Dynamic Parameter Estimation and Optimization for Batch Distillation

Seyed Mostafa Safdarnejad[a], Jonathan R. Gallacher[a], John D. Hedengren[a,*]

[a]*Department of Chemical Engineering, 350 CB, Brigham Young University, Provo, UT 84602, USA*

Abstract

This work reviews a well-known methodology for batch distillation modeling, estimation, and optimization but adds a new case study with experimental validation. Use of nonlinear statistics and a sensitivity analysis provides valuable insight for model validation and optimization verification for batch columns. The application is a simple, batch column with a binary methanol-ethanol mixture. Dynamic parameter estimation with an ℓ_1-norm error, nonlinear confidence intervals, ranking of observable parameters, and efficient sensitivity analysis are used to refine the model and find the best parameter estimates for dynamic optimization implementation. The statistical and sensitivity analyses indicated there are only a subset of parameters that are observable. For the batch column, the optimized production rate increases by 14% while maintaining product purity requirements.

Keywords: Dynamic Parameter Estimation, Nonlinear Statistics, Experimental Validation, Batch Distillation, Dynamic Optimization

1. Introduction

There are approximately 40,000 distillation columns in the US that are used to separate chemical compounds based on vapor pressure differences in industries ranging from oil and gas to pharmaceuticals. These separation columns consume 6% of the yearly US energy demand [1]. While many of the large production facilities use continuous processes, specialty and smaller-use items are often processed in batch columns [2, 3, 4]. Continuous distillation columns have been the focus of optimization work since the first column was built, but the transient nature of batch columns has caused many to remain unoptimized. The transient nature of the market for these specialty items has further hindered the optimization of batch columns [3]. As a result, little research on batch column optimization is available in the literature before 1980 [5, 6, 7, 8]. Work on batch columns has increased in the last 30 years as computers have become more sophisticated, and several studies have considered both advanced solving techniques and advanced column configurations [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23]. Terwiesch, et al. [24] and Kim and Diwekar [25] provide a detailed history of the subject and a description of current batch distillation modeling and optimization methods.

*Corresponding author. Tel.: +1 801 477 7341, Fax: +1 801 422 0151

Email addresses: safdarnejad@byu.edu (Seyed Mostafa Safdarnejad), jgallac2@byu.edu (Jonathan R. Gallacher), john.hedengren@byu.edu (John D. Hedengren)

The optimization of the batch columns can be subdivided into optimal design problems and optimal control problems. Optimal design problems generally deal with column configuration, while optimal control problems deal with column operation. These ideas are summarized well in separations textbooks such as Diwekar [10], Stichlmair and Fair [26] and Doherty and Malone [27] and will therefore not be discussed further here. Research studies on this subject follow the same general outline as presented in the textbooks [3, 28]. The models developed for batch column optimization generally fall into two categories: first-principles models and shortcut or simple models.

First-principles models are those with governing mass and energy balance equations, detailed thermodynamics, tray dynamics, system non-idealities and variable flow rates [29, 30, 31, 32]. These models are theoretically more accurate than shortcut methods, but they are only as accurate as the thermodynamic and physical property models they use [3]. The use of these models has been limited due to high computational costs. Several studies have been conducted using first-principles models and advanced solving techniques to reduce computational cost [33, 34, 35, 36, 37, 38, 39]. While these models accomplish the goal of reducing computational load, they are generally still slower than shortcut models. In addition, the lack of experimental data for batch columns makes it difficult to determine how much accuracy is lost when going from first-principles to lower-order (first-principles model with advanced or simplified numerical methods) to shortcut models [40].

The second class of models, shortcut models, has received far greater attention. These models contain less physics and are generally used for ballpark estimates and comparative studies. A typical set of assumptions for these models is as follows: constant boil-up rate, no external heat loss, ideal stages, constant relative volatility, constant molar overflow, total condenser without subcooling and no column holdup [41, 42, 28, 43, 31]. More recent shortcut models have kept most of the same assumptions while accounting for column dynamics using a non-zero column holdup [40, 36]. The primary purpose of these models is to create an accurate, computationally fast simulation for use in design and control of batch columns. While these models achieve the reduction in computational load, the lack of experimental data makes it difficult to determine the accuracy of these models [40]. The assumptions made in these models limit their use to ideal systems.

The gap between first-principles models and shortcut models is large. First-principles models can provide predictions for many systems but require thermodynamic and physical property models as inputs, while the assumptions in shortcut models make them applicable only to a small class of relatively ideal systems. In this work, a method is proposed for developing shortcut models with relaxed assumptions. The method is based on fitting parameters in place of simplifying assumptions to include system non-idealities without solving the first-principles equations. Solving for the fitting parameters requires extensive experimental data whereas first-principles models typically need less data, being based on fundamental correlations. Dynamic parameter estimation can be used to reduce the experimental load. The case study presented in this work required only one experiment to determine model parameters. As with any model containing fitting parameters, there is concern over the accuracy of the parameters. By using nonlinear statistics [44] and a model sensitivity

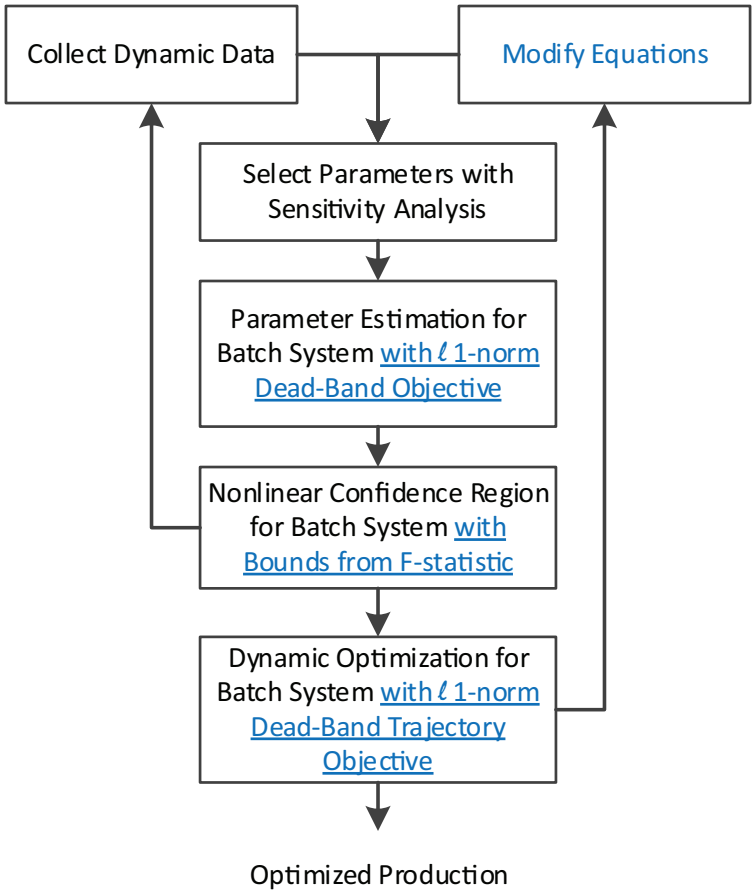


Figure 1: Overview of methodology for batch column optimization with novel contributions underlined

analysis [45], it is possible to determine how many parameters can be estimated from the collected data and the acceptable range for those parameters. These steps are shown in Figure 1 and form the heart of the method. Underlined elements of the methodology indicate the new approach to batch separation systems.

The well-known methodology shown in Figure 1 is applied to an experimental case study. The methodology includes the use of ℓ_1-norm dynamic parameter estimation, nonlinear statistics [44, 46], and a model parameter sensitivity analysis [45]. These techniques are applied together to a batch distillation column in a holistic approach to dynamic optimization. Models developed using this method account for system non-idealities not seen in typical shortcut models without sacrificing computational speed.

2. Model Development Framework

In this section, the general equations used to represent the process model, parameter estimation, nonlinear statistics and sensitivity analysis, or process optimization are reviewed.

2.1. General Process Model

In this work, a process model is developed to represent the batch relationship between adjustable process values and the production amount or product specification outcomes. The general model formulation used for batch system modeling is shown by Eq. (1).

$$0 = f(\frac{\delta x}{\delta t}, x, y, \theta, d, u) \tag{1a}$$

$$0 = g(x, y, \theta, d, u) \tag{1b}$$

$$0 \leq h(x, y, \theta, d, u) \tag{1c}$$

where x is a vector of state variables, y is a vector of measured or optimized states (outputs), p is a set of parameters, d is a time-varying trajectory of disturbance values, and u is a set of control moves. Residuals, output functions, and inequality constraints are represented by f, g, and h, respectively, with an objective function J. Continuous, binary, or integer variables can be used. Both algebraic and differential equations can be used in the general form of Eq. (1) [47, 48, 49, 50]. In the case of estimation, the objective function J is a minimization of model outputs y from measured values as $\min_{x,y,\theta}\ J\,(x, y, \theta)$. In the case of product optimization, the objective function is typically tied to maximizing economics or production targets while satisfying safety or operational constraints as $\min_{x,y,u}\ J\,(x, y, u)$. This general formulation does not imply that estimation and optimization are solved in a single problem but that both are derived from a common model and nonlinear programming formulation. These sets of equations, along with the equations described in Section 3.2, are implemented in the APMonitor Modeling Language [51, 52, 53].

2.2. Parameter Estimation

Using the ℓ_1-norm of Eq. 2 allows for dead-band (δ) noise rejection and the additional objective expressions only add linear equations to the problem.

$$\Psi = \min_{\theta,x,y} w_x^T(e_U + e_L) + w_p^T(c_U + c_L) + \Delta\theta^T c_{\Delta\theta} \tag{2a}$$

$$s.t. \qquad 0 = f(\frac{\delta x}{\delta t}, x, y, \theta, d, u) \tag{2b}$$

$$0 = g(x, y, \theta, d, u) \tag{2c}$$

$$0 \leq h(x, y, \theta, d, u) \tag{2d}$$

$$e_U \geq (y - z + \frac{\delta}{2}) \tag{2e}$$

$$e_L \geq (z - y - \frac{\delta}{2}) \tag{2f}$$

$$c_U \geq (y - \bar{y}) \tag{2g}$$

$$c_L \geq (\bar{y} - y) \tag{2h}$$

$$0 \leq e_U, e_L, c_U, c_L \tag{2i}$$

Table 1: Nomenclature for general form of the objective function with ℓ_1-norm formulation for dynamic data reconciliation

Symbol	Description
Ψ	minimized objective function result
y	model outputs $(y_0, \ldots, y_n)^T$
z	measurements $(z_0, \ldots, z_n)^T$
$\bar{y}$	prior model outputs $(\bar{y_0}, \ldots, \bar{y_n})^T$
w_x^T	measurement deviation penalty
w_p^T	penalty from the prior solution
$c_{\Delta\theta}$	penalty from the prior parameter values
δ	dead-band for noise rejection
x, u, θ, d	states (x), inputs (u), parameters (θ), or unmeasured disturbances (d)
$\Delta\theta^T$	change in parameters
f, g, h	equations residuals (f), output function (g), and inequality constraints (h)
e_U, e_L	slack variable above and below the measurement dead-band
c_U, c_L	slack variable above and below a previous model value

The nomenclature for Eq. 2 is found in Table 1.

