International Journal of Bifurcation and Chaos, Vol. 14, No. 6 (2004) 2053–2060

PARAMETER IDENTIFICATION TECHNIQUES FOR PARTIAL DIFFERENTIAL EQUATIONS

T. G. MÜLLER* and J. TIMMER
Freiburg Center for Data Analysis and Modeling (FDM), Eckerstr.1, D-79106 Freiburg, Germany
Fakultät für Physik, Albert-Ludwigs-Universität, Hermann-Herder-Str. 3, D-79104 Freiburg, Germany
**muellert@physik.uni-freiburg.de*

Received February 13, 2002; Revised May 2, 2002

Many physical systems exhibiting nonlinear spatiotemporal dynamics can be modeled by partial differential equations. Although information about the physical properties for many of these systems is available, normally not all dynamical parameters are known and, therefore, have to be estimated from experimental data. We analyze two prominent approaches to solve this problem and describe advantages and disadvantages of both methods. Specifically, we focus on the dependence of the quality of the parameter estimates with respect to noise and temporal and spatial resolution of the measurements.

Keywords: Parameter estimation; partial differential equations; spatio-temporal systems; nonlinear dynamics; complex Ginzburg–Landau equation.

1. Introduction

In recent years nonlinear pattern-forming dynamical systems have drawn much attention to describe spatiotemporal phenomena in physics, chemistry, and biology. To better understand these phenomena and to bring together experiment and theory, much has been achieved to model these systems with partial differential equations. In most of the studies, the focus was on the analysis of bifurcation points and chaos [Cross & Hohenberg, 1993; Aronson & Kramer, 2002]. While these methods are powerful tools to analyze a given theoretical model, they are unable to overcome the simulation dilemma [Timmer *et al.*, 2000]: discrepancies between simulated and measured data can be either the result of a wrong model or they can stem from inadequate dynamical parameters. Hence simulation studies for a given theoretical model cannot be used to assess the applicability of the model to a specific dataset. Therefore, to connect the theoretical work to experimental studies it is necessary to adapt the mathematical model to the experiment and to evaluate its quality. For most systems not all dynamical parameters are known with sufficient precision and thus they have to be estimated from the data.

Currently, there are two main approaches: first an approach based on regression techniques [Parlitz & Merkwirth, 2000; Bär *et al.*, 1999; Voss *et al.*, 1999] and secondly an approach based on the dynamical behavior [Bock, 1981, 1983; Baake *et al.*, 1992; Timmer *et al.*, 2000; Müller & Timmer, 2002].

The main purpose of this manuscript is to develop the principal ideas of both approaches and to highlight advantages and disadvantages. Furthermore we will compare both techniques in a simulation study which focuses on the dependence on the measurement noise and on spatial and temporal resolution of the measured data.

Throughout the manuscript we will use the following notation: Let the PDE be denoted as

$$\partial_t z(t, x) = f(p, z, \partial_x z, \partial_{xx} z, \ldots) \qquad (1)$$

with dynamical variable $z \in \mathbb{R}^Q$, parameter vector $p \in \mathbb{R}^P$, time $t \in [t_0, t_f]$ with initial and final times t_0 and t_f and spatial variable $x \in [x_{lb}, x_{rb}]$ with left and right boundaries x_{lb} and x_{rb}. Let $z^M_{ij}(p, z(t_0, x))$ denote the model trajectory, i.e. the solution of the PDE with parameter p and initial values $z(t_0, x)$, at time t_i and space x_j and let z^D_{ij} denote the experimental data at time t_i and space x_j. We assume that the true dynamical trajectory z^T_{ij} is corrupted by noise following a Gaussian distribution with zero mean and standard deviation σ_{ij}:

$$z^D_{ij} = z^T_{ij} + \eta_{ij}\,, \quad \eta_{ij} \sim N(0, \sigma^2_{ij})\,.$$

In the following we assume that all dynamical variables are observable. For more general situations including the case with unobserved components see [Müller & Timmer, 2002].

