Understanding and Applying the Extrapolation Properties of Serial Gray-Box Models

Henricus J. L. van Can
Kluyver Laboratory for Biotechnology, Delft University of Technology, 2628 BC Delft, The Netherlands

Hubert A. B. te Braake, Sander Dubbelman, Chris Hellinga, Karel Ch. A. M. Luyben, and Joseph J. Heijnen
Control Laboratory, Delft University of Technology, 2624 CD Delft, The Netherlands

There is a need for efficient modeling strategies that quickly lead to reliable models. In the serial gray-box modeling strategy the inaccurately known terms in the macroscopic balance are modeled with a black-box model structure such as a neural network. This way one efficiently obtains an accurate model with good extrapolation properties without the need to develop a rigorous white-box model in which much more knowledge would be involved and without the need to do many identification experiments. Different types of extrapolation are specified and analyzed and they are related to the different parts of the serial gray-box model structure. The strategy is demonstrated for modeling pH effects on the enzymatic conversion of penicillin G using real-life data. The resulting serial gray-box model is compared with a model from a more knowledge-driven white-box strategy and with a model from a more data-driven black-box strategy. The serial gray-box modeling strategy is especially advantageous at the medium level of process operation, which is mainly concerned with the calculation of optimal condition for (bio)chemical processes.

Introduction

Mathematical models are very useful for the design and optimization of (bio)chemical processes. Unfortunately, the model necessary for a given process usually is not directly available, so that development of the model becomes a laborious and expensive stage in the whole design or optimization procedure. Therefore, there is a need for efficient modeling strategies that quickly lead to reliable models.

Existing modeling strategies can be divided into white-box, black-box, and gray-box strategies, depending on the amount of knowledge that is used for development of the model. White-box strategies are mainly knowledge driven, which means that the time involved to obtain the necessary knowledge can be in conflict with the desire for an efficient development of the model. Black-box strategies are mainly data driven and the resulting models often do not have reliable extrapolation properties, which means that the need for many identification experiments, in order to cover the whole application domain, can be in conflict with the desire for an efficient model development. Nevertheless, black-box strategies have been applied to (bio)chemical processes, especially since convenient nonlinear black-box modeling tools like neural networks have been available (Thibault et al., 1990; Bhat and McAvoy, 1990; Psichogios and Ungar, 1991; Montague and Morris, 1994; Zhang et al., 1994; Albiol et al., 1995; Zhu et al., 1996). Gray-box modeling strategies are potentially very efficient if the black-box and white-box components are combined in such a way that the resulting models have good interpolation and extrapolation properties, without the need to develop a rigorous white-box model or the need to do many identification experiments.

Previous articles already presented two types of gray-box modeling strategies in which a neural network is combined with a white-box model: the parallel and the serial gray-box modeling strategies. In the parallel strategy the neural network is placed parallel with a white-box model. The neural network then performs as an error model, which should model the difference between the white-box model and reality. Su et al. (1992) and Côte et al. (1995) demonstrated that the

Correspondence concerning this article should be addressed to H. J. L. van Can at his current address: Melkunie, P.O. Box 222, 3440 AE Woerden, The Netherlands.

parallel approach resulted in models that had better interpolation properties than pure black-box models. Van Can et al. (1996a) demonstrated the parallel strategy for the modeling and real-time control of a pressure vessel and showed that in their case the parallel gray-box model had no reliable extrapolation properties.

In the serial gray-box modeling strategy the neural network is placed in series with a first-principles model. This seems more promising with respect to extrapolation. Various researchers (Psichogios and Ungar, 1992; Thompson and Kramer, 1994; Schubert et al., 1994; Simutis et al., 1995; Dors et al., 1995; Tholudur and Ramirez, 1996) showed the potential extrapolation properties of serial gray-box models. However, these articles made no distinction between different types of extrapolation, and no attention was paid to the origin of the extrapolation properties in the serial gray-box model structure. Still, there are two important reasons to investigate these aspects of gray-box modeling. First, as for complete black-box models, extrapolation of the black-box part of the serial gray-box model should be avoided. In order to be sure that extrapolation does not rely on the black-box part of a gray-box model, it is important to define more precise extrapolation and to find its origin in the model structure. Second, detailed knowledge about extrapolation and its origin is needed to fully benefit from the serial gray-box modeling strategy, because reliable extrapolation properties provide the basis to select a *small* set of identification experiments given a desired *large* application domain for the model. Van Can et al. (1996a,b) did distinguish between different types of extrapolation and related these to different parts of the serial gray-box model. It was shown that this is a good basis to select a *small* domain for the identification data, given a *large* application domain for the model.

It is the purpose of the present article to further specify and analyze different types of extrapolation and to relate them to the different parts of the serial gray-box model structure. It will be shown how the serial gray-box modeling strategy can be applied efficiently to yield models with good interpolation and different types of good extrapolation, without the need to develop a rigorous white-box model and without the need to do many identification experiments. This is demonstrated for the modeling of the pH effect on the enzymatic conversion of penicillin G, using real-life data. Moreover, the performance of the serial gray-box model will be compared with the performance of a model from a more knowledge-driven white-box strategy and with the performance of a model from a more data-driven black-box strategy. Based on the results in this article it can be concluded that the serial gray-box modeling strategy is especially advantageous at the medium level of process operation, which is mainly concerned with the calculation of optimal condition for (bio)chemical processes.

Outline of the Article

First, we distinguish between different levels of process operation, for which mathematical models can be used to obtain process improvements. As stated in the Introduction, for a given process the necessary model is usually not available, and therefore this article focuses on efficient modeling strategies that quickly lead to reliable models. Good extrapolation properties are an essential part of an efficient modeling strategy, because they provide the basis to select a small set of identification experiments given a large application domain for the model. Therefore, after the section on different levels of process operation we define more precisely different types of extrapolation that play a central role in the remainder of the article. After that we introduce the serial gray-box modeling strategy as a strategy in which the inaccurately known terms in the macroscopic balance are a modeled black box, for example, with a neural network. This strategy is potentially very efficient and quickly leads to reliable models with good extrapolation properties, without the need to develop a rigorous white-box model and without the need to do many identification experiments.

The remainder of the article is devoted to the application of the serial gray-box modeling strategy to the modeling of the pH effect on the enzymatic conversion of penicillin G, using real-life data. For comparison, a white-box and a black-box modeling strategy are also applied to the system under consideration. The applied model structures and their identification are reported in detail. The three model structures (gray, white, and black) are validated and compared on the basis of their interpolation and various extrapolation properties using 28 validation experiments. Finally, the discussion part of this article is devoted to the general applicability of the serial gray-box modeling strategy, including its extrapolation properties, and especially to its specific advantages for mathematical modeling at the different levels of process operation. Based on the results in this article it is argued that the serial gray-box modeling strategy is especially advantageous at the medium level of process operation, which is mainly concerned with the determination of optimal setpoint trajectories.

Levels of Process Operation

The economics of (bio)chemical processes severely depend on the choice of the (bio)catalyst, the choice of process equipment and the way of process operation. In a continuous search for improved processes, none of these aspects can be neglected. Still, in this article, we concentrate only on the construction of mathematical models for improvement of process operation. Process operations can be defined as a class of operations that has to meet production targets, ensure a constant product quality, minimize operation costs, and fully utilize the possibilities of the given catalyst and process equipment. Typically, process operations have to deal with a given situation of catalyst and process equipment. Process operation can be divided into three levels (Figure 1) and at each level mathematical models can be used: a lower level that is mainly concerned with setpoint tracking; a medium level that is mainly concerned with supervisory control tasks; and a higher level that is mainly concerned with plantwide control tasks. Generally speaking, human interference increases and mathematical precision decreases from the lower to the higher level of process operation.

The lower level of process operation should ensure a robust and stable process by reducing the effects of disturbances and by accurately tracking the predefined setpoints. Already for some decades, this level has been the domain of process control and is characterized by a high level of au-

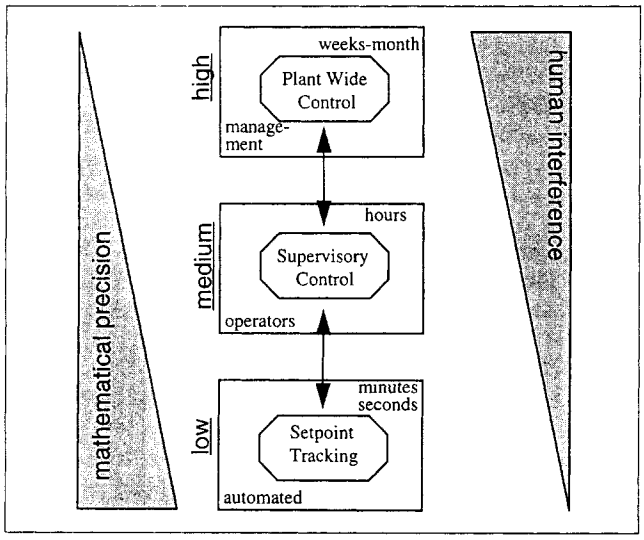


Figure 1. Three levels of process operation.

tomation. Typically, the tasks at the lower level are associated with the smaller time constants of the process. For example, adjustments for pH and temperature control are made every few seconds or minutes. Clearly, at the lower level of process operation direct on-line information of the actual state of the process is of key importance. This information can be obtained from on-line measurements or by state estimation using a model of the process. Unfortunately, for the latter case the necessary model is usually not directly available. Moreover, at the lower level of process operation improvements can be obtained by the application of advanced process control. Again, this usually requires a model that is not directly available for a given process.

The medium level of process operation is mainly concerned with supervisory control tasks. For (bio)chemical processes one could think of determining the setpoint trajectories for pH, temperature, and dissolved oxygen, and determining nutrient feeding strategies, fault detection, and checking the process for its overall performance (e.g., by off-line analyzing samples). Typically, the tasks at the medium level of process operation are associated with the larger time constants of the process. For example, samples are taken once or twice per batch process (or per day), and setpoints are seldom adjusted on-line. Traditionally, this medium level of process operation has been the role of human operators. For (bio)chemical processes the medium level of process operation is very important for the economics of the process. The performance of (bio)catalysts depends severely on their direct environmental conditions, such as substrate and product concentrations, pH, temperature, and oxygen. For example, in many fermentation processes 2–3 nutrients must be fed to the reactor in an optimal way, and also pH and temperature time-profiles can contribute to a more optimal process performance. Whereas traditionally these conditions were determined by trial and error (resulting in suboptimal conditions), nowadays the role of model-based approaches is increasing (Shioya, 1992; Lopez and Malcata, 1993; Uhlemann et al., 1994; Zhang et al., 1994; Lee and Ramirez, 1994; Chang and Hseih, 1995; Modak and Lim, 1995). In these model-based approaches the development of the model is the time-limiting step. Existing models are usually not accurate enough for one specific process, and development of new, accurate models is very time- and money-consuming.

The higher level of process operation is mainly concerned with plantwide control tasks: how much should be produced and when, and how can the available equipment be used optimally? For example, in (bio)chemical industries efficient scheduling of different batch operations can contribute to a more optimal process performance. Typically, the tasks at this level are associated with even larger time constants than the tasks at the medium level. For example, a production schedule is determined only weekly or monthly. Traditionally, this level has been the domain of management. Also at this higher level of process operation, the role of mathematical models is increasing. However, the type of models that are used at this level are mainly determined by business economics and operations research, and less by the nature of the processes at hand. Therefore, this is considered to be outside the scope of this article.

In conclusion, mathematical models are very useful for improvement of process operation. Unfortunately, for a given process the necessary model is usually not available, and therefore this article focuses on efficient modeling strategies that quickly lead to reliable models. In the next section we define different kinds of extrapolation that are essential for an efficient modeling strategy. After that, we introduce the serial gray-box modeling strategy.

Different Types of Extrapolation

There are two important reasons to distinguish between different types of extrapolation. First, as for complete black-box models, extrapolation of the black-box part of the serial gray-box model should be avoided. In order to be sure that extrapolation does not rely on the black-box part of a gray-box model, it is important to more precisely define extrapolation and to find its origin in the model structure. Second, detailed knowledge about extrapolation and its origin is needed to fully benefit from the serial gray-box modeling strategy, because reliable extrapolation properties provide the basis to select a *small* set of identification experiments given a *large* application domain for the model. We distinguish between two main types of extrapolation: amplitude and frequency extrapolation.