Many approaches can be used to find the parameters, two of which are least squares formulation and ℓ_1-norm formulation for the objective function. The least squares objective is more sensitive to bad data such as outliers as shown later in Figure 5. The ℓ_1-norm method is less sensitive to outliers and the form of the objective function used in this ℓ_1-norm formulation is smooth and continuously differentiable as opposed to using the absolute value function. A more thorough comparison of the ℓ_1-norm and least squares is provided in [54].

2.3. Confidence Intervals and Sensitivity Analysis

Reliability of the parameters is investigated by implementing an approximate nonlinear confidence interval calculation [44]. Non-linear confidence intervals can be found by solving Eq. 3 for the sets of parameters that make up the joint confidence region [55], then extracting the upper and lower bounds of that region in each dimension.

$$\frac{J(\theta) - J(\theta^*)}{J(\theta^*)} \leq \frac{p}{n-p} F_{n,n-p,1-\alpha} \tag{3}$$

In Eq. 3, $J(\theta)$ is the error between the measurements and the model prediction at a value θ of the parameters, $J(\theta^*)$ is the error between the measurements and the model prediction at the best estimates

of the parameters (θ^*), p is the number of parameters in the model, n is the number of data points, and $F_{n,n-p,1-\alpha}$ is the F-statistic at n and $n-p$ degrees of freedom with a confidence level of $1-\alpha$. The squared error objective is the only form of the nonlinear confidence interval that has a theoretical foundation. This is because the F-statistic used to define the confidence region is a ratio of χ^2 distributions that compares the equivalence of two sets of experimental results. The χ^2 distributions are intended for least square objectives instead of ℓ_1-norm objectives. According to the authors' knowledge, an equivalent F-statistic for nonlinear confidence intervals and the ℓ_1-norm has not been derived. A nonlinear confidence interval for ℓ_1-norm objectives based on the F-statistic is future work.

It is also desirable to determine the number of parameters that can be estimated or are observable given a particular model form and set of data. Large confidence intervals signal that a particular parameter may not be observable or that the effect of that parameter may be co-linearly dependent with other parameters. A well-known systematic analysis is used to determine which parameters can be estimated and rank the parameters in terms of the ability of a particular parameter to improve a particular model estimate [45, 56]. This procedure is accomplished in 3 steps: (1) efficient computation of the sensitivities, (2) scaling of the dynamic parameter sensitivities, and (3) singular value decomposition of the scaled sensitivity matrix to reveal an optimal parameter space transformation.

The first step in performing the parameter analysis is to compute the state dependencies to changes in the parameters. This can be accomplished with a variety of methods. One such method is to compute a finite difference sensitivity of the parameters with a series of perturbed simulations [57, 58]. A second method is to augment the model with adjoint equations that compute sensitivities simultaneously with the model predictions [59]. A third method is a post-processing method with time-discretized solutions to differential equation models [60, 61, 62]. This post-processing method involves efficient solutions to a linear system of equations, especially over other methods for large-scale and sparse systems [63].

The sensitivity is computed from time-discretized models that are solved by nonlinear programming solvers. At the solution, exact first derivatives of the equations with respect to variables are available through automatic differentiation. These derivatives are available with respect to the states ($\nabla f_x(x,\theta)$) and parameters ($\nabla f_p(x,\theta)$). For the objective function, objective gradients are computed with respect to states ($\nabla J_x(x,\theta)$) and parameters ($\nabla J_\theta(x,\theta)$). Sparsity in those matrices is exploited to improve computational performance, especially for large-scale systems. Sensitivities are computed by solving a set of linear equations as shown in Eq. 4 with parameter values fixed at $\bar{\theta}$ and variable solution $\bar{x}$ as nominal values.

$$\begin{bmatrix} \nabla_x f(\bar{x},\bar{\theta}) & \nabla_\theta f(\bar{x},\bar{\theta}) & 0 \\ \nabla_x J(x,\theta) & \nabla_\theta J(x,\theta) & -1 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} \nabla_\theta x \\ \nabla_\theta \theta \\ \nabla_\theta J(\bar{\theta}) \end{bmatrix} = \begin{bmatrix} 0 \\ \Delta\theta_i = 1 \\ 0 \end{bmatrix} \tag{4}$$

To further improve the efficiency of this implementation, an LU factorization of the left hand side (LHS) mass matrix is computed. This LU factorization is preserved for successive solutions of the different right hand side (RHS) vectors because the LHS does not change and successive sparse back-solves are computa-

tionally efficient in comparison with the LU factorization. Each matrix inversion computes the sensitivity of the states to a particular parameter. Each parameter is successively set equal to a change of $\Delta\theta_i = 1$. All other elements of the vector on the RHS are set to 0. The solution to this matrix inversion computes the sensitivity of all variables in the time horizon with respect to a particular parameter $S = (\nabla_\theta x)$. It also computes the sensitivity of the objective function with respect to the parameters $\left(\nabla_\theta J(\bar{\theta})\right)$.

To summarize the sensitivity analysis, an efficient method is presented to compute sensitivities as a post processing step that is efficient even for large-scale and sparse systems. The sensitivity matrix is decomposed into singular values and eigenvectors that give the relative magnitude and linear combination of parameters that are orthogonal. In this study, the transformed parameters are not estimated directly but instead used as an advisory tool to determine which parameters and how many can be estimated.

2.4. Control Optimization and Implementation

Similar to the parameter estimation developed in section 2.2, many approaches could be used in control and optimization of the dynamic systems. The form of the objective function used in this work is related to a nonlinear dynamic optimization with ℓ_1-norm formulation. In comparison to the common squared error norm, ℓ_1-norm is advantageous as it allows for a dead-band and permits explicit prioritization of control objectives. The form of the objective function with ℓ_1-norm formulation is shown in Eq. 5 [54, 64]. The nomenclature for Eq. 5 is found in Table 2.

$$\Psi = \min_{u,x,y} w_h^T\ e_h + w_l^T\ e_l + y_m^T\ c_y + u^T\ c_u + \Delta u^T\ c_{\Delta u} \tag{5a}$$

$$s.t. \qquad 0 = f(\dot{x}, x, u, d) \tag{5b}$$

$$0 = g(y, x, u, d) \tag{5c}$$

$$0 \leq h(x, u, d) \tag{5d}$$

$$\tau_c \frac{\delta z_{t,h}}{\delta t} + z_{t,h} = SP_h \tag{5e}$$

$$\tau_c \frac{\delta z_{t,l}}{\delta t} + z_{t,l} = SP_l \tag{5f}$$

$$e_h \geq (y - z_{t,h}) \tag{5g}$$

$$e_l \geq (z_{t,l} - y) \tag{5h}$$

Table 2: Nomenclature for general form of the objective function with ℓ_1-norm formulation

Symbol	Description
Ψ	minimized objective function result
y	model outputs $(y_0, \ldots, y_n)^T$
$z_t, z_{t,h}, z_{t,l}$	desired trajectory target or dead-band
w_h, w_l	penalty factors outside trajectory dead-band
$c_y, c_u, c_{\Delta u}$	cost of variables y, u, and Δu, respectively
u, x, d	inputs (u), states(x), and parameters or disturbances(d)
f, g, h	equation residuals(f), output function (g), and inequality constraints (h)
τ_c	time constant of desired controlled variable response
e_l, e_h	slack variable below or above the trajectory dead-band
SP, SP_{lo}, SP_{hi}	target, lower, and upper bounds to final set point dead-band

3. Dynamic Estimation and Optimization for a Batch Distillation Column

This established methodology is demonstrated for the first time on a binary batch distillation column. While the methods are not new, the application to this specific column is novel and gives experimental insight on issues encountered when applying dynamic optimization on applications that share common features. This section is subdivided into a brief discussion of the apparatus and experimental procedure, parameter estimation and validation, and model optimization and validation.

3.1. Apparatus and Experimental Procedure

A 38 tray, 2 inch, vacuum-jacketed and silvered Oldershaw column is used to collect all experimental data (see Figure 2). Cooling water supplies the energy sink for the total condenser at the top of the column. A 600 W reboiler heater is the only source of energy input. Reflux ratio is set using a swinging bucket and can be changed as frequently as every 5 minutes. The instantaneous distillate composition is determined using the refractive index of the solution and the total distillate collected is determined via a graduated cylinder. Cumulative distillate composition can be measured and inferred using the instantaneous compositions and a mass balance. The instantaneous distillate composition can be measured every 5 minutes. The reboiler is initially charged with 1.5 L of a 50/50 wt% mixture of methanol and ethanol for each run, with the goal being a product of 99 mol% methanol.

The non-optimized base case experiment consists of running the column at total reflux for 30 minutes, then setting the reflux ratio to a constant value, usually somewhere between 3 and 5, and letting the column run until the cumulative overhead composition reaches 99 mol% methanol. The collection time usually lasts 60 to 90 minutes, depending on the reflux ratio. The instantaneous and cumulative compositions for a typical

