Reactivity Ratio Estimation in Radical Copolymerization: From Preliminary Estimates to Optimal Design of Experiments

AUTHOR

*Niousha Kazemi[1], Benoît H. Lessard[2], Milan Marić [3], Thomas A. Duever[1], Alexander Penlidis[1]**

AUTHOR ADDRESSES

1: Institute for Polymer Research (IPR), Department of Chemical Engineering, University of Waterloo, Waterloo, Canada, N2L 3G1

2: Department of Chemical Engineering & Applied Chemistry, University of Toronto, Toronto, Canada, M5S 3E5

3: Department of Chemical Engineering, McGill University, Montreal, Canada, H3A 0C5

ABSTRACT

An error-in-variables-model (EVM) framework is presented for the optimal estimation of reactivity ratios in copolymerization systems. This framework consists of several sequential steps and practical prescriptions that can yield reliable and statistically correct reactivity ratio values. These steps include: (a) screening experiments for estimating preliminary reactivity ratios, (b) optimal design of experiments, (c) full conversion range experiments and estimation of optimal reactivity ratios, and if necessary, (d) design of sequentially optimal experiments, re-estimation of reactivity ratios and diagnostic checks. This complete methodology should become common practice for determining reactivity ratios with the highest possible level of confidence. The performance of this framework is verified experimentally with data from the controlled nitroxide-mediated copolymerization of 9-(4-vinylbenzyl)-9H-carbazole (VBK) and methyl methacrylate (MMA), a novel and largely unstudied copolymer system.

INTRODUCTION

There is no doubt about the importance of reactivity ratios in copolymerization (simplest case) and other multicomponent polymerizations, for describing chain microstructure and hence determining polymer chain composition and sequence length. At the same time, there is also no doubt that these important copolymerization parameters have been estimated incorrectly for several decades and are still being handled incorrectly (from a statistical estimation perspective) in the scientific literature.[1,2] This has resulted in a database of relatively biased and unreliable reactivity ratios.

The majority of current techniques used by polymer chemists to determine reactivity ratios are incorrect and the associated pitfalls of these techniques have been described in detail over the last three decades.[1-5] Unfortunately, despite these and other pioneering efforts,[6] what has been missing is not only a (statistically) correct estimation, but also the optimal design of experiments. Not only are reactivity ratios not correctly estimated based on collected experimental data, but also the experimental trials conducted to collect the required data are poorly designed (i.e., initial feed fractions at which to run copolymerizations are not optimally located). The issue is further compounded by the fact that the experiments conducted are not replicated, and therefore, they do not yield the necessary estimate of the underlying experimental error.

This brings us to a currently observed paradox in the scientific literature. Combining design of experiments techniques with any parameter estimation problem is highly recommended as a problem-solving approach but rarely practiced!

The parameter estimation question has been revisited in great detail in recent publications.[7,8] Taking the estimation question a step further (literally, a step before the estimation stage), Kazemi et al.[9] looked at different optimal design (of experimental trials) criteria that can lead to more reliable reactivity ratio estimates, highly superior to those from other currently implemented techniques.

In the current publication, we describe the full story of how one can go from preliminary reactivity ratio estimates to optimal estimates, i.e., we give an overview of the (statistically) correct design of experiments followed by the (statistically) correct parameter estimation. The steps are iterative, sequential and optimal, and deal with nonlinear mathematical models and estimation procedures. The overall procedure is experimentally demonstrated and verified with data from a largely unstudied and novel copolymerization system, the nitroxide-mediated controlled radical copolymerization between 9-(4-vinylbenzyl)-9H-carbazole (VBK) and methyl methacrylate (MMA).[10]

REACTIVITY RATIO DETERMINATION: THE COMPLETE EVM FRAMEWORK

In this section a complete framework for the reactivity ratio estimation problem is presented and the individual steps are described. This complete framework is based on the error-in-variables-model (EVM) method. The EVM method is a parameter estimation technique suitable for problems where all the variables are subject to error, i.e., dependent and independent variables are not distinguishable (as opposed to classical regression analysis).[11] This feature makes EVM the perfect method for estimating reactivity ratios in multicomponent polymerizations.[1,2,7]

Since collecting experimental data that result in precise parameter estimates is a resource-intensive task, there is always a need for designing experiments in an optimal fashion, thus minimizing the overall effort and maximizing the information from the process in question. Such a procedure consists of a series of alternating steps between designing experiments and using the optimally obtained experimental data for parameter estimation, which continues until the parameter estimates achieve satisfactory levels of precision. It is therefore desirable to combine the design of experiments technique with the EVM procedure in one multifaceted framework in order to improve the quality of the parameter estimation results. This structure is briefly explained in the three sections ((1) to (3)) that follow:

Parameter estimation and design of experiments

Full conversion range experiments

Recommended steps

(1) Parameter estimation and design of experiments

The basis of the EVM framework for parameter estimation was proposed originally by Reilly and Patino-Leal.[11] The algorithm was specifically applied to the problem of estimating reactivity ratios for copolymerization systems by Dube et al.[1] and Polic et al.[2], and further detailed explanations about it can be found in Kazemi et al.[8], where the latest methodology for using this algorithm and several key factors for its numerical implementation are highlighted.

At first, the vector of measurements $\underline{x}_i$ is equated to the vector of true (yet unknown) values $\underline{\xi}_i$, plus a multiplicative error term, $k\underline{\varepsilon}_i$, as shown in equation (1),

$$\underline{x}_i = \underline{\xi}_i(1+k\underline{\varepsilon}_i) \qquad \text{where } i=1,2,\ldots,n \tag{1}$$

where i refers to the trial number, k is a constant, and $\underline{\varepsilon}_i$ is a random variable, which, in the simplest case, has a uniform distribution in the interval from -1 to 1. The value of the constant k for different variables in a problem could be different, as the amount of uncertainty in data coming from different sources may be different. A log transformation of the variables in equation (1) is also necessary so that the error term becomes additive. Taking logarithms of both sides of equation (1), $\ln(1+k\varepsilon)$ can be replaced by $k\varepsilon$, provided that the magnitude of the error does not exceed 10% (k < 0.1). The error structure is chosen to be multiplicative (relative) due to the nature of the actual (physical) measurements in the reactivity ratio estimation problem. The measurement errors for all the variables appear in the variance-covariance matrix of the measurements, $\underline{V}$ (to be used shortly), which is non-singular and known. As simple algebra can show, for each variable,[12] the corresponding element in $\underline{V}$ is $k^2/3$.

