

Robust Parameter Estimation in Nonlinear Dynamic Process Models

María Rodríguez-Fernández, Antonio A. Alonso and Julio R. Banga*

^aProcess Engineering Group, IIM-CSIC (Spanish Council for Scientific Research),
C/Eduardo Cabello 6, 36208 Vigo (SPAIN)

Abstract

Parameter estimation is a key issue in the mathematical modelling of nonlinear dynamic processes. Standard (gradient-based) methods for data fitting in nonlinear dynamic systems can suffer from slow and/or local convergence, among other problems. However, this is frequently ignored, potentially leading to wrong conclusions about the validity of a model regarding a certain data set. In order to surmount these difficulties, we present alternative methods based on global optimisation and identifiability analysis.

Keywords: parameter estimation, inverse problem, global optimisation, sensitivity analysis.

1. Introduction

Building sound dynamic models is a core task in modern computer-aided process engineering. Model building is usually divided in two tasks: definition of the model structure, and parameter estimation. The latter, also known as model calibration, is a key step in the development of reliable dynamic models. Given a model structure and a set of experimental data, the objective of parameter estimation is to calibrate the model (looking for parameters which can not be measured directly) so as to reproduce the experimental results in the best possible way. This calibration is performed by minimizing a cost function which measures the goodness of the fit, like e.g. maximum likelihood or least squares criterions (Walter and Pronzato, 1997; Schittkowski, 2002).

It is well known that many process models involve coupled and highly non-linear phenomena, usually described by sets of partial and ordinary differential equations. Thus, the resulting parameter estimation problem can be very challenging to solve. In particular, complex nonlinearities might cause non-convexity, i.e. the optimisation problem may contain several local minima in the area of interest. Thus, traditional gradient-based methods, like Levenberg-Marquardt or Gauss-Newton, may fail to identify the global solution of the calibration problem. Further, when these methods find a parameter set which gives a poor fit to the experimental data, the user can not be sure if the reason is due to an incorrect model, or if it is an indication of the convergence of the optimisation solver to a local solution.

* Author to whom correspondence should be addressed: julio@iim.csic.es

In this contribution, we present a new global optimisation methodology with a number of significant advantages, namely (i) high probability of convergence to the global solution, (ii) reduced computation time by means of a hybrid (stochastic-deterministic) global optimisation approach, which increases efficiency while guaranteeing robustness, and (iii) adequate handling of measurement noise (errors) and partial observations. This methodology is complemented with several analysis tools, implemented as a Matlab toolbox, which provides: (a) parameter ranking based on dynamic sensitivities; (b) testing of a priori and a posteriori identifiability of the model, including the computation of correlation matrix; (c) evaluation of the information content of the experiments via scalar functions of the Fisher information matrix (FIM), and (d) computation of confidence intervals via the FIM and the Hessian.

It should be noted that this information (e.g. scalar functions of the Fisher information matrix) can be subsequently used to design new optimal dynamic experiments (Banga et al, 2002). The usefulness and performance of this methodology is illustrated here considering two challenging case studies.

2. Problem statement

The mathematical formulation of the parameter estimation problem is that of a nonlinear programming problem (NLP) with differential-algebraic (DAEs), or partial differential-algebraic (PDAEs) constraints. The objective is to find the parameter vector $\mathbf{p}$ of a nonlinear dynamic model in order to minimize some functional J of a weighted distance measure between a vector of measured outputs $\mathbf{y}_{\text{exp}}(t)$ and the model predictions $\mathbf{y}(\mathbf{p}, t)$:

$$J = \int_0^{t_f} \left(\mathbf{y}_{\text{exp}}(t) - \mathbf{y}(\mathbf{p}, t) \right)^T \mathbf{W}(t) \left(\mathbf{y}_{\text{exp}}(t) - \mathbf{y}(\mathbf{p}, t) \right) dt \quad (1)$$

with $\mathbf{W}(t)$ a weighting (or scaling) matrix, and subject to:

- the nonlinear dynamics (described by the DAEs or PDAEs)
- bounds for $\mathbf{p}$
- other possible constraints

Due to the nonlinear and constrained nature of the system dynamics, these problems are very often multimodal (nonconvex). Thus, traditional gradient-based methods, like Levenberg-Marquardt or Gauss-Newton, may fail to identify the global solution of the calibration problem, as discussed in e.g. Schittkowski (2002). This author presents a detailed overview of local optimization methods, and provides examples of possible difficulties that can arise, such as convergence to local solutions (with heavy dependence of these methods on the location of the starting point), or very flat objective function in the neighbourhood of a solution, among others. In order to surmount these limitations, in this contribution we will present alternative methods based on global optimisation.

