ELSEVIER

A new algorithm for the determination of protolytic constants from spectrophotometric data in multiwavelength mode: Calculations of acidity constants of 4-(2-pyridylazo)resorcinol (PAR) in mixed nonaqueous-water solvents

J. Ghasemi [a,*], Sh. Nayebi [a], M. Kubista [b], B. Sjogreen [c]

[a] Department of Chemistry, Faculty of Science, Razi University, Kermanshah 57167, Iran
[b] Department of Chemistry and Biosciences, Chalmers University of Technology, Goteborg, Sweden
[c] The Royal Institute of Technology, NADA, KTH, Stockholm, Sweden

Abstract

A new efficient, simple and versatile algorithm is presented for determination of the protolytic constants from spectrophotmetric data in multiwavelength mode based on the combining of hard and soft modeling. The algorithm was checked by determining the acidity constants of a triprotic acid from theoretical and real absorption–pH data. The real spectral data are obtained from photometric titration of different solutions of 4-(2-pyridylazo)resorcinol (PAR) by a standard base solution under an inert atmosphere. The algorithm starts the minimization process using an user supplied number of components and initial guesses of the unknown parameters and refined in a least squares manner. New algorithm is implemented in the new version of DATAN package (version 3.1). The validity of the obtained results was checked by some well known programs such as DATAN 2.1, SPECFIT/32, SQUAD, a modified version of difference spectra and a A–pH curve method. The comparison of the outputs of the DATAN 3.1 with the other programs reveals that there is a very good agreement between the obtained results and mentioned programs.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Acidity constants; DATAN; SQUAD; SPECFIT/32; Difference spectra; PAR; Water-nonaqueous binary solvents

1. Introduction

Acid dissociation constants are important parameters to indicate the extent of ionization of molecules in solution at different pH values. The acidity constants of organic reagents play a fundamental role in many analytical procedures such as acid–base titration, solvent extraction, complex formation and ion transport. It has been shown that the acid–base properties affect the toxicity, [1] chromatographic retention behavior and pharmaceutical properties [2] of organic acids and bases. Much of the theoretical foundation of modern organic chemistry is based on the observation of the effects on acid–base equilibrium of changing molecular structure [3].

It is still believed that spectrophotometric data are inherently less precise than potentiometric data [4], consequently most equilibrium constants are determined by means of potentiometric titrations, and only a few general programs are available at present for processing spectrophotometric data [5–18]. The most widespread of programs and algorithms for determining the acidity constants from absorbance data are SQUAD, SPECFIT and DATAN 2.1.

SQUAD [12–16] and SPECFIT [17,18] are based on an initial proposal of a chemical equilibrium model defining species stoichiometries and based on mass-action law and mass balance equations (hard modeling methods) and also involve least squares curve fitting procedures. DATAN 2.1,

* Corresponding author. Tel.: +98 214075448; fax: +98 8318369572.
E-mail address: jahan.ghasemi@tataa.com (J. Ghasemi).

[19–24] program calculates spectral profiles, concentrations and equilibrium constants by utilizing equilibrium expressions that are related to the components.

The difference spectra method [25] is an elegant algorithm for the analysis of the mono protic acid/base equilibria using multiwavelength detection which recover both the dissociation constants and absorption profiles spectra of organic acid/base pairs in complex mixtures using a least square method.

Here, we wish to describe an algorithm, which is applied to the determination of the acidity constants by processing of the pH–absorbance titration data. The ability of the new version of DATAN package (DATAN 3.1, Multid Analysis AB, Sweden, 2004) was compared with known programs SQUAD, DATAN 2.1, SPECFIT/32, by refining the simulated and experimental pH–absorbance data.

2. Theory

Since in all routine chemometrics the selection of the number of significant factors (chemical rank of data matrix or determining number of light absorbing species) is a crucial step, in the next section we pointed to the fundamental of the existing methods, following by a brief description of each algorithm in comparison with the new algorithm.

2.1. Principal component analysis and the rank of data matrix

Most measurements are not selective for only the constituents of interest; but the data also contain noise. In principal component analysis (PCA) the measured data are reduced to contain only the information that is relevant to the system [26–30]. The first step in PCA is to decompose the data matrix $\mathbf{A}$ into an orthonormal basis set:

$$\mathbf{A}_{n,m} = \mathbf{T}_{n,q}\mathbf{P}_{q,m}^{\mathrm{T}} = \sum_{i=1}^{q}\mathbf{t}_i\mathbf{p}_i^{\mathrm{T}} \tag{1}$$

where $\mathbf{A}_{n,m}$ contains the n recorded spectra as rows, each digitized into m data points, $\mathbf{T}_{n,q}$ the score matrix which relates to sample composition, $\mathbf{P}^{\mathrm{T}}$, where the superscript 'T' denotes transpose, is the loading matrix which relates to spectra and q is the least of n and m, which in spectroscopy usually is n. Eq. (1) is exact. The second step in PCA is to separate the eigenvectors that account for the systematic variations from those corresponding to noise:

$$\mathbf{A} = \mathbf{T}_{n,r}\mathbf{P}_{r,m}^{\mathrm{T}} + \mathbf{E}_{n,m} = \sum_{i=1}^{r}\mathbf{t}_i\mathbf{p}_i^{\mathrm{T}} + \mathbf{E}_{n,m} = \hat{\mathbf{A}} + \mathbf{E}_{n,m} \tag{2}$$

where $\hat{\mathbf{A}}$ is the predicted data matrix, $\mathbf{E}_{n,m}$ the residual matrix and r is the number of significant components or chemical species present in the solution mixture. Elbergali et al. [31] have reviewed several indicator functions, which commonly

used to assign the significant factors in PC decomposition. Meloun et al. [32] have compared the several methods of determining number of significant factors in spectroscopic data and showed the advantages and limitations of the presented methods. According to Beer's law the absorbance matrix is related to the concentrations of species and molar absorptivity, i.e. $\mathbf{A} = \mathbf{CV}$. The rank of matrix $\mathbf{A}$ is obtained from the equation $\mathrm{rank}(\mathbf{A}) = \min[\mathrm{rank}(\mathbf{C}), \mathrm{rank}(\mathbf{V}) \leq \min(m, r, n)]$. Since rank of $\mathbf{A}$ is equal rank of $\mathbf{C}$ or $\mathbf{V}$, whichever is smaller, and since $\mathrm{rank}(\mathbf{V}) \leq r$ and $\mathrm{rank}(\mathbf{C}) \leq r$, then provided $m \leq r$ and $n \leq r$, it will only be necessary to determine the rank of matrix $\mathbf{A}$ which is equivalent to the number of significant components [33,34]. All methods to identify the true dimensionality of a data set are classified into two categories: (a) precise methods based upon a knowledge of the instrumental error of the absorbance data, $s_{\mathrm{inst}}(A)$ before statistical examination; (b) approximate methods requiring no knowledge of the instrumental error of the absorbance data, $s_{\mathrm{inst}}(A)$. Many of these methods are empirical functions.

2.2. Algorithms

2.2.1. DATAN 2.1

The theory of DATa ANalysis (DATAN 2.1) method was discussed by Scarmino and Kubista in several papers [20–24].

The recorded spectra during a titration can be arranged in a data matrix, $\mathbf{A}$ which is decomposed into an orthonormal basis set by NIPALS or any PC decomposition equivalent method [35] as Eq. (2). The $\mathbf{T}$ and $\mathbf{P}^{\mathrm{T}}$ cannot be directly related to concentrations and spectral profiles. By assuming linear responses, the spectra in matrix $\mathbf{A}$ are linear combinations of the concentration $\mathbf{C}$ and spectral response $\mathbf{V}$, of the chemical components. So,

$$\mathbf{A} = \mathbf{CV} + \mathbf{E} \approx \mathbf{CV} \tag{3}$$

The $\mathbf{T}$ and $\mathbf{P}^{\mathrm{T}}$ matrices can be related to $\mathbf{C}$ and $\mathbf{V}$, respectively, as follows:

$$\mathbf{T} = \mathbf{CR} \tag{4}$$

$$\mathbf{P}^{\mathrm{T}} = \mathbf{R}^{-1}\mathbf{V} \tag{5}$$

Since, if $\mathbf{R}$ can be determined, the $\mathbf{A}$ and $\mathbf{C}$ matrices can be calculated from the Eqs. (4) and (5). The values of protolytic constants were varied to minimize the sum of square residuals:

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{r}\left(\mathbf{T}_{ij} - \sum_{k=1}^{r}\mathbf{r}_{ik}\mathbf{c}_{kj}\right)^2 \tag{6}$$

where r and n are number of light absorbing species and number of solutions, respectively. The accuracy of this fit depends crucially on the trial values of the equilibrium constants, and best fit determines their values and the elements of matrix $\mathbf{R}$. The details of this algorithm are explained in ref. [21] thoroughly.

