Filipovic, N., Scott, A. J., & Penlidis, A. (2021). Hierarchical Data Analysis for the Characterization of Polymeric Materials: Linking Measurements and Statistical Methodology. Chemical Engineering Education, 55(1), 11–22. https://doi.org/10.18260/2-1-370.660-122355

Hierarchical Data Analysis for the Characterization of Polymeric Materials: Linking Measurements and Statistical Methodology

*Nicholas Filipovic, Alison J. Scott and Alexander Penlidis**

Institute for Polymer Research, Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada
E-mail: penlidis@uwaterloo.ca

Abstract

As future chemical engineers, it is important that students be able to identify and quantify sources of error. A statistical analysis technique that is often overlooked is hierarchical design methodology, which allows for the separation of overall variance into several related components. While hierarchical methodology is relevant to many fields, we have demonstrated that it can be taught through the synthesis and characterization of polymeric materials. Basic statistical concepts are described, along with relevant examples.

Keywords

Characterization, Hierarchical Design, Polymers, Statistics

INTRODUCTION

Hierarchical or nested design methodology helps engineers to identify different sources of variation within their data. Essentially, the methodology can be viewed as a variance decomposition technique, where the overall variance is separated into several components; the goal is to locate the most significant sources of variance. For any process with multiple steps or stages, it can be useful to know whether the variance is equally a result of all operating stages, or if select process steps are contributing most of the variance.

The hierarchical design methodology and subsequent analysis is very general, and can be applied to many fields of study. However, it is often overlooked in the chemical engineering undergraduate curriculum. We would suggest that it is a valuable tool for students to add to their background, and that it can be taught alongside other chemical engineering concepts to make good use of precious teaching time. In addition to expanding their knowledge base, students can also develop improved problem analysis and investigation skills, gain laboratory experience, and advance their communication skills.

The general concept can be introduced to students with a straightforward thought experiment: consider synthesizing some material and then analyzing the material using a property characterization technique in the lab. If we replicate the synthesis process and the characterization technique several times, we will not always obtain exactly the same outcome! Common sense dictates that there will be variability observed between genuine, independent replicates. Variability can be imparted to the measured property from several possible sources of error; students can

likely identify most of these themselves. Sources of error may include random fluctuations in the operating conditions between batches/reactors, heterogeneity in the reactor as samples are collected, inconsistencies in the analytical technique, and so on.

The original motivation for integrating chemical engineering concepts (specifically polymer reaction engineering concepts) and the hierarchical design methodology came about during experimental design and data analysis in graduate student research. Each experimental stage of polymer synthesis and characterization can introduce new sources of error, and this provides a very tangible way for students to identify and quantify potential variability. Gradually and progressively, the same methodology was introduced in other settings including undergraduate student research projects, senior design projects, and lab data analysis in statistics courses. The most recent iteration of this approach was in the context of an independent research project course. As such, the instructor team and the participants had the flexibility of shifting between the academic/theoretical side and the experimental/laboratory side of the project.

This background is intended to provide some historical context, but the approaches used thus far should by no means be seen as the only methods of delivery. In fact, the methodology that is described in what follows is very versatile; it could be used as part of an undergraduate laboratory course, a lecture-based statistics course, a senior undergraduate research project, or in different stages of graduate student research. In order to ensure that readers see potential to use this approach in a variety of settings, we have kept the contextual details rather general. Of course, individual instructors could adapt the project at their discretion, especially given the diversity of student backgrounds, laboratory capabilities and course timelines.

As instructors and/or researchers, we could encourage students to explore the power of hierarchical design methodology through statistical design of experiments, synthesis of polymeric materials, and/or subsequent characterization steps. The real-world application of a seemingly complicated statistical analysis methodology can help students to understand the relevance of the approach, to recognize the methodical simplicity of the analysis steps, and (more importantly) to appreciate inherent variability in experimental work. It is our hope that the description of the methodology and the examples that follow will provide instructors with the tools that they need to integrate these important topics into undergraduate (and graduate) chemical engineering courses.

PROJECT DESCRIPTION

Hierarchical experimental designs published by Dubé et al.[1] and D'Agnillo et al.[2] have investigated the reliable measurement of error at different steps of polymer synthesis and characterization. Their studies demonstrated that important sources of error in such investigations include the polymerization process, sample heterogeneity, and inconsistencies in characterization (specifically gel permeation chromatography, GPC). Polymerizations do not necessarily occur homogeneously in a reactor; depending on which part of the reactor the sample is taken from, there may be variability. For example, a different viscosity distribution may occur due to heterogeneous mixing distribution. Furthermore, identical measurements from GPC are not expected, even for identical samples, due to random variability from test to test. Fortunately, using a hierarchical experimental design, it is possible to quantify different sources of variance by taking replicate measurements at each nested level.

The main project described herein and further illustrated in Example 1 involves the synthesis of polymeric materials and the subsequent determination of polymer molecular weight averages. The investigation includes four different experimental steps (or four "levels") where error might be introduced. See also Figure 1:

(1) The preparation of concentrated "stock solutions", which are pre-established formulations with monomers in solution.

(2) The adjustment of each "pre-polymer solution" to achieve desirable reaction conditions (pH modification, for example) and the subsequent polymerization.

(3) The collection of several samples from each polymerization.

(4) The preparation of polymer samples for molecular weight analysis via dissolution and the characterization process itself via GPC.

By convention, the lower levels are said to be nested in the higher levels. Thus, the lowest level in a nested design is usually the measurement itself; in this case it refers to the GPC analysis. As shown in Figure 1, the GPC analysis results (tests) in this study are nested within the samples, which are nested within the solution and synthesis step, which are in turn nested within the different formulations (monomer composition in the initial stock solution).

For each stock solution, at least two independent replicates are required at each step; note that Figure 1 shows three independent replicates at the GPC level. In theory, the number and nature of experimental steps ("levels") could vary as well, but the process is described here to give a sense of the project's scale. In any case, once the specific experimental steps are identified, students need

to become familiar with each process so that they can hypothesize the potential sources of error. Familiarization can be accomplished through a combination of literature searches and in-lab training; helpful resources include these references[1-4] for the statistical background and these references[5,6] for the experimental synthesis and characterization background.

