



01 Nov 2004

Optimization Of Batch Reactions In Series With Uncertainty

Keith G. Tomazi

Missouri University of Science and Technology, tomazik@mst.edu

Follow this and additional works at: https://scholarsmine.mst.edu/che_bioeng_facwork



Part of the [Biochemical and Biomolecular Engineering Commons](#)

Recommended Citation

K. G. Tomazi, "Optimization Of Batch Reactions In Series With Uncertainty," *Annals of Operations Research*, vol. 132, no. 1 thru 4, pp. 189 - 206, Springer, Nov 2004.

The definitive version is available at <https://doi.org/10.1023/B:ANOR.0000045282.19338.8e>

This Article - Conference proceedings is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Chemical and Biochemical Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.



Optimization of Batch Reactions in Series with Uncertainty

KEITH G. TOMAZI

keith.tomazi@tycohealthcare.com

*Tyco/Healthcare/Mallinckrodt, 3600 North Second Street, Building 96, PO Box 470158, St. Louis,
MO 63147, USA*

Abstract. Certain types of chemical reactions, such as the global deprotection of a polypeptide, are extremely complex. As a result, it may be very difficult or expensive to develop accurate models of these chemical reactions. Without a satisfactory kinetic model for the reaction, it is difficult to develop an optimum operating policy that will maximize the profit. Stochastic optimization is applied in this work to an example process step to obtain the optimum reaction temperature and reaction time. In the case of the “here and now” problem, the optimal conditions are a lower reaction temperature and a longer reaction time than obtained from the deterministic problem. The average reaction time for the “wait and see” problem is also longer than the deterministic case, but the average reaction temperature is very close to that of the deterministic problem. Both normally and uniformly – distributed uncertain parameters are considered.

Keywords: chemical reaction kinetics, nonlinear programming, optimization with uncertainty, peptides

1. Introduction

One important class of drug products is peptides. Peptides are biochemicals that consist of a sequence of two or more amino acids, up to a maximum molecular weight of 10,000. One example of a natural polypeptide is human insulin (Budavari, 1996). Other important peptides are synthetic molecules that may be analogues of natural biochemicals. Numerous drugs are available, or are under development for the treatment of a wide variety of illnesses, including cancer, pain management, and endocrine disorders. These products tend to be very difficult or expensive to isolate from biomass or synthesize from amino acids, and as a result, tend to be very valuable and profitable products. The total worldwide sales of peptide active pharmaceutical ingredients (API's) was \$350 million in 1999 (Watkins, 2001).

The synthesis of peptides by the Merrifield (1963) method is done in a multi-step process. First, the C-terminal amino acid is attached to a solid-phase support with a suitable linker. Synthesis of the molecule proceeds for each amino acid in the sequence using protected amino acids. All of the amino acids are protected at the amino group in order to improve the selectivity of the reaction. As a result, the activated carboxylic acid of the amino acid to be coupled can only react with the free amine on the intermediate peptide. Otherwise, polymerization of the amino acid may occur, resulting in multiple instances of that amino acid in the peptide, resulting in a multitude of unde-

sirable molecules, and extremely low yields of the desired product. The N-terminus of the intermediate peptide is deprotected, resulting in a free amine. The carboxylic acid of the next amino acid to be added is activated, and then coupled to the free amine of the intermediate peptide. The cycle of deprotection and coupling continues until the entire amino acid sequence has been synthesized. After the synthesis is complete, the crude peptide is cleaved from the support, globally deprotected, purified, and finally isolated.

A simple amino acid, such as Glycine (Gly) ($\text{H}_2\text{N}-\text{CH}_2-\text{COOH}$), has one primary amine ($-\text{NH}_2$) group, and one carboxylic acid (COOH) group. Simple amino acids such as Glycine are only protected at one place, the amine ($-\text{NH}_2$) group. However, other amino acids may contain additional potentially reactive groups that must also be protected from undesirable reactions. One example of a more complicated amino acid is Lysine (Lys) ($\text{H}_2\text{N}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}(\text{NH}_2)-\text{COOH}$). Lysine has two amine groups. Both amine groups must be protected during coupling. However, one amine group must remain protected throughout the remainder of the synthesis, otherwise, a second peptide chain may form at that location. As a result, an additional chemical reaction must follow the synthesis of the amino acid sequence and the cleavage from the support. This reaction causes the removal of the additional protecting groups, and is known as the global deprotection reaction. Since the additional protecting groups must withstand the relatively mild deprotection reactions, and the somewhat harsher cleavage reaction, the global deprotection reaction occurs in harsh conditions. However, under harsh conditions, the deprotected peptide is unstable. The selection of reaction conditions is extremely important in order to obtain product with a high yield and acceptable purity.

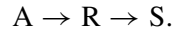
The global deprotection step can be very complex as numerous amino acids with additional protecting groups may be present in a polypeptide. An example of such a reaction with commercial significance reveals that the yield and purity are very sensitive to the process conditions (reaction time and reaction temperature). As a result, there is a strong economic incentive to obtain optimal reaction conditions. However, there are uncertainties in the kinetic model due to the complexity of the reaction.

Padmanabhan and Pitchumani (1999a, 1999b) applied stochastic analysis and modeling to the curing of thermosetting resin systems, and to Resin Transfer Molding. In light of these works, the purpose of this paper is to apply stochastic optimization to the global deprotection step of a polypeptide to obtain optimal process conditions. Two objectives are considered: (1) maximizing the profit, and (2) minimizing the variance of the profit.

This paper is organized into the following sections. (1) introduction of the problem; (2) presentation of the kinetic model; (3) sources of uncertainty into the problem; (4) the formulation of the deterministic problem and the profit function; (5) the “here and now” problem; and (6) the “wait and see” problem. Results of all cases are summarized in section 7.

2. Kinetic model

The global deprotection reaction may be represented by the following chemical steps:



Here, A represents the protected peptide, R represents the desired product, and S represents decomposed peptide.

