Dynamic transfer function-noise modelling (Some theoretical considerations)
Heuts, R.M.J.

Publication date:
1977

Citation for published version (APA):
KATHOLIEKE HOGESCHOOL TILBURG

REEKS "TER DISCUSSIE"

FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
KATHOLIEKE HOGESCHOOL TILBURG

REEKS "TER DISCUSSIE"

No. 77.053 Dec. 1976

DYNAMIC TRANSFER FUNCTION-NOISE MODELLING
(Some theoretical considerations)

R.M.J. Heuts

SUBFACULTEIT ECONOMETRIE
Contents

5.1. Introduction .................................................... 1

5.2. Technical details about the dynamic transfer function-noise modelling approach: random shock and autoregressive representations and forecasting formulas ................................................ 8

5.3. Identification of transfer function-noise models ................ 19
  5.3.1. Identification according to the Box-Jenkins approach .... 19
  5.3.2. Identification according to the Parzen-Priestley-Haugh approach ................................................. 25

5.4. Identification difficulties ..................................... 26

5.5. Estimation ...................................................... 28

5.6. Diagnostic checking applied to the residuals .................... 30

5.7. An example to illustrate the dynamic transfer function-noise modelling approach ................................................ 31

5.8. References ...................................................... 36
5.1. Introduction

There often arise situations in which it is felt that one variable is influenced by the behavior of a number of other variables. This type of relationship can generally be described as a cause and effect relationship.

A class of models which has been commonly employed in describing variable relationships is the general linear regression model

\[ x_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i z_{i,t} + \epsilon_t, \]

where \( z_{1,t}, \ldots, z_{k,t} \) are \( k \) deterministic variables at \( t = 1, \ldots, n \) different points. It is usually assumed that the \( \epsilon_t \) 's are independently and normally distributed with constant variance.

Many extensions of the basic model (5.1.1) have been made, one of which is attempting to describe the possible timing relationships between variables. For example, one could include various lagged input variables and also lagged output variables in the regression function. Some important difficulties that arise in building such distributed lag models with time series data are:

i) Some of the identification tools often used in building dynamic models, make identification of the right sort of model very difficult. For example, in using the Almon lag technique [1], one has to specify apriori the degree of a polynomial and the length of the lag structure. Jorgenson [8] recommends the use of rational distributed lag models, in which the maximum degree of two polynomials should be fixed in advance. See for this purpose chapter 15 in the book of Malinvaud [9].

ii) When lagged input variables (or regressors) are used, then usually the cross correlations between the input and output variables should give some idea about the dynamic character between them. However, as an input variable can itself be considered as a stochastic variable, which is probably autocorrelated in time, then a technique called "pre-whitening" (see section 5.3) might be more useful in finding the right dynamic structure.

iii) Even with quite complicated lag structures for the input and output variables, it may be important to consider a more general noise
structure than white noise, and then an identification procedure must also be presented to determine that noise structure (see the next sections how this can be done).

In a dynamic transfer function-noise model what is done, is to build a model which has two parts:

a) a transfer function part in which the output and input variables are related;

b) a noise part in which that part of the output series not accounted for by the input series is modelled in its own right using the univariate modelling technique as described in an other report [7].

Schematically the model may be represented as follows:

The development and use of a dynamic transfer function-noise model involves the following three important steps:

i) The identification stage: In this stage the cross correlations between the input and output series and the auto correlations and partial autocorrelations of the derived noise series are used respectively to determine the structure of the transfer function and the noise part of the model.

ii) The estimation stage: In this stage the parameters of the identified model are fitted and the model is tested for adequacy.

iii) The forecasting stage: In this stage the fitted model is used to predict future values of the output series.
We shall study three types of dynamic model formulations:

(A) In the case of one input series $z_t$ and one output series $x_t$ and a noise series $n_t$, we think in the first instance of a model formulation of the following type:

$$\sum_{i=0}^{r} \delta_i x_{t-i} = \sum_{j=0}^{s} \omega_j z_{t-j-b} + n_t, \quad t = \ldots, -1, 0, 1, \ldots,$$

where

i) $\{x_t\}$, $\{z_t\}$, and $\{n_t\}$ are assumed to be wide stationary stochastic processes;

ii) $z_t$ and $n_t$ satisfy the resp. ARMA-schemes (see definition 4, reference [7]);

$$\sum_{k=0}^{p} \pi_k n_{t-k} = \sum_{j=0}^{q} \kappa_j u_{t-j}, \quad t = \ldots, -1, 0, 1, \ldots;$$

$$\sum_{j=0}^{m} \phi_j z_{t-j} = \sum_{k=0}^{n} \theta_k u'_{t-k}, \quad t = \ldots, -1, 0, 1, \ldots;$$

iii) $b$ is a delay factor, which is an integer greater or equal to zero;

iv) the following generating functions are defined:

$$\pi(z):=\sum_{k=0}^{p} \pi_k z^k, \quad \phi(z):=\sum_{j=0}^{m} \phi_j z^j,$$

$$\kappa(z):=\sum_{j=0}^{q} \kappa_j z^j, \quad \theta(z):=\sum_{k=0}^{n} \theta_k z^k,$$

$$\delta(z):=\sum_{i=0}^{r} \delta_i z^i, \quad \omega(z):=\sum_{j=0}^{s} \omega_j z^j;$$

For identification, estimation and forecasting purposes we assume that the zeros of all the above generating functions are outside the unit circle;

v) $\{u_t\}$ is a white noise process and the $u_t$ are independent;

$v'$) $\{u'_t\}$ is a white noise process and the $u'_t$ are independent; the processes $\{u_t\}$ and $\{u'_t\}$ are independent.
For \( k \) stochastic inputs \( z_{t,i} \) (\( i = 1, \ldots, k \)) model (5.1.2) can be generalized as follows:

\[
(5.1.5) \quad r \sum_{i=0}^{k} \delta_{i} x_{t-i} = \sum_{i=1}^{k} \sum_{j=0}^{s_{i}} \omega_{i,j} z_{t-b_{i}-j,i} + n_{t},
\]

where the \( z_{t,i} \) \( i = 1, \ldots, k \) all satisfy a difference equation of the form (5.1.4).

(B) In this case the model formulation is as follows:

\[
(5.1.6) \quad r \sum_{i=0}^{r} \delta_{i} x_{t-i} = \sum_{j=0}^{s} \omega_{j} z_{t-j-b} + n_{t}, \quad t = \ldots-1,0,1,\ldots;
\]

where,

\[ d_{1}\]

i) \( \{x_{t}\}, \{w_{t} = \sum_{j=0}^{\infty} a_{j} z_{t-j}\} \), and \( \{n_{t}\} \) are assumed to be wide stationary stochastic processes, and \( z_{t} \) is the original input variable;

ii) the stochastic variables \( n_{t} \) and \( z_{t} \) satisfy the resp. ARMA- and ARIMA-schemes:

\[
(5.1.7) \quad \sum_{k=0}^{p} \pi_{k} n_{t-k} = \sum_{j=0}^{q} \kappa_{j} u_{t-j}, \quad t = \ldots-1,0,1,\ldots,
\]

\[
(5.1.8) \quad \sum_{j=0}^{\infty} \phi_{j} \sum_{k=0}^{\infty} a_{k} z_{t-k-j} = \sum_{k=0}^{\infty} \theta_{k} u_{t-k}, \quad t = \ldots-1,0,1,\ldots,
\]

(see definitions 4 and 6 in reference [7] for ARMA- and ARIMA-schemes);

iii) the generating functions are the same as under (A) and the zeros are outside the unit circle, but \( p_{1}(z) = \sum_{j=0}^{d_{1}} a_{j} z^{j} \) has all its zeros on the unit circle;

iv) the delay factor \( b \) and the processes \( \{u_{t}\} \) and \( \{u_{t}'\} \) are the same as under (A).
Remark

In case (B) we have assumed that the input process, which is non-wide stationary, can be made wide stationary by a linear operation.

The formulation under (B) can again be generalized for more input variables.

The third model formulation will be presented now.

(C) In this case the original output and input processes are non-stationary, but can be made wide-stationary by a linear operation. The difference equation is now

\[ \delta_r y_{t-i} = \sum_{i=0}^{r} \delta_i y_{t-i} = \sum_{j=0}^{s} \omega_j w_{t-j-b} + n_t, \quad t = \ldots -1, 0, 1, \ldots, \]

where,

i) \( \{y_t\} = \sum_{j=0}^{d_2} \delta_j x_{t-j} \), \( \{w_t\} = \sum_{j=0}^{d_1} \omega_j z_{t-j} \) and \( \{n_t\} \) are assumed to wide stationary processes, and \( x_t \) and \( z_t \) are the resp. original output and input series;

ii) the stochastic variables \( n_t \) and \( z_t \) satisfy the resp. ARMA- and ARIMA-schemes (5.1.7) and (5.1.8);

iii) the generating functions are the same as under (A) and the zeros are outside the unit circle, but \( p_1(z) = \sum_{j=0}^{d_1} \alpha_j z^j \) and \( p_2(z) = \sum_{j=0}^{d_2} \beta_j z^j \) have all its zeros on the unit circle;

iv) the delay factor \( b \) and the processes \( \{u_t\} \) and \( \{u_t'\} \) are the same as under (A).