3. Methods

3.1 Parameter estimation: a hybrid GO method

Initial value methods (also known as single shooting) are the most common approach: an outer non-linear optimisation (NLO) problem, with an inner initial value problem

which must be solved at each function evaluation, is solved using tailored gradient-based methods, like e.g. Gauss-Newton, or Levenberg-Mardquardt (Schittkowski, 2002). As already mentioned, for many models this procedure results in convergence to local minima, or even convergence failure if the initial guess is of bad quality. Multiple shooting methods (Bock, 1983; Müller et al, 2002) also discretize the states trajectories, resulting in larger NLOs which are solved using e.g. generalized Gauss-Newton methods. Although they are less likely to get trapped in local solutions than single shooting methods, they might fail to converge if started from bad initial guesses for the parameters. In order to surmount those difficulties, global optimisation (GO) methods should be used. GO methods can be roughly classified as deterministic, stochastic and hybrid strategies. Deterministic methods can guarantee, under some conditions and for certain problems, the location of the global optimum solution. Nevertheless, no deterministic algorithm can solve general GO problems with certainty in finite time. In fact, computational effort increases very rapidly (often exponentially) with the problem size. In the case of parameter estimation in dynamic systems using deterministic GO methods, significant advances have been made recently (Esposito and Floudas, 2000; Singer et al, 2001; Papamichail and Adjiman, 2003). However, these methods have a number of requirements about the dynamics of the system (e.g. differentiability), and currently they do not seem to be applicable to problems with a relatively large number of parameters.

Stochastic methods for GO are based in probabilistic algorithms, very often with an important heuristic component, and they rely on statistical arguments to prove their convergence in a weak way. However, many stochastic methods can locate the vicinity of global solutions in modest computational times (Banga and Seider, 1996). Additionally, stochastic methods do not require transformation of the original problem, which can be treated as a black-box. A common characteristic of the majority of stochastic global optimization methods is that they present rather slow convergence rate, particularly in the final stage of the search. This can result in excessive computation times, very especially if a large solution accuracy is required. In contrast, local deterministic methods (like those gradient-based) converge very fast if initialised properly, i.e. inside the radius of attraction of the global solution.

Hybrid GO methods combine different strategies (usually, global and local) in order to reduce their weaknesses while enhancing their strengths, i.e. conjugating their desirable features: robustness of the global method plus rapid convergence of the local method when initiated in the basin of attraction of the global solution. The overall objective is to keep efficiency without losing much robustness. Although the concept of hybrid methods is a well known idea in scientific computing, choosing the particular global and local methods and designing their interactions is a non trivial exercise. In this work, and after extensive comparative testing of many state of the art global and local solvers, we have designed a two-phase stochastic-deterministic hybrid which combines the SRES stochastic method (Runarsson and Yao, 2000) with the DN2GB local method (Dennis et al, 1981). As it will be shown below, this hybrid presents a very good compromise between performance and robustness.

3.2 Identifiability analysis

The problem we have described of determining the parameters of a system from input-output data is often called the identification problem. This is just one aspect of a larger problem, the inverse problem, which includes the study of a priori and a posteriori identifiability and identification (Walter and Pronzato, 1997). The a priori identifiability problem investigates if, under the ideal conditions of noise-free observations and error-free model structure, the unknown parameters of the postulated model can be estimated from the designed model. Although necessary, a priori identifiability is obviously not sufficient to guarantee successful parameter estimation from real data, and this is when the concept of a posteriori or practical identifiability comes into play. One still assumes that the model structure is exact, however, now the data are sparse and noisy and the question is if the unknown parameters of the postulated model can be estimated from the available data. In this work, we have implemented several procedures as a Matlab toolbox which allows to check identifiability and other related indexes, providing the following information:

