IMSOR
THE INSTITUTE OF MATHEMATICAL STATISTICS
AND OPERATIONS RESEARCH

The Technical University of Denmark
DK-2800 Lyngby - Denmark

J. No. x611
1986-11-20
HS/jkl

M A R I M A

ESTIMATION OF MULTIVARIATE TIME SERIES MODELS

User guide for a computer program
by
HENRIK SPLIID
1986

RESEARCH REPORT
No. 17/86

ISSN 0107-3826

Trykt af imsor, DTH

Content

© Henrik Spliid, 1986

Introduction to MARIMA

This program uses a fast algorithm of "pseudo-regression" type to estimate time - discrete models that can be formulated as Multivariate mixed AutoRegressive Integrated Moving Average models, from where the name MARIMA stems. The algorithm is described in detail in (2). For model considerations see (3).

The program can also perform a quick identification procedure, that is based on the same principle as used in the backwards stepwise regression analysis, where you exclude the least significant parameters one by one until some significance criterion is met. In MARIMA this criterion is a heuristic F-test criterion. This procedure is described p.8 below.

The general MARIMA (k,p,d,q) model is formulated as follows:

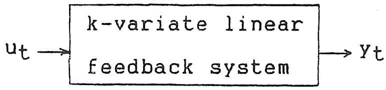
$$y_t = A_1 y_{t-1} + \dots\dots\dots + A_p y_{t-p}$$
$$+ u_t - B_1 u_{t-1} \dots\dots\dots - B_q u_{t-q}$$

where the k-dimensional timeseries y_t is obtained from some original series z_t by differencing using a pattern A^d, say, $y_t = A^d z_t$. u_t is the series of innovations. It is assumed, that u_t is a timediscrete white noise process.

In principle an arbitrary number of the coefficients in the matrices A_i and B_i can be specified equal to zero as described below.

In some situations seasonal variations are important. If this is the case, the user should consider using the routine SARMAX which enables the user to specify rather general seasonal models with regression variables.

The described model is suited to analyze the following common situation:



where one has observed the y_t process and is interested in estimating the parameters of the system and the variance-covariance matrix of the (unknown) noise process u_t. In the univariate case this model reduces to the well known Box-Jenkins ARIMA forecasting model, (2).

In the multivariate case the model is a generalization of the Box-Jenkins forecasting model, but as it is formulated here it also covers the common matrix transfer function model relating an input and an output. Such models are e.g. used in control theory. As a simple and useful example of this one could have the following model for two variables $y_{1,t}$ and $y_{2,t}$:

$$\begin{aligned} y_{1,t} = {} & a_{1,1}\, y_{1,t-1} + \dots + a_{1,p}\, y_{1,t-p} \\ & + a_{2,1}\, y_{2,t-1} + \dots + a_{2,r}\, y_{2,t-r} \\ & - b_{1,1}\, u_{1,t-1} - \dots - b_{1,q}\, u_{1,t-q} + u_{1,t} \end{aligned}$$

$y_{2,t}$ = some (perhaps uninteresting) model.

In this model $y_{1,t}$ is basicly an ordinary ARIMA process, but here it is also influenced by another process $y_{2,t}$. Such a model is often useful in econometric investigations and it is sometimes called a 'leading indicator' model - referring to the fact that the $y_{2,t}$ process can be leading in time and can actually be deciding the path of $y_{1,t}$ except for some noise $u_{1,t}$. It is easily seen that this model can be formulated as a MARIMA model with an adequate number of coefficients equal to zero. In a stochastic control problem the $y_{2,t}$ process could be an input signal while $y_{1,t}$ is output, so that the variation of

$y_{2,t}$ is not stochastic at all! In the example $y_{1,t}$ and $y_{2,t}$ are scalar processes, but they could as well be vector processes. In (3) a description of such modeling considerations is given.

Use of MARIMA

MARIMA is a subroutine and it is called from the users own program or from the routine SARMAX.

The data to be analyzed must be read into the calling program as a two dimensional vector. Suppose there are n observations and the time series has dimension k:

$y_{1,1}\ y_{2,1}\ y_{3,1} \cdots y_{k,1}$ (observation 1)
$y_{1,2}\ y_{2,2}\ y_{3,2} \cdots y_{k,2}$ (observation 2)
.
.
.
$y_{1,n}\ y_{2,n}\ y_{3,n} \cdots y_{k,n}$ (observation n)

The physical organization must be $y_{1,1}$, $y_{2,1}$, $y_{3,1} \cdots y_{k,1}$, $y_{1,2}$, $y_{2,2}$ etc. etc. until $y_{k-1,n}$ $y_{k,n}$. Or, in other words, observationwise.

MARIMA expects the array containing y to be of dimension (k,n) exactly. Thus the declaration of y must be of the form:

DIMENSION Y(k,mm) where mm $\geq$ n.

MARIMA itself requires ca. 200 bytes. A small user program as shown below adds around 35K plus storage for data and innovations to this. The total requirement is thus around 235K plus $8 \cdot n \cdot k$ bytes where n is number of observations and k is the dimension of the timeseries. If n=500 and k=2 then $8 \cdot 500 \cdot 2 = 8000 = 8K$ bytes is used for data and innovations, and a total of around 250K will be needed.

