

RESEARCH

Open Access

Filtering and identification of a state space model with linear and bilinear interactions between the states

A Al-Mazrooei^{1*}, J Al-Mutawa², M El-Gebeily² and R Agarwal^{2,3}

*Correspondence:

abo.mhmd@hotmail.com

¹Department of Mathematics,
Taibah University, Al-Madinah,
Kingdom of Saudi Arabia

Full list of author information is
available at the end of the article

Abstract

In this paper, we introduce a new bilinear model in the state space form. The evolution of this model is linear-bilinear in the state of the system. The classical Kalman filter and smoother are not applicable to this model, and therefore, we derive a new Kalman filter and smoother for our model. The new algorithm depends on a special linearization of the second-order term by making use of the best available information about the state of the system. We also derive the expectation maximization (EM) algorithm for the parameter identification of the model. A Monte Carlo simulation is included to illustrate the efficiency of the proposed algorithm. An application in which we fit a bilinear model to wind speed data taken from actual measurements is included. We compare our model with a linear fit to illustrate the superiority of the bilinear model.

Keywords: bilinear state space model; Kalman filter and smoother; maximum likelihood estimate; EM algorithm

1 Introduction

Bilinear systems are a special type of nonlinear systems capable of representing a variety of important physical processes. They are used in many applications in real life such as chemistry, biology, robotics, manufacturing, engineering, and economics [1–4] where linear models are ineffective or inadequate. They have also been recently used to analyze and forecast weather conditions [5–10].

Bilinear systems have three main advantages over linear ones: Firstly, they describe a wider class of problems of practical importance. Secondly, they provide more flexible approximations to nonlinear systems than linear systems do. Thirdly, one can make use of their rich geometric and algebraic structures, which promises to be a fruitful field of research for scientists [2] as well as practitioners.

Bilinear models were first introduced in the control theory literature in 1960s [11]. So far, the type of nonlinearity that is extensively treated and analyzed consists of bilinear interaction between the states of the system and the system input [1, 2, 12]. Aside from their practical importance, these systems are easier to handle because they are reducible to linear ones through the use of a certain Kronecker product. In this work, we treat the case where the nonlinearity of the system consists of bilinear interaction between the states of the system themselves. This means that our model will be able to handle evolutions according to the Lotka-Volterra models [6] or the Lorenz weather models [7, 8, 10], thus

enabling a wider and more flexible application of such models. To the best of our knowledge, no attempt has been made to treat such systems in the general setting presented here.

The widespread use of bilinear models motivates the need to develop their parameter identification algorithms. A lot of work exists in the literature which presents methods of estimation and parameter identification of linear and nonlinear systems [13–21]. The two most widely used techniques fall under the names of least square estimation and maximum likelihood estimation, respectively.

The maximum likelihood estimation is computed through the well-known EM algorithm [22]. It is an iterative method that tries to improve a current estimate of the system parameters by maximizing the underlying likelihood densities. The algorithm is useful in a variety of incomplete data problems, where algorithms such as the Newton-Raphson method may turn out to be more complicated. It consists of two steps called the Expectation step or the E-step and the Maximization step or the M-step; hence the name of the algorithm. This name was first coined by Dempster, Laird, and Rubin in their fundamental paper [22]. In this paper, we develop the EM algorithm for our bilinear system. This will necessitate also the development of a Kalman filter and smoother suitable for the nonlinear system at hand. The direct development of the recursions for the nonlinear filters is very complicated if not impossible altogether. Instead, we develop our recursions based on a linearization of the quadratic term that uses the most current state estimate available.

The remainder of this article is arranged as follows. In Section 2, the bilinear state space model problem is stated along with underlying assumptions. In Section 3, we derive the bilinear Kalman filter and smoother. Section 4 estimates the unknown parameters in the bilinear state space model *via* the EM algorithm. Section 5 presents a simulation example that produces very satisfactory results. A real world example is given in Section 6.

2 The bilinear state space model

In this section, we introduce a bilinear state space model and describe a generalization of the Kalman filter and smoother to this model. Our model subsumes the well-known Lorentz-96 model [7] for weather forecast, and the Lotka-Volterra evolution equations appear in many applications in chemistry, biology and control [4, 6]. Other types of bilinear models were investigated in [2, 11], where bilinearity occurs because of the interaction between the input and states of the system.

We will adopt the geometric notation as presented in [17] where the matrix inner product of two random vectors is defined by

$$\langle x, y \rangle = E(xy^T),$$

and

$$\|x\|^2 = E(xx^T) = \langle x, x \rangle.$$

We know that

$$\langle x, y \rangle = \text{cov}(x, y) + E(x)E(y)^T.$$

Given a sequence $Y = \{y_1, y_2, \dots, y_t\}$ of random vectors, the conditional expectation $E(x|Y)$ with respect to this inner product is interpreted geometrically as the orthogonal projection of the vector x in the space spanned by the vectors of Y . In particular, if x is uncorrelated with the elements of Y and if it has zero mean, then x is orthogonal to the subspace generated by Y and $E(x|Y) = 0$. We will also use the projection notation

$$\pi_t x := \pi_Y x := E(x|Y).$$

It is characterized by

$$\langle x - \pi_t x, z \rangle = 0,$$

for all $z \in \mathcal{M}(Y)$; the closed subspace of L^2 of all random vectors z which can be written as measurable functions of the elements of Y [3].

To introduce the model, let us first define the bilinear function $a : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{\frac{n(n+1)}{2}}$ by

$$a(x, y) = (x_1 y_1, x_1 y_2, \dots, x_1 y_n, x_2 y_2, x_2 y_3, \dots, x_2 y_n, \dots, x_n y_n)^T,$$

where $a(\cdot, \cdot)$ is similar to the Kronecker product function except that there is no repetition of the entries. Consider the bilinear state space model given by

$$x_{k+1} = Ax_k + Bz_k + w_k, \tag{1}$$

$$y_k = Cx_k + v_k, \quad k = 1, 2, \dots, N, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $y_k \in \mathbb{R}^p$ is the measurement vector, and $z_k = a(x_k, x_k)$ is the bilinear term given by

$$z_k = a(x_k, x_k) = (x_1^2, x_1 x_2, \dots, x_1 x_n, x_2^2, x_2 x_3, \dots, x_2 x_n, \dots, x_n^2)^T.$$

The matrices are of appropriate dimensions, *i.e.*, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times \frac{n(n+1)}{2}}$ and $C \in \mathbb{R}^{p \times p}$. The uncorrelated noise corruption signals w_k and v_k are, as usual, assumed to be white having Gaussian distribution with zero mean and covariances Q and R , respectively, *i.e.*,

$$\begin{aligned} w_k &\sim N(0, Q), & v_k &\sim N(0, R), \\ \langle w_k, w_l \rangle &= Q\delta_{kl}, & \langle v_k, v_l \rangle &= R\delta_{kl} \end{aligned}$$

and

$$\langle w_k, v_l \rangle = 0.$$

Lemma 1 $\langle w_k, z_l \rangle = 0, \langle v_k, z_l \rangle = 0, l \leq k$.

