

Bayesian Optimal Filtering in State Space Models: Empirical Example from Econometric Analysis of Economic Time Series Data

Abstract

This paper presents a methodology for optimal data cleaning and filtering of economic time series data using the Kalman filter. The Kalman filter is a recursive algorithm and as such is ideal for usage on the economic data. In this paper, we use the Kalman filter to estimate the parameters of a specified state space model which describes the economic phenomena under study. We find that the kalman filter is both predictive and adaptive, as it looks forward with an estimate of the covariance and mean of the time series one step into the future and unlike a neural network, it does not require stationary data, hence usage of the Kalman filter is proposed to predict the expected values of the measuring process in the near future and to detect the anomalies in the trend of economic time series data.

Key Words: Bayesian inference, Kalman Filter,Smoothing,High Frequency Data, State Space Model.

1 Introduction

State space models originated in engineering in the early sixties (Kalman; 1960, Kalman and Bucy; 1961,1963). Their use have since been extended to a number of applied fields like econometrics, time series analysis,genetics,spatio-temporal analysis,hydrology,climatology,speech recognition, signal processing, population dynamics etc. A state space model allow the researcher to model an observed(multiple) time series $y_t, t = 1, 2, \dots, T$ which are driven by a stochastic process.Many econometric models(time series models, in particular) can be written in state space form(see West and Harrison,1997; Petris et al,2009).The obvious flexibility of the state space approach has contributed to its immense popularity in recent years.

In this paper, we describe a selection of basic filtering procedures for estimating state space models and apply them to some economic time series data of the Nigerian economy. The rest of this paper is organized as follows: In section 2, we present a brief review of literature on application of state space models to time series and econometric analysis.Section 3 is on the model and Kalman filter method used in this paper,while section 4 borders on an application to the Nigerian economy, and section 5 concludes the paper.

2 Application of State Space Models in Time Series Econometrics

In recent years, there has been an increasing interest in the application of state- space model and its variants in econometrics and time series analysis, partly because of its flexibility and widely because of the development of modern Monte Carlo methods. Economic theory and practice is often concerned with latent or unobserved dynamic processes,hence the state space form of a dynamic system with unobserved components is a very powerful and flexible instrument because it builds on the dependence structure of a Markov chain to define more complex models for the dynamic observations (Petris et al,2009).State space models first appeared in time series literature in the seventies in the works of Goldfied and Quandt(1971;1973),Harrison and Stevens(1976) etc. They became established in the eighties and nineties,especially when various authors ex-

tended Hamilton's model to allow for time varying parameters and applied such models for investigating U.S business cycle See for instance, articles like Hutchinson(1984), Aoki(1987), Kitagawa(1987;1988), Hamilton(1988;1989), Harvey(1989), Goodwin(1993), Diebold et al(1994), Engel(1994) and Filardo(1994).

Other extensive applications of state space models in time series and econometrics can be seen in the works of Kim and Yoo(1995), Schaller and Van Norden(1997), Garcia and Perron(1996), Tutz(1994;1997), West and Harrison(1997), Kim and Nelson(1998), Huang et al(1998), Huang(1999), Soloviev et al(2011) to mention but a few.

A state space model is in principle any model that includes an observation process Y_t and a state process θ_t . It consists of a state equation and an observation equation which can be used for modeling univariate and multivariate time series in presence of non-stationarity, structural changes, irregular patterns. The linear state space model encompasses two models which are widely used in time series analysis and economics. First, the observation equation resembles a linear regression model with the distinction that the coefficients are stochastic dynamic processes rather than fixed parameters. Second, the state vector follows a first order vector autoregressive model, which in turn is a generalisation of the scalar Autoregressive Moving Average (ARMA) models. All univariate ARMA models can be written as a linear state space model (Harvey, 1989). In many real life applications, the driving force behind the evolution of economic variables are not always observable. Usually, it is assumed that there is an unobservable Markov chain θ_t called the 'state' and that Y_t is an imprecise measurement of θ_t .