The time used by MARIMA is at large proportional to n and k^2 and depends on the size of the specified model. The autoregressive part is in general not very timeconsuming in contrast to the moving average part which involves repeated computation of residuals and furthermore is posing the greatest troubles with respect to convergence. Some examples of practical use are shown later in this guide. These were all run on an IBM 3033 at NEUCC. In the examples MARIMA is loaded from a library and it is automatically linked to the users calling program.

```
//A123456  JOB (KEY,ABC,t,1),'MARIMA PROGRAM',REGION=250K *)
// EXEC FORTG,LLIB=A119007.LIB
//CSYSIN DD *
      DIMENSION Y(k,n),U(k,n),IN((p+g)*k,k)
      INTEGER P,D,Q
C Y(k,n)  = array for data
C U(k,n)  = array for estimated innovations
C k,n     = dimension of timeseries, number of obsrvations
       DO 1 J=1,n
       READ(5,2)(Y(I,J),I = 1,k
    2 FORMAT(..some suitable format..)
    1 CONTINUE
C  NOW DATA HAVE BEEN READ INTO Y. THEY WERE ORGANIZED
C  WITH ONE OBSERVATION PR. CARD
      K = k     (dimension of timeseries)
      N = n     (number of observations)
      P = p     (order of autoregressive part of model)
      D = d     (number of differencing operations wanted)
      Q = q     (order of moving average part of model)
      M = 1     (the average is subtracted from data)
        = 0     (no subtraction of average)
     IT = it    (max. number of iterations)
      F = f     (a model reduction factor - see below)
     IA = 1     (a special model is defined in IN)
        = 0     (no special model defined in IN)
     IN =      (array of indicators - see below)

      CALL MARIMA (K,N,Y,U,P,D,Q,M,IT,F,IA,IN)
C NOW U CONTAINS INNOVATIONS ORGANIZED AS Y.
      STOP
      END
//GSYSIN DD *
Y(1,1) Y(2,1) Y(3,1) ... Y(k,1)
       more data
Y(1,n) Y(2,n) Y(3,n) ... Y(k,n)
/*
```

*) small letters are to be interpreted as numbers.

The indicator M

As mentioned above it is assumed that the innovation process u_t is white noise with constant Gaussian density and zero mean. The autocorrelation function is assumed to be zero for lags > 0. Among other things this causes the y_t process to have zero mean. Therefore it is recommended to subtract the average from the data before trying to estimate a MARIMA-model causing them to vary around zero. This can be done by the program automatically by putting M = 1. Otherwise use M = 0.

Maximum number of iterations allowed IT

The program uses an iterative procedure in which new parameters and innovations (residuals) are computed in each iteration. The procedure is stopped when the trace of the innovation covariance matrix does not change significantly in 4 consecutive iterations. Computational experience shows that models with many moving average terms need more iterations than models with few such terms.

For normal estimation $(-0.04 < F < 0.04)$ IT may be put equal to $10 + k(p + 2q)$ which in most cases will be more than enough. If $q = 0$ the analysis is done in one iteration by regression analysis.

If F-test identification is used $(F > 0.04)$ the program uses an additional number of iterations which at least will be 4, but it can be more. Often $20 + k(p + 2q)$ will be sufficient.

If a result is found before IT is reached the program ends by printing out this result and returning to the users calling program. If IT is reached the program prints out the results from the two last iterations and returns to the users program. At most 100 iterations are allowed in the program.

The identification factor F

In the program a partial F-test value for each parameter in the model is computed. For pure autoregressive models this F-value is the usual partial F-test value, but for models with moving average parameters it is only approximative and in general too small especially if the parameter actually is very significant ($F >$ ca. 4). For small F-values ($F <$ ca. 4) it is a pretty good approximation and thus it can be utilized to identify insignificant parameters. Specifying $F > 0.04$ causes the program to exclude all parameters from the estimated model having F-values smaller than F. An F between 1.0 and 4.0 is often a good choice.

The indicator vector IN

In the MARIMA (k,p,d,q) model each of the k variables have $(p+q)\cdot k$ parameters, so that the array IN must be declared DIMENSION IN(r,k), where $r = (p+q)\cdot k$, if it is to be used (IA = 1). If no special model is wanted IN(1) can be used.

The array IN must contain 1's or 0's (ones or zeroes). A 1 (one) indicating a parameter that can be estimated and a 0 (zero) indicating a parameter that must be kept equal to zero.

We illustrate the use of IN by an example:

Suppose we want to estimate the following MARIMA (2,2,0,1) model:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{11,1} & a_{12,1} \\ 0 & a_{22,1} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{11,2} & a_{12,2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix}$$

$$- \begin{bmatrix} b_{11,1} & 0 \\ 0 & b_{22,1} \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

then IN is defined as follows:

```
     DIMENSION IN (6,2)
     DATA IN/1,1,1,1,1,0,  0,1,0,0,0,1/
             1st variable  2nd variable
```

where it is easily seen, that the first (2+1)*2 = 6 values of IN indicate the y1-model and the last (2+1)*2 = 6 values indicate the y2-model.

In the example it is seen, that the defined model can be written as:

$$y_{1,t} = \text{ARIMA}(2,0,1) + a_1 y_{2,t-1} + a_2 y_{2,t-2}$$

and

$$y_{2,t} = \text{ARIMA}(1,0,1)$$

That is a leading indicator model for y_1 and a pure autoregressive moving average model for y_2.

<u>Output from MARIMA</u>

is printed on unit 6 (printer) and includes estimated model, and covariance matrix of innovations and also the approximate F-values.

When MARIMA returns to the user's calling program the input series is unaltered, and the U-array contains the corresponding estimated innovations (residuals).

All results are also stored in Common's, as described page 52 and page 35.

Estimation troubles

In some situations the computations do not lead to a stationary model. This can course the program to produce very great residuals. In the present version of the program a modification of the MARIMA algorithm suggested by Stoica et al. (5) is implemented. It improves the convergence of the algorithm, but problems may still occur.If large changes in the innovation covariance matrix do occur from one iteration to the following iteration a warning is printed out. If things go very wrong the program may run into an overflow and stop without further notice. In other situations the iterative procedure just does not converge and in this case the computations end at the last iteration allowed by printing out the results from the last two iterations. Troubles can also occur with over differenced data which generally can require a noninvertible model. The above mentioned problems can occur in models with one or more moving average terms but not in pure autoregressive models. The probable reason is that the data exhibit instationarities as e.g. means far away from zero or strong (not necessarily linear) trends. Such instationarities should be removed by the user before calling MARIMA.

