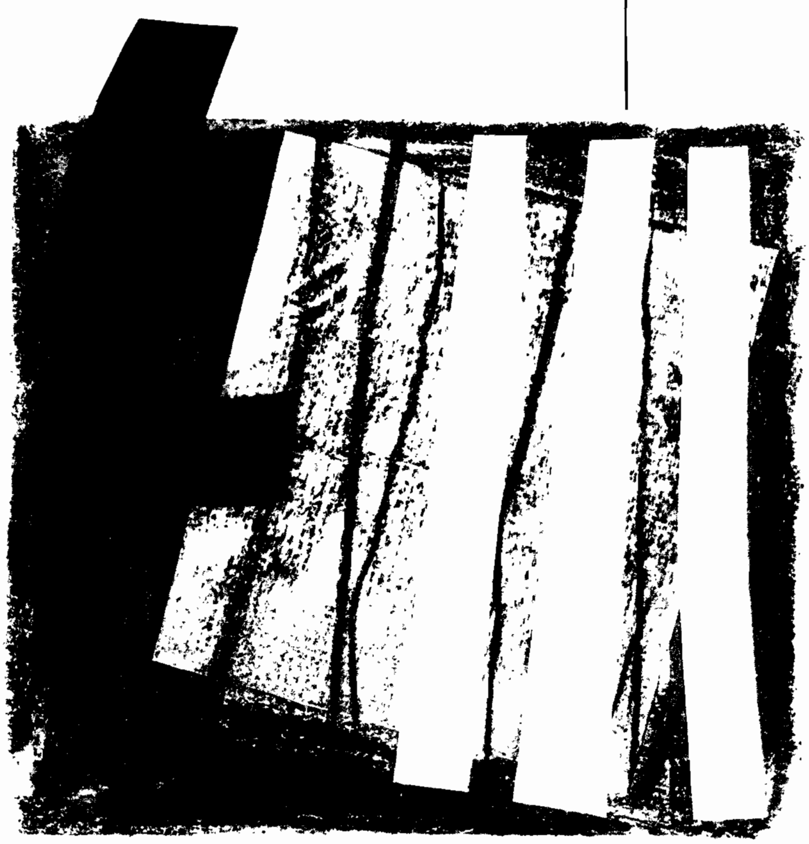
LINEAR COMBINATION OF INFORMATION IN TIME SERIES ANALYSIS

Víctor M. Guerrero and Daniel Peña

95-52



WORKING PAPERS

Universidad Carlos III de Madrid

Working Paper 95-52
Statistics and Econometrics Series 18
November 1995

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

LINEAR COMBINATION OF INFORMATION IN TIME SERIES ANALYSIS

Víctor M. Guerrero y Daniel Peña*

Abstract

An important tool in time series analysis is that of combining information in an optimal manner. Here we establish a basic combining rule of linear estimators and exemplify its use with several different problems faced by a time series analyst. A compatibility test statistic is also provided as a companion of the combining rule. This statistic plays a fundamental role for obtaining sensible results from the combination and for pointing out some possibly new directions of analysis.

Key Words
Compatibility testing; Disaggregation; Intervention analysis; Missing data; Outliers; Restricted forecasts.

*Guerrero, Instituto Tecnológico Autónomo de México; Peña, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid.

Linear Combination of Information in Time Series Analysis

An important tool in time series analysis is that of combining information in an optimal manner. Here we establish a basic combining rule of linear estimators and exemplify its use with several different problems faced by a time series analyst. A compatibility test statistic is also provided as a companion of the combining rule. This statistic plays a fundamental role for obtaining sensible results from the combination and for pointing out some possibly new directions of analysis.

KEY WORDS: Compatibility testing; Disaggregation; Intervention analysis; Missing data; Outliers; Restricted forecasts.

1. INTRODUCTION

Combining information has such a common place in the practice of statistics that the practicing statistician many times overlooks it. Hedges & Olkin (1985) presented many statistical problems that can be analyzed from this point of view. Draper et al. (1992) provided a thorough review of this field with many examples and ideas for future research. Similarly, Peña (1994) considered combining information with emphasis on understanding the structure and properties of the estimators involved in the combination.

This paper presents a basic (least squares) rule that has been frequently used by time series analysts for combining information. We consider here that some information, additional to that employed by a time series model, is available in the form of linear restrictions that have to be fulfilled exactly by an optimal estimator. Our basic concern is to obtain (conditionally) unbiased Minimum Mean Square Error Linear Estimators (MMSELE) of random vectors. Hence no distributional assumption will be required for obtaining the optimal estimators, although when normality is a reasonable assumption, the linear qualification can be dropped from MMSELE.

Each of the two sources of information is assumed to provide a linear and (conditionally) unbiased predictor. The unbiasedness assumption may be considered in some instances as unduly restrictive (see Palm & Zellner, 1992 or Min & Zellner, 1993). For our purposes and in the problems here considered, we deem such and assumption general enough and unrestrictive since debiasing can be carried out before combining. With respect to the use of only two sources of information, we remark that this is only to ease the exposition, but the ideas can be extended straightforwardly to several sources.

We assume that the models involved as well as their corresponding parameters are known, so that model building and parameter estimation are of no concern to us. On the other hand, an issue that does concern us is the practical implementation of the results provided by the combining rule. So we shall address that issue in due time. Furthermore, it should be said that throughout this paper we assume that the family of autoregressive integrated moving average (ARIMA) models is rich enough to represent the behavior of a univariate time series. Nevertheless, the results hold true for any linear time series model.

In Section 2 we establish a basic combining rule which will be used extensively in subsequent sections. A test statistic for validating the compatibility assumption between sources of information is also provided there. Section 3 applies and interprets the Basic Rule within the context of forecast updating and missing data estimation. Then, Section 4 concentrates on the problem of restricted forecasting, with several variants that respond to different states of knowledge about the future. Section 5 addresses the temporal disaggregation problem with and without preliminary series. In Section 6 we touch upon the problems of outliers and structural changes in time series. Section 7 is dedicated to the contemporaneous disaggregation of multiple time series. The final section concludes with some remarks and points out to the need of some other combining rules.

2. BASIC COMBINING RULE

In this section we present an optimal combining rule that can be employed when two basic sources of information are available. (1) A statistical model that produces the MMSELE, $\mathbf{W}$, of the random vector $\mathbf{Z}$, based on an observed set of explanatory variables X, and (2) some extra-model information $\mathbf{Y}$ given in the form of linear restrictions imposed on $\mathbf{Z}$. The model implied by (1) may be written as $\mathbf{Z} = X\beta + \mathbf{u}$, with β a fixed parameter vector and $\mathbf{u}$ a random vector, but this model form will not be used explicity in what follows. In fact, as indicated in the introduction we shall assume that the model is known, as well as its parameters. We now establish the rule and illustrate its use in different situations afterwards.

BASIC COMBINING RULE: Let us suppose that $\mathbf{Z}$, $\mathbf{W}$ and $\mathbf{Y}$ are related by

$$\mathbf{W} = \mathbf{Z} + \mathbf{e} \tag{2.1}$$

and

$$\mathbf{Y} = C\mathbf{Z}, \tag{2.2}$$

where $\mathbf{e}$ is a random vector such that $E(\mathbf{e}|X) = \mathbf{0}$, $E(\mathbf{Ze}'|X) = 0$, Cov $(\mathbf{e}|X) = \Sigma_e$ and $\mathbf{W} = E(\mathbf{Z}|X)$ is the MMSELE of $\mathbf{Z}$. If C is a known full-rank matrix and $\mathbf{W}$, $\mathbf{Y}$ and Σ_e are also known, with Σ_e nonsingular, then the MMSELE of $\mathbf{Z}$ based on $\mathbf{W}$ and $\mathbf{Y}$ is given by

$$\hat{\mathbf{Z}} = \mathbf{W} + \Sigma_e C'(C\Sigma_e C')^{-1}(\mathbf{Y} - C\mathbf{W}), \tag{2.3}$$

with MSE matrix

$$\mathrm{Cov}\left(\hat{\mathbf{Z}} - \mathbf{Z}\right) = \Sigma_e - \Sigma_e C'\left(C\Sigma_e C'\right)^{-1} C\Sigma_e . \tag{2.4}$$

Proof: Any linear estimator of $\mathbf{Z}$ based on $\mathbf{Y}$ and $\mathbf{W}$ must be of the form

$$\tilde{\mathbf{Z}} = A_1\mathbf{Y} + A_2\mathbf{W} = \left(A_1C + A_2\right)\mathbf{W} - A_1C\mathbf{e}, \tag{2.5}$$

with A_1 and A_2 some constant matrices. Then, to ensure (conditional) unbiasedness we require

$$E\left(\tilde{\mathbf{Z}} - \mathbf{Z}|X\right) = \left(A_1C + A_2 - I\right)\mathbf{W} = \mathbf{0}, \tag{2.6}$$

so that $\tilde{\mathbf{Z}} - \mathbf{Z} = (I - A_1 - C)\mathbf{e}$ and

$$\mathrm{Cov}\left(\tilde{\mathbf{Z}} - \mathbf{Z}\right) = \left(I - A_1C\right)\Sigma_e\left(I - A_1C\right)'. \tag{2.7}$$

Now, let us consider $A_1 = A + \Delta$ with $A = \Sigma_e C'(C\Sigma_e C')^{-1}$ and Δ an arbitrary constant matrix, then it follows that

$$\begin{aligned}\mathrm{Cov}(\tilde{\mathbf{Z}} - \mathbf{Z}) &= (I - AC)\Sigma_e(I - AC)' - (I - AC)\Sigma_e C'\Delta' - \Delta C\Sigma_e(I - AC)' + \Delta C\Sigma_e C'\Delta' \\ &= \Sigma_e - \Sigma_e C'(C\Sigma_e C')^{-1} C\Sigma_e + \Delta C\Sigma_e C'\Delta'. \qquad (2.8)\end{aligned}$$

Hence, the MSE matrix Cov $(\tilde{\mathbf{Z}} - \mathbf{Z})$ of any linear and (conditionally) unbiased estimator $\tilde{\mathbf{Z}}$ exceeds Cov $(\hat{\mathbf{Z}} - \mathbf{Z})$ by a positive semidefinite matrix.