In amplitude extrapolation the value of a variable during application of the model is higher than the highest, or lower than the lowest value in the set of identification data. Amplitude extrapolation can be subdivided into range and dimensional extrapolation (Van Can et al., 1996a,b). In range extrapolation (Figure 2b), a variable (x_1 or x_2) is applied outside the range within which it was varied during identification. Consequently, the range of the application domain of the model is larger than the range of the identification domain. A typical example of range extrapolation is given in Figure 3a: a variable (x_1) adopted three different values (amplitudes) during the identification of the models (solid lines), but was used at a higher or lower value (amplitude) during application of the model (dashed line). In the model development phase the boundaries of the application domain of the model are not always known accurately enough to completely avoid the necessity for range extrapolation during the application of the model. Therefore, the model should remain

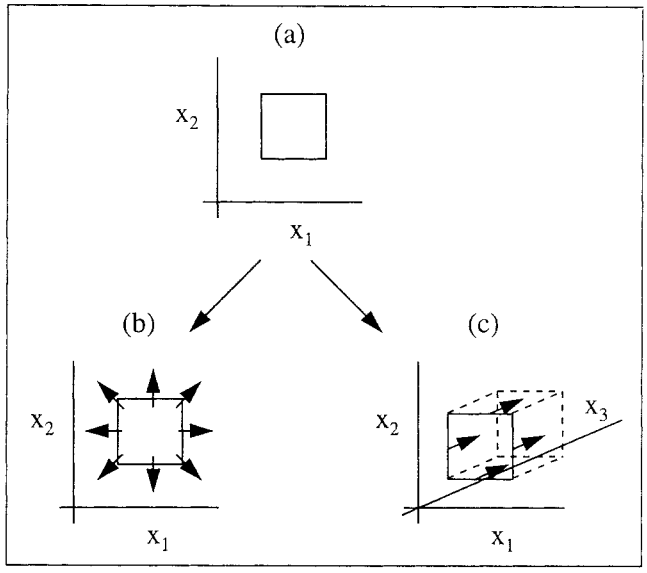


Figure 2. (a) Domain covered by the identification data; (b) extending this by range extrapolation; and (c) extending this by dimensional extrapolation.

useful when it is used just outside the range of the identification data.

In dimensional extrapolation (Figure 2c) a variable (x_3), which was kept constant during identification, varies during the application of the model. Consequently, the dimension of the application domain is larger than the dimension of the identification domain. A typical case of dimensional extrapolation is given in Figure 3b: the variable x_1 adopted only one value (amplitude) in the identification experiments (solid line), whereas it adopted different values (amplitudes) during application of the model (dashed lines). Good dimensional extrapolation properties are very important for fast model development, because the number of identification experiments can be reduced considerably when not all variables have to be changed in the identification phase of the model. Moreover, sometimes it is even impossible to vary a variable in the model development phase, because the necessary experimental setup is not available (e.g., large scale). Dimensional extrapolation can be seen as a special case of range extrapolation, in which the range is very small (or zero) in the identification phase. Still, it is important to distinguish between the two, because they can be related to different parts of the gray-box model, as is shown in this article.

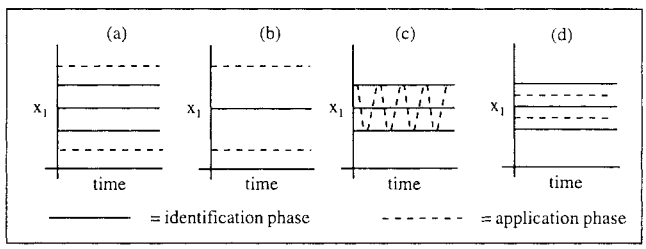


Figure 3. Typical case of (a) range extrapolation; (b) dimensional extrapolation; (c) frequency extrapolation; and (d) interpolation.

The time courses of x_1 in different identification/application experiments are plotted together in one x_1 vs. time graph.

In frequency extrapolation a variable is used at a frequency that is lower or higher than the lowest or highest frequency in the identification experiments. This means that the dynamic behavior of the system in the application phase of the model is different from the dynamic behavior of the system during the identification experiments. A typical example of frequency extrapolation is given in Figure 3c: a variable is constant in the identification experiments (solid lines), although it is used at different amplitude levels, but changes dynamically during the application of the model (dashed lines). Good frequency extrapolation properties are very important for efficient modeling of (bio)chemical processes. For example, a model that has no reliable frequency extrapolation properties and that has to be used for the calculation of optimal pH, temperature, or feed profiles in a (fed-) batch process, must be identified with data from many (fed-) batch runs that contain all kinds of possible profiles. If, on the other hand, the model has reliable frequency-extrapolation properties, the number of identification experiments can be reduced considerably, because not all of the possible kinds of profiles have to be present in the identification experiments.

For reasons of completeness, interpolation is defined as a situation in which a variable is used at frequencies and amplitudes between the highest and lowest frequency and amplitude in the identification experiments. A typical example of interpolation is given in Figure 3d: a variable (x_1) is constant during the identification and application of the model (no frequency extrapolation) and its amplitude during application is between the highest and lowest amplitude during the identification (no amplitude extrapolation).

Serial Gray-Box Modeling Strategy Based on the Macroscopic Balances

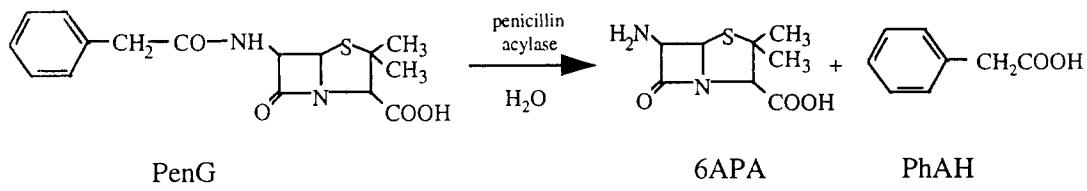
The serial gray-box modeling strategy can be combined quite naturally with the general structure of white-box dynamic models in (bio)chemical processes, which are always based on macroscopic balances, for example, of mass, energy, or momentum. These balances specify the dynamics of the relevant state variables and contain different rate terms that are associated with mechanisms such as transport and conversion. Some of these terms are associated with directly manipulated or measured variables (such as in- and outgoing flows) and, consequently, their mathematical relation with these variables is straightforward and no effort is needed to specify these relations. In that case these terms can be considered as accurately known terms. In contrast, some rate terms (e.g., conversion kinetics) have a more or less complex mathematical relation with one or more system variables, which should be modeled in order to obtain a fully specified model. In that case these terms can be considered as inaccurately known terms. If such a mathematical relation between one or more system variables and a rate term can be based on directly available or easily obtainable first principles, it can be modeled easily in a white-box way. If, on the other hand, this is not possible, detailed knowledge about the involved mechanisms is needed to obtain a complete (rigorous)

white-box model. This detailed knowledge is not needed if the complex nonlinear relation between the system variables and a rate term is modeled with a neural network. In that case, one obtains quite naturally the serial gray-box model configuration.

In this article it is shown that for a serial gray-box model the identification data only have to cover the amplitude domain of the input–output space of the inaccurately known terms without taking into account the complete dimension and frequencies of the future application domain of the complete model. Dimensional extrapolation relies on the accurately known terms in the macroscopic balances. Frequency extrapolation relies on the correct specification of the set of the macroscopic balances, which means that all relevant rate terms should be incorporated in the balance equations and which means that the relations modeled by the neural network should be static (time-independent). As a result, one obtains an accurate model with good interpolation and dimensional and frequency extrapolation properties, without the need for many identification experiments and without the need to develop a rigorous white-box model in which detailed knowledge about many more mechanisms would have to be taken into account.

Test Case

The serial gray-box modeling strategy is demonstrated with experimental data for the modeling of pH effect on the enzymatic conversion of penicillin G (PenG) to 6-aminopenicillanic acid (6APA) and phenyl acetic acid (PhAH), by the enzyme penicillin acylase:



For model identification and validation, experiments were performed at a temperature of 310 K, in the pH range of 5.5–8.5 and at initial PenG concentrations between 12.5 and 100 mM. Due to the acid–base equilibria of PenG, 6APA, and PhAH, H^+ is released during the conversion. Consequently, the amount of added base that is needed to keep the pH at setpoint during the conversion is a very valuable on-line signal, which can be used to calculate time series of concentrations and conversion rates (Appendix B). These time series provided the necessary data for model identification and validation.

The enzymatic conversion of PenG was chosen as a test case because (1) it can very well be performed in (fed/repeated) batch mode, which is very common for many (bio)chemical processes; (2) it has complex nonlinear kinetics, which is also very common for (bio)chemical processes; (3) it allows reproducible experiments, so that the results from testing the predictive properties of a model are not obscured by experimental uncertainties; and (4) the work by many researchers had already resulted in a white-box model, so that a comparison between a gray- and white-box model is relatively easy to make.

Materials and Methods

Experimental setup

The potassium penicillin G (PenG) was donated by Gist-Brocades and was at least 99% pure. The 6-aminopenicillanic acid (6APA) was purchased from Janssen Chimica (>96%), and the phenyl acetic acid (PhAH) from J. T. Backer (>99%). All compounds were used without further purification. The enzyme penicillin acylase was donated in soluble form by Gist Brocades. The activity of the available enzyme was standardized by arbitrarily defining 1.0 mL of the available enzyme solution as 1 U.

The overall experimental setup is presented in Figure 4. All experiments were performed in a thermostatted (at 310 K) reactor with a maximum volume of 1,500 cm^3 (diameter of 13 cm and a height of 11 cm), equipped with three baffles and a Rushton stirrer (250 rpm, 2×6 blades with a diameter of 5 cm positioned at 1.5 cm and 3.5 cm from the bottom of the reactor). Solutions with the required concentration PenG were prepared by dissolving known masses in a 50 mM phosphate buffer. The pH was adjusted with 1.0 N NaOH and 2.0 N HCl to the desired initial pH. The reaction was started by adding an accurately known volume of enzyme solution to the reactor. To bring the pH as close as possible to the pH setpoint, 1.0 M NaOH could be added to the reactor with a Metrohm 665 Dosimat. There was no possibility of automatically adding acid to the reactor. Time series of pH and base addition were stored for future data processing. All software (data-acquisition and controller) was custom-made in order to ensure maximum flexibility.

In most experiments samples were taken and analyzed off-line by HPLC. Samples were taken before the beginning of the conversion (before enzyme was added) and at the end of the conversion (when no base was added anymore for more than 300 s). The HPLC apparatus consisted of a Bondapack phenyl corasil precolumn, a μ Bondapack phenyl 3.9×300-mm column (Waters), a Waters 590 pump, a WISP 712 auto injector, and a Waters 996 diode array detector (ρ 254 nm), using a citrate phosphate buffer with 20% acetonitril and 10% methanol at pH 4 as the eluent.

Numerical methods

In this article, we only used feedforward artificial neural networks with one hidden layer and the tangent hyperbolic

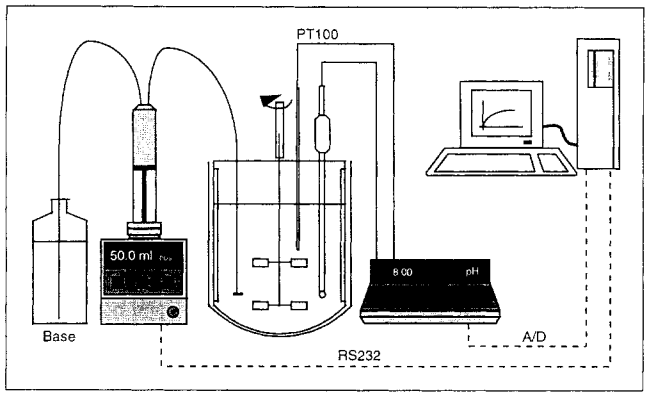


Figure 4. Experimental setup.

function as nonlinear activation function. Mathematical details and the used training algorithm are described elsewhere (Te Braake et al., 1996; Van Can et al., 1996b). The nonlinear equation, resulting from the charge balance was solved with the simplex algorithm available within MATLAB (The Mathworks Inc.). The parameters of the white-box model were estimated with the sequential quadratic programming algorithm, also available within MATLAB (The Mathworks Inc.).

Applied Models

Introduction

In this article three models were developed: a serial gray-box model (SGBM, Figure 5a), a white-box model (WBM, Figure 5b), and a black-box model (BBM, Figure 5c). The SGBM and WBM were based on the same balance equations. In the SGBM the inaccurately known terms in the balance equations were modeled black box, whereas in the white-box model these terms were modeled white box. The BBM was constructed using significantly less detailed knowledge about the system, as is explained in the section "Black-Box Model." The performance of the three models was compared on their ability to predict the base addition of a complete conversion (B_1, B_2, ..., B_k, ..., B_{kf}) given only the initial state of the system, the imposed pH (pH_0, pH_1, ..., pH_k, ..., pH_{kf}), and any in between addition of PenG or 6APA/PhAH (additions), as can be seen in Figure 5.

