

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Analysis of Chemical Processes, Determination of the Reaction Mechanism and Fitting of Equilibrium and Rate Constants

Marcel Maeder and Peter King

*Department of Chemistry, University of Newcastle, Australia  
Jplus Consulting Ltd, Perth,  
Australia*

## 1. Introduction

This chapter is intended to demonstrate some recent approaches to the quantitative determination of chemical processes based on the quantitative analysis of experimental spectrophotometric measurements. In this chapter we will discuss kinetic processes, equilibrium processes and also processes that include a combination of kinetic and equilibrium steps.

We also emphasise the advantage of 'global' multivariate (multiwavelength) data analysis which has the advantage of allowing the robust determination of more complex mechanisms than single wavelength analysis and also has the benefit of yielding the spectra of all the participating species.

Rather than dwell on the mathematical derivation of the complex numerical algorithms and a repetition of the fundamentals of non-linear regression methods and least squares fitting which are available from a wide variety of sources (Martell and Motekaitis 1988; Polster and Lachmann 1989; Gans 1992; Press, Vetterling et al. 1995; Maeder and Neuhold 2007), we aim to show the experimentalist how to obtain the results they are interested, using purpose designed global analysis software and a variety of worked examples. We will be using ReactLab, a suite of versatile and powerful reaction modelling and analysis tools developed by the authors. Other academic and commercial applications exist for multivariate and related types of analysis and the reader is encouraged to explore these for comparative purposes. All offer different features and benefits but will not be discussed here.

## 2. Spectrophotometry, the ideal technique for process analysis

Any spectroscopic technique is ideal for the analysis of chemical processes as there is no interference in the underlying chemistry by the measurement technique. This is in sharp contrast to say chromatographic analysis or other separation methods which are totally unsuitable for the analysis of dynamic equilibrium systems. Such methods are also of very limited use for kinetic studies which often are too fast on the chromatographic time scale of

typically tens of minutes to hours (except where reactions are first quenched and the intermediates stabilised). In contrast most forms of spectroscopy provide a completely non-invasive snapshot of a sample's composition at a single instant.

Amongst the different spectroscopies routinely available to the chemist, light absorption spectrophotometry in the UV-Visible (UV/Vis) is most common for several reasons: instruments are relatively inexpensive and accurate, they provide stable referenced signals as they are usually split or double beam instruments, there is a simple relationship between concentration and the measured absorbance signal (Beer-Lambert's law) and many compounds absorb somewhere in the accessible wavelength region. As a consequence there is a considerable amount of software available for the analysis of spectrophotometric data. This is the case both for kinetic and equilibrium investigations. NMR spectroscopy is a powerful method for structural investigations but it is less commonly used for quantitative analytical purposes. A theoretically very powerful alternative to UV/Vis absorption spectroscopy is FT-IR spectroscopy. The richness of IR spectra is very attractive as there is much more information contained in an IR spectrum compared with a relatively structureless UV/Vis spectrum. The main disadvantage is the lack of long term stability as FT-IR instruments are single beam instruments. Other difficulties include solvent absorption and the lack of non-absorbing cell materials, particularly for aqueous solutions. However, attenuated total reflection or ATR is a promising novel measurement technique in the IR. Near-IR spectroscopy is very similar to UV/Vis spectroscopy and is covered by the present discussions.

Of course fluorescence detection is a very sensitive and important tool particularly in kinetics studies and can yield important mechanistic information where intermediates do not possess chromophores and are therefore colourless or are studied at very low concentrations. In the main fluorescence studies are carried out at a single emission wavelength or adopting the total fluorescence approach (using cut-off filters), so there is no wavelength discrimination in the data. Whilst this type of measurement can be analysed by the methods described below and is essentially equivalent to analysing single wavelength absorption data. We will in the following discussion concentrate on the general case of processing multiwavelength measurements

### 3. The experiment, structure of the data

For kinetic investigations the absorption of the reacting solution is measured as a function of reaction time; for equilibrium investigations the absorption is recorded as a function of the reagent addition or another independent variable such as pH. Absorption readings can of course be taken at a single wavelength but with modern instrumentations it is routine and advantageous to record complete spectra vs. time or reagent addition. This is particularly prevalent with the use of photodiode array (PDA) based spectrophotometers and online detectors.

In the case of kinetics, depending on the rate of a chemical reaction the mixing of the reagents that undergo the reaction has to be done fast using a stopped-flow instrument or it can be done manually for slower reactions in the cuvette of a standard UV-Vis spectrometer with suitably triggered spectral data acquisition. A series of spectra are collected at time intervals following the mixing event to cover the reaction time of interest. The measured spectra change as the reaction proceeds from reagents to products (Wilkins 1991; Espenson 1995).

For equilibrium investigations the spectra of a series of pre-mixed and equilibrated solutions have to be recorded (Martell and Motekaitis 1988; Polster and Lachmann 1989). This is most commonly done as a titration where small amounts of a reagent are added stepwise to the solution under investigation. Titrations can be done in the cuvette, requiring internal stirring after each addition, prior to the absorption measurement, or the solutions can be mixed externally with transfer of the equilibrated solutions into the cuvette performed manually or using a flow cell and automatic pumping. In an alternative configuration optical probes can be coupled to the optical path in some spectrometers and placed into the solution contained in an external titration vessel (Norman and Maeder 2006). Often the pHs of the equilibrated titration solutions are recorded together with the absorption spectra where protonation equilibria are a feature of the mechanism.

For both kinetic and equilibrium investigations the measurement data can be arranged in a data matrix  $\mathbf{D}$  which contains row-wise the recorded spectra as a function of time or reagent addition. The number of columns of  $\mathbf{D}$  is the number of wavelengths,  $n_{lam}$ , over which the spectra are taken. For single wavelength data the matrix reduces to a single column (vector). The number of rows,  $n_{spectra}$ , corresponds to the number of spectra recorded during the process (one at each time interval for kinetics or reagent addition for an equilibrium titration). The dimensions of  $\mathbf{D}$  thus are  $n_{spectra} \times n_{lam}$ . For spectra taken on a mechanical scanning instrument, the number of wavelengths can be 1 to typically some 10 or 20 but for diode array instruments it can easily be in excess of 1000 depending on the solid state detector pixel resolution (typically these provide a resolution progression of 256, 512 and 1024 pixels). The number of spectra taken is typically much larger on a stopped-flow instrument equipped with a fast diode array detector with a typical minimum spectrum acquisition time of the order of a millisecond. Frequently a logarithmic time base is an option which enables both fast and slower events to be resolved in a single kinetic experiment. A graphical representation of a data matrix  $\mathbf{D}$  is given in Figure 1.

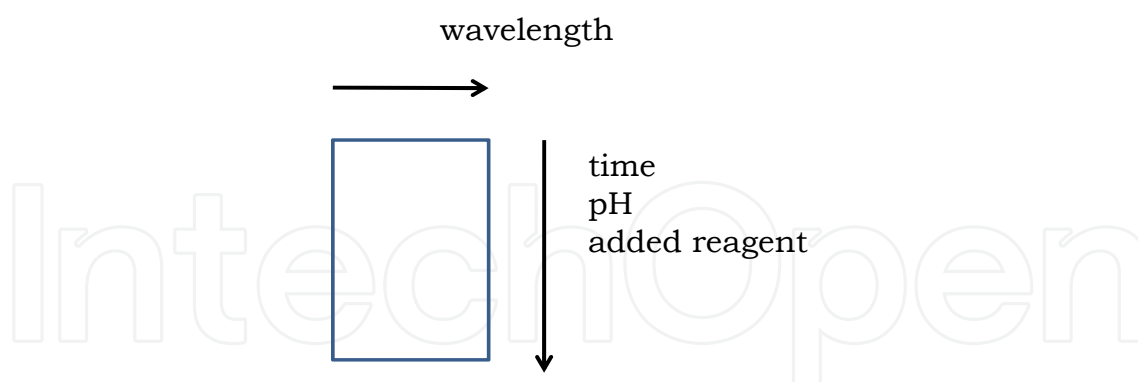


Fig. 1. Graphical representation of a data matrix  $\mathbf{D}$ , the spectra are arranged as the rows.

For both kinetic and equilibrium investigations, we obtain a series of spectra each of which represent the solution at one particular moment during the process. The spectra are taken as a function of time or reagent addition.

#### 4. Information to be gained from the measurements

The purpose of collecting this type of data is to determine the chemical reaction mechanism that describes the underlying process in terms of identifiable steps together with the

associated key parameters; the rates and/or equilibrium constants which define the interconversions and stabilities of the various species. This may initially be a purely academic exercise to characterise a novel chemical reaction for publication purposes but ultimately defines the behaviour of the participating species for any future research into this or related chemistry as well as being the foundation for commercially important applications e.g. drug binding interactions in pharmaceutical development or reaction optimisation in industrial processes.

The objective is therefore to find the chemical model which best fits the data and validate and refine this model with subsequent experiments under other conditions. The clear benefit of multi-wavelength measurements is that the model must satisfy (fit) the data at all measurement wavelengths simultaneously and this significantly helps the accurate determination of multiple parameters and also allows determination of the individual spectra of the participating species.