Chemical reaction kinetics is a subject unto itself (cf. (Levenspiel, 1972)). Although many elementary reactions follow the stoichiometry, many reactions involving multiple intermediate steps (some of which may not be easily observed) exhibit more complicated rate expressions, or rate expressions that do not directly follow from the overall stoichiometry. The typical form of a kinetic equation may be represented in the following form in terms of species A , with multiple species participating in the reaction:

$$r_A = \pm \frac{dC_A}{dt} = k_A C_A^n C_B^m \dots \quad (1)$$

Here, the sign is positive if the rate expression is written in terms of the appearance of specie A , and negative if written in terms of the disappearance of specie A . The exponents represent the order of the reaction: the reaction is of order n in species A , and order m in species B . Determination of a rate expression typically involves observing the concentration (typically expressed in moles of specie per volume of solution) of the various reactant and product species over time, then fitting these observations to a model. These experiments may be repeated at several different temperatures in order to obtain the temperature-dependence of the reaction. However, the rate expression that follows the above stoichiometry may be written as:

$$\frac{dC_A}{dt} = -k_A C_A, \quad (2)$$

$$\frac{dC_R}{dt} = k_A C_A - k_R C_R, \quad (3)$$

$$\frac{dC_S}{dt} = k_R C_R. \quad (4)$$

The differential equations may be integrated (cf. (Levenspiel, 1972)) to obtain analytical expressions for the concentrations of the three species for any given time:

$$C_A(t) = C_A^0 e^{-k_A t}, \quad (5)$$

$$C_R(t) = C_A^0 k_A \left[\frac{e^{-k_A t}}{k_R - k_A} + \frac{e^{-k_R t}}{k_A - k_R} \right], \quad (6)$$

$$C_S(t) = C_A^0 k_A \left[1 + \left(\frac{k_R}{k_A - k_R} \right) e^{-k_A t} + \left(\frac{k_A}{k_R - k_A} \right) e^{-k_R t} \right]. \quad (7)$$

The rate coefficients, k_A and k_R are typically functions of absolute temperature of the form $k = k_0 e^{-E/RT}$, which tells us that reactions tend to occur faster with increasing temperature. If the activation energy (E) is not equal in the two rate expressions,

then it may be possible to improve the selectivity of the overall reaction by adjusting the temperature. One may find that lower temperatures reduce the rate of undesirable decomposition more than the rate of the desirable deprotection reaction.

The limited observations that were possible while conducting this process step revealed that at room temperature (298 K) the deprotection reaction was typically complete in approximately 30 minutes. However, in one case, the reaction ran for an extraordinarily long time (24 hours). Very low yield of the product was obtained, approximately ten percent of theory. These observations suggest rough approximations of k_A and k_R of 10/hour and 0.1/hour, respectively, at 298 K. Other observations indicate that the degree of selectivity becomes progressively unfavorable as the reaction temperature increases. For this example, it is assumed that rate of decomposition relative to the rate of production doubles if the temperature is increased 30 degrees. That is, the rate of synthesis doubles, but the rate of decomposition quadruples. These observations and assumptions allow estimates to be made of the four kinetic parameters and provide an approximation to the reaction kinetics. Based on this model a plot of the concentrations of the different reactant species over time is presented in figure 1 for a reaction temperature of 298 K. The maximum concentration of the desired product is 9.545 moles per unit volume, and occurs at 0.465 hours.

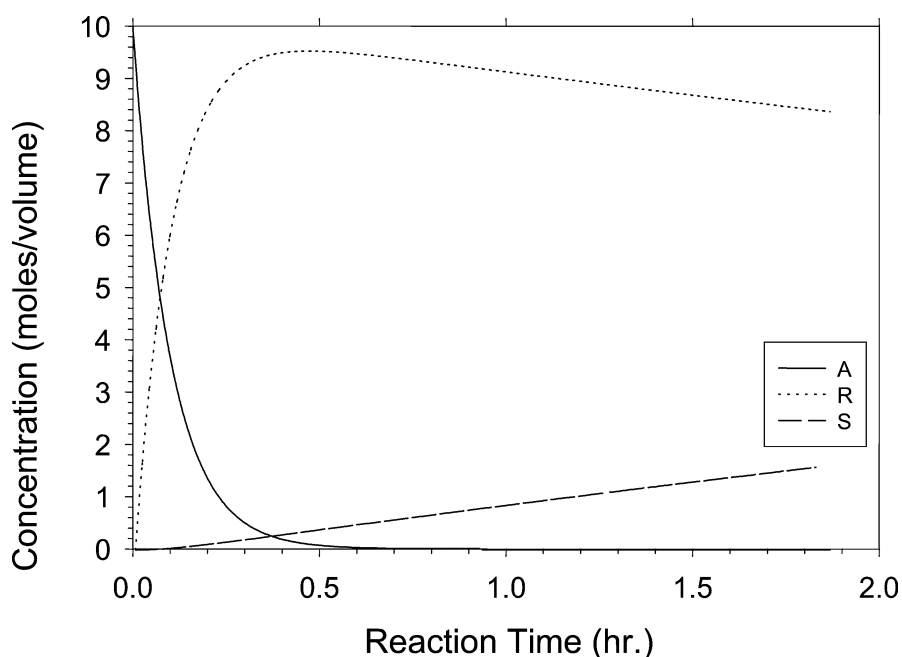


Figure 1. Concentrations of the raw material (A), desired product (R), and decomposition product (S) for the global deprotection and decomposition reactions (298 K).

3. Uncertainty

As stated in the previous section, the kinetic model and parameters may be obtained experimentally. However, this may be difficult to do if the reactions involve large numbers of intermediates, if sampling and sample preparation are dangerous, difficult or lengthy compared to the rate of the reaction, if identification of individual intermediate species is difficult, or if the cost of raw materials is prohibitive. All of these problems may be true of synthetic biochemicals.

One possible way to address this problem is to obtain an approximate model, based upon an understanding of the chemical mechanism for deprotection and decomposition. Once a simple model has been obtained, one could then apply a limited number of observations, and obtain a distribution of kinetic parameters. It may be very difficult to model the chemistry of the global deprotection reaction. Large peptides (containing approximately 40 amino acids) are fairly commonplace. These large peptides typically contain numerous amino acids with additional protecting groups. It may be extremely expensive or difficult to study the kinetics of each possible reaction in these cases (and to isolate and characterize each possible intermediate specie). As a result, there is uncertainty in the kinetic model.

