Model-based Design of Experiments for Polyether Production from Bio-based 1,3-Propanediol

Anh-Duong Dieu Vo [a]; Ali Shahmohammadi [b]; Kimberley B. McAuley [c*]

[a,c] Department of Chemical Engineering, Queen's University, Kingston, ON K7L 3N6, Canada

[b] McKetta Department of Chemical Engineering, University of Texas at Austin, Austin, TX 78712

Abstract

Sequential model-based design of experiments (MBDOE) is used to select operating conditions for new experiments in a batch-reactor that produces bio-based poly(trimethylene) ether glycol (PO3G). These Bayesian A-optimal experiments are designed to obtain improved estimates of the 70 fundamental-model parameter estimates, while accounting for the model structure and for data from eight previous industrial batch-reactor runs. Settings are selected for three decision variables: reactor temperature, initial catalyst level, and initial water concentration. If only one new experiment is conducted, it should be run at high temperature, with relatively high concentrations of catalyst and initial water. When two new runs are conducted, one should use an intermediate catalyst concentration. The effectiveness of the proposed MBDOE approach is tested using Monte-Carlo simulations, revealing that the selected experiments are superior compared to new experiments selected randomly from corners of the permissible design space.

Keywords: model-based design of experiments , sequential experimental design, polymerization model, parameter estimation, PO3G

1. Introduction

Cerenol® is a class of bio-based poly(trimethylene) ether glycol (PO3G) that has been manufactured and trademarked by Dupont since 2008.[1 - 3] The monomer used for producing Cerenol® is corn-based 1,3-propanediol derived from a glucose fermentation process (Bio-PDO®),[4] so that Cerenol® is thoroughly renewably-sourced. Cerenol® offers a variety of value-added properties (e.g., excellent biodegradability, low toxicity, high oxidative stability, etc.), and is used in a wide range of applications including automotive coatings, cosmetics, elastic fibers, and thermoplastic elastomers.[1 - 3, 5]

Cerenol® is produced in batch, semi-continuous and continuous reactors using super-acid catalyst. Operating conditions such as catalyst concentration and temperature vary depending on the desired molecular weight and the end-use properties of the final products.[3, 6 – 14] The influences of these conditions on product properties and production rates have been studied via several fundamental PO3G models.[15 - 22] Mueller *et al*. developed first PO3G models using a reaction scheme that accounts for the influence of acids on polymerization of 1,3-propanediol.[15, 16] Their models predict time-varying concentrations of monomer and water in the reactor liquid and in the vapor that evaporates from the reactor, but they ignore the formation and evaporation of oligomers. Values of kinetic parameters are not reported in their publications. Cui *et al*. extended Mueller's mechanism to account for side reactions (e.g., formation of propanal and transetherification reations), and developed a series of PO3G models that account for formation and evaporation of linear oligomers.[17 - 19] Their studies used estimability analysis techniques[23, 24] to rank the kinetic and transport parameters in their models from most estimable to least-estimable. Wu's mean-squared-error criterion [25, 26] was then used to select the parameters that are estimable from the available industrial data. Cui's parameter estimation study relied on data[3] obtained from four

batch-reactor runs conducted at 180℃ with different super-acid catalyst levels. Cui's final PO3G model and parameter estimates give a reasonable fit to the industrial data. However, their predicted overall rate of polycondensation is too low, especially during the final stage of the batch reactor experiments, resulting in a poor fit to data obtained at long reaction times.

Recently, we updated Cui's reaction scheme by removing some negligible side reactions (i.e., a reaction involving saturated end-groups and a reversible reaction involving protonation of ether links).[20 - 22] The PO3G model was then extended through multiple steps to account for: *i*) the dynamic behaviour of the overhead condenser, *ii*) revised assumptions about mass-transfer rates, *iii*) the inhibitory influence of water on polycondensation kinetics, *iv*) the generation and evaporation of cyclic oligomers, and *v*) the effects of temperature on kinetics and mass-transfer. Our most-recent PO3G model contains a total of 49 ordinary differential equations (ODEs) and 70 kinetic, transport and thermodynamic parameters.[22] We use industrial data set[3] from eight experimental runs conducted at temperatures ranging from 160 to 180℃ and super-acid catalyst levels ranging from 0.1 to 0.25 wt% to estimate parameters and assess the predictive ability of the model. We determined that 68 of a total of 70 parameters were estimable from the available data. The resulting PO3G model and parameter estimates provide a good fit to the data, but confidence intervals for many of the model parameters are very wide. [22]

Our PO3G modeling studies provide an enhanced understanding of the influences of process operating conditions on product properties and polymerization rates in Cerenol® production. However, because the available data are limited, we were not able to obtain accurate estimates of all of the kinetic, transport, and thermodynamic parameters in the models. The objective of the current study is to select operating conditions for a few additional experiments that would be helpful in achieving improved parameter estimates and more-reliable model predictions.

Conducting experimental runs can be expensive and time-consuming. To ensure that as much information as possible can be obtained from the new experiments, we use sequential model-based design of experiments (MBDOE) techniques.

MBDOE takes into account the structure of the model when selecting new experimental conditions to minimize uncertainties in parameter estimates or model predictions. [27, 28] *Sequential* MDBOE methods are particularly useful in industrial settings because they account for available data from previous experiments when selecting conditions for new experiments that will also be used for parameter estimation.[29 - 31] Common MBDOE calculations (e.g., A-, D- and E-optimal designs) rely on the inverse of the Fisher Information Matrix (FIM).[32 - 34] When models are linear in the parameters, the FIM depends only on the experimental settings and measurement uncertainties, and is independent of the parameter values.[35, 36] For nonlinear models, computation of the FIM relies on linearization of the model equations around estimated or assumed parameter values.[32, 36 – 38] As a result, the elements of the FIM depend on these parameter values.

A complication that is often encountered during sequential MBDOE for complex models, such as the PO3G model in the current article, is that the FIM may be noninvertible. A noninvertible FIM arises when there is insufficient information in the data to estimate all of the model parameters uniquely. Several approaches have been considered to overcome this problem, including parameter-subset-selection-based methods, where the unestimable parameters are left out of the MBDOE analysis.[39, 40] In this way, new experiments can be designed to improve the accuracy of the estimable parameters. An alternative approach is to use Bayesian MBDOE methods that rely on modeler-supplied prior information about some or all of the model parameters.[41 - 43] This information may be sufficient to make the augmented Bayesian FIM invertible. Shahmohammadi et al. recently proposed a simple Bayesian approach, which uses prior parameter information that

is similar to the information required for orthogonalization-based estimability ranking.[41] A simple A-optimal-design case study, involving a kinetic model for pharmaceutical production, was used to show that Shahmohammadi's approach gave superior parameter estimates compared to a parameter-subset-selection approach. In current study, we use Shahmohammadi's sequential Bayesian MBDOE method to select experimental conditions for several new PO3G batch-reactor runs. The methodology is *sequential* because the proposed new runs are designed based on information contained in old data that have already been used for preliminary parameter estimation. An A-optimal design criterion is used (rather than D- or E- optimal design) because A-optimality focuses on the accuracy of the individual parameter estimates.[41] As a result, it is relatively easy to use Monte Carlo (MC) simulations to test and verify the effectiveness of the proposed methodology.[41]

The remainder of this article is organized as follows. First, old kinetic data from industrial PO3G batch reactor experiments are described. Next, the PO3G model equations are introduced, along with the model parameters that require estimation. Available prior knowledge about these parameters is discussed. Next, the sequential Bayesian A-optimal design process is described and the resulting settings for new experiments are presented. Finally, MC simulations are used to test the effectiveness of the proposed approach.

