# AN APPLICATION OF MULTIPLE TIME SERIES METHODS

# AN APPLICATION OF MULTIPLE TIME SERIES METHODS

# TO

# CANADIAN ECONOMIC DATA

By

JILL BOOKER, B.SC.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

December 1978

MASTER OF SCIENCE (1978)
(Statistics)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: An Application of Multiple Time Series Methods to Canadian Economic Data

AUTHOR: Jill Booker, B.Sc. (McMaster University)

SUPERVISOR: Dr. J.F. MacGregor

NUMBER OF PAGES: vi, 49

ABSTRACT

This work outlines several aspects of multiple time series analysis, which are then demonstrated on a large set of data.

After introducing the general autoregressive integrated moving average model, discussion is restricted to a canonical form: the pure autoregressive process of order p (AR(p)).

Methods for identifying and fitting the AR(p) process using quasi-partial correlation matrices and Akaike's AIC criterion are discussed.

The AR model can then be used to make forecasts by taking conditional expectations at the origin time. Probability limits on the forecasts are also defined.

A method for canonical analysis of AR processes is described which can indicate possible reductions in the dimensionality of the problem.

Using computer programs developed for this project, the above methods are applied to an 11-dimensional set of Canadian economic data and the results are discussed.

ACKNOWLEDGEMENTS

I would like to thank Dr. J.F. MacGregor for his supervision of this project and for his encouragement along the way.

I would also like to thank Dr. S. Chow of the Royal Bank, who provided the data used in the project and the technical information about the Trendicator.

CONTENTS

LIST OF TABLES

## 1.0 INTRODUCTION

Much work has been done by Box and Jenkins (1976) to develop procedures for modelling univariate time series by empirical parsimonious parametric models, and for using these models for forecasting and control.

These methods can also be extended to multiple time series by considering observations at a given time to be components of a vector. A model that explains the series jointly can be identified, fitted and used to make joint forecasts.

Because of the unmanageability of a large number of time series, it is desirable to find a method for reducing the dimensionality of the problem. Box and Tiao (1977) have proposed such a method for identifying the "most forecastable" linear combinations of the original series. These combinations - hopefully significantly fewer in number than the original series - can then be used for future work.

This project attempts to demonstrate the methods mentioned above. The data chosen were those used by the Royal Bank in their Trendicator, an index of the economy published quarterly.

Eleven leading indicators are used in the Trendicator:

1. Average Hours in Manufacturing
2. Investor's Index of 134 Common Stocks
3. TSE Price-Earnings Ratio
4. Housing Starts in Urban Areas

5. Number of Residential Building Permits Issued
6. Value of Total Building Permits
7. Liabilities of Business Failures (1971$)
8. Ratio of Price to Unit Labour Cost
9. Primary Steel Production
10. Money Supply - M2 (1971$)
11. Corporate Pre-tax Profits.

The Royal Bank seasonally adjusts these series and converts them to a quarterly basis, if necessary. Any series given in current dollars is deflated to a 1971 dollar basis. Finally, each series is smoothed and indexed to a 1971 base equal to 100. The indices are averaged to get the Trendicator. These indices are used as the original data set in this project.

## 2.0 MULTIPLE TIME SERIES: MODELS AND FORECASTING

### 2.1 Multiple Time Series Models

An m-variate time series $\underline{X}_t$, t=1,2,... can, in general, be represented by the autoregressive integrated moving average (ARIMA) model

$$\underline{\phi}(B)\underline{\nabla}^d\underline{X}_t = \underline{\theta}(B)\underline{a}_t, \qquad (2.1)$$

where $\underline{a}_t$ is an m x 1 vector of white noise variables, assumed to be independent and normally distributed with mean vector $\underline{0}$ and oovariance matrix $\underline{\Sigma}_a$

B is the backward shift operator (i.e., $B\underline{X}_t = \underline{X}_{t-1}$)

$\underline{\phi}(B) = \underline{I} - \underline{\phi}_1 B - \underline{\phi}_2 B^2 - \ldots$ is the m x m matrix of autoregressive parameters ($\underline{I}$ is the m x m identity matrix)

$\underline{\theta}(B) = \underline{I} - \underline{\theta}_1 B - \underline{\theta}_2 B^2 - \ldots$ is the m x m matrix of moving average parameters, and

$\underline{\nabla}^d = \text{diag}(1-B)^{d_i}$ is the differencing operator.

A process is called stationary if its probability density function, say $f(\underline{z}_t)$, is such that

$$f(\underline{z}_{t_1},\ldots,\underline{z}_{t_r}) \equiv f(\underline{z}_{t_1+s},\ldots,\underline{z}_{t_r+s}), \text{ for all } r,s.$$

It follows that, for a stationary process, the mean vector $\underline{\mu} = E(\underline{z}_t)$ is constant over time and so is the covariance matrix $\underline{\Sigma} = E(\underline{z}_t - \underline{\mu})(\underline{z}_t - \underline{\mu})'$. The covariance matrix between any $\underline{z}_t$ and $\underline{z}_{t+k}$ depends only on the lag k and not on t. This matrix is called the autocovariance matrix of lag k and is given by

$$\underline{\Gamma}_k = E(\underline{z}_t - \underline{\mu})(\underline{z}_{t+k} - \underline{\mu})'.$$

Sample estimates of the mean and covariance can be calculated as follows:

$$\hat{\underline{\mu}} = \bar{\underline{z}}$$

$$\hat{\underline{\Sigma}} = \underline{S} = \frac{1}{n}(\underline{z}_t - \bar{\underline{z}})(\underline{z}_t - \bar{\underline{z}})'$$

The elements $\gamma_{ij}$ of $\underline{\Gamma}_k$ can be estimated by

$$c_{ij}(k) = \frac{1}{n}\Sigma_{t=1}^{n-k}(z_{i,t} - \bar{z}_i)(z_{j,t+k} - \bar{z}_j).$$

If a series $\underline{X}_t$ is not stationary in level, the series $\underline{z}_t = \underline{\nabla}^d \underline{X}_t$, for some d, may be stationary. These differences can then be represented by a stationary model.

A model of the general form given in 2.1, with full $\underline{\phi}$ and $\underline{\theta}$ matrices is not uniquely identifiable from the available data (Granger and Newbold (1977)). It is desirable, then, to choose a canonical form of the model, that is, one which has a unique representation for a given covariance structure. One such form is the pure autoregressive model of undetermined order p (AR(p)):

$\underline{\phi}(B)\underline{z}_t = \underline{a}_t$, where $\underline{z}_t$ is a stationary series and

$$\underline{\phi}(B) = \underline{I} - \underline{\phi}_1 B - \dots - \underline{\phi}_p B^p.$$

Because the model is linear in the parameters it is easy to fit using the Yule-Walker equations (described in section 2.2). The only decision required is on the order p of the process. The order of these AR(p) models may actually turn out to be greater than that of canonical forms using moving average terms, but the relative ease in fitting the AR parameters more than compensates for the lack of parsimony.

## 2.2 Identification and Estimation of Multivariate AR Models

Consider the AR(p) model given by

$\underline{\phi}(B)\underline{z}_t = \underline{a}_t$, where $\underline{z}_t$ is stationary and

$$\underline{\phi}(B) = \underline{I} - \underline{\phi}_1 B - \dots - \underline{\phi}_p B^p.$$

The problem now is to identify the order p of the process and to estimate the parameters $\underline{\phi}_1, \underline{\phi}_2, \dots, \underline{\phi}_p$.

One criterion which can be used to identify the order p of the process is Akaike's AIC (an information criterion), which is defined as follows:

$$\begin{aligned} AIC &= -2\ln(\text{maximum likelihood}) + 2\ (\text{number of free parameters}) \\ &= N \ln(|\hat{\underline{\Sigma}}_{a,k}|) + 2\ (\text{number of free parameters}), \end{aligned}$$

where N is the number of observations available and

$\hat{\underline{\Sigma}}_{a,k}$ is the covariance matrix of the residuals after fitting an AR(k) process.

AIC goes through a minimum at k=p, where p is the order of the process which maximizes the expected log likelihood or equivalently, maximizes the expected entropy of the model. (Akaike (1976))

Another aid to identification, which also produces the estimates of the $\underline{\phi}$ matrices, is the quasi-partial correlation matrix of lag k. (These matrices are analogous to the partial correlations of univariate time series methods.)

Given a stationary series $\underline{z}_t$ which can be described by an AR(p) process

$$(\underline{I} - \underline{\phi}_1 B - \dots - \underline{\phi}_p B^p)\ \underline{z}_t = \underline{a}_t,$$

where p is not known, one could fit AR(k) models of increasing order, k=1,2,.... The matrix of parameters $\hat{\underline{\phi}}_{k,k}$ in the fitted AR(k) model is the estimated k-th quasi-partial correlation matrix. Theoretically, all $\underline{\phi}_{k,k}$ matrices beyond lag p should be zero, a fact which can be used to help identify the order p of the process.

For the AR(k) model

$$(\underline{I} - \underline{\phi}_{k,1}B - \dots - \underline{\phi}_{k,k}B^k)\underline{z}_t = \underline{a}_t ,$$

the following relationships between the autocovariance matrices can be shown to hold:

$$\underline{\Gamma}_j = \underline{\Gamma}_{j-1}\underline{\phi}'_{k,1} + \dots + \underline{\Gamma}_{j-k+1}\underline{\phi}'_{k,k-1} + \underline{\Gamma}_{j-k}\underline{\phi}'_{k,k} , \quad j=1,2,\dots,k.$$

Replacing the $\underline{\Gamma}_j$ by their sample estimates $\underline{C}_j$ and the $\underline{\phi}_{k,j}$ by their estimates $\hat{\underline{\phi}}_{k,j}$, one gets the well-known Yule-Walker equations:

$$\underline{C}_j = \underline{C}_{j-1}\hat{\underline{\phi}}'_{k,1} + \dots + \underline{C}_{j-k+1}\hat{\underline{\phi}}'_{k,k-1} + \underline{C}_{j-k}\hat{\underline{\phi}}'_{k,k} , \quad j=1,2,\dots,k.$$

These k matrix equations, together with the equation

$$\hat{\underline{\Sigma}}_a = \underline{C}_0 - \underline{C}_{-1}\hat{\underline{\phi}}'_{k,1} - \dots - \underline{C}_{-k}\hat{\underline{\phi}}'_{k,k}$$

can be solved to provide approximate maximum likelihood estimates of the k+1 unknown matrices $\underline{\phi}_{k,1},\dots,\underline{\phi}_{k,k},\underline{\Sigma}_a$. These equations can be solved recursively using the following formulae (Alavi (1973)):

$$\hat{\underline{\phi}}_{1,1} = \underline{C}'_1\hat{\underline{\Sigma}}^{-1}_{b;0} \qquad \hat{\underline{\Sigma}}_{b;0} = \underline{C}_0$$

$$\hat{\underline{\phi}}_{k,k} = \{\underline{C}'_k - \Sigma_{j=1}^{k-1}\hat{\underline{\phi}}_{k-1,j}\underline{C}'_{k-j}\}\,\hat{\underline{\Sigma}}^{-1}_{b;k-1} \qquad (k \geqslant 2)$$

$$\hat{\underline{\phi}}_{k,j} = \hat{\underline{\phi}}_{k-1,j} - \hat{\underline{\phi}}_{k,k}\hat{\underline{\psi}}_{k-1,k-j} \qquad (k \geqslant 2,\ j=1,2,\dots,k-1)$$

$$\hat{\underline{\Sigma}}_{b;k-1} = \underline{C}_0 - \Sigma_{j=1}^{k-1}\hat{\underline{\psi}}_{k-1}\underline{C}'_j$$

$$\hat{\underline{\psi}}_{1,1} = \underline{C}_1\hat{\underline{\Sigma}}^{-1}_{a;0} \qquad \hat{\underline{\Sigma}}_{a;0} = \underline{C}_0$$

$$\hat{\underline{\psi}}_{k,k} = \{\underline{C}_k - \Sigma_{j=1}^{k-1}\hat{\underline{\psi}}_{k-1,j}\underline{C}_{k-j}\}\,\hat{\underline{\Sigma}}^{-1}_{a;k-1} \qquad (k \geqslant 2)$$

$$\hat{\underline{\psi}}_{k,j} = \hat{\underline{\psi}}_{k-1,j} - \hat{\underline{\psi}}_{k,k}\hat{\underline{\phi}}_{k-1,k-j} \qquad (k \geqslant 2,\ j=1,2,\dots,k-1)$$

$$\hat{\underline{\Sigma}}_{a;k-1} = \underline{C}_0 - \Sigma_{j=1}^{k-1}\hat{\underline{\phi}}_{k-1,j}\underline{C}_j$$

These sets of formulae were the basis for the FORTRAN subroutine MPARTL used in this project. The program was tested with data from a known process given in Alavi (1973), and performed well, correctly identifying the order of the process and giving good estimates of the parameters.