Let us notice that $\mathrm{Cov}(\hat{\mathbf{Z}} - \mathbf{Z})$ is always singular because the estimator (2.3) satisfies the linear restriction (2.2). Thus it will have p null eingenvectors, where p is the number of columns in C', as can be seen by postmultiplying (2.4) by C'. Also notice that (2.2) leads to the estimator $\hat{\mathbf{Z}}_1 = PC'(CPC')^{-1}\mathbf{Y}$ in which P is a nonsingular matrix satisfying the (conditional) unbiasedness condition $[PC'(CPC')^{-1}C - I]\mathbf{W} = \mathbf{0}$. Another linear and (conditional) unbiased estimator of $\mathbf{Z}$ is provided by $\hat{\mathbf{Z}}_2 = \mathbf{W}$ from (2.1). Then, given $\mathbf{W}$ and $\mathbf{Y}$ we should combine linearly $\hat{\mathbf{Z}}_1$ and $\hat{\mathbf{Z}}_2$ in such an optimal manner that the MSE matrix of the resulting estimator is minimized. In so doing we go back to (2.3), so that $P = \Sigma_e$. Therefore we conclude that the MMSELE of $\mathbf{Z}$ given $\mathbf{W}$ and $\mathbf{Y}$, can be interpreted as a linear combination of the two linear and (conditional) unbiased estimators provided by the different sources of information.

This Basic Rule allows us to combine $\mathbf{W}$ and $\mathbf{Y}$ in an optimal manner. However, it does not necessarily follow that $\mathbf{W}$ and $\mathbf{Y}$ should always be combined. In particular, it will not be sensible to combine them when they contradict each other. Then, it makes sense to test for compatibility between $\mathbf{W}$ and $\mathbf{Y}$ to see whether or not the combined predictor is reasonable. To that end, a compatibility test derived on the assumption of normality for $\mathbf{e}$ was proposed by Guerrero (1989). That is, let us consider as null hypothesis $H_0: \mathbf{Y} = CE(\mathbf{Z}|X)$. On this hypothesis $\mathbf{Y} - C\mathbf{W}$ is normally distributed with mean vector zero and covariance matrix $C\Sigma_e C'$, therefore a statistic for testing compatibility between $\mathbf{Y}$ and $\mathbf{W}$ is given by

$$K = (\mathbf{Y} - C\mathbf{W})'(C\Sigma_e C')^{-1}(\mathbf{Y} - C\mathbf{W}) \sim \chi^2_m \qquad (2.9)$$

where m is the dimension of $\mathbf{Y}$.

It is important to realize that the Basic Rule as well as its companion compatibility test, can be obtained within a more general setting in which (2.2) is replaced by $\mathbf{Y} = C\mathbf{Z} + \mathbf{u}$ with $\mathbf{u}$ a random vector such that E($\mathbf{u}$)=$\mathbf{0}$ and Cov ($\mathbf{u}$)=U. However, in all cases considered here U = 0, so there is no need of considering $\mathbf{u}$ explicity. In another work we shall consider situations in which $U \neq 0$, so that a more general combining rule will be established together with its associated relevant analysis.

Now, when talking about modelling a univariate time series $\{Z_t\}$, we shall assume that it admits an ARIMA representation. We also let $X = (Z_1,\ldots,Z_N)'$ be the observed data and $\mathbf{Z} = (Z_{N+1},\ldots,Z_{N+H})'$ be the H>1 future values to be forecasted with origin at time N. Then we know that

$$Z_{N+h} - E(Z_{N+h}|X) = \sum_{j=0}^{h-1} \Psi_j a_{N+h-j}, \text{ for } h = 1,\ldots,H \tag{2.10}$$

where the Ψ_j's are the pure moving average (MA) weights of the model and $\{a_t\}$ is a zero-mean white noise process with variance σ^2. Expression (2.10) can be rewritten in matrix notation as

$$\mathbf{Z} - E(\mathbf{Z}|X) = \Psi\mathbf{a}, \tag{2.11}$$

with $\mathbf{a}' = (a_{N-1},\ldots,a_{N+H})'$ and Ψ a lower triangular matrix with $\Psi_0 = 1$ in the main diagonal, Ψ_1 in the second diagonal, and so on.

Notice in particular that (2.11) holds true both for stationary and nonstationary time series. For a stationary series with $E(\mathbf{Z}|X) = \mathbf{0}$ we have $\mathbf{e} = \Psi\mathbf{a}$ and $\Sigma_Z = \sigma^2\Psi\Psi'$. Also $\Pi\mathbf{Z} = \mathbf{a}$, where $\Pi = \Psi^{-1}$ is a lower triangular matrix with ones in the main diagonal, $-\pi_1$ in the second diagonal and so on. The π_i's are the pure autoregressive (AR) coefficients of the ARMA process. Then $\Sigma_Z^{-1} = \sigma^{-2}\Pi'\Pi$ is the inverse autocorrelation function of the process. In general we shall call $\mathbf{W} = E(\mathbf{Z}|X)$, $\mathbf{e} = \Psi\mathbf{a}$ and $\Sigma_e = \sigma^2\Psi\Psi'$, and for a stationary process $\Sigma_Z = \Sigma_e$. It should also be stressed that even though most of the problems considered in this paper make explicit reference to univariate time series, the same ideas can be employed with multiple time series. The basic change required in that situation, from a theoretical viewpoint, will be notational.

3. MAKING EFFICIENT USE OF ALL AVAILABLE DATA

This section presents two applications of the Basic Rule. Firstly in the well known case of forecast updating and secondly in a rather simple formulation of missing data estimation.

3.1 Forecast Updating

Here we consider the problem of updating a vector of ARIMA forecasts, initially obtained for lead times h=1, ..., H, with origin at time N. In this case, as soon as we have access to a new observation Z_{N+1}, its forecast $\hat{Z}_N(1)$, becomes useless and $\hat{Z}_N(2),\ldots,\hat{Z}_N(H)$ are suboptimal by not taking into account all the available information.

To recover the optimality of the ARIMA forecasts, given Z_{N+1}, we start with (2.11) which leads us to (2.1) with $\mathbf{W} = E(\mathbf{Z}|X)$, $\mathbf{e} = \Psi\mathbf{a}$ and $\Sigma_e = \sigma^2\Psi\Psi'$. Now let C=(1, 0') with $\mathbf{0}$ a column vector of size H-1 and $\mathbf{Y} = Z_{N+1}$, so that (2.3) and (2.4) yield

$$\hat{\mathbf{Z}} = \left(\hat{Z}_N(1) + \left[Z_{N+1} - \hat{Z}_N(1)\right], \hat{Z}_N(2) + \Psi_1\left[Z_{N+1} - \hat{Z}_N(1)\right], \ldots, \hat{Z}_N(H) + \Psi_{H-1}\left[Z_{N+1} - \hat{Z}_N(1)\right]\right)'$$
$$= \left(Z_{N+1}, \hat{Z}_{N+1}(1), \ldots, \hat{Z}_{N+1}(H-1)\right)' \tag{3.1}$$

and

$$\mathrm{Cov}\,(\hat{\mathbf{Z}} - \mathbf{Z}) = \sigma^2 \begin{pmatrix} 0 & | & 0' \\ - & - & - \\ -\Psi_1 & | & \\ \cdots & | & I_{H-1} \\ -\Psi_{H-1} & | & \end{pmatrix} \Psi\Psi' = \left\{v_{ij}\right\} \tag{3.2}$$

where $v_{1j} = v_{i1} = 0$ and $v_{ij} = \mathrm{Cov}\left(\hat{Z}_N(i-1) - Z_{N+i-1}, \hat{Z}_N(j-1) - Z_{N+j-1}\right)$ for $i, j = 2, \ldots, H$. These formulas are the usual ones when updating forecasts (cf. Box & Jenkins, 1976, Ch. 5). Of course, in general we could use m $(1 \le m < H)$ additional observations at once if they arrive in a batch or else we could update the forecasts recursively, the answer will be the same.

Next, the compatibility statistic in the situation $\mathbf{Y} = Z_{N+1}$ with $\{a_t\}$ Gaussian, lead us to declare the new observation compatible with the historical record of the series, at the 100 α% significance level, when

$$K = \left[Z_{N+1} - \hat{Z}_N(1)\right]^2 / \sigma^2 < \chi_1^2(\alpha). \tag{3.3}$$

When the new data arrive in batches, the K statistic provides a Cumulative Sum (CUSUM) test.

