



Journal of Statistical Software

May 2011, Volume 41, Issue 1.

<http://www.jstatsoft.org/>

Statistical Software for State Space Methods

Jacques J. F. Commandeur Siem Jan Koopman

VU University Amsterdam

VU University Amsterdam

Marius Ooms

VU University Amsterdam

Abstract

In this paper we review the state space approach to time series analysis and establish the notation that is adopted in this special volume of the *Journal of Statistical Software*. We first provide some background on the history of state space methods for the analysis of time series. This is followed by a concise overview of linear Gaussian state space analysis including the modelling framework and appropriate estimation methods. We discuss the important class of unobserved component models which incorporate a trend, a seasonal, a cycle, and fixed explanatory and intervention variables for the univariate and multivariate analysis of time series. We continue the discussion by presenting methods for the computation of different estimates for the unobserved state vector: filtering, prediction, and smoothing. Estimation approaches for the other parameters in the model are also considered. Next, we discuss how the estimation procedures can be used for constructing confidence intervals, detecting outlier observations and structural breaks, and testing model assumptions of residual independence, homoscedasticity, and normality. We then show how ARIMA and ARIMA components models fit in the state space framework to time series analysis. We also provide a basic introduction for non-Gaussian state space models. Finally, we present an overview of the software tools currently available for the analysis of time series with state space methods as they are discussed in the other contributions to this special volume.

Keywords: ARMA model, Kalman filter, state space methods, unobserved components, software tools.

1. Promising trends

In 2001, when considering the possible drawbacks of state space models, Durbin and Koopman wrote: “In our opinion, the only disadvantages are the relative lack in the statistical and econometric communities of information, knowledge, and software regarding these models.”, see Durbin and Koopman (2001, p. 52). Ten years later, it is gratifying to see how much progress has been made in the further dissemination of these methods. Not only have state

space models been applied in a growing number of scientific fields, but – as is witnessed by this special volume of the *Journal of Statistical Software* that is completely dedicated to statistical software for state space methods – they have been implemented in **STAMP**, **R**, **MATLAB**, **REGCMPNT**, **SAS**, **EViews**, **GAUSS**, **Stata**, **RATS**, **gretl**, and **SsfPack** with links established with **S-PLUS** and **Ox**.

State space methods originated in the field of control engineering, starting with the ground-breaking paper of [Kalman \(1960\)](#). They were initially (and still are) deployed for the purpose of accurately tracking the position and velocity of moving objects such as ships, airplanes, missiles, and rockets. A riveting account of this application of state space methods to space travel can be found in the NASA (National Aeronautics and Space Administration) report of [McGee and Schmidt \(1985\)](#), a document that has clearly been written on an old-fashioned type-writer, and can be downloaded as a pdf file from <http://ntrs.nasa.gov/>. The document explains how state space methods contributed to the successful attempt of the Apollo 11 mission to land Neil Armstrong and Buzz Aldrin on the moon in 1969.

Around the eighties of the last century it was recognized by scientists involved in other fields than control engineering that these ideas could well be applied to time series analysis generally as well. Since then state space methods have been applied in a wide range of subjects, including economics, finance, political science, environmental science, road safety and medicine.

Nowadays state space methods are also used for fitting the autoregressive integrated moving average (ARIMA) models of [Box and Jenkins \(1976\)](#) as these can be put in state space form and then analysed by the Kalman filter. When fitting an ARIMA model to a time series with missing observations in SPSS version 15.0, for example, the SPSS output informs the user that “A Kalman filtering algorithm was used for estimation”, confirming the fact that missing values are easily handled in a state space framework while this is much more difficult in the Box-Jenkins approach to time series analysis.

In [Section 2](#) we first provide a general overview of state space models and unobserved component models in particular. State space methods are reviewed in [Section 3](#). [Section 4](#) introduces ARIMA and ARIMA components models, and shows how these can be put in state space form. Generalized or non-Gaussian time series models are discussed in [Section 5](#). In [Section 6](#) we introduce all the software packages currently capable of fitting state space models to time series data. [Section 7](#) concludes.

2. Linear Gaussian state space models

A time series is a set of observations which are sequentially ordered over time. In a state space analysis the time series observations are assumed to depend linearly on a *state vector* that is unobserved and is generated by a stochastically time-varying process (a dynamic system). The observations are further assumed to be subject to measurement error that is independent of the state vector. The state vector can be estimated or identified once a sufficient set of observations becomes available. In this section we concentrate on the state space model and its special cases. In [Section 3](#) we discuss methods for estimation, residual analysis and forecasting on the basis of state space models. The expositions rely mostly on the introductory textbook by [Commandeur and Koopman \(2007\)](#) and on the more advanced textbooks by [Harvey \(1989\)](#) and [Durbin and Koopman \(2001\)](#).

The general linear Gaussian state space model for the n -dimensional observation sequence y_1, \dots, y_n will be given in each contribution of this special volume by

$$y_t = Z_t \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, H_t), \quad (1)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim \text{NID}(0, Q_t), \quad t = 1, \dots, n, \quad (2)$$

where α_t is the state vector, ε_t and η_t are disturbance vectors and the system matrices Z_t, T_t, R_t, H_t and Q_t are fixed and known but a selection of elements may depend on an unknown parameter vector. Equation (1) is called the *observation* or *measurement equation*, while (2) is called the *state* or *transition equation*. The $p \times 1$ observation vector y_t contains the p observations at time t and the $m \times 1$ state vector α_t is unobserved. The $p \times 1$ irregular vector ε_t has zero mean and $p \times p$ variance matrix H_t .

The $p \times m$ matrix Z_t links the observation vector y_t with the unobservable state vector α_t and may consist of regression variables. The $m \times m$ transition matrix T_t in (2) determines the dynamic evolution of the state vector. The $r \times 1$ disturbance vector η_t for the state vector update has zero mean and $r \times r$ variance matrix Q_t . The observation and state disturbances ε_t and η_t are assumed to be serially independent and independent of each other at all time points. In many standard cases, $r = m$ and matrix R_t is the identity matrix I_m . In other cases, matrix R_t is an $m \times r$ selection matrix with $r < m$. Although matrix R_t can be specified freely, it is often composed of a selection from the first r columns of the identity matrix I_m .