This formulation under (C) can again easily be generalized for more input variables.

In the following flow-diagram the successive steps in the dynamic transfer function-noise approach are demonstrated.
MODEL STRUCTURE

CONCEPTUAL MODEL OF A SYSTEM

INPUT DATA

IDENTIFICATION

ROUGH ANALYSIS
IF NON-STATIONARITY IS PRESENT.
IF SO APPLY LINEAR OPERATIONS TO REMOVE IT

UNIVARIATE ANALYSIS OF INPUT AND OUTPUT SERIES

- CROSS CORRELATION ANALYSIS OF PREWHITENED SERIES
- AUTOCORR. AND PARTIAL CORR. ANALYSIS OF THE DERIVED NOISE SERIES

PRELIMINARY ESTIMATES

IDENTIFY THE DYNAMIC TRANSFER FUNCTION PART AND THE NOISE PART

ESTIMATION

RE-IDENTIFICATION OF MODEL STRUCTURE

INADEQUATE

CHECKING THE ADEQUACY OF THE MODEL

ESTIMATION OF ML OR LS ESTIMATORS AND ITS STANDARD DEVIATIONS

ADEQUATE

FORECASTING

GENERATE FORECASTS AND FORECASTING INTERVALS

Flow diagram 5.1.1: Flow-diagram of the dynamic transfer function-noise approach
It is our intention to stress the following points in the next sections:

(1) Like in the univariate case (ARIMA-schemes) the "random shock" form representation which Box and Jenkins give of the transfer function-noise scheme is not correct for a non-stationary output series. We shall give the correct representation, and show that a certain asymptotic result which we obtained in a previous report [7] can be extended in this case.

(2) Further, introducing matrix notation (matrix generating functions), we shall show that the forecasting formulas can be easily obtained even for the case with a delay factor $b = 0$, which case is not handled by Box and Jenkins.

(3) We shall present another identification procedure beside that of Box and Jenkins.

(4) We shall illustrate the theory with an elaborate example, taken from Box and Jenkins.
5.2. Technical details about the dynamic transfer function-noise modelling approach: random shock and autoregressive representations and forecasting formulas

Using the formulation of the dynamic model under (A) it can be shown in an analogous way as in [7], that the wide stationary solution of equation (5.1.2) can be written as

\[
(5.2.1) \quad x_t = \sum_{j=0}^{\infty} v_{j,A} u_{t-j} + \sum_{k=0}^{\infty} \psi_{k,A} u_{t-k},
\]

where the coefficients \( v_{j,A} \) and \( \psi_{k,A} \) can be determined from

\[
v_A(z) = \Theta(z)\omega(z)\frac{\phi(z)\delta(z)}{2} = \sum_{j=0}^{\infty} v_{j,A} z^j, \quad |z| \leq 1;
\]

\[
\psi_A(z) = \sum_{k=0}^{\infty} \psi_{k,A} z^k, \quad |z| \leq 1.
\]

For the model formulation under (B) the solution remains the same as in (5.2.1), however, the coefficients \( v_{j,B} \) and \( \psi_{j,B} \) are now determined as follows

\[
v_B(z) = \Theta(z)\omega(z)\frac{p_1(z)}{2} = \sum_{j=0}^{\infty} v_{j,B} z^j, \quad |z| \leq 1;
\]

\[
\psi_B(z) = \psi_A(z).
\]

The solution of the difference equation (5.1.9) under the model formulation (C) is, using the results of the ARIMA-process in [7], as follows:

\[
(5.2.2) \quad x_t = \nu_t + \sum_{j=0}^{\infty} v_{t,j,C} u_{t-j} + \sum_{j=0}^{\infty} \psi_{t,j,C} u_{t-k}, \quad t = 1, 2, \ldots,
\]

where \( \nu_t \) is in general the sum of a polynomial in \( t \) and a finite linear combination of powers of \( t \) and sinusoids.

---

1) the suffix A indicates that the coefficients belong to model formulation A.
The coefficients $v_{t,j,c}$ and $\psi_{t,j,c}$ can be determined as follows:

(a) when $j \leq t-1$, $v_{t,j,c}$ is the $j$-th Taylor coefficient of $v_B(z)$ in a neighborhood of $z = 0$, and $\psi_{t,j,c}$ is the $j$-th Taylor coefficient of $\psi_A(z)$ in a neighborhood of $z = 0$;

(b) when $j \geq t$, $v_{t,j,c}$ is the $j$-th Taylor coefficient of

$$
\left\{ \sum_{k=0}^{t-1} \beta_k z^k \right\} v_B(z), \quad |z| \leq 1,
$$

and $\psi_{t,j,c}$ is the $j$-th Taylor coefficient of

$$
\left\{ \sum_{k=0}^{t-1} \beta_k z^k \right\} \psi_A(z), \quad |z| \leq 1,
$$

and where $\beta_k$ can be found by taking the $k$-th coefficient in the power series development of $\frac{1}{p_2(z)}$ in a neighborhood of $z = 0$. 2)

(Note that Box and Jenkins formula (11.5.3) is not correct in this case).

The random shock forms in (5.2.1) and (5.2.2) can be used to determine the minimum mean square error forecasts as a function of $u_t$ and $u_t'$. We shall demonstrate it for solution (5.2.1), but it can be done on an analogous way for (5.2.2). Suppose that the point forecast $\hat{x}_{n+1,n}$ of $x_{n+1}$ made at origin $n$ is of the form

$$
\begin{align*}
\hat{x}_{n+1,n} &= \sum_{j=0}^{\infty} v_{1+j,A} u_{n-j} + \sum_{j=0}^{\infty} \psi_{1+j,A} u_{n-j},
\end{align*}
$$

Then

$$
\hat{x}_{n+1} - \hat{x}_{n+1,n} = \sum_{i=1}^{l-1} (v_{i,A} u_{n+1-i} + \psi_{i,A} u_{n+1-i}).
$$

2) In a similar way as theorem 4 in [7], it can be shown that the normalized $x_t$ in (5.2.2), $\frac{x_t}{\phi(t)}$, is under very general conditions asymptotic normally distributed.
and so

\[(5.2.4) \quad \mathcal{E} \left[ (x_{n+1} - \hat{x}_{n+1,n})^2 \right] = \sigma_u^2 \sum_{j=0}^{1-1} v_{1+j,A}^2 + \sigma_u^2 \sum_{j=0}^{1-1} \psi_{1+j,A}^2 + \sigma_u^2 \sum_{j=0}^{1-1} (v_{1+j,A} - \psi_{1+j,A})^2 \]

which is clearly minimized only if \( v_{1+j,A} = \psi_{1+j,A} \) and \( \psi_{1+j,A} = \psi_{1+j,A} \).

Thus the minimum mean square error forecast is

\[(5.2.5) \quad \hat{x}_{n+1,n} = \sum_{j=0}^{1-1} v_{1+j,A} u_{n-j} + \sum_{j=0}^{1-1} \psi_{1+j,A} u_{n-j}.\]

From the above it follows that the 1 step ahead forecast error is given by

\[x_{n+1} - \hat{x}_{n+1,n} = \sum_{i=0}^{1-1} (v_{i,A} u_{n+1-i} + \psi_{i,A} u_{n+1-i}),\]

which has expectation zero, and variance

\[(5.2.6) \quad \mathcal{E} (x_{n+1} - \hat{x}_{n+1,n})^2 = \sigma_u^2 \sum_{j=0}^{1-1} v_{j,A}^2 + \sigma_u^2 \sum_{j=0}^{1-1} \psi_{j,A}^2.\]

We shall now study how the point forecasts can be expressed as a weighted sum of previous observations. The model formulations (A), (B) and (C) shall be handled separately.