- Sensitivities, computed using a direct decoupled method, plus parameter ranking, based on dynamic sensitivities (Brun et al., 2001)
- A priori local identifiability, and a priori correlation matrix (Jacquez & Greif, 1985)
- Fisher information matrix (FIM), covariance and a posteriori correlation matrices, and FIM-based criterions (practical identifiability; Jacquez & Greif, 1985)
- Analysis of possible correlations among parameters
- Information content of the experiments via scalar functions of the FIM
- Confidence intervals (based on the FIM)

Table 1. Case 1: nominal and estimated parameters

	Nominal value	Solution A	Solution B
b ₁	34.2	34.199	35.548
b ₂	138	137.981	164.503
b ₃	6.74	6.739	7.980
b ₄	100	99.986	143.906
b ₅	200	199.969	393.524
b ₆	10	9.998	18.484
p ₁	8.73×10^{-4}	8.730×10^{-4}	12.691×10^{-4}
p ₂	1.87×10^{-2}	1.870×10^{-2}	1.843×10^{-2}

4. Case Studies

4.1 Air Drying of Foods and Bioproducts

Drying is one of the most important preservation operations used in the food and pharmaceutical industries, among others. We have considered the problem of the estimation of 8 parameters of an air drying model system of a thin slab of cellulose. The mathematical formulation of the dynamic model involves coupled nonlinear phenomena, i.e. mass and heat transfer, as described in Banga and Singh (1994). These PDEs were discretized using the method of lines, resulting in a large set of DAEs. Pseudo-experimental data were generated for 5 different experiments via simulation considering the parameter values published by Luyben et al. (1982) as the real

(nominal) values. Standard (local) methods, such as Levenberg-Marquardt, could not solve this problem, converging to bad local solutions.

In contrast, the hybrid method described above recovered the original parameter values (see solution A in Table 1) with modest computational effort (minutes using a PC), which is an order of magnitude better than using the stochastic method alone. Moreover, a second set of parameters different from the nominal values, but which also predict accurately the pseudo-experimental values, was also found (solution B in Table 1). This finding revealed a previously unknown non-identifiability problem for this model: the parameters cannot be identified in a unique way, i.e. it is possible to arrive to different solutions with the same fit to experimental data. This was confirmed with the FIM-based analysis, indicating that a reformulation of the model is needed.

4.2 Three-step Biochemical Pathway

In a recent contribution, Moles et al (2003) considered a large inverse problem regarding a three-step biochemical pathway, and tried to solve it using several deterministic and stochastic GO algorithms. Only a certain type of stochastic algorithms, evolution strategies, was able to successfully solve it, although at a large computational cost. The problem consists of the estimation of 36 kinetic parameters of a nonlinear biochemical dynamic model described by 8 nonlinear ODEs. Pseudo-measurements of the concentrations of the eight species involved in the different biochemical reactions described were the result of 16 different experiments (simulations) where the initial concentrations of the substrate and product were varied. We have considered three data sets, with relative normal distributed measurement errors of 0, 3 and 5 %. The hybrid method presented here converged to the global solution (even with noisy data), with speeds up of more than one order of magnitude with respect to the previous results.