Proof Let $l \leq k$. Then since x_l, w_k are uncorrelated and $E\{w_k\} = 0, w_k \perp x_l$. This means that $E\{w_k|x_l\} = 0$. Hence,

$$\begin{aligned} \langle w_k, z_l \rangle &= E\{E\{w_k a(x_l, x_l)|x_l\}\} \\ &= E\{E\{w_k|x_l\}a(x_l, x_l)\} \\ &= 0. \end{aligned}$$

The second equation can be shown in exactly the same way. □

The Taylor polynomial expansion of the form $a(x, x)$ at the point x_0 can be written as follows (with $z = a(x, x)$):

$$z = z_0 + z'(x_0)(x - x_0) + \frac{1}{2} \mathcal{H}(x, x_0)(x - x_0), \tag{3}$$

where $z'(x)$ is the $\frac{n(n+1)}{2} \times n$ gradient of $a(x, x)$ given by

$$z'(x) = \left[\frac{\partial x_i x_j}{\partial x_l} \right]_{i,j,l=1,2,\dots,m},$$

and $\mathcal{H}(x, x_0)$ is given by

$$\mathcal{H}(x, x_0) = \begin{bmatrix} (x - x_0)^T D_1 \\ (x - x_0)^T D_2 \\ \vdots \\ (x - x_0)^T D_m \end{bmatrix},$$

with D_k being the matrix of second-order derivatives of the entries of $a(x, x)$ and $m = \frac{n(n+1)}{2}$. That is,

$$D_k = \begin{bmatrix} \frac{\partial z'_k}{\partial x_1} & \frac{\partial z'_k}{\partial x_1} & \dots & \frac{\partial z'_k}{\partial x_1} \\ \frac{\partial z'_k}{\partial x_2} & \frac{\partial z'_k}{\partial x_2} & \dots & \frac{\partial z'_k}{\partial x_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial z'_k}{\partial x_n} & \frac{\partial z'_k}{\partial x_n} & \dots & \frac{\partial z'_k}{\partial x_n} \end{bmatrix}, \quad \text{where } k = 1, 2, \dots, m.$$

To illustrate, suppose $n = 3$, then

$$z'(x) = \begin{bmatrix} 2x_1 & 0 & 0 \\ x_2 & x_1 & 0 \\ x_3 & 0 & x_1 \\ 0 & 2x_2 & 0 \\ 0 & x_3 & x_2 \\ 0 & 0 & 2x_3 \end{bmatrix}$$

and

$$\begin{aligned}
 D_1 &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & D_2 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & D_3 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \\
 D_4 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & D_5 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & D_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.
 \end{aligned}$$

Note, for example, that the Lorentz-96 model (with $n = 3$) takes the form (1) with $A = -I$ and

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

3 A bilinear Kalman filter and smoother

In this section, we will develop a Kalman filter and smoother for the bilinear system (1) and (2).

3.1 A bilinear Kalman filter

Given a sequence of measurements $Y_t = \{y_1, y_2, \dots, y_t\}$, let

$$x_k^t = E(x_k | Y_t) := E_t(x_k),$$

$$P_k^t = \|x_k - x_k^t\|^2,$$

$$z_k^t = E_t(z_k),$$

$$\dot{P}_k^t = \langle x_k - x_k^t, z_k - z_k^t \rangle,$$

$$\ddot{P}_k^t = \|z_k - z_k^t\|^2.$$

When $x = x_k$, equation (3) becomes

$$z = z_0 + z'(x_0)(x_k - x_0) + \frac{1}{2} \mathcal{H}(x_k, x_0)(x_k - x_0). \tag{4}$$

In order to compute equation (4), we approximate the second-degree term $\mathcal{H}(x_k, x_0)$ by using the most current available state estimation for x_k ; that is,

- In the case of prediction, we take

$$x_k \approx x_k^{k-2}, \quad \text{so } \mathcal{H}(x_k, x_0) \approx \mathcal{H}(x_k^{k-2}, x_0).$$

By setting $x_0 = x_k^{k-1}$, equation (4) becomes

$$z_k \approx z_k^{k-1} + z'(x_k^{k-1})(x_k - x_k^{k-1}) + \frac{1}{2} \mathcal{H}(x_k^{k-2}, x_k^{k-1})(x_k - x_k^{k-1}).$$

- In the case of filtering, we take

$$x_k \approx x_k^{k-1}, \quad \text{so } \mathcal{H}(x_k, x_0) \approx \mathcal{H}(x_k^{k-1}, x_0).$$

By setting $x_0 = x_k^k$, equation (4) becomes

$$z_k \approx z_k^k + z'(x_k^k)(x_k - x_k^k) + \frac{1}{2} \mathcal{H}(x_k^{k-1}, x_k^k)(x_k - x_k^k).$$

- In the case of smoothing, we take

$$x_k \approx x_k^{k+2}, \quad \text{so } \mathcal{H}(x_k, x_0) \approx \mathcal{H}(x_k^{k+2}, x_0).$$

By setting $x_0 = x_k^{k+1}$, equation (4) becomes

$$z_k \approx z_k^{k+1} + z'(x_k^{k+1})(x_k - x_k^{k+1}) + \frac{1}{2} \mathcal{H}(x_k^{k+2}, x_k^{k+1})(x_k - x_k^{k+1}).$$

In summary, we have the following linearization:

$$z_k \approx z_k^t + V_k^t(x_k - x_k^t), \tag{5}$$

where

$$V_k^t = z'(x_k^t) + \frac{1}{2} \mathcal{H}(x_k^{t+1}, x_k^t).$$

We also define

$$P_{k_1, k_2}^t = \langle x_{k_1} - x_{k_1}^t, x_{k_2} - x_{k_2}^t \rangle, \tag{6}$$

$$\dot{P}_{k_1, k_2}^t = \langle x_{k_1} - x_{k_1}^t, z_{k_2} - z_{k_2}^t \rangle, \tag{7}$$

$$\ddot{P}_{k_1, k_2}^t = \langle z_{k_1} - z_{k_1}^t, z_{k_2} - z_{k_2}^t \rangle. \tag{8}$$

Theorem 2 For the bilinear state space model defined by (1) and (2), we have

$$x_{k+1}^k = Ax_k^k + Bz_k^k, \tag{9}$$

$$P_{k+1}^k = AP_k^k A^T + A\dot{P}_k^k B^T + B(\dot{P}_k^k)^T A^T + B\ddot{P}_k^k B^T + Q, \tag{10}$$

with

$$x_{k+1}^{k+1} = x_{k+1}^k + K_{k+1}[y_k - Cx_{k+1}^k],$$

$$P_{k+1}^{k+1} = [I - K_{k+1}C]P_{k+1}^k,$$

$$\dot{P}_{k+1}^{k+1} = P_{k+1}^{k+1}[V_{k+1}^{k+1}]^T,$$

$$\ddot{P}_{k+1}^{k+1} = V_{k+1}^{k+1}\dot{P}_{k+1}^{k+1},$$

$$K_{k+1} = P_{k+1}^k C^T [CP_{k+1}^k C^T + R]^{-1},$$

and

$$V_{k+1}^{k+1} = z'(x_{k+1}^{k+1}) + \frac{1}{2} \mathcal{H}(x_{k+1}^k, x_{k+1}^{k+1}), \quad k = 0, \dots, N.$$