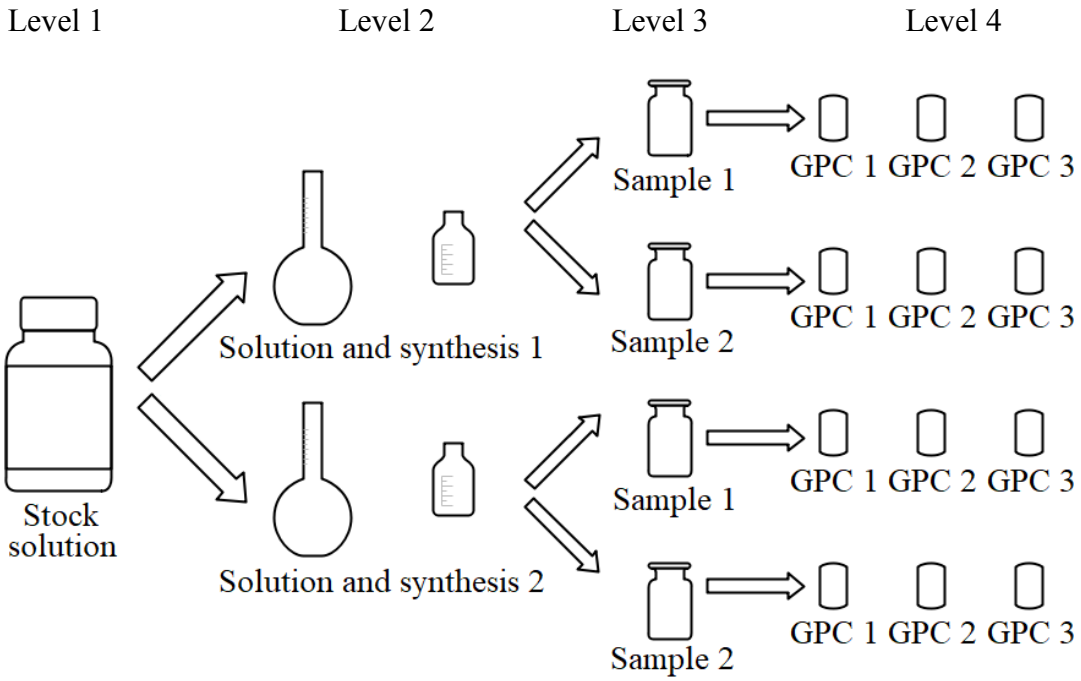

Figure 1: *Example hierarchical design for the synthesis and analysis of polymeric materials.*

In step 1, for example, students are assigned a particular stock solution formulation. They are able to prepare the solution using straightforward lab procedures including using molar concentrations and volumes to determine mass, weighing monomers, transferring monomers to volumetric flasks, and dissolving monomers in a pre-specified volume of water. Most students recognize that intentional variation may occur with varying stock solution recipes, but that unintentional, inherent error may also be introduced during the weighing and transferring of monomers into the volumetric flasks. In this case, terpolymers of 2-acrylamido-2-methylpropane sulfonic acid, acrylamide and acrylic acid are the product of choice,[5] and the monomer quantities in the stock solution are

intentionally varied between investigations; more details will be provided in Example 1. However, the same approach could be applied to any number of other polymerization studies.

Similarly, students identify sources of variation in preparing their stock solutions for synthesis (step 2), as they adjust the reaction conditions, add initiator, separate the solution into smaller aliquots, and place their samples in a warm shaker bath. Separation into several smaller aliquots allows for the synthesis of the same polymer product to occur in several different vials simultaneously. More experimental details have been provided elsewhere.[5,6] Step 3 requires students to remove samples from the water bath at pre-specified times and to stop the polymerization reaction using ice and/or an inhibitor injection. As they isolate the samples and allow them to dry, they are tasked with identifying additional sources of error in the experimental process. This step is intended to establish the consistency of the polymerization, including the equal distribution of pre-polymerization solution components and the repeatability of the polymer isolation process, across several simultaneously synthesized polymer samples.

Finally, in step 4, polymers are prepared for molecular weight analysis via gel permeation chromatography (GPC). Since the polymeric material obtained is in powder form, small quantities of the polymer must be dissolved in a pH 7 buffer liquid (mobile phase), filtered, and injected into the GPC.[6] At this final stage, students may identify long dissolution times, difficult sample filtration, randomized sampling order, and day-to-day variability as some of the potential sources of error.

Depending on time allotted for the project or lab session, students may collect experimental data themselves or the data collection may be divvied up and assigned to smaller groups. For example, instructors might consider one formulation per group, or even one "level" per group, where one group of students focuses on stock solution preparation while other students focus on GPC. Or, if time is extremely limited, students may even evaluate pre-existing data sets (see Example 3 in what follows). However, it is important for students to understand where all of the experimental information comes from, even if they do not collect the data themselves. Inevitably, if students are not solely responsible for collecting experimental data, they may try to identify the primary source of error as "group-to-group variability" or "operator error". While this is a relevant source of error, it is by no means the only contributing factor. Thus, to ensure that students fully explore the potential sources of error, a related group brainstorming activity is recommended. This discussion would best be placed after data collection (or, at least, after reviewing the experimental procedure in a case study) and before data analysis, so that sources of variability are informed by physical observations. Of course, it would also be beneficial to revisit the brainstorming activity after analysis, as time allows, to ensure that the results make physico-chemical sense.

Statistical Background

Any instructor wanting to introduce this type of project will need some background in statistics. The basic analysis steps are presented herein, but interested readers may want to refer to standard statistics textbooks[3,4] for additional information. In this section, generalized equations are provided for context, but the examples shown in what follows provide more concrete applications of the statistical analysis procedure.

In order to keep track of the experimental levels, it can be helpful to refer to each level generally from highest to lowest in alphabetical order (i.e., as per Figure 1, formulation = A, synthesis = B, sample = C and GPC = D). We can decompose or partition the total variability into the parts assignable to the various sources of error by calculating a sum of squares for each level of nesting. The variances associated with each level/step/part/component are designated herein as m_A, m_B, m_C, and m_D. Each observation is defined as y_{abc1}, y_{abc2}, …, y_{abcd}, where there are D replicated analytical tests made on the C^{th} sample, B^{th} synthesis and A^{th} formulation. The mean squared error at the lowest level of a nested design, m_D in this case, is defined as the pure error mean square, [2] and it should be calculated first (as per Eq. 1).

$$m_D = \sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\sum_{d=1}^{D}\frac{\left(y_{abcd}-\bar{y}_{abc}\right)^2}{ABC(D-1)} \quad (1)$$

In Eq. 1, $\bar{y}_{abc}$ is an average of all analytical tests at the C^{th} level. That is, $\bar{y}_{abc}$ is the average of GPC outputs (measurements) for a specific sample, which was in turn prepared from a specific stock solution and synthesis process. Since m_D is the lowest level of the design, it is an unbiased estimate of $\hat{\sigma}_D^2$, which is the component variance due to the GPC step alone. $\hat{\sigma}_D^2$ has ABC(D-1) degrees of freedom and is used in conjunction with a level of significance (related to statistical confidence), α, to obtain an error band for the instrument. As long as the data points are normally distributed, the instrument error can be expressed as $\pm 1.96\sqrt{m_D}$ at 95% confidence within the range of the experiment.