2.2.2. SPECFIT/32 and SQUAD

SPECFIT/32 program was developed by Gammp et al. [17,18] for the determination of stability constants from the spectrophotometric titration data. The mathematical features of this program have been described [17,18], and that is similar to a general non-linear least squares program used for calculation of stability constants. The SPECFIT32 is the latest version of a global analysis program for equilibrium and kinetic systems with singular value decomposition and non-linear regression modeling using the Levenberg–Marquardt method [36,37]. Gammp et al. used the factor analysis in this program and in continuation evolving factor analysis [38–41], as a powerful tool for the determination of independent components in a given data matrix.

The Stability QUotients from Absorbance Data (SQUAD) program [12] derived from SCOGS [13]. Leggett [12,13] for the first time used a factor analysis method in the program for determination of stability constants. Non-linear least-squares method used is based on the minimization of the function, $\mathbf{E}$:

$$\mathbf{E} = \sum_{j=1}^{m} (\mathbf{Y}_{calcd} - \mathbf{Y}_{obsd})_j^2 w_j \tag{7}$$

where m is the number of data points and w_j is the weight of each absorbance value. The minimization approach, which was used in SQUAD is more or less similar to the SPEC-FIT/32, i.e. the object function is a residual matrix, which should minimized with respect to the stability constants of the selected chemical equilibrium model. Running the SQUAD has different stages that were described by Leggett in detail [16]. The calculated standard deviation of absorbance $s(A)$ is used as the most important criterion for a fitness test. If, after termination of the minimization process, the condition $s_{inst}(A) \approx s(A)$ is met and the Hamilton R-factor is also less than 1%, the hypothesis of the chemical model is taken as being the most probable, and is accepted.

2.2.3. Modified version of difference spectra

For multiprotic compounds or mixture of n_a acid–base pairs, only n_a parameters and initial estimated values of pK_a, needed in order to resolve the absorption spectra of the compounds. In titration systems, a data matrix $\mathbf{A}$ from photometric titration can be presented as Eq. (2). In order to factorize $\mathbf{A}$, the rows of $\mathbf{C}$ must be linearly independent so that a unique, best fit solution, exist. One of the peculiarities of the pH-dependent data set is an interdependence of the distribution curve of acid with that of its conjugate base, where the fraction of acid and its conjugate base at a given pH can be calculated by the mass balance equations [25].

To develop a set of linearly independent vectors for the $\mathbf{C}$ matrix for a multiprotic acid system, we use the difference of the distribution curves as Frans and Harris [25] for monoprotic acids:

$$\mathbf{c}_{1j} = \alpha_{2j} - \alpha_{1j} = 1 - 2\alpha_{ij} = 10^{(pH - pKa_1)} + 10^{(2pH - pKa_1 - pKa_2)} + 10^{(3pH - pKa_1 - pKa_2 - pKa_3)}$$
$$- \frac{1}{1 + 10^{(pH - pKa_1)} + 10^{(2pH - pKa_1 - pKa_2)} + 10^{(3pH - pKa_1 - pKa_2 - pKa_3)}} \tag{8}$$

$$\mathbf{c}_{2j} = \alpha_{3j} - \alpha_{2j} = 1 - 2\alpha_{2j} = 10^{(pKa_1 - pH)}$$
$$+ 10^{(pH - pKa_2)} + 10^{(2pH - pKa_2 - pKa_3)}$$
$$- \frac{1}{10^{(pKa_1 - pH)} + 10^{(pH - pKa_2)} + 10^{(2pH - pKa_2 - pKa_3)}} \tag{9}$$

$$\mathbf{c}_{3j} = \alpha_{4j} - \alpha_{3j} = 1 - 2\alpha_{3j} = 10^{(pKa_2 - pH)}$$
$$+ 10^{(pH - pKa_3)} + 10^{(pKa_1 + pKa_2 - pH)}$$
$$- \frac{1}{1 + 10^{(pKa_2 - pH)} + 10^{(pH - pKa_3)} + 10^{(pKa_1 + pKa_2 - pH)}} \tag{10}$$

where α_1, α_2, α_3 and α_4 relate to the mole fraction of each species that forms during the titration, at the jth pH. The estimated values of pK_a used for constructing the matrix, $\mathbf{C}$, that applied to solve the equation systems as a initial matrix as follows:

$$\hat{\mathbf{V}} = \mathbf{A}\mathbf{C}^{T}(\mathbf{C}\mathbf{C}^{T})^{-1} \tag{11}$$

which the best fit has found by minimization of the value of the squared error between the actual data and the model using the following object function:

$$\chi^2 = \frac{1}{m(n-r)} \sum_i \sum_j (\mathbf{A}_{ij} - \hat{\mathbf{A}}_{ij})^2 \tag{12}$$

where m, n and r as the same values defined above and where,

$$\hat{\mathbf{A}} = \hat{\mathbf{V}} \mathbf{C} \tag{13}$$

The optimum set of pK_as to fit a given data matrix can be found by minimizing χ^2 with least-squares curve fitting minimization as such SIMPLEX or Marquardt–Levenberg algorithms in MATLAB [42–45].

2.2.4. Analysis of the A–pH curve

In the classical single wavelength method or A–pH curve analysis method, an absorbance vector, y used as experimental data, which contains the absorbance values at the maximum wavelength of the acidic or basic form spectrum, and related to concentration matrix and molar absorptivities of all species by Eqs. (2), (11) and (16).

Using initial estimates of the acidity constants and mass balance equations the concentration of all relevant species are calculated $\mathbf{C}$. Then an estimated value of the vector $\mathbf{v}$, is obtained by a linear least squares approach, Eq. (11). Now the absorbance vector $\hat{\mathbf{a}}$ (vector form) reconstructed, Eq. (13), using estimated value of $\mathbf{C}$ and $\hat{\mathbf{v}}$ (vector form), and the residual vector $\mathbf{e} = \mathbf{a} - \hat{\mathbf{a}}$ used as an objective function to refine the acidity.

2.2.5. New algorithm used by DATAN 3.1

In this study, we evaluate a new algorithm for the analysis of spectrophotometric data in multiwavelength mode using combination of hard and soft modeling. In the algorithm the data n by m matrix, $\mathbf{A}$, is decomposed into scores and loading by PCA or SVD as shown in Eqs. (1)–(5).

In the above equations, the rotation matrix $\mathbf{R}$ is not known. The matrix $\mathbf{C}$ is also unknown, but by assuming a chemical model, we can compute $\mathbf{C}$ from a small set of unknown variables:

$$\mathbf{C} = \mathbf{C}(x_1, x_2, \ldots, x_p) \tag{14}$$

In the case of protolytic equilibria, the unknown variables, x_i, are the equilibrium constants. For example, consider the equilibria of a diprotic acid:

$$H_2A \overset{K_1}{\rightleftharpoons} HA^- + H^+$$
$$HA^- \overset{K_2}{\rightleftharpoons} A^{2-} + H^+$$
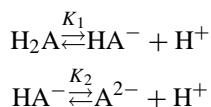
with equilibrium constants being K_1 and K_2, respectively. It is not hard to see that:

$$c_A = \frac{K_1 K_2}{c_H^2 + K_1 c_H + K_1 K_2}$$
$$c_{HA} = \frac{K_1 c_H}{c_H^2 + K_1 c_H + K_1 K_2} \tag{15}$$
$$c_{H_2A} = \frac{c_H^2}{c_H^2 + K_1 c_H + K_1 K_2}$$

where c_A denotes the analytical concentration of acid and similarly for the other species (charges were omitted for simplicity). The total concentration of acid has been normalized to one. Each row in $\mathbf{C}$ consists of the three concentrations at value of pH, which is known from the titration, so that with the knowledge of only the equilibrium constants K_1 and K_2, the entire matrix $\mathbf{C}$ can be computed. The number of unknowns has thus been reduced from $3n$ (the size of $\mathbf{C}$) to 2 (the number of equilibrium constants).

Any system, where a model of the chemical equilibria is known, can be treated in the same way as above to find an expression of the form given by Eq. (14). The introduction of a model means that the number of unknown variables is reduced from rn (the dimension of $\mathbf{C}$) to p. Ideally, p should be a small number.