## 2.3 Forecasting of AR(p) Processes

To calculate $\hat{\underline{z}}_t(\ell)$, the $\ell$-step ahead forecast at time t, the process is written as an expression for $\underline{z}_{t+\ell}$ and conditional expectations at time t are taken. These expectations are given by

$$E(\underline{z}_{t+j} | \underline{z}_t, \underline{z}_{t-1}, \ldots) = \underset{t}{E}(\underline{z}_{t+j}) = \begin{cases} \hat{\underline{z}}_t(j), & j \geq 1 \\ \underline{z}_{t+j}, & j \leq 0 \end{cases}$$

and $$\underset{t}{E}(\underline{a}_{t+j}) = \begin{cases} 0, & j \geq 1 \\ \underline{a}_{t+j}, & j \leq 0. \end{cases}$$

For the AR(p) process, then, we have

$$\underline{\phi}(B)\underline{z}_t = (\underline{I} - \underline{\phi}_1 B - \ldots - \underline{\phi}_p B^p)\underline{z}_t = \underline{a}_t.$$

This can be written as an expression for $\underline{z}_{t+\ell}$ as follows:

$$\underline{z}_{t+\ell} = \underline{\phi}_1 \underline{z}_{t+\ell-1} + \underline{\phi}_2 \underline{z}_{t+\ell-2} + \ldots + \underline{\phi}_p \underline{z}_{t+\ell-p} + \underline{a}_{t+\ell}.$$

Taking conditional expectations at time t and substituting the $\hat{\underline{\phi}}_i$ for the $\underline{\phi}_i$, i=1,2,...,p, we get the following expression from which the forecasts can be calculated:

$$\underset{t}{E}(\underline{z}_{t+\ell}) = \hat{\underline{z}}_t(\ell) = \hat{\underline{\phi}}_1 \underset{t}{E}(\underline{z}_{t+\ell-1}) + \ldots + \hat{\underline{\phi}}_p \underset{t}{E}(\underline{z}_{t+\ell-p}) + \underset{t}{E}(\underline{a}_{t+\ell}).$$

The eventual forecast function is

$$\hat{\underline{z}}_t(p+j) = \hat{\underline{\phi}}_1 \hat{\underline{z}}_t(p+j-1) + \hat{\underline{\phi}}_2 \hat{\underline{z}}_t(p+j-2) + \ldots + \hat{\underline{\phi}}_p \hat{\underline{z}}_t(j), \quad j=1,2,\ldots.$$

Forecast Error

The AR(p) model can also be expressed in a $\psi$-weight form as follows:

$$\underline{z}_{t+\ell} = \underline{\psi}(B)\underline{a}_{t+\ell} = (\underline{\psi}_0 + \underline{\psi}_1 B + \underline{\psi}_2 B^2 + \dots)\underline{a}_{t+\ell},$$

where $\underline{\psi}_0 = \underline{I}$, and the $\underline{\psi}$-weight matrices can be found by equating coefficients of $B^j$ in

$$\underline{\phi}(B)\ \underline{\psi}(B) = \underline{I}.$$

This leads to the following equations:

$$\underline{\psi}_0 = \underline{I}, \quad \underline{\psi}_j = \Sigma_{i=1}^{p} \underline{\phi}_i \underline{\psi}_{j-i}, \ j=1,2,3,\dots$$

$$\underline{\psi}_j = \underline{0}, \ j \leq 0.$$

It can be shown that the forecasts can also be expressed as

$$\hat{\underline{z}}_t(\ell) = \underline{\psi}_\ell \underline{a}_t + \underline{\psi}_{\ell+1}\underline{a}_{t-1} + \underline{\psi}_{\ell+2}\underline{a}_{t-2} + \dots .$$

The $\ell$-step ahead forecast error $\underline{e}_t(\ell)$ is given by

$$\underline{e}_t(\ell) = \underline{z}_{t+\ell} - \hat{\underline{z}}_t(\ell)$$

$$= \underline{\psi}_0\underline{a}_{t+\ell} + \underline{\psi}_1\underline{a}_{t+\ell-1} + \dots + \underline{\psi}_{\ell-1}\underline{a}_{t+1} \quad (\underline{\psi}_0 = \underline{I}).$$

The variance matrix of $\underline{e}_t(\ell)$ is

$$\underline{V}(\ell) = \underline{\Sigma} + \underline{\psi}_1\underline{\Sigma}\underline{\psi}_1' + \dots + \underline{\psi}_{\ell-1}\underline{\Sigma}\underline{\psi}_{\ell-1}' \quad (\text{since } E(\underline{e}_t(\ell)) = \underline{0}).$$

$1-\alpha$ probability limits for the individual components $z_{i,t+\ell}$, given the information available at time t, are given by

$$\hat{z}_{i,t}(\ell) \pm u(\alpha/2)\ (v_{ii}(\ell))^{\frac{1}{2}}, \ i=1,2,\dots,m,$$

where $u(\alpha/2)$ is the upper $\alpha/2$ percentile of the Normal distribution and $v_{ii}(\ell)$ is the ith diagonal element of $\underline{V}(\ell)$. Note that for any lead time $\ell$, the joint probability that all the $z_{i,t+\ell}$'s lie within their respective limits is less than $1-\alpha$.

The FORTRAN subroutine FORCST written for this project calculates the forecasts using conditional expectations and also calculates the $\psi$-weights by the recursive equations given above. The matrix $\underline{V}(\ell)$ is calculated, and diagonals are pulled out to be used in the calculation of the probability limits.

## 3.0 CANONICAL ANALYSIS OF AR TIME SERIES

In a k -dimensional process, two or more series may be a reflection of the same underlying condition. It would be useful to eliminate this redundancy, reducing the order of the process. Box and Tiao (1977) have proposed a method for doing this, by looking at linear combinations of the k original series with respect to predictability. The least predictable series can be dropped out and forecasts produced by using the most forecastable linear combinations.

Suppose $\underline{z}_t = \Sigma_{l=1}^{p} \underline{\phi}_l \underline{z}_{t-l} + \underline{a}_t$

$= \hat{\underline{z}}_{t-1}(1) + \underline{a}_t$, where $\underline{a}_t$ and $\hat{\underline{z}}_{t-1}(1)$ are independent.

Let the covariance matrices of $\underline{a}_t$, $\underline{z}_t$ and $\hat{\underline{z}}_{t-1}(1)$ be $\underline{\Sigma}_a$, $\underline{\Gamma}_0(z)$ and $\underline{\Gamma}_0(\hat{z})$, respectively. Then $\underline{\Gamma}_0(\underline{z}) = \underline{\Gamma}_0(\hat{\underline{z}}) + \underline{\Sigma}_a$.

Box and Tiao define a measure of predictability $\lambda$ as

$$\lambda = \sigma^2(\hat{z})/\sigma^2(z) = 1 - (\sigma^2(a)/\sigma^2(z)),$$

where a series is considered most predictable if it maximizes $\lambda$.

Consider now the linear combination $\dot{z}_t$ of the $z_{j,t}$'s $(j=1,\ldots,k)$

$$\dot{z}_t = \underline{m}'\underline{z}_t = m_1 z_{1t} + \ldots + m_k z_{kt}$$

$$\text{i.e.,} \dot{z}_t = \underline{m}'\hat{\underline{z}}_{t-1} + \underline{m}'\underline{a}_t = \hat{\dot{z}}_{t-1} + \dot{\underline{a}}_t.$$

So, $\sigma^2(\dot{z}) = \sigma^2(\hat{\dot{z}}) + \sigma^2(\dot{a})$, or $\underline{m}'\underline{\Gamma}_0(\underline{z})\underline{m} = \underline{m}'\underline{\Gamma}_0(\hat{\underline{z}})\underline{m} + \underline{m}'\underline{\Sigma}\,\underline{m}$.

The most predictable linear combination of the $z_{jt}$, then, is obtained by maximizing with respect to $\underline{m}$ the expression for $\lambda$,

$$\lambda = \sigma^2(\hat{\dot{z}})/\sigma^2(\dot{z}) = (\underline{m}'\underline{\Gamma}_0(\hat{\underline{z}})\ \underline{m}) \ / \ (\underline{m}'\underline{\Gamma}_0(\underline{z})\ \underline{m}).$$

Box and Tiao show that the maximum value of $\lambda$ is the largest root of the equation

$$\left|\underline{\Gamma}_0(\hat{\underline{z}}) - \lambda\underline{\Gamma}_0(\underline{z})\right| = 0 \qquad (3.1)$$

and $\underline{m}$ is obtained, except for an arbitrary scale factor, by solving the equations

$$\{\underline{\Gamma}_0(\hat{\underline{z}}) - \lambda\underline{\Gamma}_0(\underline{z})\}\ \underline{m} = \underline{0}.$$

The solutions $\lambda_1,\ldots,\lambda_k$ of (3.1) are the eigenvalues of

$$\underline{\Gamma}_0^{-1}(\underline{z})\ \underline{\Gamma}_0(\hat{\underline{z}}),$$

and the $\underline{m}$ vectors are the eigenvectors. Ordering the $\lambda$'s such that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k$, and forming the matrix $\underline{M}$ with the corresponding linearly independent eigenvectors $\underline{m}_1',\ldots,\underline{m}_k'$ as its rows, we have the transformed process

$$\dot{\underline{z}}_t = \hat{\dot{\underline{z}}}_{t-1}(1) + \dot{\underline{a}}_t,$$

with $\dot{\underline{z}}_t = \underline{M}\ \underline{z}_t$, $\dot{\underline{a}}_t = \underline{M}\underline{a}_t$, $\hat{\dot{\underline{z}}}_{t-1}(1) = \Sigma_{l=1}^{p}\dot{\underline{\phi}}_l\dot{\underline{z}}_{t-l}$, and $\dot{\underline{\phi}}_l = \underline{M}\ \underline{\phi}_l\ \underline{M}'$.

The k component series $\{\dot{z}_{1t},\ldots,\dot{z}_{kt}\}$ are

(i) ordered from least to most predictable,

(ii) are contemporaneously independent

(iii) have predictable components $\{\hat{\dot{z}}_{1(t-1)}(1),\ldots,\hat{\dot{z}}_{k(t-1)}(1)\}$ which are contemporaneously independent, and

(iv) have unpredictable components $\{\dot{a}_{1t},\ldots,\dot{a}_{kt}\}$ which are contemporaneously and temporally independent.