Proof Equation (9) is obtained by applying the conditional expectation $E_k(\cdot)$ to (1):

$$\begin{aligned} x_{k+1}^k &= E_k(x_{k+1}) \\ &= E_k(Ax_k + Bz_k + w_k) \\ &= AE_k(x_k) + BE_k(z_k) + E_k(w_k) \\ &= Ax_k^k + Bz_k^k. \end{aligned}$$

To obtain the error recursion (10), we proceed as follows:

$$\begin{aligned} P_{k+1}^k &= \|x_{k+1} - x_{k+1}^k\|^2 = \|(I - \pi_k)x_{k+1}\|^2 \\ &= \|(I - \pi_k)(Ax_k + Bz_k + w_k)\|^2 \\ &= \|A(x_k - x_k^k) + B(z_k - z_k^k) + w_k\|^2 \\ &= A\|x_k - x_k^k\|^2 A^T + B\|z_k - z_k^k\|^2 B^T + A\langle x_k - x_k^k, z_k - z_k^k \rangle B^T \\ &\quad + B\langle z_k - z_k^k, x_k - x_k^k \rangle A^T + A\langle x_k - x_k^k, w_k \rangle + \langle w_k, x_k - x_k^k \rangle A^T \\ &\quad + B\langle z_k - z_k^k, w_k \rangle + \langle w_k, z_k - z_k^k \rangle B^T + \|w_k\|^2 \\ &= AP_k^k A^T + A\dot{P}_k^k B^T + B(\dot{P}_k^k)^T A^T + B\ddot{P}_k^k B^T + Q. \end{aligned}$$

Now, when $t = k$, we derive the filtering steps. Let

$$\begin{aligned} \rho_k &= y_k - E_{k-1}(y_k) = (I - \pi_{k-1})y_k \\ &= (I - \pi_{k-1})(Cx_k + v_k) = y_k - Cx_k^{k-1} \\ &= C(x_k - x_k^{k-1}) + v_k, \quad k = 1, \dots, N. \end{aligned}$$

Then, the mean of the innovations is given by

$$E_{k-1}(\rho_k) = \pi_{k-1}(I - \pi_{k-1})y_k = 0,$$

and the variance

$$\begin{aligned} \Sigma_{k+1} &= \|\rho_{k+1}\|^2 \\ &= \|C(x_{k+1} - x_{k+1}^k) + v_{k+1}\|^2 \\ &= C\|x_{k+1} - x_{k+1}^k\|^2 C^T + \|v_{k+1}\|^2 \\ &= CP_{k+1}^k C^T + R. \end{aligned}$$

Also,

$$\begin{aligned} \langle \rho_{k+1}, y_k \rangle &= \langle y_{k+1} - y_{k+1}^k, y_k \rangle \\ &= \langle (I - \pi_k)y_{k+1}, y_k \rangle = \langle y_{k+1}, (I - \pi_k)y_k \rangle = 0, \end{aligned}$$

which means that the innovations are orthogonal to the past measurements. On the other hand,

$$\begin{aligned} \langle x_{k+1}, \rho_{k+1} \rangle &= \langle x_{k+1}, C(x_{k+1} - x_{k+1}^k) + v_{k+1} \rangle \\ &= \langle x_{k+1}, x_{k+1} - x_{k+1}^k \rangle C^T \\ &= \langle x_{k+1}, (I - \pi_k)x_{k+1} \rangle C^T \\ &= \langle (I - \pi_k)x_{k+1}, (I - \pi_k)x_{k+1} \rangle C^T \\ &= \|(I - \pi_k)x_{k+1}\|^2 C^T \\ &= P_{k+1}^k C^T. \end{aligned}$$

From these results, we conclude that x_{k+1} and ρ_{k+1} have a Gaussian joint distribution conditional on Y_k . That is,

$$\left\{ \begin{pmatrix} x_{k+1} \\ \rho_{k+1} \end{pmatrix} \middle| \{y_t\}_1^k \right\} \sim N \left\{ \begin{pmatrix} x_{k+1}^k \\ 0 \end{pmatrix}, \begin{pmatrix} P_{k+1}^k & P_{k+1}^k C^T \\ CP_{k+1}^k & \Sigma_{k+1} \end{pmatrix} \right\}.$$

Now, since Y_k, ρ_{k+1} are orthogonal,

$$\begin{aligned} x_{k+1}^{k+1} &= \pi_{k+1}(x_{k+1}) = \pi_{\{Y_k, \rho_{k+1}\}}(x_{k+1}) \\ &= \pi_{Y_k}(x_{k+1}) + \pi_{\rho_{k+1}}(x_{k+1}) \\ &= E_k(x_{k+1}) + \langle x_{k+1}, \rho_{k+1} \rangle \Sigma_{k+1}^{-1} \rho_{k+1} \\ &= x_{k+1}^k + P_{k+1}^k C^T [CP_{k+1}^k C^T + R]^{-1} \rho_{k+1} \\ &= x_{k+1}^k + K_{k+1} [y_{k+1} - Cx_{k+1}^k], \end{aligned}$$

where

$$K_{k+1} = P_{k+1}^k C^T [CP_{k+1}^k C^T + R]^{-1} = P_{k+1}^k C^T \Sigma_{k+1}^{-1}$$

represents the Kalman gain.

Next, we derive the recursion for P_{k+1}^{k+1} . Since $x_{k+1} - x_{k+1}^k = (x_{k+1} - x_{k+1}^{k+1}) + \pi_{\rho_{k+1}}(x_{k+1} - x_{k+1}^k)$ is an orthogonal decomposition,

$$\begin{aligned} P_{k+1}^{k+1} &= \|x_{k+1} - x_{k+1}^{k+1}\|^2 \\ &= \|x_{k+1} - x_{k+1}^k\|^2 - \|\pi_{\rho_{k+1}}(x_{k+1} - x_{k+1}^k)\|^2 \\ &= P_{k+1}^k - \|\langle x_{k+1} - x_{k+1}^k, \rho_{k+1} \rangle \Sigma_{k+1}^{-1} \rho_{k+1}\|^2 \\ &= P_{k+1}^k - \|P_{k+1}^k C^T \Sigma_{k+1}^{-1} \rho_{k+1}\|^2 \\ &= P_{k+1}^k - P_{k+1}^k C^T \Sigma_{k+1}^{-1} \|\rho_{k+1}\|^2 \Sigma_{k+1}^{-1} CP_{k+1}^k \\ &= P_{k+1}^k - P_{k+1}^k C^T \Sigma_{k+1}^{-1} CP_{k+1}^k \\ &= P_{k+1}^k - K_{k+1} CP_{k+1}^k \\ &= [I - K_{k+1} C] P_{k+1}^k. \end{aligned}$$