In addition to equation (1), there is a statement that relates the true (yet unknown) values of the parameters, $\underline{\theta}^*$, and variables, $\underline{\xi}_i$, via the mathematical model, represented by equation (2).

$$\underline{g}(\underline{\xi}_i, \underline{\theta}^*) = 0 \qquad \text{where} \quad i=1,2,\ldots,n \tag{2}$$

Using a Bayesian approach, the objective function for minimization in order to find the point estimates, $\underline{\hat{\theta}}$, is given by equation (3), where r_i is the number of replicates at the i^{th} trial, $\underline{\bar{x}}_i$ is the average of the r_i measurements $\underline{x}_i$, and $\underline{\hat{\xi}}_i$ denotes estimates of the true values of the variables $\underline{\xi}_i$.

$$\phi = \frac{1}{2}\sum_{i=1}^{n} r_i (\underline{\bar{x}}_i - \underline{\hat{\xi}}_i)' \underline{V}^{-1} (\underline{\bar{x}}_i - \underline{\hat{\xi}}_i) \tag{3}$$

The procedure for parameter estimation is a nested iterative one. There is a main iterative loop that searches for the parameter values (minimizing equation (3)), while an inner loop finds the true values of the variables. This procedure results in both parameter estimates and estimates of the true values of the variables. Finally, point estimates should be accompanied by joint confidence regions that reflect their level of uncertainty (for more details on these steps, see the recent discussion in Kazemi et al.[8]).

To design experiments for parameter estimation in the EVM context, we should extend the ideas from classical nonlinear regression analysis[12,13] to EVM by maximizing the determinant of the information matrix (which aims at reducing the elements of the variance-covariance matrix of the parameters).[14] The information matrix, which is the inverse of the variance-covariance matrix of the parameters, is given by matrix $\underline{G}$, in equation (4), as the second derivatives of ϕ with respect to the parameters. $\underline{G}$ is computed by equation (5). This forms the basis for locating optimal experiments.

$$\underline{G} = E\left[\frac{d^2\phi}{d\underline{\theta}_i d\underline{\theta}_j}\right] \qquad \text{i, j element} \tag{4}$$

$$\underline{G} = \sum_{i=1}^{n} r_i \underline{Z}_i^{'} (\underline{B}_i \underline{V} \underline{B}_i^{'})^{-1} \underline{Z}_i \tag{5}$$

$\underline{Z}_i$ is the vector of partial derivatives of the (model) function, $\underline{g}(\underline{\xi}_i, \underline{\theta}^*)$, with respect to the parameters, and $\underline{B}_i$ is the vector of partial derivatives of the function, $\underline{g}(\underline{\xi}_i, \underline{\theta}^*)$, with respect to the variables, given respectively by equations (6) and (7).

$$\underline{Z}_i = \left[\frac{\partial g(\underline{\xi}_i, \underline{\theta})}{\partial \theta_m} \right] \qquad \text{m}^{\text{th}} \text{ element} \tag{6}$$

$$\underline{B}_i = \left[\frac{\partial g(\underline{\xi}_i, \underline{\theta})}{\partial (\underline{\xi}_i)} \right] \tag{7}$$

Generally, there are two approaches for model-based design of experiments, namely, initial and sequential design.[14,15] The initial design refers to the problem where no prior information is available and the objective is to design initial trials in order to collect the first data set and commence the parameter estimation process. The design criterion for maximization is given by equation (8). The constraints for this problem are the function (model) itself as well as lower (L) and upper (U) bounds of the experimentally feasible region.

$$Max_{\underline{\xi}} \left| \sum_{i=1}^{n} r_i \underline{Z}_i^{'} (\underline{B}_i \underline{V} \underline{B}_i^{'})^{-1} \underline{Z}_i \right| \tag{8}$$

$$\text{subject to} \begin{cases} \underline{g}(\underline{\xi}_i, \underline{\theta}) = 0 \\ L \le \underline{\xi}_i \le U \end{cases}$$

In the sequential design of experiments scheme, on the other hand, and after having conducted a set of *n* prior experiments, it is desired to find the optimal location of the next experiment(s) which can increase the precision of the results. Such a procedure consists of a series of alternating steps between designing experiments, using the experimental data for parameter estimation, and designing further experiments. This continues until the parameter estimates achieve a satisfactory level of precision. The objective function is given by equation (9) and the constraints are similar to those discussed for equation (8).[14] In equation (9), the hat sign ($\hat{}$) denotes that the converged parameter estimates from the first set of n experiments are being used to compute $\underline{B}_i$ and $\underline{Z}_i$ (as per equations (6) and (7), respectively).

$$Max_{\underline{\hat{\xi}}_{n+1}} \left\| \left[\sum_{i=1}^{n} r_i \underline{\hat{Z}'}_n (\underline{\hat{B}}_n \underline{V} \underline{\hat{B}}_n)^{-1} \underline{\hat{Z}}_n \right] + \underline{\hat{Z}'}_{n+1'} (\underline{\hat{B}}_{n+1} \underline{V} \underline{\hat{B}}_{n+1})^{-1} \underline{\hat{Z}}_{n+1} \right| \tag{9}$$

$$\text{subject to} \begin{cases} \underline{g}(\underline{\xi}_{n+1}, \underline{\theta}) = 0 \\ L \leq \underline{\xi}_{n+1} \leq U \end{cases}$$

For designing experiments for the purpose of estimating reactivity ratios, perhaps the most cited work is by Tidwell and Mortimer.[6] The very useful (and often ignored) Tidwell-Mortimer criterion is, however, based on the assumptions of classical nonlinear regression. The criterion results in two optimal initial feed compositions, f_{11} and f_{12}, with respect to the mole fraction of monomer 1, shown in equation (10). These two optimal compositions are solely functions of (initially available) reactivity ratios.

$$f_{11} \cong \frac{2}{2 + r_1} \ , \ f_{12} \cong \frac{r_2}{2 + r_2} \tag{10}$$

But, as mentioned above, the nature of the reactivity ratio estimation problem is different from traditional nonlinear regression analysis, in that there are errors in all the variables. Because of this difference, the assumptions behind the Tidwell-Mortimer criterion do not hold for this problem and thus using equation (10) to choose optimal experiments is not appropriate within the EVM context. Implementation of the EVM design criterion, both with initial and sequential design schemes (as shown in equations (8) and (9), respectively), has recently been described in Kazemi et al.[9], with extensive details about the method, the equations, and the overall algorithm.

