MODEL-BASED EXPERIMENTAL ANALYSIS: A SYSTEMS APPROACH TO MECHANISTIC MODELING OF KINETIC PHENOMENA

Wolfgang Marquardt
Lehrstuhl für Prozesstechnik, RWTH Aachen University
D-52056 Aachen, Germany

Abstract

A lack of mechanistic understanding of kinetic phenomena in chemical process systems is still the major bottleneck for a more widespread application of model-based techniques in process design, operations and control. Kinetic phenomena are becoming of increasing importance given the rapidly developing capabilities for the numerical treatment of more complex – typically distributed parameter – models on the one hand and the need for predictive models on the other. Surprisingly, little progress has been made in recent years towards systematic approaches to kinetic model identification despite the availability of powerful measurement techniques, modeling and simulation technologies on multiple scales as well as parameter estimation and structure identification techniques. This contribution will introduce a novel concept for mechanistic modeling of complex kinetic phenomena in chemical process systems. The approach aims at the integration of high resolution measurements, modeling on multiple scales and the formulation and solution of inverse problems in a unifying framework. An incremental approach based on gradual refinement of experimental techniques and mathematical models forms the core of the suggested methodology. The foundations of the approach developed in a collaborative interdisciplinary research center at RWTH Aachen will be discussed in detail.

Keywords

Modeling, kinetic phenomena, parameter estimation, model selection, inverse problems, differential-algebraic and partial differential equations, high resolution measurement techniques

Introduction

Kinetic phenomena drive the *macroscopic behaviour of process systems*. Most notably, the kinetics of chemical and chemical reactions play a decisive role in the manufacturing of bulk and specialty chemicals, pharmaceuticals or advanced materials. In single phase systems, macro- and micro-mixing interferes with chemical conversion if the time-scales of transport and reaction overlap. The correlation between chemical kinetics and transport phenomena is even more pronounced in multi-phase reactive systems because the location and extent of reaction depends on the kinetics of transport and reaction close to the interface. The interfacial area as well as the interface morphology are decisive for heat and mass transfer across the interface and hence for the selectivity and conversion of a multi-phase reaction system. The situation is getting even more complicated if complex fluids comprising of small and large molecules (such as proteins, oligomers and polymers) have to be considered. Then, reaction and transport kinetics strongly depend on the details of the molecular structure (and possibly even on the dynamics of inter- and intra-molecular interaction).

Kinetic phenomena not only determine the behaviour of the manufacturing process but also the *properties* and hence the *quality of the manufactured product*. In case of simple fluid products, quality just relates to chemical composition which is determined by the impurities of raw materials, the selectivity of the reactions and the efficiency

of downstream separation processes. If however structured materials such as solids (e.g. particles, fibres, thin films etc.) or multi-phase systems (foams, gels, porous solids etc.) are considered, kinetic phenomena are much more important, because kinetics determine the morphology of the structured material during its formation and equilibrium is sometimes not even reached.

Last but not least, kinetic phenomena are a key to understand the *function of biological systems*. For example, the metabolism of a single cell can be cast into a complex reaction network. The stoichiometry of the network and the kinetics of the individual reactions fix the metabolic pathway which is responsible for cell growth and for secondary metabolites production. In addition, signal transduction and regulation trigger switches between alternative metabolic pathways or even cell differentiation. These mechanisms again are implemented by means of reaction networks. Transport phenomena in the cell but also between cells in a cluster or even between organs interact with the bio-reactions and together realize the overall function and behaviour of a biological system.

Issues in kinetic modelling

Kinetic modelling on different length and time scales is a key technology

- for the design, control and operation of *manufacturing processes*,
- for the design and development of *structured or functional chemical products*, and
- for a better *understanding* of the *micro-scale phenomena* in (bio-)chemical systems.

Though process systems engineering has been largely focussing on the first of these areas, the most important drivers are expected to result from the second and third area in the future. They are governed by kinetics on the micro- and the meso-scales rather than the macro-scale.

Kinetic modelling of process systems is still a challenge despite the progress we have seen in the last two decades. There is still no systematic means to derive and validate models which *capture the underlying physico-chemical mechanisms* of an observed behaviour. There are clearly *fundamental limitations*. First, according to Popper (1959) a model as well as any other theory can never be strictly verified but only falsified. Therefore, kinetic modelling has to be orientated at a certain *engineering objective* in a *pragmatic sense*. Second, models are often confined to represent phenomena on a certain scale. For example, phenomena observed on the continuum scale are driven by the mechanisms on the molecular scale. Hence, *continuum scale modelling is not truly mechanistic* since an abstraction of molecular processes is inevitable. Typically, this abstraction has to be built partly on empirical elements in order to bridge the immanent gap between both scales. Hence, any mechanistic model of a continuum could be viewed as a hybrid model combining first principles knowledge on both scales with

observations gathered during well-designed experiments. Any successful kinetic modelling requires

- a carefully designed *experiment* equipped with (possibly a combination of) *appropriate measurement techniques*,
- *modelling* and simulation on *multiple scales* including the integration between adjacent scales,
- the formulation and solution of *inverse problems* to fit a model to the data, and
- methods for *inferring the most suitable model* structure from a set of possible candidates.

This list of requirements implicitly defines a coarse-granular research agenda. *Experimentation* should directly address interacting kinetic phenomena. This is in contrast to current practice, where their isolation is attempted by a suitably designed experiment. Typically, interaction cannot completely be avoided. Hence, a largely unquantifiable level of error cannot be avoided. *Measurement techniques* have to be developed to provide information on the major state variables in an experiment – ideally at high resolution in space and time rather than measurements at a few points in time or at a few spatial locations. Consequently, large amounts of data have to be acquired and processed during model identification. *Modelling* has to address kinetic phenomena on multiple scales. Means for bridging between the inherently differing levels of abstraction on adjacent scales are needed. Inevitably, modelling will move from differential-algebraic equation systems with relatively few parameters to partial differential-algebraic models with many parameters to properly capture the kinetic mechanisms on a high level of resolution. The resulting *inverse problems* are becoming much more demanding for these types of equations. Intelligent problem formulations and adaptive solution algorithms are a key to successfully solve such estimation problems. The *generation of candidate model structures* from molecular considerations and the subsequent *selection of the best model structure* for a given purpose has to be addressed. Besides further developments within these areas, a *systematic work process* has to be defined and supported by computational tools to guide the modelling team in applying and efficiently combining the various techniques.

The collaborative research centre CRC 540

The collaborative research centre CRC 540 "Model-based Experimental Analysis in Fluid Multi-Phase Reactive Systems" (http://www.sfb540.rwth-aachen.de/) at RWTH Aachen has been addressing these issues since 1999. The interdisciplinary research is carried out by a team of about 25 researchers from 11 different research groups at RWTH Aachen with widely differing areas of expertise including measurement techniques, transport phenomena, thermodynamics, chemical and biochemical reaction engineering, process systems engineering, scientific computing and numerical analysis. The research in CRC 540 is focussing on both, the development of a

systematic work process called *Model-based Experimental Analysis* (or *MEXA* for short) and its ingredients as well as a number of challenging modelling problems on the micro- and meso-scales in the area of fluid multi-phase reactive systems. The combination of method development and benchmarking creates a fruitful push and pull situation. A selection of topics which are being studied in the context of CRC 540 and related projects include

- reactions in homogeneous and heterogeneous fluid mixtures,
- multi-component diffusion in liquids and hydrogels,
- multi-component transport and reaction in single liquid droplets levitated in a liquid phase,
- enzymatic reactions in gel particles,
- transport and reaction in falling liquid films,
- heat and mass transfer in boiling processes.

These problems are used to benchmark the development of the *MEXA work process* and to support method development by formulating requirements for the individual steps. The most important activities of the work process are shown in the block diagram of Fig. 1.
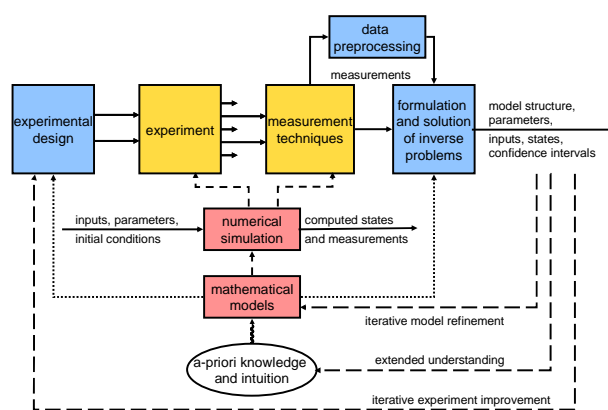

Figure 1. Model-based experimental analysis.

In a first step, a first *experiment* is designed on the basis of *a priori knowledge* and on the *intuition* of the experimentalist; a suitable *measurement system* is selected and implemented. Next, a first *mathematical model* of the experiment and the measurement system is built. *Numerical simulation* studies with this model reveal whether the measurements contain a sufficient degree of information to retrieve the kinetic process quantities of interest. Improvements of the experimental set-up and the measurement system can derived. Due to the unavoidable uncertainty in the model used, sensitivity studies with respect to the major model assumptions should be carried out. Typically, the measurements do not directly reveal those quantities which are most useful for modelling purposes. Again, the mathematical model can be employed to design the estimation techniques for inferring the quantities of interest from the measurements. Often, *data*

pre-processing has to be applied to the primary measurements before they are used in the estimation algorithm. Examples include the elimination of outliers or the calibration of spectral data to deduce concentrations (Martens and Naes, 1989). Pre-processing and estimation could, however, also be treated in an integrated manner (Albuquerque and Biegler, 1996; Tavitsainen et al., 2001). The *formulation and solution* of these *inverse problems* refers to combinations of state, parameter and unknown input estimation as well as model structure identification and model selection. Since these inverse problems are typically ill-posed in the sense of Hadamard (Engl et al., 1996; Kirsch, 1996), a simulation-based trial and error approach is employed to tune the solution algorithms to get reliable results. Last but not least, the model can be used to *design experiments* for maximum information content to obtain best results during the subsequent parameter estimation or model discrimination steps (Walter and Pronzato, 1990; Atkinson and Donev, 1992).

Typically, the first model does not reflect the real phenomena with sufficient detail and accuracy. Therefore, *iterative model refinement* intertwined with iterative *improvement of the experiment and the measurement techniques* has to be carried out to improve the predictive capabilities of the model based on the *extended understanding* gained.

A work process consisting of at least design of experiments, data interpretation and modelling dates back to at least the nineteen-seventies (e. g. Kittrell, 1970). However, its efficiency depends on the sequencing of the work process steps, on the type and quality of the methods applied during each these steps, and, in particular, on the strategies for the refinement of the experiment and the model structure. It has been only recently, that the development and benchmarking of such a work process has been formulated as an important research objective, e.g. by the CRC 540 team as well as by Asprey and Macchietto (2000). The work processes advocated by both teams consider iterative improvement. However, its power depends on the concrete strategies employed for systematically improving both, the model structure and the experimental set-up in every refinement step during model identification. While *experimental design* is the focus of the work of Asprey and Macchietto (2000), the research in CRC 540 is complementary and emphasizes the strategy for *model structure refinement*.