## 5. Beer-Lambert's law

Before we can start the possible ways of extracting the useful parameters from the measured data set, the rate constants in the case of kinetics, the equilibrium constants in the case of equilibria, we need to further investigate the structure of the data matrix **D**. According to Beer-Lambert's law for multicomponent systems, the total absorption at any particular wavelength is the sum over all individual contributions of all absorbing species at this wavelength. It is best to write this as an equation:

$$D(i,j) = \sum_{k=1}^{ncomp} C(i,k) \times A(k,j) \quad (1)$$

where:

$D(i,j)$ : absorption of the  $i$ -th solution at wavelength  $j$

$C(i,k)$ : concentration of the  $k$ -th component in the  $i$ -th solution

$A(k,j)$ : molar absorptivity of the  $k$ -th component at the  $j$ -th wavelength

$ncomp$ : number of components in the system under investigation.

Thus equation (1) represents a system of  $i \times j$  equations with many unknowns, i.e. all elements of **C** ( $n_{spectra} \times n_{comp}$ ) and all elements of **A** ( $n_{comp} \times n_{lam}$ ).

It is extremely useful to realise that the structure of Beer-Lambert's law allows the writing of Equation (1) in a very elegant matrix notation, Equation (2) and Figure 2

$$\mathbf{D} = \mathbf{C} \times \mathbf{A} \quad (2)$$

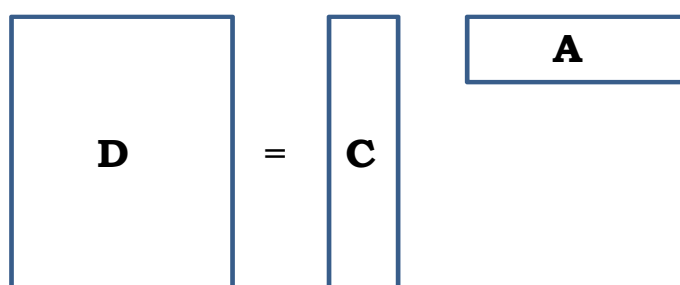


Fig. 2. Beer-Lambert's law, Equation (1) in matrix notation.

The matrix  $\mathbf{D}(n_{\text{spectra}} \times n_{\text{lam}})$  is the product of a matrix of concentrations  $\mathbf{C}(n_{\text{spectra}} \times n_{\text{comp}})$  and a matrix  $\mathbf{A}(n_{\text{comp}} \times n_{\text{lam}})$ .  $\mathbf{C}$  contains as columns the concentration profiles of the reacting components and the matrix  $\mathbf{A}$  contains, as rows, their molar absorption spectra.

Equations (1) and (2) and Figure 2 represent the ideal case of perfect absorption readings without any experimental noise. This of course is not realistic and both equations have to be augmented by an error term,  $R(i,j)$  which is the difference between the ideal value and its measured counterpart, equation (3) and Figure 3.

$$D(i,j) = \sum_{k=1}^{n_{\text{comp}}} C(i,k) \times A(k,j) + R(i,j) \quad (3)$$

$$\mathbf{D} = \mathbf{C} \times \mathbf{A} + \mathbf{R}$$

Fig. 3. Beer-Lambert's law including the residuals.

The goal of the fitting is to determine that set of matrices  $\mathbf{C}$  and  $\mathbf{A}$  for which the sum over all the squares of the residuals,  $ssq$ , is minimal,

$$ssq = \sum_{i=1}^{n_{\text{spectra}}} \sum_{j=1}^{n_{\text{lam}}} R(i,j) \quad (4)$$

At first sight this looks like a very daunting task. However, as we will see, it is manageable.

Ideally the final square sum achieved should be numerically equal to the sum of the squares of the Gaussian noise in the measurement - usually instrumental in origin. At this point the fit cannot be further improved, though this is not a guarantee that the model is the correct one.

## 6. The chemical model

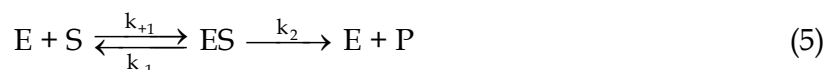
The first and central step of any data analysis is the computation of the matrix  $\mathbf{C}$  of concentration profiles based on the proposed chemical model and the associated parameters such as, but not exclusively, the rate and or equilibrium constants. Initially these key parameters may be only rough estimates of the true values.

So far the explanations are valid for kinetic and equilibrium studies. The difference between these two investigations lies in the different computations required for the calculation of the concentration profiles in the matrix  $\mathbf{C}$ .



## 7. Kinetics

The chemical model for a kinetic investigation is a set of reaction equations which describe the process under investigation. Consider as an example the basic enzymatic reaction scheme



An enzyme  $E$  reacts rapidly and reversibly with the substrate  $S$  to form an enzyme substrate complex  $ES$ . This is followed by the 1<sup>st</sup> order chemical conversion of the substrate and release of product. The free enzyme is then available to undergo another catalytic cycle.

Before proceeding to a ReactLab based mechanistic analysis it is informative to briefly outline the classical approach to the quantitative analysis of this and similar basic enzyme mechanisms. The reader is referred to the many kinetics textbooks available for a more detailed description of these methods. The scheme in equation (5) was proposed by Michaelis and Menten in 1913 to aid in the interpretation of kinetic behaviour of enzyme-substrate reactions (Menten and Michaelis 1913). This model of the catalytic process was the basis for an analysis of measured initial rates ( $v$ ) as a function of initial substrate concentration in order to determine the constants  $K_M$  (The Michaelis constant) and  $V_{max}$  that characterise the reaction. At low  $[S]$ ,  $v$  increases linearly, but as  $[S]$  increases the rise in  $v$  slows and ultimately reaches a limiting value  $V_{max}$ .

Analysis was based on the derived Michaelis Menten formula:

$$v = \frac{[E_0] [S] k_{cat}}{K_M + [S]} \quad (6)$$

Where  $V_{max} = k_{cat}[E]_0$ , and  $K_M$  is equal to the substrate concentration at which  $v = \frac{1}{2} V_{max}$ . The key to this derivation is that the enzyme substrate complex  $ES$  is in dynamic equilibrium with free  $E$  and  $S$  and the catalytic step proceeds with a first order rate constant  $k_{cat}$ . This 'turnover number'  $k_{cat}$  is represented by  $k_2$  in the scheme in equation (5).

It can be shown that under conditions where  $k_2 \ll k_{-1}$  then  $K_M$  is in fact equal to the equilibrium dissociation constant  $K_1$ ,

$$K_1 = \frac{[ES]}{[E][S]} = \frac{k_{+1}}{k_{-1}} \quad (7)$$

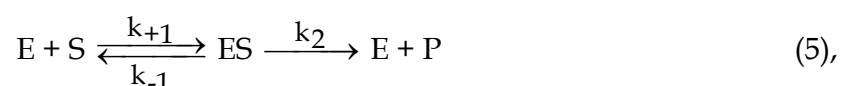
Importantly, however, the parameter  $K_M$  is not always equivalent to this fundamental equilibrium constant  $K_1$  when this constraint ( $k_2 \ll k_{-1}$ ) doesn't apply.

Furthermore though the Michealis Menten scheme can be extended to cover more complex mechanisms with additional intermediates, the  $K_M$  and  $k_{cat}$  parameters now become even more complex combinations of individual rate and equilibrium constants. The  $k_{cat}$  and  $K_M$  parameters determined by these classical approaches are therefore not the fundamental constants defining the mechanism and significant effort is required to determine the true underlying equilibrium and microscopic rate constants.

In contrast direct analysis using ReactLab to fit the core mechanism to suitable data delivers the true underlying rate and equilibrium constants in a wholly generic way that can be applied without assumptions and also to more complex models.

This involves the modelling of the entire mechanism to deliver the matrix **C** comprising the concentration profiles of all the participating species. The reaction scheme in equation (5) defines a set of ordinary differential equations, ODE's, which need to be solved or integrated (Maeder and Neuhold 2007). Reaction schemes that only consist of first order reactions can be integrated analytically in which case the concentration can be calculated directly at any point using the resulting explicit function. Most other schemes, containing one or more second order reactions, require numerical integration. Numerical integration is usually done with variable step-size Runge-Kutta algorithms, unless the system is 'stiff' (comprising both very fast and slow reactions) for which special stiff solvers, such as Gear and Bulirsch-Stoer algorithms are available (Press, Vetterling et al. 1995).

Integration, explicit or numerical, requires the knowledge of the initial conditions, in the case of kinetics the initial concentrations of all interacting species. For the above example, equation



and using the rate constants ( $k_{+1}=10^3 \text{ M}^{-1} \text{ sec}^{-1}$ ,  $k_{-1}=10^2 \text{ sec}^{-1}$ ,  $k_2=10^2 \text{ sec}^{-1}$ ) and initial concentrations, ( $[S]_0=1 \text{ M}$ ,  $[E]_0=10^{-4} \text{ M}$ ), the resulting concentration profiles generated by numerical integration and used to populate the columns of matrix **C** are shown in Figure 4.