(2) Full conversion range experiments

For the purpose of reactivity ratio estimation in copolymerization systems, there have been numerous publications in the literature over the past 50 years, during which two main approaches have established themselves. The first and main practice is to use low conversion experimental data with the instantaneous copolymer composition model (the so-called Mayo-Lewis model). The second approach, as already mentioned previously, is to use the Tidwell-Mortimer equations[6] to identify feed mole fractions for the design of experiments for reactivity ratio estimation. Both of these methodologies have drawbacks, with the former having much more severe defects than the latter. In our EVM framework, these two techniques have been modified in order to avoid their underlying problems. Similar to the explanations regarding the design of experiments problem (discussed in the previous section), other modifications, mainly related to the low conversion experimental analysis, are explained below.

Reactivity ratio estimation based on the Mayo-Lewis model has certain limitations because of the assumptions involved with this model. The model, relating the instantaneous copolymer composition, F_1, to the unreacted monomer mole fraction, f_1, is shown in equation (11), with r_1 and r_2 being the reactivity ratios for monomer 1 and 2.

$$F_1 = \frac{r_1 f_1^2 + f_1(1-f_1)}{r_1 f_1^2 + 2f_1(1-f_1) + r_2(1-f_1)^2} \tag{11}$$

Based on the assumptions and instantaneous nature of this model, only low conversion (conversion < 5%) experiments can be used for estimating reactivity ratios. If low conversion data are collected, one can "assume" that there has been no composition drift in the system, i.e., initial feed compositions, f_{10}, have remained the same and so f_{10} values can be used instead of f_1. Also, the measured copolymer composition values are in actual fact cumulative copolymer compositions, $\overline{F}_1$, which are assumed to be equal to the instantaneous copolymer composition values, F_1. These assumptions may not hold for many cases and thus using low conversion data introduces experimental error immediately. Therefore, all these become sources of bias in the reactivity ratio estimates.

Using higher conversion data, or in fact, data from the whole experimental conversion trajectory, is another approach that is more recent than the Mayo-Lewis equation. In this method, the differential equation for the unreacted monomer fraction f_1 with respect to conversion (X_n) is being integrated over the course of conversion, as shown in equation (12). The unreacted monomer mixture composition, f_1, is then used in equation (13) to evaluate the cumulative copolymer composition, $\overline{F}_1$. Also, since the measured conversion is usually on a mass (weight) basis, a weight conversion (X_w), as in equation (14), should be used to relate X_n to X_w.

$$\frac{df_1}{dX_n} = \frac{f_1 - F_1}{1 - X_n} \tag{12}$$

$$\overline{F}_1 = \frac{f_{10} - f_1(1 - X_n)}{X_n} \tag{13}$$

$$X_n = X_w \frac{Mw_1 f_{10} + (1 - f_{10})Mw_2}{Mw_1 \overline{F}_1 + (1 - \overline{F}_1)Mw_2} \tag{14}$$

The performance of this approach, which is referred to as direct numerical integration (DNI), without resorting to any analytical integration, was recently evaluated and discussed in detail in Kazemi et al.[7], with comparisons to the Mayo-Lewis model or the Meyer-Lowry analytically integrated model. It was shown that using the DNI approach and using medium and high conversion data for estimating reactivity ratios can significantly improve the quality of the results by (simply) including more information in the analysis as well as avoiding (practical) limitations of collecting low conversion data (with their inherent sources of errors). Therefore, our EVM framework based on the DNI approach can handle experimental data at any conversion level.

Given the discussion on sections (1) and (2), section (3), which follows, summarizes the prescriptions for obtaining reactivity ratios both efficiently and consistently.

(3) Recommended steps

The flowchart of figure 1 gives an overview and summary of the recommended steps of the complete methodology. Following the steps on this flowchart will increase the reliability of the reactivity ratio estimates for any copolymerization system (an example from a novel experimental system is shown in the results and discussion section). Also, these steps can be implemented during investigations of any larger multicomponent polymerization. The recommended approach is summarized in the following steps:

1. Start with the given copolymerization system.
2. Check the related literature for reasonable preliminary guesses (e.g., reactivity ratio estimates that have been estimated in similar experimental settings).
   2.1. If initial guesses for reactivity ratios are available, go to step 5.
3. If no information is available, perform preliminary experiments and collect data.
   3.1. Experiments can be performed up to medium conversion levels (say, up to 20-40%).
   3.2. Cover low, medium and high feed compositions and note any limitations in the feasible experimental region (constraints).
   3.3. Collect data on feed composition, copolymer composition, and conversion.
   3.4. Perform independent replicates (this important step is rarely done).
4. Use EVM to estimate reactivity ratios (Estimation).
   4.1. Work with the DNI approach, since it is more general.

