Environ. Sci. Technol. **2002**, *36,* 715−723

Automatic Buffer Capacity Model Building for Advanced Interpretation of Titration Curves

M. VAN DE STEENE, L. VAN VOOREN, J.-P. OTTOY, AND P. A. VANROLLEGHEM*

BIOMATH, Ghent University, Coupure Links 653, 9000 Gent, Belgium

An automatic buffer capacity model building algorithm was developed. The objective of this algorithm is to characterize multiple buffer systems from experimental titration curves. Starting from a simple mathematical model that incorporates the available prior knowledge on the buffer system, the model is sequentially extended by incorporating additional monoprotic buffer equations. To select the most appropriate model from the resulting set of automatically built models, a model structure selection technique was constructed from existing methods.

Introduction

Methods and applications based on pH titrations are used in a wide variety of fields (aerobic, anaerobic and physico-chemical wastewater treatment, food and feed applications, soil science, microbiology, aquatic chemistry, ...) (*1*−*5*). However, these applications mostly rely on the offline interpretation of titration curves and can thus not be considered "sensors". Sensors making use of pH titration curves are often referred to as titrimetric sensors. These sensors mostly work with only a few titration points and a simplified and robust data interpretation method. One of the main application fields in this area is the control of anaerobic digestion where bicarbonate and/or volatile fatty acids (VFAs) can be monitored with a titrimetric sensor (*6*−*16*).

The buffer capacity based sensor of this paper differentiates itself from the other sensors by the fact that the whole and detailed titration profile is used for model-based interpretation (software sensing). As a consequence, it is aimed to differentiate and quantify more individual buffers as compared to the application approaches mentioned above. The hardware part of the sensor described in this paper consists of a titrator unit, capable of performing acid−base titrations of aquatic samples. The raw data consists of a titration curve that is obtained by adding consecutive small amounts of NaOH to the sample and measuring the pH after each addition. A titration curve has a typical S-shape and can be transformed into a buffer capacity profile with an appropriate mathematical algorithm, in which the buffer capacity in each point of the titration curve is calculated as the derivative of the amount of base needed for a pH increase of one pH unit. The main contribution of this work is the advanced data processing of the calculated buffer capacity profiles. Therefore, the term "software sensor" originally introduced in ref *17* is applicable to this type of sensor. The most advanced data interpretation of buffer capacity profiles found in the literature is a stepwise model building approach to construct a suitable mathematical model for an unknown solution (*18, 19*). The method is based on a monoprotic approach, in which titration of a polyprotic acid can be rigorously represented as a mixture of monoprotic acids. In the proposed method, a one component model, a two component model, and so on are fitted to titration points; and this is carried out to the point where further resolution becomes meaningless. Unfortunately, the methods described in refs *18* and *19* were developed 15 years ago, and no recent literature based on this approach was found.

The buffer capacity sensor has first been evaluated for water quality monitoring (*20, 21*). Second, a manual model building algorithm for buffer capacity profiles of well-characterized samples was described (*22*). Later, this model building algorithm was modified, automated, and extended with model selection criteria (*23*). Last, the automatic model building algorithm was evaluated and implemented for the purpose of water quality monitoring (*24*).

The objective of this paper is to present and evaluate a generalized buffer capacity model building algorithm applicable for advanced interpretation of a wide variety of titration curves. It will be illustrated with experimental data of a number of selected samples that the proposed algorithm is sufficiently robust and suitable for on-line buffer characterization and quantification. The reader is also referred to (*25*) for additional information and interpretations of the results presented in this paper.

Methodology

General Linear Buffer Capacity Model. If only acid−base chemical equilibria have to be considered, a linear buffer capacity model represented by a set of equations which are linear in the concentrations is sufficient to describe observed buffer capacity data. Nonlinearities are the result of considering, among others, complexation and precipitation reactions. Also, the parameters of the equilibria are nonlinear in the model. More details on the development of the model can be found in refs *20* and *25*.

Automatic Buffer Capacity Model Building. If there is a lack of prior knowledge about the buffer systems that are to be expected in a titrated sample, it is not straightforward to construct and end up with a "satisfying" buffer capacity model. To describe efficiently the automatic buffer capacity model building algorithm, it is necessary to introduce a specific terminology:

Blind buffer: A monoprotic nonspecified buffer used for model extension, of which the acidity constant K_a is automatically determined in the automatic model building procedure.

Known buffer: A buffer of which the acidity constant(s) are exactly known or known within a minimum−maximum interval (e.g. acetic acid buffer with a pK_a between 4.6 and 5.0).

Range: The range is defined as the difference between the minimum and maximum boundary value set as an interval in which the considered pK_a value is allowed to vary.

Optimization: Nonlinear parameter estimation with the PRAXIS algorithm (*26*), to fit the buffer capacity model to the experimental buffer capacity curve. The concentrations and/or acidity constants of known and/or blind buffers that are to be estimated are user-defined.

Residual: Difference between the experimental and the simulated buffer capacity at a particular pH value (symbol ϵ).

* Corresponding author phone: +32 9 264 5932; fax: +32 9 264 6220; e-mail: peter.vanrolleghem@rug.ac.be.

10.1021/es010025d CCC: $22.00
Published on Web 01/12/2002

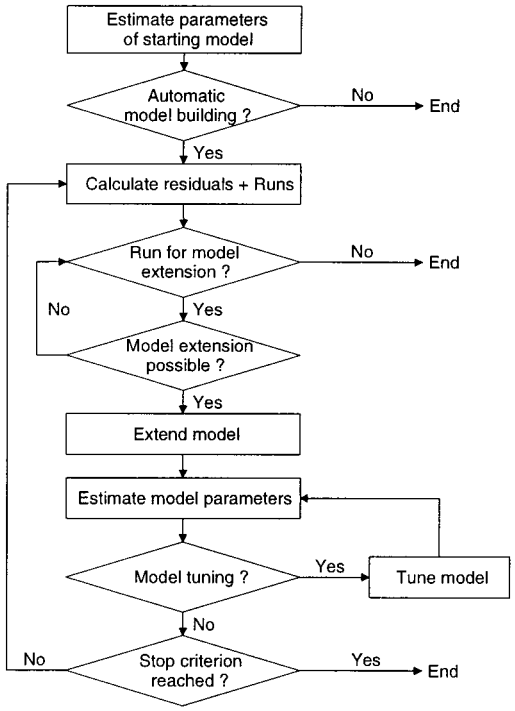

FIGURE 1. Flowchart of the automatic model building algorithm.