In general, if many iterations are needed before convergence occurs the formulated model does not fit the data very well and most often the reason is either non-stationarity or non-invertibility.

An example

The following example illustrates the use of the very simple identification and estimation procedure in MARIMA.

The data are windmeasurements from a particular point in Greenland and the purpose of the analysis is to estimate a twodimensional timediscrete model to be used in a simulation study. Such a model can of course also be used to make short term windforecasts. The data include winddirection, windspeed. windprojection on westerly direction and on northerly direction. The last two variables will be analyzed. The total card setup and the results of the analysis are shown below. In the example a MARIMA (2,2,0,1) model is used. An F-factor of 4.0 is used for identification. A preliminary analysis of the residuals from the estimation is carried out by estimating a 6th order autoregressive model for the residuals, that is a MARIMA (2,6,0,0) model. From this analysis the calculated F-values are the usual partial F-values and as so they will be very good indicators to whether the model should be of higher order or not. The analysis shows that there is a slight lack of fit, but the coefficients are in all cases rather small (!) so that the identified MARIMA (2,2,0,1) for the data is quite satisfactory.

The result of the analysis is thus

y_1 = ARIMA(1,0,1) , $y_1 = z_1 - 2.27$

y_2 = ARIMA(2,0,0) , $y_2 = z_2 - 0.41$

and that y_1 and y_2 are almost independent. z_1 and z_2 are the original data.

It can be added that the use of F = 4.0 in the model identification results in a rather simple model with few parameters. If F = 2.0 was used in the example one would get a model in which y_2 is also dependent on the u_1 innovations.

```
//A123456  JOB (***,ABC,3,1)
// EXEC FORTG
//CSYSIN DD *
      DIMENSION Y(2,600),U(2,600),IN(6,2),IM(12,2)
      INTEGER P,D,Q
C
C  AT MOST 600 OBSERVATIONS OF Y ARE READ.
C  IF END-OF-FILE IS READ BEFORE 600 THEN NTOT WILL
C  CONTAIN THE TRUE NUMBER OF OBSERVATIONS
C
      DO 1 I=1,600
      READ(5,2,END=3)Y(1,I),Y(2,I)
    2 FORMAT(16X,2F8.3)
      NTOT=I
    1 CONTINUE
C
C  WE SPECIFY A MARIMA (2,2,0,1) WITH AVERAGE SUBTRACTION.
C
    3 K=2
      P=2
      D=0
      Q=1
      MEANS=1
C
C  A MODEL IDENTIFICATION WITH F=4.0 IS WANTED
C  AND WE ALLOW 40 INTERATIONS. NO SPECIAL MODEL SPECIFIED.
C
      F=4.0
      ITER=40
      IA=0
C
      CALL MARIMA(K,NTOT,Y,U,P,D,Q,MEANS,ITER,F,IA,IN)
```

```
C
C  WE NOW USE THE AUTOCORRELATION TEST ROUTINE ON THE
C  RESIDUALS, THAT IS THE U-SERIES :
C
      LAGS=5
      IPRINT=4
C
C  SKIP THE FIRST 5 RESIDUALS, SAY, AND START WITH U(1,6) :
C
      NOBS=NTOT-5
C
      CALL AUTEST(K,NOBS,U(1,6),K,LAGS,IPRINT,TQ,NDF)
C
C  WE NOW ANALYZE THE RESIDUALS BY ESTIMATING A SIXTH
C  ORDER AUTOREGRESSIVE MODEL FOR U. U WAS CALCULATED IN
C  MARIMA (Y IS NOW USED AS INNOVATION VECTOR AND THUS
C  IT WILL BE DESTROYED IN THIS EXAMPLE).
C
      CALL MARIMA(2,NTOT,U,Y,6,0,0,0,10,4.0,0,IM)
C
      STOP
      END
//GSYSIN DD *
   20.00     2.20     2.07     0.75
    0.00     1.80     1.80     0.00
    0.00     1.80     1.80     0.00
   20.00     2.70     2.54     0.92     ETC. ETC.
```

G E N E R A L   M A R I M A   M O D E L A N A L Y S I S.
A M U L T I V A R I A T E   A R I M A - P R O C E S S
A N D   T R A N S F E R - F U N C T I O N   P R O G R A M
BY HENRIK SPLIID, I M S O R,
THE INSTITUTE OF MATHEMATICAL STATISTICS AND OPERATIONS RESEARCH, TECHNICAL UNIVERSITY OF DENMARK, DK 2800 LYNGBY, DENMARK.

DIMENSIONS OF TIMESERIES	=	2	(K)
NUMBER OF OBSERVATIONS	=	502	(N)
- - LAGS IN AUTOREGRESSIVE PART OF MODEL	=	2	(P)
- - DIFFERENCING OPERATIONS	=	0	(D)
- - LAGS IN MOVING AVERAGE PART OF MODEL	=	1	(Q)
INDICATOR FOR MEANS CORRECTION OF DATA	=	1	(M)
MAXIMUM NUMBER OF ITERATIONS	=	40	

FORMULATION OF THE AUTOREGRESSIVE MOVING AVERAGE (A R M A) MODEL:
Y(T) = A(1)*Y(T-1)+...+A(P)*Y(T-P) + U(T)-B(1)*U(T-1)- ....B(Q)*U(T-Q), A(1),...,A(P) AND B(1),...,B(Q) BEEING THE UNKNOWN COEFFICIENT MATRICES, Y(T) THE K-DIMENSIONAL TIMESERIES AND U(T) THE (UNKNOWN) INNOVATIONS.

ARMA MODEL WILL BE ESTIMATED
FOR TIMESERIES D I F F E R E N C E D 0 TIME(S).

MODEL WILL BE ESTIMATED FOR TIMESERIES CORRECTED FOR
M E A N S .
2.26607 0.41079

AN F-TEST MODEL I D E N T I F I C A T I O N WAS SPECIFIED.
PARAMETERS WITH F-VALUES LOWER THAN 4.00 ARE EXCLUDED.