The initial state vector α_1 is assumed to be generated as

$$\alpha_1 \sim \text{NID}(a_1, P_1),$$

independently of the observation and state disturbances ε_t and η_t . Mean a_1 and variance P_1 can be treated as given and known in almost all stationary processes for the state vector. For nonstationary processes and regression effects in the state vector, the associated elements in the initial mean a_1 can be treated as unknown and need to be estimated. For an extensive discussion of *initialization* in state space analysis, we refer to [Durbin and Koopman \(2001, Chapter 5\)](#)

2.1. Local level model and other unobserved component models

By appropriate choices of the vectors α_t, ε_t and η_t , and of the matrices Z_t, T_t, H_t, R_t and Q_t , a wide range of different time series models can be derived from (1) and (2). Here we discuss the class of *unobserved components time series models*. A number of special cases will be discussed in some detail. Special attention is given to the univariate *local level model*.

Letting

$$\alpha_t = \mu_t, \quad \eta_t = \xi_t, \quad Z_t = T_t = R_t = 1, \quad H_t = \sigma_\varepsilon^2, \quad Q_t = \sigma_\xi^2,$$

(all of order 1×1) for $t = 1, \dots, n$, model (1)–(2) reduces to the local level model as given by

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \end{aligned} \quad (3)$$

for $t = 1, \dots, n$. The level component μ_t can be conceived of as the equivalent of the intercept in the classical linear regression model $y_t = \mu + \varepsilon_t$ which is obtained by setting all the level

disturbances ξ_t in (3) equal to zero and with $\mu = \mu_1$. The key difference is that the intercept μ in a regression model is fixed whereas the level component μ_t in (3) is allowed to change from time point to time point.

Since the second equation in (3) defines a random walk, the local level model is also referred to as the random walk plus noise model (where the noise refers to the irregular component ε_t). It can be shown that the dynamic process for $x_t = y_{t+1} - y_t = \eta_t + \varepsilon_{t+1} - \varepsilon_t$, for $t = 1, \dots, n$, reduces to the moving average process $x_t = \varepsilon_t + \theta\varepsilon_{t-1}$ where θ relates to the *signal-to-noise ratio* $q = \sigma_\xi^2 / \sigma_\varepsilon^2$ via a quadratic function. Furthermore, the forecasting function of observations generated by the local level model is equivalent to the *exponentially weighted moving average* scheme or *exponential smoothing*.

By defining

$$\alpha_t = \begin{pmatrix} \mu_t \\ \nu_t \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad Z_t = (1 \ 0),$$

$$H_t = \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{bmatrix}, \quad \text{and} \quad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the scalar notation of (1) and (2) leads to

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\ \nu_{t+1} &= \nu_t + \zeta_t, & \zeta_t &\sim \text{NID}(0, \sigma_\zeta^2), \end{aligned} \quad (4)$$

for $t = 1, \dots, n$, and we obtain the local linear trend model.

The local linear trend model requires a 2×1 state vector α_t : one element for the level component μ_t and one element for the slope component ν_t . The slope component can be conceived of as the equivalent of the regression coefficient in the classical regression model where the observed time series y_t is regressed on the independent variable time t : $y_t = \mu + \nu t + \varepsilon_t$ with $\mu = \mu_1$ and $\nu = \nu_1$. Again, the important difference is that the regression coefficient or weight ν is fixed in classical linear regression, whereas the slope ν_t in the local linear trend model is allowed to change over time.

In the situation that the observed time series consists of quarterly or monthly data, for example, the local level and the local linear trend model can be extended with a stochastic seasonal dummy component denoted here by γ_t . In the case of a quarterly time series (the seasonal length is 4), by defining

$$\alpha_t = \begin{pmatrix} \mu_t \\ \gamma_{1,t} \\ \gamma_{2,t} \\ \gamma_{3,t} \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \omega_t \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad Z_t = (1 \ 1 \ 0 \ 0),$$

$$H_t = \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 & 0 \\ 0 & \sigma_\omega^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and expanding (1) and (2) in scalar notation, we obtain

$$\begin{aligned}
 y_t &= \mu_t + \gamma_{1,t} + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\
 \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\
 \gamma_{1,t+1} &= -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, & \omega_t &\sim \text{NID}(0, \sigma_\omega^2), \\
 \gamma_{2,t+1} &= \gamma_{1,t}, \\
 \gamma_{3,t+1} &= \gamma_{2,t},
 \end{aligned} \tag{5}$$

for $t = 1, \dots, n$, which is a local level and dummy seasonal model for a quarterly time series where the seasonal component is allowed to change over time. The seasonal dummy model is not the only approach to incorporate time-varying seasonal effects in unobserved components time series model; see [Proietti \(2000\)](#) for a review of different seasonal specifications and their properties.

When a slope component is included in (5) as well, [Harvey \(1989\)](#) refers to this model as the *basic structural time series model*. A typical application of this model is for the *seasonal adjustment* of time series. A seasonally adjusted time series is defined in this context simply by the estimate of $y_t - \gamma_t = \mu_t + \varepsilon_t$ for $t = 1, \dots, n$.