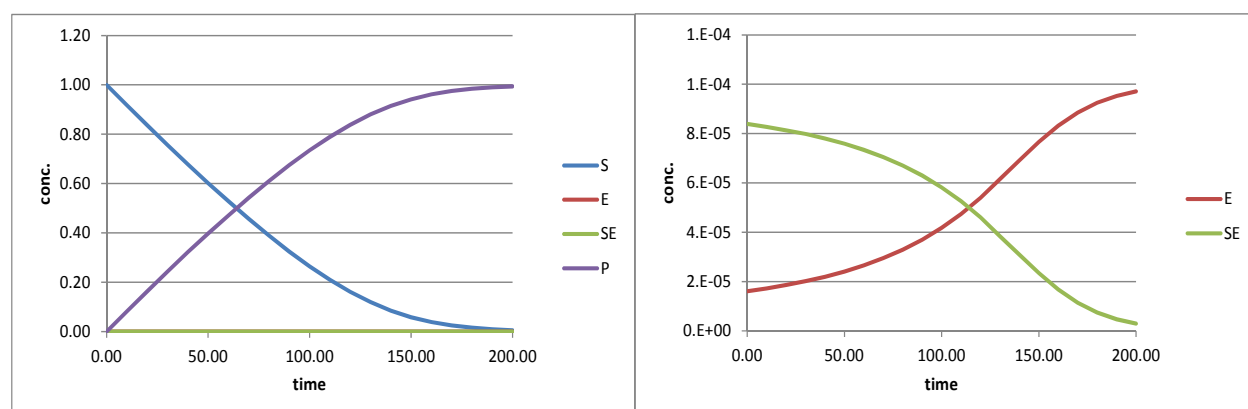


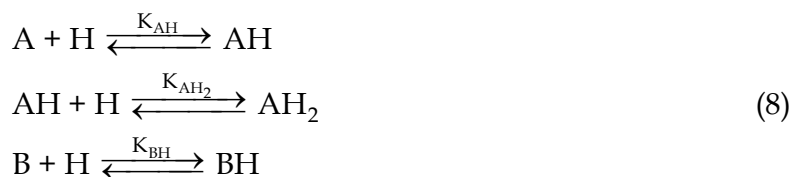
Fig. 4. Concentration profiles for the enzymatic reaction of equation (5); an expanded concentration axis is used in the right panel.

The transformation of the substrate into the product follows approximately zero-th order kinetics for most of the reaction whilst the substrate is in excess and all the enzyme sites are populated. Later in the reaction the substrate is exhausted and free enzyme released. The expanded plot in the right hand panel displays more clearly the small concentrations for the enzyme and the enzyme-substrate complex.

## 8. Equilibria

The chemical model for an equilibrium process is similar to the model of a kinetic process, only now there are exclusively equilibrium interactions, e.g.





The chemistry in this example comprises the protonation equilibria of the di-protic acid  $AH_2$  and the mono-protic acid  $BH$ . The key difference now is that the steps are defined in terms of instantaneous stability or equilibrium constants, and the fast processes of the attainment of the equilibria are not observed.

Equilibrium investigations require a titration, which consists of the preparation of a series of solutions with different but known total component concentrations. In equilibrium studies we distinguish between components and species. Components are the building blocks; in the example (6) they are  $A$ ,  $B$  and  $H$ ; species are all the different molecules that are formed from the components during the titration, the example they are  $A$ ,  $AH$ ,  $AH_2$ ,  $B$ ,  $BH$ ,  $H$  and  $OH$ . Note, the components are also species.

Instead of utilising numerical integration to compute the concentration profiles of the individual species as we did with kinetic time courses we instead use an iterative Newton-Raphson algorithm to determine the speciation based on the total component concentrations for each sample and the estimated equilibrium constants (Maeder and Neuhold 2007).

For a titration of 10ml of a solution with total component concentrations  $[A]_{tot} = 0.1M$ ,  $[B]_{tot} = 0.06 M$  and  $[H]_{tot} = 0.4M$  with 5ml of 1.0M  $NaOH$  the concentration profiles of Figure 5 result. The protonation constants are  $\log(K_{AH}) = 9$ ,  $\log(K_{AH_2}) = 3$  and  $\log(K_{BH}) = 4$ .

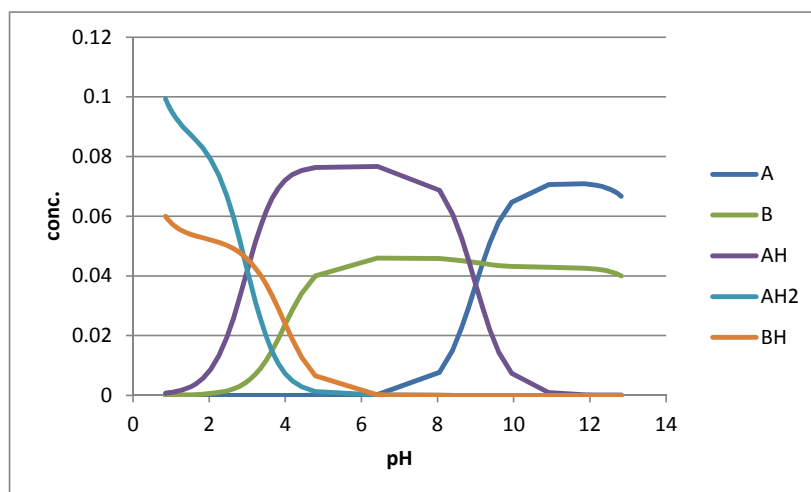
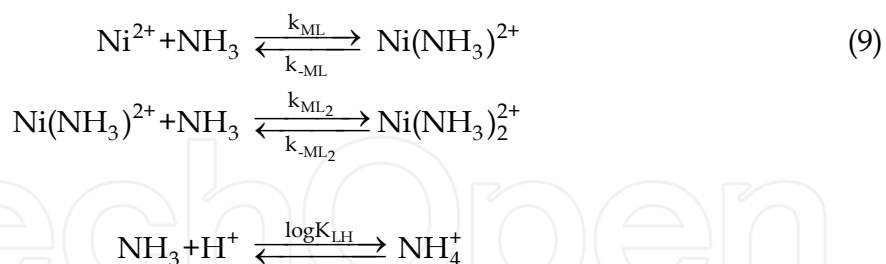


Fig. 5. Concentration profiles for the titration of an acidified solution of  $AH_2$  and  $BH$  with  $NaOH$ .

### 8.1 Kinetics with coupled protonation equilibria

A significant recent development is the incorporation of instantaneous equilibria to kinetic analyses. Careful combination of numerical integration computations alongside the Newton-Raphson speciation calculations have made this possible (Maeder, Neuhold et al. 2002). This development has made the modelling of significantly more complex and realistic

mechanisms possible. An example is the complex formation between ammonia and  $Ni^{2+}$  in aqueous solution as represented in equation (9).



$Ni^{2+}$  is interacting with  $NH_3$  to form the 1:1 and subsequently the 1:2 complexes. Importantly the ammonia is also involved in a protonation equilibrium. As a result the pH changes during the reaction and the rates of the complex formation reactions appear to change. The classical approach to this situation is to add buffers that approximately maintain constant pH and thus also the protonation equilibrium. Since buffers often interfere with the process under investigation the possibility of avoiding them is advantageous. This has only been made possible by this more sophisticated method of mechanistic analysis.

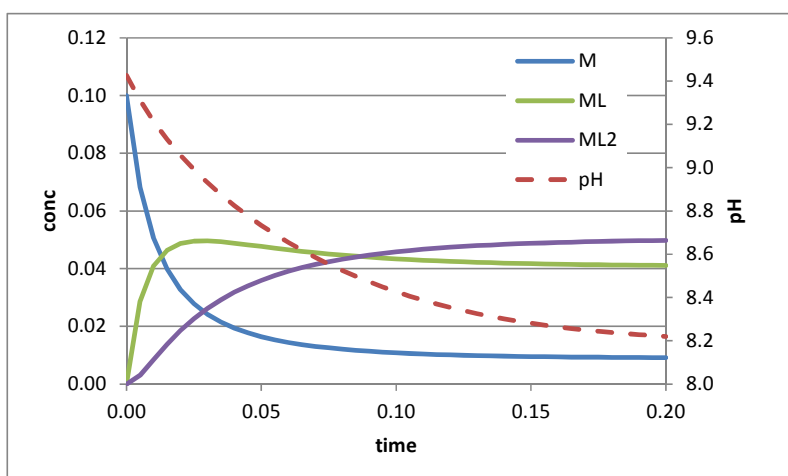


Fig. 6. The concentration profiles for the complex species in the reaction of  $Ni^{2+}$  and  $NH_3$ ; also displayed is the pH of the reacting solution.

The concentration profiles for the complex species are shown in Figure 6. The patterns for a 'normal' consecutive reaction are distorted as the reaction slows down with the drop in pH from 9.5 to 8.2. The initial concentrations for this reaction are  $[Ni^{2+}]_0=0.1$  M,  $[NH_3]_0=0.25$  M and  $[H^+]_0=0.1$  M.

## 9. Parameters

Any parameter that is used to calculate the matrix **C** of concentrations is potentially a parameter that can be fitted. Obvious parameters are the rate constants  $k$  and the equilibrium constants  $K$ ; other less obvious parameters are the initial concentrations in kinetics or the total concentrations in equilibrium titrations. Concentration determinations are of course very common in equilibrium studies (quantitative titrations); for several reasons concentrations are not often fitted in kinetic studies. For first order reactions the

concentrations are not defined at all unless there is additional spectroscopic information, i.e. molar absorptivities. For second order reactions they are in principle defined but only very poorly and thus kinetic determination is not a robust analytical technique in this case.

The parameters defining  $C$  are non-linear parameters and cannot be fitted explicitly, they need to be computed iteratively. Estimates are provided, a matrix  $C$  constructed and this is compared to the measurement according to the steps that follow below. Once this is complete it is possible to calculate shifts in these parameter estimates in a way that will improve the fit (i.e. reduce the square sum) when a new  $C$  is computed. This iterative improvement of the non-linear parameters is the basis of the non-linear regression algorithm at the heart of most fitting programs.