C O V A R I A N C E   M A T R I X FOR VARIABLES IN ANALYSIS.

3.89966D 01	-1.81924D 00
-1.81924D 00	5.94302D 00

INNOVATION COVARIANCE MATRIX TRACE IN FIRST ITERATION 3.69311D 00

G E N E R E L   M A R I M A   M O D E L   A N A L Y S I S .

RESULTS FROM ESTIMATION B E F O R E   T H E
M O D E L   I D E N T I F I C A T I O N .

NUMBER OF ITERATIONS USED = 6
MARIMA MODEL IS OF ORDER (K,P,D,Q) = (2, 2, 0, 1)

AUTOREGRESSIVE PART OF MODEL

A(1)=	0.9841	0.1832
	-0.1682	0.7042

A(2)=	0.0	-0.1280
	0.1527	0.2138

MOVING AVERAGE PART OF MODEL

B(1)=	0.2683	0.2595
	-0.2191	0.0911

I N N O V A T I O N   C O V A R I A N C E   M A T R I X

THE MATRIX IS BASED ONLY ON THE LAST (N - MAX(P,Q) - D) RESIDUALS. IT IS CALCULATED ASS SUMS OF SQUARES AND PRODUCTS DIVIDED BY (N - MAX(P,Q) - D - M) AND IT IS'NOT ADJUSTED FOR MODEL DEGREES OF FREEDOM OR AVERAGES.

2.3972D 00	-3.9544D-02
-3.9544D-02	1.2907D 00

AVERAGES FOR RESIDUALS FROM ESTIMATION WAS

-2.8141E-02   6.9030E-03

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

I N N O V A T I O N   C O R R E L A T I O N   M A T R I X
COMPUTED FROM THE ABOVE COVARIANCE MATRIX

1.0000	-0.0225
-0.0225	1.0000

M O D E L   S I G N I F I C A N C E   C O N T R O L
APPROXIMATE PARTIAL F-TEST-VALUES FOR COEFFICIENTS.
THE F-VALUES CAN BE COMPARED TO AN F(1, 493) DISTRIBUTION.

F-VALUES FOR AUTOREGRESSIVE PART OF MODEL

LAG= 1	6346.38	0.80
	4.30	23.73
LAG= 2	0.00	0.50
	3.83	2.81

F-VALUES FOR MOVING AVERAGE PART OF MODEL

LAG= 1	29.84	1.45
	6.22	0.36

G E N E R E L   M A R I M A   M O D E L   A N A L Y S I S.

RESULTS FROM F I N A L   E S T I M A T I O N
NUMBER OF UTERATIONS USED = 16
MARIMA MODEL IS OF ORDER (K,P,D,Q) = (2, 2, 0, 1)
AUTOREGRESSIVE PART OF MODEL

A(1)=	0.9809	0.0
	0.0	0.6253

A(2)=	0.0	0.0
	0.0	0.2824

MOVING AVERAGE PART OF MODEL

B(1)=	-0.2573	0.0
	0.0	0.0

I N N O V A T I O N   C O V A R I A N C E   M A T R I X

THE MATRIX IS BASED ON THE LAST (N - MAX(P,Q) - D) RESIDUALS. IT IS CALCULATED AS SUMS OF SQUARES AND PRODUCTS DIVIDED BY (N - MAX(P,Q) - D - M) AND IT IS NOT ADJUSTED FOR MODEL DEGREES OF FREEDOM OR AVERAGES.

2.4211D 00	-4.2107D-02
-4.2107D-02	1.3175D 00

AVERAGES FOR RESIDUALS FROM ESTIMATION WAS

-2.9634E-02   3.4258E-03

Result from heuristic F-test model identification procedure. The above model is the result of the heuristic F-test identification procedure.

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

I N N O V A T I O N   C O R R E L A T I O N   M A T R I X

COMPUTED FROM THE ABOVE COVARIANCE MATRIX

1.0000	-0.0236
-0.0236	1.0000

M O D E L   S I G N I F I C A N C E   C O N T R O L

APPROXIMATE PARTIAL F-TEST-VALUES FOR COEFFICIENTS.
THE F-VALUES CAN BE COMPARED TO AN F(1, 497) DISTRIBUTION.

F-VALUES FOR AUTOREGRESSIVE PART OF MODEL

LAG= 1	6475.58	0.45
	0.69	211.60
LAG= 2	0.06	1.45
	1.05	43.14

F-VALUES FOR MOVING AVERAGE PART OF MODEL

LAG= 1	27.67	1.13
	1.88	0.02

Heuristic F-TEST significance control of identified (reduced) MARIMA (2,2,0,1) model. It is noted, that only the non-zero parameter have large F-values

Output from call of AUTEST subroutine.

```
M U L T I V A R I A T E    A U T O C O R R E L A T I O N
T E S T .
            NUMBER OF VARIABELS IN INPUT VECTOR    =     2
            LENGHT OF TIME SERIES                  =   497
            NUMBER OF VARIABLES TO BE TESTED       =     2
            NUMBER OF LAGS INCORPORATED IN TEST    =     5


ORDINARY CORRELATION MATRIX (LAG ZERO)

           1.0000  -0.0258
          -0.0258   1.0000

STANDARD DEVIATIONS OF VARIABLES IN TEST

           1.5552   1.1493

INVERSE CORRELATION MATRIX

           1.0007   0.0258
           0.0258   1.0007

ORGANISATION OF CROSS-AUTO-CORRELATION MATRICES

        R(I,J) = COR( A(I,T) , A(J,T-LAG) ).

LAG =  1  CROSS-AUTO-CORRELATION-MATRIX

          -0.0003  -0.0495
           0.0611   0.0233
              CONTRIBUTION TO CHI-SQUARE-TEST =       3.08
```

```
LAG =  2   CROSS-AUTO-CORRELATION-MATRIX

           0.0315   -0.0126
           0.0981    0.0233
              CONTRIBUTION TO CHI-SQUARE-TEST =      5.75

LAG =  3   CROSS-AUTO-CORRELATION-MATRIX

           0.0742    0.0193
           0.0416    0.0518
              CONTRIBUTION TO CHI-SQUARE-TEST =      5.32

LAG =  4   CROSS-AUTO-CORRELATION-MATRIX
           0.1165   -0.0345
          -0.0155   -0.0115
               CONTRIBUTION TO CHI-SQUARE-TEST =     7.39

LAG =  5   CROSS-AUTO-CORRELATION-MATRIX
           0.0687   -0.0749
          -0.0949   -0.0250
               CONTRIBUTION TO CHI-SQUARE-TEST =     9.75
                   TOTAL CHI-SQUARE-TEST VALUE =    31.28
```