Simultaneous vs. incremental model identification

Before presenting the incremental model identification strategy of CRC 540, we will briefly summarize the commonly practiced simultaneous approach.

Simultaneous model identification

Consider for example the modelling of a two-phase (gas-liquid) stirred tank reactor where well-mixed

segregated phases can be safely assumed. A model of the reactor would require kinetic models for (i) the heat and mass transfer between the two phases, (ii) the heat transfer from the reactor to the cooling jacket and for (iii) the rates of the reactions in either of the phases or in the interface. For each kinetic phenomena several alternative model structures basing on different assumptions and theories may exist. The aggregation of such sub-models with the balance equations of both phases and the interface will inevitably lead to a multitude of candidate models. Experimental data are required to estimate model parameters, to assess model fit and to discriminate between the candidates using some measure of model validity.

Usually, the model fit is not sufficient and some improvement strategy has to be employed to compensate for the deviations between experimental data and model prediction. Typically, such *model improvement* is applied rather in an *ad hoc* than a *systematic manner*. For example, an alternative reactor model is suggested by replacing the kinetic sub-model which is suspect to be the most responsible for the poor reactor model quality. Sometimes, the complexity of the sub-models (and consequently the degree of detail they capture) is increased from low to high in an arbitrary sequence and combination.

In such an *ad hoc* iterative improvement approach the number of different model structures is quickly growing due to nested sub-models and the resulting combinatorial nature of the model selection problem. An appropriate strategy for finding the "best" model at "least" effort does not seem to be available. In addition, there are other largely unresolved issues which have been experienced by all practitioners:

- What if there are no good candidate structures for the kinetic model to start with?
- Which kinetic model contributes most to the lack of fit and should be switched subsequently?
- Is the model structure suitable?
- Is the information content in the experimental data sufficient for reliable identification?
- How to design an experiment for more valuable information to improve the model structure?
- How to deal with convergence and robustness problems of the nonlinear estimation algorithms?

This short discussion shows, that a more sensible modelling strategy is needed to gradually refine a model in a goal-oriented way, which guides the selection of candidate kinetic model structure and supports the validation of such a choice on scientific grounds.

Incremental model development and refinement

Before we continue our discussion on model identification, we want to recall that the *development of the model equations* itself can be carried out in a systematic manner (Marquardt, 1995). In a first modelling step, the balance envelopes are chosen and their interactions are determined, the intended spatio-temporal resolution of the model is decided, and the extensive quantities are selected for which balances will be formulated. In case of the illustrating two-phase reactor example, two well-mixed liquid and gas phases, interacting through a common interface, are chosen as balance envelopes. Mass and energy are selected to be balanced.

Subsequently, the balance equations are formulated and gradually refined as illustrated in Fig. 2. The decision on the structure of every balance equation on *level B* is guided by an assessment of the relevance of the physico-chemical processes occurring in the balance envelope. The balance is formulated as a sum of generalized fluxes, i.e. the hold-up variation, the inter- and intraphase transport as well as the source/sink terms. Note, that no constitutive equations are considered yet to determine these fluxes as a function of the intensive thermodynamic state variables (subsequently called states in short, i.e. the pressure, temperature and concentrations). In case of the reactor example, hold-up variation, convective flow into and out of the vessel, mass and energy transfer across the gas-liquid interface, heat transfer to the cooling jacket and chemical reaction in the liquid phase are considered.

On the next decision *level BF*, constitutive equations are specified for each flux term in the balances, i.e. the correlations for the hold-up, the interfacial fluxes, the heat loss as well as for the reaction rates in the reactor example.

Often, these correlations do not only depend on thermodynamic state functions (i.e. density, heat capacity etc.) but also on rate coefficients (i.e. reaction rate, heat and mass transfer coefficients etc. in the reactor example) which themselves depend on the states. Consequently, the decision for choosing the structure of the expression relating rate coefficients and states has to be taken on yet another *level BFR*.

This *cascaded decision process* can continue as long as the sub-models considered do not only involve constant parameters but also functions of the states.
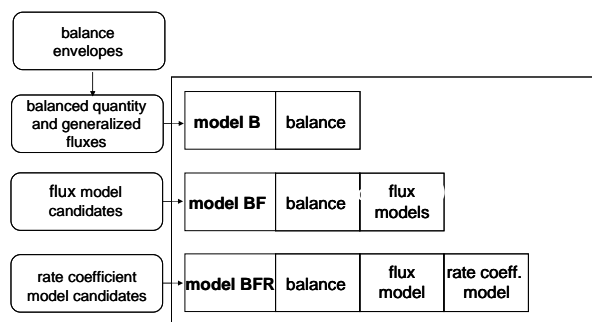

Figure 2. Gradual refinement strategy during the development of model equations.

Such a structured approach during the compilation of the essential model equations for each of the balance

envelopes renders the individual decisions in the modelling process completely transparent: the modeller is in full control of the model refinement process.

It should be noted, that the sub-models chosen on any of the decision levels do not necessarily have to be based on first principles. Rather, any mathematical correlation can be selected to fix the dependency of a flux or a kinetic coefficient as a function of intensive quantities in the sense of black-box modelling. This way, a *certain type of hybrid (or grey-box) model* (Psichogios and Ungar, 1992) arises in a natural way by combining first principles models fixed on previous decision levels with an empirical model on the current decision level.

Incremental model identification

The systematic specification of a process model and its sub-models is also a perfect starting point for devising a systematic work process for model identification as illustrated in Fig. 3.
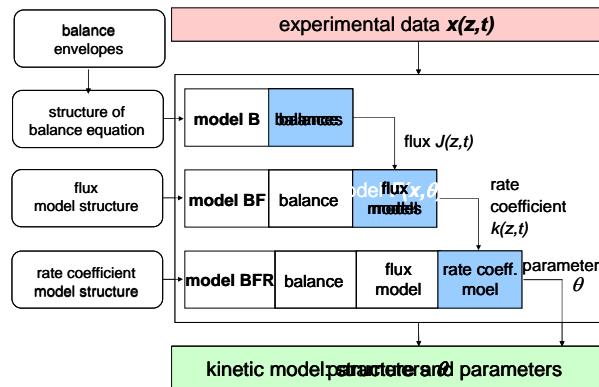

Figure 3. Incremental model refinement.

We assume to have a measurement system in place which provides *experimental data* for the states $x(z,t)$ at sufficient resolution in time t and space z. This data is sufficient – at least in principle – to *estimate an unknown flux $J(z,t)$* in the balance equation on *level B* as a function of time and space coordinates without the need for specifying a constitutive equation. This estimation problem corresponds in fact to an (approximate) inversion of the balance equation to solve for a flux $J(z,t)$ as a function of quantities which are either measured or inferred from the measurements. The flux estimates are interpreted as *inferential measurements* which, together with the real measurements, can be used in the next incremental identification step on level *BF* to *determine a rate coefficient $k(z,t)$* as a function of time and space, provided an appropriate flux model has been selected to relate the flux to rate coefficients, to measured states, and to their derivatives. Often, the flux model can directly be solved for the rate coefficient function $k(z,t)$. Finally, a model for the rate coefficients is identified on level *BFR*

which is assumed to only depend on the measured states and *constant parameters θ*. These parameters can be computed from the estimated rate coefficients $k(z,t)$ and the measured states $x(z,t)$ by solving an algebraic regression problem.

Let us again consider the two-phase reactor example. Concentration measurements (i.e. $x(t)$) are taken in both phases at a number of points in time. The liquid phase reaction and mass transfer fluxes (i.e. $J(t)$) can be estimated after differentiating the concentration measurements and solving the mass balances of both phases sequentially. Reaction rate and mass transfer coefficient data (i.e. $k(t)$) can be determined by directly solving the reaction rate model (e.g. Arrhenius' law) and the mass transfer model, respectively, using the estimated fluxes and the driving forces, which themselves can be computed from the measured concentrations. Finally, the kinetic coefficient data is correlated with the measured concentrations by some model containing constant parameters (i.e. θ).

A preliminary assessment of the incremental approach

The incremental approach has the potential to overcome a number of the disadvantages of the simultaneous approach discussed above.

Rather than postulating a potentially large number of nested model structures during model improvement in a largely *ad hoc* manner, a structured, fully transparent process is taken in the incremental model refinement strategy suggested. An uncontrolled combinatorial growth of the number of model candidates is avoided by the structured search process. The exponential growth of complexity in case of many uncertain sub-model structures is completely avoided. Any decision on the model structure relates to a single physico-chemical phenomena. Sub-model selection is guided by the previous estimation step which provides input-output data inferred from the measurements. Identifiability can be assessed on the level of the sub-model. This way, a lack of information content in the measurement system or an overparameterization of the sub-model can be discovered more easily. If there is no knowledge on good candidate models, a black-box model can be fitted first. The correlation quality obtained can act as a target for first-principles modelling on a finer scale (such as on the molecular level) to generate a suitable model structure.

The decomposition inherent to incremental model refinement also offers computational advantages. The solution of many difficult output least-squares problems with differential-algebraic or even partial differential-algebraic constraints with a potentially large number of data points resulting from high resolution measurement is completely avoided. Rather, an often linear inverse problem has to be solved first. All the following problems are nonlinear regression problems with algebraic constraints regardless of the complexity of the overall model. This decomposition not only facilitates

initialization and convergence of the estimation algorithms, but it also allows for incremental testing of model validity at every decision level for the sub-models. This transparency of the model identification process is lost in the simultaneous approach where only aggregated model mismatch can be observed. The computational effort is reduced drastically. Largely intractable estimation problems (such as those involving distributed parameter systems with many point measurements) may become computationally feasible.

Multi-step approaches to model identification have been applied rather intuitively by a couple of research groups before. Classical applications can be found in the reaction kinetics literature (e.g. Kittrell, 1970). In the so-called differential method, reaction fluxes are inferred from the experiment in order to correlate this data with measured concentrations. More recently, Tholodur and Ramirez (1996, 1999) as well as van Lith et al. (2002) have applied a two-step approach for the hybrid modelling of fermentation processes. First reaction fluxes are estimated from measured data, then neural networks and fuzzy models are employed to correlate the fluxes with the measurements. Mahoney et al. (2002) estimate the crystal growth rate directly from the population balance equations using a method of characteristics approach and indicate the possibility to correlate it with solute concentration next.

Though, the incremental refinement approach is rather intuitive, a systematic work process has not been reported and thoroughly analysed in the literature. If we consider *incremental model identification* as a *promising working hypothesis*, it is worthwhile to study the ingredients necessary for a successful implementation in more detail. In particular, we have to assess in the subsequent sections the capabilities of

- techniques for *high resolution measurement* of transient velocity, concentration and temperature fields and their calibration,
- algorithms for *model-free flux estimation* by an inversion of the balance equations,
- methodologies for the *generation, assessment and selection of the most suitable model structure* to relate fluxes to states (and their spatial gradients), and finally of
- methods for *model-based experimental design.*

Obviously, not all of these areas can be reviewed in full detail. We rather attempt to give a rough overview to assess the feasibility of the suggested incremental model identification approach and to derive new requirements.