2. The Regression Approach

To apply the regression approach to the parameter estimation problem, in a first step it is necessary to compute all terms which occur in the partial differential equations, i.e. the state variable $z_{ij} = z(t_i, x_j)$, the temporal derivative $\partial_t z_{ij}$, the spatial derivative $\partial_x z_{ij}$ and derivatives of higher order. This is done by approximating these terms with help of the measured data z^D_{ij}:

$$\begin{aligned} z_{ij} &\sim \widehat{z^D_{ij}}\,, \\ \partial_t z_{ij} &\sim \widehat{\partial_t z^D_{ij}}\,, \\ \partial_x z_{ij} &\sim \widehat{\partial_x z^D_{ij}}\,, \\ \partial_{xx} z_{ij} &\sim \widehat{\partial_{xx} z^D_{ij}}, \ldots \end{aligned}$$

$\widehat{z^D_{ij}}, \ldots$ denotes an estimate of z_{ij} computed with help of the data, e.g. after noise reduction.

Using the partial differential equation Eq. (1) at every data point at time t_i and spatial variable x_j we obtain a set of equations for the dynamical parameter vector p:

$$\partial_t z_{ij} = f(p, z_{ij}, \partial_x z_{ij}, \partial_{xx} z_{iv}, \ldots)$$

which is approximated by:

$$\widehat{\partial_t o^D_{ij}} = f(p, \widehat{z^D_{ij}}, \widehat{\partial_x z^D_{ij}}, \widehat{\partial_{xx} z^D_{ij}}, \ldots)\,.$$

By minimizing

$$\Upsilon(p) = \sum_{ij} (\widehat{\partial_t z^D_{ij}} - f(p, \widehat{z^D_{ij}}, \widehat{\partial_x z^D_{ij}}, \ldots))^2$$

it is now possible to estimate the dynamical parameter vector p which follows the concept of nonlinear regression, respectively least squares regression [Cremers & Hübler, 1987; Gouesbet *et al.*, 1996; Hegger *et al.*, 1998]. The advantages of this method are clearly the straightforward implementation and the low computational cost. Especially minimization of regression problems is a well-known problem and programs are readily available. For estimating derivatives it is possible to use a method based on the symmetrized form of the finite difference operator [Press *et al.*, 1992]. In this manuscript we use a technique based on splines [Hanke & Scherzer, 2001]. It can be shown that the approximation of the first derivative of a function with cubic splines is below a certain error bound depending on the true second derivative, the sampling interval and the noise level of the data. We found that this method yields more reliable results than the method based on the finite difference operator.

The regression approach can be easily extended to a nonparametric technique where not only parameters are estimated but also the functional relationship on the right-hand sides [Härdle, 1989; Breiman & Freeman, 1985; Voss & Kurths, 1997; Voss *et al.*, 1999].

The major disadvantage of this technique is the need to estimate temporal and spatial derivatives from noisy data [Irving & Dewson, 1997]. [Bär *et al.*, 1999] even state that "noise remains a crucial problem." It is important to note that this technique is also dependent on a sufficient resolution of the experimental data. If the resolution is too coarse the estimation of derivatives is simply not possible. This property will be investigated later in comparison with the dynamical approach.

Finally, one should also note that the estimator of the parameter vector p obtained by minimizing $\Upsilon(p)$ is not a maximum likelihood estimator due to the unknown probability distribution of the computed terms which approximate the variables occuring in the PDE. Hence, the consistency of this estimator is not guaranteed.