The purpose of doing a nested experiment is to obtain a measurement of the variance at every level where error can be introduced. To solve for the variance in the samples, the mean square must be calculated for the next level, m_C, which is expressed according to Eq. 2.

$$m_C=\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\frac{C\left(\bar{y}_{abc}-\bar{y}_{ab}\right)^2}{AB(C-1)} \tag{2}$$

In Eq. 2, $\bar{y}_{ab}$ is the average of all measurements at the B^{th} level, for any independent synthesis. Using Figure 1 as a general example, $\bar{y}_{ab}$would be the average of all GPC measurements taken for sample 1 and sample 2 from a specific synthesis. Due to the nested nature of the experiment, m_C is not an estimator of $\hat{\sigma}_C^2$ alone but needs to be corrected according to Eq. 3.

$$\hat{\sigma}_C^2=\frac{m_C-m_D}{D} \tag{3}$$

The variance associated with the polymer synthesis step is the next (higher) level in the hierarchical design. To solve for the variance at this level ($\hat{\sigma}_B^2$), m_B can be calculated as per Eq. 4.

$$m_B=\sum_{a=1}^{A}\sum_{b=1}^{B}\frac{BC\left(\bar{y}_{ab}-\bar{y}_{a}\right)^2}{A(B-1)} \tag{4}$$

In Eq. 4, $\bar{y}_a$ is the average of all replicates for each formulation. It then follows that the component variance of the solution level is expressed according to Eq. 5.

$$\hat{\sigma}_B^2 = \frac{m_B - m_C}{CD} \tag{5}$$

217

218 The highest level of variability in this experiment is quantified by the mean squared error of the

219 formulation, m_A, which is calculated according to Eq. 6.

220

$$m_A = \sum_{a=1}^{A} \frac{ABC\left(\bar{y}_a - \bar{y}\right)^2}{(A-1)} \tag{6}$$

221

222 Here, $\bar{y}$ is the grand average, or the average of all observations. To correct for $\hat{\sigma}_A^2$ we use Eq. 7.

223

$$\hat{\sigma}_A^2 = \frac{m_A - m_B}{BCD} \tag{7}$$

224

225 In principle, this approach could continue to "N" levels. However, 3 to 4 levels or stages are

226 typical. The patterns are summarized in a generalized ANOVA table for clarity; see Table 1.

TABLE 1
Generalized ANOVA Table for a Nested Design with Four Levels

Source	Sum of Squares	Degrees of Freedom	MS	Expected Value of Mean Square (MS)	Component Variance Estimates
Average	$ABCD(\bar{y}^2)$	1			
Formulation	$ABC\sum_{a=1}^{A}(\bar{y}_a-\bar{y})^2$	A-1	m_A	$BCD\hat{\sigma}_A^2+CD\hat{\sigma}_B^2+D\hat{\sigma}_C^2+\hat{\sigma}_D^2$	$\hat{\sigma}_A^2=\frac{m_A-m_B}{BCD}$
Solution	$BC\sum_{a=1}^{A}\sum_{b=1}^{B}(\bar{y}_{ab}-\bar{y}_a)^2$	A(B-1)	m_B	$CD\hat{\sigma}_B^2+D\hat{\sigma}_C^2+\hat{\sigma}_D^2$	$\hat{\sigma}_B^2=\frac{m_B-m_C}{CD}$
Sample	$C\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}(\bar{y}_{abc}-\bar{y}_{ab})^2$	AB(C-1)	m_C	$D\hat{\sigma}_C^2+\hat{\sigma}_D^2$	$\hat{\sigma}_C^2=\frac{m_C-m_D}{D}$
GPC	$\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\sum_{d=1}^{D}(y_{abcd}-\bar{y}_{abc})^2$	ABC(D-1)	m_D	$\hat{\sigma}_D^2$	$\hat{\sigma}_D^2=m_D$
Total	$\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\sum_{d=1}^{D}(y_{abcd})^2$	ABCD			

After building an ANOVA table, the next step is to determine whether or not the variance is significant at each level. A series of sequential F-tests can establish the validity of the null hypothesis, on the basis of 95% confidence, to determine whether or not the error value at a given level might be zero.

The null hypothesis for the F-test is that the ratio of two variances (as in Eq. 8) is unity, or that the variance component ($\hat{\sigma}_i^2$) at the higher level does not provide a significant contribution to the overall variability. Therefore, if the $F_{obs} < F_{crit}$, where F_{obs} may be $F_{A/B}$, $F_{B/C}$, $F_{C/D}$, etc. as shown below, we fail to reject the null hypothesis. Thus, we can conclude that $\hat{\sigma}_i^2 = 0$ and that the error associated with the level being evaluated is not significant.

The alternate hypothesis is that $m_i > m_{i+1}$ ($m_C > m_D$, for example). If the variance at a higher design level is significantly larger than the next lowest level (if m_C is significantly larger than m_D, for example), then the variance component at that upper design level provides a significant contribution to the overall variability, and $\hat{\sigma}_i^2 > 0$.

$$F_{A/B} = \frac{m_A}{m_B} \tag{8a}$$

$$F_{B/C} = \frac{m_B}{m_C} \tag{8b}$$

$$F_{C/D} = \frac{m_C}{m_D} \tag{8c}$$

The F-probes (or F_{obs} values) shown in Eq. 8 all have degrees of freedom in the numerator ($\nu 1$) and denominator ($\nu 2$) according to their mean squared values; recall Table 1. If, for any level, $F_{i/i+1}$ is larger than the critical $F_{\nu 1,\nu 2}$ distribution, then level i is identified as a significant source of variability.

It is important to note that these hypothesis tests represent an overall analysis. F-testing cannot be used to determine whether a certain subset of replicates is statistically similar. For example, if GPC analysis was performed on "D" separate days and the data on a specific day was believed to be compromised, F-testing would only show that the "D" level showed significant variability; it could not be used to identify which day was introducing bias. In such cases, it may be of interest to remove all the data from that day, i.e., changing from a "A×B×C×D" to a "A×B×C×(D-1)" resolution experiment, and repeat the analysis. Alternatively, one might consider re-evaluating the

data using blocking; all data collected on a particular day could be subdivided into a block. In such a case, variability between days could be evaluated. However, by focusing on day-to-day variability, it would not be as straightforward to quantify variability due to formulations, solutions, and samples. Therefore, there are several "what-if" scenarios that one can investigate based on a specific dataset, depending on the intended outcome.

CASE STUDIES

To demonstrate the application of this project, three specific examples are presented in different levels of detail in what follows. These case studies are intended to clarify the analysis steps, and will give instructors some additional background if they would like to incorporate such a project into their courses.

The polymerization processes described herein are relevant to a variety of important applications; the complexity of each case is representative of a real-world problem. These cases are intentionally non-trivial, and should be appropriate for upper year undergraduate students. We have highlighted multi-component polymers and polyelectrolytes (Example 1), crosslinked polymers (Example 2), and high-temperature GPC for polyolefin characterization (Example 3). Exploring such processes promotes critical thinking and provides valuable troubleshooting opportunities for students. These complications make the analysis more realistic, which we believe increases students' motivation and enhances their ability to apply these concepts in real-world situations.