3.2 Estimation of Missing Data

We now consider the problem of completing a univariate time series which has some missing values. The basic need for doing that is that most time series analysis require a data set without gaps. Of course, a highly experienced analyst may get around of those gaps without the need of filling them with estimated values, but that is the exception, not the rule. Besides, in some cases the estimation of missing observations is the main objective of the analysis.

To pose the problem in the context of Section 2, let us suppose the series has $k \ge 1$ missing values (without any specific pattern). These form the vector $\mathbf{Z}_M = (Z_{T_1}, Z_{T_2}, \ldots, Z_{T_k})'$ in wich $T_i < T_j$ if $i < j$. Let also $X = \left(Z_1, \ldots, Z_{T_1 - 1}\right)'$ be the historical data before the first missing value, $\mathbf{Z}^* = \left(Z_{T_1}, \ldots, Z_N\right)'$ be the observations from T_1 onwards and $\mathbf{Z}_G = (Z_{T_1+1}, \ldots, Z_{T_2-1}, Z_{T_2+1}, \ldots, Z_{T_k-1}, Z_{T_k+1}, \ldots, Z_N)'$ the given values observed after X. Then we can write $\mathbf{Z} = (\mathbf{Z}'_M, \mathbf{Z}'_G)' = \Omega \mathbf{Z}^*$ with Ω a permutation matrix that merely changes the chronological ordering of $\mathbf{Z}^*$. That is, Ω is an $N - T_1 + 1$ square matrix obtained from the identity matrix, in which the rows T_i, $i = 1,\ldots,k$, of the identity are placed as the first k rows of Ω. Thus we have $\mathbf{W} = E(Z|X) = \Omega E(\mathbf{Z}^*|X)$ the ARIMA forecasts of $\mathbf{Z}_M$ and $\mathbf{Z}_G$, $\mathbf{e} = \Omega\Psi\mathbf{a}$ and

$$\Sigma_e = \sigma^2 \Omega\Psi\Psi'\Omega' = \sigma^2 \begin{pmatrix} \Sigma_M & \Sigma_{MG} \\ \Sigma'_{MG} & \Sigma_G \end{pmatrix}, \tag{3.4}$$

with $\sigma^2\Sigma_M = \text{Cov}[\mathbf{Z}_M - E(\mathbf{Z}_M|X)]$, $\sigma^2\Sigma_G = \text{Cov}[\mathbf{Z}_G - E(\mathbf{Z}_G|X)]$ and $\sigma^2\Sigma_{MG} = E\{[Z_M - E(Z_M|X)][Z_G - E(Z_G|X)]'\}$. Then, by choosing $C = (0, I_{N-T_1-k+1})$ with 0 an $(N - T_1 - k + 1) \times k$ matrix, we obtain $\mathbf{Y} = \mathbf{Z}_G$ and the weighting matrix

$$\Sigma_e C'(C\Sigma_e C')^{-1} = \begin{pmatrix} \Sigma_{MG}\Sigma_G^{-1} \\ I_{N-T_1-k+1} \end{pmatrix}. \tag{3.5}$$

By the Basic Rule, the MMSELE of $\mathbf{Z}$ based on $E(Z|X)$ and $\mathbf{Z}_G$ is given by

$$\hat{\mathbf{Z}} = \left(\{E(\mathbf{Z}_M|X) + \Sigma_{MG}\Sigma_G^{-1}[\mathbf{Z}_G - E(\mathbf{Z}_G|X)]\}', \mathbf{Z}'_G\right)', \tag{3.6}$$

with MSE matrix

$$\text{Cov}(\hat{\mathbf{Z}} - \mathbf{Z}) = \sigma^2 \begin{pmatrix} \Sigma_M - \Sigma_{MG}\Sigma_G^{-1}\Sigma'_{MG} & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.7}$$

Hence, it immediately follows that

$$\hat{\mathbf{Z}}_M = E(\mathbf{Z}_M|X) + \Sigma_{MG}\Sigma_G^{-1}[\mathbf{Z}_G - E(\mathbf{Z}_G|X)] \tag{3.8}$$

with

$$\text{Cov}(\hat{\mathbf{Z}}_M - \mathbf{Z}_M) = \sigma^2(\Sigma_M - \Sigma_{MG}\Sigma_G^{-1}\Sigma'_{MG}). \tag{3.9}$$

It is interesting to show that these expressions are equivalent to the ones obtained by Peña & Maravall (1991) using the inverse autocorrelation function. The relationship between additive outliers and missing observations first indicated by Peña (1987) leads to estimating the missing values by inserting zeros at the missing points and adding to the model dummy variables for each of these zero values to represent an additive outlier. Calling D the matrix of dummy variables, it can be shown (see Peña & Maravall, 1991) that the MMSELE of $\mathbf{Z}_M$ is given by

$$\hat{\mathbf{Z}}_M = -(D'\Sigma_e^{-1}D)^{-1}D'\Sigma_e^{-1}\mathbf{Z}_c \tag{3.10}$$

where $\mathbf{Z}_c$ denotes the completed series with zeros at the missing points. The MSE matrix of this estimator is then given by

$$\text{Cov}(\hat{\mathbf{Z}}_M - \mathbf{Z}_M) = \sigma^2 D'\Sigma_e^{-1}D. \tag{3.11}$$

To show that (3.10) and (3.8) are equivalent, let us first recall that $\Sigma_e^{-1} = \sigma^{-2}\Pi'\Pi$, so that (3.10) becomes

$$\hat{\mathbf{Z}}_M = -(D'\Pi'\Pi D)^{-1} D'\Pi'\Pi\mathbf{Z}_c \qquad (3.12)$$

Now, for simplicity let us suppose that the k missing values are located at T,T+1,...,T+k-1, so that $D = (0, I_k, 0)$ where I_k is the identity matrix and let us partition Π as follows

$$\Pi = \begin{pmatrix} \Pi_1 & 0 & 0 \\ \Pi_2 & \Pi_4 & 0 \\ \Pi_3 & \Pi_5 & \Pi_6 \end{pmatrix} \qquad (3.13)$$

with Π_1, Π_4 and Π_6 square matrices of dimensions T-1, k and N-T-k+1 respectively. While Π_2, Π_3 and Π_5 are rectangular arrays of appropriate dimensions. Similarly, let $\mathbf{Z}'_6 = (\mathbf{Z}'_I, 0', \mathbf{Z}'_G)$, so that (3.12) can be rewritten as

$$\hat{\mathbf{Z}}_M = -(\Pi'_4\Pi_4 + \Pi'_5\Pi_5)^{-1}\left[(\Pi'_4\Pi_2 + \Pi'_5\Pi_3)\mathbf{Z}_I + \Pi'_5\Pi_6\mathbf{Z}_G\right]. \qquad (3.14)$$

Then, inverting (3.4) by blocks, taking into account that $\Sigma_e^{-1} = \sigma^2\Pi'\Pi$, and using (3.13) we obtain $\Sigma_{MG}\Sigma_G^{-1} = -(\Pi'_4\Pi_4 + \Pi'_5\Pi_5)^{-1}\Pi'_5\Pi_6$. On the other hand, calling $\hat{\mathbf{Z}}'_{\bullet} = (E(\mathbf{Z}_M|X)', E(\mathbf{Z}_G|X)')$ and $\mathbf{Z}'_H = (\mathbf{Z}'_I, \mathbf{0}', \mathbf{0}')$ we have

$$\hat{\mathbf{Z}}_{\bullet} = -(D'_H\Pi'\Pi D_H)^{-1} D'_H\Pi'\Pi\mathbf{Z}_H \qquad (3.15)$$

where

$$D_H = \begin{pmatrix} 0 & I_k & 0 \\ 0 & 0 & I_{N-T-k+1} \end{pmatrix}. \qquad (3.16)$$

After some algebraic manipulations and inserting expressions (3.15) and (3.16) into (3.8), we obtain (3.14).

The derivation of expressions (3.4) to (3.9) is due to Guerrero (1993), although Alvarez, Delrieu & Jareño (1993) also indicated the possibility of using restricted estimation to estimate missing values. From a practical standpoint, it is important to realize that the matrices involved in (3.8) - (3.9) may be relatively large, thus a recursive solution such as the one indicated by Guerrero (1993) becomes more efficient. In that respect, Gómez, Maravall & Peña (1993) have compared several algorithms for estimation of missing values in time series. Finally, the statistic to test for compatibility between $\mathbf{Z}_G$ and $E(\mathbf{Z}_M|X)$ becomes

$$K = \left[\mathbf{Z}_G - E(\mathbf{Z}_G|X)\right]'\Sigma_G^{-1}\left[\mathbf{Z}_G - E(\mathbf{Z}_G|X)\right]/\sigma^2 \sim \chi^2_{N-T_1-k+1}. \qquad (3.17)$$

We should notice in particular that rejection of the compatibility assumption may be expected when the missing values occurred not just by chance, but are due to some exogenous intervention or structural changes in the time series. In those cases we might consider the possibilities described in Section 4.

