PERGAMON

Automatica 37 (2001) 107–112

automatica

www.elsevier.com/locate/automatica

Brief Paper

Applying the EKF to stochastic differential equations with level effects☆

Jan Nygaard Nielsen*, Henrik Madsen

Department of Mathematical Modelling, Technical University of Denmark, Building 321, DK-2800 Lyngby, Denmark

Received 27 September 1999; revised 1 March 2000; received in final form 30 May 2000

Abstract

A transformation is introduced to effectively remove level effects, i.e. the state dependency of the diffusion function, in a restricted class of multivariate stochastic differential equations such that the general continuous–discrete-time nonlinear filtering problem may be solved using new or existing implementations of the extended kalman filter (EKF). An implementation of a quasi-maximum likelihood (QML) method for direct estimation of embedded parameters in nonlinear, multivariate stochastic differential equations using discrete-time input–output data encumbered with additive measurement noise is discussed, and its properties are compared with those provided by another software package. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Brownian motion; Continuous-time systems; Extended Kalman filters; Maximum likelihood estimation; Stochastic differential equations; Stochastic modelling

1. Introduction

Bohlin and Graebe (1995) describes a methodology for identification of nonlinear systems using a combination of statistical methods and a priori knowledge. In particular, they argue in favour of expressing the model structure in terms of a discretely, partially observed stochastic differential equation (SDE), where the measurements are encumbered with noise. A software package, *IdKit*, based on this methodology is described in Graebe (1990b) (see Bohlin & Graebe, 1995) for additional references and applications. A similar package, *CTLSM*, for linear systems, is described in Madsen (1985), and, for nonlinear systems, in Madsen and Melgaard (1991), Melgaard and Madsen (1993) and Nielsen et al. (2000a). The former is based on the Kalman Filter (KF) and the latter on the iterated extended Kalman filter (IEKF). Neither of these packages may be applied if the diffusion function depends on the state vector.

There is an increasing evidence of both theoretical and empirical nature that the level of the (process) noise depends on the state vector in a variety of applications (see Kloeden & Platen (1995)) for a survey. This will be referred to as *level effects*. In general, SDEs with level effects necessitate higher-order filters (Jazwinski, 1970; Maybeck, 1982) that offers only approximate and most likely computer intensive solutions to the filtering problem. However, for a restricted class of models, a transformation proposed here may be used to remove the level effects such that first-order filters, say, IEKF can be applied. This effectively extends the model structure for which the afore-mentioned packages are applicable.

The remainder of this paper is organized as follows: Section 2 introduces the nonlinear, multivariate and quasi-stationary stochastic state-space model and the discrete-time, multivariate measurement equation to be considered. Section 3 introduces the transformation. Section 4 provides some comments regarding currently available software packages. Section 5 describes an application, where level effects are clearly present. Finally, Section 6 concludes.

☆This paper was not presented at any IFAC meeting but is an extension of 2 conference papers presented at SYSID '97 and the 14th IFAC World Congress. This paper was recommended for publication in revised form by Associate Editor J. Schoukens under the direction of Editor Torsten Söderström.

* Corresponding author. Tel.: + 45-4525-3419; fax: + 45-4588-1397.
E-mail addresses: jnn@imm.dtu.dk (J.N. Nielsen), hm@imm.dtu.dk (H. Madsen).

PII: S0005-1098(00)00128-X

2. Nonlinear stochastic differential equations

Consider the nonlinear, multivariate, quasi-stationary stochastic differential equation (SDE) given by

$$\mathrm{d}\mathbf{X}_t = \mathbf{f}_t(\mathbf{X}_t, \mathbf{u}_t; \boldsymbol{\theta})\,\mathrm{d}t + \mathbf{g}_t(\mathbf{X}_t; \boldsymbol{\theta})\,\mathrm{d}\mathbf{W}_t, \quad t_0 \le t \le T, \tag{1}$$