NOMINAL DEGREES OF FREEDOM IS 20 WHICH SHOULD BE CORRECTED: SUBTRACT THE NUMBER OF PARAMETERS LINKING THE VARIABLES CONSIDERED IN THE TEST.
FOR RESIDUALS FROM A FULL MARIMA (K,P,D,Q) MODEL SUBTRACT K*K*(P+Q); IF THE MODEL IS REDUCED SUBTRACT THE NUMBER OF NON-ZERO PARAMETERS.

The estimated model has 4 non-zero parameters. Hence the degrees of freedom for the test value is 20-4 = 16.
The analysis of the autocorrelations of the residuals shows some lack of fit!

G E N E R E L   M A R I M A   M O D E L   A N A L Y S I S

A   M U L T I V A R I A T E   A R I M A - P R O C E S S
A N D   T R A N S F E R - F U N C T I O N   P R O G R A M
B Y   H E N R I K   S P L I I D ,   I M S O R,
THE INSTITUTE OF MATHEMATICAL STATISTICS AND OPERATIONS RESEARCH, TECHNICAL UNIVERSITY OF DENMARK, DK 2800 LYNGBY, DENMARK.

```
DIMENSIONS OF TIMESERIES                                 =   2  (K)
NUMBER OF OBSERVATIONS                                   = 502  (N)
-       -  LAGS IN AUTOREGRESSIVE PART OF MODEL          =   6  (P)
-       -  DIFFERENCING OPERATIONS                       =   0  (D)
-       -  LAGS IN MOVING AVERAGE PART OF MODEL          =   0  (Q)
INDICATOR FOR MEANS CORRECTION AF DATA                   =   0  (M)
MAXIMUM NUMBER OF ITERATIONS                             =  10
```

FORMULATION OF THE AUTOREGRESSIVE MOVING AVERAGE (A R M A) MODEL:y(T) = a(1)*Y(-1)+...+A(P)*Y(T-P) + U(T-1)-...-B(Q)*U(T - Q)A(1),...,A(P) AND B(1),...,B(Q) BEEING THE UNKNOWN COEFFICIENT MATRICES, Y(T) THE K-DIMENSIONAL TIMESERIES AND U(T) THE (UNKNOWN) INNOVATIONS.

ARMA MODEL WILL BE ESTIMATED
FOR TIMESERIES D I F F E R E N C E D 0 TIME(S)

A PURE AUTOREGRESSIVE MODEL WAS SPECIFIED.
IT WILL BE ESTIMATED DIRECTLY.

AN F-TEST MODEL I D E N T I F I C A T I O N WAS SPECIFIED.
PARAMETERS WITH F-VALUES LOWER THAN 4.00 ARE EXCLUDED.

C O V A R I A N C E   M A T R I X FOR VARIABLES IN ANALYSIS.
NO ADJUSTMENT FOR AVARAGES HAS BEEN MADE.

2.40670D 00	-4.17935D-02
-4.17935D-02	1.30970D 00

INNOVATION COVARIANCE MATRIX TRACE IN FIRST ITERATION 3.68468D 00

G E N E R E L   M A R I M A   M O D E L   A N A L Y S I S.

RESULTS FROM F I N A L   E S T I M A T I O N
NUMBER OF ITERATIONS USED = 1
MARIMA MODEL IS OF ORDER (K,P,D,Q) = (2, 6, 0, 0)

AUTOREGRESSIVE PART OF MODEL

A(1)=	0.0	0.0
	0.0	0.0
A(2)	0.0	0.0
	0.0773	0.0
A(3)	0.0	0.0
	0.0	0.0
A(4)	0.1137	0.0
	0.0	0.0
A(5)	0.0	0.0
	-0.0757	0.0
A(6)	0.0	0.0
	0.0	0.0

I N N O V A T I O N   C O V A R I A N C E   M A T R I X

THE MATRIX IS BASED ONLY ON THE LAST (N - MAX(P,Q) - D) RESIDUALS. IT IS CALCULATED AS SUMS OF SQUARES AND PRODUCTS DIVIDED BY (N - MAX(P,Q) -D - M) AND IT IS NOT ADJUSTED FOR MODEL DEGREES OF FREEDOM OR AVERAGES

2.3884D 00	-3.5646D-02
-3.5646D-02	1.2963D 00

AVERAGES FOR RESIDUALS FROM ESTIMATION WAS

-2.1002E-02	3.4942E-03

Result of estimation of 6th order autoregressive model for the residuals from the MARIMA(2,2,0,1) model for the data.

The analysis is performed by a stepwise linear regression procedure. A combined forward selection and backwards elimination algorithm is used.

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

I N N O V A T I O N   C O R R E L A T I O N   M A T R I X
COMPUTED FROM THE ABOVE COVARIANCE MATRIX

1.0000	-0.0203
-0.0203	1.0000

M O D E L   S I G N I F I C A N C E   C O N T R O L
APPROXIMATE PARTIAL F-TEST-VALUES FOR COEFFICIENTS.
THE F-VALUES CAN BE COMPARED TO AN F(1, 500) DISTRIBUTION.