Balance equations

For the SGBM and WBM, the amount of added base was calculated using the charge balance:

$$[Na^+_{base}]_{k+1} = [POSNEG]_{k+1} + F^-_{Buf,k+1}\cdot[Buf]_{k+1} + F^-_{PenG,k+1}\cdot[PenG]_{k+1} + F^-_{6APA,k+1}\cdot[6APA]_{k+1} + F^-_{PhAH,k+1}\cdot[PhAH]_{k+1}, \quad (1)$$

in which $[Na^+_{base}]$ is the concentration of the sodium ions due to the addition of base. The counterions of the buffer and the penicillin, and the ions (Na^+ or Cl^-), which were used to bring the pH to its desired initial value, were taken together in the state variable POSNEG. The contribution of water and

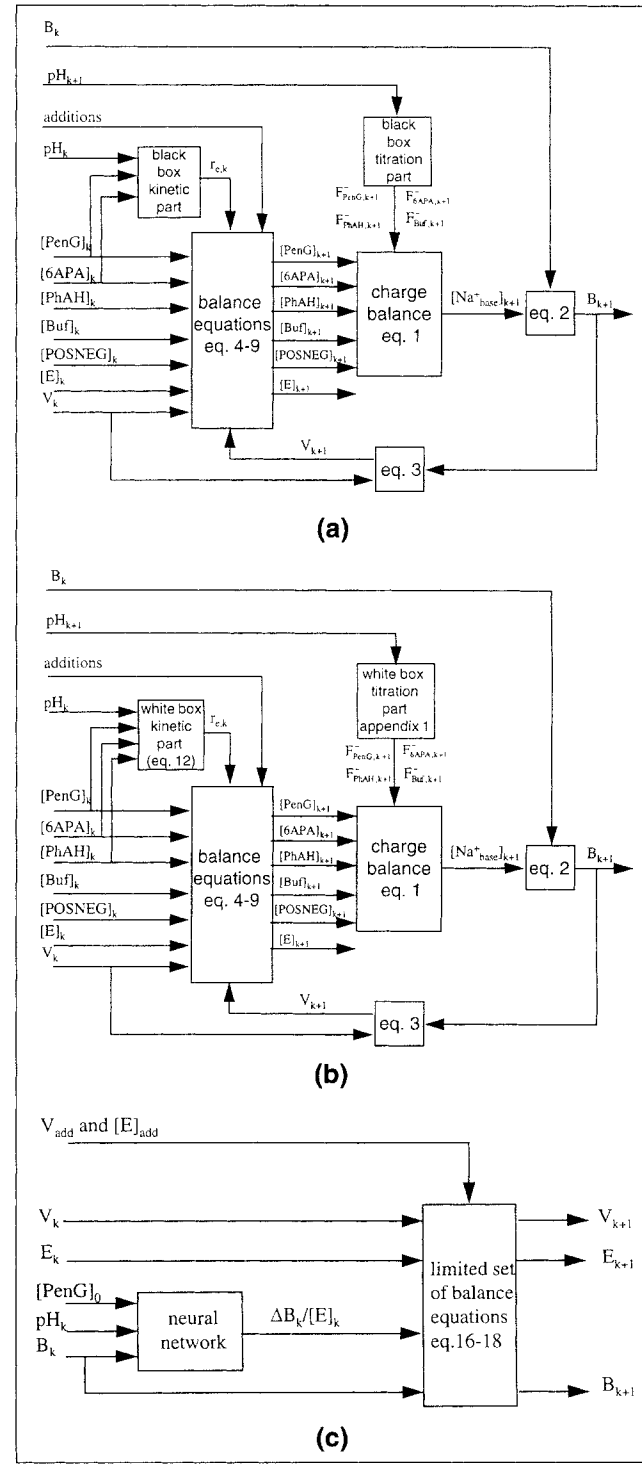


Figure 5. Structure and iterative calculation scheme for: (a) SGBM; (b) WBM; (c) BBM.

of the enzyme to the charge balance were neglected here. F^-_x is the net fraction of species x with a negative charge, which depends on the pH. These fractions are not restricted to values between 0 and 1, because components with a positive or multiple negative charge can be involved. The amount of added base (in mL) was directly related to $[Na^+_{base}]$, which was calculated from the charge balance:

$$B_{k+1} = \frac{[Na^+_{base}]_{k+1} \cdot V_{k+1}}{M_B}, \quad (2)$$

in which M_B is the molarity of the added base. In one time instant (from k to $k+1$), the volume changed due to the addition of base ($B_{k+1} - B_k$) and due to the eventual additions of buffered PenG or 6APA and PhAH solutions ($V_{k,add}$). Consequently:

$$V_{k+1} = V_k + (B_{k+1} - B_k) + V_{k,add}. \quad (3)$$

The dynamics of the state variables that were involved in the charge balance could be written down directly in the following balance equations (discrete form):

$$[PenG]_{k+1} = \frac{[PenG]_k \cdot V_k - [E]_k \cdot r_{e,k} \cdot \Delta t \cdot V_k + V_{k,add} \cdot [PenG]_{add}}{V_{k+1}} \quad (4)$$

$$[6APA]_{k+1} = \frac{[6APA]_k \cdot V_k + [E]_k \cdot r_{e,k} \cdot \Delta t \cdot V_k + V_{k,add} \cdot [6APA]_{add}}{V_{k+1}} \quad (5)$$

$$[PhAH]_{k+1} = \frac{[PhAH]_k \cdot V_k + [E]_k \cdot r_{e,k} \cdot \Delta t \cdot V_k + V_{k,add} \cdot [6PhAH]_{add}}{V_{k+1}} \quad (6)$$

$$[Buf]_{k+1} = \frac{[Buf]_k \cdot V_k + V_{k,add} \cdot [Buf]_{add}}{V_{k+1}} \quad (7)$$

$$[POSNEG]_{k+1} = \frac{[POSNEG]_k \cdot V_k + V_{k,add} \cdot [POSNEG]_{add}}{V_{k+1}} \quad (8)$$

$$[E]_{k+1} = \frac{[E]_k \cdot V_k + V_{k,add} \cdot [E]_{add}}{V_{k+1}}. \quad (9)$$

In the applied models enzyme deactivation (see Eq. 9) and ionic interactions are neglected. The balance equations (Eqs. 1, 3, 4–9) contain five terms that can be characterized as inaccurately known terms and that have a nonlinear relation with one or more system variables: the penicillin conversion-rate term (r_e), which will be referred to as the kinetic part of the model, and the four F_x^- terms, which are associated with acid–base equilibria of the involved compounds and which will be referred to as the titration curve part of the model. In the SGBM these five terms were parametrized by black-box model structures and in the WBM these terms were parametrized by white-box model structures.

r_e Term in the serial gray-box model

For (bio)chemical processes, the proper model structure for the conversion-rate term (r_e) is generally not directly known without separate experiments to reveal the relevant kinetic mechanisms. Therefore, in the SGBM the conversion rate was modeled with $r_{e,k} = f_1$ ([PenG]$_k$,[6APA]$_k$,[PhAh]$_k$,pH$_k$), which was parametrized with a neural network. In fact, the $r_{e,k}$ term was modeled as $r_{e,k} = f_1$([PenG]$_k$,[6APA]$_k$,pH$_k$), because in all identification and validation experiments the concentration 6APA was equal to the concentration phAH were equal, due to the fixed stoichiometry of the PenG conversion.

F_x^- Term in the serial gray-box model

The F_x^- terms in the SGBM were modeled in a black-box way. They were assumed to be dependent on the pH only:

$$F^-_{x,k} = f_2(pH_k). \quad (10)$$

For each component (Buf, PenG, 6APA, and PhAH) this dependence was measured by recording the titration curves (pH vs. added base) of each component. Each point in the titration curve (pH$_k$ and B_k) could be converted to a pH$_k$–$F^-_{x,k}$ point via the charge balance:

$$F^-_{x,k} = \frac{[POSNEG]_k + [Na^+_{base}]_k}{[X]_k}. \quad (11)$$

Given several pH$_k$–$F^-_{x,k}$ points for each component X, $F^-_{x,k}$ at nonmeasured pH values can be calculated by linear interpolation from the nearest higher and lower pH values.

r_e Term in the white-box model

In the WBM the parametrization of the r_e term was based on knowledge about the system. Although in general a proper model structure for the conversion rate r_e is not known *a priori* without separate experiments, for the system under consideration there have been several articles on the kinetics of the enzymatic conversion of penicillin G (excluding pH effects). Therefore, it was possible to construct a proper WBM without any separate experiments to determine a proper model structure for the conversion rate r_e (Warburton et al., 1973; Lee et al., 1982; Duan and Chen, 1996):

$$r_{e,k} = \frac{q_{max,k} \cdot \left([PenG]_k - \frac{[6APA]_k \cdot [PhAH]_k}{K^{app}_{eq,k}} \right)}{K_m + [PenG]_k + \frac{K_m \cdot [6APA]_k}{K_{i,6APA}} + \frac{K_m \cdot [PhAH]_k}{K_{i,PhAH}} + \frac{K_m \cdot [PhAH]_k \cdot [6APA]_k}{K_{i,PhAH} \cdot K_{i,6APA}} + \frac{[PenG]_k \cdot [6APA]_k}{K_{i,6APA}}}. \quad (12)$$

In this Michaelis Menten-type equation noncompetitive inhibition for 6APA and competitive inhibition for PhAH is incorporated. Substrate inhibition effects of PenG (the additional term $[PenG]^2/K_{i,PenG}$ in the denominator) can usually be neglected in the range of PenG concentrations, which were used in this work (< 100 mM). The pH dependence of the conversion rate was modeled by $q_{max,k} = f_3(pH_k)$. The pH dependence of the reaction equilibrium was taken into account by the pH dependence of the apparent equilibrium constant (K_{eq}^{app}), as is explained in detail in Appendix A.

F_x^- Term in the white-box model

In the WBM the F_x^- terms were modeled using knowledge about the system. From the acid–base equilibria of all compounds involved one can directly calculate the fraction of a particular component. For example, for the negatively charged fraction of phAH this results in

$$F_{PhAH}^- = \frac{1}{1 + \frac{[H^+]}{K_{a,PhAH}}}. \tag{13}$$

For the other fractions that were of interest to the system under consideration the formulas are given in Appendix A. The different $K_{a,x}$ values were estimated from the titration curves of the individual compounds, and they were compared with values from the literature.

Iterative calculations for the SGBM and WBM

The WBM and SGBM were validated by testing their ability to predict the base addition of a batch conversion (B_1, B_2, ..., B_k, ..., B_{kf}) given only the initial state of the system (B_0, $[PenG]_0$, $[6APA]_0$, $[PhAH]_0$, $[Buf]_0$, $[POSNEG]_0$, $[E]_0$, and V_0) the imposed pH (pH_0, pH_1, ..., pH_k, ..., pH_{kf}) and any in-between addition of PenG or 6APA/PhAH (additions). For this, an iterative calculation scheme was needed at each time step, as is visualized in Figure 5a and 5b. Given the state of the system at k and the conversion rate at k, the state at $k+1$ cannot directly be calculated, because the volume at $k+1$ is needed for the calculation of the concentrations at $k+1$. This volume at $k+1$ depends on the amount of added base and, in its turn, the amount of added base (and volume at $k+1$) depends on the concentrations at $k+1$ via the charge balance. Consequently an iterative scheme was needed to calculate the amount of added base, starting with an initial estimate for B_{k+1}. If the pH is constant, this iterative scheme can be avoided, because then a fixed stoichiometry between converted PenG and added base can be used.