where $\mathbf{X}_t \in \mathbb{R}^n$ is a stochastic state vector; $\mathbf{X}_{t_0}$ is a stochastic initial condition satisfying $E[\|\mathbf{X}_{t_0}\|^2] < \infty$; $\mathbf{u}_t \in \mathbb{R}^d$ is a vector of deterministic inputs (e.g. control signals), which is known for all t; and $\mathbf{W}_t = (W_t^1, \ldots, W_t^m)'$ is a standard Wiener process. It is assumed that the drift term $\mathbf{f}: [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^d \mapsto \mathbb{R}^n$ and the diffusion term $\mathbf{g}$: $[t_0, T] \times \mathbb{R}^n \mapsto \mathbb{R}^{n \times m}$ satisfy sufficient regularity conditions to ensure the existence of strong solutions to (1) (see Øksendal, 1995).

Remark 1. In the presence of level effects in (1), it is necessary to be precise about the interpretation of stochastic integration. In this paper the Itô interpretation is used. A discussion of the Itô and Stratonovitch interpretations may be found in, e.g. Bohlin and Graebe (1995), Kloeden and Platen (1995) and Øksendal (1995). In particular, Kloeden and Platen (1995) argue strongly that the Itô interpretation should be used for state and parameter estimation.

The real-valued discrete-time measurements $\{\mathbf{Y}_{t_k}\}$ are obtained at the sampling instants $t_0 \le t_1 < \cdots < t_k < \cdots < t_N \le T$, where N denotes the number of measurements. Let l denote $\dim(\mathbf{Y}_{t_k})$. The measurement equation is

$$\mathbf{Y}_{t_k} = \mathbf{h}_{t_k}(\mathbf{X}_{t_k}, \mathbf{u}_{t_k}; \boldsymbol{\theta}) + \mathbf{e}_{t_k}, \quad k = 1, \ldots, N, \tag{2}$$

where $\mathbf{h}$ is a nonlinear function, which is assumed to be continuously differentiable with respect to $\mathbf{X}_t$, and $\mathbf{e}_{t_k}$ is a zero mean Gaussian white noise process with covariance $\boldsymbol{\Sigma}_{t_k}$. The stochastic entities $\mathbf{X}_{t_0}$, $\mathbf{W}_t$ and $\mathbf{e}_{t_k}$ are assumed to be mutually independent for all t and t_k.

3. A multivariate transformation

In this section, a generalization of the transformation proposed by Baadsgaard et al. (1997) to a special class of multivariate SDEs is introduced. The transformation has been proposed, for univariate SDEs, by Kloeden and Platen (1995) in order to obtain closed-form solutions to some SDEs and applied by Aït-Sahalia (1999) as a means of obtaining a transition probability density function (pdf) that is closer to the normal pdf, but it also has an interesting application in nonlinear filtering theory, because it alleviates the need for higher-order filters in some applications; in particular, when the filters are utilized for parameter estimation.

Consider a bijective transformation of $\mathbf{X}_t$ given by

$$\mathbf{Z}_t = \boldsymbol{\Psi}_t(\mathbf{X}_t), \tag{3}$$

where $\boldsymbol{\Psi}_t : [t_0, T] \times \mathbb{R}^n \mapsto \mathbb{R}^n$ and $\boldsymbol{\Psi}_t$ is $C^{1,2}$, i.e. it is continuously differentiable with respect to t and twice continuously differentiable with respect to $\mathbf{X}_t$ such that, by Itô's multivariate formula, the SDE for $\mathbf{Z}_t$ is given by

$$\mathrm{d}\mathbf{Z}_t = \tilde{\mathbf{f}}_t(\mathbf{Z}_t, \mathbf{u}_t; \boldsymbol{\theta})\,\mathrm{d}t + \mathbf{G}_t(\boldsymbol{\theta})\,\mathrm{d}\mathbf{W}_t, \tag{4}$$

where the diffusion term is independent of the state $\mathbf{Z}_t$. Thus it is assumed that the dimension of the Wiener process $\mathbf{W}_t$ is preserved by transformation (3). The substitution $\mathbf{X}_t = \boldsymbol{\Psi}^{-1}(\mathbf{Z}_t)$ should be applied to the measurement equation (2).