To determine the unknowns $\mathbf{R}$ and $x_1, x_2, \ldots, x_p$, we use Eq. (4) and solve the minimization problem:

$$\text{Min}_{x, R} \left\| \mathbf{T} - \mathbf{C}(x_1, x_2, \ldots, x_p) \mathbf{R} \right\|_F \tag{16}$$

It would have been possible to minimize $\left\| \mathbf{TR}^{-1} - \mathbf{C} \right\|$, but in practice it has find out that the minimization given by Eq. (16) works basically better. Setting the derivative of the expression (Eq. (16)) to be minimized equal to zero gives the equations:

$$\mathbf{C}^T\mathbf{CR} - \mathbf{C}^T\mathbf{T} = 0$$
$$(\mathbf{T} - \mathbf{CR}) \circ \left(\frac{\partial \mathbf{C}}{\partial \mathbf{x}_i} \mathbf{R} \right) = 0 \tag{17}$$

where $\circ$ denotes element by element multiplication of matrices and $\partial \mathbf{C}/\partial x_i$ is element wise differentiation of $\mathbf{C}$ matrix with respect the unknown parameters x_i. Here, $\mathbf{R}$ is directly computed as $(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{T}$ from the upper equation. Inserting this into the second equation gives the non-linear system of p equations and p unknowns:

$$f_i(x_1, x_2, \ldots, x_p) = (\mathbf{T} - \mathbf{CR}) \circ \left(\frac{\partial \mathbf{C}}{\partial \mathbf{x}_i} \mathbf{R} \right) = 0,$$
$$i = 1, 2, \ldots, p. \tag{18}$$

To solve Eq. (18), we use Newton's method, but with numerical evaluation of the Jacobian matrix, through difference quotients, e.g.:

$$\frac{\partial f_i}{\partial x_1} = \frac{f_i(x_1 + h, x_2, \ldots, x_p) - f_i(x_1, x_2, \ldots, x_p)}{h} \tag{19}$$

where h is a small number. Since p is usually small, the extra function evaluations are not too time consuming and expensive. Now upon solution of the Eq. (18) we have a better estimation of the x_i parameters and the program controls goes up to the Eq. (14) and the calculations continue until convergence achieved. As it is clear, this algorithm has some priority over algorithm used in DATAN 2.1 in which the object function simply just evaluated for different series of acidity constants without any using of the fitting results but numerical values of residuals.

3. Experimental

The ability of the new algorithm in comparison to the previous established computer program was tested by different simulated and experimental data sets. The next two sections are the description of the way of data simulation and gathering.

3.1. Artificial data set

The various artificial or theoretical data sets were generated using the combination of the Gaussian curves to construct the spectral profiles of each species of the assumed triprotic acid molecule at the range 400–600 nm with resolution of 2 nm. Different random noises with mean zero and standard deviations 0, 1, 2 and 3% of the maximum absorbance values are loaded on the simulated spectral profiles. The concentrations of the four species of a triprotic acids were generated using SPECFIT/32 acid dissociation routine. The pH range selected was between 2.00 and 12.00 pH units.

The range of pK_as of the artificial acids in simulation were 2–5, 6–8 and 9–11 for pK_{a_1}, pK_{a_2} and pK_{a_3}, respectively.

3.2. Reagents and apparatus

The experimental details are reported in previous papers [19,46,47] and here a short explanation is presented. The analytical reagent grade 4-(2-pyridylazo)resorcinol (PAR), acetic acid, boric acid, phosphoric acid, sodium hydroxide and potassium nitrate were all from Merck Company and Fluka. Extra pure acetonitrile (AN), methanol (MeOH) and dioxane (DX) were used. The absorption spectra were recorded using a CECIL 9050 spectrophotometer with 1 nm spectral bandpass and scan speed 250 nm/min. The pH measurements were made using a Metrohm 692 pH meter equipped with a glass calomel combined electrode. To calibrate the pH meter in the various binary AN–water, MeOH–water and DX–water mixtures used, 0.01 M solutions of oxalate and succinate buffers were employed. The reference values of the pH, according to the activity scale, of these buffer solutions in different AN, MeOH and DX + water mixtures have been reported previously [48–50]. In all experiments, the ionic strengths of the solutions used, were kept constant at 0.1 M using potassium nitrate as the supporting electrolyte and the titration performed under inert atmosphere using argon gas bubbled through the titration vessel. A stock solution (2.0×10^{-5} M) of PAR was prepared. The acidic PAR solution titrated with a concentrated NaOH solution to reduce the dilution effect. The pH values in AN, MeOH and DX + water solvent mixtures were corrected using the equation pH* = pH(R) − δ, where pH* is the corrected reading and pH(R) is the pH-meter reading obtained in a partially aqueous organic solvent, determined by Douheret [49,50]. All measurements were carried out at the temperature (25.0 ± 0.5) °C.

3.3. The criteria for accepting a chemical model

Havel and Meloun presented different diagnostic devises to evaluate a chemical model [16,32]. The most important among them are partial correlation coefficients, goodness-of-fit-test, physical meaning of the spectral profiles and concentration distribution diagrams, β values which significantly different from zero (3S.D. < β value) and previous chemical experience of the understudy system.

3.3.1. Partial correlation coefficients

The partial correlation coefficient(s) values explain the quantity of the interdependency of the unknown parameters, i.e. indicates the correlation between two unknown parameters when the other(s) are kept constant. Fundamentally, correlation coefficients have values between −1 and +1. +1 or −1 mean complete dependence and zero value show totally independence. Generally, two species, which show complete correlation, could not be included in the model. These species have very relevant formation constants and the increasing

or decreasing of one component may compensate with the other.

3.3.2. Goodness-of-fit-test

This test contains the criteria for testing the correctness of the proposed chemical model. To evaluate whether or not the chemical model described the systematic variations in the experimental data adequately, the residuals errors should be analyzed. Commonly to recognize the best or real chemical model when there are several possible models the error analysis using a goodness-of-fit test is a good choice. Residuals vectors at ith wavelength are simply calculated using experimental ($\mathbf{A}_{exp,i}$) and calculated ($\mathbf{A}_{calc,i}$) absorbance values by $r_i = \mathbf{A}_{exp,i} - \mathbf{A}_{calc,i}$. The visual inspection of the residual plots may reveal the appropriateness of minimization process. The outliers and gross errors are simply identified in row (solutions) or columns (wavelengths) of the experimental absorbance data. The standard deviations of the residual values can easily be computed and compared with the standard deviations of the experimental correspondences. Detection of sign changes, any systematic variations (trend), an abrupt shift of level in the experiment in addition to detection of the outliers can be explored by visual inspection of the residual plot. The values of the mean and standard deviation of the residual close to zero and/or as much as a known experimentally limited quantities show a good fitting process and proposed chemical model. An squared sum of residual less than 0.01 is consider as a good fit. Hamilton R-factor is expressed as a relative percent fit of <0.5% is taken as excellent fit and >2% is a poor one.

3.3.3. The physical meaning of the estimated parameters, molar absorptivities and species concentrations

There are some physical constraints, which are generally applied to concentrations of species and their molar absorptivities. Non-negativity of concentrations and molar absorptivities are guaranteed experimentally and current algorithms like non-negative least squares (NNLS) routine used in SQUAD forced to zero the negative values during the iterations of the program. The log of β and molar absorptivities values should be in a regular range, neither very high nor too low. Meloun et al. [33] stated that the absolute values of standard deviation of unknown parameters gives information about the hyperparabolic error surface in the neighborhood of the pit. For well conditioned parameters, the flat error surface (residual versus parameters), is a regular ellipsoid and standard deviations are reasonably low. High standard deviations values are found with ill-conditioned parameters. The related standard deviation, $s(\beta_j) \times F_a < \beta_j$ should be fulfilled where F_a is equal to 3. The distributions of the standard deviations of the molar absorptivities must be Gaussian, otherwise, erroneous estimates of molar absorptivities are obtained.

3.3.4. Previous scientific reported quantities

It is wise before starting to analyze the present experimental data, to search for the scientific library sources to have a

good default about the number of ionizing groups, numerical values of relevant acidity constants and the probable spectral traces of all expected components. These information's help us to use very good and very close to final results values as initial guesses in minimization process. This is critical when the numbers of unknowns are high and the risk of the local minima destroys the output of the non-linear regression analysis of the spectroscopic data.

4. Results and discussion

4.1. Absorption spectra

Two set of simulated and experimental absorption spectra are prepared to examine the ability of the new algorithm in determination of acidity constants. The first set obtained using the acid/dissociation simulate function of SPECFIT/32 program with specifications mentioned in the experimental section. The quantity of added noise to the generated absorption spectra is 0, 1, 2 and 3%. A sample three dimensional plot of the generated absorption spectra loaded with a 1% constant noise is shown in Fig. 1 which has acidity constants 4.5, 7.0 and 9.5 for pK_{a_1}, pK_{a_2} and pK_{a_3}, respectively. The wavelength and pH range of the spectra are 400–600 nm and 2.00–12.00 pH units, respectively. The second set are experimental data obtained for the titrations of an acidic 2×10^{-5} M PAR solutions by a standard solution of a base in the 380–600 nm and pH range 2.00–13.50. The spectrophotometric titrations were carried out in different aqueous—AN, MeOH and DX mixtures. A typical absorption spectra of the acid–base titration of PAR in aqueous solution is shown in Fig. 2.