Black-box model

The performance of the SGBM and WBM was compared with the performance of a BBM. For the system under consideration a complete BBM using a neural network, would look like:

$$B_{k+1} = f_{nn}([PenG]_0, pH_k, B_k). \tag{14}$$

This BBM structure, in which as little knowledge as possible about the system is used, is comparable with the BBM structure that was used successfully to model the PenG conversion at constant pH and temperature (Van Can et al., 1996b). However, the BBM proposed in Eq. 14 could not be identified with the available data that were used in this article for the identification of the SGBM and WBM, because the BBM (Eq. 14) cannot cope with different initial enzyme concentrations and with the in-between enzyme addition in the identification experiments. Therefore, we used a slightly different BBM in which a little knowledge was incorporated, so that the BBM could cope with different initial enzyme concentrations and the in-between enzyme additions. In this model the base addition per amount of enzyme ($\Delta B_k/[E]_k$) was predicted by a neural network on the basis of $[PenG]_0$, pH_k, and B_k (same inputs as Eq. 14), and combined with the balances for the enzyme concentration ($[E]_k$) and the volume (V_k):

$$B_{k+1} = B_k + \frac{\Delta B_k}{[E]_k} \cdot [E]_k \tag{15}$$

$$V_{k+1} = V_k + \frac{\Delta B_k}{[E]_k} \cdot [E]_k + V_{k,add} \tag{16}$$

$$[E]_{k+1} = \frac{[E]_k \cdot V_k + V_{k,add} \cdot [E]_{add}}{V_{k+1}} \tag{17}$$

$$\frac{\Delta B}{[E]_k} = f_{nn}([PenG]_0, pH_k, B_k). \tag{18}$$

The model that is expressed in Eqs. 15–18 is in fact a kind of gray-box model because a neural network (f_{nn}, Eq. 18) is combined with white-box balance equations (Eqs. 15–17). However, compared to the SGBM model used in this article, it is much more black box in nature, because significantly fewer detailed balance equations are used. Therefore, in this article Eqs. 15–18 are referred to as the BBM.

Experimental Setup for Model Identification and Validation

Table 1 summarizes the experiments that were used for model identification and validation. Two categories of identification experiments were used (first two rows in Table 1). In the first category four titration curves (of Buf, PenG, 6APA, and PhAH) were recorded and were used for the identification of the four F_x^- terms in the SGBM and the WBM. In the second category the base addition of 16 conversion experiments were recorded and used for the identification of the r_e term in the SGBM and WBM. Moreover, these 16 conversion experiments also used for the identification of the BBM. These conversion experiments started at a PenG concentration between 12.5 and 87.5 mM, and the pH was kept constant at a value between 5.5 and 8.5. In most of the 16 conversion experiments enzyme was added to check whether or not enzyme deactivation was present.

Six categories of validation experiments were performed (rows 3–8 in Table 1). In the first category the titration curve part of the model was tested with experiments in which no conversion took place. The titration curves of eight different Buf–PenG–6APA–PhAH mixtures, which were representative for the concentration ranges in this article, were recorded

Table 1. Experimental Conditions in the Identification and Validation Experiments

Type of Experiment	No.	Concentrations	pH	Additions
Identification of F_x^-	4	50.0 mM, except [6APA] = 12.5 mM	Gradually increasing 5.5–8.5	—
Identification of r_e term	16	$[PenG]_0$ = 12.5, 37.5, 62.5 or 87.5 mM	Constant at 5.5, 6.5, 7.5, or 8.5	Enzyme
Validation of F_x^-	8	$[PenG]_0$ = 0–100 mM $[6APA]_0 = [PhAH]_0$ = 0–100 mM	Gradually increasing 5.5–8.5	—
Validation of interpolation	9	$[PenG]_0$ = 25.0, 50.0 or 75.0 mM	Constant at 6.0, 7,0 or 8.0	Enzyme
Validation of range extrapolation	3	$[PenG]_0$ = 100 mM	Constant at 6.0, 7.0, or 8.0	—
Validation of dimensional extrapolation	2	$[PenG]_0$ = 30 or 35 mM	Constant at 6.0 or 8.0	PenG or 6APA + PhAH
Validation of frequency extrapolation	4	$[PenG]_0$ = 87.5 or 37.5 mM	Changing between 5.5 and 8.5	—
Validation of combined frequency and dimensional extrapolation	2	$[PenG]_0$ = 62.5 mM	Changing between 5.5 and 8.5	PenG

and compared with the predictions of the black-box titration curve part of the SGBM and with the predictions of the white-box titration curve part of the WBM. Moreover, the titration curve part of the model was also used to calculate the ratio between the amount of added base and the amount of converted PenG (OH-stoichiometry) at constant pH values between 5.5 and 8.5.

In category 2-5 of the validation experiments the SGBM was tested for one of the following properties: interpolation (IN), range extrapolation (RE), dimensional extrapolation (DE), frequency extrapolation (FE). Finally, in the sixth category the SGBM was tested for its combined dimensional and frequency extrapolation properties (FEDE).

Identification experiments started at an initial PenG concentration ($[PenG]_0$) between 12.5 and 87.5 mM and the pH in each experiment remained constant at a value between 5.5 and 8.5. Obviously, in the interpolation experiments, similar conditions were present so that no extrapolation was needed, although the actual pH and $[PenG]_0$ values differed from the ones in the identification experiments (Table 2). Also in the range extrapolation experiments similar conditions were present. Only, extrapolation in the $[PenG]_0$ range was tested, because the $[PenG]_0$ was 14% higher than the highest concentration in the identification experiments (Table 2).

In the frequency and dimensional extrapolation experiments the experimental conditions differed significantly from the conditions in the identification experiments. In the frequency extrapolation experiments the pH dynamics were totally different from the pH dynamics in the identification experiments: in the identification experiments the pH stayed constant at a value between 5.5 and 8.5, whereas in the frequency extrapolation experiments the pH varied significantly between 5.5 and 8.5 within each experiment. In the frequency extrapolation experiments the concentration PenG, 6APA, and PhAH, and the pH stayed within the ranges of the identification experiment so that no amplitude extrapolation was needed.

In the dimensional extrapolation experiments the pH dynamics were the same as in the identification experiments (constant at a value between 5.5 and 8.5 in each experiment so that no frequency extrapolation was needed), but during the conversion buffered solutions of either PenG or 6APA-PhAH were added. All identification experiments started with a volume of 1,000 mL and the volume changed only very little due to the addition of base (less than 10%). In contrast, the dimensional extrapolation experiments did not start at 1,000 mL but at 750 mL, and the volume changed significantly from 750 mL to 1,500 mL. Consequently, in the dimensional extrapolation experiments the model was tested to perform well in a 4-dimensional input space (PenG–6APA/PhAH –pH–volume), whereas the input space in the identification

Table 2. Initial PenG Concentrations and pH Values in the Identification (ID), Interpolation (IN), and Range Extrapolation (RE) Experiments

	$[PenG]_0$ (mM)							
pH	12.5	25.0	37.5	50.0	62.5	75.0	87.5	100.0
5.5	ID 1		ID 2		ID 3		ID 4	
6.0		IN 1		IN 2		IN 3		RE 1
6.5	ID 5		ID 6		ID 7		ID 8	
7.0		IN 4		IN 5		IN 6		RE 2
7.5	ID 9		ID 10		ID 11		ID 12	
8.0		IN 7		IN 8		IN 9		RE 3
8.5	ID 13		ID 14		ID 15		ID 16	

experiments was only 3-dimensional (PenG–6APA/PhAH–pH) because the volume was nearly constant at 1,000 mL. In the combined frequency–dimensional extrapolation experiments the pH changed as in the frequency-extrapolation experiments and the solutions of PenG were added as in the dimensional-extrapolation experiments.

Theoretical Aspects of Model Identification and Validation

In the identification and validation of dynamic models one should clearly distinguish between a one-step-ahead and a multistep-ahead use of the model. In the one-step ahead use of the model (Figure 6a), the model predictions of the future system state ($\hat{y}_{k+1}$) are based on the system input (u_k) and the *measured* current state of the system (y_k). In the multistep-ahead use of the model (Figure 6b), the predictions of the future system state ($\hat{y}_{k+1}$) are based on the system input (u_k) and the *predicted* current state of the system ($\hat{y}_k$). In this article the multistep-ahead use of the model is called free run if the model has to predict the base addition of a complete batch conversion given only the initial state of the system and the imposed pH. In the free run all intermediately calculated system states are based on formerly predicted system states, so that small model errors in one step can accumulate to larger errors in the base addition at the end of the conversion. The free run is considered to be a rigorous test for the predictive power of a dynamic model, and therefore in this article all models were validated in free run.

If the parameters of a kinetic term in a dynamic model have to be estimated, one should distinguish between the so-called differential and integral modes (which are closely related to the one-step-ahead and the free-run use of a model), in which a different sum of squared-errors (SSE) criterion is applied. The differential mode was used for the identification of the r_e term in the SGBM, and the integral mode was used for the identification of the r_e term in the WBM. In the SGBM and WBM the kinetic terms describe the static relation between the system variables and the conversion rate. In its turn, the conversion rate is used in the balance equations

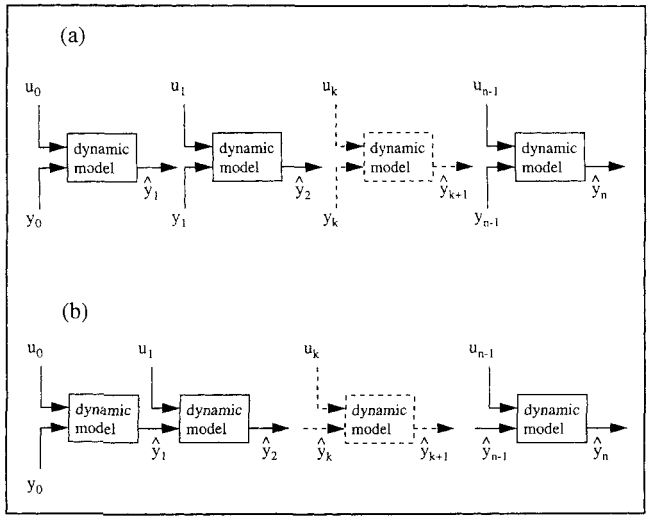


Figure 6. Dynamic model: (a) one-step-ahead use; (b) multistep-ahead use (free run).

to predict the future system state. In the differential mode, the parameters of the kinetic term are estimated by minimizing the difference between the measured conversion rate and the conversion rate that is predicted by only the kinetic term of the model. *Measured* states are used as inputs for the kinetic term. In this article, the performance criterion in the differential mode is denoted by SSE(r_e):

$$\mathrm{SSE}(r_e) = \frac{\sum_{i=1}^{N_e} (r_{e,i} - \hat{r}_{e,i})^2}{N_e} \tag{19}$$

If more than one experiment is used for identification, all data of the identification experiments are grouped together in one data set, for which SSE(r_e) is calculated. The total number of data pairs is then N_e. The differential mode is closely related to the one-step-ahead use of the model, because the predictions of the kinetic term ($\hat{r}_{e,k}$) are based on measured states of the system, and these predictions can be directly used in the balance equations to make a one-step-ahead prediction of the next system state.

In the integral mode, the parameters of the kinetic term ($r_{e,k}$) are estimated by minimizing the difference between the states of the system that are determined in the identification experiments, and their predictions in free run by the complete dynamic model. Consequently, *predicted* states are used as inputs for the kinetic term to intermediately calculate the conversion rate. In this article, the performance criterion in the integral mode is denoted by SSE(base):

$$\mathrm{SSE(base)} = \frac{\sum_{k=1}^{k_f} (B_k - \hat{B}_k)^2}{k_f} \tag{20}$$

in which k_f is the number of data points in the experiments for which the SSE(base) is calculated. If more than one experiment is used for identification, the SSE(base) of all identification experiments is summed. Obviously, the integral mode is closely related to the free run, because the predictions of system output are made in free run.

For the SGBM and WBM, first an attempt is made to estimate the parameters of the kinetic term in the differential mode, because the differential mode is computationally less intensive than the integral mode. If this results in a model with a good performance, no further effort has to be made to estimate the parameters. If it is expected that another, more optimal, parameter set will lead to a model with a better performance, an attempt is made to estimate the parameters in integral mode.

Results of the Model Identification

Identification of the F_x^- term in the SGBM

The identification of the F_x^- term in the SGBM and WBM was very straightforward. For each of the components Buf, PenG, 6APA, and PhAH one titration curve was recorded, which resulted in a pH vs. added base curve. For the SGBM these curves were directly converted to an F_x^-–pH curve by the charge balance at each point in the titration curve:

$$F_x^- = \frac{[\text{POSNEG}]_0 \cdot V_0 + M_B \cdot B_k}{[X]_0 \cdot V_0}. \quad (21)$$

At pH values that were not directly measured, the F_x^- value was calculated by linear interpolation between F_x^- values at the nearest higher and lower pH value. If the F_x^- term would depend in a complex nonlinear way on more than one state variable (not just pH), it could be worthwhile to consider a neural network for this relation.

Identification of the F_x^- term in the WBM

For the WBM the titration curves were used to estimate the K_a values of the involved components (Appendix A). In Table A1 (Appendix A), the estimated K_a values are compared with literature values. Most values were very much comparable, except for the lower K_a of 6APA. However, this K_a value was not too important, because the pH was never below 5.5.

