Computation of Optimal Temperature Policy for Batch Penicillin Synthesis Reaction via Sequential Pseudo-Uniform Design

Jyh-Shyong Chang
Department of Chemical Engineering
Tatung University
Taipei 104, TAIWAN, ROC
jschang@ttu.edu.tw

Ying-Chan Wang
Department of Chemical Engineering
Tatung University
Taipei 104, TAIWAN, ROC
tracy760923@pchome.com.tw

Abstract

The application of the uniform design (UD) to nonlinear multivariate calibration by artificial neural network (ANN) can build a model for an unknown process efficiently because it allows many levels for each factor. If the cost of each experiment is high, low partitioned levels are usually proposed first to carry out the experiments. However, if a reliable ANN model cannot be obtained based on the designed experiments, a sequential pseudo-uniform design (SPUD) method is developed to locate additional experiments in the experimental region. An information free energy index is used to validate the identified ANN model. Once the identified model is reliable, the optimal operating conditions can be determined to guide the process to achieve the desired objective. The simulation results demonstrate that an optimal temperature trajectory for the batch penicillin synthesis reaction based on the proposed method requires only a reasonable number of experiments.

1. Introduction

In the competitive market, manufacturing of innovative products, such as specialty chemicals, ceramic and composite materials, the finding of recipes, and developing new processes in due course of time are much more urgent nowadays. It is also imperative to develop a new product at the minimum cost without sacrificing quality. Traditionally, the models adopted in analyzing the experimental data designed by experimental design methods [1, 2] are either linear or quadratic. To cope with nonlinear characteristics in many processes, the uniform design method can be applied to a nonlinear multivariate system [3, 4, 5]. The main idea behind the UD is to find a set of representative points scattered uniformly and regularly in the domain to be investigated. The so-called good lattice point (*glp*) set which can be generated with the help of the number-theoretic method might be used to fulfill such a task [6]. The experimental design table based on the UD principles can be constructed by means of the *glp* set [6]. The multi-input/multi-output (MIMO) experimental data collected through the UD method can be molded by an ANN and the results had proven to be successful in learning nonlinear and complex relationships between process variables without any prior knowledge of system behavior [3]. No guidelines can be found in the related works [3, 4, 6] on how to choose suitable levels for each control factor in designing

experiments for an unknown process in advance. Because, it is usually time-consuming and costly to get experimental data in practical application, low partitioned levels are usually proposed first to carry out the experiments. If a reliable mapping model (ANN is adopted in this study) cannot be established based on the previously designed experiments, we will propose the SPUD method in this study to locate additional experiments in the experimental region. Partition of the coarse levels into fine levels resembles multigrid methods [7] adopted in solving the high-dimensional nonlinear model equations. The success of multigrid methods stems from the fact that the problem never has to be solved on the fine grid but on only relatively coarse grids with dramatically reduced computational efforts [8]. Furthermore, to determine the possible multiple optima of an unknown process, a new experimental design scheme that uses UD, SPUD, ANN, random search, fuzzy classification, and information theory is proposed in this study. The main objective of the proposed experimental design scheme is to build a reliable model and locate the optima of an unknown process simultaneously based on only a reasonable number of experiments. It is expected that the developed experimental design scheme can be applied in the real world. An alternative approach can be found in the work of Chen et al. [9]. In the experimental design scheme proposed by Chen et al.[9], an initial batch of experimental data is first collected to construct an ANN model followed by the random search to generate a number of candidates for the next batch experiments. A fuzzy classification algorithm is then used to find the cluster centers of these candidates. An information free energy index is defined to balance the need for better classification and the relevance of each class in optimization. New experiments are performed at these cluster centers to validate the model. The procedure is repeated until an optimal solution is reached. The case studies [9] reveal that this scheme focuses on locating the experiments in the investigated space toward the optima of an unknown process, resulting in necessitating too many experiments.

2. Development of the Sequential Pseudo-Uniform Design

If one applies the UD method to organizing an experimental design, the UD tables provided by Fang and Ma can be adopted accordingly. It is expressed as $U_n(q^s)$, where U stands for the uniform design, n for the number of experimental trials, q for the number of levels, and s for the maximum number of factors. In essence, for a system with n factors and each factor divided into q levels, the UD method needs q experiments. The number of experiments based on the UD method is limited. If the experimental design based on a first selection of the number of levels q_0 from a UD table cannot provide enough information to build an accurate ANN model, then we may need to interpolate sequentially another q_1 level of each factor among the original levels q_0 to run another q_1 experiment. In this study, the way to arrange these q_1 experiment such that these $q_0 + q_1$ experiments can scatter uniformly and regularly in the investigated domain is proposed.