Given that the largest p of the eigenvalues $\lambda_1,\ldots,\lambda_k$ are non-zero, the hypothesis that the remaining k-p eigenvalues are zero can be tested using the statistic

$-\{(n-k) - \frac{1}{2}(2k+1)\}\ln \Pi_{j=1}^{k-p}(1 - \lambda_j)$, where n is the number of observations available for each series. This statistic is asymptotically distributed as $\chi^2(2(k-p))$.

For an AR(1) process,

$$\dot{\underline{z}}_t = \dot{\underline{\phi}}_1 \dot{\underline{z}}_{t-1} + \dot{\underline{a}}_t,$$

so $\dot{z}_{jt} = \Sigma_{i=1}^{k} \dot{\phi}_{ji} \dot{z}_{i,t-1} + \dot{a}_{jt}$. Since $\dot{z}_{1,t-1},\ldots,\dot{z}_{k,t-1}$ and $\dot{a}_{jt}$ are independent,

$$\sigma^2_{\dot{z}_j} = \Sigma_{i=1}^{k} \dot{\phi}^2_{ji} \sigma^2_{\dot{z}_i} + \sigma^2_{\dot{a}_j}.$$

Therefore, the contributions of $\dot{z}_{1,t-1},\ldots,\dot{z}_{k,t-1}$ and $\dot{a}_{jt}$ to the variance of $\dot{z}_{jt}$ are $\dot{\phi}^2_{j1}\sigma^2_{\dot{z}_1},\ldots,\dot{\phi}^2_{jk}\sigma^2_{\dot{z}_k}$ and $\sigma^2_{\dot{a}_j}$, and the proportional contributions to $\sigma^2_{\dot{z}_j}$ are

$$\dot{\phi}^2_{j1}\sigma^2_{\dot{z}_1}/\sigma^2_{\dot{z}_j}, \ \ldots \ ,\dot{\phi}^2_{jk}\sigma^2_{\dot{z}_k}/\sigma^2_{\dot{a}_j} \text{ and } \sigma^2_{\dot{a}_j}/\sigma^2_{\dot{z}_j} = 1 - \lambda_j.$$

These can be simplified by choosing the scaling of the $\underline{m}_j$ vectors so that $\sigma^2_{\dot{z}_j} = 1$. This is done by choosing $\underline{M}$ so that $\underline{M}\ \underline{\Gamma}_0(z)\underline{M}' = \underline{I}$. Letting the transformed series scaled in this manner be called $\underline{z}^*_t$, we have $\underline{z}^*_t = \underline{\phi}^*_t \underline{z}^*_{t-1} + \underline{a}^*_t$, and the proportional contributions to the variance are calculated as

$$\phi^{*2}_{j,i},\ i=1,2,\ldots,k \text{ and } 1 - \lambda_j.$$

## 4.0 ANALYSIS OF DATA

### 4.1 Preliminary Data Adjustment

The plots of the data obtained from the Royal Bank (Appendix 1) showed clear trends, and series 10 and 11 appeared to be exponential. Series 10 and 11 were logged, and then a straight line was fitted to each set of data and residuals were calculated. These residuals (the de-trended values of the series) were used as the data in the next step.

Looking at the plots of the residuals, several values were apparently outliers. The outliers were replaced with values calculated by doing a linear interpolation with points on either side of the outlier.

The number of data points supplied by the Royal Bank varied from one series to another. All values available were used for the regressions. After correcting for outliers, the last 79 points (the number in the shortest series) of each series were used as the $\{\underline{z}_t\}$ process. These values cover the period from the first quarter of 1957 to the third quarter of 1976, inclusive.

For each series $\underline{z}_{k,t}$, $k=1,2,\ldots,11$, the mean $\bar{z}_k$ was calculated, and the series

$$\underline{z}_t = \underline{z}_t - \bar{\underline{z}}$$ was used in the analysis.

These series appeared to be stationary in level. (Plots in Appendix 1.)

## 4.2 Model Identification, Canonical Analysis and Forecasting

Autocovariance and autocorrelation matrices were calculated by subroutine MACORR, which then calls MPARTL. This routine calculates and prints the quasi-partial correlations, AIC and the $\hat{\underline{\Sigma}}_{a,k}$ matrices.

| $\underline{k}$ | $\underline{\text{AIC}}$ |
|---|---|
| 1 | -213.01 |
| 2 | -175.73 |
| 3 | -138.14 |
| 4 | -119.60 |

Table 4.1 AIC Values for AR(k) Process, k=1,2,3,4

The minimum value of AIC indicates that the process can be modelled by an AR(1) process. The estimate of $\underline{\phi}_1$ in the model is given by the quasi-partial correlation matrix $\hat{\underline{\phi}}_{1,1}$ (table 4.2). The model for the process, then, is given by

$$\underline{z}_t = \hat{\underline{\phi}}_1 \underline{z}_{t-1} + \underline{a}_t.$$

The FORTRAN subroutine CANANAL, which is optionally called by MPARTL, produced the canonical analysis. The eigenvalues and eigenvectors obtained are shown in table 4.3.

The transformed series $\underline{z}_t^*$ were also calculated. (Plots in Appendix 1) In the first series, only .2% of the variation is forecastable, in the second 1.6%, and so on, up to the last series, in which 94.25% of the variation is forecastable. This is visible in the plots of the $\underline{z}_t^*$: the first few series appear to be white noise,

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| .3749 | .0296 | .1636 | -.0687 | .1540 | .0637 | -.0901 | .1585 | .0015 | -.0659 | -.0212 |
| .2946 | .7819 | -.0114 | .0095 | .1774 | -.2101 | -.0469 | .2991 | -.1754 | .1189 | -.2578 |
| .3630 | -.1727 | .8510 | -.1060 | .2541 | -.0371 | .0403 | .2132 | -.1858 | .1039 | -.3844 |
| .2816 | .2459 | -.2293 | .1314 | .8048 | -.0626 | .0456 | .0803 | -.1788 | .1718 | -.2787 |
| .1869 | .0363 | -.0734 | -.0184 | .8331 | -.2647 | .0306 | .0238 | -.0473 | .0759 | -.1133 |
| .4067 | .3316 | -.3460 | .1886 | .2747 | -.0410 | .1342 | .1752 | -.3102 | .0982 | -.2122 |
| .0103 | .0506 | -.0314 | -.0197 | .0945 | -.0370 | -.0144 | .0219 | -.1375 | .0073 | .0183 |
| -.0382 | -.1975 | .1847 | .0898 | .0036 | .3048 | -.0196 | .4528 | -.0744 | -.2157 | .3089 |
| -.1237 | -.1043 | .2630 | -.0971 | -.3154 | .2966 | -.1584 | .2126 | .5684 | .0278 | .0580 |
| .0514 | -.0041 | -.1342 | .0676 | .6229 | -.2650 | .0184 | -.1886 | .0590 | .8259 | -.0021 |
| -.0596 | .1057 | .0035 | -.1596 | .3681 | .2106 | .0139 | .0910 | .1219 | .1428 | .7290 |

Table 4.2 Estimate of $\phi_1$ in Fitted AR(1) Process

| .0025 | .0160 | .0627 | .1238 | .2484 | .3956 | .5340 | .7201 | .8597 | .9101 | .9425 |
|---|---|---|---|---|---|---|---|---|---|---|
| -.2334 | .7895 | .2218 | -.7993 | -.1314 | .3199 | -.0081 | -.1139 | .0877 | .0178 | .0387 |
| -.1638 | .1450 | .2798 | .6045 | -.3075 | .1055 | .7148 | -.2528 | .1118 | -.1558 | -.1252 |
| .1096 | -.2980 | -.1872 | -.1341 | .4946 | -.2076 | -.3718 | .0815 | -.0209 | .2212 | .2794 |
| -.0736 | .5225 | -1.0125 | .0881 | -.3647 | -.1987 | -.2231 | -.4160 | .1683 | .0593 | -.0730 |
| .1362 | -.5060 | .6278 | .3073 | -.4109 | 1.3624 | .5209 | .5116 | -.2483 | -.0260 | -.0585 |
| .3010 | -.8666 | .4054 | -.6739 | .3357 | -.5412 | -.2925 | -.3787 | .0545 | -.0134 | .0383 |
| 1.2113 | .7164 | .1344 | -.0277 | .1457 | -.0714 | -.0318 | .0070 | -.0439 | -.0021 | -.0790 |
| -.2443 | .2471 | .4432 | .6251 | -.2520 | -.1866 | -.4405 | -.0465 | .1180 | -.0290 | .0319 |
| .3895 | -.2569 | -.0209 | -.0938 | -.7029 | -.6488 | .1545 | .4249 | .1130 | .0252 | .0670 |
| -.1299 | .2164 | .3858 | .0745 | -.2374 | -.2861 | .0722 | -.0162 | .0068 | .2671 | -.0469 |
| .0920 | -.2748 | -.4690 | -.2749 | .5932 | .2213 | -.1153 | .1144 | .2223 | .0262 | -.0579 |

Table 4.3 Eigenvalues and Corresponding Eigenvectors (in columns) from Canonical Analysis

while the later series begin to show patterns.

The $\chi^2$ - test of the eigenvalues (done at level of significance $\alpha = .05$) indicates that, given that the seven largest roots are non-zero, the remaining four are not. However, if we assume that the six largest roots are significant, there is also significance in the remaining five. A reduction of dimensionality, then, could be made from eleven to seven, by using just the seven "most forecastable" series.

| p | $\chi^2$-statistic | p-value |
|---|---|---|
| 1 | .14 | .932 |
| 2 | 1.05 | .902 |
| 3 | 4.71 | .581 |
| 4 | 12.18 | .143 |
| 5 | 28.31 | .002 |

Table 4.4 $\chi^2$ test of Significance of Remaining 11-p Eigenvalues

A table of estimated proportional contributions to the variance of the $z^*_{i,t-1}$, $i=1,2,\ldots,11$ was also produced (table 4.5). This shows that, for the "most forecastable" series, the largest contribution is from the preceding value of that series, whereas the contribution to the least forecastable series is mostly from the white noise component, $\underline{a}^*_t$.

Forecasts were produced, using subroutine FORCST, for the original AR(1) process. The forecasts were then backtracked to the Royal Bank data units (by adding the means back on, inverting the

| | $z^*_{1,t-1}$ | $z^*_{2,t-1}$ | $z^*_{3,t-1}$ | $z^*_{4,t-1}$ | $z^*_{5,t-1}$ | $z^*_{6,t-1}$ | $z^*_{7,t-1}$ | $z^*_{8,t-1}$ | $z^*_{9,t-1}$ | $z^*_{10,t-1}$ | $z^*_{11,t-1}$ | $a^*_{j,t}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| z*(1,t) | .0000 | .0001 | .0021 | .0001 | .0000 | .0000 | .0000 | .0001 | .0001 | .0000 | .0000 | .9975 |
| z*(2,t) | .0110 | .0038 | .0002 | .0001 | .0003 | .0004 | .0000 | .0002 | .0000 | .0000 | .0000 | .9840 |
| z*(3,t) | .0152 | .0322 | .0012 | .0044 | .0004 | .0003 | .0062 | .0025 | .0003 | .0001 | .0000 | .9373 |
| z*(4,t) | .0009 | .0072 | .0062 | .1071 | .0005 | .0000 | .0006 | .0007 | .0003 | .0001 | .0000 | .8762 |
| z*(5,t) | .0014 | .0110 | .0009 | .0004 | .2264 | .0045 | .0005 | .0016 | .0000 | .0018 | .0000 | .7516 |
| z*(6,t) | .0054 | .0001 | .0044 | .0001 | .0042 | .3112 | .0111 | .0527 | .0010 | .0053 | .0004 | .6044 |
| z*(7,t) | .0190 | .0167 | .0003 | .0129 | .0004 | .0202 | .4441 | .0096 | .0028 | .0020 | .0060 | .4660 |
| z*(8,t) | .0002 | .0336 | .0209 | .0015 | .0164 | .0726 | .0177 | .5357 | .0126 | .0047 | .0042 | .2799 |
| z*(9,t) | .0043 | .0073 | .0188 | .0018 | .0004 | .0135 | .0009 | .0119 | .7736 | .0164 | .0108 | .1403 |
| z*(10,t) | .0012 | .0005 | .0017 | .0009 | .0087 | .0114 | .0002 | .0001 | .0265 | .8589 | .0000 | .0899 |
| z*(11,t) | .0014 | .0011 | .0000 | .0000 | .0000 | .0000 | .0077 | .0053 | .0146 | .0007 | .9117 | .0575 |

