurable effects on variables that *can* be assessed.

The temperature measurements for typical runs of natural, linear, and optimal cooling policies are shown in Figure 12. To incorporate the physical limitations of the cooling system, the optimal control policy was subject to a cooling rate constraint of 0.1°C/min. As shown, the temperature control system was able to track the linear and optimal temperature trajectories.

Figures 13-15 present the concentration and transmittance data corresponding to the natural, linear, and optimal cooling profiles shown in Figure 12. A comparison of these figures demonstrates the relationship between the temperature profile

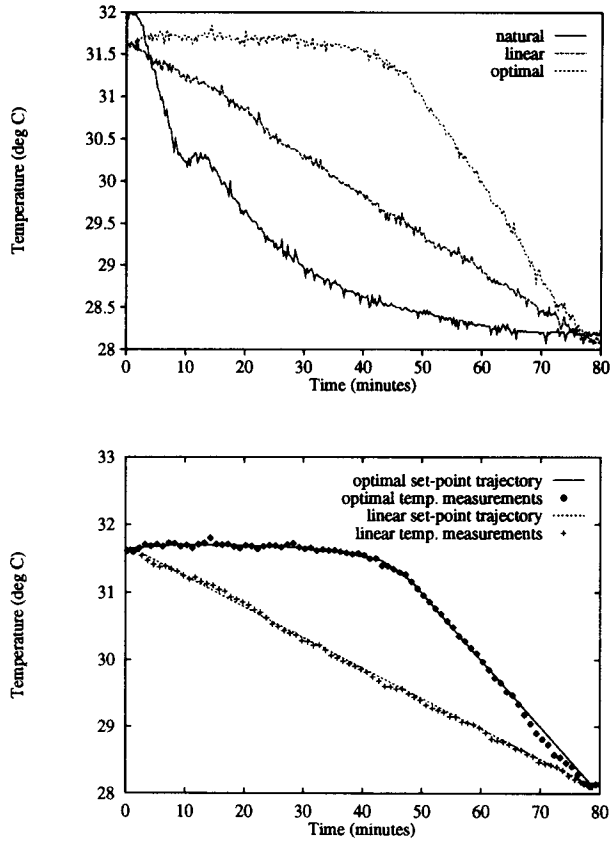

Figure 12. Experimental temperature data for natural, linear, and optimal cooling.

Setpoint trajectory tracking is also shown.

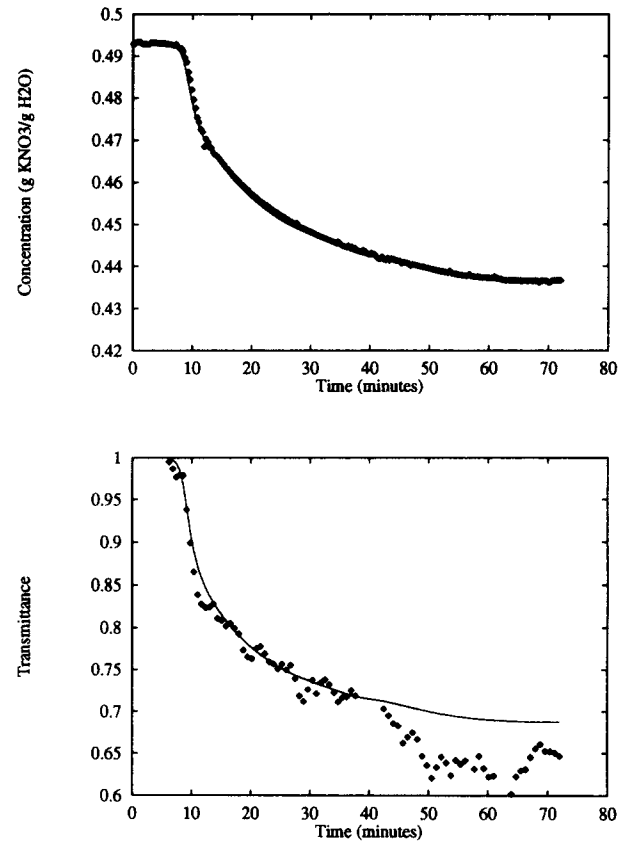

Figure 13. Concentration and transmittance measurements and model predictions for natural cooling operation.

and the system output variables. The figures also demonstrate the effectiveness of the model in predicting the behavior of the system. The model parameters used were those given in Table 4—those obtained from the experiments for model identification.