3. The Dynamical Method

The main idea of this approach is to estimate the dynamical parameters by modeling the full trajectory of the experimental data set. Thereby it is possible to circumvent the problem that normally temporal and spatial resolutions respectively the noise level do not allow for computing derivatives reliably. Moreover, if a trajectory can be found which describes the experimental data for the whole time interval, it is possible to apply statistical inference procedures to assess the quality of the model [Cox & Hinkley, 1974]. Therefore the aim is to estimate the dynamical parameters p and the initial condition $z(t, x)$ by minimizing

$$\chi^2(p, z(t, x)) = \sum_{ij} \frac{1}{\sigma_{ij}^2} (z_{ij}^D - z_{ij}^M(p, z(t_n, x)))^2 . \tag{2}$$

In ordinary differential equations this minimization problem is well known [Richter *et al.*, 1992; Schittkowski, 1995; Edsberg & Wedin, 1995; Timmer *et al.*, 1998; Schittkowski, 1999] and a special minimization routine, the multiple shooting method (MSM), is available to circumvent the problem of local minima [Bock, 1983; Baake *et al.*, 1992; Timmer *et al.*, 1998]. However, for partial differential equations, this method does not work as reliably as in the ordinary differential equation case and it has to be advanced to the extended multiple shooting method (eMSM), [Müller & Timmer, 2002]. In the following we will shortly describe the algorithmic approach to solve the minimization problem Eq. (2). Additionally, in Fig. 1 we display the idea of the method for the Lotka–Volterra system, a special ordinary differential equation [Murray, 1993].

In a first step we choose an initial guess for parameter vector p with available *a priori* knowledge or by estimating it by the regression approach. Afterwards, we divide the data set into N subsets, $n = 1, \ldots, N$, with $N = 10$ in Fig. 1. In every subset, we apply a further subdivision of the time interval $[t_{n-1}, t_n]$ into K segments (with $K = 2$ for $N = 10$ in Fig. 1) and choose initial values of the dynamical variables $z(t_{k_n}, x)$ for each segment with help of the data by using a Savitzky–Golay filter in the time domain [Savitzky & Golay, 1964]. In this way all data points may be used to generate estimates of the dynamical variables at the beginning of each segment. Afterwards, we integrate the partial differential equation in every segment using the same dynamical parameters for all segments but the computed specific initial values for each segment. Since this leads to a discontinuous trajectory consisting of many small time intervals where we integrate the PDE we additionally introduce continuity constraints which guarantee a continuous trajectory in every subset while we optimize the parameters by minimizing the distance between the model trajectory and the data. This procedure leads to an optimal trajectory for every subset with a global parameter vector but still represents a discontinuous trajectory for the whole time interval. We finally reach a continuous optimal solution of the PDE by slowly reducing the number of subsets from $N = 10$ to 1 successively applying the above procedure. This condensing scheme is displayed in Fig. 1: after starting with $N = 10$ subsets and $K = 2$ segments and obtaining convergence to an optimal parameter vector p for all subsets, we apply the same procedure to $N = 5$ and $K = 4$. Finally, for one subset with $K = 10$, this method leads to a continuous trajectory for the whole data set.

The differences between applying the eMSM to ODE and PDE are not only the computational cost, but especially the algorithmic implementation of both optimization problems. In ODE there is a fixed number of variables which implies that the number of starting values to be estimated is inherent in the problem. In contrast, the initial condition of the PDE is a function of space which has to be parameterized in some way to allow for numerical implementation. This parameterization has to be chosen carefully to ensure that the approximation is sufficiently close to the initial state of the problem. Additionally, in ODE the number of continuity constraints is fixed. In contrast, in PDE it is necessary during the optimization procedure to allow for different numbers of continuity constraints since the actual number of grid points which is used for the numerical integration may have to be increased. This may be necessary to guarantee that the numerical solution of the PDE has a negligible *a posteriori* error [Adjerid *et al.*, 1999]. Moreover, local minima occur more often in PDE than in ODE when minimizing Eq. (2). This makes the extension of the multiple shooting method to the eMSM necessary.