In addition to the uncertainty in the process kinetics, other sources of uncertainty may be in the strength of the chemicals employed for the reactions and in the size of the batch. The raw material used for the global deprotection reaction is a mixture of several different materials that promote the desired reaction while suppressing undesirable reactions. Variations in the quantities of the components in the mixture, and the properties of the individual raw materials may lead to some variations in the potency of the reactants. As a result, the selectivity and the rate of the global deprotection reaction may vary between batches. The batch size is an important variable, as when scaling up a process from the bench-top laboratory investigation to the full-scale commercial process, certain changes in the process physics may occur that can have a significant effect upon the yield, quality, and profitability of the product. Two common physical principles encountered on scaling up chemical processes are the degree of homogeneity, and the surface area available for heating or cooling a batch. The first principle, homogeneity, is related to the uniformity of mixing the reactants together. It is typically much easier to achieve uniformity of mixing in smaller scales than in larger scales. Non-homogeneity may result in localized volumes of excessively high or low concentrations of one of the reactants, or excessively high temperatures. These factors may decrease the overall yield or quality of the product. The second principle, the area available for heating or cooling, arises from the typical design of cylindrical process vessels. These vessels are normally heated or cooled through the vessel walls, but heat energy associated with the chemical reactions is generated throughout the volume of the batch. Since the ratio of the surface area of the vessel to the volume of the vessel decreases with increasing vessel size, control of the batch temperature may become more difficult in larger vessels, particularly for highly exothermic reactions such as the one presented in this work. Insufficient temperature control may also lead to reduced yield and product quality.

Table 1
Uncertain kinetic parameters.

Parameter	Normally distributed		Uniformly distributed	
	Mean	Std. deviation	Low	High
$k_{0,A}$	19531.2	1953.12	15624.96	23437.44
$k_{0,R}$	382700	38270	306160	459240
E_A/R	2258	225.8	1806.4	2709.6
E_R/R	4517	451.7	3613.6	5420.4

As a first step towards uncertainty analysis, we concentrated on the kinetic model. The uncertain variables in this problem are the kinetic parameters, $k_{0,A}$, $k_{0,R}$, E_A , E_R . Two different distributions of these variables were used: normal distributions and uniform distributions. The data are summarized in table 1. The mean values for these parameters are based upon the observation at 298 K, and the assumption of the rate increase with temperature.

4. Formulation of the deterministic optimization problem

Several possible objective functions may be used for this type of problem. One may, for example, simply maximize the yield of the desired product R. However, this objective does not consider the costs required to manufacture the product. One may also minimize the manufacturing cost per unit of product. However, this objective fails to account for the value of the product (in this case, the product is very valuable). A more sophisticated objective function may be to maximize the profit, defined as the difference in the value of the product and the cost of raw materials and labor. If impurities are difficult to resolve from the desired product, one could include the cost of purifying the product. This can be a significant addition to the cost of manufacturing the product, because certain impurities are molecularly very similar to the desired product. Removal of these impurities is so difficult that a significant amount of desired product may be lost during purification. As a result, it is important to include the purification costs in the objective function because downstream purification costs have a significant effect upon the profitability of the final product, and these impurities are generated (and can be controlled) during the global deprotection step.

The profit function of a discrete step in the manufacturing of a pharmaceutical product may be a highly complex formula that includes the fixed costs for a facility, the variable costs, and the cost of removing impurities that may be generated during that step. Fixed costs include the depreciation for the building and equipment, the utilities, and other overhead costs. These costs may be allocated over the time the facility is operational, and therefore may be expressed in an hourly basis. Variable costs include the cost of raw materials, and possibly labor (the labor cost is variable if the step represents a process bottleneck, as is the case in this example). The cost of impurity removal may be used to account for the anticipated loss in yield that will occur when removing impurities that are very difficult to resolve from the desired product. A heuristic profit function may

be written as the difference in the value of the product produced, and the costs of raw material, labor, overhead costs, and purification. The profit per batch may be written as

$$P = V_R v C_R - V_A v C_A^0 - (V_L + V_B)t - \sum p_i v C_i. \quad (8)$$

In the above expression, the reactor working volume, v , is multiplied by the concentration of each specie (expressed in moles per unit volume) to obtain the number of moles of each specie. The value of the product, and cost of the raw material, are obtained by multiplying the number of moles by the value or cost per mole. The last term in the profit function accounts for the anticipated change in yield during downstream purification (to remove impurities i that may result from the reaction at this step in the process). In this problem, there are two possible impurities: unreacted A, and the decomposition product S. Species A is very dissimilar from the desired product, and may typically be removed without significant loss of the desired product R. However, species S is chemically very similar to R. As a result, although the purification step can remove S, a considerable amount of R is lost per unit of S removed. Consider a hypothetical case of unit volume, in which the value of the product is \$100 per mole, the value of the raw material is \$10 per mole, and 0.75 moles of product is lost per mole of decomposition product removed. Multiplication of 0.75 moles of product per mole of decomposition product removed, by the value of the product (\$100 per mole) results in a cost of purification of \$75 per mole of decomposition product. The operation requires the part-time supervision of one operator at a cost of \$20 per hour, and the department burden rate is \$30 per hour. The combined labor and burden rate is therefore \$50 per hour. Substitution of numerical values into the profit function for this example case results in the following expression:

$$P = 100C_R - 10C_A^0 - 50t - 75C_S. \quad (9)$$

Reactor design problems may also be subject to certain constraints. For example, since this reaction occurs in a solution, temperature constraints include the freezing point and boiling point of the solution. There may also be additional temperature constraints based upon the mechanical integrity of the reaction vessel, and the operating limits of the process utilities used to regulate the temperature of the reactor. There may also be constraints upon the concentrations of the reactant and products imposed by the solubility limits. However, in this example case, none of the temperatures obtained were infeasible. In addition, the process chemistry of this reaction does not encounter solubility limits. As a result, none of the constraints were active.