\[\text{(A)} \quad \text{Equations (5.1.2) and (5.1.4) can be written as}
\]

\[\sum_{i=0}^{r} \delta_i x_{t-i} - \sum_{j=0}^{s} \omega_j z_{t-j-b} = \sum_{k=0}^{\infty} \xi_k u_{t-k},\]

\[\sum_{j=0}^{m} \phi_j z_{t-j} = \sum_{k=0}^{n} \theta_k u'_{t-k},\]

where the \( \xi_k \) coefficient can be determined from

\[\xi(z) = \frac{\xi(z)}{\pi(z)} = \sum_{k=0}^{\infty} \xi_k z^k, \quad |z| \leq 1.\]
Equation (5.2.7) can also be written as

\[
\begin{align*}
\sum_{i=0}^{r} \delta_i \cdot x_{t-i} - \sum_{k=b}^{s+b} \omega_{k-b} \cdot z_{t-k} &= \sum_{k=0}^{\infty} \xi_k \cdot u_{t-k} \\
\sum_{j=0}^{m} \phi_j \cdot z_{t-j} &= \sum_{k=0}^{n} \Theta_k \cdot u'_{t-k}
\end{align*}
\]
(5.2.8)

or in matrix notation

\[
\sum_{k=0}^{\infty} A_k \cdot y_{t-k} = \sum_{j=0}^{\infty} B_j \cdot u_{t-j}
\]
(5.2.9)

where

\[
A_k = \begin{pmatrix}
\delta_k & -\omega_{k-b} \\
0 & \phi_k
\end{pmatrix}, \quad B_j = \begin{pmatrix}
\xi_j & 0 \\
0 & \Theta_j
\end{pmatrix}
\]

and \(\delta_j = 0, \) for \(j < 0\) and \(j > r; \) \(\omega_j = 0, \) for \(j < 0\) and \(j > s; \) \(\phi_j = 0, \)

for \(j < 0\) and \(j > m; \) \(\Theta_j = 0, \) \(j > n.\)

The matrix generating function \(A(z)\) defined by \(A(z) = \sum_{k=0}^{\infty} A_k \cdot z^k\) exists

for all \(z,\) and the matrix generating function \(B(z)\) defined by

\[
B(z) = \sum_{j=0}^{\infty} B_j \cdot z^j = \begin{pmatrix}
\xi(z) & 0 \\
0 & \Theta(z)
\end{pmatrix}, \text{ exists for } |z| \leq 1,
\]

when \(\pi(z) = 0\) has all its roots outside the unit circle. The matrix generating function \(C(z) = A(z) \cdot B(z)^{-1} = \sum_{k=0}^{\infty} C_k \cdot z^k, \) exists for \(|z| \leq 1,\) when

\(\kappa(z) = 0\) and \(\Theta(z) = 0\) have its roots outside the unit circle. The matrix generating function \(C(z)\) is as follows
Under the above conditions equation (5.2.9) can be written as

\[ (5.2.10) \quad U_t = \sum_{k=0}^{\infty} C_k Y_{t-k} \]

or

\[ (5.2.11) \quad C_0 Y_t = -\sum_{k=1}^{\infty} C_k Y_{t-k} + U_t \]

where the \( C_k \) are determined from \( C(z) \).

From the following it will become clear that to determine the best point forecast for \( x_{n+1} \) at point of time \( n \), it is useful to consider two cases, namely \( b \neq 0 \) and \( b = 0 \).

**Case \( b \neq 0 \)**

When \( \pi_0 = \kappa_0 = \delta_0 = \phi_0 = \theta_0 = 1 \), then \( C_0 \) is the identity matrix and then the best point forecast for \( Y_{n+1} \) at point of time \( n \) is

\[ (5.2.12) \quad \hat{Y}_{n+1,n} = \sum_{k=1}^{\infty} C_k Y_{n+1-k} \]

from which it follows

\[ (5.2.13) \quad \hat{X}_{n+1,n} = \sum_{j=1}^{\infty} p_{j,A} X_{n+1-j} + \sum_{j=1}^{\infty} q_{j,A} z_{n+1-j} \]

where the coefficients \( p_{j,A} \) and \( q_{j,A} \) can be determined from

\[ \frac{\pi(z) \delta(z)}{\kappa(z)} = \sum_{j=0}^{\infty} p_{j,A} z^j, \quad |z| < 1 \]

\[ \frac{\omega(z) \pi(z) z^b}{\kappa(z)} = \sum_{j=0}^{\infty} q_{j,A} z^j, \quad |z| < 1. \]
(Note that when \( b \geq 1 \), the coefficients \( q_0, A \cdots, q_{b-1}, A \) are all zero).

The best forecast for \( x_{n+2} \) at point of time \( n \) is

\[
\hat{x}_{n+2, n} = p_{1, A} x_{n+1, n} + \sum_{j=2}^{\infty} p_{j, A} x_{n+2-j} + \sum_{j=b}^{\infty} q_{j, A} z_{n+2-j} \quad \text{when } b \geq 2
\]

\[
= q_{b, A} z_{n+1, n} + \sum_{j=b+1}^{\infty} q_{j, A} z_{n+2-j} \quad \text{when } b = 1,
\]

where \( \hat{x}_{n+1, n} \) can be determined via

\[
\hat{x}_{n+1, n} = \sum_{j=1}^{\infty} \eta_{j, A} z_{n+1-j}
\]

where the \( \eta_{j, A} \) coefficient can be determined from the power series development of \( \frac{\phi(z)}{\theta(z)} \), \( |z| \leq 1 \).

On an analogous way as above the \( \hat{x}_{n+1, n} (1 \geq 3) \) can be determined.

Case \( b = 0 \)

When \( \pi_0 = \kappa_0 = \delta_0 = \phi_0 = \theta_0 = 1 \), then \( C_0 \) is the following matrix:

\[
C_0 = \begin{pmatrix} 1 & -\omega_0 \\ 0 & 1 \end{pmatrix}
\]

and the best point forecast for \( y_{n+1} \) at point of time \( n \) is

\[
\begin{pmatrix} \hat{x}_{n+1, n} - \omega_0 \hat{z}_{n+1, n} \\ \hat{z}_{n+1, n} \end{pmatrix} = - \sum_{k=1}^{\infty} c_k y_{n+1-k}
\]

or written out

\[
\hat{x}_{n+1, n} = \sum_{j=1}^{\infty} p_{j, A} x_{n+1-j} + \sum_{j=1}^{\infty} q_{j, A} z_{n+1-j}
\]

\[
\hat{z}_{n+1, n} = \sum_{j=1}^{\infty} \eta_{j, A} z_{n+1-j}
\]

where the coefficients \( p_{j, A} \) and \( q_{j, A} \) can be found in (5.2.14) and \( \delta_j \) are the same as in (5.2.16).

From (5.2.18) it follows that in this case we have

\[
\hat{x}_{n+1, n} = \sum_{j=1}^{\infty} p_{j, A} x_{n+1-j} + \sum_{j=1}^{\infty} q_{j, A} z_{n+1-j} + \omega_0 \sum_{j=1}^{\infty} \eta_{j, A} z_{n+1-j}
\]
The best point forecast for $Y_{n+2}$ at point of time $n$ is

$$
\begin{align*}
\hat{Y}_{n+2,n} &= \hat{Y}_{n+1,n} + \sum_{j=1}^{\infty} p_{j,B} \hat{X}_{n+1-j,n} + \sum_{j=1}^{\infty} q_{j,B} \hat{Z}_{n+1-j,n} \\
&+ \sum_{j=2}^{\infty} q_{j,A} \hat{Z}_{n+2-j,n} + w_0 \{ n_{1,A} \hat{X}_{n+1,n} + \sum_{j=2}^{\infty} n_{j,A} \hat{Z}_{n+2-j,n} \}.
\end{align*}
$$

From (5.2.20) it follows that

$$
\hat{X}_{n+2,n} = \hat{X}_{n+1,n} + \sum_{j=2}^{\infty} q_{j,A} \hat{Z}_{n+2-j,n} + w_0 \{ n_{1,A} \hat{X}_{n+1,n} + \sum_{j=2}^{\infty} n_{j,A} \hat{Z}_{n+2-j,n} \}.
$$

In an analogous way we can compute $\hat{X}_{n+1,n}$ ($1 \geq 3$).

Using the formulation of the dynamic model under (B) the formula's (5.2.1) and (5.2.6) remain the same, except that the coefficients $v_{j,A}$ and $\psi_{k,A}$ have to be replaced by $v_{j,B}$ and $\psi_{k,B}$.

We shall now study how the forecasts under model formulation (B) can be expressed as a weighted sum of previous observations. Equations (5.1.6), (5.1.7), and (5.1.8) can be written as

$$
\begin{align*}
\hat{X}_{n+1,n} &= \sum_{i=0}^{r} \delta_i X_{t-i} - \sum_{j=0}^{s} \omega_j W_{t-j-b} = \sum_{k=0}^{m} \varepsilon_k u_{t-k} \\
&+ \sum_{j=0}^{m} \theta_j W_{t-j} = \sum_{k=0}^{n} \theta_k u'_{t-k}.
\end{align*}
$$

This is an analogous set of equations as in (5.2.7). For forecasting we shall again consider two cases.