In econometric applications, we think of the state θ_t as an auxiliary time series that is assumed to follow a Markov process which facilitates the task of specifying the probability distribution of the observable time series Y_t . State-space models consider a time series as the output of a dynamic system perturbed by random disturbances. They have an elegant and powerful probabilistic structure, offering a flexible framework for a very wide range of applications. Computations can be implemented by recursive algorithms. While the state equation formulates the dynamics of the state variables, the observation equation relates the observed variables to the unobserved state vector. The state vector can contain trend,

seasonal, cycle and regression components plus an error term. Models that relate the observations over time to different components, which are usually modeled as individual random walks, are referred to as structural time series models. The stochastic behavior of the state variable, its relationship to the data and the covariance structure of the errors depend on parameters that are also generally unknown.

The state variable and the parameters have to be estimated from the data, hence the maximum likelihood estimates of the parameters can be obtained by applying the Kalman filter. In the Box and Jenkins methodology, observations are first made stationary. This is accomplished by using various data transformations like taking logarithms and time differences before fitting ARMA models to the transformed data. Wide availability of computer software and a well developed mathematical theory have contributed to the enduring popularity of this approach. The state space methodology offers a number of important advantages over the popular class of ARMA models in so many ways. First is the fact that the state space method does not require the data to be stationary, which eliminates a problem which the Box-Jenkins theory offers minimum guidance. Another fact is that state space methods can accommodate a multivariate framework and they are able to deal with irregular patterns, interventions, sudden jumps and regime shifts in data. The problems of estimation and forecasting are solved by recursively computing the conditional distribution of the quantities of interest, given the available information. In this sense, they are quite naturally treated from a Bayesian approach.

3 Model Specification and Parameter Estimation

3.1 Linear- Gaussian State Space Model

Consider a linear-Gaussian state space model for an $m \times 1$ -dimensional time series y_t consisting of a measurement equation relating the observed time series to a $p \times 1$ -dimensional unobserved state vector θ_t and a Markovian transition equation that describes the evolution of the state vector over time. The model takes the following general form.

$$y_t = F_t \theta_t + v_t \quad v_t \sim N(0, V_t) \quad (1)$$

$$\theta_t = G_t \theta_{t-1} + w_t \quad w_t \sim N_p(0, W_t) \quad (2)$$

$$\theta_0 \sim N_p(m_0, C_0) \quad (3)$$

where y_t is a vector of dimension $m \times 1$

Equation (1) is known as the observation equation while equation (2) is a first order Markov process called the evolution equation. G_t and F_t are known matrices of order $p \times p$ and $m \times p$ respectively that determine how the observation and state equations evolve in time (see West and Harrison, 1997).

The matrices F_t , V_t , G_t and W_t are known as the system matrices and contain non-random elements. If they do not depend deterministically on t , the state space system is time invariant, otherwise they are time varying. The initial state distribution is assumed to be Normally distributed with parameters m_0 and C_0 as shown in (3) where $E(v_t \theta'_0) = 0, E(w_t \theta'_0) = 0$ for $t = 1, \dots, T$. If some or all the elements of θ_t are assumed to be covariance stationary, the corresponding elements of m_0 and C_0 can be solved analytically from the elements of the system matrices. However, for non-stationary elements of θ , it is customary to set the corresponding elements of C_0 to a very large positive number.

3.2 Linear Regression With Time-Varying Parameters

A dynamic linear regression model with time varying slope and intercept parameters has the form

$$y_t = \alpha_t + \beta_t X_t + v_t \quad v_t \sim N(0, \sigma_v^2) \quad (4)$$

$$\alpha_t = \alpha_{t-1} + w_{\alpha,t} \quad w_{\alpha,t} \sim N_p(0, \sigma_\alpha^2) \quad (5)$$

$$\beta_t = \beta_{t-1} + w_{\beta,t}, \quad w_{\beta,t} \sim N_p(0, \sigma_\beta^2) \quad (6)$$

3.3 Optimal State Estimation and Prediction Using Kalman Filter

In a state space model, the unobserved state vector θ is the signal and the measurement error w_t is the noise. The objective of filtering is to track the

dynamic evolution of unobservable state variables using (typically) noisy measurements of observables(DeJong et al,2008). This requires the computation of integrals over the unobserved states, which in general must be approximated numerically. Given observed data $y_1, \dots, y_2$, the goals of state space estimation is twofold:

- (1)Optimal signal extraction(estimation) and
- (2) Optimal h-step ahead prediction of states and data.

