Working Paper 92-40
Statistics and Econometrics Series 28
September 1992

División de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

MISSING OBSERVATIONS AND ADDITIVE OUTLIERS IN TIME SERIES MODELS

Agustín Maravall and Daniel Peña*

Abstract

The paper deals with estimation of missing observations in possible nonstationary ARIMA models. First, the model is assumed known, and the structure of the interpolation filter is analyzed. Using the inverse or dual autocorrelation function it is seen how estimation of a missing observation is analogous to the removal of an outlier effect; both problems are closely related with the signal plus noise decomposition of the series. The results are extended to cover, first, the case of a missing observation near the two extremes of the series; then to the case of a sequence of missing observations, and finally to the general case of any number of sequences of any length of missing observations. The optimal estimator can always be expressed, in a compact way, in terms of the dual autocorrelation function or a truncation thereof; is mean squared error is equal to the inverse of the (appropriately chosen) dual autocovariance matrix. The last part of the paper illustrates a point of applied interest: When the model is unknown, the additive outlier approach may provide a convenient and efficient alternative to the standard Kalman filter-fixed point smoother approach for missing observations estimation.

Key words:
ARIMA models, interpolation, inverse autocorrelations.

*Maravall, European University Institute, Florence, Italy; Peña, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. Daniel Peña acknowledges support from DGICYT, project PB 90-0266.

1 Introduction

In this paper we deal with interpolation of missing observations in time series that are the outcome of Autoregressive Integrated Moving Average (ARIMA) processes. For a stationary time series, the problem of interpolating missing values given an infinite realization of the (known) stochastic process was solved by Kolmogorow and Wiener [see, for example, Grenander and Rosenblatt (1957), or Whittle (1963)]. The interpolator is obtained as the expected value of the missing observation given the observed infinite realization of the series. For many years, however, their result was not extended to the more general problem of interpolation in a finite realization of a (possibly) nonstationary time series, generated by a model with unknown parameters. A first step connecting the classical result on interpolation with estimation of missing values in nonstationary series with unknown model parameters is found in Brubacher and Wilson (1976). In their approach, the unobserved values are treated as unknown parameters, and are estimated by a least squares method. The missing observation estimator obtained can be interpreted as a symmetric weighted combination of the observed data before and after the gap, where the symmetric weights are the elements of the Inverse or Dual Autocorrelation Function DACF of the process, a function introduced in Cleveland (1972). The authors also noticed how their result was a straightforward extension of the classical result on stationary series.

For some years, however, the important contribution of Brubacher and Wilson went mostly unnoticed. To quote an example, in a review paper on inverse autocorrelation by Chatfield (1979), no mention is made of the key role this autocorrelation plays in the field of interpolation, nor is the work of Brubacher and Wilson mentioned. Perhaps the relatively small impact of their work was due to the fact that, contrary to standard procedure, in their approach the missing values were treated as parameters, and not computed as the conditional expectation of the unknown random variable. Moreover, they dealt with nonstationary series, and the properties of missing observations estimators for this class of series were not well-understood at the time.

Of the several approaches to the problem of interpolation in time series, possibly the one that offers at present the best-known and most complete solution is based on the Kalman filter. Jones (1980) used Akaike's state space representation of an ARMA model to compute its likelihood function in the case of missing observations. Shumway and Stoffer (1982) proposed using the EM algorithm in conjuntion with the conventional Kalman smoothed estimators for estimating the parameters by maximum likelihood allowing for missing data. Computation of the estimates by a modified Newton–Raphson routine was discussed by Wincek and Reinsel (1986). Harvey and Pierse (1984) extended the work of Jones to deal with nonstationary time series, and used the fixed–point algorithm to estimate the missing values. The important contribution of Harvey and Pierse had a limitation, requiring no missing values at the beginning or the end of the series. Kohn and Ansley (1986) obtained a general solution to the problem of interpolation in finite nonstationary series with unknown model parameters. In their approach, in order to define the likelihood, the data is transformed to eliminate dependence on the starting values. Next, a modified Kalman filter is used to compute the likelihood, and a modified fixed–point smoothing algorithm interpolates the missing observations. Both are generalizations of the ordinary Kalman filter and fixed–point smoother for handling a partially diffuse initial state vector. The powerful approach of Kohn and Ansley, developed over a sequence of papers, possibly represents the present state of the art. Examples of additional contributions are found in De Jong (1991), where an alternative modification of the Kalman filter handles diffuse initial states in a numerically safe way, and in Bell and Hillmer (1991), where it is shown how suitable initialization of the ordinary Kalman filter can provide the same results as the "transformation" approach of Kohn and Ansley. Finally, Gómez and Maravall (1992a) develop a methodology based on a standard state–space representation of the series and on the ordinary Kalman and fixed–point smoothing filters, which is seen to yield the same results of Kohn and Ansley (1986) and of Harvey and Pierse (1984), when the latter is applicable.

It is worth noticing that the Kalman filter–fixed point smoother method mentioned in the previous paragraph does not refer to the work by Brubacher and Wilson. More in line with the regression approach of these authors, an alternative approach to missing values in time series takes into account the relationship between estimation of outliers and interpolation. Peña (1987) showed that, for stationary autoregressive models, missing value estimation was asymptotically equivalent to additive outlier estimation. In particular, the likelihood is in both cases the same, apart from a determinant whose effect will tend to zero as the length of the series increases (relative to the number of missing observations). Ljung (1989) extended the additive outlier approach to blocks of missing data, and analysed the likelihood in these cases. Peña and Maravall (1991) used the additive outlier–missing observation relationship to find the optimal interpolator for any pattern of missing data in an infinite realization of a possibly nonstationary series, and showed how the vector of interpolators could be expressed using the DACF. Further extensions of the DACF approach to missing observation interpolation are found in Battaglia and Bhansali (1987).

Whatever the approach, estimation of missing observations in ARIMA time series requires two distinct steps. First, maximization of a well–defined likelihood yields estimators of the model parameters. Second, once the parameters have been estimated, interpolators of the missing values are obtained by computing the conditional expectation of the missing observations given the available data. This paper centers mostly on the second step: the filter that yields the conditional expectation of interest for the general case of any pattern of missing observations in a possibly nonstationary time series. The main purpose of the paper is to provide a better understanding of the structure of this filter, and how it relates to the stochastic structure of the series and to other statistical problems such as outlier removal and signal extraction. In particular, the relationship with estimation of outlier effects is seen to provide an implication of considerable applied interest.

Section 2 provides some background material and considers the case of a single missing observation for a complete realization of the series. Section 3 discusses some properties of the estimator and relates missing observation interpolation to the problem of removing an outlier effect. Section 4 considers the relationship between interpolation and the problem of decomposing a time series into signal plus noise. Section 5 presents an interesting alternative derivation of the optimal estimator, which is then used in section 6 to consider the case of a missing observation near one of the extremes of the series (i.e., the case of a finite realization). Section 7 generalizes the results to a vector of missing observations, first when they are consecutive, and second to the general case of any number of sequences of any length of missing observations in a finite series. Finally, section 8 presents the empirical application, in which estimation of missing observations by the standard Kalman filter–fixed point smoother approach and by an additive outlier approach are compared using a well–known example.

2 Optimal Interpolation of a Missing Value

In order to establish some terminology and assumptions that will be used throughout the paper, let the series in question follow the general ARIMA model

$$\phi(B)\, z_t = \theta(B)\, a_t, \tag{2.1}$$

where $\phi(B)$ and $\theta(B)$ are finite polynomials in the lag operator B, and a_t is a Gaussian white–noise process with variance V_a. Without loss of generality, we set $V_a = 1$; thus, in the following pages, all variances and mean–squared errors will be implicitly expressed in units of the one–step–ahead forecast error (or innovation) variance. The polynomial $\phi(B)$ may contain any number of unit roots and hence the process can be nonstationary; we assume, however, that the model is invertible, so that the roots of $\theta(B)$ lie outside the unit circle. Thus, the model (2.1) can alternatively be expressed in autoregressive form as

$$\pi(B)\, z_t = a_t, \tag{2.2}$$

where $\pi(B)$ is the convergent polynomial

$$\pi(B) = \phi(B)\, \theta(B)^{-1} = (1 - \pi_1 B - \pi_2 B^2 - \cdots).$$

Define the "inverse or dual model" of (2.1) as the one that results from interchanging the AR and MA polynomials; therefore the dual model is given by

$$\theta(B)\, z_t^D = \phi(B)\, a_t, \tag{2.3}$$

or

$$z_t^D = \pi(B)\, a_t, \tag{2.4}$$

Since (2.1) is invertible, model (2.3) will be stationary; its autocorrelation generating function (ACGF) will be given by

$$\rho^D(B) = \pi(B)\,\pi(F)/V^D, \tag{2.5}$$

where $F = B^{-1}$ denotes the forward operator, and V^D is the variance of the dual process, equal to

$$V^D = \sum_{i=0}^{\infty} \pi_i^2, \quad (\pi_0 = 1), \tag{2.6}$$

which will always be finite. The function (2.5) has been often referred to as the inverse autocorrelation function [Cleveland (1972)]. Since, in the next sections, we shall use autocorrelation matrices, and the inverse of the inverse autocorrelation matrix is not equal, in general, to the autocorrelation matrix, to avoid awkward expressions, we shall refer to (2.5) as the dual autocorrelation function (DACF). This is also in line with the duality properties of autoregressive and moving average polynomials in ARIMA models; see, for example, Pierce (1970). Trivially, from the ARIMA expression of the model, the DACF is immediately available.

Consider first the case of a series z_t which has a missing value for $t = T$, and denote by $z_{(T)}$ the vector of observed values. For a linear stationary series, the minimum mean-squared error (MMSE) estimator of z_T is given by

$$\hat{z}_T = E(z_T / z_{(T)}),$$

that is

$$\hat{z}_T = Cov(z_T,\, z_{(T)})'\, Var^{-1}\,(z_{(T)})\, z_{(T)},$$

where $Cov(z_T, z_{(T)})$ is a vector with the i–th element given by $Cov(z_T,\, z_i), i \neq T$, and $Var(z_{(T)})$ is the covariance matrix of $z_{(T)}$. Therefore, $\hat{z}_T$ is a filter given by a linear combination of the observed values, where the weights depend on the covariance structure of the process. As the series approaches ∞ in both directions, the filter becomes centered and symmetric, and it is well known [see, for example, Grenander and Rosenblatt (1957)] that its weights are the dual autocorrelations of z_t; thus the optimal estimator of the missing value can be expressed as

$$\hat{z}_T = -\sum_{k=1}^{\infty} \rho_k^D \, (z_{T-k} + z_{T+k}), \tag{2.7}$$

where ρ_k^D is the coefficient of B^k in (2.5). It is also well-known [see, for example, Brubacker and Wilson (1976) or Ljung (1989)] that the result (2.7) remains unchanged if the stationarity assumption is dropped and the process (2.1) becomes a nonstationary ARIMA model. The filter (2.7) will be finite for a pure AR model, and will extend to ∞ otherwise; invertibility of the model guarantees, however, its convergence in this last case.

Since (2.7) can be rewritten as

$$\hat{z}_T = (1 - \rho^D(B))\, z_T, \tag{2.8}$$

it follows that

$$E(z_T - \hat{z}_T)^2 = E(\rho^D(B)\, z_T)^2 = E(\pi(F)\, a_T)^2/(V^D)^2,$$

and hence the Mean-Squared Error (MSE) of $\hat{z}_T$ is

$$\text{MSE}\,(\hat{z}_T) = 1/V^D. \tag{2.9}$$

To illustrate (2.7), consider first the AR(1) model

$$z_t = \phi\, z_{t-1} + a_t.$$

Its dual model is $z_t^D = a_t - \phi\, a_{t-1}$, with variance $V^D = 1 + \phi^2$ and autocorrelations $\rho_1^D = -\phi/(1+\phi^2)$, $\rho_k^D = 0$ for $|\, k \,| > 1$. The missing observation estimator is then given by

$$\hat{z}_T = \frac{\phi}{1+\phi^2}\, (z_{T-1} + z_{T+1}),$$

in agreement with the result in Gourieroux and Monfort (1989, p. 734); moreover

$$\text{MSE}\,(\hat{z}_T) = 1/(1+\phi^2).$$

As a second example, we use the more complicated nonstationary model:

$$\Delta\,\Delta_{12}\,z_t = (1-\theta_1 B)\,(1-\theta_{12}\,B^{12})\,a_t, \tag{2.10}$$

(the so-called Airline Model), popularized by Box and Jenkins (1970, chap. 9). The model has been found useful for many monthly economic series that display trend and seasonal behavior. (Values of θ_1 close to 1 imply relatively stable trends and, similarly, large values of θ_{12} represent relatively stable seasonality.) The Airline Model has also become a standard example in the literature on missing observations: see, for example, Harvey and Pierse (1984) and Kohn and Ansley (1986). We shall follow their tradition and the Airline Model will be used as an example throughout the paper. Figure 1 displays the two-sided symmetric filters that yield the estimator of a missing value in the middle of the series for 3 sets of parameter values. It is seen how stable components induce long filters, while unstable ones place practically all weight on recent observations. Table 1 presents the root mean squared error (RMSE) of the final estimator $\hat{z}_T$ for different values of θ_1 and θ_{12} (the units have been standarized by setting $V_a = 1$). Table 1 is practically symmetric for θ_1 and θ_{12}. As θ_1 and θ_{12} tend to 1, the RMSE of the estimator tends also to 1. This is sensible, since, in the limit, the two differences in (2.10) would cancel out, and, ignoring deterministic components, the series z_t would simply be the white-noise a_t, with variance 1. On the contrary, as the series approaches noninvertibility, the estimation error tends to zero, but the filter $\rho^D(B)$ tends then towards nonconvergence and, in the limit, the estimator ceases to exist.