The design principles of the UD method [10, 11] are:

(a) The occurrence of each level of each factor in the experiments is once only.
(b) The number of experiments equals the level number of factors.
(c) The selected experimental points are distributed uniformly in the experimental domain.

These three principles are also considered in the developed SPUD method. There are $k = \prod_{l=1}^{q_1} (q_1^{s-1} - (l-1))$ possible combinations that satisfy the above constraints (a) and (c) in locating these q_1 interpolated experiments. To fit the requirement (d), the proposed SPUD method suggests that these sequential q_1 experiments be arranged by the following maximini problem.

$$\underset{\forall\ k\ \text{combinations of the interpolated}\ \mathbf{x}^*}{\text{Maximize}}\quad J_{k'=1,k} =$$

$$\min\sum_{i=1}^{q_1}\left(\sum_{j=1}^{q_0}\left\|\mathbf{x}^*_{ik'}-\mathbf{x}_{jk'}\right\|+\sum_{l=1,l\neq i}^{q_1}\left\|\mathbf{x}^*_{ik'}-\mathbf{x}^*_{lk'}\right\|\right) \quad (1)$$

If $q_t = q_0 + q_1$ experiments cannot provide enough information for building a reliable model (ANN used in this study), then the procedure discussed above can be repeated.

3. Optimization of Products and Process via SPUD Method

The proposed algorithm to determine the optimum operating conditions for producing a product or a process via the SPUD method is summarized as follows:

(1) Choose a suitable level q_0 for each control factor based on the available UD table.

(2) Build an FNN model based on the experimental data from (1).

(3) Check the adequacy of the identified FNN model by the interpolated experiments chosen based on the SPUD method (recycle from (4)) or the UD method (recycle from (6)).

(4) Find the optimal conditions by the random search method if the identified model is adequate; go to (5). Otherwise, do the chosen experiments provided by the SPUD method and identify a new FNN model based on the available experiment data; go to (3).

(5) Check the adequacy of the optimal conditions experimentally; if the optimal conditions determined are reliable, stop the procedure.

(6) If the condition given in (5) cannot be met, carry out more experiments around the assumed optimal conditions based on the UD table again and identify a new FNN model; go to (3).

The objective function is defined based on the identified FNN model:

$$\underset{\mathbf{l}\le\mathbf{x}\le\mathbf{u}}{\text{Minimize}}\ f(\mathbf{y}(\mathbf{x})) \quad (2)$$

where $f(\mathbf{y}(\mathbf{x}))$ is the objective function based on the identified FNN model $\mathbf{y}(\mathbf{x})$. To carry the algorithm, each element of $\mathbf{x}$ is rescaled to be in a span of [-1 1]. There are $2^{\dim(\mathbf{x})}$ combinations of the elements of $\mathbf{x}$ with either a positive or a negative sign. In the same figure, it is termed "one Run" each time the chosen r_{Run} test experiments based on the SPUD are carried out. The model error based on the test experiments is defined as:

$$\varepsilon_{Run} = \frac{\sum_{i=1}^{r_{Run}}\left(\left|\frac{Y(\mathbf{x}_i)-y(\mathbf{x}_i)}{Y(\mathbf{x}_i)}\right|\right)}{r_{Run}} \quad (3)$$