High resolution measurement techniques

Measurements of the states should be carried out at high resolution in time and space to facilitate the first step of incremental model identification, i.e. model-free flux estimation. The requirements on the resolution depend on the concrete modelling problem. They can be assessed analytically (e. g. Bardow and Marquardt, 2004b) or by simulation studies.

Recent developments have been focussing on accessing data from fluid mixtures at *high resolution in space and time*. These techniques are applied *in-situ and* (if any possible) *non-invasive* to observe the system while the kinetic phenomena are occurring without disturbances due to the presence of an invasive probe. A selection of methods are summarized in the following.

Temperature measurements can be taken at high resolution in time by micro-thermocouples (e.g. Buchholz et al., 2004). Simultaneous temperature measurements at many spatial locations are possible but disturb the flow conditions in the sample as any other invasive technique. Infrared thermography is an alternative non-invasive measurement principle to access surface temperatures (e.g. Groß et al., 2004) at reasonable spatio-temporal resolution. The spatial resolution is correlated with the sample size due to the fixed number of pixels of the CCD chip used for signal acquisition.

Optical spectroscopy can provide *concentration measurements* with good accuracy at high resolution in time while simultaneous measurements at different locations in space are more difficult to achieve. However, concentration measurements are possible in fluids along a line (e. g. Bardow et al. 2003) or even on a plain (Krytsis et al., 2000) by means of Raman spectroscopy at good resolution. Concentration measurements can also be done in three spatial dimensions by means of nuclear magnetic resonance imaging (MRI), but the achievable accuracy and resolution is still limited (Maiwald et al., 2004).

Velocities can be measured in three dimensions by MRI. An appropriate design of the pulse sequences allows a wide range of resolutions in time and space (Fukushima, 1999, Gladden, 2003a,b). Alternatively, particle image velocimetry can be employed to get two- or even three-dimensional velocity measurements (Prasad, 2000). Tracer particles are added to the fluid and tracked by optical means. Particle trajectories and velocities can be computed from the primary signal. These methods can provide good accuracy and high spatial resolution. Averaging over time is unavoidable to get a reasonable signal to noise ratio.

The *disperse state of multi-phase systems* – in particular the interfacial area and its distribution – can be assessed by tomography methods such as MRI (Gladden, 2003b), ultrasound or electrical impedance imaging (Reinecke and Mewes, 1997). The accuracy and spatio-temporal resolution of these methods is yet limited, but very good visualization of the qualitative nature of the multi-phase flow is possible. Alternatively, probe detectors can be used if transient information at higher accuracy is required at a single spatial location. Such invasive techniques have been applied to particulate systems (Barrett, 2003) as well as to fluid dispersions (Cartellier and Achard, 1991) with great success.

Simultaneous measurement of all intensive quantities at high spatio-temporal resolution would obviously be desirable. There are opportunities for future research to

start with one of the established measurement principles and to work towards this ambitious goal. Raman spectroscopy as well as MRI are two principles which could be evolved into methods providing simultaneous information on more than one intensive quantity in space and time (e.g Goldbrunner et al., 2003; Gladden, 2003a).

The primary signal provided by these complex measurement systems is not the information one is interested in. Therefore, *calibration procedures* have to be applied to infer the quantities of interest. A prominent example is the conversion of spectral data into concentrations. Recently, chemometrics (or soft-modeling) methods (Martens, Naes, 1989) have attained a lot of attention. They are routinely applied to address all kinds of calibration problems. However, due to their linear nature, accuracy and range of validity are often limited. So-called hard modelling techniques try to capture some *a priori* knowledge on the measurement principle in order to overcome these problems (e.g. Alsmeyer et al., 2004).

From our experience, the accuracy achieved by current high resolution measurement techniques is often not sufficient for model discrimination. This is mainly due to the fact that the conversion of the raw data into useful measurement information is often not the core interest of the experimentalist. Highly accurate calibration and the quantification of the systematic and statistical errors (upper bounds or even statistical error distributions) are therefore often not attempted. Such information is also difficult to obtain, since it may require the proper modelling of the measurement equipment and the underlying processes which is a demanding and time-consuming task calling for process systems rather than experimental skills.

Calibration and the subsequent steps of model identification do not necessarily have to be tackled sequentially as indicated in Fig. 3. Rather, it is sometimes necessary to integrate both if calibration is not possible or difficult to achieve. For example, if the spectra of all the pure components are not available in a highly reactive mixture, simultaneous calibration and reaction model identification is a viable alternative (Amrhein et al., 1999; Taavitsainen et al., 2001, 2003).

Future efforts should focus on further improving the resolution of the measurements in time, space and chemical scales and on a high precision calibration in a large region of the state space. There seem to be enormous opportunities at the interface between measurement technology and mathematical modelling with the objective of facilitating proper physical interpretation of the raw measurement data and quantifying measurement errors.

Model-free estimation of generalized fluxes

For generalized flux estimation lumped and distributed parameter models should be distinguished. In lumped parameter models, generalized fluxes $J(t)$ only depend on time t and occur as an additive term in the balance equations. In distributed parameter models, generalized fluxes $J(z,t)$ depend on time and spatial coordinates. They can occur either in the interior or on the boundary of a spatial domain and hence show up as an additive term in the balance equation or in a boundary condition. Examples include (i) the time varying reaction and mass transfer fluxes in the gas-liquid reactor, (ii) the diffusive flux in a stagnant liquid which enters the differential mass balance as a function of time and space, or (iii) the heat flux from a falling liquid film to the surrounding gas, which enters the boundary condition and depends on time and the surface coordinates on the wavy film.

Flux estimation – an ill-posed inverse problem

Regardless of the concrete application context, the problems for estimating unknown generalized fluxes can be cast into an operator theoretic setting. Let $T_{w\Pi y}$ be the operator mapping an unknown flux $w \in \mathbf{W}$ into the observed quantities $y \in \mathbf{Y}$ according to

$$T_{w \to y} w = y \quad . \tag{1}$$

The sets $\mathbf{W}$ and $\mathbf{Y}$ are appropriate function spaces. The operator $T_{w\Pi y}$ is implicitly given by the model equations. Since measurements are always corrupted with error, the observed quantities y in Eq. (1) have to be replaced by the measurements

$$\tilde{y}(t) = y(t) + n(t) \quad , \tag{2}$$

before we can attempt to solve the model for the unknown generalized fluxes w. The estimation problem boils down to solving the operator equation (1) with measurements $\tilde{y}$

$$T_{w \to y} w = \tilde{y} \tag{3}$$

for the unknowns w. This problem is in general ill-posed in the sense of Hadamard (Engl et al., 1996; Kirsch, 1996), because, for all admissible data $\tilde{y}$, a solution $\hat{w}$ may either (i) not exist (ii), it may not be unique or (iii) it may be unstable in the sense that small variations in $\tilde{y}$ cause large variations in the solution $\hat{w}$.

In order to guarantee existence and uniqueness of the estimate $\hat{w}$ a generalized solution of the operator equation (3) has to be determined. This solution is given by

$$\hat{w} = \hat{T}_{w \to y} \tilde{y} \quad , \tag{4}$$

where $\hat{T}$ is the generalized inverse of the operator T which is a minimum norm least-squares solution of Eq. (3) if the L_2-norm is used (Engl et al., 1996, Kirsch, 1996). The quality of the solution, however, depends heavily on the continuity properties of $\hat{T}$.

In general, $\hat{T}$ is unbounded such that stability cannot be guaranteed. Regularization methods can be used to recover

stability. A regularization method is a family of operators T_α such that

$$\lim_{\alpha \to 0} T_\alpha y = \hat{T}_{w \to y} y \quad , \tag{5}$$

where α is a regularization parameter (Engl et al., 1996, Kirsch, 1996). The regularization operator is obviously an approximation of the generalized inverse for small α. As any approximation, regularization introduces an extra error to the estimate. If the observation error is bounded by some error level ε such that

$$\|\tilde{y} - y\| \leq \varepsilon \quad , \tag{6}$$

the error in the regularized solution is given by

$$T_\alpha \tilde{y} - \hat{T} y = T_\alpha (\tilde{y} - y) + (T_\alpha - \hat{T}) y \quad . \tag{7}$$

The first term is the so-called data error which is due to measurement errors whereas the second term accounts for the regularization error. In the limit of a vanishing regularization (i.e. $\alpha \prod 0$), the approximation error tends to zero but the data error tends to infinity because of the unboundedness of the generalized inverse. This simple result shows that any choice of T_α results in a trade-off between regularization error (usually leading to a biased estimate) and data error (contributing to the variance for zero mean errors $n(t)$ or to both, bias and variance in the general case). The choice of an appropriate regularization level α is the key for good estimation quality.

These inherent properties of any inversion problem in function space render the estimation of generalized fluxes a difficult undertaking. Only carefully designed inversion schemes will result in estimates $\hat{w}$ which can be used in subsequent steps of incremental identification.

Many types of regularization operators have been suggested in the literature. Regularization can be introduced by filtering (Tikhonov and Arsenin, 1977), by truncated singular value decomposition (Engl et al., 1996) or Tikhonov regularization which stems from a penalized least-squares minimization of the equation residual of Eq. (3) (Engl et al., 1996). A completely different type of regularization results from the projection of the unknown function w from the function space $\mathbf{W}$ into some vector space $\mathbf{V}$ (Kirsch, 1996). Such a projection is necessary in any case, since all numerical techniques employed to solve the estimation problem require some finite-dimensional parameterization (or discretization) of the unknown function w. In all practical cases, the interplay between the unavoidable discretization and additional regularization in the problem formulation has to be carefully exploited (Binder et al., 2002; Ascher and Haber, 2001).

Any regularization operator has to be complemented by a *regularization parameter choice method* which establishes an optimal trade-off between data and regularization error. Most of the available strategies are based on residual norms. Some of them exploit *a priori* knowledge on the error level (cf. Eq. (6)) such as Mozorov's discrepancy principle (Kirsch, 1996), others do not require such information like the L-curve criterion (Hansen, 1998) or generalized cross validation (Engl et al., 1996). Often, more than one regularization parameter has to be introduced. For example, in case of a discretized solution of an inverse problem stabilized by means of some Tikhonov regularization during the numerical solution, the discretization level as well as the penalty parameter become regularization parameters. Unfortunately, little knowledge on systematic approaches to the choice of multiple parameters is yet available (Belge et al., 2002).

Unknown input estimation in lumped parameter systems

Since the unknown generalized flux shows up as an additive term in the balances of a lumped parameter model, the estimation problem is equivalent to an *unknown input* or *disturbance estimation problem* which has been extensively studied in systems and control theory. A quite general problem formulation is given by the input-affine nonlinear system

$$\begin{aligned} \dot{x} &= A(x) + B(x)w, \quad x(0) = x_0, \\ y &= C(x) + D(x)w, \end{aligned} \tag{8}$$

with state vector $x(t) \in \Re^n$, output vector $y(t) \in \Re^p$ and unknown input vector $w(t) \in \Re^q$. A, B, C and D are matrix functions of appropriate dimension. This model defines the operator $T_{w \prod y}$ in Eq. (1) mapping the unknown inputs $w(t)$ to the outputs $y(t)$ in case of zero initial conditions $x_0 = 0$.