Remark 2. Note that (4) contains the same parameters as (1) and describes a relation between the same input–output variables as the originating continuous-discrete state-space model (1)–(2).

Assumption 3. Assume that the diffusion terms are strictly nonzero, i.e.

$$g_t^{ij}(\mathbf{X}_t; \boldsymbol{\theta}) \neq 0, \quad i = 1, \ldots, n, \; j = 1, \ldots, m. \tag{5}$$

Assumption 4. Assume that for each i there exists only one g^{ij} as a function of one and only one state variable $X_t^{v(i)}$, where $v(i)$ should be different for each i, i.e.

$$g_t^{ij}(\mathbf{X}_t; \boldsymbol{\theta}) = g_t^{ij}(X_t^{v(i)}; \boldsymbol{\theta}), \quad i = 1, \ldots, n, \; j = 1, \ldots, m. \tag{6}$$

Assume further that $g_t^{ij}(X_t^{v(i)}; \boldsymbol{\theta})$ is bijective and that the function $[g_t^{ij}(x; \boldsymbol{\theta})]^{-1}$ is integrable with respect to x.

Given these assumptions, the main result is given in the following theorem:

Theorem 5. *Let* $\mathbf{X}_t$ *be a solution to* (1). *Then Assumptions* 3 *and* 4 *provide necessary and sufficient conditions for the existence of a transformation* (3) *given by*

$$\psi_t^k(X_t^{v(i)}) = \int \frac{\mathrm{d}x}{g_t^{ij}(x; \boldsymbol{\theta})}\Big|_{x = X_t^{v(i)}}, \quad k, i = 1, \ldots, n, \; j = 1, \ldots, m \tag{7}$$

such that (4) *is fulfilled.*

Proof. Applying Itô's multivariate formula to (3) yields a new Itô SDE with the kth component Z_t^k, $k = 1, \ldots, n$, satisfying

$$\mathrm{d}Z_t^k = \left(\frac{\partial \psi_t^k(\mathbf{X}_t)}{\partial t} + \sum_{i=1}^{n} \frac{\partial \psi_t^k(\mathbf{X}_t)}{\partial x_i} f_t^i(\mathbf{X}_t, \mathbf{u}_t; \boldsymbol{\theta}) + \frac{1}{2} \sum_{j=1}^{m} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \frac{\partial^2 \psi_t^k(\mathbf{X}_t)}{\partial x^{i_1} \partial x^{i_2}} g_t^{i_1 j}(\mathbf{X}_t; \boldsymbol{\theta}) g_t^{i_2 j}(\mathbf{X}_t; \boldsymbol{\theta}) \right) \mathrm{d}t + \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{\partial \psi_t^k(\mathbf{X}_t)}{\partial x_i} g_t^{ij}(\mathbf{X}_t; \boldsymbol{\theta})\,\mathrm{d}W_t^j. \tag{8}$$

To obtain a constant diffusion term (unity for reasons of parameter identifiability), it immediately follows that the following should hold:

$$\frac{\partial \psi^k(\mathbf{X}_t)}{\partial x^i} g_t^{v(i)j}(\mathbf{X}_t;\boldsymbol{\theta}) = 1 \quad \text{and} \quad \frac{\partial \psi^k(\mathbf{X}_t)}{\partial x^i} g_t^{ij}(\mathbf{X}_t;\boldsymbol{\theta}) = 0 \tag{9}$$

for $k,i = 1,\dots,n$; $j = 1,\dots,m$; $v(i) \neq i$. Assumption 4 ensures that there exists only one $X_t^{v(i)}$ for each ψ^k. Under Assumption 3, Eq. (7) follows immediately from (9).