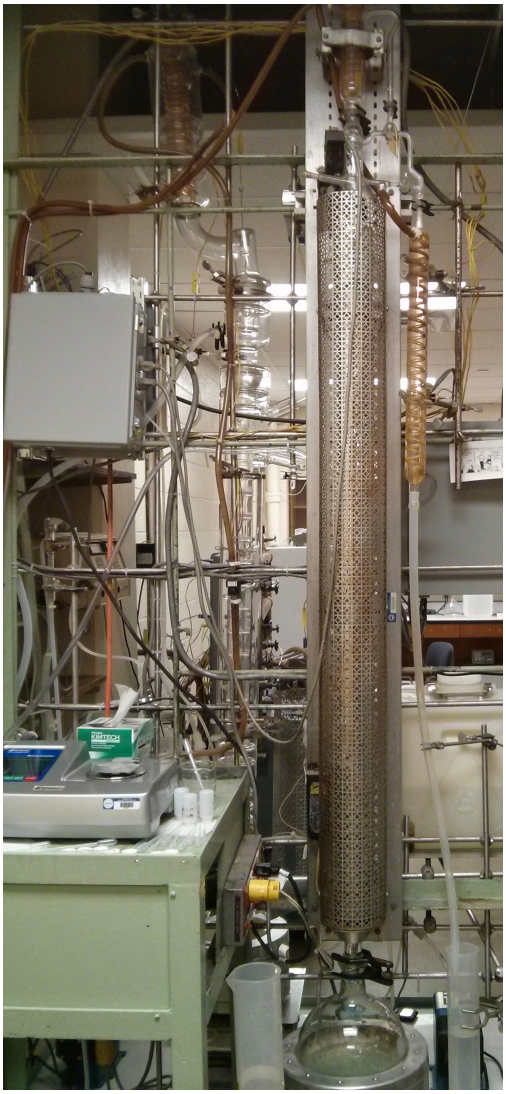
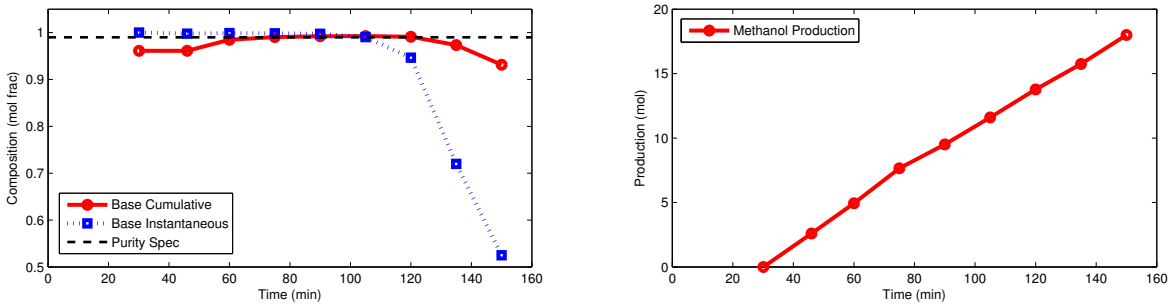


Figure 2: Apparatus used for the experiments

run, as well as the amount of product collected, can be seen in Figures (3a) and (3b), respectively. In this case, running the column at total reflux for 30 minutes, then using a constant reflux ratio of 4 for the next 90 minutes resulted in 13.7 moles of 99 mol% methanol.



(a) Instantaneous and cumulative product mole fraction

(b) Methanol production

Figure 3: Non-optimized base case where the final required purity (> 99 mol% ethanol) is not met

3.2. Equations for the Simplified Process Model

Distillation is an inherently complex process involving mass and energy transfer, thermodynamics, and often reaction kinetics. Models that describe these phenomena do not have to be complex, however. The model developed here is used to describe the separation of a 50/50 wt% mixture of methanol and ethanol, and is simple by design to illustrate this point.

The VLE model used here is found in the CHEMCAD database [65] and is shown in Eq. 6:

$$y_n^* = -2.016x_n^4 + 0.6861x_n^3 - 1.206x_n^2 + 1.721x_n + 0.0003984 \tag{6}$$

where x_n is the liquid mole fraction of methanol and y_n^* is the vapor mole fraction of methanol in equilibrium with the liquid. The subscript n denotes the stage for which the mole fraction is being calculated. An adjustment to the equilibrium vapor mole fraction is used because equilibrium is not often achieved during column operation. This adjustment is in the form of a Murphree efficiency and is shown in Eq. 7:

$$y_n = y_{n+1} - E_{MV}(y_{n+1} - y_n^*) \tag{7}$$

where y_n is the actual mole fraction and E_{MV} is the efficiency. The efficiency is a fitting parameter used to account for system non-idealities and is found using the data collected as part of this work.

The liquid mole fraction for each stage is found by performing a material balance at each stage, n, as shown by Eq. 8 where V is the vapor flow through the column, L is the liquid return flow, and N_{tray} is the number of moles of liquid on the stage. The number of moles and the composition in the reboiler (N_{reb} and x_{reb}) change with time and are represented by Eqs. 9 - 10. The number of moles in the condenser (N_{cond}) is assumed constant while the composition of the condenser (x_{cond}) varies throughout the run (see Eq. 11). Variation of the number of moles and composition of the product with time are represented by Eqs. 12 and

13. The liquid holdup for the condenser and trays are also design variables and are described in Eqs. 14 and 15, where f_{tray} and f_{cond} are the fitting parameters representing the fraction of the initial reboiler charge on each tray and in the condenser, respectively. The tray holdup is assumed constant across all stages. The stages are numbered from 1 to 40 with the top being 1 (condenser).

$$\frac{dx_n}{dt} = \frac{L(x_{n-1} - x_n) - V(y_n - y_{n+1})}{N_{tray}} \tag{8}$$

$$x_{reb}\frac{dN_{reb}}{dt} + N_{reb}\frac{dx_{reb}}{dt} = Lx_{39} - Vy_{reb} \tag{9}$$

$$\frac{dN_{reb}}{dt} = L - V \tag{10}$$

$$N_{cond}\frac{dx_{cond}}{dt} = V\ (y_2 - x_{cond}) \tag{11}$$

$$\frac{dn_p}{dt} = D \tag{12}$$

$$x_p\frac{dn_p}{dt} + n_p\frac{dx_p}{dt} = D\ x_{cond} \tag{13}$$

$$N_{cond} = N_{reb.init}\ f_{cond} \tag{14}$$

$$N_{tray} = N_{reb.init}\ f_{tray} \tag{15}$$

The vapor flow rate is found using the energy balance shown in Eq. 16:

$$V = \frac{h_{dot}\ h_f}{H_{vap}} \tag{16}$$

where h_{dot} is the heat input from the heater, H_{vap} is the heat of vaporization for the methanol/ethanol system, and h_f is a fitting parameter representing the heating efficiency. The heat of vaporization is approximated as a weighted average of the pure component heats of vaporization obtained from the DIPPR Database [66]. The liquid flow rate, the reflux ratio, and the distillate rate are found using an overall mass balance and the definition of the reflux ratio, shown in Eqs. 17 and 18, respectively:

$$V = L + D \tag{17}$$

$$R = \frac{L}{D} \tag{18}$$

where R is the reflux ratio and D is the distillate rate. Constant molar overflow is assumed throughout the model and applies to the equations shown above.

3.3. Equations for the Detailed Process Model

A more detailed (although not completely from first-principles) model [67] with energy balance equations validates the simplified model developed in Section 3.2. A similar notation as the simplified model is used for the detailed model with a distinction in the stage number in which the material and energy balances are developed. Vapor and liquid leaving each stage are noted as V_n and L_n, respectively. The equations used in the detailed model are based on the following assumptions:

- constant molar hold up for the condenser and trays
- fast heat transfer throughout the column
- liquid temperature on each tray at the mixture bubble point
- vapor liquid equilibrium relationships based on temperature dependent vapor pressures
- pressure drop across each tray is 1 $mmHg = \Delta P$
- temperature dependent density, heat capacity, vapor pressure, and heat of vaporization

The overall and component mole balances as well as the energy balance equation for a control volume over the condenser and accumulator lead to Eqs. 19 to 22.

$$V_2 = L_1 + D \tag{19}$$

$$L_1 = R\ D \tag{20}$$

$$N_{cond}\frac{\mathrm{d}x_{cond}}{\mathrm{d}t} = V_2\ y_2 - (L_1 + D)\ x_{cond} \tag{21}$$

$$Q_{cond} = V_2\ h_{V_2} - (L_1 + D)\ h_{L_1} \tag{22}$$

A component and overall mole balance over the trays result in Eqs. 23 and 24. Eq. 25 also represents an energy balance for each tray in the column.

$$N_{tray}\frac{\mathrm{d}x_n}{\mathrm{d}t} = L_{n-1}\ x_{n-1} - L_n\ x_n + V_{n+1}\ y_{n+1} - V_n\ y_n \tag{23}$$

$$0 = V_{n+1} - V_n + L_{n-1} - L_n \tag{24}$$

$$V_{n+1}\ (h_{V_{n+1}} - h_{L_n}) = V_n\ (h_{V_n} - h_{L_n}) - L_{n-1}\ (h_{L_{n-1}} - h_{L_n}) \tag{25}$$

A component mole balance and the associated energy balance equation for the reboiler are presented by Eqs. 26 and Eq. 27. The reboiler heating rate, Q_{reb}, is 600 W to drive the separation together with the cooling of the condenser, Q_{cond}. The overall mole balance for this model is similar to the simplified model (Eq. 10).

$$x_{reb}\frac{\mathrm{d}N_{reb}}{dt} + N_{reb}\frac{\mathrm{d}x_{reb}}{dt} = L_{39}\ x_{39} - V_{40}\ y_{reb} \tag{26}$$

$$Q_{reb}\ h_f = V_{40}\ (h_{V_{40}} - h_{L_{40}}) - L_{39}\ (h_{L_{39}} - h_{L_{40}}) \tag{27}$$

Accumulation of product and the change in composition of the product with respect to changes in product moles are shown in Eqs. 12 and 13. The enthalpy of mixture for both liquid and gas phases is a mole average of the enthalpy of each component. Enthalpy of each component is obtained by integrating the heat capacity for liquid and adding the heat of vaporization for vapor. The temperature profile in the column is also a function of the equilibrium composition of each stage. The relationship between temperature and liquid composition of each stage is based on vapor pressure and the pressure on each tray (P_n) as shown in 28 with $n_s = 2$.

$$P_1 = 0.86\ atm\ (\text{Ambient Pressure in Provo, UT}) \tag{28a}$$

$$P_n = P_{n-1} - \Delta P \tag{28b}$$

$$P_n = \sum_{i=1}^{n_s} \gamma_i\ x_i\ P_i^{sat}\,(T_i) \tag{28c}$$

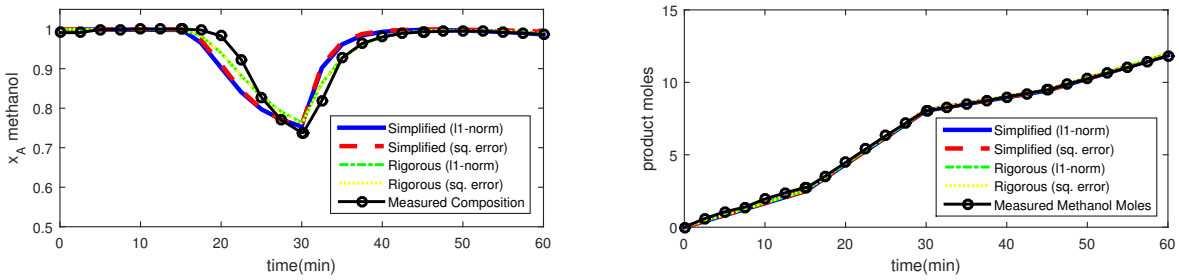
The vapor composition at each tray is determined by the vapor liquid equilibrium correlation shown in Eq. 29 and is combined with the previous Eq. 7 to relate the equilibrium composition (y_n^*) to the actual tray composition (y_n) based on the Murphree efficiency.

$$y_n^*\ P_n = \gamma\ x_n\ P_n^{sat}(T_n) \tag{29}$$

A full listing of the model equations, data, and Python source code is given in Appendix A. The more sophisticated model demonstrates that the simpler and less rigorous model is able to adequately predict the batch column performance for the purpose of optimization. The model validation is shown in the subsequent section.