Table 4.5 Estimated Proportional Contributions to the Variance of $z^*_{j,t}$, j=1,2,...,11



| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| .0003 | -.0120 | .0461 | .0078 | .0065 | .0020 | .0002 | .0076 | -.0075 | -.0029 | -.0002 |
| -.1049 | -.0617 | -.0143 | -.0119 | .0165 | .0192 | -.0015 | -.0134 | -.0002 | .0020 | .0051 |
| -.1233 | .1795 | .0340 | .0663 | -.0205 | .0161 | -.0787 | .0496 | .0181 | .0081 | -.0024 |
| .0306 | -.0846 | -.0789 | .3273 | .0229 | -.0001 | -.0252 | -.0266 | -.0185 | .0094 | -.0039 |
| .0369 | .1049 | -.0292 | -.0191 | .4758 | .0668 | .0227 | -.0402 | .0052 | .0429 | -.0051 |
| .0733 | -.0077 | -.0660 | .0074 | -.0646 | .5578 | .1055 | .2295 | -.0310 | -.0725 | .0199 |
| -.1378 | .1293 | .0173 | .1137 | -.0195 | -.1421 | .6664 | .0979 | .0527 | .0450 | .0774 |
| .0155 | -.1833 | -.1445 | -.0383 | .1280 | -.2694 | -.1332 | .7319 | .1122 | .0685 | .0651 |
| .0654 | -.0857 | .1372 | .0423 | -.0202 | .1161 | -.0307 | -.1090 | .8795 | .1282 | .1040 |
| .0345 | -.0219 | .0411 | -.0300 | -.0933 | .1069 | -.0132 | -.0084 | -.1627 | .9268 | .0021 |
| .0371 | .0333 | .0031 | -.0017 | .0019 | -.0014 | -.0878 | -.0727 | -.1209 | -.0257 | .9548 |

Table 4.6 $\hat{\underline{\phi}}^*$

regression equation and exponentiating where necessary).

Forecasts were also produced for the seven most forecastable $\underline{z}^*$ series. The forecasts of the remaining four series were set to zero (their expected value, since these series are essentially white noise). These forecasts were then put back in terms of the $\underline{z}_t$ (by calculating $\underline{z}_t = \underline{M}^{-1} \underline{z}_t^*$ ), and again backtracked to the original Royal Bank units.

These forecasts, along with the actual values and probability limits, are shown in table 4.7.

| $\ell$ \ k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 95.94 | 103.73 | 91.39 | 102.21 | 101.33 | 104.30 | 101.10 | 97.29 | 104.86 | 124.09 | 109.65 |
| | 94.27 | 101.37 | 88.71 | 99.54 | 99.16 | 102.43 | 99.40 | 94.89 | 103.16 | 121.53 | 107.65 |
| 1 | 97.61 | 105.70 | 93.12 | 103.57 | 102.40 | 105.80 | 102.28 | 98.84 | 106.03 | 126.50 | 112.23 |
| | 95.77 | 100.99 | 90.01 | 101.30 | 100.88 | 103.69 | 99.60 | 95.21 | 103.37 | 124.29 | 107.60 |
| | 96.68 | 104.45 | 90.92 | 101.30 | 100.70 | 103.90 | 100.74 | 96.73 | 104.90 | 121.89 | 110.57 |
| | 96.29 | 104.63 | 91.99 | 102.44 | 101.45 | 104.35 | 101.02 | 96.86 | 105.34 | 124.64 | 109.97 |
| | 94.31 | 101.61 | 88.48 | 99.24 | 98.95 | 102.15 | 99.30 | 93.95 | 103.32 | 121.16 | 107.37 |
| 2 | 98.28 | 107.28 | 94.53 | 104.32 | 102.87 | 106.18 | 102.22 | 98.93 | 106.83 | 128.00 | 113.73 |
| | 95.90 | 99.84 | 88.93 | 100.53 | 100.53 | 103.40 | 100.34 | 96.67 | 103.97 | 126.24 | 110.68 |
| | 97.33 | 105.98 | 91.83 | 101.31 | 100.59 | 104.28 | 100.71 | 97.02 | 105.59 | 120.88 | 110.71 |
| | 96.49 | 105.15 | 91.81 | 101.85 | 101.03 | 104.31 | 100.75 | 92.26 | 105.42 | 125.23 | 110.99 |
| | 94.29 | 101.91 | 88.15 | 99.03 | 98.92 | 102.11 | 99.28 | 93.44 | 103.49 | 121.12 | 107.25 |
| 3 | 98.69 | 108.39 | 95.47 | 104.67 | 103.14 | 106.51 | 102.21 | 99.09 | 107.36 | 129.48 | 114.85 |
| | 95.77 | 99.56 | 88.48 | 100.76 | 100.55 | 103.46 | 98.71 | 97.02 | 103.31 | 127.82 | 109.72 |
| | 97.64 | 106.74 | 92.29 | 101.38 | 100.63 | 104.54 | 100.69 | 97.20 | 106.07 | 120.65 | 110.81 |
| | 96.61 | 105.77 | 91.99 | 101.97 | 101.15 | 104.52 | 100.76 | 96.15 | 105.74 | 125.99 | 111.45 |
| | 94.24 | 102.25 | 87.81 | 99.01 | 98.98 | 102.22 | 99.29 | 93.10 | 103.66 | 121.26 | 107.23 |
| 4 | 98.98 | 109.29 | 96.17 | 104.92 | 103.33 | 106.82 | 102.23 | 99.20 | 107.82 | 130.92 | 115.83 |
| | 96.40 | 100.17 | 88.67 | 101.17 | 100.27 | 102.74 | 98.90 | 96.23 | 104.14 | 128.74 | 109.92 |
| | 97.82 | 107.32 | 92.62 | 101.46 | 100.70 | 104.74 | 100.67 | 97.20 | 106.48 | 120.94 | 110.95 |

AR(1) forecast
lower limit of AR(1) forecast
upper limit of AR(1) forecast
actual value
forecast using 7 $z^*$ series

Table 4.7 $\ell$-step ahead Forecasts of Series k, k=1,2,...,11

## 5.0 DISCUSSION

When fitting the model, it was assumed that the series were all aligned properly, but this may not have been true - all the series are leading indicators of the economy, but they may not all have the same lead time. A few re-alignments were tried, but did not seem to make any difference ( $|\Sigma_a|$ and AIC should be reduced if the shift makes an improvement since the fit will be better). A deeper study of the indicators may have made suitable shifts obvious. In any case, a shift in some series may have led to a greater dimensionality reduction in the canonical analysis.

The lack of significant reduction of dimensionality in the canonical analysis suggests that, although there may be some underlying relationship reflected by some of the series, most of them are, in fact, bringing some new information.

With a certain amount of economic knowledge, possible relationships among the series could be suggested and tests, described in Box and Tiao (1977) can be done to see if these relationships lie in the subspace defined by the least forecastable components. In this problem, for instance, there is possibly a relationship between the series relating to building permits.

From the table of estimated proportional contributions to the variance of $z^*_{jt}$ (table 4.5), it appears that the first four series are essentially white noise - there is little contribution from past history. The last three combinations ( the most forecastable series) depend greatly on their own past. Looking at the $\underline{\phi}^*$ matrix (table 4.6) it can be seen that these three series ($z^*_{9t}, z^*_{10,t}, z^*_{11,t}$) can nearly be expressed as independent univariate AR(1) processes:

$$z^*_{9t} = .88z^*_{9,t-1} + a^*_{9t}$$
$$z^*_{10,t} = .93z^*_{10,t-1} + a^*_{10,t}$$
$$z^*_{11,t} = .95z^*_{11,t-1} + a^*_{11,t} ,$$

and, in fact, $z^*_{10,t}$ and $z^*_{11,t}$ are nearly non-stationary (since the autoregressive parameter is nearly one).

The forecasts produced using only the seven most forecastable components and back-tracked to the original units came out very close to those produced using the full AR(1) model. This tends to confirm the canonical analysis theory - by taking only the most forecastable linear combinations we have not lost any significant amount of information.

The actual values of the series lie within the confidence limits in 53 out of 66 cases (80.3%). The forecasts for series 2 (Investor's Index of 134 Common Stocks), however, seem to be particularly poor.

## 6.0 SUMMARY AND CONCLUSIONS

In this project, several methods for multiple time series analysis have been applied to an 11-dimensional set of Canadian economic data.

Computer programs were developed to identify, fit and forecast autoregressive processes, and to do a canonical analysis on the processes as they are fitted.

Using these programs, an AR(1) process was identified for the Royal Bank data set. A canonical analysis showed that no substantial reduction of dimensionality was possible. This suggests that the Royal Bank's eleven leading indicators were well chosen, since for the most part they are bringing new information.

Forecasts were produced both for the original eleven series and for a reduced set of seven linear combinations of them. The forecasts (except for series two) agreed well with each other and with the actual values.

## References


Akaike, H. (1976). Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. System Identification Advances and Case Studies, Mehra and Lainiotis editors. New York: Academic Press.

Alavi, A.S. (1973). ASARIMA(b;p,d,s,D,q)Models: Some Multivariate Extensions of Box-Jenkins Forecasting, Ph.D. thesis, Lancaster.

Baguley, R.W. (1976). The Royal Bank Trendicator Index of Leading Indicators. Technical Memo.

Box and Jenkins(1976). Time Series Analysis: Forecasting and Control (Revised Edition). San Fransisco: Holden-Day.

Box and Tiao(1977). A Canonical Analysis of Multiple Time Series. Biometrika, 64, 355-365.

Granger and Newbold(1977). Forecasting Economic Time Series. New York: Academic Press.