The equation for \dot{P}_{k+1}^{k+1} is obtained as follows:

$$\begin{aligned} \dot{P}_{k+1}^{k+1} &= \langle x_{k+1} - x_{k+1}^{k+1}, z_{k+1} - z_{k+1}^{k+1} \rangle \\ &= \langle x_{k+1} - x_{k+1}^{k+1}, V_{k+1}^{k+1} (x_{k+1} - x_{k+1}^{k+1}) \rangle \\ &= \langle x_{k+1} - x_{k+1}^{k+1}, x_{k+1} - x_{k+1}^{k+1} [V_{k+1}^{k+1}]^T \rangle \\ &= P_{k+1}^{k+1} [V_{k+1}^{k+1}]^T. \end{aligned}$$

Finally, for \ddot{P}_{k+1}^{k+1} we have

$$\begin{aligned} \ddot{P}_{k+1}^{k+1} &= \langle z_{k+1} - z_{k+1}^{k+1}, z_{k+1} - z_{k+1}^{k+1} \rangle \\ &= V_{k+1}^{k+1} \langle x_{k+1} - x_{k+1}^{k+1}, x_{k+1} - x_{k+1}^{k+1} [V_{k+1}^{k+1}]^T \rangle \\ &= V_{k+1}^{k+1} P_{k+1}^{k+1} [V_{k+1}^{k+1}]^T \\ &= V_{k+1}^{k+1} \dot{P}_{k+1}^{k+1}. \end{aligned}$$

This completes the proof. □

We summarize the bilinear Kalman filter as follows:

$$\begin{aligned} x_k^{k-1} &= Ax_{k-1}^{k-1} + Bz_{k-1}^{k-1}, \\ P_k^{k-1} &= AP_{k-1}^{k-1}A^T + A\dot{P}_{k-1}^{k-1}B^T + B(\dot{P}_{k-1}^{k-1})^T A^T + B\ddot{P}_{k-1}^{k-1}B^T + Q, \\ x_k^k &= x_k^{k-1} + K_k[y_k - Cx_k^{k-1}], \end{aligned} \tag{11}$$

$$P_k^k = [I - K_k C]P_k^{k-1}, \tag{12}$$

$$\dot{P}_k^k = P_k^k [V_k^k]^T,$$

$$\ddot{P}_k^k = V_k^k \dot{P}_k^k,$$

$$K_k = P_k^{k-1} C^T [CP_k^{k-1} C^T + R]^{-1},$$

and

$$V_k^k = z'(x_k^k) + \frac{1}{2} \mathcal{H}(x_k^{k-1}, x_k^k), \quad k = 0, \dots, N.$$

Also, note that the bilinear Kalman filter algorithm is a generalization of the Kalman filter for the linear case which is given in [17].

3.2 A bilinear Kalman smoother

In this subsection, we will develop a Kalman smoother for the bilinear system (1) and (2).

We will use the following notation:

$$P_{k_1, k_2}^N = \langle x_{k_1} - x_{k_1}^N, x_{k_2} - x_{k_2}^N \rangle,$$

$$\dot{P}_{k_1, k_2}^N = \langle x_{k_1} - x_{k_1}^N, z_{k_2} - z_{k_2}^N \rangle,$$

$$\ddot{P}_{k_1, k_2}^N = \langle z_{k_1} - z_{k_1}^N, z_{k_2} - z_{k_2}^N \rangle.$$

Lemma 3 *Let*

$$\epsilon_{k+1} = \{v_{k+1}, \dots, v_N, w_{k+2}, \dots, w_N\}. \tag{13}$$

Then for $1 \leq k \leq N - 1$ and with the approximation (5),

$$L\{y_m\}_1^N = L\{y_m\}_1^k, x_{k+1} - x_{k+1}^k, \epsilon_{k+1}\}, \tag{14}$$

where $L\{\cdot\}$ denotes the subspace spanned by $\{\cdot\}$.

Proof Recall that

$$z_m = z_m^N + V_m^N(x_m - x_m^N),$$

that is,

$$(z_m - z_m^N) \in L\{x_m - x_m^N\}.$$

Since

$$\begin{aligned} y_m - y_m^N &= C(x_m - x_m^N) + v_m, \\ L\{y_m\}_1^N &= L\{y_m\}_1^{N-1}, y_N\} \\ &= L\{y_m\}_1^{N-1}, y_N - y_N^{N-1}\} \\ &= L\{y_m\}_1^{N-1}, x_N - x_N^{N-1}, v_N\}. \end{aligned}$$

Similarly, since

$$\begin{aligned} y_{m+2} - y_{m+2}^N &= CA(x_{m+1} - x_{m+1}^N) + CB(z_{m+1} - z_{m+1}^N) + Cw_{m+1} + v_{m+2}, \\ L\{y_m\}_1^N &= L\{y_m\}_1^{N-1}, y_N\} \\ &= L\{y_m\}_1^{N-1}, y_N - y_N^{N-1}\} \\ &= L\{y_m\}_1^{N-1}, x_{N-1} - x_{N-1}^{N-2}, v_{N-1}, z_{N-1} - z_{N-1}^{N-2}, w_{N-1}, v_N\} \\ &= L\{y_m\}_1^{N-1}, x_{N-1} - x_{N-1}^{N-2}, v_{N-1}, v_N, w_{N-1}\}. \end{aligned}$$

Continuing in this manner, we get (14). □

We state the bilinear Kalman smoother in the following theorem.

Theorem 4 *Consider the bilinear state space model (1) and (2) with x_N^N and P_N^N as given in (11) and (12). Then for $k = N - 1, \dots, 1$, we have*

$$x_k^N = x_k^k + J_k(x_{k+1}^N - x_{k+1}^k), \tag{15}$$

$$P_k^N = P_k^k - J_k P_{k+1}^k J_k^T, \tag{16}$$

where

$$J_k = [P_k^k A^T + \dot{P}_k^k B^T][P_{k+1}^k]^{-1}.$$

Proof Noting the mutual orthogonality of $\{y\}_1^k$, $\{x_{k+1} - x_{k+1}^k\}$ and ϵ_{k+1} and the orthogonality of x_k and ϵ_{k+1} ,

$$\begin{aligned} x_k^N &= \pi_N x_k = \pi_k x_k + \pi_{(x_{k+1} - x_{k+1}^k)} x_k \\ &= x_k^k + \langle x_k, x_{k+1} - x_{k+1}^k \rangle \|x_{k+1} - x_{k+1}^k\|^{-2} (x_{k+1} - x_{k+1}^k) \\ &= x_k^k + \langle x_k, x_{k+1} - x_{k+1}^k \rangle [P_{k+1}^k]^{-1} (x_{k+1} - x_{k+1}^k). \end{aligned}$$

Now,

$$\begin{aligned} \langle x_k, x_{k+1} - x_{k+1}^k \rangle &= \langle x_k, Ax_k + Bz_k + w_k - x_{k+1}^k \rangle \\ &= \langle x_k, A(x_k - x_k^k) + B(z_k - z_k^k) \rangle \\ &= \langle x_k - x_k^k, x_k - x_k^k \rangle A^T + \langle x_k - x_k^k, z_k - z_k^k \rangle B^T \\ &= P_k^k A^T + \dot{P}_k^k B^T. \end{aligned}$$