3.4. Model Validation

Model validation is accomplished through dynamic parameter estimation. The parameter estimation experiment was similar to a doublet test, with reflux ratios set to 3.5, 1, 7 and 3.5. The column was allowed to come to steady state at infinite reflux before starting data collection; the reflux ratio was adjusted every 15 minutes thereafter. The parameters found by fitting the model with experimental data are heater efficiency



(a) Instantaneous distillate composition

(b) Methanol production

Figure 4: Model validation for initial parameter estimation

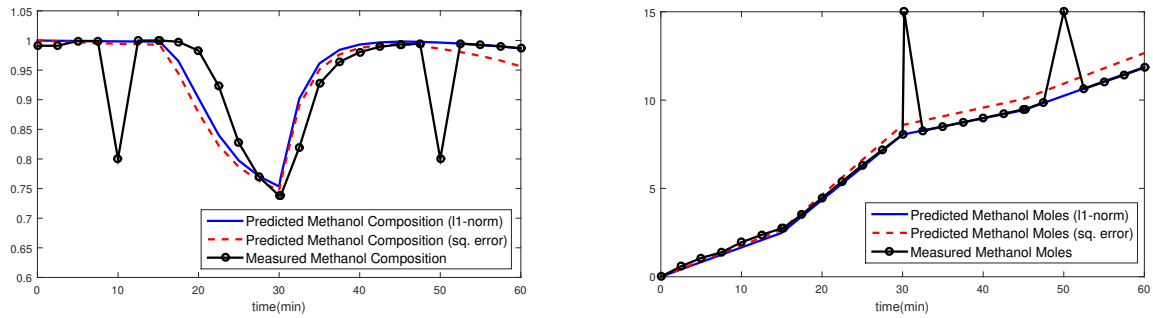
(h_f), vaporization efficiency (E_{MV}), condenser molar holdup as a fraction of initial reboiler charge (f_{cond}), and tray molar holdup as a fraction of initial reboiler charge (f_{tray}). The parameter best estimates are shown in Table 3.

Table 3: Confidence interval calculation for the four parameter case

Parameter	Best Estimate	Upper 95% CI	Lower 95% CI
h_f	0.719	0.799	0.639
E_{MV}	0.691	2.420	0
f_{cond}	0.029	0.254	0
f_{tray}	5.077e-4	0.142	0

The instantaneous distillate composition from the experimental run and the associated simplified and detailed model predictions using optimized parameters are shown in Figure (4a). The maximum error between the simplified model predictions and the experimental values is 10%. The maximum error between the more detailed model and experimental composition data is 4.8% for the ℓ_1-norm objective and 5.3% for the squared error objective. Cumulative methanol production is shown in Figure (4b). The error between model and prediction is almost non-existent using both an ℓ_1-norm or squared error objective. The simplified model parameter estimation has 3,510 equations with the squared error objective and 3,780 equations with the ℓ_1-norm objective and requires less than 10 CPU seconds to solve. The more detailed model parameter estimation has 11,644 equations with the squared error objective and 11,972 equations with the ℓ_1-norm objective and requires 89.4 (ℓ_1-norm) and 53.1 (squared error) CPU seconds to solve. All calculations are performed on a Intel Core i7-2760QM CPU operating at 2.4 GHz with the APOPT solver. Because the simplified model produces similar results to the detailed model and solves sufficiently fast for online real-time optimization, it is selected for the batch column optimization.

If artificial outliers are introduced in both the composition (80 mol% ethanol at $t = 10\ min$ and $t = 50\ min$) and cumulative production (15 moles at $t = 30\ min$ and $t = 50\ min$), the squared error predictions deviate while the ℓ_1-norm estimates do not (see Figure 5).



(a) Instantaneous distillate composition

(b) Methanol production

Figure 5: Insensitivity of the ℓ_1-norm estimation to outliers compared to the squared error objective

While this particular example did not include significant outliers, many industrial applications of batch distillation may have instruments that report values with drift, noise, or outliers [68]. While gross error detection can resolve many of these data quality issues, it is also desirable to have estimation methods that are less sensitive to bad data as shown in this example.

3.5. Testing the Reliability of the Estimated Parameters

Nonlinear confidence intervals are calculated for four potential parameters. Confidence regions are typically reported as upper and lower limits on a particular parameter. This work extends the nonlinear confidence region to multivariate analysis that improve co-linearity assessment for batch distillation processes beyond a singular value decomposition or linear analysis. However, a look at the confidence interval for each individual parameter is useful to illustrate the procedure for model validation. A wide confidence interval suggests that there is insufficient structure in the model (observability) to determine the parameters from available measurements. Another insight that is gained from the confidence intervals is a test of the data diversity that leads to tight confidence regions. A tighter confidence region implies that a smaller deviation of the parameter from an optimal value is not statistically likely given a set of data to which the model is reconciled. Table 3 shows the expected value and 95% confidence interval for each parameter. As seen in the table, the interval for heater efficiency is narrow and in the range of values expected for a heater. The intervals for the other three parameters are large enough to include zero and the interval for vapor efficiency includes physically impossible values. Although the fit between model and data is excellent there are large parameter confidence intervals. One possible explanation for the large intervals is that the model is over-parameterized and thus has too many degrees of freedom. Thus, a sensitivity analysis is implemented to investigate the correct parameterization of the model.

The scaled sensitivity is shown graphically in Figure 6. The sensitivity is scaled by solution values as $\hat{S}_{i,j} = \left(\nabla_{\theta_j} x_i\right) \frac{\bar{\theta}_i}{\bar{x}_i}$ to show relative effects with a unitless transformation. The scaling is applied with parameters $\bar{\theta}$ and variables $\bar{x}$ at solution values.

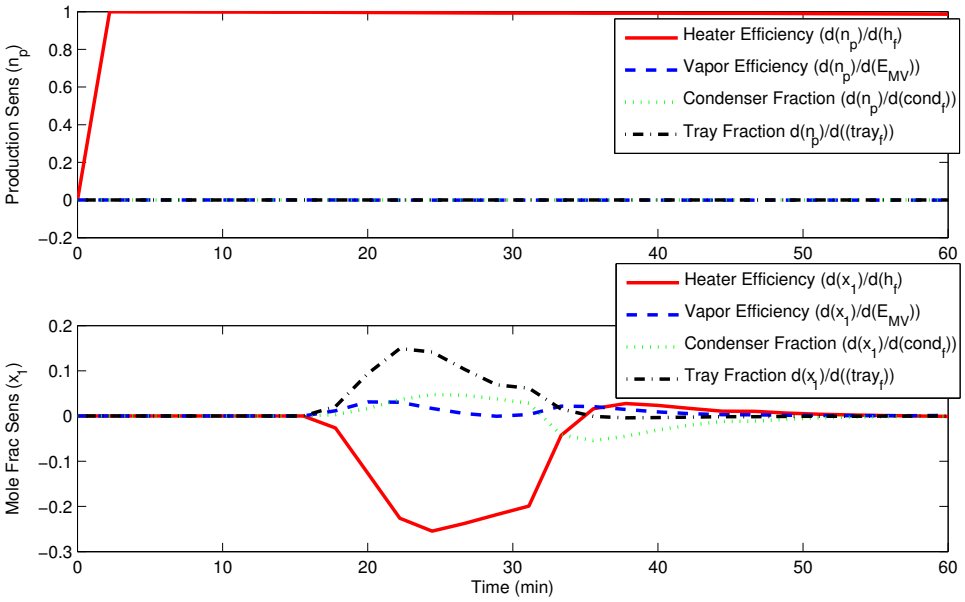


Figure 6: Scaled variable sensitivities to the parameters

One clear result from this sensitivity study is that the total production (n_p) is dependent on the heat input to the batch column and that other parameters have little effect on the total production. As expected, a higher heating rate (h_f) vaporizes additional liquid and increases the flow to the condenser. With a specified reflux rate, the total production rate increases proportionally. In other words, a 1% increase in heating produces 1% additional product. This scaled sensitivity is shown as a value of 1.0 in the top subplot of Figure 6. The sensitivities of instantaneous product composition to the parameters are nearly co-linear as seen by the bottom subplot of Figure 6. For example, heater efficiency (h_f) and tray holdup fraction (f_{tray}) can be increased and decreased, respectively, to produce nearly the same final answer. Other parameters also show a high degree of co-linearity.

While sensitivity plots such as Figure 6 are instructive, it can be difficult for large-scale systems to detect co-linearity or the number and selection of parameters that can be estimated from the data. An alternative way to show the same information is to decompose the sensitivity matrix with a singular value decomposition to reveal magnitudes of singular values (relative importance of transformed linear combinations of parameters) and eigenvectors (orthogonal vectors for the parameter space transformation). The singular value decomposition is applied to the dynamic sensitivity analysis to show that there is one principle parameter (h_f) that can be used to match production data (n_p) as shown in Figure 7.

In this application, the parameter h_f is principally used to match n_p. For selecting a next parameter, f_{tray} or E_{MV} are feasible candidates with similar effect on the model. Estimating a third parameter is likely not needed as seen by the magnitude of the singular values. The singular value analysis gives a linear combination of the parameters estimated in transformed parameter space as given by the eigenvectors.

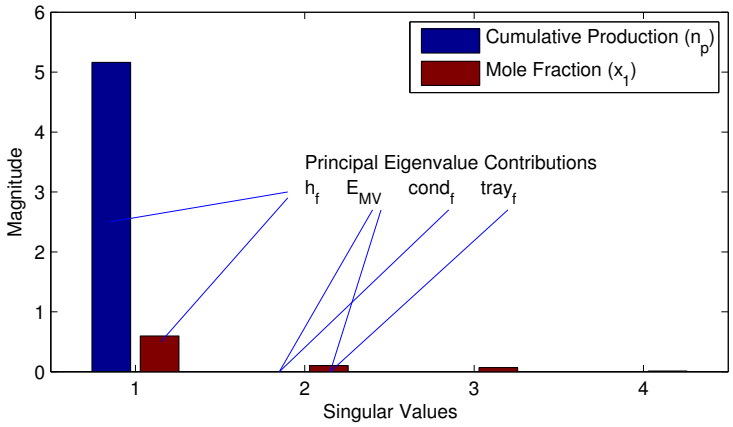


Figure 7: Singular values reveal independent linear combinations of parameters to reconcile data

This analysis is useful even for the non-transformed parameter estimation where the parameter estimates have physical meaning and constraints are enforced to reflect physical realism. For example, in the case of h_f, a value greater than 1.0 is not likely because it represents the fraction of reboiler heater duty that enters the liquid. It is expected that some of the heat escapes due to lack of insulation or conduction. In transformed space, the physical connection to the parameters is lost.