    4.2. If measurement error levels are known, specify k (as per equation (1)) and populate the variance-covariance matrix of the measurements. If not enough replicates are available to calculate k, typical values which are based on the usual amount of error in common polymerization measurements can be used (i.e., k as 1% for gravimetric analysis and 5% for NMR analysis).
    4.3. Run the EVM parameter estimation algorithm as presented in Kazemi et al.[8]
    4.4. Construct joint confidence regions for reactivity ratios.[8]
5. Locate optimal feed compositions (Design).
    5.1. With copolymerizations, there are two parameters to be estimated (r_1 and r_2) and thus at least two optimal feed compositions to be calculated.[16]
    5.2. Note infeasible and feasible experimental regions. The feasible range for the optimal experiments can be chosen from a relatively small mole fraction (below 0.1) up to 1. Using the EVM design criterion, the first feed composition is always at the lowest limit of the feasible region, while the second one is a feed composition within the range 0 to 1, depending on the reactivity ratio values. This rule of thumb also applies for the sequential design scheme.
    5.3. Use the EVM design criterion as shown in equation (8) for the initial design scheme and equation (9) for the sequential design scheme. Refer to Kazemi et al.[9] for more information.
    5.4. It is preferable to use a global optimization technique.[9,16]
6. Make sure that the optimal points are experimentally feasible (i.e., they do not violate any process constraints). If not feasible, return to step 5 and adjust accordingly. For

copolymerizations, the lower bound of the feasible region is always chosen as one of the optimal points. So, set/adjust this value at the lowest possible practical feed composition.

7. Perform the optimal experiments (Experimentation).
   7.1. Initial feed compositions (f_{10}) are at the optimal points.
   7.2. Run experiments up to "moderate/high" conversion. That is, for each designed initial feed composition, experimental data on copolymer composition can be collected at conversion levels of 5%, 10%, 20%, 30% and 40%, even up to approximately 60% (i.e., the EVM approach does not have any low conversion restrictions, as common practice stipulated before).
   7.3. Perform replicated experiments; collecting independent replicates for each run increases the accuracy of the results significantly and it is therefore highly recommended.
8. Using EVM, re-estimate reactivity ratios, as per step 4.
9. If results are satisfactory, proceed with step 10. If not, use the new point estimates and the information matrix from step 7 to carry out the next sequentially designed experiment(s).
   9.1. Find the next optimal trial, using equation (9) and go to step 7.
      - Repeat steps 7 to 9 until the outcome is satisfactory.
10. Check/plot the r_1 / r_2 joint confidence region.
11. Finish.