Stoichiometry between converted PenG and added base (OH-stoichiometry)

For the remainder of the identification procedure, we needed the OH-stoichiometry (OH-stoich), which refers to the stoichiometry between the amount of converted PenG (mol) and the amount of added base (OH, mol). This stoichiometry depends on the the pH and could directly be calculated from the F_x^- terms. During the conversion the increase in negative charge should be compensated by the addition of base that contains the positively charged Na^+ ions. Consequently, the OH stoichiometry can be calculated as:

$$\text{OH-stoich(pH)} = F^-_{\text{6APA}}(\text{pH}) + F^-_{\text{PhAH}}(\text{pH}) - F^-_{\text{PenG}}(\text{pH}), \quad (22)$$

because PenG is converted to 6APA and PhAH. Table 3 gives the OH-stoich(pH) values at different pH values. The OH-stoich(pH), calculated with the black-box F_x^- terms and those calculated with the white-box F_x^- terms are very much comparable. In addition, OH-stoich(pH) was calculated with the measured added base and off-line HPLC analyses of PenG at t_f:

$$\text{OH-stoich(pH)} = \frac{B_{t_f} \cdot M_B}{[\text{PenG}]_0 \cdot V_0 - [\text{PenG}]_{\text{HPLC},t_f} \cdot V_{t_f}} \quad (23)$$

Unfortunately, the HPLC results were not accurate enough to fully verify the theoretical calculated OH-stoich(pH) (Eq. 22). Still, there is no reason to believe that the theoretical stoich(pH) values (black and white) were wrong, and they were therefore used in the remainder of the identification procedure.

Table 3. Comparison of OH-Stoich(pH) Values Calculated with the White-Box Model, with the Black-Box Model and from Measurements

pH	OH-Stoich(pH) Black Box	OH-Stoich(pH) White Box	OH-Stoich(pH) HPLC Eq. 23 + Std. Dev.
5.5	0.88	0.87	0.85 ± 0.02
6.0	0.95	0.96	0.98 ± 0.03
6.5	0.98	0.99	0.98 ± 0.02
7.0	0.99	1.00	0.99 ± 0.03
7.5	0.99	1.00	0.98 ± 0.02
8.0	1.00	1.00	0.99 ± 0.01
8.5	1.00	1.00	0.99 ± 0.01

Enzyme deactivation

Little is known about the stability of the enzyme penicillin acylase at pH values other than 8. Therefore, in 13 of 16 identification experiments (ID; Table 2) enzyme was added to check for possible enzyme deactivation. From these additions it could be concluded that in the pH range 5.5–8.5 and for the duration of the experiments in this article (maximum 3 h), enzyme deactivation was neglectable.

Identification of the r_e term in the SGBM

The pH and $[\text{PenG}]_0$ values in the identification experiments are given in Table 2. The identification experiments were equally spread over the pH range between 5.5 and 8.5, and over the $[\text{PenG}]_0$ range between 12.5 and 87.5 mM. For the calculation of the parameters of the neural network in the SGBM, we first had to calculate $[\text{PenG}]_k$, $[\text{6APA}]_k$ and $r_{e,k}$ from the available data. Appendix B explains in detail how this was done using the OH-stoich between converted PenG (mol) and added base (mol) (Eq. 22). With a sample interval of 60 s, the 16 IDs (Table 2) resulted in 1,600 data pairs, which were randomly divided into two sets: a training set of 1,200 data pairs (75%) for the calculation of the parameters; and a test set of 400 data pairs (25%) used for cross validation in the identification phase of the models. All data were scaled between -1 and 1 for the calculation of the parameters of the neural networks.

In the SGBM a neural network was used to describe the kinetics of the conversion. The number of hidden nodes and the related parameters had to be determined in the identification procedure. The identification for the neural network was straightforward, because the complexity of the neural network model was only increased stepwise by adding an additional neuron to the hidden layer. Moreover, the parameters of the neural network were estimated in the relatively simple differential mode. Thus, the parameters of neural networks, varying the in the number of hidden neurons from 1 to 25, were calculated and the resulting neural network models were tested on the test set (cross validation). When more than five neurons in the hidden layer were used the $\text{SSE}(r_e)$ in this cross-validation test no longer decreased. So, a neural network with five neurons in the hidden layer was chosen as the black-box component to describe the conversion kinetics in the SGBM.

Identification of the r_e term in the WBM

In the WBM a Michaelis Menten-type equation (Eq. 12) described the conversion kinetics. First, the relation between $q_{\max}$ and the pH $[q_{\max} = f_3(\text{pH}_k)]$ was determined. To find

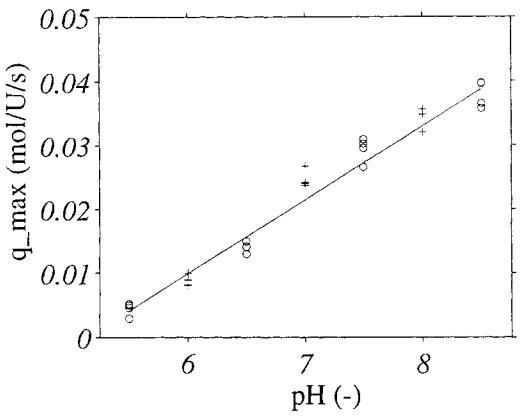


Figure 7. q_{max} values calculated with Eq. 24 compared with experimentally determined q_{max} values.

+ = Identification experiments, ○ = interpolation experiments.

this relation, the initial rate of each experiment was plotted against the pH in that experiment. The initial rate was equal to q_{max} because K_m was expected (and proved) to be much lower than the initial PenG concentration. From Figure 7, it is clear that there is a linear relation between q_{max} and pH. Therefore, $q_{max} = f_3(pH_k)$ was parametrized as

$$q_{max,k} = q_{max,0} + q_{max,r} \cdot pH_k. \quad (24)$$

Linear regression resulted in $q_{max,0} = -0.059 \pm 0.007$ (10^{-4} mol/U/s) and $q_{max} = 0.012 \pm 0.001$ (10^{-6} mol/U/s). After the initial estimation of $q_{max,0}$ and $q_{max,r}$, an attempt was made to estimate the six parameters of Eqs. 12 and 24 ($q_{max,0}$, $q_{max,r}$, K_m, and $K_{i,6APA}$, $K_{i,PhAH}$, and K_{eq}) in the differential mode. The initial parameter set consisted of $K_m = 3.26$ mM, $K_{i,6APA} = 104.2$ mM, $K_{i,PhAH} = 22.6$ mM (values from Duan and Chen, 1996), $K_{eq} = 4.75e\text{-}8$ M (Tewari and Goldberg, 1988), and of course $q_{max,0} = -0.0059$ mol/U/s and $q_{max,r} = 0.012$ mol/U/s (estimated from the initial conversion rates as was explained earlier). This initial parameter set already had a reasonable performance, but in the differential mode it did not converge to a parameter set that had a better performance [lower SSE(r_e)] in the cross-validation test on the 400 data pairs of the test set. Several attempts were made to estimate a subset of the parameters with only a subset of the data. However, none of these attempts resulted in a parameter set that had a better performance [lower SSE(r_e)] in the cross-validation test than the initial parameter set. It was therefore decided to estimate the parameters in the integral mode, starting with initial values cited earlier.

The parameters of Eqs. 12 and 24 were estimated in integral mode, starting with the initial parameter set as was just explained. Starting at a total SSE(base) on the identification experiments of 514 mL^2, we converged to a total SSE(base) of 174 mL^2 after 150 iterations. The final parameter values for the WBM are given in Table 4. Comparison with parameter values that were found for the same enzyme in prior work (Van Can et al., 1996b) is very difficult due to the large confidence intervals. Only the q_{max} at pH = 8.0 can be compared: the q_{max} in this article is 15% lower than the q_{max} determined in Van Can et al. (1996b). The (slow) enzyme deactivation during the storage of the enzyme is probably the reason for this because the experiments in this article were performed 8–10 months after the experiments described in Van Can et al. (1996b). A comparison of the parameter values found in this work and values found by others (Table 4) is not very useful, because in most cases no indication is given for the 95% confidence level of the estimated parameters, and the different origins of the used acylase might be a major source for different parameter values.

Table 4. Parameters (Incl. 95% Confidence Level) for the White-Box Model and Compared with Values Found by Others

Parameter	This Work	Van Can et al., 1996b	Values Found by Others
$q_{max,0}$ (1e-3 mmol/U/s)	67.2 ± 55	—	—
$q_{max,r}$ (1e-3 mmol/U/s)	12.5 ± 8	—	—
q_{max} at pH = 8 (1e − 3 mmol/U/s)	32.8	37.4 ± 0.8	—
K_M (mM)	0.6 ± 3	0.23 ± 0.5	0.01–1.1*
$K_{i,6APA}$ (mM)	109 ± 148	203 ± 45	0.24–40.9*
$K_{i,PhAH}$ (mM)	1.5 ± 0.9	0.9 ± 1.7	0.01–1.34*
K_{eq} (1e-5 mM)	12.9 ± 0.2	—	4.75**

*Gathered by Duan and Chen, 1996.
**Tewari and Goldberg, 1988.

Identification of the BBM

The BBM consisted of a neural network, which predicted the amount of added base per amount of enzyme, and of the balance equations for the enzyme concentration and volume (see Eqs. 15–18). The inputs ($[PenG]_0$, pH_k, B_k) for training the neural network were directly available from the identification experiments. The outputs ($\Delta B_k/[E]_k$), which were used for training the neural network, could be calculated straightforwardly from the identification experiments after the enzyme concentration at each time instant k ($[E]_k$) was calculated. The identification procedure for the neural network of the BBM was very much comparable with the identification of the neural-network part in the SGBM. Thus, the parameters of the neural network were estimated in the relatively simple differential mode and the same identification experiments were used (Table 2). Also here 1,200 data pairs were used for training the neural networks varying in number of hidden nodes from 1 to 20, and 400 data pairs were used for the cross-validation test of these neural networks. In this case SSE($\Delta B_k/[E]_k$) was minimized for calculation of the neural network parameters, and this criterion also indicated the performance of a neural network in the cross-validation test. After reaching eight neurons in the hidden layer the SSE($\Delta B_k/[E]_k$) in the cross-validation test no longer decreased. Therefore this neural network was chosen as the BBM.

Results of the Model Validation

Titration model

The titration-curve part of the SGBM and WBM was tested by comparing the predicted and measured titration curves of eight different Buf–PenG–6APA–PhAH mixtures going from pH 5.5 to pH 8.5. In these TITRA experiments no conversion

Table 5. Results of the Validation of the Titration Curve Part of the Model

Experiment	$[Buf]_0$ (mM)	$[PenG]_0$ (mM)	$[6APA]_0 = [PhAH]_0$ (mM)	SSE Black Box F_x^- Term	SSE White Box F_x^- Term
TITRA 1	50	0	50	0.46	0.34
TITRA 2	50	25	25	1.18	0.79
TITRA 3	50	50	0	0.92	1.50
TITRA 4	50	0	100	1.34	0.98
TITRA 5	50	25	75	1.49	1.50
TITRA 6	50	50	50	0.12	0.14
TITRA 7	50	75	25	1.32	0.90
TITRA 8	50	100	0	1.62	2.68
Total SSE				8.45	8.83

took place, and therefore $r_{e,k}$ and $[E]_k$ in the balance equations for the SGBM and WBM (Eqs. 1–9) were zero. Consequently, model predictions for the TITRA experiments relied on the correctness of the balance equations (Eqs. 1–8) and, of course, on the correctness of the F_x^- terms (black or white). The Buf, PenG, 6APA, and PhAH concentrations in the TITRA experiments were comparable with these concentrations in the other validation experiments in which conversion took place: the buffer concentration was always 50 mM, PenG concentration between 0 and 100 mM, and 6APA and PhAH were equal and between 0 and 100 mM.

Table 5 gives the performance of the titration-curve part of the SGBM and the WBM for the TITRA experiments. The performance was calculated as the SSEs between the measured and predicted base addition. In addition, for the SGBM Figure 8 compares the predicted and measured base addition for experiment TITRA 8. Given the performance of the SGBM model on TITRA 8 (Figure 8), and given the fact that the performance on the TITRA experiments is better (lower SSE in Table 5), it can be concluded that the black-box titration-curve part of the SGBM could accurately predict the base addition in titration curves in the relevant concentration range. The fact that the titration-curve part of the WBM was based on K_a values, which were also found by other researchers (see Table A1, Appendix A) and the fact that the titration-curve part of the SGBM and WBM had a comparably good performance (Table 5) provided an extra indication that the titration curve part was correct for both models.

