



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Statistics & Probability Letters 73 (2005) 71–78

**STATISTICS &  
PROBABILITY  
LETTERS**

[www.elsevier.com/locate/stapro](http://www.elsevier.com/locate/stapro)

# A note on prediction and interpolation errors in time series

Pedro Galeano, Daniel Peña\*

*Departamento de Estadística, Universidad Carlos III de Madrid, C/Madrid 126, Getafe, 28903 Madrid, Spain*

Received 28 July 2004

Available online 8 April 2005

---

## Abstract

In this note, we analyze the relationship between one-step ahead prediction errors and interpolation errors in time series. We obtain an expression of the prediction errors in terms of the interpolation errors and then we show that minimizing the sum of squares of the one-step ahead standardized prediction errors is equivalent to minimizing the sum of squares of standardized interpolation errors.

© 2005 Elsevier B.V. All rights reserved.

*Keywords:* Fixed-point smoothing; Interpolation error; Kalman filter; Prediction error

---

## 1. Introduction

It is well known that the likelihood function of an ARMA( $p, q$ ) process can be written in terms of the one-step ahead prediction errors using the conditional distribution of each observation given the previous data. This is called the prediction error decomposition. The maximum likelihood estimate (MLE) of the parameters can be computed by minimizing the concentrated likelihood function, which depends on the one-step ahead prediction errors. The interpolation problem consists in the estimation of a missing observation by using the past and future values of the time series. The interpolator which minimizes the mean-squared error criterion is computed by the expected value of the observation given the rest of the sample. The interpolation error is the difference between the interpolated value and the true value of the observation. In the state-space form of ARMA models, the interpolator is obtained with some smoothing algorithm, such as the

---

\*Corresponding author.

*E-mail address:* [daniel.pena@uc3m.es](mailto:daniel.pena@uc3m.es) (D. Peña).

fixed point smoothing (FPS) (see Anderson and Moore, 1979). The aim of this note is to show the relationship between prediction errors and interpolation errors and to prove that the parameter values which minimize the mean squared prediction error are the same as those which minimize the mean-squared interpolation errors. This note is organized as follows. In Section 2 we introduce the notation and briefly review the FPS algorithm. In Section 3, we first obtain an expression of the one-step ahead prediction error in terms of the interpolation errors, second we derive the covariances between interpolation errors and third we show that minimizing the sum of squares of the one-step ahead standardized prediction errors leads to the same result as minimizing the sum of squares of the standardized interpolation errors. Section 4 illustrates the result in the simplest case of a first-order autoregressive process.

## 2. Kalman filter and fixed point smoothing

Let  $\{z_t\}$  be a process following a zero mean stationary and invertible ARMA( $p, q$ ) model,

$$\phi(B)z_t = \theta(B)u_t, \quad (1)$$

where  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ ,  $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  and  $\{u_t\}$  is a sequence of independent  $N(0, \sigma^2)$  variables. We denote the vector of ARMA parameters in (1) by  $\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$  and a sample generated by this process by  $z = (z_1, \dots, z_T)'$ , where  $T$  is the sample size. Let  $\Sigma_z$  be the covariance matrix of  $z$ , then the likelihood function is

$$L(z|\beta, \sigma^2) = (2\pi)^{-T/2} |\Sigma_z|^{-1/2} \exp\left(-\frac{z' \Sigma_z^{-1} z}{2}\right).$$

Let  $z_{t|t-1} = E[z_t | z_{t-1}, \dots, z_1]$  for  $t = 1, \dots, T$ , be the one-step ahead predictions obtained by minimizing the mean squared errors, where  $z_{1|0} = E[z_1]$ , and let  $e_t = z_t - z_{t|t-1}$  be the corresponding one-step ahead prediction errors with variances  $E[(z_t - z_{t|t-1})^2] = \sigma^2 v_{t|t-1}^2$ , and where  $\text{var}(z_1) = \sigma^2 v_{1|0}^2$ . The log-likelihood,  $\ell(z|\beta, \sigma^2) = \log L(z|\beta, \sigma^2)$ , can be written as