**Case $b \neq 0$**

The best point forecast for $X_{n+1}$ at point of time $n$ is

$$
\hat{X}_{n+1,n} = \sum_{j=1}^{\infty} p_{j,B} X_{n+1-j,n} + \sum_{j=1}^{\infty} q_{j,B} Z_{n+1-j,n}.
$$
where the coefficients $p_j,B$ and $q_j,B$ can be determined from

\[
\begin{align*}
-\pi(z)\delta(z) = \sum_{j=0}^{\infty} p_j,B z^j, \quad |z| \leq 1 \\
\omega(z)\pi(z)z^b p_1(z) = \sum_{j=0}^{\infty} q_j,B z^j, \quad |z| \leq 1.
\end{align*}
\]

(5.2.23)

The point forecasts $\hat{x}_{n+2,n}$ and $\hat{x}_{n+1,n}$ are the same as (5.2.15) and (5.2.16), except that the coefficients $p_j,A$, $q_j,A$ and $n_j,A$ have to be replaced by $p_j,B$, $q_j,B$ and $n_j,B$, where the last coefficient can be determined from

\[
(5.2.24) \quad \frac{\phi(z)}{\theta(z)} p_1(z) = \sum_{j=0}^{\infty} n_j,B z^j, \quad |z| \leq 1.
\]

The rest of the point forecasts can be found on an analogous way as under model formulation (A).

Case $b = 0$

In the same way as case $b = 0$ under (A) the best point forecast for $x_{n+1}$ at point of time $n$ is

\[
(5.2.25) \quad \hat{x}_{n+1,n} = \sum_{j=1}^{\infty} p_j,B \hat{x}_{n+1-j,n} + \sum_{j=1}^{\infty} q_j,B \hat{z}_{n+1-j,n} + w_0 \sum_{j=1}^{\infty} n_j,B \hat{z}_{n+1-j,n}.
\]

A general forecast formula is as follows:

\[
(5.2.26) \quad \hat{x}_{n+k,n} = \sum_{j=1}^{k-1} p_j,B \hat{x}_{n+1,n} + \sum_{j=k}^{\infty} p_j,B \hat{x}_{n+k-j,n} + \\
\sum_{j=1}^{k-1} q_j,B \hat{z}_{n+1,n} + \sum_{j=k}^{\infty} q_j,B \hat{z}_{n+k-j,n} + \\
\eta_{1,B} \hat{z}_{n+2,n} + \eta_{2,B} \hat{z}_{n+1,n} + \eta_{3,B} \hat{z}_{n}, \quad k = 3
\]

\[
\eta_{1,B} \hat{z}_{n+k-1,n} + \eta_{2,B} \hat{z}_{n+k-2,n} + \eta_{3,B} \hat{z}_{n+k-3,n}, \quad k \geq 4.
\]
We shall now study how the forecasts under model formulation (C) can be expressed as a weighted sum of previous observations. Equations (5.1.8) and (5.1.9) can be written as

\[
\begin{align*}
\sum_{i=1}^{r} \delta_i Y_{t-i} - \sum_{j=0}^{s} \omega_j w_{t-j-b} &= \sum_{j=0}^{\infty} \xi_k u_{t-k} \\
\sum_{j=0}^{m} \phi_j w_{t-j} &= \sum_{k=0}^{n} \theta_k u_{t-k},
\end{align*}
\]

(5.2.28)

where \(\{Y_t\}\) and \(\{w_t\}\) are the differenced stochastic processes of the resp. original stochastic processes \(\{x_t\}\) and \(\{z_t\}\) to get stationarity.

This is an analogous set of equations as in (5.2.7). For forecasting purposes we again consider two cases.

**Case \(b \neq 0\)**

The best point forecast for \(x_{n+1}\) at point of time \(n\) is

\[
\begin{align*}
\tilde{x}_{n+1,n} &= \sum_{j=1}^{\infty} p_{j,C} x_{n+1-j} + \sum_{j=1}^{\infty} q_{j,C} z_{n+1-j},
\end{align*}
\]

(5.2.29)

where the coefficients \(p_{j,C}\) and \(q_{j,C}\) can be determined from

\[
\begin{align*}
\frac{-\pi(z)\delta(z) p_{0}(z)}{\kappa(z)} &= \sum_{j=0}^{\infty} p_{j,C} z^j, \quad |z| \leq 1 \\
\frac{\omega(z)\pi(z)z^b p_{1}(z)}{\kappa(z)} &= \sum_{j=0}^{\infty} q_{j,C} z^j, \quad |z| \leq 1.
\end{align*}
\]

(5.2.30)

The point forecasts \(\tilde{x}_{n+2,n}\) and \(\tilde{x}_{n+1,n}\) are the same as in (5.2.15) and (5.2.16), except that the coefficients \(p_{j,A}\), \(q_{j,A}\) and \(\eta_{j,A}\) have to be replaced by \(p_{j,C}\), \(q_{j,C}\) and \(\eta_{j,C}\), where the coefficient \(\eta_{j,C}\) is equal to \(\eta_{j,B}\) in (5.2.24). The rest of the point forecasts can be determined in an analogous way as for model (A).
Case \( b = 0 \)

In this case the forecasting formulas are the same as in (5.2.25) and (5.2.26), except that the coefficients \( p_j, q_j, r_j \) must be replaced by \( p_j, q_j, r_j \). Updating the forecasts

It would be somewhat tedious to have to recompute forecasts from the above formulas whenever new information becomes available, but it turns out that a recursive updating procedure can be used. (We shall only demonstrate it for model (A) as it can be done analogously for (B) and (C)).

Namely from (5.2.5) we have

\[
(5.2.31) \quad \hat{x}_{n+1, n} = \sum_{j=0}^{\infty} v_{1+j, A} u_{n-j} + \sum_{j=0}^{\infty} \psi_{1+j, A} u_{n-j}
\]

where the coefficients \( v_{j, A} \) and \( \psi_{j, A} \) can be determined from the developments under formula (5.2.1).

Further we have

\[
(5.2.32) \quad \hat{x}_{n+1, n-1} = \sum_{j=0}^{\infty} v_{1+j, A} u_{n-j-1} + \sum_{j=0}^{\infty} \psi_{1+j, A} u_{n-j-1}
\]

Thus, by subtraction, we get

\[
(5.2.33) \quad \hat{x}_{n+1, n} = \hat{x}_{n+1, n-1} + v_{1, A} u_{n} + \psi_{1, A} u_{n}
\]

In words, the origin \( n-1 \) forecast of \( x_{n+1} \) can be updated to become the origin \( n \) forecast of \( x_{n+1} \) by adding a constant multiple of \( u_{n} \) and a constant multiple of \( u_{n} \). In this way the forecasts \( \hat{x}_{n+1, n-1} \), \( l = 1, 2, \ldots, h \) can be updated to yield the forecasts \( \hat{x}_{n+1, n}, l = 1, 2, \ldots, h \).

It remains to determine \( u_{1} \) and \( u_{n} \) in (5.2.33).

From (5.2.1) and (5.2.5) we know that the one step ahead forecast error is

\[
(5.2.34) \quad x_n - \hat{x}_{n, n-1} = v_{0, A} u_{n} + u_{n}
\]
and when \( b > 1 \), this formula becomes simpler, as \( v_{0,A} = 0 \) (see formulas under equation (5.2.1)).

On a similar way the one step ahead forecast error for the input series \( z_t \) can be found using the results in our previous report [7]

\[
(5.2.35) \quad \hat{z}_n - \hat{z}_{n,n-1} = \frac{u'_n}{n}.
\]

Formulas (5.2.34) and (5.2.35) are based on the assumption that

\( x_0 = \pi_0 = \delta_0 = \phi_0 = \theta_0 = 1. \)

Using (5.2.34) and (5.2.35) then (5.2.33) is equal to

\[
(5.2.36) \quad \hat{x}_{n+1,n} = \hat{x}_{n+1,n-1} + v_{1,A}(\hat{z}_n - \hat{z}_{n,n-1}) +
\]

\[
\psi_{1,A} \{(x_n - \hat{x}_{n,n-1}) - v_{0,A}(z_n - \hat{z}_{n,n-1})\}.
\]

The above updating forecasting formulas can be used to calculate the forecasts in an easily recursive way.
5.3. Identification of transfer function-noise models

5.3.1. Identification according to the Box-Jenkins approach

Under model formulation (A) the following representation can be obtained

\[(5.3.1.1)\] \( x_t = \sum_{j=0}^{\infty} v_{j,A} z_{t-j} + \sum_{j=0}^{\infty} \psi_{j,A} u_{t-j} \)

under model formulation (B) we can get

\[(5.3.1.2)\] \( x_t = \sum_{j=0}^{\infty} v_{j,B} w_{t-j} + \sum_{j=0}^{\infty} \psi_{j,B} u_{t-j} \)

and under formulation (C)

\[(5.3.1.3)\] \( x_t = \sum_{j=0}^{\infty} v_{j,C} w_{t-j} + \sum_{j=0}^{\infty} \psi_{j,C} u_{t-j} \)

where \( x_t \) and \( w_t \) are the transformed output and input series to get wide stationarity, and the coefficients \( v_{j(A,B,C)} \) and \( \psi_{j(A,B,C)} \) for the respective models (A), (B) and (C) can be obtained by

\[(5.3.1.4)\]

\[
\begin{align*}
\varphi(A,B,C)(z) &= \frac{\omega(z) z^b}{\delta(z)} = \sum_{j=0}^{\infty} v_{j(A,B,C)} z^j, \quad |z| \leq 1 \\
\psi(A,B,C)(z) &= \frac{\kappa(z)}{\delta(z)\pi(z)} = \sum_{j=0}^{\infty} \psi_{j(A,B,C)} z^j, \quad |z| \leq 1.
\end{align*}
\]

The identification of the transfer function part can be done using the following two important steps:

i) Derive rough estimates \( \hat{v}_{j(A,B,C)} \) of the so-called impulse response weights \( v_{j(A,B,C)} \) in (5.3.1.1), (5.3.1.2) and (5.3.1.3).