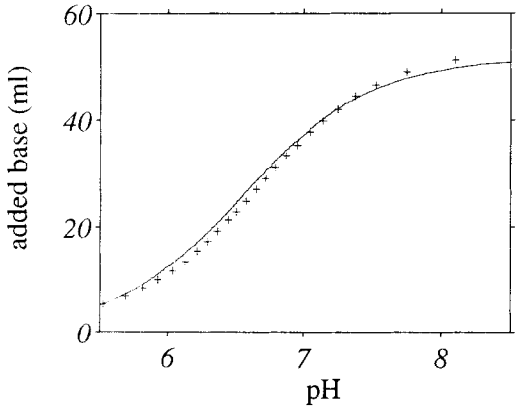


Figure 8. Performance of the black-box titration-curve part of the SGBM for titration curve TITRA 8.

Measured base addition (−) and predicted base addition (+).

Interpolation

The pH and $[PenG]_0$ values in the interpolation experiments (IN) are given in Table 2 and are compared with those values in the identification experiments (ID). It can be seen that the interpolation experiments were carried out at pH and $[PenG]_0$ values that were not used in the identification experiments, but that each interpolation experiment was surrounded by four identification experiments. The performance of the models was tested by predicting the base addition of the IN experiments in free run, which meant that only the initial state and the imposed pH were given to the model. All intermediate states were calculated. In Table 6 the performance [SSE(base)] of the three models (SGBM, WBM, and BBM) on the interpolation experiments is given. Furthermore, Figure 9 visualizes the performance of the SGBM on experiment IN 8. Given the performance of the SGBM for IN 8 in Figure 9 and given the fact that the performance of the other interpolation experiments was better [lower SSE(base) in Table 6], it can be concluded that the SGBM had good interpolation properties over the whole range of pH–$[PenG]_0$ values that were used in this article. Compared to the SGBM, the BBM resulted in slightly higher SSE(base) for most IN experiments. Still the BBM is considered to have good interpolation properties over the pH–$[PenG]_0$ range used, because no interpolation experiment was predicted very badly. In contrast, the performance of the WBM was clearly worse for some IN experiments, especially for the higher PenG concentration (IN 3, IN 6, and IN 9) and for two experiments at pH = 6.0 (IN 2 and IN 3). This poor performance of the WBM could already be expected from the relatively

Table 6. Results of the Validation of the Interpolation Properties of the SGBM, WBM, and BBM

Exp.	SSE (B_k) of SGBM	SSE (B_k) of WBM	SSE (B_k) of BBM
IN 1	0.49	3.20	1.85
IN 2	0.15	20.21	1.79
IN 3	0.28	43.77	1.37
IN 4	0.08	0.61	1.16
IN 5	2.22	2.01	3.77
IN 6	0.55	12.98	3.52
IN 7	0.82	1.31	0.57
IN 8	2.67	3.48	3.16
IN 9	1.86	17.38	2.45
Total SSE (B_k)	9.12	104.95	19.64

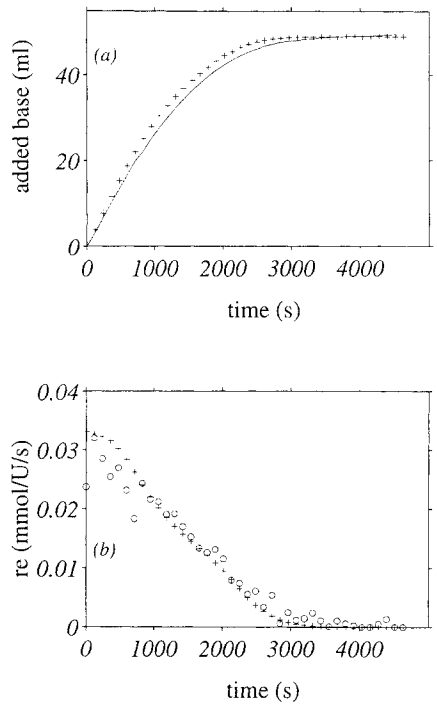


Figure 9. Performance of the SGBM for interpolation experiment IN 8.

(a) Base addition predicted by the SGBM (+) compared with the measured base addition (−); (b) the "measured" $r_{e,k}$, calculated from the base addition (○), compared with the $r_{e,k}$ predicted by the neural network (+).

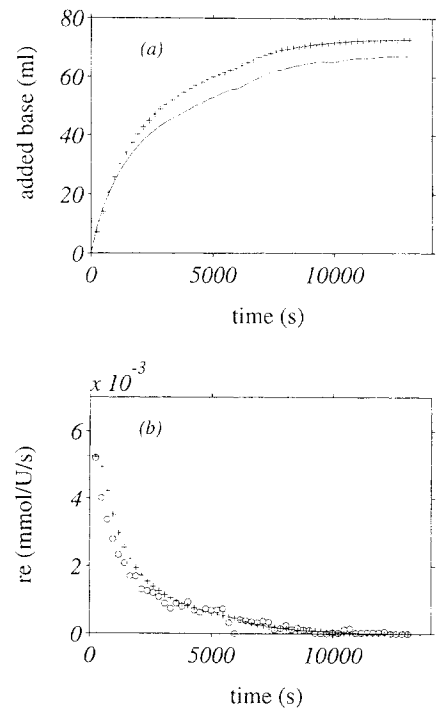


Figure 10. Performance of the SGBM for range-extrapolation experiment RE 1.

(a) Base addition predicted by the SGBM (+) compared with the measured base addition (−); (b) The "measured" $r_{e,k}$, calculated from the base addition (○), compared with the $r_{e,k}$ predicted by the neural network (+).

large SSE(base) that was obtained as a minimum after estimation of the parameters in integral mode. Clearly, in contrast to the SGBM and BBM, the WBM had no good interpolation properties over the whole range of pH–$[PenG]_0$ values that were used in this article.

Range extrapolation

In the range extrapolation experiments, the $[PenG]_0$ was 14% higher than the highest $[PenG]_0$ in the identification experiments (Table 2). In Table 7 the performance [SSE(base)] of the three models (SGBM, WBM, and BBM) on the range-extrapolation experiments is given. Furthermore, Figure 10 visualizes the performance of the SGBM for experiment RE 1. Although the range-extrapolation properties were clearly worse than the interpolation properties, the prediction of the SGBM and the WBM were not completely out of range. Moreover, in contrast to what might be expected, the WBM had no advantage over the SGBM. The predictions of the BBM for RE 2 and RE 3 were not so bad, but the prediction for RE 1 was very bad [resulting in a very high SSE(base)].

Table 7. Results of the Validation of the Dimensional (DE), Frequency (FE), and Combined Frequency/Dimensional (FEDE) Extrapolation Properties of the SGBM, WBM, and BBM

Exp.	SSE (B_k) of SGBM	SSE (B_k) of WBM	SSE (B_k) of BBM
RE 1	35.46	78.73	224.55
RE 2	14.42	4.64	8.51
RE 3	10.45	53.54	25.19
DE 1	12.93	39.52	—
DE 2	1.32	4.64	—
FE 1	25.93	7.8	87.0
FE 2	3.89	64.24	2260
FE 3	3.02	2.51	277.76
FE 4	3.44	13.03	317.37
FEDE 1	14.00	18.46	—
FEDE 2	30.86	52.43	—

Frequency and dimensional extrapolation

The experimental conditions in the dimensional-extrapolation (DE), frequency-extrapolation (FE) and combined frequency–dimensional (FEDE) experiments are given in Table 8. In the DE (and FEDE) experiments solutions of PenG or 6APA–PhAH were added to the reactor. Consequently, the model structures had to be able to deal with these additions. Moreover, in the DE experiments the volume changed over a much wider range (750–1,500 mL) than in the identification and interpolation experiments, in which the volume changed only a little due to the addition of base (1,000 mL to maximally 1,087 mL). In the frequency-extrapolation experiments (FE and FEDE) the pH was not constant at a value between 5.5 and 8.5, as in the identification and interpolation experiments, but it changed dynamically within each experiment. In Figure 11a this is shown for experiment FE 1 and in Figure 12a this is shown for experiment FEDE 2.

The BBM could not be used for the dimensional extrapolation experiments, because the inputs of the BBM provided

Table 8. Experimental Conditions in the Dimensional (DE), Frequency (FE), and Combined Frequency/Dimensional (FEDE) Extrapolation Experiments

Exp.	$[PenG]_0$	pH_0	pH vs. Time	Additions
DE 1	30	6.0	Constant	3×250 mL 60 mM PenG
DE 2	35	8.0	Constant	3×250 mL 65 mM 6APA + PhAH
FE 1	87.5	8.5	Figure 13a	—
FE 2	87.5	5.5	8.5 – 7.0 – 8.5	—
FE 3	37.5	7.0	6.5 – 8.5 – 8.0 – 8.5	—
FE 4	37.5	7.0	8.5 – 7.5 – 8.5	—
FEDE 1	62.5	8.5	6.5 – 8.5 – 7.5 – 8.5	2×250 mL 100 mM PenG
FEDE 2	62.5	5.5	Figure 14a	2×250 mL 100 mM PenG

no possibilities to deal with the addition of PenG or 6APA–PhAH. In contrast, for the frequency-extrapolation experiments, the BBM contained all the necessary inputs and even no extrapolation in the amplitude domain of these inputs was needed. However, using the BBM in the frequency-extrapolation experiments in fact means that the extrapolative nature of these experiments is not recognized. Evidently, this results in a bad performance, as can be seen in Table 7 [high SSE(base) for FE 1–4]. Consequently, the performance of the BBM for the FE experiments should be regarded as an absolute minimum performance that should be improved significantly by a model with good frequency-extrapolation properties. The performance of the SGBM, WBM, and BBM for the FE, DE, and FEDE experiments is given in Table 7. Indeed, the SGBM and the WBM did have a significantly better performance for the frequency extrapolation experiments than the BBM. For the SGBM Figure 13 visualizes the performance on experiment DE 1, Figure 11 visualizes the performance on experiments FE 1, and Figure 12 visualizes the performance on experiment FEDE 2. In these figures, the experiment with the worst performance for that type of experiment is displayed. Given the reasonable performance of the SGBM on experiments DE 1, FE 1, and FEDE 2 (Figures 11–13), and given the fact that it had a significantly better performance for the other experiment (Table 7), it can be concluded that the SGBM had reasonably good extrapolation properties. However, it is noticed that the extrapolation experiments are not predicted with the same accuracy as the interpolation experiments.

Discussion

For the presented case the serial gray-box modeling strategy was successful, especially showing reliable dimensional and frequency extrapolation to regions far outside the do-

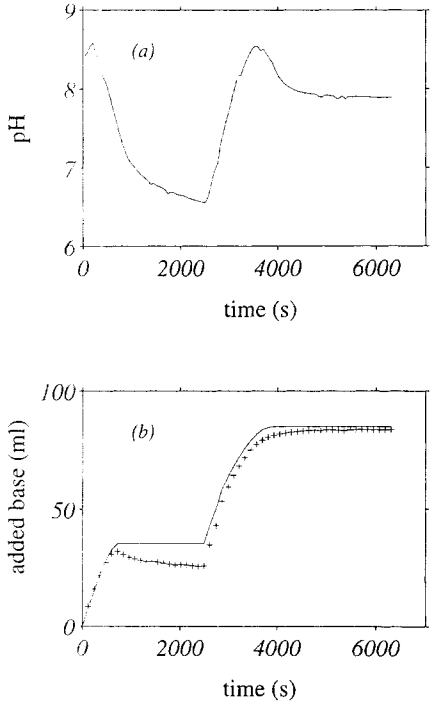


Figure 11. Performance of the SGBM for frequency-extrapolation experiment FE 1.

(a) The dynamic behavior of the pH that was used as input for the model; (b) Base addition predicted by the SGBM (+) compared with the measured base addition (–).

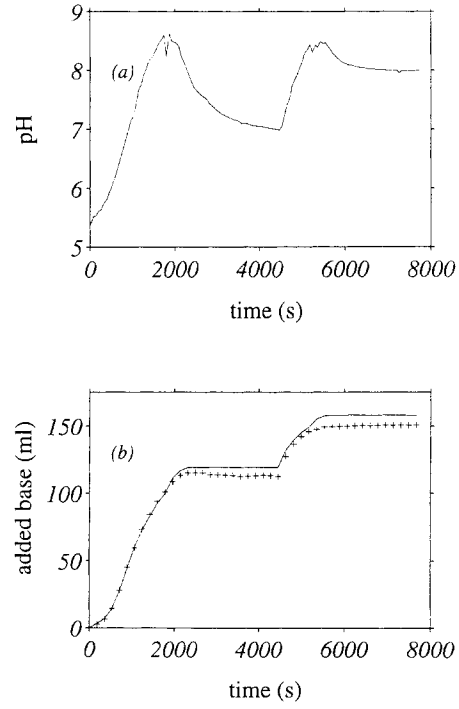


Figure 12. Performance of the SGBM for combined frequency/dimensional-extrapolation experiment FEDE 2.