F-VALUES FOR AUTOREGRESSIVE PART OF MODEL

LAG= 1	0.04	1.55
	2.76	0.02
LAG= 2	0.35	0.30
	5.57	0.43
LAG= 3	3.20	0.12
	0.88	2.29
LAG= 4	6.56	0.59
	0.12	0.00
LAG= 5	2.28	2.42
	5.33	0.48
LAG= 6	1.46	1.62
	0.71	1.73

Significance control for 6th order autoregressive model for the residuals.

REFERENCES

(1) Box, G.E.P. and Jenkins, G.M., "Time Series Analysis, Forecasting and Control", Holden-Day, San Francisco 1970.

(2) Spliid, H., " A Fast Estimation Method for the Vector Autoregressive Moving Average Model with Exogeneous variables", Journal of the American Statistical Association, 1983, Vol. 78, pp. 843-849.

(3) Tiao, G.C. and Box, G.E.P., "Modeling Multiple Time Series", Journal of the American Statistical Association, 1981, Vol. 76, pp. 802-816.

(4) Priestley, M.B., "Spectral Analysis of time Series", Vol.1 and 2, Academic Press, London 1981.

(5) Stoica, Söderström, Ahlen and Solbrand, "On the Convergence of Pseudo-Liniar Regression Algorithms", International Journal of Control, 1985, vol. 41, pp. 1429-1444.

APPENDIX

Additional routines

This appendix describes some more advanced additional routines that may be utilized by the user of MARIMA.

Generally the data organization in these routines is compatible with that of MARIMA such that they can be used before or after MARIMA without much further programming in the same job.

Standardization of variables.

Often it can be beneficial to scale the variables in the input series to MARIMA. This may be accomplished by the subroutine STANDZ as follows:

```
CALL STANDZ (K,N,Y,IPRINT,AVER,STDV)
```

where Y is the K-dimensional series of length N, dimensioned Y(K,N) (the same organization as in the call to MARIMA).

```
IPRINT = 0  : no printed output from STANDTZ
IPRINT = 1  : averages and standard deviations
              are printed
AVER    = vector of length K : contains averages
          at return
STDV    = vector of length K : contains standard
          devitations at return
```

The vector Y is transformed to (Y(i,j)-AVER(i))/STDV(i), i=1,K, j=1,N.

When calling MARIMA use the MEANS parameter equal to 1, then the means correction is taken into account, although the computations as such are superfluous.

Residual autocorrelation test

After returning from MARIMA (or perhaps before calling) it can be of interest to analyze the residuals (or perhaps the data series itself) for whiteness. This is done by means of the autocorrelation function.

The subroutine is called in the following way:

```
CALL AUTEST(K,N,A,KTEST,LAGS,IPRINT,TQ,NDF)
```

Where A is the series of dimension (K,N) to be checked. If only the first KTEST of the K variables in A are to be checked then give this number in KTEST.

If all variables in A are to be included in the test use KTEST=K.

LAGS is the number of lags to be checked, LAGS>0.

IPRINT controls print : 0=no print, 1=minimal print i.e. only the result TQ and NDF, 2=like 1 plus correlation matrix and standard deviations of variables, 3=like 2 plus contributions to test statistic up to til lag no. LAGS.

TQ is Chi-square value computed and NDF is the uncorrected degrees of freedom.

The test is according to : Li & McLeod : J. Roy. Stat. Soc., Ser.B, Vol. 43, 231-239.

Estimation and plotting of bivariate statistics

For identification and model control a subroutine is called in the following format

```
        CALL TWOSPE
      * (K,N,Y,Y,I1,I2,LAGCOR,LAGWIN,IWIN,LINES,J1,J2,J3,J4)
```

Y(K,N) is an input K-dimensional time series of length N.
Note that Y occurs twice in the list of calling arguments.

I1 = No. of first variable ($1 \leq I1 \leq K$).

I2 = - - second -- ($1 \leq I2 \leq K$).

LAGCOR = No. of lags in crosscorrelation function plot for the two variables I1 and I2. (0 = no plot).

LAGWIN = length (half length of twosided) weight window for smoothing cross covariance function in estimating the cross spectral density between the two variables I1 and I2. The same window is used for the univariate spectra.

IWIN = type of smoothing window:

1. : Bartlett window of length LAGWIN
2. : Tukey - - - -
3. : Parzen - - - -

LINES = No. of spectral lines estimated in spectral plot in linear frequency scale. If a negative value of LINES is used a geometric frequency scale is used in the plot.

J1 and J2 = scaling parameters for spectral plot for first variable. If J1=J2 an automatic scaling is used. If J1<J2 the scale of the plot will range from 10^{J1} to 10^{J2} (10-logarithmic scale used).

J3 and J4 = scaling parameters for spectral plot for second variable.
If J1 = J2 = J3 = J4 is used the same scaling is used for both spectral plots and it is found by TWOSPE.

Output from the routine includes bivariate crosscorrelation function, univariate spectra for the two variables and coherence and phase spectrum between the two variables. (see Priestley (1981) pp 654-681.)

The routine can be usefull as an identification tool in detection of lead-lag relations between the variables in the series Y. If the routine is used on the residuals it can detect lack of fit.

Estimation and plotting of bivariate statistics for series versus estimated values or residuals.

Suppose the following model for the series Y(K,N) has been obtained by MARIMA (CALL MARIMA(K,N,Y,A, ...other specifications):

$$Y_t = \text{MARIMA} + A_t$$, where also A is dimensioned A(K,N)

Then one can use :

```
  CALL MULSPE
(K,N,Y,A,NVAR,LAGCOR,LAGWIN,LINES,IWIN,J1,J2,J3,J4)
```

For the NVAR'th variable in Y(K,N), that is Y(NVAR,1)....Y(NVAR,N), the estimated values Y(NVAR,1)-A(NVAR,1)....Y(NVAR,N)-A(NVAR,N) are generated. The Y-series and the estimated Y-series are treated like in the above described call of TWOSPE. At return Y and A are unaltered.