Another extension is to include one or more cycles to any of the special models within the class of unobserved components time series models. By defining

$$\begin{aligned}
 \alpha_t &= \begin{pmatrix} \mu_t \\ c_t \\ c_t^* \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \kappa_t \\ \kappa_t^* \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho \cos(\lambda_c) & \rho \sin(\lambda_c) \\ 0 & -\rho \sin(\lambda_c) & \rho \cos(\lambda_c) \end{bmatrix}, \quad Z_t = (1 \quad 1 \quad 0), \\
 H_t &= \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_c^2(1 - \rho^2) & 0 \\ 0 & 0 & \sigma_c^2(1 - \rho^2) \end{bmatrix}, \quad \text{and} \quad R_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},
 \end{aligned}$$

in (1) and (2), we obtain the following local level plus cycle model as given by

$$\begin{aligned}
 y_t &= \mu_t + c_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\
 \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\
 c_{t+1} &= \rho[\cos(\lambda_c)c_t + \sin(\lambda_c)c_t^*] + \kappa_t, & \kappa_t &\sim \text{NID}(0, \sigma_c^2(1 - \rho^2)), \\
 c_{t+1}^* &= \rho[-\sin(\lambda_c)c_t + \cos(\lambda_c)c_t^*] + \kappa_t^*, & \kappa_t^* &\sim \text{NID}(0, \sigma_c^2(1 - \rho^2)),
 \end{aligned} \tag{6}$$

for $t = 1, \dots, n$, where $0 < \rho \leq 1$ is the *damping factor* and λ_c is the frequency of the cycle measured in radians so that $2\pi / \lambda_c$ is the *period* of the cycle. In case $\rho = 1$, the cycle reduces to a fixed sine-cosine wave but the component is still stochastic since the initial values c_1 and c_1^* are stochastic variables with mean zero and variance σ_c^2 . A typical application of this model is for the signal extraction of *business cycles* from macro-economic time series.

2.2. Regression and intervention effects

Another extension of the local level and local linear trend models concerns the incorporation of fixed explanatory and intervention variables. In the case of one regression variable x_t and one intervention variable w_t , for example, we have $y_t = \mu_t + \beta x_t + \lambda w_t + \varepsilon_t$ for the local level

model and a state vector of three elements is required: one for the level component μ_t , one for the regression coefficient β , and one for the regression coefficient λ . The substitution of

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \\ \lambda_t \end{pmatrix}, \quad \eta_t = \xi_t, \quad T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Z_t = (1 \quad x_t \quad w_t),$$

$$H_t = \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad R_t = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

in (1) and (2) results in

$$\begin{aligned} y_t &= \mu_t + \beta_t x_t + \lambda_t w_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\ \beta_{t+1} &= \beta_t, \\ \lambda_{t+1} &= \lambda_t, \end{aligned} \tag{7}$$

where $\beta = \beta_1 = \beta_t$ and $\lambda = \lambda_1 = \lambda_t$ for $t = 1, \dots, n$. This is the local level model with one continuous explanatory variable x_t and one intervention variable w_t . By adding disturbance terms to the state equation for β_t in (7), this regression weight is effectively subjected to a random walk, thus allowing for the estimation of time-varying regression effects (see [Harrison and Stevens \(1976\)](#) for an early application of such time-varying regression effects).

Letting τ denote the time point at which an intervention effect occurred, variable w_t can either be coded as a pulse intervention:

$$w_t = \begin{cases} 0, & t < \tau, \quad t > \tau \\ 1, & t = \tau \end{cases}$$

(to model an outlier observation), or as a level intervention:

$$w_t = \begin{cases} 0, & t < \tau, \\ 1, & t \geq \tau, \end{cases}$$

(to model a structural break in the level of the series), or as a slope intervention:

$$w_t = \begin{cases} 0, & t < \tau, \\ 1 + t - \tau, & t \geq \tau, \end{cases}$$

(to model a structural break in the slope of the series). Other types of intervention effects can be modelled as well, see [Box and Tiao \(1975\)](#).

2.3. Structural time series analysis

From the previous sections, a structural approach to time series analysis emerges which is facilitated by the state space framework. In this approach, different unobserved components or building blocks responsible for the dynamics of the series such as trend, seasonal, cycle, and the effects of explanatory and intervention variables are identified separately before being

put together in a state space model. It is the responsibility of the researcher to decide what components are required in a specific situation, and then to consider whether they apply to the time series under consideration. This explains why state space models are often referred to as *structural time series models*. The monograph of Harvey (1989) has been instrumental in the dissemination of state space models outside the field of control engineering.

2.4. Multivariate models

The structural approach to time series analysis is easily extended to multivariate time series. For example, when we denote y_t as a $p \times 1$ vector of observations, a multivariate local level model can be adopted for modelling the p time series simultaneously:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \Sigma_\varepsilon), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \Sigma_\xi), \end{aligned} \tag{8}$$

for $t = 1, \dots, n$, where μ_t , ε_t , and ξ_t are $p \times 1$ vectors and Σ_ε and Σ_ξ are $p \times p$ variance matrices. In what is known as the *seemingly unrelated time series equations model* (8), the series are modelled as in the univariate situation, but the disturbances driving the level components are allowed to be instantaneously correlated across the p series. When slope, seasonal, or cycle components are also included in the model, each of these three components has an associated $p \times p$ variance matrix for their disturbance vectors, allowing for correlation between the series. If the rank r of Σ_ξ in (8) is assumed smaller than p , then the model implies that the p series have r *common levels*. Such common factors may not only have a nice and interesting interpretation, but may also result in more efficient inferences and forecasts since the number of parameters reduces.

3. State space analysis

For given values of all system matrices – and for given initial conditions a_1 and P_1 – the state vector can be estimated in three different ways, yielding what are known as the *filtered*, the *predicted*, and the *smoothed* estimates of the state vector. Depending on the types of state estimates required in the analysis, the estimates of the state vector can be obtained by performing one or two passes through the observed time series:

1. a *forward* pass, from $t = 1, \dots, n$, using a recursive algorithm known as the *Kalman filter* enables the computation of predicted states, based on y_1, \dots, y_{t-1} , filtered states, based on y_1, \dots, y_t , and observation prediction errors;
2. a *backward* pass, from $t = n, \dots, 1$, using output of the Kalman filter and using recursive algorithms known as *state and disturbance smoothers* enables the computation of smoothed estimates of states and disturbances that are based on all observations y_1, \dots, y_n .