APPENDIX 1

PLOTS OF SERIES

A1.1 Plots of Original Royal Bank Data

x(1,t) Average Hours in Manufacturing

x(2,t) Investor's Index of 134 Common Stocks

x(3,t) TSE Price-Earnings Ratio

x(4,t) Housing Starts in Urban Areas

x(5,t) Number of Residential Building Permits Issued

x(6,t) Value of Total Building Permits

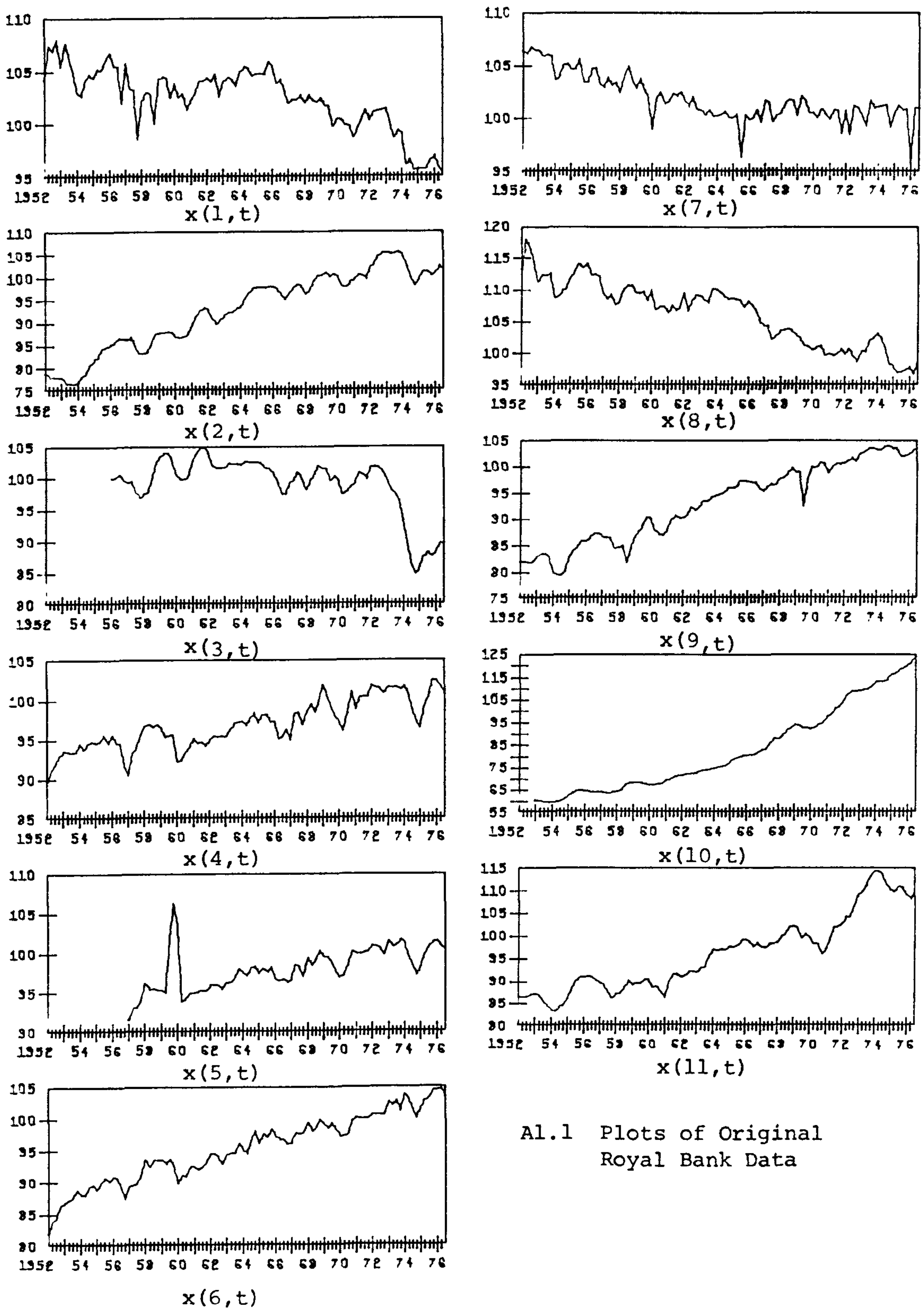
x(7,t) Liabilities of Business Failures

x(8,t) Ratio of Price to Unit Labour Cost

x(9,t) Primary Steel Production

x(10,t) Money Supply - M2

x(11,t) Corporate Pre-tax Profits




A1.1 Plots of Original Royal Bank Data

## A1.2 Plots of Detrended Series

These are the $z_{k,t}$ series defined in section 4.1. The original series from which they are obtained are listed below .

z(1,t) Average Hours in Manufacturing

z(2,t) Investor's Index of 134 Common Stocks

z(3,t) TSE Price-Earnings Ratio

z(4,t) Housing Starts in Urban Areas

z(5,t) Number of Residential Building Permits Issued

z(6,t) Value of Total Building Permits

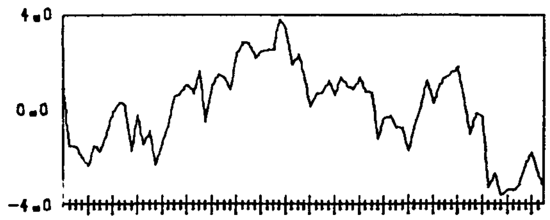
z(7,t) Liabilities of Business Failures

z(8,t) Ratio of Price to Unit Labour Cost

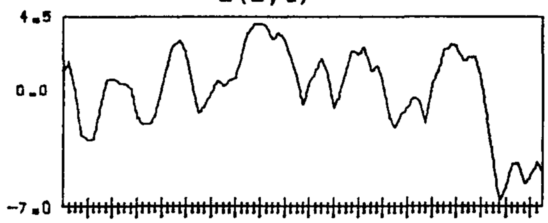
z(9,t) Primary Steel Production

z(10,t) Money Supply - M2

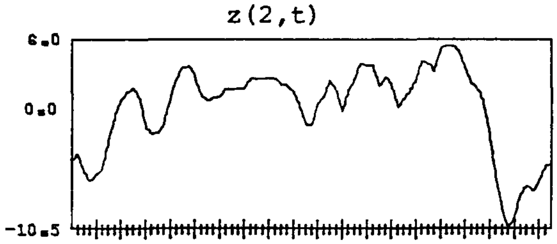
z(11,t) Corporate Pre-tax Profits



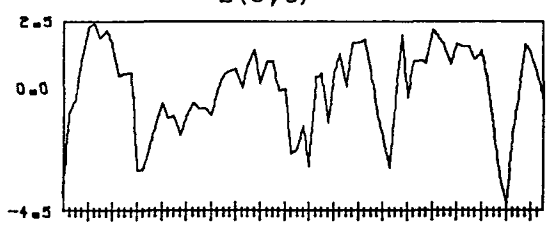

z(1,t)



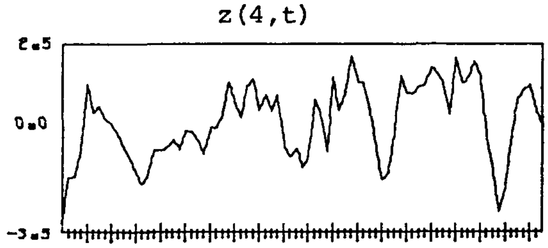

z(2,t)



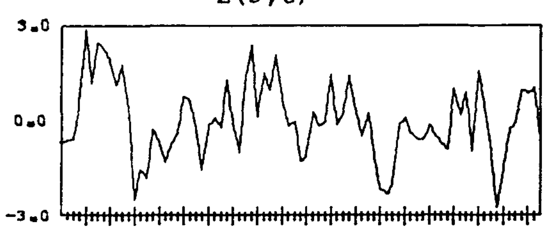

z(3,t)




z(4,t)

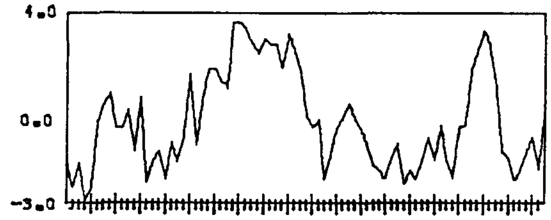



z(5,t)

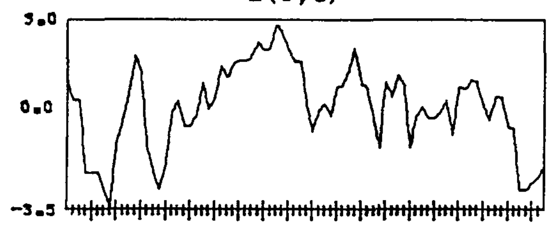



z(6,t)

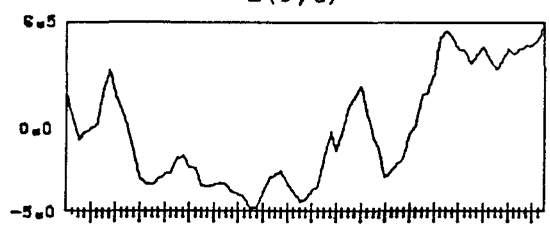

z(7,t)



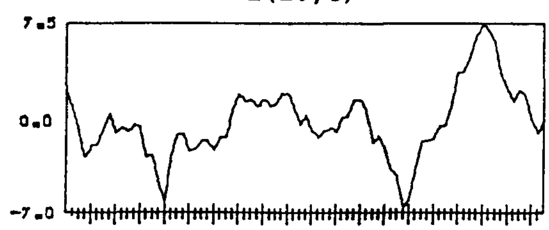

z(8,t)




z(9,t)

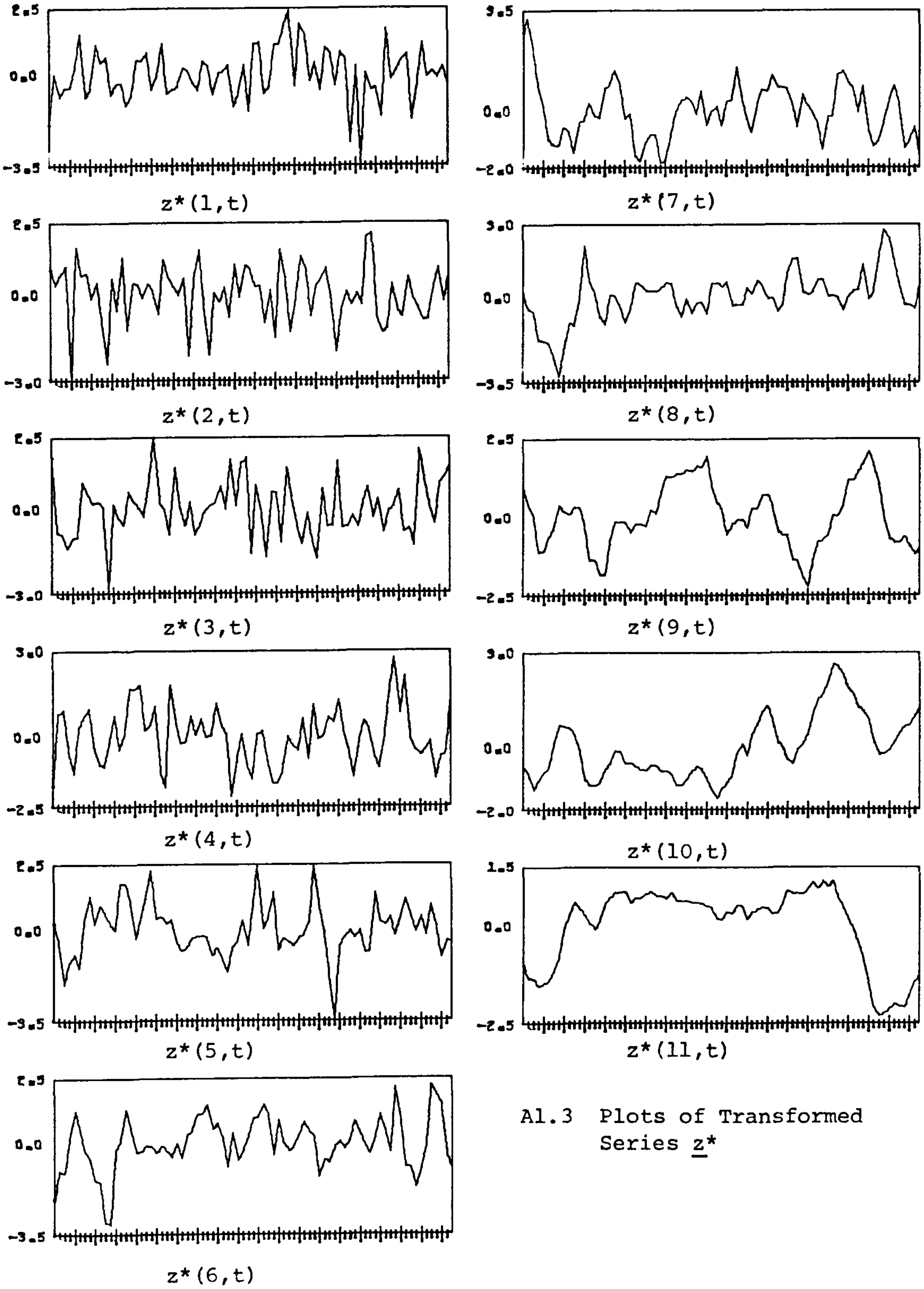



z(10,t)




z(11,t)

A1.2 Plots of Detrended Series




A1.3 Plots of Transformed Series z*

APPENDIX 2: COMPUTER PROGRAMS

Two sets of FORTRAN subroutines were written for this project: one to identify, fit and do canonical analysis of multiple AR time series (subroutines MACORR, MPARTL, MADD3, MSUB3, MPROD3, MTRA3, CANANAL and MMULT), and one to forecast AR series (subroutines FORCST and MPROD31).