Example 1

The first example highlights the terpolymerization of 2-acrylamido-2-methylpropane sulfonic acid (AMPS), acrylamide (AAm) and acrylic acid (AAc). AMPS/AAm/AAc is a water-soluble polymer that can be used as a viscosity modifier in chemical enhanced oil recovery, and the effectiveness of the viscosity modification is dependent on the molecular weight averages of the polymeric material. Thus, there is real-world motivation to obtain accurate molecular weight averages for the materials produced; it is important to know which steps of the synthesis and characterization process are introducing the most error.

As described generally earlier, the polymerization of AMPS/AAm/AAc can be broken down into four main steps: stock solution preparation for a pre-specified formulation, pre-polymerization solution preparation and synthesis, sampling, and characterization (GPC). As shown in Figure 2, the project included three unique formulations, which are arbitrarily labeled J, K and L. The synthesis of each formulation was independently replicated (synthesis replicates are designated by "R"), and two samples were taken from each synthesis. Finally, the molecular weight average of each sample was characterized via GPC three times. For each GPC characterization, an aliquot was dissolved in the mobile phase (pH 7 buffer) over several days, filtered, and transferred into a single GPC vial. The entire sample preparation process, from taking an aliquot to filling the GPC vial, was repeated for each test. Thus, each GPC injection was from a unique GPC vial; three GPC vials were used for each sample, and twelve vials were used for each formulation. Characterization occurred in random order over the course of three days, with daily recalibration of the system using well-characterized standards.

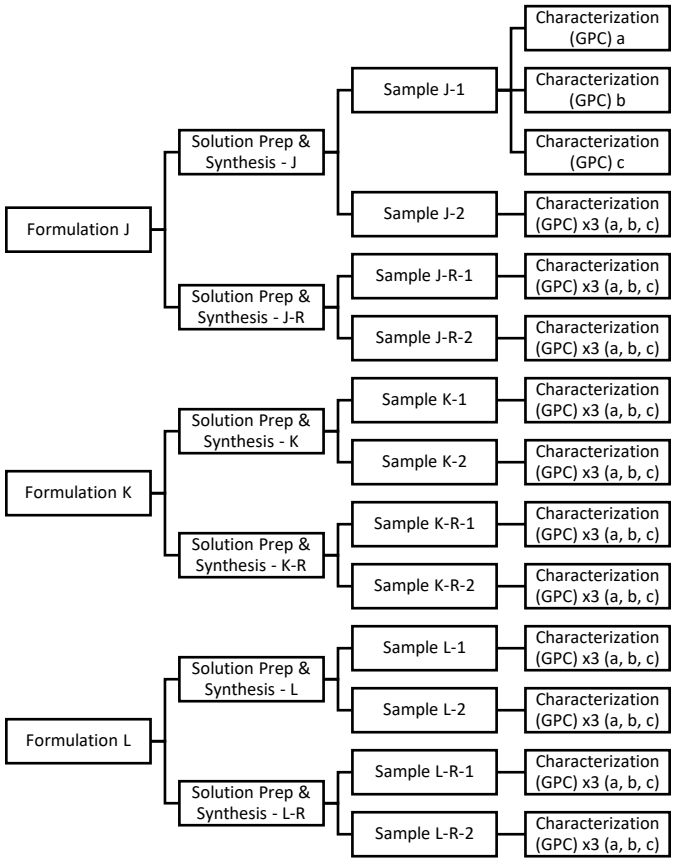

Figure 2: *Four-stage nested design for the terpolymerization of AMPS/AAm/AAc.*

For this investigation, the formulations (at the highest level) were intentionally varied, as shown in Table 2. Varying formulations provided information about how the initial concentrations of component monomers might affect the molecular weight (or other properties not discussed herein) of the resultant terpolymer. However, all subsequent steps, namely synthesis, sampling and GPC, were kept consistent to the extent possible. Experimental details have been provided elsewhere.[5,7]

TABLE 2
Experimental Conditions for Terpolymerization Formulations

Formulation	$f_{AMPS,0}/f_{AAm,0}/f_{AAc,0}$	[M] (mol/L)	[I] (mol/L)
J	0.20/0.40/0.40	1.0	0.004
K	0.21/0.69/0.10	1.5	0.009
L	0.10/0.75/0.15	1.5	0.009

$f_{i,0}$ = initial mole fraction of monomer i, [M] = overall monomer concentration, [I] = initiator (4,4'-azobis(4-cyanopentanoic acid)) concentration

Once the synthesis of all samples and the subsequent characterization was completed, students were tasked with selecting which dataset or datasets to work with. Unlike typical chemicals whose molecules all have the same molecular weight, the molecular weights of polymers are typically not uniform; polymerization reactions create chains which generally have different lengths and configurations leading to different molecular weights. Thus, GPC analysis provides the determination of several molecular weight averages, including number-average molecular weight ($\bar{M}_n$), weight-average molecular weight ($\bar{M}_w$), and peak molecular weight (M_p), as well as the polydispersity index (PDI) and the bulk intrinsic viscosity. Therefore, students investigated the relevance of each variable before selecting which dataset to work with.

One student justified their decision to analyze M_p as follows:

> "The terpolymer AMPS/AAm/AAc is known to have a relatively broad molecular weight distribution. The molecules in the very high molecular weight tail of the distribution may not even elute from the column, thus leading to an underestimation of $\bar{M}_w$ and PDI. [The underestimation] is due to electrostatic interactions between polyelectrolytes and GPC column internals, which were also observed for the copolymer AAm/AAc.[8] Since $\bar{M}_n$ emphasises the number of molecules in the injected samples (which is not changing), it is not the most reliable average. Hence, the most trusted representation was the peak molecular weight, M_p."

In general, most students recognized that M_p would provide the most useful data in this case, especially based on prior work in the area.[6-8] However, the same statistical analysis could be

performed on any of the other variables. A sample data set for M_p is used for the remainder of this example, but results would of course vary from one project/group to the next.

The next step was to evaluate the data from the 3×2×2×3 hierarchical characterisation of M_p, as per Figure 2. As described in the discussion surrounding Table 1, the generalized ANOVA table and related F-tests were employed. As shown in Tables 3 and 4, significant differences in variances were detected only at the formulation level.

TABLE 3
ANOVA Table for AMPS/AAm/AAc Study (A×B×C×D = 3×2×2×3)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Average	9.11×10^{13}	1	
Formulation	5.03×10^{10}	2	2.51×10^{10}
Solution	3.26×10^{9}	3	1.09×10^{9}
Sample	3.47×10^{10}	6	5.79×10^{9}
GPC	1.52×10^{11}	24	6.35×10^{9}
Total	9.13×10^{13}	36	

TABLE 4
F-Testing Results for AMPS/AAm/AAc Study (A×B×C×D = 3×2×2×3)

Type of Test	F_{obs}	F_{crit}	Reject null?
Sample/GPC	0.91	2.51	Fail to reject
Solution/Sample	0.19	4.76	Fail to reject
Formulation/Solution	23.13	9.55	Reject

The F-testing results in Table 4 provided initial evidence that significant differences in the polymer molecular weight were only caused/determined by the formulation recipe; this was as expected/predicted, since formulations J, K and L were intentionally varied.