For the deterministic case, the problem may be written in standard form as

$$\text{Maximize } P(\mathbf{x}) \quad (10)$$

subject to:

- temperature range:

$$\begin{aligned} T &> T_{\min}, \\ T &< T_{\max}; \end{aligned} \quad (11)$$

- and the kinetic model:

$$C_A(t) = C_A^0 e^{-k_A t}, \quad (12)$$

$$C_R(t) = C_A^0 k_A \left[\frac{e^{-k_A t}}{k_R - k_A} + \frac{e^{-k_R t}}{k_A - k_R} \right], \quad (13)$$

$$C_S(t) = C_A^0 k_A \left[1 + \left(\frac{k_R}{k_A - k_R} \right) e^{-k_A t} + \left(\frac{k_A}{k_R - k_A} \right) e^{-k_R t} \right], \quad (14)$$

$$k_A = k_{0,A} e^{-E_A/RT}, \quad (15)$$

$$k_R = k_{0,R} e^{-E_R/RT}. \quad (16)$$

Given the number of constraints and number of decision variables, the degree of freedom for this optimization problem is found to be two. The two operating parameters selected for this study are (1) reaction time, (2) operating temperature. Surface and contour plots for the case of an initial concentration of A of 10 moles per volume are presented below. See figure 2. These plots indicate that the profit becomes relatively insensitive to the decision variables within a region around the maximum profit of \$817.0, which occurs at a reaction time of 0.6415 hours, and a reaction temperature of 280.42 K.

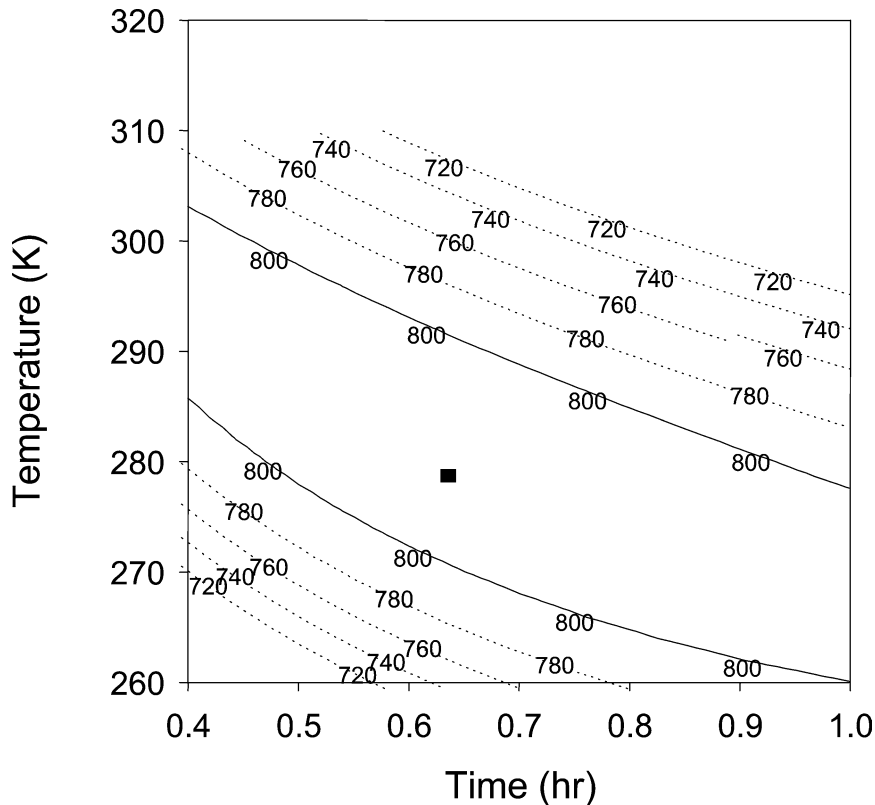


Figure 2. Contour plot of the profit function (deterministic case). The point of maximum profit is indicated.

As a result, an operator or designer will find that the profit does not decrease by more than a few percent as long as the process operates within this region. An operator may therefore compensate for high or low process temperature by adjusting the amount of reaction time.

The optimum set of decision variables (reaction time and reaction temperature) were obtained using Newton's method with a BFGS updating scheme (Press et al., 1999). The optimum set of decision variables is a reaction time of 0.6415 hours, and a reaction temperature of 280.42 K (see figure 2). The maximum profit is \$817.0 per batch. The average values of the uncertain parameters were used in the kinetic model. See table 1. The solution was tested by an endgame strategy, in which random points were selected adjacent to the optimum. All of these random points led to less than the optimal profit. In addition, the problem was restarted with different initial values, and convergence always led to the same set of decision variables.

5. Optimization with uncertainty (here and now case)

An operator of a peptide manufacturing plant needs to know the operating conditions for the reaction. These decisions must sometimes be made with limited available information about the uncertainty of the kinetic parameters of the reaction. In other words, the decision must be made with the information that is available "here and now." As a result, an operator may operate the reactor in such a way as to maximize the average profit of the reaction over some number of batches. Or, the operator may choose to operate the reactor in such a way as to minimize the one-sided variance of the profit. For the "here and now" cases the problem may be written in the standard forms as:

$$\text{maximize } E[P(\mathbf{x}, \mathbf{u})] \quad (17)$$

or

$$\text{minimize } \text{Var}[P(\mathbf{x}, \mathbf{u})] \quad (18)$$

subject to:

$$\text{Probability}(T > T_{\min}) = 1, \quad (19)$$

$$\text{Probability}(T < T_{\max}) = 1, \quad (20)$$

$$C_A(t) = C_A^0 e^{-k_A t}, \quad (21)$$

$$C_R(t) = C_A^0 k_A \left[\frac{e^{-k_A t}}{k_R - k_A} + \frac{e^{-k_R t}}{k_A - k_R} \right], \quad (22)$$

$$C_S(t) = C_A^0 k_A \left[1 + \left(\frac{k_R}{k_A - k_R} \right) e^{-k_A t} + \left(\frac{k_A}{k_R - k_A} \right) e^{-k_R t} \right], \quad (23)$$

$$k_A = k_{0,A} e^{-E_A/RT}, \quad (24)$$

$$k_R = k_{0,R} e^{-E_R/RT}. \quad (25)$$

Table 2
Results for the “here and now case” with normally-distributed uncertain parameters.