Thus,

$$\begin{aligned} x_k^N &= x_k^k + [P_k^k A^T + \dot{P}_k^k B^T][P_{k+1}^k]^{-1} (x_{k+1} - x_{k+1}^k) \\ &= x_k^k + J_k (x_{k+1} - x_{k+1}^k). \end{aligned}$$

Equation (15) now follows by taking the projection π_N again of both sides and noting that $k \leq N$. To derive (16), we compute

$$\begin{aligned} P_k^N &= \|x_k - x_k^N\|^2 = \|x_k - x_k^k - J_k (x_{k+1} - x_{k+1}^k)\|^2 \\ &= \|x_k - x_k^k\|^2 - \langle x_k - x_k^k, x_{k+1} - x_{k+1}^k \rangle J_k^T - J_k \langle x_{k+1} - x_{k+1}^k, x_k - x_k^k \rangle + J_k P_{k+1}^k J_k^T \\ &= P_k^k - \langle (1 - \pi_k)x_k, x_{k+1} \rangle J_k^T - J_k \langle x_{k+1}, (1 - \pi_k)x_k \rangle + J_k P_{k+1}^k J_k^T \\ &= P_k^k - \langle (1 - \pi_k)x_k, Ax_k + Bz_k \rangle J_k^T - J_k \langle Ax_k + Bz_k, (1 - \pi_k)x_k \rangle + J_k P_{k+1}^k J_k^T \\ &= P_k^k - (P_k^k A^T + \dot{P}_k^k B^T) J_k^T - J_k (AP_k^k + B\dot{P}_k^k) + J_k P_{k+1}^k J_k^T \\ &= P_k^k - J_k P_{k+1}^k J_k^T - J_k P_{k+1}^k J_k^T + J_k P_{k+1}^k J_k^T = P_k^k - J_k P_{k+1}^k J_k^T, \end{aligned}$$

which completes the proof. □

The next theorem states the bilinear lag-one recursions.

Theorem 5 Consider the bilinear state space model (1) and (2). Then

$$\begin{aligned} P_{k+1,k}^N &= AP_k^N + B(\dot{P}_k^N)^T, \\ \dot{P}_{k+1,k}^N &= P_{k+1,k}^N [V_k^N]^T. \end{aligned}$$

Proof Using the definitions in (6) and (7),

$$\begin{aligned} P_{k+1,k}^N &= \langle x_{k+1} - x_{k+1}^N, x_k - x_k^N \rangle = \langle (1 - \pi_N)x_{k+1}, (1 - \pi_N)x_k \rangle \\ &= \langle x_{k+1}, (1 - \pi_N)x_k \rangle = \langle Ax_k + Bz_k + w_k, (1 - \pi_N)x_k \rangle \\ &= A \langle x_k, (1 - \pi_N)x_k \rangle + B \langle z_k, (1 - \pi_N)x_k \rangle \\ &= AP_k^N + B(\dot{P}_k^N)^T. \end{aligned}$$

Also,

$$\begin{aligned} \dot{P}_{k+1,k}^N &= \langle x_{k+1} - x_{k+1}^N, z_k - z_k^N \rangle \\ &= \langle x_{k+1} - x_{k+1}^N, V_k^N(x_k - x_k^N) \rangle \\ &= \langle x_{k+1} - x_{k+1}^N, x_k - x_k^N \rangle [V_k^N]^T \\ &= P_{k+1,k}^N [V_k^N]^T. \end{aligned} \quad \square$$

4 The bilinear EM algorithm

The unknown parameter set $\theta = \{A, B, C, Q, R, V, \mu\}$ is estimated by the EM algorithm that iteratively updates the current estimate $\theta(i)$ of θ by maximizing the log-likelihood function

$$\begin{aligned} &\log \mathcal{L}(\theta, X_N, Y_N) \\ &= \operatorname{argmin}_{\theta} \left\{ \log f_0(x_0 - \mu) + \sum_{k=1}^N \log (f_w(x_k - Ax_{k-1} - Bz_{k-1}) \times f_v(y_k - Cx_k)) \right\}, \end{aligned} \quad (17)$$

where

- $f_0(\cdot)$ represents the n -variate normal density of the initial state x_0 with mean μ and the covariance matrix V .
- $f_v(\cdot)$ represents the p -variate normal density with zero mean and the covariance matrix R .
- $f_w(\cdot)$ represents the n -variate normal density function with zero mean and the covariance matrix Q .

The conditional expectation step (E-step) finds the missing data, *i.e.*, X_N , given the observed data and current estimated parameters, and then substitutes these expectations for the missing data. Specifically, let $\theta(i-1)$ be the current estimate of the parameter θ , then the E-step finds the conditional expectation $E\{\cdot\}$ of the complete-data log-likelihood given $\theta(i-1)$:

$$q(\theta|\theta(i-1)) = E\{\log \mathcal{L}(\theta, X_N, Y_N) | Y_N, \theta(i-1)\}. \quad (18)$$

The M-step determines $\theta(i)$ by maximizing the expected complete-data log-likelihood

$$q(\theta(i)|\theta(i-1)) \geq q(\theta|\theta(i-1)), \quad \forall \theta.$$

The following theorem accomplishes the expectation step.

Theorem 6 For the bilinear state space model (1) and (2),

$$\begin{aligned}
 q(\theta(i)|\theta(i-1)) = & -\frac{1}{2} \log |V| - \frac{1}{2} \text{Tr}\{V^{-1}(\Delta - \widehat{x}_0\mu^T - \mu\widehat{x}_0^T + \mu\mu^T)\} \\
 & - \frac{N}{2} \log |Q| - \frac{1}{2} \text{Tr}\{Q^{-1}(\Theta - \Psi A^T - \Pi B^T - A\Psi^T \\
 & + A\Phi A^T - B\Pi^T + B\Lambda B^T)\} \\
 & - \frac{N}{2} \log |R| - \frac{1}{2} \text{Tr}\{R^{-1}(\delta - \Omega C^T - C\Omega^T + C\Phi C^T)\} + \text{const},
 \end{aligned}$$

where

$$\begin{aligned}
 \Delta &= E_N(x_0 x_0^T), & \widehat{x}_0 &= E_N(x_0), \\
 \Theta &= \sum_{k=1}^N (x_k^N (x_k^N)^T + P_k^N), & \Psi &= \sum_{k=1}^N (x_k^N (x_{k-1}^N)^T + P_{k,k-1}^N), \\
 \Pi &= \sum_{k=1}^N (x_k^N (z_{k-1}^N)^T + \dot{P}_{k,k-1}^N), & \Phi &= \sum_{k=1}^N (x_{k-1}^N (x_{k-1}^N)^T + P_{k-1}^N), \\
 \Gamma &= \sum_{k=1}^N E_N(x_{k-1} z_{k-1}^T), & \Lambda &= \sum_{k=0}^{N-1} (z_k^N (z_k^N)^T + \ddot{P}_k^N), \\
 \Omega &= \sum_{k=1}^N x_k^N y_k^T, & \delta &= \sum_{k=1}^N \|y_k\|^2.
 \end{aligned}$$