Two different approaches may be distinguished for estimating the unknown input $w(t)$ from available (erroneous) measurements $\tilde{y}(t)$. One class of methods postulates a simple dynamic model for $w(t)$. Often, a simple integrator is chosen, because the variation of $w(t)$ is assumed to be small compared to the dominant time constant of the process. A state estimator (i.e. a Kalman filter or a Luenberger observer) is then applied to determine the state of the extended model (e.g. Kurtz and Henson, 1998). The estimation result, however, depends on the choice of the disturbance model structure (de Valliere and Bonvin, 1990). Alternatively, the unknown input estimation problem can be interpreted as an inversion problem: a dynamical system Σ_I is the inverse of another dynamical system Σ if the output of Σ acting as the input of Σ_I produces the input of Σ. A general solution to the inversion problem has been given by Silverman (1971) for linear systems, by Hirschorn (1979) for nonlinear systems, and by Daoutidis and Kravaris (1991) for nonlinear differential-algebraic systems.

Inversion theory clearly shows the inherent limitations and challenges. Invertibility requires $T_{w \prod y}$ to be of full rank. In simple terms, inversion is only feasible if the number of measured outputs p is at least as large as the

number of unknown inputs q. For example, if we want to estimate the reaction fluxes for each of the species in a (homogeneous) reactive mixture, type and number of concentration measurements have to be chosen such that the rank of the stoichiometric matrix of the measured species equals the number of independent reactions (Brendel et al., 2003).

If the necessary invertibility condition is fulfilled, a symbolic representation of the inverse of the dynamic system can easily be computed. However, it involves differentiation of each observed output $\tilde{y}_i(t)$. The order of differentiation corresponds to the relative degree β_i of output i. Hence, ill-posedness of the inversion problem can be related to (potentially higher order) differentiation of the measurements. This itself is a special linear inversion problem, where the input to a chain of β_i integrators is estimated from the measured output (Tikhonov and Arsenin, 1977). Hence, any nonlinear inversion problem (8) can be solved properly if a satisfactory solution for the (higher order) differentiation of a measured signal is available.

In the mathematical literature, a variety of numerical methods have been reported for the *differentiation of measured data*: (i) the classical method of Reinsch (1967) uses spline smoothing of the measurements and applies subsequent differentiation (Wahba, 1990; Hanke and Scherzer, 2001); (ii) different kinds of regularization methods have been reported in the inverse problems literature (e.g. Tikhonov and Arsenin (1977), Kirsch (1996), Engl et al. (1996)); (iii) recursive filters are widely employed in signal processing (Oppenheim and Schafer, 1999); (iv) multi-resolution methods have been reported recently (Abramovich and Silverman, 1998). In all cases, a proper choice of the (sometimes implicit) regularization parameter is a key for good estimation quality.

In process systems engineering, unknown input estimation has not yet been widely employed for the *model-free estimation of mass, heat or momentum fluxes in ideally mixed systems.* Most often, attention has been on the determination of reaction or heat of reaction flux profiles from transient concentration and temperature measurements. Applications include reaction kinetics identification (e.g. Hosten, 1979), reaction calorimetry (e.g. de Valliere and Bonvin, 1989), or reactor monitoring (e.g. Schuler and Schmidt, 1992). In reaction kinetics, straight forward differentiation of measured concentration data (e.g. by finite differences) is frequently employed by practitioners. Special smoothing techniques have been reported to avoid amplification of unavoidable errors in the concentration measurements in the reaction fluxes (e.g. Hosten, 1979; Kamenski and Dimitrov, 1993; Gonzalez-Tello et al., 1996; Yeow et al. 2003). State estimation techniques have been alternatively employed for this purpose (e.g. de Valliere and Bonvin, 1989, 1990; Schuler and Schmidt, 1992; Bastin and Dochain, 1990; Elicabe et al., 1995; Farza et al., 1998). The trade-off between bias and variance has not been addressed systematically. Rather, the choice of numerical parameters in a heuristic smoothing algorithm or the covariances in Kalman filtering has been largely based on trial and error using simulation studies.

More recently, the theory of inverse problems has been applied to better deal with the trade-off between bias and variance in the estimate of a reaction flux or a heat of reaction. Tikhonov-Arsenin filtering (Tikhonov and Arsenin, 1977) and spline smoothing (Reinsch, 1967) have been successfully employed to estimate reaction fluxes in simulated (semi-)batch chemical reaction experiments (Mhamdi and Marquardt, 1999, Brendel et al., 2003, Bardow and Marquardt, 2004) and to estimate growth and substrate uptake rates in simulated continuous fermentation experiments (Mhamdi and Marquardt, 2003). A Tikhonov regularization approach combined with adaptive discretization of the unknown input has been used for the estimation of the heat release in a simulated stirred tank reactor (Binder et al., 2002).

Boundary flux estimation in distributed parameter systems

In case of distributed parameter systems, the problem of unknown flux estimation cannot be cast into a single general problem formulation as it has been the case for lumped parameter systems. Rather, specific problem classes have to be introduced. Some example problems and associated solution approaches are presented subsequently.

Unknown flux estimation has been employed for a long time in the heat transfer community. The heat flux between a heat conducting body and a fluid cannot be measured directly. Instead, temperature measurements are taken inside the heat conducting body to infer the heat flux at the surface as a function of time t and of the spatial surface coordinates z_Γ (Beck et al., 1985). The unknown flux enters the boundary condition of the heat conduction equation. In this case, the operator $T_{w\Pi y}$ maps the surface heat flux $w(z_\Gamma,t)$ for example to the temperatures $y_i(t)=T(z_i,t)$ measured at p points in the interior of the heat conducting body by means of micro-thermocouples. The measurements are predicted from a solution of the three-dimensional heat conduction equation

$$\frac{\partial T}{\partial t} = a\Delta T, \quad T(z,0) = T_0(z), \quad \begin{aligned} -\lambda\nabla T\mid_\Gamma &= w(z_\Gamma,t), \\ -\lambda\nabla T\mid_\Lambda &= q(z_\Lambda,t), \end{aligned} \quad (9)$$

with Γ and Λ being the parts of the surface with unknown and with known boundary heat fluxes $w(z_\Gamma,t)$ and $q(z_\Gamma,t)$, respectively. Obviously, if the temperature field would be completely accessible by measurement, the three components of the unknown boundary heat flux vector $w(z_\Gamma,t)$ could be computed by differentiating the measured temperature field with respect to the three spatial coordinates. Hence, differentiation is again at the heart of this inverse problem.

Inverse heat transfer problems are ill-posed (Alifanov et al., 1995). They are particularly difficult to solve

because the estimated heat flux is not only a function of time but also of two spatial surface coordinates. If temperature measurements can be taken at a few spatial locations only, the achievable local spatial resolution of the surface heat flux estimate is limited. This can be easily understood by an analysis of an approximate heat conduction model which results from a method of lines discretization of the spatial coordinates, for example by means of finite elements (Lüttich et al., 2004). This discretization results in a high-order linear state space model of type (8) with $w(t)$ representing q time-varying parameters in the surface heat flux discretization and with $y(t)$ comprising p point measurements. Invertibility requires $p \geq q$. Hence, the spatial resolution of the surface heat flux estimate, manifested in the number of parameters q employed during discretization, is limited by the number of measured temperature time series p. This number is determined by the number of pixels of the CCD chip in thermography or by the number of thermocouples mounted in the heat conducting body. In addition to the unavoidable bias introduced by regularization due to discretization, the amplification of measurement noise has to be controlled in a systematic manner.

A large variety of *methods* have been developed for *inverse heat transfer problems* in the engineering (e.g. Beck et al., 1985) as well as in the applied mathematics communities (e.g. Alifanov et. al., 1995). This problem class may be considered as one of the major benchmark applications tackled by the latter to illustrate the potential of a novel method for the solution of an inverse problem. The available solution techniques are specifically designed to the problem characteristics (e.g. Raynaud, Bransier, 1986), or they are employing generic methods comparable to those used for flux estimation in lumped parameter systems such as extended state estimation (e.g. Marquardt and Auracher, 1990), Tikhonov regularization (Alifanov, 1994), or Tikhonov-Arsenin filtering (Blum and Marquardt, 1997, Lüttich et al., 2004). Most often, simple configurations are studied, where one point measurement is available to estimate a single spatially averaged time varying surface heat flux (see e.g. Marquardt, Auracher, 1990, Beck et al., 1996, for experimental studies). Only few papers deal with multi-dimensional problems, where (many) temperature measurements are available for surface heat flux estimation (e.g. Huang, Wang, 1999). These multi-dimensional inversion problems are numerically challenging even in the linear case. We have experienced the limitations of numerical techniques based on fixed discretization techniques (Lüttich et al., 2004) during our investigations of boiling processes. Current work is therefore focusing on the development of a fully adaptive numerical scheme to solve inverse heat transfer problems with Tikhonov regularization by means of a combination of multigrid and conjugate gradient techniques (Groß et al., 2004). The interplay between adaptive discretization and the choice of regularization requires special attention (e.g. Ascher and Haber, 2001).

Interior flux estimation in distributed parameter systems

Unknown fluxes may not only occur on the boundary but also in the interior of a distributed parameter system. In case of heat conduction, the heat flux in a solid body may be considered. The balance equation reads as

$$\frac{\partial T}{\partial t} = -\frac{1}{\rho c_p} \left[\frac{\partial q_1}{\partial z_1} + \frac{\partial q_2}{\partial z_2} + \frac{\partial q_3}{\partial z_3} \right] + \sigma \qquad (10)$$

in Cartesian coordinates $z=[z_1,z_2,z_3]^T$ with conductive heat flux $q=[q_1,q_2,q_3]^T$ and source σ. Assume first that $\sigma=0$ and that the temperature $T(z,t)$ could be perfectly measured as a function of space and time. Then, the left hand side of the heat equation follows from differentiation with respect to time t. The resulting partial differential equation can not uniquely be solved for the three components of the heat flux vector $q(z,t)$. Hence, model free flux estimation is not possible from the balance equation only. If a constitutive equation is chosen for the heat flux, such as e.g. Fourier's law, the equation reads as

$$\rho c_p \frac{\partial T}{\partial t} = \nabla \lambda \cdot \nabla T \quad , \qquad (11)$$

which, after differentiation of the measured $T(z,t)$ with respect to time t and to the spatial coordinates $z,$ becomes a partial differential equation for λ (provided ρc_p is known). In this case, two levels of the incremental refinement strategy in Fig. 3, namely level B of the balance equation and BF of the constitutive flux equations, have to be merged in order to determine the heat conductivity $\lambda(z,t)$ without specifying a model for this rate coefficient. The resulting estimation problem is not linear anymore as in case of an isolated treatment of level B. Instead, it is often of quasi-linear type, because the unknown rate coefficients parameterize the linear differential operator.

