Towards a Better Understanding of the Epoxy Polymerization Process Using Multi-Objective Evolutionary Computation

Kalyanmoy Deb[1]*, Kishalay Mitra[2], Rinku Dewri[3] and Saptarshi Majumdar[4]

[1]Professor, Mechanical Engineering Department, Indian Institute of Technology Kanpur, Kanpur 208016, India, deb@iitk.ac.in

[2]Manufacturing Practice, Tata Consultancy Services, 54B Hadapsar Industrial Estate, Pune 411013, India, kmitra@pune.tcs.co.in

[3]Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur 721302, India, rinku@webteam.iitkgp.ernet.in

[4]Tata Research Development and Design Centre, 54B Hadapsar Industrial Estate, Pune 411013, India, smajumdar@pune.tcs.co.in

KanGAL Report Number 2004001

Abstract

The optimization of the epoxy polymerization process involves a number of conflicting objectives and more than twenty decision parameters. In this paper, the problem is treated truly as a multi-objective optimization problem and Pareto-optimal solutions corresponding to two and three objectives are found using a commonly-used evolutionary algorithm (the elitist non-dominated sorting GA or NSGA-II). A combination of minimization or maximization of various objectives, such as the number average molecular weight, polydispersity index and reaction time, are considered here. The first two objectives are related to the properties of a polymer, whereas the third objective is related to productivity of the polymerization process. The decision variables are discrete addition rates of various ingredients e.g. the amount of addition for bisphenol-A (a monomer), sodium hydroxide and epichlorohydrin at different time steps (modeled in a semibatch operation), whereas the satisfaction of all species balance equations is treated as constraints. First, a substantial improvement over the benchmark condition (reported by performing a one-time ingredient addition in the beginning in a batch process) is observed. Secondly, the use of the real-coded NSGA-II is found to perform much better than the binary-coded NSGA-II. Thirdly and importantly, this study brings out a salient aspect of using an evolutionary approach to multi-objective problem solving. The availability of multiple optimal trade-off solutions allows an user to search for meaningful similarities (and dissimilarities) among the obtained solutions. Changes in distribution of various polymer species in the course of polymerization process as observed among various Pareto-optimal solutions are identified and explained. Such information should provide important information about optimal solutions corresponding to different trade-offs among objectives, which are otherwise difficult to obtain. Moreover, a study with three objectives shows the optimal-objective-map showing the trade-off among objectives for high-performance solutions. The systematic approach of starting with two-objective problems to capture the essential features of interesting optimal operating conditions to finally solving the three-objective problem associated with

*Corresponding author, Tel: +91-512-2597205, Fax: +91-512-2597408, email: deb@iitk.ac.in

the epoxy polymerization problem for discovering the optimal trade-off interactions should motivate further such studies on other chemical process optimization problems.

Keywords: Epoxy polymerization, multi-objective optimization, Pareto-optimal solutions, genetic algorithms, domination.

1 Introduction

Real-world optimization problems often demand to cater the need of solving more than one objective simultaneously. Multi-objective optimization problems lead to a set of optimal solutions, known as *Pareto-optimal* solutions, as opposed to the single solution provided by any single-objective optimization task. Although only one solution must be chosen at the end of the optimization task and this often must be performed with the guidance of a decision-maker, it is a better practice to first find a set of Pareto-optimal solutions to have an idea of the extent of trade-offs possible among the underlying objectives before focusing on a particular solution (Deb, 2003). There are mainly two approaches to find a set of Pareto-optimal solutions: (i) single-objective preference based methods, in which multiple objectives are scalarized into a single objective by means of certain preference information (Chankong and Haimes, 1983; Mietinnen, 1999) and multiple applications of the approach are conducted to find a set of Pareto-optimal solutions and (ii) evolutionary multi-objective optimization (EMO) methods, in which multiple Pareto-optimal solutions are found simultaneously in a single simulation run (Deb, 2001). It is needless to write that the second approach may be preferred simply due to the time and effort saved in simultaneous computation of Pareto-optimal solutions.

Although the field of research and application on multi-objective optimization is not new, the use of EMO techniques in various engineering and business applications is a recent phenomenon. Polymerization processes, being quite complex in nature, offer themselves as an extremely challenging candidate for multi-objective optimization studies. In modeling the polymerization system, several molecular parameters, such as the number or weight average molecular weights (M_n or M_w, respectively), the polydispersity index (PDI), concentration of different functional groups etc., can all be predicted quite accurately using various experimentally measured indices such as strength and stiffness of the final product. Moreover, the desired objectives in a polymerization process often exhibit conflicting relationships and therefore become an ideal problem for multi-objective optimization studies. In this paper, multi-objective optimization of a semibatch epoxy polymerization system which is often used to manufacture high-strength composites, reinforced plastics, adhesives, protective coatings in appliances, etc. is thoroughly investigated.

A recent review by Bhaskar et al. (2000) reveals that several studies are carried out on multi-objective optimization of polymerization reactors. Tsoukas et al. (1982), Fan et al. (1984), Farber (1986, 1989), Butala et al. (1988), and Choi and Butala (1991) carried out multi-objective optimization of copolymerisation reactors. These studies used Pontryagin minimum principle – a pattern search technique – coupled with the ϵ-constraint method (Chankong and Haimes, 1983) to obtain a Pareto-optimal set of non-dominated solutions. Wajge and Gupta (1994) studied multi-objective optimization of the nylon-6 batch reactor and obtained different optimal temperature histories corresponding to different solutions on the Pareto-optimal set using the same technique. Sareen and Gupta (1995) extended that work and studied the nylon-6 semibatch reactor and obtained different optimal pressure histories and optimal jacket fluid temperature corresponding to different solutions on the Pareto-optimal set. The solution technique used in these studies requires as many simulation runs as the desired number of

distinct Pareto-optimal solutions, since each simulation run leads to a single Pareto-optimal solution. To alleviate this difficulty, a number of studies using an EMO approach are carried out on the multi-objective optimization of nylon-6 and polymethyl methacrylate (PMMA) reactors (Chakravarthy et al., 1997; Mitra et al., 1998; Garg and Gupta, 1999; Garg et al., 1999; Gupta and Gupta, 1999). These studies are mainly based on adapted version of the non-dominated sorting genetic algorithm (NSGA) developed by Srinivas and Deb (1995). In all these studies, constraints were handled using the standard penalty function method. It is well-known in the optimization literature that the tuning of penalty parameters is a crucial issue in penalty-based constraint handling methods. The NSGA procedure is modified by its developers to remedy this problem and to introduce a few other important features in the development of NSGA-II (Deb et al., 2002). The use of a novel constraint-handling approach requiring no penalty parameter is a major improvement, which has made the NSGA-II quite efficient in solving constrained multi-objective optimization problems. We shall give a brief overview of the NSGA-II procedure in Section 3.

In the Taffy process (Kumar and Gupta, 1987), the most popular industrial process for preparing epoxy polymers, bisphenol-A (monomer) and epichlorohydrin, in excess, are reacted in presence of sodium hydroxide (NaOH) leading to polymer formation that is having glycidyl ether end group (building block) at both the ends. In the Advancement process (McAdams and Gannon, 1986), the other route, pulverized sodium hydroxide is added in steps to the reaction mixture of bisphenol-A and epichlorohydrin dissolved in a solvent. It is well established that in all commonly used industrial processes of epoxy polymerization, alkali has a key role, which is usually added in semi-batch mode. Alkali controls oligomeric impurities in the advancement process (Ray, 1972; Kumar and Gupta, 1987). But the role of addition of other ingredients (bisphenol-A and epichlorohydrin) is not well established. Experimental and theoretical studies are very few in open literatures for the epoxy polymerization process. Studies on optimization of epoxy reactors are very rare according to the knowledge of the authors. Raha and Gupta (1998) used species balance and equation of moments approach to study the process. They gave special importance to build the entire modeling framework as well as the effect of kinetic parameters and reactant's effect over the performance of the reaction process. This work was extended elsewhere (Majumdar et al., 2003; Raha et al., 2003; Mitra et al., 2003) with single and multi-objective optimization studies showing the effect of different addition amounts and patterns over the performance of the process. These past studies motivated us to launch this present detail study for a better understanding of the true nature of interactions among three conflicting objectives associated with the epoxy polymerization process.

In the reminder of the paper, we formulate the epoxy polymerization process as a couple of two-objective optimization problem leading to a three-objective optimization problem. The formulation of associated constraints and the choice of decision variables are explained. Thereafter, a brief introduction to the NSGA-II procedure is given highlighting the constraint handling strategy. The simulation results of NSGA-II are then described and supported by comparing them with a number of single-objective preference-based optimized solutions. The interesting task of revealing important insights about the polymerization process is illustrated next by analyzing the obtained Pareto-optimal solutions. Finally, all the information gathered by the two and three-objective optimization tasks are combined together to understand the true trade-off among the three conflicting objectives considered in this study, and important and useful conclusions are derived.

2 Formulation

In this section, we formulate the epoxy polymerization problem as multi-objective optimization problems of various degrees of difficulty.

2.1 Model

The complete kinetic scheme for the above mentioned polymerization system is given in Appendix A and can also be found elsewhere (Batzer and Zahir, 1977). Raha and Gupta (1998) and Majumdar et al. (2003) have validated the model with the available experimental data. Using the species balance approach, ODEs corresponding to the initial value problem (IVP) are derived for various species and their moments. These equations are solved by an explicit numerical integration routine and kinetic parameters appearing in the equations are estimated with the help of a genetic algorithm (GA). Details on the modeling aspect, the solution procedure and the parameter estimation can be found in Majumdar et al. (2003). Here, we briefly describe the mathematical model.

The "state" of the reaction scheme can be well described by a set of 48 state variables ($\mathbf{x} = (x_1, x_2, \ldots, x_{48})^T$), including all species balance and moment balance equations. Equations for these can be written using mass balance equations and by obtaining balance of moments of various orders. The state variable equations, in general, can be written in the form:

$$\frac{dx_i}{dt} = f_i(\mathbf{x}, \mathbf{U}), \quad i = 1, 2, \ldots, 48, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{U}$ are vectors of the state and manipulated variables (such as the amount of intermediate additions for different ingredients at different times). Manipulated variable vector consists of three discrete histories, namely, discrete history for NaOH addition (described here with $U_1(t)$), discrete history for epichlorohydrin addition ($U_2(t)$) and discrete history for bisphenol-A addition ($U_3(t)$).

The various molecular species including the monomer considered for the modeling exercise are also presented in Appendix B. Mole balance equations for the end groups and lower oligomers in the batch reactor, the moment equations for various molecular species and the equations for number as well as average molecular weights are derived for this system. Appendix C shows the species balance of zeroth, first and second moment equations for a typical species AB_n. Similar equations exist for all other species mentioned in Appendix B. Appendix C also gives equations for the end group concentration and the monomeric species. Given three discrete profiles ($\mathbf{U}$ at time zero and at other time steps) and the initial conditions of all state variables ($\mathbf{x}$ at time zero), the reaction scheme model can be solved by using an explicit integrator (RK-type method (Walas, 1991)). This simulation procedure can then be combined with the NSGA-II optimization code for performing a multi-objective optimization.

2.2 Defining the Optimization Problems

Three different multi-objective optimization problems are studied here. The first problem (problem 1) is related to the quality of polymer produced, whereas the second problem (problem 2) addresses the productivity issue also. The objective for problem 1, $\mathbf{I}_1$, is a vector and consists of two objectives: The first (I_{11}) is the maximization of M_n and the second (I_{12}) is minimization of PDI. Thus, the simultaneous attainment of the two objectives has to be obtained with satisfying the mass and moment balance equations. A scheme of the reactions is shown in Appendix A and a few such reaction equations are given in Appendix C. All 48 equations can

be found elsewhere (Majumdar et al., 2003; Raha and Gupta, 1998). Simultaneous attainment of these two objectives is going to produce high quality polymers with different trade-offs between M_n and PDI values:

$$\textbf{Problem1}: \begin{cases} \text{Maximize} & I_{11} = M_n, \\ \text{Minimize} & I_{12} = \text{PDI}, \\ \text{subject to} & \text{satisfying mass and moment balance equations}, \\ & u_i^{\min} \leq u_i \leq u_i^{\max} \quad i = 1, \ldots, 21. \end{cases} \tag{2}$$

One computer simulation runs from zero (initial condition, $t = 0$) to $t = t_{\text{sim}}$ (7 hr used here). Each of the three profiles ($U_1(t)$, $U_2(t)$, $U_3(t)$) are, therefore, discretized into seven equally spaced points ($N_{\text{Digit}} = 21$ discrete values, u_1, u_2, ..., u_{21}). This means that the decision variables u_1 to u_7 represent $U_1(t)$ (addition rates of NaOH), u_8 to u_{14} represent $U_2(t)$ (addition rates of EP) and u_{15} to u_{21} represent $U_3(t)$ (addition rates of AA_0) at seven different times ($t = 0, 1, 2, 3, 4, 5, 6$ hr). Each of these variables is bound to lie between a lower bound ($u_i^{\min}$) and a upper ($u_i^{\max}$) bound. No restriction on the M_n and PDI values are used as constraints, as the satisfaction of variable bounds will ensure limiting values on M_n and PDI.

The objective for problem 2, $\mathbf{I}_2$, is a vector of two objective functions: The first (I_{21}) is the maximization of M_n and the second (I_{22}) is minimization of the overall reaction time, t_{sim}. Simultaneous attainment of these two objectives has to be obtained with satisfying the mass and moment balance equations. As the resultant product is supposed to obey the chemical principles due to the satisfaction of constraints, the attainment of the second objective is going to ensure the optimal productivity (by reducing the reaction time). In short, this formulation improves productivity:

$$\textbf{Problem2}: \begin{cases} \text{Maximize} & I_{21} = M_n, \\ \text{Minimize} & I_{22} = t_{\text{sim}}, \\ \text{subject to} & \text{satisfying mass and moment balance equations}, \\ & u_i^{\min} \leq u_i \leq u_i^{\max} \quad i = 1, \ldots, 21, \\ & t_{\text{sim}}^{\min} \leq t_{\text{sim}} \leq t_{\text{sim}}^{\max}. \end{cases} \tag{3}$$

The reaction time t_{sim} is also a decision variable in this problem, thereby making the total number of variables to $N_{\text{Digit}} = 22$. It is understood here that depending on the value of the reaction time variable, t_{sim}, the simulation time of the reaction scheme is going to be determined. The reaction time variable is also restricted to be varied within a lower and a upper bound.