3.1. Kalman filter for prediction, filtering and forecasting

The forward pass through the data with the well-known Kalman (1960) filter provides all estimates that are relevant for the filtered and the predicted state. The main purpose of the

Kalman filter is to obtain optimal estimates of the state at time point t , *only* considering the observations $\{y_1, y_2, \dots, y_{t-1}\}$. A key property of the predicted state and its related estimates is therefore that they are only based on *past values* of the observed time series. The recursive formulas for the Kalman filter are

$$\begin{aligned} v_t &= y_t - Z_t a_t, & F_t &= Z_t P_t Z_t^\top + H_t, \\ K_t &= T_t P_t Z_t^\top F_t^{-1}, & L_t &= T_t - K_t Z_t, \\ a_{t+1} &= T_t a_t + K_t v_t, & P_{t+1} &= T_t P_t L_t^\top + R_t Q_t R_t^\top, \end{aligned} \quad (9)$$

for $t = 1, \dots, n$. The values of a_t in (9) represent the predicted state, while the values of P_t quantify the estimation error variance matrix of the predicted state a_t . Under the assumption of normality, the latter variances are useful for the construction of *confidence intervals* for the predicted state, which – assuming that we are interested in their 90% confidence limits for example – can be calculated as

$$a_t \pm 1.64\sqrt{P_t},$$

for $t = 1, \dots, n$. A modification of the Kalman filter also allows the computation of the filtered estimate of the state vectors, that is

$$a_{t|t} = a_t + P_t Z_t^\top F_t^{-1} v_t, \quad P_{t|t} = P_t - P_t Z_t^\top F_t^{-1} Z_t P_t, \quad t = 1, \dots, n,$$

where $a_{t|t}$ is the optimal estimate of the state at time point t by considering the observations $\{y_1, y_2, \dots, y_t\}$ while $P_{t|t}$ is the state filtered estimation error variance matrix. The values of v_t in (9) are called the *one-step ahead prediction* or *forecast errors*, since they quantify the lack of accuracy of a_t in predicting the observed value of y_t at time point t ; the values of F_t are the variances of these one-step ahead prediction errors v_t .

One of the convenient features of state space methods is the ease with which they deal with two important aspects of time series analysis – forecasting and missing observations: by treating them in exactly the same way. Missing observations are handled by setting K_t and v_t in (9) equal to 0. Forecasts for y_{n+1}, \dots, y_{n+k} given y_1, \dots, y_n are simply obtained by applying the Kalman filter for $t = 1, \dots, n, n+1, \dots, n+k$ and by treating y_{n+1}, \dots, y_{n+k} as missing observations.

3.2. State and disturbance smoothing

The backward pass through the data is only required for smoothing that leads to estimates such as the smoothed states and smoothed disturbances. The main purpose of state and disturbance smoothing is to obtain estimated values of the state and disturbance vectors at time point t , considering *all* available observations $\{y_1, y_2, \dots, y_n\}$. The recursive formulas for state smoothing are

$$r_{t-1} = Z_t^\top F_t^{-1} v_t + Z_t^\top r_t, \quad N_{t-1} = Z_t^\top F_t^{-1} Z_t + L_t^\top N_t L_t, \quad (10)$$

$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \quad V_t = P_t - P_t N_{t-1} P_t, \quad (11)$$

for $t = n, \dots, 1$. The recursive formulas for smoothing (10) are initialized with $r_n = 0$ and $N_n = 0$. The state smoothing equations (11) yield the smoothed state estimate $\hat{\alpha}_t$ and is defined as the optimal estimate of α_t using the full set of observations $\{y_1, y_2, \dots, y_n\}$;

the state smoothing equations also yield the corresponding smoothed state estimation error variance matrix V_t .

Analogous to the predicted state, under the assumption of normality the smoothed state estimation error variance matrix V_t is useful for the construction of confidence intervals for the smoothed state components, which – should we happen to be interested in their 90% confidence limits for example – can be calculated as

$$\hat{\alpha}_t \pm 1.64\sqrt{V_t},$$

for $t = 1, \dots, n$.

The recursions for r_{t-1} and N_{t-1} in (10) also enable the computation of the smoothed estimates of the disturbances ε_t and η_t in the following way,

$$\hat{\varepsilon}_t = H_t \left(F_t^{-1} v_t - K_t^\top r_t \right), \quad \text{Var}(\hat{\varepsilon}_t) = H_t \left(F_t^{-1} + K_t^\top N_t K_t \right) H_t, \quad (12)$$

$$\hat{\eta}_t = Q_t R_t^\top r_t, \quad \text{Var}(\hat{\eta}_t) = Q_t R_t^\top N_t R_t Q_t, \quad (13)$$

for $t = n, \dots, 1$. The equations (12) and (13) compute the smoothed observation disturbances $\hat{\varepsilon}_t$, the smoothed state disturbances $\hat{\eta}_t$, and their corresponding smoothed estimation error variance matrices $\text{Var}(\hat{\varepsilon}_t)$ and $\text{Var}(\hat{\eta}_t)$.

3.3. Diagnostic checking

All significance tests in linear Gaussian state space models – and the construction of confidence intervals – are based on three assumptions concerning the residuals of the analysis. The residuals should satisfy independence, homoscedasticity, and normality, in this order of importance. Whether the residuals satisfy these three assumptions can be established by diagnosing what are known as the *standardized prediction errors*. They are defined as

$$v_t^* = \frac{v_t}{\sqrt{F_t}}, \quad (14)$$

for $t = 1, \dots, n$. For the computations of the one step-ahead prediction errors v_t and their variances F_t in (14), we refer to the recursive formulas for the Kalman filter given in (9). In case y_t is a vector of time series, the prediction error v_t^* is defined by $v_t^* = G_t v_t$ where G_t is constructed such that $G_t G_t^\top = F_t^{-1}$.