4.2. Number of absorbing species

The number of light absorbing species in the two sets of absorption spectra are obtained by different approximate
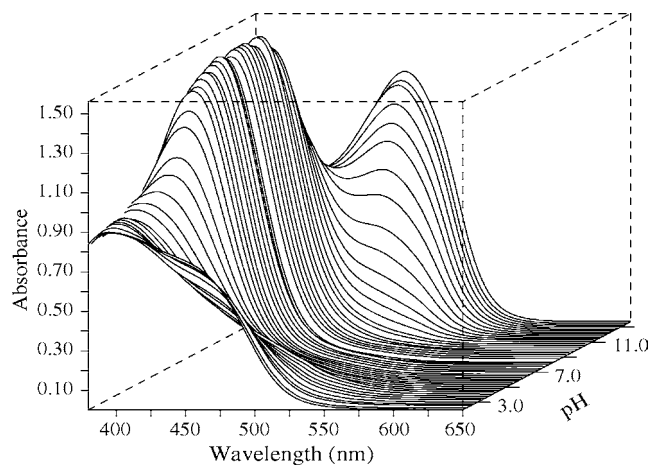
Fig. 2. The absorption spectra of [PAR] = 2.0×10^{-5} M in aqueous solution in different pH.

methods requiring no knowledge of the instrumental error of the absorbance data, $s_{inst}(A)$. Many of these methods are empirical functions. The most common utilized functions are eigenvalues (EV), reduced eigenvalues (Red. EV), residual standard deviation (R.S.D.), root mean square error (rms), residual percentage variance (RPV), imbedded error (IE), Malinowski indicator function (M Indicator Fcn), third derivative of IE function (TD), factor indicator functions (IND) and real error indicator. The details of these empirical functions are presented in [31,32,51]. The application of the indicator functions to the experimental data set is shown in Fig. 3. As it is expected for a three protic acids with four spectroscopically distinct components, some of the utilized functions distinguish four significant factors for two sets of data.

4.3. Determination of acidity constants of simulated and experimental data sets

The acidity constants of the simulated absorption spectra loaded with different noise levels are calculated by different programs in conjunction with the new algorithm. All results are shown in the Table 1. The general outputs of these programs are, calculated spectral profiles based on the optimum values of the acidity constants, distribution diagram based on the known pH values and the current acidity constants. The reconstructed absorption spectra and residual plot based on the calculated and original absorption spectra, acidity constants and their standard deviation and error surface plots are another important output of the mentioned programs. The outputs of the new algorithm, are the spectral profiles of each species and concentration distribution diagram of simulated absorbance data, from the computer fitting of the simulated absorption spectra with a 3% constant noise for an assumed acid with pK_{a_1}, pK_{a_2} and pK_{a_3} equal to 4.50, 7.00 and 9.50, respectively, are shown in Fig. 4. As it can be seen from Table 1 and the results of the new algorithm are consistent with the other programs and the comparison of these results
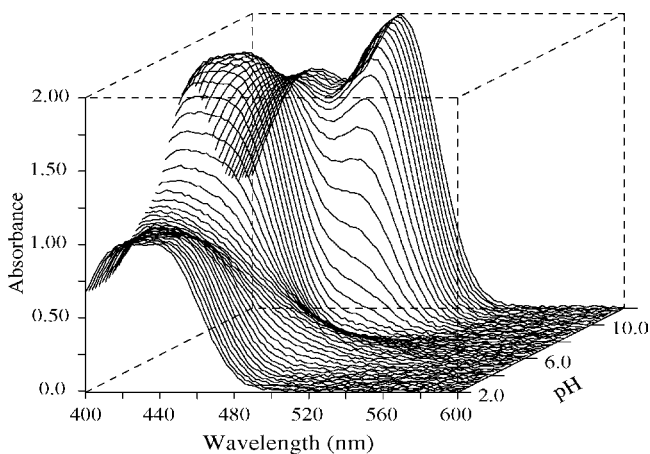

Fig. 1. The simulated absorption spectra of an acid with $pK_{a_1} = 4.5$, $pK_{a_2} = 7.0$ and $pK_{a_3} = 9.5$ loaded with 1% constant noise.

Table 1
The pK_a values and statistics for simulated data using different programs

Data set	Method	pK_{a_1}	pK_{a_2}	pK_{a_3}	r.s.s.[a]	s.d.r.[b]	R-factor (%)	NSF[c] EV	REI	MIF	REV	TD[e]	PCC[d] r_{12}	r_{13}	r_{23}
0% Noise	SPECFIT/32	4.500 ± 0.00039	7.00 ± 0.00011	9.500 ± 0.00007	9.25e−7	4.85e−5	0.095	3	4	4	3	4	−0.163	−0.028	−0.192
	SQUAD[f]	4.50 ± 0.0003	7.00 ± 0.0001	9.50 ± 0.0001	0.0001	0.0003	0.081						0.386	0.944	0.167
	New algorithm	4.50 ± 0.0629	6.99 ± 0.0921	9.50 ± 0.1157	1.05e−8	9.7e−6	0.001								
	Single λ	4.499	6.999	9.501	0.169	0.058	0.147								
	Difference spectra	4.500	7.000	9.500	1.07e−8	9.76e−6	0.001								
	DATAN 2.1	4.50	6.99	9.50	0.001	0.0004	0.01								
1% Noise	SPECFIT/32	4.501 ± 0.005	6.999 ± 0.001	9.4993 ± 0.0007	0.0026	0.0014	1.371	3	4	4	3	4	−0.162	−0.028	−0.190
	SQUAD	4.499 ± 0.008	7.000 ± 0.002	9.501 ± .001	0.053	0.005	1.202						0.609	0.559	0.70
	New algorithm	4.490 ± 0.011	6.99 ± 0.012	9.49 ± 0.010	0.0002	0.001	0.158								
	Single λ	4.483	7.002	9.499	5.5e−5	0.001	0.074								
	Difference spectra	4.497	7.001	9.499	2.65e−4	0.0014	0.201								
	DATAN 2.1	4.49	7.00	9.49	0.0004	0.004	0.245								
2% Noise	SPECFIT/32	4.498 ± 0.0082	6.996 ± 0.0024	9.498 ± 0.00137	0.001	0.003	1.405	3	4	4	3	4	−0.165	−0.028	−0.192
	SQUAD	4.73 ± 0.073	7.19 ± 0.058	9.31 ± 0.05	0.49	0.18	1.050						0.605	0.556	0.701
	New algorithm	4.49 ± 0.0130	6.99 ± 0.0269	9.50 ± 0.0346	6.62e−4	0.002	0.317								
	Single λ	4.424	7.002	9.488	2.19e−4	0.002	0.147								
	Difference spectra	4.498	6.996	9.498	0.0011	0.003	0.405								
	DATAN 2.1	4.49	6.99	9.49	0.0045	0.005	0.78								
3% Noise	SPECFIT/32	4.489 ± 0.0137	7.003 ± 0.038	9.500 ± 0.002	0.0025	0.005	1.621	4	4	4	4	4	−0.161	−0.028	−0.193
	SQUAD	4.495 ± 0.0237	6.949 ± 0.007	9.490 ± 0.004	0.45	0.018	1.303						0.616	0.568	0.70
	New algorithm	4.480 ± 0.020	6.990 ± 0.020	9.490 ± 0.020	0.0014	0.003	0.459								
	Single λ	4.123	6.993	10.490	0.001	0.004	0.338								
	Difference spectra	4.508	7.001	9.501	0.0025	0.005	0.621								
	DATAN 2.1	4.5	7.0	9.5		0.007	1.200								

The exact values are 4.5, 7.0 and 9.5 for pK_{a_1}, pK_{a_2} and pK_{a_3}, respectively.