$$\ell(z|\beta, \sigma^2) = -\frac{T}{2} \log 2\pi\sigma^2 - \frac{1}{2} \sum_{t=1}^T \log v_{t|t-1}^2 - \frac{1}{2\sigma^2} \sum_{t=1}^T \frac{e_t^2}{v_{t|t-1}^2}$$

and the MLE of  $\sigma^2$  is given by

$$\hat{\sigma}_{\text{MLE}}^2 = \frac{1}{T} \sum_{t=1}^T \frac{e_t^2}{v_{t|t-1}^2} \quad (2)$$

and, using (2), the MLE of  $\beta$ ,  $\hat{\beta}_{\text{MLE}}$ , maximizes the concentrated log-likelihood given by

$$S(\beta) = \frac{1}{T} \sum_{t=1}^T \log v_{t|t-1}^2 + \log \left( \sum_{t=1}^T \frac{e_t^2}{v_{t|t-1}^2} \right).$$

The state-space representation for ARMA( $p, q$ ) models proposed by Jones (1980) is obtained by defining  $r = \max\{p, q + 1\}$ , with

$$\begin{aligned} z_t &= H'x_t, \\ x_t &= Fx_{t-1} + Gu_t, \end{aligned} \tag{3}$$

where  $H = (1, 0, \dots, 0)'$ ,  $x_t = (z_t, z_{t+1|t}, \dots, z_{t+r-1|t})'$ ,  $G = (1, \psi_1, \dots, \psi_{r-1})'$ , and

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ \phi_r & \phi_{r-1} & \phi_{r-2} & \cdots & \phi_1 \end{bmatrix},$$

where  $\psi(B) = \phi(B)^{-1}\theta(B) = 1 + \sum_{i=1}^{\infty} \psi_i B^i$ , and  $z_{t+j|t} = E[z_{t+j}|z_t, \dots, z_1]$  with  $E[(z_{t+j} - z_{t+j|t})^2] = \sigma^2 v_{t+j|t}^2$ . With this representation, the Kalman filter computes the log-likelihood through the recursions

$$\begin{aligned} x_{t|t-1} &= Fx_{t-1|t-1}, \\ \Sigma_{t|t-1} &= F\Sigma_{t-1|t-1}F' + Q, \\ K_t &= \Sigma_{t|t-1}H(v_{t|t-1}^2)^{-1}, \\ x_{t|t} &= x_{t|t-1} + K_t(z_t - z_{t|t-1}), \\ \Sigma_{t|t} &= (I - K_tH')\Sigma_{t|t-1}, \end{aligned} \tag{4}$$

for  $t = 1, \dots, T$ , where  $z_{t|t-1} = H'x_{t|t-1}$ ,  $v_{t|t-1}^2 = H'\Sigma_{t|t-1}H$ ,  $Q = GG'$  and,

$$\begin{aligned} x_{t|s} &= E[x_t|z_1, \dots, z_s], \\ \sigma^2 \Sigma_{t|s} &= cov[x_t|z_1, \dots, z_s]. \quad s, t = 1, \dots, T. \end{aligned}$$

The initial conditions are  $x_{1|0} = x_{0|0} = 0$  and  $\sigma^2 \Sigma_{1|0} = \sigma^2 \Sigma_{0|0} = cov(x_0)$  and  $\sigma^2$  is estimated with (2). Now, suppose that we want to interpolate the observation at time  $t = h$ . The interpolated value,  $E[z_h|z_{(h)}] = z_{h|T}^{(h)}$ , where  $z_{(h)} = \{z_i : i = 1, \dots, T, i \neq h\}$  is obtained in two steps. First, we assume that the value  $z_h$  is missing and compute the estimation of the state variables with the Kalman filter under this condition. Second, we compute the interpolated value by going backwards with the FPS algorithm. The situation in which  $z_h$  is not observed can be represented by the state-space model,

$$\begin{aligned} z_t &= (1 - I_t^{(h)})H'x_t + I_t^{(h)}w_t, \\ x_t &= Fx_{t-1} + Gu_t, \end{aligned} \tag{5}$$

where  $I_t^{(h)}$  is a dummy variable such that  $I_t^{(h)} = 0$ ,  $t \neq h$  and  $I_h^{(h)} = 1$  and  $w_t$  represents independent  $N(0, 1)$  variables, independent of  $z_h$ . The Kalman filter applied to this situation is given by