Run: A sequence of consecutive positive or negative residuals as function of the pH.

Model degree: The degree of a model is defined as the number of blind buffers incorporated in the model.

Model Building Algorithm. An overview of the different steps in the algorithm is given in Figure 1. The algorithm starts with the optimization of a user defined starting model. The residuals and Runs obtained with the starting model are calculated. This information is then used to define the candidate pH ranges for model extension. A number of criteria to define the best position for model extension in the buffer capacity profile are described and investigated in ref *23*. The criterion considered here is the sum of the residuals in each Run. This criterion combines the number of points per Run and the surface of the Run. The best candidate pH (and corresponding pK_a) for model extension is found in the Run with the highest positive sum of residuals (experimental buffer capacity > simulated buffer capacity). The pH with the maximum residual value is selected as the pK_a for an extra monoprotic buffer (*23*). This is illustrated in Figure 2.

The Runs are classified toward their priority as candidate positions for model extension. Then, the algorithm will evaluate for the best ranked Run if model extension is possible at that pH position. The basic idea of model extension is that the ranges (defined by respectively a minimum and maximum boundary) of all acidity constants to be estimated may never overlap with each other (e.g. if the pK_a of ammonium has a range from 9.0 to 9.4, then the algorithm will not allow any other buffer with a range that overlaps with the already occupied pK_a range). If model extension is not possible at the first ranked pH position (e.g. if that position and its neighborhood is occupied by another buffer) the algorithm proceeds to the second ranked Run for model extension etc. When an extended model is optimized, the algorithm proceeds, if necessary, with one or several model tuning cycles. Model tuning is the process in which the boundaries (minimum and maximum) for acidity constants of blind buffers that have been estimated are moved if the estimated value is too close to one of its boundaries. In this way some flexibility is given to the boundaries that may have been predefined too strictly. After each model tuning cycle, the tuned model is optimized again, and the tuning cycle is eventually repeated several times. Finally, after each tuning session, the model building stop criteria are calculated and evaluated (see next section).

Optimal Buffer Capacity Model Selection. When the automatic model building algorithm is applied, one ends up with a set of mathematical models that have all been fitted to the experimental buffer capacity data. The next logical step is a selection of the most appropriate model for the purpose the model will be used for. Model structure selection techniques, also called model structure characterization methods, are widely available from literature (*27*−*30*). Six useful model selection criteria are presented below. The first two criteria are statistical tests; the four next criteria include information about the model "complexity".

The Runs-Test or Testing Changes of Sign. Let R be the number of changes of sign in the residual sequence $\epsilon(1)$, $\epsilon(2)$, ... $\epsilon(N)$, with N the number of experimental data points. For the illustrated example in Figure 2, $R = 6$. Under the assumption of the null hypothesis H_0: $\epsilon(t)$ is zero mean white noise (i.e. the residuals are independent, have an expectation of 0 and are equally likely to be positive or negative), a test statistic can be constructed (*29*):

$$u = \frac{R - (N-1)/2}{\sqrt{N-1}/2} \rightarrow \mathcal{N}(0, 1) \qquad (1)$$

The test statistic is asymptotically distributed as a standard normal distribution. Hence, a 95% asymptotic confidence interval for u is given by $|u| \leq 1.96$ for critical significance level $\alpha = 0.05$. Practically, if N is large (e.g. $N > 30$), $(N - 1)$

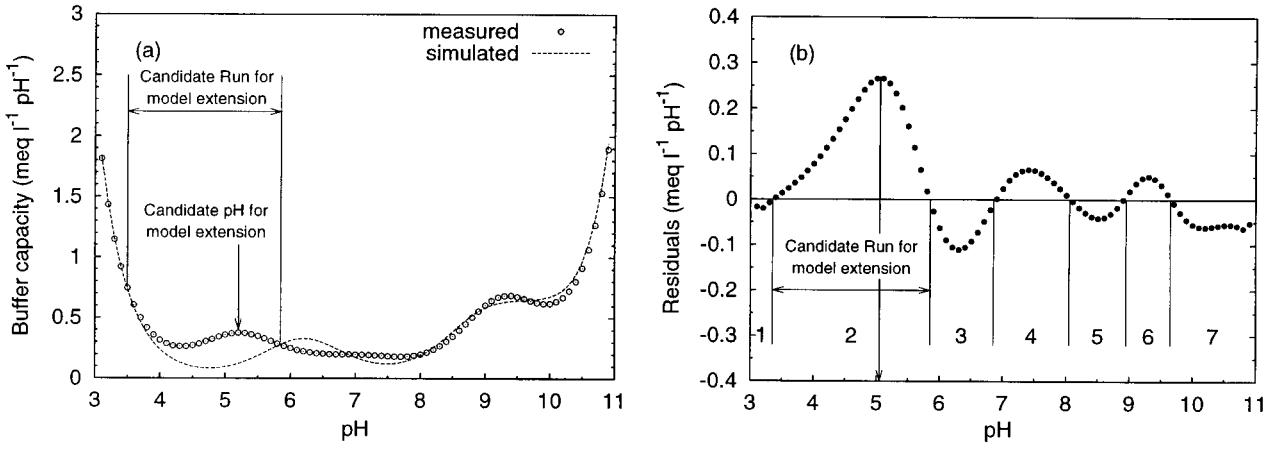

FIGURE 2. Experimental and simulated buffer capacity curves (a) and calculated residuals (b), with indication of the Runs, the best candidate Run for model extension, and the best candidate pH within that Run (downward arrow).

might be approximated by N. If several models have to be compared with each other, the model with the lowest number of parameters for which $|u|$ is lower than u_α is selected.

The F-Test for Comparison of Model Structures. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two model structures, such that $\mathcal{M}_1 \subset \mathcal{M}_2$ (for example $\mathcal{M}_1$ corresponds to a lower-order model than $\mathcal{M}_2$). In such a case they are called hierarchical model structures (*29*). Further, let SSE_i denote the sum of squared errors (or residuals) in the structure $\mathcal{M}_i$ ($i = 1, 2$) and let $\mathcal{M}_i$ have p_i parameters. The test statistic

$$F_w = \frac{(SSE_1 - SSE_2)/(p_2 - p_1)}{SSE_2/(N - p_2)} \quad (2)$$

is used to compare the model structures $\mathcal{M}_1$ and $\mathcal{M}_2$. If F_w is "large" ($F_w > F_{\alpha;\ p_2 - p_1;\ N - p_2}$), one concludes that the decrease in loss function from SSE_1 to SSE_2 is significant and, hence, that the model structure $\mathcal{M}_2$ is significantly better than $\mathcal{M}_1$. On the other hand, when F_w is "small", the conclusion is that $\mathcal{M}_1$ and $\mathcal{M}_2$ are almost equivalent, and according to the parsimony principle the smaller model structure $\mathcal{M}_1$ should be chosen as the most appropriate one.