Objective	Time (h)	Temperature (K)	Profit
Maximum average profit	0.7719	273.94	\$700.8
Minimum variance	0.9388	265.50	\$697.6

Table 3
Results for the “here and now case” with uniformly-distributed uncertain parameters.

Objective	Time (h)	Temperature (K)	Profit
Maximum average profit	0.8324	271.34	\$667.1
Minimum variance	1.0822	264.21	\$665.0

In the above representation, \mathbf{u} represents a vector of uncertain parameters (the kinetic parameters). In this problem, the constraints must be obeyed at all times and not just on the average or with probability less than 1.

Here we used the generalized representation of the “here and now” problem presented by Diwekar (2003). This procedure involves two loops (1) the optimizer loop and (2) the sampling loop. Sampled data (200 sets each) for the two distributions were obtained using Hammersley Sequence Sampling¹ (Diwekar and Kalagnanam, 1997). The computational method for the first case was to find the set of decision variables (time and temperature) that obtained the maximum profit averaged over all sets of sampled kinetic parameters. In other words, the optimizer selected a set of decision variables, and the model returned the average profit for all sets of sampled data. The optimizer continued to select sets of decision variables until the problem converged. The computational method for the second case was similar to the first, except that the objective was to minimize the one-sided variance in profit around a specified mean. In this case, the “mean” selected was \$900, which was greater than the profit obtained for any cases in the “wait and see” problem. See section 6. It is not suitable to simply minimize the variance of the profit, since under unfavorable conditions in which the profit becomes very small for all sets of kinetic parameters, the variance approaches zero. By minimizing the one-sided variance below a high profit, the objective function is penalized for low profits. As a result, the optimizer converges to a point that has the most consistently high profits for all of the sets of sampled data.

Results of the computations are presented in table 2 for normally-distributed uncertain parameters, and in table 3 for uniformly-distributed uncertain parameters.

6. Optimization with uncertainty (wait and see case)

If an operator has sufficient knowledge of the kinetic parameters prior to each batch, that is, a complete understanding of the potency of the raw materials, and of the kinetic model, the operator may adjust the reactor conditions to maximize the profit of each

Table 4
Results for the “wait and see” case.

Kinetic parameters	Time (h)	Temperature (K)	Profit
Normally distributed	0.7230 (avg.)	281.55 (avg.)	\$778.7 (avg.)
Uniformly distributed	0.7378 (avg.)	282.08 (avg.)	\$765.7 (avg.)

batch. For example, a thorough investigation into the reaction may reveal that the kinetic parameters of the reaction are most strongly influenced by the properties of one of the raw materials. This may lead to the development of an analytical test to quantify this property that may be performed on each lot of raw material prior to use. (However, no action is taken to eliminate the source of variability in the raw material). An operator may then take these results, and obtain computed values of the optimum reaction time and temperature for the batch. The formulation of this problem is to obtain the maximum profit for each set of sampled kinetic parameters. There are a total of n sets of sampled parameters.

$$\text{Maximize } P_i(\mathbf{x}_i, \mathbf{u}_i), \quad i = 1, \dots, n, \quad (26)$$

subject to:

$$T_i > T_{\min}, \quad (27)$$

$$T_i < T_{\max}, \quad (28)$$

$$C_A(t_i) = C_A^0 e^{-k_{A,i} t_i}, \quad (29)$$

$$C_R(t_i) = C_A^0 k_{A,i} \left[\frac{e^{-k_{A,i} t_i}}{k_{R,i} - k_{A,i}} + \frac{e^{-k_{R,i} t_i}}{k_{A,i} - k_{R,i}} \right], \quad (30)$$

$$C_S(t_i) = C_A^0 k_{A,i} \left[1 + \left(\frac{k_{R,i}}{k_{A,i} - k_{R,i}} \right) e^{-k_{A,i} t_i} + \left(\frac{k_{A,i}}{k_{R,i} - k_{A,i}} \right) e^{-k_{R,i} t_i} \right], \quad (31)$$

$$k_{A,i} = k_{0,A,i} e^{-E_{A,i}/RT_i}, \quad (32)$$

$$k_{R,i} = k_{0,R,i} e^{-E_{R,i}/RT_i}.$$

In the above formulation, n is the number of samples.

The average values of the decision variables and profits for each distribution of kinetic parameters are presented in table 4. These results tell us that an increase in profit of the reaction with a priori knowledge of the reaction kinetics (the “wait and see” case) may be obtained over the case of limited knowledge (the “here and now” case). The increase in profit, or Expected Value of Perfect Information (EVPI) is \$77.9 for normally-distributed uncertain parameters, or \$98.6 for uniformly-distributed uncertain parameters.

Graphs of the distribution of decision variables (reaction time and temperature) and profits are presented in figures 3, 4, and 5, respectively. The distributions of the reaction times appear essentially log-normally distributed for normally-distributed uncertain parameters, and somewhat log-uniformly distributed (or log-normally distributed with a broader distribution) for uniformly-distributed uncertain parameters. The optimal