2. Background

2.1. Kinetic Data

Figure 1 shows the batch-reactor system used by DuPont to conduct the experimental runs that were previously used for model development and parameter estimation.[3, 17 - 22] At the start of each experiment, the desired operating temperature was selected, and monomer (1,3-propanediol) was charged to the reactor, which is sparged with nitrogen to help remove water generated by the main

polycondensation reaction. Each batch reactor run starts when a specified amount of super-acid catalyst (triflic acid) is added to the liquid phase. Over the course of the run, monomer is consumed to produce oligomer and polymer chains, which accumulate in the liquid phase. N_2 bubbles containing volatile by-products of the reactions (e.g., water, propanal, and some oligomers) travel through the liquid to the reactor headspace. The overhead gas flows to a condenser and liquid condensate is collected. During each run, samples of both the liquid in the reactor and the condensate entering the condensate collector are collected at various times. These samples were subsequently analyzed by gas chromatography and proton nuclear magnetic resonance (NMR). The available industrial data set consists of eight experimental runs, which were conducted at five different temperatures between 160 and 180℃ and four levels of super-acid catalyst between 0.10 and 0.25 wt%.[3] **Table 1** shows the experimental settings for these eight runs. No initial water was used in these experiments. **Table 2** summarizes the measurements available from the monomer/polymer liquid phase and from the condensate during each run. In total, 2046 data values were collected from the 8 experimental runs. In the current study, A-optimal settings for new additional experiments will be selected, given that these 2046 prior data values are already available for use in parameter estimation.

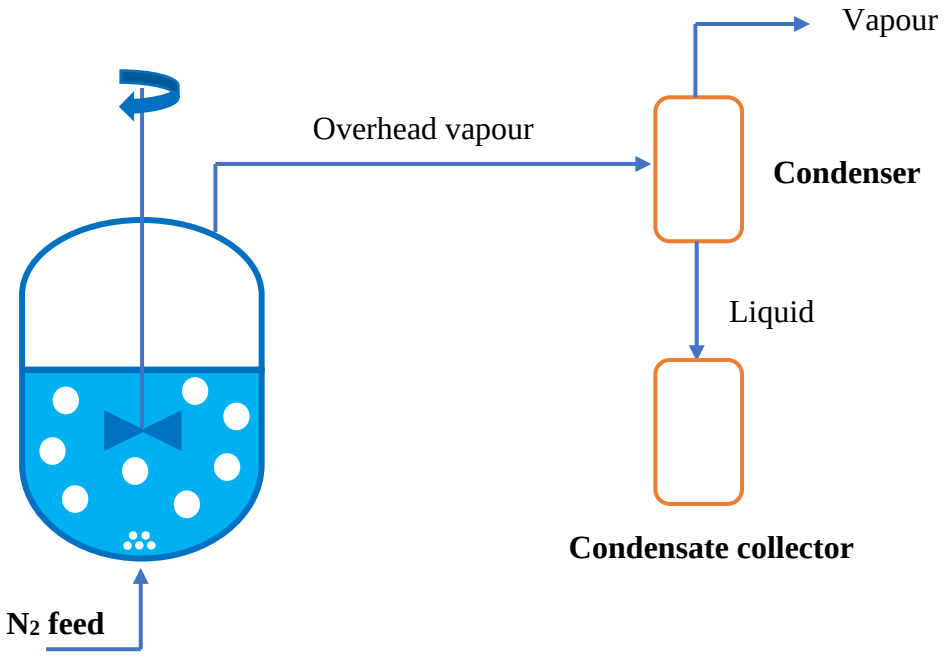


Figure 1. PO3G reactor and condenser system[3, 17 - 22]

Table 1: Xie's experimental settings[3]

Temperature (℃)	180	180	180	180	160	165	170	175
Cat. Concentration (wt%)	0.10	0.15	0.20	0.25	0.15	0.15	0.15	0.15
Initial Water Concentration (wt%)	0	0	0	0	0	0	0	0

Table 2: Available data from Xie's comprehensive data set[3]

	Liquid phase	Condensate
Concentration of monomer (ppm by weight)	✓	✓
Concentration of oligomers (ppm by weight)	✓	✓
Concentration of water (wt%)	✓	✓
Concentration of unsaturated ends (mmol/kg)	✓	
Total mass of condensate accumulated in the condensate collector (g)		✓

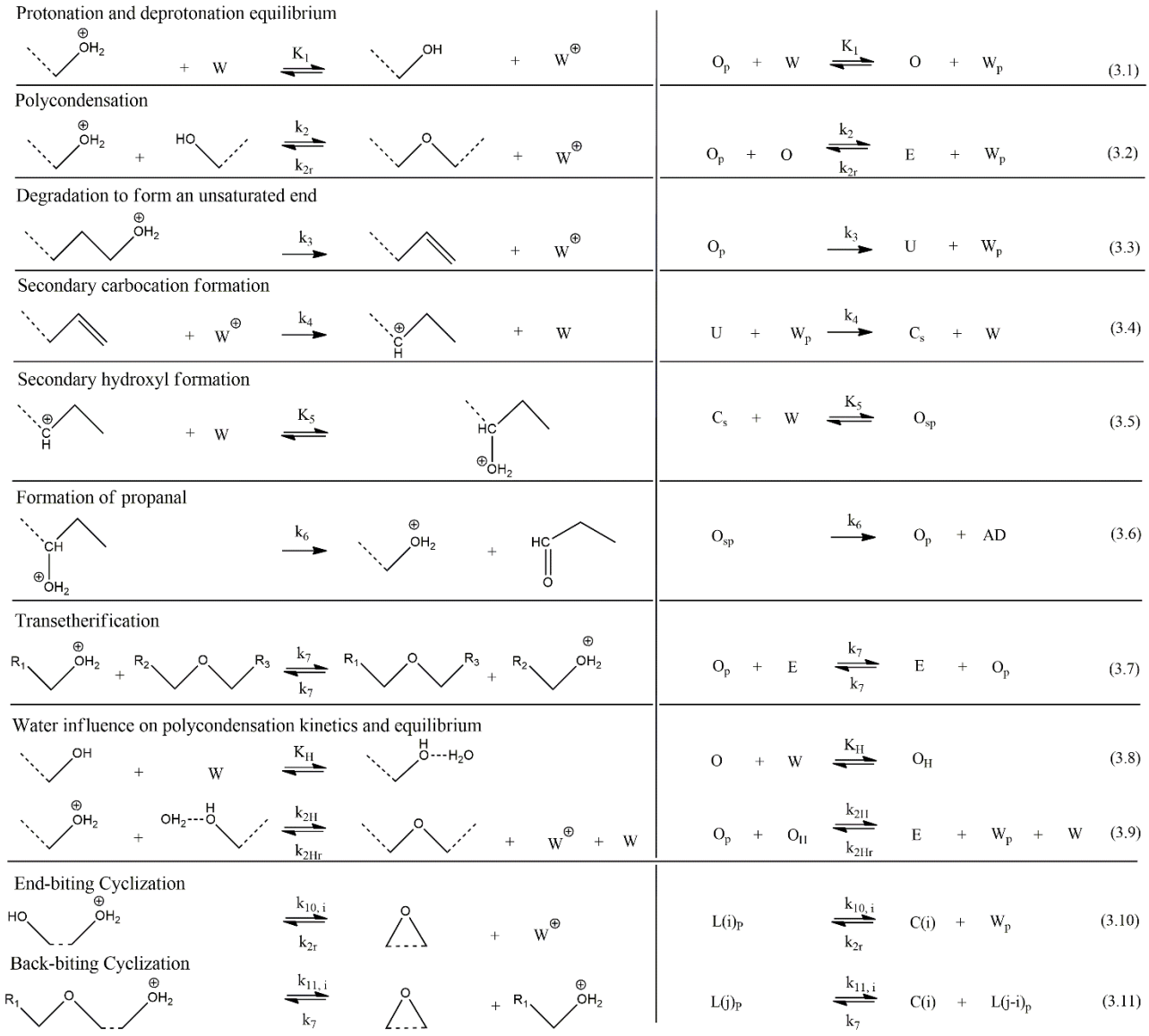
2.2. Model Equations and Model Parameters

The PO3G model used for designing new experiments in the current article was developed based on the reaction mechanism in **Table 3**.[21, 22] The mechanism includes eleven reactions, which are shown in two ways. The column on the left in Table 3 provides structural information for the molecules participating in the reactions, focusing on polymer end groups and ether links. The same reactions are presented in the column on the right using symbols that appear in the model equations. Hydroxyl ends are shown as *O*, ether links as *E*, unsaturated ends as *U*, secondary carbocation ends as C_s, water as *W* and propanal as *AD*. A linear chain with *i* repeat units is denoted as *L*(*i*). A cyclic oligomer with *i* repeat units is denoted as *C*(*i*). Protonated functional groups and protonated molecules are indicated using the subscript *p*.