Akaike's Information Criterion (AIC). Another approach to model structure selection consists of using a criterion that in some way penalizes the decrease of the loss function with increasing model complexity. A widely used criterion is Akaike's information criterion (*29, 30*)

$$AIC = N \log\left(\frac{SSE}{N}\right) + 2p \quad (3)$$

with SSE and N and p as defined before. The first term in eq 3 decreases with increasing p (increasing complexity), while the second term penalizes too complex (overparametrized) models. The model structure with the smallest criterion value is selected.

Final Prediction Error (FPE). The FPE criterion is similar to the AIC criterion, but with a different penalizing term (*29, 30*):

$$FPE = \frac{SSE}{N}\left(1 + \frac{2p}{N - p}\right) \quad (4)$$

Schwarz or Bayesian Information Criterion (SIC or BIC). An example of a consistent criterion is the Schwarz Information Criterion (*30, 31*):

$$SIC = N \log\left(\frac{SSE}{N}\right) + p \log(N) \quad (5)$$

If $N \geq 8$, SIC will tend to favor models of lower complexity than those chosen by AIC.

Corrected Akaike's Information Criterion (AIC_C). The AIC criterion is a biased criterion, leading to overfitting or the selection of overdimensioned models. Therefore, a bias corrected form of the AIC criterion was introduced (*32, 33*):

$$AIC_C = N \log\left(\frac{SSE}{N}\right) + N\,\frac{1 + p/N}{1 - (p + 2)/N} \quad (6)$$

The bias correction is of particular use when the sample size is small, or when the number of parameters p is a moderate to large fraction of the sample size N (*33*).

The choice of the critical significance level α in the Runs-test and F-test strongly affects the selected optimal model and has been investigated in ref *25*. It was found that a classical choice of α of e.g. 0.01 or 0.05 for this type of application (i.e. buffer capacity modeling) results in overfitting compared to the "human expert advice". Similar observations were found in another field, where oxygen uptake rate (OUR) models are fitted to respirometric data (*34*). Therefore, throughout the presented work, an arbitrary but well-equilibrated value of $\alpha = 0.0001$ was chosen for both the Runs-test and the F-test.

TABLE 1. Chemicals for the Preparation of Samples of Which Titration Curves Were Collected

chemical	code	protonicity: theoretical	protonicity: practical[b]	pK_{a1}	pK_{a2}	pK_{a3}
oxalic acid	oxa	2	1	1.23	4.19	
sodium acetate	ace	1	1	4.75		
ammonium chloride	amm	1	1	9.20		
malonic acid	mal	2	2	2.83	5.69	
sodium tripoly-phosphate	pho	3	2	2.12[a]	7.21[a]	12.67[a]
citric acid	cit	3	3	3.14	4.77	6.39

[a] Values of ortho-phosphate. The real values have to be determined experimentally. [b] pK_a values are not more than 2 units outside the considered pH simulation interval (3.5−10.5).

Software Implementation. The program *bomb* (buffer capacity optimal model builder) is developed in C++ and has been compiled on different computer platforms. The algorithm for Gauss-Jordan elimination was adopted from the library "Numerical Recipes in C" (*35*), and the PRAXIS algorithm in C is available from the author at no cost (*26*).

Experimental Data Collection. A number of titration curves (N typically around 60) with known composition were collected in the framework of adequate buffer capacity model development. The titration data collection is described in ref *22*. The database contains 146 titration curves, originating from 66 different samples, each titrated in 2- or 3-fold. Four different concentrations of each chemical were used for the preparation of the samples: 0.5, 1, 2, and 5 mequiv L^{-1}. The chemicals that were used to prepare the samples are presented in Table 1. One specific sample, containing 1 mequiv L^{-1} citric acid and 0.5 mequiv L^{-1}acetic acid, was selected as test case and for illustration purposes.

The titration curves of the samples with well-known composition were classified in four groups:

Group 1: 27 titration curves with one monoprotic buffer; 1 pK_a value to be estimated;

Group 2: 27 titration curves with one di- or triprotic buffer; 2 or 3 pK_a values to be estimated;

Group 3: 60 titration curves with two buffers (mono-, di- or triprotic); 2−5 pK_a values to be estimated;

Group 4: 32 titration curves with more than two buffers; 3−7 pK_a values to be estimated.

Results

One of the samples containing citric acid and acetic acid (cit-ace) with concentrations 1−0.5 mequiv L^{-1} was selected as a test case to compare two modeling approaches. The first approach is based on the availability of a priori knowledge about the buffers present in the sample, while in the second approach, it is assumed that no information about the composition of the sample is available. These two approaches are presented and compared in the following paragraphs.

Automatic Model Building in the Presence of Prior Knowledge. The buffer capacity model that was fitted to the experimental data included the water buffer, an inorganic carbon (IC) buffer, a citric acid buffer, and an acetate buffer. In previous work (*25*), it was found that it can be useful to allow some flexibility on the pK_a values of the considered buffer systems. Such flexibility corrects for small deviations of the real pK_a compared to the experimental pK_a, due to measurement errors, ionic strength effects, temperature effects, ... Therefore, the buffer capacity model was fitted twice to the experimental data, without and with extra

TABLE 2. Buffer Capacity Model Specifications for a Simulation Interval between pH 3.5 and pH 10.5

buffer	variable	initial guess or value	estimated?	lower limit	upper limit
water	pka_water	15.74	yes	15.70	15.82
	conc_water	55.5 mol L^{-1}	no		
IC	pka1_carbon	6.37	no		
	pka2_carbon	10.25	no		
	conc_carbon	0.5 mg CO_2 L^{-1}	yes	0	5.5
citric acid	pka1_citric	3.14	no/yes[a]	3.01	3.27
	pka2_citric	4.77	no/yes[a]	4.6	4.9
	pka3_citric	6.39	no/yes[a]	6.2	6.52
	conc_citric	0.1 mmol L^{-1}	yes	0	1
acetic acid	pka_acetic	4.75	no/yes[a]	4.6	4.9
	conc_acetic	0.1 mmol L^{-1}	yes	0	1

[a] "Yes" for the incorporation of extra flexibility on the considered pK_a.