INSERT	Table 1

3 Missing Observation and Additive Outlier

Consider now a series which follows model (2.1), but with an additive outlier (instead of a missing value) at time T. The effect of the outlier can be estimated in the following way. Express the observed series Z_t as

$$\begin{aligned} Z_t &= z_t, \qquad t \neq T \\ Z_T &= z_T + \omega \end{aligned} \tag{3.1}$$

where ω is the outlier effect. Construct then the dummy variable d_t, such that $d_t = 0$ for $t \neq T$ and $d_T = 1$, and write model (2.2) as

$$\pi(B)\,(Z_t - \omega\, d_t) = a_t,$$

or equivalently,

$$\pi(B)\, Z_t = \omega\, \pi(B)\, d_t + a_t. \tag{3.2}$$

Defining the variables $y_t = \pi(B)\, Z_t$ and $x_t = \pi(B)\, d_t$, (3.2) is seen to be the simple regression model

$$y_t = \omega\, x_t + a_t,$$

with x_t deterministic and a_t white–noise; therefore the MMSE estimator of ω is given by

$$\hat{\omega} = \Sigma\, y_t\, x_t / \Sigma\, x_t^2. \tag{3.3}$$

Using results from the Appendix in Box and Tiao (1975), after simplification, it is found that, for a complete realization of the series,

$$\Sigma\, y_t\, x_t = \Sigma\, \pi(B)\, Z_t\, \pi(B)\, d_t = \pi(B)\, \pi(F)\, Z_T \tag{3.4}$$

and

$$\Sigma x_t^2 = \Sigma \pi(B)\, d_t\, \pi(B)\, d_t = \Sigma \pi_i^2 = V^D \tag{3.5}$$

so that expression (3.3) becomes

$$\hat{\omega} = (\pi(B)\, \pi(F)/V^D)\, Z_T, \tag{3.6}$$

[in agreement with the result in Chang, Tiao and Chen (1988)] and, from (2.5),

$$\hat{\omega} = \rho^D(B)\, Z_T. \tag{3.7}$$

The estimator of the series, once the outlier effect has been removed, is

$$\hat{z}_T = Z_T - \hat{\omega}, \tag{3.8}$$

and, using (3.7), it can be expressed as

$$\hat{z}_T = (1 - \rho^D(B))\, Z_T = - \sum_{k=1}^{\infty} \rho_k^D\, (z_{T-k} + z_{T+k}),$$

identical to expression (2.7). As a consequence, when the model is known, the relationship between interpolation of a missing observation and estimation of an additive outlier can be summarized in two alternative ways: On the one hand, removal of the outlier effect at period T is equivalent to estimating a missing observation for T. Alternatively, estimation of a missing observation can be seen as the result of the following procedure: First, fill the "hole" in the series with an arbitrary number Z_T; then treat Z_T as an additive outlier. Removing the estimated outlier effect from Z_T, the missing observation estimator is obtained.

Some properties of the estimators (3.7) and (3.8) — or, equivalently, (2.7) — are worth noticing:

(1) The derivation remains unchanged when the autoregressive polynomial of model (2.1) contains nonstationary roots. As for the MSE, since $z_T - \hat{z}_T =$

$\hat{\omega} - \omega$, expression (3.5) yields $\text{MSE}(\hat{z}_T) = \text{MSE}(\hat{\omega}) = (V^D)^{-1}$, in agreement with (2.9). Thus, even for nonstationary series, the MSE of the estimator is finite. Since $V^D > 1$, it will always be smaller than the one-period-ahead forecast error variance, as should happen. As the process approaches noninvertibility, then $\text{MSE}(\hat{z}_T) \rightarrow 0$; in the limit, the problem degenerates, however, because the filter $\rho^D(B)$ becomes nonconvergent.

(2) The procedure yields implicitly an estimated pseudo-innovation for T, equal to the difference between $\hat{z}_T$, obtained with the two-sided filter (2.7), and $\hat{z}_{T-1}(1)$, the one-period-ahead forecast of z obtained at $(T-1)$ using a one-sided filter. This pseudo-innovation can be expressed as a linear combination of all innovations for periods $T+k$, $k > 0$.

(3) If the model (2.1) contains some difference of the series (and hence is nonstationary), it will be that $\pi(1) = 0$, and hence, from (2.5),

$$\rho^D(1) = 1 + 2\Sigma\, \rho_k^D = 0,$$

where the summation sign extends from 1 to ∞. Therefore, $-\Sigma\, \rho_k^D = \frac{1}{2}$ and the sum of the weights in (2.7) is one; the estimator $\hat{z}_T$ is, in this case, a weighted average of past and future values of the series. If the process is stationary, then $\pi(1) > 0$, from which it follows that

$$-\sum \rho_k^D = \frac{1}{2} \left| \frac{\Sigma \pi_i^2 - \pi(1)^2}{\Sigma \pi_i^2} \right| < \frac{1}{2},$$

and hence the estimator $\hat{z}_T$ represents a shrinkage towards zero, the mean of the process.

4 Relationship with Signal Extraction

Consider model (2.1) and assume we wish to decompose the series z_t into signal plus noise, as in

$$z_t = s_t + u_t, \tag{4.1}$$

where $u_t \sim$ niid$(0, V_u)$, and s_t and u_t are mutually orthogonal. For period T, the MMSE estimator of the noise is the conditional expectation of u_t given the series z_t. for a complete realization of the series, this estimator is given by [see, for example, Box, Hillmer and Tiao (1978)]

$$\hat{u}_t = V_u \, \pi(B) \, \pi(F) \, z_T, \tag{4.2}$$

and comparing (4.2) with (3.6) it is seen that, except for a scale factor, the filter that provides the estimator of the noise is identical to the filter that yields the estimator of the outlier effect. Furthermore, from (2.8) and (4.2) it is obtained that the estimator of the missing observation will satisfy the equality

$$\hat{z}_T = z_T - k \, \hat{u}_T, \tag{4.3}$$

where $k = (V_u \, V^D)^{-1}$. Let $V_{\hat{u}}$ denote the variance of $\hat{u}_T$. From (4.2) and (2.2),

$$\hat{u}_t = V_u \, \pi(F) \, a_t,$$

and hence $V_{\hat{u}} = (V_u)^2 \, V^D$. Therefore, the constant k can be alternatively expressed as

$$k = V_u / V_{\hat{u}},$$

i.e., as the ratio of the variances of the (theoretical) noise component and of its MMSE estimator. Since the estimator $\hat{u}_T$ has always a smaller variance than the theoretical component u_t [see, for example, Maravall (1987)], the ratio k is always larger than one. Thus the smoothing implied by the estimation of a missing observation is equivalent to extracting from the series a multiple of its noise component. In this sense, the missing observation estimator can be seen to be an underestimation of the signal.

Assume that z_T is properly generated by (2.1) but that it is nevertheless treated, first, as an outlier and, second, as a missing observation. The estimators of the outlier effect, of the noise, of the missing observation and of the signal can be expressed as

$$\begin{aligned} \hat{w} &= \rho^D(B)\, z_T; \qquad \hat{u}_T = \frac{1}{k}\,\rho^D(B)\, z_T \\ \hat{z}_T &= (1-\rho^D(B))\, z_T; \qquad \hat{s}_T = 1 - \frac{1}{k}\,\rho^D(B)\, z_T. \end{aligned} \tag{4.4}$$

Thus, estimation of an additive outlier, of a missing observation, of the signal and of the noise are performed, up to a scale factor, with a similar filtering procedure. In order to illustrate the relationship among the filters in (4.4), consider the same example of section 2, the Airline Model given by equation (2.10), with parameter values $\theta_1 = .4$ and $\theta_{12} = .6$ (the particular values of the original Box and Jenkins example). The series can be decomposed into mutually orthogonals trend, seasonal, and white–noise irregular component [see Hillmer and Tiao (1982)]. In terms of the signal plus noise decomposition we are considering, the signal will be the sum of the first two components, and the noise will be the irregular component. The decomposition is identified by setting the variance of the noise equal to its maximum possible value, in which case the canonical decomposition is obtained. Let f denote frequency in radians, and $g(f)$ the (pseudo)spectrum of z_t [see, for example, Harvey (1989)]. The signal in the series will be associated with the peaks in $g(f)$ for the trend and seasonal frequencies, and the spectrum of the noise is a constant, equal to .314 V_a. Figure 2 displays $g(f)$ and the frequency domain representation of the filters used to obtain the signal and the missing observation, and figure 3 displays the spectrum of the inverse model [equal to $1/g(f)$], and the frequency domain representation of the filters that provide the estimator of the noise component and of the outlier effect (of course, the maxima of the inverse model spectrum correspond to the minima of $g(f)$ and vice–versa). It is seen that, as should be expected, the estimator of the signal filters the frequencies for which there is a large signal, and the estimator of the noise those for which the noise contribution is relatively more important (i.e., the minima of $g(f)$). In particular, for the trend and seasonal frequencies, the signal filters entirely the frequency, while the filter for the noise is zero. From the figures it is seen how the filters for estimating the missing observation and the outlier effect follow exactly the same principle: the missing observation is estimated by filtering the signal, while the outlier effect is obtained by filtering the noise.

Notwithstanding the similarities between the filters, figures 2 and 3 clearly evidence a difference: more of the series variation is assigned to the signal than to the missing observation and, accordingly, less is assigned to the noise than to the outlier effect (despite the fact that the canonical decomposition has maximized the variance of the noise). This is a general result as is immediately seen by combining the first two expressions in (4.4), to yield $\hat{\omega} = k\,\hat{u}_t$, and hence,

$$V(\hat{\omega}) = k^2 V_{\hat{u}} = k\,V_u > V_u.$$

This has an interesting implication: Since model (2.1) is invertible, V_u and $V_{\hat{u}}$ are positive. Assume that $\omega = 0$ but z_T is treated nevertheless as an outlier. Then the estimator of the (nonexistent) outlier effect would still be a multiple of the noise component that can be extracted from z_T. (For the previous example, this multiple k is close to 2, although for other models it may take much larger values). In this sense one can speak of structural underestimation of the signal by the missing observation estimator and of overestimation of the outlier effect. This is reflected in the negative value of the transfer function for the missing observation estimator in some frequency ranges (see figure 2), and has the effect of introducing a phase shift of π radians in the gain function of the interpolation filter for those frequency ranges, as evidenced in figure 4.

5 The Optimal Interpolator as a "Pooled" Estimator

Consider the problem of estimating a missing observation at time T for a series that follows the AR(2) model

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t. \tag{5.1}$$

An obvious estimator of z_T is the one–period–ahead forecast of the series $(\cdots, z_{T-2}, z_{T-1})$. Denoting this estimator by z_T^0,

$$z_T^0 = \phi_1 z_{T-1} + \phi_2 z_{T-2}, \tag{5.2}$$

and its MSE is given by $\text{MSE}(z_T^0) = M_0 = V_a = 1$. This estimator ignores the information z_{T+k}, $k > 0$. An alternative estimator that uses this information can be obtained by backcasting z_T in the sequence $(z_{T+1}, z_{T+2}, \cdots)$. This second estimator z_T^2 is given by

$$z_T^2 = (z_{T+2} - \phi_1 z_{T+1})/\phi_2, \tag{5.3}$$

with associated MSE $M_2 = 1/\phi_2^2$.

While z_T^0 is computed as the last value of z in (5.1), i.e. by setting $T = t$, z_T^2 is computed by setting T equal to the first element in (5.1), i.e. $T = t - 2$. Equation (5.1) still offers another possibility, namely, when z_T is in the middle. This will happen when $t = T + 1$ in (5.1) and, solving for z_T, a third estimator is obtained:

$$z_T^1 = (z_{T+1} - \phi_2 z_{T-1})/\phi_1, \tag{5.4}$$

with MSE $M_1 = 1/\phi_1^2$.

Since the three estimation errors are functions of a_T, a_{T+1}, and a_{T+2}, respectively, the three estimators are independent. A pooled estimator of z_T can be obtained as a weighted average of them, where the weights are proportional to their precision. If z_T^p denotes the pooled estimator,

$$z_T^p = h(z_T^0/M_0 + z_T^1/M_1 + z_T^2/M_2),$$

where $h^{-1} = 1/M_0 + 1/M_1 + 1/M_2$. Considering (5.2)-(5.4) and the values of M_0, M_1 and M_2, after simplification, it is found that

$$z_T^p = \frac{1}{1+\phi_1^2+\phi_2^2}\left[\phi_1(1-\phi_2)\,(z_{T-1}+z_{T+1}) + \phi_2(z_{T-2}+z_{T+2})\right] \tag{5.5}$$

or, considering the DACF of an AR(2) process,

$$z_T^p = -\rho_1^D(z_{T-1}+z_{T+1}) - \rho_2^D(z_{T-2}+z_{T+2}),$$

the same as expression (2.7).

The previous result for the AR(2) model generalizes to any linear invertible (possibly nonstationary) model of the type (2.1). To see this, consider the pure autoregressive representation of the model:

$$z_t = \pi_1\, z_{t-1} + \pi_2\, z_{t-2} + \cdots + a_t, \tag{5.6}$$

or, for $t = T + j, \quad (j = 0, 1, 2, \cdots)$,

$$z_{T+j} = \pi_1\, z_{T+j-1} + \pi_2\, z_{T+j-2} + \cdots + \pi_j\, z_T + \cdots + a_{T+j}. \tag{5.7}$$

Using a notation similar to that used in the AR(2) example, the estimator z_T^j is given by

$$\begin{aligned} z_T^j &= (1/\pi_j)\,(z_{T+j} - \pi_1\, z_{T+j-1} - \cdots) = \\ &= (1/\pi_j)\,(\pi(B)\, F^j + \pi_j)\, z_T, \end{aligned} \tag{5.8}$$

(for $j = 0$ we adopt the convention $\pi_0 = -1$), and its MSE is $M_j = 1/\pi_j^2$. Letting $j = 0, 1, 2, \cdots$, the pooled estimator, z_T^p, is given by (all summation signs extend from $j = 0$ to $j = \infty$)

$$z_T^p = h \sum_j z_T^j / M_j, \tag{5.9}$$

where $h^{-1} = \sum_j (1/M_j) = \sum_j \pi_j^2 = V^D$. Thus, using (5.8),

$$\begin{aligned} z_t^p &= (1/V^D) \sum_j \pi_j (\pi(B) F^j + \pi_j) z_T = \\ &= (1/V^D) (\sum_j \pi_j^2) z_T + (1/V^D) \sum_j \pi_j F^j \pi(B) z_T = \\ &= z_T - (1/V^D) \pi(B) \pi(F) z_T = (1 - \rho^D(B)) z_T, \end{aligned}$$

and, considering (2.8), $z_T^p = \hat{z}_T$, as claimed. Therefore, the optimal estimator of the missing observation can be seen as a weighted average of the estimators that are obtained by assuming that the missing observation occupies all possible different positions for z in the autoregressive equation (2.2).