[a] Residual sum of squares.
[b] Standard deviation of residuals.
[c] Number of significant factors.
[d] Partial correlation coefficient.
[e] Third derivative of IE function ref. [31].
[f] Number of wavelength is 37 and number of solution is 49, EV, eigenvalue; REI, real error indicator; MIF, Malinowski indicator function, REV, reduced eigenvalue.
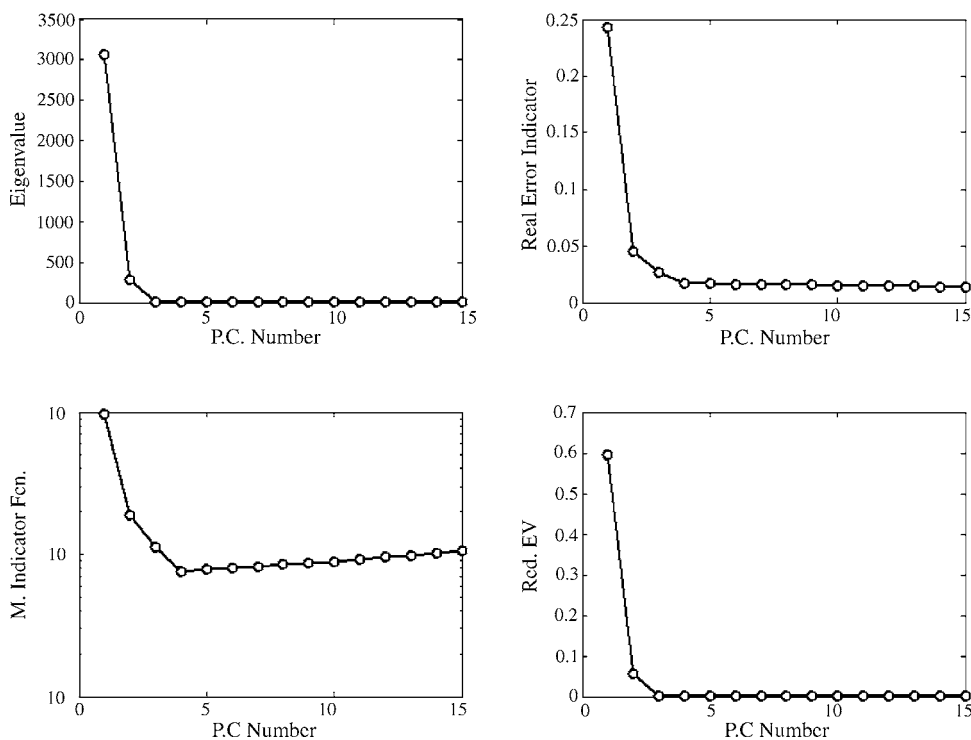
Fig. 3. The plot of different indicator functions for PAR absorption spectra in aqueous solution.

verify the good ability of the new algorithm to deconvolution of the absorption data matrix with respect to the other programs. The *A*–pH curve fitting is depicted in Fig. 5 at a selected wavelength. As it can be seen from Fig. 5 the distribution pattern of the experimental points around the calculated curve with zero mean and standard deviation. Its magnitude is the same as the original noise added to the absorbance in the simulation process, and reveals that the selected chemical model is a good model to span all of the systematic variation in the data.

All experimental data set were refined by different computer programs and also with the new algorithm. The acidity constants and some other useful statistical parameters are also calculated and listed in Table 2. The pure spectral profile of each species and concentration distribution diagrams were calculated by computer refinement of absorption data in all solvent systems used and the results of the PAR in aqueous solution is shown in Fig. 6 for new algorithm. The computer plot of the data refinement at a selected wavelength shows the fitting process by *A*–pH curve method at Fig. 7. The more or less equal distribution of the experimental points around the calculated curve indicate the correctness of the proposed chemical model to span all systematic variations. The three dimensional residual plots, results of the fitting (non-linear curve fitting) of the experimental data of PAR in aqueous solution to a triprotic solution equilibria using SQUAD, new algorithm and SPECFIT/32 are shown in Fig. 8. The interval made by upper and lower limits of the new algorithm residual plot is narrower than the other two residual plots of SQUAD and SPECFIT/32. This is also indicates the successful appli-

cation of the new algorithm in fitting process of the chemical model with experimental data.

The other practical and useful indicators for goodness of fit test are the calculated Hamilton *R*-factor, squared sum of the residuals, mean of residuals, standard deviation of the residuals, standard deviation of the calculated acidity constants and partial correlation coefficients from the minimization of the simulated and experimental data are included in Table 2. The results are in acceptable limit of these types of outputs, confirm the chemical model in one hand, and verify the proposed algorithm with respect to other programs. The comparison of the statistical parameters calculated from a residual analysis reveals, which the proposed algorithm has significant priorities over other previously presented such as SPECFIT/32 and/or SQUAD.

The partial correlation coefficients show the interrelations of the unknown parameters. The SPECFIT/32 and SQUAD programs calculate the partial correlation coefficients and are listed in Table 2. The obtained values did not show severe interrelations, the unknown parameters are almost independent, and the unknown values can be obtained without ambiguity of the absolute values. The standard deviation of the calculated parameters have significant difference to the parameters and all fulfill the condition; $s(\beta_j) \times F_a < \beta_j$, where F_a is 3 and the $s(\beta_j)$ and β_j. A one tailed student *t*-test at 95% confidence level and one degree of freedom ($t \geq 6.314$) using Table 2 showed the significance difference between β values and their associated standard deviations. Hamilton *R*-factor has the lowest values for the new algorithm with respect to the remaining programs which again shows the quality of

Table 2
The pK_a values and goodness of fit statistics for PAR experimental data in different mixed solvents using different programs

Solvent mixture	Method	pK_{a_1}	pK_{a_2}	pK_{a_3}	r.s.s.[a]	s.d.r.[b]	R-factor (%)	NSF[c]					PCC[d]		
								EV	REI	MIF	REV	TD[e]	r_{12}	r_{13}	r_{23}
Pure water	SPECFIT/32	3.04 ± 0.04	5.47 ± 0.014	12.25 ± 0.007	0.008	0.006	1.41	3	4	4	3	4	−0.215	0.027	−0.079
	SQUAD[f]	3.03 ± 0.01	5.48 ± 0.01	12.25 ± 0.01	0.053	0.005	1.2						0.327	0.173	0.264
	New algorithm	3.04 ± 0.01	5.48 ± 0.01	12.26 ± 0.01	0.020	0.0005	0.39								
	Single λ	2.95	5.49	12.26	0.010	0.014	1.29								
	Difference spectra	3.04	5.47	12.25	0.008	0.006	1.42								
	DATAN 2.1	3.07	5.50	12.24	0.043	0.004	0.89								
10% AN[g]	SPECFIT/32	2.7 ± 0.2	5.6 ± 0.1	12.16 ± 0.06	0.071	0.014	2.025	4	4	4	4	4	−0.216	0.0169	−0.0467
	SQUAD	2.79 ± 0.07	5.48 ± 0.07	12.24 ± 0.04	1.469	0.0436	1.300						0.565	0.540	0.782
	New algorithm	2.73 ± 0.01	5.59 ± 0.01	12.16 ± 0.03	0.0012	0.0025	1.059								
	Single λ	2.63	5.62	12.17	0.018	0.025	5.165								
	Difference spectra	2.73	5.59	12.16	0.071	0.014	8.026								
	DATAN 2.1	2.73	5.59	12.16	0.0045	0.007	1.89								
20% AN	SPECFIT/32	2.7 ± 0.2	5.6 ± 0.1	13.08 ± 0.08	0.081	0.017	2.955	3	4	4	3	4	−0.211	0.0117	−0.0315
	SQUAD	2.67 ± 0.06	5.70 ± 0.06	12.95 ± 0.03	0.930	0.0347	2.401						0.562	0.551	0.665
	New algorithm	2.72 ± 0.02	5.89 ± 0.04	13.07 ± 0.01	0.0016	0.0027	1.244								
	Single λ	1.88	6.01	13.0	0.025	0.029	6.125								
	Difference spectra	2.73	5.89	13.08	0.081	0.017	8.953								
	DATAN 2.1	2.72	5.89	13.07	0.064	0.045	2.01								
30% AN	SPECFIT/32	2.4 ± 0.1	6.18 ± 0.08	12.91 ± 0.06	0.033	0.010	2.731	3	4	4	3	4	−0.141	0.0196	−0.0691
	SQUAD	5.11 ± 0.09	8.63 ± 0.08	13.07 ± 0.06	1.278	0.0406	2.122						0.634	0.599	0.813
	New algorithm	2.37 ± 0.09	6.17 ± 0.09	12.91 ± 0.10	0.0005	0.0017	0.738								
	Single λ	1.86	6.27	12.86	0.014	0.022	4.573								
	Difference spectra	2.38	6.18	12.91	0.033	0.010	5.735								
	DATAN 2.1	2.37	6.17	12.91	0.0043	0.033	1.980								
40% AN	SPECFIT/32	2.4 ± 0.2	6.4 ± 0.1	11.03 ± 0.1	0.083	0.016	3.443	3	4	4	3	4	−0.114	0.0305	−0.0122
	SQUAD	3.4 ± 0.1	7.35 ± 0.08	11.40 ± 0.05	1.245	0.0401	2.560						0.498	0.311	0.497
	New algorithm	2.29 ± 0.15	6.35 ± 0.14	11.03 ± 0.15	0.0015	0.0024	1.278								
	Single λ	2.74	6.27	11.62	0.025	0.031	6.183								
	Difference spectra	2.30	6.35	11.03	0.083	0.016	9.441								
	DATAN 2.1	2.39	6.35	11.03	0.042	0.006	2.00								
50% AN	SPECFIT/32	2.22 ± 0.08	6.48 ± 0.04	12.99 ± 0.04	0.011	0.006	3.329	4	4	4	4	4	−0.090	0.0127	−0.0648
	SQUAD	4.2 ± 0.1	7.38 ± 0.08	10.9 ± 0.1	1.090	0.0375	0.976						0.528	0.503	0.473
	New algorithm	2.22 ± 0.08	6.47 ± 0.05	12.98 ± 0.07	0.0003	0.0014	0.608								
	Single λ	1.85	6.48	12.95	0.004	0.012	2.351								
	Difference spectra	2.22	6.48	12.98	0.011	0.006	3.344								
	DATAN 2.1	2.22	6.46	12.98	0.004	0.0089	1.23								
10% MeOH	SPECFIT/32	3.0 ± 0.2	6.05 ± 0.07	12.19 ± 0.04	0.114	0.020	4.869	3	4	4	3	4	−0.219	0.0216	−0.0678
	SQUAD	3.06 ± 0.03	5.96 ± 0.01	12.18 ± 0.01	0.499	0.024	1.890						0.470	0.319	0.308
	New algorithm	2.99 ± 0.03	6.05 ± 0.08	12.19 ± 0.07	0.003	0.004	0.857								
	Single λ	3.01	6.04	12.22	0.072	0.052	4.403								
	Difference spectra	2.99	6.05	12.19	0.114	0.020	4.87								
	DATAN 2.1	3.01	5.8	12.11	0.0067	0.078	1.45								