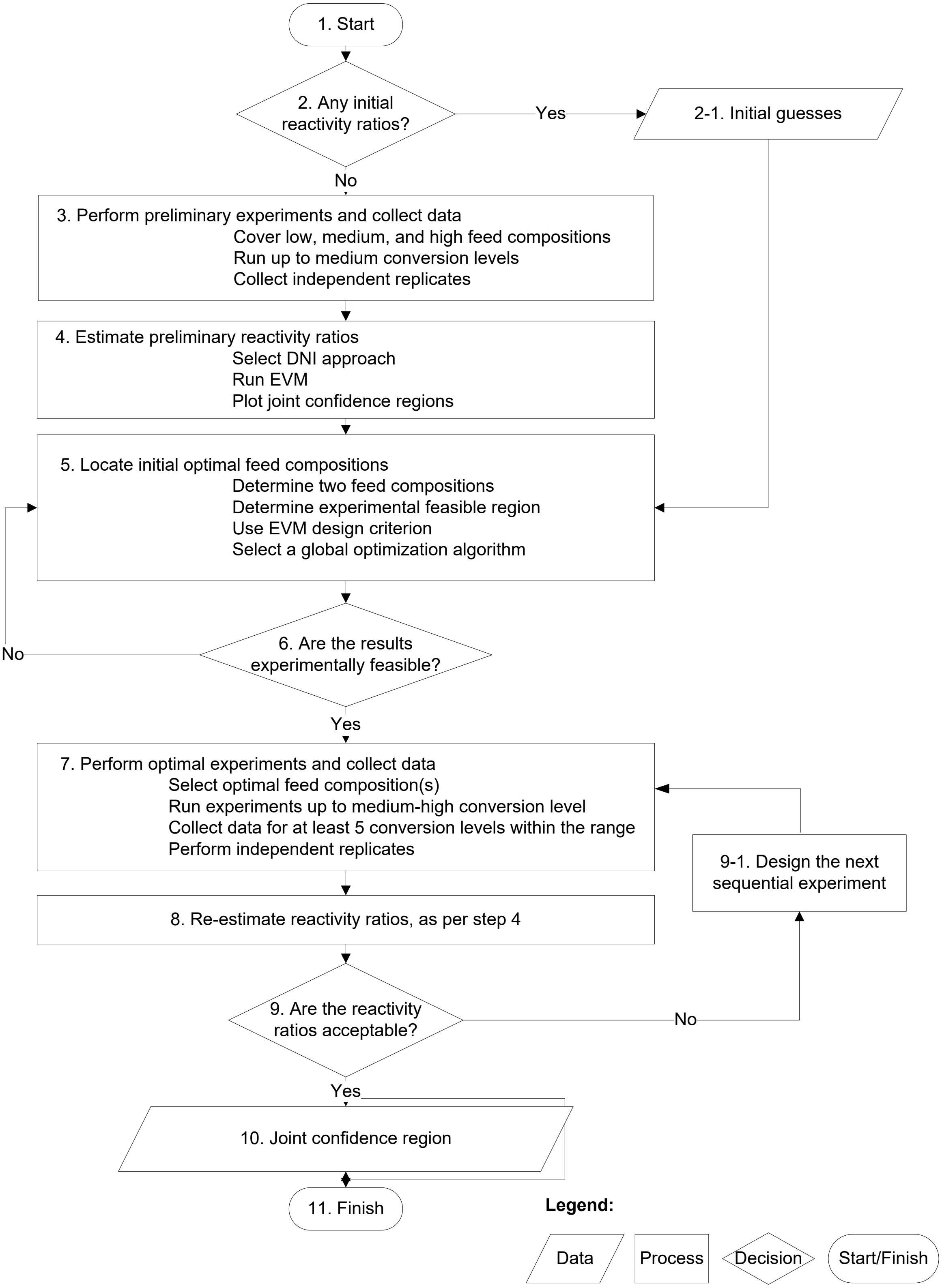

Figure 1. EVM framework flowchart for reactivity ratio estimation

EXPERIMENTAL

Materials

N,N-Dimethylformamide (DMF, 99.8%) and methyl methacrylate (MMA, 99%) were obtained from Aldrich, while methanol (99.8%), and tetrahydrofuran (THF, >99.9%) were obtained from Fisher Scientific. Prior to polymerizations, the MMA was purified by passage through a mixed column of calcium hydride (Aldrich) and basic alumina (Aldrich) (the mixture was 1:20 by weight, respectively) and stored in a sealed flask in a refrigerator under a head of nitrogen. Synthesis of VBK was performed in accordance with previous literature.[10] 2-({tert-butyl[1-(diethoxyphosphoryl)-2,2-dimethylpropyl]amino}oxy)-2-methylpropionic acid (BlocBuilder, 99%) was obtained from Arkema, while {tert-butyl[1-(diethoxyphosphoryl)-2,2-dimethylpropyl]amino} nitroxide (SG1, >85%) was kindly donated by Noah Macy of Arkema. No additional purifications were performed, unless otherwise stated.

Synthesis of MMA/VBK copolymers

The MMA/VBK copolymerizations were carried out under similar conditions to those reported by Lessard et al.[10] In a typical copolymerization, a stir bar along with a mixture of BlocBuilder, SG1, DMF, VBK and MMA were added to a sealed three-neck round-bottom flask fitted with a temperature well, condensing column, and mounted on a heating mantle with stir plate. The ratio of BlocBuilder to monomer was determined to give a target molecular weight at complete conversion of 20-25 kg.mol-1 while using an initial molar ratio of $[SG1]_0/[BlocBuilder] \approx 0.1$. In all cases, DMF added to the mixture gave a ≈20wt% solution. The mixture was bubbled with ultra-pure nitrogen and stirred for 30 min at room temperature prior to heating to 80 °C. A series of samples were then drawn by syringe. Half of each sample was directly introduced into a 5mm

NMR tube for determination of conversion and copolymer composition, while the second half was allowed to evaporate until constant mass and further analyzed by 1H NMR spectroscopy as a second determination of copolymer composition. The polymerization was allowed to proceed for 8-24 hours depending on the experiment and the final polymer was precipitated in methanol.

Copolymer characterization

Copolymer composition was determined by 1H NMR spectroscopy using the characteristic resonances for the MMA (δ = 3.75 ppm) and the VBK (δ = 5.64 ppm) units. Copolymer conversion (X_W) was determined by the following equation: $X_W = X_{MMA}.n_{MMA,0} + X_{VBK}.n_{VBK,0}$ where X_{MMA} and X_{VBK} are the individual comonomer conversions and $n_{MMA,0}$ and $n_{VBK,0}$ are the initial molar weight% of MMA and VBK, respectively. As previously reported,[17] X_{MMA} for MMA and X_{VBK} for VBK were determined by comparing the integrated peaks corresponding to the vinyl protons (δ = 6.09 and 5.55 ppm for MMA and δ = 5.71, 5.17, and 6.64 ppm for VBK) of the respective monomers, to the methoxy group (δ = 3.75 ppm, corresponding to MMA) and the methylene group (δ = 5.64 ppm, corresponding to VBK) corresponding to the respective bound monomer units in the copolymer.

RESULTS AND DISCUSSION: EXPERIMENTAL VERIFICATION OF COMPLETE SCHEME

The following subsections illustrate the complete procedure with the EVM framework, for estimating the reactivity ratios of the 9-(4-vinylbenzyl)-9H-carbazole (VBK) and methyl methacrylate (MMA) copolymerization system. This is a novel and largely unstudied controlled radical copolymerization (due to the electron-donating properties of VBK), and hence fertile ground to illustrate the procedure with it. VBK has been identified as a controlling comonomer for a variety of methacrylates, such as oligo(ethylene glycol) methyl ether methacrylate (OEGMA)[17], 2-N-morpholinoethyl methacrylate (MEMA)[18], and methacrylic acid (MAA)[19]. The VBK reactivity ratios with the corresponding methacrylates have been identified as a contributing factor to this control.[20] A preliminary investigation concerning the reactivity ratios of VBK/MMA was described in Lessard et al.[10] In summary, experimental data from Lessard et al.[10] were analyzed first to obtain preliminary reactivity ratios, which were subsequently used for the optimal design of experiments. In the second step, experiments were performed based on these optimal feed compositions. In the third step, reactivity ratios were re-estimated based on the designed data. Finally, these updated reactivity ratios were used once more in order to design the next optimal experiment but now in a sequential fashion. After performing sequential experiments, the reactivity ratios were re-estimated. The results showed a clear improvement in the quality and precision of the reactivity ratios obtained.

Preliminary experimentation (steps 1-4, figure 1)

Preliminary attempts at determining the reactivity ratios for this system involved low conversion experiments and a nonlinear fitting procedure for estimating reactivity ratios.[10] The low conversion data are shown in table 1, where the X_W, f_0, and $\overline{F}$ columns show the mass conversion, initial feed composition, and cumulative copolymer composition, respectively. The reference reactivity ratios from Lessard et al.[10] with their confidence intervals are shown in table 2.

Table 1. Low conversion data for VBK/MMA copolymerization (Lessard et al.[10])

$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$
17	0.01	0.03
13	0.03	0.06
16	0.05	0.14
10	0.10	0.37
10	0.10	0.29
8	0.30	0.71
8	0.30	0.55
4	0.40	0.79
4	0.40	0.59
7	0.50	0.74
7	0.50	0.62
4	0.80	0.91
4	0.80	0.89

To obtain reactivity ratios for the subsequent design of experiments, low conversion data points, as in table 1, were analyzed with the DNI approach[7,8] and new reactivity ratios were obtained, as shown in the second row of table 2. The Mayo-Lewis model, which is used most commonly for estimating reactivity ratios in the literature, was not employed in the current work, as the conversion levels of the data points (shown in table 1) were much higher than 5%; hence, this is a perfect application of the DNI approach, as recently described in Kazemi et al.[7,8] Figure 2 shows joint confidence regions (JCR) for the reactivity ratios in table 2. The reference reactivity ratios are not contained in the new JCR (dashed line). The new JCR, reflecting the level of uncertainty in the estimates, is quite large (especially with respect to the uncertainty in the r_{VBK} (r_1) value).