It is noteworthy that, in principle, this optimization procedure slowly reduces the degrees of freedom of our theoretical model. While in the beginning the model comprises the possibility of choosing starting values in every segment for every subset to obtain a trajectory close to the data,

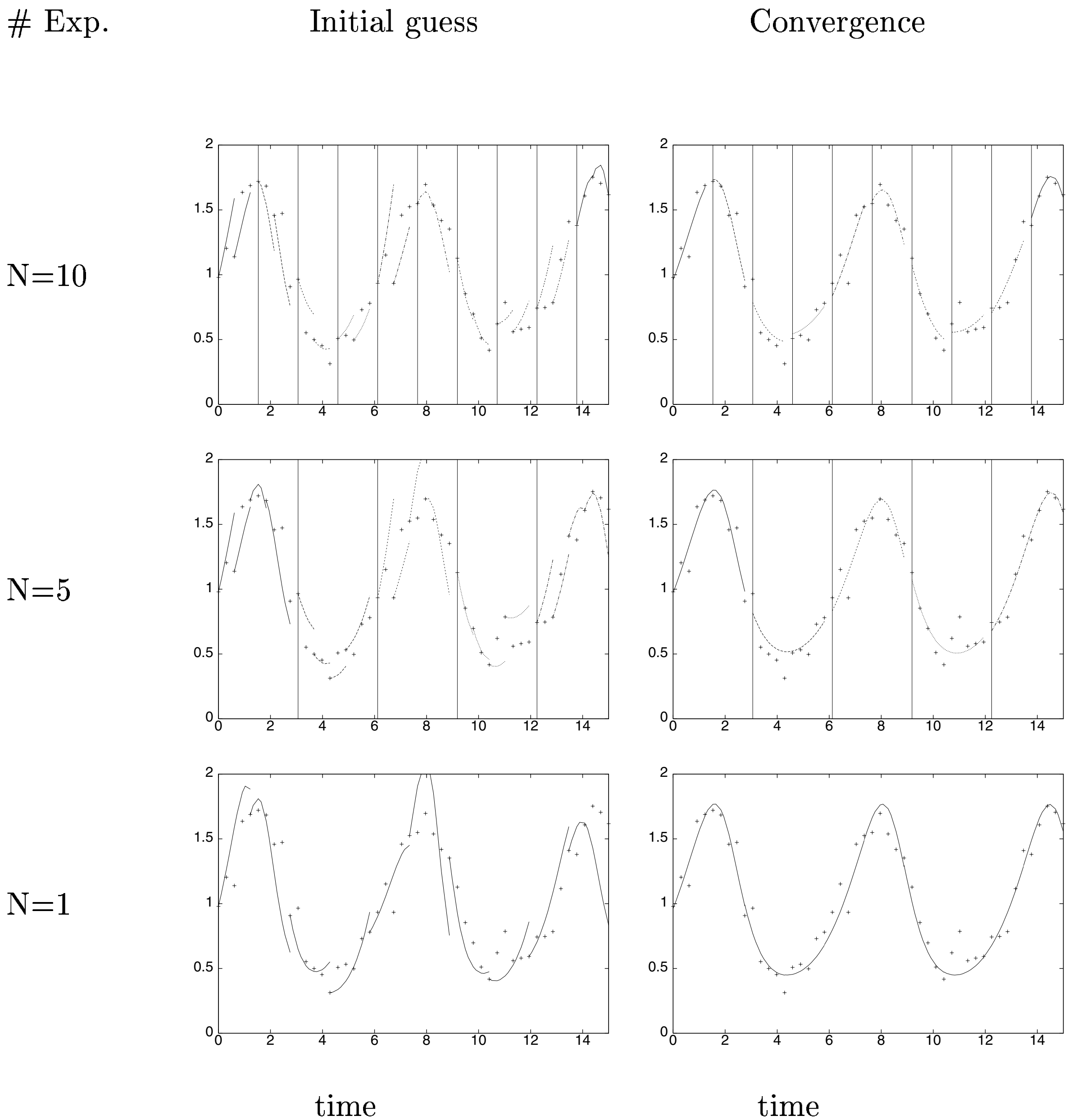

Fig. 1. Application of the extended multiple shooting method (eMSM) to an ODE example: for a decreasing number of subsets the MSM is applied in each subset leading to piecewise continuous trajectories after convergence. By reducing the number of subsets and thus the number of degrees of freedom during the fitting procedure, we obtain a continuous trajectory for the whole data set.

in the end only the initial values for the first time point remain.