The objective for problem 3, $\mathbf{I}_3$, is a vector of three objective functions: The first (I_{31}) is the maximization of M_n, the second (I_{32}) is minimization of t_{sim}, and the third (I_{33}) is the minimization of PDI. Simultaneous attainment of these three objectives has to be obtained with satisfying the mass and moment balance equations. On the one hand, the product quality is going to be assured by the first and third objectives and the constraints assigned on them, on the other hand, the attainment of second objective ensures productivity. In this problem, the reaction time t_{sim} is also a decision variable. This problem also has $N_{\text{Digit}} = 22$ decision variables:

$$\textbf{Problem3}: \begin{cases} \text{Maximize} & I_{31} = M_n, \\ \text{Minimize} & I_{32} = t_{\text{sim}}, \\ \text{Minimize} & I_{33} = \text{PDI}, \\ \text{subject to} & \text{satisfying mass and moment balance equations}, \\ & u_i^{\min} \leq u_i \leq u_i^{\max} \quad i = 1, \ldots, 21, \\ & t_{\text{sim}}^{\min} \leq t_{\text{sim}} \leq t_{\text{sim}}^{\max}. \end{cases} \tag{4}$$

3 Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II)

The non-dominated sorting GA or NSGA-II procedure (Deb et al., 2002) for finding multiple Pareto-optimal solutions in a multi-objective optimization problem is described here.

Like in a genetic algorithm (GA), NSGA-II starts with a population of N_{pop} random solutions. In the N-th iteration of NSGA-II, the offspring population $\mathcal{Q}_N$ is first created by using the parent population $\mathcal{P}_N$ and the usual genetic operators – reproduction, recombination, and mutation (Goldberg, 1989). Thereafter, both populations are combined together to form $\mathcal{R}_N$ of size $2N_{\text{pop}}$. Then, a non-dominated sorting procedure (Deb, 2001) is applied to classify the entire population $\mathcal{R}_N$ into a number of hierarchical non-dominated fronts. The non-dominated sorting procedure is described in the following paragraph.

The definition of *domination* between two solutions is given as follows (Deb, 2001; Miettinen, 1999):

Definition 1 *A solution* $\mathbf{x}^{(1)}$ *is said to dominate the other solution* $\mathbf{x}^{(2)}$*, if both the following conditions are true:*

1. *The solution* $\mathbf{x}^{(1)}$ *is no worse than* $\mathbf{x}^{(2)}$ *in all objectives.*
2. *The solution* $\mathbf{x}^{(1)}$ *is strictly better than* $\mathbf{x}^{(2)}$ *in at least one objective.*

For a given set of solutions (for example, those shown in Figure 1(a)), a pair-wise comparison can be made using the above definition and whether one solution dominates the other can be established. All solutions which are not dominated by any other members of the set are called the non-dominated solutions of level one. These solutions (3, 5, and 6 in the figure) are also said to lie on the best non-dominated front. By temporarily discounting these solutions, the above principle can be followed again and the next non-dominated front (solutions 1 and 4 in the figure) can be identified. Figure 1(b) shows how the given set of six solutions are classified into three non-dominated fronts. Notice how both minimization and maximization of objectives

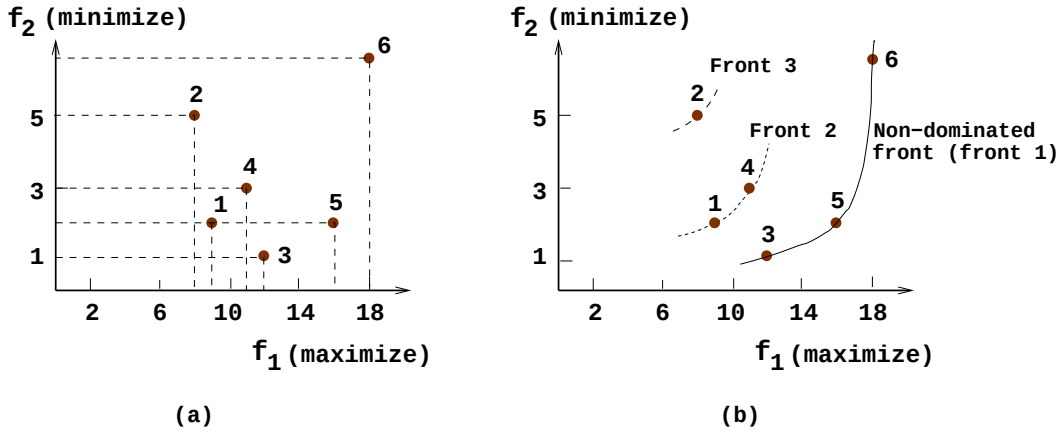

Figure 1: A set of solutions and the best non-dominated front.

can be handled by using the domination definition 1 and without using any additional fix-ups, such as conversion of one type of objective to another.

Once the non-dominated sorting of the set $\mathcal{R}_N$ is over, the new population is filled with solutions of different non-dominated fronts, one at a time. The filling starts with the best non-dominated front and continues with solutions of the second non-dominated front, followed by the third non-dominated front, and so on. Since the overall population size of $\mathcal{R}_N$ is $2N_{\text{pop}}$, not all fronts may be accommodated in N slots available in the new population. All fronts which could not be accommodated are simply deleted. When the last allowed front is being considered, there may exist more solutions in the last front than the remaining slots in the new population. This scenario is illustrated in Figure 2. Instead of arbitrarily discarding some members from the last front, the solutions which will make the *diversity* of the selected solutions the highest are chosen. In this step, the crowding-sorting of the solutions of front

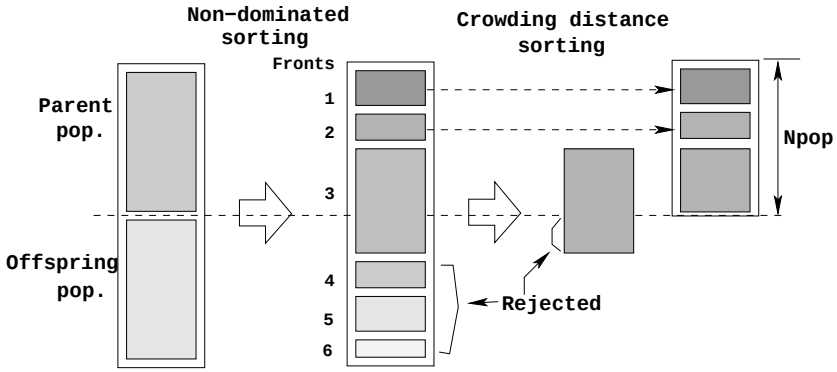

Figure 2: Schematic of the NSGA-II procedure.

i (the last front which could not be accommodated fully) is performed by using a *crowding distance metric*, which we shall described next. The population is arranged in descending order of magnitude of the crowding distance values.

The crowded comparison operator ($<_c$) compares two solutions and returns the winner of the tournament. It assumes that every solution i has two attributes:

1. A non-domination rank rnk_i in the population.
2. A local crowding distance (d_i) in the population.

The crowding distance d_i of a solution i is a measure of the search space around i which is not occupied by any other solution in the population. Based on these two attributes, the crowded tournament selection operator is described as follows.

Definition 2 *Crowded Tournament Selection Operator: A solution i wins a tournament with another solution j if any of the following conditions are true:*

1. *If solution i has a better rank, that is, $rnk_i < rnk_j$.*
2. *If they have the same rank but solution i has a better crowding distance than solution j, that is, $rnk_i = rnk_j$ and $d_i > d_j$.*

The first condition makes sure that chosen solution lies on a better non-dominated front. The second condition resolves the tie of both solutions being on the same non-dominated front by deciding on their crowded distance. The one residing in a less crowded area (with a larger

crowding distance d_i) wins. The crowding distance d_i can be computed in various ways. However, NSGA-II uses a crowding distance metric (we describe in the next paragraph) which requires only $O(MN_{\text{pop}} \log N_{\text{pop}})$ computations, where M is the number of objectives and N_{pop} is the number of solutions in the GA population.

To get an estimate of the density of solutions surrounding a particular solution i in the population, the average distance of two solutions on either side of solution i along each of the objectives is computed. This quantity d_i serves as an estimate of the perimeter of the cuboid formed by using the nearest neighbors as the vertices (this is called the *crowding distance*). In Figure 3, the crowding distance of the i-th solution in its front (marked with filled circles) is the average side-length of the cuboid (shown by a dashed box). The side-length in each

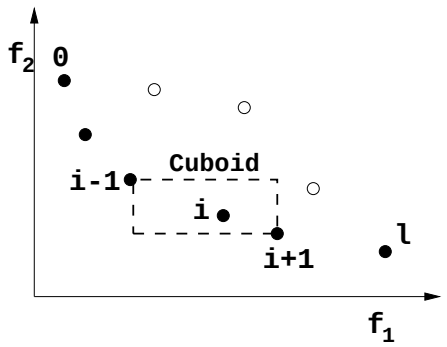

Figure 3: The crowding distance calculation.

objective is calculated using the following normalized difference in objectives values:

$$d_{\mathcal{I}_j^m} = \frac{f_m^{(\mathcal{I}_{j+1}^m)} - f_m^{(\mathcal{I}_{j-1}^m)}}{f_m^{\max} - f_m^{\min}}. \tag{5}$$

The index $\mathcal{I}_j^m$ denotes the solution index of the j-th member in the sorted list using the m-th objective. Thus, for the m-th objective, $\mathcal{I}_1^m$ and $\mathcal{I}_l^m$ denote the lowest and highest objective function values, respectively. For the boundary solutions, an infinite distance value is used, meaning that they are most important solutions to be selected for the next iteration. Thus, this metric denotes half of the perimeter of the enclosing cuboid with the nearest neighboring solutions placed on the vertices of the cuboid (Figure 3). The parameters $f_m^{\max}$ and $f_m^{\min}$ can be set as the population-maximum and population-minimum values of the m-th objective function.

3.1 Constraint-Handling in NSGA-II

The constraint handling method modifies the crowded tournament selection operator, where two solutions are picked from the population and the better solution is chosen. In the presence of constraints, each solution can be either feasible or infeasible. Thus, there may exist at most three situations: (i) both solutions are feasible, (ii) one is feasible and other is not, and (iii) both are infeasible. We consider each case by simply redefining the domination principle as follows for any two solutions $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$) (Deb, 2001):

Definition 3 *A solution $\mathbf{x}^{(i)}$ is said to 'constrain-dominate' a solution $\mathbf{x}^{(j)}$ (or $\mathbf{x}^{(i)} \preceq_c \mathbf{x}^{(j)}$), if any of the following conditions are true:*

1. *Solution* $\mathbf{x}^{(i)}$ *is feasible and solution* $\mathbf{x}^{(j)}$ *is not.*

2. *Solutions* $\mathbf{x}^{(i)}$ *and* $\mathbf{x}^{(j)}$ *are both infeasible, but solution* $\mathbf{x}^{(i)}$ *has a smaller constraint violation.*

3. *Solutions* $\mathbf{x}^{(i)}$ *and* $\mathbf{x}^{(j)}$ *are feasible and solution* $\mathbf{x}^{(i)}$ *dominates solution* $\mathbf{x}^{(j)}$ *in the usual sense (Definition 1).*

The above change in the definition allows a minimal change in the NSGA-II procedure described earlier. Figure 4 shows the non-dominated fronts on a six-membered population. In the absence of the constraints, the non-dominated fronts (shown by dashed lines) would have been `((1,3,5), (2,6), (4))`, but in their presence, the new fronts are `((4,5), (6), (2), (1), (3))`. The first non-dominated front is constituted with the best feasible solutions in the

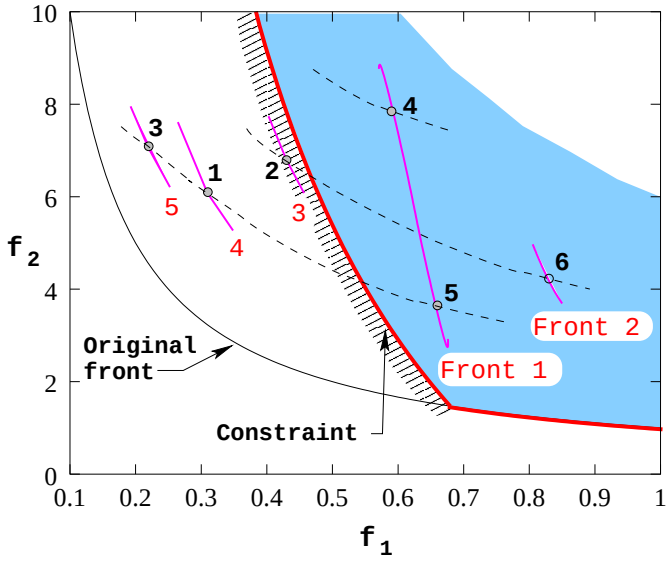

Figure 4: Non-constrain-dominated fronts for minimization of f_1 and f_2.

population and any feasible solution lies on a better non-dominated front than an feasible solution.

The highlight of this procedure is that it makes a simple adjustment on the definition of the domination and does not require any additional parameter, such as the penalty parameter.