4. RESTRICTED FORECASTING WITH BINDING CONSTRAINTS

The problem considered now has several variants which may be deemed as relevant possibilities derived from compatibility testing. These are: compatible sources (Section 4.1) and incompatible sources (Sections 4.2 - 4.4).

4.1 Restricted Forecasting without Uncertainty in the Restrictions

This case occurs when some restrictions to be imposed on the time series forecast are known to be true in advance. For instance we may consider imposing budget constraints, or else we may view this kind of application as a scenario (or what if) analysis. For instance, Guerrero (1989) mentions the problem of forecasting the monthly Financing Granted by the Mexican Bank System when **Y**, the total annual financing, was known in advance.

To apply the Basic Rule to this problem we use (2.11) again so that $\mathbf{W} = E(\mathbf{Z}|X)$, $\mathbf{e} = \Psi\mathbf{a}$ and $\Sigma_e = \sigma^2\Psi\Psi'$. Furthermore, the $m \geq 1$ restrictions in $\mathbf{Y} = C\mathbf{Z}$ are assumed to be linearly independent and coming from outside information (external to the model). On these conditions we readily obtain.

$$\hat{\mathbf{Z}} = E(\mathbf{Z}|X) + \Psi\Psi'C'(C\Psi\Psi'C')^{-1}[\mathbf{Y} - CE(\mathbf{Z}|X)] \qquad (4.1)$$

with

$$\mathrm{Cov}(\hat{\mathbf{Z}} - \mathbf{Z}) = \sigma^2[I - \Psi\Psi'C'(C\Psi\Psi'C')^{-1}C]\Psi\Psi'. \qquad (4.2)$$

The corresponding test statistic becomes

$$K = [\mathbf{Y} - CE(\mathbf{Z}|X)]'(C\Psi\Psi'C')^{-1}[\mathbf{Y} - CE(\mathbf{Z}|X)]/\sigma^2 \sim \chi^2_m. \qquad (4.3)$$

We emphasize here the importance of compatibility testing, since rejecting this hypothesis may lead to different relevant formulations, for instance to assume that **Y** is true and that $E(\mathbf{Z}|X)$ is not a valid forecast, because a structural change is foreseen during the forecast horizon (see Sections 4.2, 4.3 and 4.4).

4.2 Change Foreseen in the Deterministic Structure of the Model

Let us now suppose that a structural change in the deterministic structure of the time series model is foreseen to occur during the forecast horizon of interest. This idea may come from subject matter considerations, for example when an intervention is anticipated. This case may be considered as an ex-ante intervention analysis, in which the whole effect of the intervention is presumably accounted for by way of some linear restrictions on the future values of the series. Guerrero (1991) considered this situation and provided the solution and an example drawn from the Mexican Economy of the past decade.

We start first with a formulation that allows us to take into account the intervention effects, that is

$$Z_t^{(D)} = Z_t + D_{t,T}, \qquad (4.4)$$
where $\delta(B)D_{t,T} = \omega(B)S_{t,T}$ is a dynamic function of the intervention effects, with $\delta(B) = 1 - \delta_1 B - \ldots - \delta_r B^r$ and $\omega(B) = \omega_0 + \omega_1 B + \ldots + \omega_s B^s$ polynomials in the backshift operator B, $S_{t,T}$ is a step function that takes on the value 1 when $t \geq T$ and 0 otherwise, with T the time point at which the intervention effects start, so that $N < T \leq H$. Without loss of generality we assume T=N+1 (see Guerrero, 1991). Now let us notice that in order to determine the r+s+1 parameters involved in specifying $D_{t,T}$ we would require at least as many linear independent restrictions to be provided by **Y**. So, since in practice we usually have access only to one or two such restrictions, we are forced to postulate at most a first order linear dynamic model
$$(1 - \delta B)D_{t,N+1} = \omega S_{t,N+1} \qquad (4.5)$$
which is still a very flexible model.

To employ the Basic Rule let $\mathbf{Z^*} = (Z_{N+1}, \ldots, Z_{N+H})'$ be the vector of future values without intervention effects, $\mathbf{D} = (D_{N+1,N+1}, \ldots, D_{N+H,N+1})'$ the deterministic effects and $\mathbf{Z}^{(D)} = (Z_{N+1}^{(D)}, \ldots, Z_{N+H}^{(D)})' = \mathbf{Z}^* + \mathbf{D}$. (2.11) is again assumed to hold with $\mathbf{Z}^{(D)}$ in place of $\mathbf{Z}$, in such a way that $\mathbf{W} = E(\mathbf{Z}^*|X) + \mathbf{D}$, $\mathbf{e} = \Psi\mathbf{a}$ and $\Sigma_e = \sigma^2\Psi\Psi'$ in (2.1). Then, when $\mathbf{Y} = C\mathbf{Z}^{(D)}$ imposes m linear independent restrictions, it follows that
$$\hat{\mathbf{Z}}^{(D)} = E(\mathbf{Z}^*|X) + \mathbf{D} + \Psi\Psi'C'(C\Psi\Psi'C')^{-1}\left[\mathbf{Y} - CE(\mathbf{Z}^*|X) - C\mathbf{D}\right] \qquad (4.6)$$
with MSE matrix given by (4.2).

A difficulty with (4.6) is that **W**, and hence **D**, is assumed known, but in practice it has to be determined from the available data. Moreover, all the information about **D** is that corresponding to the forecast horizon, that is, **Y** and $E(\mathbf{Z}^*|X)$. Therefore **D** can be specified by solving

$$C\hat{\mathbf{D}} = \mathbf{Y} - CE(\mathbf{Z}^*|X), \qquad (4.7)$$
which is assumed to be a system of consistent equations (i.e. any linear relationship existing among the rows of C also exists among the rows of $\mathbf{Y} - CE(\mathbf{Z}^*|X)$). The solution to (4.7) is
$$\hat{\mathbf{D}} = C^{-}\left[\mathbf{Y} - CE(\mathbf{Z}^*|X)\right] + (I - C^{-}C)\mathbf{v} \qquad (4.8)$$
with C^{-} a generalized inverse of C and **v** an arbitrary H-dimensional constant vector. Hence (4.6) takes the form
$$\hat{\mathbf{Z}}^{(D)} = E(\mathbf{Z}^*|X) + \hat{\mathbf{D}} \qquad (4.9)$$

On the other hand, the compatibility statistic (2.9) is always cero in this case because of (4.7), so that **W** and **Y** are necessarily compatible by construction of $\hat{\mathbf{D}}$.

4.3 Change Foreseen in the Stochastic Structure of the Model

The structural change that we now consider is again due to an intervention, whose effects are of a stochastic nature and occur during the forecast horizon of the series. This case was also considered and exemplified by Guerrero (1991). A formulation that takes into account the intervention effects is

$$Z_t^{(S)} = Z_t + v_t S_{t,T}, \tag{4.10}$$

where $\{v_t\}$ follows a stationary autoregressive moving average (ARMA) process $\phi(B)v_t = \theta(B)\varepsilon_t$ with AR and MA polynomials $\phi(B)$ and $\theta(B)$ respectively and $\{\varepsilon_t\}$ a zero-mean white noise process uncorrelated with $\{Z_t\}$. $S_{t,T}$ is the step function defined in the previous case and T=N+1 is also assumed without loss of generality.

In order for (4.10) to be applicable in practice, we should realize that the (stochastic) intervention effects have to be assessed from the information on the future values of the series $\mathbf{Z}^{(S)} = \left(Z_{N+1}^{(S)},\ldots,Z_{N+H}^{(S)}\right)'$, which is contained in the ARIMA forecasts and in some (one or two) linear restrictions. Thus we should restrict the ARMA model for $\{v_t\}$ to have a very simple structure. In what follows we assume that it is a zero-mean white noise process with variance σ_v^2. Now let $\mathbf{Z^*} = (Z_{N+1},\ldots,Z_{N+H})'$ be the vector of future values without intervention effects and $\mathbf{v} = (v_{N+1},\ldots,v_{N+H})'$ the stochastic effects, so that $\mathbf{Z}^{(S)} = \mathbf{Z}^* + \mathbf{v}$. Then let us replace $\mathbf{Z}$ by $\mathbf{Z}^*$ in (2.11), and let $\mathbf{W} = E\left(\mathbf{Z}^{(S)}|X\right) = E\left(\mathbf{Z}^*|X\right)$, $\mathbf{e} = \Psi\mathbf{a} + \mathbf{v}$ and $\Sigma_e = \sigma_a^2\Psi\Psi' + \sigma_v^2 I$ in (2.1). Hence, if $\mathbf{W}, \Sigma_e$ and $\mathbf{Y} = C\mathbf{Z}^{(S)}$ are known, the MMSELE of $\mathbf{Z}^{(S)}$ is given by

$$\hat{\mathbf{Z}}^{(S)} = E\left(\mathbf{Z}^*|X\right) + \left(\sigma_a^2\Psi\Psi' + \sigma_v^2 I\right)C'\left(\sigma_a^2 C\Psi\Psi'C' + \sigma_v^2 CC'\right)\left[\mathbf{Y} - CE\left(\mathbf{Z}^*|X\right)\right], \tag{4.11}$$

with MSE matrix

$$Cov\left(\hat{\mathbf{Z}}^{(S)} - \mathbf{Z}^{(S)}\right) = \left[I - \left(\sigma_a^2\Psi\Psi' + \sigma_v^2 I\right)C'\left(\sigma_a^2 C\Psi\Psi'C' + \sigma_v^2 CC'\right)^{-1}C\right]\left(\sigma_a^2\Psi\Psi' + \sigma_v^2 I\right). \tag{4.12}$$