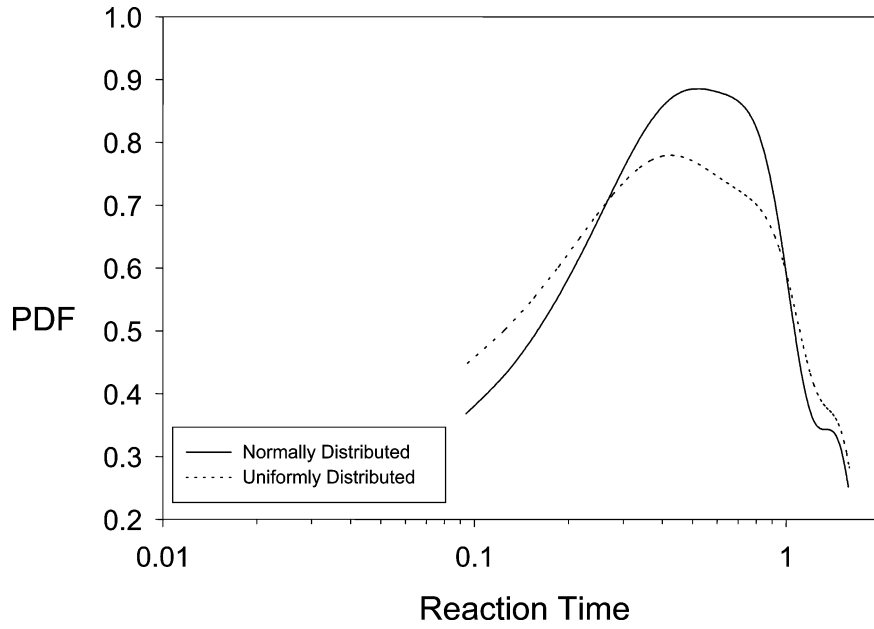


Figure 3. PDF of optimal reaction times for the “wait and see” problem, for both normally and uniformly distributed uncertain parameters.

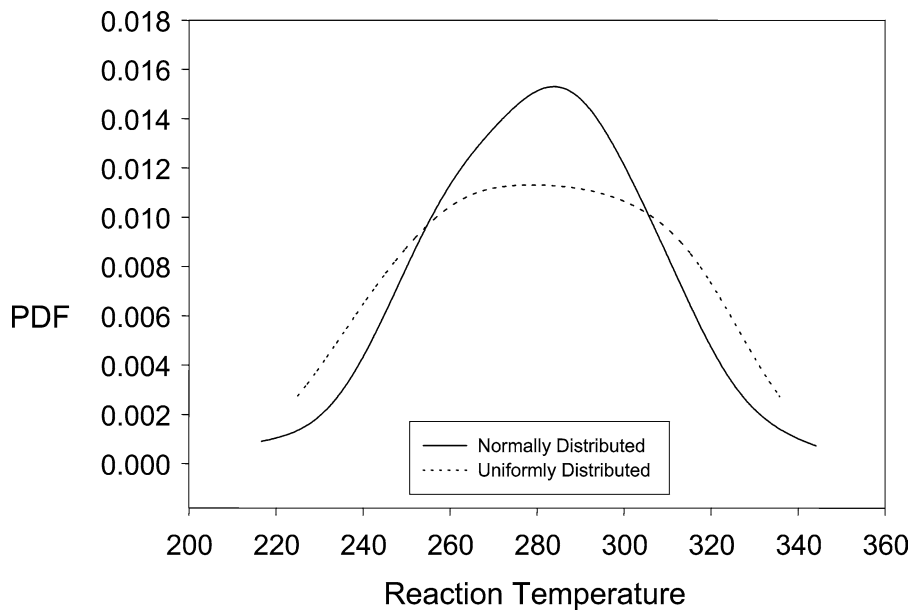


Figure 4. PDF of optimal reaction temperatures for the “wait and see” problem, for both normally and uniformly distributed uncertain parameters.

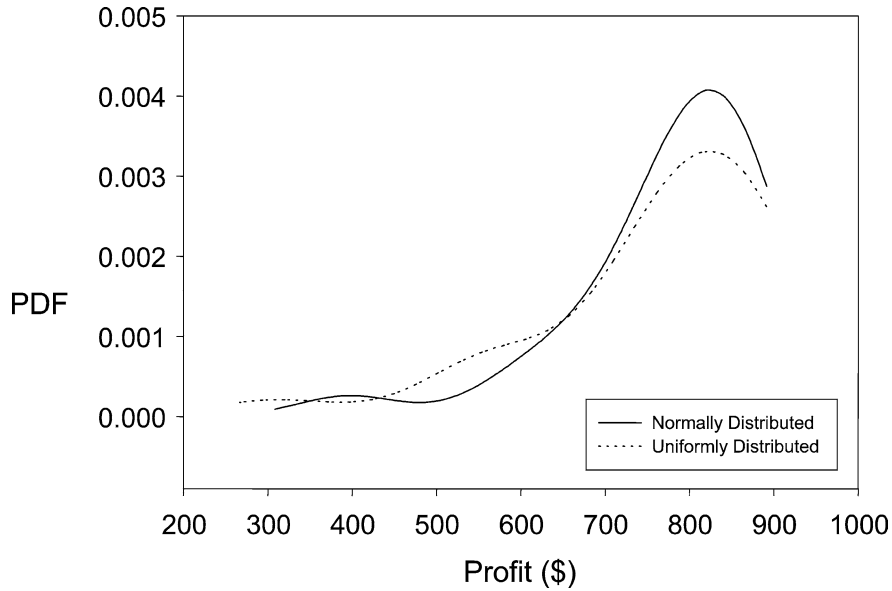


Figure 5. PDF of optimum profits for the “wait and see” problem, for both normally and uniformly distributed uncertain parameters.

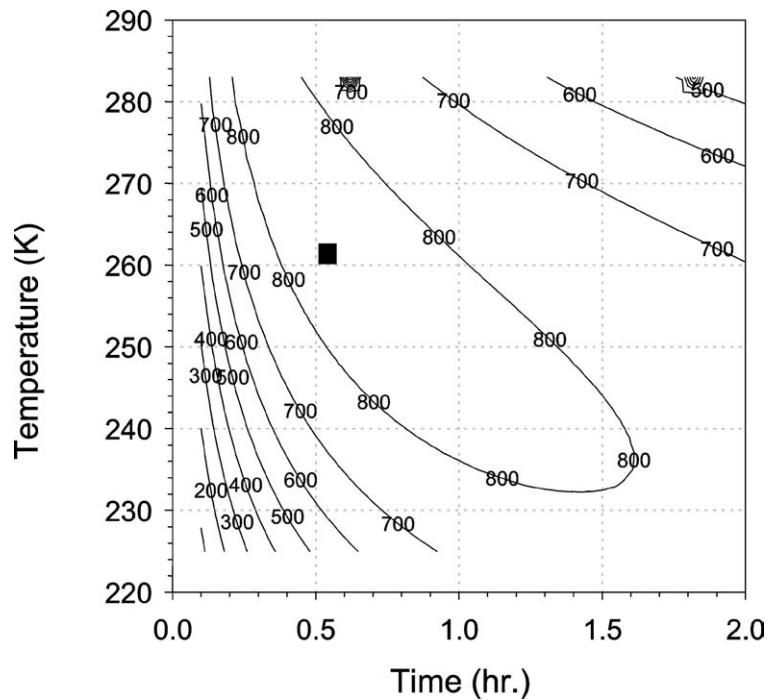


Figure 6. Contour plot of the profit function of a high-probability case of the “wait and see” problem. The point of maximum profit is indicated.