3.2 NSGA-II for the Epoxy Polymerization Problem

Each solution is represented as a real-valued vector of 21 (or 22 for problems 2 and 3) values indicating the addition of NaOH, EP, and AA_0. Thus, a solution with known values of these quantities can be evaluated by following the simulation procedure described earlier. The evaluation will produce two (or three for problem 3) objective values and will indicate the violation of constraints, if any. NSGA-II allows both minimization and maximization of different objectives by using the domination definition 1. Thus, the objectives are used straightway (without any transformation) as they are described in equations 2 to 4. The constraint-handling procedure described above is used to produce the non-dominated ranking, which implicitly de-emphasizes the infeasible solutions. For the real-coded NSGA-II, we use the simulated binary crossover

(SBX) and the polynomial mutation operators (Deb and Agrawal, 1995). For the binary-coded NSGA-II (mainly used here for comparison purposes), we use a single-point crossover and a bitwise mutation operator (Goldberg, 1989). It is interesting to note that since solutions are always created within the specified lower and the upper bounds, the variable bounds will never be violated by both NSGA-IIs.

When a pre-specified maximum iteration count ($N = N_{\max}$) is reached, NSGA-II is terminated and the non-dominated solutions of the final population are declared as the obtained Pareto-optimal solutions. In problems 1 and 2, $N_{\max} = 200$ and a population size of $N_{\text{pop}} = 250$ are used. Since problem 3 deals with a three-dimensional Pareto-optimal front, we have chosen a larger population size and run NSGA-II for more iterations: $N_{\text{pop}} = 1,000$ and $N_{\max} = 500$. The crossover and mutation probabilities of $p_c = 0.9$ and $p_m = 0.1$ are used for the real-coded NSGA-II and of $p_c = 0.9$ and $p_m = 0.001$ for the binary-coded NSGA-II. For the SBX operator mentioned above the distribution index of 0.01 and for the polynomial mutation operator the distribution index of 0.01 are used.

4 Results and Discussions

As mentioned earlier, the availability of experimental data for the epoxy polymerization processes is really scarce. As far as authors' information is concerned, very little or no experimental or plant data is available for polydispersity index (PDI) and number averaged molecular weight (M_n) for the epoxy polymerization system. This makes the entire modeling analysis extremely difficult, as validation of the modeling exercise cannot be done in terms of M_n and PDI. Raha and Gupta (1998) addressed this issue in an indirect way. They validated their moment-based model with the data of Batzer and Zahir (1977) for the monomer concentration profile. This is quite reasonable as after some hours of polymerization reactions, if the monomer concentration profile can be matched closely with the experimental data, that itself proves the validity of the proposed mechanism as well as the entire modeling exercise. In the present work the same approach is followed. Kinetic parameters are estimated from a parameter optimization exercise using a simple GA which is used to minimize the error between the simulated values and experimental data on certain monomer concentrations. Figure 5 shows the match between the simulated results and experimental data of a primary oligomer (FF_0). These parameters are used for the models shown in the appendices to simulate various scenarios. Batzer and Zahir (1977) used 0.4 kmol/m^3 for NaOH, 1.280 kmol/m^3 for epichlorohydrin (EP), and 0.4 kmol/m^3 for bisphenol-A (AA_0) as initial concentrations for the three ingredients. Based on this, following variable bounds are set here to allow the optimizer (NSGA-II) a substantial search space to look for the optimized solutions:

1. The addition rate of NaOH varies between 0.2 and 1.0 kmol/m^3 at $t = 0$ hr and between zero and 1.0 kmol/m^3 for $t > 0$ hr.

2. The addition rate of EP varies between 0.2 and 2.0 kmol/m^3 at $t = 0$ hr and between zero and 2.0 kmol/m^3 for $t > 0$ hr.

3. The addition rate of AA_0 varies between 0.2 and 1.0 kmol/m^3 for $t = 0$ hr and between zero and 1.0 kmol/m^3 for $t > 0$ hr.

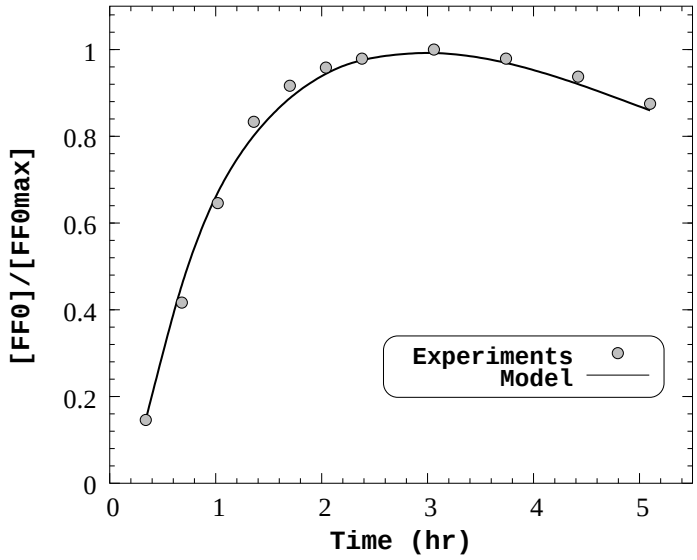

Figure 5: Comparison of experimental data and the mathematical model in which kinetic parameters are obtained using a genetic algorithm.

4.1 Discussion on Problem 1

A simulation result corresponding to the initial condition of Batzer and Zahir (1977) (with M_n=633.2 kg/kg-mole and PDI=1.61) is considered as the benchmark performance data with which our optimized solutions will be compared later. After the multi-objective optimization with NSGA-II, Pareto-optimal solutions are generated for the problem 1 and are marked as "NSGA-II" solutions in Figure 6. The solutions termed as "Initial" denote the objective vectors with which the NSGA-II search process is started (created by using random values of 21 decision variables within their respective bounds). The upper and lower bounds of the decision variables are used as stated earlier. The figure indicates that the random solutions (within the chosen variable bounds) are far from being close to the true Pareto-optimal front, particularly producing solutions with large M_n values. Although a maximum of 200 iterations are scheduled, the real-coded NSGA-II took only 34 iterations to get the distribution shown in Figure 6. The figure clearly shows that a wide range of distribution in both M_n and in PDI values are obtained. The trade-off obtained between the two objectives is also clear from the figure. For the chosen variable bounds, this is the maximum trade-off possible between M_n and PDI values. The following observations can be made from the obtained results:

1. The resulting M_n-PDI trade-off is *non-convex*, a phenomenon which is rare in real-world multi-objective optimization problems. Recall that here M_n is maximized and PDI is minimized.

2. The NSGA-II progresses quite well from the initial population (which consists of inadequate solutions having very large PDI values and small M_n values) to reach the final front.

3. Investigating the solutions of Figure 6 further, it can be inferred that compared to the benchmark solution (marked as 'Benchmark' in the figure) there exist better solutions

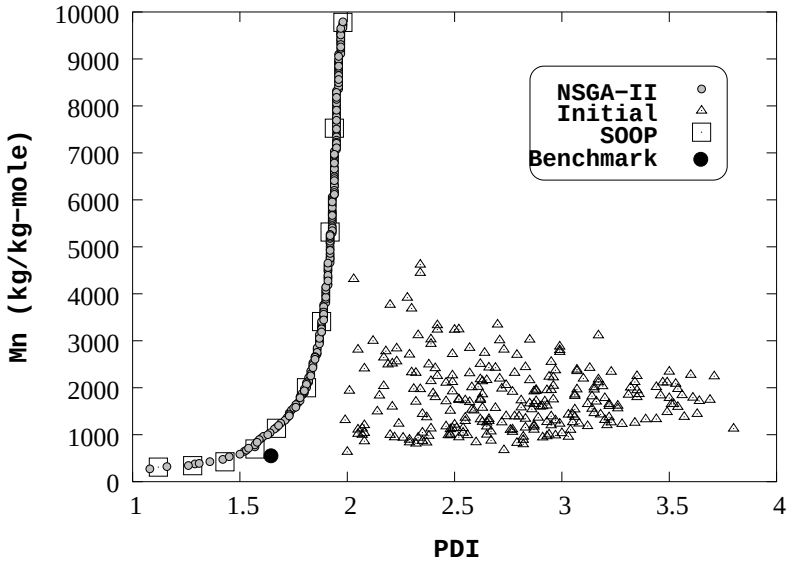

Figure 6: Obtained NSGA-II solutions for the M_n and PDI optimization are compared with single-objective optimization (SOOP) solutions. Initial solutions of NSGA-II and the benchmark solution are also shown.

in the system which can produce larger M_n (ranging from 645 to 966 kg/kg-mole) and less PDI (ranging from 1.52 to 1.61, respectively) value, leading to polymers with better properties as compared to benchmark solution. In fact, the figure also clearly indicates that the benchmark solution is *dominated* by some NSGA-II solutions. If more than 1.61 PDI value is accepted in an application, M_n values much larger than the benchmark value can be achieved and once the PDI value increases beyond 1.8 or so, the corresponding M_n value is observed to increase at a fast rate.

4.1.1 Comparing with Single-Objective Optimization Results

In order to verify whether the obtained NSGA-II solutions can actually be close to the true Pareto-optimal front of this problem, we use a single-objective preference based method next. Since, the above observation indicates that the Pareto-optimal front is non-convex, the commonly-used weighted-sum approach (Chankong and Haimes, 1983, Mietinnen, 1999) may not be the right approach for this problem, as the weighted-sum approach is known to fail in finding the Pareto-optimal solutions in the non-convex region (Deb, 2001). Thus, we use the the ϵ-constraint method here (Mietinnen, 1999). In this case, we convert the second objective (Minimization of PDI) into an additional constraint as PDI $\leq$ PDI$_\epsilon$ and maximize only the first objective. Other constraints and variable bounds, as given in problem 1, are kept the same. To obtain different Pareto-optimal solutions, we simply choose a different value for PDI$_\epsilon$ and optimize the resulting single-objective optimization problem using a single-objective GA. The constraints are handled using a penalty parameter less procedure (Deb, 2000). The GA parameters, such as the population size, operator parameters, etc., are kept the same as those used in the above NSGA-II study. This approach must be applied as many times as the desired distinct Pareto-optimal solutions, since each single-objective optimization solution leads to only one solution on the Pareto-optimal front. Figure 6 marks these solutions as

'SOOP' solutions obtained by 10 independent GA runs, each performed with a different PDI_ϵ value. Since these solutions are found to lie on or near the Pareto-optimal front obtained by the NSGA-II, it can be stated that the Pareto-optimal front found by NSGA-II is probably the true Pareto-optimal front.

4.1.2 Comparing with Binary-Coded NSGA-II

Next, all 21 decision variables are coded in binary strings of seven bits each and the same NSGA-II is used, except that the real-coded crossover and mutation operators are replaced by single-point and bit-wise mutation operators. The comparative performance between real-coded and binary-coded NSGA-II is presented in Figure 7, where it is observed that the real-coded NSGA-II performs slightly better than the binary-coded NSGA-II for higher values of M_n. Thus, the agreement of the results obtained from three procedures give us enough confi-

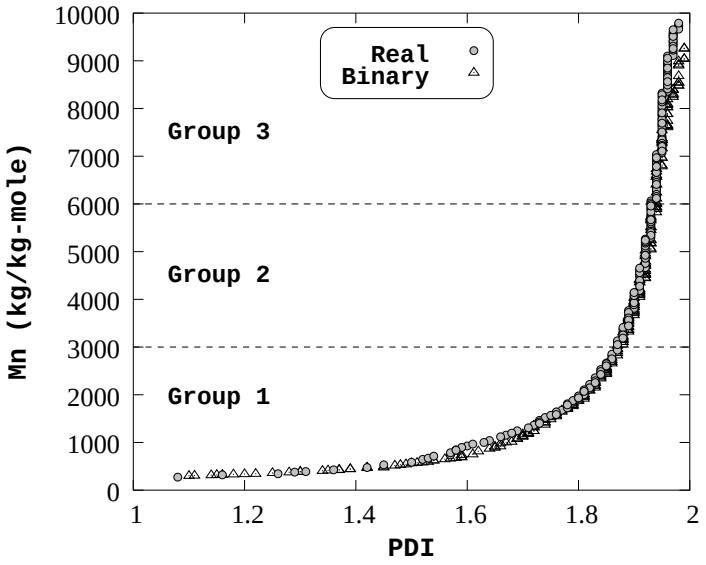

Figure 7: Real-coded and binary-coded NSGA-II solutions for the M_n and PDI optimization.

dence on the optimality of the obtained solutions and more importantly on the performance of the real-coded NSGA-II on the two-objective epoxy polymerization problem. We now discuss an interesting and important aspect of the obtained solutions advocated in this paper.

4.1.3 Searching for Salient Properties of Pareto-Optimal Solutions

Each of the solutions on Pareto-optimal front carries information about 21 decision variables. The Fritz-John necessary condition for Pareto-optimality conditions are well established (Deb, 2001; Mietinnen, 1999) and is presented below:

Theorem 1 *(Fritz–John necessary condition). A necessary condition for* $\mathbf{x}^*$ *to be Pareto-optimal is that there exist vectors* $\boldsymbol{\gamma} \geq \mathbf{0}$ *and* $\mathbf{v} \geq \mathbf{0}$ *(where* $\boldsymbol{\gamma} \in \mathbb{R}^M$, $\mathbf{v} \in \mathbb{R}^J$ *and* $\boldsymbol{\gamma}, \mathbf{v} \neq \mathbf{0}$*) such that the following conditions are true:*

1. $\sum_{m=1}^{M} \gamma_m \nabla f_m(\mathbf{x}^*) - \sum_{j=1}^{J} v_j \nabla g_j(\mathbf{x}^*) = \mathbf{0}$, *and*

2. $v_j g_j(\mathbf{x}^*) = 0$ *for all* $j = 1, 2, \ldots, J$,

where the underlying multi-objective optimization problem (with M objectives f_m to be minimized) is assumed to have J inequality constraints: $g_j(\mathbf{x}) \geq 0$ for $j = 1, \ldots, J$. The above conditions, although cannot be applied to our problem directly due to the inability to compute the gradients of the objective functions and constraints, suggest certain common properties which all Pareto-solutions must satisfy. This leads us to believe that the obtained solutions, if close to the Pareto-optimal solutions, will share some common properties among them and, of course, will have some differences in order to have trade-offs among them. If such properties exist, they would be worth searching for in a real-world application problem, as the sheer knowledge of them will provide important and useful information about the optimal trade-off among objectives (Deb, 2003).