In contrast to the regression approach, this method is not critically dependent on high resolution and low noise and works also in the case of unobserved components, see [Müller & Timmer, 2002]. Furthermore the estimator of the dynamical parameters is a maximum likelihood estimator and guarantees the consistency of the estimator. This also enables comparing different models using statistical inference methods.

A major disadvantage of this method is its computational cost. One has to perform many integrations of the partial differential equations since it is necessary to compute derivatives of the solution of the PDE with respect to initial values and parameters. This leads roughly to a factor of 10^3 in comparison to the regression method.

4. Comparison of Both Methods

In the following simulation study we will compare both approaches and will focus on the dependency on temporal and spatial resolutions and on noise level. As an example problem we take a typical example of the complex Ginzburg–Landau equation [Bär *et al.*, 1999; Argentina & Coullet, 1997]. The PDE of this system reads

$$\partial_t z_1 = z_2$$
$$\partial_t z_2 = (\mu - z_1^2)z_2 - z_1 - a z_1^2 - z_1^3 + \partial_{xx} z_1 + \kappa \partial_{xx} z_2$$

with parameter values as in [Bär *et al.*, 1999], $\mu = 0.2$, $a = 2.08$, $\kappa = 1.0$. A typical solution of this PDE is displayed in Fig. 2. For numerical integration of the PDE we used the method of lines [Schiesser, 1991; Ames, 1992] with sufficiently accurate spatial discretizations so that the *a posteriori* error was below 0.01 using a method of [Adjerid *et al.*, 1999]. For a more detailed discussion of this system and the simulation technique see [Müller & Timmer, 2002]. The system length is 100.0 ($x_{lb} = 0.0$, $x_{rb} = 100.0$) while the integration time is 20.0 ($t_0 = 0.0$, $t_f = 20.0$). The system comprises periodic boundary conditions which reflect a common situation for many experiments where annular containers are used [Kolodner, 1992; Andersen *et al.*, 2002]. To analyze the dependency of the results on different spatial and temporal resolutions of the observed data and on the noise level, we investigated the performance of both methods for resolutions ranging from 64×100, i.e. 64 observed spatial grid points and 100 observed time points, to 32×40. The noise levels of the added observational noise are $\sigma = \sigma_0 \sigma_{\text{signal}}$ with $\sigma_0 = 0.02$, 0.1, 0.2, 0.3 and σ_{signal} the standard deviation of the data. These settings resemble realistic situations observed in different experiments [Fullana *et al.*, 1997; Valette *et al.*, 1997; Voss *et al.*, 1999]. The initial values for the dynamical parameters in the dynamical approach were $\mu = 1.0$, $a = 1.5$, $\kappa = 5.0$, while for the regression approach we used the true parameters as initial values. Different initial parameter values were chosen for two reasons: first to show that the extended multiple shooting method is independent of initial values and secondly to demonstrate that the parameter estimates of the regression approach are biased despite true initial parameter values. For every setting we simulated 100 different data sets and obtained estimated parameters for every data set. To display the results in an easily accessible form, we operationally introduced an area in the parameter space close to the true parameter and counted only those estimated parameters within this acceptance area. An estimated parameter vector was accepted if it was within the 10% range of the true parameter, i.e. $(\sum_i |p_i^E - p_i^T|^2/|p_i^T|^2)^{1/2} < 0.10$ with p_i^T and p_i^E the ith component of the true and the estimated parameter vector. In Table 1 we list results for both approaches.