In Table 3, reaction (3.1) shows the equilibrium for protons exchanging between hydroxyl ends and water with the equilibrium constant K_1. Reaction (3.2) is the main polycondensation reaction, in which hydroxyl ends are consumed with forward rate constant k_2 (kg mmol^{-1} h^{-1}). The reverse of reaction (3.2) is a hydrolysis reaction with rate constant k_{2r} (kg mmol^{-1} h^{-1}). Reaction (3.3) is a side reaction in which unsaturated ends are produced from degradation of hydroxyl ends (with rate constant k_3 (h^{-1})). In reaction (3.4) secondary carbocation end groups are produced from protonation of unsaturated ends with rate constant k_4 (kg mmol^{-1} h^{-1}). Reactions (3.5) and (3.6) account for formation of propanal with a lumped rate constant K_5k_6 (h^{-1}). This lumped rate constant is used because protonated secondary hydroxyl end O_{sp} is treated as a short-lived intermediate that does not accumulate appreciably in the reacting mixture. Reaction (3.7) accounts for transetherification between oligomeric and polymer chains with rate constant k_7 (kg mmol^{-1} h^{-1}). Transetherification reactions determine the rate at which linear oligomers are regenerated from long polymer chains to replenish the oligomer molecules that evaporate from the reaction mixture.

Reactions (3.8) and (3.9) account for the inhibitory influence of water on polycondensation kinetics and equilibrium. Hydrated hydroxyl ends O_H are assumed to be in equilibrium with regular hydroxyl ends O, with equilibrium constant K_H. Polycondensation involving hydrated hydroxyl ends, with forward rate constant k_{2H} and reverse rate constant k_{2Hr}, has a slower reaction rate compared to the main reaction (3.2). Reactions (3.10) and (3.11) account for the formation of cyclic oligomers via end-biting and back-biting reactions, respectively. End-biting reactions that produce rings with i repeat units have forward rate constant $k_{10,\,i}$ (h^{-1}). Similarly, back-biting reactions that generate rings with i repeat units have forward rate constant $k_{11,\,i}$ (h^{-1}). The reverse of reaction (3.10) is a hydrolysis reaction, which is assumed to have the same rate constant k_{2r} as the reverse of the main polycondensation reaction (3.2). The reverse of reaction (3.11) is a ring-opening polyaddition reaction, which contributes to the growth of polymer chains and is assumed to have the same rate constant k_7 as the transetherification reaction.[21, 22]

Table 3. Mechanism for PO3G production in batch reactor[21,22]



Our most-recent PO3G model, which was developed to account for the influence of temperature on the reactions in Table 3, has a total of 49 ODEs and 70 kinetic, transport, and thermodynamic parameters.[22] **Table 4** shows three of the 49 ODEs. The first is a material balance on the mass *m* (kg) of the reaction mixture, the second is a material balance on unreacted monomer *L(1)* in the reactor, and the third is a material balance on the monomer in the overhead condenser $L(1)_c$. The

remaining 46 ODEs, along with a complete list of model equations and all assumptions that were used when developing the model, are provided in **Tables S1** to **S5** in the Supplementary Information and in a previous publication.[22] Equation (4.1) describes dynamic changes in m due to evaporation of volatile species such as water, propanal, monomer and oligomers into nitrogen bubbles that pass through the liquid phase and exit into the headspace. Mass-transfer coefficients of these species are k_W, k_{AD}, $k_{L(i)}$ with i=1..5 and $k_{C(i)}$ with i=2..7 (in kg m^{-2} h^{-1}). The interfacial area A_b (m^2 kg^{-1}) of the nitrogen bubbles per unit mass of reactor liquid is calculated using a correlation that accounts for changes in liquid viscosity as the polymerization proceeds.[20 - 22] In the model, all species and end-group concentrations have units of mmol per kg of liquid. Hypothetical concentrations of water, propanal, monomer and oligomers that would be in equilibrium with their average partial pressures in the bubbles (i.e., $[W]^*$, $[AD]^*$, $[L(i)]^*$ and $[C(i)]^*$), are obtained from temperature-dependent VLE correlations.[20 - 22] Note that because the concentrations of propanal and oligomers in the nitrogen bubbles are assumed to be very low, and no VLE data are available for these species, a value of zero is used for $[AD]^*$, $[L(i)]^*$ with i=2..5 and $[C(i)]^*$ with i=2..7. Uncertainty about whether the nitrogen bubbles tend to recirculate in the liquid due to stirring or they travel mostly upward through the liquid phase is accounted by an adjustable bubble-backmixing parameter γ (i.e., γ=1.0 corresponds to bubbles that are well-mixed within the liquid and $\gamma \approx 0.5$ corresponds to bubbles that rise without significant recirculation). M_W, M_{AD}, $M_{L(i)}$ and $M_{C(i)}$ are molar masses of the corresponding species (in kg $mmol^{-1}$).

Equation (4.2) is a dynamic material balance on the monomer. The first term on the right-hand side accounts for consumption of monomer by polycondensation. In equation (4.2), factor f_1 accounts for the probability that a hydroxyl end in the reaction mixture is protonated. The overall rate constant k_2^W (which depends on k_2, K_H and $[W]$) accounts for the inhibitory influence of water

on polycondensation kinetics. Algebraic expressions for k_2^W and for the equilibrium constant K_2^W are provided in Table S3 in the Supplementary Information. The second term in equation (4.2) accounts for generation of monomer via hydrolysis (reverse of reaction (3.2)). The factor f_2 is used to account for the probability of a water molecule in the reaction mixture being protonated. Expressions for f_1, f_2 and other probability factors that appear in the ODEs are provided in Table S3 in the Supplementary Information. The third and fourth terms in equation (4.2) account for consumption and regeneration of monomer via transetherification (reaction (3.7)). The fifth term accounts for the consumption of monomer via ring-opening reactions (reverse of reaction (3.11)). The sixth term accounts for generation of monomer via back-biting reactions of oligomer chains (reaction (3.11)). The seventh term accounts for removal of water from the liquid phase by evaporation. The final term accounts for the change in concentration of monomer associated with the reduction in total mass of the reaction mixture.

Equation (4.3) describes the time-varying concentration of monomer in the overhead condenser, $[L(1)_c]$. In the model, the condenser, which was operated at 10℃ and 1 atm, is treated as an instantaneous equilibrium flash combined with a well-mixed tank to account for the liquid accumulation within the condenser. In equation (4.3), $F_{gtot}L_f$ is the flow rate of the liquid flowing from the instantaneous flash into the tank, and $x_{L(1)}$ is the mole fraction of monomer in this liquid stream. The average molecular weight of this liquid is M_{L_f}(kg mmol^{-1}). Algebraic expressions for F_{gtot}, L_f, $x_{L(1)}$, M_{L_f}, and other terms that appear in the condenser model are provided in the Supplementary Information. In equation (4.3), the total mass of liquid that accumulates in the condenser, m_{Lc} (kg) is assumed to be constant and is treated as a model parameter. The model includes additional 33 ODEs that describe time varying concentrations within the reactor (i.e., for water [*W*], propanal [*AD*], linear oligomers [*L*(*i*)], cyclic oligomers [*C*(*i*)] , catalyst [*cat*], hydroxyl

ends [O], unsaturated ends [U] and secondary carbocation ends [C_s], along with the zeroth and first moments of the total chain-length distribution for linear species).[22] Additional 13 ODEs describe time-varying concentrations of water [W_c], propanal [AD], and volatile oligomers ([$L(i)_c$] and [$C(i)_c$]) in the condenser liquid, along with total mass of liquid that has accumulated in the condensate collector.[22]

Table 4: Representative differential equations in the PO3G model

$$\frac{dm}{dt} = -mk_W A_b([W]-\gamma[W]^*)M_W - mk_{AD}A_b([AD]-\gamma[AD]^*)M_{AD} - m\sum_{i=1}^{5} k_{L(i)}A_b([L(i)]-\gamma[L(i)]^*)M_{L(i)} - m\sum_{i=2}^{7} k_{C(i)}A_b([C(i)]-\gamma[C(i)]^*)M_{C(i)}$$	(4.1)
$$\frac{d[L(1)]}{dt} = -4k_2^W f_1[L(1)][O] + k_{2r}f_{E1}[E]f_2[W] - k_7 f_1 2[L(1)][E] + \frac{1}{2}k_7 f_1[O]f_{E1}[E] - k_7 2f_1[L(1)]\sum_{i=2}^{7} iC(i) + \sum_{i=3}^{7} k_{11,i-1}2f_1[L(i)] - k_{L(1)}A_b([L(1)]-\gamma[L(1)]^*) - \frac{[L(1)]}{m}\frac{dm}{dt}$$	(4.2)
$$\frac{d[L(1)_c]}{dt} = (F_{gtot}L_f x_{L(1)} - F_{gtot}L_f M_{L_f}[L(1)_c])\frac{1}{m_{Lc}}$$	(4.3)