An investigation is performed next to identify whether the obtained M_n-PDI solutions bear some kind of similarity in terms of the associated decision variables. Interesting trends are discovered and shown in Figure 8. Though the decision variables are discrete addition rates

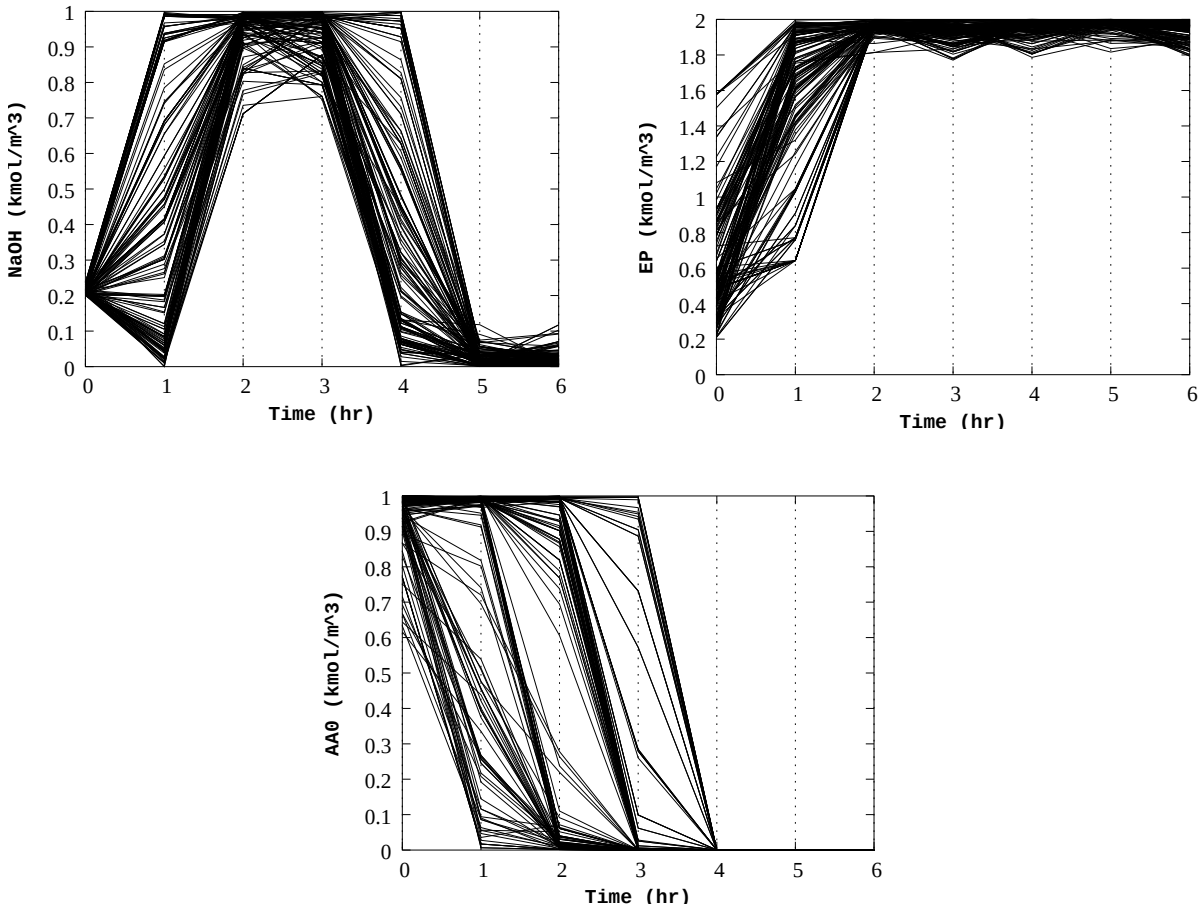

Figure 8: Time-variant addition rates of NaOH (top-left), EP (top-right), and AA_0 (bottom) show common patterns for real-coded NSGA-II solutions on problem 1.

at various time steps, they are joined with straight lines to show the trends. A casual look at the plots will reveal some interesting patterns followed in all obtained solutions. Although

the additions at every hour could have been anywhere on the vertical axis at each time step, all obtained solutions seem to follow some patterns. These patterns reveal important insights about the optimal working characteristics of the epoxy polymerization problem, some of which we have deciphered and are discussed in the following:

1. The general trend captured in NaOH addition rate is to start from the lower bound, reduce or increase the addition rate at the first hour, increase close to the upper bound in the next hour, then continue with the same rate for some more time and finally reduce close to the lower bound. This phenomenon can be explained as follows. A high amount of NaOH addition is required in the first phase of reaction process for a better initiation of polymerization (first three steps are responsible for chain initiation, as presented in Appendix A) and the amount of NaOH can be kept low in the later part of the process as mainly the growth of chain length (in Appendix A, last two steps are responsible for chain growth) occurs at that part of time and NaOH is produced as a by-product. The reason for NaOH to come down at the lower bound after the addition at the zero-th hour for some solutions is due to the fact that the initiation steps require NaOH as a reactant and also produce as a by-product. In case of M_n value less than 7.0×10^3 kg/kg-mole, NaOH is found to be added to the system in a controlled fashion. For M_n value greater than 7.0×10^3 kg/kg-mole, NaOH, instead of decreasing at time step of the first hour after the first addition, goes up and rest of the trend remains same as it is stated earlier. In these cases, NaOH takes part not only in the initiation but also in growth of chain length by helping to form certain species of polymer (BE_n) which contributes significantly towards the high value of M_n.

2. The trend found for epichlorohydrin is straightforward. It must be started at different levels depending on the required M_n value, but must be quickly increased close to the upper limit and must be continued at that rate till the end. The reason for small required value of epichlorohydrin initially is due to it having a less contribution in polymer chain initiation. However, epichlorohydrin should be supplied maximally in the later stages due to its major contribution in chain growth mechanisms. Also, more epichlorohydrin results in more BF_n, EF_n, and FF_n kind of polymer species in the ultimate polymer product. It is experienced by the authors that if the amount of epichlorohydrin is less (especially in the later part of the reaction process), one can expect more growth in AB_n and BB_n type of polymer species. So, based on the requirement of the user, controlled species distribution can be achieved by different kind of intermediate additions.

3. Bisphenol-A must be started in large amount and then must be reduced with time. This is because bisphenol-A takes part actively in polymerization initiation step and its requirement is reduced at the later part (in Appendix A, last two steps are responsible for chain growth). The addition rate is observed to prolong for longer time steps for higher values of M_n. This happens because a large M_n value is achieved by adding more amount of bisphenol-A in the system. It has been seen that the added bisphenol-A is consumed almost completely (i.e. not added more than the amount required) in most of the cases.

It is clear from these observations that different polymers (with different M_n and PDI values) must be produced optimally with a different addition pattern of ingredients. Although this is intuitive, Figure 8 shows a particular pattern of achieving them *optimally*.

To better understand the characteristics of solutions at different parts of the Pareto-optimal front, we divide the entire Pareto-optimal region into three groups and then discuss the proper-

ties of each group by choosing one representative solution from each group. This may be used to predict the addition patterns to be maintained when a user picks up a solution from Pareto-optimal front and wants to manufacture the polymer of that specification. As the solutions on Pareto-optimal front spans M_n from 0.4×10^3 kg/kg-mole to almost 9.5×10^3 kg/kg-mole, each region roughly spans 3.0×10^3 kg/kg-mole in M_n axis: $0.0 - 3.0 \times 10^3$ kg/kg-mole M_n (Group 1), $3.0 - 6.0 \times 10^3$ kg/kg-mole M_n (Group 2), and $6.0 - 9.0 \times 10^3$ kg/kg-mole M_n (Group 3). The representative solution in each group are shown in Table 1 and the regions for each group is also marked in Figure 7. The trend in addition rates for each group is shown in

Table 1: Three representative solutions picked from the Pareto-optimal front obtained by real-coded NSGA-II in each of the three problems are shown.

Source	PDI (kg/kg)	M_n (kg/kg-mole)	Reaction time (hr)
Benchmark	1.61	633.2	7.00
Problem 1			
Group 1	1.59	926.5	7.00
Group 2	1.87	3186.1	7.00
Group 3	1.97	9507.8	7.00
Problem 2			
Group 1	1.70	946.6	2.57
Group 2	1.90	3219.5	4.00
Group 3	1.99	9507.6	6.60
Problem 3			
Group 1	1.61	943.71	3.80
Group 2	1.88	3218.02	5.17
Group 3	1.97	9508.45	6.98

Figure 9. Each group has a distinct pattern of adding the three constituents. Although in all

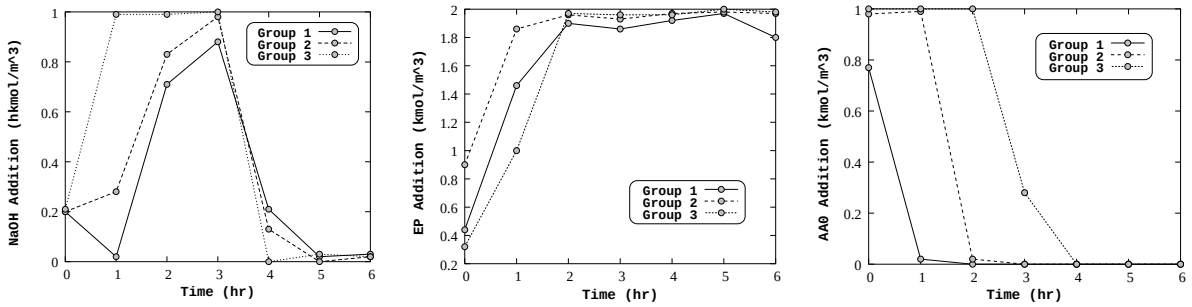

Figure 9: Time-variant addition rates of NaOH (left), EP (middle), and AA_0 (right) are shown for three representative solutions from each group on the Pareto-optimal front obtained using the real-coded NSGA-II for problem 1.

cases the NaOH addition must be high in the middle of the reaction time span and must be drastically reduced soon thereafter, polymers with very small M_n values need to be produced with almost zero addition of NaOH at $t = 1$ hr. Similarly, for producing polymers having very large M_n value, the comparatively small dosage of epichlorohydrin must be added initially

$t = 0$. Interesting trends are observed with the bisphenol-A addition rates. With more M_n requirement, more bisphenol-A must be continually added in the first few hours and then must be drastically reduced to zero. Although some such conclusions can be made by understanding the chemistry of the process, the above plots presents the trend quantitatively for a specific requirement of a polymer. We summarize these findings in the following:

Group 1: NaOH addition needs a delayed rise accompanied by a dip in first hour after starting from the low initial (lower bound) concentration; Epichlorohydrin addition needs a fast rise from low starting concentration; AA_0 addition needs a very quick fall from a somewhat high initial (near to upper bound) concentration.

Group 2: NaOH addition needs a delayed rise without any fall in between after starting from the low initial (lower bound) concentration; Epichlorohydrin additions needs a faster rise from relatively higher starting concentration; AA_0 addition needs a fall at the second hour starting from a high initial (upper bound) concentration.

Group 3: NaOH addition needs a sharp rise after starting from the low initial (lower bound) concentration; Epichlorohydrin addition needs a comparatively delayed rise from the low starting concentration; AA_0 addition needs a much late (after the second hour) fall from the high initial (upper bound) concentration.

The information above conveys that there lies a relationship between the solutions in the Pareto-optimal front and the system under consideration. The relationship can be regarded as the 'blue-print' of the system. Given a set of objectives, certain properties emerge from the system, not arbitrarily, but following some basic properties of the system. This relationship between the property of the system and the solutions of the Pareto-optimal front would be of tremendous importance to practitioners. Such observation has also been observed in other engineering design problems such as gearbox design, truss-structure design, etc. (Deb, 2003) and in other chemical process optimization problems, such as in multi-objective optimization of various processes like continuous casting of steel (Mitra and Ghosh, 2003), grinding of lead zinc ore (Mitra and Gopinath, 2003), iron ore sintering process (Mitra and Nath, 2003). Various rules that are evolved while characterizing the Pareto-optimal front can be used as innovative and intelligent thumb rules for epoxy polymerization. Such intelligent rules, though appeared simple, cannot be predicted easily for a complex system like epoxy polymerization and requires a thorough study and can be proved to be beneficial for most of the industrial systems where operating rules are generally slowly developed on the basis of experience. Pareto characterization can help an operator to control the operation of an epoxy polymerization reactor for maintaining different production schedules as well. Besides utilizing the mathematical optimality conditions outlined in standard texts (Deb, 2001; Mietinnen, 1999), there exist no other methods to decipher such vital information about a complex problem.