## 10. Calculation of the absorption spectra

The relationship between the matrix  $C$  and the measurement is based on equation (3). The matrix  $A$  contains the molar absorptivity for each species at each measured wavelength. All these molar absorptivities are unknown and thus also parameters to be determined. When spectra are collected the number of these parameters can be very large, but fortunately they are linear parameters and can be dealt with differently to the non-linear parameters discussed above.

Once the concentration profiles have been calculated, the matrix  $A$  of absorption spectra is computed. This is a linear least-squares calculation with an explicit solution

$$A = C^+D \quad (10)$$

$C^+$  is the pseudo-inverse of the concentration matrix  $C$ , it can be calculated as  $(C^tC)^{-1}C^t$ , or better using a numerically superior algorithm (Press, Vetterling et al. 1995).

## 11. Non-linear regression: fitting of the non-linear parameters

Fitting of the parameters requires the software to systematically vary all non-linear parameters, the rate and equilibrium constants as well as others such as initial concentrations, with the aim of minimising the sum of squares over all residuals, as defined in equation (4).

There are several algorithms for that task, the simplex algorithm which is relatively easy to program and features robust convergence with a high price of slow computation times particularly for the fitting of many parameters. Significantly faster is the Newton-Gauss algorithm; additionally it delivers error estimates for the parameters and with implementation of the Marquardt algorithm it is also very robust (Gans 1992; Maeder and Neuhold 2007).

As mentioned earlier non-linear regression is an iterative process and, provided the initial parameter estimates are not too poor and the model is not under-determined by the data, will converge to a unique minimum yielding the best fit parameters. With more complex models it is often necessary to fix certain parameters (either rate constants, equilibrium constants or complete spectra) particularly if they are known through independent investigations and most fitting applications will allow this type of constraint to be applied.

## 12. ReactLab analysis tools

ReactLab™ (Jplus Consulting Ltd) is a suite of software which is designed to carry out the fitting of reaction models to either kinetic or equilibrium multiwavelength (or single wavelength) data sets. All the core calculations described above are handled internally and the user simply provides the experimental measurements and a reaction scheme that is to be fitted to the data. A range of relevant supporting options are available as well as comprehensive graphical tools for visualising the data and the result of the analysis. To facilitate this all data, models and results are provided in pre-formatted Excel workbooks to allow post processing of results or customised plots to be added after the main analysis is complete.

## 13. Representing the chemical model

As has been discussed, at the root of the analysis is the generation of the species concentration matrix  $C$ .

Fitting a proposed reaction mechanism, or part of it, to the data therefore requires the determination of  $C$  from a reaction scheme preferably as would be written by a chemist. The ReactLab representation of the  $Ni^{2+}/NH_3$  complexation mechanism of Equation (9) is shown in Figure 7. Forward arrows,  $>$ , are used to represent rate constants and the equal sign,  $=$ , represents an instantaneous equilibrium, e.g. a protonation equilibrium.

Reactants	Reaction Type	Products	Label	Parameters k / log K	±	Fit <input checked="" type="checkbox"/>
M+L	>	ML	k <sub>ML</sub>	5.845E+02		<input checked="" type="checkbox"/>
ML	>	M+L	k <sub>-ML</sub>	1.114E+00		<input type="checkbox"/>
ML+L	>	ML <sub>2</sub>	k <sub>ML<sub>2</sub></sub>	2.965E+02		<input checked="" type="checkbox"/>
ML <sub>2</sub>	>	ML+L	k <sub>-ML<sub>2</sub></sub>	2.051E+00		<input type="checkbox"/>
L+H	=	LH	logK <sub>LH</sub>	9.250E+00		<input type="checkbox"/>

Fig. 7. The ReactLab definition of the mechanism in Equation (7) with rate and equilibrium constants used to compute the concentration profiles of Figure 6.

To get from this scheme to our intermediate matrix  $C$  involves a number of key computational steps requiring firstly the dissection of the mechanism into its fundamental mathematical building blocks. Many analysis tools require the user to take this initial step manually, and therefore understand some fairly sophisticated underlying mathematical principles. Whilst this is no bad thing it does complicate the overall process and provide a significant barrier to trying and becoming familiar with this type of direct data fitting using numerical integration based algorithms. A significant advance has been the development of model editors and translators which carry out this process transparently. These can be found in a variety of data fitting applications including ReactLab.

## 14. Examples

In the following we will guide the reader through the steps required for the successful analysis of a number of example data sets. This section consists of two examples each from kinetics and equilibrium studies.

Example 1: Consecutive reaction scheme  $A \xrightarrow{k_1} B \xrightarrow{k_2} C$

The data set comprises a collection spectra measured as a function of time. These are arranged as rows of the 'Data' worksheet in Figure 8, the first spectrum in the cells D6:X6, the second in D7:X7, and so on. For each spectrum the measurement time is required and these times are collected in column C. The vector of wavelengths at which the spectra were acquired is stored in the row 5 above the spectra. The two inserted figures display the data as a function of time and of wavelength.

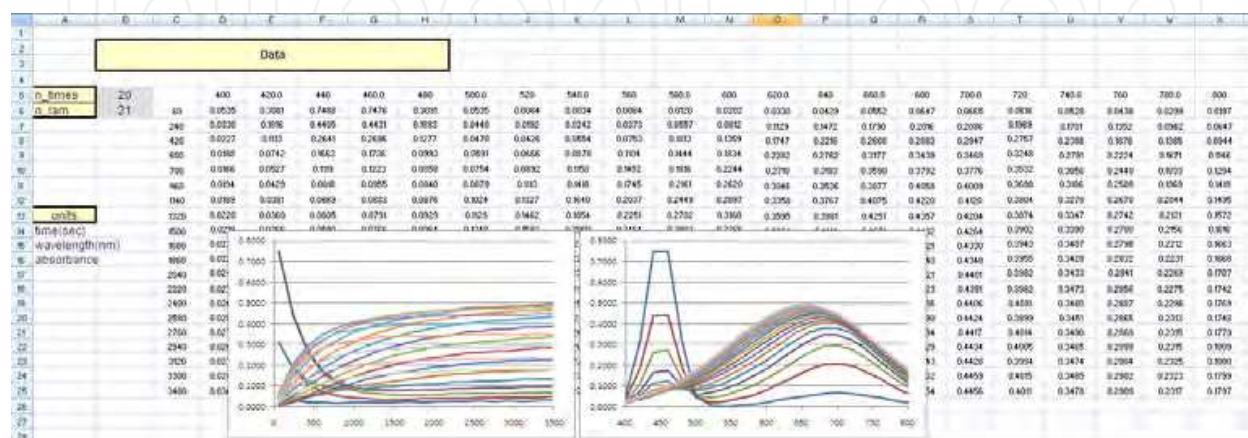
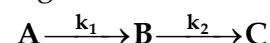


Fig. 8. The data arranged in an excel spreadsheet; the figure on the left displays the kinetic traces at all wavelengths, the figure on the right displays the measured spectra.

Prior to the fitting, the chemical reaction model on which the analysis will be based needs to be defined. As mentioned above ReactLab and other modern programs incorporate a model translator that allows the definition in a natural chemistry language and which subsequently translates automatically into internal coefficient information that allows the automatic construction of the mathematical expressions required by the numerical and speciation algorithms. Note for each reaction an initial guess for the rate constant has to be supplied. The ReactLab model is for this reaction is shown in Figure 9.

Reactants	Reaction Type	Products	Label	Parameters k / log K
A	>	B	k1	1.000E-02
B	>	C	k2	1.000E-04

Fig. 9. The definition of the chemical model for the consecutive reaction scheme



The 'compiler' recognises that there are 3 reacting species, A, B, C, and 2 rate constants. For the initial concentrations the appropriate values have to be supplied by the user. In the example  $[A]_{\text{init}}=0.001 \text{ M}$   $[B]_{\text{init}}$  and  $[C]_{\text{init}}$  are zero. Further the spectral status of each species needs to be defined, in the example all 3 species are 'colored' i.e. they do absorb in the wavelength range of the data, see Figure 10. The alternative 'non-absorbing' indicates that the species does not absorb in the measured range. Advanced packages including ReactLab also allow the implementation of 'known' spectra which need to be introduced elsewhere in the workbook.



n_species	3
n_par	2
n_aux_par	0

Species	A	B	C
init [ ]	1.00E-03	0.00E+00	0.00E+00
Spectrum	colored	colored	colored

Fig. 10. For the consecutive reaction scheme there are 3 reaction species for which initial concentrations need to be given.

The program is now in a position to first calculate the concentration of all species as a function of time and subsequently their absorption spectra. The results for the present initial guesses for the rate constants are displayed in Figure 11.

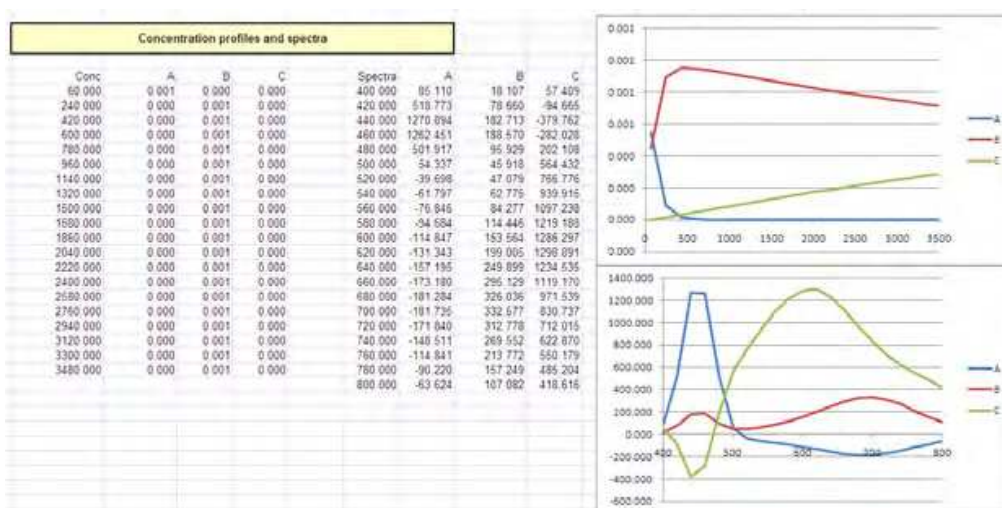


Fig. 11. The concentration profiles and absorption spectra, as calculated with initial guesses for the rate constants shown in Figure 9.