As mentioned at the beginning of the section, for a long enough series, an obvious (though inefficient) estimator of the missing observation z_T is the one–period–ahead forecast of the series $[\cdots, z_{T-2}, z_{T-1}]$, i.e., of the series truncated at $(T-1)$. Denote this forecast by $z_{T-1}^f(1)$. Similarly,, another obvious estimator is the one–period–behind backcast, obtained with the representation in F of process (2.1):

$$\phi(F) z_t = \theta(F) e_t,$$

where e_t is a sequence of independent, identically, normally distributed variables, with zero mean and variance $V_e = V_a = 1$, applied to the series $[z_{T+1}, z_{T+2}, \cdots]$. Denote this estimator by $z_{T+1}^b(-1)$.

Since the two estimators combined are based on the set of all available observations, Abraham (1981) proposed to use as interpolator a convex combination of the two:

$$\hat{z}_T = \alpha\, z_{T-1}^f(1) + (1-\alpha)\, z_{T+1}^b(-1), \tag{5.10}$$

where α is chosen so as to minimize the MSE of the forecast [for a related approach, see also Damsleth (1980)]. Except for AR(1) model case, expression (5.10) will differ from (5.9), and does not provide, as a consequence, the minimum MSE estimator of z_T. For the AR(2) example of equation (5.1), expression (5.10) eventually yields

$$\hat{z}_T = \phi_1 (z_{T-1} + z_{T+1}) + \phi_2 (z_{T-2} + z_{T+2}),$$

different from the optimal estimator (5.5).

The "pooling" interpretation of the estimator permits to decompose its MSE in an interesting way. Considering (5.7), the number of nonzero autoregressive coefficients determines the number of independent interpolators that can be pooled in (5.9), and the MSE of each interpolator z_T^j is given by $(\pi_j^2)^{-1}$. Broadly, thus, large AR processes with large coefficients (in absolute value) will provide interpolators with small estimation error. For example, for an AR(1) model, the minimum MSE is obtained for $\phi = 1$, in which case MSE $(\hat{z}_T) = \frac{1}{2}$. For an AR(2) model, the minimum MSE becomes $\frac{1}{6}$, and is obtained when the two roots of the AR polynomial are both equal to 1.

Notice that the information about the missing point contained in the forecast and in the backcast can be considerably different. For an AR(1) model, the information about the missing observation z_T contained in the forecast (equal to 1) is larger than or equal to the information contained in the backcast (equal to ϕ^2). For an AR(2) model, however, the information contained in the backcast (equal to $\phi_1^2 + \phi_2^2$) could be much larger than that contained in the forecast (still equal to 1).

6 Missing Observation Near the Two Extremes of the Series

6.1 Preliminary Estimator

The optimal estimator of a missing observation at time T, given by (2.7), is a symmetric filter centered at T. Although it extends theoretically from $-\infty$ to $+\infty$, invertibility of the series guarantees that the filter will converge towards zero, and hence that it can be truncated and applied to a finite length series. However, for T close enough to either end of the series, (2.7) cannot be used since observations needed to complete the filter will not be available.

Let the missing observation be z_T, and the last observed value z_{T+n}. Assume that n is small enough so that the filter has not converged in the direction of the future and, in order to simplify the discussion, that the series is long enough so that the filter can be safely truncated in the direction of the past. To derive the optimal estimator of z_T we use the method employed in section 5. From expression (5.6), since z_{T+j} for $j > n$ has not been observed yet, only $(n+1)$ equations of the type (5.7) can be obtained, namely those corresponding to $j = 0, 1, \cdots, n$. Therefore, expression (5.9) remains valid with the summation sign extending now from $j = 0$ to $j = n$, and $h^{-1} = \Sigma_{j=0}^{n} \pi_j^2$. Denote by V_n^D the truncated variance of the dual process,

$$V_n^D = \Sigma_{j=0}^{n} \pi_j^2,$$

and by $\pi_n(F)$ the truncated AR polynomial

$$\pi_n(F) = (1 - \pi_1 F - \cdots - \pi_n F^n).$$

Then, if $\hat{z}_{T,n}$ represents the estimator of a missing observation n periods before the end of the series,

$$\hat{z}_{T,n} = (1/V_n^D) \sum_{j=0}^{n} \pi_j (\pi(B) F^j + \pi_j) z_T =$$

$$= \; z_T - (1/V_n^D)\, \pi(B)\, (\sum_{j=0}^{n} \pi_j\, F^j)\, z_T,$$

or

$$\hat{z}_{T,n} = (1 - (1/V_n^D)\, \pi(B)\, \pi_n(F))\, z_T, \qquad (6.1)$$

where $\pi(B)\, \pi_n(F)$ is a "truncated" DACF, to be denoted $\rho_n^D(B)$.

Following a derivation similar to the one in section 2, it is straightforward to find that, if an additive outlier is assumed n periods before the end of the series, the estimator of the corresponding dummy variable coefficient is given by

$$\hat{\omega}_n = (1/V_n^D)\, \pi(B)\, \pi_n(F)\, Z_T = \rho_n^D(B)\, Z_T \qquad (6.2)$$

Since expression (6.1) does not depend on the value of the series at T, the estimator $\hat{z}_{T,n}$ can be rewritten

$$\hat{z}_{T,n} = Z_T - \hat{\omega}_n. \qquad (6.3)$$

Expressions (6.2) and (6.3) are the analogue of expressions (3.7) and (3.8) for the case of a missing observation near the end of the series. Expression (6.1) provides an asymmetric filter. When $n = 0$ it yields the one–period–ahead forecast of the series and when $n \to \infty$ it becomes the historical or final estimator given by (2.7).

To illustrate the effect that the truncation induces on the filter, for the Airline Model example of section 4, figure 5 compares the complete symmetric filter for the final estimator with the one–sided filter of the one–period–ahead forecast (i.e., the filter for $\hat{z}_{T,0}$), and with the filter of the preliminary estimator after 12 additional periods have been observed (i.e., the filter for $\hat{z}_{T,12}$). The effect of the truncation is remarkable.

If the missing observation is near the beginning of the series (n periods after the first observation) the previous derivation remains unchanged, applied to the reversed series. In this case expression (6.2) becomes

$$\hat{\omega}_n = (1/V_n^D)\, \pi_n(B)\, \pi(F)\, Z_T,$$

which, for $n = 0$, provides the one–period–behind backcast of the series.

6.2 Mean–Squared Error and Revisions

When the last observation is for period $(T+n)$, and for small enough n, the estimator $\hat{z}_{T,n}$ given by (6.1) is a preliminary estimator, that will be revised as new observations become available. Eventually, as n increases, the historical or final estimator $\hat{z}_T$, given by (2.7) will be obtained. Let δ and δ_n denote the error in the historical and in the preliminary estimator, respectively. Thus:

$$\begin{aligned} \delta &= z_T - \hat{z}_T \;= \hat{\omega} - \omega, \\ \delta_n &= z_T - \hat{z}_{T,n} = \hat{\omega}_n - \omega, \end{aligned}$$

and, from (6.1),

$$\delta_n = (1/V_n^D)\, \pi(B)\, \pi_n(F)\, z_T = (1/V_n^D)\, \pi_n(F)\, a_T,$$

where the last equality makes use of (2.2). Considering that

$$E\,[\pi_n(F)\, a_T]^2 = V_n^D,$$

the MSE of the preliminary estimator is found to be

$$\text{MSE}\,(\hat{z}_{T,n}) = \text{MSE}\,(\hat{\omega}_n) = 1/V_n^D = 1/\sum_{j=0}^{n} \pi_j^2$$

and hence equal to the inverse of the (appropriately) truncated variance of the dual process.

"Concurrent" estimation of a missing value (i.e., when the missing observation occurs for the last period in the series) is obtained when $n = 0$ and, of course, is equal to the one–period–ahead forecast, with estimation error variance $V_a = 1$. As time passes and n increases, the MSE of the estimator will decrease from 1 to $1/V^D$ and, if r_n denotes the difference between the preliminary estimator and the final one

$$r_n = \hat{z}_T - \hat{z}_{T,n} = \hat{\omega}_n - \hat{\omega},$$

then

$$\text{MSE}\,(r_n) = \frac{1}{V_n^D} - \frac{1}{V^D}.$$

Starting with concurrent estimation and moving to the final one, the variance of the total revision the estimator will undergo is equal to $1 - 1/V^D$.

To give an idea of the magnitude of the revision, table 2 displays its variance (as a fraction of the innovation variance V_a) for the Airline Model and the parameter values considered in table 1. It is seen that for large negative values of θ_1 and θ_{12}, historical estimation reduces drastically the uncertainty of the one–period–ahead forecast. On the contrary, as θ_1 and θ_{12} approach 1, historical estimation improves little upon the one–period–ahead forecast. This was to be expected since, in the limit, when θ_1 and θ_{12} are 1, as noted earlier, the series becomes white–noise and hence no "future" observation z_{T+k} can be informative for estimating the missing value z_T.

INSERT	Table 2

Besides the magnitude of the revision, it is of interest to know how long it takes for it to be completed; or, in other words, how distant the missing observation has to be from the end of the series for its estimator to be considered as (approximately) final. Table 3 exhibits the number of periods it takes to remove 95% of the total revision variance in table 2. For the vast majority of cases, this percentage is reached in less than 3 years and, except for some cases associated with close to noninvertible parameters, if the missing observation or the outlier are more than two years "old", the estimator can safely be taken as final.

INSERT	Table 3

It is worth noticing that, comparing tables 2 and 3, a somewhat comforting result emerges: the revision lasts long when the revision error is small and hence of little importance; inversely, when the revision error is large, convergence to the final estimator tends to be fast.

The symmetric and centered character of the filter that yields the estimator $\hat{\omega}$ of the outlier effect or, equivalently, of the associated dummy variable coefficient, and the existence, thus, of revisions in this estimator has some implication of interest in applied econometric work. First, what may seem at first an outlier may turn out not to be one, and viceversa; early detection of outliers can be considerably unreliable. Moreover, innovations are used in dynamic economic models to measure unanticipated changes. Often these models contain dummy variables to reflect, for example, "structural breaks" [see for example Winder and Palm (1989)]. Even if the model is assumed known and the period at which the structural break happens is instantly identified by the agent, the relevant series of innovations that approximate the agent's forecast error should be computed using the preliminary estimate of ω and its successive revisions, and not as the residuals of the model with the final estimator of ω superimposed.

7 A Vector of Missing Observations

7.1 Consecutive Periods

Consider, first, a time series z_T, generated by model (2.1), with $k+1$ consecutive missing observations at $t = T, T-1, \cdots, T-k$. We can always fill the holes with arbitrary numbers Z_T, Z_{T-1}, $\cdots$, Z_{T-k}, and define the observed series Z_t as

$$\begin{aligned} Z_t &= z_t, & t &\neq T, \cdots, T-k \\ Z_{T-j} &= z_{T-j} + \omega_j, & j &= 0, 1, \cdots, k, \end{aligned}$$

with unknown ω_j. For the rest of this section, let j take the values $0, 1, \cdots, k$. Then, the set of dummy variables

$$d_t^j = 0 \quad \text{for} \quad t \neq T+j; \qquad d_{T+j}^j = 1;$$

together with (2.2), yield the model

$$\pi(B)\,(Z_t - \sum_j \omega_j\, d_t^j) = a_t.$$

The regression equation becomes

$$y_t = \sum_j \omega_j\, x_{jt} + a_t, \tag{7.1}$$

where $y_t = \pi(B)\, Z_t$ and $x_{jt} = \pi(B)\, d_t^j$. Let $\hat{\omega}$ denote the vector of estimators $(\hat{\omega}_0 \cdots \hat{\omega}_k)$, x_j the column vector with element (x_{jt}), and x the matrix $(x_0\, x_1 \cdots x_k)$. From (7.1)

$$\hat{\omega} = (x'\, x)^{-1} x'\, y. \tag{7.2}$$

Since, summing over t, it is obtained that

$$\begin{aligned} \Sigma\, x_{jt}\, y_t &= \pi(B)\, \pi(F)\, Z_{t-j} \\ \Sigma\, x_{jt}^2 &= V^D \\ \Sigma\, x_{j-h,t}\, x_{jt} &= -\pi_h + \sum_{i=1}^{\infty} \pi_i\, \pi_{i+h}\, \gamma_h^D, \end{aligned}$$

where γ_h^D denotes the lag–h dual autocovariance, the matrix $(x'\,x)$ is the (symmetric) dual autocovariance matrix Ω^D:

$$\Omega^D = x'\,x = \begin{bmatrix} V^D & \gamma_1^D & \gamma_2^D & \cdots & \gamma_k^D \\ & V^D & \gamma_1^D & \cdots & \gamma_{k-1}^D \\ & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & & \cdot \\ & & & & \gamma_1^D \\ & & & & V^D \end{bmatrix}, \tag{7.3}$$

truncated to be of order $k+1$. Let R^D denote the corresponding dual autocorrelation matrix

$$R^D = \begin{bmatrix} 1 & \rho_1^D & \rho_2^D & \cdots & \rho_k^D \\ & 1 & \rho_1^D & \cdots & \rho_{k-1}^D \\ & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & & \cdot \\ & & & & \rho_1^D \\ & & & & 1 \end{bmatrix}; \tag{7.4}$$

considering that $\Omega^D = V^D\, R^D$, if Z denote the vector of arbitrary numbers $(Z_T, \cdots, Z_{T-k})'$, the estimator (7.2) can be expressed as

$$\hat{\omega} = (R^D)^{-1}\, \rho^D(B)\, Z. \tag{7.5}$$

If $\hat{z}$ denotes the estimator of the vector of missing observations, $(\hat{z}_T, \cdots, \hat{z}_{T-k})'$, it can be then obtained through

$$\hat{z} = Z - \hat{\omega}. \tag{7.6}$$

Equations (7.5) and (7.6) are the vector generalizations of (3.7) and (3.8). The missing observation estimators can be seen as the outcome of a similar procedure: First, filling the holes in the series with arbitrary numbers, which then are treated as additive outliers. Removing from the arbitrary numbers the outlier effects, the missing observation estimators are obtained.