As several polymer species are involved in this polymerization study, the next interest would be to investigate the distribution of various species during the course of polymerization. Figure 10 (left plot) shows EE_n as a major product for polymers of lower dispersity and as dispersity increases, BE_n, another species, evolves as the nearest competitor. As the first order moment is directly related to the growth in chain length, variation of different polymer species plotted against the first-order moment (the right plot in Figure 10) shows the same trend. Polymer of higher M_n can be achieved by the growth of various other polymer species (BF_n, EF_n, BB_n), but as the number of species increases in the final distribution, the dispersity (PDI) also increases. The only way, therefore, to achieve higher M_n with a lower PDI is to

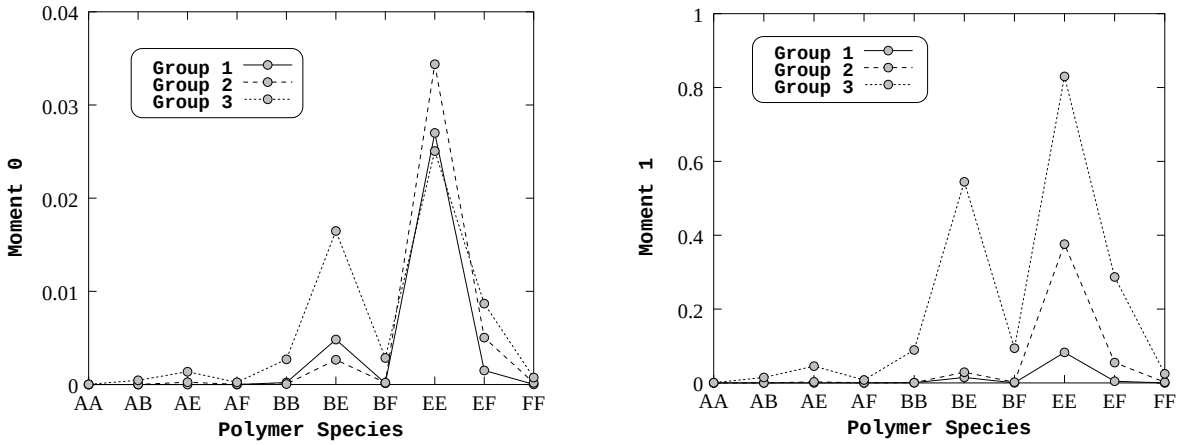

Figure 10: Change in concentration (left) and first-order moment of different polymer species are shown for representative solutions from each group obtained using real-coded NSGA-II for the M_n and PDI optimization.

have growth of any particular polymer species in course of polymerization. Here, for polymers having a higher PDI value, like one in Group 3, as more than one species start competing with the major species (EE_n), a higher M_n is achieved at the cost of a higher PDI value. For a user, if a higher polydispersity value is desired, a polymer with a higher M_n value will be desired and otherwise not. Even, based on the type of the species, the user can have choices among various options from the Pareto-optimal front solutions.

4.1.4 NSGA-II Parametric Study

The effect of varying GA parameters on the obtained Pareto-optimal front is studied next. If the population size is decreased from 250 to 100, the convergence rate is marginally faster (converging at iteration 27, instead of 34 iterations required with 250 population members), but the convergence speed is achieved at the cost of a poor spread of solutions on the front, especially towards the smaller PDI values. The effect of reducing the crossover probability (from 0.9 to 0.7) again leads to no significant change in converging and maintaining spread of solutions on the Pareto-optimal front. The effect of a change in the mutation probability (from 0.1 to 0.01) indicates no significant change as well. However, when the distribution index of the SBX operator is changed gradually from 0.01 to 100, a slightly inferior spread of Pareto-optimal solutions is obtained. When the distribution index for polynomial mutation operator is changed gradually from 0.01 to 100, a better spread of Pareto-optimal solutions is observed. In all cases, the performance is compared by keeping the same number of overall solution evaluations. Since the chosen NSGA-II parameter values (as described in problem 1) produce reasonably good convergence and maintenance of spread of solutions, we continue the rest of the simulations with these values.

4.2 Discussion on Problem 2

Next, we turn our attention to the optimization of M_n and the reaction time. The reaction time is also treated as an additional decision variable. The same upper and lower bounds for ingredients are kept for this case also, only addition being the reaction time which is varied

as follows: $0 \leq t_{\text{sim}} \leq 7$ hr. After the multi-objective optimization with NSGA-II, Pareto-optimal solutions (marked as 'NSGA-II') are plotted in Figure 11. To show the progress

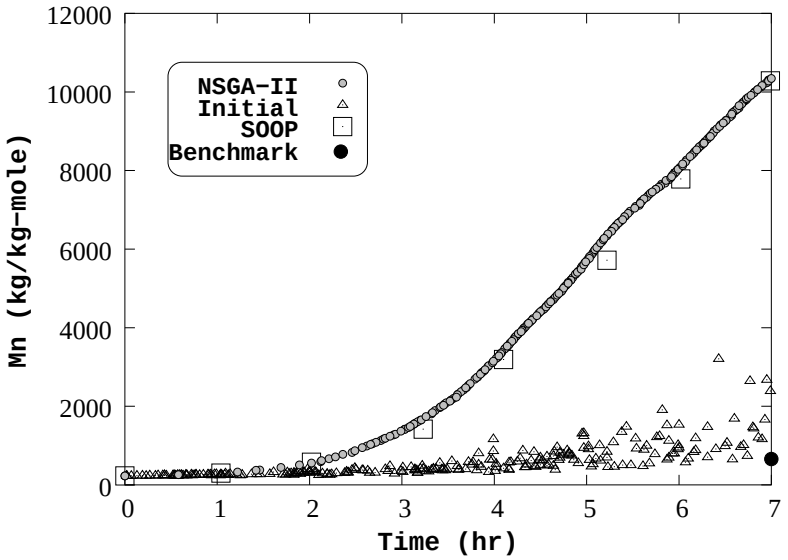

Figure 11: Obtained NSGA-II solutions for the M_n and reaction time optimization are compared with single-objective optimization (SOOP) solutions. Initial solutions of NSGA-II and the benchmark solution are also shown.

made by NSGA-II over iterations, the randomly created initial solutions are also shown as "Initial" in the figure. Once again, the random solutions within the chosen variable bounds are far from being close to the true Pareto-optimal front and NSGA-II is able to find a wide range of solutions starting with these random solutions. The generic observation is that a smaller reaction time produces polymers having a smaller M_n value. Once again, the M_n-time optimality front seems to be non-convex.

Like before, the ϵ-constraint method is used to constrain the reaction time values and M_n is maximized. A number of solutions (marked as 'SOOP') obtained with this repeatedly-applied single-objective formulation are also shown in Figure 11. Both the NSGA-II and the single-objective GA solutions agree at different places on the Pareto-optimal front, thereby establishing confidence in the optimality of the obtained solutions. However, the solutions obtained with a binary-coded NSGA-II in this problem are not as sparse, as shown in Figure 12. The solutions are inferior to those found by the real-coded NSGA-II. The discreteness introduced by the binary coding of the decision variables may have turned out to be too difficult for the binary-coded NSGA-II to maintain a good spread of solutions.

The NSGA-II solutions are much better than the benchmark solution, as shown in Figure 11. Among many solutions from the Pareto-optimal front obtained using the real-coded NSGA-II, one solution is chosen for demonstrating how the same quality polymer (having the same order of magnitude of M_n and PDI as compared to the benchmark solution) can be obtained in a much less processing time. A polymer having M_n value of 616 kg/kg-mole and PDI value of 1.57 is compared with the benchmark solutions (M_n of 633.2 kg/kg-mole and PDI of 1.61 obtained with 7 hr of reaction time). The obtained NSGA-II solution requires only 2.13 hr of reaction time, thereby requiring less than one-third time of completion from

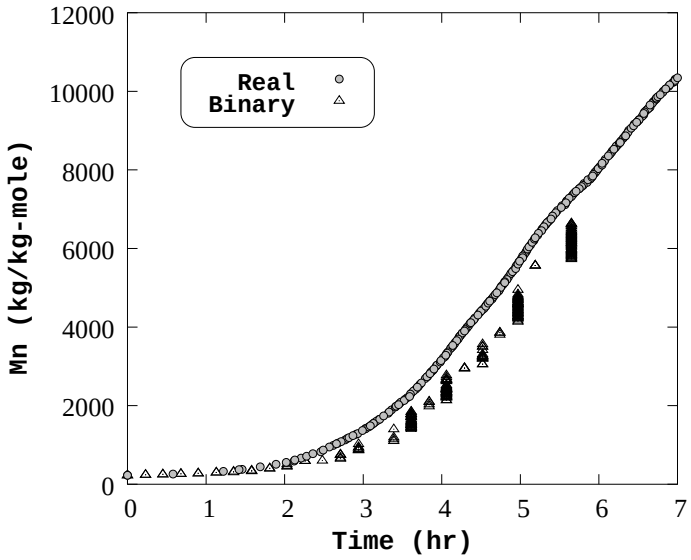

Figure 12: Real-coded and binary-coded NSGA-II solutions for the M_n and reaction time optimization.

the benchmark solution. This is equivalent to a 300% improvement in the productivity. In general, in case of problem 2, it has been observed that by manipulating several ingredient addition rates, a similar quality polymer (as compared to that obtained in problem 1) can be obtained in a much less processing time. But since the PDI minimization is not performed in problem 2, for similar M_n values in both cases, PDI values for the M_n-time Pareto-optimal solutions (problem 2) are higher as compared to the same for problem 1. This calls for an optimization of all three objectives for a better solution, a matter which we deal in Section 4.3.

4.2.1 Analyzing the Solutions of Problem 2

In problem 2, each of the solutions on the Pareto-optimal front carries information of 22 decision variables. A similar investigation, as done in problem 1, is intended to carry out next to identify whether these solutions on the time Pareto-optimal front have some commonality among themselves. As the reaction time is a variable here, it is not easy to perform a trend analysis similar to that used in problem 1. But it can be stated that in case of problem 2, though the patterns for addition of epichlorohydrin and bisphenol-A are almost similar to those in case of problem 1, the reduction in time is mainly achieved by adding more NaOH into the system right from the start of the reaction process. That is, for preparing a polymer having a similar value of M_n as that obtained problem 1, the reaction time is found to be reduced drastically in problem 2 as compared to the fixed reaction time of seven hours (for problem 1) by increasing the initial addition rate of NaOH, but at a cost of a slightly higher PDI value. As more and more NaOH, EP and AA_0 are added to the system, secondary species like BE_n, EE_n, BF_n, EF_n get generated (as compared to the primary species BB_n, AB_n) in the system contributing significantly in increasing the M_n value and by parallely increasing the PDI value. This is due to the fact that an increase in number of species increases the dispersity value.

Figure 13 compares the characteristics of equivalent solutions (having similar M_n values) taken from the Pareto-optimal front of problems 1 and 2. Table 1 shows these chosen solutions

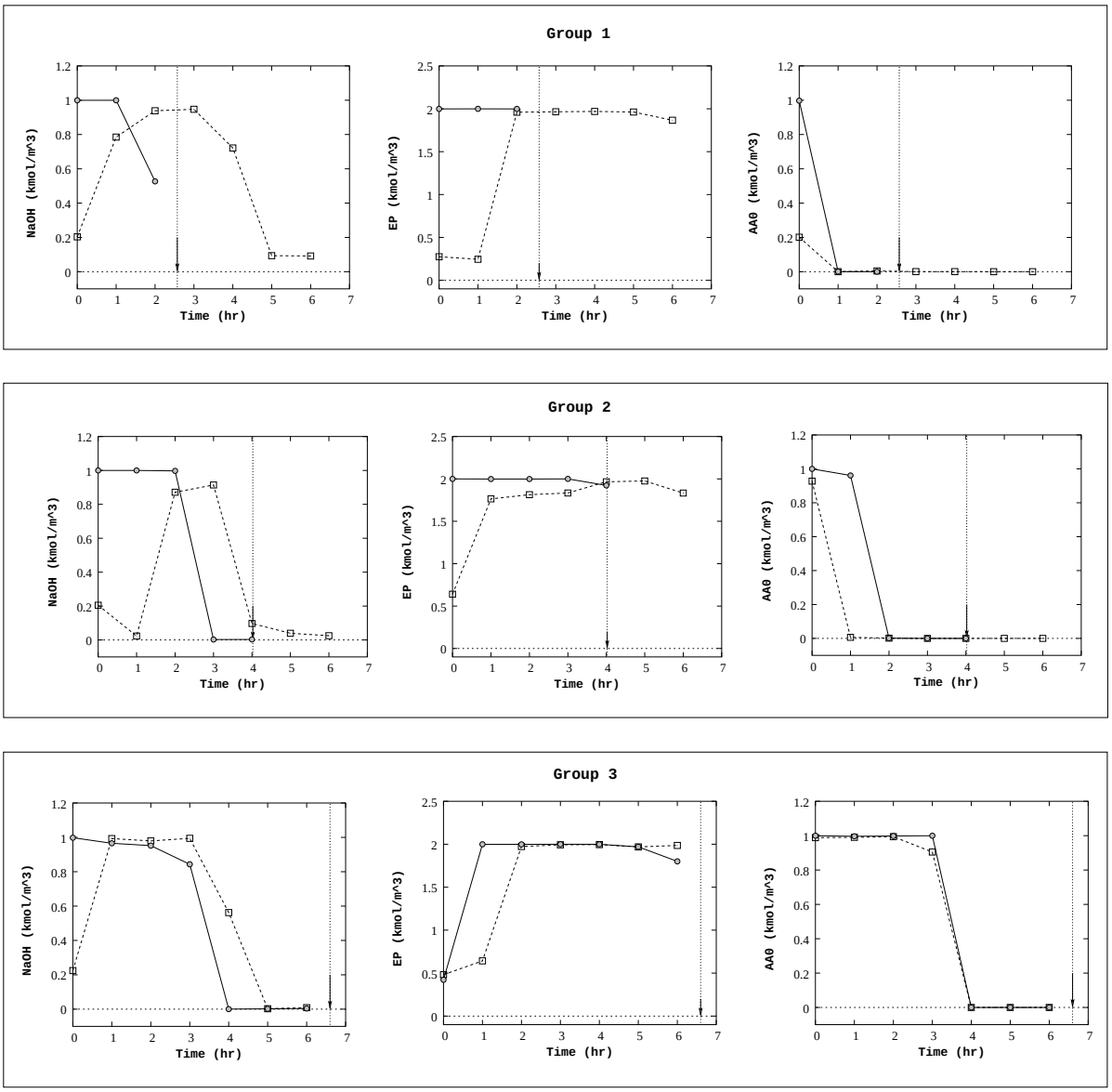

Figure 13: Effect of reaction time minimization on three equivalent optimal solutions from problems 1 and 2. Each row shows the comparison on a particular group. The dashed line is taken from problem 1 and solid line is taken from problem 2.

from the three groups in both fronts. The figure clearly shows that although similar quality polymers are produced by both optimizations, their operating principles are quite different. The top row shows the addition trend of NaOH, EP and AA_0 for two equivalent solutions from the low M_n group (Group 1). The middle row shows the same for the intermediate M_n group (Group 2) and the bottom row shows the same for the high M_n group (Group 3). The dashed pattern is for the 7 hr long solution obtained in problem 1, whereas the solid pattern is obtained in problem 2. For the latter case, the end of reaction time is shown by a vertical downward arrow.