By way of illustration, an example with two factors is adopted. Consider a first arrangement of UD experiments by $U_4(4^2)$: C_1=(1, 3)、C_2=(3, 7)、C_3=(5, 1)、C_4=(7, 5), 4 experiments ($q_0 = 4$) are required (Figure 1). As shown in Figure 1, the coordinates x_1 and x_2 of these experiments are set to be integers. Transforming the integral coordinates into a real dimension is required in a real application. If the experiments with $q_0 = 4$ levels cannot provide enough information to build a reliable model, then we suggest new experiments at the points where the x_1 and x_2 coordinates are located at 2, 4, and 6 as shown in Figure 1. Now, $q_t(= q_0 + q_1 = 7)$ levels are established. In this example, we want to determine these sequential q_1(=3) experiments among the six $(k = 6)$ possible interpolated experiment combinations, which are located at [(2, 2), (4, 4), (7, 7)], [(2, 2), (4, 6), (6, 4)], [(2, 4), (4, 2), (6, 6)], [(2, 4), (4, 6), (6, 2)], [(2, 6), (4, 2), (6, 4)], and [(2, 6), (4, 4), (6, 2)], respectively. Because $k = 6$ is a small number, the maximization-minimization problem (eq 1) can be carried out over these six possible interpolated experiment combinations exhaustively to find out which combination could achieve the maximum value of the objective function as defined in eq 1. The resulting SPUD experiments will be the same as those designed using the UD method. However, the possible interpolated experiment combinations will be increased rapidly as shown in the equation $k = \prod_{l=1}^{q_1}(q_1^{\,s-1} - (l-1))$. For example, if a problem has two factors ($s = 2$) and $q_0 = 10$, then there

will be 362,880 possible interpolated experiment combinations ($q_1 = 9$). Therefore, in this study, if $k \le 1\times10^5$, the maximini problem (eq 1) was done over all the possible interpolated experiments exhaustively; otherwise, a maximum of 1×10^5 experiment combinations were chosen randomly from the interpolated coordinates of the factors that satisfy the design principles of the UD method. A maximum value of the objective function (eq 1) can almost be determined over these 1×10^5 randomly chosen experiment combinations. In this way, a sequential pseudo-uniform design of new experiments can be worked out properly.

Liang et al. [4] proposed the term "discrepancy" for use in measuring the uniformity of the distributed experiment points. Let $\Re = \{\mathbf{x}_k, k = 1,..., n\}$ be a set of points on C^s, where $\mathbf{x}_k (= [x_{k1}, x_{k2},..., x_{ks}]^T)$ is an s-dimensional column vector. For any $\boldsymbol{\gamma} \in C^s$, let $N(\boldsymbol{\gamma}, \Re)$ be the number of points that satisfy $\boldsymbol{x}_k \le \boldsymbol{\gamma}$. Then

$$D(n, \Re) = \sup\left|(N(\boldsymbol{\gamma}, \Re)/n - v([0, \boldsymbol{\gamma}]\right| \qquad (3)$$

is called the discrepancy of $\Re$, where $v([0, \boldsymbol{\gamma}]) = \gamma_1 \times \cdots \times \gamma_s$ denotes the volume of the super rectangle $[0, \boldsymbol{\gamma}]$.

To show the applicability of the SPUD method, we give an example with two factors, each factor being partitioned into 29 levels. The value of $D(n, \Re)$ designed by the UD method is 0.01517 as shown in Figure 2(a). For the same problem, if we first choose 15 levels based on the UD method and then proceed to use the SPUD method to locate the additional 14 levels, the same discrepancy of $\Re$ ($D(n, \Re)$ =0.01517) can also be obtained as shown in Figure 2(b).

4. Artificial Neural Network

Artificial neural networks are known to be a powerful tool to approximate complex multivariable functions [12, 13]. In this study, we adopt the most traditional feedforward neural network (FNN) to build the relationships between any nonlinear inputs and outputs of an unknown process. An FNN is usually composed of three layers of the network structure: input layer, hidden layer, and output layer. The routines provided in the Neural Network Toolbox (MathWorks) were used to build such a neural network.

5. Random Search

Two optimization methods are usually available to find the optimum of an objective function constrained by a nonlinear functional model such as FNN. The first method is the gradient-based optimization technique such as the Newton method. The other method is the non-gradient based optimization technique such as the random search method [14]. The random search method [15] explores sequentially the parameter space of an objective function in a seemingly random fashion to find the optimal condition for minimizing (or maximizing) the objective function. Although the random search is a direct application method without the need of differential information, the optimal point obtained is quite dependent on the initial guess. Therefore, a large amount of initial guess will be initiated to locate the most representative candidate points for the best performance. The calculation strategy of the random search method proposed by Solis and Wets [15] was adopted.

6. Information Index

In this study, the information index developed by Chen et al. [9] was not used to locate the additional experiments to be explored but was adopted as a quality evaluation of the identified model. The required techniques for retrieving the information index based on the experimental data can be found in the article of Chen et al. [9] except that the term T_{max} in equation 16 of Chen et al. [9] shall be corrected as f_{max}.