The pure error variance, which is estimated by the mean squared error at the lowest level, was estimated here to be 6.35×10^{9} for the GPC instrument; see Table 3. Since replicate GPC

measurements of the same sample are assumed to be normally distributed, the pure error variance corresponds to an error for aqueous GPC of ±156,186 g/mol at 95% confidence in this experiment.

The polymer formulations K and L were more similar to each other than to formulation J, which was richer in acrylic acid. Also, formulation J had a lower total monomer molarity and initiator molarity; recall Table 2. In order to determine if smaller formulation differences could still be detected and to see if the solution and sample levels remained insignificant, students chose to repeat the analysis using a 2×2×2×3 experiment (with formulation J removed). Tables 5 and 6 show the ANOVA and F-testing for the reduced data set to detect the variation of M_p across experimental levels.

TABLE 5
ANOVA Table for AMPS/AAm/AAc Study (A×B×C×D = 2×2×2×3; Formulation J Removed)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Average	6.11×10^{13}	1	
Formulation	4.82×10^{10}	1	4.82×10^{10}
Solution	2.75×10^{9}	2	1.37×10^{9}
Sample	1.39×10^{10}	4	3.47×10^{9}
GPC	9.31×10^{10}	16	5.82×10^{9}
Total	6.13×10^{13}	24	

TABLE 6
F-Testing Results for AMPS/AAm/AAc Study (A×B×C×D = 2×2×2×3; Formulation J Removed)

Type of Test	F_{obs}	F_{crit}	Reject null?
Sample/GPC	0.60	3.01	Fail to reject
Solution/Sample	0.40	6.94	Fail to reject
Formulation/Solution	35.07	18.51	Reject

The results shown in Table 6 indicate that when similar formulations were being compared, there were still significant differences at the formulation level, but not in the solution/synthesis or the sample levels. The analysis results suggest that the initial monomer composition (i.e., the quantity

of each comonomer in the initial recipe) was a significant factor, since formulations K and L had the same total monomer concentration and the same initiator concentration (see Table 2).

As demonstrated in Tables 3 and 5, the mean squared error was fairly high at the GPC level. Thus, students hypothesized that the high GPC error may have masked the error in the solution and sample levels. Therefore, to further investigate error at the GPC level, calibration constants were revisited.

To account for any stochastic drift during characterization, the GPC had been recalibrated daily using well-characterized standards. While the calibration constants were similar from day to day, some fluctuation was observed. Therefore, as an alternative to applying different calibration constants each day (as had been done for the results reported thus far), all calibration constants measured over the course of about three days were averaged to allow for a more consistent calibration from day to day. The pooled calibration reduced the day-to-day variability that would be hidden within the GPC replicates.

With the new pooled calibration data, ANOVA tables were reproduced and F-testing was revisited. The full 3×2×2×3 experiment (recall Tables 3 and 4) and the reduced data set (where formulation J was excluded to leave a 2×2×2×3 resolution design; recall Tables 5 and 6) were both re-evaluated using the pooled calibration dataset.

The results of the analysis are not included herein for the sake of brevity, but students found that repeating their ANOVA table calculations using pooled calibration data for both the full 3×2×2×3

386 experimental design and the reduced 2×2×2×3 experimental design led to a large reduction of error
387 at the GPC level (GPC mean square). Using the pure error variance from the GPC level, the error
388 band for aqueous GPC was determined to be ±138,737 g/mol at 95% confidence, which was
389 approximately 10% less than the error band obtained with the daily recalibrated (original) data set.
390 The decrease in variance obtained using pooled calibration data suggests that recalibrating the
391 GPC daily introduced error; daily calibration may have been overcorrecting for day-to-day
392 variation, since there should not have been any considerable drift in the laboratory at that time.
393 The analysis of pooled calibration data still confirmed the results obtained earlier, as the only
394 significant variance was observed between formulations.
395
396 ***Potential Extension: Sensitivity Analyses***
397
398 An interesting extension would be to use the collected data to confirm that hierarchical design
399 strategies are capable of detecting differences in molecular weight for different experimental
400 levels, not just the formulation level that was observed from the AMPS/AAm/AAc experimental
401 data. Therefore, one might add a range of biases to a subset of the experimental data. For
402 demonstration purposes, a molecular weight bias was added to all data stemming from the first
403 solution/synthesis of formulation L (recall Figure 2). The goal here was to determine at which
404 point our hierarchical design would be able to detect a significant difference at the
405 solution/synthesis level when experimental data from all three formulations were included.
406
407 As shown in Figure 3, it is possible to compare the calculated F-probe (F_{obs}) with different
408 molecular weight biases, constantly comparing the resulting values to the critical F value (recall

Eq. 8); this analysis was first performed using the original, daily recalibrated data from the full 3×2×2×3 experiment. Figure 3 shows that if data coming from the first solution/synthesis of the L formulation had peak molecular weights approximately 150,000 g/mol higher than what was observed experimentally, there would be statistically significant differences at the solution/synthesis level. Such a molecular weight difference could easily occur experimentally, especially if the solution preparation process and subsequent synthesis are not carefully handled. Consider, for example, the impact of a miscalculated reaction time or an incorrectly set temperature controller. The simulation confirms that the hierarchical design of experiments would identify such sources of error if they impacted the peak molecular weight by at least 150,000 g/mol.

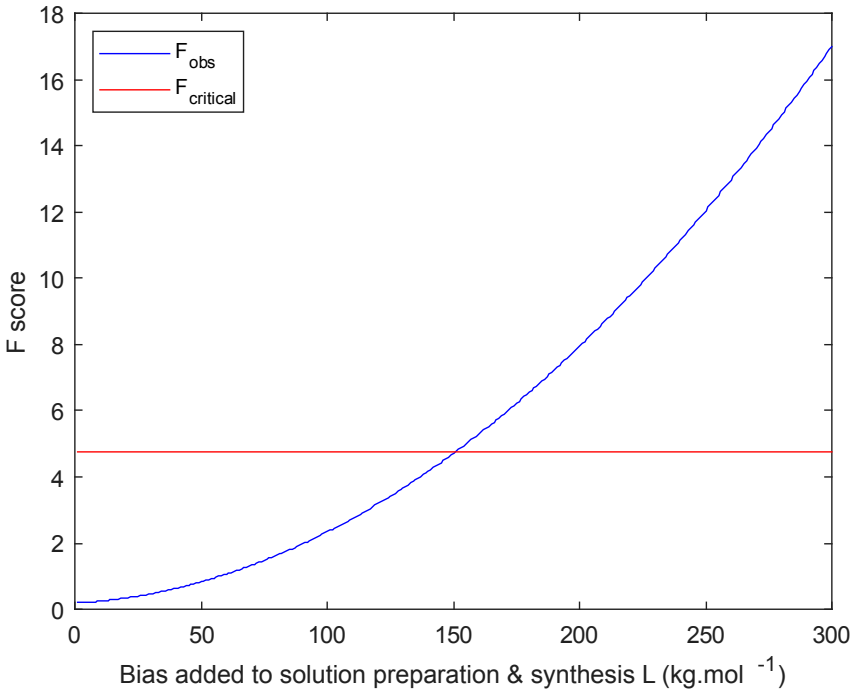

Figure 3: *Sensitivity analysis where the first solution preparation/synthesis data of formulation L are intentionally biased (3×2×2×3).*