The concentration profiles indicate a fast first reaction  $A \rightarrow B$  and a much slower subsequent reaction  $B \rightarrow C$ . However, the calculated, partially negative absorption spectra clearly demonstrate that there is 'something wrong', that the initial guesses for the rate constants are obviously not correct. In this example the deviations are not too severe indicating the model itself is plausible.

Clicking the Fit button initiates the iterative improvement of the parameters and after a few iterations the 'perfect' results are evident. This of course is supportive of the validity of the model itself. If the scheme is wrong and cannot account for the detail in the data, a good fit will be unobtainable. The ReactLab GUI at the end of the fit is given in Figure 12.

On the other hand an over complex model has to be carefully avoided as any data can usually be fitted with enough parameters (including artefacts!). Occam's razor should be assiduously applied accepting the simplest model that fits the data as the most likely.

Of course it is also a risk that a model is underdetermined by the data. Put simply the information in the measurement is not sufficient to deliver a unique solution, and the program will not converge properly and usually oscillate delivering one of an infinite number of solutions (usually combinations of rates). Whilst this does not imply the model is



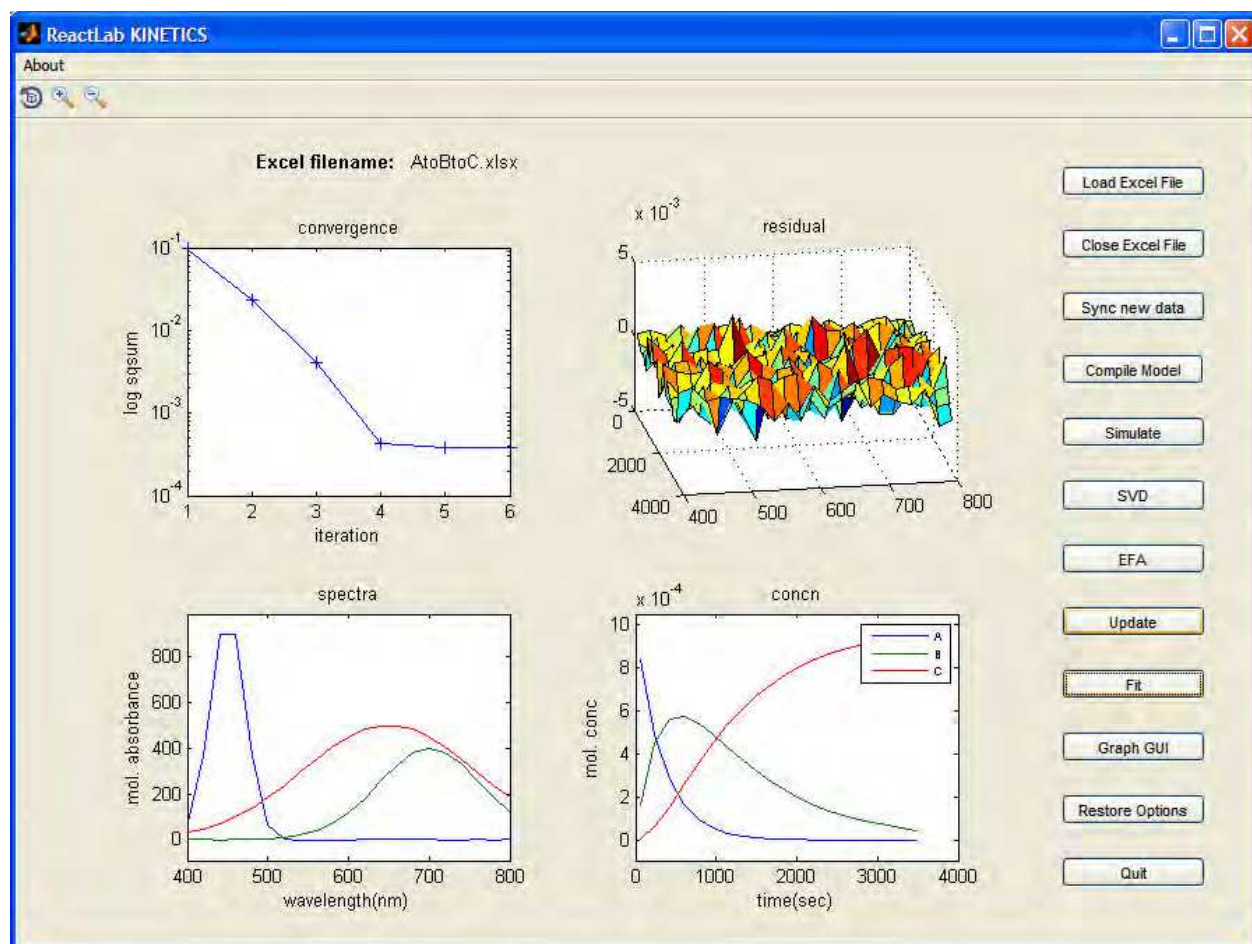


Fig. 12. The ReactLab GUI displaying the progress of  $ssq$ , the residuals, the absorption spectra and the concentration profiles after the fitting.

incorrect, further work will be required to determine and fix key parameters or spectra in order to resolve the problem.

Example 2: Kinetic analysis of the reaction of Ellmans reagent (*DTNB*) and thioglycerol *RS*.

This example illustrates the direct fitting of a simplified model followed by the correct and more complex model to a data set collected using a PDA on a stopped flow (data courtesy of TgK Scientific Ltd, UK).

Ellmans reagent, 5,5'-Dithio-bis(2-nitrobenzoic acid) or *DTNB* is a commercially available reagent for quantifying thiols both in pure and biological samples and measuring the number of thiol groups on proteins. The reaction yields a colored thiolate, *RS-TNB*, ion which absorbs at 412nm and can be used to quantify the original thiol concentration. In this particular case the reaction with thioglycerol, *RS*, leads to a 2 step disulphide exchange reaction and is particularly suited for establishing the dead-time of stopped flow instruments (Paul, Kirschner et al. 1979). The reaction is represented in Figure 13. The model in the ReactLab definition is given in Figure 16.

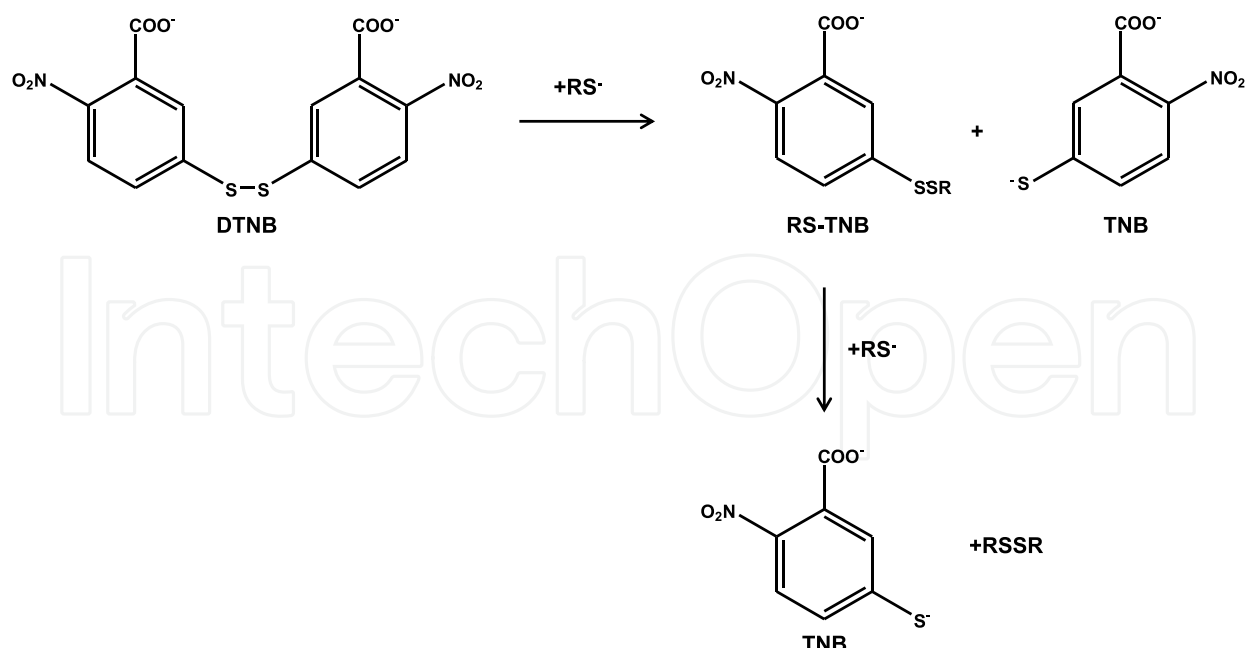


Fig. 13. The 2-step reaction of *DTNB* with a thiolate, *RS<sup>-</sup>*.