Table 2 (*Continued*)

Solvent mixture	Method	pK_{a_1}	pK_{a_2}	pK_{a_3}	r.s.s.[a]	s.d.r.[b]	R-factor (%)	NSF[c]					PCC[d]		
								EV	REI	MIF	REV	TD[e]	r_{12}	r_{13}	r_{23}
20% MeOH	SPECFIT/32	2.87 ± 0.06	5.37 ± 0.04	12.66 ± 0.01	0.015	0.008	2.732	3	4	4	3	4	−0.246	0.0235	−0.056
	SQUAD	2.87 ± 0.02	5.37 ± 0.01	12.66 ± 0.01	0.160	0.013	0.977						0.794	0.767	0.953
	New algorithm	2.86 ± 0.03	5.37 ± 0.04	12.66 ± 0.07	0.0004	0.0016	0.480								
	Single λ	2.87	5.40	12.53	0.007	0.016	2.011								
	Difference spectra	2.87	5.37	12.66	0.015	0.008	2.73								
	DATAN 2.1	2.9	5.36	12.61	0.003	0.0047	0.98								
30% MeOH	SPECFIT/32	2.6 ± 0.2	5.6 ± 0.1	12.23 ± 0.06	0.179	0.029	3.192	3	4	4	3	4	−0.207	0.0205	−0.0273
	SQUAD	2.62 ± 0.08	5.62 ± 0.04	12.23 ± 0.02	0.297	0.062	2.308						0.382	0.203	0.538
	New algorithm	2.61 ± 0.04	5.61 ± 0.03	12.23 ± 0.03	0.002	0.0034	0.873								
	Single λ	2.58	5.55	12.20	0.030	0.035	4.008								
	Difference spectra	2.62	5.62	12.23	0.179	0.030	8.190								
	DATAN 2.1	2.67	5.67	12.22	0.056	0.045	2.11								
40% MeOH	SPECFIT/32	2.74 ± 0.06	6.27 ± 0.04	13.79 ± 0.04	0.040	0.009	2.881	4	4	4	4	4	−0.106	0.0159	−0.0537
	SQUAD	2.73 ± 0.02	6.28 ± 0.01	13.72 ± 0.01	0.340	0.0205	1.361						0.556	0.364	0.565
	New algorithm	2.73 ± 0.02	6.27 ± 0.01	13.72 ± 0.02	0.0002	0.0009	0.196								
	Single λ	1.89	6.34	13.79	0.025	0.032	3.432								
	Difference spectra	2.74	6.27	13.78	0.040	0.009	2.882								
	DATAN 2.1	2.8	6.11	13.77	0.001	0.003	0.567								
50% MeOH	SPECFIT/32	2.69 ± 0.11	6.37 ± 0.07	13.06 ± 0.06	0.120	0.0186	2.203	3	4	4	3	4	−0.137	0.0159	−0.0158
	SQUAD	2.67 ± 0.05	6.37 ± 0.03	13.07 ± 0.02	0.689	0.0314	2.089						0.670	0.474	0.620
	New algorithm	2.68 ± 0.05	6.37 ± 0.06	13.06 ± 0.05	0.0009	0.0017	0.449								
	Single λ	1.84	6.28	13.44	0.008	0.017	2.094								
	Difference spectra	2.69	6.37	13.06	0.120	0.018	5.204								
	DATAN 2.1	2.72	6.28	13.04	0.003	0.067	0.980								
10% DX	SPECFIT/32	2.96 ± 0.13	5.72 ± 0.09	12.3 ± 0.1	0.30	0.018	2.445	3	4	4	3	4	−0.320	−0.025	0.121
	SQUAD	2.8 ± 0.1	5.42 ± 0.07	12.11 ± 0.04	0.806	0.0323	2.000						0.478	0.412	0.504
	New algorithm	2.96 ± 0.04	5.72 ± 0.06	12.17 ± 0.07	0.0035	0.0022	0.695								
	Single λ	3.19	4.93	12.19	0.054	0.035	6.973								
	Difference spectra	2.97	5.7	12.18	0.297	0.018	6.448								
	DATAN 2.1	2.96	5.7	12.17	0.034	0.024	1.650								
20% DX	SPECFIT/32	2.82 ± 0.12	5.62 ± 0.09	12.24 ± 0.04	0.183	0.015	4.736	3	4	4	3	4	−0.421	−0.017	−0.056
	SQUAD	2.690 ± 0.03	5.91 ± 0.03	12.24 ± 0.02	0.479	0.0176	1.340						0.478	0.437	0.581
	New algorithm	2.81 ± 0.04	5.61 ± 0.04	12.24 ± 0.03	0.004	0.0023	0.667								
	Single λ	1.88	6.08	12.23	0.027	0.025	4.663								
	Difference spectra	2.82	5.61	12.24	0.183	0.015	4.736								
	DATAN 2.1	2.81	5.61	12.24	0.009	0.045	1.123								
30% DX	SPECFIT/32	2.77 ± 0.12	5.83 ± 0.09	12.32 ± 0.05	0.224	0.018	3.404	4	4	4	4	4	−0.419	−0.004	0.038
	SQUAD	2.66 ± 0.04	6.06 ± 0.04	12.31 ± 0.02	0.781	0.0230	2.718						0.466	0.412	0.504
	New algorithm	2.76 ± 0.09	5.83 ± 0.05	12.31 ± 0.01	0.006	0.0026	0.883								
	Single λ	2.59	6.21	12.32	0.028	0.026	4.587								
	Difference spectra	2.77	5.84	12.32	0.224	0.016	5.409								
	DATAN 2.1	2.76	5.83	12.31	0.012	0.0045	1.456								

	SPECFIT/32	2.51 ± 0.18	6.08 ± 0.10	12.19 ± 0.05	0.300	0.020	3.221	3	4	4	3	4	−0.245	0.0258	−0.0212
	SQUAD	2.53 ± 0.04	6.20 ± 0.03	12.19 ± 0.02	0.520	0.018	3.011						0.446	0.408	0.484
40% DX	New algorithm	2.43 ± 0.03	6.11 ± 0.06	12.19 ± 0.04	0.007	0.003	0.953								
	Single λ	2.23	6.28	12.19	0.036	0.029	4.907								
	Difference spectra	2.43	6.12	12.19	0.611	0.025	8.984								
	DATAN 2.1	2.43	6.11	12.19	0.045	0.045	2.222								
	SPECFIT/32	2.47 ± 0.06	6.56 ± 0.06	12.29 ± 0.04	0.156	0.013	2.641	3	4	4	3	4	−0.123	0.0185	−0.036
	SQUAD	2.51 ± 0.03	6.58 ± 0.03	12.29 ± 0.02	0.481	0.015	1.222						0.460	0.392	0.494
50% DX	New algorithm	2.47 ± 0.02	6.56 ± 0.03	12.29 ± 0.03	0.004	0.002	0.763								
	Single λ	2.35	6.59	12.26	0.018	0.019	3.047								
	Difference spectra	2.48	6.56	12.29	0.157	0.012	4.656								
	DATAN 2.1	2.47	6.56	12.29	0.005	0.045	1.455								

[a] Residual sum of squares.

[b] Standard deviation of residuals.

[c] Number of significant factors.

[d] Partial correlation coefficient.

[e] Third derivative of IE function ref. [31].

[f] Number of wavelength is 43 and number of solutions is 45.