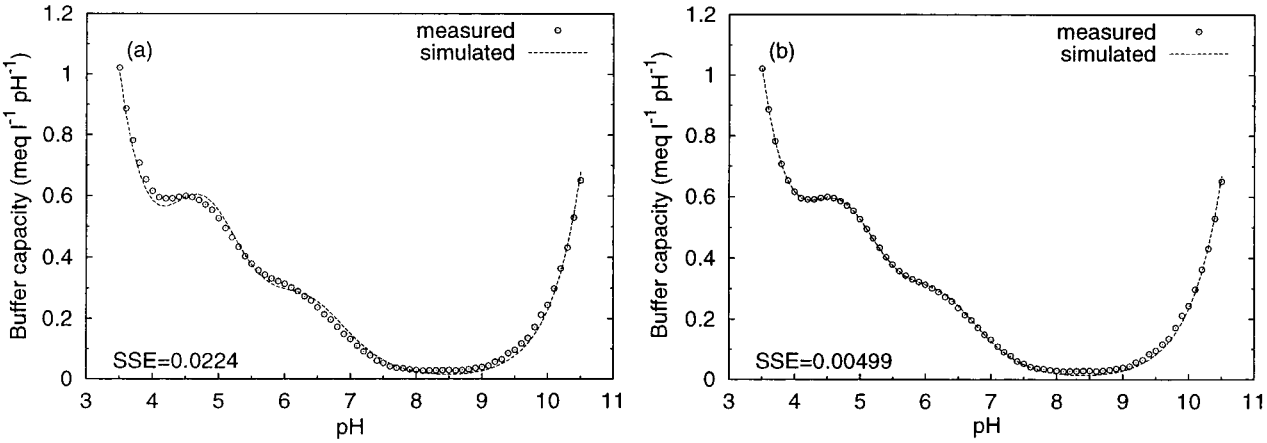

FIGURE 3. Experimental and simulated buffer capacity curves of the testcase. Results are shown for the buffer capacity model without (a) and with (b) extra flexibility on the p$K_a$values of the citric and acetic acid buffer.

TABLE 3. Simulation Results of the Selected Test Case for the Buffer Capacity Model without (a) and with (b) Extra Flexibility on the pK_a Values of the Citric and the Acetic Acid Buffer

		model (a)		model (b)		
	unit	estimate	SD	estimate	SD	expected
pK_a H_2O		15.79	0.012	15.81	0.0078	
pK_{a1} citric				3.27	0.11	3.14
pK_{a2} citric				4.88	0.09	4.77
pK_{a3} citric				6.25	0.03	6.39
pK_a acetic				4.60	0.07	4.75
C_{IC}	mg CO_2 L^{-1}	1.69	0.72	3.03	0.49	
C_{citric}	mmol L^{-1}	0.35	0.016	0.33	0.014	0.33
C_{acetic}	mmol L^{-1}	0.56	0.023	0.55	0.039	0.5

flexibility on the theoretical pK_a values, respectively, called models (a) and (b). The model specifications are summarized in Table 2.

The experimental and simulated buffer capacity curves of the selected sample, without and with extra flexibility on the pK_a values are respectively shown in Figure 3(a) and (b). The corresponding results for the estimations of the pK_a values and the concentrations are given in Table 3 together with their standard deviations.

The applied buffer capacity models (a) and (b) have respectively four and eight parameters to be estimated. Both models succeeded in a satisfying estimation of the concentrations of respectively the citric and the acetic acid buffer. As can be expected, the fit of the simulated to the experimental buffer capacities, however, is worse for model (a) compared to model (b) (see Figure 3). With model (b) the pK_{a1} value of the citric buffer and the pK_a value of the acetic buffer are estimated respectively at their maximum and minimum allowed boundary value. Allowing a wider flexibility range for these pK_a values resulted in concentrations of these two buffers slightly more deviating from their true values (results not shown). An advantage of the flexible pK_a approach compared to the fixed pK_a approach is that the flexible pK_a approach allows to more easily detect additional unexpected buffers (e.g. interferences) in the buffer capacity profile. For instance, for this particular example in the situation (b), it can be seen that the fitted mathematical model describes the experimental data fairly well, except that between pH 8 and pH 9 probably another buffer is present that is not accounted for in the model. On the other hand, in the situation (a) it is not clear whether extra buffers are necessary and at what pH they should be included in order to fit the experimental buffer capacity profile more closely. Therefore, the extra buffer between pH 8 and pH 9 in situation (a) remains unnoticed due to the overall lack of fit.

The above presented approach is based on the prior knowledge that the selected sample contains at least a monoprotic and a triprotic acid (respectively acetic and citric acid). In the next approach, it is considered that this prior information is not available.

Automatic Model Building in the Absence of Prior Knowledge. The second modeling approach uses the automatic model building algorithm included in the software *bomb*. The same well-known sample is considered again, and the stepwise model building process is initiated, starting from a model that only contains the water buffer. The algorithm will systematically add new buffers until the situation that the model cannot further be extended (e.g. due to the restriction that pK_a ranges may never overlap with each other). At this stage, an appropriate model selection criterion (e.g. AIC, Runs-test, ...) is not considered yet.