As mentioned, f_{tray} and E_{MV} have a similar effect on the model. In this study, E_{MV} is selected as the second parameter. It was therefore determined to first solve for all four parameters using ℓ_1-norm analysis, then fix both holdups and re-solve for the heater efficiency and the vaporization efficiency. The resulting confidence region and parameter best estimates are shown in Figure 8.

With only two parameters, the confidence regions are able to be graphically visualized. Instead of confidence intervals with lower and upper bounds, the 95% confidence region is a given by any point within the area on the contour plot that falls within the boundary. Both the ℓ_1-norm and squared error objectives are included in this plot to demonstrate that slightly different optimal solutions and confidence regions are reported for differing objectives that align model and measured values. One notable issue is that the objective function is relatively insensitive to vapor efficiency (E_{MV}), especially as the vapor efficiency is above 0.4. The 95% confidence region suggests that values between 0.37 and 1.0 are valid parameter estimates for E_{MV} and that only one parameter is required for parameter estimation. The objective function is very sensitive to heater efficiency (h_f) but not to E_{MV}. One possible explanation for this is that this is a high purity column where a difference of 0.01 in the mole fraction is of approximate equal importance to about 1.0 mole of production. Although the objective is scaled to account for this discrepancy, parameters such as h_f greatly influence both the predicted moles produced and the product composition. The additional parameter E_{MV} is required to achieve an acceptable fit for product composition although it is less influential than the value of h_f. The objective function contours confirm the observations from the sensitivity analysis and singular value decomposition shown previously in Figures 6 and 7. The fit to the parameter estimation experiment

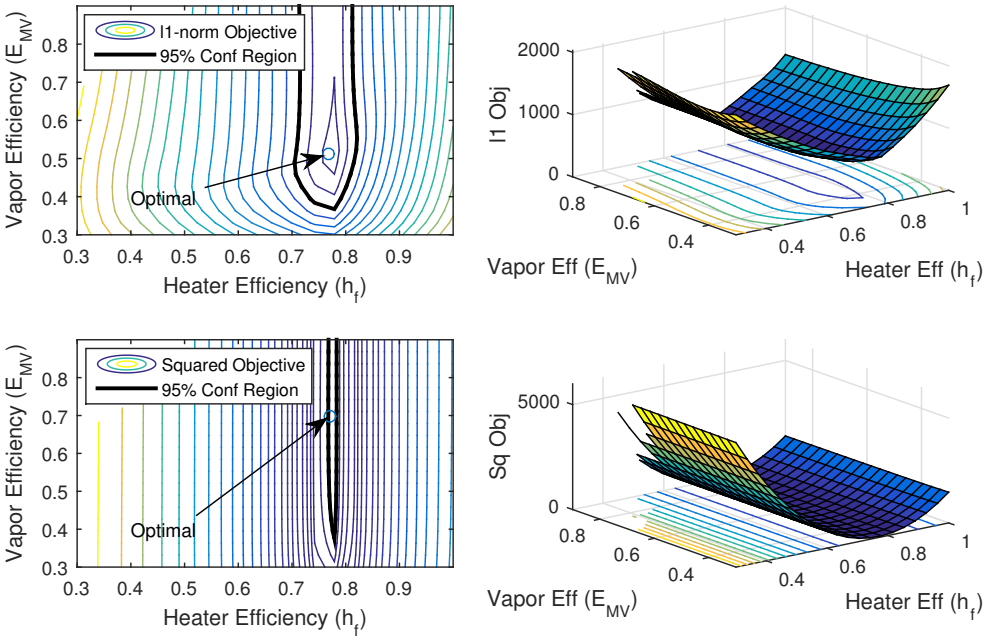
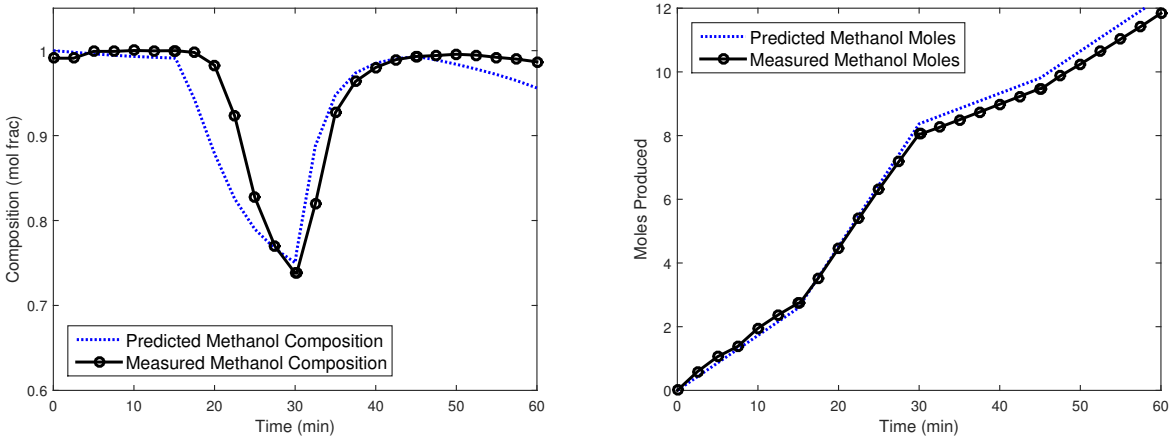


Figure 8: Contour and surface plots of the objective function value for values of heater efficiency (h_f) and vapor efficiency (E_{MV}). The 95% confidence interval for the ℓ_1-norm is not correct (future work) and the confidence interval for the squared error is an approximation.

is shown in Figures (9a) and (9b). With the model sufficiently validated, the next step is to optimize the column control scheme.



(a) Instantaneous distillate composition

(b) Methanol production

Figure 9: Model validation for final parameter estimates

3.6. Model Optimization and Validation

The objective in this case study is to maximize the amount of methanol produced in the column during a 90 minute run. The non-optimized base case production over a 90 minute run is 9.5 moles of 99.2 mol% methanol at a constant reflux ratio of 4 (see Section 3.1). The design variable in this study is reflux ratio, with the option to change the reflux ratio every 5 minutes. The control scheme for the optimized run is shown in Figure 10; the base case profile is shown for comparison purposes. The optimized reflux ratio scheme starts low before increasing in a nominally linear pattern. This is done to take advantage of the initially high concentration of methanol in the condenser after the startup period.

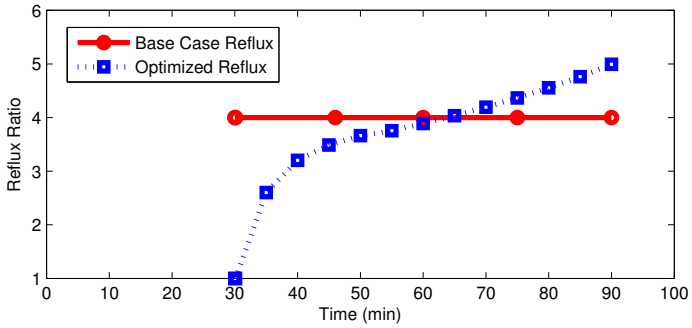
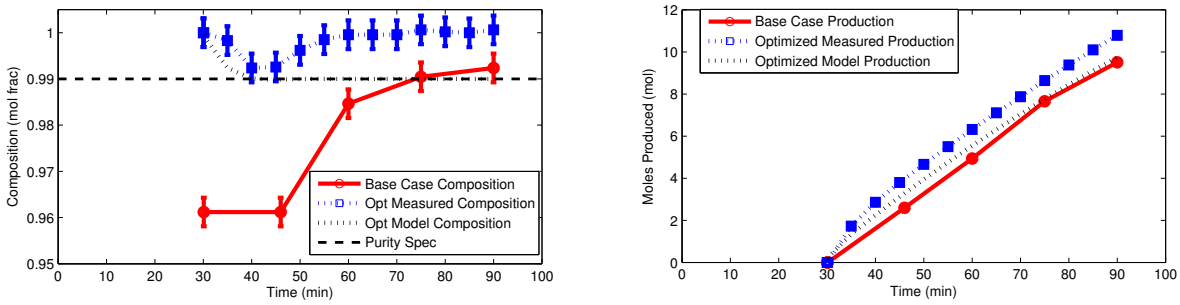


Figure 10: Reflux ratio for optimized control scheme compared to the non-optimized base case

The cumulative composition and total production are shown in Figure 11a and Figure 11b, respectively, with parameter values of $h_f = 0.8$, $E_{MV} = 0.37$, $f_{tray} = 0.0009$, and $f_{cond} = 0.006$. Also shown in the figures are the model predictions and the non-optimized base case results. The optimized control scheme resulted in 10.8 moles of 99.8 mol% methanol. This change represents a 14% increase in column production over the base case. Given the high concentration, it is possible to collect more product throughout the optimized run and still meet the purity specification. However, given the error associated with experimental measurements, the prediction was left at a slightly conservative estimate to ensure the purity specification was achieved. The success of this effort is seen in the fact that the error bars on the optimized composition measurements stay above the purity requirement while those for the non-optimized base case do not.



(a) Cumulative composition

(b) Methanol production

Figure 11: Optimized control scheme compared to the non-optimized base case and to the model prediction

Also seen in the figures are the model predictions. The model predicts 9.75 moles of 99.0 mol% methanol will be produced during the run. The difference between model prediction and experiment is 10% and 0.8% for overall production and product composition, respectively. The agreement between model and experiment is excellent and reflects the work done to validate the model.

4. Conclusions

Models of batch distillation are typically either first-principles and computationally expensive or simple and valid for ideal systems. In this work, a well-known methodology for parameter estimation, uncertainty quantification, and dynamic optimization is used to develop a simplified model for optimization of a batch distillation column. This methodology uses experimental data to solve for model fitting parameters and validates the results with nonlinear confidence intervals. This allows the models to include system non-idealities and be applicable for real-time analysis. This is accomplished using dynamic data with ℓ_1-norm error minimization. A dynamic sensitivity analysis reduces batch experimental data requirements by determining *a priori* which parameters can be estimated. Nonlinear statistics are applied to quantify *a posteriori* the accuracy of those same parameters. The results from the simplified model also agree with a first-principles model but the simplified model solves 5-10 times faster than a first-principles model. While the methodology is not novel, the application to this specific case study with experimental data is demonstrated for the first time with insight into practical implications of working with real data.

The case study involves optimizing the control scheme for an existing batch column. A 38 tray, 2 inch, vacuum-jacketed and silvered Oldershaw batch distillation column was used to collect experimental data. One experiment was performed to collect data for model validation and another experiment was performed to validate the optimized control scheme. The optimized control scheme resulted in a 14% production increase over the base case while still meeting the purity requirements. The model predictions for the optimized run are within 10% of the experimental data.