Proof Since the system is Markovian, we may use Bayes' rule successively to get

$$p(\theta, X_N, Y_N) = p(y_1, \dots, y_N, x_0, \dots, x_N) \tag{19}$$

$$= p(x_0) \prod_{k=1}^N p(y_k | x_k) \prod_{k=0}^{N-1} p(x_{k+1} | x_k). \tag{20}$$

From the assumptions on x_0 , w_k , and v_k , the density functions $p(x_0)$, $p(y_k | x_k)$, and $p(x_{k+1} | x_k)$ are given by

$$\begin{aligned}
 p(x_0) &= \frac{1}{(2\pi)^{\frac{n}{2}} |V|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x_0 - \mu)^T V^{-1}(x_0 - \mu)\right\}, \\
 p(x_{k+1} | x_k) &= \frac{1}{(2\pi)^{\frac{n}{2}} |Q|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x_{k+1} - Ax_k - Bz_k)^T Q^{-1}(x_{k+1} - Ax_k - Bz_k)\right\},
 \end{aligned}$$

and

$$p(y_k | x_k) = \frac{1}{(2\pi)^{\frac{m}{2}} |R|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(y_k - Cx_k)^T R^{-1}(y_k - Cx_k)\right\}.$$

Now, substituting these densities in (17) and taking the logarithm of both sides, we get

$$\mathcal{L}(\theta, X_N, Y_N) = -\frac{1}{2} \log |V| - \frac{1}{2}(x_0 - \mu)^T V^{-1}(x_0 - \mu)$$

$$\begin{aligned}
 & -\frac{N}{2} \log |Q| - \frac{1}{2} \sum_{k=0}^{N-1} (x_{k+1} - Ax_k - Bz_k)^T Q^{-1} (x_{k+1} - Ax_k - Bz_k) \\
 & -\frac{N}{2} \log |R| - \frac{1}{2} \sum_{k=1}^N (y_k - Cx_k)^T R^{-1} (y_k - Cx_k) + \text{const.}
 \end{aligned}$$

The result follows upon taking the expectation conditional on Y_N , making use of

$$\begin{aligned}
 E_N(x^T Ay) &= \text{Tr}[AE_N(xy^T)], \\
 E_N(x_k z_k^T) &= 0, \\
 E_N(x_{k+1} z_k^T) &= BE_N(z_k z_k^T),
 \end{aligned}$$

and simplifying. The middle equality follows from the fact that odd moments of Gaussian random variables vanish. \square

The computation of Θ , Φ , Ψ , Π and Λ given a current estimate $\theta(i)$ of θ involves the use of the bilinear Kalman filter and smoother introduced in Sections 3.1 and 3.2. For this purpose, we introduce

$$\begin{aligned}
 \Theta &= \sum_{k=1}^N E_t(x_k x_k^T) = \sum_{k=1}^N (x_k^t (x_k^t)^T + P_k^t), \\
 \Phi &= \sum_{k=1}^N (x_{k-1}^t (x_{k-1}^t)^T + P_{k-1}^t), & \Psi &= \sum_{k=1}^N (x_k^t (x_{k-1}^t)^T + P_{k,k-1}^t), \\
 \Pi &= \sum_{k=1}^N (x_k^t (z_{k-1}^t)^T + \dot{P}_{k,k-1}^t), & \Lambda &= \sum_{k=0}^{N-1} (z_k^t (z_k^t)^T + \ddot{P}_k^t).
 \end{aligned}$$

The next step of the EM algorithm is to maximize the function $q(\theta(i)|\theta(i-1))$ with respect to θ .

Theorem 7 *The maximizer of $q(\theta(i)|\theta(i-1))$ is obtained for the parameter vector θ given by*

$$\mu = \hat{x}_0, \quad V = \Delta, \quad A = \Psi \Phi^{-1}, \tag{21}$$

$$B = \Pi \Lambda^{-1}, \quad Q = \frac{1}{N} (\Theta - \Psi^T \Phi^{-1} \Psi - \Pi \Lambda^{-1} \Pi^T), \tag{22}$$

$$C = \Omega \Phi^{-1}, \quad R = N^{-1} \{ \delta - \Omega \Phi^{-1} \Omega^T \}. \tag{23}$$

Proof Let

$$\begin{aligned}
 q_1(\mu, V) &= -\frac{1}{2} \log |V| - \frac{1}{2} \text{Tr} \{ V^{-1} (\Delta - \hat{x}_0 \mu^T - \mu \hat{x}_0^T + \mu \mu^T) \}, \\
 q_2(A, B, Q) &= -\frac{N}{2} \log |Q| - \frac{1}{2} \text{Tr} \{ Q^{-1} (\Theta - \Psi A^T - \Pi B^T - A \Psi^T \\
 & \quad + A \Phi A^T - B \Pi^T + B \Lambda B^T) \}, \\
 q_3(C, R) &= -\frac{N}{2} \log |R| - \frac{1}{2} \text{Tr} \{ R^{-1} (\delta - \Omega C^T - C \Omega^T + C \Phi C^T) \}.
 \end{aligned}$$

Then

$$q(\theta(i)|\theta(i-1)) = q_1(\mu, V) + q_2(A, B, Q) + q_3(C, R) + \text{const},$$

which means that $q(\theta(i), \theta(i-1))$ is maximized by separately minimizing q_1, q_2, q_3 . This is done by setting the partial derivative of q with respect to each parameter equal to zero (i.e., $\frac{\partial q}{\partial x} = 0$) and solving the resulting system of equations. \square

The EM algorithm for a bilinear state space model is summarized as follows.

Bilinear EM algorithm

1. Initialize the EM algorithm by choosing initial values of $\theta(0)$.
2. Calculate the incomplete-data likelihood, $\log L(Y_n; \theta)$.
3. Execute the E-step by using the bilinear Kalman filter and smoother in (9)-(10) and (15)-(16), respectively.
4. Execute the M-step using (21)-(23) and update the estimates of θ using (M-step) to obtain $\theta(i)$.
5. Repeat Steps 2 to 4 until convergence.

5 Simulation results

A 1,000 Monte Carlo simulation is performed to illustrate the utility of the bilinear algorithm. The observed data are generated according to the second-order bilinear state space model

$$x_{k+1} = \begin{bmatrix} 0.4 & 0.1 \\ -0.1 & 0.2 \end{bmatrix} x_k + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} z_k + w_k, \quad k = 1, \dots, 1,000, \tag{24}$$

$$y_k = \begin{bmatrix} 0 & 1 \end{bmatrix} x_k + v_k, \tag{25}$$

where w_k and v_k are independent identically distributed (i.i.d.) Gaussian noises such that

$$w_k \sim N(0, 0.01 \times I_2),$$

$$v_k \sim N(0, 0.01).$$

In all simulations, the number of iterations for the EM algorithm is fixed and its value set to $J = 100$.