The point at which the solution/synthesis level becomes significant at 95% confidence is 150 kg/mol (150,000 g/mol). Graphically, this is the crossover point. Interestingly, this value is almost

exactly the same as the instrumental error (at 95% confidence) that we obtained for the GPC step, which was ±156,186 g/mol. The result may be coincidental, but the fact that it is of the same order of magnitude as the instrument error is further evidence that hierarchical design strategies not only handle noise extremely robustly, but also detect true changes very efficiently. This type of extension allows students to think about their results in a meaningful way, and encourages brainstorming among students. A similar analysis could be performed for the other subsets of data described herein. Students might choose to look only at formulations K and L, or at the dataset obtained from the pooled calibration. Alternatively, the same type of sensitivity analysis could be applied to other data from the literature. Open-ended extensions like the ones described herein give students some additional autonomy over their work, which should further motivate their investigation.

The sensitivity analysis illustrated here has shown that even though the differences between the solution preparation/synthesis steps were insignificant for these AMPS/AAm/AAc syntheses, they could very quickly become significant factors. Had we not been able to see solution/synthesis level significance until we increased the molecular weight averages by several million (knowing that this polymer is not likely to experience that magnitude of variation in the lab), we may have drawn conclusions about the high error in the GPC and its ability to mask other sources of variability. However, this was not the case. Error at the GPC level was low enough to allow for detection of reasonable variation in the solution and synthesis level; however, these differences were simply not observed experimentally.

Example 2

Experiments conducted in this second study involved the nitroxide-mediated radical copolymerization of styrene (STY) and divinyl benzene (DVB) using N-tert-butyl-N-(2-methyl)-1-phenylpropyl)-O-(1-phenylethyl) hydroxylamine (TIPNO) as a unimolecular initiator; experimental details are provided elsewhere.[9] It is well-known that systems involving DVB are prone to crosslinking, involving the formation of gel materials that are difficult to deal with in the laboratory. Crosslinking and gelation could introduce a considerable amount of error in subsequent property characterization techniques.

As part of a systematic and comprehensive polymerization kinetic study, number- and weight-average molecular weights of the produced polymers were measured at various conversion levels using GPC. In this case, the hierarchical design was used to investigate the total variation in the molecular weight measurement from three important sources: error associated with the GPC measurement itself (analytical error), error related to the polymerization (reactor or process, carried out under identical conditions), and the variability in the measurements corresponding to different sampling times. In general, the same nested design approach was applied as in Example 1, but with three levels rather than four, and with said levels defined differently. Using a single formulation and focusing instead on sampling time, which is related to conversion level and hence to gel formation, allowed for a somewhat reduced experimental load for students.

Take, for example, the 3×2×2 experiment illustrated in Figure 4. Three samples taken out at 60 minutes, 420 minutes and 480 minutes (A = 3) were reproduced in a replicated polymerization/

synthesis (B = 2). For a sample from each polymerization, two independent GPC measurements were carried out (C = 2).

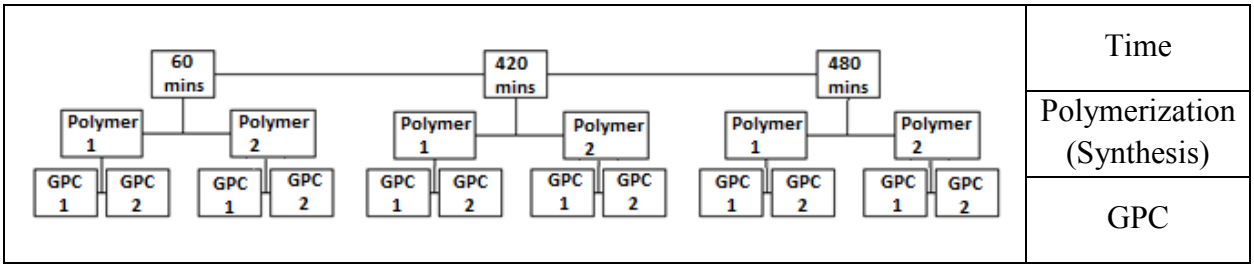

Figure 4: *Three-stage nested design for the copolymerization of STY/DVB.*

As with Example 1, students were tasked with learning about the experimental steps and identifying potential sources of error, then collecting the data themselves. In this case, since organic solvent-based GPC was used for the characterization (with tetrahydrofuran as solvent), both number-average molecular weight ($\bar{M}_n$) and weight-average molecular weight ($\bar{M}_w$) data were of interest; students could choose to analyze one or both of the data sets.

As shown in Tables 7 and 8, the pure error variance associated with only the GPC measurements was estimated to be 1.21×10^8 for $\bar{M}_n$ and 8.32×10^9 for $\bar{M}_w$. On the basis of a 95% confidence interval, this translates into an analytical error of ±21,553 g/mol for $\bar{M}_n$ and an error of ±178,815 g/mol for $\bar{M}_w$. This is the error solely based on the GPC measurements.

TABLE 7 ANOVA Table for STY/DVB Study (using $\bar{M}_n$ data)				
Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**Component Variance**
Average	2.64×10^{10}	1		
Time	1.26×10^{10}	2	6.30×10^{9}	1.37×10^{9}
Polymerization	2.46×10^{9}	3	8.20×10^{8}	3.49×10^{8}
GPC	7.26×10^{8}	6	1.21×10^{8}	1.21×10^{8}
Total	4.22×10^{10}	12		

487

TABLE 8 ANOVA Table for STY/DVB Study (using $\bar{M}_w$ data)				
Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	**Component Variance**
Average	2.59×10^{12}	1		
Time	2.60×10^{12}	2	1.30×10^{12}	2.59×10^{11}
Polymerization	7.92×10^{11}	3	2.64×10^{11}	1.28×10^{11}
GPC	4.99×10^{10}	6	8.32×10^{9}	8.32×10^{9}
Total	6.03×10^{12}	12		

488

489 Similarly, the error related to polymerization was found to be 3.49×10^{8} and 1.28×10^{11} for the
490 number- and weight-average molecular weights, respectively. These results are indicative of the
491 variability in the two polymers that were prepared, and thus, reflects the degree of inconsistency
492 in the preparation techniques. Finally, the error in the molecular weight measurements
493 corresponding to different times or conversion levels was 1.37×10^{9} and 2.59×10^{11} for number-and
494 weight-average molecular weights, respectively.