Equation (7.5) provides another interesting expression for $\hat{z}$. Let $\omega_j^{(1)}$ denote the estimator of ω_j obtained by assuming that, in the series Z_t, only Z_{T-j} is arbitrary, and using the method of section 2 for the scalar case. Define the vector $\omega^{(1)} = (\omega_0^{(1)}, \cdots, \omega_k^{(1)})'$. Then, considering (3.7), (7.5) can be rewritten as

$$\hat{\omega} = (R^D)^{-1}\,\omega^{(1)}, \tag{7.7}$$

from which it is seen that, for the vector case, the estimator of the missing observation is a weighted average of the estimators obtained by treating each missing observation as the only one that is missing; i.e., by applying the DACF to the arbitrarily filled series. The weights are the elements of the inverse dual autocorrelation matrix. [For stationary series, this inverse matrix may provide a crude approximation to the autocorrelation matrix; see Battaglia (1983)].

To see that expression (7.6) does not depend on the arbitrary vector Z, write

$$\rho^D(B)\, Z_T = \begin{bmatrix} Z_T + \sum_i \rho_i^D (Z_{T+i} + Z_{T-i}) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ Z_{T-k} + \sum_i \rho_i^D (Z_{T-k+i} + Z_{T-k-i}) \end{bmatrix} =$$

$$= Z + (B_1\, B_2\, B_3) \begin{bmatrix} Z^- \\ Z \\ Z^+ \end{bmatrix} = (I + B_2)\, Z + B_1\, Z^- + B_3\, Z^+,$$

where Z^- and Z^+ contain observations prior to $T-k$ and posterior to T, respectively. (Thus Z^- and Z^+ are the available observations in the series z_t). The matrix B_2 is easily seen to be equal to $R^D - I$, thus

$$\rho^D(B)\, Z_T = R^D Z + B_1\, Z^- + B_3\, Z^+,$$

and, from (7.5),

$$\hat{\omega} = Z + (R^D)^{-1}\,(B_1 Z^- + B_3 Z^+).$$

Plugging this expression in (7.6) it follows that the estimator $\hat{z}$ does not depend on Z, the vector of arbitrary numbers.

Finally, since the MSE of $\hat{\omega}$ in (7.2) is the matrix $(x'\,x)^{-1}$, from (7.3) it follows that

$$\text{MSE}(\hat{z}) = \text{MSE}(\hat{\omega}) = (\Omega^D)^{-1},$$

where Ω^D is the dual autocovariance matrix.

As an example, table 4 presents the MSE of the estimators of the missing observations in an AR(1) model for the case of a block of 3 and a block of 4 missing values. In the latter case the estimators have, naturally, larger MSE. As expected, the largest uncertainty (MSE) corresponds to the center observations. Also, as in the single missing observation case, the MSE are smallest and the estimator most precise when $\phi = 1$.

INSERT	Table 4

7.2 Finite Series; the General Case

Equations (7.5) and (7.6) were derived for a complete realization of the series z_t, with missing observations at periods $T, T-1, \cdots, T-k$. Assume now that, similarly to section 5, the last observation available is for period $T+n$. Equations (7.6) and (7.7) remain unchanged except that $\omega^{(1)}$ becomes $\omega_n^{(1)}$, and contains now the vector of estimators obtained by assuming successively that each missing observation is the only missing one and applying equation (6.2). The matrix R_n^D is given by

$$R_n^D = \begin{bmatrix} 1 & \rho_{1,n}^D & \rho_{2,n}^D & \cdots & \rho_{k,n}^D \\ \rho_{-1,n+1}^D & 1 & \rho_{1,n+1}^D & \cdots & \rho_{k-1,n+1}^D \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \rho_{-k,n+k}^D & \rho_{-k+1,n+k}^D & \cdots & \cdots & 1 \end{bmatrix}, \tag{7.8}$$

where $\rho_{i,j}^D$ is the coefficient of B^i in $(1/V_j)\ \pi(B)\ \pi_j(F)$. The MSE of the vector of missing observations becomes

$$\Omega_n^D = \begin{bmatrix} V_n^D & \gamma_{1,n}^D & \gamma_{2,n}^D & \cdots & \gamma_{k,n}^D \\ \gamma_{-1,n+1}^D & V_{n+1}^D & \gamma_{1,n+1}^D & \cdots & \gamma_{k-1,n+1}^D \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \gamma_{-k,n+k}^D & \gamma_{-k+1,n+k}^D & \cdots & \cdots & V_{n+k}^D \end{bmatrix},$$

where $\gamma_{i,j}^D = V_j^D\ \rho_{i,j}^D$. The matrix Ω_n^D is a symmetric matrix since $\gamma_{-i+j,n+i}^D = \gamma_{-j+i,n+j}^D$ for $i,\ j = 0,\ 1,\ \cdots,\ k$.

Finally, assume in all generality, that the series z_t has $k+1$ missing observations for periods T, $T-m_1$, $T-m_2$, $\cdots$, $T-m_k$, where $m_1 < m_2 < \cdots < m_k$. Proceeding as before, that is, by arbitrarily filling the holes in the series, treating these arbitrary numbers as outliers and removing their effect, the same equations (7.5) and (7.6) are obtained. The matrix R^D of (7.4) becomes

$$R^D = \begin{bmatrix} 1 & \rho_{m_1}^D & \rho_{m_2}^D & \cdots & \rho_{m_k}^D \\ & 1 & \rho_{m_2-m_1}^D & \cdots & \rho_{m_k-m_1}^D \\ & & 1 & \cdots & \rho_{m_k-m_2}^D \\ & & \cdot & & \cdot \\ & & & \cdot & \cdot \\ & & & & \cdot \\ & & & & 1 \end{bmatrix}, \tag{7.9}$$

where ρ_j^D denotes the coefficient of B^j in the polynomial $\rho^D(B)$, and the subindices of the dual autocorrelations in (7.9) reflect the time distances between each pair of missing observations. The MSE of the estimator is equal to the inverse of the dual autocovariance matrix associated with (7.9). If the last observation of the series is for period $T+n$, the autocorrelations ρ_i^D in row j of the matrix R^D in (7.9) would be replaced by $\rho_{i,n+j-1}^D$, the coefficient of B^i in the expression $(1/V_{n+j-1}^D)\ \Pi(B)\ \Pi_{n+j-1}(F)$.

To illustrate (7.9), assume the series z_t has missing observations for $t = T, T+1$ and $T+4$. The matrix R^D is then equal to

$$R^D = \begin{pmatrix} 1 & \rho_1^D & \rho_4^D \\ \rho_1^D & 1 & \rho_3^D \\ \rho_4^D & \rho_3^D & 1 \end{pmatrix},$$

and $\hat{\omega} = (\hat{\omega}_0, \hat{\omega}_1, \hat{\omega}_2)$ is given by

$$\hat{\omega} = (R^D)^{-1}\, \rho^D(B) \begin{bmatrix} Z_T \\ Z_{T+1} \\ Z_{T+4} \end{bmatrix}. \tag{7.10}$$

Dropping, for notational simplicity, the superscript "D" from the dual autocorrelations, the estimator $\hat{\omega}_0$ is found to be $\hat{\omega}_0 = |R|^{-1}\,[(1-\rho_3^2)\,\rho(B)\,Z_T - (\rho_1 - \rho_3\,\rho_4)\,\rho(B)\,Z_{T+1} + (\rho_1\,\rho_3 - \rho_4)\,\rho(B)\,Z_{T+4}]$,

where

$$|R| = 1 + 2\rho_1\,\rho_3\,\rho_4 - \rho_1^2 - \rho_3^2 - \rho_4^2.$$

Since the coefficient of Z_T in $\rho(B)\,Z_T$, $\rho(B)\,Z_{T+1}$, and $\rho(B)\,Z_{T+4}$ is, respectively, 1, ρ_1 and ρ_4, it is easily seen that the coefficient of Z_T in (7.10) is 1. Similarly, the coefficients of Z_{T+1} and Z_{T+4} are seen to be zero, so that the estimator of z_T

$$\hat{z}_T = Z_T - \hat{\omega}_0$$

does not depend on the three arbitrary numbers Z_T, Z_{T+1}, and Z_{T+4}.

As a final example, a particular case of estimating sequences of missing observations is the problem of interpolation when there is only available one observation at equally spaced intervals. Consider interpolation of quarterly data generated from a random walk when only one observation per year is available. The models for the series and its dual are given by

$$\Delta z_t = a_t; \qquad z_t^D = (1 - B)\, a_t,$$

so that the dual autocorrelations are $\rho_1 = -.5$ and $\rho_k = 0,\ k \neq 0, 1$. The matrix R^D of (7.9) is seen to be block diagonal, where the blocks are all equal to the (3×3) symmetric matrix

$$R_1^D = \begin{pmatrix} 1 & -.5 & 0 \\ & 1 & -.5 \\ & & 1 \end{pmatrix}.$$

Expression (7.5) consists of a set of uncoupled systems of 3 equations, corresponding to the 3 holes in each year. Let Z_0 and Z_4 denote two successive annual observations (i.e., $Z_0 = z_0,\ Z_4 = z_4$), and $Z = (Z_1, Z_2, Z_3)'$ denote the vector of arbitrary numbers that fill the unobserved quarters. Each system of equations is of the form

$$\hat{\omega} = (R_1^D)^{-1}\,(Z - v),$$

where v is a vector with the jth element given by $(Z_{j-1}+Z_{j+1})/2,\ j = 1, 2, 3$. From $\hat{z} = Z - \hat{\omega}$ it is then obtained:

$$\begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_3 \end{bmatrix} = \begin{bmatrix} Z_1 & - & \hat{\omega}_1 \\ Z_2 & - & \hat{\omega}_2 \\ Z_3 & - & \hat{\omega}_3 \end{bmatrix} = \begin{bmatrix} 3/4\, z_0 & + & 1/4\, z_4 \\ 1/2\, z_0 & + & 1/2\, z_4 \\ 1/4\, z_0 & + & 3/4\, z_4 \end{bmatrix},$$

which is the linear interpolation formula obtained by Nerlove, Grether and Carvalho (1979, pp. 101–102). Since the variance of z_t^D is $V^D = 2$, the MSE of $\hat{z}$, equal to $(V^D\, R_1^D)^{-1}$, becomes the symmetric matrix

$$\text{MSE}\,(\hat{z}) = \begin{bmatrix} .75 & .50 & .25 \\ & 1 & .50 \\ & & .75 \end{bmatrix}.$$

8 An Application

When the model is known, estimation of the missing observations by regression with additive outliers, as described in the previous sections, can be seen as a method to compute the conditional expectation of the missing value given the available observations. It provides thus an alternative procedure to the fixed point smoother used in the standard approach to missing observations estimation [see Anderson and Moore (1979), and Harvey and Pierse (1984) for its extension to nonstationary series]. In practice, when the model is not known, the regression parameters associated with the outlier effects are typically concentrated out of the likelihood. As a consequence, one may wonder whether, when the model is not known, the two approaches:

(a) Maximization of an appropriately defined likelihood function with the Kalman filter and application of the fixed point smoother;

(b) Estimation of missing observations by regression, filling the holes with additive outliers;

may still yield results that are reasonably close. Notice that the outlier approach is a particularly simple case of the so–called Intervention Analysis models of Box and Tiao (1975).

Differences between the two procedures would be mostly due to differences between the "missing observation" likelihood and the "additive outlier" likelihood. Comparing the two likelihoods [Ljung (1989), Peña (1987)], the term comprising the sum of squares can be seen to be, in both cases, the same; what differs is a determinant. This difference, however, becomes smaller and smaller as the length of the series increases relative to the number of missing observations. Moreover, since the determinant in question is readily obtained, the additive outlier likelihood can be corrected by this factor, so as to obtain the likelihood of the missing observations case.

To compare the two approaches, we consider the same series as Harvey and Pierse (1984) and Kohn and Ansley (1986): the series of airlines passengers analysed by Box and Jenkins (1970). It consists of 144 monthly observations, for which a model of the type (2.10) is appropriate for the logs. Our aim is to compare the standard approach to missing observations estimation represented by the method of Kohn and Ansley (1986), with the additive outlier regression approach with and without the correction in the determinant mentioned above. The three approaches will be denoted, respectively, the Fixed–Point–Smoother/Missing Observation (FPS/MO) approach, the Additive Outlier/Missing Observation (AO/MO) approach, and the Additive Outlier/Regression (AO/REG) approach.

In order to homogenize comparisions, all computations have been made with a program named TRAM ("Time Series Regression with Arima Noise and Missing Observations"), written in Fortran, and described in Gómez and Maravall (1992b). (The program, together with the necessary documentation, is available from the authors upon request.) Very briefly the three approaches of interest are handled by TRAM in the following way:

(a) The FPS/MO method produces the missing observations estimators of Kohn and Ansley (1986) and of Harvey and Pierse (1984), when the latter is applicable. Only the available observations are used to define the likelihood and, once the model has been estimated, missing observations are obtained through the fixed point smoother. The method in TRAM is based on an alternative definition of the likelihood, which permits a direct and standard state space representation of the (original) nonstationary series. In this way, the ordinary Kalman filter and ordinary fixed–point–smoother are efficiently used for estimation, forecasting, and interpolation. The methodology is described in Gómez and Maravall (1992a); Bell and Hillmer (1991) have also shown how suitable initialization of the ordinary Kalman filter can yield the same results as the complex approach of Kohn and Ansley (1986).

(b) The AO/MO method fills the holes in the series corresponding to the missing observations with initial values. Each one of these values is then treated as an additive outlier, that is, as a regression dummy variable. The fitted value in the regression is the missing observation estimator. The regression parameters are concentrated out of the likelihood, and are estimated by using, first, a Cholesky decomposition of the error covariance matrix to transform the regression equation (the Kalman filter provides an efficient algorithm to compute the variables in this new regression). Then, the resulting least–squares problem is solved by orthogonal matrix factorization using the Householder transformation. This procedure yields a numerically stable method to compute GLS estimators of the regression parameters, which avoids matrix inversion. At each iteration, the likelihood is computed with the ordinary Kalman filter, and then corrected by the appropriate determinantal factor, so that it becomes the missing observation likelihood.