Basically, there are two types of state estimation, namely (a)Filtering and (b)Smoothing. Filtering implies getting optimal estimates of θ_t given information available at time t, $D_t = (y_1, \dots, y_t)$, so that $E(\theta_t|D_t)$ is the filtered estimate of θ .On the other hand, smoothing implies getting optimal estimates of θ_t given information available at time T, $D_t = (y_1, \dots, y_T)$, so that $E(\theta_t|D_T)$ is the smoothed estimate of θ .

3.4 The Kalman Filter

A filter is a device which separates entities into their constituting components.The process of finding the best estimate from noisy data is known as filtering out the noise.It does not just clean up the data measurements but also projects these measurements into the state estimate.The Kalman Filter was developed in 1960 by Rudolf E. Kalman (Kalman,1960) as an algorithm used for aerospace guidance applications and to solve state space models in the linear case.A Kalman filter is an optimal estimator which infers parameters of interest from indirect, sometimes inaccurate and uncertain observations.It is an optimal process because if all noise is gaussian, the Kalman filter minimizes the mean square error of the estimated parameters. Named after Kalman (1960, 1963) the Kalman filter is a recursive algorithm that computes estimates for the unobserved components at time t , based on the available information at the same date. The Kalman filter calculates the mean and variance of the unobserved state θ_t , given the observations. It provides an exact and complete solution to the Bayesian filtering problem which is very economical in computer calculation and storage requirements.It is a recursive algorithm i.e the current best estimate is updated whenever a new observation is obtained.It is an optimal process because if all noise is Gaussian, the Kalman filter minimizes the mean square error of the estimated parameters. The variance of the measurements equation needs to be known before implementing a Kalman filter.Given the initial state and covariance, we have

sufficient information to find the optimal state estimate using the Kalman filter equations.

The Kalman filter is a set of recursion equations for determining the optimal estimates of the state vector θ_t , given information available at time t , D_t . The filter consists of two sets of equations:

(i) Prediction Equations (ii) Updating Equations

The filter prediction and update steps require a few basic matrix calculations of which only the conditional means and variances of the filtering and prediction density need to be stored in each step of the iteration. To describe the filter, let

$$m_t = E(\theta_t | D_t)$$

be the optimal estimator of θ_t based on D_t and let

$$C_t = E((\theta_t - m_t)(\theta_t - m_t)^T | D_t)$$

be the mean square error matrix of m_t .

Prediction Equations

Let $\theta_{t-1} | y_{1:t-1} \sim N(m_{t-1}, C_{t-1})$, where $y_{1:t-1}$ denotes all observations up to time $t-1$. Then the one-step-ahead predictive density $\theta_t | y_{1:t-1}$ is Gaussian with parameters:

$$\begin{aligned} E(\theta_t | y_{1:t-1}) &= m_{t-1} \equiv A_t \\ \text{Var}(\theta_t | y_{1:t-1}) &= C_{t-1} + W_t \equiv R_t \end{aligned}$$

The one-step-ahead predictive density of $Y_t | y_{1:t-1}$ is Gaussian with parameters:

$$\begin{aligned} f_t &= E(Y_t | y_{1:t-1}) = X_t A_t \\ Q_t &= \text{Var}(Y_t | y_{1:t-1}) = X_t R_t X_t' + V \end{aligned}$$

The filtering density of θ_t given $y_{1:t}$ is Gaussian with parameters:

$$\begin{aligned} m_t &= E(\theta_t|y_{1:t}) = A_t + R_t X_t' Q_t^{-1} e_t \\ C_t &= Var(\theta_t|y_{1:t}) = R_t - R_t X_t' Q_t^{-1} X_t R_t \end{aligned}$$

where $e_t = Y_t - f_t$ is the forecast error.

Given m_{t-1} and C_{t-1} at time $t - 1$, the optimal predictor of θ_t and its associated MSE matrix are

$$\begin{aligned} m_{t/t-1} &= E(\theta_t|D_{t-1}) = G_t m_{t-1} \\ C_{t/t-1} &= E((\theta_t - m_{t-1})(\theta_t - m_{t-1})'|D_{t-1}) \\ &= G_t C_{t-1} G_t' + W_t \end{aligned}$$

.