Table 2. Reactivity ratio estimates for VBK (M_1) /MMA (M_2) copolymerization

	Copolymerization model	Data set/Conversion level	r_1	r_2
Lessard et al.[10]	Mayo-Lewis	Lessard et al.[10]	2.7 ± 1.5	0.24 ± 0.14
Current work	DNI	Lessard et al.[10]	4.13	0.29
Current work	DNI	EVM initial optimal data	1.64	0.32
Current work	DNI	EVM initial and sequential optimal data	1.53	0.32
Current work	DNI	Tidwell-Mortimer initial optimal data	2.66	0.48

Note that for the EVM implementation of the DNI approach and calculation of the JCR, the k values (see equation (1)) for feed mole fraction and conversion measurements were 0.005, corresponding to a typical error level of 0.5% (coming from gravimetry). The k value for copolymer composition measured via NMR was 0.10 (i.e., errror level of 10%).

The next step was to calculate initial optimal feed compositions. As mentioned earlier, since there are two parameters under study, two optimal feed compositions (mole fractions), f_{11} and f_{12}, should be located (f_{11} refers to the mole fraction of M_1 (VBK) for the first trial and f_{12} is the mole fraction of VBK for the second trial). The values of these points depend on the prior values of reactivity ratios, and in this case, we used the reactivity ratios estimated from the low conversion data (second row of table 2). The EVM design criterion[9] resulted in two feed compositions, with $f_{11} \rightarrow 0$ and $0 < f_{12} < 1$. In order to keep the design procedure optimal and practical, a (practically) small value of f_{11} should be the lower limit of the feasible region that is determined for each system based on operating conditions and experimental settings. Therefore, in this case, feed compositions of f_{11}=0.04 and f_{12}=0.30 were suggested by the optimal design criterion for the initial optimal trials.

Optimal design of experiments (steps 5-11, figure 1)

New optimal experiments were performed at the specified feed compositions in 20% DMF (by weight). The early stages of copolymerization are dominated by VBK incorporation, since r_{VBK} is larger than r_{MMA}. As conversion increases, the amount of VBK in the copolymer decreases. To capture this information, experiments should be run up to relatively medium-high conversion (otherwise, this information is completely lost at low conversion level trials with copolymer chains rich in VBK). The maximum achievable conversion for the new trials was 40% and six samples per experiment were collected at different conversion values. Table 3 shows the new optimal data set. Independent replicates were also conducted and are included in the table as well.

Using the inital optimal data set, the reactivity ratios were re-estimated and the results are shown in table 2 (third row), as well as in figure 2. The JCR of the new reactivity ratios (solid line) is considerably smaller. The optimally designed experiments are expected to provide the best set of observations for the purpose of parameter estimation and, therefore, the point estimates would show a lower amount of uncertainty (smaller JCR). This part of the analysis clearly proves our point that implementing the EVM framework increases the certainty of the results and in turn our understanding about the true (yet unknown) values of these reactvity ratios.

Table 3. Optimally designed data (initial design scheme)

$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$	$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$
5.6	0.04	0.08	2.9	0.29	0.55
17.2	0.04	0.07	14.8	0.29	0.42
22.5	0.04	0.05	15.4	0.29	0.51
38	0.04	0.04	16.1	0.29	0.49
32.6	0.04	0.05	19	0.29	0.50
36.5	0.04	0.05	28.9	0.29	0.52
1.1	0.04	0.57	4.3	0.3	0.61
6.9	0.04	0.13	15.6	0.3	0.36
7.9	0.04	0.11	18.3	0.3	0.41
11.3	0.04	0.10	20.3	0.3	0.41
14.7	0.04	0.09	23.5	0.3	0.42
			26.8	0.3	0.42

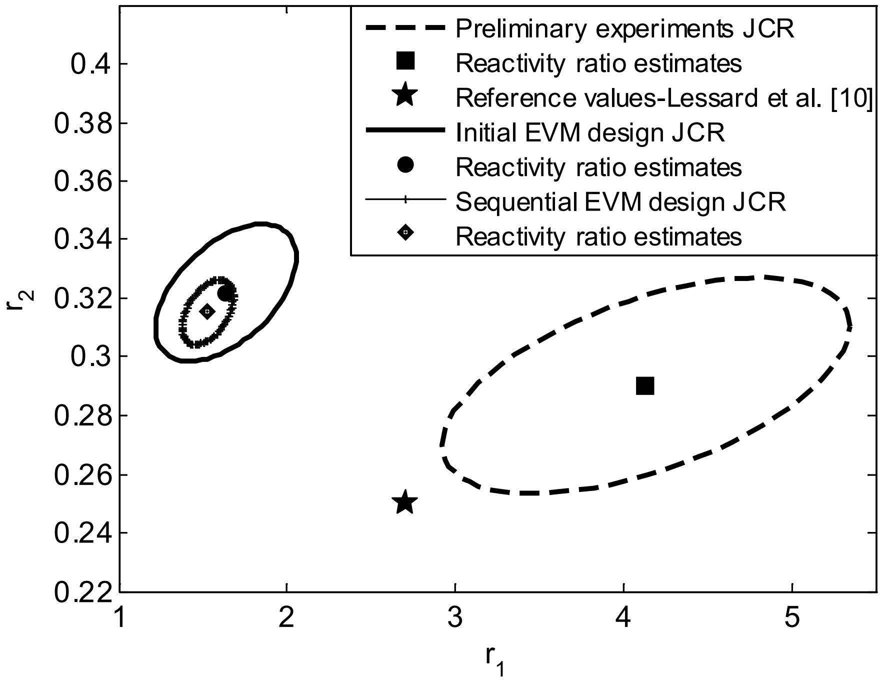

Figure 2. Preliminary, initial and sequential optimal reactivity ratio estimates (M_1=VBK, M_2=MMA)

To complete the analysis, the next step was taken in order to study the effect of including the next sequentially optimal experiment as well. The next sequentially optimal trial, given by the EVM design criterion,[9] was at the feed composition of f_1=0.45. Very similar to the initial set of optimal experiments, six samples per experiment at different conversion levels were run. Table 4 shows these data points. These data points were added to those of initial optimal experiments and the combined data set (tables 3 and 4) was re-analyzed by the DNI approach for updating the reactivity ratios.