(c) The AO/REG method for estimating missing observations is the same as the AO/MO one, except that no correction to the likelihood is made, and hence the additive outlier likelihood is maximized.

Some comments are in order:

The Additive Outlier formulation would a priori seem inefficient since the addition of regression variables increases the size of the model. Besides, the Additive Outlier approach requires the specification of initial values for the missing observations, which is not required in the FPS/MO approach. On the other hand, since it only implies the estimation of (impulse) dummy variables, it offers the advantage of its simplicity. Moreover, since by filling the holes in the series with initial values it becomes possible to difference the series, the algorithm of Morf, Sidhu and Kailath (1974) can be employed, which implies a gain in computational efficiency. Furthermore, one by–product of the Additive Outlier approach is the computation of the entire matrix of MSE for the vector of missing observations estimators, and not simply the MSE of each individual interpolator. This full matrix of MSE is of applied importance since, for example, it is required in order to compute confidence intervals for the rates of growth of the interpolated series, when there are several missing observations that are not too distant. The ordinary fixed point smoother does not offer this possibility since the covariances between estimators are not obtained; this limitation can be overcome by, for example, using the results on the matrixes of MSE obtained from the DACF, as explained in the previous sections. Doing so, however, increases the complexity of the FPS/MO approach.

Back to the Airline Model example, the first case we consider consists of one isolated missing observation for period $T = 103$ (July 1957). Table 5 presents the estimation results obtained with the three methods. In the two AO methods, the initial value of the missing observation has been set equal to .5 of the sum of the two adjacent observations. It is seen that the two methods FPS/MO and AO/MO yield the same results, which are very close to those obtained with the AO/REG method. The column "time" indicates the time needed for a 486 PC with 33 Mh to run the program (compiled with Microsoft Fortran compiler). Although an important percentage of this time is spent on additional operations that the program TRAM performs; these were practically identical for the three methods under comparison. In summary, for the case of a single missing observation, the Additive Outlier approach is as precise as the FPS/MO one, and certainly faster.

INSERT	Table 5

An application of the results obtained in the previous sections concerns the selection of the initial value when an Additive Outlier approach is used. Obviously, an optimal choice would be to use expression (2.7) for $\hat{z}_{103}$, with the DACF estimated from the available series. This procedure, however, involves nontrivial additional computations and, since the variability of the series is heavily dominated by the nonstationary autoregressive roots, a reasonable approximation, trivial to compute, is to simply use the filter associated with those unit roots. In this case, the function $\rho^D(B)$ becomes that of the model

$$z_t^D = (1 - B)\,(1 - B^{12})\,a_t, \tag{8.1}$$

and hence the filter has only a few nonzero terms and does not involve any unknown parameter. This procedure is equivalent to running the fixed point smoother on the model

$$\Delta\,\Delta_{12}\,z_t = a_t.$$

For the first example, however, the selection of the initial value had practically no effect on the estimation results.

Example 2 is the same as the one called Data Set 3 in Kohn and Ansley (1986). From the airline passenger series, five observations are removed for periods $T = 7$, 102, 103, 104, and139 (July 1949, June, July and August 1957, and July 1960). Table 6 presents the estimation results using the three methods. In the Additive Outlier cases, the initial values have been set equal to .5 of the sum of the two closest observations at both sides (the "naive" initialization). As in example 1, the FPS/MO and AO/MO methods yield identical interpolators, associated MSE, and parameter estimates (identical also to the values reported by Kohn and Ansley). These values are again very close to the ones obtained with the AO/REG method. As in example 1, the Additive Outlier approach is as precise and considerably faster than the standard (FPS/MO) approach.

INSERT	Table 6

The third example is the same as Data Set 4 in Kohn and Ansley (1986), and is as example 2 with all the July values removed. As seen in Kohn and Ansley, in this case the first missing observation (z_7) cannot be estimated and becomes a free parameter. All the July interpolations depend on this free parameter; the only estimable missing obervations are those for $T = 102$ and $T = 104$. Table 7 displays the estimation results. The 14 missing values (all the months of July, plus z_{102} and z_{104}) are filled with the naive initialization (one half of the sum of the closest values at both sides). As before, the FPS/MO and AO/MO methods yield the same results, equal also to those reported by Kohn and Ansley. The AO/REG method provides results that are considerably close. However, the increase in the number of missing observations and hence in the number of regression variables in the Additive Outlier approach implies that the use of a corrected or uncorrected likelihood has an effect (although small) on parameter estimation. As for computational efficiency, the Additive Outlier approach becomes now slower than the FPS/MO approach.

INSERT	Table 7

The fourth example is similar to Data Set 2 of Kohn and Ansley (1986) [it is also the example considered by Harvey and Pierse (1984)], although the total number of missing observations has been reduced. It consists of the airline passenger series with the months February to November removed from the last two years of the series (1959 and 1960). There are, thus, 20 missing observations: two arrays of 10 consecutive missing observations, separated by December and January values.

As mentioned earlier, the Additive Outlier approach requires initial values to fill the missing observations holes. In the AO/MO case, since the likelihood is that of the missing observations case (and hence equal to the FPS/MO likelihood), and the regression parameters are concentrated out of the likelihood, the parameters of the ARIMA model will not depend on the chosen initial values. Further, since the conditional expectation that provides the missing observations estimators is a function of the ARIMA model parameters, it follows that the interpolators will not be affected by the choice of the initial values. It can be seen that, for the AO/REG case (that is, when the likelihood is not corrected), the effect of using better initial values (such as the ones obtained from the DACF expressions) is negligible. Thus, in the AO/MO and AO/REG methods, naive initialization is used: the February to November values for 1959 are set equal to the average of the January and December 1959 values; similarly, the missing observations for 1960 are filled with the average of the January and December 1960 values.

Figure 6 displays the 20 interpolators obtained with the FPS/MO method, the 95% confidence interval, and the actual values of the (removed) series. The interpolator is seen to perform well, and all 20 values of the series lie comfortably within the confidence interval. Table 8 presents the results obtained with the three methods (the last column displays the standard error of the FPS/MO interpolator; differences in the standard errors computed with the three methods were minor). It is seen how the FPS/MO method and the AO/MO method yield exactly the same results. Use of the (uncorrected) Additive Outlier likelihood (i.e., the AO/REG method) yields slightly different estimates of the ARIMA model parameters, which translates into very small differences in the missing observations interpolators (although the root-mean-squared error remains practically unchanged).

INSERT	Table 8

Figure 7 displays the three series of interpolators: they are virtually indistinguishable. However, as evidenced in table 8, for this example with 20 missing observations, the FPS/MO method is markedly faster.

In summary, the examples we have discussed suggest the following:

(a) the standard approach to missing observations estimation, based on the Kalman filter computation of a likelihood function defined for the observed values, and on the fixed point smoother, and

(b) the Additive Outlier approach to missing observations estimation,

yield interpolators with very similar degrees of precision; this is particularly true when the likelihood in the Additive Outlier case is corrected by the determinant factor, so that it becomes equal to the missing observation likelihood.

When the number of missing observations is small, the Additive Outlier approach provides a computationally faster procedure. However, as the number of missing observations increases, the standard (Kalman filter–fixed point smoother) approach becomes relatively faster.

Since the differences in computing time are nevertheless moderate and would not be a major concern in most applications, the Additive Outlier approach seems to offer a valid alternative to the standard Kalman filter–fixed point smoother approach to missing observations estimation in time series. (Incidentally, the Additive Outlier method can be enforced with the widely available Intervention Analysis methodology.)

An advantage of the Additive Outlier approach is that, as mentioned previously, it provides an estimator of the full matrix of MSE for the estimators; this information is important in order to construct, for example, confidence intervals for the rates of growth of the interpolated series. Besides, unless one has available proper software (such as the program TRAM), the Additive Outlier specification is conceptually simpler. For example, estimating coefficient of dummy variables in (stationary or not) autoregressive models, which ultimately can be done simply by OLS, is certainly easier than moving to a state space representation, setting up the proper initialization of the filter, running the Kalman filter, maximizing the likelihood, and using a fixed point smoother.

9 Summary

We have considered the problem of estimating missing observations in time series that follow general nonstationary ARIMA models. Section 1 presents some background and a review of the literature most relevant to our discussion. In the first part of the paper the parameters of the ARIMA model are assumed known. The optimal estimator is the conditional expectation of the missing observations given the available ones and we concern ourselves with obtaining expressions for that expectation that explicitly show its dependence on the stochastic structure of the series; its relationship with other important statistical problems is also considered.

Section 2 presents the case of a single missing observation in a complete realization of the series and relates the optimal interpolator and its mean–squared–error to the Inverse or Dual Autocorrelation Function of the series. Section 3 shows how the filter that yields the missing observation estimator is identical to the one that removes the effect of an additive outlier, and in section 4 it is seen how, up to a proportionality factor, the filter that estimates the outlier effect is the same as the one that estimates the noise. Accordingly, the missing observation estimator is obtained by filtering the signal, in the signal plus noise decomposition of the series.

Section 5 presents an alternative derivation of the conditional expectation as a pooled estimator, and this is used in section 6 to obtain expressions for the estimator and its mean–squared error for the case of an observation near one of the extremes of the series (i.e., the case of a finite realization). Preliminary estimation and revisions are then discussed. It is seen, for example, how preliminary estimators that will suffer large revisions tend to converge fast to the final estimator, while slow convergence is associated with small revision errors. Section 7 extends the results, first, to a vector of consecutive observations and, finally, to the general case of any number of sequences of any length of missing observations (a particular case is interpolation of high frequency data when only low frequency data is observed).

It is shown how the optimal estimator can always be expressed, in a compact way, in terms of the (perhaps truncated) dual autocorrelation function; the mean–squared estimation error is equal to the inverse of the (appropriately chosen) dual autocovariance matrix. The estimator can also be seen as the result of the following procedure: First, fill the holes in the series with arbitrary numbers; then estimate each missing observation as if it was the only missing value in the arbitrarily filled series; and finally compute a weighted average of those estimates, where the weights are elements of the inverse dual autocorrelation matrix.

The last part of the paper — section 8 — considers an application where the ARIMA model parameters are not known. For a well-known example, three ways of estimating different patterns of missing observations are compared; two of the methods are based on an Additive Outlier (regression) approach, and the third one is the standard approach whereby the Kalman filter is used to compute an appropriately defined likelihood, and the fixed-point-smoother provides the interpolators. The comparison indicates that the three methods have similar precision in estimating missing values. When the number of missing observations is relatively small, the Additive Outlier methods provide a more efficient procedure, while the opposite is true when the number of missing values becomes large. Some additional advantages/disadvantages of the different approaches are also discussed.

References

Abraham, B. (1981), "Missing Observations in Time Series", *Communications in Statistics: Theory and Methods* A, 10, 1643–1653.

Anderson, B. and Moore, J. (1979), *Optimal Filtering*, New Jersey: Prentice Hall.

Battaglia, F. (1983), "Inverse Autocovariances and a Measure of Linear Determinism for a Stationary Process", *Journal of Time Series Analysis* 4, 2, 79–87.

Battaglia, F. and Bhansali, R.J. (1987), "Estimation of the Interpolation Error Variance and an Index of Linear Determinism", *Biometrika* 74, 771–779.

Bell, W.R. and Hillmer, S.C. (1991), "Initializing the Kalman Filter for Nonstationary Time Series Models", *Journal of Time Series Analysis* 4, 283–300.

Box, G.E.P., Hillmer, S.C. and Tiao, G.C. (1978), "Analysis and Modeling of Seasonal Time Series", in A. Zellner (ed.), *Seasonal Analysis of Economic Time Series*, Washington, D.C.: U.S. Dept. of Commerce — Bureau of the Census, 309–334.

Box, G.E.P. and Jenkins, G.M. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden–Day.

Box, G.E.P. and Tiao, G.C. (1975), "Intervention Analysis with Applications to Economic and Environmental Problems", *Journal of the American Statistical Association* 70, 71–79.

Brubacher, S.R. and Wilson, G.T. (1976), "Interpolating Time Series with Application to the Estimation of Holiday Effects on Electricity Demand", *Applied Statistics* 25, 2, 107–116.

Chang, I., Tiao, G.C. and Chen, C. (1988), "Estimation of Time Series Models in the Presence of Outliers", *Technometrics* 30, 2, 193–204.

Chatfield, C. (1979), "Inverse Autocorrelations", *Journal of the Royal Statistical Society* A, 142, 363–377.

Cleveland, W.P. (1972), "The Inverse Autocorrelations of a Time Series and their Applications", *Technometrics* 14, 277–298.

Damsleth, E. (1980), "Interpolating Missing Values in a Time Series", *Scandinavian Journal of Statistics* 7, 33–39.

De Jong, P. (1991), "Stable Algorithms for the State Space Model", *Journal of Time Series Analysis* 12, 143–154.

Gómez, V. and Maravall, A. (1992a), "Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter", EUI Working Paper ECO No. 92/80, Department of Economics, European University Institute.

Gómez, V. and Maravall, A. (1992b), "Time Series Regression with ARIMA Noise and Missing Observations: Progam TRAM", EUI Working Paper ECO No. 92/81, Department of Economics, European University Institute.

Gourieroux, C. and Monfort, A. (1990), *Séries Temporelles et Modèles Dynamiques*, Paris: Economica.

Grenander, U. and Rosenblatt, M. (1957), *Statistical Analysis of Stationary Time Series*, New York: John Wiley.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Harvey, A.C. and Pierse, R.G. (1984), "Estimating Missing Observations in Economic Time Series", *Journal of the American Statistical Association* 79, 125–131.

Hillmer, S.C. and Tiao, G.C. (1982), "An ARIMA-Model Based Approach to Seasonal Adjustment", *Journal of the American Statistical Association* 77, 63–70.

Jones, R. (1980) "Maximum Likelihood Fitting of ARMA Models to Time Series With Missing Observations", *Technometrics* 22, 389–395.

Kohn, R. and Ansley, C.F. (1986), "Estimation, Prediction and Interpolation for ARIMA Models with Missing Data", *Journal of the American Statistical Association* 81, 751–761.

Ljung, G.M. (1989), "A Note on the Estimation of Missing Values in Time Series", *Communications in Statistics, Simulation and Computation* 18, 2, 459–465.

Maravall, A. (1987), "On Minimum Mean Squared Error Estimation of the Noise in Unobserved Component Models", *Journal of Business and Economic Statistics* 5, 115–120.