An even more interesting and closely related problem is the *determination of diffusive fluxes* in complex systems. Examples include diffusive transport in multi-component liquids (even without, but in particular with accompanying chemical reaction), in electrolytes mixtures, or in hydrogel matrices. In all these cases, the transport mechanisms are still not well understood. Model-free estimation of unknown diffusive fluxes from high resolution concentration measurements in a solid or a stagnant fluid is not possible in the most general case for the reasons just presented for heat flux estimation. The problem can readily be solved, however, if an experiment is designed where diffusion only occurs in a single spatial direction. Then, only one non-zero mass flux component has to be determined from differentiated concentrations measured along a line in the direction of the diffusive flux. Such a strategy has been followed by Bardow et al. (2003). Diffusion in the axial direction of a tube is observed by high resolution concentration measurements using Raman

spectroscopy. These measurements have been first differentiated with respect to time by means of spline smoothing (Reinsch, 1967) and subsequently integrated over the spatial coordinate to render an estimate of the diffusive flux as a function of time and the axial coordinate of the tube without specifying a diffusion model. This technique also carries over to multi-component diffusion (Bardow and Marquardt, 2004a) provided concentration measurements are available for every species.

Another case of interior flux estimation occurs if the source σ is non-vanishing and unknown. This source can be determined from temperature measurements if a model for the flux q is available. This problem has been extensively studied in the context of heat conduction (Fatullayev, 2002) as well as diffusion (Reeve and Spivack, 1994) problems. Obviously, parameterization is mandatory (Abou Khachfe and Jarny, 2001), if both, the transport and the source flux functions are unknown. They are not identifiable without hypothesizing a model structure due to their additive occurrence in Eq. (10).

A sequence of two different experiments can be carried out alternatively to avoid such parameterization. For example, in case of diffusion in a homogeneous fluid, the mass source (typically due to reaction) can always be determined from an experiment in a well-mixed vessel before an investigation of diffusive transport is attempted.

Open issues in model-free flux estimation

The studies on model-free *flux estimation for lumped parameter systems* have been focusing on the identification of not directly measurable mass and heat fluxes caused by reaction. The exemplary case studies cited above clearly reveal the advantages of employing a mathematically sound regularization approach (i.e. a regularization operator and an associated parameter choice method) in order to obtain a high estimation quality. In particular, no assumptions have to be made regarding a mathematical model underlying the unknown flux to be estimated as a function of time. The most pressing open theoretical problems for lumped parameter systems include parameter choice methods for multiple regularization parameters, methods for higher order differentiation and the rigorous treatment of unknown initial conditions. Further, benchmarking by means of experimental studies on reactive and multi-phase systems is yet largely lacking.

Model-free *boundary flux estimation in distributed parameter systems* have been focussing on heat transfer problems. For example, the mechanistic modelling of boiling heat transfer has been one of the drivers. There, the interactions between phase transition at a solid-fluid interface, heat and mass transfer and the complex fluid dynamics of the vapour-liquid flow close to the boiling surface are not yet understood. Still, local heat flux estimation is possible on a heater surface, since level *B* in Fig. 3 is independent of all these physical phenomena. The

heat flux estimates are still not sufficient in resolution to adequately support model structure generation on level *BF*. Similar surface flux estimation problems arise with any other heat, mass or even momentum transfer problem between the phases of a multi-phase process system. In addition, the geometrical structure of the interface itself is subject to debate in case of fluid interfaces. Notable examples include the irregular surface wave structure of falling liquid films or of liquid drops under heat and mass transfer conditions. These problems are much more difficult than the inverse heat transfer problem because of the nonlinear and more complex nature of the balance equations for a multi-component transport problem in a fluid.

There is little experience yet on model-free *interior flux estimation in distributed parameter models.* Most papers deal with source flux terms (e.g. Reeve and Spivack, 1994, Fatullayev, 2002) and very few papers consider transport flux terms (e.g. Mahoney et al., 2002). Usually, interior transport fluxes are estimated only after introducing a constitutive equation (cf. Eq. (11)) resulting in a quasi-linear identification problem for the estimation of the space and time dependent rate coefficient (e.g. Hanke and Scherzer, 2001; Ascher and Haber, 2001). It seems that our own work in diffusion modelling (Marquardt and Bardow, 2004a, Bardow et al., 2004) is the only attempt to systematically address this issue and to exploit the linear structure of the flux estimation problem.

Inverse heat conduction problems, though difficult in themselves, are the simplest types of problems which are largely amenable to theoretical analysis and which are the most tractable for numerical experimentation. It seems to be the ideal problem class to reveal the inherent difficulties and to study solution alternatives for unknown flux estimation in distributed parameter systems. Most importantly, the "dimensionality" of the available measurements has to match the "dimensionality" of the estimated quantities to enable approximate inversion. In particular, the estimation of a multi-dimensional function (such as a surface heat flux) from a few measured time series (such as temperatures at some interior points) will typically not result in accurate estimates, even if state of the art techniques for the solution of inverse problems are employed. The always limited number of measurements accommodated by an experimental set-up constrains the achievable resolution in the estimate in principle. However, also the level of discretization of numerical solution algorithms limits the resolution in the estimate, because any discretization acts as a kind of regularization and therefore tends to smooth the estimated functions. Nevertheless, unknown flux estimation is possible with sufficient accuracy, if the problem formulation is carefully chosen such that the capabilities of both, the experiment and the inversion algorithm are pushed to their limits.

This discussion obviously reveals the challenges for future research: on the one hand the resolution and accuracy of measurements of temperature, concentrations and velocities on a surface and in a volume has to be

improved, on the other adaptive solution methods for properly regularized estimation problems have to be developed for three-dimensional transport problems. More experience on inverse transport problems in fluids has to be built up.

Kinetic model selection

As a result of unknown flux estimation on *level B* of the incremental model identification strategy, rectified measurements as well as estimated fluxes are available with high resolution (sampled in time and potentially in space). These data can be interpreted as a set of (inferential) measurements for the identification of a flux equation on the subsequent *levels BF* and *BFR* in Fig. 3. These constitutive equations relate the (inferred) fluxes to the (measured) states and to their (estimated) spatial gradients in case of distributed parameter models by means of a *purely algebraic relation*. Models for the flux as well as for the rate coefficient have to be selected on *levels BF* and *BFR*. The model selection problem consists of finding a single or a sub-set of models, which is consistent with the measured state and inferred state gradient and flux data, and which is best suited for a particular purpose (Verheijen, 2003). Model selection includes discrimination between different candidate kinetic model structures and the estimation of their parameters on both levels *BF* and *BFR*. These models are *always algebraic* regardless of the nature of the *process systems model* by construction of the incremental model identification procedure.

On *level BF,* a (set of) suitable constitutive flux equation(s) has to be selected to relate the flux to a rate coefficient, the states and the state gradients. No additional model for the rate coefficient is specified yet. Rather, the rate coefficient is determined as a function of time and space from the available data. For example, a diffusion coefficient can be computed after conjecturing the validity of Fick's law (Bardow and Marquardt, 2004a) or a reaction rate coefficient can be calculated after the structure of the reaction rate correlation has been fixed (Bardow and Marquardt, 2004b) at all sampling points (in space and time) for which measured or estimated information is available. Let us consider a first order reaction to be more specific: the reaction rate has been estimated on *level B* from concentration measurements. The rate coefficient $\hat{k}(t)$ can be estimated from

$$\hat{k}(t_i) = \frac{\hat{r}(t_i)}{\tilde{c}(t_i)} \quad , \quad i = 1,...n_s \quad , \tag{12}$$

at n_s sampling points t_i without postulating Arrhenius' law employing the estimated reaction flux $\hat{r}(t)$ and the measured or estimated concentration $c(t)$. Obviously, there is no redundancy in this step, if no identical repetitive experiments with concentration and reaction flux data at the same sampling points are available. Measurement

errors could therefore unacceptably affect the estimated rate coefficient. In particular, error amplification will show up for small concentration levels. However, this step can be very useful for the generation of model candidates, because it reveals (at least) the qualitative nature of rate coefficient function.

On *level BFR*, a model for the rate coefficient has to be selected. The model parameters (which are assumed to be constants) can be computed from the measured state and the estimated rate coefficient data. For example, if Arrhenius' law is postulated in case of the identification of a first order reaction, the constant parameters k_0 and E can be estimated from the estimated rate coefficient and measured temperature using

$$\hat{k}(t_i) = k_0 \exp(-E / R\tilde{T}(t_i)) \quad , \quad i = 1,...n_s \quad , \tag{13}$$

as constraints in the usual weighted least-squares problem. Typically, the rate coefficient models are of a nested nature, because they refer to quantities which themselves are state functions. For example, Maxwell-Stefan diffusion models depend on an activity coefficient model which itself depends on a saturation pressure model which depends on temperature. In principle, iterative refinement could be continued in a cascaded manner until only constant parameters show up in a correlation.

The levels *BF* and *BFR* can also be merged. This could be favourable for the sake of a better estimation quality. However, two types of models, one for the flux and another for the rate coefficients would have to be postulated at the same time. In case of the reaction rate example, merging both levels would result in an estimation of the constants k_0 and E from $n_s > 2$ measurements.

Regardless of a separate or a combined treatment of levels *BF* and *BFR,* a model selection problem has to be solved. Model selection requires a pre-selection of a set of candidate model structures, parameter fitting and the assessment of the predictive quality of all candidate models. These issues will be discussed next in more detail.

Generation of candidate model structures

The most difficult task crucial for the success of the modelling process in its entirety (Ljung, 1987) is the identification of candidate model structures to relate the fluxes to the state variables (and their spatial gradients). Since these structures are most often nested to account for the various state-dependent quantities, the generation of model structures has a combinatorial element. An experienced modeller with intuition and a profound understanding of the physico-chemical domain will suggest such nested candidate models after a careful inspection of the experimental results. This task of qualitative reasoning is typically carried out by a human expert in practice. Schaich et al. (2001) suggest however a computer-assisted method for the pre-selection of reaction kinetics models based on qualitative simulation.

Different strategies are employed to guide the *pre-selection of model structures.* They may result from (i) some molecular theory (e.g. Maxwell-Stefan for multi-component diffusion or reaction rate expressions for elementary reactions), from (ii) well-established empirical models (e.g. Arrhenius' law for the temperature dependence of the reaction rate constant or Ergun's equation for the pressure drop in a porous media), or from (iii) an aggregation approach (e.g. for composing a formal reaction kinetics expression from an understanding of the elementary reaction steps) to mention just some examples.

The model structures often result from simplifying and abstracting the kinetic phenomena on a more detailed level. For example, molecular modelling and simulation have often been used to derive correlations for continuum properties of materials. A classical example is the work of Barrett and Prausnitz (1975) who seem to have initiated this line of research for equations of state. Molecular simulation has been employed more recently to successfully derive correlation structures for diffusive transport (Liu et al., 1998; Merzliak and Pfennig, 2003). This research strategy seems to be very promising for the future to support the generation of novel model structures.