Updating Equations

When new observations y_t become available, the optimal predictor $m_{t|t-1}$ and its MSE matrix are updated using

$$\begin{aligned} m_t &= m_{t|t-1} + C_{t|t-1} F_t' Q_t^{-1} (y_t - F_t m_{t|t-1}) \\ &= m_{t|t-1} + C_{t|t-1} F_t' Q_t^{-1} v_t \\ C_t &= C_{t|t-1} - C_{t|t-1} F_t' Q_t^{-1} F_t C_{t|t-1} \\ &= C_{t|t-1} - K_t v_t \end{aligned}$$

where

$$Q_t = F_t C_{t|t-1} F_t' + V_t$$

and

$$K_t = C_{t|t-1} F_t' Q_t^{-1}$$

is the Kalman gain matrix which gives the weight on new information $e_t = y_t - F_t m_{t|t-1}$ in the updating equation for m_t .

Our model estimation is done by the method of Kalman filter by computing the predictive and filtering distributions of θ_t inductively starting from $\theta_0 \sim N(m_0, C_0)$.

Kalman Smoother

Once all data D_T are observed, the optimal estimates $E(\theta|D_T)$ can be computed using the backwards kalman smoothing recursions

$$E(\theta_t|D_T) = m_{t|T} = m_t + C * _t (m_{t+1}|T - G_{t+1}m_t)$$

and

$$\begin{aligned} E(\theta_t - m|T)(\theta_t - m_t|T)'|D_T) &= C_{t|T} \\ &= C_t + C * _t (C_{t+1|T} - C_{t+1|t})C * _t' \\ C * _t &= C_t G_{t+1}' C_{t+1|t}^{-1} \end{aligned}$$

The algorithm starts by setting $m_T|T = m_T$ and $C_T|T = C_T$ and then proceeds backwards for $t = T-1, \dots, 1$. Proofs of this algorithm and detailed procedure can be found in Petris et al (2009).

4 Empirical Example

4.1 Dynamic Linear Modeling of High Frequency Economic Data

Economic indicators are usually complex high frequency data(Fuquene et al,2013).

Dynamic linear models have recently been used for modeling economic indicators(see for example Petris et al(2009), West and Harrison (1997), Harvey (1989), Durbin and Koopman (2001), and the recent overviews by Ku nsch (2001) and Migon et al. (2005)).State space models consider a time series as the output of a dynamic system perturbed by random disturbances,hence they represent the role of history differently in a finite-dimensional state.In this section, we implement and apply our proposed dynamic linear model, with diffuse priors for the distribution of the state and observational precision errors of economic indicators of Nigeria.We fit a linear-Gaussian univariate dynamic linear model presented in section (3.1) to eight economic indicator of the Nigerian economy.In the proposed model, we set $C_0 = 1000$. We obtained convergence of all parameters using 1000 iterations after a burn-in period of 200. We consider two large sized examples: a simulation study highlighting the properties of our approach, and a real-world example concerning high frequency economic data.