It can be seen that for all data sets the extended multiple shooting method yields much more reliable results than the regression approach. As expected, for a given spatial and temporal resolution there is a critical noise level for which the regression approach fails completely due to the problem

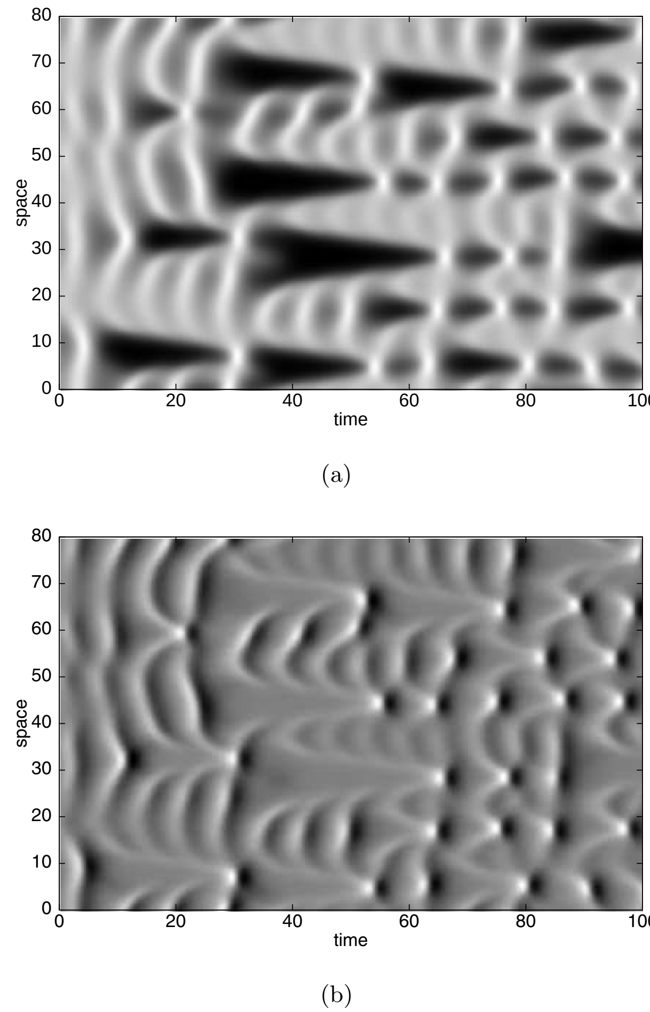

Fig. 2. Typical solution of the partial differential equation used in the simulation study. Variables z_1 and z_2 are displayed in (a) and (b).

Table 1. Percentage of accepted parameter estimates for the (a) regression approach, and (b) eMSM-approach. $L = 1, 2, 3, 4$ corresponds to a temporal resolution of 40, 60, 80 and 100 time points, $K = 1, 2$ corresponds to a spatial resolution of 32 and 64 spatial grid points.

(a)	Percentage of Accepted Estimates							
	$K = 1$				$K = 2$			
σ_0	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 1$	$L = 2$	$L = 3$	$L = 4$
0.02	13	20	22	26	36	68	71	75
0.10	7	9	10	14	13	19	21	25
0.20	0	1	1	4	0	1	2	5
0.30	0	0	0	0	0	0	1	1

(b)	Percentage of Accepted Estimates							
	$K = 1$				$K = 2$			
σ_0	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 1$	$L = 2$	$L = 3$	$L = 4$
0.02	61	62	74	83	65	68	79	91
0.10	38	35	48	65	50	54	60	82
0.20	33	36	38	53	42	52	58	67
0.30	30	32	40	45	34	37	44	58

of estimating derivatives from noisy data whereas the eMSM converges to the true parameter vector even for high noise levels. Nevertheless, it has to be noted that the computational cost of the eMSM is a factor of 1.5×10^3 higher.