The problem now is that σ_v^2 is unknown and we must be able to determine its value from the known quantities $\mathbf{W}, \Sigma_e$ and $\mathbf{Y}$. To that end, let us notice that those values as well as σ_v^2, are involved in the coresponding compatibility statistic

$$K = \left[\mathbf{Y} - CE\left(\mathbf{Z}^*|X\right)\right]'\left(\sigma_a^2 C\Psi\Psi'C' + \sigma_v^2 CC'\right)^{-1}\left[\mathbf{Y} - CE\left(\mathbf{Z}^*|X\right)\right] \tag{4.13}$$

which, on the assumption of Gaussianity for both $\{Z_t\}$ and $\{v_t\}$, has a χ_m^2 distribution with m the number of (linearly independent) restrictions imposed by $\mathbf{Y}$. Next, since a reasonable selection of σ_v^2 is one that makes $\mathbf{Y}$ and $E\left(\mathbf{Z}^*|X\right)$ compatible with each other, we should select σ_v^2 in such a way that $K < \chi_m^2(\alpha)$ at a prespecified significance level α. Such a value, say $\hat{\sigma}_v^2$, is then replaced in (4.11) and (4.12) to obtain a working solution.

4.4 Change in Parameter Values due to an Intervention in the Forecast Horizon

Here we address the problem of combining ARIMA forecasts with some linear restrictions, when an intervention is anticipated and we fear its effects will change the original values of the AR and MA parameters. We start again with expression (2.11), which is assumed to hold for the future values without intervention effects $\mathbf{Z}^* = (Z_{N+1},\ldots,Z_{N+H})'$, and will be written as

$$\mathbf{Z}^* - E^0(\mathbf{Z}^*|X) = \Psi^0\mathbf{a} \qquad (4.14)$$

to stress the fact that the original parameters, say $\mathbf{b}^0 = (b_1^0,\ldots,b_k^0)'$ with $k \leq m$, are being used (let us recall that Ψ^0 is derived from $\mathbf{b}^0$).

A formulation that takes into account possible changes in the parameter values is

$$Z_t^{(P)} = Z_t + \sum_{j=1}^{t-T}\left(\hat{\Psi}_j - \Psi_j^0\right)a_{t-j}S_{t,T}, \qquad (4.15)$$

where the $\hat{\Psi}_j$s are the new pure MA weights of the ARIMA model, which must be obtained from a set of estimated parameter values $\hat{\mathbf{b}}$ that takes into account the future values of the series $\mathbf{Z}^{(P)} = \left(Z_{N+1}^{(P)},\ldots,Z_{N+H}^{(P)}\right)'$. Again let $S_{t,T}$ be the step function defined in Section 4.2 and let us assume, without loss of generality, that T=N + 1. Then let us notice that (4.15) leads us to $\mathbf{Z}^{(P)} = \mathbf{Z}^* + \left(\hat{\Psi} - \Psi^0\right)\mathbf{a}$, with $\hat{\Psi}$ defined in similar fashion as Ψ^0, so that (4.14) implies

$$\mathbf{Z}^{(P)} - E^0(\mathbf{Z}^*|X) = \hat{\Psi}\mathbf{a}. \qquad (4.16)$$

Then the Basic Rule can be applied with $\mathbf{W} = E^0(\mathbf{Z}^*|X)$, $\Sigma_e = \sigma^2\hat{\Psi}\hat{\Psi}'$ and $\mathbf{Y} = C\mathbf{Z}^{(P)}$, to obtain

$$\hat{\mathbf{Z}}^{(P)} = E^0(\mathbf{Z}^*|X) + \hat{\Psi}\hat{\Psi}'C'\left(C\hat{\Psi}\hat{\Psi}'C'\right)^{-1}\left[\mathbf{Y} - CE^0(\mathbf{Z}^*|X)\right]. \qquad (4.17)$$

The practical problem now is to get $\hat{\Psi}$. A solution to this problem was given by Guerrero (1990a) who derived a formula for the new parameter estimators involving the information provided by way of the restrictions. He also showed that a test statistic for the hypothesis $H_0: \mathbf{b} - \mathbf{b}^{(0)} = \mathbf{0}$ (i.e. null overall change in parameter values), when $E^0(\mathbf{Z}^*|X)$, $\sigma^2\Psi^0\Psi^{0\prime}$ and $\mathbf{Y}$ are known, is given by

$$K = \left[\mathbf{Y} - CE^0(\mathbf{Z}^*|X)\right]'\left(C\Psi^0\Psi^{0\prime}C'\right)^{-1}\left[\mathbf{Y} - CE^0(\mathbf{Z}^*|X)\right]/\sigma^2 \sim \chi_k^2. \qquad (4.18)$$

This test statistic has exactly the same form as that in (4.3). However, when k<m (4.18) will be more sensitive than (4.3), because we are being now more explicit about the reason why $\mathbf{Y}$ may not be compatible with $E^0(\mathbf{Z}^*|X)$. Moreover, Guerrero (1990a) provided some examples in which the parameter changes are interpreted as changes in the unobserved structural components of the time series (trend and/or seasonality).

5. TEMPORAL DISAGGREGATION OF UNIVARIATE TIME SERIES

The problem of temporal disaggregating a time series is that of estimating an unobserved random vector $\mathbf{Z} = (Z_1, ..., Z_{mn})'$ on the basis of knowing some linear aggregates $Y_i = \sum_{j=1}^{n} c_j Z_{n(i-1)+j}$, with i=1,...,m. Here n denotes the intraperiod frecuency of observation (i.e. if $\{Y_i\}$ is observed annualy and $\{Z_t\}$ is a monthly series, n=12) m is the number of whole-period observations and $\mathbf{c} = (c_1, ..., c_n)' \neq \mathbf{0}$. Some usual forms of **c** are: $\mathbf{c} = (0,0,...,0,1)'$ for interpolating a stock series, $\mathbf{c} = (1,1,...,1)'$ for distributing a flow series and $\mathbf{c} = (1/n, 1/n, ..., 1/n)'$ for distributing an index series.

5.1 Temporal Disaggregation Without Auxiliary Data

Let us assume that $\{Z_t\}$ admits an ARIMA representation, so that (2.11) holds true, with $\mathbf{a} = \mathbf{a}_Z = (a_{Z.1}, ..., a_{Z.mn})'$ such that $E(\mathbf{a}_Z | X) = \mathbf{0}$ and $E(\mathbf{a}_Z \mathbf{a}'_Z | X) = \sigma_Z^2 I$, where $X = (..., Z_{-1}, Z_0)'$ denotes the infinite past of the series. In fact, for an ARIMA model to be reasonable in this situation we assume that the process started at a finite time point with fixed initial values. The vector $\mathbf{Y} = (Y_1, ..., Y_m)'$ can be written in the form **Y** = **CZ** by defining $C = I \otimes \mathbf{c}'$, where $\otimes$ denotes Kronecker product. Then the problem is posed as in the Basic Rule, so that the MMSELE of **Z**, when E(**Z**|X), **Y** and $\sigma_Z^2 \Psi\Psi'$ are known, is of the form (4.1). However, in practice such a formula is useless because X is unknown, as well as E(**Z**|X) and $\sigma^2 \Psi\Psi'$. Some approaches that have been followed to overcome these difficulties are the following.

(1) Assume a priori that E(**Z**|X) and $\Psi\Psi'$ have some simple structures, say E(**Z**|X)=**0** and Ψ is derived as if an integrated process of order one or two were an adequate representation for $\{Z_t\}$, then calculate

$$\hat{\mathbf{Z}} = \Psi\Psi' C' (C\Psi\Psi' C')^{-1} \mathbf{Y}. \qquad (5.1)$$

Such a solution was essentially proposed by Lisman & Sandee (1964) and by Boot, Feibes & Lisman (1967) when n=4, and by Cohen, Müller & Padberg (1971) for arbitrary n.

(2) Assume that the model for $\{Z_t\}$ can be somehow known to the analyst, perhaps by assuming that some disaggregated observations exist. This allows the specification of the matrix Ψ, so that we can calculate $\hat{Z}$ if E(**Z**|X) is also assumed to be known. That was the approach of Harvey & Pierse (1984).

(3) Derive the disaggregate ARMA model of the stationary series $\{(1-B)^d Z_t\}$ from that of $\{(1-B)^d Y_i\}$ using the theoretical relationship that links those two series. Thus, obtain the autocovariance matrix of $\{(1-B)^d Z_t\}$ and use it in an expresion similar to (5.1). Then

obtain the estimator of $\mathbf{Z}$ from the previous one, by applying a linear operator that essentialy serves to transform $\left((1-B)^d\hat{\mathbf{Z}}_{d+1},\ldots,(1-B)^d\hat{\mathbf{Z}}_{mn}\right)'$ into $\hat{\mathbf{Z}}=\left(\hat{\mathbf{Z}}_1,\ldots,\hat{\mathbf{Z}}_{mn}\right)'$. This solution was developed by Wei & Stram (1990); see also Stram & Wei (1986).