495

496 From this hierarchical analysis, it was clear that error caused by the GPC (lower level) was of the
497 lowest magnitude when compared to the other variables. Hypothesis testing was also conducted to
498 determine the impact of the different variables using an F-test. As described previously, the null
499 hypotheses were used to check if σ_A^2 (related to sampling time) and σ_B^2 (related to polymerization
500 replicates) were equal to zero. In this case, the hypothesis testing on σ_A^2 failed to reject the null
501 hypothesis of $\sigma_A^2 = 0$ for both $\bar{M}_n$ and $\bar{M}_w$. The hypothesis test outcome suggests that the error

associated with the different reaction times does not significantly contribute to overall variability. On the other hand, the null hypothesis of $\sigma_B^2 = 0$ was rejected for both $\overline{M}_n$ and $\overline{M}_w$. Therefore, there is strong evidence to conclude that the polymerization error contributes significantly to the overall error.

These results suggest that synthesis steps and/or conditions may be introducing variability. For the copolymerization of STY/DVB, crosslinking and gelation are known to introduce inaccuracies in molecular weight determination; this may have contributed to the error. Although the samples are drawn at the same times during polymerization, the conversion may have varied from run to run, and the synthesized polymer characteristics may vary considerably as a result. This would be of particular concern if samples were collected at/near the gel point.

In any case, students must be called upon to think critically about their results, reconciling physico-chemical explanations with their analysis results. The physico-chemical piece is an important aspect of the project; students should enhance their statistical background as well as improve their understanding of polymerization processes and related characterization steps.

Example 3

Our final example uses data from the literature, originally reported by D'Agnillo et al.[2] This type of case study, as illustrated in Figure 5, may be used in one of two ways. When instructors have dedicated ample time to this type of project, the data may be analyzed as a "first step" to confirm that the statistical analysis approach is well-understood; this can be done in parallel to experimental work for another study. Alternatively, if time is more limited, the case study alone would be

sufficient to introduce the concept of hierarchical data analysis and polymer characterization. However, to ensure that students can appreciate sources of experimental error, some exposure to laboratory techniques would still be extremely beneficial.

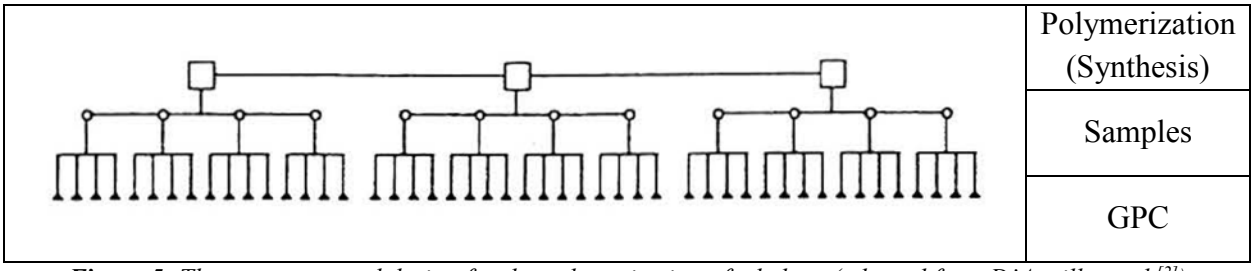

Figure 5: *Three-stage nested design for the polymerization of ethylene (adapted from D'Agnillo et al.[2])*

The $\bar{M}_w$ data from D'Agnillo et al.[2] could be given to students for a preliminary analysis, and then students could compare their analysis results to the published ANOVA table, F-testing results, and so on. The analysis was recently confirmed by one of our students, and results were in excellent agreement with the original publication. For the sake of brevity, the interested reader can consult the specific reference.[2]

LESSONS LEARNED

The hierarchical design methodology described herein, along with the examples and experiences we cite, has been shared with graduate and undergraduate students. The graduate students used the methodology as part of their research, primarily as a tool to analyze data and gain significant insights into the process behaviour from which they were collecting data. The undergraduate students used the methodology to complement what they had learned in their Applied Statistics course (2nd year) and their Design of Experiments course (3rd or 4th year); it was a helpful tool as they analyzed the data collected during group design projects or individual research projects in

their senior year. Although typical course evaluations were not solicited from these students, we have compiled several comments and anecdotal information. These remarks were received from students who participated in these design/research projects over the past couple of years and have since made use of the statistical tools in other settings.

From a student who graduated and is currently gainfully employed:

> "…You won't be surprised to hear that hierarchical designs had immediate application in the workplace! There is a [company name and process description] production facility in [location in USA], which sent us two samples of … powder from the top of the reactor and two from the bottom. I made three [specimens] with each sample. We will soon be running [tests] for determining specific properties. This is a 2×2×3 hierarchical design, as you taught us! If we didn't take the replicates from different locations in the reactor and it was just natural heterogeneity, we could spin our wheels for months developing mechanistic theories based on artifacts of statistical variance...".

Another comment:

> "For me it is difficult to talk about hierarchical experiments without using examples to explain, but it is clear that as a student in this design project group I have grasped the concept to a degree that it is now natural for me to always consider not only the measurement error, but also the steps along the way. This is no trivial thing and actually I have noticed that very few people (even within ChE BASc, MASc/MEng or PhDs!) think like this. To be fair, neither did I before taking on this project, which demonstrates that I have in fact learned a great deal. Perhaps…the

reactor example that we did is the best and most natural place to start for chemical engineering students…".

Another student provided the following comment:

"At [company name] we made [specimens] and ran conductivity tests on the outlet flow with different inlet solution concentrations. The issue we had was evaluating the performance based on these different inlet concentrations. It is true that one [specimen] may have performed better with one solution over another but these units were all handmade (even the inlet solutions were mixed by us) so it was very difficult to say if it was performing better or not due to these lurking variables. A hierarchical design strategy could be used to control for these variables perhaps using different solution recipes at the top level, then pooling solution mixing and [specimen] number for the second level and measurement variance of the conductivity at the bottom. Hierarchical design can basically be used whenever there is a measurement and a true change affected by lurking variables, which is quite often the case...".