Morf, M., Sidhu, G.S. and Kailath, T. (1974), "Some New Algorithms for Recursive Estimation on Constant, Linear, Discrete–Time Systems", IEEE *Transactions on Automatic Control*, AC — 19, 315–323.

Nerlove, M., Grether, D.M. and Carvalho, J.L. (1979), *Analysis of Economic Time Series: A Synthesis*, New York: Academic Press.

Peña, D. (1990), "Influential Observations in Time Series", *Journal of Business and Economic Statistics* 8, 235–241.

Peña, D. (1987), "Measuring the Importance of Outliers in ARIMA Models", in M. Puri et al. (eds), *New Perspectives in Theoretical and Applied Statistics*, New York: John Wiley.

Peña, D. and Maravall, A. (1991), "Interpolation, Outliers and Inverse Autocorrelations", *Communications in Statistics, Theory and Methods* 20, 3175–3186.

Pierce, D.A. (1970), "A Duality Between Autoregressive and Moving Average Processes Concerning their Least Squares Estimates", *Annals of Mathematical Statistics* 41, 442–462.

Shumway, R.H. and Stoffer, D.S. (1982), "An Approach to Time Series Smoothing and Forecasting with the EM Algorithm", *Journal of Time Series Analysis* 4, 253–264.

Whittle, P. (1963), *Prediction and Regulation by Linear Least-Square Methods*, London: English Universities Press.

Wincek, M.A. and Reinsel, G.C. (1986), "An Exact Maximum Likelihood Estimation Procedure for Regression–ARMA Time Series Models with Possibly Nonconsecutive Data", *Journal of the Royal Statistical Society* 48, 3, 303–313.

Winder, C.A. and Palm, F. (1989), "Intertemporal Consumer Behavior under Structural Changes in Income", *Econometric Reviews* 8, 1–87.

Table 1: RMSE of a Missing Observation Estimator(*); Airline Model

θ_1	θ_{12}						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
-0.9	0.068	0.130	0.165	0.189	0.205	0.216	0.222
-0.6	0.100	0.200	0.265	0.317	0.361	0.400	0.436
-0.3	0.132	0.265	0.350	0.418	0.477	0.529	0.577
0.0	0.158	0.316	0.418	0.500	0.570	0.632	0.689
0.3	0.180	0.361	0.477	0.570	0.650	0.721	0.786
0.6	0.200	0.400	0.529	0.632	0.721	0.800	0.872
0.9	0.215	0.431	0.571	0.684	0.781	0.869	0.949

(*) as a fraction of the innovation standard error

Table 2: Variance of the Total Revision in the Preliminary Estimator (*); Airline Model

θ_1	θ_{12}						
	-0.9	-0.6	-.3	0.0	0.3	0.6	0.9
-0.9	0.995	0.983	0.973	0.964	0.958	0.953	0.950
-0.6	0.990	0.960	0.930	0.900	0.870	0.840	0.810
-0.3	0.982	0.930	0.877	0.825	0.772	0.720	0.667
0.0	0.975	0.900	0.825	0.750	0.675	0.600	0.525
0.3	0.967	0.870	0.772	0.675	0.577	0.480	0.382
0.6	0.960	0.840	0.720	0.600	0.480	0.360	0.240
0.9	0.954	0.814	0.674	0.532	0.390	0.246	0.099

(*) as a fraction of the innovation variance

Table 3: Length of the Revision (in months); Airline Model

θ_1	θ_{12}						
	-0.9	-0.6	-0.3	0.0	0.3	0.6	0.9
-0.9	12	7	5	5	4	4	4
-0.6	13	13	13	13	13	5	2
-0.3	24	13	13	13	13	13	2
0.0	25	13	13	13	13	24	1
0.3	36	24	13	13	13	24	36
0.6	36	24	13	13	24	26	72
0.9	45	24	13	17	27	36	132

Table 4: MSE of the Estimator for Blocks of Missing Observations(*); AR(1) Model

	Block of 3			Block of 4			
	1st MO	2nd MO	3rd MO	1st MO	2nd MO	3rd MO	4th MO
AR(1) $\phi = 1$	0.750	1	0.750	0.8	1.2	1.2	0.8
AR(1) $\phi = 0.5$	0.988	1.176	0.988	0.997	1.232	1.232	0.997

(*) as a fraction of the innovation variance

Table 5: Example 1 (One Missing Observation). Estimation Results

Removed observation	FPS/MO		AO/MO		AO/REG	
Period Value						
103 6.142	6.156	(0.028)	6.156	(0.028)	6.156	(0.028)
Model parameters						
θ_1 = 0.402 (0.080)	0.401	(0.080)	0.401	(0.080)	0.399	(0.080)
θ_{12} = 0.557 (0.084)	0.556	(0.084)	0.556	(0.084)	0.555	(0.085)
V_a = 0.00137	0.00138		0.00138		0.00138	
Time (in sec.)	16.3		7.8		7.3	

(The standard errors are given in parenthesis)

Table 6: Example 2 (Five Missing Observations). Estimation Results

Removed observations		FPS/MO		AO/MO		AO/REG	
Period	Value						
7	4.997	5.013	(0.031)	5.013	(0.031)	5.013	(0.031)
102	6.045	6.024	(0.030)	6.024	(0.030)	6.024	(0.030)
103	6.142	6.147	(0.031)	6.147	(0.031)	6.148	(0.031)
104	6.146	6.148	(0.030)	6.148	(0.030)	6.148	(0.030)
139	6.433	6.409	(0.032)	6.409	(0.032)	6.409	(0.032)
Model parameters							
θ_1		0.405		0.405		0.397	
θ_{12}		0.566		0.566		0.562	
V_a		0.00140		0.00140		0.00140	
Time (in sec.)		18		10.2		9.8	

(The standard errors are given in parenthesis)

Table 7: Example 3 (Fourteen Missing Observations; Two Estimable ones). Estimation Results

Removed observations	FPS/MO	AO/MO	AO/REG
Period Value			
102 6.045	6.023 (0.030)	6.023 (0.030)	6.024 (0.030)
104 6.146	6.147 (0.030)	6.147 (0.030)	6.148 (0.030)
Model parameters			
θ_1	0.430	0.430	0.393
θ_{12}	0.573	0.573	0.571
V_a	0.00140	0.00140	0.00140
Time (in sec.)	15	19.4	21.6

(The standard errors are given in parenthesis)

Table 8: Example 4 (Twenty Missing Observations). Estimation Results

Removed observations		FPS/MO	AO/MO	AO/REG	SE of interpolator
Period	Value				
122	5.835	5.836	5.836	5.837	0.036
123	6.006	5.988	5.988	5.989	0.041
124	5.981	5.967	5.967	5.968	0.044
125	6.040	6.001	6.001	6.001	0.046
126	6.157	6.175	6.175	6.174	0.047
127	6.306	6.294	6.294	6.294	0.047
128	6.326	6.308	6.308	6.307	0.046
129	6.138	6.142	6.142	6.143	0.044
130	6.009	6.017	6.017	6.017	0.041
131	5.892	5.887	5.887	5.887	0.036
....	....	....	....	....	....
134	5.969	5.980	5.980	5.981	0.040
135	6.038	6.125	6.125	6.126	0.045
136	6.133	6.097	6.097	6.098	0.049
137	6.157	6.123	6.123	6.123	0.051
138	6.282	6.290	6.290	6.289	0.053
139	6.433	6.402	6.402	6.401	0.053
140	6.407	6.409	6.409	6.408	0.052
141	6.231	6.236	6.236	6.236	0.050
142	6.133	6.104	6.104	6.103	0.046
143	5.966	5.966	5.966	5.966	0.041
RMSE		0.0275	0.0275	0.0276	
Model parameters					
θ_1		0.356	0.355	0.334	
θ_{12}		0.557	0.557	0.570	
V_a		0.00140	0.00140	0.00140	
Time (in sec.)		16.1	22.6	22.2	

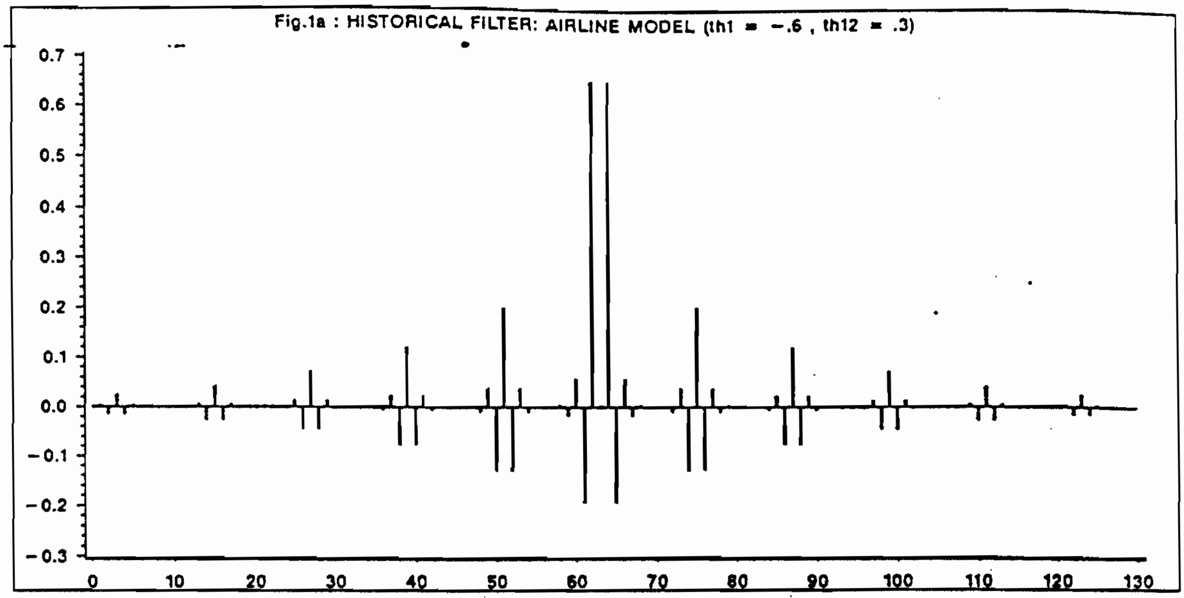
Fig.1a : HISTORICAL FILTER: AIRLINE MODEL (th1 = -.6 , th12 = .3)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

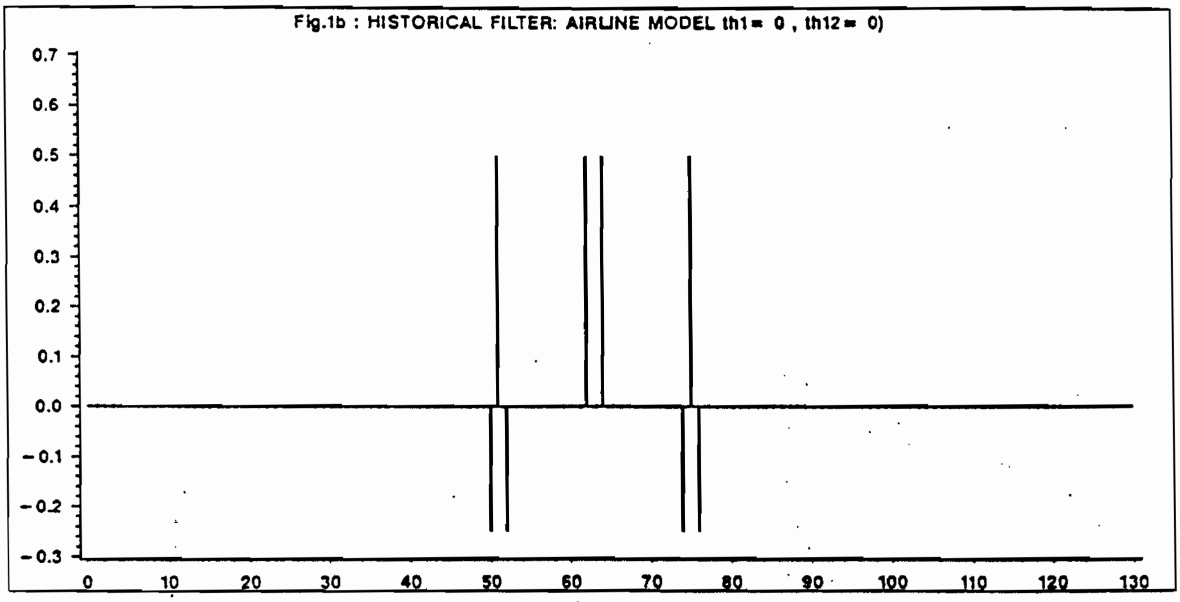
Fig.1b : HISTORICAL FILTER: AIRLINE MODEL th1= 0 , th12= 0)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

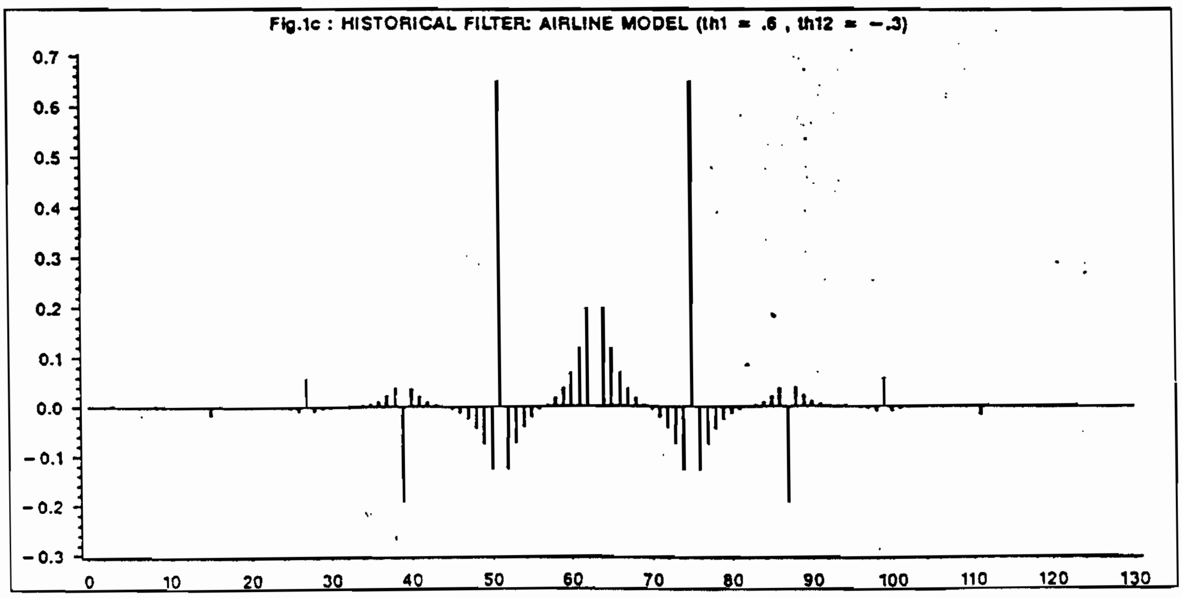
Fig.1c : HISTORICAL FILTER: AIRLINE MODEL (th1 = .6 , th12 = -.3)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