$$\begin{aligned} x_{t|t-1}^{(h)} &= Fx_{t-1|t-1}^{(h)}, \\ \Sigma_{t|t-1}^{(h)} &= F\Sigma_{t-1|t-1}^{(h)}F' + Q, \\ K_t^{(h)} &= (1 - I_t^{(h)})\Sigma_{t|t-1}^{(h)}H(v_{t|t-1}^{2,(h)})^{-1}, \\ x_{t|t}^{(h)} &= x_{t|t-1}^{(h)} + K_t^{(h)}(z_t - z_{t|t-1}^{(h)}), \\ \Sigma_{t|t}^{(h)} &= (I - (1 - I_t^{(h)})K_t^{(h)}H')\Sigma_{t|t-1}^{(h)}, \end{aligned} \quad (6)$$

for  $t = 1, \dots, T$ , where  $z_{t|t-1}^{(h)} = (1 - I_t^{(h)})H'x_{t|t-1}^{(h)} + I_t^{(h)}w_t$ ,  $v_{t|t-1}^{2,(h)} = (1 - I_t^{(h)})H'\Sigma_{t|t-1}^{(h)}H + I_t^{(h)}$ , and,

$$\begin{aligned} x_{t|s}^{(h)} &= E[x_t|z_1, \dots, z_s] \\ \Sigma_{t|s}^{(h)} &= cov[x_t|z_1, \dots, z_s]. \end{aligned} \quad s, t = 1, \dots, T.$$

All the values have the superscript  $h$  in order to distinguish between the Kalman filter with the observation at  $t = h$  and without it. Of course, for  $t < h$ ,  $x_{t|t-1}^{(h)} = x_{t|t-1}$ . Note that for  $t = h$ ,  $v_{h|h-1}^{2,(h)} = 1$ . Second, we use the FPS algorithm to obtain the interpolated value, that can be derived as follows. Consider the augmented process  $y_t = [x_t' \ x_t^a]'$ , such that,  $x_t^a = x_{t-1}^a$  and  $x_h^a = x_h$  with state-space form given by

$$\begin{aligned} \begin{bmatrix} x_t \\ x_t^a \end{bmatrix} &= \begin{bmatrix} F & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-1}^a \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} u_t, \\ z_t &= \begin{bmatrix} (1 - I_t^{(h)})H' & 0 \end{bmatrix} \begin{bmatrix} x_t \\ x_t^a \end{bmatrix} + I_t^{(h)}w_t. \end{aligned}$$

Now, applying the Kalman filter to the augmented process  $y_t$ , for  $t \geq h$ , with the initial condition  $\Sigma_{h|h-1}^a = \Sigma_{h|h-1}^{(h)}$  and the FPS works as follows (see Gómez and Maravall (1994) for more details):

$$\begin{aligned} K_t^a &= (1 - I_t^{(h)})(\Sigma_{t|t-1}^a)'H(v_{t|t-1}^{2,(h)})^{-1}, \\ x_{h|t}^{(h)} &= x_{t|t-1}^{(h)} + K_t^a(z_t - z_{t|t-1}^{(h)}), \\ \Sigma_{h|t}^{(h)} &= \Sigma_{h|t-1}^{(h)} - (1 - I_t^{(h)})\Sigma_{t|t-1}^a H(K_t^a)', \\ \Sigma_{t+1|t}^a &= \Sigma_{t|t-1}^a (F - (1 - I_t^{(h)})FK_t^a H')'. \end{aligned} \quad (7)$$

We note that for  $t = h$ , the FPS gives  $K_h^a = 0$ ,  $x_{h|h}^{(h)} = x_{h|h-1}^{(h)}$ ,  $\Sigma_{h|h}^{(h)} = \Sigma_{h|h-1}^{(h)}$  and  $\Sigma_{h+1|h}^a = \Sigma_{h|h-1}^a F'$ . The interpolated value for  $z_h$  is  $z_{h|T}^{(h)} = H'x_{h|T}^{(h)}$  and the corresponding interpolated error is given by

$$i_h = z_h - z_{h|T}^{(h)}, \quad (8)$$

with  $E[i_h^2] = \sigma^2(H'\Sigma_{h|T}^{(h)}H)$ .