The subroutines have been written generally, and should be able to handle up to an eleven-dimensional process. (A larger process would require changes to FORMATs and DIMENSIONs throughout.)

A call to MACORR will initiate the identification, fitting and canonical analysis of the series.

FORCST should be called to forecast an AR process.

A2.1 Subroutine MACORR(Z,NOB,NL,NSERIES,IC)

Identification, fitting and canonical analysis are started by a call to MACORR.

This subroutine calculates and prints the autocovariance and autocorrelation matrices and the determinants of the autocorrelation matrices up to lag NL for the series in Z.

MACORR then calls MPARTL.

Note that any data transformations must be done before calling MACORR and that the array Z must have exactly NOB rows and NSERIES columns in the calling program.

```
      SUBROUTINE MACORR(Z,NOB,NL,NSERIES,IC)
      COMMON /PARTL/C(11,11,11)
      DIMENSION Z(NOB,NSERIES)
      DIMENSION AC(11,11,11),ZBAR(11)
      DIMENSION DUM(11,11)
      COMMON AC,ZBAR,DUM
C Z(I,J)         ITH OBSERVATION ON JTH SERIES
C AC(I,J,K)      (I,J)TH ELEMENT OF AUTOCORRELATION MATRIX
C                OF LAG K
C C(I,J,K)       (I,J)TH ELEMENT OF AUTOCOVARIANCE MATRIX
C                AT LAG K
C DUM            WORK AREA
C LINV3F         IMSL SUBROUTINE, USED HERE TO CALCULATE
C                DETERMINANTS
C NOB            NUMBER OF OBSERVATIONS IN EACH SERIES
C NSERIES        NUMBER OF SERIES
C NL             FITTING IS TO BE DONE FOR AR(K),K=1,2,...,NL
C IC             CANONICAL ANALYSIS TO BE DONE FOR ALL AR(K)
C                K LESS THAN OR EQUAL TO IC
C NOTE *** ARRAY Z IN CALLING PROGRAM MUST BE DIMENSIONED
C                NOB BY NSERIES
C CALCULATE THE MEAN OF EACH SERIES
      DO 1 I=1,NSERIES
      ZBAR(I)=0.0
      DO 3 J= 1,NOB
3     ZBAR(I)=ZBAR(I)+Z(J,I)
      ZBAR (I)=ZBAR(I)/FLOAT(NOB)
1     CONTINUE
C CALCULATE THE AUTOCOVARIANCE MATRICES FOR LAGS 0,1,...,NL
      NL1=NL+1
      WRITE(6,201)
201   FORMAT("1AUTOCOVARIANCE MATRICES C(K)")
      DO 5 K=1,NL1
      WRITE(6,203)K-1
203   FORMAT("0LAG ",I2,/" ")
      L1=NOB-K+1
      DO 7 I=1,NSERIES
      DO 17 J=1,NSERIES
      C(I,J,K)=0.0
      DO 9 L=1,L1
      C(I,J,K)=C(I,J,K)+(Z(L,I)-ZBAR(I))*(Z(L+K-1,J)-
     +ZBAR(J))
9     CONTINUE
      C(I,J,K)=C(I,J,K)/FLOAT(NOB)
17    CONTINUE
      WRITE(6.205)(C(I,J,K),J=1,NSERIES)
205   FORMAT(" ",11(F9.4,2X))
7     CONTINUE
5     CONTINUE
C CALCULATE AUTOCORRELATION MATRICES FOR LAGS 0,1,...,NL
      WRITE(6,207)
207   FORMAT("1AUTOCORRELATION MATRICES R(K)")
      DO 11 K=1,NL1
      WRITE(6,202) K-1
202   FORMAT("0LAG ",I2)
      IJOB=4
      D1=2.
      DO 13 I=1.NSERIES
      DO 15 J=1,NSERIES
      DENOM=SQRT (C(I,I,1)*C(J,J,1))
      AC(I,J,K)=C(I.J,K)/DENOM
      DUM(I,J)=AC(I,J,K)
15    CONTINUE
      WRITE(6,205)(AC(I,J,K),J=1,NSERIES)
13    CONTINUE
```

```
C GET THE DETERMINANT OF THE AUTOCORRELATION MATRIX
      CALL LINV3F(DUM,B,IJOB,NSERIES,NSERIES,D1,D2,ZBAR,
     +IER)
C NOTE ZBAR HAS BEEN DESTROYED
      WRITE(6,209)K-1,D1*2.0**D2
209   FORMAT("0DETERMINANT OF R(",I2,")",5X,E13.6)
11    CONTINUE
C CALL MPARTL TO CALCULATE QUASI-PARTIAL CORRELATION
C MATRICES AND COVARIANCE MATRICES OF RESIDUALS FOR AR(K)
C PROCESSES.  CANONICAL ANALYSIS WILL AUTOMATICALLY
C BE DONE FOR AR(K),K LESS THAN OR = IC
      CALL MPARTL(IC,Z,NOB,NL,NSERIES)
      RETURN
      END
```

A2.2 Subroutine MPARTL

MPARTL calculates the quasi-partial correlation matrices using the recursive formulae in Alavi (1973). MPARTL also optionally calls CANANAL to do the canonical analysis for the order of the model currently being fitted.

Output includes $\hat{\underline{\Sigma}}_{a;k}$ and its determinant, the $\hat{\underline{\phi}}_{k,k}$ matrices and AIC for each order of AR process fitted.

```
      SUBROUTINE MPARTL(IC,Z,NOB,NL,NSERIES)
C NOTATION CORRESPONDS TO ALAVI"S
C A 1 FOLLOWING A MATRIX NAME IMPLIES THE INVERSE
C SIGBL,SGBL          MATRIX SIGMA B (K-2)
C SIGBT,SGBT          MATRIX SIGMA B (K-1)
C SIGAL,SGAL          MATRIX SIGMA A(K-2)
C SIGAT,SGAT          MATRIX SIGMA A(K-1)
C PHIL(I,J,L)         (I,J)TH ELEMENT OF PHI(K-1,L)
C PHIT(I,J,L)         (I,J)TH ELEMENT OF PHI(K,L)
C PSIL(I,J,L)         (I,J)TH ELEMENT OF PSI(K-1,L)
C PSIT(I,J,L)         (I,J)TH ELEMENT OF PSI(K,L)
C C(I,J,K+1)          (I,J)TH ELEMENT OF MATRIX C(K)
C SUM,DUM,DUM1,WKAR WORK MATRICES
C LINV2F              IMSL SUBROUTINE TO GET MATRIX INVERSE
      COMMON /PARTL/C(11,11,11)
      COMMON /CANL/ PHIL,SIGAT
      COMMON CT,PHIT,PSIL,PSIT,SUM,DUM,DUM1
      DIMENSION Z(NOB,NSERIES)
      DIMENSION SIGBL(11,11),SGBL(11,11,1),SIGAL(11,11),
     +SIGBL1(11,11),SGBL1(11,11,1),SIGAL1(11,11)
      DIMENSION SIGBT(11,11),SGBT(11,11,1),SIGAT(11,11),
     +SIGBT1(11,11),SGBT1(11,11,1),SIGAT1(11,11)
      DIMENSION SGAL(11,11,1),SGAL1(11,11,1),
     + SGAT(11,11,1),SGAT1(11,11,1),PSIL(11,11,10)
      DIMENSION CT(11,11,11),PHIL(11,11,10),PHIT(11,11,10
     +),PSIT(11,11,10),SUM(11,11,1),DUM(11,11,1),DUM1(11,11)
      DIMENSION WKAR(160)
      COMPLEX W(11)
      EQUIVALENCE (SIGBL(1,1),SGBL(1,1,1)),
     +            (SIGBL1(1,1),SGBL1(1,1,1))
      EQUIVALENCE (SIGAL(1,1),SGAL(1,1,1)),
     +            (SIGAL1(1,1),SGAL1(1,1,1))
      EQUIVALENCE (SIGAT(1,1),SGAT(1,1,1)),
     +            (SIGAT1(1,1),SGAT1(1,1,1))
      EQUIVALENCE (SIGBT(1,1),SGBT(1,1,1)),
     +            (SIGBT1(1,1),SGBT1(1,1,1))
      WRITE(6,601)
601   FORMAT(1H1)
C SIGMA A(0) = SIGMA B(0) = C(0)
      DO 1 I=1,NSERIES
      DO 1 J=1,NSERIES
      DUM1(I,J)=C(I,J,1)
      SIGBL(I,J)=C(I,J,1)
1     SIGAL(I,J)=C(I,J,1)
C GET DETERMINANT OF SIGMA A(0)
      IJOB=4
      D1=2.
      CALL LINV3F(DUM1,B,IJOB,NSERIES,NSERIES,D1,D2,DUM,IER)
      K=0
      WRITE (6,303)K,D1*2.0**D2
      WRITE(6,603)K
603   FORMAT("0SIGMA A ",I2)
      DO 17 I=1,NSERIES
17    WRITE(6,203)(SIGAL(I,J),J=1,NSERIES)
C GET EIGENVALUES OF SIGMA A(0)
      DO 16 I=1,NSERIES
      DO 16 J=1,NSERIES
16    DUM1(I,J)=C(I,J,1)
      IJOB=0
      CALL EIGRF(DUM1,NSERIES,NSERIES,IJOB,W,Z,IZ,DUM,IER)
      WRITE(6,301)(W(I),I=1,NSERIES)
301   FORMAT("0EIGENVALUES",3(/" ",5(F10.4,1X,F10.4,2X)))
```

```
      IDGT=0
      IA=11
C GET INVERSE OF SIGMA B(0) IN SIGBL1
      CALL LINV2F(SIGBL,NSERIES,IA,SIGBL1,IDGT,WKAR,IER)
C SET SIGAL1=SIGBL1
      DO 10 I=1,NSERIES
      DO 10 J=1,NSERIES
10    SIGAL1(I,J)=SIGBL1(I,J)
C GET MATRICES C TRANSPOSE FOR C(K),K=1,2,...,NL
      NL1=NL+1
      DO 2 J=2,NL1
      CALL MTRA3(C,J,CT,NSERIES)
2     CONTINUE
C GET STARTING POINTS PHI(1,1),PSI(1,1) FOR RECURSIVE
C FORMULAE GIVEN BY ALAVI
      CALL MPROD3(CT,2,SGBL1,1,PHIL,1,NSERIES)
      CALL MPROD3(C,2,SGAL1,1,PSIL,1,NSERIES)
      DO 12 I=1,NSERIES
      DO 12 J=1,NSERIES
12    DUM1(I,J)=PHIL(I,J,1)
      WRITE(6,201)
201   FORMAT("0PHI(1,1)")
      DO 4 I=1,NSERIES
4     WRITE(6,203)(PHIL(I,J,1), J=1,NSERIES)
203   FORMAT(" ",11(F10.4,2X))
      K=1
      CALL EIGRF(DUM1,NSERIES,NSERIES,IJOB,W,Z,IZ,DUM,IER)
      WRITE(6.301)(W(I),I=1,NSERIES)
C NOW START RECURSIVE CALCULATION OF THE PHIS
      DO 3 K=2,NL
      K1=K-1
      DO 24 I=1,NSERIES
      DO 24 J=1,NSERIES
24    SUM(I,J,1)=0.0
C CALCULATE SIGMA B(K-1)
      DO 26 J=1,K1
      CALL MPROD3(PSIL,J,CT,J+1,DUM,1,NSERIES)
26    CALL MADD3(DUM,1,SUM,1,SUM,1,NSERIES)
      CALL MSUB3(C,1,SUM,1,SGBT,1,NSERIES)
C CALCULATE SIGMA A(K-1)
      DO 28 I=1,NSERIES
      DO 28 J=1,NSERIES
28    SUM(I,J,1)=0.0
      DO 30 J=1,K1
      CALL MPROD3(PHIL,J,C,J+1,DUM,1,NSERIES)
30    CALL MADD3(DUM,1,SUM,1,SUM,1,NSERIES)
      CALL MSUB3(C,1,SUM,1,SGAT,1,NSERIES)
C GET DETERMINANT OF SIGMA A(K-1)
      DO 14 I=1,NSERIES
      DO 14 J=1,NSERIES
14    DUM1(I,J)=SIGAT(I,J)
      IJOB=4
      D1=2.
      CALL LINV3F(DUM1,B,IJOB,NSERIES,NSERIES,D1,D2,DUM,IER)
      WRITE(6,303)K-1,D1*2.0**D2
303   FORMAT("0DETERMINANT OF SIGMA A ",I2,5X,F12.5)
      XK1=K1
```

```
C CALCULATE AIC
      XNS=NSERIES
      AIC=FLOAT(NOB)*ALOG(D1*2.**D2)+2.*(XK1*XNS**2)
      WRITE(6,605)AIC
605   FORMAT(" AIC= ",F10.4)
      WRITE(6,603)K-1
      DO 19 I=1,NSERIES
19    WRITE(6,203)(SIGAT(I,J),J=1,NSERIES)
C DO CANONICAL ANALYSIS IF WANTED
      IF(K-1 .LE. IC)CALL CANANAL(K-1,NSERIES,Z,NOB)
      WRITE(6,205)K
205   FORMAT("0PHI(",I2,",J)")
C NOW CALCULATE THE PHI(K,J),J=1,...,K
      DO 20 I=1,NSERIES
      DO 20 J=1,NSERIES
20    SUM(I,J,1)=0.0
      DO 5 J=1,K1
      CALL MPROD3(PHIL,J,CT,K-J+1,DUM,1,NSERIES)
5     CALL MADD3(SUM,1,DUM,1,SUM,1,NSERIES)
      CALL MSUB3(CT,K+1,SUM,1,DUM,1,NSERIES)
      IDGT=0
      CALL LINV2F(SIGBT,NSERIES,IA,SIGBT1,IDGT,WKAR,IER)
      CALL MPROD3(DUM,1,SGBT1,1,PHIT,K,NSERIES)
      WRITE(6,207)K
207   FORMAT ("0J= ",I2)
      DO 6 I=1,NSERIES
6     WRITE(6,203)(PHIT(I,J,K),J=1,NSERIES)
      DO 7 J=1,K1
      CALL MPROD3(PHIT,K,PSIL,K-J,DUM,1,NSERIES)
      CALL MSUB3(PHIL,J,DUM,1,PHIT,J,NSERIES)
      WRITE(6,207)J
      DO 8 I=1,NSERIES
8     WRITE (6,203)(PHIT(I,J1,J),J1=1,NSERIES)
7     CONTINUE
C CALCULATE THE PSI(K,J),J=1,...,K
      DO 22 I=1,NSERIES
      DO 22 J=1,NSERIES
22    SUM(I,J,1)=0.0
      DO 9 J=1,K1
      CALL MPROD3(PSIL,J,C,K-J+1,DUM,1,NSERIES)
9     CALL MADD3(DUM,1,SUM,1,SUM,1,NSERIES)
      CALL MSUB3(C,K+1,SUM,1,DUM,1,NSERIES)
      IDGT=0
      CALL LINV2F(SIGAT,NSERIES,IA,SIGAT1,IDGT,WKAR,IER)
      CALL MPROD3(DUM,1,SGAT1,1,PSIT,K,NSERIES)
      DO 11 J=1,K1
      CALL MPROD3(PSIT,K,PHIL,K-J,DUM,1,NSERIES)
      CALL MSUB3(PSIL,J,DUM,1,PSIT,J,NSERIES)
11    CONTINUE
      DO 13 I=1,NSERIES
      DO 13 J=1,NSERIES
      SIGBL(I,J)=SIGBT(I,J)
      SIGAL(I,J)=SIGAT(I,J)
      DO 15 L=1,K
      PSIL(I,J,L)=PSIT(I,J,L)
15    PHIL(I,J,L)=PHIT(I,J,L)
13    CONTINUE
3     CONTINUE
      RETURN
      END
```