7. Application to a Batch Penicillin Synthesis Reaction

In order to prove the ability of the proposed SPUD method in determining an optimal temperature trajectory of a batch reactor where penicillin synthesis reaction is carried out. For the case under studied, the SPUD method is applied to build a reliable model and to locate the optima of an unknown process simultaneously based on

only a reasonable number of experiments.

The differential equations describing the state of the batch penicillin synthesis system are as follows:

The cell-mass production rate:

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = b_1 y_1 - \frac{b_1}{b_2} y_1^2 \tag{4}$$

The penicillin synthesis rate:

$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = b_3 y_1 \tag{5}$$

with

$$b_1 = w_1 \left[\frac{1.0 - w_2(\theta - w_3)^2}{1.0 - w_2(25 - w_3)^2} \right] \tag{6}$$

$$b_2 = w_4 \left[\frac{1.0 - w_2(\theta - w_3)^2}{1.0 - w_2(25 - w_3)^2} \right] \tag{7}$$

$$b_3 = w_5 \left[\frac{1.0 - w_2(\theta - w_6)^2}{1.0 - w_2(25 - w_6)^2} \right] \tag{8}$$

where $w_1 = 13.1$, $w_2 = 0.005$, $w_3 = 30°\mathrm{C}$, $w_4 = 0.94$, $w_5 = 1.71$, and $w_6 = 20°\mathrm{C}$. y_1 and y_2 represent respectively the dimensionless concentrations of cell mass and penicillin; t is the dimensionless time, $0 \le t \le 1$; and θ is the temperature in degrees Celsius. The parameters b_i , $i = 1 \sim 3$, are functions of temperature, $b_i \ge 0$. The initial concentrations are specified as $y_1(0) = 0.3$ and $y_2(0) = 0$.

The way of designing the temperature function of this case is to use the orthogonal collocation method. There are three orthogonal collocation points along the dimensionless time axis, which are located at 0, 0.667, and 1 respectively. The locations of temperatures ($\theta_1, \theta_2, \theta_3$) are taken in the temperature range of $19℃ \le \theta \le 30℃$. The temperature curves formed by these 3 orthogonal collocation points are shown in Figure 3. Therefore, our objective is to find a temperature trajectory such that the concentration of penicillin at the end of the fermentation (i.e., $t = 1$) is maximum.

$$J = \max_{\theta(t)} y_2(t = 1) \tag{9}$$

4 Runs are made totally in this case (Figure 4), 33 experiments are required and the relationship between the numbers of network nodes are the Run of each time is as shown on Table 1. Temperature collocation points of the minimum objective value found by the model, which is built at the last time of Run, ($\theta_1, \theta_2, \theta_3$) is located at (25.94℃, 24.73℃, 24.00℃) (as shown in Figure 5), the model value is 1.4736. When this point is put into the original process, its value is 1.4733. The model error is $\frac{|43.31 - 43.99|}{43.31} = 1.55\%$.

8. Conclusions

In industrial production, a desired product recipe or an optimal operation approach is usually sought. Because the system under study is usually unknown, it is hoped that the best product recipe or the optimal operating path of a process can be obtained with minimum experiments. In this research, the experiment design is first based on the UD method. Then the developed SPUD method is applied to locate additional experiments in the experimental region. An information free energy index is used to validate the identified ANN model. Once the identified model is reliable, the optimal operating conditions can be determined to guide the process to achieve the desired objective.

The batch penicillin synthesis reaction process of the case study could be regarded as the operational optimization problem of an unknown process. Application of the SPUD method required 33 batches, we could obtain the optimal temperature trajectory for this case study without resorting to a physical model based approach.

Acknowledgments

We thank the National Science Council (Grant NSC 92-2214-E-036-002) and Dr. T. S. Lin, President of Tatung University, Taipei, Taiwan, R.O.C. for all the support conducive to the completion of this work.

References

[1] Logothetis, N.; Wynn, H. P. Quality Through Design: Experimental Design, Off-line Quality Control, and Taguchi's Contributions, Oxford University Press Inc.: New York, 1994.

[2] Myers, R. H.; Montgomery, D. C. Response Surface Methodology: Process and Product Optimization Using Designed Experiments, John Wiley & Sons. Inc.: New York, 1995.