### 3. A relationship between prediction and interpolation errors in ARMA processes

In this section, we analyze the relationship between interpolation and prediction errors and obtain an expression for the prediction error in terms of the interpolation errors. This relationship allows us to obtain the covariance matrix of the interpolation errors. We also show that the parameter values which minimize the sum of squares of the standardized one-step ahead prediction errors are the same as that which minimize the sum of squares of the standardized interpolation errors. This main result is summarized in the following theorem:

**Theorem 1.** *Let  $z = (z_1, \dots, z_T)'$  be a time series generated by the stationary and invertible ARMA( $p, q$ ) process in (1). Let  $i_h$  be the interpolation error of the observation at  $t = h$  and let  $e_h, \dots, e_T$  be one-step ahead prediction errors assuming that all the observations are known. Then,  $i_h$  can be written as follows:*

$$i_h = c_h^{(h)} e_h + c_{h+1}^{(h)} e_{h+1} + \dots + c_T^{(h)} e_T, \tag{9}$$

where the coefficients  $c_h^{(h)}, \dots, c_T^{(h)}$  are given by

$$c_h^{(h)} = 1 - H' \sum_{i=h+1}^T K_i^a b_i^h \quad c_t^{(h)} = -H' \left( K_t^a - \sum_{i=t+1}^T K_i^a b_i^h \right), \quad t = h + 1, \dots, T, \tag{10}$$

and the coefficients  $b_t^i$  are given by

$$b_t^{h+i} = H' F^{t-h-i} K_h - H' F \sum_{j=1}^{t-h-i} K_{h+j}^{(h)} b_{h+j}^{h+i}, \quad b_t^h = H' F^{t-h} K_h - H' F \sum_{j=1}^{t-h-1} K_{h+j}^{(h)} b_{h+j}^{h+i}, \tag{11}$$

for  $i = 0, \dots, t - h - 1$ .

**Proof.** Let,

$$e_t^{(h)} = z_t - z_{t|h-1}^{(h)}, \quad t = h + 1, \dots, T \tag{12}$$

be the prediction errors assuming that observation at  $t = h$  is missing. The relationships in (7) provide the following expression for the interpolated value  $z_{h|T}^{(h)}$ :

$$z_{h|T}^{(h)} = z_{h|h-1}^{(h)} + H' K_{t+1}^a (z_{h+1} - z_{h+1|h}^{(h)}) + \dots + H' K_T^a (z_T - z_{T|T-1}^{(h)})$$

and by (8) and (12),

$$i_h = e_h - H' K_{t+1}^a e_{h+1}^{(h)} - \dots - H' K_T^a e_T^{(h)}. \tag{13}$$

We will obtain an expression of the errors  $e_t^{(h)}$  in terms of the prediction errors  $e_t$ , for  $t > h$ . For (12), we have

$$e_t^{(h)} = z_t - z_{t|h-1}^{(h)} + z_{t|h-1} - z_{t|h-1}^{(h)} = e_t + z_{t|h-1} - z_{t|h-1}^{(h)}.$$

Using the Kalman filters in (4) and (6), it can be shown that

$$z_{t|h-1} - z_{t|h-1}^{(h)} = H' F^{t-h} K_h e_h + H' F^{t-h-1} (K_{h+1} e_{h+1} - K_{h+1}^{(h)} e_{h+1}^{(h)}) + \dots + H' F (K_{t-1} e_{t-1} - K_{t-1}^{(h)} e_{t-1}^{(h)}),$$

and, therefore,

$$e_t^{(h)} = H'F^{t-h}K_h e_h + H'F^{t-h-1}(K_{h+1}e_{h+1} - K_{h+1}^{(h)}e_{h+1}^{(h)}) + \dots + H'F(K_{t-1}e_{t-1} - K_{t-1}^{(h)}e_{t-1}^{(h)}) + e_t. \tag{14}$$