The effect of the temperature profiles on the product of the crystallizer is demonstrated by the weight mean sizes given in Table 7; these are results from sieve analysis for the runs corresponding to Figures 13-15. The model predicted mean sizes are also given.

The sieve analysis results of several runs are given in Table 8. The initial concentrations of the runs were all approximately 0.493 g KNO_3/g H_2O. With the exception of the case with higher seed loading (noted in the table), the only difference in the operating conditions for the runs was the cooling policy used. Because the final temperatures for the runs were not all equal, the yields were not consistent. To assess the effect of the cooling policy, runs with similar yields should be compared. (The crystallizer yields shown in the table were computed from the initial and final concentration measurements.)

There are several sources of error in sieve analysis. From observations made during these experiments, the most likely source of error was agglomeration of small crystals during drying, which would be consistent with the unexpectedly high values of L_{wm} and low values of the coefficient of variation.

There is error in the sieve analysis results, and there is undoubtedly some error in the model, so it is not surprising that there is some disparity between the measured and model-predicted values of L_{wm}. Nevertheless, the qualitative effect of the cooling policy is the same: the weight mean size can be increased by using a linear cooling profile instead of allowing natural cooling, and a further increase can be realized by using the optimal cooling policy.

As discussed, the increase in L_{wm} obtained by using cooling policies other than natural cooling is attained by avoiding the strong peak in supersaturation that is characteristic of natural cooling. The supersaturation data for natural, linear, and optimal cooling runs are given in Figure 16.

Table 8 includes the results from the implementation of an optimal control policy in which the seed loading was higher than that of the other runs—loading was higher by a factor of 10. The difference between the optimal temperature trajectories implemented for the lower and the higher seed loading was not very great. However, as shown by a comparison of Figures 15 and 17, the difference in the system behavior was significant. The high seed loading led to an immediately perceivable transmittance decrease. The seed load was high enough that the solute deposition on the seed crystals resulted in a decrease in the concentration early in the run, while there was a relatively long induction period for the case of low seed loading. It should be noted that the model is able to predict the system behavior for the high seed loading experiment, even though model identification was performed using data from low seed loading experiments.

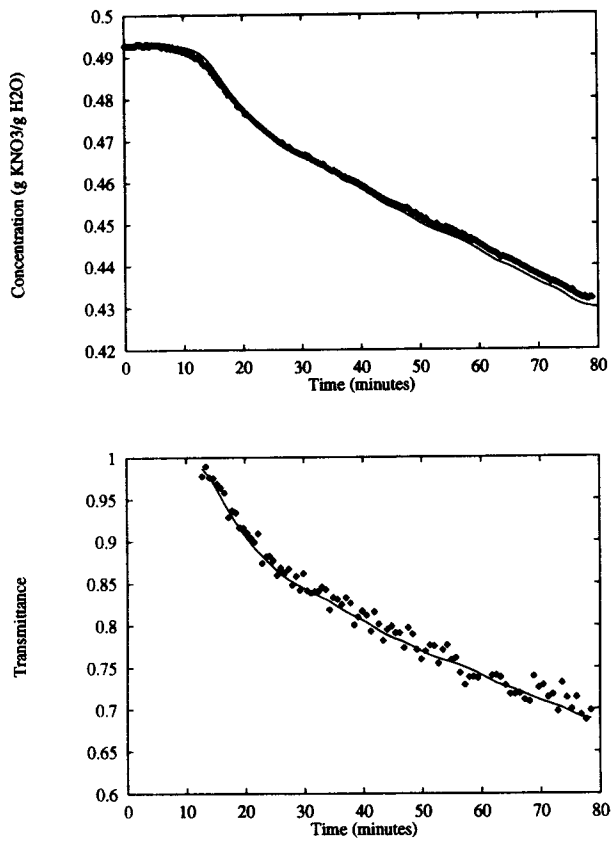

Figure 14. Concentration and transmittance measurements and model predictions for linear cooling operation.