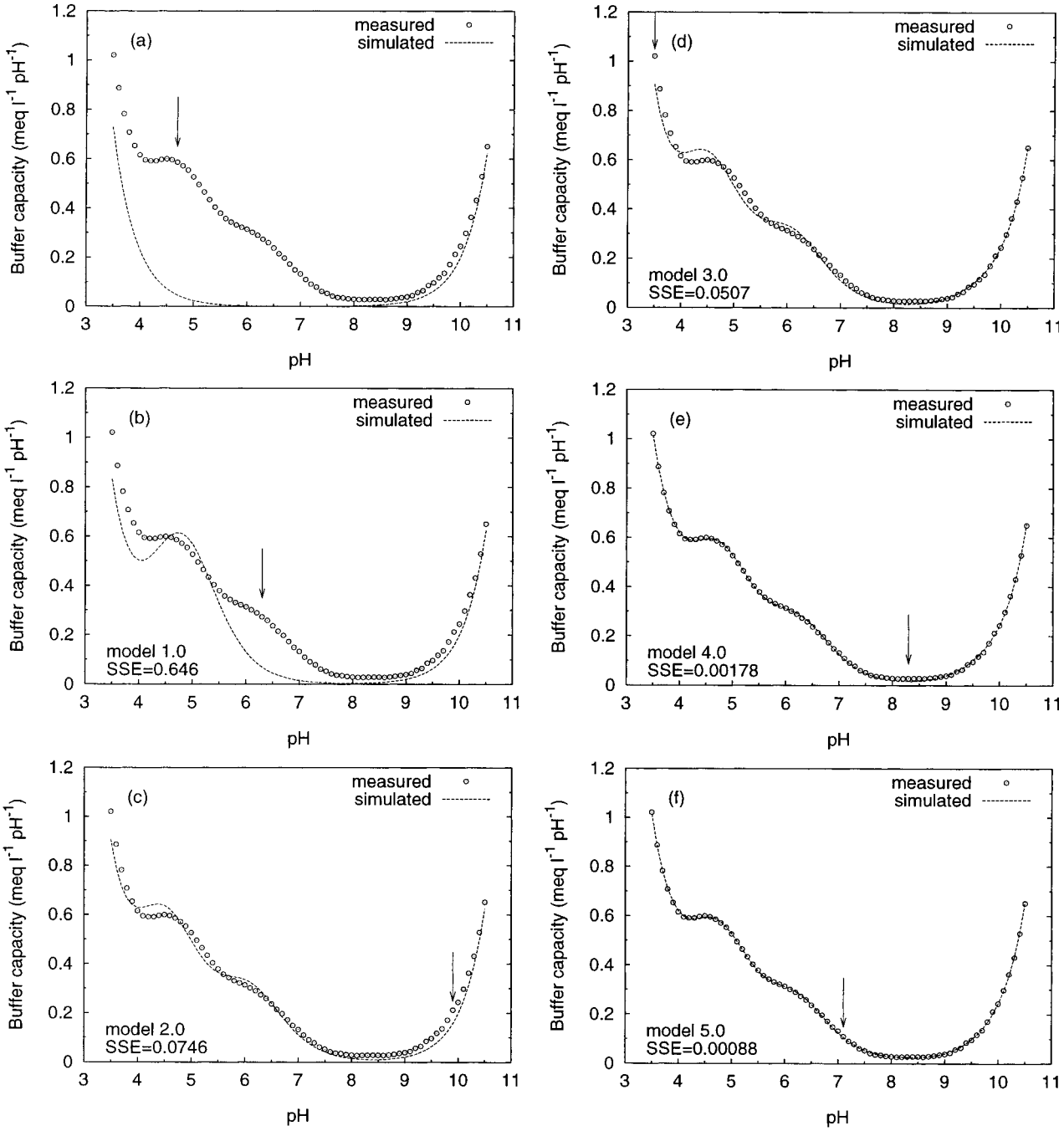

FIGURE 4. Experimental and simulated buffer capacity curves of a well-known sample. The stepwise model building process starts with plot (a) (the starting model) and ends with plot (f) (the final model). The arrows indicate the automatically proposed pK_a positions for model extension.

The stepwise model building process with the test case is illustrated in Figure 4. The pH simulation interval was set between pH 3.5 and pH 10.5, and the pK_a of the water buffer was set to 15.81, based on previous models where the pK_a of water was allowed to be estimated. At the last modeling step (i.e. a sixth model extension at pH 7.1 of the fitted model shown in Figure 4(f)), the parameter estimation routine PRAXIS ran into a local minimum because it did not find an objective function value that was lower than the previous, less complicated model (which is theoretically not possible in case of correct optimization, because the less complicated model is always a subset of the more complicated model). Apparently, the optimization problem became overparametrized, and thus it was not useful to proceed the model building process. The simulation results with the models illustrated in Figure 4(b)−(f) are summarized in Table 4. The columns in Table 4 show the simulation results for the different automatically built models. The buffer introduced in model 1 represents the combination of the dissociation of acetic acid (pK_a = 4.75) and the second dissociation step of citric acid (pK_a = 4.77). The buffers introduced in models 2 and 4, respectively, represent the third and first dissociation step of citric acid. The estimated pK_a of the first dissociation of citric acid step is 3.50 which is the minimum pH value of the considered pH interval, and the model building algorithm does not allow the introduction of buffers with pK_a outside the interval. The buffer introduced in model 3 is an unexpected buffer, and further investigation indicated that this buffer originates from interfering silicates entering the titration vessel through the NaOH titrant (*25*). The buffer introduced in model 5 is not considered very important, despite the fact that its concentration is significantly different

TABLE 4. Simulation Results of the Test Case for the Automatically Built Buffer Capacity Models Illustrated in Figure 4, Plots (b) until (f)[a]

	model 1 plot (b)	model 2 plot (c)	model 3 plot (d)	model 4 plot (e)	model 5 plot (f)
pK_a blind1	4.78 ± 0.06	4.51 ± 0.03	4.50 ± 0.02	4.74 ± 0.01	4.74 ± 0.007
pK_a blind2		6.08 ± 0.06	6.08 ± 0.05	6.24 ± 0.01	6.22 ± 0.008
pK_a blind3			9.88 ± 0.21	9.90 ± 0.04	9.99 ± 0.03
pK_a blind4				3.50 ± 0.05	3.50 ± 0.03
pK_a blind5					8.21 ± 0.13
C_{blind1}	1.00 ± 0.05	0.93 ± 0.02	0.93 ± 0.02	0.83 ± 0.008	0.83 ± 0.006
C_{blind2}		0.47 ± 0.02	0.47 ± 0.02	0.40 ± 0.004	0.40 ± 0.003
C_{blind3}			0.08 ± 0.02	0.08 ± 0.003	0.08 ± 0.002
C_{blind4}				0.32 ± 0.007	0.32 ± 0.005
C_{blind5}					0.02 ± 0.002

[a] The concentrations are expressed as mmol L^{-1}, and the table entries are estimate ± SD.

TABLE 5. Model Selection Criteria Calculated for the Consecutive Models in the Automatic Model Building Process of the Test Case[a]

	model 1	model 2	model 3	model 4	model 5	model 6
AIC	−330	−479	−502	−736	**−782**	−729
AIC_C	−256	−405	−428	−660	**−705**	−649
FPE	0.0096	0.0012	0.00085	0.000031	**0.000016**	0.000035
SIC	−325	−470	−489	−718	**−760**	−701
Runs-test (*p*)	$<10^{-6}$	$<10^{-6}$	0.000001	**0.00013**	0.00047	0.00047
F-test (*p*)	$<10^{-6}$	$<10^{-6}$	0.000003	$<10^{-6}$	$<10^{-6}$	**1.00000**[b]

[a] The bold items mark the selected models based on each considered criterion [b] An insignificant *p*-value for model 6 means that model 5 is selected.

from 0 (*t*-test, $\alpha = 0.01$). The model building process finally tried to introduce a buffer with p$K_a = 7.1$ (model 6, results not shown), but then the optimization problem became overparametrized and thus the algorithm stopped at this stage.