5.2 Temporal Disaggregation on the Basis of a Prelimnary Series

The problem to be considered now is the same as in Section 5.1, except that X will no longer be considered as the infinite past of the series. Rather $\mathbf{W}=E(\mathbf{Z}|X)$ will be considered a preliminary series, derived perhaps from a set of information on auxiliary variables X which help to explain the behavior of $\{Z_t\}$ or else $\{W_t\}$ is observed directly as a preliminary series. In both cases, a preliminary vector $\mathbf{W}=E(\mathbf{Z}|\mathbf{W})$ is known and it will be assumed to be the MMSELE of $\mathbf{Z}$. Now let us also assume that (2.11) holds true with $\mathbf{W}$ in place of X and $\mathbf{a}=(a_1,\ldots,a_{mn})'$ such that $E(\mathbf{a}|\mathbf{W})=\mathbf{0}$ and $E(\mathbf{aa}'|\mathbf{W})=\sigma^2P$, where P is a positive definite matrix. This assumption may be justified by assuming that $\{Z_t\}$ and $\{W_t\}$ follow the same ARIMA model, with the same AR and MA parameters, but different generating white noise processes. Such an assumption makes sense when $\mathbf{W}$ is indeed a preliminary estimate of $\mathbf{Z}$ and it allows to derive the matrix Ψ from observed data.

Thus, if $\mathbf{W}$, $\Sigma_e=\sigma^2\Psi P\Psi'$ and $\mathbf{Y}$ are known, we obtain

$$\hat{\mathbf{Z}}=\mathbf{W}+\Psi P\Psi'C'(C\Psi P\Psi'C')^{-1}(\mathbf{Y}-C\mathbf{W}) \qquad (5.2)$$

and

$$\mathrm{Cov}(\hat{\mathbf{Z}}-\mathbf{Z})=\sigma^2\left[I-\Psi P\Psi'C'(C\Psi P\Psi'C')^{-1}C\right]\Psi P\Psi'. \qquad (5.3)$$

These formulas were obtained by Guerrero (1990b) who showed that the corrresponding expressions proposed by Denton (1971) are particular cases of (5.2) - (5.3) when $\mathbf{W}$ is directly observed. Similarly, when comparing Chow and Lin's (1971) solution with (5.2) - (5.3) it is clear that they did not condition on $\mathbf{W}$ given and focused their attention on simultaneously estimating $\mathbf{Z}$ and the regression parameters linking the auxiliary information $\mathbf{X}$ with $\mathbf{Z}$. By doing that, they did not pay much attention to the potential autocorrelation structure in the regression errors. In addition, the covariance matrix of their solution $\hat{\mathbf{Z}}$, is (5.3) plus another term that can be related to the fact that only $\mathbf{X}$ (not $\mathbf{W}$) was assumed to be given.

To apply (5.2) - (5.3) in practice, we require knowing not just Ψ (which is obtainable from the ARIMA model for $\{W_t\}$), but P as well. Guerrero (1990b) suggested to use a two-step procedure which is akin to using estimated generalized least squares. Then a compatibility test between $\mathbf{Y}$ and $\mathbf{W}$ is given by

$$K=(\mathbf{Y}-C\mathbf{W})'(C\Psi P\Psi'C')^{-1}(\mathbf{Y}-C\mathbf{W})/\sigma^2\sim\chi^2_n. \qquad (5.4)$$

We notice however, that Guerrero (1990b) suggested a statistic different from this one, which does not involve P. Such a statistic can be applied even before the temporal disaggregation procedure has been carried out, to see if **W** can indeed be considered a preliminary estimate of **Z**. On the other hand, a recursive procedure that avoids manipulation of $\Psi P \Psi'$, which in practice may be a very large matrix, was proposed by Guerrero & Martínez (1993).

6. OUTLIERS AND STRUCTURAL CHANGES IN TIME SERIES

6.1 Detecting and Measuring the Effect of Influential Outliers

We consider here the single outlier case with known time of occurrence T. So, we start with the two basic mechanisms that may generate an Additive Outlier (AO) or an Innovational Outlier (IO). These were considered for instance, by Tsay (1986) and Peña (1990). That is, let

$$Z_t^{(AO)} = Z_t + \omega_A P_{t,T} \tag{6.1}$$

or

$$Z_t^{(IO)} = Z_t + \sum_{j=0}^{t-T} \Psi_j \omega_I P_{t-j,T} \tag{6.2}$$

where $P_{t,T}$ denotes the pulse function that takes on the value 1 when t=T and it is zero otherwise.

The AO situation follows from the results of Section 3.2 by letting $\mathbf{Z}_M = Z_T$, $X = (Z_1,\dots,Z_{T-1})'$ and $\mathbf{Z}_G = (Z_{T+1},\dots,Z_N)'$. Therefore, if we let $X^* = (X', Z'_G)'$, $\Sigma_{TG} = (\Psi_1,\dots,\Psi_{N-T})$ and $\Sigma_G = C\Psi\Psi'C'$, with $C = (\mathbf{0}, I_{N-T})$, we have

$$\hat{Z}_T = E(Z_T|X^*) = E(Z_T|X) + \Sigma_{TG}(C\Psi\Psi'C')\left[\mathbf{Z}_G - E(\mathbf{Z}_G|X)\right] \tag{6.3}$$

and

$$\mathrm{Var}\left(\hat{Z}_T - Z_T\right) = \sigma^2\left[1 - \Sigma_{TG}(\mathbf{C\Psi\Psi'C'})^{-1}\Sigma'_{TG}\right] \tag{6.4}$$

This estimator can also be written using equation (3.10) as

$$\hat{Z}_T = \mathbf{1}'_T \Pi'\Pi \mathbf{Z}_c \tag{6.5}$$

where $\mathbf{1}_T$ is a vector with one at position T and zero elsewhere, and $\mathbf{Z}_c$ has a zero at position T and the observed values otherwise. Now, calling ρ_i^I to the inverse autocorrelation coeficients, (6.5) can be rewritten as

$$\hat{Z}_T = -\Sigma\rho_i^I\left(Z_{T-i} + Z_{T+i}\right) \tag{6.6}$$

which is the well-known equation obtained by Brubacher & Wilson (1976) for the optimal interpolator of a time series.

If we now consider $Y = Z_T^{(AO)}$ known, the Basic Rule leads to the obvious result

$$\hat{Z}_T = Z_T^{(AO)} \tag{6.7}$$

with $\mathrm{Var}\left(\hat{Z}_T - Z_T\right) = 0$. Expression (6.7), as related to (6.1), implies that

$$\hat{\omega}_A = Z_T^{(AO)} - E(Z_T|X^*) \tag{6.8}$$

with $\mathrm{Var}(\hat{\omega}_A)$ given by (6.4). Moreover, on the Gaussianity assumption for $\{a_t\}$ we obtain the following statistic for testing compatibility between $Z_T^{(AO)}$ and $E(Z_T|X^*)$

$$\begin{aligned} K &= \left[Z_T^{(AO)} - E(Z_T|X^*)\right]^2 / \mathrm{Var}(\hat{\omega}_A) \\ &= \hat{\omega}_A^2 / \left[\sigma^2\left(1 - \Sigma_{TG}\Sigma_G^{-1}\Sigma'_{TG}\right)\right] \sim \chi_1^2. \end{aligned} \tag{6.9}$$

Noting that $\left(1 - \Sigma_{TG}\Sigma_G^{-1}\Sigma'_{TG}\right)^{-1}$ is the (1,1) element of the matrix Σ_e^{-1} of (3.4) and since $\Sigma_e^{-1} = \sigma^{-2}\Pi'\Pi$, it is clear that

$$1 - \Sigma_{TG}\Sigma_G^{-1}\Sigma'_{TG} = \sum_{i=0}^{N-T}\pi_i^2. \tag{6.10}$$

Therefore the compatibility statistic (6.9) can be written as $\hat{\omega}_A^2 / \left(\sigma^2\Sigma\pi_i^2\right)$ which is the likelihood ratio test proposed by Chang, Tiao & Chen (1988) for testing for additive outliers. This statistic is also related to the influence measures derived by Peña (1990, 1991).