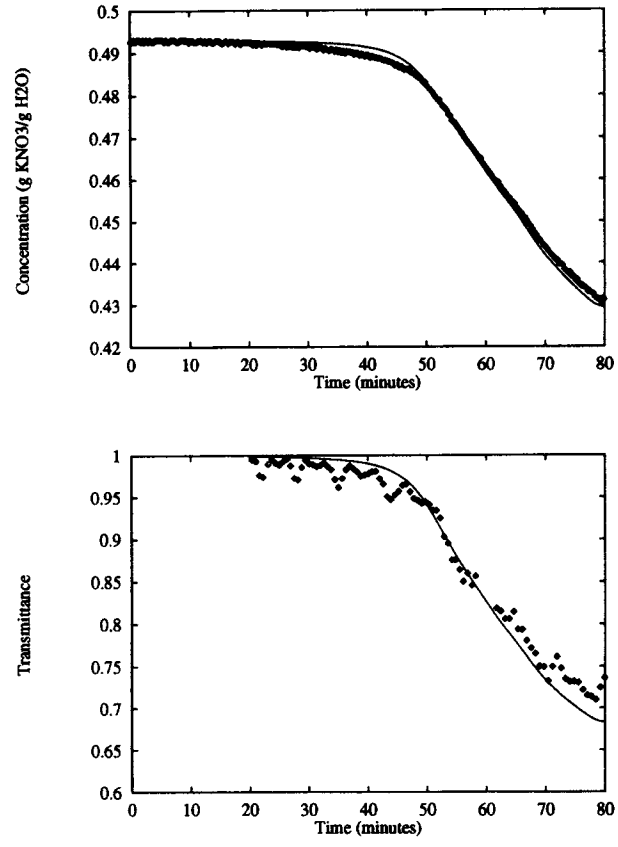

Figure 15. Concentration and transmittance measurements and model predictions for optimal cooling policy implementation.

Performance indices other than that used in the above discussion may be determined to be more beneficial, and the use of temperature as the lone manipulated variable may be deemed to be inadequate for a particular system. The technique is flexible enough to accommodate other performance indices and manipulated variable choices.

Conclusions

Supersaturation can be achieved by chemical reaction, temperature change, solvent evaporation, feed rate changes, and introduction of an additive. The size and CSD of a crystallizer product are governed by the rates of nucleation and growth, both of which are driven by supersaturation. Because nucleation and growth compete for a resource (the solute) and are driven by the same variable (supersaturation), the control problem is challenging and requires a careful characterization of the kinetics of the two phenomena.

Of the few studies that have focused on control of batch cooling crystallizers, most have dealt with determining a cooling profile that optimizes a performance index that is stated in terms of the moments of the CSD. As demonstrated by Bohlin and Rasmuson (1992) via simulation and experimentally by Chianese et al. (1984), even the qualitative benefits expected from optimal cooling profiles are not necessarily realized, at least not reproducibly. Poorly characterized kinetics and inaccurate process control (that is, poor implementation of profiles) are the probable causes of the shortcomings.

The results given in this article demonstrate that there are effective techniques for computing and implementing an optimal manipulated variable trajectory, the objective function for which allows consideration of downstream process efficiency. The interplay between model identification and the model-based control strategy were discussed, and an illustration was given of the difficulties that can arise if the system kinetics are poorly characterized. It has been shown that with careful experimental design, the kinetic parameters and their confidence regions can be estimated from measurements of solute concentration and light transmittance through a crystal slurry sample. The parameter confidence regions obtained in this study are elliptical regions based on linear statistics, but a Monte Carlo study was used to demonstrate that the elliptical regions provide a reasonable representation of the actual confidence regions for the KNO_3-H_2O test system.