If a fundamental understanding of the kinetic phenomena of interest is lacking and if there are no resources available to acquire such understanding, purely *hypothetical mathematical structures* can be selected to correlate the experimental flux-state data. Obviously, this is the general problem of approximating a set of data by means of suitable functions. This problem has been tackled in different communities from various perspectives (e.g. Bates and Watts, 1988; Norgaard et al., 2000; Hastie et al., 2001). In this case, model structure selection typically requires two decisions, namely one on an appropriate class of regression models and another on the desired level of resolution. Examples for the first include the various types of neural networks or polynomial, hierarchical and wavelet bases. The level of resolution is the most critical decision to be taken. As in any inverse problem, a compromise has to be found to balance bias and variance. If many degrees of freedom are included in the regression model, the noise in the data is represented in addition to the true process response. If in contrast too few degrees of freedom are provided, the fine detail of the response cannot be resolved. This trade-off can be accomplished by an appropriate choice of the degrees of freedom in the regression model (e.g. the nodes in a neural net).

The choice of the *structure of the regression model* can be supported most elegantly if the residual norm decreases monotonically in a predictable manner if degrees of freedom are added in a certain sequence. Wavelet bases (Dahmen, 1997) and hierarchical bases (Yserentant, 1992) provide an estimate of the improvement of the goodness of fit in case of uniform refinement, where the basis functions on the next scale are added to capture more detail. Overfitting can be largely avoided if an appropriate level of resolution is chosen.

While the theoretical properties of wavelet bases are more favourable than those of hierarchical bases, the latter are preferred because of they provide straight forward generalization to high dimensions. Both types of functions have been employed in the past. For example, Bakshi and Stephanopoulos (2001) or Amato and Antoniadis (2001) discuss the use of wavelet bases, whereas Garcke et al. (2001) employ hierarchical bases on sparse grids to avoid the "curse of dimensionality", i.e. the exponential growth of the number of parameters for high-dimensional models.

Parameter estimation

After the selection of a set of candidate model structures, every model has to be fitted to the data. *Identifiability of the model parameters* cannot be taken for granted. Roughly speaking, we have to make sure that model responses differ significantly for different sets of parameters (Walter, Pronzato, 1990). Local methods check the rank of some information matrix prior to model fitting for an initial parameter estimate, whereas structural methods attempt to decide on identifiability for a large range of parameters. Functional expansion (Walter, Pronzato, 1990) and semi-infinite programming (Asprey, 2003) have been suggested. The latter approach is promising though improvement in global semi-infinite optimization is required to facilitate application even for the algebraic problems in incremental identification.

If the parameters in the algebraic models are identifiable, some measure of the distance between the measured or previously estimated data and the model predictions has to be chosen as the objective function to be minimized. Pragmatically, either a weighted L_1-, L_2- or L_∞-norm of the sum of the differences between the model predictions and the (state) measurements or (flux) estimates over all samples, or, alternatively, some likelihood function or a one of its derivates is minimized (Bates and Watts, 1988). Since both, the measurements as well as the estimated fluxes or rate coefficients contain errors, an error-in-variables (or orthogonal distance regression) approach is recommended in a balanced way (Britt and Lücke, 1978) though the error in the flux estimate is often much larger than the measurement error. An efficient solution algorithm for such estimation problems is available (Boggs et al., 1992).

In case of sub-models with a large number of parameters (e.g. black-box models) the formulation and the solution of the parameter estimation problem has to address the objective of balancing bias and variance in the estimate. The strategies employed match the ones we already have discussed in the context of model-free flux estimation. The most common regularization approaches include early stopping of the iterative solution algorithm and the introduction of some penalty term into the objective which corresponds to Tikhonov regularization. Generalized cross validation is a preferred method of choosing an appropriate level of regularization.

Alternatively, regularization can be accomplished by an adaptive choice of the level of detail in the regression model during parameter fitting. Various techniques have been reported in the neural network literature. Constructive learning algorithms gradually build up the number of network nodes and links between them (e.g. Chen et al., 1992; Kwok and Yeung, 1997, Marsland et al., 2002) whereas pruning methods start with a large number of nodes and eliminate those which are redundant and hence potentially introduce overfitting (e.g. Bärmann and Biegler-König, 1992; Psichogios and Ungar, 1994; Norgaard et al., 2000). A sensitivity based approach for adaptive selection of hierarchical basis functions in a sparse grid regression model combined with an L-curve stopping criterion has been reported by Brendel and Marquardt (2003), who extend the results of Garcke and Griebel (2001). Adaptive methods have shown to be favourable to achieve the desired bias-variance trade-off. More work needs to be done to improve computational efficiency, model fit and to provide a theoretical basis for adaptation for both neural network architectures and multi-scale bases.

Model adequacy tests and model ranking

There are different ways of testing the *adequacy* of the candidate models and to come up with a *ranking* of their relative prediction qualities (Verheijen, 2003). The selection criteria are of a multi-objective nature; they should reflect (i) the appropriateness of the model structure for reflecting the mechanisms underlying the physico-chemical phenomena observed, (ii) the intended use and the potential reuse of the model, and (iii) the capability of the model to fit the experimental data available. Only the latter goodness of fit criterion is left after an appropriate set of candidate models has been pre-selected and requirements on prediction quality have been formulated. The selection of a measure for quantifying the distance between the data and model predictions is crucial. According to the principle of parsimony, the model with the least complexity should be favoured. Various measures are available such as error variance, Akaike's information criterion, shortest data descriptor, Bayesian posterior probabilities etc. (Ljung, 1987). Still the choice of one of them is almost an art since it heavily influences the decision on the most adequate candidate model.

Model adequacy tests can be classified into inference, Bayesian and optimization approaches (Verheijen, 2003).

In *inference approaches,* a sequence of tests and decision is carried out to rank the models based on one or more statistical criteria and engineering insight. Verheijen (2003) presents a decision tree for this purpose. Bayesian approaches base model selection on the posterior probability for the model correctly predicting the next experimental observation. Such techniques have been reported by Stewart et al. (1998) for example for multi-response models.

In case of *optimization based approaches* model selection and ranking are carried out simultaneously. The model candidates are linked into a superstructure to result in a mixed-integer optimization problem. It can be solved either by genetic (McKay et al., 1997) or by mathematical programming methods (Skrifvars et al., 1998) to simultaneously identify the most suitable model from the set of candidates and to estimate its parameters. Though computationally and conceptually different, the adaptive refinement method of Brendel and Marquardt (2003) is related, since the functions of a given orthogonal multi-scale basis are also decided upon during the fitting process.

Discussion

Kinetic model selection problems are computationally much less demanding for the incremental identification strategy as compared to the common simultaneous approach. Incremental identification not only avoids the combinatorial character of model selection but also requires largely model selection for algebraic models. Hence, the wealth of methods available for model selection, identifiability analysis, model regression and validation are directly applicable and no generalization to differential equation problems is necessary.

Numerical parameter estimation problems are often difficult to solve due to the ill-conditioning of the optimization problem. This ill-conditioning has various sources including non-convexity, problem size and poor identifiability due to inappropriate model structures and lacking quality measurements. Robust parameter estimation methods are required to address this issue. For example, Arora and Biegler (2004) report a new trust region algorithm for parameter estimation in differential-algebraic models with favourable properties. Incremental identification suffers less from ill-conditioning because of a smaller problem size and full transparency with respect to the selection of suitable model structures and appropriate measurements.

Model-based experimental design

Regardless the identification strategy, the experiments have to be designed such that the information content in the data is maximized to facilitate model selection as well as parameter estimation. Model-based experimental design techniques have been developed for this purpose (Walter, Pronzato, 1990; Atkinson, Donev, 1992; Asprey, 2003). One has to distinguish between techniques for obtaining the best parameters in a given model structure and for discriminating best between alternative candidate models during model selection. Asprey and Macchietto (2000) present a work process for model identification which heavily relies on these concepts.

If the experiment is designed for *parameter precision* some measure of goodness of fit is maximized with the experimental conditions as the degrees of freedom and the

process model as the constraints. Typically, the objective function is defined as a measure (such as the determinant or the trace) of an information matrix related to the covariances of the estimated parameters. This criterion is local and assumes a reasonable reference for the parameter values. Alternatively, global criteria such as Markov chain Monte Carlo methods have been advocated (cf. Haario and Turunen (2003) for a demonstrating example).

An experimental design to best *discriminate between candidate models* tries to maximize some discrimination criterion by appropriately choosing the experimental conditions subject to the model constraints. Often, the objective function comprises a norm of the distance between all candidate models or employs a Bayesian argument (Verheijen, 2003).

Optimal experimental design problems are computationally demanding due to nonlinearity and nonconvexity. Further, second order derivatives are required in any gradient based algorithm. Though an efficient numerical algorithm has been reported for differential-algebraic models (Bauer et al., 2000) and a commercial product is available as part of gPROMS, more work is necessary to improve reliability and efficiency.

Though methods and tools are available for all kinds of optimal experimental design, their application to the incremental identification strategy is not straight forward. Since there is no model involved, experimental design does not make any sense if the flux estimation problem on *level B* in Fig. 3 is considered in isolation. Rather, the experiment should be designed such that the model selection on *level BF* and *BFR* is supported. Appropriate flux data has to be generated on level *B* to achieve a reliable model selection on subsequent levels. New problem formulations are required for this purpose. Further, experimental design for data-driven nonlinear regression models does not seem to have obtained sufficient attention yet. In particular in this case, the design should not only consider goodness of fit but also the capability of the model to interpolate between clusters of measurements and to extrapolate beyond the region where measurements have been taken.

Summary and future perspectives

We have introduced a *promising methodology for the incremental identification of kinetic models* of process systems. The discussion of the state of the art in the various areas clearly shows that such a strategy has become feasible largely due to the advances in high resolution measurement techniques and in the theory and the computational tools for the solution of inverse problems. Still, a lot of work needs to be done in the individual areas and in particular in the integration of the methods and tools to a powerful, ultimately computer-assisted work process of model-based experimental analysis.

In the next two subsections, we will point to open problems with the incremental approach and discuss future research opportunities.

A critique of the incremental approach

The major *advantages* of the incremental approach are (i) the full transparency of the modelling identification process and (ii) the computational efficiency of the tailored estimation algorithms. These advantages are a direct consequence of the exploitation of the inherent structure of the identification problem by decomposition.

However, the incremental approach also reveals at least two rather obvious *drawbacks*, the potentially poor statistical properties of the estimates and the need for measurements with high resolution. Both are a direct consequence of decomposition in the incremental approach. Since less *a priori* knowledge is introduced in the early stages of identification, more measurement information is required. This is a desired property of incremental identification in order to reduce the risk of fitting inappropriate model structures and to support the discovery of more suitable structures from the experimental data. Since high resolution measurement technology is steadily maturing, this drawback will ultimately vanish.