5. Discussion

We compared two currently applied methods for parameter estimation in partial differential equations. As a benchmark we used an example equation of the complex Ginzburg–Landau-type which is an important class of equations for describing nonequilibrium complex systems.

The first approach redefines the problem as a regression-problem. By computing all terms which occur in the partial differential equation from the data, it is possible to obtain a set of equations for the parameter vector p which leads to a least-squares-minimization problem. While being a fast and easy to implement technique, we showed that estimating derivatives from noisy data is a general drawback of this method. The second approach, the dynamical approach, aims at obtaining a solution of the PDE for the full trajectory of the experimental data set. This also leads to a minimization problem which can be reliably solved by implementing a special minimization technique, the extended multiple shooting method, which was introduced. In a simulation study we showed the superiority of the latter approach.

It is difficult to assess the general superiority of this method in comparison to the regression approach. Nevertheless we think that in analogy to parameter estimation in ODE the more complex a spatiotemporal system is the more will the extended multiple shooting method be superior to the regression approach. Future work will have to concentrate on investigating specific real-world settings.

We would like to emphasize that these techniques require a parameterized model and do not lead to equations of motions from data, hence some *a priori* knowledge is needed. At least for complex systems from physics this holds for most cases. Very often the aim is to find a "best" of a few possible modeling approaches or even to compare two models. In this setting the eMSM is a powerful tool which may solve questions that cannot be answered by regression techniques since it allows for statistical evalution [Valette *et al.*, 1997; Voss *et al.*, 1999; Coca & Billings, 2001].

References

Adjerid, S., Flaherty, J. & Babuska, I. [1999] "A posteriori error estimation for the finite element method-of-lines solution of parabolic problems," *Math. Mod. Meth. Appl. S* **9**, 261–286.

Ames, W. F. [1992] *Numerical Methods for Partial Differential Equations* (Academic Press, San Diego).

Andersen, K., Abel, M., Krug, J., Ellegard, C., Soendergaard, L. & Udesen, J. [2002] "Pattern dynamics of vortex ripples in sand: Nonlinear modeling and experimental validation," *Phys. Rev. Lett.* **88**, 234–302.

Argentina, M. & Coullet, P. [1997] "Chaotic nucleation of metastable domains," *Phys. Rev.* **E56**, R2359–R2362.

Aronson, I. & Kramer, L. [2002] "The world of complex Ginzburg–Landau equation," *Rev. Mod. Phys.* **74**, 99–143.

Baake, E., Baake, M., Bock, H. & Briggs, K. [1992] "Fitting ordinary differential equations to chaotic data," *Phys. Rev.* **A45**, 5524–5529.

Bär, M., Hegger, R. & Kantz, H. [1999] "Fitting differential equations to space-time dynamics," *Phys. Rev.* **E59**, 337–342.

Bock, H. [1981] "Numerical treatment of inverse problems in chemical reaction kinetics," *Modelling of Chemical Reaction Systems*, eds. Ebert, K., Deuflhard, P. & Jäger, W. (Springer, Berlin), pp. 102–125.

Bock, H. G. [1983] "Recent advances in parameter identification techniques for ODE," *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, eds. Deuflhard, P. & Hairer, E., Progress in Scientific Computing, Vol. 2 (Birkhäuser, Basel), pp. 95–121.

Breiman, L. & Freeman, J. [1985] "Estimating optimal transformations for multiple regression and correlation," *J. Am. Stat. Assoc.* **80**, 580–619.

Coca, D. & Billings, S. [2001] "Identification of coupled map lattice models of complex spatio-temporal patterns," *Phys. Lett.* **A287**, 65–73.

Cox, D. & Hinkley, C. [1974] *Theoretical Statistics* (Chapman and Hall, London).

Cremers, J. & Hübler, A. [1987] "Construction of differential equations from experimental data," *Z Naturforsch* **A42**, 797–802.