A 3-D representation of the spectra measured at different time intervals for a total of 1.5 sec. on a stopped-flow instrument is shown in Figure 14.

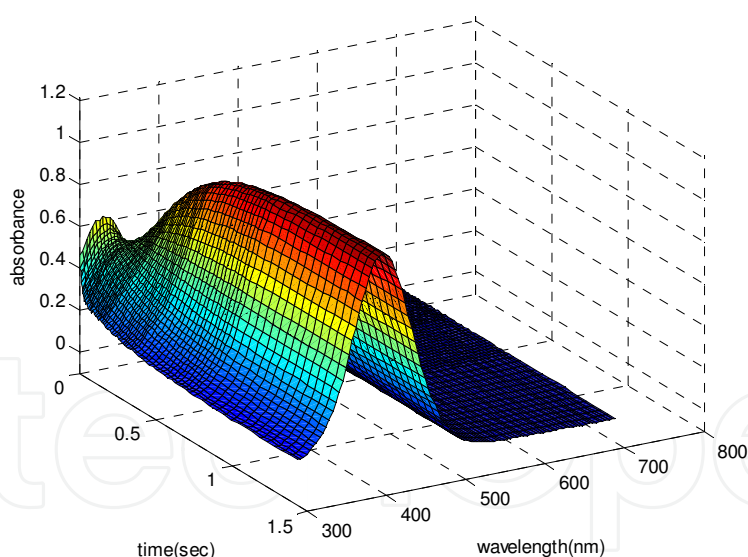


Fig. 14. Spectral changes resulting from the reactions of Ellmans reagent (*DTNB*) and thioglycerol (*RS<sup>-</sup>*).

In the experiment a large excess of thioglycerol was used and thus the two second order reactions can be approximated with a two-step pseudo first order sequential mechanism. Thus, we first attempt to fit this biphasic reaction with a simple consecutive reaction scheme with three colored species  $A \xrightarrow{k_1} B \xrightarrow{k_2} C$  (Figure 15). The fitted rates are  $75\text{sec}^{-1}$  and  $3.9\text{sec}^{-1}$ .

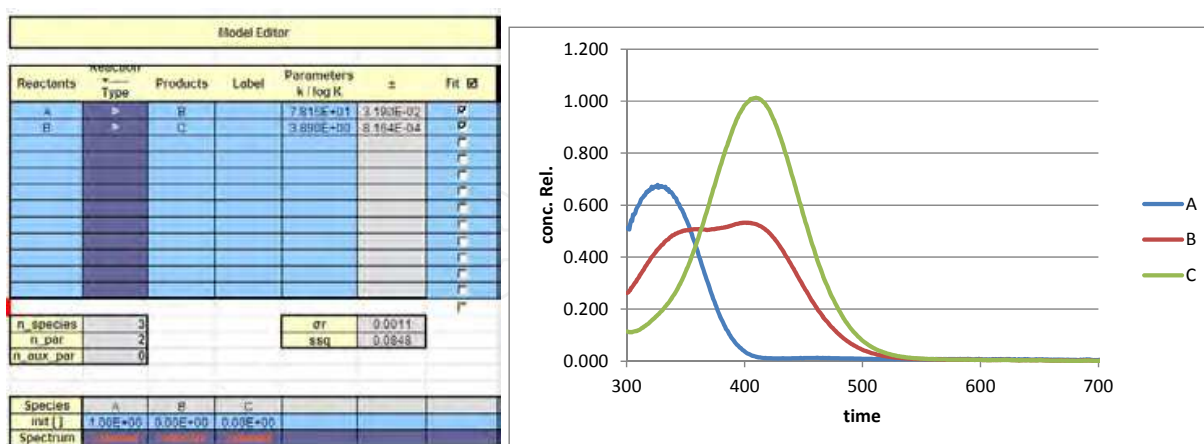


Fig. 15. A simple consecutive reaction model with best fit rates and spectra.

The problems with this fit are two-fold. First we know the reactions are second order and that the intermediate spectrum for species B is implausible with two peaks.

The complete model as defined in ReactLab, together with the fit at 412 nm, is displayed in Figure 16. Whilst the square sum is not significantly improved the spectra are now correct according to literature sources and the corresponding rates for the 2 steps are  $1.57 \times 10^4 \pm 6$   $M^{-1}sec^{-1}$  and  $780 \pm 2$   $M^{-1}sec^{-1}$ .

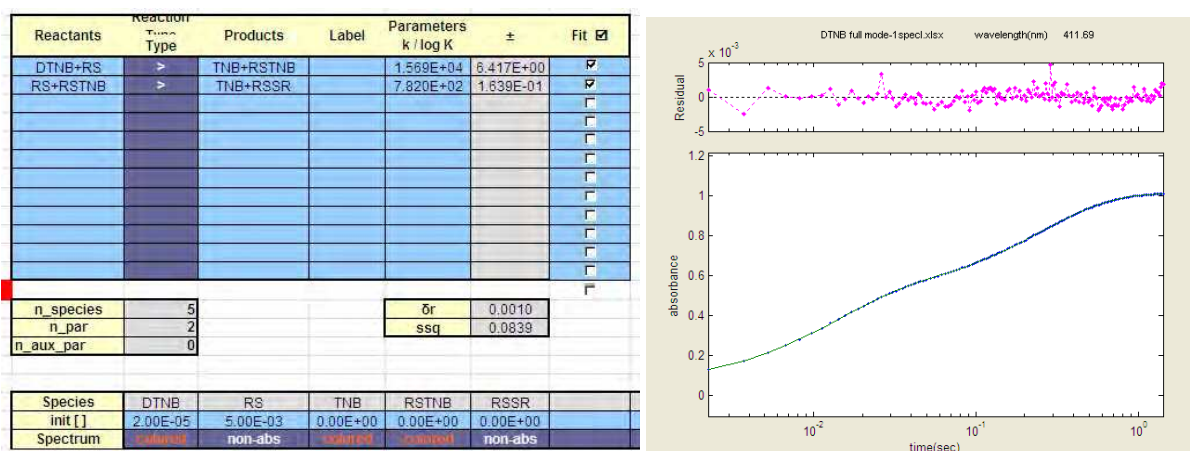


Fig. 16. Fitting the complete mechanism to the data yields the correct rates. The right panel displays the quality of the fit at 412 nm on a logarithmic time axis.

The concentration profiles and spectra resulting from a fit to the correct model are now as in Figure 17.

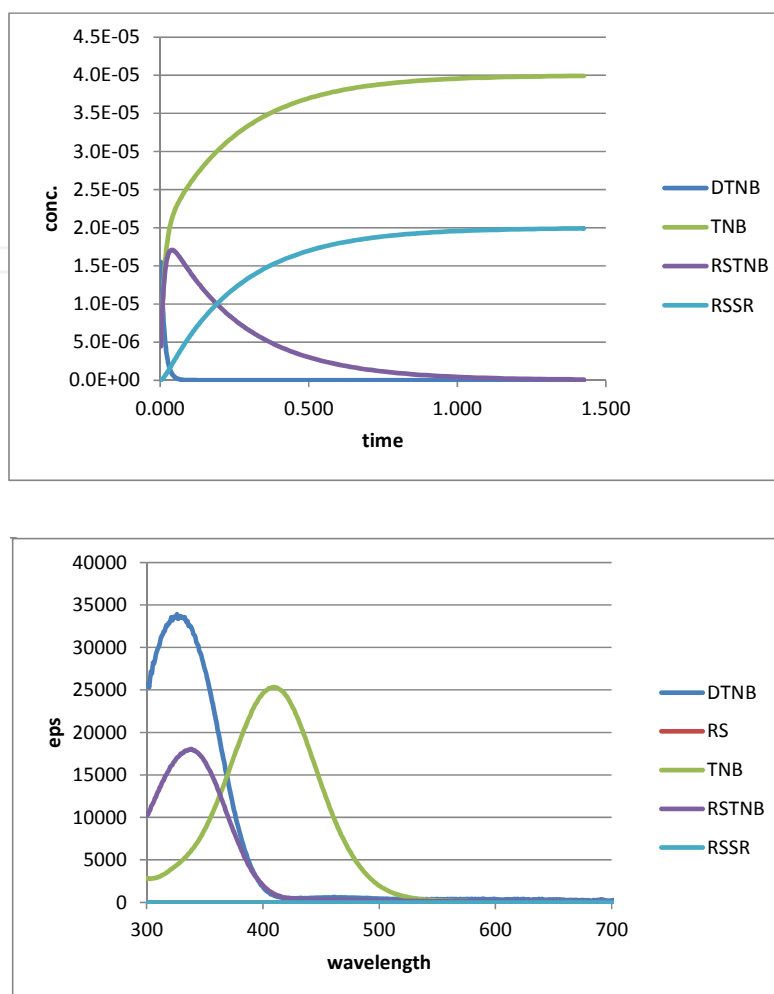


Fig. 17. Concentration profiles and molar absorption spectra for the analysis based on the complete reaction scheme.

This example does serve to demonstrate that good fits can be obtained with an incorrect or simplistic model and that some insight and care is required to establish the correct mechanism and obtain genuine parameters. What is certainly true is that the second model could only be fitted because of the numerical integration of the more complex second order mechanisms. This was a trivial change of model configuration in ReactLab and could not have been achieved using classical analysis approaches. Secondly the importance of dealing with whole spectra is highlighted in that the spectra resulting for the fit provide important insight into the underlying chemistry and must make sense in this respect. Single wavelength kinetic analysis has no such indirect reinforcement.