We first start by writing (5.1.4) in the so-called autoregressive form

\[(5.3.1.5)\] \( u_t = \sum_{j=0}^{\infty} e_{j,A} z_{t-j} \)
and for model (B) and (C) the same can be done:

$u'_t = \sum_{j=0}^{\infty} e_j(B,C) w_{t-j}$

where the coefficients $e_j(A,B,C)$ can be obtained from

$E(A,B,C)(z) = \frac{\phi(z)}{\theta(z)} = \sum_{j=0}^{\infty} e_j(A,B,C) z^j, |z| \leq 1.$

So, to a close approximation (when (5.3.1.5) or (5.3.1.6) is an infinite sum) the autocorrelated input series $z_t$ or $w_t$ are transformed to the white noise series $u'_t$. If we now apply the same transformation to $x_t$ or $y_t$, calling it $a'_t$, then

$a'_t = \sum_{j=0}^{\infty} \psi_j(A,B,C) u'_{t-j} + \sum_{j=0}^{\infty} \psi_j(A,B,C) u'_{t-j}$

where the coefficient $\psi_j(A,B,C)$ can be determined from

$\psi_j(A,B,C)(z) = \frac{\kappa(z)\phi(z)}{\delta(z)\theta(z)} = \sum_{j=0}^{\infty} \psi_j(A,B,C) z^j, |z| \leq 1.$

On multiplying (5.3.1.8) on both sides by $u'_{t-j}$ and taking expectations, we obtain

$\gamma_{u'a}(k) = \nu_k(A,B,C) \sigma_u'^2$

where

$\gamma_{u'a}(k) = \text{cov}(a'_t, u'_{t-k})$

or in terms of correlations

$\nu_k(A,B,C) = \rho_{u'a}(k) \frac{a}{\sigma_u'}, k = 0,1,2,\ldots$

where $\rho_{u'a}(k)$ is the cross correlation function between the prewhitened input and the correspondently transformed output series.

In order to get estimates of the $\nu_k$'s, the population values in (5.3.1.10) must be replaced by corresponding sample values.
Writing,

\[ s_{a}^2 = \frac{1}{n-1} \sum_{t=1}^{n} (s_t - \bar{s})^2 \]
\[ s_{u'}^2 = \frac{1}{n} \sum_{t=1}^{n} (u'_t)^2 \]

and

\[ r_{u'a}(k) = \frac{\frac{1}{n} \sum_{t=1}^{n-k} (s_{t+k} - \bar{s})u'_t}{s_a s_{u'}} \]

we find estimates of the \( v_k \)'s given by

\[ (5.3.1.11) \tilde{v}_k(A,B,C) = r_{u'a}(k) \frac{s_a}{s_{u'}} \]

\( k = 0,1,2,... \)

ii) These estimates can be used to give some idea of the values of \( s \) and \( r \), the orders of the respective polynomials \( \omega(z) \) and \( \delta(z) \) which might be appropriate, and also to select a tentative value for the delay parameter \( b \).

From (5.3.1.4) we have

\[ (5.3.1.12) \delta(z) v(A,B,C)(z) = \omega(z) z^b. \]

This implies that \( v_0, ..., v_{b-1} \) (for convenience we neglect the suffixes \( A, B, \) and \( C \) here) are all zero when \( b \geq 1 \) - and hence their estimates close to zero.

The polynomial \( \delta(z) \) determines the eventual pattern followed by the impulse weights \( v_k \) and the polynomial \( \omega(z) \) determines initial values for these weights. This can be seen by equating coefficients of \( z \) in the identity

\[ (5.3.1.13) (1+\delta_1 z+\delta_2 z^2 +...+ \delta_r z^r) (v_0+v_1 z+v_2 z^2 + ...) = \]

\[ = (\omega_0+\omega_1 z +...+ \omega_s z^s)z^b \]
from which we find

\[
\begin{align*}
  v_j &= 0 & j < b \\
  v_j &= -\delta_1 v_{j-1} - \cdots - \delta_r v_{j-r} + \omega_0 & j = b \\
  v_j &= -\delta_1 v_{j-1} - \cdots - \delta_r v_{j-r} + \omega_{j-b} & j = b+1, \ldots, b+s \\
  v_j &= -\delta_1 v_{j-1} - \cdots - \delta_r v_{j-r} & j > b+s.
\end{align*}
\]

(5.3.1.14)

We shall give an example of the impulse response weights behavior for a certain transfer function form.

Example

For the transfer function

\[
(5.3.1.15) \quad x_t - \delta_1 x_{t-1} - \delta_2 x_{t-2} = \omega_0 z_{t-b} - \omega_1 z_{t-b-1} + u_t
\]

The impulse response weights \(v_j\) are:

\[
\begin{align*}
  v_j &= 0 & j < b \\
  v_j &= \omega_0 & j = b \\
  v_j &= \delta_1 \omega_0 - \omega_1 & j = b+1 \\
  v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} & j > b+1.
\end{align*}
\]

In practice it is usually the case that an adequate model can be found without taking either \(r\) or \(s\) to be greater than two. A more detailed discussion of the properties of the impulse response weights \(v_k\) can be found in Box and Jenkins book [4].

As the impulse response weights are directly proportional to the cross correlations, we can also use the cross correlations for identification, because as we shall see, more is known about the asymptotic properties of
their estimators.

Given that a preliminary estimate \( \hat{v}_j(A, B, C) \) has been obtained, then an estimator of the noise series is provided by (see (5.3.1.1), (5.3.1.2) and (5.3.1.3))

\[
\hat{n}_t = x_t - \sum_{j=0}^{\infty} v_{j, A} z_{t-j} - E v_{j, A} z_{t-j}
\]

or

\[
\hat{n}_t = x_t (\text{or } x_t) - \sum_{j=0}^{\infty} v_{j(B, C)} w_{t-j}.
\]

The autocorrelations and partial autocorrelations of the noise series \( \hat{n}_t \) are then used to identify a suitable model for the noise structure (see [7] for further analysis).

We shall conclude this section indicating how a guess can be obtained for the order parameters \( r \) and \( s \) and for the delay factor \( b \). The delay factor may be guessed from the lag value between the two series at which the first large cross correlation occurs.

A guess of the number of \( \omega \) and \( \delta \) parameters needed can be determined by comparing the sample cross correlations of the appropriately differenced series with the theoretical cross correlations corresponding to particular values of \( r \) and \( s \). Figure 5.3.1.1 (which can be determined from table 10.1 in Box and Jenkins [4]) shows some theoretical cross correlation patterns obtained for different values of \( r \) and \( s \). This figure can be very helpful for the identification.
Figure 5.3.1.1: Cross correlations for different values of the r and s order parameters.
5.3.2. Identification according to the Parzen - Priestley - Haugh approach

The idea of the identification approach in this section can be found in Parzen [10], Priestley [12], and Haugh [6]. To identify the relationship existing between the input series $z_t$ or $w_t$, and the output series $x_t$ or $y_t$, we first start by constructing a univariate ARMA- or ARIMA-scheme for the input and output series separately, and then determine the cross correlation function of the associated univariate white noise input series, $u'_t = \sum_{j=0}^{\infty} c_j z_{t-j}$ or $u'_t = \sum_{j=0}^{\infty} c_j y_{t-j}$, and the univariate white noise output series, $a'_t = \sum_{j=0}^{\infty} c_j x_{t-j}$ or $a'_t = \sum_{j=0}^{\infty} c_j y_{t-j}$. This cross correlation function will be denoted as $\rho_{u'a'}(k)$, $k = 0, 1, 2, \ldots$. Studying the pattern of $\rho_{u'a'}(k)$, $k = 0, 1, 2, \ldots$ will give an indication of the dynamic transfer function behavior of the univariate white noise input and the univariate white noise output. This identified pattern may then be used to select the orders in the transfer function model, relating the two univariate residual series $u'_t$ and $a'_t$ (see Haugh [6]). Such a model may be written as

\[
\begin{align*}
\sum_{i=0}^{r} \delta_i a'_{t-i} &= \sum_{j=0}^{s} \omega_j u'_{t-j-b} + n_t, \quad t = \ldots, -1, 0, 1, \\
\sum_{k=0}^{p} \pi_k n'_{t-k} &= \sum_{j=0}^{q} \kappa_j u'_{t-j}, \quad t = \ldots, -1, 0, 1, 
\end{align*}
\]

(5.3.2.1)

where $u'_t$ is a white noise series uncorrelated with $u'_t$.