In total, the PO3G model has70 kinetic, transport, and thermodynamic parameters. **Table 5** shows a list of the kinetic parameters that appear in the reaction mechanism in Table 3 (including activation energies and reaction enthalpies, which are used to account for the influence of temperature). **Table 6** shows a list of the mass-transfer parameters along with their activation energies. Tables 5 and 6 show values of the parameter estimates obtained in our previous parameter estimation study using the available industrial data, along with their estimated standard deviation

(SD). Estimates of the rate constants, equilibrium constants and mass-transfer coefficients shown in Tables 5 and 6 correspond to a reference temperature of 180 °C. Also included in Table 6 are VLE parameters that are used to calculate $[W]^*$,[22] along with the bubble back-mixing parameter γ and the mass of liquid in the condenser m_{Lc}. Using the available data we were able to estimate 68 out of the 70 parameters in our recent parameter estimation study. The two parameters that could not be estimated and were left fixed at their initial guesses are $E_{k_{AD}}$ and $f_{P_W^{185}(0.02)}$ (values shown in *italic*).

Equation (1) shows the objective function that was used for parameter estimation. The first five terms on the right-hand side account for deviations between measured and predicted concentrations of monomer, linear oligomers, cyclic oligomers, water, propanal, and unsaturated ends in the reactor liquid. The next four terms account for deviations between measured and predicted concentrations of monomer, linear oligomer, cyclic oligomers, water, and propanal in the condensate. The final term is related to the mass of liquid accumulated in the condensate collector. Measured values are denoted by subscript m. The weighting factors $s_{L\,ppm}$ (ppm), $s_{C\,ppm}$ (ppm), s_W (wt%), $s_{y\,U}$ (mmol/kg), $s_{Lc\,ppm}$ (ppm), $s_{Cc\,ppm}$ (ppm), s_{Wc} (wt%), and $s_{m_{cc}}$ (g) account for uncertainties in measurements of the corresponding species. Values of these weighting factors are provided in **Table 7**. **Table 8** shows initial guesses and the lower and upper bounds used during parameter estimation for 6 of the 70 parameters. The lower and upper bounds for all of the parameters, along with their initial guesses, are provided in the Supplementary Information. As shown in Tables 5 and 6, only 36 parameter estimates obtained using the available data are significantly different from zero (values shown in **bold**), based on their approximate 95% confidence levels. An important objective of the current study is to use a simplified Bayesian MBDOE method to select conditions for new experiments. These experiments will be selected to

provide narrower confidence intervals for the parameters that have already been estimated and to estimate the two parameters that are currently unestimable.

$$J = \frac{1}{s^2_{L\,ppm}} \sum_{i=1}^{7} \sum ([L(i)]_m - [L(i)])^2 + \frac{1}{s^2_{C\,ppm}} \sum_{i=2}^{7} \sum ([C(i)]_m - [C(i)])^2$$
$$+ \frac{1}{s^2_W} \sum ([W]_m - [W])^2 + \frac{1}{s^2_{L\,ppm}} \sum ([AD]_m - [AD])^2$$
$$+ \frac{1}{s^2_{y\,U}} \sum ([U]_m - [U])^2 + \frac{1}{s^2_{Lc\,ppm}} \sum_{i=1}^{5} \sum ([L(i)_c]_m - [L(i)_c])^2$$
$$+ \frac{1}{s^2_{Cc\,ppm}} \sum_{i=2}^{7} \sum ([C(i)_c]_m - [C(i)_c])^2 + \frac{1}{s^2_{Wc}} \sum ([W_c]_m - [W_c])^2$$
$$+ \frac{1}{s^2_{Lc\,ppm}} \sum ([AD_c]_m - [AD_c])^2 + \frac{1}{s^2_{m_{cc}}} \sum (m_{cc_m} - m_{cc})^2 \quad (1)$$

Table 5: Kinetic parameters in PO3G model

Parameter	**Units**	**Estimate from previous study** [22]	**SD**	**Parameter**	**Units**	**Estimate from previous study** [22]	**SD**
K_1	-	1.014	27.458	ΔH_{K_1}	kJ mol^{-1}	**-79.460**	32.061
k_2	kg mmol^{-1} h^{-1}	0.053	1.370	E_{k_2}	kJ mol^{-1}	**115.340**	2.666
$\frac{K_{2\,app}}{K_1}$	-	0.500	1.690	$\Delta H_{\frac{K_{2\,app}}{K_1}}$	kJ mol^{-1}	**-157.663**	4.532
k_3	h^{-1}	0.933	8.274	E_{k_3}	kJ mol^{-1}	**122.844**	18.605
k_4	kg mmol^{-1}h^{-1}	412.889	2593.823	E_{k_4}	kJ mol^{-1}	451.338	362.957
K_5k_6	h^{-1}	122.637	1.67×105	$\Delta H_{K_5k_6}$	kJ mol^{-1}	189.364	1.78×10^5
k_7	kg·mmol^{-1}·h^{-1}	1.222	1.714	E_{k_7}	kJ mol^{-1}	**408.458**	0.810
K_H	mmol^{-1} kg	0.017	1.699	ΔH_{K_H}	kJ mol^{-1}	-17.793	18.713
$k_{10,\ 2}$	h^{-1}	37.416	2.985	$E_{k_{10,\ 2}}$	kJ mol^{-1}	**747.092**	1.844

$f_{k_{10,\ 3}}$	-	2.975	2.782	$E_{k_{10,\ 3}}$	kJ mol^{-1}	**408.809**	0.881
$f_{k_{10,\ 4}}$	-	2.319	3.167	$E_{k_{10,\ 4}}$	kJ mol^{-1}	**1606.830**	25.044
$f_{k_{10,\ 5}}$	-	1.658	4.026	$E_{k_{10,\ 5}}$	kJ mol^{-1}	**2565.871**	504.303
$f_{k_{10,\ 6}}$	-	1.445	4.517	$E_{k_{10,\ 6}}$	kJ mol^{-1}	**2240.207**	243.956
$f_{k_{10,\ 7}}$	-	1.230	5.202	$E_{k_{10,\ 7}}$	kJ mol^{-1}	**1620.730**	60.087
$k_{11,\ 2}$	h^{-1}	1.354	3.052	$E_{k_{11,\ 2}}$	kJ mol^{-1}	**299.089**	1.094
$f_{k_{11,\ 3}}$	-	3.958	2.703	$E_{k_{11,\ 3}}$	kJ mol^{-1}	**388.655**	1.521
$f_{k_{11,\ 4}}$	-	**31.171**	2.885	$E_{k_{11,\ 4}}$	kJ mol^{-1}	**350.279**	0.889
$f_{k_{11,\ 5}}$	-	**18.533**	3.083	$E_{k_{11,\ 5}}$	kJ mol^{-1}	**346.575**	0.939
$f_{k_{11,\ 6}}$	-	**11.737**	3.258	$E_{k_{11,\ 6}}$	kJ mol^{-1}	**331.612**	1.024
$f_{k_{11,\ 7}}$	-	6.955	3.691	$E_{k_{11,\ 7}}$	kJ mol^{-1}	**314.906**	1.226