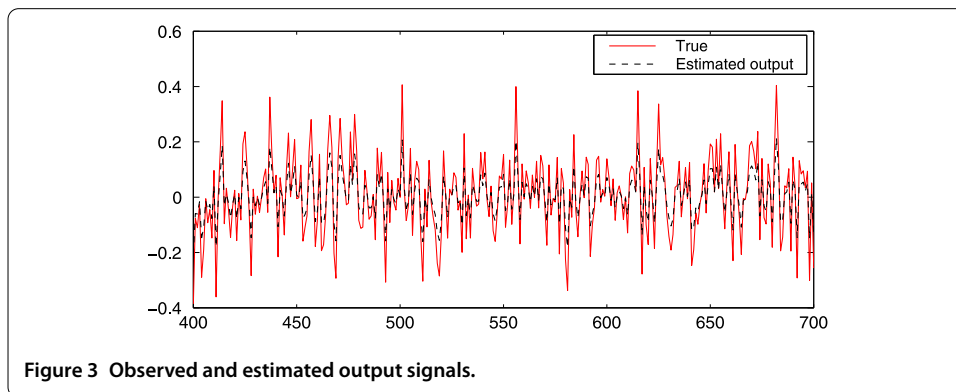
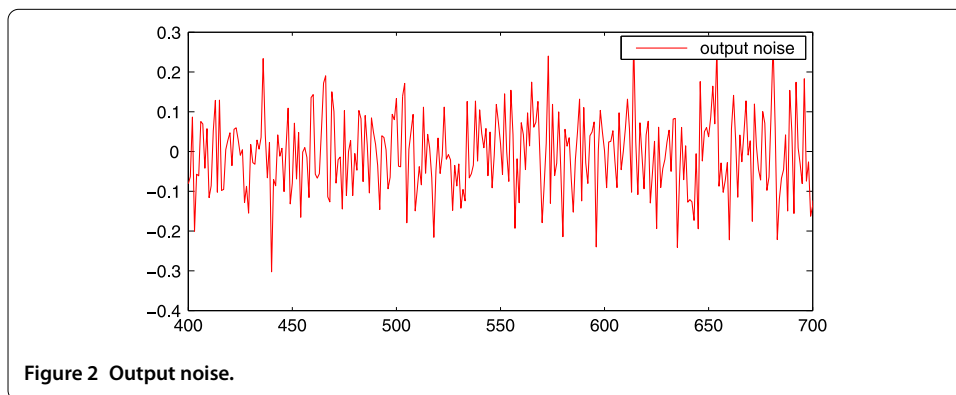
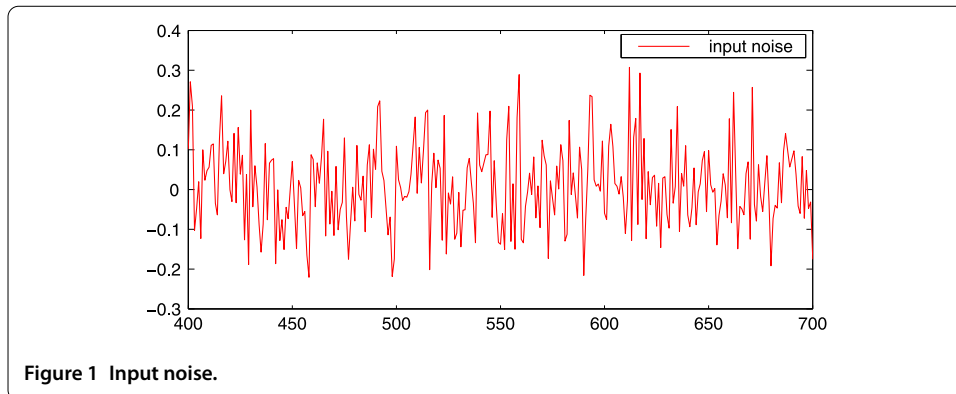
Figure 1 shows a sample of realizations of the input noise w_k , and Figure 2 shows the output noise v_k , respectively. Figure 3 compares the observed output signals and the estimated output signal. The average estimates of the parameters are

$$A = \begin{bmatrix} 0.3891 & 0.1015 \\ -0.09812 & 0.2117 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0012 & 1.1016 & 0 \\ 0 & -0.0345 & 0.9761 \end{bmatrix},$$

$$C = \begin{bmatrix} -0.0927 & 1.014 \end{bmatrix}.$$

The mean square error (MSE) is defined as

$$E_N = \frac{1}{N} \sum_{k=1}^N (y_k - Cx_{k|k-1})^2,$$



and its value for 1,000 run for different values of $\text{cov}(R)$ and $\text{cov}(Q)$ is kept constant, which is shown in Table 1.

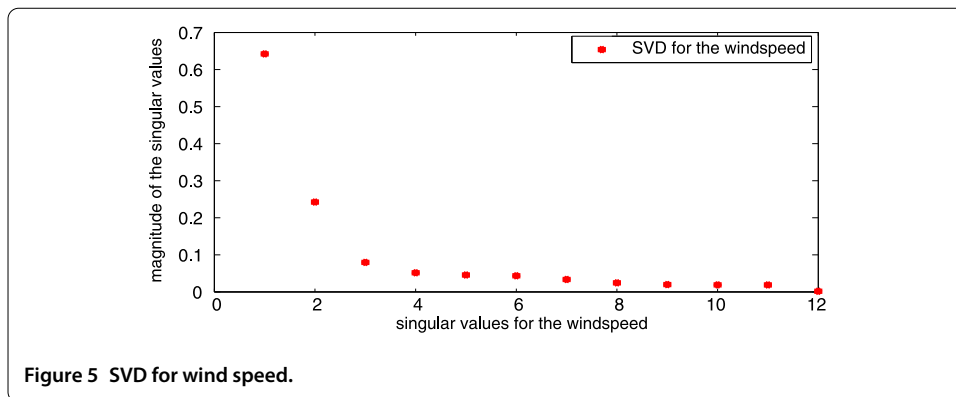
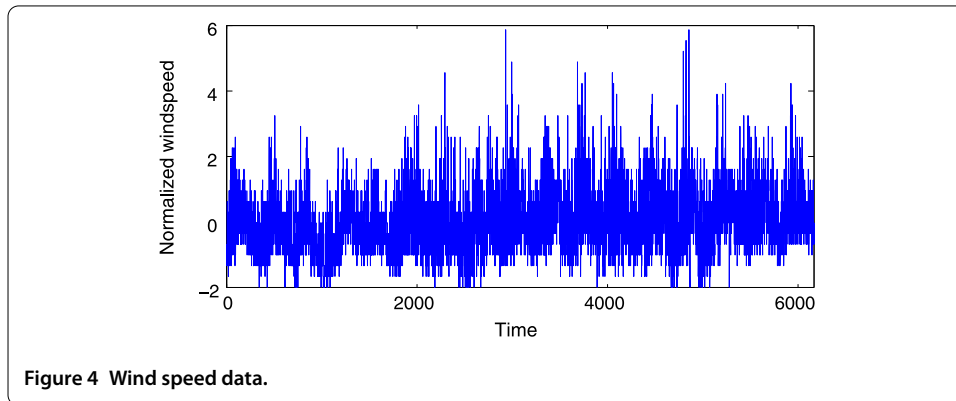
6 Application to wind speed

In this section, we apply the proposed bilinear algorithm to the daily averaged wind speed data for Arar, a city located in the north eastern region of the kingdom of Saudi Arabia for a period of $16\frac{1}{2}$ years as shown in Figure 4. It should be noted that all the calculations are carried out on normalized time series data.