Once a model is identified relating $u'_t$ and $a'_t$, we may substitute in (5.3.2.1) the previously fitted univariate models for $z_t$ or $w_t$ and $x_t$ or $y_t$, to obtain a relation between the original wide stationary input and output series. To determine the relationship between the wide stationary input and output series in practice, one has to develop ratios of polynomials which probably can be simplified by cancelling like factors, to obtain a simpler form. Once a model has been identified, parameters are estimated simultaneously and further it will be checked if the model is adequate.

However in this report we shall use the identification approach as described in section 5.3.1, as it works good in practice and computer programs are available.
5.4. Identification difficulties

In the identification stage it is most important to have methods available which will clearly demonstrate what sort of model is applicable to the data. It has long been recognized in particular that the cross correlation function between two time series can be difficult to interpret because of the autocorrelation which will usually be present in each series [2].

Such autocorrelation can inflate the variance of cross correlation estimators above that expected when cross correlating two white noise series for example. Also, the cross correlation estimators at different lags will be correlated (possibly to a great extent). In fact many series arising in practice are either nonstationary or so highly autocorrelated, that cross correlating such series would not appear to have any value as an identification procedure.

To see what can happen in general consider the following expression for the asymptotic covariance between two cross correlation estimators 1):

\[
\text{cov}\{r_{xy}(k), r_{xy}(k+1)\} = (n-k)^{-1} \sum_{v=-\infty}^{\infty} \rho_x(v)\rho_y(v+1) + \\
+ \rho_{xy}(-v)\rho_{xy}(v+2k+1) + \rho_{xy}(k)\rho_{xy}(k+1)[\rho_{xy}(v)^2 + \frac{1}{2}\rho_x(v)^2] + \\
+ \frac{1}{2}\rho_y(v)^2 - \rho_{xy}(k)[\rho_x(v)\rho_{xy}(v+k+1) + \rho_{xy}(-v)\rho_y(v+k+1)] - \\
- \rho_{xy}(k+1)[\rho_x(v)\rho_{xy}(v+k) + \rho_{xy}(-v)\rho_y(v+k)]
\]

where the following notation will be used:

1) $x_t$ and $y_t$ are assumed to be jointly stationary and normally distributed.
\[
\begin{align*}
\hat{c}_{xy}(k) &= \left\{ \begin{array}{ll}
n^{-1} \sum_{t=1}^{n-k} (x_t - \overline{x})(y_{t+k} - \overline{y}), & k \geq 0 \\
n^{-1} \sum_{t=1}^{n+k} (x_{t-k} - \overline{x})(y_t - \overline{y}), & k < 0
\end{array} \right.
\end{align*}
\]

(5.4.2)

It should be noted that cross correlation estimators such as \( r_{xy}(k) \) above are consistent and asymptotically normal under reasonable conditions on \( x_t \) and \( y_t \) \([5]\). It can be seen from (5.4.1) that the covariance pattern of the cross correlations can be quite complicated, depending both on the true cross correlation function \( \rho_{xy}(k) \) and the two autocorrelation functions \( \rho_x(k), \rho_y(k) \). In particular if one were to judge the size of the individual estimators by comparison to an approximate standard error of \( n^{-1/2} \) (or \( (n-k)^{-1/2} \)), one could be greatly misled.

A special case which is of interest occurs when two processes are not cross correlated and one is white noise. Suppose \( y_t = \eta_t \) is a white noise process, but \( x_t \) is autocorrelated, then using formula (5.4.1) we get the following asymptotic result

\[
\begin{align*}
\text{Cov}(r_{xy}(k), r_{xy}(k+1)) &= (n-k)\rho_{xx}(1) \\
\text{Var}(r_{xy}(k)) &= (n-k)^{-1}.
\end{align*}
\]

(5.4.3)

So for large samples we may under the null hypothesis (= no cross correlation) use the approximate standard deviation of \( n^{-1/2} \) (or \( (n-k)^{-1/2} \)). As the cross correlations \( r_{ux}(k) \) are proportional to the impulse response weights, these cross correlations can also be used for identification and they have the advantage that something is known about their asymptotic distribution.
5.5. Estimation

Using the results of section 5.2 equations (5.1.2), (5.1.6) and (5.1.9) can be written as

\[
(5.5.1) \quad u_t = \sum_{j=0}^{\infty} p^*_j(A,B,C) x_{t-j} + \sum_{j=0}^{\infty} q^*_j(A,B,C) z_{t-j}
\]

where for model (A)

\[
\begin{align*}
 p^*_j, A &= -p_j, A \quad \text{and} \quad q^*_j, A &= -q_j, A \quad \text{(see (5.2.14)),}
\end{align*}
\]

and for model (B)

\[
\begin{align*}
 p^*_j, B &= -p_j, B \quad \text{and} \quad q^*_j, B &= -q_j, B \quad \text{(see (5.2.23)),}
\end{align*}
\]

and for model (C)

\[
\begin{align*}
 p^*_j, C &= -p_j, C \quad \text{and} \quad q^*_j, C &= -q_j, C \quad \text{(see (5.2.30)).}
\end{align*}
\]

Denote a set of guessed values for the parameters by

\[
\beta_0 = (\delta_1, 0, \ldots, \delta_r, 0; \omega_0, 0, \ldots, \omega_s, 0; \pi_1, 0, \ldots, \pi_p, 0; \kappa_1, 0, \ldots, \kappa_q, 0)
\]

and let \(u_{t, 0}\) be that value of \(u_t\) in (5.5.1) for the guessed values \(\beta_0\) (assuming that the orders of the schemes are known; and for convenience neglecting the suffixes \(A, B, \text{and } C\)),

\[
(5.5.3) \quad u_{t, 0} = \sum_{j=0}^{\infty} p^*_j, 0 x_{t-j} + \sum_{j=0}^{\infty} q^*_j, 0 z_{t-j},
\]

where the \(p^*_j, 0\) and \(q^*_j, 0\) are the \(p^*_j\) and \(q^*_j\) with known \(\beta_0\) vector.

Then a first order Taylor series expansion of \(u_t = u_t(\beta)\) about the parameter values \(\beta = \beta_0\) gives

\[
(5.5.4) \quad u_t \approx u_{t, 0} + \sum_{i=1}^{r} (\delta_i - \delta_i, 0) \left(\frac{\partial u_t}{\partial \delta_i}\right)_{\beta_0} + \sum_{h=1}^{s} (\omega_h - \omega_h, 0) \left(\frac{\partial u_t}{\partial \omega_h}\right)_{\beta_0} + \sum_{g=1}^{p} (\pi_g - \pi_g, 0) \left(\frac{\partial u_t}{\partial \pi_g}\right)_{\beta_0} + \sum_{f=0}^{q} (\kappa_f - \kappa_f, 0) \left(\frac{\partial u_t}{\partial \kappa_f}\right)_{\beta_0}
\]
or rearranging

\[(5.5.5) \quad u_{t,0} = - \sum_{i=1}^{r} (\delta_i - \delta_{i,0}) \frac{\partial u_t}{\partial \delta_i} \beta_0 - \sum_{h=1}^{s} (\omega_h - \omega_{h,0}) \frac{\partial u_t}{\partial \omega_h} \beta_0 - \sum_{p=1}^{p} (\pi_p - \pi_{p,0}) \frac{\partial u_t}{\partial \pi_p} \beta_0 - \sum_{g=1}^{g} (\kappa_f - \kappa_{f,0}) \frac{\partial u_t}{\partial \kappa_f} \beta_0 + u_t.\]

The initial guesses can then be modified by applying ordinary least squares to this model, and this procedure is repeated until convergence is achieved. One can use a non-linear least squares routine where the required partial derivatives are evaluated numerically at each step of the iteration procedure.