[g] The percent is volume percent EV, eigenvalue; REI, real error indicator; MIF, Malinowski indicator function; REV, reduced eigenvalue.
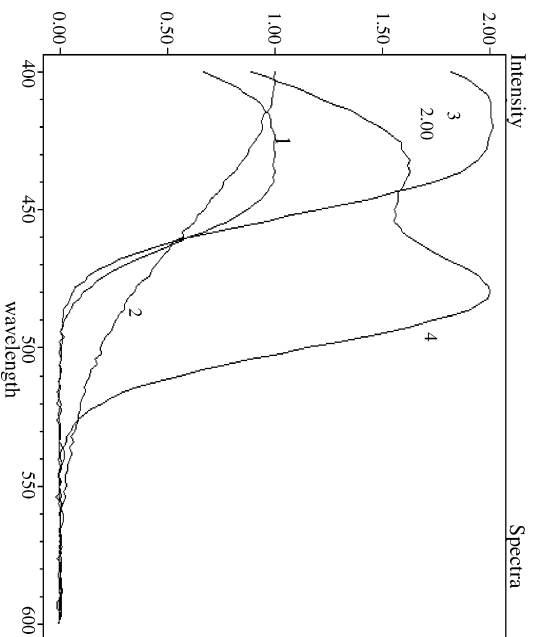
Fig. 4. Pure spectral profiles of H_3L^+ (1), H_2L (2), HL^- (3) and L^{2-} (4) for a theoretical acid by new algorithm and simulated spectra with 3% noise.
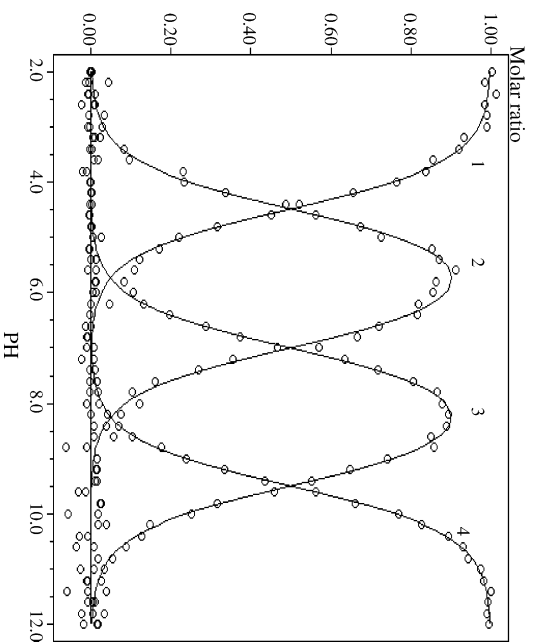

Fig. 5. The plot of experimental (points) and calculated absorbance data (line) vs. pH at 422 nm by A–pH curve method and 1% added noise.
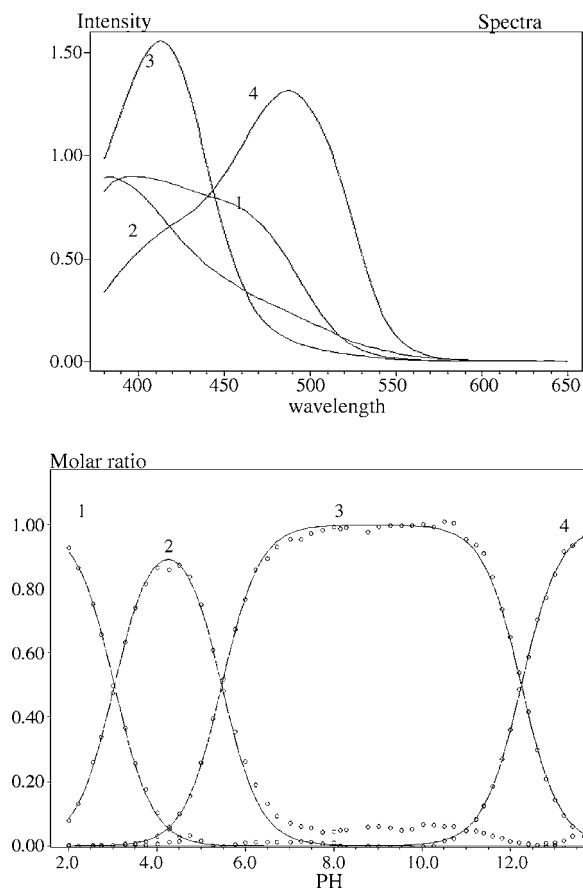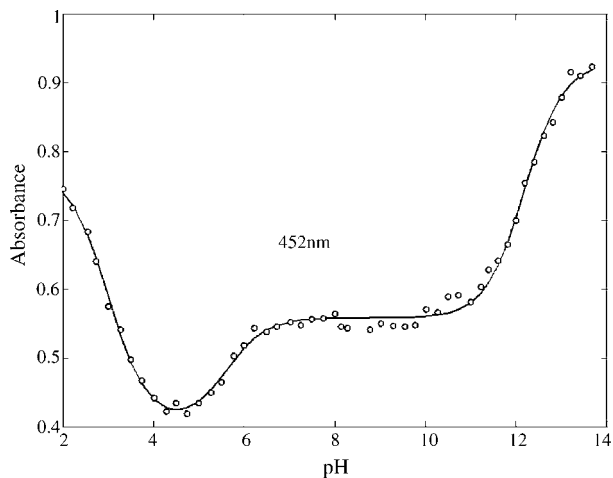
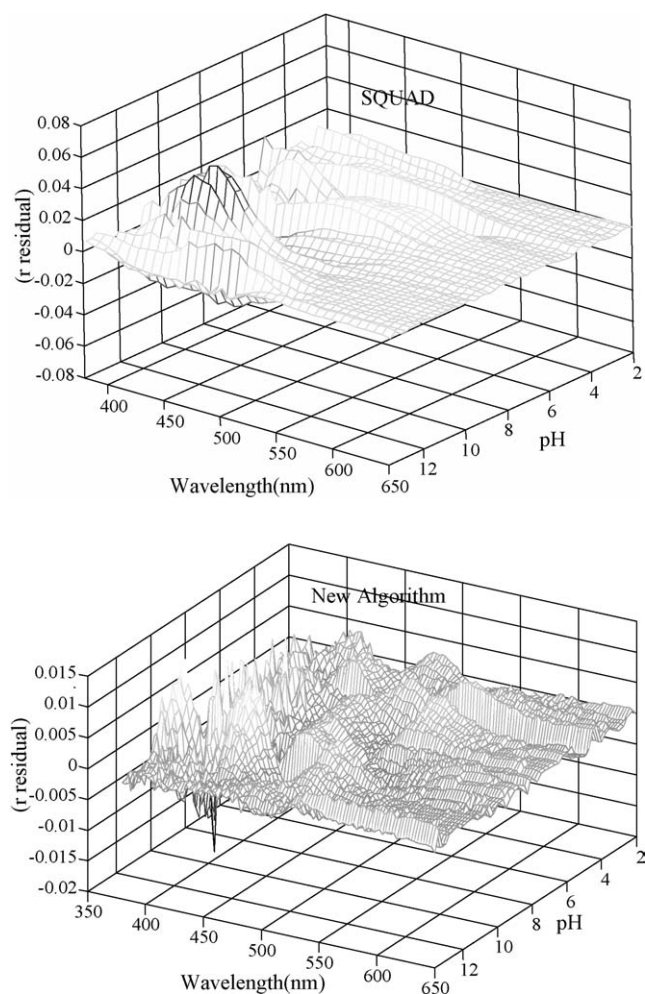Fig. 6. Pure spectral profiles and distribution diagram of PAR in aqueous solution H_3L (1), H_2L (2), HL (3) and L (4) obtained by new algorithm.

the fitting process which return to the quality of the experimental data and also the deconvolution of the spectral data to their pure spectra of the involve species at each pH value. Residual sum of squares (r.s.s.) and standard deviation of residuals (s.d.r.) are calculated for new algorithm along with the other programs for all experimental data sets and included

Fig. 8. Three dimensional residual plots of the absorption spectra of PAR in aqueous solution by SQUAD, New algorithm and SPECFIT/32.

Fig. 7. The plot of experimental (points) and calculated absorbance data (line) vs. pH of PAR in aqueous solution at 452.5 nm by A–pH curve method.

Table 3
The physical properties of utilized water and non-aqueous solvents

Solvent	Dielectric constant (D)	Donor number (DN)
Water	78	33
Methanol (MeOH)	32.6	19
Acetonitrile (AN)	36	14
Dioxane (DX)	2.2	14.8

in Table 2. The calculated values is an order of magnitude comparable to the other established programs and indicate the quality of the fitting of the new algorithm.

4.4. Effect of the solvent media on the acidity constants

The physical properties of the used solvents as co-solvent with water in order to study the effect of addition of non-aqueous solvents are listed in Table 3.