Consequently, starting from  $e_{h+1}^{(h)} = e_{h+1} + H'FK_h e_h$ , we obtain the values of  $e_t^{(h)}$  in terms of  $e_h, \dots, e_t$  as:

$$e_t^{(h)} = e_t + b_t^{t-1}e_{t-1} + \dots + b_t^h e_h,$$

where the coefficients  $b_t^s$  are obtained recursively from (13) and (14), and are given by,

$$b_t^{h+i} = H'F^{t-h-i}K_h - H'F \sum_{j=1}^{t-h-i} K_{h+j}^{(h)} b_{h+j}^{h+i}, \quad b_t^h = H'F^{t-h}K_h - H'F \sum_{j=1}^{t-h-1} K_{h+j}^{(h)} b_{h+j}^{h+i},$$

for  $i = 0, \dots, t - h - 1$ , which shows (9) with the coefficients in (10).

Some consequences are as follows. First,  $E[i_h] = 0$ . Second, the variance of  $i_h$  is given by

$$var(i_h) = E(i_h^2) = E \left[ \left( \sum_{t=h}^T c_t^{(h)} e_t \right)^2 \right] = \sigma^2 \sum_{t=h}^T (c_t^{(h)})^2 v_{t|h}^2.$$

Third, if  $m > h$ , then,

$$cov(i_h, i_m) = \sigma^2 \sum_{t=m}^T c_t^{(h)} c_t^{(m)} v_{t|h}^2.$$

Let  $e = (e_1, \dots, e_T)'$  and  $i = (i_1, \dots, i_T)'$  be the vectors of prediction and interpolation errors. The vector  $e$  has a diagonal covariance matrix  $\Sigma_e$  with elements  $\sigma^2 v_{t|t-1}^2$ ,  $t = 1, \dots, T$ . The vectors  $e$  and  $i$  are related by  $i = Ce$ , where  $C$  is an upper triangular  $T \times T$  matrix with elements  $c_{uv} = c_u^{(v)}$ ,  $u, v = 1, \dots, T$ . Consequently, the covariance matrix of  $i$  is  $\Sigma_i = C\Sigma_e C'$ , and taking into account that  $i = Ce$ , we also have that

$$i' \Sigma_i^{-1} i = (Ce)' (C\Sigma_e C')^{-1} (Ce) = e' \Sigma_e^{-1} e.$$

As a consequence, the parameters that minimize the sum of squares of standardized interpolation errors,  $i' \Sigma_i^{-1} i$ , are the ML estimates, that is the parameters which minimize the sum of squares of standardized prediction errors,  $e' \Sigma_e^{-1} e$ . As a by-product, we get that,

$$\frac{\partial(e' \Sigma_e^{-1} e)}{\partial(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)} = \frac{\partial(i' \Sigma_i^{-1} i)}{\partial(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma^2)}.$$

#### 4. Illustration

As an illustration, consider an stationary AR(1) model with zero mean and autoregressive parameter  $\phi$ . Running the Kalman filter for a realization of this process,  $z = (z_1, \dots, z_T)'$ , with initial conditions  $x_{1|0} = x_{0|0} = 0$  and  $\Sigma_{1|0} = \Sigma_{0|0} = (1 - \phi^2)^{-1}$ , we get the prediction errors  $e = (e_1, \dots, e_T)'$  and their conditional variances  $\sigma^2 v_{1|0}^2 = \sigma^2(1 - \phi^2)^{-1}$  and  $\sigma^2 v_{t|t-1}^2 = \sigma^2$ ,  $t > 1$ . Running

the Kalman filter assuming that the observation  $t = h$  is missing, and then, the FPS algorithm, we get,

$$\Sigma_{t+1|t}^a = \begin{cases} \phi & t = h, \\ 0 & t > h, \end{cases} \quad K_t^a = \begin{cases} 0 & t \neq h + 1, \\ \frac{\phi}{1+\phi^2} & t = h + 1, \end{cases}$$

$$x_{h|t}^{(h)} = \begin{cases} \phi x_{h-1} & t = h, \\ \frac{\phi}{1+\phi^2}(z_{h-1} + z_{h+1}) & t > h, \end{cases} \quad \Sigma_{h|t}^{(h)} = \begin{cases} 1 & t = h, \\ \frac{1}{1+\phi^2} & t > h, \end{cases}$$