Table 6: Mass-transfer parameters and their activation energies

Parameter	**Units**	**Estimate from previous study** [22]	**SD**	**Parameter**	**Units**	**Estimate from previous study** [22]	**SD**
γ	-	0.207	2.278	m_{Lc}	kg	0.020	10.506
k_W	kg m^{-2} h^{-1}	**1540.990**	0.241	E_{k_W}	kJ mol^{-1}	**86.469**	0.417
k_{AD}	kg m^{-2} h^{-1}	**2.109×10^{6}**	2.03×10^{5}	$E_{k_{AD}}$	kJ mol^{-1}	*41.84*	-
$k_{L(1)}$	kg m^{-2} h^{-1}	**14.547**	0.664	$E_{k_{L1}}$	kJ mol^{-1}	**192.280**	0.324
$k_{L(2)}$	kg m^{-2} h^{-1}	1.516	1.265	$E_{k_{L2}}$	kJ mol^{-1}	**119.491**	1.586
$k_{L(3)}$	kg m^{-2} h^{-1}	0.472	4.265	$E_{k_{L3}}$	kJ mol^{-1}	**131.420**	5.602
$k_{L(4)}$	kg m^{-2} h^{-1}	0.379	5.404	$E_{k_{L4}}$	kJ mol^{-1}	**262.627**	12.608
$k_{L(5)}$	kg m^{-2} h^{-1}	0.238	8.697	$E_{k_{L5}}$	kJ mol^{-1}	**452.796**	35.580
$k_{C(2)}$	kg m^{-2} h^{-1}	**138.207**	2.653	$E_{k_{C2}}$	kJ mol^{-1}	13.971	44.506
$k_{C(3)}$	kg m^{-2} h^{-1}	**43.126**	1.092	$E_{k_{C3}}$	kJ mol^{-1}	**90.064**	1.902
$k_{C(4)}$	kg m^{-2} h^{-1}	**9.301**	1.124	$E_{k_{C4}}$	kJ mol^{-1}	**57.180**	2.279
$k_{C(5)}$	kg m^{-2} h^{-1}	0.586	2.950	$E_{k_{C5}}$	kJ mol^{-1}	6.506	61.627
$k_{C(6)}$	kg m^{-2} h^{-1}	0.363	5.847	$E_{k_{C6}}$	kJ mol^{-1}	**117.757**	18.783

$P_W^{160}(0.0005)$	mmHg	10.007	497.204	$f_{P_W^{185}(0.0005)}$	-	1.504	497.667
$P_W^{160}(0.02)$	mmHg	**1012.427**	284.040	$f_{P_W^{185}(0.02)}$	-	*4*	-

Table 7. Measurement **s**tandard deviations used as weighting factors in the objective function

Weighting factor	**Units**	**Value**
$s_{L\ ppm}$	ppm	2.80×10^{4}
$s_{C\ ppm}$	ppm	3.03×10^{2}
s_W	wt%	2.5×10^{-2}
$s_{Lc\ ppm}$	ppm	8.45×10^{3}
$s_{Cc\ ppm}$	ppm	9.10×10^{2}
s_{Wc}	wt%	3.9
$s_{m_{cc}}$	g	73.64

Table 8: Lower bounds and upper bounds of selected parameters

Parameter	**Lower bound**	**Upper bound**	**Initial guess**
k_2	8×10^{-4}	10	0.02
E_{k_2}	0	2000	200
$\frac{K_{2\ app}}{K_1}$	0.10	10	8.90
$\Delta H_{\frac{K_{2\ app}}{K_1}}$	-250	0	-40
$k_{L(1)}$	0	22500	15.69
$E_{k_{L1}}$	0	2000	40

2.3. Simplified A-Optimal Bayesian MBDOE

Recently, Shahmohamadi and McAuley developed a simplified Bayesian A-optimal objective function for use in sequential MBDOE: [41]

$$J_A = \text{trace}((\mathbf{Z^T Z} + \mathbf{I})^{-1}) \qquad (2)$$

In equation (2), $\mathbf{Z}$ is a $N \times p$ scaled sensitivity matrix, and $\mathbf{I}$ is a $p \times p$ identity matrix, where N is the total number of data values that will be used for parameter estimation and p is the number of model parameters. Elements of $\mathbf{Z}$ are computed using:

$$Z_{ij} = \frac{\partial g_i}{\partial \theta_j} \frac{s_{\theta_j}}{s_{y_i}} \qquad (3)$$

where g_i is the model prediction of the i^{th} data value and θ_j is the j^{th} parameter requiring estimation. The scaling factor s_{θ_j} is a user-supplied standard deviation that reflects the modeler's prior uncertainty about the j^{th} parameter value. The scaling factor s_{y_i} is a user-supplied standard deviation that reflects the uncertainty in the measurement of the i^{th} data value. When designing experiments to improve the parameter estimates in the current model, partial derivatives $\frac{\partial g_i}{\partial \theta_j}$ were obtained using finite difference approximations. Each parameter was perturbed, one at a time, away from its initial parameter value (i.e., by 5% of its initial value). The corresponding change in the value of each model prediction, Δg_i, was determined via simulation and was used to calculate $\frac{\partial g_i}{\partial \theta_j} \approx \frac{\Delta g_i}{\Delta \theta_j}$. In the current study, parameter estimates shown in Tables 5 and 6, which were obtained using industrial data in our previous study,[22] were used to compute $\frac{\Delta g_i}{\Delta \theta_j}$. Scaling factors s_{θ_j} were computed using the parameter bounds shown in Table 8 (and in the Supplementary Information). Prior information about each parameter was specified using a normal distribution whose standard deviation s_{θ_j} is 1/6 of the distance between the corresponding lower and upper bounds.

In equation (2), $\mathbf{Z^T Z} + \mathbf{I}$ is a simplified Bayesian Fisher information matrix, where $\mathbf{Z^T Z}$ corresponds to the information contained in the data and the identity matrix reflects the modeler's

prior knowledge about plausible parameter values.[41] In sequential A-optimal MBDOE calculations, like those in the current article, the scaled sensitivity matrix contains two parts:

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z_{old}} \\ \mathbf{Z_{new}} \end{bmatrix} \qquad (4)$$

In equation (4), the elements of $\mathbf{Z_{old}}$ are computed based on experimental settings and data from past experiments. This top portion of **Z** does not change during MBDOE calculations. The bottom part in $\mathbf{Z_{new}}$ is computed using the proposed settings for new experiments. The elements of $\mathbf{Z_{new}}$ are updated when the optimizer updates its decision variables, which are proposed experimental settings for new reactor run(s). In the current study, $\mathbf{Z_{old}}$ was computed using the experimental conditions shown in Table 1. It has 2046 rows, which correspond to data values in the industrial data set used for preliminary parameter estimation.[22] The elements of $\mathbf{Z_{new}}$ are computed using the decision variables for the proposed new run(s). For example, when one new run is designed $\mathbf{Z_{new}}$ has 580 rows, corresponding to predictions of the 580 data values that would be collected during the proposed new run. When two new runs are designed simultaneously, $\mathbf{Z_{new}}$ has 1160 rows. The elements of $\mathbf{Z_{new}}$ are updated when the optimizer selects updated values for the decision variables (i.e., the reactor temperature T, the initial concentration of catalyst $[Cat]_{0\ wt\%}$, and the initial concentration of water $[W]_{0\ wt\%}$) because these decision variables influence the model predictions g_i used to compute $\mathbf{Z_{new}}$.

3. Selected Settings for New Experiments Obtained using MBDOE

Objective function (2) was used to select operating conditions for one new experiment (and then for two new experiments) aimed at improving the accuracy of the parameter estimates. **Figure 2** shows the design space for the new experiments. Notice that the ranges of permissible temperatures and catalyst concentrations are wider than the corresponding ranges of experimental

settings for the old data reported in Table 1. For all of the experiments in Table 1, only monomer and catalyst were initially added to the reactor, and water was generated over time. The proposed design space considers adding some water, up to 10 wt%, at the start of the new designed experiments. Each new experiment will have a fixed duration of 10 hours, with samples of the reactor liquid and condensate collected for analysis every 0.5 hours.