[3] Zhang, L.; Liang, Y. Z.; Jiang, J. H.; Yu, R. Q.; Fang, K. T. Uniform Design applied to Nonlinear Multivariate Calibration by ANN. *Analytica Chimica Acta* **1998**, 370, 65.

[4] Liang, Y. Z.; Fang, K. T.; Xu, Q. S. Uniform Design and its Applications in Chemistry and Chemical Engineering. *Chem. Int. Lab. Syst.* **2001**, *58*, 43.

[5] Cheng, B.; Zhu, N.; Fan, R.; Zhou, C.; Zhang, G.; Li, W.; Ji, K. Computer Aided Optimum Design of Rubber Recipe using Uniform Design. *Polymer Testing* **2002**, *21*, 83.

[6] Fang, K. T.; Wang, Y. *Number-theoretic Methods in Statistics*, Chapman & Hall, 2-6 Boundary Row, London SE1 8HN, UK, 1994.

[7] Wesseling, P. *An Introduction to Multigrid Methods*; Wiley & Sons: Chichester, U. K., 1991.

[8] Briesen, H.; Marquardt, W. An Adaptive Multigrid Method for Steady-State Simualtion of Petroleum Mixture Separation Processes, *Ind. Eng. Chem. Res.* **2003**, 42, 2334.

[9] Chen, J.; Wong, D. S. H.; Jang, S. S. Product and Process Development using Artificial Neural-Network Model and Information Analysis. *AIChE J.* **1998**, *44*, 876.

[10] Fang, K. T. and Ma, Z. X. *Orthogonal and Uniform Experimental Design*; Hong Kong Baptist University: HK, 2000.

[11] Fang, K. T. Uniform Design: Application of Number-Theoretic Methods in Experimental Design. *Acta Math. Appl. Sin.* **1980**, *3*, 363.

[12] Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Neural Networks are Universal Approximators. *Neural Networks*. **1989**, *2*, 359.

[13] Hornik, K.; Stinchcombe, M.; White, H. Universal Approximation of an Unknown Mapping and Its Derivatives Using Multilayer Feedforward Networks. *Neural Networks*. **1990**, 3, 551.

[14] Jang, J. S. R.; Sun, C. T.; Mizutani, E. *Neural-Fuzzy and Soft Computing*; Prentice-Hall International: USA, 1997.

[15] Solis, F. J.; Wets, J. B. Minimization by Random Search Technique. *Math. Opera. Res.* **1981**, *6*, 19.

Table 1. Relationship between the Run and the Node

Run	Node
1	2
2	3
3	4
4	5

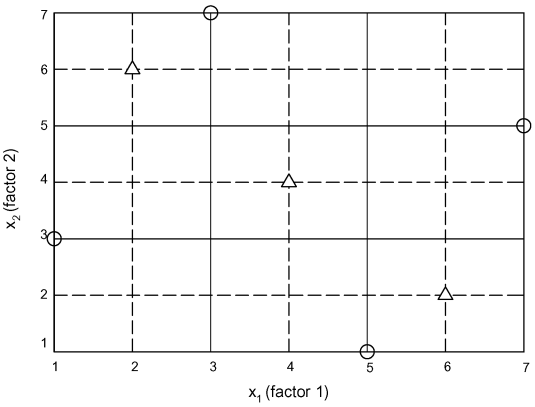

Figure 1. The sequential pseudo-uniform design (SPUD) for the case of two factors (◯: initial uniform design points; △: sequential pseudo-uniform design points).

.