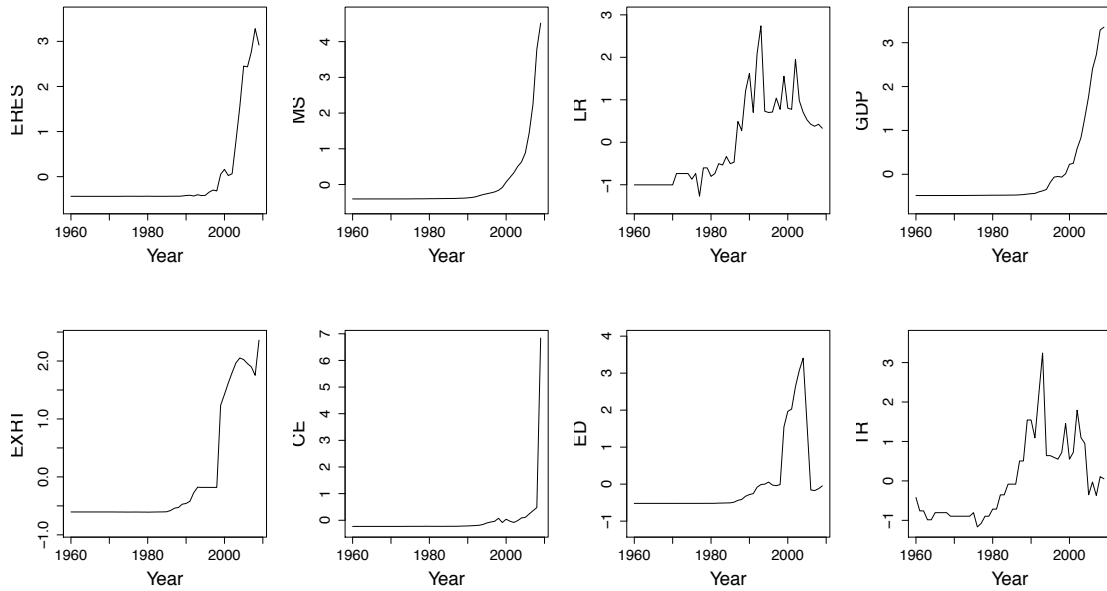


Figure 1: Annual time-series data on external reserve (ERES), money supply (MS), lending rate (LR), gross domestic product (GDP), exchange rate (EXRT), capital expenditure (CE), external debt (ED) and treasury bill rate (TR). These data were standardized prior to analysis.

In this linear model, y_t is the value of the time series we are trying to capture and model with respect to θ_t . More specifically, it is an estimate of the covariance, or co-movement between the unobserved state and the time series value. Basically, we start out estimating our guess of θ , the average and covariance of the series based upon measurements of the underlying series, which in this case are simply the normal parameters $N(\text{mean}, \text{std})$ used to generate the random walk. We compute the filtering and predictive densities of the eight economic time series using the recursive Kalman filter algorithm, starting from $\theta_0|D_0 \sim N(m_0, C_0)$, then computing $p(\theta_1|D_1)$ and proceeding recursively. It can be seen from figure 2 that our linear model performs well on the economic data under study. We can see that the model is capable of predicting the future variance of the signals with respect to the series. Note the conspicuous traces in the forecasts of exchange rate, external debt, treasury bill rate and lend-

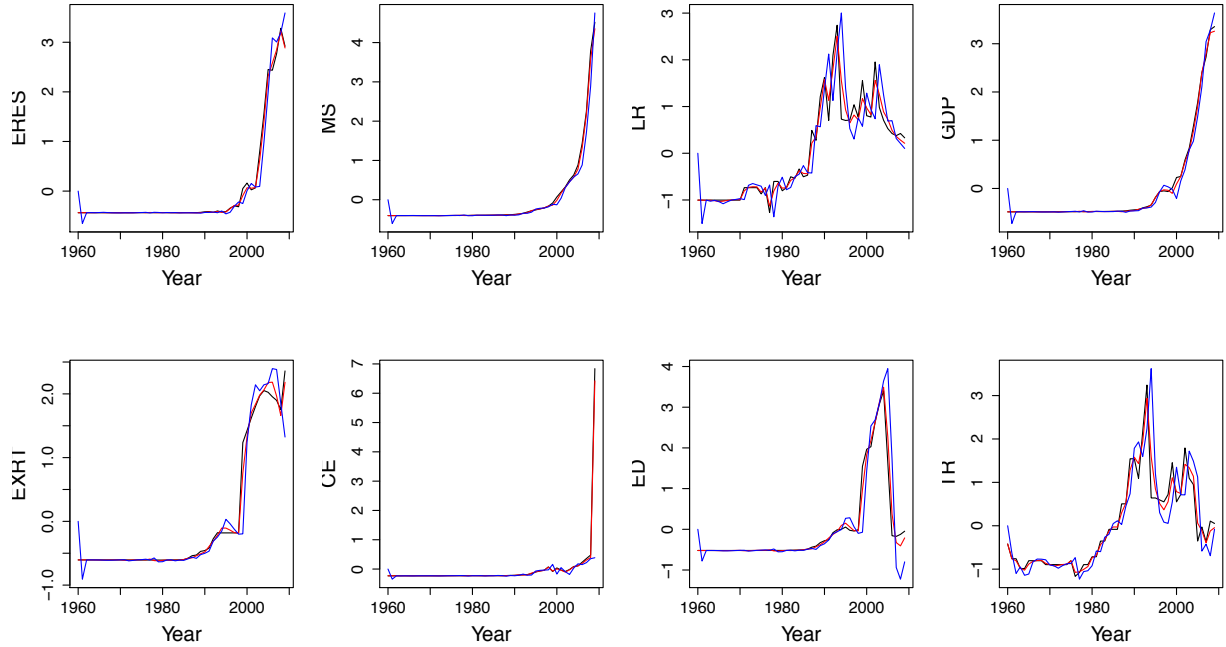


Figure 2: Filtered states, & forecasts. Black lines are the data, red the filtered states, & blue the forecasts