Table 4. Optimally designed data (sequential design scheme)

$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$
3.7	0.45	0.42
5.4	0.45	0.61
10.2	0.45	0.61
7.7	0.45	0.60
6.5	0.45	0.76
7.8	0.45	0.67

The final set of estimated reactivity ratios is included in table 2 (fourth row) and also in figure 2 along with the corresponding JCR. Based on the results in figure 2, it can be clearly seen that using f_1=0.45, as the next sequential experiment, resulted in the smallest JCR (i.e., the most reliable reactivity ratio estimates among all analyzed cases). This was also in complete agreement with the previously obtained reactivity ratios. For all these cases, the sizes of JCRs are considerably smaller than JCRs obtained from initially designed (or non-designed) experiments, thus validating our premise that performing sequentially optimal experiments improves the quality of reactivity ratios, thus making it a very useful step in the process of determining reliable reactivity ratios.

An important aside: comparison between the EVM and Tidwell-Mortimer design criteria

As the last part of the analysis for this case study, we decided to compare experimentally the performance of the relatively new EVM design criterion with the Tidwell-Mortimer design criterion. For this purpose, the Tidwell-Mortimer optimal feed compositions were calculated using equation (10) and then experiments were performed with those feed compositions at f_{11} =0.13 and f_{12}=0.33 (again, f_{11} refers to the mole fraction of VBK for the first trial, whereas f_{12} refers to the mole fraction of VBK for the second trial). The data points from this stage of experimentation are shown in table 5. Using these data points, the reactivity ratios were re-estimated and the results are shown in Table 2 (last row) as well as in figure 3, along with their JCR (dashed line). This figure also shows the previously obtained reactivity ratios from the initial EVM optimal experiments with their JCR (solid line). As can clearly be seen in this figure, the Tidwell-Mortimer design reactivity ratios have a larger JCR than the EVM design reactivity ratios, indicating that these point estimates are not as precise as the EVM design reactivity ratios. This first-ever comparison perfectly supports our claims about the superiority of the EVM design criterion for the purpose of estimating reactivity ratios, which is related to the fact that the EVM statistical framework is the most appropriate methodology for the problem of reactivity ratio estimation.

Table 5. Designed data using Tidwell-Mortimer criterion (initial design scheme)

$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$	$X_W(wt\%)$	$(f_0)_{VBK}$	$\overline{F}_{VBK}$
5.43	0.13	0.29	2.38	0.33	0.74
12.04	0.13	0.23	4.22	0.33	0.59
19.40	0.13	0.19	6.21	0.33	0.57
20.22	0.13	0.21	29.24	0.33	0.26
44.83	0.13	0.18	9.08	0.33	0.48
6.33	0.13	0.31	24.60	0.33	0.33
8.71	0.13	0.25	3.56	0.33	0.72
9.25	0.13	0.24	5.68	0.33	0.46
23.15	0.13	0.17	3.74	0.33	0.57
15.81	0.13	0.25	3.59	0.33	0.63
19.53	0.13	0.27	5.64	0.33	0.57
			33.95	0.33	0.30

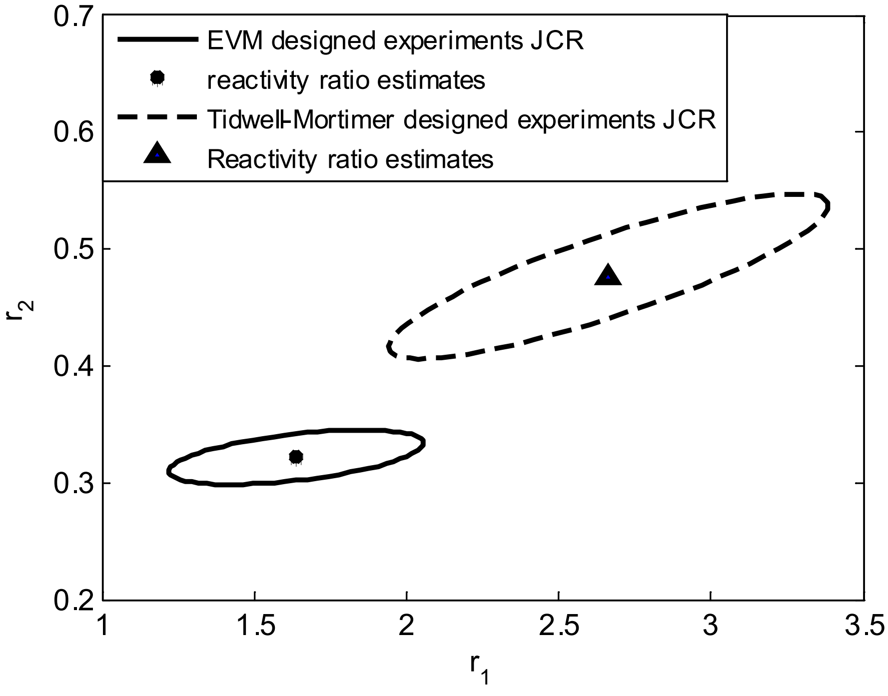

Figure 3. EVM vs. Tidwell-Mortimer reactivity ratio estimates (M_1=VBK, M_2=MMA)

CONCLUSIONS

We have discussed statistically correct EVM estimation procedures and extended them to include optimal design of experiments for reactivity ratio estimation in copolymerization. The methodology is applicable to terpolymerizations as well. For a comprehensive approach to yield reliable reactivity ratio values, several sequential steps and practical prescriptions have been suggested: (a) reactivity ratio estimation from screening experiments (typically but not necessarily at low conversion levels), (b) optimal design of experiments, (c) reactivity ratio estimation from full conversion range experiments, and (d) sequential design of experiments and re-estimation of reactivity ratios and diagnostic checks. This complete methodology, which should become the norm, especially in copolymerization or terpolymerization kinetic studies, has subsequently been experimentally demonstrated and verified with the new and largely unstudied VBK/MMA nitroxide-mediated copolymerization system. The estimated optimal reactivity ratios are r_{VBK}=1.53 and r_{MMA}=0.32, whereas the 95% joint confidence region for these values is shown in figure 2.