The drift term $\tilde{\mathbf{f}}_t(\mathbf{Z}_t,\mathbf{u}_t;\boldsymbol{\theta})$ in the transformed system (4) follows immediately from (8). Thus the transformation (3) may introduce additional nonlinearities in the drift $\tilde{\mathbf{f}}$. The implications with respect to parameter identifiability must be analysed in each particular case.

The transformation does not alter the interpretation of the model parameters, and the substitution $\mathbf{X}_t = \boldsymbol{\Psi}^{-1}(\mathbf{Z}_t)$ may be used to obtain estimates of the states of the original SDE (1), but it is, in general, difficult to obtain the associated variance of the state estimates. However, using a Taylor expansion of the inverse transformation approximate variances may be obtained.

Given the fact that the Gaussian pdf is completely characterized by the first two (conditional) moments, filters are most often based on the conditional mean and conditional variance (Jazwinski, 1970; Maybeck, 1982). This implies that the transformation need not be applied if model (1) takes on a particular form such that explicit expressions for the conditional mean and conditional variance are available. For instance, this holds for the square-root process proposed by Feller (1951), i.e.

$$\mathrm{d}X_t = \alpha(\beta - X_t)\,\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}W_t, \tag{10}$$

where α, β and σ are positive, real constants. It is, of course, still an explicit assumption that the first two conditional moments provide sufficient information about the SDE (10) and the associated non-Gaussian transition pdf.

Remark 6. Explicit expressions for the conditional mean and variance are only available for very few models.

Clearly, by writing (4) in stochastic integral form, see Remark 1, the expectation of the stochastic integral is zero under some technical conditions ensuring that the expectation exists. This implies that existing software packages based on first-order filters, such as the IEKF and simplications thereof, may be applied to the transformed system.

4. Software implementations

It is outside the scope of the present paper to discuss all available software packages, but some comments will be put forth in this section. Emphasis will be placed on the two software packages, *IdKit* and *CTLSM*, mentioned in the Introduction.[1] Both packages may be applied to the transformed model proposed in Section 3. However, they are not developed with models like (10) for which explicit expressions for the conditional mean and variance are available in mind.

IdKit and *CTLSM* are both based on a prediction error decomposition (PED) that provides the residuals for which a likelihood function is specified assuming that the residuals are Gaussian. The residuals are obtained using the KF and the IEKF, respectively, in the following way: *IdKit* computes a deterministic reference trajectory by numerical integration of (8) under the assumption that the diffusion is zero and applies the KF to a linear perturbation model (linearised about the reference trajectory) to obtain also the covariances, whereas *CTLSM* solves a linearized version of (1) by means of the exponential matrix using the IEKF. The perturbation approach used in *IdKit* is only feasible provided that the level of the process noise is sufficiently "small" (cf. Graebe, 1990a). This assumption that may be too restrictive for SDEs with level effects is not made in *CTLSM* which is based entirely on a stochastic model specification and implementation of the filter as argued in Mortensen (1969). To decrease the sensitivity of outliers *CTLSM* uses a transformation of the residuals in the optimization.

Both packages allows for imposing constraints on the parameters. In *CTLSM* the mean and covariance of a Gaussian a priori pdf of the parameters may be specified. In the optimization *IdKit* uses a finite difference approximation to the derivative of the residuals w.r.t. the parameters and uses that to compute the gradient and the Hessian, and finds the optimum by Gauss–Newton iterations. *CTLSM* computes finite difference approximations to the gradient and the Hessian, and a quasi-Newton method based on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) updating formula for a secant approximation of the Hessian and soft-line search to find the optimum. For model validation *IdKit* uses the weighted square of scores, while *CTLSM* computes the portmanteau-lack-of-fit test statistic, the autocorrelation function of the residuals, crosscorrelation functions between the residuals and the input variables and a cumulative residual periodogram.

Both packages rely on a large amount of provided application-independent code and a small amount of application-dependent code. However, *IdKit* and the associated user's shell, *IKUS*, constitutes a more modern environment for system identification.

[1] Other packages include *Cypros* and $Matrix_x$.