reaction temperatures appear normally distributed for normally-distributed uncertain parameters, and very broadly normally-distributed for uniformly-distributed uncertain parameters.

The profits appear to be somewhat log-normal in shape, with the profits of the uniformly-distributed uncertain parameters somewhat flatter and broader than the profits of the normally-distributed parameters. There also appear to be some outliers of low profit. These may be effects of the nonlinearity of the model. The rate coefficients are exponential in temperature and activation energy, and the concentrations of the species are exponential in time and in the rate coefficients. In addition, the rate coefficients appear in the numerator and denominator in the concentration equations. Contour plots of the profit for the “wait and see” problem with normally-distributed parameters are presented for a high-probability example, and for an outlier example. See figures 6 and 7. These plots suggest that the sensitivity of the profit to time or temperature may shift, depending upon the kinetic parameters, since the slope of the “ridge” is closer to horizontal in figure 7 than in figure 6. In other words, the profit in figure 6 is insensitive over a temperature range of 30 degrees, and a time of 0.6 hours; but the profit in figure 7

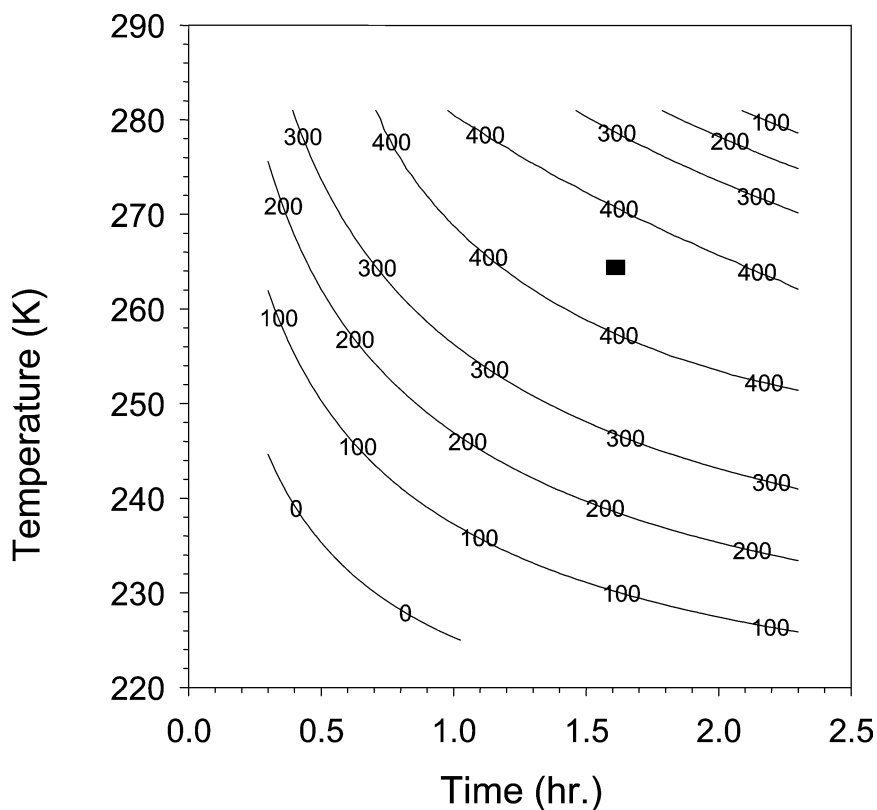


Figure 7. Contour plot of the profit function of an outlier case of the “wait and see” problem. The point of maximum profit is indicated.

is insensitive over a narrower range in temperature (15 degrees) but a wider range in time (1 hour).

7. Results summary

The following formulations of the problem were considered:

1. The deterministic problem, in which there are no uncertainties considered. The average values of the kinetic parameters were used. See table 1.
2. The “here and now” problem, in which there is uncertainty in the kinetic parameters. The optimizer finds the set of decision variables that maximize the average profit for all sets of sampled data.
3. The “wait and see” problem, in which there is variability of the kinetic parameters, but the true value of these parameters is known a priori. A separate optimization exercise is performed for all sets of sampled data, and the average values of the profit and decision variables are obtained.
4. The “minimum variance” problem, which is similar to the “here and now” problem, except that the optimizer finds the set of decision variables that minimizes the one-sided variance of the profit around a specified mean. In this case, the “mean” value selected was \$900, which is greater than the maximum profit obtained in any of the cases in the “wait and see” problem.
5. The “sampled” profit, in which the decision variables from the deterministic case are used for all sets of sampled data. This case may represent decision variables obtained from a perfect knowledge of the kinetic model, but uncontrolled variability of the process.

The results of the optimization problems are presented in table 5 for normally-distributed uncertain parameters, and in table 6 for uniformly-distributed uncertain parameters.

These results tell us that a plant operator may expect to obtain a profit of \$817.0 per batch in the deterministic case. Given that uncertainties exists, the actual expected value of this situation is \$669.5 for normally distributed uncertain parameters, or \$664.2 for uniformly distributed uncertain parameters. If the reaction kinetics are understood,

Table 5
Results for normally-distributed uncertain parameters.

