

Advances in the modelling and analysis of complex and industrial processes

Marcel Maeder^{a,*}, Yorck-Michael Neuhold^a, Graeme Puxty^a, Paul Gemperline^b

^a Department of Chemistry, University of Newcastle, Newcastle, NSW 2308, Australia

^b Department of Chemistry, East Carolina University, Greenville, NC 27858, USA

Received 12 March 2004; accepted 4 June 2005

Available online 2 November 2005

Abstract

Data fitting is an important tool for the analysis of chemical processes. Limitations in the traditional fitting programs require strict control of external parameters such as temperature, pH, etc. Recent developments in fitting programs combine the analysis of non-ideal data with the global analysis of several different measurements. There are several advantages, the methods circumvent the necessity of external control of these parameters (thermostating, buffering), they simplify experimental design, and they deliver additional information such as activation parameters and reaction enthalpies. Several practical examples and applications are given.

Crown Copyright © 2005 Published by Elsevier B.V. All rights reserved.

Keywords: Data modelling; Data fitting; Experimental design; Global analysis; Activation parameters; Reaction enthalpies

1. Introduction

Data fitting is an old and very well-established tool for the analysis of chemical processes [1]. Before the advent of inexpensive computing power, linear relationships were ‘fitted’ with ruler and graph paper. Nonlinear relationships had to be linearised in some appropriate way. Subsequent analysis involved a manual straight-line fit and slope and intercept were interpreted according to the linearisation used. While it is possible to computerise this approach, it is not adequate at all. Error analysis is seriously hampered by the distortions imposed by the linearisation function used. Nonlinear least-squares fitting is superior as there are no distortions of the noise structure of the data. Computer programs for basic nonlinear least-squares fitting are readily available, e.g. solver in Excel or similar functions in Matlab, MathCAD, Maple, etc.

Modelling the process under investigation is the core of any fitting algorithm. In this context, we understand by modelling the quantitative mathematical description of the process. Generally, this means calculating the concentrations of all relevant chemical species as a function of the independent variable such as time in kinetics and chromatography or added

reagent in equilibrium investigations. Models are often based on the law of mass action. Until very recently, in order to be able to model the data it has been essential to ensure ideal behaviour, i.e. temperature, ionic strength, pH, etc., have to be kept constant. Under such ideal conditions and provided the process under investigation is relatively simple, there are often analytical solutions for the equations describing the process. For more complex processes, standard numerical methods are available, provided they still adhere to ideal behaviour. Otherwise, the computations are significantly more complicated and methods for the numerical analyses only appeared very recently [2–6].

If secondary processes such as the temperature dependence of a chemical reaction are to be investigated, the traditional approach has been to determine the rate at several fixed temperatures and a secondary analysis of the individual results is performed, based on the Arrhenius or Eyring equations [7]. Similarly, analyses of pH dependences result in the understanding of the protonation equilibria and their influences on the process under investigation [8]. Such investigations are performed by acquiring data at a series of fixed (buffered) pH values, followed by a secondary analysis of the pH dependence of the primary constants determined at each pH. Another comparable task is to extrapolate to zero ionic strength to yield thermodynamically properly defined parameters. This requires the analysis of a series of data acquired at different, constant

* Corresponding author.

E-mail address: Marcel.Maeder@newcastle.edu.au (M. Maeder).

ionic strengths, followed by a secondary analysis based on one of the several approximations for the activity coefficients as a function of the ionic strength [8].

Such investigations are expensive and require careful planning, in particular to properly take into account the limitations of the external control (pH buffers, thermostating, etc.).

Industrially relevant processes are often radically different from academic investigations in many respects: conditions are highly non-ideal as processes have to be performed on a large scale with no or incomplete thermostating, in highly concentrated or even neat reagents due to cost and environmental problems related to solvent disposal. Incomplete and slow mixing is an additional problem in large reactors, this aspect is not discussed in this contribution. Model-based analyses of such data based on ideal behaviour are not reliable.

An alternative to modelling analyses of such non-ideal processes is to apply model-free analysis methods such as EFA [9], ITTFA [10], ALS [11] and similar approaches. There are several examples of successful application [3,12]; however, these methods tend to suffer from a lack of robustness and from inherent rotational ambiguities [13]. Nevertheless, these methods can deliver valuable initial insight and can guide the researcher in the process of finding the correct model.

In our research laboratories, we have been developing alternative paths to address the difficulties encountered in the modelling of non-ideal data: rather than imposing external control to enforce ideal behaviour, we enhanced the modelling algorithms to encompass the non-ideal behaviour. Temperature changes [2,3], pH changes [5], changes in ionic strength [6], etc. during the process, are modelled. Initially, applications have been restricted to academic investigations but recently we have started to exploit these capabilities for the investigation of industrially relevant processes [4]. In this contribution, we present the basic mathematical concepts required for the analysis of such non-ideal data sets. We will also present examples of applications on recent chemical investigations.

2. Nonlinear data fitting

In order to explain the methodology of analysing data sets of the complexity described above, we need to recapitulate the principles of nonlinear least squares fitting. The task is to determine the best set of parameters for a given measurement and a pre-defined model. The quality of the fit is a function of the data, the model and the parameters:

$$\text{quality of fit} = f(\text{Data}, \text{Model}, \text{Parameters}). \quad (1)$$

For statistical and computational reasons, the quality of the fit is usually defined as the sum over all the squares (ssq) of the deviations between the measured data and their computed, modelled representation.

The first step is the establishment of the correct model, e.g. the chemical reaction scheme in a kinetic investigation or the expected equilibrium species in equilibrium studies. This is generally the most difficult task and requires the knowledge and experience of the researcher. Here, model-free methods such as

EFA and or ALS can play a very useful role as the results of these methods can guide the investigator towards the correct model. The objective of the present contribution, however, is the description of model-based fitting procedures and thus we will assume the correct model has been determined.

We will start with the description of multi-wavelength absorption data, as this is by far the most commonly used data type for the investigation of chemical processes. Please note that this includes the UV–VIS as well as the near and mid-infrared wavelength ranges. Absorption data are governed by Beer–Lambert’s law and as a consequence measurements of this kind are best described by a matrix equation [14]

$$\mathbf{Y} = \mathbf{CA} + \mathbf{R} \quad (2)$$

where \mathbf{Y} is a matrix, the rows of which are formed by the absorption spectra measured as a function of the progress of the process.¹ The columns of \mathbf{Y} are the absorption traces measured at the different wavelengths. According to Beer–Lambert’s law, this matrix can be decomposed into the product of a matrix \mathbf{C} containing, column wise, the concentration profiles of the absorbing species and a matrix \mathbf{A} containing, row wise, their molar absorptivities. The matrix \mathbf{R} is a collection of the residuals, the difference between the measurement \mathbf{Y} and its calculated representation \mathbf{CA} .

The task is to determine that set of parameters for which the sum over all the squares of the elements of the matrix \mathbf{R} of residuals is minimal [1].

$$\text{ssq} = \sum \sum R_{ij}^2. \quad (3)$$

Given the model and the measurement, the sum of squares is a function of the parameters to be fitted. Initially, this number of parameters seems to be very large as it includes the nonlinear parameters (e.g. rate or equilibrium constants), which define the matrix \mathbf{C} and the large number of molar absorptivities which form the matrix \mathbf{A} . Thus, for spectral series taken at many wavelengths the number of parameters to be fitted can be very high. It is of utmost importance to recognize that the molar absorptivities are linear parameters, which can be effectively eliminated from the list of parameters that need to be fitted iteratively. This is described elsewhere [14,15] and we will not further pursue this aspect here. Summarising, given the model and the measurement \mathbf{Y} , Eqs. (1) and (3) can be combined and written as

$$\text{ssq} = f(\mathbf{k}) \quad (4)$$

where \mathbf{k} is the vector of nonlinear parameters, as defined by the chosen model.

Any method that performs this minimisation task can be applied. The simplex algorithm is probably the most straightforward one; it suffers, however, from very slow progress for even modest numbers of parameters. More advanced fitting algorithms such as the one used in ‘solver’ in Excel tend to outperform simplex based algorithms. We prefer the Newton–

¹ Note, we use upper case bold characters for matrices, lower case bold characters for vectors and italic characters for scalars.

Download English Version:

<https://daneshyari.com/en/article/1180227>

Download Persian Version:

<https://daneshyari.com/article/1180227>

[Daneshyari.com](https://daneshyari.com)