ing rate. The model does not seem to capture or describe any features in less volatile series like Gross Domestic Product(GDP),money supply, and capital expenditure.However, the one-step ahead forecasts of external reserves and external debt suggests that the external reserve increases in the following year while external debt reduces further.

5 Concluding Remarks

In this paper, we proposed a methodology for data cleaning of the economic data that requires no complex data processing. The Kalman filter method has been successfully tested on an macroeconomic dataset.It relies solely on optimal filtering and smoothing of the data measurements. As future work, we propose to use the Kalman filtering results (its prediction and corresponding variance) for detection of outliers in higher frequency

economic data. From our experience, we find that the Kalman filter is not a complex algorithm, but it can be difficult to adjust all the required initial conditions and parameters, which demand either an expert user or a statistical method to adjust them. With optimal tuning of the parameters, very good results can be achieved. Some advantages to the Kalman filter are that it is predictive and adaptive, as it looks forward with an estimate of the covariance and mean of the time series one step into the future and unlike a Neural Network, it does not require stationary data.

References

- Anderson, B. and Moore, J. (1979). Optimal Filtering, Prentice-Hall, Englewood Cliffs.
- Barra, J. and Herbach, L. H. (1981). Mathematical Basis of Statistics, Academic Press.
- Bawens, L., Lubrano, M. and Richard, J.-F. (1999). Bayesian inference in dynamic econometric models, Oxford University Press, N.Y.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. Published posthumously in Phil. Trans. Roy. Stat. Soc. London, 53, 370 418 and 54, 296-325. Reprinted in Biometrika 45 (1958), 293 315, with a biographical note by G.A. Barnard. Reproduced in Press (1989), 185 217.
- Berger, J. (1985). Statistical Decision Theory and Bayesian Analysis, Springer, Berlin.
- Bernardo, J. and Smith, A. (1994). Bayesian Theory, Wiley, Chichester.
- Box, G., Jenkins, G. and Reinsel, G. (1994). Time Series Analysis, Forecasting and Control, 3rd edn, Holden Day, San Francisco.
- Caines, P. (1988). Linear Stochastic Systems, Wiley, New York.
- Campbell, J., Lo, A. and MacKinlay, A. (1996). The econometrics of financial markets, Princeton University Press.
- Carmona, R. A. (2004). Statistical analysis of financial data in Splus, Springer
- Carter, C. and Kohn, R. (1994). On Gibbs sampling for state space models, Biometrika 81: 541 – 553.
- Campbell, J., Lo, A. and MacKinlay, A. (1996). The econometrics of financial markets, Princeton University Press.
- Carmona, R. A. (2004). Statistical analysis of financial data in S-plus, Springer Verlag, New York.
- Carter, C. and Kohn, R. (1994). On Gibbs