(a) The dynamic behavior of the pH that was used as input of the model; (b) Base addition predicted by the SGBM (+) compared with the measured base addition (–).

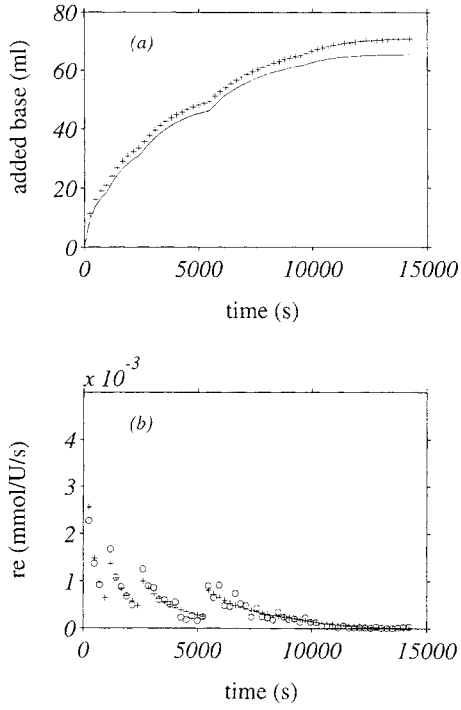


Figure 13. Performance of the SGBM for dimensional-extrapolation experiment DE 1.

(a) Base addition predicted by the SGBM (+) compared with the measured base addition (−); (b) The "measured" $r_{c,k}$, calculated from the base addition (○), compared with the $r_{c,k}$ predicted by the neural network (+).

main of the identification data. Of course, the correctness of the structure of the white-box part in the serial gray-box model is of key importance. The question for the more general case is whether or not suitable white-box knowledge is available. It has already been argued that the serial gray-box modeling strategy can be easily combined with the concept of macroscopic balances in (bio)chemical processes. These balances specify the dynamics of the relevant state variables and contain different rate terms that are associated with mechanisms such as transport and conversion. Some of these terms can be associated reliably with directly manipulated or measured variables (such as in- and outgoing flows) and, consequently, their mathematical relation with these variables is straightforward and no effort is needed to specify these relations. In that case these terms can be considered to be accurately known terms. In contrast, some rate terms (e.g., conversion kinetics) have a mathematical relation with one or more system variables that should be modeled in order to obtain a fully specified model. If these relations are modeled in a black-box way, for example, with a neural network, one obtains quite naturally the SGBM configuration. In such a configuration the identification data only have to cover the amplitude domain of the input–output space of the inaccurately known terms without taking into account the complete dimension and frequencies of the future application domain, which considerably reduces the number of identification experiments. The dimensional extrapolation relies on the accurately known term in the macroscopic balances, and the frequency extrapolation relies on the correct specification of the set of macroscopic balances, which means that all relevant rate terms should be static (time-independent).

The case presented here is a good example of this approach. The used macroscopic balances contained five inaccurately known terms (the four F_x^- terms and the conversion kinetics r_c). The complex static relation between the kinetic term and the state variables ([PenG], [6APA], and pH) was modeled by a neural network and the F_x^- terms were modeled in a simple black-box way (linear interpolation). The limited set of identification data covered the input–output space of the inaccurately known terms sufficiently so that no extrapolation on these terms was needed in the application phase of the model. In the dimensional-extrapolation experiments, the addition of PenG or 6APA–PhAH resulted in additional terms in the macroscopic balance (terms with subscripts added in Eqs. 4–9). Although the input–output range of these terms was not covered by the identification experiments, they could be used for reliable dimensional extrapolation, because they could be considered as accurately known terms.

In the identification experiments the pH was used at only one frequency because in each experiment the pH was constant at a value between 5.5 and 8.5, whereas in the frequency-extrapolation experiments it was used at other frequencies because the pH changed dynamically between 5.5 and 8.5 in each experiment. Nevertheless, the model could be used for reliable frequency extrapolation, because the set of macroscopic balances was specified correctly (all relevant rate terms incorporated) and the identification data covered the amplitude domain of the inaccurately known rate terms sufficiently. If, however, the set of macroscopic balances is not specified correctly, frequency extrapolation will be unreliable. For example, an ill-specified set of macroscopic balances in which one F_x^- term is neglected will not result in accurate frequency extrapolation, although this knowledge is not necessary to obtain good interpolation properties (see BBM).

It is believed that the here-presented serial gray-box modeling strategy can be applied to a wide range of (bio)chemical processes, because in (bio)chemical engineering dynamic models are always based on macroscopic balances, and often information is available to choose the inaccurately known terms that should be modeled in the black-box way. However, two warnings are in place here, especially with respect to frequency extrapolation. First, in the serial gray-box modeling strategy, the correct specification of the set of macroscopic balances is essential, which means that all relevant rate terms should be incorporated in the macroscopic balances, and that the relations modeled by the neural network are static (time-independent). Given only a limited amount of knowledge about the application system, one might not be able to specify the correct set of macroscopic balances. In that case it might be necessary to model dynamic relations with the neutral network in order to obtain an acceptable fit to the identification data, which means that not only current system states but also former system states are used as input for the neural network. Although the resulting model can have good interpolation and dimensional-extrapolation properties, it becomes doubtful whether it has good frequency-extrapolation properties. Second, in the serial gray-box modeling strategy, the identification data have to cover the amplitude do-

main of *all* relevant, inaccurately known rate terms. This means that one should be careful with identification experiments under mild dynamic conditions, if frequency extrapolation is needed. Under mild dynamic conditions a mechanism and its associated rate term in the macroscopic balances might be neglected, whereas under more dynamic conditions (e.g., when frequency extrapolation is required) this mechanism and its associated rate term cannot be neglected. Therefore, taking into account these two warnings, it is advisable to validate potential frequency-extrapolation properties with at least one or a few experiments in which dynamic conditions relevant for the application of the model are present.

The performance of the SGBM in the range of extrapolation experiments was worse than its performance on the interpolation experiments, although its predictions were not completely out of range. Moreover, its performance was comparable with the performance of the WBM (see Table 7). In the serial gray-box modeling strategy the identification data cover the input domain of the black-box part of the model. Consequently, range extrapolation of an SGBM is always related to the black-box part of the model, and an SGBM should therefore be applied only with great care in range-extrapolation situations. It is believed that WBMs generally have good range-extrapolation properties and are therefore potentially advantageous in this respect. However, this is only true when the parameters of the WBM are carefully identified, including a thorough experimental design program. This is often not the case. Moreover, valuable statistical information from the identification phase (domain of the identification data, confidence interval, correlation coefficients) is often absent or gets lost when after a while the model and its parameters are applied in situations that differ from the situations in the identification phase. Therefore, in practice WBMs should also be applied with great care in range-extrapolation situations, and compared to SGBMs they are not necessarily advantageous in this respect.

In this article, the performance of the SGBM was compared with the performance of a WBM. With respect to interpolation properties, the SGBM was more accurate, and with respect to extrapolation properties they had a more or less comparable performance. Both models were identified with the same data, but for the WBM significantly more knowledge was needed to construct the model, such as complicated kinetics (Eq. 12) and equilibrium thermodynamics (Appendix A). Moreover, instead of relatively simple differential mode (used for the BBM and SGBM) the computationally more complex integral mode was needed to estimate the parameters of the WBM. The performance of the WBM might be improved by incorporating more detailed knowledge about the system, for example, making the inhibition parameters ($K_{i,6APA}$ and $K_{i,PhAH}$) pH dependent or using the concentration of charged PenG, 6APA, or PhAH components that bind to the enzyme instead of overall concentrations. However, this exactly indicates the advantages of an SGBM: also in a situation where there is no detailed knowledge of the system, the serial gray-box modeling strategy can lead efficiently to a model with a good performance. Of course, this does not mean that *a priori* WBMs should be neglected when a model is required. In practice, it is often possible to construct a first version of a WBM with relatively limited effort. If this simple first version is sufficiently accurate for the intended purpose, no alternative modeling approaches have to be considered. However, if this first version is not accurate enough, a serial gray-box modeling strategy is an attractive alternative to a white-box strategy.

In this article, the performance of the SGBM was also compared with the performance of a BBM. With respect to interpolation properties, the two models had a comparable performance, but for dimensional and frequency extrapolation, the BBM could not be used, whereas the SGBM had a good performance. Only four additional experiments (four titration curves) were needed to obtain these extrapolation properties. The BBM structure provided no possibility of dealing with additions of PenG or 6APA–PhAH, and therefore the BBM had no dimensional extrapolation properties. Many identification experiments with all kinds of additions would be needed to identify a BBM that could deal with additions as in the dimensional-extrapolation experiments. Moreover, the BBM structure provided no possibility of using the information of the titration curves, and therefore the BBM was not reliable for frequency extrapolation. Many identification experiments with many pH profiles would be needed to identify a BBM that could deal with pH profiles as used in the frequency-extrapolation experiments.

The lack of good extrapolation properties of BBMs means that a black-box modeling strategy is only advantageous when the application domain of the model can be restricted *a priori* and when the input variables can be varied freely to generate identification data that cover the application domain sufficiently. Typically, an advantageous situation for black-box modeling will occur for already existing processes at the lowest level of process operation, which is mainly concerned with accurate setpoint tracking. At this level the domain is restricted by the situation at hand (given equipment and operating ranges) and process input variations are allowed, because these also happen when the process has to be controlled. At the medium level of process operation, which is mainly concerned with determination of optimal setpoint trajectories, an advantageous situation for black-box modeling will only occur when the routine data acquisition of daily processes provides sufficiently rich identification sets.

More typically, at the medium level of process operation an advantageous situation for black-box modeling will not occur and many experiments with many variations at the actual (large-scale) process are needed to generate the necessary data for black-box modeling. This might be unwanted or even impossible in a production environment. However, for SGBMs the experiments at the actual full scale might not be needed (or be reduced considerably) because of the reliable dimensional extrapolation properties of SGBMs, or the number of experiments at the actual full-scale process, and especially the variations to be applied in these experiments might be reduced considerably because of the reliable frequency-extrapolation properties of SGBMs. In conclusion, especially at the medium level of process operation the properties of serial gray-box modeling strategy are expected to show to full advantage.

Conclusions

For the presented case of an enzymatically catalyzed reaction, it was shown that in the serial gray-box strategy a model could be constructed efficiently by using a straightforward

balance equation in which the inaccurately known terms could be modeled in a black-box way. Neural networks provided a convenient tool to model the complex relation between the system variables and conversion kinetics. The identification data only had to cover the input space of the inaccurately known terms in the macroscopic balances. The accurately known terms could be used for reliable dimensional extrapolation, and the correct balance equations provided a good basis for reliable frequency extrapolation. Therefore, the SGBM had a short development time and showed good interpolation, and frequency and dimensional extrapolation properties, which were tested in several experiments.

Compared to a more knowledge-driven white-box strategy, the SGBM structure is only based on readily available or easily obtainable knowledge, so that the development time of SGBMs will also be short in a situation where there is no detailed knowledge of the system available. Compared to a more data-driven black-box strategy, the serial gray-box strategy leads to models with good frequency and dimensional-extrapolation properties, so that with the same number of identification experiments the model can be applied to a much wider range of conditions. Especially at the medium level of process operation, which is concerned with the calculation of optimal setpoint trajectories for the process, the properties of the SGBM are expected to show to full advantage.

Acknowledgments

The authors thank Gist Brocades N.V. for providing the Penicillin G and the Penicillin acylase. A. J. Krijgsman, J. M. Schuffelen, and H. B. Verbruggen are thanked for valuable discussion. J. Houwers and C. Ras are thanked for their experimental support in this project. A. J. J. Straathof and L. A. M. van der Wielen are thanked for sharing with us their expertise on the penicillin case.

Notation

add = subscript denoting external addition
[Buf] = concentration buffer, mol/L
f_x = fraction of a charged species of component X (Appendix A)
$[H^+]$ = concentration H^+
k_f = last time instant of an experiment
$K_{x,a}$ = dissociation constant of component X, mol/L
K_m = Michaelis Menten coefficient, mol/L
q_{max} = maximum enzyme-specific conversion rate, mol/L/s
Δt = time interval, s

Literature Cited

Albiol, J., C. Campmajo, C. Casas, and M. Poch, "Biomass Estimation in Plant Cell Cultures: A Neural Network Approach," *Biotechnol. Prog.*, **11**, 88 (1995).