The 6 model selection criteria, applied to the test case, are summarized in Table 5. If the automatic model selection was activated, all criteria would select model 5 as the final model, except for the Runs-test, that would favor model 4. Of course, the latter finding completely depends on the choice of the critical significance level α. Also, model 6 ended up in a wrong optimization result, thus forcing all criteria (including the *F*-test) to select the simpler model 5. The first four criteria (AIC, AIC_C, FPE, and SIC) show their highest decrease in value from models 1 to 2 and from models 3 to 4. This points to very important increases in model adequacy at those two stages. This is also nicely reflected in Figure 4.

Evaluation of Six Model Selection Criteria. All 146 titration curves were used as input data to the automatic model building environment in the software *bomb*. It was assumed that no prior knowledge was available about the buffers to be expected in the samples. Consequently, the starting model only contained the water buffer. The model building process was repeated for each of the 6 model selection criteria.

Table 6 presents summarizing frequency tables with the counts of titration curves corresponding with the selected model degree tabulated in function of the amount of practical pK_a values to be found. In general, the first four criteria (AIC, AIC_C, PFE, and SIC) select the same model. However, large differences in the selected model are noticed for the Runs-test and the *F*-test. A more detailed analysis per group of titration curves is given in the next paragraphs.

In group 1 (containing 1 monoprotic buffer), the automatic model building algorithm perfectly detects the single monoprotic buffer in all experimental buffer capacity profiles and incorporates this buffer as the first blind buffer in the model. The estimated concentrations were not significantly different from the theoretically expected concentrations. However, the model building process does not stop at this stage in most cases but continues with the incorporation of new blind buffers. In order of importance, first a buffer around pH 10 with a typical concentration of 0.1 mmol L^{-1}, and second a buffer around pH 6.5 with a typical concentration between 0.05 and 0.1 mmol L^{-1} are added to the model. As mentioned in the previous sections, these two buffers are assumed to be silicate and inorganic carbon, respectively. Their concentrations are mostly found to differ significantly from 0. The Runs-test and *F*-test based model building generally stop at this stage. However, based on the AIC and related criteria, the model is mostly extended with one or even two extra buffers of unknown origin (around pH 7 and pH 4). The concentrations of these extra buffers are very low and sometimes not significantly different from 0. Practically, it is concluded that the AIC and related criteria go rather far in the model building process. In seven out of the 27 cases, the *F*-test selected the model with only one buffer as the final model, thus not recognizing the silicate and/or the IC buffer. This behavior could be attributed to an optimization problem. It was noticed that for some "incomplete" models (thus models still deviating a lot from the final model), the parameter estimation routine ended in a local minimum at the second modeling step (five cases out of the 27). Consequently, the simpler model was selected by the *F*-test. The AIC and related criteria suffered from this flaw too. Only the Runs-test was not influenced by local minima problems, because this test only examines the randomness of the residuals. As a preliminary conclusion, the Runs-test was found to perform most "realistically" for this group of titration curves.

For the titration curves of group 2 (containing a di- or triprotic buffer), similar findings as for group 1 can be formulated. The selected models mostly contain, as expected, one or two supplementary buffers compared to the models selected in group 1. The AIC and related criteria often select a model with five or six blind buffers. This is again a higher model degree than what should be considered practically. Further, these high degree models illustrate certain weaknesses in the model building algorithm. More particularly, sometimes buffers are initially added in the model (interfering

TABLE 6. Frequency Tables of the Number of Titration Curves Classified toward the Selected Model Degree and the Applied Model Selection Criterion in Function of the Number of Expected pK_a Values To Be Found with the Model Building Algorithm

group[a]				no. of pK_a values[b]	final model degree	criterion for model selection					
1	2	3	4			AIC	AIC_C	FPE	SIC	Runs-test	*F*-test
27	0	0	0	1	1	4	4	4	5	1	7
					2	2	2	2	4	7	5
					3	2	2	2	4	19	9
					4	14	14	14	11	0	6
					5	5	5	5	3	0	0
0	8	20	0	2	1	1	1	1	1	0	4
					2	0	0	0	2	1	0
					3	1	1	1	1	16	2
					4	8	8	9	8	10	8
					5	15	15	14	15	1	14
					6	3	3	3	3	0	0
0	19	23	8	3	1	0	0	0	0	0	1
					2	0	0	2	0	0	10
					3	1	1	2	1	4	0
					4	9	9	9	9	27	10
					5	24	24	25	24	15	21
					6	16	16	12	16	2	8
					7	0	0	0	0	0	0
					8	0	0	0	0	2	0
0	0	12	0	4	1	0	0	0	0	0	2
					2	0	0	1	0	0	5
					3	0	0	0	0	0	0
					4	3	3	4	3	5	4
					5	9	9	7	9	2	1
					6	0	0	0	0	1	0
					7	0	0	0	0	0	0
					8	0	0	0	0	4	0
0	0	5	11	5	3	1	1	1	1	0	1
					4	2	2	3	2	1	8
					5	4	4	3	4	11	3
					6	7	7	7	7	4	4
					7	2	2	2	2	0	0
0	0	0	9	6	4	1	1	1	1	0	1
					5	4	4	4	4	7	3
					6	3	3	3	3	2	5
					7	1	1	1	1	0	0
0	0	0	4	7	4	0	0	0	0	2	0
					5	3	3	3	3	2	4
					6	1	1	1	1	0	0

[a] Overview of the number of titration curves belonging to each of the four groups (see "Experimental data collection"). [b] E.g. number of pK_a values = 2 includes titration curves with one diprotic (group 2) or two monoprotic buffers (group 3).

buffers, in small but significant concentrations), that in a later modeling stage become insignificant. As the algorithm is now, it does not allow for removing a buffer from the model in case the concentration of this buffer is no longer significantly different from 0. It is suggested to include this feature in a next version of the model building algorithm (e.g. based on a similar *F*-test based method that is also implemented in stepwise linear regression algorithms (*36*)). Local minima problems with "incomplete models", as described earlier, occurred less frequently (only one out of the 27 cases). In groups 3 and 4, this particular problem was no longer noticed. Thus, it is concluded that only the simplest type of titration curves sometimes leads to a too low model degree due to an optimization problem. Again, the Runs-test was found to select the most realistic final model.