The pK_a values of PAR obtained in AN, MeOH and DX–water mixtures increase with percentage of AN, MeOH and DX. These variations could be explained by the fact that there is preferential solvation in these media that is related to the structural features of these mixtures. The composition of the immediate surroundings of a solute may differ from the composition of the bulk mixture. Preferential solvation is attributable to an excess or deficiency of molecules of one of the solvents in these surroundings [52]. If the solute displays no preference for the solvent molecules, the solvent composition in the primary coordination shell, in the immediate neighborhood of the solute, is the same as that in the bulk.

The data shown in Table 2 clearly illustrate the important influence of the nature of the solvent on the dissociation reactions. It has been shown that the solvating ability [53] (as expressed by the Guttmann donicity scale) and dielectric constant of the solvent play a fundamental role in dissociation reactions. Water is a solvent of high solvating ability (i.e. donor number, DN = 33.0) and dielectric constant ($\varepsilon = 78$), which can dissociate the acid and stabilize the produced anion and hydrogen ion.

It is interesting to note that there is actually a positive correlations of pK_a's of three dissociation steps (first step decreases and second and third steps increases) and the mole fraction of AN, MeOH and DX in the binary mixed solvents used. It has been reasonably assumed that preferential solvation of the charged particles by water is mainly responsible for such a monotonic dependence of the acidity constants of PAR on the solvent composition. It is clear that the dissociation of an uncharged acid in a solvent requires the separation of two ions of opposite charges. The work required to separate these charges is inversely proportional to the dielectric constant of the solvent. The energy required for dissociation is supplied by solvation of the ions, and also the proton transfer from acid to the solvent molecule supplies an additional energy. If the dielectric constant and the solvating ability of the solvent are decreased, more energy will be required to separate the anion and cation, and consequently the extent of dissociation of the acid will be lowered. Therefore, the decrease in the pK_a of the first step and the increase in those of the second and third steps are due to increasing the mole fraction of AN, MeOH and DX in the binary mixed solvents.

It is not surprising to mention that, the variations of the acidity constants are according to variation in the donor number and also dielectric solvents. As it is clear by addition of all organic solvents to water the second and third ionization steps increased and also the first step decreased. The degree of ariations is fully depending on the extent of the differences of physical parameters (donor number and dielectric constant) of the added organic solvents with water.

5. Conclusion

The acidity constants of a simulated tri-protic acid and also some experimental data of photometric titration of acidic solution of 4-(2-pyridylazo)resorcinol (PAR) with standard base solution were determined with some well known computer programs and new algorithm. The applicabilities of the new algorithm were compared with the other programs by using refined final acidity constants, pure spectra and distribution diagrams for all assumed species and error analysis of the residuals. According to statistical parameters, results from residual analysis show good superiority observed for new algorithm over previous tested programs. It is concluded that the new algorithm may be used for more complex thermodynamic or kinetic systems. Of course more successful examples are needed for generalization of the new algorithm.

References

[1] Y.H. Zhao, L.-H. Yuan, L.S. Wang, Bull. Environ.Contam. Toxicol. 57 (1996) 242.
[2] H. Rochester, Acidity Functions, Academic Press, New York, 1971.
[3] L.P. Hammet, Physical Organic Chemistry, McGraw-Hill, New York, 1940.
[4] D.J. Leggett, S.L. Kelly, L.R. Shiue, Y.T. Wu, D. Chang, K.M. Kadish, Talanta 30 (1983) 579.
[5] R.M. Alcock, F.R. Hartley, D.E. Rogers, J. Chem. Soc. Dalton (1978) 115.
[6] H. Gampp, M. Maeder, A.D. Zuberbuhler, Talanta 27 (1980) 1037.
[7] K. Nagano, D.E. Metzler, J. Am. Chem. Soc. 89 (1976) 2891.
[8] J.R. Siefker, Anal. Chim. Acta 52 (1970) 545.
[9] J.J. Kankare, Anal. Chem. 42 (1970) 1322.
[10] A.D. Zuberbuhler, Th.A. Kaden, Talanta 26 (1979) 1111.
[11] M. Briellmann, A.D. Zuberbuhler, Helv. Chim. Acta 65 (1982) 46.
[12] D.J. Leggett, W.A.E. McBryde, Anal. Chem. 47 (1975) 1065.
[13] D.J. Leggett, Anal. Chem. 49 (1977) 276.
[14] D.J. Leggett, S.L. Kelly, L.R. Shiue, Y.T. Wu, D. Chang, K.M. Kadish, Talanta 90 (1983) 579.
[15] M. Meloun, J. Havel, Folia Fac. Sci. Natl. Univ. Purk. Brun. 25, Chemia 17, Part 7, 1984.
[16] J. Havel, M. Meloun, in: D.J. Leggett (Ed.), Computation Methods for the Determination of Formation Constants, Plenum Press, New York, 1985.
[17] H. Gammp, M. Maeder, Ch. J. Meyer, A.D. Zuberbuhler, Talanta 32 (1985) 95.

[18] H. Gammp, M. Maeder, Ch. J. Meyer, A.D. Zuberbuhler, Talanta 32 (1985) 257.
[19] J. Ghasemi, A. Niazi, M. Kubista, A. Elbergali, Anal. Chim. Acta 455 (2002) 335.
[20] I. Scarmino, M. Kubista, Anal. Chem. 65 (1993) 409.
[21] M. Kubista, R. Sjoback, J. Nygren, Anal. Chim. Acta 302 (1995) 121.
[22] J. Nygren, A. Elbergali, M. Kubista, Anal. Chem. 70 (1998) 4841.
[23] L. Antonov, G. Gorgov, V. Petrov, M. Kubista, J. Nygren, Talanta 49 (1999) 106.
[24] J. Nygren, J.M. Andrade, M. Kubista, Anal. Chem. 68 (1996) 1706.
[25] S.D. Frans, J.M. Harris, Anal. Chem. 56 (1984) 466.
[26] M.A. Sharaf, D.L. Ilman, B.R. Kowalski, Chemometrics, Wiley, Chichester, 1986.
[27] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[28] D.R. Cox, D. Oakes, Analysis of Survival Data, Chapman and Hall, London, 1984.
[29] D.L. Massart, R.G. Brereton, R.E. Dessy, P.K. Hopke, C.H. Spiegelman, W. Wegscheider (Eds.), Chemometrics Tutorials, Elsevier, Amsterdam, 1990.
[30] D.L. Massart, W. Wegscheider, B.G. Vandeginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1990.
[31] A.K. Elbergali, J. Nygren, M. Kubista, Anal. Chim. Acta 379 (1999) 143.
[32] M. Meloun, J. Čapek, P. Mikšík, R.G. Brereton, Anal. Chim. Acta 423 (2000) 51.
[33] M. Meloun, J. Havel, E. Högfeldt, Computation of Solution Equilibria, Ellis Horwood, Chichester, 1988.

[34] J.J. Kankare, Anal. Chem. 42 (1970) 1322.
[35] M. Kubista, R. Sjoback, B. Albinsson, Anal. Chem. 65 (1993) 994.
[36] (a) H. Gampp, M. Maeder, Ch. J. Meyer, A. Zuberbühler, Talanta 32 (1985) 1133;
     (b) H. Gampp, M. Maeder, Ch. J. Meyer, A. Zuberbühler, Talanta 33 (1986) 943.
[37] M. Meader, Anal. Chem. 59 (1987) 527.
[38] A. Lorber, Anal. Chem. 56 (1984) 1004.
[39] G.H. Golub, F. Vanloan, Matrix Computations, John Hopkins University Press, Baitimore, MD, 1983.
[40] M. Maeder, A.D. Zuberbuhler, Anal. Chem. 52 (1990) 2220.
[41] J.A. Nelder, R. Mead, Comput. J. 7 (1965) 308.
[42] R. O' Nelll, Appl. Stat. 20 (1971) 338.
[43] S.N. Deming, S.L. Morgan, Anal. Chem. 45 (1973) 278A.
[44] S.L. Morgan, S.N. Deming, Anal. Chem. 46 (1974) 1170.
[45] P. Horst, Chemom. Int. Lab. Syst. 14 (1992) 5.
[46] J. Ghasemi, Sh. Ahmadi, M. Kubista, A. Forootan, J. Chem. Eng. 48 (2003) 1178.
[47] A. Rouhollahi, F.M. Kiaie, J. Ghasemi, M. Kubista, A. Niazi Pol., J. Chem., in press.
[48] R.G. Bates, Determination of pH, Theory and Practice, Wiley, New York, 1973.
[49] G. Douheret, Bull. Soc. Chim. Fr. (1967) 1412–1419.
[50] G. Douheret, Bull. Soc. Chim. Fr. (1968) 3122–3131.
[51] E.R. Malinowski, Factor Analysis in Chemistry, second ed., Wiley, Chichester, 1991.
[52] Y. Marcus, Y. Migron, Phys. Chem. 95 (1991) 400.
[53] V. Guttmann, Coordination Chemistry in Nonaqueous Solutions, Springer, New York, 1960.