Berezin, I. V., A. A. Klesov, A. L. Margolin, P.C. Neis, and V. K. Savirskaya, *Antibiotiki*, **21**, 519 (1976).

Bhat, N., and T. J. McAvoy, "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems," *Comput. Chem. Eng.*, **14**, 573 (1990).

Chang. J.-S., and W.-Y. Hseih, "Optimization and Control of Semibatch Reactors," *Ind. Eng. Chem. Res.*, **34**, 545 (1995).

Côte, M., B. P. A. Grandjean, P. Lessard, and J. Thibault, "Dynamic Modeling of the Activated Sludge Process: Improving Prediction Using Neural Networks," *Water Res.*, **29**, 995 (1995).

Dors, M., R. Simutis, and A. Lübbert, "Advanced Supervision of Mammalian Cell Cultures Using Hybrid Process Models," *Preprints of the 6th International Conference on Computer Applications in Biotechnology*, A. Munack and K. Schügerl, eds., p. 72 (1995).

Duan, G., and J. Y. Chen, "Kinetic Analysis of the Effect of Product Removal on the Hydrolysis of Penicillin G by Immobilised Penicillin Acylase," *Process Biochem.*, **31**, 27 (1996).

Lee, S. B., D. Dewey, and D. Y. Ryu, "Reaction Kinetics and Mechanisms of Penicillin Amidase: A Comparative Study by Computer Simulation," *Enzyme Microb. Technol.*, **4**, 35 (1982).

Lee, J., and W. F. Ramirez, "Optimal Control of Induced Foreign Protein Production by Recombinant Bacteria in Fed-Batch Reactors," *AIChE J.*, **40**, 899 (1994).

Lopez, T. I., and F. X. Malcata, "Optimum Temperature Policy for Batch Reactors Performing Biochemical Reactions in the Presence of Enzyme Deactivation," *Bioprocess Eng.*, **9**, 129 (1993).

Modak, J. M., and H. C. Lim, "Optimal Operation of Fed-Batch Bioreactors with Two Control Variables," *Chem. Eng. J.*, **42**, B15 (1989).

Montague, G., and J. Morris, "Neural-Network Contributions in Biotechnology," *Trends Biotechnol.*, **12**, 312 (1994).

Psichogios, D. C., and L. H. Ungar, "Direct and Indirect Model Based Control Using Artificial Neural Networks," *Ind. Eng. Chem. Res.*, **30**, 2564 (1991).

Psichogios, D. C., and L. H. Ungar, "A Hybrid Neural Network-First Principles Approach to Process Modeling," *AIChE J.*, **38**, 1499 (1992).

Schubert, J., R. Simutis, M. Dors, I. Havlik, and A. Lübbert, "Bioprocess Optimization and Control: Application of Hybrid Modeling," *J. Biotechnol.*, **35**, 51 (1994).

Shioya, S., "Optimization and Control in Fed-Batch Bioreactors," *Adv. Biochem. Eng.-Biotechnol.*, **46**, 112 (1992).

Simutis, R., I. Havlik, M. Dors, and A. Lübbert, "Artificial Neural Networks of Improved Reliability for Industrial Process Supervision," *Preprints of the 6th International Conference on Computer Applications in Biotechnology Sarmisch Parten Kirchen*, A. Munack and K. Schügerl, eds., p. 59 (1995).

Smolyakov, V. S., and M. P. Primanchuk, *Russ. J. Phys. Chem.*, **40**, 463 (1966).

Su, H.-T., N. Bath, P. A. Minderman, and T. J. McAvoy, "Integrating Neural Networks with First Principles Models for Dynamic Modeling," IFAC Dynamics and Control of Chemical Reactors, Maryland, p. 327 (1992).

Te Braake, H. A. B., H. J. L. van Can, G. van Straten, and H. B. Verbruggen, "Regulated Activation Weights Neural Network (RAWN)," *ESANN*, Brugge, Belgium, p. 19 (1996).

Tewari, Y. B., and R. N. Goldberg, "Thermodynamics of the Conversion of Penicillin G to Phenylacetic Acid and 6-Aminopenicillanic Acid," *Biophys. Chem.*, **29**, 245 (1988).

Thibault, J., V. C. van Breusegem, and A. Cheruy, "On-Line Prediction of Fermentation Variables Using Neural Networks," *Biotechnol. Bioeng.*, **43**, 1041 (1990).

Thompson, M. L., and M. A. Kramer, "Modeling Chemical Processes Using Prior Knowledge and Neural Networks," *AIChE J.*, **40**, 1328 (1994).

Tholudur, A., and F. Ramirez, "Optimization of Fed-Batch Bioreactors Using Neural Network Parameter Function Models," *Biotechnol. Prog.*, **12**, 302 (1996).

Uhlemann, J., M. Cabassud, M. Le Lann, E. Borredon, and G. Casamatta, "Semi-Batch Reactor Optimization and Control for the Epoxidation of Furfural," *Chem. Eng. Sci.*, **49**, 3169 (1994).

Van Can, H. J. L., H. A. B. te Braake, C. Hellinga, K. Ch. A. M. Luyben, and J. J. Heijnen, "A Strategy for Dynamic Process Modeling Based on Neural Networks in Macroscopic Balances," *AIChE J.*, **42**, 3403 (1996a).

Van Can, H. J. L., H. A. B. te Braake, C. Hellinga, K. Ch. A. M. Luyben, and J. J. Heijnen, "An Efficient Model Development Strategy for Bioprocesses Based on Neural Networks in Macroscopic Balances: Part I," *Biotechnol. Bioeng.*, **54**, 549 (1997).

Van der Wielen, L. A. M., "A Countercurrent Adsorptive Fluidized Bed Reactor for Heterogeneous Bioconversions," PhD Thesis, Delft Univ. of Technology, Delft, The Netherlands (1997).

Warburton, D., P. Dunnill, and M. D. Lilly, "Conversion of Benzylpenicillin to 6-Aminopenicillanic Acid in a Batch Reactor and Continuous Feed Stirred Tank Reactor Using Immobilized Penicillin Amidase," *Biotechnol. Bioeng.*, **15**, 13 (1973).

Zhang, Q., F. R. Reid, J. B. Lichtfield, J. Ren, and S.-W. Chang, "A Prototype Neural Network Supervised Control System for *Bacillus thuringiensis* Fermentations," *Biotechnol. Bioeng.*, **43**, 483 (1994).

Zhu, Y.-H., T. Rajalahti, and S. Linko, "Application of Neural Networks to Lysine Production," *Chem. Eng. J.*, **62**, 207 (1996).

Table A1. Dissociation Constants (at 310 K) for all Relevant Components

Species	This Article	Literature
Phosphate (Buf)	6.7	6.7 (Van der Wielen, 1997)
PhAH	4.2	4.2 (Van der Wielen, 1997) 4.3 (Smoyakov et al., 1966)
PenG	2.7	2.6 (Van der Wielen, 1997) 2.6 (Tewari et al., 1988)
$6APA_1$	4.4	4.4 (Van der Wielen, 1997) 5.4 (Tewari and Goldberg, 1988) 4.6 (Berezin et al., 1976)
$6APA_2$	1.8	2.6 (Van der Wielen, 1996) 2.6 (Tewari et al., 1988)

Appendix A: Application of Equilibrium Thermodynamics to the PenG Conversion

In this appendix equilibria thermodynamics are applied to the PenG conversion. Only the resulting equations are presented. For a more comprehensive description, refer to Tewari and Goldberg (1988) and to Van der Wielen (1997).

For the system under consideration the following acid–base equilibria are involved:

$$\begin{aligned}
&\mathrm{HPenG} \overset{K_{a,\mathrm{PenG}}}{\leftrightarrow} \mathrm{PenG^-} + \mathrm{H^+}\\
&6\mathrm{APAH^{\pm}} \overset{K_{a1,\mathrm{APA}}}{\leftrightarrow} 6\mathrm{APA^-} + \mathrm{H^+}\\
&6\mathrm{APAH_2^+} \overset{K_{a2,6\mathrm{APA}}}{\leftrightarrow} 6\mathrm{APAH^{\pm}} + \mathrm{H^+}\\
&\mathrm{PhAH} \overset{K_{a,\mathrm{PhAH}}}{\leftrightarrow} \mathrm{PhA^-} + \mathrm{H^+}\\
&\mathrm{H_2PO_4^-}\ \mathrm{H_2PO_4^{-2}} \overset{K_{a,\mathrm{Buf}}}{\leftrightarrow} \mathrm{HPO_4^{-2}} + \mathrm{H^+}. \qquad \text{(A1)}
\end{aligned}$$

The overall negative fraction (F_x^-) of PenG, 6APA, PhAH, and Buf is calculated from the fraction of the individual species (f_x) of those components that have a charge, resulting in the following equations:

$$F^-_{\mathrm{PenG}} = f_{\mathrm{PenG^-}} = \frac{1}{1+\dfrac{[H^+]}{K_{a,\mathrm{PenG}}}} \qquad \text{(A2)}$$

$$\begin{aligned}
F^-_{6\mathrm{APA}} &= f_{6\mathrm{APA^-}} - f_{6\mathrm{APAH_2^+}}\\
&= \frac{1}{1+\dfrac{[H^+]^2}{K_{a,\mathrm{APAH}}\cdot K_{a,\mathrm{APAH2}}}+\dfrac{[H^+]}{K_{a,\mathrm{APAH}}}}\\
&\quad - \frac{1}{\dfrac{K_{a,\mathrm{APAH}}\cdot K_{a,\mathrm{APAH2}}}{[H^+]^2}+1+\dfrac{K_{a,\mathrm{APAH2}}}{[H^+]}} \qquad \text{(A3)}
\end{aligned}$$

$$F^-_{\mathrm{PhAH}} = f_{\mathrm{PhA^-}} = \frac{1}{1+\dfrac{[H^+]}{K_{a,\mathrm{PhAH}}}} \qquad \text{(A4)}$$

$$\begin{aligned}
F^-_{\mathrm{Buf}} &= 2\cdot f_{\mathrm{H_2PO_4^{-2}}} + f_{\mathrm{H_2PO_4^-}}\\
&= 2\cdot\frac{1}{1+\dfrac{[H^+]}{K_{a,\mathrm{Buf}}}} + \frac{1}{1+\dfrac{K_{a,\mathrm{Buf}}}{[H^+]}}. \qquad \text{(A5)}
\end{aligned}$$

If $\mathrm{PenG^-} \leftrightarrow 6\mathrm{APA^-} + \mathrm{PhA^-} + \mathrm{H^+}$ is taken as the reference reaction, the apparent equilibrium constant can be calculated as:

$$K_{\mathrm{eq}}^{\mathrm{app}} = \frac{K_{\mathrm{eq}}\cdot f_{\mathrm{PenG^-}}}{[H^+]\cdot f_{6\mathrm{APA^-}}\cdot f_{\mathrm{PhA^-}}}. \qquad \text{(A6)}$$

Appendix B: Calculation of $r_{e,k}$ and Concentrations from the Measured Base Addition

Starting with the known initial condition (V_0, $[E]_0$, $[\mathrm{PenG}]_0$, $[6\mathrm{APA}]_0$, and $[\mathrm{PhAH}]_0$) and the measured base addition (B_k), the state of the system could be reconstructed based on the stoichiometry of PenG:6APA:$\mathrm{PhAH^-}$ = 1:1:1:1 and the OH-stoichiometry that is pH-dependent (Table 2):

$$V_{k+1} = V_0 + B_k \qquad \text{(B1)}$$

$$[\mathrm{PenG}]_{k+1} = \frac{[\mathrm{PenG}]_0\cdot V_0 - \dfrac{M_B\cdot B_k}{\text{OH-stoich(pH)}}}{V_k} \qquad \text{(B2)}$$

$$[6\mathrm{APA}]_{k+1} = \frac{[6\mathrm{APA}]_0\cdot V_0 + \dfrac{M_B\cdot B_k}{\text{OH-stoich(pH)}}}{V_k} \qquad \text{(B3)}$$

$$[\mathrm{PhAH}]_{k+1} = \frac{[\mathrm{PhAH}]_0\cdot V_0 + \dfrac{M_B\cdot B_k}{\text{OH-stoich(pH)}}}{V_k} \qquad \text{(B4)}$$

$$[E]_{k+1} = \frac{[E]_k\cdot V_k}{V_{k+1}}. \qquad \text{(B5)}$$

The conversion rate at each time instant k was calculated as

$$r_{e,k} = \frac{M_B\cdot([\mathrm{PenG}]_{k+1} - [\mathrm{PenG}]_k)}{[E]_k\cdot V_k\cdot\Delta t}. \qquad \text{(B6)}$$

Manuscript received Mar. 21, 1997.