By way of a final comment on this example; we noted that the data was collected under pseudo first order conditions i.e. one reagent in excess. This ubiquitous approach was essential to enable the determination of second order rate constants using a first order fit by classical analysis using explicit functions (usually sums of exponentials). In the pseudo first order simplification a 2<sup>nd</sup> order rate constant is calculated from the observed pseudo first order rate constant.

Numerical integration methods eliminate the need for this constraint and therefore any requirement to work under pseudo first order conditions (or indeed the comparable constraint of keeping the reactant concentrations equal).

Example 3: Equilibrium investigation, concentration determination of a diprotic and a strong acid

Titration can be used for the determination of equilibrium constants and insofar the analysis is very similar to a kinetic investigation. Titrations are also an important analytical tool for the determination of concentrations, in real terms this is probably the more common application.

Let us consider a titration of a solution of the diprotic acid  $AH_2$  in the presence of an excess of a strong acid, e.g.  $HCl$ . The equilibrium model only includes the diprotic acid as the strong acid is always completely dissociated, see the ReactLab model in Fig. 18.

Reactants	Reaction Type	Products	Label	Parameters k / log K
A+H	=	AH		8.000E+00
AH+H	=	AH <sub>2</sub>		3.000E+00

Fig. 18. The model for a diprotic acid that undergoes two protonation equilibria.

The components are  $A$  and  $H$ , the species are  $A$ ,  $AH$ ,  $AH_2$ ,  $H$ ,  $OH$ . For the components the total concentrations have to be known in each solution during the titration. They are collected in the columns E and F of a spreadsheet. The measured spectra are collected from the cell N7 on, see Figure 19

		A	B	C	D	E	F	G	H	M	N	O	P	Q
1										Expand				
2		Data and Component concentrations												
3														
4														
5						Total component []					lam			
6	n_spectra	151	Vadd(ml)	Vtot(ml)	A	H				400.0	410.0	420.0	430.0	
7	n_lam	21	0.000	10.000	0.100	0.250				0.000	0.000	0.001	-0.000	
8			0.100	10.100	0.099	0.244				0.000	0.001	0.000	0.000	
9			0.200	10.200	0.098	0.239				0.001	0.001	-0.001	0.000	
10			0.300	10.300	0.097	0.234				-0.001	-0.001	0.001	0.000	
11			0.400	10.400	0.096	0.229				-0.001	0.000	0.000	-0.000	
12			0.500	10.500	0.095	0.224				0.000	-0.001	0.000	0.000	

Fig. 19. Data entry for a titration, the crucial total concentrations are stored in the columns E and F, the spectra to the right (incomplete in this Figure).

In the example 10ml of a solution containing  $A$  and  $H$  are titrated with 0.1 ml aliquots of base. The concentrations  $[A]_{tot}$  and  $[H]_{tot}$  are computed from the volumes and concentrations of the original solution in the 'beaker' and in the 'burette'. These concentrations are stored in the main sheet in the rows 37 and 38 as seen in Figure 20.

The definition of the total concentrations of  $A$  and  $H$  in the 'beaker' are defined in the cells C37, D37, the same component concentrations in the burette solution in the row below. Often these concentrations are known and then there is nothing to be added. In the example the component concentrations in the 'beaker' are to be fitted. This is achieved by defining



Reactants	Reaction Type	Products	Label	Parameters k / log K	±	Fit <input type="checkbox"/>
A+H	=	AH		8.000E+00		<input type="checkbox"/>
AH+H	=	AH2		3.000E+00		<input type="checkbox"/>

Label	Auxiliary Parameters	±	Fit <input type="checkbox"/>
[A] beaker	8.000E-02		<input checked="" type="checkbox"/>
[H] beaker	2.000E-01		<input checked="" type="checkbox"/>

n_species	5	δr		logKw	-1.400E+01
n_par	2	ssq	8.93E-02		
n_aux_par	2				

Species	A	H	AH	AH2	OH
Spectrum	colored	non-abs	colored	colored	non-abs
Init[ ]	A	H			
Beaker	0.0800	0.2000			
Burette	0.0000	-0.3000			
Vtot (ml)	10.0000				

Fig. 20. Model entry and information on concentrations and spectral status.

them as auxiliary parameters in the cells K7, K8; the contents of cells C37, D37, are now references to the auxiliary parameters which are fitted.

Fitting results in values for the concentrations and their error estimates, Figure 21.

Label	Auxiliary Parameters	±	Fit <input checked="" type="checkbox"/>
[A] beaker	1.000E-01	2.126E-04	<input checked="" type="checkbox"/>
[H] beaker	2.499E-01	2.152E-04	<input checked="" type="checkbox"/>

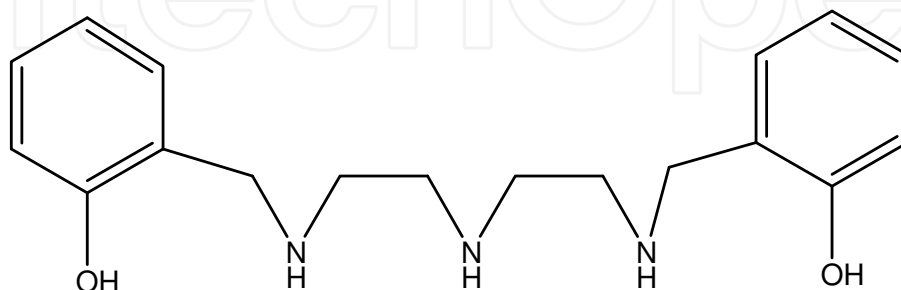
Fig. 21. The result of the fitting of the concentrations, complete with error analysis.

Example 4: Equilibrium Interaction between  $Cu^{2+}$  and PHE (1,9-Bis(2-hydroxyphenyl)-2,5,8-triazanonane)

In this example we demonstrate the analysis of 'pH-metered' titrations, a mode of titration that is common in for the investigation of equilibria in aqueous solution. In this kind of titration the independent variable is the pH, rather than the added volume of reagent as has been the case in all previous examples. As before the measured spectra are the dependent variables. An important rationale for 'pH-metered' titrations is the fact that it is often difficult to completely exclude  $CO_2$  during the whole titration. Unknown amounts of  $CO_2$  result in the addition of unknown amounts of acid via formation of carbonate species. In 'pH-metered' titrations the effect of this impurity is minimal as long as none of the carbonate species interfere with the process under investigation; the effect in the 'default mode' can be much more pronounced. The price to pay for that advantage is more complex data acquisition as the pH has to be measured and recorded together with the spectra after each addition of reagent.



The example is the titration of *PHE*, 1,9-Bis(2-hydroxyphenyl)-2,5,8-triazanonane, with  $\text{Cu}^{2+}$  in aqueous solution. (Gampp, Haspra et al. 1984) The structure of the ligand is shown below. It forms several complexes: *ML*, where the ligand is penta-coordinated presumable via all three secondary amine groups as well as the deprotonated phenolates; and two partially protonated species *MLH* and *MLH<sub>2</sub>*, in these complexes one or both of the phenolates are protonated and most likely not or only very weakly coordinated.



In this titration a solution of  $7.23 \times 10^{-4}$  M  $\text{Cu}^{2+}$  and  $1.60 \times 10^{-3}$  M *PHE* with an excess *HCl* were titrated with a total of approx.  $750 \mu\text{L}$  *NaOH* solution. After each addition of the base the pH and the spectrum were measured. The total concentrations of metal and ligand are entered for each sample in the 'Data' worksheet. Note that the columns for the total concentration of the protons is left empty: the measured pH in column M is defining the free proton concentration which in turn is used to compute all species concentrations in conjunction with the total concentrations of in this case the metal ion and the ligand provided, see Figure 22.

Data and Component concentrations					lam						
n_spectra	36	Vadd(ml)	Vtot(ml)	L	H	Cu	pH	750	740	730	720
n_lam	25	0.000		1.60E-03		7.23E-04	2.391	0.0249	0.0247	0.0245	0.0245
		0.021		1.60E-03		7.22E-04	2.411	0.0250	0.0249	0.0246	0.0246
		0.043		1.60E-03		7.22E-04	2.441	0.0252	0.0252	0.0249	0.0249
		0.064		1.60E-03		7.21E-04	2.475	0.0253	0.0253	0.0251	0.0250
		0.086		1.60E-03		7.21E-04	2.513	0.0255	0.0255	0.0251	0.0251
		0.107		1.60E-03		7.20E-04	2.556	0.0256	0.0257	0.0254	0.0254
Units		0.129		1.60E-03		7.19E-04	2.602	0.0256	0.0257	0.0255	0.0254
addition (ml)		0.149		1.59E-03		7.19E-04	2.653	0.0258	0.0259	0.0256	0.0256
wavelength		0.171		1.59E-03		7.18E-04	2.710	0.0260	0.0259	0.0256	0.0257
Absorbance		0.193		1.59E-03		7.17E-04	2.778	0.0261	0.0262	0.0258	0.0259
		0.213		1.59E-03		7.17E-04	2.856	0.0262	0.0263	0.0260	0.0261

Fig. 22. Only the total concentrations of the metal and ligand are required, the column for the protons is left empty. Column M contains the pH values and the entry 'pH' in cell M6 to indicate a 'pH-metered' titration.