And another student had this to say:

"At [company name] we varied powder formulations for creating different plastics. These were batch processes so there were many entry points of error for new runs. In one instance, we used an additive to try to achieve tailored properties which were manifest in the [specific property] testing of the finished product; however, without replicates, we were in the dark about the error. After hearing about this strategy, we should have taken replicates at two levels: making multiple plastic sheets for each run and fusing multiple identical sheets to different plaques. Doing this

591 would not only have incidentally given valuable information about the error in our process but
592 also definitively established the significance of adding said additive...".
593
594 And finally, we received insight from a previous co-op student experience:
595 "…we were extracting DNA, RNA and protein from mouse liver and the amounts of the specific
596 protein were quantified with [various characterization techniques]. So in this case, hierarchical
597 design could have been used by taking replicates from the same piece of mouse tissue, then on
598 the extraction process and the test. This would have helped a lot in determining if the results
599 were significant…".
600
601 The comments here speak for themselves; the project was of value to students and
602 confirmed/solidified key concepts that had been targeted. Not only did students learn about the
603 technical aspects of polymer characterization and statistical hierarchical design and analysis, but
604 they were also able to articulate the importance of the analysis technique and reflect on its potential
605 application in industry. There is an inherent mindset shift that has occurred for these students, and
606 they have become much more aware of error sources in each step of the process that they are
607 evaluating. As one student wrote, "…The beauty of the methodology is that it teaches you a certain
608 way of thinking. This way of thinking, where we can easily assess entry points of error and
609 quantitatively state that, for example, the top of the reactor or resin bin produces better product
610 than the bottom (or mid-point) and then attack why this is the case, is so valuable…".
611
612 While we prefer to emphasize the lessons learned by students, we should also comment on the
613 insights gained by the instructional team; this will ensure that the implementation becomes even

more effective in future course offerings. Overall, we felt that the relationships between statistical design of experiments, polymer synthesis and polymer characterization were well-established, and that the subsequent analyses were at a suitable level of difficulty for students. As with any group project, it is important to ensure equitable distribution of work; this is critical in both the laboratory setting and during the statistical analysis steps. To ensure that all students are motivated to contribute, it may be beneficial to assign a "lab participation" grade and/or assign a "group reflection" piece near the end of the project.

One additional comment is related to the selection of polymerization processes described herein. The first two examples, polyelectrolyte terpolymerization and crosslinking copolymerization, were both fairly complex processes. As such, both materials presented some challenges during the experimental steps. The synthesis of the AMPS/AAm/AAc terpolymer was difficult for some students, especially in terms of pH adjustment prior to synthesis. Occasionally, the exothermic titration increased the temperature of the pre-polymerization solution too substantially, which resulted in some premature polymerization. Also, for both the AMPS/AAm/AAc terpolymer and the STY/DVB copolymer, there were some issues with sample preparation prior to GPC analysis. AMPS/AAm/AAc can take a very long time to dissolve completely in the mobile phase, and undissolved material may be inadvertently filtered out prior to analysis if students are impatient. Similarly, the STY/DVB copolymer crosslinked under some conditions, leading to the formation of insoluble gel. Given the limitations of GPC characterization, the molecular weight averages of these insoluble portions could not be accurately measured.

However, as mentioned earlier, the case studies selected were intentionally complicated, as they mirror real-world situations that students may face. Exploring properties of polyelectrolytes and crosslinked polymers provide important troubleshooting opportunities for future chemical engineers. That said, this paper is intended to provide instructors with the tools needed to develop a similar project in their own courses; each instructor will inevitably choose their own polymerization processes to work with. For instructors with limited polymerization background, a homopolymerization process may be more suitable. Consider, for example, the synthesis of polystyrene (in either solution or emulsion): it is a fairly straightforward process, but one might still vary the recipe and/or the sampling time before characterization via GPC. Thus, it could be an interesting and relatively simple study based on the prescriptions described herein.

CLOSING REMARKS

Using polymer property characterization studies to teach hierarchical design statistics provides students with exposure to several topics that they may not otherwise discover. This type of project can be used to integrate general principles related to polymer science (understanding polymerization processes and molecular weight distributions, for example) with advanced laboratory skills (including sample preparation, instrument operation and data collection), while simultaneously ensuring that students are able to identify relevant sources of error and are able to quantify them using hierarchical data analysis techniques.

From the students' perspective, this type of project provides them with more opportunities to appreciate experimental design principles, complementary to their lab sessions and/or to their

senior design projects. They are also encouraged to spend time in the lab, gaining valuable hands-on experience. While they will immediately see how sources of error persist in polymer synthesis and characterization, they will also be able to carry the statistical methodology with them to other aspects of chemical engineering. Undergraduate course work in chemical engineering programs can often seem far removed from industrial applications. Technical courses can be very theoretical in nature, and it can be difficult for students to appreciate the real-world relevance. This type of project gives students the opportunity to see how technical concepts apply to industrial problem-solving (designing experiments, identifying sources of error, troubleshooting, etc.). Creating links between the classroom and the workplace will ultimately strengthen the skills that students will require in industry.

From an instructor's perspective, this intersection of several relevant topics makes it possible to achieve a wide range of learning outcomes. Students will not only expand their technical knowledge base, but they will also gain experience handling complex, open-ended problems. The project might also include a review of relevant literature, collaboration with classmates, and communication of results. Thus, students will benefit immensely from this type of project.

ACKNOWLEDGEMENTS

The authors would like to thank several graduate and undergraduate students who worked with these techniques over various projects and were kind enough to provide us with useful feedback about their experience.

682 **REFERENCES**

683

1. Dubé M and Penlidis A (1996) Hierarchical data analysis of a replicate experiment in emulsion terpolymerization. *AIChE J.* 42(7): 1985-1994.
2. D'Agnillo L, Soares J and Penlidis A (1999) A hierarchical data analysis of a replicate experiment in polyethylene synthesis with high-temperature gel permeation chromatography. *Polym. React. Eng.* 7(2): 259-281.
3. Box G, Hunter W and Hunter J (1978) *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building.* Wiley. Toronto, Canada.
4. Montgomery D (2012) Ch 14: Nested and Split-Plot Designs in *Design and Analysis of Experiments*. 8th edition. Wiley. Toronto, Canada.
5. Scott A, Kazemi N and Penlidis A (2017) AMPS/AAm/AAc terpolymerization: Experimental verification of the EVM framework for ternary reactivity ratio estimation. *Processes.* 5(1): 9.
6. Scott A, Duever T and Penlidis A (2019) The role of pH, ionic strength and monomer concentration on the terpolymerization of 2-acrylamido-2-methylpropane sulfonic acid, acrylamide and acrylic acid. *Polymer.* 177: 214-230.
7. Scott A and Penlidis A (2020) Designing optimal terpolymers for enhanced oil recovery (polymer flooding). *Ind. Eng. Chem. Res.* 59(16): 7426-7437.
8. Riahinezhad M, McManus N and Penlidis A (2015) Effect of monomer concentration and pH on reaction kinetics and copolymer microstructure of acrylamide/acrylic acid copolymer. *Macromol. React. Eng.* 9(2): 100-113.

9. Scott A, Nabifar A, Hernandez-Ortiz J, McManus N, Vivaldo-Lima E and Penlidis A (2014) Crosslinking nitroxide-mediated radical polymerization of styrene with divinylbenzene. *Eur. Polym. J.* 51: 87-111.

684