Cross, P. C. & Hohenberg, M. C. [1993] "Pattern formation outside of equilibrium," *Rev. Mod. Phys.* **65**, 853–1088.

Edsberg, L. & Wedin, P. [1995] "Numerical tools for parameter estimation in ODE-systems," *Optim. Meth. Softw.* **6**, 193–217.

Fullana, J., Gal, P. L., Rossi, M. & Zaleski, S. [1997] "Identification of parameters in amplitude equations describing coupled wakes," *Physica* **D102**, 37–56.

Gouesbet, G., Le Sceller, L., Letellier, C., Brown, R., Buchler, J. & Kolláth, Z. [1996] "Reconstructing a dynamics from a scalar time series," *Ann. NY Acad. Sci.* **808**, 25–50.

Hanke, M. & Scherzer, O. [2001] "Inverse problems light: Numerical differentiation," *Am. Math. Mon.* **108**, 512–521.

Härdle, W. [1989] *Applied Nonparametric Regression* (Cambridge University Press, Cambridge).

Hegger, R., Kantz, H., Schmüser, F., Diestelhorst, M., Kapsch, R. & Beige, H. [1998] "Dynamical properties of a ferroelectric capacitor observed through nonlinear time series analysis," *Chaos* **8**, 727–736.

Irving, A. & Dewson, T. [1997] "Determining mixed linear-nonlinear coupled differential equations from multivariate discrete time series sequences," *Physica* **D102**, 15–36.

Kolodner, P. [1992] "Counterpropagating quasilinear wave packets in binary-fluid convection," *Phys. Rev. Lett.* **69**, 2519–2522.

Müller, T. & Timmer, J. [2002] "Fitting parameters in partial differential equations from partially observed noisy data," *Physica* **D171**, 1–7.

Murray, J. [1993] *Mathematical Biology* (Springer, Heidelberg).

Parlitz, U. & Merkwirth, C. [2000] "Prediction of spatiotemporal time series based on reconstructed local states," *Phys. Rev. Lett.* **84**, 1890–1893.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. [1992] *Numerical Recipes* (Cambridge University Press, Cambridge).

Richter, O., Nörtersheuser, P. & Pestemer, W. [1992] "Non-linear parameter estimation in pesticide degradation," *Sci. Total Env.* **123/124**, 435–450.

Savitzky, A. & Golay, M. [1964] "Smoothing and differentiation of data," *Anal. Chem.* **36**, 1627–1639.

Schiesser, W. [1991] *The Numerical Method of Lines* (Academic Press, San Diego).

Schittkowski, K. [1995] "Parameter estimation in differential equations," in *Recent Trends in Optimization Theory and Applications*, ed. Agarwal, R., WSSIAA (World Scientific, Singapore), pp. 353–370.

Schittkowski, K. [1999] "PDEFIT: A FORTRAN code for data fitting in partial differential equations," *Optim. Meth. Softw.* **10**, 539–582.

Timmer, J., Müller, T. & Melzer, W. [1998] "Numerical methods to determine calcium release flux from calcium transients in muscle cells," *Biophys. J.* **74**, 1694–1707.

Timmer, J., Rust, H., Horbelt, W. & Voss, H. [2000] "Parametric, nonparametric and parametric modelling of a chaotic circuit time series," *Phys. Lett.* **A274**, 123–134.

Valette, D., Jacobs, G. & Gollub, J. [1997] "Oscillations and spatiotemporal chaos of one-dimensional fluid fronts," *Phys. Rev.* **E55**, 4274–4287.

Voss, H. & Kurths, J. [1997] "Reconstruction of nonlinear time delay models from data by the use of optimal transformations," *Phys. Lett.* **A234**, 336–344.

Voss, H. U., Kolodner, P., Abel, M. & Kurths, J. [1999] "Amplitude equations from spatiotemporal binary-fluid convection data," *Phys. Rev. Lett.* **83**, 3422–3425.