Corresponding Author

Institute for Polymer Research (IPR), Department of Chemical Engineering, University of Waterloo, Waterloo, Canada, N2L 3G1

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally. Niousha Kazemi, Benoît H. Lessard, Milan Marić, Thomas A. Duever, and Alexander Penlidis*

ACKNOWLEDGEMENTS

We kindly acknowledge OMNOVA Solutions USA/United Way Worldwide, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Canada Research Chair (CRC) program for financial support. BHL would also like to thank the Government of Canada for a Banting post-doctoral fellowship. MM would also like to thank the CFI New Opportunities Fund and Arkema Inc. for donating the BlocBuilder and SG1.

REFERENCES

(1) Dube, M.; Sanayei, R. A.; Penlidis, A.; O'Driscoll, K. F.; Reilly, P. M. A Microcomputer Program for Estimation of Copolymerization Reactivity Ratios. *J. Polym. Sci., Part A: Polym. Chem.* **1991**, 29, 703.

(2) Polic, L.; Duever, T. A.; Penlidis, A. Case Studies and Literature Review on the Estimation of Copolymerization Reactivity Ratios. *J. Polym. Sci., Part A: Polym. Chem.* **1998**, 36, 813.

(3) Van Der Meer, R.; Linssen, H. N.; German, A. L. Improved Methods of Estimating Monomer Reactivity Ratios in Copolymerization by Considering Experimental Errors in Both Variables. *J. Polym. Sci., Polym. Chem. Ed.* **1978**, 16, 2915.

(4) Patino-Leal, H.; Reilly, P. M.; O'Driscoll, K. F. On the Estimation of Reactivity Ratios. *J. Polym. Sci., Polym. Letter Ed.* **1980**, 18, 3, 219.

(5) O'Driscoll, K. F.; Reilly, P. M. Determination of Reactivity Ratios in Copolymerization. *Makromol. Chemie. Macromol. Symp.* **1987**, 10-11, 355.

(6) Tidwell, P. W.; Mortimer, G. A. An Improved Method of Calculating Copolymerization Reactivity Ratios. *J. Polym. Sci., Part A: General Papers* **1965**, 3, 369.

(7) Kazemi, N.; Duever, T. A.; Penlidis, A. Reactivity Ratio Estimation from Cumulative Copolymer Composition Data. *Macromol. React. Eng.* **2011**, 5, 385.

(8) Kazemi, N.; Duever, T. A.; Penlidis, A. A Powerful Estimation Scheme with the Error-In-Variables-Model for Nonlinear Cases: Reactivity Ratio Estimation Examples. *Comp. & Chem. Eng.* **2013**, 48, 200.

(9) Kazemi, N.; Duever, T. A.; Penlidis, A. Design of Experiments for Reactivity Ratio Estimation in Multicomponent Polymerizations Using the Error-In-Variables-Model Approach. *Macromol. Theory Simul.* **2013**, 22, 5, 261.

(10) Lessard, B. H.; Ling, E. J. Y.; Morin, M. S.; Marić, M. Nitroxide-Mediated Radical Copolymerization of Methyl Methacrylate Controlled with a Minimal Amount of 9-(4-Vinylbenzyl)-9H-Carbazole. *J. Polym. Sci., Part A: Polym. Chem.* **2011**, 49, 1033.

(11) Reilly, P. M.; Patino-Leal, H. A Bayesian Atudy of the Error-In-Variables Model. *Technometrics*. **1981**, 23, 3, 221.

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

(13) Box, G. E. P.; Lucas, H. L. Design of Experiments in Non-Linear Situations. *Biometrika*. 1959, 46, 77.

(14) Keeler, S. E. PhD Thesis, Department of Chemical Engineering, University of Waterloo, 1989.

(15) Box, M. J. Some Experiences with a Nonlinear Experimental Design Criterion. *Technometrics*. **1970**, 12, 569.

(16) Edgar, T. F.; Himmelblau, D. M.; Lasdon, L. S. *Optimization of chemical processes*, McGraw-Hill, New York, 2001.

(17) Lessard, B. H.; Ling, E. J. Y.; Marić, M. Fluorescent, Thermoresponsive Oligo (ethylene glycol) Methacrylate/9-(4-Vinylbenzyl)-9 H-carbazole Copolymers Designed with Multiple LCSTs via Nitroxide Mediated Controlled Radical Polymerization. *Macromolecules,* **2012**, 45, 1879.

(18) Lessard, B. H.; Savelyeva, X.; Marić, M. Smart Morpholine-Functional Statistical Copolymers Synthesized by Nitroxide Mediated Polymerization. *Polymer*, **2012**, 53, 5649.

(19) Lessard, B. H.; Marić, M. Water-Soluble/Dispersible Carbazole-Containing Random and Block Copolymers by Nitroxide-Mediated Radical Polymerisation. *Can. J. Chem. Eng.* **2013**, 91, 618.

(20) Lessard, B. H.; Guillaneuf, Y.; Mathew, M.; Liang, K.; Clement, J. L.; Gigmes, D.; Hutchinson, R. A.; Marić, M. Understanding the Controlled Polymerization of Methyl Methacrylate with Low Concentrations of 9-(4-Vinylbenzyl)-9 H-carbazole Comonomer by Nitroxide-Mediated Polymerization: The Pivotal Role of Reactivity Ratios. *Macromolecules* **2013**, 46, 805.