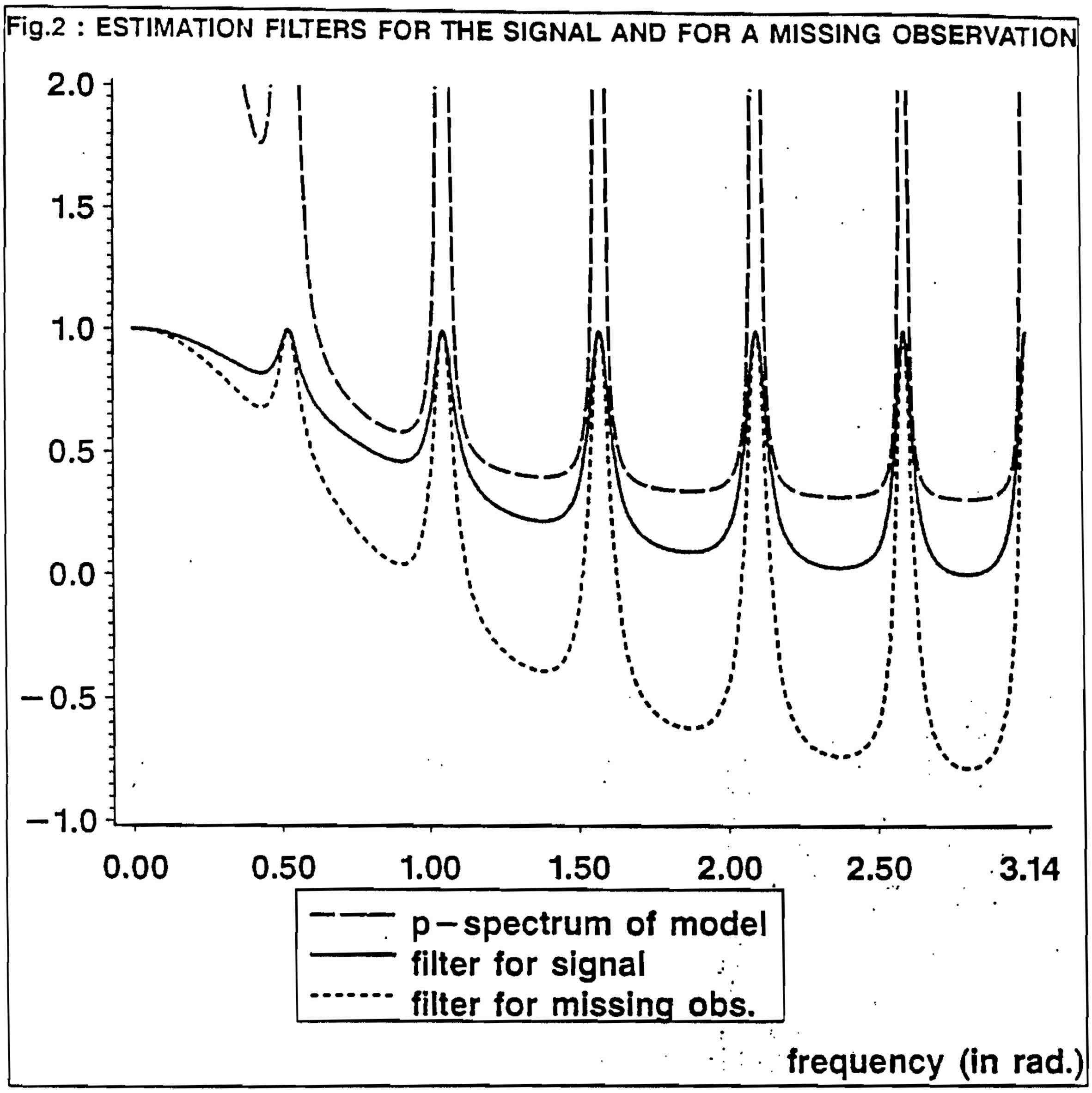
Fig.2 : ESTIMATION FILTERS FOR THE SIGNAL AND FOR A MISSING OBSERVATION
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0.00
0.50
1.00
1.50
2.00
2.50
3.14
p−spectrum of model
filter for signal
filter for missing obs.
frequency (in rad.)

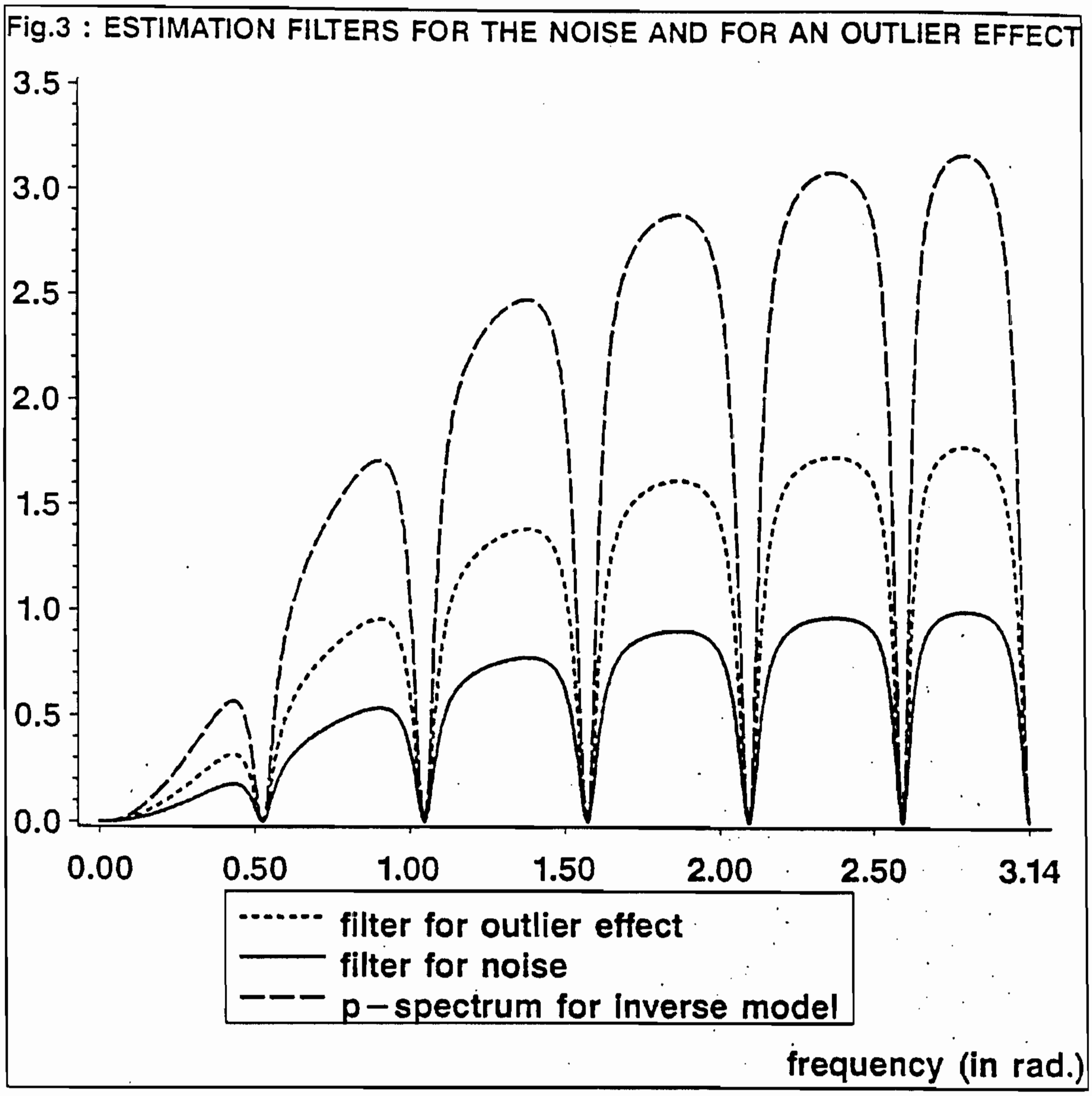
Fig.3 : ESTIMATION FILTERS FOR THE NOISE AND FOR AN OUTLIER EFFECT
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.00
0.50
1.00
1.50
2.00
2.50
3.14
filter for outlier effect
filter for noise
p−spectrum for inverse model
frequency (in rad.)

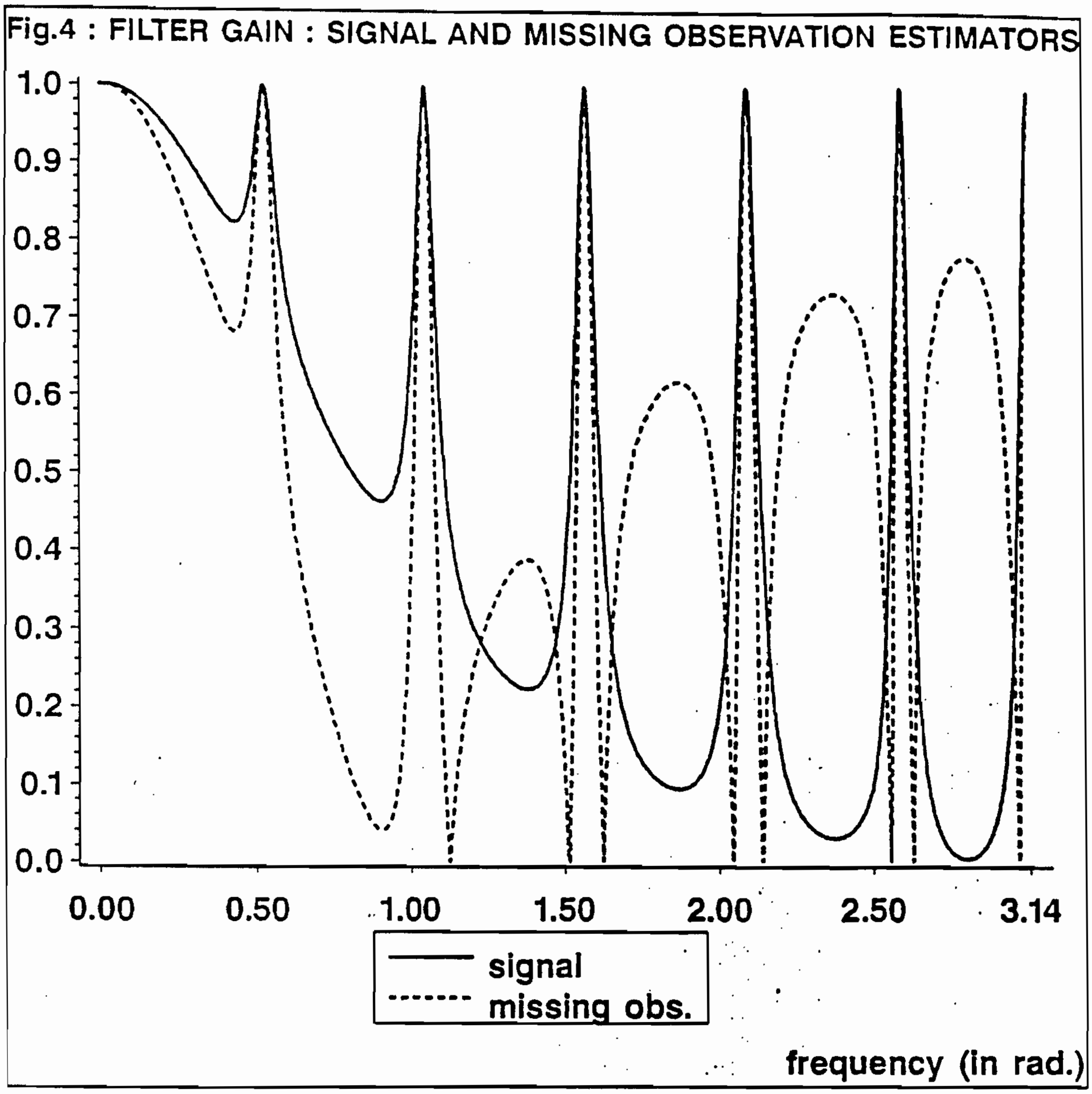
Fig.4 : FILTER GAIN : SIGNAL AND MISSING OBSERVATION ESTIMATORS
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00
0.50
1.00
1.50
2.00
2.50
3.14
signal
missing obs.
frequency (in rad.)