Decomposition of the identification problem and incremental refinement favour error propagation. Hence, the method unavoidably leads to biased and non-consistent estimates in contrast to the simultaneous approach which results in asymptotically unbiased and consistent estimates, if the model structure is correct. The lack of favourable statistical properties has been pointed out before for the differential method in reaction kinetics (Hosten, 1979). A thorough theoretical analysis on this issue has been carried out recently (Bardow and Marquardt, 2004b) for the simple example of identifying the reaction kinetics of a first order reaction. Theoretical analysis and computational experience shows that error propagation can indeed become a serious issue for fast reactions and large variance of the reaction flux estimates. Incremental identification should therefore be considered as a kind of bootstrap approach. Promising model structures and very good initial guesses for the parameters can be obtained in a fully transparent process at little engineering effort and low computational requirements. A final parameter estimation run with the most favourable model structure and the parameters available from the incremental procedure as initial guesses converges within few iterations and improves the statistical properties of the estimate. Further, incremental identification is advantageous in those cases where knowledge on appropriate model structures is lacking.

Mechanistic vs. data-driven modelling

The clarification and representation of kinetic mechanisms in reactive multi-phase systems has been the starting point of the research in CRC 540. However, if

heuristic nonlinear regression models are employed on levels *BF* and *BFR,* the incremental approach does not lead to a rigorous first principles model, but rather results in a certain type of hybrid (or grey-box) model (Psichogios and Ungar, 1992) which combines first principles and data-driven model constituents (Agarwal, 1997). While the development of hybrid models has largely been driven by the high effort of black-box modelling and the lack of extrapolation capabilities of black-box models (Psichogios and Ungar, 1992; Agarwal, 1997; Mogk et al., 2002; Oliveira, 2004), incremental identification aims at an extension of the mechanistic understanding of a process systems and a representation of its behavior by a first-principles model. Both research directions may ultimately merge to benefit from the respective experiences.

Research opportunities

Method development is carried out in all of the above mentioned areas in CRC 540 driven by a set of modelling problems which are summarized subsequently. Obviously, more experience should also be built up through applications on a variety of other problems.

The full methodology has been worked out already for reaction kinetics in homogeneous systems (Brendel et al., 2003), a benchmark problem of industrial relevance involving lumped-parameter models. Extensions to more complex reaction mechanisms and to multi-phase systems are currently being considered. As a first example, esterification reactions catalyzed by immobilized enzymes are studied (Heinemann, 2003). The exploration of the potential and the limitations of incremental identification applied to metabolic pathways seems to be attractive given the relevance of improving our understanding of the function and behaviour of biological systems.

A benchmark problem in the area of distributed parameter systems is the modelling of diffusion in potentially reactive multi-component systems. Liquids but also more complex systems like hydrogels are under investigation. The method has been successfully applied to binary and ternary diffusion in liquids (Bardow et al., 2003, 2004; Bardow and Marquardt, 2004a). Our future work will focus on multi-component diffusion with special emphasis on electrolytes and hydrogels.

The experimental setup chosen allows a spatially one-dimensional treatment of the diffusion problem. In case of pool boiling and heat and mass transfer to a potentially reactive falling film or droplet in a dispersion, a multi-dimensional treatment is required. Our focus has been so far only on the solution of model-free flux estimation algorithms (Lüttich et al., 2004) and their application to experimental data (Groß et al., 2004). These inverse problems are computationally involved and will require a lot more attention before fully satisfactory solutions will be available for transport problems in falling films and levitated drops.

The methodological focus of our work is the improvement, analysis and validation of the suggested work process of model-based experimental analysis.

While the presented incremental refinement strategy fixes the location of the uncertainty by construction, Kristensen et al. (2004) report a method which facilitates the discovery of model deficiencies in a stochastic modelling framework. A stochastic process is introduced in selected model equations which are subject to possible uncertainty. After the major sources of uncertainty have been identified, a regression model is fitted to correlate the unknown terms with the states. An adaptation of this approach to reveal unanticipated uncertainties in the incremental approach could be a worthwhile research objective.

Another important open issue is the clarification of the number and type of steps tackled sequentially. Besides the possible combination of either *levels B* and *BF* or *BF* and *BFR* in Fig. 3 discussed above, the calibration of primary measurements (such as spectroscopic data) could be merged with flux estimation to possibly improve estimation quality. Identifiability as well as achievable quality should be analyzed rigorously for the various alternatives. Both strategies for incremental model refinement and iterative optimal experimental design and validation (cf. Fig. 1) have to be improved.

Last but not least, the discovery of new model structures by a combination of data-driven kinetic modelling and modelling and simulation on the molecular level as part of the *MEXA* work process constitutes a particularly interesting problem.

References

Abou Khachfe, R. and Y. Jarny (2001). Determination of heat sources and heat transfer coefficient for two-dimensional heat flow – numerical and experimental study. *Int. J. Heat Mass Transfer*, **44**, 1309-1322.

Abramowich, F. and B. W. Silverman (1998). Wavelet decomposition approaches to statistical inverse problems. *Biometrika*, **85**, 115-129.

Agarwal, M. (1997). Combining neural and conventional paradigms for modelling, prediction and control. *Int. J. Systems Science*, **28**, 65-81.

Alifanov, O. M. (1994). *Inverse Heat Transfer Problems.* Springer, Berlin.

Alifanov, O., E. Artyukhin and S. Rumyantsev (1995). *Extreme Methods for Solving Ill-posed Inverse Problems with Applications to Inverse Heat Transfer Problems.* Begell House, New York.

Alsmeyer, F., H.-J. Koß and W. Marquardt (2004). Spectral hard modeling for the analysis of reactive mixtures. *Applied Spectroscopy*, in press.

Amato, U. and A Antoniadis (2001). Adaptive wavelet series estimation in separable nonparametric regression models. *Statistics and Computing,* **11**, 373-394.

Amrhein, M., B. Srinivasan, D. Bonvin and M. M. Schumacher (1999). Calibration of spectral reaction data. *Chemometrics Intelligent Lab. Systems*, **46**, 249-264.

Arora, N. and L. T. Biegler (2004). A trust region SQP algorithm for equality constrained parameter estimation with simple parameter bounds. *Computational Optimization and Applications*, **28**, 51-86.

Ascher, U. M. and E. Haber (2001). Grid refinement and scaling for distributed parameter estimation problems. *Inverse Problems,* **17**, 571-590.

Asprey, S. P. and S. Macchietto (2000). Statistical tools in optimal model building. *Comput. Chem. Engg.*, **24**, 1261-1267.

Asprey, S. P. (2003). Statistical dynamic model building,: applications of semi-infinite programming. In: S. P. Asprey, S. Macchietto (Eds.), *Dynamic Model Development.* Elsevier, Amsterdam, 85-104.

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. Oxford University Press, Oxford.

Bakshi, R. B. and G. Stephanopoulos (2001). Wave-Net: A multi-resolution hierarchical neural network with localized learning. *AIChE J.*, 39, 57-81.

Bärmann, F., and F. Biegler-König (1992). On a class of efficient learning algorithms for neural networks. *Neural Networks,* **5**, 139-144.

Bauer, I., H. G. Bock, S. Körkel and J. P. Schlöder (2000). Numerical methods for optimum experimental design in DAE systems. *J. Computational and Applied Mathematics,* **120**, 1-25.

Bastin, G. and D. Dochain (1990). *On-line Estimation and Adaptive Control of Bioreactors.* Elsevier, Amsterdam.

Bardow, A., W. Marquardt, V. Göke, H.-J. Koss and K. Lucas (2003). Model-based measurement of diffusion using Raman spectroscopy. *AIChE J.*, **49**, 323-333.

Bardow, A. and W. Marquardt (2004a). Identification of diffusive transport by means of an incremental approach. *Comput. Chem. Engg.,* **28**, 585-595.

Bardow, A. and W. Marquardt (2004b). Incremental and simultaneous identification of reaction kinetics: methods and comparison. *Chem. Eng. Sci.*, in press.

Bardow, A., V. Göke, H. J. Koß, K. Lucas and W. Marquardt (2004). Concentration-dependent diffusion coefficients from a single experiment using model-based Raman spectroscopy. *Int. Conf. Properties and Phase Equilibria for Product and Process Design, PPEPPD'04.* Submitted to *Fluid Phase Equilibria*.

Barrett, S. and J. M. Prausnitz (1975). Perturbed hard-chain theory: an equation of state for fluids containing small or large molecules. *AIChE J.*, **21**, 1123-1132.

Barrett, P. (2003). Selecting in-process particle-size analyzers. *Chem. Eng. Progress*, **99,** (8):26-32.

Bates, D. M. and D. G. Watts (1988). *Nonlinear Regression and its Applications.* Wiley, New York.

Belge, M., M. E. Kilmer and E. L. Miller (2002). Efficient determination of multiple regularization parameters in a generalized L-curve framework. *Inverse Problems,* **18**, 1161-1183.

Beck, J. V., B. Lackwell and C. R. St. Clair (1985). *Inverse Heat Conduction: Ill-posed Problems.* Wiley, New York.

Beck, J. V., B. Blackwell and A. Haji-Sheikh (1996). Comparison of some inverse heat conduction methods using experimental data. *Int. J. Heat Mass Transfer*, **39**, 3649-3657.

Binder, T., L. Blank, W. Dahmen and W. Marquardt (2002). On the regularization of dynamic data reconciliation problems. *Journal of Process Control*, 12, 557-567.

Boggs, P. T., R. H. Byrd, J. E. Rogers and R. B. Schnabel (1992). *User's Reference Guide for ODRPACK Version 2.01. Software for Weighted Orthogonal Distances Regression.* National Institute of Standards and Technology, NISTIR 92-4834.

Brendel, M. and W. Marquardt (2003). Stepwise refinement of sparse grids in data mining applications. *Proc. IFAC Symp. System Identification, SYSID 2003,* 1667-1672.

Brendel, M., A. Mhamdi, D. Bonvin and W. Marquardt (2003). An incremental approach for the identification of reaction kinetics. *Proc. IFAC Symp. Advanced Control of Chemical Processes, ADCHEM 2003,* 177-182.

Britt, H. I. and R. H. Luecke (1973). The estimation of parameters in nonlinear implicit models, *Technometrics*, **15**, 233–247.

Buchholz, M., T. Lüttich, H. Auracher and W. Marquardt (2004). Experimental investigation of local processes in pool boiling along the entire boiling curve. *Int. J. Heat and Fluid Flow,* **25,** 243–261.

Cannon, J. R. and P. DuChateau (1998). Structural identification of an unknown source term in a heat equation. *Inverse Problems*, **14**, 535-551.

Cartellier, A. and J. L. Achard (1991). Local phase detection probes in fluid/fluid two-phase flows. *Rev. Sci. Instruments*, **62**, 279-303.

Chen, S., C. F. N. Cowan and P. M. Grant (1992). Orthogonal least squares algorithm for radial basis function networks. *IEEE Trans. Neural Networks*, **2**, 302-309.

Dahmen, W. (1997). Wavelet and multiscale methods for operator equations. *Acta Numerica*, **215**, 55-228.