Appendix A. Batch Distillation Model

The binary batch distillation column is represented by 42 constants, 252 variables, 595 explicit equations (intermediates), and 241 implicit differential and algebraic equations (DAEs). The equations are discretized over the startup (30 minutes) and measurement time horizon (60 minutes) with added objective to produce a final nonlinear programming problem with 11,972 equations and 11,976 variables (for the ℓ_1-norm). The following model in Listing 1 is expressed in the APMonitor Modeling Language. The software is freely available at APMonitor.com as a MATLAB, Python, or Julia package for dynamic simulation and optimization. This particular set of files can be accessed from the following GitHub archive [69].

Listing 1: Binary Distillation Column Model in APMonitor Modeling Language

```
% Binary Batch Distillation Column
% Component 1 = methanol
% Component 2 = ethanol
Constants
  n = 40 % stages
  x0 = 0.59 % initial composition
  % Constants for heat of vaporization
  A_m = 3.2615e7
  B_m = -1.0407
  C_m = 1.8695
  D_m = -0.60801
  A_e = 6.5831e7
  B_e = 1.1905
  C_e = -1.7666
  D_e = 1.0012
  % Critical temperatures (K)
  Tc_m = 512.5
  Tc_e = 514
  % Density coefficients
  rho_m_1 = 2.3267
  rho_m_2 = 0.27073
  rho_m_3 = 512.05
  rho_m_4 = 0.24713
  rho_e_1 = 1.6288
  rho_e_2 = 0.27469
  rho_e_3 = 514
  rho_e_4 = 0.23178
  % Heat capacity coefficients
  cp_m_liq_1 = 2.5604E5
  cp_m_liq_2 = -2.7414E3
  cp_m_liq_3 = 1.4777E1
  cp_m_liq_4 = -3.5078E-2
  cp_m_liq_5 = 3.2719E-5
  cp_e_liq_1 = 1.0264E5
  cp_e_liq_2 = -1.3963E2
  cp_e_liq_3 = -3.0341E-2
  cp_e_liq_4 = 2.0386E-3
  cp_e_liq_5 = 0
  % Standard heats of formation (J/kmol)
  h_form_std_m = -2.391E8
  h_form_std_e = -2.7698E8
  % Vapor pressure coefficients
  vpm[1] = 82.718
  vpm[2] = -6904.5
  vpm[3] = -8.8622
  vpm[4] = 7.4664E-06
  vpm[5] = 2
  vpe[1] = 73.304
  vpe[2] = -7122.3
  vpe[3] = -7.1424
```

```
  vpe[4] = 2.8853E-06
  vpe[5] = 2
End Constants

Parameters
  rr = 3.5 % reflux ratio
  hf = 0.8 % fractional heat loss fraction
  vf = 0.45 % tray efficiency
  tray_hol = 0.07 % tray holdup
  condenser_hol = 0.144 % condenser holdup
  heat_rate = 36000 , > 0 % 36000 J/min = 600 W
  gamma = 1.0  % activity coefficient

Variables
  x[1:n] = x0 , >= 0 , <= 1
  y[2:n] = x0 , >= 0 , <= 1
  L[1:n-1] = 0.36 , > 0  % mol/min
  V[2:n] = 0.72 , > 0  % mol/min
  D = 0.36 , > 0  % mol/min
  boil_hol = 28 , >0 % mol
  Q_cond = 0
  np = 0 , >= 0 % mol
  xp = 0.99 , >= 0 , <= 1
  T[1:n] = 320 % tray temperature
  ystar[2:n] = x0 % theoretical vapor composition

Intermediates
  % tray pressures
  P[1] = 101325 * 0.86  % local atmospheric pressure
  P[2:n] = P[1:n-1] + 101325/760  % pressure drop
  % pure component and mixture vapor pressure (Pa)
  vp1[1:n] = exp(vpm[1]+vpm[2]/T[1:n]+vpm[3]*LOG(T[1:n])+vpm[4]*(T[1:n]^vpm[5]))
  vp2[1:n] = exp(vpe[1]+vpe[2]/T[1:n]+vpe[3]*LOG(T[1:n])+vpe[4]*(T[1:n]^vpe[5]))
  vp[1:n] = x[1:n] * vp1[1:n] + (1-x[1:n]) * vp2[1:n]
  % pure component and mixutre density (kmol/m3 or mol/L)
  rho_meth[1:n-1] = rho_m_1 / (rho_m_2^(1+(1-T[1:n-1]/rho_m_3)^rho_m_4))
  rho_etha[1:n-1] = rho_e_1 / (rho_e_2^(1+(1-T[1:n-1]/rho_e_3)^rho_e_4))
  rho_mix[1:n-1]  =  rho_meth[1:n-1] * x[1:n-1] + rho_etha[1:n-1] * (1-x[1:n-1])
  % pure component heat of vaporization (J/mol)
  Hvap_m[1:n] = A_m*(1-T[1:n]/Tc_m)^(B_m+C_m*(T[1:n]/Tc_m)+D_m*(T[1:n]/Tc_m)^2)/1000
  Hvap_e[1:n] = A_e*(1-T[1:n]/Tc_e)^(B_e+C_e*(T[1:n]/Tc_e)+D_e*(T[1:n]/Tc_e)^2)/1000
  % pure component liquid enthalpies (J/mol)
  h_liq_m[1:n] = (cp_m_liq_1 * (T[1:n])     + cp_m_liq_2 * (T[1:n])^2/2 + &
                  cp_m_liq_3 * (T[1:n])^3/3 + cp_m_liq_4 * (T[1:n])^4/4 + &
                  cp_m_liq_5 * (T[1:n])^5/5)/1000
  h_liq_e[1:n] = (cp_e_liq_1 * (T[1:n])     + cp_e_liq_2 * (T[1:n])^2/2 + &
                  cp_e_liq_3 * (T[1:n])^3/3 + cp_e_liq_4 * (T[1:n])^4/4 + &
                  cp_e_liq_5 * (T[1:n])^5/5)/1000
  % pure component vapor enthalpies (J/mol)
  h_gas_m[2:n] = h_liq_m[2:n] + Hvap_m[2:n]
  h_gas_e[1:n] = h_liq_e[1:n] + Hvap_e[1:n]
  % tray vapor and liquid enthalpies (J/mol)
  h_gas[2:n] = y[2:n] * h_gas_m[2:n] + (1-y[2:n])*h_gas_e[2:n]
  h_liq[1:n] = x[1:n] * h_liq_m[1:n] + (1-x[1:n])*h_liq_e[1:n]

Equations
  % tray bubble point temperature
  P[1:n] = vp[1:n]
  % vapor liquid equilibrium
  ystar[2:n] * P[2:n] = gamma * x[2:n] * vp1[2:n]
  % non-ideal separation with tray efficiency
  y[n] = ystar[n]
  y[2:n-1] = y[3:n]-vf*(y[3:n]-ystar[2:n-1])
  % reflux ratio = L/D
  L[1] = rr * D
  % Condenser mole balance (methanol)
  condenser_hol * x[1] = - (L[1]+D) * x[1] + V[2] * y[2]
  % Tray mole balance (methanol)
  tray_hol * $x[2:n-1] = L[1:n-2] * x[1:n-2] - (L[2:n-1]) * x[2:n-1] &
                - V[2:n-1] * y[2:n-1] + y[3:n] * V[3:n]
  % Reboiler mole balance (methanol)
  boil_hol * $x[n] + $boil_hol * x[n] = L[n-1] * x[n-1] - V[n] * y[n]
```

```
% Overall condenser mole balance
V[2] = D * (rr+1)
% Overall tray mole balance
0 = V[3:n] + L[1:n-2] - V[2:n-1] - L[2:n-1]
% Energy balance (no dynamics)
0 = (V[2]* (h_gas[2] - h_liq[1]) - Q_cond)
0 =  V[3:n]   * (h_gas[3:n]   - h_liq[2:n-1]) - V[2:n-1] * (h_gas[2:n-1] - h_liq[2:n-1]) &
   - L[1:n-2] * (h_liq[1:n-2] - h_liq[2:n-1])
0 = heat_rate * hf - V[n]   * (h_gas[n]-h_liq[n]) - L[n-1] * (h_liq[n-1]-h_liq[n])
% Production rate equations
$boil_hol = -D
$np = D
xp * $np + np * $xp = x[1] * D
```

The following Python script shown in Listing 2 is the list of commands necessary to reproduce the dynamic batch distillation case presented in this paper. The parameter estimation uses two external files including the model file (distill.apm) and a data file (data.csv) that are also shown in this Appendix.

Listing 2: Python Dynamic Estimation

```
from apm import *
s = 'http://byu.apmonitor.com'
a = 'distill_l1_norm'
apm(s,a,'clear all')
apm_load(s,a,'distill.apm')
csv_load(s,a,'data.csv')
apm_option(s,a,'nlc.imode',5)
apm_option(s,a,'nlc.max_iter',100)
apm_option(s,a,'nlc.nodes',2)
apm_option(s,a,'nlc.time_shift',0)
apm_option(s,a,'nlc.ev_type',1)
apm_info(s,a,'FV','hf')
apm_info(s,a,'FV','vf')
apm_info(s,a,'FV','tray_hol')
apm_info(s,a,'FV','condenser_hol')
apm_info(s,a,'CV','x[1]')
apm_info(s,a,'CV','np')
output = apm(s,a,'solve')
print(output)
apm_option(s,a,'hf.status',1)
apm_option(s,a,'vf.status',1)
apm_option(s,a,'tray_hol.status',1)
apm_option(s,a,'condenser_hol.status',1)
apm_option(s,a,'x[1].fstatus',1)
apm_option(s,a,'np.fstatus',1)
apm_option(s,a,'x[1].wsphi',10000)
apm_option(s,a,'x[1].wsplo',10000)
apm_option(s,a,'np.wsphi',10)
apm_option(s,a,'np.wsplo',10)
apm_option(s,a,'x[1].meas_gap',1e-4)
apm_option(s,a,'np.meas_gap',0.01)
apm_option(s,a,'hf.lower',0.001);
apm_option(s,a,'hf.upper',1.0);
apm_option(s,a,'vf.lower',0.001);
apm_option(s,a,'vf.upper',0.6);
apm_option(s,a,'tray_hol.lower',0.01);
apm_option(s,a,'tray_hol.upper',0.1);
apm_option(s,a,'condenser_hol.lower',0.1)
apm_option(s,a,'condenser_hol.upper',0.5)
output = apm(s,a,'solve')
print(output)
y = apm_sol(s,a)
print('hf: ' + str(y['hf'][-1]))
print('vf: ' + str(y['vf'][-1]))
print('tray_hol: ' + str(y['tray_hol'][-1]))
print('cond_hol: ' + str(y['condenser_hol'][-1]))
```

```
print('np: ' + str(y['np'][-1]))
print('xp: ' + str(y['xp'][-1]))

import matplotlib.pyplot as plt
import pandas as pd
data_file = pd.read_csv('data_for_plotting.csv')

plt.figure(1)
plt.subplot(3,1,1)
plt.plot(y['time'],y['np'],'bx-',linewidth=2.0)
plt.plot(data_file['time'],data_file['np'],'ro')
plt.legend(['Predicted','Measured'])
plt.ylabel('Moles')

ax = plt.subplot(3,1,2)
plt.plot(y['time'],y['x[1]'],'bx-',linewidth=2.0)
plt.plot(data_file['time'],data_file['x[1]'],'ro')
plt.plot(y['time'],y['xp'],'k:',linewidth=2.0)
plt.legend(['Predicted','Measured','Cumulative'])
plt.ylabel('Composition')
ax.set_ylim([0.6, 1.05])

plt.subplot(3,1,3)
plt.plot(y['time'],y['x[1]'],'bx-',linewidth=2.0)
plt.plot(y['time'],y['x[2]'],'k:',linewidth=2.0)
plt.plot(y['time'],y['x[5]'],'r--',linewidth=2.0)
plt.plot(y['time'],y['x[10]'],'m.-',linewidth=2.0)
plt.plot(y['time'],y['x[20]'],'y-',linewidth=2.0)
plt.plot(y['time'],y['x[30]'],'g-.',linewidth=2.0)
plt.plot(y['time'],y['x[40]'],'k-',linewidth=2.0)
plt.legend(['x1','x2','x5','x10','x20','x30','x40'])
plt.ylabel('Composition')

plt.savefig('results_l1.png')
plt.show()
```