The covariance matrix of the estimates may be obtained from the converged value of the matrix \((R'R)^{-1} \sigma_u^2\) as described in [4,7].
5.6. Diagnostic checking applied to the residuals

Suppose that a transfer function-noise model has been fitted by least squares and the residuals $\hat{u}_t$'s calculated by substituting least squares estimates for the parameters, the estimated autocorrelation function $r_{\hat{u}}(k)$ of these residuals is computed and also the cross correlation function $r_{\hat{u}'\hat{u}}(k)$ between the estimated prewhitened input $\hat{u}_t'$ and the residuals $\hat{u}_t$. Then Box and Jenkins show that if the autocorrelation function $r_{\hat{u}}(k)$ or the cross correlation function $r_{\hat{u}'\hat{u}}(k)$ show marked patterns, this suggests model inadequacy. However, it is dangerous to read too much into the behaviour of the individual correlations. A helpful overall check is based on two chi-squared tests [11]. The first of these runs as follows. Consider the first $K$ estimated autocorrelations $r_{\hat{u}}(1),...,r_{\hat{u}}(K)$ and $K$ be sufficiently large, then, if the form of the model is correct, the quantity

$$Q = m \sum_{k=1}^{K} r_{\hat{u}}^2(k)$$

is approximately distributed as chi-square with $K - p - q$ degrees of freedom, where $m$ is the number of estimated residuals which is available. We see that the chi-square statistic only depends on the number of parameters ($p + q$) in the noise model. The second test is employed as follows. The cross correlations $r_{\hat{u}'\hat{u}}(k)$ for $k = 0,1,2,...,K$ are estimated for $K$ sufficiently large. When the functional form of the model is correct, the quantity

$$S = m \sum_{k=0}^{K} r_{\hat{u}'\hat{u}}^2(k)$$

is approximately distributed as chi-square with $K + 1 - (r + s + 1)$ degrees of freedom, where $(r + s + 1)$ is the number of parameters in the transfer function part of the model.
We shall use an example taken from the book of Box and Jenkins [4] to illustrate the dynamic transfer function-noise modelling approach. Suppose that we wish to describe an input series $z_t$ and the corresponding output series $x_t$ from some physical system. The physical process is a gas furnace in which air and methane are combined to form a mixture of gases containing CO$_2$ (carbon dioxide). The air feed was kept constant, but the methane feed rate could be varied in any desired manner and the resulting CO$_2$ concentration in the off gases measured. Every nine seconds data were collected of the in- and output series to provide information about the dynamics of the system over a region of interest. The input series was put in at level zero and could during the experiment vary between an upper and lower bound. The number of observations of both series is 296. The autocorrelations of the output and input series damp out fairly quickly, indicating that no differencing operation is required.

Using the identification and estimation procedures as described in [7] the input series $z_t$ is well described by a third-order autoregressive scheme:

$$z_t - 1.97 z_{t-1} + 1.37 z_{t-2} - 0.34 z_{t-3} = u'_t,$$

with $s_{u'}^2 = 0.0353$; and the output series $x_t$ by an ARMA (4,2)-scheme:

$$x_t - 2.42 x_{t-1} + 2.38 x_{t-2} - 1.16 x_{t-3} + 0.23 x_{t-4} =$$

$$\epsilon_t - 0.31 \epsilon_{t-1} + 0.47 \epsilon_{t-2}$$

with $s_{\epsilon}^2 = 0.1081$.

Following the remarks in section 5.3.1 the input series is transformed to white noise by:

$$u'_t = z_t - 1.97 z_{t-1} + 1.37 z_{t-2} - 0.34 z_{t-3}$$
and the same transformation is applied to the output series:

$$(5.7.4) \quad a_t = x_t - 1.97 x_{t-1} + 1.37 x_{t-2} - 0.34 x_{t-3},$$

with $s_{u'} = 0.188$ and $s_a = 0.358$.

The estimated cross correlation function between $u_t'$ and $a_t$ is shown in Table 5.7.1, together with the estimate $(5.3.1.11)$ of the impulse response weights $\hat{v}_k = \frac{0.358}{0.188} r_{u'a}(k)$.

Table 5.7.1: Estimated cross correlations after prewhitening, approximate standard errors for the cross correlations, and approximate impulse response weights for the gas furnace data

<table>
<thead>
<tr>
<th>k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{u'a}(k)$</td>
<td>-0.01</td>
<td>0.05</td>
<td>-0.03</td>
<td>-0.28</td>
<td>-0.46</td>
<td>-0.27</td>
<td>-0.17</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(r_{u'a}(k))$</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>$\hat{v}_k$</td>
<td>-0.02</td>
<td>0.10</td>
<td>-0.06</td>
<td>-0.53</td>
<td>-0.63</td>
<td>-0.88</td>
<td>-0.52</td>
<td>-0.32</td>
<td>-0.06</td>
<td>0.06</td>
<td>-0.10</td>
</tr>
</tbody>
</table>

Remarks by Table 5.7.1.
The approximate standard errors for the cross correlations $r_{u'a}(k)$ shown in Table 5.7.1 are the square roots of the variances obtained from $(5.4.3)$. The estimated cross correlations together with two standard error limits are plotted in Figure 5.7.1.
Using the results in Figure 5.3.1.1 the pattern of the $\hat{\Psi}$'s might be accounted for by a scheme with $(r, s, b)$ equal to $(1, 2, 3)$:

\begin{equation}
X_t + \delta_1 X_{t-1} = \omega_0 Z_{t-3} + \omega_1 Z_{t-4} + \omega_2 Z_{t-5} + n_t
\end{equation}

Using (5.3.1.16) and the previous computed impulse response weights the noise series can be estimated as follows:

\begin{equation}
\hat{n}_t = x_t - \sum_{j=0}^{\infty} \hat{\Psi}_j z_{t-j}.
\end{equation}

Studying the autocorrelations and partial autocorrelations of $\hat{n}_t$ leads to identification of the noise series. In doing so the scheme for $n_t$ would be:

\begin{equation}
\hat{n}_t + \pi_1 \hat{n}_{t-1} + \pi_2 \hat{n}_{t-2} = \hat{u}_t.
\end{equation}

The simultaneous estimation of the parameters as described in section 5.5 gives

\begin{equation}
x_t - 0.57 x_{t-1} = -0.53 z_{t-3} - 0.37 z_{t-4} - 0.51 z_{t-5} + n_t
\end{equation}

\begin{tabular}{llll}
& (0.21) & (0.08) & (0.15) & (0.16)
\end{tabular}
and

\[(5.7.9) \quad n_t = 1.53 \, n_{t-1} + 0.63 \, n_{t-2} = u_t, \quad (0.05) \quad (0.05)\]

with \(s_u^2 = 0.0561\), where the figures between brackets are approximate standard errors of the estimated parameters. 2)

The diagnostic checking criteria as described in section 5.6 provide no grounds for questioning the model adequacy (for details see Box and Jenkins [4]), and so scheme (5.7.8) and (5.7.9) can be used for forecasting.

Using (5.2.13) and (5.2.14) the best one-period ahead forecast is

\[(5.7.10) \quad \hat{x}_{n+1, n} = \sum_{j=1}^{\infty} p_j \hat{x}_{n+1-j, n} + \sum_{j=1}^{\infty} q_j \hat{z}_{n+1-j, n},\]

where

\[
\begin{align*}
p_1 &= 2.1, \quad p_2 = -1.5, \quad p_3 = 0.36, \quad p_{j \geq 4} = 0, \\
q_1 &= q_2 = 0, \quad q_3 = -0.53, \quad q_4 = 0.44, \quad q_5 = -0.28, \quad q_6 = 0.55, \quad q_7 = -0.32, \quad q_{j \geq 8} = 0.
\end{align*}
\]

The best forecast for \(x_{n+2}\) at point of time \(n\) is (see (5.2.15))

\[(5.7.12) \quad \hat{x}_{n+2, n} = p_1 \hat{x}_{n+1, n} + \sum_{j=2}^{\infty} p_j \hat{x}_{n+2-j, n} + \sum_{j=3}^{\infty} q_j \hat{z}_{n+2-j, n}.
\]

The best forecast for \(x_{n+1}\) (1 > 3) at point of time \(n\) is (see (5.2.15) and (5.2.16))

\[(5.7.13) \quad \hat{x}_{n+1, n} = p_1 \hat{x}_{n+1-1, n} + p_2 \hat{x}_{n+1-2, n} + p_3 \hat{x}_{n+1-3, n} +
\]