The general trend captured in the NaOH addition rate is that NaOH must be added at a high rate to start with, then must continue with the a similar rate for some time and then must be reduced. Here, NaOH is realized to take part not only in the initiation process, but also in the growth of the chain length. Epichlorohydrin (EP), on the other hand, is found to be needed more (as compared to problem 1) in the beginning for low and medium M_n requirement, but for high M_n values, EP addition needs to be slowed down in the beginning. Trends for bisphenol-A (AA_0) remain almost the same as in problem 1 (that is, adding more in the beginning and using the same for polymerization for the rest of the time). Barring the details, it can be concluded that although all three ingredients are required to be added more in the beginning (in comparison to a 7 hr long polymerization process) to achieve a similar polymer in a smaller time, the addition patterns of ingredients, requiring the formation of the desired polymer are similar.

Next, the species distribution for all three groups mentioned in Table 1 for problem 2 is plotted in terms of moment zero (left plot in Figure 14) and the first-order moment (right plot in Figure 14) carrying information on concentration of species and growth of chain length, respectively. From the left plot in Figure 14, it can be seen that the species BB_n (AB_n being its

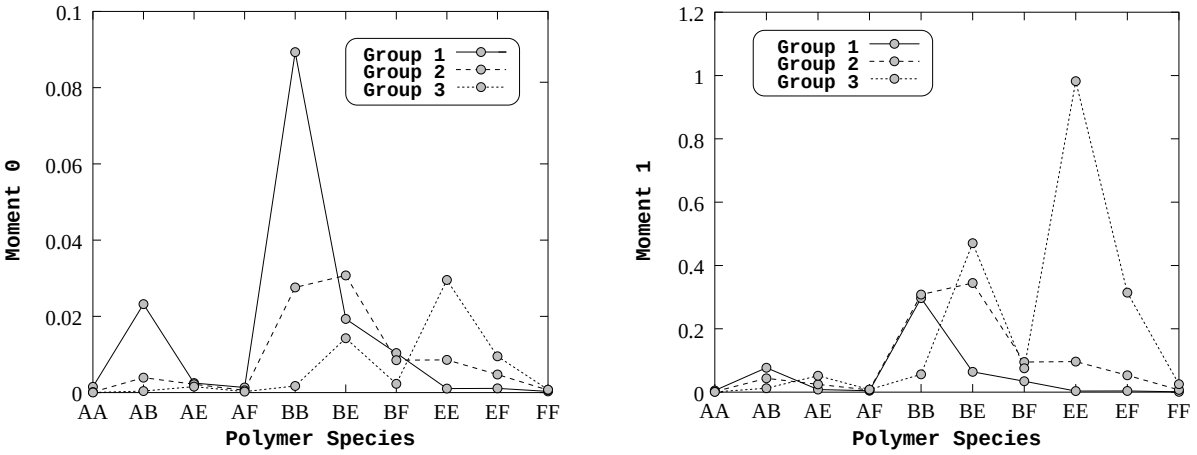

Figure 14: Change in concentration (left) and first-order moment of different polymer species are shown for representative solutions from each group obtained using real-coded NSGA-II for the M_n and reaction time optimization for problem 2.

close competitor) is a primary species and there is a clear shift towards BE_n, BF_n, EE_n, EF_n for the cases with higher time and M_n. In fact, with more addition of NaOH, epichlorohydrin and higher bisphenol-A affect the process for producing more BE_n, BF_n, EE_n, EF_n type of polymer species. Similarly first moments for all the cases have also indicated a large growth in chain length for BE_n, BF_n, EE_n, EF_n type of species for a longer time and a higher M_n value (right plot in Figure 14). Based on an operator's requirement, the species distribution

in the final product can also be controlled more directly by adding additional constraints in the problem formulation.

4.3 Discussion on Problem 3

The Pareto-optimal solutions for three objective optimization problem (henceforth called as 'three-dimensional Pareto-optimal front') are plotted in Figure 15.

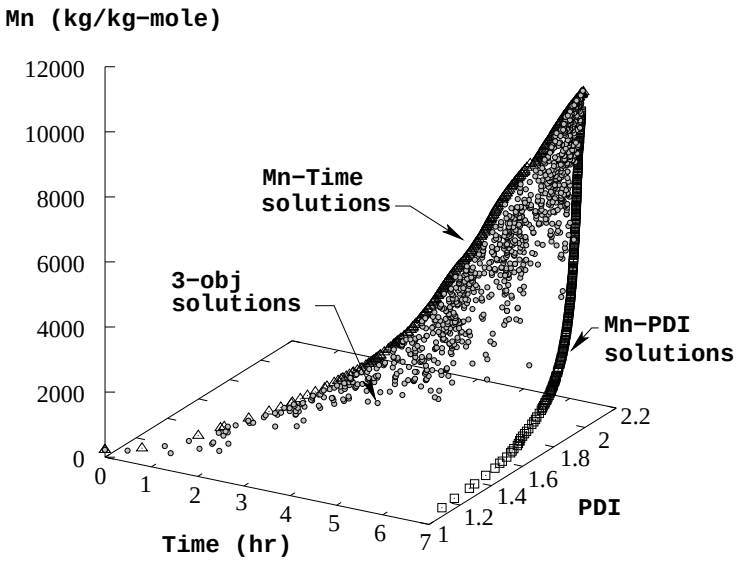

Figure 15: NSGA-II solutions obtained from the three-objective optimization are shown. The fronts obtained in the previous two problems are found to lie on two edges of the obtained three-dimensional front.

When the previously obtained two fronts (from problems 1 and 2) are drawn on this plot, they are interestingly found to be the limiting fronts lying on the two edges of the three-dimensional front. First, we deduce the following conclusions from the three-dimensional front:

1. The three-dimensional Pareto-optimal front is a non-convex front.

2. For low values of M_n, polymers can be prepared in much less than 7 hours. Since the reaction time is also minimized in problem 3, there exists almost no solution requiring as large as 7 hr to produce a polymer having a low M_n.

3. For high values of M_n, more reaction time is needed. As in the polymerization process, the repeat units get added with the monomer to form a longer chain-length, a polymer having a high value of M_n requires more time to form than a polymer having a low value of M_n. But for a desired M_n and PDI combination, one can find a solution requiring smaller than 7 hours to do the job, but the occurrence of such a quick operation gets reduced with the requirement of higher and higher M_n values.

4. Finally, for the maximum M_n requirement, there exists only one solution (with M_n=10,402.12 kg/kg-mole), requiring 7 hours (the maximum allowed) of reaction time, but also producing the largest PDI value (1.985). Such is the trade-off often observed in a multi-objective

problem and Figure 15 shows many such trade-off solutions producing different values of M_n and PDI and requiring different amount of reaction time.

5. Another interesting aspect is that for any fixed reaction time, the M_n-time optimization (problem 2) solutions produced the maximum M_n value and there exists no other solution producing a better (smaller) PDI value and an identical M_n value. But, the three-dimensional Pareto-optimal front provides more information about the trade-off than both two-objective Pareto-optimal fronts, discussed earlier.

On the three-dimensional Pareto-optimal front, there seems to be a larger concentration of solutions towards the M_n-time front (in other words, there are more solutions requiring a smaller processing time). To understand this trend better, we repeat problem 1 for different reaction times. We force the reaction time to end at $t_{\text{sim}} = 1$ hr, 2 hr and so on, and collect all the obtained Pareto-optimal solutions. Thereafter, we perform a non-domination check considering all three objectives (maximization of M_n, minimization of PDI, and minimization of t_{sim}) and the resulting solutions are plotted in Figure 16. An interesting aspect is revealed. For identical PDI values, there exists a smaller reaction time solution outperforming a larger time solution. This feature of the problem produces a *third* boundary on this three-dimensional Pareto-optimal front, thereby showing the complete bounded trade-off surface of interactions. The figure also shows the projection of the three-dimensional trade-off surface on the three

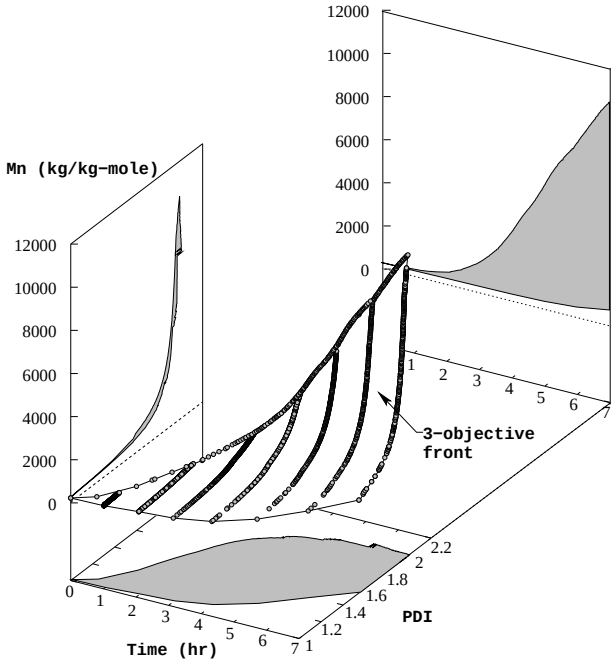

Figure 16: Several optimizations of problem 1 with different t_{sim} values help find the third boundary of true three-dimensional Pareto-optimal front.

two-objective planes. It is clear that although there are some rooms for trade-offs among M_n-time and PDI-time combinations, in terms of M_n-PDI combinations, there is almost a straightforward trade-off. This fact also indicates that the consideration of minimization of the reaction time is important in this problem to reveal interesting time-saving trade-off solutions.

To understand the three-objective interactions further, we consider the PDI-time and M_n-time projections (as shown shaded in Figure 16) and plot them again in Figure 17 in the left and right plots, respectively. On the PDI-time plot, we show contours of fixed M_n values.

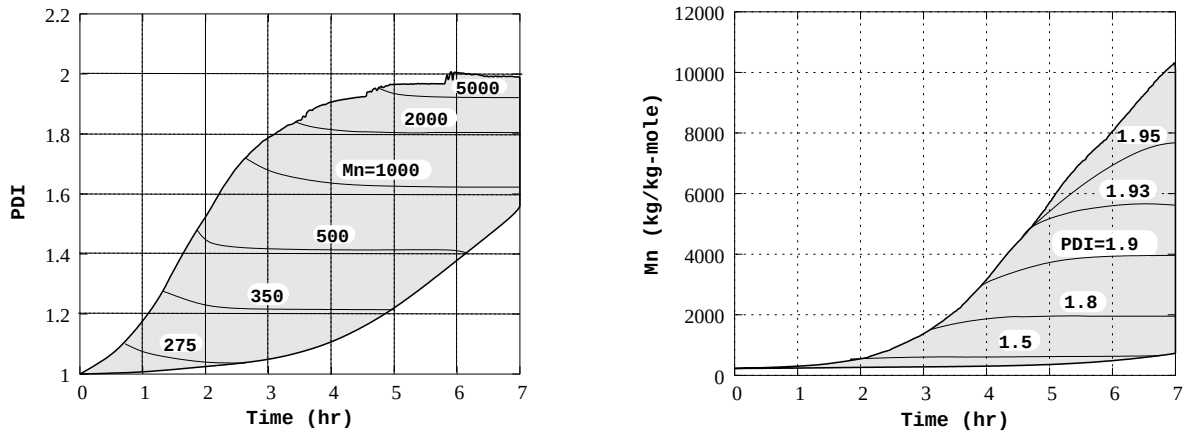

Figure 17: PDI and reaction time interaction (left) and $M_n$and reaction time interaction (right) reveal interesting properties about the optimal operating conditions of the polymerization problem. In the left figure, M_n is shown in kg/kg-mole.

For each contour line, the trade-off between PDI and reaction time is clear. However, what is more interesting is that the trade-off becomes marginal away from the M_n-time Pareto-optimal front (the upper boundary). To achieve a small advantage in the PDI value, a large reaction time is necessary. Since the slope of the three-objective Pareto-optimal front is quite small at these regions, the NSGA-II has found very few solutions away from the M_n-time boundary. It can then be concluded that for a fixed M_n requirement, it is better to consider the trade-off solutions close to the M_n-time Pareto-optimal boundary. Another aspect is the rate at which the 'region of optimality' reduces for an increasing M_n value. For example, if a polymer with M_n greater than 5,000 kg/kg-mole is desired, there does not exist too many options in terms of the reaction time. The right plot in Figure 17 shows contour lines for fixed values of PDI on the three-objective Pareto-optimal front. Once again the trade-off between M_n and reaction time is clear from the plot. Although a similar conclusion (about choosing solutions close to the M_n-time optimal boundary) can be made for smaller PDI values, for large PDI values (such as PDI=1.95) the trade-off between M_n and reaction time is quite substantial. These two plots can be chosen to determine a suitable operating condition for the epoxy polymerization problem.