The data file includes time, reflux ratio, the instantaneous product composition, and the total product moles. The data file includes the first 30 minutes with nearly infinite reflux when the batch column approaches a steady state. At 30 minutes, the reflux ratio is changed to collect dynamic data at regular intervals as shown in Table A.4.

Figure A.12 shows the results of the ℓ_1-norm parameter estimation that are computed with the Python script in Listing 2. The first subplot shows the predicted and measured total moles. Note that the data collection starts after 30 minutes when the column is initially brought to steady state. The second and third subplots show the tray and product compositions. Over the first 30 minutes, there is insignificant total production. The product composition is below the 99% target but quickly reaches the desired purity once the reflux ratio is changed to allow production. Some of the individual tray compositions are shown in the final subplot. However, these compositions are not measured directly, only predicted from the model fit to the produced moles and product composition measurements.

Appendix B. References

[1] A. Lucia, B. R. McCallum, Energy targeting and minimum energy distillation column sequences, Computers & Chemical Engineering 34 (6) (2010) 931 – 942. doi:10.1016/j.compchemeng.2009.10.006.

[2] P. Li, H. A. Garcia, G. Wozny, E. Reuter, Optimization of a semibatch distillation process with model

Table A.4: Data File (data.csv) in Tabular Form

Time (s)	Reflux Ratio	Composition	Product Moles
0	10000	x	x
2.5	10000	x	x
5	10000	x	x
7.5	10000	x	x
10	10000	x	x
12.5	10000	x	x
15	10000	x	x
15.2	10000	x	x
17.5	10000	x	x
20	10000	x	x
22.5	10000	x	x
25	10000	x	x
27.5	10000	x	x
29.9	10000	x	x
30.1	3.5	x	x
32.5	3.5	0.991101728	0.590780164
35	3.5	0.999141476	1.060146
37.5	3.5	0.999141476	1.381291045
40	3.5	1.000195093	1.949735921
42.5	3.5	0.999773764	2.369812428
45	3.5	0.999984448	2.740502759
45.2	1	0.999984448	2.740502759
47.5	1	0.997875836	3.530570903
50	1	0.982790672	4.46255546
52.5	1	0.923087024	5.394604756
55	1	0.827170789	6.312827944
57.5	1	0.769688475	7.187411473
60	1	0.738080225	8.051023448
60.2	7	0.738080225	8.051023448
62.5	7	0.819332828	8.256959561
65	7	0.927134663	8.496361625
67.5	7	0.963811731	8.739591512
70	7	0.980221149	8.984573896
72.5	7	0.98940185	9.230547824
75	7	0.993222995	9.476936818
75.2	3.5	0.993222995	9.476936818
77.5	3.5	0.994282141	9.871343681
80	3.5	0.995763281	10.24134224
82.5	3.5	0.994705521	10.63582289
85	3.5	0.992162858	11.0298608
87.5	3.5	0.990039603	11.42352977
90	3.5	0.986847248	11.84121504

validation on the industrial site, Industrial & Engineering Chemistry Research 37 (4) (1998) 1341–1350. doi:10.1021/ie9706951.

[3] S. Ulas, U. M. Diwekar, M. A. Stadtherr, Uncertainties in parameter estimation and optimal control in batch distillation, Computers & Chemical Engineering 29 (8) (2005) 1805 – 1814. doi:10.1016/j.compchemeng.2005.03.002.

[4] K. Wang, T. Löhl, M. Stobbe, S. Engell, A genetic algorithm for online-scheduling of a multiproduct polymer batch plant, Computers & Chemical Engineering 24 (2) (2000) 393–400.

[5] E. Robinson, The optimisation of batch distillation operations, Chemical Engineering Science 24 (11) (1969) 1661 – 1668. doi:10.1016/0009-2509(69)87031-4.

[6] D. Mayur, R. May, R. Jackson, The time-optimal problem in binary batch distillation with a recycled

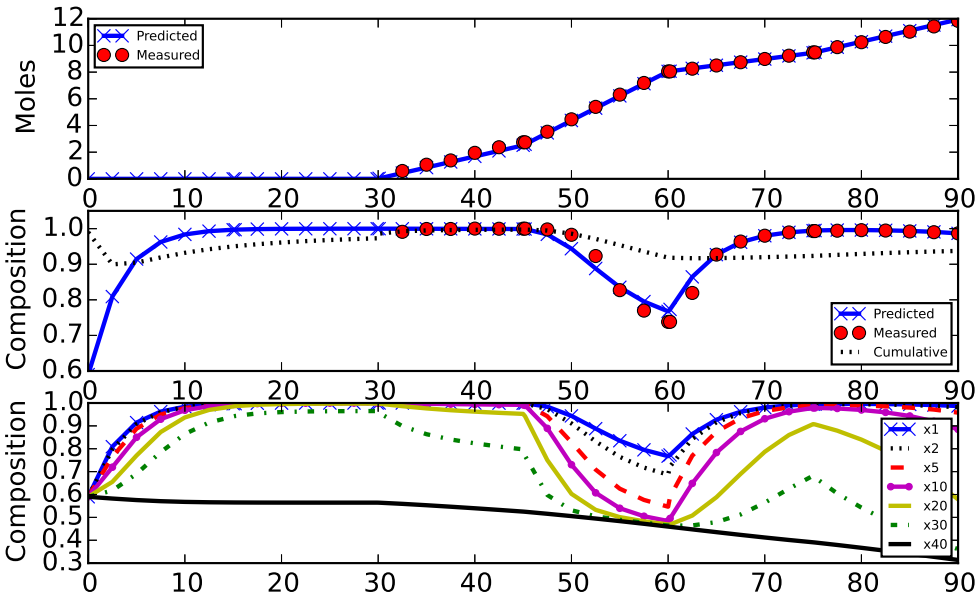


Figure A.12: Results of the ℓ_1-norm Estimation with the Detailed Model

601 waste-cut, The Chemical Engineering Journal 1 (1) (1970) 15 – 21, the Chemical Engineering Journal.
602 doi:10.1016/0300-9467(70)85026-2.

603 [7] E. Robinson, The optimal control of an industrial batch distillation column, Chemical Engineering
604 Science 25 (6) (1970) 921 – 928. doi:10.1016/0009-2509(70)85037-0.

605 [8] W. L. Luyben, Some practical aspects of optimal batch distillation design, Industrial & Engineering
606 Chemistry Process Design and Development 10 (1) (1971) 54–59. doi:10.1021/i260037a010.

607 [9] Y.-S. Kim, Optimal control of time-dependent processes (design, batch distillation, catalyst deactiva-
608 tion).

609 [10] U. Diwekar, Batch Distillation: Simulation, Optimal Design, and Control, CRC Press, 1995.

610 [11] B. T. Safrit, A. W. Westerberg, Improved operational policies for batch extractive distillation columns,
611 Industrial & Engineering Chemistry Research 36 (2) (1997) 436–443. doi:10.1021/ie960343z.

612 [12] K. H. Low, E. Srensen, Simultaneous optimal configuration, design and operation of batch distillation,
613 AIChE Journal 51 (6) (2005) 1700–1713. doi:10.1002/aic.10522.

614 [13] L. Bonny, Multicomponent batch distillations campaign: Control variables and optimal recycling policy,
615 Industrial & Engineering Chemistry Research 45 (26) (2006) 8998–9009. doi:10.1021/ie0609057.

[14] L. Bonny, Multicomponent batch distillations campaign: Control variables and optimal recycling policy. a further note, Industrial & Engineering Chemistry Research 52 (5) (2013) 2190–2193. doi:10.1021/ie3018302.

[15] J. Kim, D. Ju, Multicomponent batch distillation with distillate receiver, Korean Journal of Chemical Engineering 20 (3) (2003) 522–527. doi:10.1007/BF02705559.

[16] S.-S. Jang, Dynamic optimization of multicomponent batch distillation processes using continuous and discontinuous collocation polynomial policies, The Chemical Engineering Journal 51 (2) (1993) 83 – 92. doi:10.1016/0300-9467(93)80014-F.

[17] V. A. Varma, G. V. Reklaitis, G. Blau, J. F. Pekny, Enterprise-wide modeling & optimizationan overview of emerging research challenges and opportunities, Computers & Chemical Engineering 31 (5) (2007) 692–711.

[18] C. Loeblein, J. Perkins, B. Srinivasan, D. Bonvin, Performance analysis of on-line batch optimization systems, Computers & Chemical Engineering 21 (1997) S867–S872.

[19] H. Shi, F. You, A novel adaptive surrogate modeling-based algorithm for simultaneous optimization of sequential batch process scheduling and dynamic operations, AIChE Journal.

[20] W. Dai, D. P. Word, J. Hahn, Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process, Control Engineering Practice 22 (2014) 231–241.

[21] Y. Chu, F. You, Integration of scheduling and dynamic optimization of batch processes under uncertainty: Two-stage stochastic programming approach and enhanced generalized benders decomposition algorithm, Industrial & Engineering Chemistry Research 52 (47) (2013) 16851–16869.

[22] B. Srinivasan, S. Palanki, D. Bonvin, Dynamic optimization of batch processes: I. characterization of the nominal solution, Computers & Chemical Engineering 27 (1) (2003) 1–26.