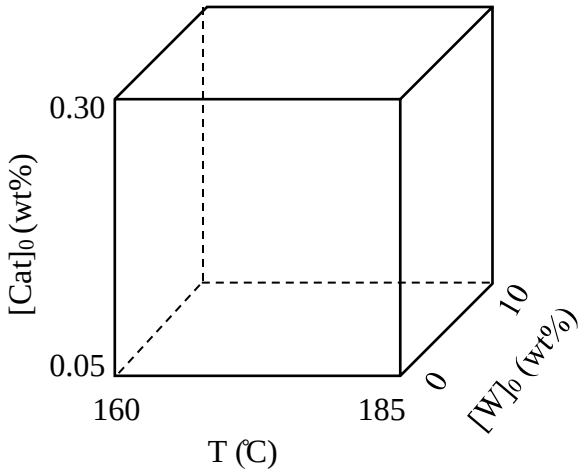


Figure 2. Design space for new experiment(s)

Table 9 shows the selected optimal settings obtained using objective function (2) when either one new experiment or two new experiments are designed. These results suggest that if only one new experiment can be performed, it should be conducted using the highest possible temperature and a high catalyst level, with an initial water concentration near the upper bound. This result makes sense because no data have previously been collected from runs such a high temperature and high catalyst concentration (see Table 1). Furthermore, it makes sense that water is added to the reaction mixture because one of the problems encountered when estimating the model parameters from the old data is that several VLE parameters for water were difficult to estimate due to relatively low water concentrations in all of the runs.[22] Data collected from an experiment conducted with added

initial water should be helpful in estimating some or all of the water VLE parameters and for estimating kinetic parameters for reactions that involve water. Table 9 also shows that, if two new experiments can be performed, both should be conducted at high temperatures with a high initial water concentration. One of the two experiments should use a high catalyst concentration and the other should use a moderate catalyst concentration. To reduce the chances of identifying local minima, the optimization was repeated several times, using different corners of the design space as initial guesses. We confirmed the optimality of the results in Table 9 (and the suitability of numerical tolerances used for the *fmincon* optimizer in MATLAB) by plotting the value of J_A in equation (2) vs. each of the decision variables, with the other decision variables held constant at their converged values. These plots (Figures S1 and S2 in the Supplementary Information) confirm that the results reported in Table 9 are indeed minima.

Table 9: Selected operating conditions for new experiments

No. of new experiments	**Temperature (°C)**	**Catalyst Concentration (wt% on dry basis)**	**Water Concentration (wt%)**
1 Experiment	185.0	0.28	9.10
2 Experiments	183.5	0.18	10.00
	185.0	0.26	9.75

If new experimental conditions were selected in different ways (i.e., using the proposed MBDOE approach and a more conventional approach based on the experimenter's judgment), it would be difficult to use the resulting experiments to assess which settings resulted in the most improvement to the parameter estimates. The reasons for this difficulty are: i) if the new experiments are only performed once, then which experiment(s) provide better results depends somewhat on luck due to the different random measurement errors encountered in different experiments, and ii) the

underlying true values of the parameters would still be unknown. To overcome these issues a MC simulation study has been conducted and is described below.

4. Using MC Simulations to Test the Effectiveness of the Proposed MBDOE Method

In this section, the effectiveness of the proposed A-optimal Bayesian MBDOE method is assessed using MC simulations. Synthetic data were generated for the PO3G batch-reactor system shown in Figure 1, using the PO3G model equations described in section 2.2. Parameter values estimated from industrial data (shown in Tables 5 and 6) were used as true parameter values $\boldsymbol{\theta^{true}}$. The standard deviations in Table 7 were used as true standard deviations for normally-distributed measurement noise included in the synthetic data.

In the current study, eight sets of synthetic old data were generated using the experimental settings in Table 1. For each synthetic old data set, a corresponding set of parameter estimates $\boldsymbol{\hat{\theta}^{old}}$ was obtained using objective function (1). Lower and upper bounds shown in Table S6 of the Supplementary Information were used during parameter estimation. For each set of synthetic old data, the parameters were ranked from most-estimable to least-estimable, and Wu's MSE criterion was used to determine which parameters are estimable, following the same process as in our PO3G parameter-estimation studies.[17 - 22] The resulting eight sets of $\boldsymbol{\hat{\theta}^{old}}$ values are reported in Table S7 in the Supplementary Information. A sum of squared deviations for each set of $\boldsymbol{\hat{\theta}^{old}}$ values was calculated using:

$$SSD_{\theta} = \sum_{j=1}^{70} \left(\frac{\hat{\theta}_j - \theta_j^{true}}{s_{\theta_j}} \right)^2 \qquad (5)$$

where the scaling factor s_{θ_j} was set at 1/6 of the distance between the lower and upper bound for each of the 70 model parameters. Part a) of **Figure 3** shows a plot of the resulting SSD_{θ} values for

the eight sets of $\widehat{\boldsymbol{\theta}}^{\boldsymbol{old}}$. The scatter in the results arises due to the different random-number sequences used when generating the 8 old data sets.

Next, each set of simulated old data (and the corresponding parameter estimates) was used to design one new experimental run using the proposed sequential Bayesian A-optimal MBDOE approach. Objective function (2) was used to select settings for T, $[Cat]_{0\ wt\%}$, and $[W]_{0\ wt\%}$ for a new experiment, as was done using the industrial data in section 3. The resulting eight sets of selected settings are reported in Table S8 in the Supplementary Information. As expected, the resulting experimental settings are similar to, but not exactly the same as, settings in Table 9 obtained using the real data. The highest possible temperature was selected for all eight proposed experiments. A high catalyst level and a high initial water concentration were also selected for most of the eight proposed experiments. These eight sets of experimental settings were then used to generate new synthetic data. Data from each new synthetic run was then combined with the corresponding synthetic old data so that parameters could be re-estimated. The resulting eight sets of parameter estimates $\widehat{\boldsymbol{\theta}}^{\boldsymbol{new}}$ are reported in Table S9 of the Supplementary Information. The SSD_θ values for these eight sets of $\widehat{\boldsymbol{\theta}}^{\boldsymbol{new}}$ were calculated using equation (5) and are shown in part b) of Figure 3. These results reveal that using additional data from one new designed experiment resulted in parameter estimates that are noticeably closer to the true parameter values than when only the synthetic old data was available for parameter estimation. Results in part c) of Figure 3 summarize the improvement in the parameter estimates when two new experiments are designed and simulated based on the eight old synthetic data sets. As expected, the synthetic data from two new experiments resulted in superior parameter values compared with data from only one new synthetic experiment.

Results in parts d) and e) of Figure 3 are shown to compare the effectiveness of the proposed MBDOE approach with what might happen if an experimentalist selected new settings randomly from among the corners of the design space shown in Figure 2. Eight sets of synthetic corner data were generated using the eight corner settings and were combined with the eight sets of synthetic old data. To generate part e) of Figure 3, two randomly-selected corner data sets were combined with each set of synthetic old data. When performing these random assignments, we ensured that each synthetic old data set was paired with two distinct corners and that each corner was paired with two different old data sets. As shown in parts d) and e) of Figure 3, using data from one or two corner experiments improved the accuracy of the parameter estimates compared to when only old data was used for parameter estimation (part a)). However, the results in parts d) and e) are considerably worse than the corresponding results in parts b) and c), indicating the that proposed MBDOE methodology is superior to selecting random corners of the design space for new experiments for this case study.

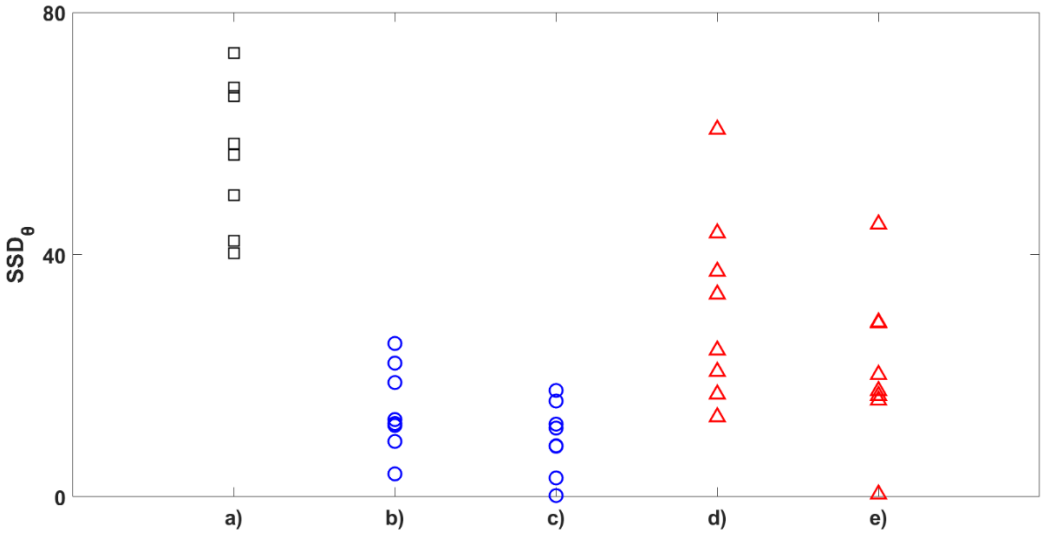


Figure 3. SSD_θ for parameter estimates obtained using a) old settings alone, b) old settings plus 1 MBDOE experiment, c) old settings plus 2 MBDOE experiments, d) old settings plus 1 corner experiment, and e) old settings plus 2 corner experiments.