To estimate the dimension of the state in the state space model, we apply the stochastic subspace system identification algorithm described in [23]. This is done by constructing the singular value diagram of the block Hankel matrix for the normalized wind speed

Table 1 Comparison of the mean square errors

cov(<i>R</i>)	Linear MSE [3]	Bilinear MSE
0.01	0.39988	0.049985
0.02	0.58788	0.080165
0.04	0.97789	0.104683



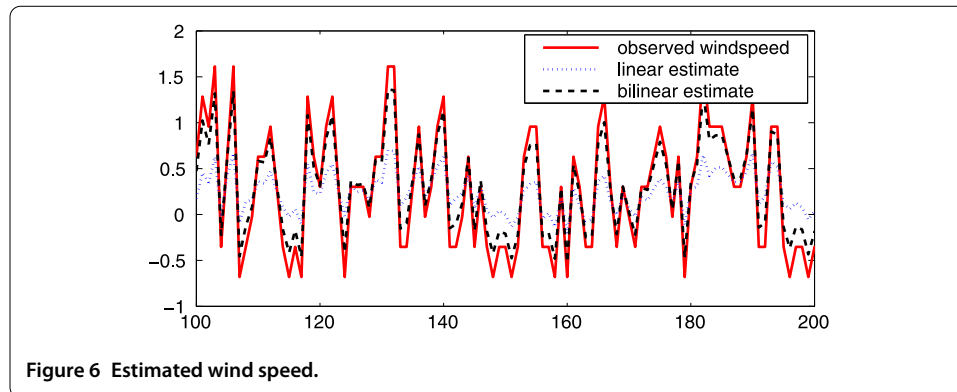
data as shown in Figure 5. That is, the dimension of the state equals the number of the significant singular values; here $n = 2$. For clarity, we compare the observed wind speed values with the estimated ones using a linear model [3] with our proposed algorithm for a period of 100 days as shown in Figure 6. The estimated parameters for the linear state space model are

$$A = \begin{bmatrix} 0.9475 & 0.1991 \\ 0.1968 & 0.0935 \end{bmatrix}, \quad C = \begin{bmatrix} 0.5355 & 0.3492 \end{bmatrix},$$

and for the bilinear state space model, they are

$$A = \begin{bmatrix} 0.9475 & 0.1991 \\ 0.1968 & 0.0935 \end{bmatrix}, \quad C = \begin{bmatrix} 0.5355 & 0.3492 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 1.092 & -1.092 \\ 0.01092 & 0.0557 & 0.1092 \end{bmatrix}.$$



The MSE for the estimated wind speed data using the linear EM algorithm is 0.3864, and 0.0197 for the bilinear EM algorithm.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors have achieved equal contributions to each part of this paper. All authors read and approved the final version of the manuscript.

Author details

¹Department of Mathematics, Taibah University, Al-Madinah, Kingdom of Saudi Arabia. ²Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Kingdom of Saudi Arabia. ³Department of Mathematics, Texas A&M University-Kingsville, Kingsville, TX 78363-8202, USA.

Acknowledgements

The first author was supported by Tayyebah University. The second and third authors would like to thank King Fahd University for the excellent research facilities they provide.

Received: 22 April 2012 Accepted: 24 September 2012 Published: 9 October 2012

References

1. Krener, AJ: Bilinear and nonlinear realizations of input-output maps. *SIAM J. Control* **13**(4), 827-834 (1975)
2. Pardalos, PM, Yatsenko, V: *Optimization and Control of Bilinear Systems: Theory, Algorithms and Applications*. Springer, Berlin (2008)
3. Shumway, R, Stoffer, D: An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Ser. Anal.* **3**(4), 253-264 (1982)
4. Strogatz, SH: *Nonlinear Dynamics and Chaos, with Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Books, New York (1994)
5. Galanis, G, Anadranistakis, M: A one-dimension Kalman filter for correction of near surface temperature forecasts. *Meteorol. Appl.* **9**, 437-441 (2002)
6. Goel, NS: *On the Volterra and Other Nonlinear Models of Interacting Populations*. Academic Press, San Diego (1971)
7. Lorenz, EN, Emanuel, KE: Optimal sites for supplementary weather observations: simulations with a small model. *J. Atmos. Sci.* **55**, 399-414 (1998)
8. Lorenz, EN: Designing chaotic models. *J. Atmos. Sci.* **62**, 1574-1588 (2005)
9. Monbet, V, Ailliot, P, Prevosto, M: Survey of stochastic models for wind and sea state time series. *Probab. Eng. Mech.* **22**, 113-126 (2007)
10. Roy, D, Musielak, ZE: Generalized Lorenz models and their routes to chaos. III. Energy-conserving horizontal and vertical mode truncations. *Chaos Solitons Fractals* **33**, 1064-1070 (2007)
11. Priestley, MB: *Non-Linear and Non-Stationary Time Series Analysis*. Academic Press, San Diego (1989)
12. Gibson, S, Wills, A, Ninness, B: Maximum-likelihood parameter estimation of bilinear systems. *IEEE Trans. Autom. Control* **50**, 1581-1596 (2005)
13. Anderson, JL, Anderson, SL: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.* **127**, 2741-2758 (1999)
14. Bendat, J: *Nonlinear System Analysis and Identification from Random Data*. Wiley-Interscience, New York (1990)
15. Daum, FE: Exact finite dimensional nonlinear filters. *IEEE Trans. Autom. Control* **31**(7), 616-622 (1986)
16. Ha, QP, Trinh, H: State and input simultaneous estimation for a class of nonlinear system. *Automatica* **40**, 1779-1785 (2004)
17. Kailath, T, Sayed, A, Hassibi, B: *Linear Estimation*. Prentice Hall, New York (2000)
18. Kerschen, G, Worden, K, Vakakis, AF, Golinval, J: Past, present and future of nonlinear system identification in structural dynamics. *Mech. Syst. Signal Process.* **20**, 505-592 (2006)
19. Ljung, L: *System Identification: Theory for the User*, 2nd edn. Prentice Hall, Upper Saddle River (1999)

20. Norgaard, M, Poulsen, NK, Ravn, O: New developments in state estimation for nonlinear systems. *Automatica* **36**, 1627-1638 (2000)
21. Wiener, N: *Nonlinear Problems in Random Theory*. MIT Press, Boston (1958)
22. Dempster, A, Laird, N, Rubin, D: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **39**, 1-38 (1977)
23. Tanaka, H, Katayama, T: A stochastic realization algorithm via block LQ decomposition in Hilbert space. *Automatica* **42**, 741-746 (2006)

doi:10.1186/1687-1847-2012-176

Cite this article as: Al-Mazrooei et al.: **Filtering and identification of a state space model with linear and bilinear interactions between the states.** *Advances in Difference Equations* 2012 2012:176.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