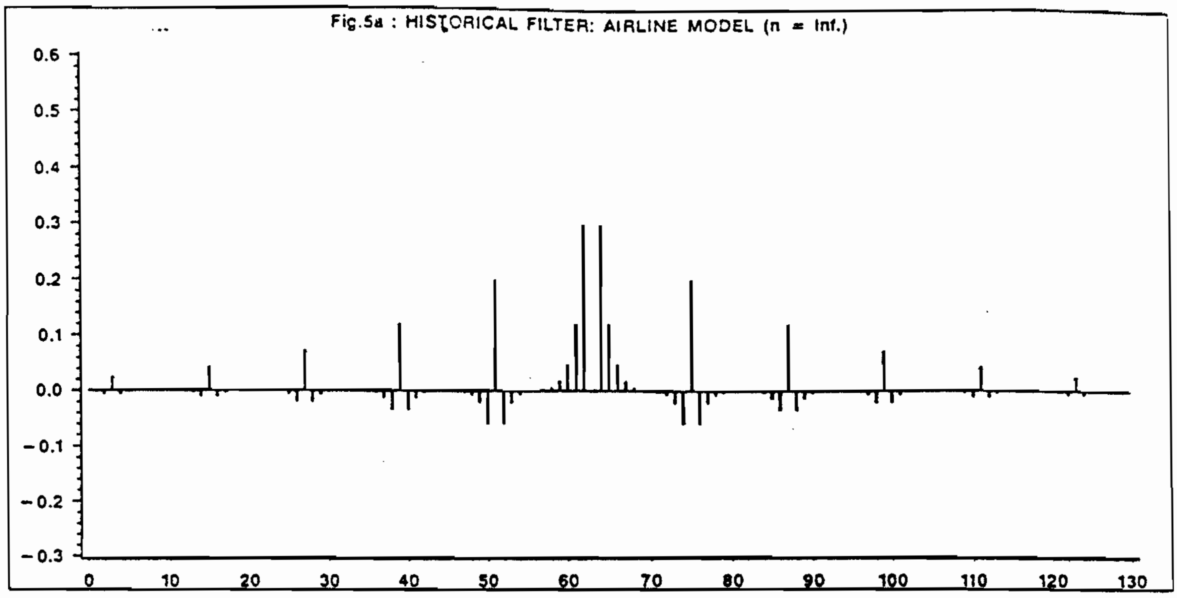
Fig.5a : HISTORICAL FILTER: AIRLINE MODEL (n = Inf.)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

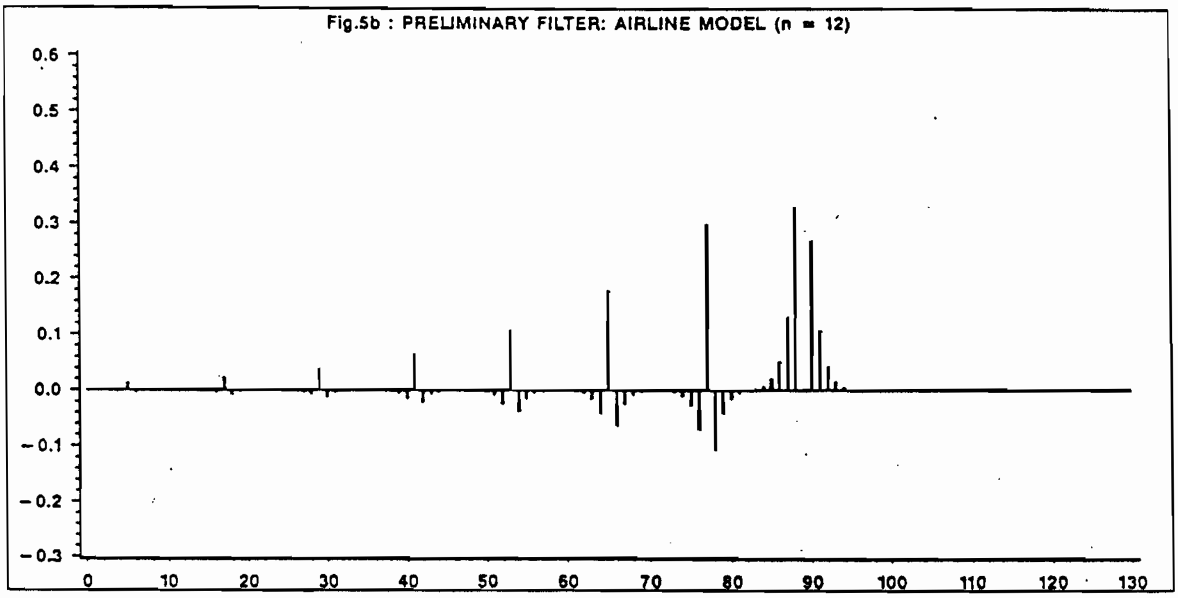
Fig.5b : PRELIMINARY FILTER: AIRLINE MODEL (n = 12)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

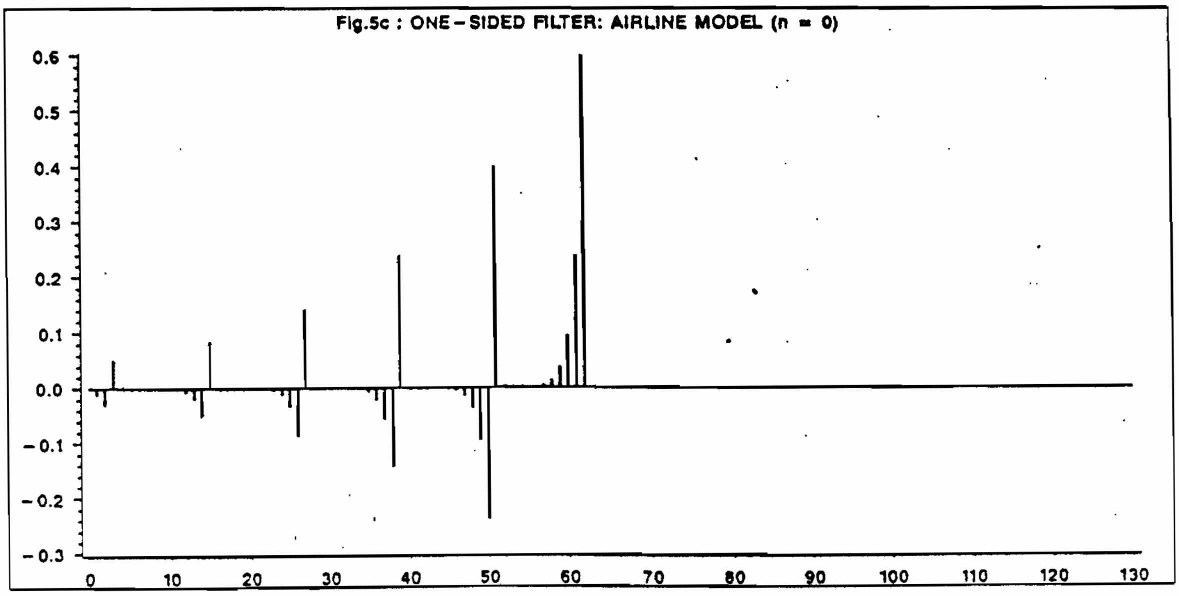
Fig.5c : ONE−SIDED FILTER: AIRLINE MODEL (n = 0)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0
10
20
30
40
50
60
70
80
90
100
110
120
130

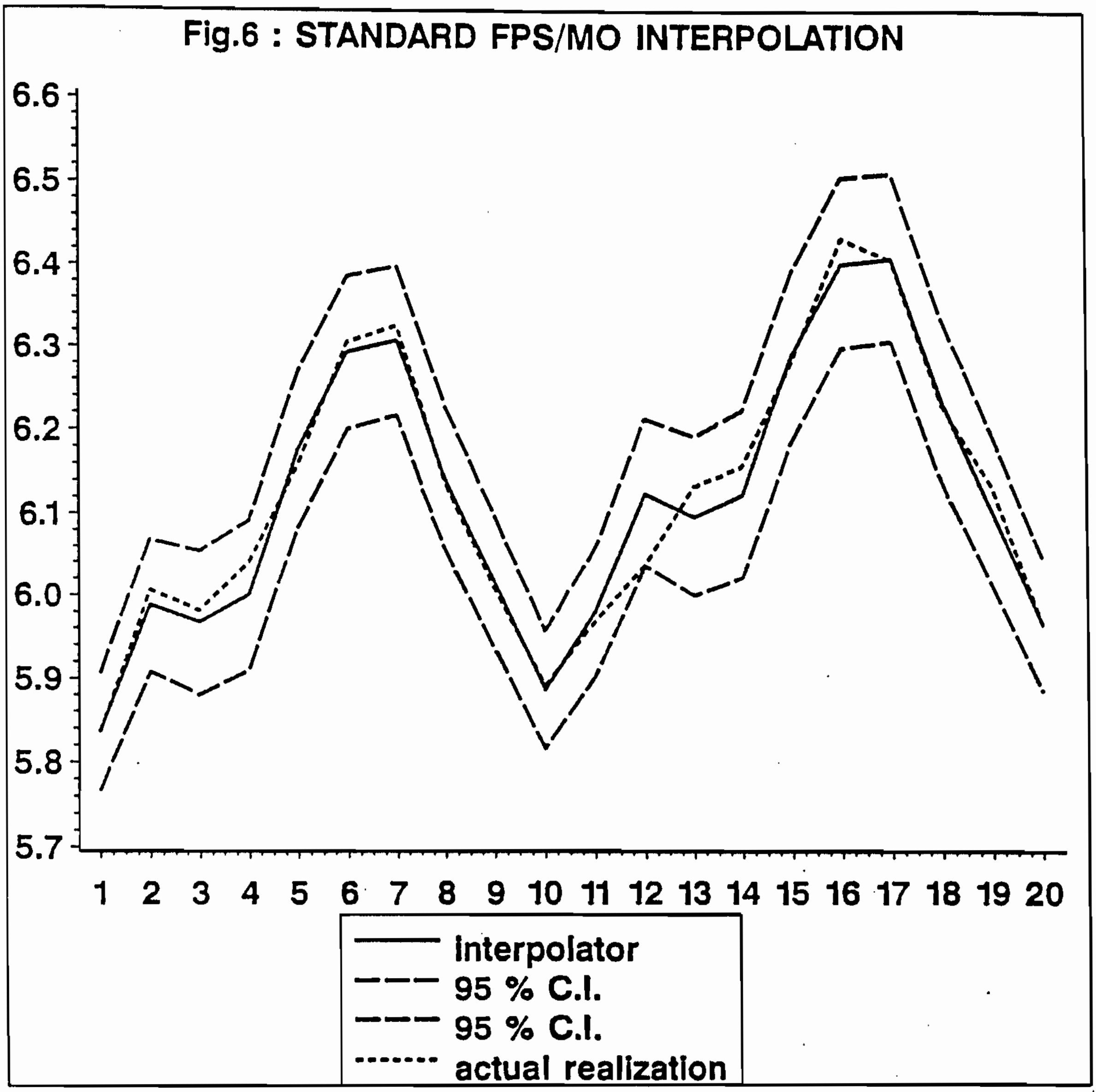
Fig.6 : STANDARD FPS/MO INTERPOLATION
6.6
6.5
6.4
6.3
6.2
6.1
6.0
5.9
5.8
5.7
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
interpolator
95 % C.I.
95 % C.I.
actual realization

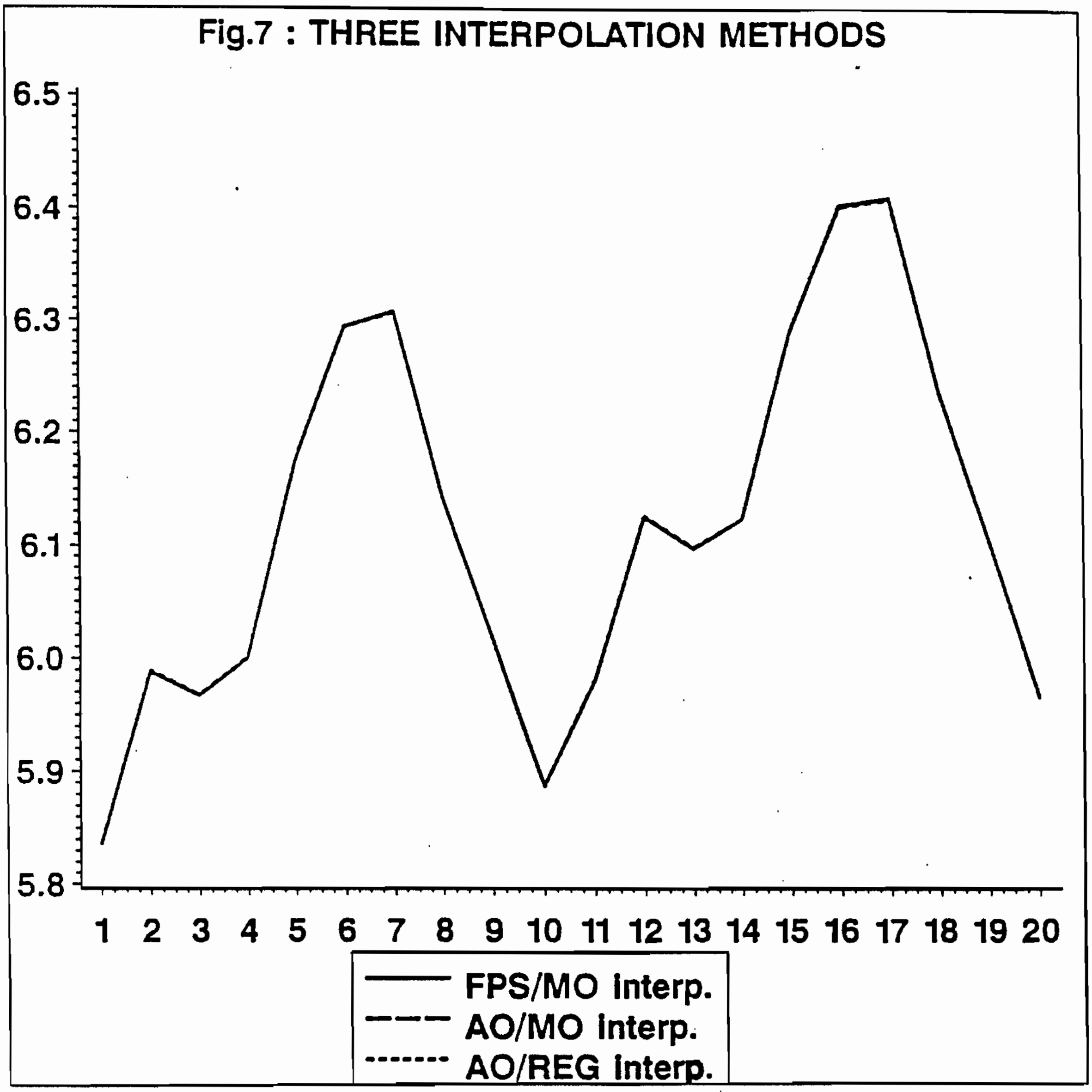
Fig.7 : THREE INTERPOLATION METHODS
6.5
6.4
6.3
6.2
6.1
6.0
5.9
5.8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
FPS/MO interp.
AO/MO interp.
AO/REG interp.