Alternatively one can use :

```
   CALL CROSPE
(K,N,Y,A,NVAR,LAGCOR,LAGWIN,IWIN,LINES,J1,J2,J3,J4)
```

For the NVAR'th variable in Y(K,N) the Y-series and the corresponding A-series in A(K,N) are treated in the above call of the routine TWOSPE. This routine may also be used to analyze two univariate series Y(N) and A(N), say, in which case K=1 and where A is not necessarily the residuals of Y. In this situation NVAR=1 also.

Computation of residuals or forecasts using a MARIMA model

One-step-ahead residuals for another series than used in the estimation can be found by the routine FOREC3.

```
      CALL FOREC3 (K,NNEW,YNEW,ANEW)
```

where YNEW(K,NNEW) is the new series and A(K,NNEW) is the residuals for the YNEW series to be calculated.

The routine FOREC3 uses the content of three labelled commons (see next section)

```
      COMMON / RESULT / RDATA(710),DDATA(100)
      DOUBLE PRECISION DDATA
      COMMON / SPEC / IDATA (7)
      COMMON / CLAGS / JDATA (77)
```

as they have been created by MARIMA. Thus if FOREC3 is used in an other job where the MARIMA-analysis is not performed these commons must be created separately before calling FOREC3. This can be done by saving them in the job where MARIMA created them and retrieving them before FOREC3 is called.

The following illustrates the method:

Job 1 : Declare the COMMONS RESULT, SPEC and CLAGS.
CALL MARIMA (K,N,Y,A,...)
save RESULT, SPEC and CLAGS at some medium.
etc.

Job 2 : Declare the COMMONS RESULT, SPEC and CLAGS.
Read RESULT, SPEC and CLAGS from the medium.
CALL FOREC3 (K,NNEW,YNEW,ANEW)
etc.

Retrieval of MARIMA residuals for other purposes

When a MARIMA job is finished all model definitions and results are stored in labelled COMMONS:

```
      COMMON/RESULT/AVERA(10),COEF(70,10),UCOV(10,10)
      DOUBLE PRECISION UCOV
      COMMON/SPEC/INI,LAGY,NROT,KVARY,LAGU,MEANX,NDIF
      COMMON/CLAGS/MMV,MMLAG(3,10),MAXLAG,MMQ,MPA,
     *MMDIF(4,10),MAXSEA,KKA,IGET
```

or for short

```
      COMMON/RESULT/RDATA(710),DDATA(100)
      DOUBLE PRECISION DDATA
      COMMON/SPEC/IDATA(7)
      COMMON/CLAGS/JDATA(77)
```

that is 710 REAL*4's, 100 REAL*8's, 7 INTEGER's and 77 INTEGER's.

The content of these commons is described in the section describing commons in general.

If the routine FOREC3 , described above, is to be used the common RESULT, SPEC and CLAGS must be correctly remembered from the MARIMA analysis in which they were created.

If the routine DIFEQU is to be used these commons are also needed.

Differencing Y(K,N) one variable at a time

Before calling MARIMA it may necessary to difference one of the variables in the Y(K,N) series. This differencing may be ordinary or seasonally and can be carried out by the routine UNIDIF for one variable in Y at a time. General differencing of a multivariate series is described below (next routine).

If a seasonal model is wanted the user should consider the routine SARMAX described below.

Suppose an s-period seasonal differencing is wanted, $s \geq 1$. Then the following transformation is made

$$z_i = y_i - y_{i-s} \qquad \text{if } i-s > 1$$
$$z_i = y_i \qquad \text{if } i-s < 0$$

where y_i is the i'th observation and z_i is the new differenced i'th observation by which y_i is replaced. Note that the first s values of the series is left unchanged (alternatively one may say that they are differenced using leading zeroes).

If more than one variable in Y(K,N) has to be differenced UNIDIF may be used for each variable, but in this case it may be more convenient to use GENDIF (multiple mixed ordinary and seasonally differencing), which is a general routine. Also if the routine DIFEQU is to be used later for calculation of forecasts in integrated form it is easier to use GENDIF, since the calling format for this routine is useable when DIFEQU is called.

UNIDIF is called as follows:

```
      CALL UNIDIF(K,N,Y,KD,NDIF,ISUM)
```

where

Y = time series dimensioned Y(K,N)

KD = no. variable in Y(K,N), $1 \leq KD \leq K$.
NDIF = length of differencing period (season), $NDIF \geq 1$
ISUM = -1 the time series is differenced
+1 the time series is integrated (reverse of -1).

Example : CALL UNIDIF(K,N,Y,3,12,-1) will cause the 3rd variable in Y(K,N) to be seasonally differenced with s=12. The 12 first values of the variable are left unchanged.

If MARIMA is called after the differencing of the series, the first s values should in general be left out of the analysis. The easiest way to accomplish this is to make MARIMA start at observation 1+s.

Example: Suppose the 3rd variable is differenced with s=12 and MARIMA is called afterwards:

```
CALL UNIDIF(K,N,Y,3,12,-1)
CALL MARIMA(K,N-12,Y(1,1+12), A(1,1+12),... )
```

such that MARIMA starts at observation 1+12=13 and only N-12 observations are analyzed.

Note the parallel specification for the series of residuals, A(K,N). It has the effect that the A's and the Y's are aligned all the time, also if the original series is recreated by using, say,

```
CALL UNIDIF(K,N,Y,3,12,+1).
```

The first 12 values of A will be undefined unless the user has initialized them before the call to MARIMA while the remaining N-12 values will be found by MARIMA.

Mixed ordinary and seasonal differencing using a pattern.

Before using the routine of this section the user should be familiar with the routine UNIDIF described previously.

Note, that if seasonal models are analyzed it may be advantageous to use the more general routine SARMAX described below.