We were interested to find out which parameter estimates benefitted most from the new designed experiments. **Table 10** shows a list of the nine parameters whose estimates were improved the most (by more than 90%) by using additional data from one new MBDOE experiment. The root-mean-squared-error (RMSE) value for each of the parameters (based on all eight synthetic data sets) was calculated using:

$$RMSE_j = \sqrt{\frac{\sum_{k=1}^{8}\left(\frac{\hat{\theta}_j - \theta_j^{true}}{s_{\theta_j}}\right)_k}{8}} \qquad (6)$$

As shown in Table 10, improvements in these MBDOE parameter estimates are considerably better than those obtained using additional random corner runs. As expected, estimates of most parameters were improved more when data from two MBDOE runs was used compared to when only one MBDOE run was used to re-estimate the parameters. When all 70 parameters are

considered, the average improvement in the RMSE resulting from one MBDOE experiment is 49.6%, whereas the average improvement resulting from one corner-point experiment is only 25.2%.

Table 10. Percentage improvement in RMSE for the most-improved parameter estimates

Parameter	Δ_{RMSE} (%)			
	MBDOE		Corner	
	1 run	2 runs	1 run	2 runs
$k_{C(5)}$	93.0	94.7	5.9	93.3
$E_{k_{10,\ 4}}$	96.7	96.7	61.1	89.5
$E_{k_{10,\ 5}}$	97.6	99.2	41.3	87.2
$E_{k_{10,\ 6}}$	96.4	98.1	66.9	81.8
$E_{k_{10,\ 7}}$	94.2	92.9	50.1	94.1
$E_{k_{11,\ 6}}$	90.5	94.2	90.1	92.0
$E_{k_{11,\ 7}}$	91.4	93.5	87.4	91.9
$P_W^{160}(0.0005)$	96.0	99.2	41.1	94.9
$P_W^{160}(0.02)$	98.0	99.3	0.7	7.3

Figure 4 shows the number of parameter estimates that are significantly different from zero at the 95% confidence level for all old and new sets of synthetic parameter estimates. A repeated value is indicated by a symbol with a dot. As shown in Figure 4, the number of significant parameter estimates increased considerably when additional data from one MBDOE run were used to aid parameter estimation. This number increased even further when data from two MBDOE experiments were used. Using additional data from one and two corner experiments also helped to increase the number of parameter estimates that are significantly different from zero, but the MBDOE results are better on average.

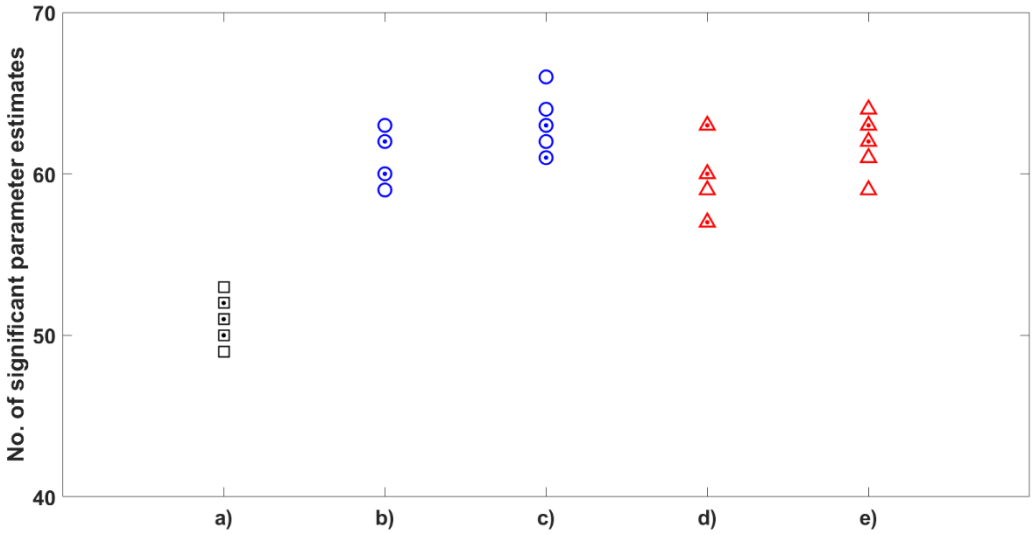


Figure 4. Number of significant parameter estimates at the 95% confidence level when parameter estimates were obtained using a) old settings alone, b) old settings plus 1 MBDOE experiment, c) old settings plus 2 MBDOE experiments, d) old settings plus 1 corner experiment, and e) old settings plus 2 corner experiments.

5. Conclusions

In this study, a simplified sequential A-optimal MBDOE method is used to design new experimental runs for a PO3G batch-reactor system, using a fundamental PO3G model and industrial data. Operating conditions selected for each new run include: reactor temperature, initial catalyst level, and initial concentration of water. The permissible ranges of these experimental settings are wider than those used in previous industrial experiments. A Bayesian objective function is used to account for prior knowledge about physically-realistic parameter values, thereby overcoming problems with a singular Fisher Information Matrix, which would need to be inverted using traditional MBDOE methods.

The MBDOE results indicate that, if one new experimental run can be performed, it should be conducted at the highest permissible temperature (i.e., 185℃), with a high level of catalyst (i.e., 0.28 wt%) and a high-water concentration (i.e., 9.10 wt%). If two new experiments can be performed, both experiments should be conducted at high temperatures (i.e., 183.5℃, 185℃), with high initial water concentrations (i.e., 10.00 wt%, 9.75 wt%), and both moderate and high catalyst concentrations (i.e., 0.18 wt%, 0.26 wt%).

Monte-Carlo simulations are used to compare the effectiveness of the proposed MBDOE method with and alternative more-traditional approach where the modeler selects corner points (at random) from the permissible design space. Eight sets of synthetic old data are generated using the experimental settings corresponding to the industrial data, resulting in eight sets of synthetic old parameter estimates. The proposed MBDOE approach is then used to design one new experiment for each set of the synthetic old data and then two new experiments based on the synthetic old data. The resulting designed experiments are then simulated and used along with corresponding synthetic old data to re-estimate the parameters. A performance measure (the scaled sum-of-squared-deviations between the true parameter values and their estimates) is used to confirm that the new parameter estimates are considerably more accurate compared than those obtained using the synthetic old data alone. We show that the number of parameter estimates that are significantly different from zero (at the 95% confidence level) increases when additional MBDOE runs are used for parameter estimation. As expected, the parameter estimation results improve when additional data from two MBDOE experimental runs is avalailable, compared with the situation when only one new MBDOE run is performed.

Also in this MC study, results obtained using the MBDOE approach are compared with those obtained when new experimental runs are randomly selected from among the corners of the

permissible design space. Additional data from randomly-selected corner experiments also improves the parameter estimates, as expected, but the simulation results confirm that experiments designed using MBDOE are superior. For example, the parameter estimates obtained using one new MBDOE experiment are 24.4% closer, on average, to their true values than parameter estimates obtained using one new corner-point experiment.

Acknowledgements

The authors acknowledge financial support from Mitacs Globalink and the Natural Sciences and Engineering Research Council of Canada. Data provided by E.I. du Pont Canada are gratefully acknowledged. Technical advice provided by Dr. T. Xie and Dr. R.E. Spence from E.I. du Pont Canada is also gratefully acknowledged.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Sunkara HB, Ng H DuPont™ Cerenol™ - A New Family of High Performance Bio-Based Polymers. Presented at: Sarnia Bio-Based Polyols; October 19, 2006; Sarnia, ON, Canada.
2. Sunkara HB, Miller RW. DuPont™ Cerenol™ - A New Family of High Performance Polyether Glycol. Presented at: 5th Annual World Congress on Industrial Biotechnology & Bioprocessing; April 27-30, 2008; Chicago, IL, USA
3. Xie T. Development of Bio-Based Polymers: Fundamentals, Process Scale up, and Technology Challenges. Presented at: Polymer Reaction Engineering VII; May 3-8 2009; Niagara Fall, Ontario, Canada