## A2.3 Subroutines MPROD3, MSUB3, MTRA3 and MADD3

These subroutines are used by MPARTL to do matrix operations on arrays with three dimensions. The third dimension corresponds to the subscripts on the matrices used in the notation of this project.

```
      SUBROUTINE MPROD3(A,K,B,L,C,M,NSERIES)
      DIMENSION A(11,11,1),B(11,11,1),C(11,11,1)
C MATRIX C(M) = A(K) * B(L), WHERE A(I,J,K) IS THE IJ-TH
C ELEMENT OF A(K)
C SIMILARLY FOR B,C
      DO 1 I1=1,NSERIES
      DO 1 I2=1,NSERIES
      C(I1,I2,M)=0.0
      DO 3 I3=1,NSERIES
3     C(I1,I2,M)=C(I1,I2,M)+A(I1,I3,K)*B(I3,I2,L)
1     CONTINUE
      RETURN
      END


      SUBROUTINE MSUB3 (A,K,B,L,C,M,NSERIES)
C C(M)=A(K)-B(L)
      DIMENSION A(11,11,1),B(11,11,1),C(11,11,1)
      DO 1 I1=1,NSERIES
      DO 1 I2=1,NSERIES
1     C(I1,I2,M)=A(I1,I2,K)-B(I1,I2,L)
      RETURN
      END


      SUBROUTINE MTRA3(A,K,AT,NSERIES)
C PUTS TRANSPOSE OF A(K) IN AT(K)
C I.E., AT(I,J,K)=A(J,I,K)
      DIMENSION A(11,11,1),AT(11,11,1)
      DO 1 I=1,NSERIES
      DO 1 J=1,NSERIES
1     AT(I,J,K)=A(J,I,K)
      RETURN
      END


      SUBROUTINE MADD3(A,K,B,L,C,M,NSERIES)
C C(M)=A(K)+B(L)
      DIMENSION A(11,11,1),B(11,11,1),C(11,11,1)
      DO 1 I=1,NSERIES
      DO 1 J=1,NSERIES
1     C(I,J,M)=A(I,J,K)+B(I,J,L)
      RETURN
      END
```

## A2.4 Subroutine CANANAL

This subroutine is called by MPARTL to do the canonical analysis as described in Box and Tiao (1977).

Output includes the matrix $\underline{C}_0(\hat{z})$, $\underline{C}_0^{-1}\underline{C}_0(\hat{z})$ and its eigenvalues and eigenvectors, the scaled matrix $\underline{M}$, the $\chi^2$ test of the remaining eigenvalues ($\alpha$=.05), a listing of the $\underline{z}^*$ series, the $\hat{\underline{\phi}}^*$ matrices and, for AR(1) processes, the table of estimated proportional contributions to the variance of the $z^*_{jt}$.

```
      SUBROUTINE CANANAL(IP,NSERIES,Z,NOB)
C THIS SUBROUTINE DOES CANONICAL ANALYSIS FOR AR(IP) PROCESS
C NOTATION CORRESPONDS TO BOX AND TIAO
C IP             ORDER OF THE PROCESS
C Z(I,J)         ITH OBSERVATION ON SERIES J.DIMENSION NOB
C                BY NSERIES
C M              MATRIX WITH SCALED EIGENVECTORS AS ROWS
C PHI,SIGA       ESTIMATES OF PHI AND SIGMA A FOR AR(IP) PROCESS
C ZSTAR          TRANSFORMED SERIES = M * Z
C C              AUTOCOVARIANCE MATRICES
C CZHAT          AUTOCOVARIANCE OF ZHAT(DEFINED BY BOX AND TIAO)
      COMMON /PARTL/C(11,11,11)
      COMMON /CANL/ PHI(11,11,10),SIGA(11,11)
      REAL M
      DIMENSION Z(NOB,NSERIES)
      DIMENSION WKAR(400)
      DIMENSION C01(11,11)
      DIMENSION INDEX(11)
      DIMENSION DUM(11,11),CZHAT(11,11),C0(11,11)
      DIMENSION DUMMY(11,11),ZSTAR(79,11),EIGEN(11),M(11,11)
      EQUIVALENCE (C0(1,1),C(1,1,1))
      COMPLEX DUM1(11,11)
      COMPLEX W(11)
      WRITE(6,614)IP
614   FORMAT("1CANONICAL ANALYSIS FOR AR(",I2,") PROCESS")
4     DO 1 I=1,NSERIES
      DO 1 J=1,NSERIES
1     CZHAT(I,J)=C0(I,J)-SIGA(I,J)
      WRITE(6,606)
606   FORMAT("0MATRIX C0(ZHAT)")
      DO 13 I=1,NSERIES
13    WRITE(6,603)(CZHAT(I,J),J=1,NSERIES)
      IDGT=0
      CALL LINV2F(C0,NSERIES,NSERIES,DUM,IDGT,WKAR,IER)
C C0 INVERSE IS NOW IN DUM. CALCULATE C0INV*C0(ZHAT)
      DO 9 I=1,NSERIES
      DO 9 J=1,NSERIES
      C01(I,J)=0.
      DO 11 K=1,NSERIES
11    C01(I,J)=C01(I,J)+DUM(I,K)*CZHAT(K,J)
9     CONTINUE
C PRINT C01=C0INV*C0(ZHAT)
      WRITE(6,605)
605   FORMAT("0MATRIX C0 INVERSE * C0(ZHAT)")
      DO 45 I=1,NSERIES
45    WRITE(6,603)(C01(I,J),J=1,NSERIES)
C NOW GET THE EIGENVALUES AND EIGENVECTORS OF THE MATRIX
C IN C01
      IJOB=1
      CALL EIGRF(C01,NSERIES,NSERIES,IJOB,W,DUM1,NSERIES,
     +WKAR,IER)
C CHECK THAT ALL EIGENVALUES ARE REAL
      DO 15 I=1,NSERIES
      IF(ABS(AIMAG(W(I))) .GT. 1.0E-6) GO TO 6
15    INDEX(I)=I
```

```
C SORT THE EIGENVALUES INTO ASCENDING ORDER
      N1=NSERIES-1
      DO 17 I=1,N1
      I1=I+1
      DO 19 J=I1,NSERIES
      IF(REAL(W(INDEX(I))) .LT. REAL(W(INDEX(J))))GO TO 19
      ITEMP=INDEX(I)
      INDEX(I)=INDEX(J)
      INDEX(J)=ITEMP
19    CONTINUE
17    CONTINUE
      DO 21 I=1,NSERIES
21    EIGEN(I)=REAL(W(INDEX(I)))
      WRITE(6,602)(EIGEN(I),I=1,NSERIES)
602   FORMAT("0SORTED EIGENVALUES",/" ",11(F9.4,2X))
C SORT THE EIGENVECTORS CORRESPONDINGLY AND PUT THEM INTO
C ROWS OF M
      DO 23 I=1,NSERIES
      DO 23 J=1,NSERIES
23    M(I,J)=REAL(DUM1(J,INDEX(I)))
      WRITE(6,604)
604   FORMAT("0SORTED MATRIX OF EIGENVECTORS(ROWS)")
      DO 47 I=1,NSERIES
47    WRITE(6,603)(M(I,J),J=1,NSERIES)
C SCALE M SO THE M*C(0)*MT = I
      CALL MMULT(M,C0,DUM,0,NSERIES)
      CALL MMULT(DUM,M,DUMMY,1,NSERIES)
      DO 25 I=1,NSERIES
      TEMP=SQRT(DUMMY(I,I))
      DO 25 J=1,NSERIES
25    M(I,J)=M(I,J)/TEMP
      WRITE(6,612)
612   FORMAT("0SCALED MATRIX M",/" ")
      DO 49 I=1,NSERIES
49    WRITE(6,603)(M(I,J),J=1,NSERIES)
C DO TESTS OF SIGNIFICANCE OF P SMALLEST EIGENVALUES
      WRITE(6,627)
627   FORMAT("0",15X,"CHI-SQUARE",/3X,"P",13X,"STATISTIC",
     +5X,"P-VALUE",/" ")
      CONST=-((NOB-NSERIES)-0.5*(2*NSERIES+1))
      CPROD=1.
      DO 41 I=1,NSERIES
      CPROD=CPROD*(1.-EIGEN(I))
      CHI=CONST*ALOG(CPROD)
      DF=2*I
      CALL MDCH(CHI,DF,PROB,IER)
      WRITE(6,629)I,CHI,1.-PROB
629   FORMAT(2X,I2,12X,F10.4,5X,F7.3)
      IF(1.-PROB .LT. .05)GO TO 43
41    CONTINUE
C CALCULATE THE TRANSFORMED SERIES
43    DO 27 I=1,NOB
      DO 27 J=1,NSERIES
27    ZSTAR(I,J)=0.
      DO 29 I=1,NOB
      DO 31 I1=1,NSERIES
      DO 31 I2=1,NSERIES
31    ZSTAR(I,I1)=ZSTAR(I,I1)+M(I1,I2)*Z(I,I2)
29    CONTINUE
      WRITE(6,902)
902   FORMAT("1TRANSFORMED SERIES Z*")
      DO 51 I=1,NOB
51    WRITE(6,603)(ZSTAR(I,J),J=1,NSERIES)
```

```
C NOW GET THE PHISTAR(L),L=1,2,...,IP
C PHISTAR(L) = M*PHI(L)*M(INVERSE)
      WRITE(6,610)
610   FORMAT("1PHI STAR(L)")
      DO 37 L=1,IP
      DO 39 I=1,NSERIES
      DO 39 J=1,NSERIES
39    DUM(I,J)=PHI(I,J,L)
      CALL MMULT(M,DUM,DUMMY,0,NSERIES)
      IDGT=0
      CALL LINV2F(M,NSERILS,NSERIES,DUM,IDGT,WKAR,IER)
      CALL MMULT(DUMMY,DUM,CZHAT,0,NSERIES)
37    WRITE(6,611)L
611   FORMAT("0L= ",I3)
      DO 53 I=1,NSERIES
53    WRITE(6,603)(CZHAT(I,J),J=1,NSERIES)
603   FORMAT(" ",11(F10.4,2X))
C PRINT TABLE OF ESTIMATED PROPORTIONAL CONTRIBUTIONS
C IF THIS IS AN AR(1) PROCESS
      IF(IP .NE. 1)GO TO 2
      WRITE(6,616)(L,L=1,NSERIES)
616   FORMAT("0ESTIMATED PROPORTIONAL CONTRIBUTIONS",/8X,
     +11("Z*",I2,"(T-1)",1X),2X,"A*JT",/" ")
      DO 55 L=1,NSERIES
55    WRITE(6,613)L,(CZHAT(L,J)**2,J=1,NSERIES),1.-EIGEN(L)
613   FORMAT(          " Z*",I2,"T",2X,12(F8.4,2X))
2     WRITE(6,615)
615   FORMAT(1H1)
      RETURN
6     WRITE(6,625) I
625   FORMAT("0***** EIGENVALUE ",I2," IS NOT REAL",
     +/" CANONICAL ANALYSIS TERMINATES HERE")
      RETURN
      END
```