5. Empirical work

In this section a financial application of the proposed transformation and estimation method is given using simulated data. Some comparisons are made with a truncated second-order filter (Maybeck, 1982). Finally, a limitation of the transformation method is illustrated.

5.1. Short-term interest rates

To compare the transformation approach using the EKF with a truncated second-order filter, a model of short-term interest rates is considered. A univariate SDE is considered for clarity. In financial econometrics, the following model is often considered, (see e.g. Chan et al., 1992):

$$\mathrm{d}X_t = \alpha(\theta - X_t)\,\mathrm{d}t + \sigma X_t^{\gamma}\,\mathrm{d}W_t, \tag{11}$$

where X_t is the continuous-time short-term interest rate. Many of the term structure models found in the literature may be nested within this model class by imposing appropriate parameter constraints (see Chan et al., 1992) for a survey. For $\theta \equiv 0, \alpha < 0$, Eq. (11) may be used to model biological growth in a single-species population with unlimited ressources. A biological application is considered in Wang (1994).

By inserting the diffusion term from (11) in transformation (7) and applying Itô's Lemma the following transformed process is obtained:

$$\mathrm{d}Z_t = \left[\alpha\theta\{(1-\gamma)Z_t\}^{\gamma/(\gamma-1)} - \alpha(1-\gamma)Z_t - \frac{\gamma\sigma^2}{2(1-\gamma)Z_t}\right]\mathrm{d}t + \sigma\,\mathrm{d}W_t. \tag{12}$$

Applying the inverse transformation to the measurement equation

$$Y_{t_k} = X_{t_k} + e_{t_k}$$

yields the transformed measurement equation

$$Y_{t_k} = \{(1-\gamma)Z_t\}^{1/(1-\gamma)} + e_{t_k}. \tag{13}$$

5.1.1. A Monte Carlo study

Model (11) is solved numerically using the Euler discretization scheme (Kloeden & Platen, 1995). Each sampling interval $[t_{k-1}, t_k)$ is divided into $S = 200$ small time steps of length $\Delta = 1/S$ and independent N(0,Δ) distributed random variables $\Delta W_{t_{k-1}+s\Delta}$, $s = 1, \ldots, S-1$, are simulated. A discrete time approximation to (11) is then generated by the Euler scheme, i.e.

$$\tilde{X}_{t_{k-1}+s\Delta} = \tilde{X}_{t_{k-1}+(s-1)\Delta} + \alpha(\theta - \tilde{X}_{t_{k-1}+(s-1)\Delta})\Delta + \sigma\tilde{X}^{\gamma}_{t_{k-1}+(s-1)\Delta}\Delta W_{t_{k-1}+s\Delta}. \tag{14}$$

Table 1
Results for 50 sample sequences from the short rate model (11) with large variations

Parameter	True values	Mean	*t*-value	*F*-value
Extended Kalman filter with transformation				
α	0.0250	0.0261	1.4409	0.9862
θ	10.0000	10.1700	2.5767	0.7881
σ^2	0.0100	0.0099	− 0.1314	0.9922
γ	0.7500	0.7515	0.1945	0.8150
σ_e^2	0.0500	0.0562	4.9328	1.1299
Truncated second-order filter				
α	0.0250	0.0263	1.7078	0.9467
θ	10.0000	10.1680	2.5468	1.2039
σ^2	0.0100	0.0111	1.3518	0.6872
γ	0.7500	0.7508	0.0633	0.8984
σ_e^2	0.0500	0.0482	− 1.3724	0.9624

Table 2
Results for 50 sample sequences from the short rate model (11) with small variations

Parameter	True values	Mean	*t*-value	*F*-value
Extended Kalman filter with transformation				
α	0.0250	0.0272	3.0053	0.8707
θ	10.0000	9.9778	− 0.6862	1.4721
σ^2	0.0500	0.0712	2.1411	0.8816
γ	0.5000	0.5332	0.9866	0.8210
σ_e^2	0.0500	0.0518	4.4624	1.0029
Truncated second-order filter				
α	0.0250	0.0272	2.9910	1.1234
θ	10.0000	9.9763	− 0.7212	0.6715
σ^2	0.0500	0.0296	2.3626	0.6217
γ	0.5000	0.5230	0.4348	0.7644
σ_e^2	0.0500	0.0497	− 0.5570	1.0642

Using this scheme 50 stochastic independent time series consisting each of $N = 2000$ observations are generated.