The assumptions of independence and normality of the residuals can be diagnosed using the Box-Ljung test statistic and the Bowman and Shenton test statistic, respectively. The assumption of homoscedasticity can be checked by testing whether the variance of the standardized prediction errors in the first third part of the series is equal to the variance of the errors corresponding to the last third part of the series. For further details concerning these diagnostic tests, we refer to Harvey (1989), Durbin and Koopman (2001) and Commandeur and Koopman (2007).

A second diagnostic tool for determining the appropriateness of a model is provided by inspection of what are known as the *auxiliary residuals*. As already mentioned above, the disturbance smoothing filters applied in the backward pass through the data yield, amongst others, estimates of the smoothed observation and state disturbances, and of their variances.

The auxiliary residuals are obtained by dividing the smoothed observation and state disturbances with the square root of their corresponding variances, as follows:

$$e_t^* = \frac{\hat{\varepsilon}_t}{\sqrt{\text{Var}(\hat{\varepsilon}_t)}} \quad \text{and} \quad r_t^* = \frac{\hat{\eta}_t}{\sqrt{\text{Var}(\hat{\eta}_t)}}, \quad (15)$$

for $t = 1, \dots, n$, resulting in *standardized* smoothed disturbances. Visual inspection of the standardized smoothed observation disturbances e_t^* allows for the detection of possible *outlier* observations in a time series, while inspection of the standardized smoothed state disturbances r_t^* enables the detection of *structural breaks* in the underlying development of a time series.

Each auxiliary residuals can be considered as a t test for the null hypothesis that there was no outlier observation (when inspecting the auxiliary residuals at the left of (15)) or as a t test for the null hypothesis that there was no structural break in the corresponding unobserved component of the observed time series (when inspecting the auxiliary residuals at the right of (15)). Applying the usual 95% confidence limits of ± 1.96 corresponding to a two-tailed t test, possible outlier observations or structural breaks in the unobserved components making up the state vector are thus easily detected. A more detailed discussion on outlier and break detection in a state space analysis is provided by [Harvey and Koopman \(1992\)](#) and [de Jong and Penzer \(1998\)](#).

3.4. Parameter estimation

So far, we have presented all of the results that can be obtained with state space methods as if the disturbance variances, the fixed regression effects, the parameters ρ and λ_c associated with cycles, etcetera, are given and known. In practice, of course, these parameters are unknown, and have to be estimated.

It can be shown that the Kalman filter presented in (9) also provides the necessary ingredients required for evaluating the log-likelihood function, which is given by

$$\log L(y|\psi) = -\frac{np}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \left(\log |F_t| + v_t^\top F_t^{-1} v_t \right), \quad (16)$$

where the v_t are the one-step ahead prediction errors, the F_t are their variances for $t = 1, \dots, n$ defined in (9), and ψ denotes the vector of unknown parameters. The log-likelihood (16) is maximized with respect to ψ numerically using the score vector or the EM algorithm, see [Durbin and Koopman \(2001\)](#) and [Commandeur and Koopman \(2007\)](#) for details. In this way we obtain the maximum likelihood estimate of ψ .

4. ARIMA and ARIMA components models

In this section we consider more traditional approaches to the analysis of time series proposed in [Box and Jenkins \(1976\)](#). We will show how the *autoregressive integrated moving average* (ARIMA) model can be put into state space form. An essential feature of the Box-Jenkins approach to time series analysis is that the observed times series is assumed to be *stationary*, meaning that its means and covariances should be invariant when the series is shifted or translated through time. If an observed time series satisfies stationarity, then the corresponding series can be analysed with an *autoregressive moving average* (ARMA) model. However, if

the observed time series also contains non-stationary features like a trend and a seasonal, for example, then stationarity has to be enforced before the analysis can start by *differencing* the observed time series: key features of the series such as a trend and a seasonal will be removed from the series y_t by transforming it into

$$\Delta y_t = y_t - y_{t-1},$$

to remove the trend in the series, and

$$\Delta_s y_t = y_t - y_{t-s},$$

to remove a seasonal with periodicity s from the series. In some cases a combined removal of trend and seasonal is necessary and is achieved by

$$\Delta \Delta_s y_t = (y_t - y_{t-s}) - (y_{t-1} - y_{t-s-1}).$$

Should the differenced time series still not be stationary, then the differencing procedure can be continued by taking second differences such as

$$\Delta^2 y_t, \quad \Delta_s^2 y_t, \quad \Delta^2 \Delta_s y_t, \quad \Delta \Delta_s^2 y_t, \quad \Delta^2 \Delta_s^2 y_t,$$

or even combinations with third differences, etcetera.

We denote the stationary time series by y_t^* which can be equal to the series y_t but can also be the series y_t after it is appropriately differenced. The Box-Jenkins ARMA(p, q) model for the analysis of univariate time series is given by

$$y_t^* = \phi_1 y_{t-1}^* + \cdots + \phi_p y_{t-p}^* + \zeta_t + \theta_1 \zeta_{t-1} + \cdots + \theta_q \zeta_{t-q}, \quad \zeta_t \sim \text{NID}(0, \sigma_\zeta^2), \quad (17)$$

where p and q are non-negative integers, and ζ_t is a series of independent disturbances (also referred to as *white noise*) that is normally distributed here. In (17) the values of p and q denote the number of autoregressive and moving average terms required to model the observed series, respectively. If the observed series already satisfies stationarity, then $y_t^* = y_t$ in (17), and the corresponding model is called an ARMA model. Should this not be the case, then y_t^* is obtained after some differencing operations and we are dealing with an ARIMA model.