The titration curves of group 3 (containing two buffers) have an increased complexity and should theoretically contain at least between 2 (e.g. amm-ace) and 5 (e.g. cit-mal) pK_a values. For the simplest type of combinations (e.g. the combination of a monoprotic and a diprotic buffer), the quantified buffers corresponded with the buffers present, both for the pK_a positions and the respective concentrations. Again, supplementary buffers were sometimes found around pH 10 (silicates). Another particularity related to the model building algorithm was discovered. In the example cit-mal, the $pK_{a2} = 5.69$ of malonic acid is only 0.7 different from the $pK_{a3} = 6.39$ of citric acid. In most examples of this type, a blind buffer was incorporated in the model somewhere between these two theoretical pK_a values and was accounting for both buffer systems. This points to a certain "limit of resolution" which is further discussed in the next paragraph. For the simplest type of curves in group 3, the Runs-test was slightly favored for practical purposes. However, for the more complicated titration curves, the Runs-test was not found most appropriate in a number of situations. More particularly, for six out of 60 titration curves, the Runs-test selected a model with eight blind buffers, whereas all other criteria considered only six blind buffers as the maximum number of buffers needed in the final model. The maximum number of buffers that are currently available in the automatic model building algorithm is 8. Therefore, in those cases where the Runs-test selected eight blind buffers, the necessary Runs-test criterion was never reached. This phenomenon only occurred in examples with phosphate as one of the buffers. It can be pointed out that the uncertain position of the pK_{a2} value of phosphate could be the reason an acceptable fit was not found with the Runs-test criterion. An illustrative titration example of pho-mal is given in Figure 5. The fitted buffer capacity model contains four blind buffers (with pK_a values in order of importance 5.7, 9.3, 3.5, and 6.8). A further model extension proposed at pH 8.1 results in an insignificant *F*-test result ($p = 0.26$), whereas the Runs-test result is still very significant ($u = -4.5$ or $p = 0.000007$). Further model extensions do not make the fit better compared to Figure 5. Thus, when the Runs-test criterion is selected for this particular example (with $\alpha = 0.0001$), the model building algorithm will proceed until model extension is not possible any more and will finally not find any model that fulfills the requested criterion.

The final model results of the easiest type of titration curves in group 4 (e.g. the combination of two monoprotic buffers and a diprotic buffer) could easily be related with the buffers to be expected. The *F*-test criterion mostly selected the same model as the AIC and related criteria. The Runs-test criterion often selected models of lower complexity. The results obtained with the *F*-test criterion were found to correspond most closely to reality. The interfering buffer at pH 10 (silicates) was still detected and quantified in most cases. The "limit of resolution" mentioned in the previous paragraph was further investigated. Again, it was found that two buffers with neighboring pK_a values are pooled together in the model building algorithm. It was found that the practical limit of resolution for the considered examples is between 0.5 and 1 pH units. Furthermore, the intervals of the pK_a values are not allowed to overlap, thus this automatically limits the number of buffers that can enter in the model around the same pK_a. A comparison with the Gordon algorithm described in refs *18* and *19* was made, and the defined criterion for the limit of resolution was compared with the obtained results. Gordon found that a minimal pK_a separation of 0.1−0.2 is borderline for most applications, whereas 0.5 is a common value. However, it should be noted that the Gordon algorithm only allows variation on the pK_a value that is entered in the model last, together with the neighboring pK_a value, whereas in the *bomb* algorithm, variation may be allowed on any pK_a. It can be concluded that the limit of resolution obtained here is somewhat lower than what is described in refs *18* and *19* but still in an acceptable range. It is also expected that the approach presented in this paper is more robust compared to the

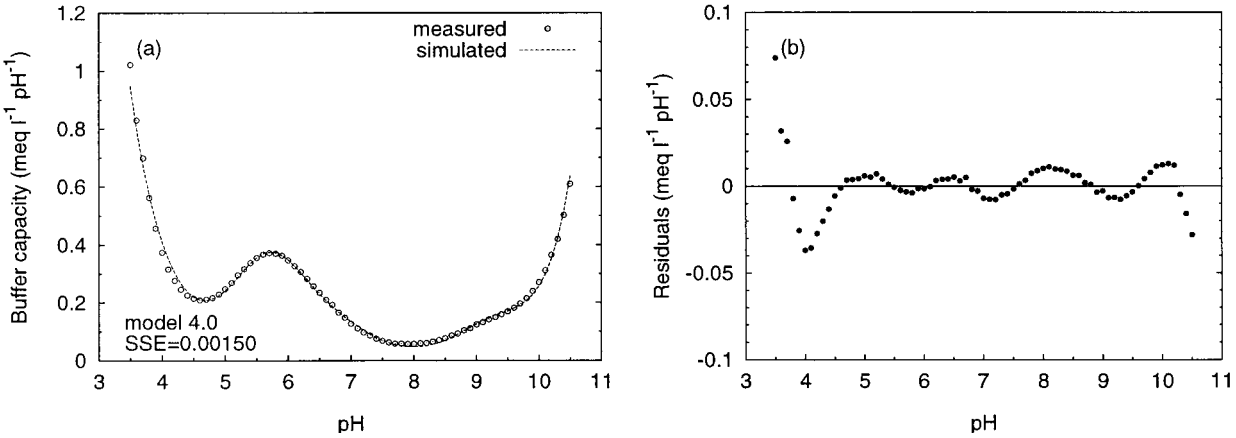

FIGURE 5. Experimental and simulated buffer capacity curves (a) and calculated residuals (b) of a sample pho-mal.

Gordon algorithm, because of the higher flexibility on the pK_a values, the built-in model "tuning" in case an estimated pK_a value touches one of its boundary values, and the use of model selection criteria.

Discussion

The automatic model building algorithm that is implemented in the software *bomb* has been applied to many titration curves of diverse samples, with the objective to develop tailor-made buffer capacity models. A number of shortcomings that were detected and described in the first development stage of this algorithm (*23*) are now completely solved. At present, the model building algorithm has been evaluated as sufficiently robust and fail-safe.