Once again, like problems 1 and 2, three similar M_n solutions are chosen from the three-dimensional Pareto-optimal front to investigate the effect. These solutions are presented in Table 1 as well. It can be easily observed that in case of problem 3, similar M_n (943.71 kg/kg-mole as compared to 946.6 kg/kg-mole of problem 2) can be achieved with a smaller PDI (1.61 as compared to 1.70 of problem 2), but at the cost of a little more reaction time (3.80 hr as compared to 2.57 hr). This was one difficulty in problem 2 stated earlier: similar M_n is achieved at the cost of more PDI, but with the smallest possible reaction time. This trade-off between PDI and time can be very effectively used by an operator for different kinds of operations, i.e. the operator can choose a solution from the results of problem 3 if the product quality is extremely important for a client (of course, at the cost of productivity) or choose a solution from the results of problem 2 if the product quality requirement is not so stringent, but the

productivity is of utmost importance. For example, one solution of problem 3 recommends M_n=630 kg/kg-mole, PDI=1.54 and reaction time equal to 2.27 hr, but for the benchmark case, they are as follows: M_n=631, PDI=1.61 and reaction time equal to 7 hr. So, a much better quality polymer (with reduced polydispersity) can be produced in less than one-third of the actual reaction time (recommending more than 300% productivity improvement). Although the problem 2 finds two similar polymers (with M_n=616 and 666 kg/kg-mole) in slightly less reaction times (2.13 and 2.21 hr, respectively), the PDI values (1.57 and 1.60, respectively) are slightly inferior. This shows the effectiveness of using all three objectives in a study, where all of them are important in the decision making process.

The extreme solutions are extracted from the above studies and are presented in Table 2. For a fixed reaction time, the trade-off between M_n and PDI is clear from the table. Interestingly, Figure 16 also indicates that there exist polymers with a continuous variation of M_n and PDI between the two extreme cases for each reaction time shown in the table.

Table 2: Extreme polymers extracted from the three-objective Pareto-optimal front (Figure 16) for different reaction times of the epoxy polymerization process.

Reaction	Maximum M_n		Minimum PDI	
time (hr)	M_n (kg/kg-mole)	PDI	M_n (kg/kg-mole)	PDI
1	294.524	1.147	238.519	1.005
2	561.872	1.522	236.751	1.008
3	1379.881	1.788	268.711	1.034
4	3162.372	1.906	282.337	1.069
5	5333.817	1.934	289.212	1.075
6	7949.621	1.966	279.515	1.096
7	10402.122	1.985	271.734	1.083

The three-objective optimization study also leads to the investigation of species distribution for the groups presented in Table 1 (problem 3). Both the zero-th and the first-order moments are shown for these groups in left and right plots in Figure 18, respectively. It is clear from the figures that in problem 3, there is no development of new species in the course of the polymerization (as also observed in problem 2) and distribution of species is much organized in these two cases. Here, BE_n and EE_n are the primary species when M_n requirement is low. With a larger requirement of M_n, secondary species, like BB_n, BF_n, and EF_n, get generated, a phenomenon which was observed with solutions from problem 1 as well.

5 Conclusions and Extensions

A well-validated model consisting of a large number of moment-based ordinary differential equations has been utilized for the multi-objective optimization of epoxy polymerization process using an evolutionary algorithm (NSGA-II). The aim of the study has been to extract the discrete addition patterns of the ingredients for optimizing various objectives simultaneously (e.g. problem 1: maximization of M_n with minimization of PDI, problem 2: maximization of M_n with minimization of reaction time and problem 3: maximization of M_n, minimization of PDI, and minimization of reaction time). The salient features which resulted from this work are:

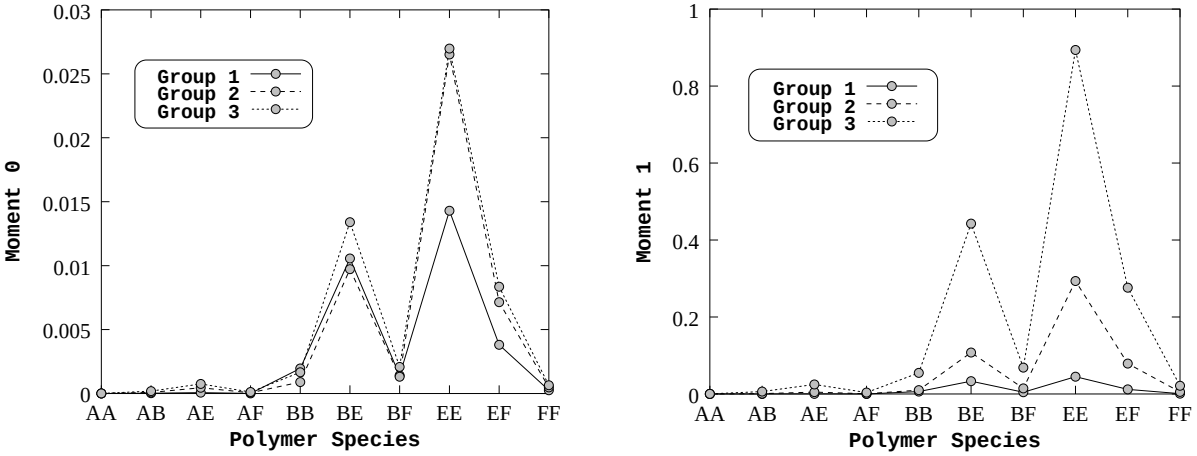

Figure 18: Change in concentration (left) and first-order moment of different polymer species are shown for representative solutions from each group obtained using real-coded NSGA-II for the M_n, PDI, and reaction time optimization.

1. All three problems considered here have displayed a non-convex Pareto-optimal front. NSGA-II has been able to find solutions on or near the true Pareto-optimal front of the problems. This has been validated by solving the multi-objective problems using a single-objective preference based optimization method (ϵ-constraint method). The advantage of using the NSGA-II is that it has found multiple (as many as 250) Pareto-optimal solutions in a single simulation run.

2. In all cases, the real-coded NSGA-II has been observed to perform better than a binary-coded NSGA-II in terms of locating the true Pareto-optimal front, as well as in maintaining a good diversity of Pareto-optimal solutions.

3. Solutions to problem 1 have led to better M_n (higher) and PDI (lower) values as compared to the available benchmark data on the problem. Solutions to problems 2 and 3 have provided optimal operating conditions of producing polymers with a similar M_n, but with a smaller PDI value in as much as one-third of the reaction time (as compared to benchmark solution).

4. Importantly, the multi-objective optimization of the epoxy polymerization process has led to the discovery of certain operating principles (addition time-pattern of three ingredients (NaOH, Epichlorohydrin, and Bisphenol-A)) for all high-performance solutions. The trade-off between the objectives have been clearly characterized by showing and contrasting a representative additive pattern of all three ingredients. Such information brings out the 'blue-print' of the optimal operating conditions of a chemical process and are often important from the point of view of an operator.

5. The distribution of various species for all three problems has been studied to show how a change in ingredient addition pattern has changed different species distribution during the course of polymerization. This study has also showed the change in the distribution of various species for different degrees of polymerization, which has provided further insight into the underlying process.

6. Semi-batch mode of operations for epoxy processes has been found to be much superior

to one-time addition processes (batch mode).

7. The study of multi-objective optimization on all three objectives has demonstrated a systematic procedure of identifying the overall trade-off relationship among the objectives. Since the NSGA-II procedure is capable of finding multiple Pareto-optimal solutions with a good spread in them, it has been possible to perform such a study. Once such a trade-off relationship is identified and a mapping of every solution from the trade-off surface to the corresponding decision variables is established, such information not only provides a deep insight to the problem but also can be used an 'operating chart' for executing an optimal task. For a desired requirement in any combination of objectives, a suitable operating solution can be found from such a chart.

The unavailability of experimental data for M_n and PDI in this problem has been one of the shortcomings of this study. Only data for some primary oligomer concentration profiles were available as experimental data with which the present model has been validated. With the availability of more experimental data, the mathematical model can further verified. In fact, using the recommendations of this optimization study, experiments can be performed and effectiveness of the current recommendations can be validated as well. Specific growth of a particular polymer species out of several considered can also be performed depending on the requirements of the industry. Nevertheless, this extensive study has clearly demonstrated the usefulness of a multi-objective optimization procedure in understanding a complex process, such as the epoxy polymerization process, and outlined a systematic search procedure for arriving at the optimal operating chart with trade-off information about salient objectives of the process – a matter which would ultimately have a tremendous practical importance.

Acknowledgments

Authors acknowledge the support of TRDDC as well as TCS management in course of this work. K. Deb acknowledges the Bessel Research award from Alexander von Humboldt Foundation Germany.

Nomenclature

AA_0	=	Bisphenol-A (monomer)
AB_m	=	Molecular species defined in Appendix B
B	=	Sodium phenoxide end group
d_i	=	Crowding distance of i-th population member
E	=	Glycidyl ether group
EP	=	Epichlorohydrin
F	=	Chlorohydrin end group
F_i	=	i-th fitness function
I_{ij}	=	j-th objective function for the i-th problem
I_i	=	Objective function vector of the i-th problem
$\mathcal{I}_j^m$	=	Index of j-th population member in the ordered set using m-th objective
K_i	=	Reaction rate constant
M_n	=	Number average molecular weight (kg/kg-mole)
M_w	=	Weight average molecular weight (kg/kg-mole)
NaOH	=	Sodium hydroxide
N_{Digit}	=	Number of u_i values used in NSGA-II
N_{Gen}	=	Iteration number
$N_{\text{Gen,max}}$	=	Maximum number of iterations
N_{Pop}	=	Number of solutions in the NSGA-II population
PDI	=	Poly-Dispersity index (M_w/M_n)
$\mathcal{P}_N$	=	Parent population at N-th iteration
p_m	=	Mutation probability
p_c	=	Crossover probability
$\mathcal{Q}_N$	=	Offspring population at N-th iteration
$\mathcal{R}_N$	=	Combined population at N-th iteration
rnk_i	=	Non-dominated ranking of the i-th solution
t_{sim}	=	Reaction time (hr)
$\mathbf{U}$	=	Vector of control variables, $U_1(t)$, $U_2(t)$, $U_3(t)$
$U_{\text{pop,j}}^{(i)}$	=	Value of control variable at the end of k-th time interval for i-th solution
$U_{\text{pop,j}}^{\max}$, $U_{\text{pop,j}}^{\min}$	=	Upper and lower bounds on control variable at the end of k-th time interval
$\mathbf{v}$	=	Lagrange multiplier vector used in Fritz-John necessary condition
$\mathbf{x}$	=	Vector of state variables, x_i
$\boldsymbol{\gamma}$	=	Lagrange multiplier vector used in Fritz-John necessary condition
λ_j^k	=	k-th moment of the j-th species, for $k = 0, 1, 2$
δ_{m0}	=	A factor (0 for molecular weight determination without epichlorohydrin and 1 with epichlorohydrin)
[ ]	=	Concentrations
$<_c$	=	Crowded comparison operator used in NSGA-II
$\preceq_c$	=	Constrain-domination operator

References

Batzer, H. and S. S. Zahir, "Studies in the molecular weight distribution of epoxide resins. IV. Molecular weight distribution of epoxide resins made from bisphenol-A and epichlorohydrin", J. App. Poly. Sci., 21, 1843 (1977).

Bhaskar, V., S. K. Gupta and A. K. Ray, "Applications of Multi-objective Optimization in Chemical Engineering", Reviews in Chemical Engineering, 16(1), 1 (2000).

Butala, D., K. Y. Choi, and M. K. H. Fan, "Multiobjective Dynamic Optimization of a Semibatch Free-Radical Copolymerization Process with Interactive CAD Tools" Comput. Chem. Eng.,12, 1115 (1988).

Chakravarthy, S. S. S., D. N. Saraf, and S. K. Gupta, "Use of Genetic Algorithms in the Optimization of Free Radical Polymerizations Exhibiting the Trommsdorff Effect," J. Appl. Poly. Sci., 63, 529 (1997).

Chankong, V. and Y. Y. Haimes, Multiobjective Decision Making - Theory and Methodology, Elsevier, New York (1983).

Choi, K. Y., and D. N. Butala, "An Experimental-Study of Multi-objective Dynamic Optimization of a Semibatch Copolymerization Process," Poly. Eng. Sci.,31, 353 (1991).

Deb, K., Multi-objective Optimization Using Evolutionary Algorithms, Wiley, Chichester, U. K. (2001).

Deb, K., "An efficient constraint handling method for genetic algorithms", Computer Methods in Applied Mechanics and Engineering, 186 (2 - 4), 311 (2000).

Deb, K. and Agrawal, R. B., "Simulated binary crossover for continuous search space". Complex Systems, 9 (2), p 115 - 148, (1995).

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multi-objective Genetic Algorithms", IEEE Transactions on Evolutionary Computation, 6 (2), 182 (2002).

Deb, K., "Unveiling Innovative Design Principles By Means of Multiple Conflicting Objectives". Engineering Optimization, 35 (5), p 445–470 (2003).

Fan, L. T., C. S. Landis, and S. A. Patel, Frontiers in Chemical Reaction Engineering, L. K. Doraiswamy and R. A. Mashelkar, eds., Wiley Eastern, New Delhi, p 609 (1984).

Farber, J. N., "Steady State Multiobjective Optimization of Continuous Copolymerization Reactors," Poly. Eng. Sci., 26, 499 (1986).

Farber, J. N. Handbook of polymer Science and Technology, Vol. 1, N. P. Cheremisinoff, ed., Dekker, New York, page 429 (1989).

Garg, S., and S. K. Gupta, "Multiobjective Optimization of a Free Radical Bulk Polymerization Reactor Using Genetic Algorithm," Macromol. Theor. Simul., 8, 46 (1999).

Garg, S., S. K. Gupta and D. N. Saraf, "On-Line Optimization of Free Radical Bulk Polymerization Reactors in the Presence of Equipment Failure," J. Appl. Poly. Sci., 71, 2101 (1999).

Goldberg, D. E., Genetic Algorithms in Search, Optimization and Machine Learning, Addition-Wesley, Reading, MA (1989).