At least for the KNO_3-H_2O system, the kinetic characterization is sufficiently reliable to assure the desired system be-

Table 7. Effects of Temperature Profiles on Weight Mean Size of Crystallizer Product

Cooling Policy	Measured L_{wm} (μm)	Model-Predicted L_{wm} (μm)
Natural	449	448
Linear	630	493
Optimal	666	514

Table 8. Weight Mean Size and Coefficient of Variation of Crystallizer Product of Several Runs

Cooling Policy	T_f (°C)	Yield (g)	Measured L_{wm} (μm)	Predicted L_{wm} (μm)	Measured *c.v.*	Predicted *c.v.*
	28.2	98	449	448	0.40	0.47
Natural	28.3	99	460	450	0.36	0.47
	28.4	93	465	448	0.37	0.47
	28.1	101	630	493	0.41	0.46
Linear	28.1	103	564	492	0.39	0.47
	28.5	92	532	492	0.36	0.47
	28.1	102	666	514	0.49	0.50
Optimal	28.4	91	507	481	0.44	0.50
	*27.8	107	574	470	0.26	0.22

*High seed loading used in run.

havior. In fact, the model was shown to accurately predict the behavior of the system, even in operating regions significantly different from the experiments used for model identification. Implementation of a model-based optimal cooling policy on the KNO_3-H_2O system led to increases in the weight mean size of up to 48% compared to that of natural cooling.

As shown from a simulation study of a hypothetical chemical system, the impact of the optimal temperature profile on the final CSD is strongly dependent on the growth and nucleation parameters. In other words, an optimal temperature profile could be more or less effective than was realized on the KNO_3-H_2O system, depending on the kinetics of the system. For chemical systems for which temperature manipulation is determined to be ineffective for affecting the CSD, the consideration of additional or alternative manipulated variables may be necessary. For example, Jones et al. (1984) have investigated the feasibility of using fines dissolution. The framework of the control strategy presented in this article is flexible with respect to the choice of manipulated variables.

Due to the accuracy with which the model characterizes the system and the ability to precisely follow a given manipulated variable trajectory, feedback offers no real advantage for the KNO_3-H_2O system used in this study. Nevertheless, for a particular system, the plant/model mismatch, the deviations from the nominal trajectory, or the system disturbances may be significant. Possible disturbances of an industrial process include improper seeding (for example, adding incorrect size or amount of seeds, or seeding at the incorrect system state) and batch-to-batch variations (usually due to impurities and solvent

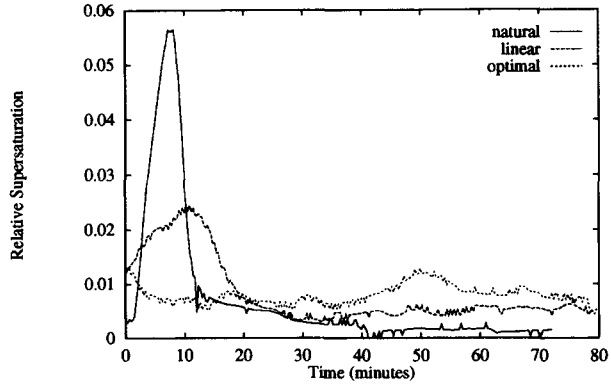

Figure 16. Experimental relative supersaturation data for natural, linear, and optimal cooling.

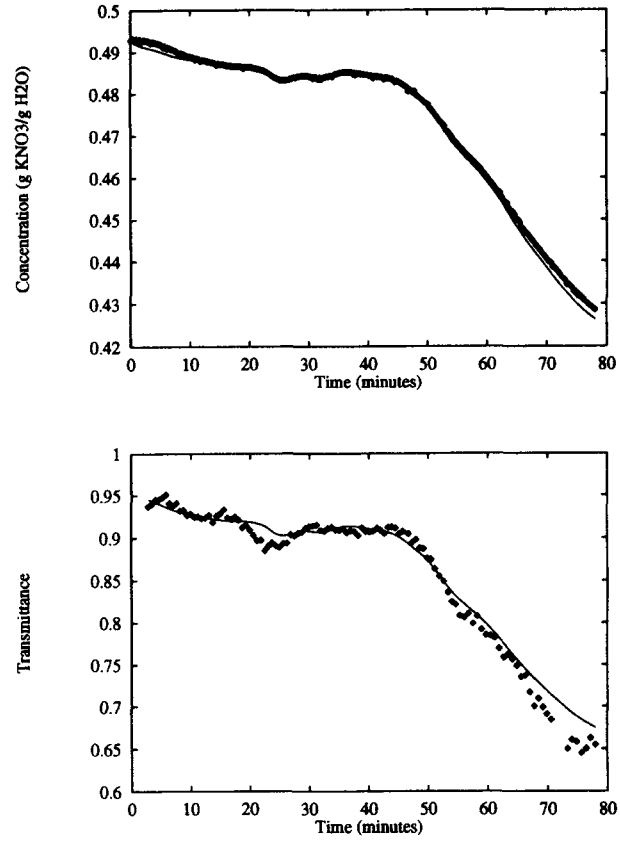

Figure 17. Concentration and transmittance measurements and model predictions for optimal cooling policy implementation with high seed loading.

effects). In other words, there are cases in which feedback compensation would be beneficial. For such cases, Rawlings et al. (1993) provide suggestions for incorporating feedback.