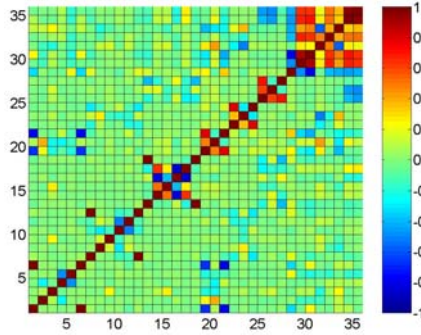


Figure 1. Correlation matrix for checking a posteriori identifiability

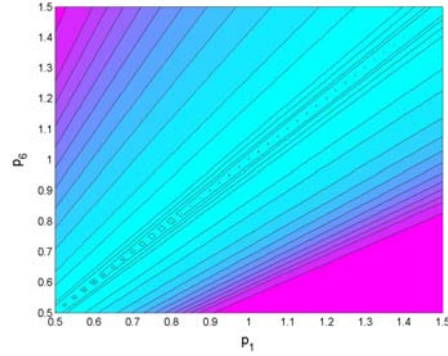


Figure 2. Contour plots of J in the parameter plane

Despite it was able to recover all the parameters, the correlation matrix (Figure 1) showed a high correlation between certain pairs of parameters, indicating a significant ill-conditioning. To further illustrate this situation, Figure 2 shows the contour plot of the error functional in the parameter plane for one of the pairs of highly correlated parameters (p_1 and p_6). Thus, in order to improve the identifiability, new experiments must be designed (Banga et al, 2002).

5. Conclusions

In this contribution, we have presented a novel stochastic-deterministic method for the robust and efficient solution of difficult parameter estimation problems. This method has been complemented with a set of analysis tools in order to evaluate identifiability and related issues. The capabilities and fine performance of this method are illustrated by solving two challenging case studies which can not be solved with standard approaches. Further, the analysis tools discovered previously unknown properties of these problems regarding their identifiability and ill-conditioning, suggesting ways to improve these situations.

References

- Banga, J. R. and Seider W.D., 1996, Global optimization of chemical processes using stochastic algorithms. In "State of the Art in Global Optimization", C.A. Floudas and P.M. Pardalos (Eds.), 563-583. Kluwer Academic Pub.
- Banga J.R. and Singh R.P., 1994, Optimization of Air Drying of Foods. *J. Food Eng.* 23, 189-221.
- Banga J.R., Versyck K.J. and Van Impe J.F., 2002, Computation of optimal identification experiments for nonlinear dynamic process models: an stochastic global optimization approach. *Industrial & Engineering Chemistry Research*, 41:2425-2430.
- Bock H., 1983, Recent advances in parameter-identification for ordinary differential equations. In: Deuffhard, P., Hairer, E. (Eds.), *Progress in Scientific Computing*, vol. 2. Birkhäuser, Boston, 95-121.
- Brun R., Reichert P. and Kunsch H.R., 2001, Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37 (4): 1015-1030
- Dennis, J.E., D.M. Gay and R.E. Welsch, 1981, An Adaptive Nonlinear Least-Squares Algorithm, *ACM Trans. Math. Software*, 7:369-383.
- Esposito, W.R. and C.A. Floudas, 2000, Global optimization for the parameter estimation of differential-algebraic systems, *Ind. Eng. Chem. Res.* 39: 1291-1310.
- Jacquez J.A. and Greif P., 1985, Numerical parameter identifiability and estimability: integrating identifiability, estimability and optimal sampling desing. *Mathematical Biosciences*, 77:201-227.
- Luyben K.C.A.M., Liou J.K. and Bruin S., 1982, Enzyme degradation during drying. *Biotechnology and Bioengineering*, 24, 533-552.
- Moles C.G., Mendes P. and Banga J.R., 2003, Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Research*, 13(11):2467-2474.
- Müller, T., Noykova, N., Gyllenberg, M. and Timmer, J., 2002. Parameter identification in dynamical models of anaerobic wastewater treatment. *Mathematical Biosci.* **177/178**, 147-160
- Papamichail, I. and Adjiman C.S., 2002, A rigorous global optimisation algorithm for problems with ordinary differential equations, *J. Global Optimization* 24:1-33.
- Runarsson, T. and Yao X., 2000, Stochastic ranking for constrained evolutionary optimization, *IEEE Transactions on Evolutionary Computation* 4:284-294.
- Schittkowski K., 2002, Numerical data fitting in dynamical systems – A practical introduction with applications and software. Kluwer Academic.
- Singer, A.B., J.K. Bok and P.I. Barton, 2001, Convex underestimators for variational and optimal control problems, *Computer Aided Chem. Eng.* 9:767-772.
- Walter, E., Pronzato L. *Identification of Parametric Models from Experimental Data*. Springer. 1997.

Acknowledgements

We thank the Spanish Ministry of Science and Technology (MCyT AGL2001-2610-C02-02) and Xunta de Galicia (PGIDIT02PXIC40211PN) for financial support.