(2) Note that the residual variance of the univariate output model is 0.1081 and that of the transfer function-noise model is 0.056, so a reduction of approximately 50%.
where \( \hat{z}_{n+1,n} \) is the one-period ahead point forecast of the input can be determined via

\[
\tilde{z}_{n+1,n} = \sum_{j=1}^{\infty} \eta_j \tilde{z}_{n+1-j,n},
\]

and \( \hat{z}_{n+2,n} \) via

\[
\hat{z}_{n+2,n} = \eta_1 \tilde{z}_{n+1,n} + \sum_{j=2}^{\infty} \eta_j \tilde{z}_{n+1-j,n},
\]

and the \( \eta_j \) coefficients are:

\[
\eta_1 = -1.97, \quad \eta_2 = 1.37, \quad \eta_3 = -0.34, \quad \eta_{j>4} = 0.
\]

In our example the updating formula (5.2.36) is equal to

\[
(5.7.14) \quad \hat{z}_{n+1,n} = \tilde{z}_{n+1,n-1} + v_{1,A} \{ z_n - \tilde{z}_{n,n-1} \} + \psi_{1,A} \{ x_n - \tilde{z}_{n,n-1} \}.
\]

Box and Jenkins [4] have shown that for the gas furnace example the model with the leading indicator (transfer function-noise scheme) produced forecasts of considerable greater accuracy than those with the univariate scheme (5.7.2).
5.8. References


In de Reeks ter Discussie zijn verschenen:

<table>
<thead>
<tr>
<th>Nummer</th>
<th>Auteur</th>
<th>Titel</th>
<th>Datum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H.H. Tigelaar</td>
<td>Spectraalanalyse en stochastische lineaire differentievergelijkingen.</td>
<td>juni '75</td>
</tr>
<tr>
<td>2</td>
<td>J.P.C. Kleijnen</td>
<td>De rol van simulatie in de algemene econometrie.</td>
<td>juni '75</td>
</tr>
<tr>
<td>3</td>
<td>J.J. Kriens</td>
<td>A stratification procedure for typical auditing problems.</td>
<td>juni '75</td>
</tr>
<tr>
<td>4</td>
<td>L.R.J. Westermann</td>
<td>On bounds for Eigenvalues</td>
<td>juni '75</td>
</tr>
<tr>
<td>5</td>
<td>W. van Hulst</td>
<td>Investment/financial planning with endogenous lifetimes: a heuristic approach to mixed integer programming.</td>
<td>juli '75</td>
</tr>
<tr>
<td>6</td>
<td>M.H.C. Paardekooper</td>
<td>Distribution of errors among input and output variables.</td>
<td>augustus '75</td>
</tr>
<tr>
<td>7</td>
<td>J.P.C. Kleijnen</td>
<td>Design and analysis of simulation Practical statistical techniques.</td>
<td>augustus '75</td>
</tr>
<tr>
<td>8</td>
<td>J. Kriens</td>
<td>Accountantscontrole met behulp van steekproeven.</td>
<td>september '75</td>
</tr>
<tr>
<td>9</td>
<td>L.R.J. Westermann</td>
<td>A note on the regula falsi</td>
<td>september '75</td>
</tr>
<tr>
<td>10</td>
<td>B.C.J. van Velthoven</td>
<td>Analoge simulatie van ekonomische modellen.</td>
<td>november '75</td>
</tr>
<tr>
<td>11</td>
<td>J.P.C. Kleijnen</td>
<td>Het economisch nut van nauwkeurige informatie: simulatie van ondernemingsbeslissingen en informatie.</td>
<td>november '75</td>
</tr>
<tr>
<td>12</td>
<td>F.J. Vandamme</td>
<td>Theory change, incompatibility and non-deductibility.</td>
<td>december '75</td>
</tr>
<tr>
<td>13</td>
<td>A. van Schaik</td>
<td>De arbeidswaardeleer onderbouwd?</td>
<td>januari '76</td>
</tr>
<tr>
<td>14</td>
<td>J. van Lieshout, J. Ritzen, J. Roemen</td>
<td>Input-ouputanalyse en gelaagde planning.</td>
<td>februari '76</td>
</tr>
<tr>
<td>15</td>
<td>J.P.C. Kleijnen</td>
<td>Robustness of multiple ranking procedures: a Monte Carlo experiment illustrating design and analysis techniques.</td>
<td>februari '76</td>
</tr>
<tr>
<td>16</td>
<td>J.P.C. Kleijnen</td>
<td>Computers and operations research: a survey.</td>
<td>april '76</td>
</tr>
<tr>
<td>17</td>
<td>J.P.C. Kleijnen</td>
<td>Statistical problems in the simulation of computer systems.</td>
<td>mei '76</td>
</tr>
<tr>
<td>18</td>
<td>F.J. Vandamme</td>
<td>Towards a more natural deontic logic.</td>
<td>juni '76</td>
</tr>
<tr>
<td>19</td>
<td>J.P.C. Kleijnen</td>
<td>Design and analysis of simulation: practical, statistical techniques.</td>
<td>juli '76</td>
</tr>
<tr>
<td>20</td>
<td>H.H. Tigelaar</td>
<td>Identifiability in models with lagged variables.</td>
<td>augustus '76</td>
</tr>
<tr>
<td>21</td>
<td>J.P.C. Kleijnen</td>
<td>Quantile estimation in regenerative simulation: a case study.</td>
<td>augustus '76</td>
</tr>
<tr>
<td>22</td>
<td>W. Derks</td>
<td>Inleiding tot econometrische modellen van landen van de E.E.G.</td>
<td>september '76</td>
</tr>
<tr>
<td>23</td>
<td>B. Diederer, Th. Reijs, W. Derks</td>
<td>Principles of Economics for computers.</td>
<td>augustus '76</td>
</tr>
<tr>
<td>24</td>
<td>J.P.C. Kleijnen</td>
<td>Hybride simulatie van ekonomische modellen.</td>
<td>augustus '76</td>
</tr>
<tr>
<td>25</td>
<td>B. van Velthoven</td>
<td>Hybride simulatie van ekonomische modellen.</td>
<td>augustus '76</td>
</tr>
<tr>
<td>Author</td>
<td>Title</td>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>26. F. Cole</td>
<td>Forecasting by exponential smoothing, the Box and Jenkins procedure and spectral analysis. A simulation study.</td>
<td>september '76</td>
<td></td>
</tr>
<tr>
<td>27. J. Heuts</td>
<td>Some reformulations and extensions in the univariate Box-Jenkins time series analysis.</td>
<td>juli '76</td>
<td></td>
</tr>
<tr>
<td>28. W. Derks</td>
<td>Vier econometrische modellen.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>29. J. Frijns</td>
<td>Estimation methods for multivariate dynamic models.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>30. P. Meulendijks</td>
<td>Keynesiaanse theorieën van handelsliberalisatie.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>31. W. Derks</td>
<td>Structuuranalyse van econometrische modellen met behulp van Grapentheorie. Deel I: inleiding in de Grapentheorie.</td>
<td>september '76</td>
<td></td>
</tr>
<tr>
<td>32. W. Derks</td>
<td>Structuuranalyse van econometrische modellen met behulp van Grapentheorie. Deel II: Formule van Mason.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>33. A. van Schaik</td>
<td>Een direct verband tussen economische veroudering en bezettingsgraadverliesen.</td>
<td>september '76</td>
<td></td>
</tr>
<tr>
<td>34. W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van Grapentheorie. Deel III: De graaf van dynamische modellen met één vertraging.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>35. W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van Grapentheorie. Deel IV: Formule van Mason en dynamische modellen met één vertraging.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>36. J. Roemen</td>
<td>De ontwikkeling van de omvangsverdeling in de levensmiddelenindustrie in de D.D.R.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>37. W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van Grapentheorie. Deel V: De graaf van dynamische modellen met meerdere vertragingen.</td>
<td>oktober '76</td>
<td></td>
</tr>
<tr>
<td>38. A. van Schaik</td>
<td>Een direct verband tussen economische veroudering en bezettingsgraadverliesen. Deel II: gevoeligheidsanalyse.</td>
<td>december '76</td>
<td></td>
</tr>
<tr>
<td>39. W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van Grapentheorie. Deel VI. Model I van Klein, statisch.</td>
<td>december '76</td>
<td></td>
</tr>
<tr>
<td>40. J. Kleijnen</td>
<td>Information Economics: Inleiding en kritiek</td>
<td>november '76</td>
<td></td>
</tr>
<tr>
<td>41. M. v.d. Tillaart</td>
<td>De spectrale representatie van multivariate zwak-stationaire stochastische processen met discrete tijdparameter.</td>
<td>november '76</td>
<td></td>
</tr>
<tr>
<td>42. W. Groenendaal  Th. Dunnewijk</td>
<td>Een econometrisch model van Engeland</td>
<td>december '76</td>
<td></td>
</tr>
<tr>
<td>43. R. Heuts</td>
<td>Capital market models for portfolio selection.</td>
<td>september '76</td>
<td></td>
</tr>
</tbody>
</table>