Acknowledgments

Financial support from Eastman Kodak Company and the National Science Foundation under grant CTS-8957123 is gratefully acknowledged.

Notation

A = linear constant matrix
A_c = heat-transfer area for cooling crystallizer
b_l = vector of lower bounds of linear constraints of NLP problem
b_u = vector of upper bounds of linear constraints of NLP problem
B = rate of crystal formation mechanisms
B_j = matrix for linearized model at jth sampling instant
B^o = rate of crystal nucleation at size L_0
c_p = heat capacity
C = solute concentration (mass solute/mass of solvent)
C_{sat} = solute saturation concentration (mass of solute/mass of solvent)
c.v. = coefficient of variation of weight-size distribution
e_{ij} = error between measurement and prediction, $y_{ij} - \bar{y}_{ij}$
$e_i(s_i)$ = energy of scattered light measured by detector of inner radius s_i
E_b = nucleation rate activation energy
E_g = growth rate activation energy
f = crystal distribution function
g = crystal growth rate order
g = vector of inequality constraints

G = crystal growth rate
h = conversion factor, volume of slurry per mass of solvent
$\boldsymbol{h}$ = vector of equality constraints
ΔH_c = heat of crystallization
j = slurry density power law coefficient for nucleation rate expression
k_a = crystal surface area shape factor (surface area equals $k_a L^2$)
k_b = nucleation rate coefficient
k_v = crystal volume shape factor
l = light path length
L = characteristic particle or crystal size
$\boldsymbol{L}$ = linear matrix operator for sensitivity analysis
L_{wm} = weight mean size
N_d = number of measurement samples during an experiment
N_m = number of quantities measured
N_p = number of parameters estimated
$R_{\max}$ = maximum allowable cooling rate
S = supersaturation, $(C-C_{sat})/C_{sat}$
t = time
t_f = final time of batch run
T = temperature
$\boldsymbol{T}$ = vector of temperature values at n points in time between $t=0$ and $t=t_f$
T_c = coolant temperature
$T_{\min}$ = lower bound on crystallizer temperature
$T_{\max}$ = upper bound on crystallizer temperature
U = heat-transfer coefficient
V = volume of the crystallizer's contents
$\boldsymbol{V}$ = measurement covariance matrix
$\boldsymbol{V}_\theta$ = parameter covariance matrix for the linearized problem
$\boldsymbol{x}$ = vector of system states
y_{ij} = ith measurement taken at the jth sampling instant
$\hat{y}_{ij}$ = model prediction for the ith measurement taken at the jth sampling instant

Greek letters

α = joint confidence limit
$\chi^2_{N_p}$ = chi-square statistic with N_p degrees of freedom
$\delta\boldsymbol{\theta}_{\max}$ = maximum possible deviation from $\boldsymbol{\theta}^*$ within confidence region—major axis endpoints
$\delta\boldsymbol{\theta}_{\text{worst}}$ = parameter deviation direction that has largest effect on optimal control policy
γ = scaling factor
$\boldsymbol{\theta}$ = vector of parameters
$\boldsymbol{\theta}^*$ = nominal parameters set
λ_i = ith eigenvalue
μ_k = kth moment of the CSD $= \int_0^\infty f L^k dL$
$\hat{\mu}_3$ = third moment of the CSD with units cm^3/g solvent $(h\int_0^\infty f L^3 dL)$
ρ = density of the crystallizer contents
ρ_c = crystal density
σ_i^2 = variance of ith measured variable
Φ = objective function

Literature Cited

Bard, Y., *Nonlinear Parameter Estimation*, Academic Press, New York (1974).