To illustrate how an ARMA model can be formulated in state space form, we consider the ARMA(2, 1) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \zeta_t + \theta_1 \zeta_{t-1}, \quad \zeta_t \sim \text{NID}(0, \sigma_\zeta^2),$$

for $t = 1, \dots, n$. By considering the state space model (1)-(2) and by defining

$$\alpha_t = \begin{pmatrix} y_{t+1} \\ \phi_2 y_t + \theta_1 \zeta_{t+1} \end{pmatrix}, \quad \eta_t = \zeta_{t+1}, \quad T_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix}, \quad Z_t = (1 \ 0), \\ H_t = 0, \quad Q_t = \sigma_\zeta^2, \quad R_t = \begin{bmatrix} 1 \\ \theta_1 \end{bmatrix},$$

for $t = 1, \dots, n$, we have represented the ARMA(2, 1) model in the state space framework. For a general overview of how any of the general ARMA (p, q) models can be represented in the state space framework, we refer to [Durbin and Koopman \(2001, Section 3.3\)](#).

A more general approach to time series analysis can be based on the ARIMA components model as given by

$$y_t = \sum_{j=1}^m \mu_t^{(j)}, \quad \mu_t^{(j)} \sim \text{ARIMA}, \quad j = 1, \dots, m, \quad (18)$$

for $t = 1, \dots, n$, where the $\mu_t^{(j)}$'s follow mutually independent ARIMA processes with different levels of differencing and different values for p and q for $j = 1, \dots, m$. The dynamic properties of y_t can be derived from the different ARIMA processes while the time series y_t can be expressed as an ARIMA process itself. Unobserved components time series models can be regarded as special cases of the ARIMA components model. For example, the local level model (3) can be expressed as an ARIMA components model with $m = 2$, $\mu_t^{(1)} = \mu_t$ and $\mu_t^{(2)} = \varepsilon_t$. We take $\mu_t^{(1)}$ as an ARMA(0,0) process after first differencing (with $p = 0$ and $q = 0$) and we take $\mu_t^{(2)}$ as an ARMA(0,0) process without differencing. A more general and more detailed discussion of ARIMA components models can be found in Bell (2004).

5. Non-Gaussian state space models

The state space model (1)–(2) is for Gaussian observations y_t . In case the observations y_t are discrete or have other distributional properties than Gaussian, we can consider the generalized state space model

$$y_t \sim p(y_t | Z_t \alpha_t), \quad (19)$$

for $t = 1, \dots, n$, where $p(\cdot)$ is a particular non-Gaussian density function, possibly for discrete observations, and with the state vector α_t and system matrix Z_t as defined in the previous section. Conditional on the state α_t , we assume that the observations are serially independent. This assumption is similar to the linear Gaussian model. The state vector α_t evolves stochastically over time according to the linear Gaussian state equation (2).

Time series observations which come from the exponential family distribution require density function

$$p(y_t | \theta_t) = \exp \left[y_t^\top \theta_t - b_t(\theta_t) + c_t(y_t) \right], \quad (20)$$

where $\theta_t = Z_t \alpha_t$, $b_t(\cdot)$ is a twice differentiable function and $c_t(\cdot)$ is a function of y_t only. The modelling framework (19)–(2)–(20) makes it possible to analyse time series observations coming from, among others, Poisson, binary, binomial and multinomial distributions. For example, in case a Poisson density is considered, we have $b_t(\theta_t) = \exp \theta_t$ and $c_t(y_t) = \log(y_t!)$.

The generalized state space model (19)–(2)–(20) is more general than the exponential family only. We can consider continuous density functions for $p(\cdot)$ such as those related to the heavy-tailed distributions Student's t , mixture of normals and general error. In these cases, we can often return to the linear observation equation (19) but written as

$$y_t = \theta_t + \varepsilon_t, \quad \varepsilon_t \sim p(\varepsilon_t),$$

where $p(\varepsilon_t) = p(y_t | \theta_t)$. When we take $p(\cdot)$ as the Gaussian density with mean zero and variance H_t , the generalized model reduces to the Gaussian state space model (1)–(2).

A particular example of a non-Gaussian model of interest for financial markets is the stochastic volatility model where the observations are financial returns with a constant mean and a stochastically time-varying variance. The univariate stochastic volatility model is given by

$$y_t = \exp\left(\frac{1}{2}\theta_t\right)\varepsilon_t, \quad \varepsilon_t \sim p(\varepsilon_t), \quad (21)$$

where $p(\varepsilon_t)$ can be Gaussian, Student's t or mixture of normals. The log-variance θ_t can be generally specified as $\theta_t = Z_t\alpha_t$ but it is usually considered as a stationary autoregressive process. For more detailed discussions concerning the stochastic volatility model (21), see [Harvey, Ruiz, and Shephard \(1994\)](#) and [Shephard \(2005\)](#).

The statistical analysis based on generalized state space models cannot rely on linear methods such as the Kalman filter and smoothing recursions of Section 3. A number of statistical approaches can be adopted to treat non-Gaussian and nonlinear features of the model adequately. [Durbin and Koopman \(2000\)](#) have developed an estimation methodology based on importance sampling and simulated maximum likelihood methods. Bayesian estimation methods for generalized state space models rely typically on Markov chain Monte Carlo (MCMC) methods and are developed by [Frühwirth-Schnatter \(1994\)](#), [Carter and Kohn \(1994\)](#) and [Shephard and Pitt \(1997\)](#).

More recent developments on the analysis of generalized or nonlinear state space models are reported in the engineering literature. These developments are based on methods collectively known as particle filtering and originate from the work by [Gordon, Salmond, and Smith \(1993\)](#), [Kitagawa \(1996\)](#) and with a noteworthy contribution by [Pitt and Shephard \(1999\)](#). Particle filtering and related Monte Carlo methods are still in development as can be learned from the recent review article of [Creal \(2011\)](#).

Textbook treatments of generalized state space models from both classical and Bayesian perspectives are presented by, among others, [Harvey \(1989, Section 6.6\)](#), [West and Harrison \(1997\)](#), [Doucet, deFreitas, and Gordon \(2000\)](#) and [Durbin and Koopman \(2001, Part II\)](#).