In the IO situation we realize that (6.2) induces a deterministic change in the structure of the time series. So, we follow the ideas proposed in Section 4.2. Thus, let (2.11) hold true with $X = (Z_1,\dots,Z_{T-1})'$, $\mathbf{Z} = (Z_T,\dots,Z_n)'$ and $\mathbf{Z}^{(IO)} = \mathbf{Z} + \Psi\boldsymbol{\omega}$, $\boldsymbol{\omega} = (\omega_I, \mathbf{0}')'$. Then if we write $\mathbf{W} = E(\mathbf{Z}|X) + \Psi\boldsymbol{\omega}$, $\mathbf{e} = \Psi\mathbf{a}$, $\Sigma_e = \sigma^2\Psi\Psi'$ and $\mathbf{Y} = \mathbf{Z}^{(IO)}$, the Basic Rule leads us to $\hat{\mathbf{Z}}^{(IO)} = \mathbf{Y}$ and $\mathrm{Cov}\left(\hat{\mathbf{Z}}^{(IO)} - \mathbf{Z}^{(IO)}\right) = \mathbf{0}$, when $\boldsymbol{\omega}$ is known. In the present case it is unknown, but can be estimated from the available data as

$$\Psi\hat{\boldsymbol{\omega}} = \mathbf{Y} - E(\mathbf{Z}|X), \tag{6.11}$$

from which it follows that

$$\hat{\omega}_I = Z_T^{(IO)} - E(Z_T|X). \tag{6.12}$$

The statistic for testing compatibility between $\mathbf{Y}$ and $E(\mathbf{Z}|X)$, when $\{a_t\}$ is Gaussian, becomes

$$\begin{aligned} K &= \left[\mathbf{Y} - E(\mathbf{Z}|X)\right]'(\Psi\Psi')^{-1}\left[\mathbf{Y} - E(\mathbf{Z}|X)\right]/\sigma^2 \\ &= \hat{\boldsymbol{\omega}}'\Psi'(\Psi\Psi')^{-1}\Psi\hat{\boldsymbol{\omega}}/\sigma^2 \\ &= \hat{\omega}_I^2/\sigma^2 \sim \chi_1^2. \end{aligned} \tag{6.13}$$

This is equivalent to the likelihood ratio test statistic proposed by Chang, Tiao & Chen (1988), and also employed by Peña (1990,1991) to express a measure of outlier influence.

6.2 Reallocation Outliers

A situation considered and illustrated by Wu, Hosking & Ravishanker (1993), basically consists in restricting a block of consecutive observations affected by outliers to produce the same sum as if no outliers were present. So, let us consider the multiple additive outlier formulation

$$Z_t^{(RO)} = Z_t + \sum_{i=0}^{m} \lambda_i P_{t,c+i} \tag{6.14}$$

where $\{Z_t^{(RO)}\}$ is the series with Reallocation Outliers (RO), $P_{t,T}$ is the pulse function defined as in Section 6.1 and λ_i is the effect associated with the outlier occurring at time c+i. The values of c and m (timing and duration of reallocation) are assumed known, although Wu, Hosking & Ravishanker (1993) also addressed the case in which c and/or m are unknown.

Now let us define $\mathbf{Z}^{(RO)} = \left(Z_c^{(RO)}, Z_{c+1}^{(RO)}, \ldots, Z_{c+m}^{(RO)}\right)'$ and $\mathbf{Z}_M = (Z_c, Z_{c+1}, \ldots, Z_{c+m})'$ the vectors of observations with and without outlier effects. Similarly let $X = (Z_1, \ldots, Z_{c-1})'$ and $\mathbf{Z}_G = (Z_{c+m+1}, \ldots, Z_N)'$ be the vectors of past and future observations with respect to $Z^{(RO)}$, and $X^* = (X', \mathbf{Z}'_G)'$. Thus, from (3.8) and (3.9) we know that $\mathbf{Z}_M$ can be expressed as $\mathbf{Z}_M = \mathbf{W}_M + \mathbf{e}$, with

$$\mathbf{W}_M = E(\mathbf{Z}_M | X^*) = E(\mathbf{Z}_M | X) + \Sigma_{MG}\Sigma_G^{-1}\left[\mathbf{Z}_G - E(\mathbf{Z}_G | X)\right] \tag{6.15}$$

and $\mathbf{e}$ a random vector such that $E(\mathbf{e}|X^*) = \mathbf{0}$ and

$$\Sigma_e = \sigma^2\left(\Sigma_M - \Sigma_{MG}\Sigma_G^{-1}\Sigma'_{MG}\right). \tag{6.16}$$

Then, if we consider $Y = \mathbf{1}'\mathbf{Z}^{(RO)}$ given, with $\mathbf{1}$ an m-vector of ones, the Basic Rule yields

$$\hat{\mathbf{Z}}_M = \mathbf{W}_M + \Sigma_e \mathbf{1}(\mathbf{1}'\Sigma_e\mathbf{1})^{-1}\mathbf{1}'(\mathbf{Z}^{(RO)} - \mathbf{W}_M), \tag{6.17}$$

with

$$\mathrm{Cov}\left(\hat{\mathbf{Z}}_M - \mathbf{Z}_M\right) = \Sigma_e - \Sigma_e\mathbf{1}(\mathbf{1}'\Sigma_e\mathbf{1})^{-1}\mathbf{1}'\Sigma_e. \tag{6.18}$$

Next, from (6.14) we have $\mathbf{Z}^{(RO)} = \mathbf{Z} + \lambda$ with $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)'$, so that

$$\hat{\lambda} = \mathbf{Z}^{(RO)} - \hat{\mathbf{Z}}_M = \left(\mathbf{Z}^{(RO)} - \mathbf{W}_M\right) - \Sigma_e\mathbf{1}(\mathbf{1}'\Sigma_e\mathbf{1})^{-1}\mathbf{1}'\left(\mathbf{Z}^{(RO)} - \mathbf{W}_M\right) \tag{6.19}$$

and $\mathrm{Cov}(\hat{\lambda}) = \mathrm{Cov}(\hat{\mathbf{Z}}_M - \mathbf{Z}_M)$. Furthermore, a statistic for testing compatibility between $Y = \mathbf{1}'\mathbf{Z}^{(RO)}$ and $\mathbf{W}_M$, on the normality assumption for $\mathbf{e}$, is given by (2.9) as

$$K = \left(\mathbf{1}'\mathbf{Z}^{(RO)} - \mathbf{1}'\mathbf{W}_M\right)'(\mathbf{1}'\Sigma_e\mathbf{1})^{-1}\left(\mathbf{1}'\mathbf{Z}^{(RO)} - \mathbf{1}'\mathbf{W}_M\right)$$

$$= \left[\mathbf{1}'\left(\mathbf{Z}^{(RO)} - \mathbf{W}_M\right)\right]^2 / (\mathbf{1}'\Sigma_e\mathbf{1}) \sim \chi_1^2. \tag{6.20}$$

Expressions (6.18), (6.19) and (6.20) are easily seen to be the same as those deduced by Wu, Hosking & Ravishanker (1993), by recognizing that in their notation σ^2A^{-1} and $A^{-1}\mathbf{b}$, are Σ_e and $\mathbf{Z}^{(RO)} - \mathbf{W}_M$ in ours. In particular (6.20) is useful for testing whether the additive outliers are reallocation or not. Moreover, they also proposed a statistic for testing the null hypothesis of no outliers versus a reallocation, which in our context becomes a test for compatibility between $\mathbf{Z}^{(RO)}$ and $\hat{\mathbf{Z}}_M$, that is

$$K^* = \left(\mathbf{Z}^{(RO)} - \hat{\mathbf{Z}}_M\right)'\Sigma_e^{-1}\left(\mathbf{Z}^{(RO)} - \hat{\mathbf{Z}}_M\right) \sim \chi_m^2. \tag{6.21}$$

6.3 Detecting and Modelling Structural Changes in Time Series

The underlying ideas employed in Section 4 can also be exploited in the case in which a structural change is feared or known ex-post, to have occurred at time T, during the observation period of a time series. To be able to appreciate this, let us assume now that (2.11) holds true with historical data $X = (Z_1, \ldots, Z_{T-1})'$ and future values without intervention effects $\mathbf{Z}^* = (Z_T, Z_{T+1}, \ldots, Z_N)'$. Let also $\mathbf{Y} = \mathbf{Z}^{(1)}$ be the future values contaminated by intervention effects (actually observed), then (4.1) and (4.2) yield the obvious results $\hat{\mathbf{Z}}^{(1)} = \mathbf{Y}$ and $\mathrm{Cov}(\hat{\mathbf{Z}}^{(1)} - \mathbf{Z}^{(1)}) = \mathbf{0}$. Nevertheless, in this case

$$K = [\mathbf{Y} - E(\mathbf{Z}^*|X)]'(\Psi\Psi')^{-1}[\mathbf{Y} - E(\mathbf{Z}^*|X)]/\sigma^2 \sim \chi^2_{N-T+1} \qquad (6.22)$$

is still useful for validating empirically the inclusion of intervention effects into the model. In fact, Box & Tiao (1976) suggested using this statistic as an overall check and provided a guide about its use to decompose it "into components associated with various relevant alternatives".

When (6.22) indicates incompatibility between $\mathbf{Y}$ and $E(\mathbf{Z}^*|X)$, some alternatives are the following.

(1) Perform a Box-Tiao intervention analysis on the assumption that the structural change in the model is of a deterministic nature. Thus, in accordance with Section 4.2, the intervention effects must be represented by a dynamic function capable of reproducing the forecast errors $\mathbf{Y} - E(\mathbf{Z}^*|X)$, as indicated in expression (4.7). The practical problem is that the dynamic function intervention parameters should be simultaneously estimated with all the other parameters in the model (see Box & Tiao, 1975).

(2) Approach the intervention analysis as in Section 4.3, in such a way that the stochastic structure is assumed to be affected. Notice that in this case, even if the series of stochastic effects follows an ARMA, not just a white noise process, (4.11) produces $\hat{\mathbf{Z}}^{(1)} = \mathbf{Y}$ and (4.12) is the zero matrix. Hence the contaminating model has to be deduced from (4.13), perhaps on the basis of subject matter considerations about its ARMA structure (e.g. white noise). Then a procedure like the one described in Section 4.3 may be employed.