implying that the interpolated value is  $z_{h|T}^{(h)} = \frac{\phi}{1+\phi^2}(z_{h-1} + z_{h+1})$  with interpolation error  $i_h = z_h - \frac{\phi}{1+\phi^2}(z_{h-1} + z_{h+1})$ . From (13) and (11), we get,

$$i_h = \begin{cases} (1 - \phi^2)e_1 - \phi e_2 & h = 1, \\ \frac{1}{1+\phi^2}e_h - \frac{\phi}{1+\phi^2}e_{h+1} & h = 2, \dots, T - 1, \\ e_T & h = T, \end{cases} \tag{15}$$

which gives the variance of  $i_h$ ,

$$var(i_h) = \begin{cases} \frac{\sigma^2}{1+\phi^2} & h = 2, \dots, T - 1, \\ \sigma^2 & h = 1, T, \end{cases}$$

and the covariances between interpolation errors,

$$cov(i_h, i_m) = \begin{cases} \frac{-\sigma^2\phi}{(1+\phi^2)} & m = h + 1, h = 1, T - 1, \\ \frac{-\sigma^2\phi}{(1+\phi^2)^2} & m = h + 1, h = 2, \dots, T - 2, \\ 0 & m > h + 1, h = 1, \dots, T, \end{cases}$$

implying that the interpolation errors are uncorrelated if  $m - h > 1$ . Finally, we show the equality  $i' \Sigma_i^{-1} i = e' \Sigma_e^{-1} e$  in the case of an AR(1) model. For that, we note that  $\Sigma_e$  can be written as

$$\Sigma_e = \sigma^2 \left( I + \frac{\phi^2}{1 - \phi^2} U \right),$$

where  $U$  is a matrix which all its elements are 0 except the (1,1) element that is 1. From (15) we have that  $i = Ce$  where the matrix  $C$  has elements,

$$C(i, j) = \begin{cases} (1 - \phi^2) & (i, j) = (1, 1), \\ -\phi & (i, j) = (1, 2), \\ \frac{1}{1+\phi^2} & j = i, 2 \leq i \leq T - 1, \\ -\frac{\phi}{1+\phi^2} & j = i + 1, 2 \leq i \leq T - 1, \\ 1 & (i, j) = (T, T), \\ 0 & \text{otherwise.} \end{cases}$$

As  $i' \Sigma_i^{-1} i = e' C' (C \Sigma_e C')^{-1} C e$ , we only need to show that  $C' (C \Sigma_e C')^{-1} C = \Sigma_e^{-1}$ . For that

$$\begin{aligned} C \Sigma_e C' &= C \sigma^2 \left( I + \frac{\phi^2}{1 - \phi^2} U \right) C' = \sigma^2 \left( C C' I + \frac{\phi^2}{1 - \phi^2} C U C' \right) \\ &= \sigma^2 C C' \left( I + \frac{\phi^2}{1 - \phi^2} (C')^{-1} U C' \right) = \sigma^2 C C' \left( I + \frac{\phi^2}{1 - \phi^2} U \right). \end{aligned}$$

Therefore,

$$\begin{aligned} C' (C \Sigma_e C')^{-1} C &= \frac{1}{\sigma^2} C' \left( I + \frac{\phi^2}{1 - \phi^2} U \right)^{-1} (C C')^{-1} C = \frac{1}{\sigma^2} C' (I - \phi^2 U) (C')^{-1} \\ &= \frac{1}{\sigma^2} (I - \phi^2 C' U (C')^{-1}) = \frac{1}{\sigma^2} (I - \phi^2 U) = \Sigma_e^{-1}. \end{aligned}$$

## Acknowledgements

We acknowledge the financial support by project SEJ2004-03303 and by Fundación BBVA, Spain.

## References

- Anderson, B., Moore, J., 1979. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ.
- Gómez, V., Maravall, A., 1994. Estimation, prediction, and interpolation for nonstationary series with the Kalman filter. *J. Amer. Statist. Assoc.* 89, 611–624.
- Jones, R.H., 1980. Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics* 22, 389–395.