6. Software packages

An overview of all the software packages covered in this special volume is provided in Table 1, including their current version number, and indicating whether they currently can handle univariate unobserved component models, multivariate unobserved component models, the univariate treatment of multivariate time series presented in [Durbin and Koopman \(2001, Chapter 6\)](#), the exact initialization procedures described in [Durbin and Koopman \(2001, Chapter 5\)](#), and non-Gaussian state space models.

7. Structure of the papers

When inviting the authors to contribute to this special volume, in order to obtain some degree of uniformity we asked them all to adhere to the following structure in their papers:

- Title: State space methods in <name of software package>
- Author: <author>

	UM	MM	EI	UTMTS	NLNGM
EViews 6	√	√			
gretl 1.9.1	√	√		√	
MATLAB 7.0 R14	√	√	√		√
R base 2.11.1	√				
R + d1m 1.0-2	√	√			
R + KFAS 0.5.1	√	√	√	√	√
RATS 8	√	√	√		√
REGCMPNT 1.16	√		√		
SAS 9.2	√	√	√		
S-PLUS	√	√	√	√	√
SsfPack Basic 3.0	√	√			√
SsfPack Extended 3.0	√	√	√	√	√
STAMP 8.2	√	√	√	√	
Stata 11	√	√	√	√	

Table 1: Overview of state space software packages and their options. The options are univariate models (UM), multivariate models (MM), exact initialization (EI), univariate treatment of multivariate time series (UTMTS), and non-linear and non-Gaussian models (NLNGM).

- Abstract
- 1. Introduction
- 2. Case 1: The local level model applied to the Nile data
- 3. Case 2: ...
- 4. Case 3 (possibly)
- 5. Conclusions (including possibilities and limitations of the software at hand)

For each case study in their contribution, we asked them to present and discuss the (important parts of the) required code and the analysis results. Moreover, for case study 1 we asked the contributors to show how to apply the local level model to the Nile data (a series of readings of the annual flow volume at Aswan from 1871 to 1970 also used as an illustration in [Durbin and Koopman 2001](#)) with the software at hand, and when discussing this analysis in their paper we also asked them to include the following results:

- the ML estimates of the two variances including asymptotic standard errors
- the plot with data and smoothed state vector with 90% CI
- the plot with standardized prediction errors
- the two plots with auxiliary residuals and 95% CI
- a plot with forecasts until and including 1980 with 50% CI

After discussing the code and results of the analysis of the Nile data with the local level model in Section 2 of their paper, we finally asked them to add one or two examples of their own design. Of these added examples we mention the following highlights, including references to the corresponding papers in this volume:

- illustrations on ocean and climate time series, specifically El Niño and a space-time data set on sea surface temperature (with **STAMP** in [Mendelssohn 2011](#));
- a basic structural model with cycle for the Italian industrial production and a stochastic volatility model for the FTSE100 data (with **Ox/SsfPack** in [Pelagatti 2011](#));
- a Bayesian analysis of the Nile data and a multivariate dynamic capital asset pricing model for four assets (with **R** in [Petris and Petrone 2011](#));
- an affine model for the term structure of interest rates (with **S-PLUS** in [Zivot 2011](#));
- a cubic spline analysis for the motorcycle acceleration data of Silverman (1985) and some non-Gaussian illustrations (with **MATLAB** in [Peng and Aston 2011](#));
- a time series model with a sampling error component (with **REGCMPNT** in [Bell 2011](#));
- a bivariate latent risk model for the simultaneous analysis of mobility and road traffic fatalities (with **EViews** in [Van den Bossche 2011](#));
- an illustration of the dynamic stochastic general equilibrium (DSGE, with **RATS** in [Doan 2011](#));
- a vector autoregressive moving average (VARMA) model, and a dynamic factor model (with **Stata** in [Drukker and Gates 2011](#));
- a multivariate structural time series model for the labour market applied to the number of employees and the number of hours worked (with **gretl** in [Lucchetti 2011](#));
- the use of diffuse priors in the state vector with an application for a two-way random effects panel model (with **SAS** in [Selukar 2011](#));
- a Bayesian analysis of time varying volatility for Standard & Poors 500 returns (with **Ox** in [Bos 2011](#)).

8. Conclusions

We have reviewed state space methods for linear Gaussian state space models and discussed their generalizations for models with nonlinear non-Gaussian features. In this special volume of the *Journal of Statistical Software* twelve contributions are given that discuss the currently available software implementations of state space methods and their extensions. The most important aim of this volume is to highlight the presence of such implementations for different statistical software platforms. We hope it will encourage more applied researchers to use these software tools in order to improve the quality of their time series analyses.

Acknowledgments

We would like to thank all the authors for their enthusiastic response to our invitation to contribute a paper to this special volume: Filip A. M. Van den Bossche (**EViews**), Matteo M. Pelagatti (**Ox/SsfPack**), Rajesh Selukar (**SAS**), David Drukker and Richard Gates (**Stata**), Giovanni Petris and Sonia Petrone (**R**), Jyh-Ying Peng and John A. D. Aston (**MATLAB**), Roy Mendelsohn (**STAMP**), Riccardo (Jack) Luchetti (**gretl**), Charles S. Bos (**Ox-Bayes-MCMC**), William R. Bell (**REGCMPNT**), Thomas Doan (**RATS**), and Eric Zivot (**S-PLUS**). We are also grateful for the referees of these papers: Michael Massman, Stefano Grassi, John A. D. Aston, Charles S. Bos, Matteo M. Pelagatti, Giovanni Petris, David Drukker, William R. Bell, Riccardo Luchetti, Filip A. M. Van den Bossche, and Roy Mendelsohn. We finally are indebted to Achim Zeileis for his extensive editorial help.