In Tables 1 and 2 the estimation results for two different parameter sets in (11) for the EKF and the truncated second-order filter are shown. Table 1 shows the results for a model with large variations, whereas the parameters in Table 2 represents a model with less variation. In each table the results are listed in three columns. In the first column the mean of the estimated values are given. In the second column the *t*-statistics given by $\sqrt{n}(\bar{x} - x_{\mathrm{sim}})/\sigma_x$ are stated, where n is the number of simulated series, x_{sim} is the true value and σ_x is the empirical standard deviation of the estimated parameters. In the third column the *F*-statistics given by $z = s_x^2/\bar{s}^2$ are stated, where $\bar{s}^2$ is the mean of the estimated variance of the parameters. The *t*-values in Table 1 show that unbiased estimates of the four parameters in (11) except for the long-term mean θ, which is overestimated for both methods. With respect to σ_e^2 the transformation approach gives a biased estimate in this case, contrary to the second-order filter. This is most likely due to the highly nonlinear measurement equation caused by the transformation.

For the estimates in Table 2, the parameter α becomes slightly biased for both methods. Similar simulation studies in Baadsgaard (1996) show that this bias frequently occur when the data do not excite the model sufficiently well. However, due to the fact the variations in the data are small, the two diffusion parameters σ^2 and γ become almost perfectly correlated, which obviously gives rise to some estimation problems. Again it is seen that the measurement noise becomes biased when the EKF is applied.

5.2. Stochastic volatility models

Another recent application of the nonlinear filtering approach is univariate stochastic volatility models,

$$\mathrm{d}X_t = \alpha X_t\,\mathrm{d}t + \sigma_t X_t\,\mathrm{d}W_t^1, \quad X_{t_0} = X_0, \tag{15}$$

where X_t denote the price of a stock at time t, $\alpha > 0$ is the rate-of-return, σ_t is the stochastic volatility and W_t^1 is a Wiener process. The famous Black–Scholes model (Black & Scholes, 1973) is obtained for $\sigma_t = \sigma$, i.e. constant volatility. Empirical studies show that a SDE should be specified for σ_t in order to model the dynamics of stock prices,

$$\mathrm{d}\psi(\sigma_t) = a(\sigma_t)\,\mathrm{d}t + b(\sigma_t)\,\mathrm{d}W_t^2, \tag{16}$$

where the functions $\psi(\sigma_t)$, $a(\cdot)$ and $b^2(\sigma_t)$ should be identified from the data, and (W_t^1, W_t^2) are correlated Wiener processes with correlation coefficient ρ. This generalization is considered in Nielsen et al. (2000b), see also the references therein. However the transformation proposed in Section 3 is not applicable if σ_t is described by the SDE (16) according to Assumption 4.

6. Conclusion

For a limited class of SDEs with level effects a transformation is proposed such that first-order filters, like the IEKF, may be used to obtain an approximate solution to the continuous–discrete filtering problem. Hence, the numerical problems that are most often associated with higher-order filters are avoided, but at the cost of a more complicated drift term and measurement equation. For the transformation to exist, some restrictions must be imposed on the diffusion term in the model specification. In addition to these restrictions, it is also recommendable to parameterize the diffusion term such that a likelihood ratio test may be carried out as a means of testing for a statistically significant level effect. The method is evaluated in a Monte Carlo study using statistical tests.

References

Aït-Sahalia, Y. (1999). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approach. Working paper. Department of Economics, Princeton University.