(3) Consider the possibility that the parameters have changed their values, by comparing (6.22) not with a χ^2_{N-T+1} but with a χ^2_k distribution, where k is the number of parameters of the ARIMA model (cf. Box & Tiao, 1976).

7. CONTEMPORANEOUS DISAGGREGATION OF MULTIPLE TIME SERIES

The contemporaneous disaggregation problem arises naturally when balancing national accounts. Solomon & Weale (1993) considered this problem and emphasized "the idea that discrepancies in the national accounts should be removed by least squares adjustment". So, they posed the problem as that of estimating a true vector of observations

$\mathbf{Z} = (\mathbf{Z}'_1, \ldots, \mathbf{Z}'_N)'$ where $\mathbf{Z}_t = (Z_{1t}, \ldots, Z_{kt})'$ is an observation on k variables at time t. There exist preliminary vectors of data $\mathbf{W}_t$ for each t, corresponding to $\mathbf{Z}_t$, so that $\mathbf{W} = (\mathbf{W}'_1, \ldots, \mathbf{W}'_N)'$ is available. The discrepancy between $\mathbf{Z}$ and $\mathbf{W}$ is assigned to an additive error $\mathbf{e}$ such that $E(\mathbf{e}) = \mathbf{0}$ and $Cov(\mathbf{e}) = \Sigma_e$. The true values satisfy a linear accounting constraint $C\mathbf{Z} = \mathbf{0}$, which the estimate $\hat{\mathbf{Z}}$ should also satisfy.

To aply the Basic Rule in this situation, let $\mathbf{W} = E(\mathbf{Z}|\mathbf{W})$ and Σ_e be known, so that

$$\hat{\mathbf{Z}} = \mathbf{W} - \Sigma_e C'(C\Sigma_e C')^{-1} C\mathbf{W} \quad (7.1)$$

and

$$Cov(\hat{\mathbf{Z}} - \mathbf{Z}) = \Sigma_e - \Sigma_e C'(C\Sigma_e C')^{-1} C\Sigma_e. \quad (7.2)$$

Furthermore, a statistic for testing compatibility between $\mathbf{W}$ and the accounting constraint, is given by

$$K = \mathbf{W}'C'(C\Sigma_e C')^{-1} C\mathbf{W} \sim \chi^2_{kN}. \quad (7.3)$$

Thus, the theoretical problem, as noticed by Solomon & Weale (1993), is solved when Σ_e is known, which in practice is not. So, they proposed a specific form of Σ_e, by taking into account an assumed autocorrelation structure derived from the way in which the data were constructed. A second difficulty is the practical problem associated with the size of the matrix Σ_e, which is typically very large.

Guerrero and Nieto (1994) proposed a different solution to these problems. Firstly, Σ_e is derived from the observed preliminary data by assuming that $\{\mathbf{W}_t\}$ and $\{\mathbf{Z}_t\}$ share the same autocorrelation structure, but have different generating white noise processes. Secondly, the manipulation of Σ_e is solved through a recursive approach.

8. CONCLUSIONS

We have shown that a Basic Rule for combining information from two different sources is a very useful tool for solving time series problems. Such a rule produces in fact a weighted average of the two different predictors coming from each source of information. Its optimality is easily revelead by the corresponding MSE matrix which is not only minimum for the class of linear and unbiased estimators considered, but because it shows that using the extra-model information reduces the original variability in the model estimator.

Realizing that many statistical procedures are derived by combining information is important from a unifying point of view. Besides we advocate the use of compatibility tests in order to appreciate whether the combination makes sense or not. Some of these tests have already appeared in the time series literature, associated mainly with likelihood-based inferences.

Many other problems faced by a time series analyst can also be posed in the context of combining information. For instance when combining forecasts from different models or when benchmarking time series. In those cases, another more general combining rule may be employed, in which the linear restrictions to be incorporated into the combination impose unbinding constraints on the estimators. We shall consider those topics in another paper.

REFERENCES

Alvarez, L.J., Delrieu, J.C. and Jareño, J. (1993) "Tratamiento de Predicciones Conflictivas: Empleo eficiente de información extramuestral." *Estadística Española* **35**, 439-461.

Boot, J.C.G., Feibes, W. and Lisman, J.H.C. (1967). "Further Methods of Derivation of Quarterly Figures from Annual Data." *Appl. Statist.* **16**, 65-75

Box, G.E.P. and Tiao, G.C. (1975). "Intervention Analysis with Applications to Environmental and Economic Problems." *J. Am. Statist. Assoc.* **70**, 70-79

Box, G.E.P. and Tiao, G.C. (1976). "Comparison of Forecast and Actuality." *Appl. Statist.* **25**, 195-200.

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis, Forecasting and Control.* Rev. ed. San Francisco: Holden-Day.

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

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

Chow, G.C. and Lin, A. (1971). "Best Linear Interpolation, Distribution and Extrapolation of Time Series by Related Series." *Rev. Econ. Statist.* **53**, 372-375.

Cohen, K.J., Müller, W. and Padberg, M.W. (1971). "Autoregressive Approaches to Disaggregation of Time Series Data." *Appl. Statist.* **20**, 119-129

Denton, F.T. (1971). "Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization." *J. Am. Statist. Assoc.* **66**, 99-102.

Draper, D., Gaver, D.P. Jr., Goel, P.K., Greenhouse, J.B., Hedges, L.V., Morris, C.N., Tucker, J.R. and Waternaux, C.M. (1992) *Combining Information. Statistical Issues and Opportunities for Research.* Washington: National Academy Press.

Gómez, V., Maravall, A. and Peña, D. (1993). "Computing Missing Values in Time Series." Working Paper. Universidad Carlos III de Madrid.

Guerrero, V.M. (1989). "Optimal Conditional ARIMA Forecasts." *J. Forecasting* **8**, 215-229.

Guerrero, V.M. (1990a). "Restricted ARIMA Forecasts which Account for Parameter Changes." *ESTADISTICA* **42**, 17-31.

Guerrero, V.M. (1990b). "Temporal Disaggregation of Time Series. An ARIMA-based Approach." *Int. Statist. Rev.* **58**, 29-46.

Guerrero, V.M. (1991). "ARIMA Forecasts with Restrictions Derived from a Structural Change." *Int. J. Forecasting* 7, 339-347.

Guerrero, V.M. (1993). "Restricted Forecasts of Missing Observations in Univariate Time Series Models." Working Paper. Instituto Tecnológico Autónomo de México.

Guerrero, V.M. and Martínez, J. (1993). "A Recursive ARIMA-based Procedure for Disaggregating a Time Series Variable Using Concurrent Data." Working Paper. Instituto Tecnológico Autónomo de México.

Guerrero, V.M. and Nieto, F.H. (1994). "Temporal and Contemporaneous Disaggregation of Multiple Time Series." Working Paper. Instituto Tecnológico Autónomo de México.

Harvey, A.C. and Pierse, R.G. (1984). "Estimating Missing Observations in Economic Time Series." *J. Am. Statist. Assoc.* **79**, 125-131.

Hedges, L.V. and Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. New York: Academic Press.

Lisman, J.H.C. and Sandee, J. (1964). "Derivation of Quarterly Figures from Annual Data." *Appl. Statist.* **13**, 87-90

Min, Ch. and Zellner, A. (1993). "Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates." *J. Econometrics* **56**, 89-118.

Palm, F.C. and Zellner, A. (1992). "To Combine or not to Combine? Issues of Combining Forecasts." *J. Forecasting* **11**, 687-701.

Peña D. (1987). "Measuring the Importance of Outliers in ARIMA Models." In *New Perspectives in Theoretical and Applied Statistics*, M. Puri et al. (eds) New York: Wiley, 109-118.

Peña, D. (1990). "Influential Observations in Time Series." *J. Bus. Econ. Statist.* **8**, 235-241.

Peña, D. (1991). "Measuring Influence in Dynamic Regression Models." *Technometrics* **33**, 93-101.

Peña, D. (1994). "Combining Information in Statistical Modelling." Working Paper. Universidad Carlos III de Madrid.

Peña, D. and Maravall, A. (1991). "Missing Observations, Additive Outliers and Inverse Autocorrelation Function." *Comm. Statist.* (Theory and Methods) **A-20**, 3175-3186.

Solomou, S. and Weale, M. (1993). "Balanced Estimates of National Accounts when Measurement Errors are Autocorrelated: the UK, 1920-38." *J. Roy. Statist. Soc.* **A-156**, 89-105.

Stram, D.O. and Wei, W.W.S. (1986). "A Methodological Note on the Disaggregation of Time Series Totals." *J. Time Ser. Anal.* **7**, 293-302

Tsay, R.S. (1986). "Time Series Model Specification in the Presence of Outliers." *J. Amer. Statist. Assoc.* **81**, 132-141.

Wei, W.W.S. and Stram, D.O. (1990). "Disaggregation of Time Series Models." *J. Roy Statist. Soc.* **B-52**, 453-467

Wu, L.S.Y., Hosking, J.R.M. and Ravishanker, N. (1993). "Reallocation Outliers in Time Series." *Appl. Statist.* **42**, 301-313.