References

- Bell WR (2004). “On RegComponents Time Series Models and Their Applications.” In AC Harvey, SJ Koopman, N Shephard (eds.), *State Space and Unobserved Components Models*. Cambridge University Press, Cambridge.
- Bell WR (2011). “**REGCMPNT** – A Fortran Program for Regression Models with ARIMA Component Errors.” *Journal of Statistical Software*, **41**(7), 1–23. URL <http://www.jstatsoft.org/v41/i07/>.
- Bos CS (2011). “A Bayesian Analysis of Unobserved Component Models Using Ox.” *Journal of Statistical Software*, **41**(13), 1–24. URL <http://www.jstatsoft.org/v41/i13/>.
- Box GEP, Jenkins GM (1976). *Time Series Analysis*. Holden-Day, San Francisco.
- Box GEP, Tiao GC (1975). “Intervention Analysis with Applications to Economic and Environmental Problems.” *Journal of the American Statistical Association*, **70**, 70–79.
- Carter CK, Kohn R (1994). “On Gibbs Sampling for State Space Models.” *Biometrika*, **81**, 541–53.
- Commandeur JJF, Koopman SJ (2007). *An Introduction to State Space Time Series Analysis*. Oxford University Press, Oxford.
- Creal DD (2011). “A Survey of Sequential Monte Carlo Methods for Economics and Finance.” *Econometric Reviews*, **29**, forthcoming.
- de Jong P, Penzer J (1998). “Diagnosing Shocks in Time Series.” *Journal of the American Statistical Association*, **93**, 796–806.
- Doan T (2011). “State Space Methods in **RATS**.” *Journal of Statistical Software*, **41**(9), 1–16. URL <http://www.jstatsoft.org/v41/i09/>.
- Doucet A, deFreitas JFG, Gordon NJ (eds.) (2000). *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York.

- Drukker DM, Gates RB (2011). “State Space Methods in Stata.” *Journal of Statistical Software*, **41**(10), 1–25. URL <http://www.jstatsoft.org/v41/i10/>.
- Durbin J, Koopman SJ (2000). “Time Series Analysis of Non-Gaussian Observations Based on State Space Models from Both Classical and Bayesian Perspectives.” *Journal of the Royal Statistical Society B*, **62**, 3–56.
- Durbin J, Koopman SJ (2001). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.
- Frühwirth-Schnatter S (1994). “Data Augmentation and Dynamic Linear Models.” *Journal of Time Series Analysis*, **15**, 183–202.
- Gordon NJ, Salmond DJ, Smith AFM (1993). “A Novel Approach to Non-Linear and Non-Gaussian Bayesian State Estimation.” *IEE-Proceedings F*, **140**, 107–13.
- Harrison PJ, Stevens CF (1976). “Bayesian Forecasting.” *Journal of the Royal Statistical Society B*, **38**, 205–247.
- Harvey AC (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.
- Harvey AC, Koopman SJ (1992). “Diagnostic Checking of Unobserved Components Time Series Models.” *Journal of Business and Economic Statistics*, **10**, 377–89.
- Harvey AC, Ruiz E, Shephard N (1994). “Multivariate Stochastic Variance Models.” *Review of Economic Studies*, **61**, 247–64.
- Kalman RE (1960). “A New Approach to Linear Filtering and Prediction Problems.” *Journal of Basic Engineering, Transactions ASMA, Series D*, **82**, 35–45.
- Kitagawa G (1996). “Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models.” *Journal of Computational and Graphical Statistics*, **5**, 1–25.
- Lucchetti R (2011). “State Space Methods in gretl.” *Journal of Statistical Software*, **41**(11), 1–22. URL <http://www.jstatsoft.org/v41/i11/>.
- McGee LA, Schmidt SF (1985). “Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry.” *Technical Report NASA-TM-86847*, NASA.
- Mendelssohn R (2011). “The **STAMP** Software for State Space Models.” *Journal of Statistical Software*, **41**(2), 1–19. URL <http://www.jstatsoft.org/v41/i02/>.
- Pelagatti MM (2011). “State Space Methods in Ox/**SsfPack**.” *Journal of Statistical Software*, **41**(3), 1–25. URL <http://www.jstatsoft.org/v41/i03/>.
- Peng JY, Aston JAD (2011). “The State Space Models Toolbox for MATLAB.” *Journal of Statistical Software*, **41**(6), 1–25. URL <http://www.jstatsoft.org/v41/i06/>.
- Petris G, Petrone S (2011). “State Space Models in R.” *Journal of Statistical Software*, **41**(4), 1–24. URL <http://www.jstatsoft.org/v41/i04/>.

- Pitt MK, Shephard N (1999). “Filtering via Simulation: Auxiliary Particle Filter.” *Journal of the American Statistical Association*, **94**, 590–9.
- Proietti T (2000). “Comparing Seasonal Components for Structural Time Series Models.” *International Journal of Forecasting*, **16**, 247–260.
- Selukar R (2011). “State Space Modeling Using SAS.” *Journal of Statistical Software*, **41**(12), 1–13. URL <http://www.jstatsoft.org/v41/i12/>.
- Shephard N (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press, Oxford.
- Shephard N, Pitt MK (1997). “Likelihood Analysis of Non-Gaussian Measurement Time Series.” *Biometrika*, **84**, 653–67.
- Van den Bossche FAM (2011). “Fitting State Space Models with EViews.” *Journal of Statistical Software*, **41**(8), 1–16. URL <http://www.jstatsoft.org/v41/i08/>.
- West M, Harrison J (1997). *Bayesian Forecasting and Dynamic Models*. 2nd edition. Springer-Verlag, New York.
- Zivot E (2011). “State Space Modeling Using **SsfPack** in **S+FinMetrics** 3.0.” *Journal of Statistical Software*, **41**(5), 1–26. URL <http://www.jstatsoft.org/v41/i05/>.

Affiliation:

Jacques J. F. Commandeur
Department of Econometrics FEWEB
VU University Amsterdam
De Boelelaan 1105
NL-1081 HV Amsterdam
E-mail: jcommandeur@feweb.vu.nl
URL: <http://staff.feweb.vu.nl/jcommandeur/>