Gupta, R. R., and S. K. Gupta, "Multi-objective Optimization of an Industrial Nylon-6 Semibatch Reactor System Using Genetic Algorithm", J. Appl. Poly. Sci., 73, 729 (1999).

Kumar, A. and S. K. Gupta, Reaction Engineering of Step Growth Polymerization, Plenum, New York (1987).

Majumdar, S., K. Mitra, and S. Raha, "Analysis of Semibatch Operation in Epoxy Polymerization", in press (2003).

McAdams, L. V. and J. A. Gannon, Encyclopedia of Polymer Science and Engineering (2nd ed.), Wiley, New York, Vol. 6, page 322 (1986).

Miettinen, K., Nonlinear Multiobjective Optimization, Kluwer, Boston (1999).

Mitra, K., K. Deb, and S. K. Gupta, "Multi-objective Dynamic Optimization of an Industrial Nylon 6 Semibatch Reactor Using Genetic Algorithm", J. App. Poly. Sci., 69, 69 (1998).

Mitra, K. and R. Gopinath, "Multiobjective Optimization of an Industrial Grinding Operation Using Elitist Nondominated Sorting Genetic Algorithm", accepted in Chem. Eng. Sci. (2003).

Mitra, K. and S. Ghosh, "Multiobjective optimization of Spray Cooling in an Industrial Thin Slab Casting Machine using Genetic Algorithm", under review (2003).

Mitra, K., S. Raha, and S. Majumdar, "Multiobjective dynamic optimization of Epoxy Polymerization Process using Nondominated Sorting Genetic Algorithm II", under review (2003).

Mitra, K., and N. Nath, "Multiobjective Pareto Optimization of Two-layer Sintering for different bed heights by Genetic Algorithm", under review (2003).

Raha, S. and S. K. Gupta, "A General Kinetic Model for Epoxy Polymerization", J. App. Poly. Sci., 70, 1859 (1998).

Raha, S., S. Majumdar and K. Mitra, "Effect of NaOH Addition in Epoxy Polymerization Process: A Single and Multi-objective Approach through Genetic Algorithm, Macromolecular Theory and Simulation", accepted in Macromolecular Theory Simulation, in press (2003).

Ray, W. H., "On the Mathematical Modeling of Polymerization Reactors", J. Macromol. Sci., Revs. Macromol. Chem., C8, 1(1972).

Sareen, R., and S. K. Gupta, "Multiobjective Optimization of an Industrial Semibatch Nylon 6 Reactor," J. Appl. Poly. Sci., 58, 2357 (1995).

Srinivas, N., and K. Deb, "Multiobjective Function Optimization Using Nondominated Sorting Genetic Algorithms", Evol. Comput., 2, 3 (1995).

Tsoukas, A., M. V. Tirrell, and G. Stephanopoulos, "Multi-objective Dynamic Optimization of Semibatch Copolymerization Reactors," Chem. Eng. Sci., 37, 1785 (1982).

Wajge, R. M., and S. K. Gupta, "Multiobjective Dynamic Optimization for a Nonvaporizing Nylon-6 Batch Reactor," Poly. Eng. Sci., 34, 1161 (1994).

Walas, S. M. Modeling with Differential Equations in Chemical Engineering. Butterworth-Heinemann, Boston (1991).

A Reaction Scheme for Epoxy Polymerization

Here, we present five reaction schemes of the epoxy polymerization process. The details can be found in Majumdar et al. (2003) and in Raha and Gupta (1998).

1. $$\text{—OH} + \text{NaOH} \xrightarrow{k_1} \text{—ONa} + \text{H}_2\text{O}$$
   $$\text{—A} + \text{NaOH} \xrightarrow{k_1} \text{—B} + \text{W}$$

2. $$\text{—ONa} + \text{CH}_2\text{—CHCH}_2\text{Cl} \xrightarrow{k_2/\text{H}_2\text{O}} \text{—OCH}_2\text{CHCH}_2\text{Cl} + \text{NaOH}$$
   $$\text{—B} + \text{EP} \xrightarrow{k_2/\text{H}_2\text{O}} \text{—F} + \text{NaOH}$$

3. $$\text{—OCH}_2\text{CHCH}_2\text{Cl} + \text{NaOH} \xrightarrow{k_3} \text{—OCH}_2\text{CH—CH}_2 + \text{NaCl} + \text{H}_2\text{O}$$
   $$\text{—F} + \text{NaOH} \xrightarrow{k_3} \text{—E} + \text{NaCl} + \text{W}$$

4. $$\text{—ONa} + \text{CH}_2\text{—CHCH}_2\text{O—} \xrightarrow{k_4/\text{H}_2\text{O}} \text{—OCH}_2\text{CHCH}_2\text{O—} + \text{NaOH}$$
   $$\text{—B} + \text{E—} \xrightarrow{k_4/\text{H}_2\text{O}} \text{—(BE)—} + \text{NaOH}$$

5. $$\text{—ONa} + \text{ClCH}_2\text{CHCH}_2\text{O—} \xrightarrow{k_5} \text{—OCH}_2\text{CHCH}_2\text{O—} + \text{NaCl}$$
   $$\text{—B} + \text{F—} \xrightarrow{k_5} \text{—(BF)—} + \text{NaCl}$$

The end groups present in the reaction mass are as follows:

—A	—OH
—B	—ONa
—E	—CH$_2$CH—CH$_2$
—F	—OCH$_2$CHOHCH$_2$Cl

B Different Molecular Species Present in Reaction Mass

Here, we describe different molecular species (monomer and polymer) present in the reaction mass:

AA_n $\mathrm{H(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROH}$

AB_n $\mathrm{H(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nORONa}$

AE_n $\mathrm{H(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2CH{-}CH_2}$ (O)

AF_n $\mathrm{H(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2Cl}$

BB_n $\mathrm{Na(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nORONa}$

BE_n $\mathrm{Na(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2CH{-}CH_2}$ (O)

BF_n $\mathrm{Na(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2Cl}$

EE_n $\mathrm{CH_2{-}CHCH_2(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2CH{-}CH_2}$ (O, O)

EF_n $\mathrm{CH_2{-}CHCH_2(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2Cl}$ (O)

FF_n $\mathrm{ClCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2(OROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2)_nOROCH_2\underset{\displaystyle OH}{\underset{|}{C}}HCH_2Cl}$

EP $\mathrm{CH_2{-}CHCH_2Cl}$ (O)

where, $\mathrm{R} = -\mathrm{H_4C_6}-\underset{\displaystyle CH_3}{\underset{|}{\overset{\displaystyle CH_3}{\overset{|}{C}}}}-\mathrm{C_6H_4}-$ and $n = 0, 1, 2, 3, \ldots$.

C Model Equations

Here, we present a few example model equations.

1. Equation for concentration of AB_n:

$$\begin{aligned}\frac{d[\mathrm{AB}_m]}{dt} &= k_1[NaOH]\,(2[\mathrm{AA}_m]-[\mathrm{AB}_m]) - k_2[\mathrm{EP}][\mathrm{AB}_m] - k_4[\mathrm{AB}_m]\sum_{n=0}^{\infty}([\mathrm{AE}_n]+\\ &\quad [\mathrm{BE}_n]+2[\mathrm{EE}_n]+[\mathrm{EF}_n]) - k_5[\mathrm{AB}_m]\sum_{n=0}^{\infty}([\mathrm{AF}_n]+[\mathrm{BF}_n]+[\mathrm{EF}_n]+2[\mathrm{FF}_n])\\ &\quad + \left(2k_4\sum_{i=0}^{m-1}[\mathrm{BB}_i][\mathrm{AE}_{m-i-1}] + k_4\sum_{i=0}^{m-1}[\mathrm{AB}_i][\mathrm{BE}_{m-i-1}] + 2k_5\sum_{i=0}^{m-1}[\mathrm{AB}_i][\mathrm{BF}_{m-i-1}]\right.\\ &\quad \left. +2k_5\sum_{i=0}^{m-1}[\mathrm{BB}_i][\mathrm{AF}_{m-i-1}]\right)\delta_{m0},\end{aligned}$$

where $\delta_{m0} = \begin{cases} 0, & \text{for } m=0,\\ 1, & \text{for } m>0.\end{cases}$

2. The zero-th, first, and second moment equations of species AB are as follows:

$$\begin{aligned}\frac{d[\lambda^0_{\mathrm{AB}}]}{dt} &= -k_1[\mathrm{NaOH}]\lambda^0_{\mathrm{AB}} + 2k_1[\mathrm{NaOH}]\lambda^0_{\mathrm{AA}} - k_2[\mathrm{EP}]\lambda^0_{\mathrm{AB}} - k_4\lambda^0_{\mathrm{AB}}(\lambda^0_{\mathrm{AE}}+2\lambda^0_{\mathrm{EE}}+\lambda^0_{\mathrm{EF}})\\ &\quad -k_5\lambda^0_{\mathrm{AB}}(\lambda^0_{\mathrm{AF}}+\lambda^0_{\mathrm{EF}}+\lambda^0_{\mathrm{FF}}) + 2k_4\lambda^0_{\mathrm{BB}}\lambda^0_{\mathrm{AE}} + 2k_5\lambda^0_{\mathrm{BB}}\lambda^0_{\mathrm{AF}},\\ \frac{d[\lambda^1_{\mathrm{AB}}]}{dt} &= -k_1[\mathrm{NaOH}]\lambda^1_{\mathrm{AB}} + 2k_1[\mathrm{NaOH}]\lambda^1_{\mathrm{AA}} - k_2[\mathrm{EP}]\lambda^1_{\mathrm{AB}} - k_4\lambda^1_{\mathrm{AB}}(\lambda^0_{\mathrm{AE}}+2\lambda^0_{\mathrm{EE}}+\lambda^0_{\mathrm{EF}})\\ &\quad -k_5\lambda^1_{\mathrm{AB}}(\lambda^0_{\mathrm{AF}}+\lambda^0_{\mathrm{EF}}+2\lambda^0_{\mathrm{FF}}) + 2k_4(\lambda^1_{\mathrm{BB}}\lambda^0_{\mathrm{AE}}+\lambda^0_{\mathrm{BB}}\lambda^1_{\mathrm{AE}}) + 2k_5(\lambda^1_{\mathrm{BB}}\lambda^0_{\mathrm{AF}}+\lambda^0_{\mathrm{BB}}\lambda^1_{\mathrm{AF}})\\ &\quad +k_4\lambda^0_{\mathrm{AB}}\lambda^1_{\mathrm{BE}} + k_5\lambda^0_{\mathrm{AB}}\lambda^1_{\mathrm{BF}},\\ \frac{d[\lambda^2_{\mathrm{AB}}]}{dt} &= -k_1[\mathrm{NaOH}]\lambda^2_{\mathrm{AB}} + 2k_1[\mathrm{NaOH}]\lambda^1_{\mathrm{AA}} - k_2[\mathrm{EP}]\lambda^2_{\mathrm{AB}} - k_4\lambda^2_{\mathrm{AB}}(\lambda^0_{\mathrm{AE}}+2\lambda^0_{\mathrm{EE}}+\lambda^0_{\mathrm{EF}})\\ &\quad -k_5\lambda^2_{\mathrm{AB}}(\lambda^0_{\mathrm{AF}}+\lambda^0_{\mathrm{EF}}+2\lambda^0_{\mathrm{FF}}) + 2k_4(\lambda^2_{\mathrm{BB}}\lambda^0_{\mathrm{AE}}+\lambda^0_{\mathrm{BB}}\lambda^2_{\mathrm{AE}}+\lambda^1_{\mathrm{BB}}\lambda^1_{\mathrm{AE}}) + 2k_5(\lambda^2_{\mathrm{BB}}\lambda^0_{\mathrm{AF}}+\\ &\quad \lambda^0_{\mathrm{BB}}\lambda^2_{\mathrm{AF}}+2\lambda^1_{\mathrm{BB}}\lambda^1_{\mathrm{AF}}) + k_4(\lambda^0_{\mathrm{AB}}\lambda^2_{\mathrm{BE}}+2\lambda^1_{\mathrm{AB}}\lambda^1_{\mathrm{BE}}) + k_5(\lambda^0_{\mathrm{AB}}\lambda^2_{\mathrm{BF}}+2\lambda^1_{\mathrm{AB}}\lambda^1_{\mathrm{BF}}).\end{aligned}$$

3. Equations for end groups, epichlorohydrin, and NaOH are as follows:

$$\begin{aligned}\frac{d[\mathrm{A}]}{dt} &= -k_1[\mathrm{NaOH}][\mathrm{A}],\\ \frac{d[\mathrm{B}]}{dt} &= -k_1[\mathrm{NaOH}][\mathrm{A}] - k_2[\mathrm{EP}][\mathrm{B}] - k_4[\mathrm{E}][\mathrm{B}] - k_5[\mathrm{F}][\mathrm{B}],\\ \frac{d[\mathrm{E}]}{dt} &= -k_1[\mathrm{NaOH}][\mathrm{A}] - k_4[\mathrm{E}][\mathrm{B}],\\ \frac{d[\mathrm{F}]}{dt} &= -k_2[\mathrm{EP}][\mathrm{B}] - k_5[\mathrm{F}][\mathrm{B}] - k_3[\mathrm{F}][\mathrm{NaOH}],\\ \frac{d[\mathrm{EP}]}{dt} &= -k_2[\mathrm{EP}][\mathrm{B}],\\ \frac{d[\mathrm{NaOH}]}{dt} &= -k_1[\mathrm{NaOH}][\mathrm{A}] + k_2[\mathrm{EP}][\mathrm{B}] - k_3[\mathrm{F}][\mathrm{NaOH}] - k_4[\mathrm{E}][\mathrm{B}].\end{aligned}$$

All 48 equations describing the complete mathematical model can be found elsewhere (Majumdar et al., 2003; Raha and Gupta, 1998).