Case	Time (h)	Temperature (K)	Profit/Batch
Deterministic/w. uncert.	0.6415	280.42	\$817.0/\$699.5
Wait and see	0.7230 (avg.)	281.55 (avg.)	\$778.7 (avg.)
Here and now	0.7719	273.94	\$700.8
Minimum variance	0.9388	265.50	\$697.6

Note. VSS = \$1.3 per batch, EVPI = \$77.9 per batch.

Table 6
Results for uniformly-distributed uncertain parameters.

Case	Time (h)	Temperature (K)	Profit
Deterministic/w. uncert.	0.6415	280.42	\$817.0/\$664.2
Wait and see	0.7378 (avg.)	282.08 (avg.)	\$765.7 (avg.)
Here and now	0.8324	271.34	\$667.1
Minimum variance	1.0822	264.21	\$665.0

Note. VSS = \$2.9 per batch, EVPI = \$98.6 per batch.

and all process conditions are known a priori (the “wait and see” case), then the profit may be \$778.7 for normally-distributed uncertain parameters, or \$765.7 for uniformly-distributed parameters. If the reaction kinetics are only understood in terms of their probability distributions (the “here and now” case), the profit may be \$700.8 for normally distributed uncertain parameters, or \$667.1 for uniformly-distributed uncertain parameters. There is little difference in profit if the “here and now” problem is changed to minimize the profit variance, rather than to maximize the average profit. This would lead to a reduction in profit of \$3.2 per batch, or \$2.1 per batch, if the uncertainties were distributed normally or uniformly, respectively. There is still a slight improvement in profit if the stochastic solution to the “here and now” problem is adopted instead of using the reaction conditions obtained from the deterministic solution, but without reducing the process uncertainty (this is the sampled case). The value of the stochastic solution (VSS) is \$1.3 for normally-distributed uncertain parameters, and \$2.9 for uniformly-distributed uncertain parameters. In other words, if the effort and expense is made to obtain a good understanding of the reaction kinetics, but the variability of the process is not improved, then the profit will still be less than if a stochastic solution is employed. It is possible that the small difference in profit for these last three cases may be due to the relatively flat objective surface adjacent to the true maximum.

These results also tell us that if the problem is uncertain (excluding the “wait and see” case), the set of decision variables obtained favor lower temperatures and longer reaction times. This is particularly true if an operator desires to minimize the variance of the profit. The lower temperatures produce a more favorable selectivity in the reactions (the ratio of the rate coefficients of decomposition, to production of the desired product, decreases as the temperature is reduced). The improvement in selectivity should reduce the purification costs downstream of the global deprotection step. However, the improvement in reaction selectivity requires longer reaction time, which increases the labor and overhead costs.

8. Conclusions

This paper presented model of a complex reaction for obtaining a polypeptide as a product. Uncertainties due to model approximations are characterized from the experimental experience. A profit function based on the difference between the value of the product and the manufacturing and downstream purification costs is formulated to obtain opti-

mal operating conditions. Results and analysis is presented for deterministic as well as stochastic optimization problems. The here and now problem was analyzed using two different objective functions and robust solutions are presented. In wait and see solutions, the outliers of the output distributions are attributed to nonlinearities and changing sensitivity of the objective function with respect to the decision variables.

9. Notations

A	Raw material.
C_i	Concentration of species i , moles/volume.
E	Activation energy of a reaction, energy/mole.
k_i	Rate coefficient of reaction i , 1/time.
k_0	Rate constant, 1/time.
m, n	Order of reaction with respect to a certain species, dimensionless.
P	Profit, units of currency.
p_i	Cost of removing impurity i , units of currency per mole.
R	Desired product.
R	Gas constant, energy/(mole · K).
r_i	rate of appearance or disappearance of species i , moles/(volume · time).
S	Undesirable decomposition product.
t	Reaction time.
T	Reaction temperature, K.
\mathbf{u}	Vector of uncertain variables.
V_A	Value of raw material, units of currency per mole.
V_R	Value of product, units of currency per mole.
V_L	Cost of labor, units of currency per time.
V_B	Cost of burden, units of currency per time.
v	Batch volume.
\mathbf{x}	Vector of decision variables.

Acknowledgments

The author greatly appreciates the support of Professor Urmila Diwekar, who provided the author with a copy of her forthcoming book on optimization, for supplying the source code for the HSS sampling program and for her thoughtful comments and suggestions. The author also greatly appreciates her and Dr. Kemal Sahin for providing source code for the probability density function. The author appreciates all of the fine work that was done by his company's Peptide Development Center in identifying the decomposition reaction of the peptide. The author also appreciates the thoughtful suggestions of Dr. Mel Bagley.

Note

1. Source code for this sampling technique was provided by Diwekar (2003).

References

- Budavari, S. (1996). *The Merck Index*, 12th ed. Whitehouse Station: Merck & Co.
- Diwekar, U.M. (1997). "Code for Hammersley Sequence Sampling." Personal communication.
- Diwekar, U.M. (2003). *Introduction to Applied Optimization*. Norwell, MA: Kluwer Academic Publishers.
- Diwekar, U.M. and J. Kalagnanam. (1997). "An Efficient Sampling Technique for Optimization under Uncertainty." *AIChE Journal* 43, 440–449.
- Diwekar, U.M. and K.H. Sahin. (2002). "PDF Using Kernel Functions: Computer Code." Personal communication.
- Levenspiel, O. (1972). *Chemical Reaction Engineering*, 2nd ed. New York: Wiley.
- Merrifield, R.B. (1963). "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide." *Journal of American Chemical Society* 85, 2149–2154.
- Padmanabhan, S.K.R. and Pitchumani. (1999a). "Stochastic Analysis of Isothermal Cure of Resin Systems." *Polymer Composites* 20, 72–85.
- Padmanabhan, S.K.R. and Pitchumani. (1999b). "Stochastic Modeling of Nonisothermal Flow during Resin Transfer Molding." *International Journal of Heat and Mass Transfer* 42, 3057–3070.
- Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. (1999). *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, Vol. 1, 2nd ed. New York: Cambridge University Press.
- Watkins, K. (2001). "Peptides: A Boom in the Making." *Chemical and Engineering News*, January 8.