Bohlin, M., and Ä. C. Rasmuson, "Application of Controlled Cooling and Seeding in Batch Crystallization," *Can. J. Chem. Eng.*, **70**, 120 (1992).

Boxman, A., "Particle Size Measurement for Control of Industrial Crystallizers," PhD Thesis, Delft Univ. of Technology (1992).

Brook, R. J., and G. C. Arnold, *Applied Regression Analysis and Experimental Design*, Marcel Dekker, New York (1985).

Bryson, A. E., and Y. Ho, *Applied Optimal Control*, Hemisphere, New York (1975).

Chianese, A., S. D. Cave, and B. Mazzarotta, "Investigation on Some Operating Factors Influencing Batch Cooling Crystallization," *Industrial Crystallization 84*, S. J. Jančić, and E. J. de Jong, eds., Elsevier Science, p. 443 (1984).

Denn, M. M., *Process Modeling*, Longman, White Plains, NY (1986).

Dongarra, J. J., C. B. Moler, J. R. Bunch, and G. W. Stewart, *LINPACK User's Guide*, SIAM, Philadelphia (1979).

Gill, P. E., W. Murray, M. A. Saunders, and M. H. Wright, "User's Guide for SOL/NPSOL (Version 4.0): A Fortran Package for Nonlinear Programming," Technical report SOL 86-2, Systems Optimization Laboratory, Dept. of Operations Research, Stanford Univ. (1986).

Hulburt, H., and S. Katz, "Some Problems in Particle Technology," *Chem. Eng. Sci.*, **19**, 555 (1964).

Jager, J., S. de Wolf, W. Klapwijk, and E. J. de Jong, "A New Design for On-Line Product-Slurry Measurements," *Industrial Crystallization 87*, J. Nývlt and S. Žáček, eds., Elsevier Science Publishers p. 415 (1987).

Jones, A. G., "Optimal Operation of a Batch Cooling Crystallizer," *Chem. Eng. Sci.*, **29**, 1075 (1974).

Jones, A. G., A. Chianese, and J. W. Mullin, "Effect of Fines Destruction on Batch Cooling Crystallization of Potassium Sulphate Solutions," *Industrial Crystallization 84*, S. J. Jančić and E. J. de Jong, eds., Elsevier Science Publishers, p. 191 (1984).

Karpinski, P., J. Budz, and Z. Naruc, "Effect of Seeding Moment on Quality of CSD from Batch Cooling Crystallizer," *Science Papers of the Institute of Chemical Engineering and Heat Systems*, D. Belina, ed., Book No. 38, Symp. No. 5, Wroclaw, Poland, Technical Univ. of Wroclaw, Wroclaw, Poland, p. 172 (1980).

Miller, S. M., "Modelling and Quality Control Strategies for Batch Cooling Crystallizers," PhD Thesis, Univ. of Texas at Austin (1993).

Randolph, A. D., and M. A. Larson, "Transient and Steady State Size Distributions in Continuous Mixed Suspension Crystallizers," *AIChE J.*, **8**(5), 639 (1962).

Rawlings, J. B., S. M. Miller, and W. R. Witkowski, "Model Identification and Control of Solution Crystallization Processes: a Review," Accepted for publication in *I&EC Res.* (1993).

Rippin, D. W. T., "Statistical Methods for Experimental Planning in Chemical Engineering," *Comput. Chem. Eng.*, **12**, 109 (1988).

Sargent, R. W. H., and G. R. Sullivan, "The Development of an Efficient Optimal Control Package," in *Lecture Notes in Control and Information Sciences 7*, J. Stoer, ed., 8th IFIP Conference on Optimization Techniques, Berlin, Springer-Verlag, p. 158 (1978).

Seinfeld, J. H., "Identification of Parameters in Partial Differential Equations," *Chem. Eng. Sci.*, **24**, 65 (1969).

Witkowski, W. R., S. M. Miller, J. B. Rawlings, "Light Scattering Measurements to Estimate Kinetic Parameters of Crystallization," *Crystallization as a Separations Process*, A. S. Myerson and K. Toyokura, eds., American Chemical Society, Washington, DC (1990).

Manuscript received May 25, 1993, and revision received Oct. 13, 1993.