Baadsgaard, M. T. (1996). *Estimation in stochastic differential equations.* Master's thesis, Department of Mathematical Modelling. Lyngby, Denmark.

Baadsgaard, M., Nielsen, J. N., Spliid, H., Madsen, H., & Preisel, M. (1997). Estimation in stochastic differential equations with a state dependent diffusion term. In Y. Sawaragi & S. Sagara (Eds.), *SYSID '97 — 11th IFAC symposium on system identification*, IFAC.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*, 637–654.

Bohlin, T., & Graebe, S. F. (1995). Issues in nonlinear stochastic grey-box identification. *International Journal of Adaptive Control and Signal Processing*, *9*(6), 465–490.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, *47*, 1209–1227.

Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics*, *54*(1), 173–182.

Graebe, S. F. (1990a). IDKIT: *A software gray box identification*: *Mathematical reference.* Technical Report, TRITA-REG 90/03, Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Graebe, S. F. (1990b). *IDKIT: A software guide for gray box identification*, *user's guide*, *version RT-1.0.* Technical Report, TRITA-REG 90/04, Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory.* New York: Academic Press.

Kloeden, P. E., & Platen, E. (1995). *Numerical solutions of stochastic differential equations* (2nd ed.). Heidelberg: Springer.

Madsen, H. (1985). *Statistically determined dynamical models for climate processes.* Ph.D. Thesis, IMSOR, DTH, Lyngby.

Madsen, H., & Melgaard, H. (1991). *The mathematical and numerical methods used in CTLSM — a program for ML-estimation in stochastic*, *continuous time dynamical models*. Technical Report, Institute of Mathematical Statistics and Operations Research.

Maybeck, P. S. (1982). *Stochastic models, estimation and control.* London: Academic Press.

Melgaard, H., & Madsen, H. (1993). *CTLSM version 2.6 — a program for parameter estimation in stochastic differential equations*. Technical Report, IMSOR. No. 1.

Mortensen, R. E. (1969). Mathematical problems of modeling stochastic nonlinear dynamic systems. *Journal of Statistical Physics*, *1*, 271–296.

Nielsen, J. N., Madsen, H., & Melgaard, H. (2000a). *Estimating parameters in discretely*, *partially observed stochastic differential equations*. Technical Report 2000-7, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Nielsen, J. N., Vestergaard, M., & Madsen, H. (2000b). Estimation in continuous-time stochastic volatility models using nonlinear filters. *International Journal of Theoretical and Applied Finance*, *3*(2), 279–308.

Øksendal, B. (1995). *Stochastic differential equations* (4th ed.). Heidelberg: Springer.

Wang, C. (1994). *Stochastic differential equations and a biological system.* Ph.D. Thesis, Institute of Mathematical Modelling, The Technical University of Denmark. Lyngby, Denmark.

Henrik Madsen received the M.Sc. in Engineering in 1982, and then Ph.D. in Statistics in 1986, both at the Technical University of Denmark. He was appointed Assistant Professor in Statistics in 1986, Associate Professor in 1989, and Professor in Statistics with a special focus on Stochastic Dynamic Systems in 1999. He has been external lecturer at a number of universities. He is involved in a large number of cooperative projects with other universities, research organizations and industrial partners. His main research interest is related to analysis and modelling of stochastic dynamic systems. This includes signal processing, time series analysis, identification, estimation, grey-box modelling, prediction, optimization and control. The applications are mostly related to Environmetrics, Energy systems, Informatics and Finance. He has authored or co-authored approximately 190 papers and technical reports, and about 10 educational texts. He is in charge of three courses at the Technical University of Denmark.

Jan Nygaard Nielsen was born in Ringkøbing, Denmark, in 1969. He received his M.Sc. in statistics (financial engineering) from the Technical University of Denmark in 1996. He is scheduled to defend his Ph.D. Thesis in August 2000 at the Department of Mathematical Modelling there. His research interests are in time series analysis, multivariate analysis, stochastic filtering theory and financial engineering.