Six different model selection criteria were evaluated for the purpose of selecting the most appropriate model from the resulting set of automatically built models for each individual sample. The AIC, AIC_C, FPE, and SIC criteria were found to perform very similarly. The Runs-test and *F*-test criteria often have a different behavior compared to each other and compared to the AIC and related criteria. There is no "best" criterion for general purposes, because all criteria have their advantages and disadvantages. An interesting perspective for further research could be to use a combination of the results with the different model selection criteria to make the proposed model building algorithm more robust. The model selected by each of the criteria could be introduced in an "expert system" that can make a choice based on the advantages and disadvantages known for each criterion.

The benefits of automatic model building compared to a fixed model approach are different for various applications. In situations where prior knowledge about the sample is already high, the main benefit of the automatic model building is that extra information about the sample can be obtained. This includes the detection of unexpected or interfering buffers (e.g. silicates) and the detection of experimental problems in the titration system. In the other situations, where the prior knowledge is not high, the automatic model building environment can be applied as a support tool for a quick characterization of unknown buffer capacity profiles. In practical situations where a buffer capacity model has to be developed, applying the automatic modeling approach on a number of preliminary samples can help the expert to define the most appropriate model.

Acknowledgments

This research was part of the project 3G010297, which was financially supported by the Fund for Scientific Research Flanders (FWO-VL) and a scholarship from the Flemish Institute for the Improvement of Scientific-Technological Research in the Industry (IWT). The authors also want to thank Carlo Mussche for the titration data collection and express their gratitude to all reviewers for their constructive comments.

Literature Cited

(1) Gibs, J.; Schoenberger, R. J.; Suffet, I. H. *Water Res.* **1982**, *16*, 699.
(2) Hill, A. R.; Irvine, D. M.; Bullock, D. H. *J. Food Sci.* **1985**, *50*, 733.
(3) Husted, S.; Jensen, L. S.; Jørgensen, S. S. *J. Sci. Food Agric.* **1991**, *57*, 335.
(4) Lucey, J. A.; Gorry, C.; Fox, P. F. *Milchwissenschaft* **1993**, *48*(4), 183.
(5) Sawyers, D. E.; Dentel, S. K. *Water Sci. Technol.* **1992**, *26*(9−11), 2265.
(6) Anderson, G. K.; Yang, G. *Water Environ. Res.* **1992**, *64*(1), 53.
(7) Bernard, O.; Polit, M.; Hadj-Sadok, Z.; Pengov, M.; Dochain, D.; Estaben, M.; Labat, P. *Water Sci. Technol.* **2001**, *43*(7), 175.
(8) Bisogni, J. J., Jr. *Water Environ. Res.* **1994**, *66*(1), 16.
(9) Bisogni, J. J., Jr.; Witzmann, S. W.; Stedinger, J. R. *Water Environ. Res.* **1998**, *70*(7), 1303.
(10) Buchauer, K. *Water SA* **1998**, *24*(1), 49.
(11) de Haas, D. W.; Adam, N. *Water SA* **1995**, *21*(4), 307.
(12) Moosbrugger, R. E.; Wentzel, M. C.; Ekama, G. A.; Marais, G.v. R. *Water Sci. Technol.* **1993**, *28*(2), 237.
(13) Powell, G. E.; Archer, D. B. *Biotechnol. Bioeng.* **1989**, *33*, 570.
(14) Rozzi, A. In *Biomethane, Production and Uses*; Buvet, R., Fox, M. F., Picken, D. J., Eds.; Turret-Wheatland; Rickmansworth, 1984; p 79.
(15) Rozzi, A.; Di Pinto, A. C.; Brunetti, A. *Environ. Technol. Lett.* **1985**, *6*, 594.
(16) von Sachs, J.; Feitkenhauer, H.; Meyer, U. *Proceedings, 5th international symposium on systems analysis and computing in water quality management*; Watermatex: Ghent, Belgium, 2000; pp 2.43−2.46.
(17) Bastin, G.; Dochain, D. *On-line estimation and adaptive control of bioreactors*; Elsevier: Amsterdam, 1990.
(18) Gordon, W. E. *J. Phys. Chem.* **1979**, *83*(11), 1365.
(19) Gordon, W. E. *Anal. Chem.* **1982**, *54*(9), 1595.
(20) Van Vooren, L.; Lessard, P.; Ottoy, J.-P.; Vanrolleghem, P. A. *Environ. Technol.* **1999**, *20*, 547.
(21) Van Vooren, L.; Willems, P.; Ottoy, J.-P.; Vansteenkiste, G. C.; Verstraete, W. *Water Sci. Technol.* **1996**, *33*(1), 81.
(22) Mussche, C. M. Sc. Dissertation, Ghent University, Belgium, 1997.
(23) Van De Steene, M. M. Sc. Dissertation, Ghent University, Belgium, 1998.
(24) Van Vooren, L.; Van De Steene, M.; Ottoy, J.-P.; Vanrolleghem, P. A. *Water Sci. Technol.* **2001**, *43*(7), 105.
(25) Van Vooren, L. Ph.D. Dissertation, Ghent University, Belgium, 2000.
(26) Gegenfurtner, K. R. *Behaviour Research Methods, Instrum., Comput.* **1992**, *24*(4), 560.
(27) Gustafsson, F.; Hjalmarsson, H. *Automatica* **1995**, *31*(10), 1377.
(28) Ljung, L. *System Identification: Theory for the User*; Prentice Hall: New Jersey, 1987.
(29) Söderström, T.; Stoica, P. In *System Identification*; Prentice Hall: New Jersey, 1989; Chapter 11.

(30) Vanrolleghem, P. A.; Dochain, D. In *Advanced instrumentation, Data interpretation, and control of biotechnological processes*; Van Impe, J. F., Vanrolleghem, P. A., Iserentant, D., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; Chapter 10.
(31) Hurvich, C. M.; Tsai, C.-L. *Biometrika* **1991**, *78*(3), 499.
(32) Hurvich, C. M.; Shumway, R.; Tsai, C.-L. *Biometrika* **1990**, *77*(4), 709.
(33) Hurvich, C. M.; Tsai, C.-L. *Biometrika* **1989**, *76*(2), 297.
(34) Vanrolleghem, P. A. Ph.D. Dissertation, Ghent University, Belgium, 1994.
(35) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C. The art of scientific computing*, 2nd ed.; Cambridge University Press: Cambridge, 1995.
(36) Neter, J.; Kutner, M. H.; Nachtsheim, C. J.; Wasserman, W. *Applied linear statistical models*, 4th ed.; Irwin: London, 1990.

Received for review January 22, 2001. Revised manuscript received October 23, 2001. Accepted October 24, 2001.

ES010025D