A2.5 Subroutine MMULT

This subroutine is used by CANANAL to do matrix multiplications, optionally using the transpose of the second matrix.

```
      SUBROUTINE MMULT(A,B,C,IFLAG,NSERIES)
      DIMENSION A(11,11),B(11,11),C(11,11)
C IF IFLAG=1, C=A*BT
C IF IFLAG=0, C=A*B
      IF(IFLAG .EQ. 1)GO TO 2
      DO 1 I=1,NSERIES
      DO 1 J=1,NSERIES
      C(I,J)=0.
      DO 1 K=1,NSERIES
1     C(I,J)=C(I,J)+A(I,K)*B(K,J)
      RETURN
2     DO 3 I=1,NSERIES
      DO 3 J=1,NSERIES
      C(I,J)=0.
      DO 3 K=1,NSERIES
3     C(I,J)=C(I,J)+A(I,K)*B(J,K)
      RETURN
      END
```

A2.6 Subroutine FORCST(IP,IT,NFORC,NSERIES,NOB,Z,ZHAT,PHI,SIGMA,V)

This subroutine calculates and prints NFORC forecasts and probability limits for an AR(IP) process ($\hat{\underline{\phi}}$ matrices in PHI, $\hat{\underline{\Sigma}}_a$ in SIGMA). Forecasts are returned in ZHAT and diagonals of $\underline{V}(l)$ in V.

```
      SUBROUTINE FORCST(IP,IT,NFORC,NSERIES,NOB,Z,ZHAT,
     +PHI,SIGMA,V)
C FORCST PRODUCES NFORC FORECASTS OF AR(IP) PROCESS
C PROBABILITY LIMITS ALSO CALCULATED
      DIMENSION Z(NOB,NSERIES),ZHAT(10,NSERIES),
     +PHI(NSERIES,NSERIES,10),SIGMA(NSERIES,NSERIES),
     +V(10,NSERIES),DUM(11),PSI(11,11,10),DUM1(11,11),
     +DUM2(11,11)
C IP             ORDER OF THE PROCESS
C IT             ORIGIN FROM WHICH FORECASTS TO BE MADE
C NFORC          NUMBER OF FORECASTS WANTED (LESS THAN 11)
C NSERIES        NUMBER OF SERIES (LESS THAN 12)
C NOB            NUMBER OF OBSERVATIONS
C Z(I,J)         ITH OBSERVATION ON THE JTH SERIES
C ZHAT(I,J)      I-STEP AHEAD FORECAST OF SERIES J
C PHI(I,J,K)     I-J TH ELEMENT OF ESTIMATE OF PHI(K),
C                K=1,2,...,IP
C SIGMA          COVARIANCE MATRIX OF RESIDUALS AFTER
C                FITTING AR(IP) PROCESS
C DUM,DUM1,DUM2   WORK AREAS
C PSI(I,J,K)     I-J TH ELEMENT OF PSI(K),(THE PSI-WEIGHT
C                MATRICES)
C V(I,J)         THE ITH DIAGONAL ELEMENT OF VARIANCE
C                MATRIX V(J)
C INITIALIZE THE PSI MATRICES TO ZERO
      DO 11 I=1,NSERIES
      DO 11 J=1,NSERIES
      DO 11 K=1,10
11    PSI(I,J,K)=0.
      DO 1 L=1,NFORC
      DO 3 I=1,NSERIES
      V(L,I)=0.
3     ZHAT(L,I)=0.0
      DO 5 J=1,IP
C CALCULATE FORECASTS USING CONDITIONAL EXPECTATIONS
      IF(L-J .LE. 0)GO TO 2
      CALL MPROD31(PHI,J,ZHAT,10,L-J,DUM,NSERIES)
      GO TO 4
2     CALL MPROD31(PHI,J,Z,NOB,IT+L-J,DUM,NSERIES)
4     DO 7 I=1,NSERIES
7     ZHAT(L,I)=ZHAT(L,I)+DUM(I)
C CALCULATE PSI(L-1)
      IF(L .NE. 1)GO TO 6
C L=1, SO PSI(L-1) IS THE IDENTITY MATRIX
      IF(J .GT. 1)GOTO 5
      DO 29 K=1,NSERIES
29    PSI(K,K,1)=1.0
      GOTO 5
C ADD ON CURRENT TERM OF PSI(L-1), GIVEN BY
C  PSI(I) * PSI(L-J-1)
6     DO 31 K1=1,NSERIES
      DO 31 K2=1,NSERIES
      DO 31 K3=1,NSERIES
31    PSI(K1,K2,L)=PSI(K1,K2,L)+PHI(K1,K3,J)*
     +PSI(K3,K2,L-J)
5     CONTINUE
      WRITE(6,605)L-1
605   FORMAT("0MATRIX PSI(",I2,")")
      DO 41 K1=1,NSEPIES
41    WRITE(6,607)(PSI(K1,K2,L),K2=1,NSERIES)
607   FORMAT(" ",11(F9.4,2X))
```

```
C CALCULATE VARIANCE MATRIX IN DUM1 AND PULL OUT
C DIAGONALS
      IF(L .GT. 1)GO TO 8
C FORMULA FOR VARIANCE REDUCES TO V(L)=SIGMA
C PULL OUT THE DIAGONALS
      DO 33 I=1,NSERIES
      DO 33 J=1,NSERIES
      DUM1(I,J)=SIGMA(I,J)
      IF(I .EQ. J)V(L,I)=SIGMA(I,I)
33    CONTINUE
      GO TO 1
C ADD ON LATEST TERM TO VARIANCE - PSI(J)*SIGMA*PSIT(J)
8     DO 35 I=1,NSERIES
      DO 35 J=1,NSERIES
      DUM2(I,J)=0.
      DO 35 K=1,NSERIES
35    DUM2(I,J)=DUM2(I,J)+PSI(I,K,L)*SIGMA(K,J)
      DO 37 I=1,NSERIES
      DO 37 J=1.NSERIES
      DO 37 K=1,NSERIES
37    DUM1(I,J)=DUM1(I,J)+DUM2(I,K)*PSI(J,K,L)
      WRITE(6,609)L
609   FORMAT("0MATRIX V(",I2,")")
      DO 43 I=1,NSERIES
43    WRITE(6.607)(DUM1(I,J),J=1,NSERIES)
C NOW PULL OFF THE DIAGONALS AND PUT INTO V
      DO 39 I=1,NSERIES
39    V(L,I)=DUM1(I,I)
1     CONTINUE
      WRITE(6,601) (I,I=1,NSERIES)
601   FORMAT("1FORECASTS",/5X,"SERIES",/" LAG",2X,
     +11(3X,I2,6X))
      DO 9 L=1,NFORC
9     WRITE(6.603)L,(ZHAT(L.J),J=1,NSERIES),(SQRT(V(L,J))
     +,J=1,NSERIES)
603   FORMAT(1X,I3,11(2X,F9.4),/6X.11(2X,F9.4))
      RETURN
      END
```

A2.7 Subroutine MPROD31

This subroutine is used by FORCST to multiply an array with three dimensions by an array with two dimensions and put the result in a one-dimensional array.

```
      SUBROUTINE MPROD31(A,K,B,N,L,C,NS)
C USED BY FORCST FOR MATRIX MULTIPLICATIONS
      DIMENSION A(NS,NS,1),B(N,NS),C(NS)
      DO 1 I=1,NS
      C(I)=0.0
      DO 3 KK=1,NS
3     C(I)=A(I,KK,K)*B(L,KK)+C(I)
1     CONTINUE
      RETURN
      END
```