The measurement is displayed in Figure 23, each curve is the measured absorption at one particular wavelength.

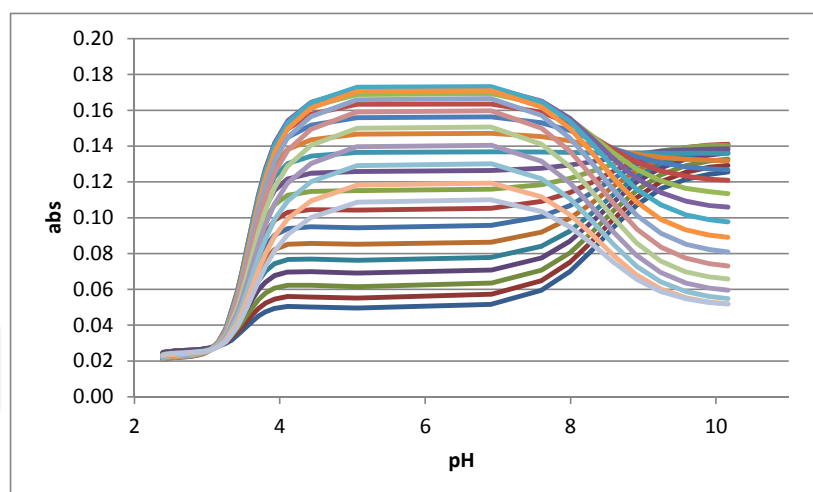


Fig. 23. The measurement, here displayed as a series of titration curves at the different wavelengths.

The ligand *PHE* has five protonation constants which have to be determined independently. The successive  $\log K$  values are 10.06, 10.41, 9.09, 7.94 and 4.18. Note that in the model the protonations are defined as overall stabilities, see Figure 24.

The results of the analysis are summarised in Figure 24 and Figure 25. Again the protonation equilibria for the complex species are defined as formation constants, the  $\log K$  values for the protonation equilibria  $ML+H \rightleftharpoons MLH$  and  $MLH+H \rightleftharpoons MLH_2$  are 8.42 and 3.92.

Reactants	Reaction Type	Products	Label	Parameters $\log K/\log \beta$	$\pm$	Fit <input type="checkbox"/>
L+H	=	LH		11.060		<input type="checkbox"/>
L+2H	=	LH2		21.470		<input type="checkbox"/>
L+3H	=	LH3		30.560		<input type="checkbox"/>
L+4H	=	LH4		38.500		<input type="checkbox"/>
L+5H	=	LH5		42.680		<input type="checkbox"/>
Cu+L	=	CuL		22.563	0.011	<input checked="" type="checkbox"/>
Cu+L+H	=	CuLH		30.979	0.011	<input checked="" type="checkbox"/>
Cu+L+2H	=	CuLH2		34.895	0.002	<input checked="" type="checkbox"/>
						<input type="checkbox"/>
						<input type="checkbox"/>
						<input type="checkbox"/>
						<input type="checkbox"/>
						<input type="checkbox"/>
<b>n_species</b>		<b>12</b>		<b><math>\sigma</math></b>	<b>4.63E-04</b>	
<b>n_par</b>		<b>8</b>		<b>ssq</b>	<b>1.76E-04</b>	
<b>n_aux_par</b>		<b>0</b>				

Fig. 24. The fitted equilibrium constants for the formation of the *ML*, *MLH* and *MLH2* complexes.

The concentration profiles are represented in two different modes, the left part has the measured pH as the x-axis and only the metal species are shown, the right part shows all species concentrations as a function of the added volume of base. This figure reveals that a substantial excess of acid has been added to the initial solution and the first 0.2 mL of base are used to neutralise this excess

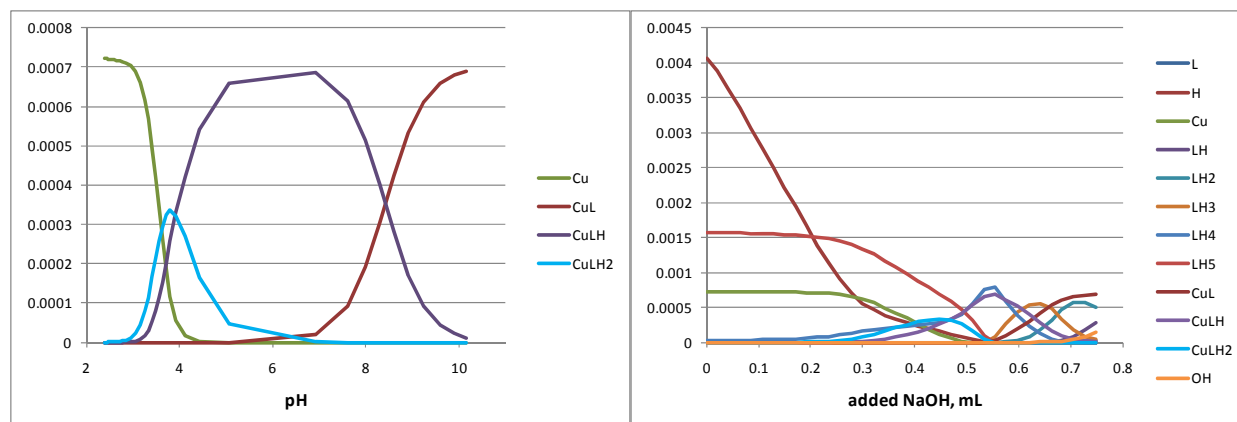


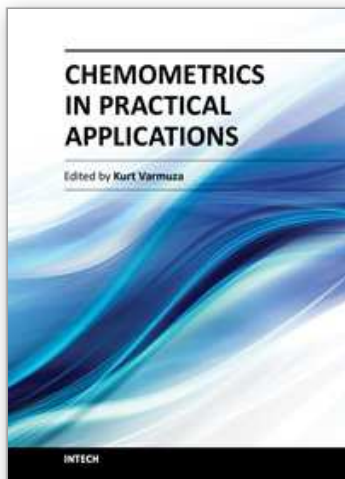
Fig. 25. The concentration profiles of the complexes as a function of pH and all species as a function of the volume of added base.

## 15. Conclusion

Data fitting is a well-established method that has been extensively used for the analysis of chemical processes since the beginnings of instrumental methods. Generally, increasing sophistication of instrumentation has inspired parallel developments in numerical methods for appropriate analysis of ever better and more plentiful data. This chapter concentrates on spectroscopic methods for the investigation of chemical processes; it details the structure of the data and the principles of model based data fitting. It is rounded off by a collection of typical and illustrative examples using a modern data analysis software package. The aim is to demonstrate the power and effortlessness of modern data analysis of typical data sets.

## 16. References

- Espenson, J. H. (1995). *Chemical Kinetics and Reaction Mechanisms*. New York, McGraw-Hill.
- Gampp, H., D. Haspra, et al. (1984). "Copper(II) Complexes with Linear Pentadentate Chelators." *Inorganic Chemistry* 23: 3724-4730.
- Gans, P. (1992). *Data Fitting in the Chemical Sciences*, Wiley.
- Maeder, M. and Y.-M. Neuhold (2007). *Practical Data Analysis in Chemistry*. Amsterdam, Elsevier.
- Maeder, M., Y. M. Neuhold, et al. (2002). "Analysis of reactions in aqueous solution at non-constant pH: No more buffers?" *Phys. Chem. Chem. Phys.* 5: 2836-2841.
- Martell, A. E. and R. J. Motekaitis (1988). *The Determination and Use of Stability Constants*. New York, VCH.
- Menten, L. and M. I. Michaelis (1913). "Die Kinetik der Invertinwirkung." *Biochem Z* 49: 333-369.
- Norman, S. and M. Maeder (2006). "Model-Based Analysis for Kinetic and Equilibrium Investigations." *Critical Reviews in Analytical Chemistry* 36: 199-209.
- Paul, C., K. Kirschner, et al. (1979). "Calibration of Stopped-Flow Spectrophotometers Using a Two-Step Disulfide Exchange Reaction." *Analytical Biochemistry* 101: 442-448.
- Polster, J. and H. Lachmann (1989). *Spectrometric Titrations: Analysis of Chemical Equilibria*. Weinheim, VCH.
- Press, W. H., W. T. Vetterling, et al. (1995). *Numerical Recipes in C*. Cambridge, Cambridge University Press.
- Wilkins, R. G. (1991). *Kinetics and Mechanism of Reactions of Transition Metal Complexes*. Weinheim, VCH.



## **Chemometrics in Practical Applications**

Edited by Dr. Kurt Varmuza

ISBN 978-953-51-0438-4

Hard cover, 326 pages

**Publisher** InTech

**Published online** 23, March, 2012

**Published in print edition** March, 2012

In the book "Chemometrics in practical applications", various practical applications of chemometric methods in chemistry, biochemistry and chemical technology are presented, and selected chemometric methods are described in tutorial style. The book contains 14 independent chapters and is devoted to filling the gap between textbooks on multivariate data analysis and research journals on chemometrics and chemoinformatics.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Marcel Maeder and Peter King (2012). Analysis of Chemical Processes, Determination of the Reaction Mechanism and Fitting of Equilibrium and Rate Constants, Chemometrics in Practical Applications, Dr. Kurt Varmuza (Ed.), ISBN: 978-953-51-0438-4, InTech, Available from:

<http://www.intechopen.com/books/chemometrics-in-practical-applications/analysis-of-chemical-processes-determination-of-the-reaction-mechanism-and-fitting-of-equilibrium-an>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen