A Review of Algorithms and Trends in Kinetic Model Identification for Chemical and Biochemical Systems

G. Maria
Laboratory of Chemical & Biochemical Reaction Engineering,
University Politehnica Bucharest, P.O. 15-253 Bucharest, Romania
Email: gmaria99m@hotmail.com

Original scientific paper
Received: February 2, 2003
Accepted: June 15, 2004

Simulation of complex (bio)chemical reactions plays an important role in a process kinetics characterisation. However, detailed kinetic modelling is a difficult task because the model has to reflect the process complexity under variate operating conditions, starting from a limited number of observed variables, (non-)conventional data recorded with a limited sampling frequency, and often with a low reproducibility. Extensive investigations can lead to structured models of complexity depending on the utilisation scope. To overcome weak results, the identification problem must be well formulated, data consistent, numerical estimation appropriate and effective, and the estimate quality analysis adequate. While statistical estimation theory has been extensively developed in terms of objective function choice and solution analysis, numerical algorithm application to (bio)chemical kinetic systems presents particularities and difficulties. The present paper aims to review the main steps and trends in solving the kinetic model identification problem. Rules for a suitable problem formulation vs. modelling objectives, advanced numerical algorithms for obtaining a reliable solution, and an estimate of suitable analysis are shortly summarised.

Key words:
(bio)chemical kinetic model identification, estimation algorithms

Introduction

Reliable and sufficiently accurate mechanistic kinetic models are very effective tools in understanding a chemical/biological process, its influential variables, and physical meaning of parameters. This will eventually lead to significant benefits in safety and optimal plant design and monitoring and will facilitate similar process analyses. Mechanistic based models are preferred to the empirical ones, offering the advantage of a deep understanding of the process occurrence, a better quality interpolation in the experimental domain, physical/physiological meaning of parameters, and confidence in model predictions. Physical meaning of model parameters also allows a better evaluation and interpretation of the estimate quality, and result storage for further process analyses, process and plant optimization. Reaction paths, comprising hundreds and thousands of elementary steps, have been constructed and stored in databanks for a large variety of processes, like catalytic surface reactions, gas phase radicalic reactions (oxidation, pyrolysis),[1–3] combustion,[4] cell metabolic processes and enzymatic reactions.[5–9] Simulation of spatially homogeneous (zero spatially dimensional) and one- or two-dimensional chemical systems, with detailed reaction mechanisms, has caused no problems in recent years. The current trend in all mentioned applications is to develop integrated identification and simulation platforms able to incorporate geometric system and transport characteristics together with detailed kinetics.

Complex kinetic modelling is justified not only by the research trend to increase the knowledge about the process occurrence and characteristics, but also well motivated by the economic benefits deriving from kinetic process simulation: i) more accurate and economic plant design and re-design; ii) more accurate process optimization; iii) improvement of product quality due to market pressure; iv) improved process monitoring and safety indices; v) robust and effective process control; vi) minimization of by-products and waste, and reduction of ecological impact; vii) improved predictions of process performances for variations in feed-stocks, catalyst, biomass, operating conditions; viii) more precise process scheduling and production planning for multi-product plants. Identification and simulation of extended reaction paths become possible due to improved and accurate reaction intermediate measurement techniques, improved data analysis and storage capacity, increased effectiveness of numerical algorithms and computing capacity, offering a quick process analysis and real-time simulation. Massive parallel computing allows today com-

plex kinetic simulations. This is the case, for instance, of simulating high-level complex cell processes in molecular biology, like sub-cellular communications, reactions involving proteins and enzymes during cell exo-/endo-cytosis, synaptic transmission, transport and cell signal reception (E-Cell simulator of Tomita et al.[10]; M-Cell simulator of Stiles et al.[11]; V-Cell simulator of Schaff et al.[12,13]).

The kinetic model identification problem consists of: i) formulation of an extended/reduced reaction schema (including elementary and overall reactions, intermediates, products, and lumped species), ii) identification of the most adequate rate expressions vs. available information and data, iii) estimate significant kinetic parameters vs. observed data and, iv) adapting the model structure and parameters according to available information.

However, complex (bio)chemical process characterisation via detailed kinetic models requires extensive off/on-line process investigations under a wide range of operating conditions at different scales. Chemical processes are often very sensitive to feed quality and catalyst activity changes, while bioprocesses, implying dynamic changes of cell populations, are very sensitive to input-flow oscillations, operating conditions and surroundings. Because (bio)chemical reaction paths usually involve a large number of intermediates, by-products, (nutrients) and additives, they are very complex and difficult to be kinetically described in detail. Moreover, at a catalyst-surface or biocell molecular level, it is very difficult to obtain standard kinetic data, e.g. species concentrations vs. process time, and the kinetic model identification has to be often based on a mixture of qualitative – quantitative information. Problem multimodality, non-convexity, high interrelated model parameters, variable observability of species, particular kinetic model form or incomplete data, all these can cause serious convergence or solution reliability problems when an incomplete numerical analysis is applied.

To cover such a difficult experimental and numerical task, it is advisable to follow the kinetic modelling conventional steps: kinetic data acquisition and preliminary data analysis (mass balance, error matrix, stoichiometry consistency, and data reconciliation checks); propose a reaction path and rate expressions (based on experimental evidence, thermodynamic evaluations and simplifying hypotheses); formulate an appropriate statistical estimation criterion and establish process constraints; generate an initial guess of model parameters by means of an appropriate shortcut method; solve the estimation nonlinear-programming (NLP), or mixed--integer-nonlinear-programming problem (MINLP) with an effective optimization algorithm; analyse the problem solution, e.g. model adequacy, prediction capacity, and parameter significance; if necessary, reduce (by lumping of reactions and/or species) or extend the model structure and discriminate among alternatives; improve the model quality based on new planned experiments.

The scope of this paper is to summarize the main conventional estimation steps and the systematic approach to be followed for obtaining a satisfactory kinetic model. Appropriate formulations of the (non-convex) estimation problem and effective algorithms can help in generating a good quality solution. Reduced kinetics discrimination via a MINLP criterion and binary decision variables is also presented. The paper also aims to summarize the main trends in numerical algorithms for handling kinetic identification difficulties and non-conventional routes to characterize the high-complex kinetic systems, when lack of conventional information exists.

Kinetic data preliminary analysis: data error and balance

Before developing a kinetic model, an analysis of the data quality is necessary. To be consistent, data must include enough information about the process vs. the target model structure. Data must be obtained on the whole reaction time domain, scattered to characterize satisfactorily the "picks" and "inflexions" of observed curves, independent on outliers, and recorded with an acceptable experimental error (which will be reflected in the model quality). The experimental error (i. e. the "noise") depends on the measurement possibilities. For a detailed process modelling, larger amount and precise data are necessary, while for quick process characterization and approximate predictions, rough data completed with qualitative information are usually sufficient.

Standard kinetic data implies a set of observed kinetic curves $\boldsymbol{C}(t)$ of species concentrations ($\boldsymbol{C}$-vector) over the reaction time (t), under various operating conditions. However, each data set must be recorded under iso-thermal experimental conditions that make the mass and heat transport resistances negligible.[14,15]

Classical statistical analysis can be applied to determine the data error structure, data consistency, type of dependencies among observations and species inter-connectivities. The experimental error in a u-th experiment of the observed variables $\boldsymbol{C}$ is usually assumed to be normally distributed $N(0,\Sigma_u)$, with the covariance matrix $\Sigma_u = [\sigma^2_{uij}]$ ($i,j = 1,\ldots n_s$, observed species; $u = 1,\ldots,n$, number of runs[16,17]). An estimate of the noise matrix ele-

ments, denoted with s^2_{uij}, can be obtained from *m*-replicate experiments of each *u*-th run, by using the relationships:[18]

$$s^2_{ii} = \frac{\sum_{k=1}^{m}(C_{ki} - \overline{C}_i)^2}{m-1};$$

$$s^2_{ij} = \frac{\sum_{k=1}^{m}(C_{ki} - \overline{C}_i)(C_{kj} - \overline{C}_j)}{m-1}; \qquad (1)$$

$$\overline{C}_i = \frac{\sum_{k=1}^{m} C_{ki}}{m}.$$

This matrix allows evaluation of inter-correlations among observed errors. The square of correlation coefficients, e.g. $r_{jk} = s_{jk} \Big/ \sqrt{s_{jj} s_{kk}} \in [-1,1]$, represents the percentage of the *j*-th observed variable variance, which can be explained by its association with the *k*-th variable.[19] By assuming the normality of errors for observed variables, they can be scaled and directly compared via the transformation $(C_i - \overline{C}_i)/\sigma_{ii}$. To check the normal distribution of an observed random variable, classical tests compare the observed and predicted event occurrence frequencies, or the normal cumulative probabilities. Among distribution moments, the first (mean, $m_1 = \overline{C}$), the second (variance, $m_2 = \sigma^2$), the third (skewness, or "asymmetry", $m_3 \Big/ \sqrt{m_2^3}$), and the forth (kurtosis, or "peakeness", $m_4 \Big/ m_2^2$) are the most used in characterising the distribution properties.[18,19]

Advanced statistical techniques can determine similarities or redundancy among data sets, principal components that affect the process (the so-called principal component analysis[20]), or existence of linear dependencies among observations.[19] Former test is important when some "observations" are indirectly measured, being deducted from other experiments. The use of such data with an inappropriate estimator can lead to a biased and poor kinetic estimate. Linear dependencies among data can be detected through eigenvectors corresponding to the zero eigenvalues of the matrix $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$, where[21] $\boldsymbol{H} = [C_{ui} - \overline{C}_i]$. Approximate linear dependencies among data are similarly detected for small eigenvalues $\lambda < (n-1)\overline{\sigma}^2$ (where $\overline{\sigma}^2$ denotes the average value for the error variance over n_s observations, see also below paragraph). Dependencies among observations must be eliminated from the (determinant) estimation criterion in order to avoid singularities, degenerescence, and poor estimate quality. Such a situation can also be detected from the corresponding small eigenvalues of the residual matrix $\boldsymbol{C} - \hat{\boldsymbol{C}}$ ($\hat{\boldsymbol{C}}$ denotes the predicted concentrations by a model).[22]

If the noise level is significant, several techniques can be used to detect the gross errors and to correct (reconciliate) the data. Advanced reconciliation techniques are based on the several system constraints, as molar and atomic mass balance, electric charge balance, thermodynamic constraints and, if known and consistent, on the reaction stoichiometry. These methods usually use factor analysis and chemometric techniques to detect redundant measurements, to correct data, to extract stoichiometric information, and to estimate un-observed states and reaction rates.[20,23,24]

Molar or mass balance in an experimental reactor, over a certain time interval $[t_0, t_u]$, can be written as a linear set of equations:[25]

$$D_{ui} = V(t_u)\,C_i(t_u) - V(t_0)\,C_i(t_0) + \int_{t_0}^{t_u} [F_{\mathrm{out}}(t)\,C_{i,\mathrm{out}}(t) - F_{\mathrm{in}}(t)\,C_{i,\mathrm{in}}(t)]\,\mathrm{d}t, \qquad (2)$$

where V denotes the volume of the system, F_{in} and F_{out} are the input and output volumetric flow rates, C_{i} denotes molar or mass concentration of species $i = 1,\ldots, n_s$, while $u = 1,\ldots, n$, indexes the number of considered runs. Data can be checked for balance vs. various constraints, as atomic species or electric charge conservation, each considered molar species in eq. (2) being associated with a constraint matrix $\boldsymbol{M}$ ($n_s \times$ no. of constraints), fulfilling the balance set:[26]

$$\boldsymbol{DM} = 0. \qquad (3)$$

For an atomic species balance check, $\boldsymbol{M}$ represents the atomic matrix of the involved species. Set eq. (3) can be used to correct the recorded concentrations by means of the least squares estimator,[25] $\hat{\boldsymbol{D}} = \boldsymbol{D}(1 - \boldsymbol{M}\boldsymbol{M}^{+})$ (superscript '+' denotes the pseudoinverse), or a maximum likelihood estimator.[27] When applying the data correction techniques, a normally distributed measurement error is assumed, with zero mean and known (diagonal) variance matrix, while data scaling by means of data matrix column normalization can avoid weak results.[25]

Building-up kinetic models

To kinetically model a (bio)chemical process on a mechanistic (structured) basis, several alternatives can be considered:

– continuous variable models
– discrete (Boolean) variable models
– mixed continuous-discrete variable models
– stochastic variable models

Conventional ordinary-differential equations (ODE) kinetic models, with a mechanistic description of the process, have been proved to be effective in characterizing continuous processes and perturbations, especially when systems are large and when molecular details are of little importance for the analysis. The Boolean approach, even if less realistic, is computationally tractable for complex biosystems, at a molecular/cell level, involving networks of genes that are either "on" or "off" according to defined Boolean relationships.[28,29] Mixed models realise a promising compromise among continuous and discrete representation. Stochastic (bio)chemical kinetic models replace the 'average' ODE model solution by a more detailed random-based simulator.[5,30–32] In stochastic models, the species concentrations are replaced by individual molecular species, and Monte Carlo methods are used to predict their interactions, with the expense of a considerable computational effort. Rate equations are replaced by individual reaction probabilities, while the model output is stochastic in nature. Stochastic representation is useful when a large number of species has to be accounted, of which the spatial location become important. This is the case, for instance, of cell process simulation, when the small number of molecules for a certain species is more sensitive to stochasticity of a reactive process than the species present in larger amounts. In such cases, simulation via continuous models can lead to only average process predictions, lacking of accuracy for random process representation (as cell signalling, gene mutation, etc.).

Successful kinetic modelling can be realised only in a strategy that includes experimental and computational loops. In such a way, the developed kinetic model complexity depends on the amount of available information and on the utilisation scope: extended models are used for a detailed description of the process kinetics and transport phenomena; moderate reduced models are used for process design and optimization; reduced models are used for safety analysis, control and real-time process monitoring. Extended models require a steady experimental and computational effort to identify and verify all parameters and reaction steps. However, due to the variability of materials, procedure, conditions, catalyst or biomass characteristics, the kinetic information may not always be generalised. Low data reproducibility, often present in biological systems, or reduced kinetic information require a dynamic process modelling based on an evolutive kinetic model structure and up-dated parameters, e.g. the so-called "tendency modelling".[4,25,33–37] In such dynamic models, only essential reactions are retained, based on measurable variables. An important problem to be considered during kinetic modelling is the distinction between the qualitative and quantitative process knowledge, stability and instability of involved species, the dominant fast and slow modes of process dynamics, the macroscopic and microscopic state of the process, and the non-biological and biological elements of the state vector.

A classical route to elaborate a reaction schema is based on experimental measurements of intermediates and products, prior information from kinetic databanks, stoichiometric and thermodynamic checks. The proposed reaction network can contain all possible elementary steps from databanks,[2] supplemented with quantum chemistry methods to evaluate thermochemical and reaction parameters.[1,38–40] However, the resulted large reaction schemes and over-parameterization require not only extensive experiments to check all intermediates and elementary steps, but also a large computational effort to identify and refine parameters, which may still often result in poor estimate quality.[41]

A second possibility is to increase continuously the model complexity through step-by-step interpretation of the intermediate and product trajectories, by using previous information to discriminate among model alternatives on a numerical/statistical basis.[42,43] Reaction mechanism is investigated under various operating conditions and the kinetic model is derived based on a sensitivity and principal component analysis of species, variables, and reaction terms.[4,44–47]

Current trend in kinetic modelling is to use more structured and complex strategies, by taking into account constraint representation, algorithms for model development and path synthesis, by using all types of information, conventional or not, and assembling suitable reactions and kinetic modules from databanks. Complex software is now able to realize integrated platforms for model synthesis, parameter estimation, model reduction and discrimination. The modular approach and automatic generation of ODE, differential-algebraic (DAE), or stochastic models, allow simulation of complex chemical systems,[48–57] or biochemical systems (MPS, MetaModel, GEPASI, ESSYNS, METASIM, ProMoT/DIVA, BioSpice, Cellerator, Dbsolve, Jarnac, StochSim software, see review of Hucka et al.[9]). Oriented and unified programming languages have been developed (CellML of Hedley et al.;[58] SBML of Hucka et al.[9]) to include the bio-system organization and complexity in integrated platforms for cellular system simulation. These platforms include representation of cells, neurons, bio-informa-

tic sequences, bio-polymer sequences, complex molecular structures, gene expression, gene-finding (E-Cell of Tomita et al.;[8,10] V-Cell of Schaff et al.;[12,13] M-Cell of Stiles et al.;[11] A-Cell of Ichikawa[59]). Integrated modelling and simulation platforms tend to use a large variety of chemical and biological databanks, including physico-chemical properties, species biodegradability,[60,61] reactions in solutions,[62,63] enzymes, proteins, genes, metabolic reactions (CRGM-database;[64] NIH-database[65]).

If the model structure is too extended vs. the available information, a model reduction can be applied by using various experimental and computational methods (see below 'model reduction' paragraph). Experimental rules can point-out or "mask" intermediate species and/or steps. Computational rules imply kinetic model estimation, study of the effect of parameter changes on the solution (sensitivity analysis), and identification of the redundant parts of the model or variables.

Kinetic model formulation

Stoichiometric and invariance relation checks

Mechanistic kinetic models are based on a proposed, elementary or overall, reaction path derived from the experimental data and databank information. Experimental isolation techniques can be used to step-by-step elucidate the process mechanism. Moreover, when proposing an elementary reaction schema, several shortcut numerical techniques can be used to check feasibility and consistency of the stoichiometry and reaction path. The current trend in model identification is to increase the importance of the preparative steps, including data evaluation, analysis of elementary reactions, stoichiometry and species inter-connectivity, as follows:

– check and correct reaction stoichiometry

– determine the number of linear independent reactions

– determine the number of overall reactions

– determine the number of dependent species

– correct data through dependencies among species concentrations, and known stoichiometry

– determine species interconnectivities and 'chemical distance'

Once the reaction network is elucidated, the stoichiometric matrix $\boldsymbol{v} = [v_{ij}]$, defined over all reactions $i = 1,.., n_r$ (rows) and species $j = 1,.., n_s$ (columns) is checked based on the recorded experimental data, e.g. matrix $\boldsymbol{D}$ of $(n \times n_s)$ dimensions. Data accounts species moles N_j over a number of runs ($u = 1,\ldots, n$) and extents of reaction (X_{ui}):

$$\boldsymbol{D} = \boldsymbol{X}\boldsymbol{v}; \quad X_{ui} = (N_{jo} - N_{ju})/\boldsymbol{v}_{ji}, \ \forall j. \quad (4)$$

If enough precise data are available, Bonvin & Rippin[66] propose a chemometric method called "target factor analysis" to check the stoichiometry consistency, or even to estimate unobserved variables by means of known stoichiometry (see also applications of Maria & Rippin[45]). To check the proposed stoichiometry $\boldsymbol{v}_{\text{tar}}$ consistency vs. data $\boldsymbol{D}$, a comparison with the "experimental" stoichiometry is performed. This is realised by pointing-out the real factors influencing the data, by means of singular value decomposition of $\boldsymbol{D}$, e.g. $\boldsymbol{D} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}} \approx \boldsymbol{U}_{\mathbf{r}}\boldsymbol{S}_{\mathbf{r}}\boldsymbol{V}_{\mathbf{r}}^{\mathrm{T}}$. The reduced $\boldsymbol{S}_{\mathbf{r}}$ diagonal matrix contains only the significant singular values vs. the noise level, e.g. $[\boldsymbol{S}_{\mathbf{r}}]_{ii} > \sigma$. Consistent proposed stoichiometry of a reaction $v_{\text{tar},i}$ is validated if their projection onto the stoichiometric space of data (of rank corresponding to the number of independent reactions) will not differ significantly from itself, thus holding the inequality:

$$\left\| v_{\text{tar},i}\boldsymbol{P} - v_{\text{tar},i} \right\|_{\infty} < 0.3; \quad \boldsymbol{P} = (\boldsymbol{V}_{\mathbf{r}}^{\mathrm{T}})^{+}(\boldsymbol{V}_{\mathbf{r}}^{\mathrm{T}}). \quad (5)$$

Biokinetic data can thus be (on-line) checked and corrected by using the so-called "incremental target factor analysis".[34,67]

Vice-versa, well-defined stoichiometric coefficients (or "stoichiometric numbers" in biochemistry) can be used to correct data by using eq. (4) and a least squares estimator accounting various extents of reactions $\boldsymbol{X}$:[48]

$$\boldsymbol{D} = \boldsymbol{X}\boldsymbol{v}; \quad \hat{\boldsymbol{v}} = \min \left\| \boldsymbol{D}_{\text{exp}} - \boldsymbol{X}\boldsymbol{v} \right\|_2^2 \quad (6)$$

The minimum reaction set of independent reactions, necessary to characterize the analysed chemical system of n_r reactions, is given by the rank of $\boldsymbol{v}^{\mathrm{T}}$. If $rank(\boldsymbol{v}^{\mathrm{T}}) < n_r$, there are n_r-$rank(\boldsymbol{v}^{\mathrm{T}})$ linear dependent reactions.[68] Linear dependencies can be expressed by the columns of matrix $\boldsymbol{N}$, that is the null-space (or kernel) of the matrix $\boldsymbol{v}^{\mathrm{T}}$ (eq. 7). The independent columns $\boldsymbol{n}_i$, which span the matrix $\boldsymbol{N}$, are also solutions of the steady-state (QSS) reaction rate system $\boldsymbol{r} = [r_i]$, that is:

$$v^{\mathrm{T}}\boldsymbol{N} = 0; \quad v^{\mathrm{T}}\boldsymbol{r}|_{\text{QSS}} = 0. \quad (7)$$

The columns $\boldsymbol{n}_i$ of matrix $\boldsymbol{N}$ are the steady-state flux vectors compatible with the system stoichiometry. A sub-space $\boldsymbol{n}_{\text{m},i}$ of $\boldsymbol{n}_i$ can be used to construct overall reactions, that is $v_{\text{m}}^{\mathrm{T}}\boldsymbol{n}_{\text{m},i}$.[69]

If the number of species in the model is greater than the number of steady-state mass balance equations, species can be divided in independent $\boldsymbol{C}_{\text{ind}}$ and dependent $\boldsymbol{C}_{\text{dep}}$, e.g. $\boldsymbol{C}^{\mathrm{T}} = (\boldsymbol{C}_{\text{dep}}\,\boldsymbol{C}_{\text{ind}})$. The number of dependent species can be calculated from the rank of matrix $\boldsymbol{B}$ to which $\boldsymbol{v}^{\mathrm{T}}$ is the null space:

$$\boldsymbol{B}\boldsymbol{v}^{\mathrm{T}} = 0^{\mathrm{T}} \Rightarrow \boldsymbol{v}\boldsymbol{B}^{\mathrm{T}} = 0; \quad \mathrm{size}(\boldsymbol{C}_{\mathrm{dep}}) = \mathrm{rank}(\boldsymbol{B}). \quad (8)$$

The independent molecular species conservation vectors of $\boldsymbol{B}^{\mathrm{T}}$ cause some rows of the stoichiometry matrix to be linearly dependent. The number of independent conservation relations is given by the n_{s}-$rank(\boldsymbol{v}^{\mathrm{T}})$. If $\boldsymbol{v}^{\mathrm{T}}$ is full rank, e.g. $rank(\boldsymbol{v}^{\mathrm{T}}) = n_{\mathrm{s}}$, the system has no conservation relation.

Matrix $\boldsymbol{B}$ can be used to check the steady-state mass balance equations of an open system, written for the feed (initial state, index 'o') and a subsequent system state vector:[69]

$$\boldsymbol{B}[\boldsymbol{v}^{\mathrm{T}}\boldsymbol{r} + \boldsymbol{k}_0(\boldsymbol{C}_0 - \boldsymbol{C})] = 0 \Rightarrow$$
$$\Rightarrow \boldsymbol{B}(\boldsymbol{C}_0 - \boldsymbol{C}) = 0 \Rightarrow \boldsymbol{B}\boldsymbol{C} = \text{constant}. \quad (9)$$

(where $\boldsymbol{k}_0$ is a known constant accounting system flow characteristics). Invariance relationship (9) can be written in a different way, if one introduces the matrix $\boldsymbol{H} = [C_{iu} - \overline{C}_i]$ (of $n_{\mathrm{s}} \times n$ dimensions[21]). Matrix $\boldsymbol{B}$ $(m \times n_{\mathrm{s}})$, including all linear relations among species concentrations, can be partitioned as $\boldsymbol{B}^{\mathrm{T}} = [\boldsymbol{B}_1\ \boldsymbol{B}_2]$, by placing in the first $m_1 < m$ rows of $\boldsymbol{B}_1$ the exact linear dependencies among observations, that is:

$$\boldsymbol{B}_1\boldsymbol{H}_u = 0, \quad u = 1,\ldots, n \quad (10)$$

The exact linear dependencies of eq. (10) correspond to the zero-eigenvalues $\lambda_j = 0$, $j = 1,\ldots, m_1$ of the $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$ matrix, associated with the z_j eigenvectors. If $\boldsymbol{Z}_1$ is the $(m_1 \times n_{\mathrm{s}})$ matrix whose rows consist of these normalized vectors $(z_k^{\mathrm{T}} z_k = 1)$, then a non-singular transformation exists, leading to exact linear dependencies:[21]

$$\boldsymbol{B}_1 = \boldsymbol{T}\boldsymbol{Z}_1 \Rightarrow \boldsymbol{T}\boldsymbol{Z}_1\boldsymbol{C}_u = \text{constant}. \quad (11)$$

($\boldsymbol{T}$ = transformation matrix; $u = 1,...,$ n the number of runs). Due to the experimental noise, it also exists $m - m_1$, approximate linear dependencies among observations, similarly to eq. (11), of which eigenvectors correspond to $\lambda_k < (n-1)\overline{\sigma}^2$ ($\overline{\sigma}^2$ = average error variance over n_{s} observations).

In a more systematic way, by arranging the stoichiometric matrix such that the upper $\mathrm{rank}(\boldsymbol{v}^{\mathrm{T}})$ rows are linearly independent, e.g. $\boldsymbol{v}^{\mathrm{T}} = (\boldsymbol{v}_{\mathrm{ind}}^{\mathrm{T}} \cdot \vdots\ \boldsymbol{v}_{\mathrm{dep}}^{\mathrm{T}})$, one can define a link matrix $\boldsymbol{L}$ between dependent and independent species such that, in steady-state conditions, one can write:[68]

$$\boldsymbol{v}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{v}_{\mathrm{ind}}^{\mathrm{T}}; \quad \frac{\mathrm{d}\boldsymbol{C}_{\mathrm{ind}}}{\mathrm{d}t} = \boldsymbol{L}'\frac{\mathrm{d}\boldsymbol{C}_{\mathrm{ind}}}{\mathrm{d}t};$$
$$\boldsymbol{C}_{\mathrm{dep}} = \boldsymbol{L}'\,\boldsymbol{C}_{\mathrm{ind}} + \text{constant}. \quad (12)$$

The link matrix can be evaluated from eq. (12) by using the pseudoinverse, e.g. $\boldsymbol{L} = \boldsymbol{v}^{\mathrm{T}}(\boldsymbol{v}_{\mathrm{ind}}^{\mathrm{T}})^{+}$. To avoid possible near-singularities, Maria[70] proposes to refer the pseudoinverse to $(\boldsymbol{v}_{\mathrm{ind}}^{\mathrm{T}})^{+} = (\boldsymbol{v}_{\mathrm{ind}}\boldsymbol{v}_{\mathrm{ind}}^{\mathrm{T}} + \alpha\boldsymbol{I})^{-1}\boldsymbol{v}_{\mathrm{ind}}$ instead of minimum Euclidean norm formula, by using a small $a < \overline{\sigma}^2$ to avoid poor conditioning.

Conservation relations and reaction invariants can be derived in various ways. In general, such relations can be written for species, atoms, electric charge, or can include thermodynamic non-negative constraints.[68] For instance, the atomic species conservation can be formulated as:

$$\boldsymbol{v}\boldsymbol{E}^{\mathrm{T}} = 0 \quad (13)$$

($\boldsymbol{E} = [E_{ij}]$ the atomic matrix, written for all molecular species $i = 1,...,$ n_{s}, and atomic species $j = 1,...,$ n_{a}).

Additionally, Vance et al.[71] review and propose several quick experimental-computational rules to check a reaction schema via species inter-connectivities. By inducing experimental perturbations to a (bio)chemical system, by means of tracers, or by fluctuating the inputs of the system, one can measure the propagation through hypothetical consecutive / parallel reaction path. Then, various techniques can determine the "distance" among observed species, and rules to include this information in elaborating a reaction schema.

The current trend in kinetic estimation is to pay an increased attention to the data preliminary analysis: noise characterization, data reconciliation, stoichiometry checks, detection of species inter-connectivities, data redundancy and co-linearity, independent reactions and invariants. More structured and systematic strategies are used to elaborate a kinetic model, including constraint representation, rules of assembling elementary reactions and reaction modules, in integrated platforms for model identification and process simulation. The modular approach is very suitable for complex biological system representation, by using databank information and specialized programming languages. Kinetic analysis tends to also include disparate information, not all time present in a standard form.

Propose the rate expressions and formulate the kinetic model

Once the stoichiometry and data checks are fulfilled, the next modelling step is to propose kinetic expressions for reaction rates, and to write a mass balance ODE set for all involved species. By adopting the hypothesis of separable influences of temperature and concentrations in rate expressions, and if only elementary reactions are included in the ki-

netic schema, then the partial orders of reactions β become equal with the stoichiometric coefficients. The resulted kinetic model, characterizing a (bio)chemical process dynamics and species concentration evolution over time, is:[14]

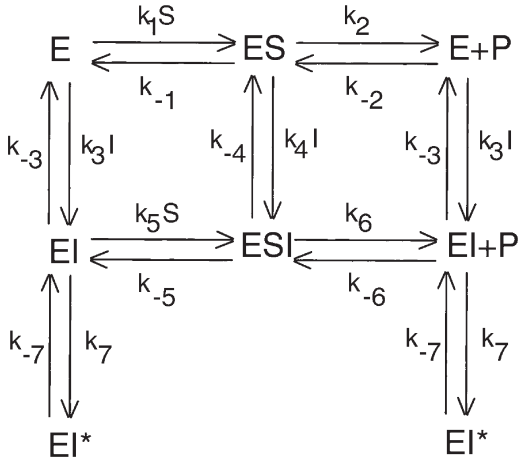
$$r_j = \frac{1}{V}\frac{\mathrm{d}N_j}{\mathrm{d}t} = \sum_{i=1}^{n_r} v_{ij} k_i \prod_{m=1}^{n_s} (N_m/V)^{\beta_{mi}};$$
$$N_j(t_0) = N_{j0}, \tag{14}$$

(r = reaction rate; t = time; V = volume; β = reaction order; N_j = moles of species j). The assumption concerning elementary reactions can be checked by experimentally analysing evolution of the derivatives $\partial r_i / \partial(N_j/V)$ over time. The formulated model eq. (14), written under variable volume conditions *V(t)*, can be solved if the volume evolution is defined by an additionally balance equation.

If the system volume is constant over time, the kinetic model can be formulated in terms of concentrations:

$$r_j = \frac{\mathrm{d}C_j}{\mathrm{d}t} = \sum_{i=1}^{n_r} v_{ij} k_i \prod_{m=1}^{n_s} C_m^{\beta_{mi}} = f_j(\boldsymbol{C}, \boldsymbol{k}, t);$$
$$C_j(t_0) = C_{j0}; \quad j = 1,\ldots,n_s \quad \text{(no. of species)};$$
$$C_{mu} = f_m^*(\hat{\boldsymbol{C}}(t_u, \boldsymbol{k}), \boldsymbol{k}) + \varepsilon_{mu}; \quad m = 1,\ldots,n'_s$$
$$\text{(no. of observed species)}; \tag{15}$$
$$u = 1,\ldots, n \text{ (runs)};$$
$$\text{subjected to the constraints:}$$
$$\boldsymbol{g}(\boldsymbol{C}, \boldsymbol{k}) \geq 0; \quad \alpha_i \leq k_i \leq \beta_i, \quad i = 1,2,\ldots,p$$
$$\text{(no. of parameters)},$$

(ε = measurement error matrix ($n'_s \times n$); '^' denotes the model predicted value). The number of observed species is less or equal with the total number of species participating to the reactions. The rate constants $\boldsymbol{k}$ are dependent on the operating conditions, e.g. temperature, pressure, pH, and also on the (bio)catalyst, or isotope nature.[15,72,73] The model constraints include thermodynamic and physical limitations, or some reaction path hypotheses (quasi-steady-state relationships, Wegscheider principle in reversible cyclic reactions, etc.[74,75]). For instance, if one considers the enzymatic kinetics for substrate (S) consumption in the presence of an enzyme (E) and an inhibitor (I), to form a product (P), a possible reaction mechanism for mixed competitive inhibition with slow enzyme binding is given by Szedlacsek & Duggleby[76] [Schema 1].



Schema 1 – *Reaction path for a considered enzymatic reaction*

The kinetic schema contains closed cycles, leading to relationships between the equilibrium and rate constants:

$$K_1 K_4 = \frac{k_{-1}}{k_1}\frac{k_{-4}}{k_4} = K_3 K_5 = \frac{k_{-3}}{k_3}\frac{k_{-5}}{k_5} \Rightarrow$$
$$g_1(\boldsymbol{k}) = \frac{k_{-1}}{k_1}\frac{k_{-4}}{k_4} - \frac{k_{-3}}{k_3}\frac{k_{-5}}{k_5} = 0. \tag{16}$$

If one applies the steady-state assumption for the complex (ES), reached instantaneously, and by neglecting k_{-2} and k_{-6}, then another constraint relationship is imposed on the kinetic model:

$$g_2(\boldsymbol{C}, \boldsymbol{k}) = \frac{\mathrm{d}(C_{\mathrm{ES}})}{\mathrm{d}t} = k_1 C_{\mathrm{S}} C_{\mathrm{E}} + k_{-4} C_{\mathrm{ESI}} -$$
$$-(k_{-1} + k_2 + k_4 C_1) C_{\mathrm{ES}} = 0. \tag{17}$$

Species concentrations are positive values, and eq. (17) can then be easily transformed in inequality relationships.

In practice, kinetic models can contain not only elementary but also overall reactions, with apparent rate constants and partial orders of reactions. Moreover, when data are recorded under mass transport limitations, the estimated rate constants are biased due to the inference with transport parameters. Simplificatory hypotheses lead to power law or hyperbolic-type rate equations. This is the case of eliminating intermediate concentrations with quasi-steady-state assumptions (QSSA), when considering a limiting step for the overall reaction, or when applying pre-equilibrium hypotheses for consecutive reactions.[73,77,78] As an example, the classical Michaelis-Menten rate expression for an enzymatic process includes a reduced number of parameters depending on the adopted hypotheses.[79] Such simplificatory hypotheses lead to reduced rate expres-

sions of hyperbolic type (review of Segel[79] for enzymatic kinetics, and of Froment[80,81] for catalytic LHHW kinetics).

In general, there are various numerical procedures to reduce dimensionality of a kinetic model by lumping species and/or reactions. In fact, by applying such a rule one reduces the number of parameters to be identified with the expense of loss of kinetic information on certain reactions and species, thus leading to biased model predictions, and a loss in physical meaning of rate constants.[4,82,83]

Once the kinetic model is complete, before starting the rate constant estimation step, it is advisable to try to reduce the model non-linearity in order to avoid further possible convergence problems for the applied optimization algorithms. Such a model regularization can be realised for instance, by reducing intercorrelations between Arrhenius parameters (*A,E*), for instance through the Pritchard & Bacon[84] transformation in a vicinity of the average experimental temperature (denoted with T_{ref}):

$$k = A\exp\left(-\frac{E}{RT_{\text{ref}}}\right)\exp\left[-\frac{E}{R}\left(\frac{1}{T}-\frac{1}{T_{\text{ref}}}\right)\right]. \quad (18)$$

Ratkowsky[85,86] and Espie[87] offer valuable reviews and suggest nonlinear transformations for kinetic model rate expressions.

Formulate the estimation problem

Once the rate expressions and orders of reactions have been established, based on the product-intermediate experimental analysis and prior information, the next step is to determine the kinetic model parameters, which is the rate constant (vector $\boldsymbol{k}$). Standard methods start from filtered, reconcilied, and smoothed experimental kinetic curves, i.e. lumped or individual species concentration trajectories over the reaction time. Kinetic parameters are estimated based on a statistical estimation criterion that minimizes the residual differences ($\boldsymbol{e}$) between data and model predictions in terms of output variables (that is $e_{ju} = C_{ju} - \hat{C}_j(t_u, \boldsymbol{k})$, for species j = 1,..., n_s and runs u = 1,..., n).

According to rate equation form, a better estimate can be obtained under certain experimental conditions (see for instance reviews of Holmberg,[88] Schügerl,[89] Moser,[15] Nielsen & Villadsen,[90] for extended Monod or Michaelis-Menten kinetics). Otherwise, due to the noised and often incomplete data, several estimates can be obtained under similar operating conditions, and the estimation methods can present convergence problems to locate a global feasible solution. The estimation objective function is linked with the statistical estimation methods because the observed data are always subjected to experimental errors, and multiple physico-chemical-biological constraints are usually imposed to the kinetic parameters. Thus, the estimation objective function depends on the error matrix structure and knowledge, noise magnitude, on the observed variable and error distribution, on prior information about parameters, and regularization functions.[17]

In fact, the estimation problem is formulated as a numerical nonlinear programming problem (NLP), or a mixed-integer NLP (denoted with MINLP). By applying the estimation theory, the resulted problem consists of searching the maximum of the posterior distribution (P) of the estimated kinetic parameters ($\hat{\boldsymbol{k}}$), that is:

$$\hat{\boldsymbol{k}} = \arg\max_{\text{k}} \Phi = (\text{constant})P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})P(\boldsymbol{k}_0) \quad (19)$$

In the relationship (19), $L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$ denotes the likelihood function, e.g. the probability of obtaining $\boldsymbol{C}$ by accounting all the experimental set of independent variables $\boldsymbol{x}$ and conditioned by the use of the parameters $\boldsymbol{k}$ in the mathematical model. The independent variables consist in measured experimental conditions, such as time, temperature, pH, etc. Notation $P(\boldsymbol{k}_0)$ denotes the apriori distribution of parameter initial solution $\boldsymbol{k}_0$, usually adopted uniform or of normal type with the covariance matrix $\boldsymbol{V}_0$. After substitution, the NLP estimation problem minimizes a cost function Φ, dependent on the residues e_{ju}, the noise level ε (usually normal distributed, of noise matrix Σ), and on prior information:

$$\hat{\boldsymbol{k}} = \arg\min_{\text{k}} \Phi = \Phi\,[e,P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x}),P(\varepsilon),\boldsymbol{k}_0,P(\boldsymbol{k}_0)];$$
$$\varepsilon \sim N(0,\Sigma) \quad \text{subjected to:}$$
$$\boldsymbol{g}(\boldsymbol{k}) \geq 0; \quad \alpha_i \leq k_i \leq \beta_i, \quad i = 1,2,\ldots,p \quad (20)$$
$$\text{(no. of parameters)},$$

Particular forms of estimators are presented in Table 1, the optimization problem being usually formulated for a normal noise and the logarithmic form of the likelihood function, ln $L(\boldsymbol{k})$. When including prior knowledge on parameters, the biased criterion penalises the movement from the initial guess of the most insensitive parameters in the model, thus preventing estimates with large variances.[91]

The formulated NLP objective function is subjected to parameter constraints and mass balance equations of the kinetic model. The most used estimator is the weighted least squares (WLS) with/without including prior knowledge on parameters (in the Bayes sense or in the Tikhonov regular-

Table 1 – *Statistical estimators (eq. 20; the normal loglihelihood function is used for models with: n_s= no. of observations; p = no. of parameters; n = no. of runs; the heteroscedasticity parameter is $\gamma_j = 0$, for constant absolute error, and $\gamma_j = 2$ for constant relative error; w_{reg},λ_l = weights of Tikhonov criterion).*

Estimator	Objective function (Φ)	Remarks
Bayes (BE)	$\max\Phi = (\text{constant})P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})P(\boldsymbol{k}_0)$; $\ln L = -\frac{n\cdot n_s}{2}\ln(2\pi) - \frac{\pi}{2}\ln(\det(\Sigma)) - \frac{1}{2}\sum_{u=1}^{n} e_u^{\mathrm{T}}\Sigma^{-1}e_u$	$L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$, known; $P(\boldsymbol{k}_0)$, known; $e_{iu} = C_{iu} - \hat{C}_{iu}$.
	$\min\Phi = \frac{1}{2}(\boldsymbol{k}-\boldsymbol{k}_0)^{\mathrm{T}}\boldsymbol{V}_0^{-1}(\boldsymbol{k}-\boldsymbol{k}_0) + \frac{1}{2}\sum_{u=1}^{n} e_u^{\mathrm{T}}\Sigma^{-1}e_u$	Σ known, independent on run (u = 1,..., n).
	$\min\Phi = \frac{1}{2}(\boldsymbol{k}-\boldsymbol{k}_0)^{\mathrm{T}}\boldsymbol{V}_0^{-1}(\boldsymbol{k}-\boldsymbol{k}_0) + \frac{1}{2}\ln\left[\prod_{u=1}^{n}\det\Sigma_u\right] +$ $\frac{1}{2}\sum_{u=1}^{n}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s}\frac{(C_{ui}-\hat{C}_{ui})(C_{uj}-\hat{C}_{uj})}{\sigma_{uij}^2}$	Σ_u known, dependent on run (u = 1,..., n).
Maximum Likelihood (MLE)	$\max\Phi = L(\boldsymbol{k})$	$L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$, known; Σ known.
	$\min\Phi = \frac{1}{2}\ln\prod_{u=1}^{n}\prod_{i=1}^{n_s}\sigma_{uii}^2 + \frac{1}{2}\prod_{i=1}^{n_s}\prod_{u=1}^{n}\left[\frac{C_{ui}-\hat{C}_{ui}}{\sigma_{uii}}\right]^2$	Σ known, diagonal.
	$\min\Phi = \frac{n}{2}\ln\left(\det\sum_{u=1}^{n}(C_{ui}-\hat{C}_{ui})/(C_{uj}-\hat{C}_{uj})\right)$	Σ unknown; (dependent observations have to be removed).
	$\min\Phi = \frac{1}{2}\sum_{j=1}^{n_s}\left[n_j[\ln(2\pi)+1] + n_j\ln\left[\frac{1}{n_j}\sum_{u=1}^{n_j}\frac{(C_{uj}-\hat{C}_{uj})^2}{\hat{C}_{uj}^{\gamma_j}}\right] + \gamma_j\sum_{u=1}^{n_j}\ln(\hat{C}_{uj})\right]$	Σ known; various data sets, of n_j size
Weighted least squares (WLS)	$\min\Phi = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{u=1}^{n}\left[\frac{C_{ui}-\hat{C}_{ui}}{\sigma_{uii}}\right]^2$	$L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$, known; Σ known, diagonal.
	$\min\Phi = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{u=1}^{n}\left[\frac{C_{ui}-\hat{C}_{ui}}{\sigma_{uii}}\right]^2 + \frac{1}{2}\sum_{l=1}^{p}\left(\frac{k_l-k_{l,0}}{[\boldsymbol{V}_0]_{ll}}\right)^2$	$\boldsymbol{V}_0$ known, $P(\boldsymbol{k}_0)$ normal (Bayes).
	$\min\Phi = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{u=1}^{n}\left[\frac{C_{ui}-\hat{C}_{ui}}{\sigma_{uii}}\right]^2 + w_{\mathrm{reg}}\sum_{l=1}^{p}\lambda_l\left(\frac{k_l-k_{l,0}}{k_{l,0}}\right)^2$	$\boldsymbol{k}_0$, known; Tikhonov regularisation.
Least squares (NLS)	$\min\Phi = \frac{1}{2\sigma^2}\sum_{i=1}^{n_s}\sum_{u=1}^{n}(C_{ui}-\hat{C}_{ui})^2$	$L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$, known; Σ known, diagonal and constant ($\sigma^2\boldsymbol{I}$).
	$\min\Phi = \frac{1}{2\sigma^2}\sum_{i=1}^{n_s}\sum_{u=1}^{n}(C_{ui}-\hat{C}_{ui})^2 + \frac{1}{2}\sum_{l=1}^{p}\left(\frac{k_l-k_{l,0}}{[\boldsymbol{V}_0]_{ll}}\right)^2$	$\boldsymbol{V}_0$ known, $P(\boldsymbol{k}_0)$ normal (Bayes).
χ^2-estimator	$\min\Phi = \sum_{u=1}^{n}\left[\frac{C_u-\hat{C}_u}{\sigma_u}\right]^2$, (single observation)	$L = P(\boldsymbol{C}/\boldsymbol{k},\boldsymbol{x})$, known; Σ known, diagonal.
MinMax	$\min\Phi = \max_u(e_u)$	Σ unknown.
Others (biased)	$\min\Phi = \sum_{u=1}^{n}\lvert C_u-\hat{C}_u\rvert$; $\min\Phi = \frac{1}{n}\sum_{u=1}^{n}\left\lvert\frac{C_u-\hat{C}_u}{C_u}\right\rvert$; $\min\Phi = \max_u\left\lvert\frac{C_u-\hat{C}_u}{C_u}\right\rvert$	(single observation).

ized form, Table 1). The former criterion is used in kinetic identification when parameters present very different sensitivities in the model functions.[91] Special estimation cases, such as mismatched data or noise matrix, correlated observations, implicit models, are extensively discussed in the literature.[16,17,41,92–97]

The current trend is to perform model synthesis and estimation in one step. Search is conducted on simultaneously determining ***k***-s and multiplicative binary parameters ***y***, which decide on the degree of extension of the model. Model solution is thus obtained by using a MINLP problem formulation, where some kinetic parameters (or groups of them, or kinetic terms) are determined simultaneously with ***y*** variables:[46,47]

$$\hat{\boldsymbol{k}} = \arg\min_k \Phi\,[\boldsymbol{k}\boldsymbol{y}^{\mathrm{T}}, \boldsymbol{F}(\boldsymbol{C},\boldsymbol{k})\boldsymbol{y}^{\mathrm{T}}];$$

subjected to:

$$\boldsymbol{g}(\boldsymbol{k}\boldsymbol{y}^{T}, \boldsymbol{F}\boldsymbol{y}^{\mathrm{T}}) \geq 0; \quad \alpha_i \leq k_i \leq \beta_i; \quad \gamma_j \leq y_j \leq \xi_j;$$

$$i = 1,2,\ldots,p_c, \quad j = 1,2,\ldots,p_i \qquad (21)$$

(where ***y*** is a vector of binary variables; ***F(C,k)*** vector includes possible redundant kinetic lumps; (α_i, β_i) are bounds of parameter k_i; (γ_j, ξ_j) are bounds of parameter y_j, e.g. {0,1} for binary variables). An adequate model that includes the minimum necessary number of parameters is validated as solution (e.g. the so-called "principle of parsimony").

As an example, the overall rate of an enzymatic process with multiple binding on non-cooperative sites can be written in a MINLP formulation as:[79]

$$r = [E]_t \frac{k_{p1}\frac{[S]}{K_{S_1}} + k_{p2}\frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}} + \ldots}{1 + \frac{[S]}{K_{S_1}} + \frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}} + \ldots} \qquad (22)$$

$$\ldots \frac{+k_{p3}\frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}}\frac{y_3[S]}{K_{S_3}} + k_{p4}\frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}}\frac{y_3[S]}{K_{S_3}}\frac{y_4[S]}{K_{S_4}}}{+\frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}}\frac{y_3[S]}{K_{S_3}} + \frac{[S]}{K_{S_1}}\frac{y_2[S]}{K_{S_2}}\frac{y_3[S]}{K_{S_3}}\frac{y_4[S]}{K_{S_4}}}$$

(S = substrate, E = enzyme, index 't' denotes total value). Thus, together with the rate and equilibrium constants (k_{p1}, k_{p2}, k_{p3}, k_{p4}, K_{S_1}, K_{S_2}, K_{S_3}, K_{S_4}), multiplicative binary parameters [y_2, y_3, y_4] have also to be estimated in order to retain the significant kinetic terms, which ensure a satisfactory model adequacy vs. experimental data.

Finally, it is to remark that, due to lack of standard kinetic data and disparate qualitative-quantitative information, not all identification techniques are based on statistical estimators. Non-conventional estimation objective functions can also be formulated, even if the obtained solution is biased, presenting no statistical property. This is the case, for instance, of biological systems when, as mentioned by Holmberg,[88] kinetic estimation has to account low data reproducibility, small number of observed variables of variate observability, a limited sampling frequency, complex dynamics with a wide range of time constants, high nonlinearities, imprecision and data irreproducibility, substantial stability punctuated by abrupt failures, and a sensitive, readily adaptable community of micro-organisms. Kinetic estimation problem becomes even more difficult when applied at a microscopic cellular level, due to the near astronomical process complexity and disparate kinetic information organized in a non-standard qualitative-quantitative form. A large variety of non-conventional estimation criteria can be tried, such as: apparent regulatory properties of the system, maximization of system recovering rate after a dynamic perturbation,[98] smallest amplitude of the recovering path, smallest sensitivity vs. perturbations,[6] system steady-state stability properties,[99] system oscillatory properties,[69] physical constraints, system flexibility vs. external conditions,[100] imposed succession of reactions (or events into a cell), etc. It is probable that, for complex (bio)chemical systems, multi-objective identification criteria have to be used, with weights in accordance with the main system functions. All types of kinetic information tend to be included in estimation criteria, in a more complex way, through binary decision variables, and a larger set of objective functions.

Solving the kinetic estimation problem

To solve the NLP estimation problem eq. (20), several numerical techniques can be used: (i) indirect (or exact) methods, which ensure minimization of the estimation objective function via an iterative optimization procedure with repeatedly ODE kinetic model solution;[16] (ii) direct methods based on ODE kinetic model transformation into an algebraic one and problem approximate solution in one step;[101,102] (iii) recursive estimators, of Kalman or non-Kalman type, that on-line adapt the system states (concentrations, temperature, etc.) and parameters, based on on-line acquired experimental information.[103]

A rough kinetic initial estimate can be evaluated by using few experimental data and a shortcut direct estimator. A good initial estimate can avoid further convergence problems for a subsequent exact estimator, especially for multimodal, or poor-conditioned cases (e.g. poor kinetic data, over-pa-

rameterized or poor-estimable models). Direct methods replace the repeated kinetic model solution with solving an over-determined algebraic set obtained after the discretization (DP) or integral transformation (IP) of the ODE model. In spite of a large number of DP and IP variants (see below paragraph), the direct estimate is biased, very sensitive to the data noise level, of poor quality, and with any statistical property. The DP or IP estimate cannot fully use the prior kinetic information, and present frequent problems of solution multiplicity, stability and non-feasibility. A modification of IP, that is the MIP proposed by Maria & Rippin,[102] offers the possibility to include prior information in generating a more reliable initial estimate of rate constants for relatively low complex kinetic models. The MIP solution is more resistant to poor-conditioned cases, due to the used information about similar processes.[104,105]

When the purpose of the kinetic model is the process control, relative simple models can be identified via recursive extended Kalman filters (EKF), based on the on-line acquired experimental information, previous estimate, and successive model linearizations. However, in some cases, on-line identification is a difficult task because the process variable characteristics have to be reflected in up-dated reduced model structure and parameters.[88] Such recursive estimators are very sensitive to the model structure, data noise, model linearizations, variations in species/reaction observability, and the quality of the prior parameters and their variance.[106,107] Sometimes, exact NLS steps are periodically applied in order to reinitialise the EKF prior knowledge.[108] When on-line adapting the model structure, changes in the model complexity can be detected by combining chemometric and estimation techniques.[25,45,67]

If consistent and complete data are available, a refined estimate is obtained by using an exact (indirect) estimator solved with a suitable optimization algorithm. A complete statistical analysis of the estimate quality and model adequacy in respect to the available prior information allows reducing or extending the model structure and improvement in the estimate quality. A MINLP formulation of an exact estimator can thus consider a larger number of reduced kinetic model alternatives, and can retain the adequate model, that includes the minimum necessary number of parameters.

Shortcut techniques for initial estimate generation

Shortcut estimators, applied to ODE kinetic models, are quick rules to generate an initial estimate of rate constants, and can be applied to data obtained under isothermal, iso-pH, iso-catalyst operating conditions, and for only observed species variables. Experimental conditions and reactors can be chosen to fulfil such requirements and to avoid mass transport limitations.[14,15,75] Even if the reactor is not isothermally operated, special procedures can reconstruct isothermal data from non-isothermal experiments.[109,110] The main shortcut estimation methods are: (i) direct analytical procedures; (ii) discretization procedures (DP); (iii) integral transformation procedures (IP); (iv) combinations of DP or IP with other numerical/statistical rules.[102,111]

Before applying a shortcut estimator, a supplementary numerical treatment of kinetic data is necessary. For instance, the kinetic curves can be approximated with empirical algebraic functions to obtain a more accurate evaluation of the "experimental" integrals of concentrations required by the IP. Usually cubic spline approximations of $\boldsymbol{C}(t)$ give satisfactory results.[23,111] When spline approximation curves are obtained using a smoothing factor, supplementary regularity properties induce a favourable effect on evaluation of the "experimental" reaction rates required by DP (i.e. $\boldsymbol{C}(t)$ curve slopes in the experimental points t_u). As the smoothing factor is larger as the approximated kinetic curve is closer to a straight line across the experimental points. For low noised data, by slightly smoothing the kinetic curves, distorted reaction rate evaluation can be avoided and the DP estimate can be improved.[111]

If the ODE kinetic model is of low complexity, their algebraic analytical solution can be obtained. As a consequence, simple regression rules can be applied to determine the rate constants from the over-determined (non)linear set of integral equations written for the observed variables (C_{mu} in eq. 15). Collections of integral forms of kinetic models of various complexities are reported in the literature.[73,78,79,112]

DP are applied, in a graphical or numerical form, to the experimental kinetic curves $C(t)$ recorded in a differential reactor/manner.[15,75] The DP estimator replaces the iterative ODE model solution (which improve the estimation function) with solving an overdetermined algebraic set derived from the basic model for all observed species j:

$$G_{ju}(\boldsymbol{C}_0, \boldsymbol{C}, \boldsymbol{r}, \boldsymbol{k}, t) = 0; \quad j = 1,..., n_s; \quad u = 1,..., n; \quad n > p; \quad \boldsymbol{g}(\boldsymbol{C}, \boldsymbol{k}) \geq 0. \tag{23}$$

All DP variants replace the differential model terms (e.g. reaction rates) with "observed" rates, numerically evaluated by means of various discretization schemes.[113] More evoluated schemes are combined with smoothed approximation functions to reduce the noise effect on the DP estimate

and to confer regularity properties to kinetic curves.[43,111,114] Even if the DP accounts all the available data, it is proved that most of information is concentrated in the reaction time domains where the parameter sensitivity coefficients [i.e. $\partial \boldsymbol{f}/\partial \boldsymbol{k}$, $\boldsymbol{J} = \partial \boldsymbol{C}/\partial \boldsymbol{k}$ from eq. (15)], and the observed reaction rates are large. Thus, the so-called information matrix $E(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{J})$ becomes well-conditioned. Several DP variants have been developed:[14] initial rate variation; algebraic sets written for all observations and experimental points with/without liniarized expressions and suitable transformations (for instance Walker plots, Lineweaver-Burk and Gates linearizations in enzymatic kinetics[15]). The (non)linear set can be solved numerically, under certain parameter constraints, by using classical algorithms,[115,116] or by transforming in a NLP problem.

Although simple, DP suffer from an important number of disadvantages: biased estimate, inconsistent, and with no statistical property, because the estimation matches the reaction rates as observed variables instead of the observed species concentrations. Thus, the "experimental" error of rate variables is amplified, DP solution being frequently inadequate, infeasible, and imprecise. DP are very sensitive to the presence of "outliers" in data, to poor-conditioned cases (incomplete or degenerated data), degenerated model terms or model over-parameterization. DP do not use in estimation the prior information concerning parameters or similar processes.

IP estimators are applied to kinetic curves $\boldsymbol{C}(t)$ recorded in an integral reactor/manner.[14,15] IP replace the repeated ODE model solution with solving an overdetermined algebraic set obtained from the integrated reaction rates of observed species (eq. 15). The simplest variant of IP is the basic algorithm of Himmelblau et al.,[117] which lead to the (nonlinear) algebraic set written for all observations $j = 1,..., n_s$, and runs $u = 1,..., n$, and subjected to certain constraints in parameters:

$$C_j(t_u) = C_j(t_0) + \int_{t_0}^{t_u} f_j(\boldsymbol{C}(\boldsymbol{k},t),\boldsymbol{k},t)\mathrm{d}t;$$
$$\boldsymbol{g}(\boldsymbol{C},\boldsymbol{k}) \geq 0. \tag{24}$$

Concentration integrals are evaluated with satisfactory precision if data are enough scattered on the reaction time domain (being more numerous around the intermediate curve "picks"). IP estimate is influenced by the experimental errors in a smaller degree compared to DP, especially when kinetic curves are regularized by means of smoothed spline functions used to evaluate the integrals in eq. (24).[111]

Several IP variants have been developed:[14,15,73,101] graphical and/or numerical variants of "initial concentration variation", or halving time concentration measurements; weighted IP; IP from the liniarized, reduced (with eliminated time and/or pre-equilibrium assumptions), or Laplace transformed ODE model; analytical IP by using the model secular equation. Similarly to DP, the IP estimate presents the same mentioned disadvantages. However, IP solution is slightly superior to DP and, for some variants, consistent because estimation refers to the original observed species concentrations.

Although imprecise and biased, the DP and IP are simple rules to quickly generate an initial guess for kinetic model parameters. The rules require 3-6 times less model evaluations than an exact NLS estimator. However, classical DP and IP can lead to a poor estimate for poor-conditioned estimation problems.

A considerable improvement in the shortcut estimator reliability and effectiveness was realised by Maria & Rippin[102] by modifying the IP with including the transfer of prior information from kinetic databanks. The proposed MIP is applicable to relatively low complex (non)linear kinetic models, being simple, rapid, and reliable even for poor-conditioned estimation cases. MIP does not have convergence problems, and does not require tuning factors or model linearizations. MIP transforms the ODE set model into an algebraic one by using classical IP. Then, prior information from a similar process, modelled by the same type of kinetic equations, with known kinetic curves $(\boldsymbol{C}',\boldsymbol{t}')$, parameters and variance $(\hat{\boldsymbol{k}}',\hat{\boldsymbol{V}}')$ (the so-called "historic process") is used concomitantly with the "current process" data [of known kinetic curves $(\boldsymbol{C}, \boldsymbol{t})$ and unknown parameters and variance $(\hat{\boldsymbol{k}},\hat{\boldsymbol{V}})$]. Selection of similar data is based on a similarity index, accounting the same number of observed species, evidence of similar curve shapes, and quasi-constancy of the rate ratios for the "current" and "historic" processes in a reduced concentration – time domain. Prior information $(\boldsymbol{C}',\boldsymbol{t}')$ and $(\hat{\boldsymbol{k}}',\hat{\boldsymbol{V}}')$ is included in estimation by dividing the IP relationships of observed similar species over a common reduced time domain Δt^*:

$$\frac{\dfrac{\Delta C_j}{\Delta t}\dfrac{\Delta t^*}{\Delta t^*}}{\dfrac{\Delta C'_j}{\Delta t'}} = \frac{M_{t^*}\left(\sum_z D_z G_z\right)}{M_{t^*}\left(\sum_z \widetilde{D}_z \widetilde{G}_z\right)} = \frac{\sum_z D_z M_{t^*}(G_z)}{\sum_z \widetilde{D}_z M_{t^*}(\widetilde{G}_z)} =$$
$$= \varphi \frac{D_d M_{t^*}(G_d)}{\widetilde{D}_d M_{t^*}(\widetilde{G}_d)}; \tag{25}$$

[$M_t(C) = (\int_{t_1}^{t_2} C\,\mathrm{d}t)/\Delta t$, denotes the integral mean of the concentration; $\Delta t = (t_2 - t_1)$; $j = 1,..., n_s$; $D, \widetilde{D}, G, \widetilde{G}$ = complex kinetic terms including $\boldsymbol{C}$, $\boldsymbol{C'}$, $\boldsymbol{k}$, $\boldsymbol{k'}$ variables]. From the resulted ratio of sums, the dominant term (index d) is identified over Δt^*, and then retained together with the left term of (25) in the MIP algebraic set. The parameters included in the dominant terms, corresponding to high rates, are thus estimated first in time-domains where their sensitivity and estimability is higher. The degree of dominance of a term may differ in various time sub-intervals, and an average of the kinetic constants coming from various observations and intervals is taken as a final MIP solution. The proportionality factor φ is fited to ensure a satisfactory process "superposition" on the same time-scale. Parameters included in poor conditioned terms can be adjusted according to their ratio from similar processes. MIP, similarly to EKF estimators, is also effective in on-line adapting of model parameters when changes in species observability occur,[36] or from preliminary process investigations.[118,119]

The current trend in using shortcut estimation techniques includes: i) the use of a combination of methods (DP, IP, principal component regression PCR, ridge selection analysis RSA); ii) problem decomposition for detecting sources of degeneracy, poor-conditioning, and over-parameterization; iii) extended sets of prior kinetic information from databanks.

Exact estimation solved with optimization methods

To solve an exact kinetic estimation problem, the objective function to be minimized (maximized) is formulated as a NLP eq. (20), or a MINLP eq. (21) problem. Optimization algorithms are of two types: gradient methods (using objective function and constraint derivatives), and gradientless methods (usually random algorithms, without gradient evaluations). Many (bio)chemical systems are difficult to be estimated by using classical gradient-based algorithms, due to the problem multimodality and nonlinearity, highly intercorrelated model parameters, species variable observability, particular model forms, incomplete data sets, or large number of explicite/implicite/mixed-integer nonlinear constraints. All these can cause serious convergence problems to the estimation rule, amplified by particular form of functions/constraints inducing problem non-convexities.[120] To overcome such difficulties, the use of an effective and suitable optimization algorithm, together with an initial data and model form analysis, is crucial in locating a global feasible solution.

Gradient methods (GM) for solving NLP use an iterative formula, which improves the previous (initial) estimate $\hat{\boldsymbol{k}}^{\text{iter}}$:

$$\hat{\boldsymbol{k}}^{\text{iter}+1} = \hat{\boldsymbol{k}}^{\text{iter}} + \lambda^{\text{iter}} \boldsymbol{s}^{\text{iter}}, \qquad (26)$$

where search direction vector ($\boldsymbol{s}^{\text{iter}}$) and step length λ^{iter} are adjusted iteratively according to each optimization algorithm. GM can be classified in methods of 1-st order (which use the first derivatives, or the Jacobian of objective function and constraints), methods of 2-nd order (which use first and second order derivatives, or the Hessian), and variable-metric methods (which use an approximation of the Jacobian and Hessian matrices). Among GM, it is to mention the classical methods of steepest descent, Gauss-Newton, Newton, Marquardt, conjugate gradients, Broyden, Powell, Davidon-Fletcher-Powell, Broyden-Fletcher-Goldfarb-Shanno, sequential quadratic programming, etc.[115,116,121,122] During the iterative search, the problem constraints are accounted in an augmented objective function (called Lagrange function), or during evaluation of the search directions and step lengths. GM are knew as being very rapid, requiring a reduced number of objective function evaluations, but also local convergent and very dependent on the initial solution choice. An increased reliability in locating a global feasible solution can be realized if GM search is re-started from various initial guesses (randomly generated in the feasible domain). Moreover, GM encounter difficulties in poor-conditioned, highly constrained problems, due to the necessity to accurately evaluate the function first/second order derivatives. Advanced implementations of GM try to overcome such drawbacks.[115]

For handling non-convex MINLP cases of estimation problems by applying GM, a series of sub-problems obtained from an appropriate decomposition of the original problem have to be solved. Algorithms based on the generalised Bender's decomposition[123] or the outer approximation[124,125] try to identify the sources of non-convexities. The original problem is decomposed in a "primal" (the original NLP problem solved for a fixed set of integer variables), and the "master" (a mixed-integer linear programming problem solved for continuous variables to provide new integer variable values). Partitioning is made such that the optimum integer variables can be determined independently of continuous variables. Since a number of constraints must be evaluated prior to the solution, the master problem is solved as a series of relaxed sub-problems. To find the global solution, Kocis & Grossmann[126] localise first the non-convexities by local/global

tests and then a relaxation is imposed on the invalid solutions. Floudas et al.[127] suggest partitioning of the original problem to ensure that global solutions to the primal and master sub-problems are attained for all iterations. By partitioning the variables responsible for non-convexities, the master and the primal problems are iterated until no improvement can be obtained. Grossmann & Sargent[128] use a branch-and-bound procedure, by solving consecutive NLP sub-problems with appropriate branching criteria. In any variant (branch-and-bound; outer--approximation;[125] outer-approximation with equality constraint relaxation and augmented penalty;[126,129] generalized outer-approximation;[130] generalized cross decomposition[131]), the gradient methods need to identify and eliminate the sources of non-convexities by analysing the objective function/constraints, by performing transformations, or by solving a perturbed NLP problem around the local solution. Difficulties in solving poor-conditioned or highly constrained cases, and risk to eliminate feasible sub-domains, persist to most of GM.

Today, with the growing availability of powerful computational means, the optimization routine efficiency is expressed more and more on their reliability and robustness to reach the global solution rather than on the computational cost (usually expressed as the number of objective function evaluations). Some other routine characteristics as simplicity, easy-to-use, amount of complementary calculations for search adjustment, independence on the initial solution guess, can be decisive in choosing the right optimization algorithm.

From this point of view, the gradientless *random searches* (RS, review of Maria[104]) become more and more attractive. Even if slower convergent than GM, they are able to by-pass most of the mentioned GM's difficulties in a simple way, being very reliable in solving complicated multimodal optimization problems, and being not very dependent on the initial solution choice. Recent RS reported highly effectiveness in solving non-convex MINLP problems with an increased reliability in finding the global solution. Published works reported very good results of using RS for solving kinetic identification problems, even for poor-conditioned or over-parameterized cases.[46,47,104,120,132,133]

RS use the same iterative relationship (26) to improve the estimate $\boldsymbol{k}^{\text{iter}}$, but search directions and step lengths are iteratively adjusted by means of random distributions. The random sampling of trial points can be fixed or variable, independent, uniform, or following the search progress, on a sphere or hypercube in the parameter-space. Step lengths can be fixed, of optimum-size, or adaptive, by considering the search history in the new point generation. MINLP problems of eq. (21) can be easily handled by completing eq. (26) with a random iterative formula also for integer variables, for instance:[120]

$$\begin{aligned} \boldsymbol{y}^{\text{iter+1}} &= \gamma + \text{INT}[\boldsymbol{Z}^{\text{iter}} \boldsymbol{M}^{\text{iter}} + \boldsymbol{Q}]; \\ \boldsymbol{M}^{\text{iter+1}} &= \boldsymbol{M}^{\text{iter}}(1-\varepsilon); \quad \boldsymbol{M}^{0} = \xi - \gamma, \end{aligned} \tag{27}$$

where: $\boldsymbol{Z}$ is a square-diagonal $(p\text{-}p_c) \times (p\text{-}p_c)$ matrix consisting of random numbers in the interval [0,1]; $\boldsymbol{Q}$ is a constant $(p\text{-}p_c) \times 1$ vector with elements equal to 0.5; $\boldsymbol{M}$ is the current search domain; ξ,γ are the domain upper and lower limits respectively; ε is a search contraction factor; p is the number of search variables; p_c is the number of continuous variables; INT denotes the integer operator returning the largest integer, less or equal to the operand. Such a search strategy provides an opportunity for the integer variables to span their range of possible values while the search regions for the continuous variables contract or expand.

Stop criteria of RS iterations are based on sufficient evidence of a global optimum, or on a cost decision for search interruption. The RS can be classified in several classes: pure RS; adaptive random search (ARS); simulated annealing (SA); genetic algorithms (GA); clustering algorithms (CA); evolutionary algorithms (EA).

Adaptive random searches iteratively modify the search (step-length, direction, checked domain, random point generator distribution) based on the failure/success of the previous steps. This strategy is usually completed with a periodic expansion and contraction of the search domain in order to refine the solution and to overcome local optima. The most important ARS sub-classes are:[104] RS with centroid generation, Luus-Jaakola's ARS class, and adaptive step length ARS, or combinations of these. Among effective ARS should be mentioned MWL,[134] Luus method,[135] ARDS,[136] ICRS,[137] M-SIMPSA,[138,139] SGA.[140] Banga & Seider[132] propose an effective ARS for handling complex NLP problems in dynamic process optimal control when gradient methods fail. Salcedo[120] and Cardoso et al.[138] present an effective ARS or combinations with SA for solving MINLP problems.

Mihail & Maria[141] and Maria[133] propose an effective multimodal-multilevel ARS (called MMA and respectively MMAMI) for handling non-convex NLP and MINLP estimation problems. The basic MMA increases the convergence rate and global solution reliability by using two strategies: a local pseudo-one-dimensional search, and a global pseudo-multimodal search. MMA presents only four procedure parameters that control the weight of the global/local search, contraction/expansion rate of the search domain according to the search his-

tory, and the degree of local solution refinement. The MMAMI adds the integer-variable search relationship (27) to the basic MMA.

Simulated annealing methods are based on the Markov chain theory, accounting only the last step information in directing the random search. In spite of their very low convergence rate, the reliability in solving complex global optimization problems is very high. The main difference vs. other RS consists in the possibility to accept, in some conditions, a detrimental search step with a Boltzmann distribution probability, thus surpassing local optima.[142,143,144] Cardoso et al.[138,139,145] propose a reliable SA coupled with the simplex method for solving non-convex MINLP problems, even if the computational effort is 2–3 orders of magnitude higher comparatively with those of ARS.

Genetic algorithms are RS in which the iterative random point generation presents similarities with biogenetic mutation and natural selection.[146] GA is conducted using information from a population of candidate solutions. Crossover and mutation operators generate new offsprings, while a fitness function controls the search progress. A GA iteration usually implies the following steps: i) starting from the current solution, random uniformly generate a certain number ("population size") of feasible points ("individuals"); ii) "individuals" are randomly divided in two or several subsets ("parents"); iii) from the "parental" subsets, two "individuals" are selected by means of a random or adaptive (proportional, ranking, etc.) rule; iv) "parental" vectors are rewritten in a certain code (for instance binary) becoming a "chromosome" of a certain length; v) a new "individual" is created by crossover of two "parental chromosomes"; vi) the "offspring" suffers several random "mutations" of parts of them, with a certain frequency; such a step allows exploring the whole search space, what only selection and crossover cannot fully guarantee; vii) finally, the mutant "individual" is decoded and used to check the search progress (objective function evaluation); viii) steps iii-vii are repeated a certain number of times ("number of generations") and the best iterative point is retained. GA improvements by using orthogonal crossover, effective "crowding" operators, and combinations with SA and CA, avoid common GA defect of early convergence, and increase the global search reliability in solving non-convex NLP/MINLP problems.[147,148]

Evolutionary algorithms (EA) are RS that mimic the evolution of the species in natural systems.[149] Like GA, EA random generate the "population" of tried points based on a "mutation" operator, and use only the objective function and constraints in ranking and eliminating tried points. Selection procedure, subsequent to a generation step, can be applied in two ways. $(\mu + \lambda)$ EA generates from μ parents, by mutation, λ-offsprings and then, from the $\mu + \lambda$ sorted members, the best μ become the parents of the next generation. (μ, λ) EA generates from μ parents, by mutation, $\lambda > \mu$ offsprings and then, from the λ sorted members, the μ best of them become the future parents. Costa & Oliveira[150] report effective EA for solving MINLP problems. Pham[151] proposes a general competitive EA, in which search results from several "populations" or "families" (using different search strategies) allow to set the procedure operators and parameters to be used in the next step. Both "best" and "stalled" populations are thus allowed to evolve but in a different manner. The analysis is exemplified for a combined EA and GA, while Wong & Wong[152] study an evolutive hybrid of GA and SA.

There are several analogies among ARS, GA and EA to be mentioned. For instance, GA population size can be assimilated with the ARS iterative population size; GA mutation frequency corresponds to EA offspring mutation frequency and to ARS control of local/global convergence; ARS periodic domain expansion and contraction is analogous to the increasing diversity of the GA population by using crowding schemes or dissolving "niche/clusters";[148] competition-cooperation among "families" in EA is equivalent with avoiding the "elitism" induced by the "fitness" function in GA, with the continuous switching between local and global search in ARS, or with the multistart local searches in clusters in CA.

There are frequent combinations among RS, or of RS with GM, in order to improve the global optimum search reliability or the local convergence rate. As an example, Maria[133] propose a combination of ARS and EA for solving non-convex MINLP. Parallel search is conducted in a certain number of "families", each family following a complete MMAMI search cycle. In general, there are two types of competition levels in an EA. One is the survival competition inside a "family", while the second level is the competition among "families" due to the search results separately obtained during a search cycle. Thus, after a cycle, families are line-up ranked as following: the "best family" (presenting the best estimation objective function) is placed on the top position while the "worst family" is placed on the bottom position. The aim is to confer to the families different mutation characteristics for the next search cycle: as a family is better ranked, as fewer mutations will be performed, and search will be focused on refining the local optimum; as a family is lower ranked, as larger mutations will be performed, allowing a global search on the whole space. In fact, this competition-cooperation strategy ensures for the bottom placed families

a higher chance to surpass local optima and to move to the top positions. For the up-ranked families, by keeping the same policy in the next cycle, would be possible to be moved to the rear and then another search policy will be set to surpass local optima.

Current trends in solving exact estimation problems is to use effective combinations of optimization algorithms to avoid local convergence, to reduce impact of initial solution choice, and to make the procedure more tractable and easier-to-use for a wider class of estimation problems.

Estimate quality analysis

For an identifiable kinetic system (see conditions of Vajda et al.[82]), the obtained kinetic estimate has to be checked for unicity and statistical quality. At the same time, the estimate must be checked, according to the available information, for their physical meaning and induced model predictions in variate reaction conditions (dynamic, oscillating, non-isothermal, stable quasi-steady-states, with/without mass transport limitations, under ideal/nonideal fluid flow, in the presence of various amount of inert or other components).

Estimate statistical quality is analysed from two perspectives: parameter inference based on a sensitivity analysis, and model adequacy based on comparison of residuals and experimental noise in direct relationship with the data quality and model structure uncertainty.[17,23,153] Estimate quality is rather influenced by the model structure and available data quality, than the optimization method effectiveness.[154] Because in most of (bio)chemical processes there is an excess of degrees of freedom in adjustable parameters than in observed and manipulated variables; adequate modelling can lead to multiple solutions, even if a reduced model structure is checked. Although, a complete theoretical analysis of solution unicity, stability and multiplicity is possible for ODE systems,[77] model discrimination via optimization routines and supplementary experiments, are preferred when coupled with the structure improvement.[41]

Statistical tests for estimate analysis are largely discussed in the literature.[16,17,23,45,116,155] The main tests are presented in Table 2 for the NLP estimation problem eq. (20). Detailed estimate checks can orient further decisions for model improvement via performing supplementary experiments, model extension/reduction, change of strategy for parameter estimation, or change of the estimation procedure. If the statistical analysis indicates high parameter intercorrelations (e.g. parameters which cannot be separately estimated precisely), this effect can be diminished by changing the model structure, or by performing new experiments (the estimate standard deviation is proportional with $n^{-0.5}$).[16] This can be realised by carrying out isothermal experiments to avoid simultaneously estimation of Arrhenius constants,[156] by performing experiments that reduce parameter intercorrelations,[16,157] by rejecting some parameters,[45] by performing parameter lumping,[82] by reducing the estimation vector dimensionality and fixing some insensitive parameters,[16,91,158] or by performing nonlinear transformations in model equations.[87]

There are frequent situations when the estimate significance tests indicate different parameters presenting approximately the same effect over model predictions, or a low effect. In this case, the estimation solution is ambiguous, even if the experimental data are satisfactory. An immediate effect is the poor-conditioning of the estimation problem (i.e. poor-conditioning of the Jacobian matrix, and of the estimate variance-covariance matrix), high parameter intercorrelations, and poor estimate significance. Several sources can be explored:[16]

i) if the parameters are uncertain and residues large but acceptable vs. the noise level, the estimate quality can be improved by adding new experiments in variate operating conditions;

ii) if the estimated parameters are uncertain but the residues small, the poor-conditioning is due to the model or data degenerated form (i.e. additive and/or multiplicative parameter terms, large noise, near-collinear data, non-scattered data). In this case, the information is insufficient to determine all the parameters and separate experiments have to be carried out for reducing parameter intercorrelations and for improving the data quality; reduction of the model structure can be also explored;

iii) if the estimate is uncertain and residues unacceptable, the model has to be rejected and reformulated.

Kinetic model reduction and discrimination

A complete estimate quality analysis (model adequacy and parameter inference) in direct relationship with the data quality offers information concerning model uncertainty and can indicate directions for structure extension or reduction. They are frequent cases of analysed (bio)chemical systems presenting a structural low kinetics identifiability, due to system characteristics (biomass changes, catalyst variability) and few observations vs. the high order internal state vector.[159] To overcome these difficulties, data from several experimental reactors, collected under variate operating

Table 2 – *Estimate analysis tests (n_s = no. of observations; p = no. of parameters; n = no. of runs; SSR = $\sum_{i}^{n_s}\sum_{u}^{n} e_{iu}^2$ = sum of squares of residuals; $e_{iu} = C_{iu} - \hat{C}_{iu}$, or $e_{iu} = \overline{C}_{iu} - \hat{C}_{iu}$; $i = 1,\ldots, n_s$; $u = 1,\ldots, n$)*

Test type	Relationship and remarks
Estimator tests:	
local/global solution tests	– see Floudas[177]
optimization convergence tests	– convergence rate, search progress tests[97]
efficient estimator	– $E(\hat{\boldsymbol{V}}) = \boldsymbol{G}^{-1}$; $\boldsymbol{G} = \sum_{u=1}^{n} \boldsymbol{J}_u^{\mathrm{T}} \Sigma_u^{-1} \boldsymbol{J}_u$; $\boldsymbol{J}_u = \left\lfloor \partial C_{iu} / \partial k_j \right\rfloor$ = Jacobian matrix
unbiased estimator	– $E(\hat{\boldsymbol{k}}(\boldsymbol{C}, x)) = \boldsymbol{k}$
asymptotic consistent estimator	– $\lim_{n\to\infty}(\hat{\boldsymbol{k}}(\boldsymbol{C}, x)) = \boldsymbol{k}$
invariant estimator	– does not depend on the k-vector dimensionality
sufficient estimator	– $P(\boldsymbol{k}_0)$, $P(\boldsymbol{k})$, $L(\boldsymbol{k})$, $P(\boldsymbol{C})$, $P(\varepsilon)$ offer maximum of information
robust estimator	– does not depend on the $\boldsymbol{k}_0$, $P(\boldsymbol{k}_0)$
Model adequacy tests:	
F-test	– $s^2 / s_{\mathrm{e}}^2 < F(n \cdot n_{\mathrm{s}} - p, n_{\mathrm{e}} - 1; \alpha)$; $s^2 = \mathrm{SSR}/(n \cdot n_{\mathrm{s}} - p)$; s_{e}^2 = approximated error dispersion from n_{e} replicates
χ^2-test	– (see Appendix)
Q-test	– $P(\Phi > \hat{\Phi}) = 1 - Q((n \cdot n_{\mathrm{s}} - p)/2, \hat{\Phi}/2) > 0.1$; (Q-statistics)[18,178]
r = correlation coefficient of (x, C)	– $\lvert r \rvert < 1$ (see Fogiel[155])
residual's plots (lack-of-fit)	– check prediction quality: C_{ui} vs. x; $\hat{C}_{ui}$ vs. x; [95 % HPD band of $\hat{C}_{ui}(x)$];
	– check residual randomness (runs test): e_{ui} vs. x;
	– check error for normality and uniformity; 'outliers' detection: $\hat{C}_{ui}$ vs. C_{ui}; C_{ui} vs. e_{ui}; $\hat{C}_{ui}$ vs. $e_{ui}/s_{e_{ui}}^2$; $\left\lfloor e_{ui}/\sigma_{uii}^2 \right\rfloor$ vs. $[e_{ui}/s_{e_{ui}}^2]$; where $s_{e_{ui}}^2 = \sigma_{uii}^2 \sqrt{1 - h_{uii}}$; $\boldsymbol{H}_u = \boldsymbol{J}_u (\boldsymbol{J}_u^{\mathrm{T}} \boldsymbol{J}_u)^{-1} \boldsymbol{J}_u^{\mathrm{T}}$; $\boldsymbol{H}_u = \left\lfloor h_{ij} \right\rfloor_u$ = "hat" matrix[17]; $\boldsymbol{J}_u = \left\lfloor \partial C_{iu} / \partial k_j \right\rfloor$ = Jacobian matrix
	– check model nonlinearity and error intercorrelations: $e_{ui}/s_{e_{ui}}^2$ vs. x;
	– check observation correlations: $\det\left\{\left[\sum_{u=1}^{n} e_{ui} e_{uj}\right]_{ij}\right\}$ eigenvalues plots;
	– 2D and 3D response surface plots, etc.
Estimate significance (inference) tests:	
t_j-tests	– k_j parameter Student test (see Appendix)
R_{ij}-tests	– (k_i, k_j) parameter's inter-correlation coefficient $\lvert R_{ij} \rvert < 1$ (see Appendix)
$(\lambda_j / \tilde{\sigma}^2)$-test	– (see Appendix); $\tilde{\sigma}^2 = \min(\sigma_{uii}^2)$
$k_j = \hat{k}_j \pm \mathrm{st.dev.}(k_j)$	– k_j parameter confidence interval (see Appendix);
parameter's joint confidence region at the $(1-\alpha)$ probability level	– $(\hat{\boldsymbol{k}} - \boldsymbol{k})^{\mathrm{T}} \boldsymbol{V}^{-1} (\hat{\boldsymbol{k}} - \boldsymbol{k}) \le p s^2 F(p, n \cdot n_{\mathrm{s}} - p; 1-\alpha)$, for Σ known; $\Phi(\boldsymbol{k}) \le \Phi(\hat{\boldsymbol{k}}) \left\lfloor \exp(\chi^2(n_{\mathrm{s}} - 1; 1-\alpha))/n \right\rfloor$, for Σ unknown;[41]
Special model analysis tests:	
discrimination among rival models	– Bard[16]; Froment & Hosten[41]
error-in-variable estimate tests	– Kemeny et al.[179]; Ricker[95]; Valko & Vajda[180]; Kim et al.[181,182]
model reduction tests	– QSSA tests for intermediate elimination;[4] species lumping tests;[4,165–168,170,172] parameter rejection tests;[45,83] parameter lumping tests;[82] parameter sensitivity tests;[16,91,158] global sensitivity tests.[44,164]

conditions, have to be used together with reduced mechanistic models and adaptive structures and parameters.[91] When kinetic information is insufficient to all precise model parameters and terms, estimate quality tests, parameter sensitivity analysis, physical restrictions, principal component and ridge parameter selection, it can all suggest model reductions in terms of reactions or variables.[4,45,83,160] Key-parameter subset selection and on-line model, updating with an effective estimation procedure has to compensate the system, data and model mismatch for leading to satisfactory predictions accounting the estimate uncertainty.[161] All the most influential terms in the kinetic and transport relationships (pH, temperature, nutrients, additives, (bio-)catalyst, mixing characteristics) have to be accounted even if a reduced kinetic model is adopted.

Insensitive parameters to the input data changes can cause intrinsic poor-conditioning of the estimator, biased and poor-quality solution (large estimate covariance matrix and high parameter intercorrelations), ambiguous estimate, and often optimization algorithm failure (especially for GM[158]), even if the experimental data are satisfactory. Several sources of difficulties can be explored:[16,82] insufficient experiments, non-scattered, very noised, or degenerated, near-collinear data (quantitative identifiability problem); model degenerated form, with additive and/or multiplicative terms for parameters (qualitative identifiability problem); model over-parameterization in respect to the insufficient experimental kinetic information about the process. If no supplementary information is available, a considerable increase in estimate quality is obtained by model reduction. The model reduction cost is a loss of information on certain species and reactions, a loss in model generality, prediction capabilities, and physical meaning of rate constants, and a biased estimate.

Kinetic model reduction is realised by using experimental methods and computational algorithms. The computational techniques perform lumping in concentration variables (states), in parameters, in reactions, or can reduce equation terms and remove side-reactions. From the large variety of techniques are mentioned the followings.

(i) Elimination of the reaction time variable from kinetic expressions.[73] In such a way, a reduced differential set of reduced reaction orders are obtained with the expense of lack of information when including parameters and species lumps in the model.

(ii) The use of pre-equilibrium assumption in mixed equilibrium/irreversible kinetic schemes.[73] This method is applicable to the consecutive reactions when at least one step is an equilibrium reaction.

(iii) Elimination of low concentration reaction intermediates by applying the steady-state approximation (QSSA).[4,73,162] In such a way, a reduction of system stiffness, due to low observable species present in small amounts, is realised with the cost of a loss in prediction capabilities. Species subjected to elimination are identified based on small product of target species i lifetime ($LT_i = -1/J_{ii}$) and their production rate.[4,163] Model simplification is realised with the expense of a corresponding bias in prediction capabilities for species i that is: $\Delta C_i^{\mathrm{s}} = \dfrac{1}{J_{ii}} \dfrac{\mathrm{d}C_i}{\mathrm{d}t}$; $J_{ij} = \dfrac{\partial f_i(\boldsymbol{C},\boldsymbol{k})}{\partial C_j}$.

(iv) Lumping of species in various ways:

(iv.1) Redundant species in a model are detected based on insignificant effect of concentration change on the rate of production of key species. Sensitivity coefficients $s_{im}(t) = \partial \ln(r_m)/\partial \ln(C_i)$ of species i in respect to reactions m are used to construct a global sensitivity measure $B_i = \sum_{m=1}^{n_r} s_{im}^2(t)$, by including all direct and indirect effects, in order to decide on redundant species (small B_i)[4,164]

(iv.2) An elaborated technique for linear and nonlinear species lumping in linear kinetic models have been proposed by Wei & Kuo[165] and Li & Rabitz.[166–168] Lumping of species is performed when there is insufficient information to characterize the dynamics of all compounds, or when by-products and intermediate separate prediction is not crucial for the process analysis. In other terms, the information is condensed in a smaller set including groups of species represented as single variables. The new lumped species $\hat{\boldsymbol{C}}$ are related to the original ones $\boldsymbol{C}$ by a lumping function h, which can be linear or non-linear:

$$\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t} = \boldsymbol{f}(\boldsymbol{C},\boldsymbol{k}) \Rightarrow \frac{\mathrm{d}\hat{\boldsymbol{C}}}{\mathrm{d}t} = \hat{\boldsymbol{f}}(\hat{\boldsymbol{C}},\hat{\boldsymbol{k}}); \quad \hat{\boldsymbol{C}} = \boldsymbol{h}(\boldsymbol{C}); \quad \mathrm{size}(\hat{\boldsymbol{C}}) < \mathrm{size}(\boldsymbol{C}). \tag{28}$$

For a linear lumping, $\hat{\boldsymbol{C}} = \boldsymbol{M}\boldsymbol{C}$, Wei and Kuo[165] and Li and Rabitz[166–168] indicate the necessary and sufficient conditions for an exact or approximate lumping, $\hat{\boldsymbol{f}}(\hat{\boldsymbol{C}}) = \boldsymbol{M}\boldsymbol{f}(\boldsymbol{C})$. For an exact lumping of kinetic model eq. (28), these conditions require that $\boldsymbol{M}\boldsymbol{f}(\boldsymbol{C})$ to be function of $\hat{\boldsymbol{C}}$, and to exist an inverse (or a generalised inverse) of $\boldsymbol{M}$, because $\boldsymbol{C} = \boldsymbol{M}^{+}\hat{\boldsymbol{C}}$.

For a linear kinetic model, with $\boldsymbol{f}(\boldsymbol{C},\boldsymbol{k}) = \boldsymbol{J}^{\mathrm{T}}\boldsymbol{C}$ in eq. (28), the *exact* linear lumping matrix $\boldsymbol{M}$ can be constructed from the eigenvectors of $\boldsymbol{J}^{\mathrm{T}}$ (e.g. $\boldsymbol{X} = [\boldsymbol{x}_j]$), because any subspace spanned by a subset of the eigenvectors is an $\boldsymbol{J}^{\mathrm{T}}$ invariant. As a con-

sequence, [span{0}, span{x_1},...] give a 1-dimensional lumping matrix $\boldsymbol{M}$; [span{x_1, x_2}, span{x_2, x_3},...] give a 2-dimensional lumping matrix $\boldsymbol{M}$ with rows formed with the $\boldsymbol{X}$-columns, etc. For an *approximate* linear lumping, there are several accepted lumping errors.[4] Although necessary and sufficient conditions for species nonlinear lumping have been derived,[169] the problem high complexity indicates no general rule to be followed (see some trials based on canonical forms of $\boldsymbol{J}^{\mathrm{T}}$ decomposition, reported by Tomlin et al.[4]).

(iv.3) Species lumping by applying the stochastic Markov chain theory is based on the complex reaction association to a Markov chain process moving toward a stationary state. Species presenting a comparable entropy index in this chain can be lumped in the same class.[170]

(iv.4) Species lumping based on a priori fuzzy information matrix concerning species similarities. A fuzzy matrix max-min composition rule is repeatedly applied leading to lump similar species. The loss of information by lumping is evaluated through a fuzzy entropy index.[170]

(v) Rejection of side-reactions by using the ridge selection analysis of rate constants (RSA).[45,83,111] This method analyses the parameter significance through a $\lambda_j / \tilde{\sigma}^2$ test (see Appendix), completed with the estimate plots vs. the Hoerls' RSA factor (α, see Appendix), *t*-tests, and parameter intercorrelations analysis. For linear kinetic systems, as the Hoerl factor increases as the estimate quality improves, becoming more and more independent on the experimental information, and biased.[111,158]

(vi) Lumping of kinetic parameters by means of principal component analysis (PCA).[20,82] The method is based on the ($p \times p$) information matrix $\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}$, defined in logarithmic terms, e.g. $\boldsymbol{S}^{\mathrm{T}} = [\boldsymbol{S}_1\ \boldsymbol{S}_2 \ldots \boldsymbol{S}_n]$, and $\boldsymbol{S}_u = \left[\dfrac{\partial \ln C_{iu}}{\partial \ln k_m}\right]$. Small eigenvalues λ_i of $\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}$ matrix, e.g. $\lambda_i < \sigma^2$, correspond to eigenvectors with proportional terms, from which are derived parameter multiplicative lumps of the form k_i / k_j^c (c = constant).

(vii) Reduction of the estimation vector size by fixing some of the low-sensitive/low-estimable parameters to their apriori values (PCR, principal component regression).[16,158,171] PCR identifies redundant parameters corresponding to small (near-zero) eigenvalues of the model Hessian matrix $2\boldsymbol{J}^{\mathrm{T}}\boldsymbol{J}$, $\boldsymbol{J} = [\partial \boldsymbol{f} / \partial \boldsymbol{k}]$. Low estimable parameters are fixed to their recommended apriori values, with an immediate effect of reducing the estimate dispersion, while the estimate becomes biased (see also an application given by Lei & Jorgensen[91]). As the size of the search parameter vector is smaller as the estimate is more independent on the data, reflecting the increased inability in determining the "true" parameter values. However, the PCR can be a very useful technique in determining an approximate estimate for highly intercorrelated parameters (for instance Arrhenius constants).

(viii) Reduction of the estimation vector size by using a detailed parameter sensitivity analysis. Insignificant parameters (index j) are rejected or fixed according to sensitivity individual rejection tests:

$$\Delta_{p,j} < 1(3); \qquad \text{where: } \Delta_{p,j} = \frac{1}{n}\sum_{u=1}^{n}\sum_{i=1}^{n_s}\frac{k_j}{C_{iu}}\left|\frac{\partial C_i}{\partial \boldsymbol{k}_j}\right|_{x_u} \tag{29}$$

($\boldsymbol{x}_u$ = independent variable vector for the run u). Parameter $\boldsymbol{k}_j$ sensitivity tests are defined vs. every species i and run u, leading to a complementary rejection test and to a global sensitivity index S_{gj}:[44,88,91,164]

$$s^*_{ij,u} < 0.01; \quad S_{g,j} = \sum_{u=1}^{n}\sum_{i=1}^{n_s} s^*_{ij,u}, \tag{30}$$

where:

$$s^*_{ij,u} = \frac{|C_i(t_u,(k_1,\ldots,k_j,\ldots,k_p)) - C_i(k_1,\ldots,0,k_p))|}{C_i(t_u,(k_1,\ldots,k_j,\ldots,k_p))}.$$

Elaborated software can automatically evaluate the sensitivity coefficients.[172]

When developing several reduced model alternatives, an important problem is to determine how to select the appropriate one. Several statistical tests have been developed in this respect, and a sequential experimental program can be designed to discriminate among rival models (see below paragraph). Discrimination tests are based on the goodness of fit and model parameter significance. For instance, if one selects between two rival models (M1) and (M2) from *n*-runs, one evaluates the ratio of the maximum likelihood functions for the two identified models:[41]

$$\frac{L_{\max,M1}}{L_{\max,M2}} = \sqrt{\left(\frac{s^2_{M1}}{s^2_{M2}}\right)^n}. \tag{31}$$

Rejection of one of the models can be based on a developed statistics. Thus, by choosing the positive numbers $A = \xi/(1-\alpha)$, and $B = (1-\xi)/\alpha$, $0 < B < 1 < A$, one can define the model acceptance levels $(1-\alpha)$ and $(1-\xi)$ (usually α, ξ are small numbers, like 0.05 or 0.1). Then, M1 is preferred for $L_{\max,M1}/L_{\max,M2} > A$, and rejected for $L_{\max,M1}/L_{\max,M2} < B$.

The same procedure can be repeated accounting several models, by taking as reference the model with the smallest predicted variance s^2 (e.g. proportional with Φ defined as a sum of squares of residuals, Table 1). Bates & Watts[17] propose an F-statistics and test to accept a reduced model (or "nested" model, including p_r parameters) vs. an extended one (with p parameters), over n experimental runs, and with a confidence level $(1-\alpha)$:

$$\frac{[\Phi_r(\hat{\boldsymbol{k}}_r)-\Phi(\hat{\boldsymbol{k}})]/(p-p_r)}{[\Phi(\hat{\boldsymbol{k}})]/(n-p)} \le F(p-p_r, n-p; 1-\alpha). \quad (32)$$

Sometimes the reduction tests are applied concomitantly with the estimation rule,[83] or by formulating the model estimation-reduction problem in terms of a MINLP problem eq. (21). Even if such a procedure can lead to the most adequate model, the reduced solution still has to be checked for parameter significance (Table 2), physical meaning, and requirement to preserve sufficient information from the process. In fact, application of the MINLP criterion for model reduction re-constructs the individual parameter sensitivities $s^*_{ij,u}$ of eq. (30), by successively trying to delete parameters from the model through the binary multiplicators $\boldsymbol{y}$.

Current trends in model reduction and discrimination concern the use of combined tests in order to diminish the risk of a false decision due to artificially inflated tests by some low sensitive parameters. Sequential model reduction and application of a MINLP criterion are modern tools increasingly used over the last decade. Another trend consists in application of a large variety of procedures to adapt the structure and model parameters according to the acquired information from the process and utilization scope, which results in a semi-automatically model up-grading rule (the so-called "tendency modelling").

Design of experiments for a precise parameter estimation and model discrimination

There are various rules to design experiments for better kinetic model identification. For instance, special methods try to reduce the problem dimensionality, in order to step-by-step disclose various features of the process mechanism, as followings:[73]

(i) isolation techniques (or pseudo-order techniques) which maintain a reactant to a constant (high) concentration, thus decreasing the apparent reaction order;

(ii) study of the reversible reactions close to the equilibrium conditions, leading to apparent reduced reaction orders;

(iii) neglect terminal stages of a process when slow-reactions (leading to by-products) become significant;

(iv) neglect initial stages of a process when only the final successive reaction steps are of interest;

(v) use an intermediate as a starting reactant to 'isolate' portion of the reaction schema; this route is very effective when applied to large reaction schemes with variable species observability during reaction;

(vi) mimic a reactant with another compound that cannot undergo the full reaction complexity;

(vii) change the reaction conditions that minimize kinetic complexity (experiments at very low or very high conversions, variate temperatures, pH, solvent/inert, etc.);

(viii) special choice of initial conditions that simplify the kinetic interpretation of data;

(ix) exploit analytical selectivity to simplify kinetic interpretation.

In general, an experimental program is designed to allow acquisition of maximum of information on the process over the experimental domain.[173] Oriented experiments try to improve the estimate quality or to make model discrimination easier. Experiments to reduce parameter inter-correlations and to increase estimate precision are based on the estimate covariance matrix $\boldsymbol{V}(\hat{\boldsymbol{k}})$ (see Appendix). For instance, the new experimental point $\hat{\boldsymbol{x}}$ is D-optimal designed by minimizing the determinant of $\boldsymbol{V}$ (i.e. minimum volume of the parameter confidence region). The A-optimal experiment design minimizes the trace of $\boldsymbol{V}$, that is the sum of individual parameter variances. The E-optimal experiment design minimizes the largest eigenvalue of the $\boldsymbol{V}$ matrix, thus making more spherical the parameter confidence region.

Other experiments are designed to increase the discrepancy in predictions among rival models, thus allowing a better discrimination based on a defined index.[16,174,175]

Conclusions

By reviewing the estimation rules applied over the last decade to identify kinetic models from (bio)chemical kinetic data, several trends can be pointed-out.

The current trend in the identification rules is to increase the importance of the preparative steps of data numerical analysis, including error matrix characterization, data reconciliation, stoichiometric checks, detection of species inter-connectivities, data redundancy, co-linearity, independent reactions, and reaction invariants.

Concerning kinetic modelling, the current trend is to use more structured and complex strategies to generate the model, accounting constraint representation, all types of information (in a standard or non-standard form), model synthesis by assembling elementary reactions and kinetic modules, through integrated platforms for model identification and process simulation. The modular approach and automatic generation of ODE, DAE, or stochastic kinetic models allow simulation of complex (bio)chemical systems by using information stored in databanks. Oriented and unified programming languages have been developed in this respect to include the system organization and complexity (for instance biocell representation, neurons, bio-informatic sequences, bio-polymer sequences, complex molecular structures, gene expression, gene-finding, etc).

Still popular quick shortcut estimators can offer an approximate solution by processing a quite large amount of information. Estimation solution reliability can be increased by using a combination of shortcuts (DP, IP, PCR, RSA) and exact estimators.[104,176] The trend in estimation is also to approach model/data degeneracy, poor-conditioning, and over-parameterization cases, by decomposing the model equations, and by using non-conventional process information to re-construct low identifiable terms. Transfer of prior information rules from databanks and analogous processes can lead to a better estimate, and can prevent further estimation convergence problems, infeasible or local solutions with no physical meaning.

Selection of an appropriate statistical estimation objective function, and search of the feasible global extreme with an effective optimization routine can lead to a satisfactory kinetic model solution. A MINLP formulation of the estimation problem allows a concomitant model estimation-reduction based on an adequacy criterion. Such a rule can save computational time if the procedure is completed with a parameter significance check. When rigorous statistical formulation is not possible, current estimation criteria tend to include all types of kinetic information (conventional or not), in a more complex way, through binary decision variables and a larger set of objective functions. The non-conventional estimations try to exploit the system global properties, such as regulatory effectiveness, process periodicity, or sensitivity to perturbations.

When solving the exact estimation problems, the trend is to use effective combinations of optimization algorithms to avoid local convergence, to reduce impact of initial solution choice, and to make the procedure more tractable and easier-to-use for a wider class of estimation problems. Beside reported improvements in using gradient methods (GM), the trend is to increasing use of random searches (RS) from various classes: adaptive, evolutive, genetic algorithms, or simulated annealing. Even if slower convergent, their combined use (sometimes with GM) presents the advantage of simplicity, better 'resistance' to poor-conditioned cases, independence vs. derivative accurate evaluation, increased reliability in obtaining the global feasible solution, easy handling of non-convex cases with complicated constraints and mixed variables, less dependence on the initial guess choice. Today, with the growing availability of powerful computational means, the optimization routine efficiency is expressed more and more on their robustness to reach the global solution rather than the computational cost. Routine availability and easy-to-use, low amount of complementary calculations, automatic procedure parameter adjustment, all these characteristics can be decisive in choosing the right optimization algorithm.

Concerning the estimate quality analysis, and the possibility to reduce or to extend the kinetic model, the current trend is to use a combination of tests for diminishing the risk of false decisions due to artificially inflated tests of low sensitive parameters. Detailed estimate quality analysis can detect low estimable cases, suggesting possibilities to reduce the kinetic model, by using rules for species and reaction lumping, or by fixing/rejecting by-reaction parameters. Sequential model reduction and application of a MINLP estimation criterion are modern tools to detect model redundancy, being increasingly used over the last decade. A large variety of procedures can adapt the structure and model parameters according to the acquired information from the process, which results in a continuous model up-grading ("tendency modelling").

Oriented experiment design techniques can lead to a considerable improvement in estimate quality and to a better discrimination among rival models.

Nomenclature

A – frequency Arrhenius factor, eq. (18); constant, below eq. (31)

$\boldsymbol{B}, B$ – matrix defined in eq. (8), $[m \times n_s]$; constant, below eq. (31)

$\boldsymbol{B}_i$ – species i global sensitivity measure, defined in 'model reduction' (iv.1) chapter

$\boldsymbol{C}$ – species concentration vector, $[n_s]$

det – determinant

$\boldsymbol{D}, D$ – molar balance matrix eq. (2), $[n \times n_s]$; kinetic terms defined in eq. (25)

$e_{ju} = C_{ju} - \hat{C}_{ju}$ – concentration residual matrix, $[n_s \times n]$

E – Arrhenius activation energy, eq. (18)

$E()$ – expected value

$\boldsymbol{E} = [E_{ij}]$ – matrix of atomic species, eq. (13), $[n_s \times n_a]$

f,f^* – functions defined in eq. (15)
$F,\boldsymbol{F}$ – volumetric flow rate in eq. (2); model functions; Fischer statistics; vector defined in eq. (21)
G – model functions in eq. (23); kinetic terms in eq. (25)
$\boldsymbol{g}$ – constraint functions defined in eq. (15)
$\boldsymbol{h}$ – lumping functions defined in eq. (28)
$\boldsymbol{H}$ – $|C_{ui} - \overline{C}_i|$ matrix below eq. (9), $[n_s \times n]$; "hat" matrix in Table 2
$\boldsymbol{I}$ – identity matrix
$\boldsymbol{J}$ – Jacobian matrix ($\boldsymbol{J} = \partial \boldsymbol{C}/\partial \boldsymbol{k}$, Table 2)
k_0 – constant in eq. (9)
$\boldsymbol{k}$ – kinetic rate constants
$\boldsymbol{K}$ – equilibrium constants
$L, \boldsymbol{L}$ – likelihood function, eq. (19); link matrix defined in eq. (12)
m – number of replicated runs, eq. (1); dimension of $\boldsymbol{B}$ matrix
m_j – j-th moment of a distribution
$\boldsymbol{M}, M$ – atomic or constraint matrix, eq. (3), $[n_s \times n_a]$; integral mean in eq. (25); search domain matrix (eq. 27); lumping matrix below eq. (28)
n – number of experimental runs
n_a – number of atomic species
n_e – number of replicates (Table 2)
$\boldsymbol{n}_i$ – independent columns of matrix $\boldsymbol{N}$
n_r – number of reactions
n_s – number of (observed) species
$\boldsymbol{N}$ – null space of $\boldsymbol{v}^T$ matrix (eq. 7)
N_j – moles of species j
p – number of parameters
$\boldsymbol{P}, P$ – projection matrix in eq. (5); probability distribution function, eq. (19)
Q – Q-statistics (Table 2)
$\boldsymbol{Q}$ – constant vector with elements ½ (in eq. 27)
$r, \boldsymbol{r}$ – correlation coefficients (below eq. 1, Tab.2); reaction rate vector in eq. (14), $[n_r]$
$\boldsymbol{R}, R$ – parameter intercorrelation matrix in eq. (A5); universal gas constant (eq. 18)
$\boldsymbol{s}, s$ – search direction vector in eq. (26); species sensitivity coefficients (in 'model reduction' (iv.1) chapter and eq. 30); square root of s^2
s^2 – estimate of the experimental error variance σ^2 (eq. 1); model error variance (eq. A2)
$\boldsymbol{S}$ – matrix of singular values from s.v.d. of $\boldsymbol{D}$ (before eq. 5); sensitivity matrix defined in 'model reduction' (vi) chapter
S_g – global sensitivity index (eq. 30)
SSR – Sum of Squares of Residuals (Table 2)
st.dev. – standard deviation
s.v.d. – singular value decomposition of a matrix
t – time; Student statistics
$\boldsymbol{T}, T$ – transformation matrix defined in eq. (11); temperature (absolute) in eq. (18)
$\boldsymbol{U}$ – matrix derived from s.v.d. of $\boldsymbol{D}$ (before eq. 5); matrix defined in eq. (A4)
$\boldsymbol{V}, V$ – matrix derived from s.v.d. of $\boldsymbol{D}$ (before eq. 5); parameter covariance matrix in eq. (A3), $[p \times p]$; volume
$\boldsymbol{x}$ – independent variable vector, eq. (19); eigenvectors of Jacobian $\boldsymbol{J}^T$ matrix, below eq. (28)
$\boldsymbol{X}$ – extent of reaction vector (eq. 4), $[n \times n_r]$; eigenvectors of Jacobian $\boldsymbol{J}^T$ matrix, below eq. (28)
$\boldsymbol{y}$ – binary variable vector, eq. (21)
w_{reg} – weights in the Tikhonov criterion (Table 1)
$\boldsymbol{z}$ – eigenvectors of $\boldsymbol{H}^T\boldsymbol{H}$ matrix, eq. (10)
$\boldsymbol{Z}$ – matrix containing $\boldsymbol{z}$ eigenvectors, eq. (11), $[m \times n_s]$; square diagonal random matrix (eq. 27)

Greeks

α – lower limit of parameters, eq. (15); confidence level; Hoerl factor, eq. (A4); constant below eq. (31)
β – reaction orders, eq. (14); upper limit of parameters, eq. (15)
Δ – difference; average sensitivity coefficients (eq. 29)
ε – measurement error matrix defined in eq. (15); contraction factor in eq. (27)
φ – multiplicative scaling factor used by the MIP (eq. 25)
χ^2 – χ^2 statistics
Φ – objective function, eqns. (19,20,21)
γ – lower limit of integers $\boldsymbol{y}$, eq. (21); lower limits in eq. (27); heteroscedasticity parameter (Table 1)
$\nu = [\nu_{ij}]$ – stoichiometric matrix, $[n_r \times n_s]$
λ – eigenvalue; search step length, eq. (26); number of "offsprings" in EA; weights in the Tikhonov criterion (Table 1)
μ – number of "parents" in EA
ξ – upper limit of integers $\boldsymbol{y}$, eq. (21); upper limits in eq. (27)
σ; σ^2 – error standard deviation; error variance
$\Sigma_u = [\sigma^2_{uij}]$ – error covariance matrix of run u, $[n_s \times n_s]$
$\|\bullet\|_2$ – Euclidean norm
$\|\bullet\|_\infty$ – infinite norm

Index

dep – dependent
in – input
ind – independent
out – output
0 – initial
ref – reference
tar – target

Superscipts

$\bar{\bullet}$ – mean value
$\hat{}$ – predicted or estimated value; lumped
T – transpose
+ – pseudoinverse
iter – iteration

APPENDIX
Some statistical tests to check model adequacy and parameter significance

For a kinetic data set with normal distributed noise $N(0, \Sigma_u)$, of diagonal covariance matrix $\Sigma_u = [\sigma^2_{uii}]$ (i,j = 1,… n_s, observed species; u = 1,..., n, number of runs), an applied weighted least squares (WLS) estimation criterion (Table 1) leads to the estimated parameters $\hat{\boldsymbol{k}}$ by minimizing the objective function:[17]

$$\min_{\boldsymbol{k}} \Phi(\boldsymbol{k}) = \sum_{u=1}^{n}\sum_{i=1}^{n_s}\left(\frac{\hat{C}_i(t_u,\boldsymbol{k}) - C_i(t_u)}{\sigma_{uii}}\right)^2,$$

subjected to:

$$\hat{\boldsymbol{k}} > 0; \quad g_j(\boldsymbol{C},\boldsymbol{k}) \geq 0; \quad j = 1,2,\ldots,m; \qquad \text{(A1)}$$

$$\alpha_i \leq k_i \leq \beta_i; \quad i = 1,2,\ldots,p.$$

Supplementary constraints are imposed to the kinetic parameters from physical meaning reasons and to fulfil mass balance constraints (for instance thermodynamical equilibrium constraints, reversible cyclic reaction constraints, etc.). The kinetic estimate $\hat{\boldsymbol{k}}$ can be used to analyse the model prediction capabilities, model adequacy, and parameter significance.

The sum of squares of residuals $SSR = \sum_{u=1}^{n}\sum_{i=1}^{n_s}(\hat{C}_i(t_u,\boldsymbol{k}) - C_i(t_u))^2$ indicates the global model adequacy, for instance through the χ^2 statistical test. If the following inequality holds:

$$\frac{s^2}{\tilde{\sigma}^2} = \frac{SSR/(n\cdot n_s - p)}{\tilde{\sigma}^2} < \chi^2(n\cdot n_s - p; 0.95),$$
$$\tilde{\sigma}^2 = \min(\sigma^2_{uii}), \quad u = 1,..., n; \quad i = 1,..., n_s, \qquad \text{(A2)}$$

then the model is adequate with a 95% probability. The adequacy is completed with the residual plots (Table 2). Such plots can reveal deviations from the constant noise hypothesis, presence of "outliers" in the experimental error distribution, quality of the model adequacy, systematic positive or negative residuals, and the residual magnitude compared with observations. The estimate significance is checked by means of the following matrices, numerically evaluated from the liniarized model around the estimate:[17]

$$E(\boldsymbol{V}(\hat{\boldsymbol{k}})^{-1}_{[p\times p]}) = \sum_{u=1}^{n}\sum_{i=1}^{n_s}\frac{1}{\sigma^2_{uii}}\left[\frac{\partial C_{iu}}{\partial \boldsymbol{k}}\right]^{\mathrm{T}}_{[p\times 1]}\left[\frac{\partial C_{iu}}{\partial \boldsymbol{k}}\right]_{[1\times p]},$$

(parameter covariance matrix); (A3)

$$\boldsymbol{U}(\hat{\boldsymbol{k}})_{[p\times p]} = \sum_{u=1}^{n}\sum_{i=1}^{n_s}\left[\frac{\partial C_{iu}}{\partial \boldsymbol{k}}\right]^{\mathrm{T}}_{[p\times 1]}\left[\frac{\partial C_{iu}}{\partial \boldsymbol{k}}\right]_{[1\times p]} + \alpha \boldsymbol{I},$$

(RSA matrix); (A4)

$$\boldsymbol{R}(\hat{\boldsymbol{k}}) = [R_{ij}]_{[p\times p]}; \quad R_{ij} = \frac{[\boldsymbol{V}(\hat{\boldsymbol{k}})]_{ij}}{\sqrt{[\boldsymbol{V}(\hat{\boldsymbol{k}})]_{ii}[\boldsymbol{V}(\hat{\boldsymbol{k}})]_{jj}}} \in [-1,1],$$

(parameter intercorrelations). (A5)

In eq. (A3), the formula is valid for the case of an efficient estimator (Table 2), $\boldsymbol{V}(\hat{\boldsymbol{k}})$ representing the lowest theoretically attainable variance.[16] In eq. (A4), α denotes the Hoerl factor, proposed as being $\alpha = \tilde{\sigma}^2$ by Maria and Rippin[45] ($\tilde{\sigma}^2$ denotes the $\min \sigma^2_{ii}$).

An estimated parameter $\hat{\boldsymbol{k}}_j$ can be considered significant in the kinetic model (with a probability of 95%), if its Student test is fulfilled, that is:

$$t_j = \left|\frac{\hat{\boldsymbol{k}}_j}{\sqrt{[\boldsymbol{V}(\hat{\boldsymbol{k}})]_{jj}}}\right| > t(n\cdot n_s - p; 97.5\,\%), \qquad \text{(A6)}$$

together with the "ridge selection" test:

$$\lambda_j / \tilde{\sigma}^2 > 1(3), \qquad \text{(A7)}$$

where λ_j are eigenvalues of matrix $\boldsymbol{U}$. The (A7) test, proposed by Maria & Rippin,[45] starts from the ridge-analysis of Hoerl, and the observation that, for low estimability cases of poor-conditioned $\boldsymbol{J}_i^{\mathrm{T}}\boldsymbol{J}_i$ matrix (where $\boldsymbol{J}_i^{\mathrm{T}} = [\partial C_u/\partial \boldsymbol{k}]$), the estimate presents large variances because:

$$E((\hat{\boldsymbol{k}} - \boldsymbol{k})^{\mathrm{T}}(\hat{\boldsymbol{k}} - \boldsymbol{k})) \approx \sum_{m=1}^{p}\frac{\sigma^2_{ii}}{\lambda^*_m} \geq \frac{\sigma^2_{ii}}{\lambda^*_p}, \qquad \text{(A8)}$$

(where λ^*_p is the smallest eigenvalue of $\boldsymbol{J}_i^{\mathrm{T}}\boldsymbol{J}_i$). As a consequence, small vales of λ_j in relationship (A7), inferring with the minimum noise level, corresponds to low sensitive parameters in the model.

A pair of less estimable, high correlated parameters $[\hat{\boldsymbol{k}}_i, \hat{\boldsymbol{k}}_j]$ (for instance, due to the insufficient available data, model over-parameterization, degeneracy in the data or model form) corresponds to high inter-correlation coefficients $|R_{ij}| \in [0.95, 1]$, leading to large confidence intervals of parameters:

$$k_j = \hat{k}_j \pm \{\sqrt{[\boldsymbol{V}(\hat{\boldsymbol{k}})]_{jj}}\}\, t(n\cdot n_s - p; 97.5\,\%); \; j = 1,\ldots,p. \qquad \text{(A9)}$$

References

1. *Senkan, S. M.,* Detailed chemical kinetic modeling: Chemical reaction engineering of the future, in: Wei, J., Anderson, J. L., Bischoff, K. B., Seinfeld, J. H. (Eds.), Advances in chemical engineering, Vol. 18, Academic Press, New York, 1992.
2. *Tsang, W.,* Chemical kinetic data base for hydrocarbon pyrolysis, Ind. Eng. Chem. Res. **31** (1992) 3–8.
3. *Tomlin, A. S., Pilling, M. J., Merkin, J. H., Brindley, J., Burgess, N., Gough, A.,* Reduced Mechanisms for Propane Pyrolysis, Ind. Eng. Chem. Res. **34** (1995) 3749–3760.
4. *Tomlin, A. S., Turanyi, T., Pilling, M. J.,* Mathematical tools for the construction, investigation and reduction of combustion mechanisms, in: Pilling, M. J., Hancock, G. (Eds.), Low temperature combustion and autoignition, Elsevier, Amsterdam, 1997, p. 293–437.
5. *Shimizu, T. S., Bray, D.,* Computational cell biology – The stochastic approach, MIT Press Math6X9, Cambridge, 2002.
6. *Maria, G.,* Evaluation of protein regulatory kinetics schemes in perturbed cell growth environments by using sensitivity methods, Chemical and Biochemical Engineering Quarterly (Croatia) **17** (2003b) 99–117.
7. *Bower, J. M., Bolouri, H.,* Computational modeling of genetic and biochemical networks, MIT Press, Cambridge, 2001.
8. *Tomita, M.,* Whole-cell simulation: a grand challenge of the 21-st century, Trends in Biotechnology **19** (2001) 205–210.
9. *Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J. H., Hunter, P. J., Juty, N. S., Kasberger, J. L., Kremling, A., Kummer, U., Le Novere, N., Loew, L. M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E. D., Nakayama, Y., Nelson, M. R., Nielsen, P. F., Sakurada, T., Schaff, J. C., Shapiro, B. E., Shimizu, T. S., Spence, H. D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., Wang, J.,* The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models, Bioinformatics **19** (2003) 524–531.
10. *Tomita, M., Hashimoto, K., Takahashi, K., Shimizu, T., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J. C., Hutchinson III, C. A.*, E-Cell: Software environment for whole-cell simulation, Bioinformatics **15** (1999) 72–84.
11. *Stiles, J. R., Bartol, T. M. Jr., Salpeter, E. E., Salpeter, M. M.,* Monte Carlo simulation of neurotransmitter release using MCell, a general simulator of cellular physiological processes, in: Computational Neuroscience, Bower, J. M. (ed.), Plenum Press, New York, 1998, pp. 279–284.
12. *Schaff, J. C., Loew, L. M.,* The Virtual Cell, Pacific Symposium on Biocomputing **4** (1999) 228–239.
13. *Schaff, J. C., Slepchenko, B. M., Choi, Y. S., Wagner, J. M., Resasco, D., Loew, L. M.,* Analysis of non-linear dynamics on arbitrary geometries with the Virtual Cell, Chaos **11** (2001) 115–131.
14. *Levenspiel, O.,* Chemical reaction engineering, Wiley, New York, 1999.
15. *Moser, A.,* Bioprocess technology: Kinetics and reactors, Springer-Verlag, New York, 1988.
16. *Bard, Y.,* Nonlinear parameter estimation, Academic Press, New York, 1974.
17. *Bates, D. M., Watts, D. G.,* Nonlinear regression analysis and its applications, Wiley, New York, 1988.
18. *Perry, R. H., Green, D. W., Maloney, J. O.,* Perry's chemical engineers' handbook, McGraw-Hill, New York, 1997.
19. *Middleton, G. V.,* Data analysis in the earth sciences using Matlab, Prentice Hall, New Jersey, 2000.
20. *Malinowski, E. R.,* Factor analysis in chemistry, Wiley, New York, 1991.
21. *Box, G. E. P., Hunter, W. G., MacGregor, J. F., Erjavec, J.,* Some problems associated with the analysis of multiresponse data, Technom. **15** (1973) 33–51.
22. *Bates, D. M., Watts, D. G.,* Multiresponse estimation with special application to linear systems of differential equations, Technom. **27** (1985) 329–339.
23. *Otto, M.,* Chemometrics. Statistics and computer application in analytical chemistry, Wiley-VCH, Weinheim, 1999.
24. *Stephanopoulos, G., San, K. Y.,* Studies on on-line bioreactor identification. 1. Theory, Biotechnol. Bioengng. **26** (1984) 1176–1188.
25. *Harmon, J. L., Duboc, P., Bonvin, D.,* Factor analytical modeling of biochemical data, Comp. & Chem. Eng. **19** (1995) 1287–1300.
26. *Narasimhan, S., Jordache, C.,* Data reconciliation and gross error detection: An intelligent use of process data. Gulf Publishing Co., Houston, 2000.
27. *Maquin, D., Adrot, O., Ragot, J.,* Data reconciliation with uncertain models, ISA Transactions **39** (2000) 35–45.
28. *Somogyi, R., Sniegoski, C. A.,* Modeling the complexity of genetic networks: understanding multigenetic and pleiotropic regulation, Complexity **1** (1996) 45–63.
29. *Akutsu, T., Kuhara, S., Maruyama, O., Miyano, S.,* A system for identifying genetic networks from gene expression patterns produced by gene disruptions and overexpressions, 9th Workshop on Genome Informatics, Tokyo, Dec. 10, 1998.
30. *Gillespie, D. T.,* Exact stochastic simulation of coupled chemical reactions, J. Physical Chemistry **81** (1977) 2340–2361.
31. *Gillespie, D. T., Mangel, M.,* Conditioned averages in chemical kinetics, J. Chemical Physics **75** (1981) 704–709.
32. *Stiles, J. R., Bartol, T. M.,* Monte Carlo modeling of realistic cellular structure and physiology, in: Computational Neuroscience: Realistic Modeling for Experimentalists, DeSchutter, E. (ed.), CRC Press, Boca Raton, 2000.
33. *Filippi-Bossy, C., Bordet, J., Villermaux, J., Marchal-Brassely, S., Georgakis, C.,* Batch reactor optimization by use of tendency models, Comp. & Chem. Eng. **13** (1989) 35–47.
34. *Harmon, J. L., Duboc, P., Bonvin, D.,* Application of factor analysis to the resolution of biochemical reaction networks, ESCAPE3 Symp., Graz (Austria), 1993.
35. *Terwiesch, P.*, Dynamic optimization of batch process operations with imperfect modeling, Ph. D. Diss 10857, ETH Zürich, 1994.
36. *Maria, G., Terwiesch, P., Rippin, D. W. T.,* Reliable on-line kinetics identification for a semi-batch process by parallel use of shortcut transfer of information rule and recursive estimators, Chem. Engng. Commun. **143** (1996) 133–147.
37. *Fotopoulos, J.,* Quantification of model uncertainty and its impact on the optimization and state estimation of batch processes via tendency modelling, Ph. D. Diss., Lehigh University, 1996.

38. *Hillewaert, L. P., Dierick, J. L., Froment, G. F.,* Computer generation of reaction schemes and rate equations for thermal cracking, AIChE Jl. **34** (1988) 17–24.
39. *Willems, P. A., Froment, G. F.,* Kinetic modeling of the thermal cracking of hydrocarbons. 1. Calculation of frequency factors. 2. Calculation of activation energies, Ind. Eng. Chem. Res. **27** (1988) 1959–1971.
40. *Park, T. Y., Froment, G. F.,* Kinetic modeling of the methanol to olefins process, Ind. Eng. Chem. Res. **40** (2001) 4172–4196.
41. *Froment, G. F., Hosten, L. H.,* Catalytic kinetics: modelling, in: Anderson, L., Boudart, M. (Eds.), Catalysis: Science and technology, Springer-Verlag, Berlin, 1981.
42. *Brendel, M., Mhamdi, A., Bonvin, D., Marquardt, W.*, An incremental approach for the identification of reaction kinetics, Preprints of the ADCHEM 2003, International Symposium on Advanced Control of Chemical Processes, Hong Kong, 2003. (submitted for publication, RWTH-LPT Aachen).
43. *Bardow, A., Marquardt, W.,* Incremental and simultaneous identification of reaction kinetics: methods and comparison, Chem Eng. Sci. 2004 (accepted).
44. *Vajda, S., Valko, P., Turanyi, T.,* Principal component analysis of kinetic models, Int. J. Chem. Kinetics **17** (1985) 55–81.
45. *Maria, G., Rippin, D. W. T.,* (1993) A note concerning two techniques for complex kinetic pathway analysis, Chem. Eng. Sci. **48** (1993) 3855–3864.
46. *Edwards, K., Edgar, T. F.,* General methods of kinetic model reduction, DYCORD – 4th IFAC Symp. Dynamics & Control of Chem. Reactors, Distillation Columns & Batch Processes, Copenhagen, June 7, 1995.
47. *Edwards, K., Edgar, T. F., Manousiouthakis, V. I.,* Reaction mechanism simplification using mixed-integer nonlinear programming. Comp. & Chem. Eng. 24 (2000) 67–69.
48. *Filippi, C., Greffe, J. L., Bordet, J., Villermaux, J., Barnay, J. L., Bonte, P., Georgakis, G.,* Tendency modeling of semibatch reactors for optimization and control, Chem. Eng. Sci. **41** (1986) 913–920.
49. *McCroskey, P. S.,* A knowledge-based approach for the development of complex kinetic mechanisms, Ph. D. Diss., Carnegie-Mellon Univ., Pittsburg, 1988.
50. *Maria, G., Rippin, D. W. T.,* Recursive robust kinetics estimation by using mechanistic shortcut technique and a pattern-recognition procedure, Comp. & Chem. Eng. **20** (1996) S587-S592.
51. *Valdes-Perez, R. E.,* Machine discovery of chemical reaction pathways, Ph. D. Diss., Carnegie-Mellon Univ., Pittsburgh, 1990.
52. *Valdes-Perez, R. E.,* On the concept of stoichiometry of reaction mechanisms, J. Phys. Chem. **95** (1991) 4918–4921.
53. *Valdes-Perez, R. E.,* A necessary condition for catalysis in reaction pathways, J. Phys. Chem. **96** (1992) 2394–2396.
54. *Knight, J. P.,* Computer-Aided Design Tools to Link Chemistry and Design in Process Development, Ph. D. Diss., Massachusetts Institute of Technology, 1995.
55. *Mavrovouniotis, M. L., Bonvin, D.,* Towards design of reaction paths, FOCAPD94 Conference, Snowmass Village (CO), July 10, 1994.
56. *Mavrovouniotis, M. L.,* Product and process design with molecular-level knowledge, in: Stephanopoulos, G. (Ed.), Artificial intelligence approaches in product and process engineering, Academic Press, New York, 1995.
57. *Prickett, S. E., Mavrovouniotis, M. L.,* Construction of complex reaction systems. I. Reaction description language. II. Molecule manipulation and reaction application algorithms, Comp. & Chem. Eng. **21** (1997) 1219–1254.
58. *Hedley, W. J., Nelson, M. R., Bullivant, D. P., Nielson, P. F.,* A short introduction to CellML, Philosophical Transactions of the Royal Society of London A **359** (2001) 1073–1089.
59. *Ichikawa, K.,* A-Cell: graphical user interface for the construction of biochemical reaction models, Bioinformatics Application Note **17** (2001) 483–484.
60. *Pedersen, F., Tyle, H., Niemela, J. R., Guttmann, B., Lander, L., Wedebrand, A.,* Environmental hazard classification, TemaNord Report 1994:589, NCM, Copenhagen, 1994.
61. *Xu, S., Nirmalakhandan, N.,* Use of QSAR models in predicting joint effects in multicomponent mixtures of organic chemicals, Water Res. **32** (1998) 2391–2399.
62. *Buxton, G. V., Greenstock, C. L., Helman, W. P., Ross, A. B.,* Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals in aqueous solution, J. Phys. Chem. Ref. Data **17** (1988) 513–759.
63. *Hoigne, J., Bader, H.,* Rate constants of reactions of ozone with organic and inorganic compounds in water, Water Res. **17** (1983) 173–194.
64. *CRGM*-Database, Centro di Ricerca in Genetica Molecolare, Istituto di Istologia e Embriologia Generale, University of Bologna, 2002. http://apollo11. isto. unibo. it/Cells. htm
65. NIH Database, National Institute of Health database, 2004. http:// molbio. info. nih. gov/ molbio/ db. html
66. *Bonvin, D., Rippin, D. W. T.,* Target factor analysis for the identification of stoichiometric models, Chem. Eng. Sci. **45** (1990) 3417–3426.
67. *Prinz, O., Bonvin, D.,* Monitoring discontinuous reactors using factor analytical techniques, IFAC DYCORD+ Symp., College Park, Maryland, 1992.
68. *Heinrich, R., Schuster, S.,* The regulation of cellular systems, Chapman & Hall, New York, 1996.
69. *Tsuchiya, M., Ross, J.,* Application of genetic algorithm to chemical kinetics: systematic determination of reaction mechanism and rate coefficients for a complex reaction network, J. Phys. Chem. A **105** (2001) 4052–4058.
70. *Maria, G.,* Relations between apparent and intrinsic kinetics of "programmable" drug release in human plasma, Chemical Engineering Science (submitted for publishing) (2004).
71. *Vance, W., Arkin, A., Ross, J.,* Determination of causal connectivities of species in reaction networks, PNAS **99** (2002) 5816–5821.
72. *Scacchi, G., Bouchy, M., Foucaut, J. F., Zahraa, O.,* Cinetique et catalyse, TecDoc, Paris, 1996.
73. *Connors, K. A.,* Chemical kinetics: The study of reaction rates in solution, VCH Publ., New York, 1990.
74. *Schuster, S., Schuster, R.,* Detecting strctly detailed balanced subnetworks in open chemical reaction networks, J. Math. Chem. **6** (1991) 17–40.
75. *Froment, G. F., Bischoff, K. B.,* Chemical Reactor Analysis and Design, Wiley, New York, 1990.
76. *Szedlacsek, E. S., Duggleby, R. G.,* Kinetics of slow and tight-binding inhibitors, Methods in Enzymology **249** (1995) 144–180.
77. *Compton, R. G.,* Comprehensive chemical kinetics. Vol. 32: Kinetic models of catalytic reactions, Elsevier, Amsterdam, 1991.

78. *Szabo, Z. G.,* Kinetic characterization of complex reaction systems, in: Bamford, C. H., Tipper, C. F. H. (Eds.), Comprehensive chemical kinetics, Elsevier, Amsterdam, 1969.
79. *Segel, I. H.,* Enzyme kinetics: Behaviour and analysis of rapid equilibrium and steady-state enzyme systems, Wiley, New York, 1975.
80. *Froment, G. F.,* The kinetics of complex catalytic reactions, Chem. Eng. Sci. **42** (1987) 1073–1087.
81. *Froment, G. F.,* Fundamental kinetic modeling of complex processes, in: Sapre, A. V., Krambeck, F. J. (Eds.), Chemical Reactions in Complex Systems, Van Nostrand Reinhold, New York, 1991.
82. *Vajda, S., Rabitz, H., Walter, E., Lecourtier, Y.,* Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models, Chem. Eng. Commun. **83** (1989) 191–219.
83. *Maria, G.,* An adaptive strategy for solving kinetic model concomitant estimation-reduction problems, Canadian J. Chem. Eng. **67** (1989) 825–832.
84. *Pritchard, D. J., Bacon, D. W.,* Statistical assessment of chemical kinetics models, Chem. Eng. Sci. **30** (1975) 567–574.
85. *Ratkowsky, D. A.,* A statistically suitable general formulation for modelling catalytic chemical reactions, Chem. Eng. Sci. **40** (1985) 1623–1628.
86. *Ratkowsky, D. A.,* Handbook of nonlinear regression models, Marcel Dekker Inc., New York, 1990.
87. *Espie, D. M.,* The use of nonlinear parameter estimation for dynamic chemical reactor modelling, Ph. D. Diss., Imperial College of Science and Technology, London, 1986.
88. *Holmberg, A.,* A systems engineering approach to biotechnical processes – experiences of modelling estimation and control methods, Acta Polytechnica Scandinavica (Mathematics & Computer Science Series), no. 33 (1981) 1–45.
89. *Schügerl, K.,* Bioreaction engineering: Reactions involving microorganisms and cells, Vol. 1,. Wiley, New York, 1985.
90. *Nielsen, J., Villadsen, J.,* Bioreaction engineering principles, Plenum Press, New York, 1994.
91. *Lei, F., Jorgensen, S. B.,* Estimation of kinetic parameters in a structured yeast model using regularisation, J. Biotech. **88** (2001) 223–237.
92. *Box, M. J., Draper, N. R.,* Estimation and design criteria for multiresponse nonlinear models with non-homogeneous variance, Appl. Statistics **21** (1972) 13–24.
93. *Mezaki, R., Draper, N. R., Johnson, R. A.,* On the violation of assumptions in nonlinear least squares by interchange of response and predictor variables, Ind. & Eng. Chem. Fundam. **12** (1973) 251–260.
94. *Stewart, W. E., Sorensen, J. P.,* Bayesian estimation of common parameters from multiresponse data with missing observations, Technom. **23** (1981) 131–141.
95. *Ricker, N. L.,* Comparison of methods for nonlinear parameter estimation, Ind. & Eng. Chem. Process Des. Dev. **23** (1984) 283–286.
96. *Seber, G. A. F., Wild, C. J.,* Nonlinear regression, Wiley, New York, 1989.
97. *Iordache, O., Maria, G., Corbu, S.,* Statistical modelling and estimation of chemical process models, Romanian Academy Publ. House, Bucharest, 1991.
98. *Yang, Q., Lindahl, P., Morgan, J.,* Dynamic responses of protein homeostatic regulatory mechanisms to perturbations from steady state, J. theor. Biol. **222** (2003) 407–423.
99. *Hudder, B., Yang, Q., Bolfing, B., Maria, G., Morgan, J. J., Lindahl, P. A.,* Computational modeling of iron metabolism in mitochondria, NIST Conference on Systems Biology Approaches to Healt Care: Mitochondrial Proteomics, Sept. 17, Gaithersburg (MD), 2002.
100. *Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E. D*., Metabolic network structure determines key aspects of functionality and regulation, Nature **420** (2002) 190–193.
101. *Hosten, L. H.,* A comparative study of shortcut procedures for parameter estimation in differential equations, Comp. & Chem. Eng. **3** (1979) 117–126.
102. *Maria, G., Rippin, D. W. T.,* Modified integral procedure MIP as a reliable shortcut method for kinetic model estimation: Isothermal, non-isothermal and (semi-)batch process cases, Comp. & Chem. Eng. **21** (1997) 1169–1190.
103. *Goodwin, G. C., Sin, K. S.,* Adaptive filtering prediction and control, Prentice-Hall Inc., New Jersey, 1984.
104. *Maria, G.,* Adaptive random search and short-cut techniques for process model identification and monitoring, FOCAPO98 Int. Conf. Found. Computer Aided Process Operations, Snowbird (US), July 5, 1998.
105. *Maria, G., Constantinescu, C., Ozil, P.,* Quick identification of the wastewater biological treatment process by using shortcut techniques and previous plant operation data, ESCAPE10 Conference, Florence (Italy), May 2000, p. 841–846.
106. *de Valliere, P., Bonvin, D.,* Application of estimation techniques to batch reactors. 3. Modelling refinements which improve the quality of state and parameter estimation, Comp. & Chem. Eng. **14** (1990) 799–808.
107. *Baratti, R., Alvarez, J., Morbidelli, M.,* Design and experimental verification of a nonlinear catalytic reactor estimator, Chem. Eng. Sci. **48** (1993) 2573–2585.
108. *Jang, S. S., Joseph, B., Mukai, H.,* Comparison of two approaches to on-line parameter and state estimation of nonlinear systems, Ind. & Eng. Chem. Process Des. Dev. **25** (1986) 809–814.
109. *Wojciechowski, B. W.,* The temperature scanning adiabatic plug-flow reactor, Canad. J. Chem. Eng. **70** (1992) 721–726.
110. *Wojciechowski, B. W., Rice, N. M.,* Temperature scanning plug flow reactor, Chem. Eng. Sci. **48** (1993) 2881–2887.
111. *Maria, G., Muntean, O.,* Model reduction and kinetic parameters identification for the methanol conversion to olefins, Chem. Eng. Sci. **42** (1987) 1451–1460.
112. *Fremaux, B., Lepoutre, G.,* Elements de cinetique et de catalyse, Technique et Documentation Lavoisier, Paris, 1989.
113. *Zwillinger, D., Krantz, S. G., Rosen, K. H.,* Standard mathematical tables and formulae, CRC Press, Boca Raton, 1996.
114. *Glowinsky, J., Stocki, J.,* Estimation of kinetic parameters – initial guess generation method, AIChE J. **27** (1981) 1041–1048.
115. *Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.,* Numerical recipes in Fortran: The art of scientific computing, Cambridge University Press, 1992.
116. *Constantinides, A., Mostoufi, N.,* Numerical methods for chemical engineers with Matlab applications, Prentice Hall, New Jersey, 1999.
117. *Himmelblau, D. M., Jones, C. R., Bischoff, K. B.,* Determination of rate constants for complex kinetics models, Ind. & Eng. Chem. Fundamentals **6** (1967) 539–547.

118. *Maria, G., Heinzle, E.,* Kinetic system identification by using shortcut techniques in early safety assessment of chemical processes, Jl. Loss Prevention in the Process Industries **11** (1998) 187–206.
119. *Maria, G., Heinzle, E.,* Testing novel shortcut methods for complex kinetic characterisation in early safety assessment of a chemical process, Comp. & Chem. Eng. **23** (1999) S71-S74.
120. *Salcedo, R. L.,* Solving non-convex nonlinear programming and mixed-integer nonlinear programming problems with adaptive random search, Ind. & Eng. Chem. Res. **31** (1992) 262–273.
121. *Fletcher, R.,* Practical methods of optimization, Vol. 1–2, Wiley, New York, 1980.
122. *Coleman, T., Branch, M. A., Grace, A.,* Optimization toolbox for use with Matlab, MathWorks Inc., Natick, 1999.
123. *Geoffrion, A. M.,* Generalized Benders decomposition, J. Optim. Theory and Appl. **10** (1972) 237.
124. *Duran, M. A., Grossmann, I. E.,* Mixed integer nonlinear programming approach for process systems synthesis, AIChE J. **32** (1986a) 592–606.
125. *Duran, M. A., Grossmann, I. E.,* An outer-approximation algorithm for a class of mixed-integer nonlinear programs, Math. Program. **36** (1986b) 307–339.
126. *Kocis, G. R., Grossmann, I. E.,* Global optimization of non-convex mixed-integer nonlinear programming (MINLP) problems in process synthesis, Ind. & Eng. Chem. Res. **27** (1988) 1407–1421.
127. *Floudas, C. A., Aggarwal, A., Ciric, A. R.,* Global optimum search for non-convex NLP and MINLP problems, Comp. & Chem. Eng. **13** (1989) 1117–1132.
128. *Grossmann, I. E., Sargent, R. W. H.,* Optimal design of multipurpose chemical plants, Ind. & Eng. Chem. Process Des. Dev. **18** (1979) 343–348.
129. *Viswanathan, J., Grossmann, I. E.,* A combined penalty function and outer-approximation method for MINLP optimization, Comput. & Chem. Eng. **14** (1990) 769–782.
130. *Fletcher, R., Leyffer, S.,* Solving mixed-integer nonlinear programs by outer approximation, Math. Prog. **66** (1994) 327–335.
131. *Holmberg, K.,* On the convergence of the cross decomposition, Math. Progr. **47** (1990) 269–281.
132. *Banga, J. R., Seider, W. D.,* Global optimization of chemical processes using stochastic algorithms, SIAM Conference on State of the Art in Global Optimization, Princeton Univ., April 28, 1995.
133. *Maria, G.,* ARS combination with an evolutionary algorithm for solving MINLP optimization problems, Proceedings 22nd IASTED International Conference on Modelling, Identification and Control, Innsbruck (Austria), February 10, 2003a, p. 112–118.
134. *Wang, B. C., Luus, R.,* Reliability of optimization procedures for obtaining global optimum, AIChE J. **24** (1978) 619–626.
135. *Luus, R.,* Determination of the region sizes for LJ optimization procedure, Hung. J. Ind. Chem. **26** (1998) 281–286.
136. *Martin, D. L., Gaddy, J. L.,* Process optimization with the adaptive randomly directed search, AIChE Symp. Series **78** (1982) 99–107.
137. *Casares, J. J., Banga, J. R.,* Analysis and evaluation of a wastewater treatment plant model by stochastic optimization, Appl. Math. Model. **13** (1989) 420–424.
138. *Cardoso, M. F., Salcedo, R. L., Feyo de Azevedo, S.,* The simplex-simulated annealing approach to continuous non-linear optimization, Comp. & Chem. Eng. **20** (1996) 1065–1080.
139. *Cardoso, M. F., Salcedo, R. L., Feyo de Azevedo, S., Barbosa, D.,* A simulated annealing approach to the solution of MINLP problems, Comp. & Chem. Eng. **21** (1997) 1349–1364.
140. *Salcedo, R., Goncalves, M. J., Feyo de Azevedo, S.,* An improved random search algorithm for nonlinear optimization, Comp. & Chem. Eng. **14** (1990) 1111–1126.
141. *Mihail, R., Maria, G.,* A modified Matyas algorithm MMA for random process optimization, Comp. & Chem. Eng. **10** (1986) 539–544.
142. *Schoen, F.,* Stochastic techniques for global optimization: A survey of recent advances, J. Global Opt. **1** (1991) 207–228.
143. *Boender, C. G., Romeijn, H. E.,* Stochastic methods, in Horst, R., Pardalos, P. M. (Eds.), Handbook of global optimization, Kluwer Acad. Publ., Dordrecht, 1995.
144. *Kalivas, J. H.,* Adaptation of simulated annealing to chemical optimization problems, Elsevier, Amsterdam, 1995.
145. *Cardoso, M. F., Salcedo, R. L., Feyo de Azevedo, S.,* Nonequilibrium simulated annealing: a faster approach to combinatorial minimization, Ind. Eng. Chem. Res. **33** (1994) 1908–1918.
146. *Goldberg, D. E.,* Genetic algorithms in search, optimization and machine learning, Addison Wesley, New York, 1989.
147. *Yu, H., Fang, H., Yao, P., Yuan, Y.,* A combined genetic algorithm / simulated annealing algorithm for large scale system energy integration, Comp. & Chem. Eng. **24** (2000) 2023–2035.
148. *Hanagandi, V., Nikolaou, M.,* A hybrid approach to global optimization using a clustering algorithm in a genetic search framework, Comp. & Chem. Eng. **22** (1998) 1913–1925.
149. *Schwefel, H. P.,* Evolution and optimal seeking, Wiley, New York, 1985.
150. *Costa, L., Oliveira, P.,* Evolutionary algorithms approach to the solution of mixed integer non-linear programming problems, Comp. & Chem. Eng. **25** (2001) 257–266.
151. *Pham, Q. T.,* Competitive evolution: a natural approach to operator selection, in: Yao, X. (Ed.), Progress in evolutionary computation, Springer Verlag, Berlin, 1995.
152. *Wong, K. P., Wong, Y. W.,* Development of hybrid optimization techniques based on genetic algorithms and simulated annealing, in: Yao, X. (Ed.), Progress in evolutionary computation, Springer Verlag, Berlin, 1995.
153. *Ziegel, E. R., Gorman, J. W.,* Kinetic modelling with multiresponse data, Technom. **22** (1980) 139–151.
154. *Stewart, W. E., Henson, T. L., Box, G. E. P.,* Model discrimination and criticism with single-response data, AIChE J. **42** (1996) 3055–3062.
155. *Fogiel, M.,* The statistics problem solvers, Research & Edducation Association, Piscataway, New Jersey, 1998.
156. *Pritchard, D. J., Bacon, D. W.,* Prospects for reducing correlations among parameter sstimates in kinetic models, Chem. Eng. Sci. **33** (1978) 1539–1543.
157. *Dodge, Y., Fedorov, V. V., Wynn, H. P.,* Optimal design and analysis of experiments, North-Holland, Amsterdam, 1988.

158. *Le Roux, C., Antonio, G., Joulia, X.,* Estimation de paramètres de systèmes algébro-differentiels, Récents Progrès en Génie des Procedés (Grenoble) **7** (1993) 43–50.

159. *Beck, M. B.,* Identification, estimation and control of biological wastewater treatment processes, IEE Proc. **133** (1986) 254–264.

160. *Treitz, G., Maria, G., Giffhorn, F., Heinzle, E.,* Kinetic model discrimination via step-by-step experimental and computational procedure in the enzymatic oxidation of D-glucose, J. Biotechnology 85 (2001) 271–287.

161. *Forbes, J. F., Marlin, T. E., MacGregor, J. F.,* Model adequacy requirements for optimizing plant operations, Comp. & Chem. Eng. **18** (1994) 497–510.

162. *Srinivasan. V., Aiken, R. C.,* Stage-wise parameter estimation for stiff differential equations, AIChE J. **32** (1986) 195–199.

163. *Turanyi, T., Tomlin, A. S., Pilling, M. J.,* On the error of the quasi-steady-state approximation, J. Phys. Chem. **97** (1993) 163–172.

164. *Turanyi, T.,* Review on sensitivity analysis of complex kinetic systems. Tools and applications, J. Mathematical Chemistry **5** (1990a) 203–248.

165. *Wei, J., Kuo, J. C. W.,* A lumping analysis in monomolecular reaction systems. 1. Analysis of the exactly lumpable system. 2. Analysis of approximately lumpable system, Ind. & Eng. Chem. Fundam. **8** (1969) 114–133.

166. *Li, G., Rabitz, H.,* A general analysis of approximate lumping in chemical kinetics, Chem. Eng. Sci. **45** (1990) 977–1002.

167. *Li, G., Rabitz, H.,* Determination of constrained lumping schemes for nonisothermal first-order reaction systems, Chem. Eng. Sci. **46** (1991a) 583–596.

168. *Li, G., Rabitz, H.,* New approaches to determination of constrained lumping schemes for a reaction system in the whole composition space, Chem. Eng. Sci. **46** (1991b) 95–111.

169. *Li, G., Rabitz, H., Toth, J.,* A general analysis of exact nonlinear lumping in chemical kinetics, Chem. Eng. Sci. **49** (1994) 343–361.

170. *Iordache, O., Maria, G., Pop, G.,* Lumping analysis for the methanol conversion to olefins kinetic model, Ind. & Eng. Chem. Research **27** (1988) 2218–2224.

171. *Marquardt, D. W.,* Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation, Technom. **12** (1970) 591–599.

172. *Turanyi, T.,* Reduction of large reaction mechanisms, New J. Chem. **14** (1990b) 795–803.

173. *Box, G. E. P., Hunter, W. G., Hunter, J. S.,* Statistics for experimenters, Wiley, New York, 1978.

174. *Buzzi-Ferraris, G., Forzatti, P., Emig, G., Hofmann, H.,* Sequential experimental design for model discrimination in the case of multiple responses, Chem. Eng. Sci. **39** (1984) 81–85.

175. *Buzzi-Ferraris, G., Forzatti, P., Canu, P.,* An improved version of a sequential design criterion for discriminating among rival multiresponse models, Chem. Eng. Sci. **45** (1990) 477–481.

176. *Vajda, S., Valko, P., Yermakova, A.,* A direct-indirect procedure for estimation of kinetic parameters. Comp. & Chem. Eng. **10** (1986) 49–58.

177. *Floudas, C. A.,* Nonlinear and mixed-integer optimization, Oxford Univ. Press, Oxford, 1995.

178. *Abramowitz, M., Stegun, I. A.,* Handbook of mathematical functions, National Bureau of Standards, Washington, 1964.

179. *Kemeny, S., Manczinger, J., Skjold-Jorgensen, S., Roth, K.,* Reduction of thermodynamic data by means of the multiresponse maximum likelihood principle, AIChE J. **28** (1982) 20–29.

180. *Valko, P., Vajda, S.,* An extended Marquardt-type procedure for fitting error-in-variables models, Comp. & Chem. Eng. **11** (1987) 37–43.

181. *Kim, I. W., Liebman, M. J., Edgar, T. F.,* Robust error-in-variables estimation using nonlinear programming techniques, AIChE J. **36** (1990) 985–993.

182. *Kim, I. W., Edgar, T. F., Bell, N. H.,* Parameter estimation for a laboratory water-gas-shift reactor using a nonlinear error-in-variable method, Comp. & Chem. Eng. **15** (1991) 361–367.