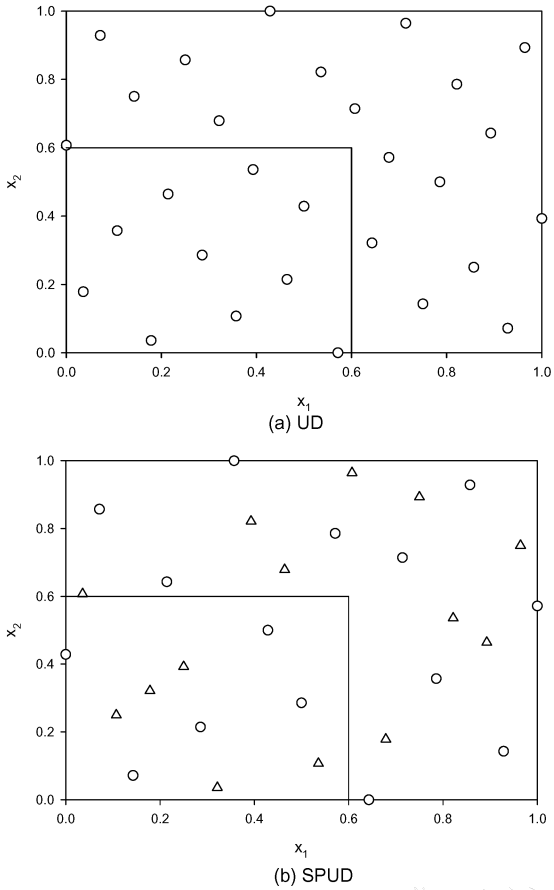

Figure 2. . Comparison of the discrepancy measure designed by (a) the UD method (discrepancy $= \left|10/29 - 0.6 \times 0.6\right| = 0.01517$) and (b) the SPUD method (discrepancy $= \left|10/29 - 0.6 \times 0.6\right| = 0.01517$; ◯ : initial uniform design points; △ : sequential pseudo-uniform design points).

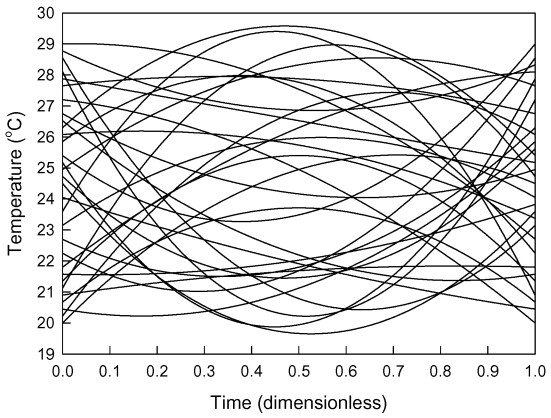

Figure 3. The temperature trajectories based on the three collocation points of temperature.

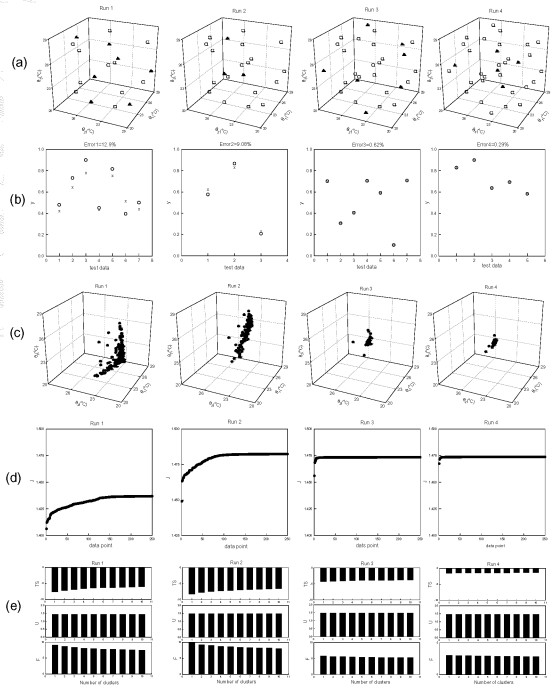

Figure 4. Optimal experimental design of the batch penicillin synthesis reaction ((a) experimental arrangement points: training ○ and testing ▲ (b) testing point error (c) location of the best model input value (d) objective value of the best model input value (e) information index).

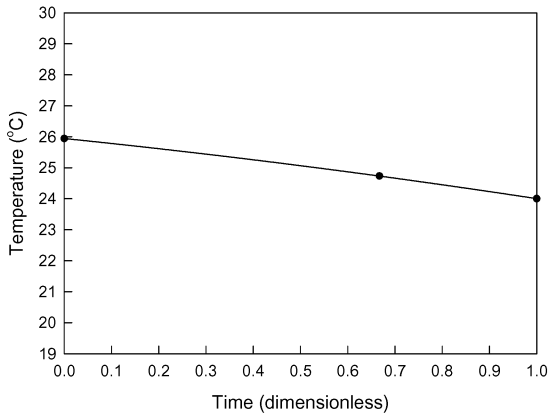

Figure 5. The optimal temperature trajectory of three temperature collocation points of the batch penicillin synthesis reaction.