4. Mark E, Sharon LH, Lisa AL, Jeff PP, Gregory W, inventor; Danisco US Inc., DuPont US Holding LLC, assignee. Process for the biological production of 1,3-propanediol with high titer. U.S. Patent No. 6514733B1. February 04, 2003.
5. Hofer R, ed., *Sustainable Solutions for Modern Economies*. Cambridge, UK: Royal Society of Chemistry; 2009: 436 – 478
6. Harmer MA, Confer DC, Hoffman CK, Jackson SC, Liauw AY, Minter AR, Murphy ER, Spence RE, Sunkara HB. Renewably sourced polytrimethylene ether glycol by superacid catalyzed condensation of 1,3-propanediol. *Green Chem.* 2010; 12(8): 1410-1416
7. Sunkara HB, Marchildon EKA, Ng H, Manzer LE, inventor; DuPont Canada Inc., DuPont Industrial Biosciences USA LLC, assignee. Continuous process for the preparation of polytrimethylene ether glycol. U.S. Patent No. 20020010374A1. April 13, 2004
8. Sunkara HB, Manzer LE, inventor; DuPont Industrial Biosciences USA LLC, assignee. Production of polytrimethylene ether glycol and copolymers thereof. U.S. Patent No. 6977291B2. December 20, 2005
9. Sunkara HB, Parmpi P, inventor; EI Du Pont de Nemours and Co, assignee. Process for preparation of polytrimethylene ether glycols. U.S. Patent No. 7074969B2. July 11, 2006
10. Sunkara HB, Do HQ, inventor; EI Du Pont de Nemours and Co, assignee. Poly(trimethylene-ethylene ether) glycols. U.S. Patent No. 7244810B2. July 17, 2007
11. Sunkara HB, Do HQ, inventor; DuPont Industrial Biosciences USA LLC, assignee. Processes for removal of color bodies from polytrimethylene ether glycol polymers. U.S. Patent No. 7294746B2. November 13, 2007

12. Sunkara HB, Ng HC, inventor; DuPont Industrial Biosciences USA LLC, assignee. Polytrimethylene ether glycol and polytrimethylene ether ester with excellent quality. U.S. Patent No. 7323539B2. January 29, 2008
13. Sunkara HB, Seapan M, Diffendall GF, Gallagher GF, inventor; EI Du Pont de Nemours and Co, assignee. Hydrogenation of polytrimethylene ether glycol. U.S. Patent No.7342142 B2. March 11, 2008
14. Sunkara HB, Ng HC, inventor; DuPont Industrial Biosciences USA LLC, assignee. Manufacture of polytrimethylene ether glycol. U.S. Patent No. 7388115B2. June 17, 2008
15. Mueller PA, Murphy ER, Rajagopalan B, Congalidis JP, Minter AR. Polymerization Reactor Modeling in Industry. *Macromol. Symp*. 2011; 5: 261 - 277
16. Mueller PA, Rajagopalan B, Congalidis JP, Murphy ER. Mathematical Modeling of Acid-Catalyzed 1,3-Propanediol Polymerization. *Macromol. React. Eng.* 2012; 6: 126 - 152
17. Cui WJ, McAuley KB, Spence RE, Xie T. Mathematical Model of Polyether Production from 1,3-Propanediol. *Macromol. React. Eng*. 2013; 7: 237-253
18. Cui WJ, McAuley KB, Spence RE, Xie T. Assessment of Mass-Transfer Effects during Polyether Production from 1,3-Propanediol. *Macromol. React. Eng*. 2014; 8: 476-492
19. Cui WJ, McAuley KB, Spence RE, Xie T. Mathememattical Modeling of Polyether Production from 1,3-Propanediol: Accounting for Linear Oligomers. *Macromol. React. Eng*. 2015; 9: 186 - 204
20. Vo ADD, Elraghy A, Spence RE, McAuley KB. An Improved Model for Polyether Production from 1,3-Propanediol. *Macromol. React. Eng*. 2020; 14, 1900045.
21. Vo ADD, Cui WJ, McAuley KB. An Improved PO3G Model–Accounting for Cyclic Oligomers. *Macromol. Theory Simul.* 2020; 29, 2000023.

22. Vo ADD, Spence RE, McAuley KB. Parameter Estimation in a PO3G Model with Temperature Effects. *Macromol. Theory Simul*. 2021; 30, 2000061.

23. Thompson DE, McAuley KB, McLellan PJ. Parameter estimation in a simplified MWD model for HDPE produced by a ziegler-natta catalyst. *Macromol React Eng*. 2009;3(4):160-177. doi:10.1002/mren.200800052

24. Yao KZ, Shaw BM, Kou B, McAuley KB, Bacon DW. Modeling Ethylene/Butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polym React Eng*. 2003;11(3):563-588. doi:10.1081/PRE-120024426

25. Wu S, McLean K a. P, Harris TJ, McAuley KB. Selection of optimal parameter set using estimability analysis and MSE-based model-selection criterion. *Int J Adv Mechatron Syst*. 2011;3(3):188. doi:10.1504/IJAMECHS.2011.042615

26. Wu S, Mcauley KB, Harris TJ. Selection of simplified models: II. Development of a model selection criterion based on mean squared error. *Can J Chem Eng.* 2011;89(2):325-336. doi:10.1002/cjce.20479

27. John RCS, Draper NR. D-Optimality for regression designs: A review. *Technometrics*. 1975;17(1):15-23. doi:10.1080/00401706.1975.10489266

28. Welch WJ. Computer-aided design of experiments for response estimation. *Technometrics [Internet]*. 1984; 26: 217–224.

29. Ford I, Silvey SD. A sequentially constructed design for estimating a nonlinear parametric function. *Biometrika*. 1980;67(2):381-388. doi:10.1093/biomet/67.2.381

30. Issanchou S, Cognet P, Cabassud M. Sequential experimental design strategy for rapid kinetic modeling of chemical synthesis. *AIChE J*. 2005;51(6):1773-1781. doi:10.1002/aic.10439

31. Bauer I, Bock HG, Körkel S, Schlöder JP. Numerical methods for optimum experimental design in DAE systems. *J Comput Appl Math*. 2000;120(1):1-25. doi:http://dx.doi.org/10.1016/S0377-0427(00)00300-9
32. Box GEP, Lucas HL. Design of Experiments in Non-Linear Situations. Vol 46. *[Oxford University Press, Biometrika Trust]*; 1959.
33. Shirt RW, Harris TJ, Bacon DW. Experimental Design Considerations for Dynamic Considerations. *Ind. Eng. Chem. Res*. 1994; 33: 2656-2667
34. Walter E, Pronzato L. Qualitative and quantitative experiment design for phenomenological models-A survey. *Automatica*. 1990;26(2):195-213. doi:10.1016/0005-1098(90)90116-Y
35. Ford I, Titterington D, Kitsos CP. Recent Advances in Nonlinear Experimental Design. *Technometrics*. 1989; 31(1): 49-60. doi:10.2307/1270364
36. Atkinson AC. Developments in the Design of Experiments, Correspondent Paper. *Int Stat Rev / Rev Int Stat.* 1982;50(2):161-177. doi:10.2307/1402599
37. Pinto JC, Lobão MW, Monteiro JL. Sequential experimental design for parameter estimation: a different approach. *Chem Eng Sci.* 1990;45(4):883-892. doi:http://dx.doi.org/10.1016/0009-2509(90)85010-B
38. Box GEP, Hunter JS, Hunter WG. *Statistics for experimenters. Design, innovation, and discovery*. NJ: John Wiley & Sons, Inc.; 2005
39. Shahmohammadi A, McAuley KB. Sequential Model-Based A- and V-Optimal Design of Experiments for Building Fundamental Models of Pharmaceutical Production Processes. *Comput Chem Eng*. 2019

40. Shahmohammadi A, McAuley KB. Sequential Model-Based A-Optimal Design of Experiments When the Fisher Information Matrix Is Noninvertible. *Ind Eng Chem Res.* 2019;58(3):1244-1261. doi:10.1021/acs.iecr.8b03047
41. Shahmohammadi A, McAuley KB. Using Prior Parameter Knowledge in Model-Based Design of Experiments for Pharmaceutical Production. *AIChE J*. 2020; 66: e17021. https://doi.org/10.1002/aic.17021
42. Chaloner K, Verdinelli I. Bayesian experimental design: A review. Stat Sci. 1995;10(3):273-304. doi:10.1214/ss/1177009939
43. Ruggoo A, Vandebroek M. Bayesian sequential script design optimal model-robust designs. Comput Stat Data Anal. 2004;47(4):655-673. doi:10.1016/j.csda.2003.09.014