- sampling for state space models, *Biometrika* 81: 541 553.
- Chatfield, C. (2004). *The analysis of time series*, sixth edn, CRC - Chapman & Hall, London.
- D'Agostino, R. and Stephens, M. (eds) (1986). *Goodness-of-fit techniques*, M. Dekker, New York.
- Dalal, S. and Hall, W. (1983). Approximating priors by mixtures of conjugate priors, *J. Roy. Statist. Soc. Ser. B* 45(2): 278 286.
- DeJong, D. N., Dharmarajan, H., Liesenfeld, R., and Richard, J. F. (2008). Efficient filtering in state-space representations. Available at SSRN 1337092.
- Diebold, F. and Li, C. (2006). Forecasting the term structure of government bond yields, *Journal of Econometrics* 130: 337 364.
- Diebold, F., Rudebusch, G. and Aruoba, S. (2006). The macroeconomy and the yield curve: a dynamic latent factor approach, *Journal of Econometrics* 131: 309 338.
- Doucet, A., De Freitas, N. and Gordon, N. (eds) (2001). *Sequential Monte Carlo methods in practice*, Springer.
- Durbin, J. and Koopman, S. (2001). *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford.
- Engle, R. and Granger, C. (1987). Co-integration and error correction: representation, estimation, and testing, *Econometrica* 55: 251 276.
- Fru wirth-Schnatter, S. (1994). Data augmentation and dynamic linear models, *Journal of Time Series Analysis* 15: 183 202.
- Gilks, W. and Berzuini, C. (2001). Following a moving target Monte Carlo inference for dynamic Bayesian models, *J. Royal Statist. Soc. B* 63 : 127146.
- Gourieroux, C. and Monfort, A. (1997). *Time series and dynamic models*, Cambridge University Press, Cambridge.
- Granger, C. (1981). Some properties of time series data and their use in econometric model specification, *Journal of Econometrics* 16: 150 161.
- Hannan, E. and Deistler, M. (1988). *The Statistical Theory of Linear Systems*, Wiley, New York.
- Harrison, P. and Stevens, C. (1976). Bayesian forecasting (with discussion), *J. Royal Statist. Soc. B* 38: 205 247.
- Harvey, A. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.
- Jacquier, E., Polson, N. and Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion), *Journal of Business and Economic Statistics* 12: 371 417.

- Kalman, R. (1960). A new approach to linear filtering and prediction problems, Trans. of the AMSE - Journal of Basic Engineering (Series D) 82: 35 45.
- Kalman, R. and Bucy, R. (1963). New results in linear filtering and prediction theory, Trans. of the AMSE Journal of Basic Engineering (Series D) 83.
- Kolmogorov, A. (1941). Interpolation and extrapolation of stationary random sequences, Bull. Moscow University, Ser. Math. 5.
- Lauritzen, S. (1981). Time series analysis in 1880: A discussion of contributions made by T.N. Thiele, International Statist. Review 49: 319 331.
- Lauritzen, S. (1996). Graphical models, Oxford University Press, Oxford.
- Lindley, D. and Smith, A. (1972). Bayes estimates for the linear model, Journal of the Royal Statistical Society. Series B (Methodological) 34: 1 41.
- Liu, J. (2001). Monte Carlo strategies in scientific computing, Springer.
- Liu, J. and West, M. (2001). Combined parameter and state estimation in simulation -based filtering, in A. Doucet, N. De Freitas and N. Gordon (eds). Sequential Monte Carlo methods in practice, Springer.
- Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models, Biometrika 65: 297 303.
- Lu tkepohl, H. (2005). New introduction to multiple time series analysis, Springer
- Prakasa Rao, B. (1999). Statistical Inference for Diffusion Type Processes, Oxford University Press, N.Y.
- Rabiner, L. and Juang, B. (1993). Fundamentals of Speech Recognition, Prentice - Hall, Englewood Cliffs, New Jersey.
- Reinsel, G. (1997). Elements of Multivariate Time Series Analysis, 2nd edn, Springer -Verlag, New York.
- Robert, C. (2001). The Bayesian choice, 2nd edn, Springer -Verlag, New York.
- Robert, C. and Casella, G. (2004). Monte Carlo statistical methods, 2nd edn, Springer, New York.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility, in D. Cox, D. Hinkley and O. Barndorff-Nielsen (eds), Time series models with econometric, finance and other applications, Chapman and Hall, London, pp. 1 67.
- Shumway, R. and Stoffer, D. (2000). Time Series analysis and its Applications, SpringerVerlag, New York.

Storvik, G. (2002). Particle filters for StateSpace models with the

presence of unknown static parameters, *IEEE Transactions on Signal Processing* 50(2): 281–289.

Uhlig, H. (1994). On singular wishart and singular multivariate beta distributions, *Annals of Statistics* 22: 395–405.

Wang, L., Liber, G. and Manneback, P. (1992). Kalman filter algorithm based on singular value decomposition, *Proc. of the 31st conf. on decision and control*, pp. 1224–1229.

West, M. and Harrison, J. (1997). *Bayesian forecasting and dynamic models*, second edition. first edition: 1989 edn, Springer, New York.

Wiener, N. (1949). *The Extrapolation, Intepolation and Smoothing of Stationary Time Series*, John Wiley & Sons, New York.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons. New York.