[23] T. Bhatia, L. T. Biegler, Dynamic optimization in the design and scheduling of multiproduct batch plants, Industrial & Engineering Chemistry Research 35 (7) (1996) 2234–2246.

[24] P. Terwiesch, M. Agarwal, D. W. Rippin, Batch unit optimization with imperfect modelling: a survey, Journal of Process Control 4 (4) (1994) 238–258.

[25] K. Kim, U. Diwekar, New era in batch distillation: Computer aided analysis, optimal design and control, Reviews in Chemical Engineering 17 (2) (2001) 111 – 164.

[26] J. Stichlmair, J. R. Fair, Distillation: principles and practices, Vch Verlagsgesellschaft Mbh, 1998.

[27] M. F. Doherty, M. F. Malone, Conceptual design of distillation systems, McGraw-Hill Science/Engineering/Math, 2001.

[28] U. Diwekar, R. Malik, K. Madhavan, Optimal reflux rate policy determination for multicomponent batch distillation columns, Computers & Chemical Engineering 11 (6) (1987) 629 – 637. doi:10.1016/0098-1354(87)87008-4.

[29] M. Barolo, G. Guarise, Batch distillation of multicomponent systems with constant relative volatilities, Chemical Engineering Research and Design 74 (8) (1996) 863 – 871. doi:http://dx.doi.org/10.1205/026387696523166.

[30] Y. H. Kim, Optimal design and operation of a multi-product batch distillation column using dynamic model, Chemical Engineering and Processing: Process Intensification 38 (1) (1999) 61 – 72. doi:http://dx.doi.org/10.1016/S0255-2701(98)00065-8.

[31] I. Mujtaba, M. Hussain, Optimal operation of dynamic processes under process-model mismatches: Application to batch distillation, Computers & Chemical Engineering 22, Supplement 1 (1998) S621 – S624, european Symposium on Computer Aided Process Engineering-8. doi:http://dx.doi.org/10.1016/S0098-1354(98)00109-4.

[32] S. Gruetzmann, G. Fieg, Startup operation of middle-vessel batch distillation column: modeling and simulation, Industrial & Engineering Chemistry Research 47 (3) (2008) 813–824. doi:10.1021/ie070667v.

[33] M. Hanke, P. Li, Simulated annealing for the optimization of batch distillation processes, Computers & Chemical Engineering 24 (1) (2000) 1 – 8. doi:http://dx.doi.org/10.1016/S0098-1354(00)00317-3.

[34] M. Diehl, J. Gerhard, W. Marquardt, M. Mnnigmann, Numerical solution approaches for robust nonlinear optimal control problems, Computers & Chemical Engineering 32 (6) (2008) 1279 – 1292. doi:http://dx.doi.org/10.1016/j.compchemeng.2007.06.002.

[35] V. Prasad, B. Bequette, Nonlinear system identification and model reduction using artificial neural networks, Computers & Chemical Engineering 27 (12) (2003) 1741 – 1754. doi:http://dx.doi.org/10.1016/S0098-1354(03)00137-6.

[36] P. Li, H. P. Hoo, G. Wozny, Efficient simulation of batch distillation processes by using orthogonal collocation, Chemical Engineering & Technology 21 (11) (1998) 853–862. doi:10.1002/(SICI)1521-4125(199811)21:11<853::AID-CEAT853>3.0.CO;2-2.

[37] S. Jain, J.-K. Kim, R. Smith, Operational optimization of batch distillation systems, Industrial & Engineering Chemistry Research 51 (16) (2012) 5749–5761. doi:10.1021/ie201844g.

[38] M. Leipold, S. Gruetzmann, G. Fieg, An evolutionary approach for multi-objective dynamic optimization applied to middle vessel batch distillation, Computers & Chemical Engineering 33 (4) (2009) 857 – 870. doi:http://dx.doi.org/10.1016/j.compchemeng.2008.12.010.

[39] L. Biegler, X. Yang, G. Fischer, Advances in sensitivity-based nonlinear model predictive control and dynamic real-time optimization, Journal of Process Control (2015) –doi:http://dx.doi.org/10.1016/j.jprocont.2015.02.001.

[40] A. Bonsfills, L. Puigjaner, Batch distillation: simulation and experimental validation, Chemical Engineering and Processing: Process Intensification 43 (10) (2004) 1239 – 1252. doi:10.1016/j.cep.2003.11.009.

[41] M. Barolo, F. Botteon, Calculation of parametric sensitivity in binary batch distillation, Chemical Engineering Science 53 (10) (1998) 1819 – 1834. doi:http://dx.doi.org/10.1016/S0009-2509(98)00017-7.

[42] A. O. Converse, G. D. Gross, Optimal distillate-rate policy in batch distillation, Industrial & Engineering Chemistry Fundamentals 2 (3) (1963) 217–221. doi:10.1021/i160007a010.

[43] M. M. Lopes, T. W. Song, Batch distillation: Better at constant or variable reflux?, Chemical Engineering and Processing: Process Intensification 49 (12) (2010) 1298 – 1304. doi:http://dx.doi.org/10.1016/j.cep.2010.09.019.

[44] J. V. Beck, K. J. Arnold, Parameter estimation in engineering and science, James Beck, 1977.

[45] S. Wu, K. A. McLean, T. J. Harris, K. B. McAuley, Selection of optimal parameter set using estimability analysis and MSE-based model-selection criterion, International Journal of Advanced Mechatronic Systems 3 (3) (2011) 188–197. doi:10.1504/IJAMECHS.2011.042615.

[46] B. Srinivasan, D. Bonvin, E. Visser, S. Palanki, Dynamic optimization of batch processes: Ii. role of measurements in handling uncertainty, Computers & Chemical Engineering 27 (1) (2003) 27–44.

[47] S. M. Safdarnejad, J. D. Hedengren, L. L. Baxter, L. Kennington, Investigating the impact of cryogenic carbon capture on the performance of power plants, in: Proceedings of the American Control Conference (ACC), Chicago, IL, 2015, pp. 5016–5021. doi:10.1109/ACC.2015.7172120.

[48] K. M. Powell, J. D. Hedengren, T. F. Edgar, Dynamic optimization of a hybrid solar thermal and fossil fuel system, Solar Energy 108 (2014) 210 – 218. doi:http://dx.doi.org/10.1016/j.solener.2014.07.004.

[49] K. M. Powell, J. D. Hedengren, T. F. Edgar, Dynamic optimization of a solar thermal energy storage system over a 24 hour period using weather forecasts, in: American Control Conference (ACC), 2013, IEEE, 2013, pp. 2946–2951.

[50] S. M. Safdarnejad, J. D. Hedengren, N. R. Lewis, E. L. Haseltine, Initialization strategies for optimization of dynamic systems, Computers & Chemical Engineering 78 (2015) 39 – 50. doi:10.1016/j.compchemeng.2015.04.016.

[51] J. D. Hedengren, APMonitor Modeling Language, http://www.apmonitor.com, (accessed November 2015).

[52] B. J. Spivey, J. D. Hedengren, T. F. Edgar, Constrained nonlinear estimation for industrial process fouling, Industrial & Engineering Chemistry Research 49 (17) (2010) 7824–7831.

[53] R. Asgharzadeh Shishavan, C. Hubbell, H. Perez, J. Hedengren, D. Pixton, Combined rate of penetration and pressure regulation for drilling optimization by use of high-speed telemetry, SPE Drilling & Completion (SPE-170275-PA). doi:10.2118/170275-PA.

[54] J. D. Hedengren, R. A. Shishavan, K. M. Powell, T. F. Edgar, Nonlinear modeling, estimation and predictive control in APMonitor, Computers & Chemical Engineering 70 (2014) 133 – 148. doi:10.1016/j.compchemeng.2014.04.013.

[55] G. Seber, C. Wild, Nonlinear Regression, John Wiley & Sons, Inc., Hoboken, New Jersey.

[56] N. R. Lewis, J. D. Hedengren, E. L. Haseltine, Hybrid dynamic optimization methods for systems biology with efficient sensitivities, Processes 3 (3) (2015) 701. doi:10.3390/pr3030701.
URL http://www.mdpi.com/2227-9717/3/3/701

[57] Z. K. Nagy, R. D. Braatz, Worst-case and distributional robustness analysis of finite-time control trajectories for nonlinear distributed parameter systems, Control Systems Technology, IEEE Transactions on 11 (5) (2003) 694–704.

[58] P. N. Brown, A. C. Hindmarsh, L. R. Petzold, Using krylov methods in the solution of large-scale differential-algebraic systems, SIAM Journal on Scientific Computing 15 (6) (1994) 1467–1488.

[59] Y. Cao, S. Li, L. Petzold, R. Serban, Adjoint sensitivity analysis for differential-algebraic equations: The adjoint dae system and its numerical solution, SIAM Journal on Scientific Computing 24 (3) (2003) 1076–1089.

[60] J. Renfro, A. Morshedi, O. Asbjornsen, Simultaneous optimization and solution of systems described by differential/algebraic equations, Computers and Chemical Engineering 11 (5) (1987) 503 – 517. doi:10.1016/0098-1354(87)80025-X.

[61] L. Sun, J. D. Hedengren, R. W. Beard, Optimal trajectory generation using model predictive control for aerially towed cable systems, Journal of Guidance, Control, and Dynamics 37 (2) (2014) 525–539. doi:10.2514/1.60820.

[62] J. Tolsma, P. I. Barton, DAEPACK: An open modeling environment for legacy models, Industrial & Engineering Chemistry Research 39 (6) (2000) 1826–1839.

[63] J. T. Leppävuori, M. M. Domach, L. T. Biegler, Parameter estimation in batch bioreactor simulation using metabolic models: Sequential solution with direct sensitivities, Industrial & Engineering Chemistry Research 50 (21) (2011) 12080–12091. doi:10.1021/ie201020g.

[64] S. M. Safdarnejad, J. D. Hedengren, L. L. Baxter, Plant-level dynamic optimization of cryogenic carbon capture with conventional and renewable power sources, Applied Energy 149 (2015) 354 – 366. doi:http://dx.doi.org/10.1016/j.apenergy.2015.03.100.

[65] Chemcad version 6.5.0, owned by Chemstations, Inc. in Houston, TX. Copyright 2015.

[66] R. L. Rowley, W. V. Wilding, J. L. Oscarson, T. A. Knotts, N. F. Giles, DIPPR data compilation of pure chemical properties, Design Institute for Physical Properties, AIChE, New York, NY, 2013.

[67] J. D. Seader, E. J. Henley, Separation Process Principles, John Wiley & Sons, 2006.

[68] J. D. Hedengren, A. N. Eaton, Overview of estimation methods for industrial dynamic systems, Optimization and Engineeringdoi:10.1007/s11081-015-9295-9.

[69] J. D. Hedengren, APMonitor Modeling Language: Batch Distillation Files, https://github.com/APMonitor/batch_distillation, (accessed November 2015).