Daoutidis, P. and C. Kravaris (1991). Inversion and zero dynamics in nonlinear multivariable control. *AIChE J.*, **37**, 527-538.

De Valliere, P. and D. Bonvin (1989). Application of estimation techniques to batch reactors. II. Experimental studies in state and parameter estimation. *Comput. Chem. Engg.* **13**, 11-20.

De Valliere, P. and D. Bonvin (1990). Application of estimation techniques to batch reactors. III. Modeling refinements which improve the quality of state and parameter estimation. *Comput. Chem. Engng.,* **14**, 799-808.

Elicabe, G. E., E. Ozdeger and C. Georgakis (1995). On-line estimation of reaction rates in semi-continuous reactors. *Ind. Eng. Chem. Res.*, **34**, 1219-1227.

Engl, H. W., M. Hanke and A. Neubauer (1996). Regularization of Inverse Problems. Kluwer.

Farza, M., K. Busawon and H. Hammouri (1998). Simple nonlinear observers for on-line estimation of kinetic rates in bioreactors. *Automatica*, **34**, 312-318.

Fukushima, E. (1999). Nuclear magnetic resonance as a tool to study flow. *Ann. Rev. Fluid Mechanics,* **31**, 95-123.

Garcke, J., M. Griebel and M. Thess (2001). Data mining with sparse grids. *Computing,* **67**, 225-253.

Gladden, L. F. (2003a). Magnetic resonance: ongoing and future role in chemical engineering research. *AIChE. J.*, **49**, 2.

Gladden, L. F. (2003b). Recent advances in MRI studies of chemical reactors: ultrafast imaging of multiphase flows. *Topics in Catalysis*, **24** (October 2003)**,** 19-28.

Goldbrunner, M., J. Karl and D. Hein (2003). Concentration and temperature measurements by means of Raman spectroscopy in case of condensation with non-condensable gas. *Kerntechnik*, **68**, 106-113.

Groß, S., M. Soemers, A. Mhamdi, F. Al Sibai, W. Marquardt, A. Reusken and U. Renz (2004). Optimization-based identification of boundary conditions in falling film experiments using high resolution temperature measurements. *Int. J. Heat Mass Transfer,* submitted.

Haario, H. and I. Turunen (2003). Methodological aspects in the modelling of novel unit operations. In: S. P. Asprey, S. Macchietto (Eds.), *Dynamic Model Development..* Elsevier, Amsterdam, 1-20.

Hanke, M. and Scherzer, O. (2001). Inverse problems light: numerical differentiation. *Amer. Math. Monthly*, **108**, 512-521.

Hastie, T., R. Tibshirani and J. Friedman (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics. Corrected 3rd printing, 2003, Springer, New York.

Heinemann, M. (2003) *Experimental Analysis, Modeling and Dynamic Simulation of Thermodynamic and Kinetic Phenomena in Gel-stabilized Enzyme Carriers.* VDI-Verlag, Düsseldorf.

Hirschorn, R. M. (1979). Invertibility of nonlinear control systems. *SIAM J. Control and Optimization, 17,* 289-297.

Hosten, L. H. (1979). A comparative study of short cut procedures for parameter estimation in differential equations. *Comput. Chem. Engg.*, **3**, 117-126.

Huang, C.-H. and S.-P. Wang (1999). A three-dimensional inverse heat conduction problem in estimating surface heat flux by conjugate gradient method. *Int. J. Heat and Mass Transfer, 42*, 3387-3403.

Kirsch, A. (1996). *An Introduction to the Mathematical Theory of Inverse Problems.* Springer, New York.

Kittrell, J. R. (1970). Mathematical modelling of chemical reactions. *Adv. Chem. Eng.*, **8**, 97-183.

Kurtz, M. J. and M. A. Henson (1998). State and disturbance estimation for nonlinear systems affine in the unmeasured variables. *Comput. Chem. Engg.*, **22**, 1441-1459.

Kristensen, N. R., H. Madsen and S. B. Jorgensen (2004). A method for the systematic improvement of stochastic grey box models. *Comput. Chem. Engg., in press.*

Kyritsis, D. C., P. G. Felton, Y. P. Huang and F. V. Bracco (2000). Quantitative two-dimensional instantaneous Raman concentration measurements in a laminar methane jet. *Applied Optics*, **39**(36), 6771-6780.

Liu, H., C. Silva and M. E. Macedo (1998). Unified approach to the self-diffusion coefficients of dense fluids over wide ranges of temperature and pressure. Hard-sphere, square-well, Lennard-Jones and real substances. *Chem. Eng. Sci.*, **53**, 2403-2422.

Ljung, L. (1987). *System Identification. Theory for the User.* Prentice Hall, Englewood Cliffs.

Lüttich, T., A. Mhamdi and W. Marquardt (2004). Design formulation and solution of multi-dimensional inverse heat conduction problems. *Numerical Heat Transfer,* accepted for publication.

Maiwald, M., H. H. Fischer, Y. K. Kim (2004).Quantitative high-resolution on-line NMR spectroscopy in reaction and process monitoring. *J. Magn. Res.,* **166**, 135-146.

Mahoney, A. W., F. J. Doyle and D. Ramkrishna (2002). Inverse problems in population balances: growth and nucleation from dynamic data. *AIChE J.*, **48**, 981-990.

Marquardt, W. and H. Auracher (1990). An observer-based solution of inverse heat transfer problems. *Int. J. Heat Mass Transfer,* **33**, 1545-1562.

Marquardt, W. (1995). Towards a process modeling methodology. In R. Berber (Ed.) *Methods of Model-based Control.* NATO-ASI Series, Kluwer, 3-41.

Martens, H. and T. Naes (1989). Multivariate Calibration. Wiley, New York.

Marsland, S., J. Shapiro, U. Nehmzow (2002). A self-organizing network that grows when required. *Neural Networks,* **15**, 1041-1058.

McKay, B., M. Willis and G. Barton (1997). Steady-state modeling of chemical process systems using genetic programming. *Comput. Chem. Engg.*, **21**, 981-996.

Merzliak, T. and A. Pfennig (2003). Development of a model for the description of intra-diffusion in homogeneous liquid mixtures. Submitted to *Molecular Simulation.*

Mhamdi, A. and W. Marquardt (1999). An inversion approach to the estimation of reactions rates in chemical reactors. European Control Conference, ECC'99, F1004-1.

Mhamdi, A. and W. Marquardt (2003). Estimation of reaction rates by nonlinear system inversion. *Proc. IFAC Symp. Advanced Control of Chemical Processes, ADCHEM 2003,* 931-936.

Mogk, G., T. Mrziglod and A. Schuppert (2002). Application of hybrid models in chemical industry. In: J. Grievink, J. van Schijndel (Eds.) *Proc. Europ. Symp. on Computer-Aided Process Engg., ESCAPE-12.* Elsevier, 931-936.

Norgaard, M., O. Ravn, N. K. Poulsen and L. K. Hansen (2000). *Neural Networks for Modelling and Control of Dynamic Systems: A Practitioners Handbook.* Springer.

Oppenheim, A. and R. Schafer (1999). *Discrete-Time Signal Processing.* Prentice-Hall, Englewood Cliffs.

Oliveira, R. (2004). Combining first principles modelling and artificial neural networks: a general framework. *Comput. Chem. Engg.*, **28**, 755-766.

Popper, K. (1959). *The Logic of Scientific Discovery.* Hutchinson, London.

Prasad, A. K. (2000). Particle image velocimetry. *Current Science*, **79**, 51-60.

Psichogios, D. C. and L. H. Ungar (1992). A hybrid neural network – first principles approach to process modeling. *AIChE J.* **38**, 1499-1511.

Psichogios, D. C. and L. H. Ungar (1994). SVD-NET: An algorithm that automaticall selects network structure. *IEEE Trans. Neural Networks,* **5**, 513-515.

Raynaud, M. and J. Bransier (1986). A new finite-difference method for the nonlinear inverse heat-conduction problem. *Numer. Heat Transfer*, **9,** 27-42.

Reeve, D. E. and M. Spivack (1994). Determination of a source term in the linear diffusion equation. Inverse Problems, 10, 1335-1344.

Reinecke, N. and D. Mewes (1997). Multielectrode capacitance sensors for the visualization of transient two-phase flows. *Experimental Thermal and Fluid Science,* **15**, 253-266.

Reinsch, C. H. (1967). Smoothing by spline functions. Num. Math., 10, 177-183.

Schaich, D., R. Becker and R. King (2001). Qualitative modelling for automatic identication of mathematical models of chemical reaction systems. *Control Engineering Practice* , **9**, 1373–1381.

Schuler, H. and C. Schmidt (1992). Calorimimetric state estimators for chemical reactor diagnosis and control: review of methods and applications. *Chem. Eng. Sci.* **47**, 899-915.

Silverman, L. M. (1969). Inversion of multivariable linear systems. *IEEE Trans. Autom. Control,* **14**, 270-276.

Skrifvars, H., S. Leyffer and T. Westerlund (1998). Comparison of certain MINLP algorithms when applied to a model structure determination and parameter estimation problem. *Comp. Chem. Engg.* **22**, 1829-1835.

Stewart, W. E., Y. Shon and G. E. P. Box (1998). Discrimination and goodness of fit of multiresponse mechanistic models. *AIChE J.*, **44**, 1404-1412.

Taavitsainen, V. M., H. Haario and M. Laine (2003). Rapid estimation of chemical kinetics by implicit calibration – II. *J. Chemometrics*, **17,** 140-150.

Taavitsainen, V. M. and H. Haario (2001). Rapid estimation of chemical kinetics by implicit calibration – I. J. Chemometrics, **15**, 215-239.

Tholudur, A. and W. F. Ramirez (1996). Optimization of fed-batch bioreactors using neural network parameter function models. *Biotechnol. Prog.*, **12**, 302-309.

Tholudur, A. and W. F. Ramirez (1999). Neural network modelling and optimization of induced foreign protein production. *AIChE J.*, **45**, 1660-1670.

Tikhonov, A. N. and V. Y. Arsenin (1977). *Solution of Ill-Posed Problems.* V. H. Winston and Sons.

Walter, E. and L. Pronzato (1990). Qualitative and quantitative experiment design for phenomenological models – a survey. *Automatica,* **26**, 195-213.

Walter, E. and L. Pronzato (1997). *Identification of Parametric Models from Experimental Data.* Springer, Berlin.

Van Lith, P.F., B.H.L. Betlem and B. Roffel (2002). A structured modeling approach for dynamic hybrid fuzzy first-principles models. *J. Proc. Cont.,* **12**, 605-615.

Verheijen, P. J. T. (2003). Model selection. An overview of practices in chemical engineering. In: S. P. Asprey, S. Macchietto (Eds.), *Dynamic Model Development.* Elsevier, Amsterdam, 85-104.

Wahba, G. (1990). *Spline Models for Observational Data.* SIAM, Philadelphia.

Yserentant, H. (1992). Hierarchical bases. *Proc. ICIAM'91.* SIAM, Philadelphia.