If general differencing of the variables in the time series Y(K,N) is wanted one may use:

```
CALL GENDIF(K,N,Y,MDIF,NPAT,ISUMS,MAXDS)
```

where

Y = time series dimensioned Y(K,N)
MDIF = integer array dimensioned MDIF(K,NPAT). MDIF contains the (seasonal) patterns to be used.
NPAT = no. of (seasonal) patterns to be used.
ISUMS = -1 seasonal differencing is used
= 1 summing is used (reverse effect of -1)
MAXDS = no. of observations lost as a result of the differencing. Output value found by GENDIF.

Example : If the series Y contains 3 variables which have to be seasonally differenced with seasonalities 0, 12 and 6 the array MDIF(3,1)=(0,12,6) and NPAT=1. MAXDS=12 will be returned.

Example : If the series Y contains 3 variables which have to be differenced as follows :

Variable	Y_1	differenced	2 times	with	seasons	12 and 3.
-	Y_2	-	1 time	-	season	6.
-	Y_3	-	2 times	-	seasons	1 and 1.

then MDIF(3,2)=((12,6,1),(3,0,1)) such that the two seasonal patterns (12,6,1) and (3,0,1) are used. NPAT=2. Again the sequence of differencing is unimportant. In the example one might use MDIF(3,2)=((3,6,1),(12,0,1)) to give the same result. In the example MAXDS=15 will be found by GENDIF.

The routine may be used instead of UNIDIF or together with it . There are no restrictions on the calling sequence. All routines are fully reversible. It is recommended to use GENDIF alone since it is a more general routine.

Computation of forecasts and MARIMA model in aggregated forms.

After calling MARIMA a subroutine DIFEQU can be called. It is able to compute K-step predictions and confidence intervals for predictions in differenced as well as non-differenced forms.It also prints out seasonal models in integrated form (useful if SARMAX was used). The calling of the routine is as follows:

```
      CALL DIFEQU(K,NS,Y,A,KA,IGET,YFOR,NFY,IPRY,VFOR,
     +            NFV,IPRV,MDIF,NPAT,LPR,INIT,PROB)
```

where

K = dimension of Y and A, both are dimensioned (K,N). Note that N $\geq$ NS

NS = starting point for forecasts. Forecasts will be for observations NS+1, .., NS+NFY.

Y = series dimensioned Y(K,N), where N $\geq$ NS

A = series of residuals A(K,N), where N $\geq$ NS (as for Y) which must be computed up till at least NS (for example by MARIMA or by the routine FOREC3).

KA = No. of variables among the K variables which are random. If some of the variables among the K variables are exogeneous variables and deterministic they are placed in the Y vector on the last places. For these K-KA variables no model is estimated and they serve as regression variables for the first KA variables, If KA=K all K variables are treated as variables which may be forecasted.

IGET = indicator. IGET=1 : the deterministic variables for the forecasting are taken from the Y series. If IGET=0 : then the det.var.'s are placed in the YFOR vector by the user before calling DIFEQU.
Note: If NS=N then IGET=0 must be used.

YFOR = Vector of dimension YFOR(K,NFY) for forecasts. The last K-KA places are taken from Y (IGET=1) or defined by the user (IGET=0) (before calling for all forecasting points from 1 to NFY. (output also)).

NFY = length of forecast (starting from observation NS)

IPRY = output indicator. 0: no printing of forecasts. 1 : only printing of forecasts. 2 : as for 1 plus title. 3 : as for 2 plus confidence bands for all non-detetministic variables

VFOR = vector of dimension VFOR(K,K,NFV) for covariance matrices for forecasts up till NFV steps ahead. (output also).

NFV = length of forecast for covariances.

IPRV = output indicator for VFOR. 0 : no printed output. 1 : print of covariance matrices. 2 : print of standard deviations for forecasts. 3 : as for 1 plus 2.

MDIF = array containing differencing pattern as used by GENDIF. If the Y series is given to DIFEQU in differenced form MDIF should not used. (that is NPAT=0).

NPAT = is number of patterns each of length K in MDIF. NPAT=0 indicates that MDIF is empty or not to be used. Thus the array MDIFS(K,NPAT) is expected by DIFEQU.

LPR = indicator. LPR = 0 : no printed output of model used. LPR = 1 : the MARIMA model in difference equation form is printed out. If differencing is specified via MDIF the model in aggregated form is printed.

INIT = indicator. INIT = 1 must be used when DIFEQU is called after MARIMA. This will generate the VFOR vector and other results which may be re-used. INIT = 0 can be used in subsequent calls which will save many unnecessary recomputations. NOTE that in the first call to DIFEQU NFV=NFY=(max. length of forecast) should be used.

PROB = probability for confidence bands for forecasts. If PROB=0.95, for example, 2.5% and 97.5% limits are found.

Example : 5 variables in a series consisting of 110 observations.

```
   DIMENSION Y(5,110),A(5,110),MDIFS(5,2),YFOR(5,10),
  +          VFOR(5,5,10),IN(..)
   DATA MDIF/1,1,0,0,0, 0,1,0,0,0/
   DO 10 I=1,110
10 READ(5,11)(Y(J,I),J=1,5)
11 FORMAT(..)
```

Now data are read into Y. N=110 observations.
We now proceed to difference the series according to MDIFS using GENDIF:

```
   N=110
   K=5
   NPAT=2
   CALL GENDIF (K,N,Y,MDIF,NPAT,-1,MAXDS)
```

MAXDS=2 will be found by GENDIF. We now call MARIMA for estimation of a model. The first MAXDS observations are left out :

```
   CALL MARIMA(K,N-MAXDS,Y(1,1+MAXDS),A(1,1+MAXDS),....,IN)
```

we now recreate the original series by using:

```
   CALL GENDIF(K,N,Y,MDIF,NPAT,+1,MAXDS)
```

Finally a, say, 10-step ahead forecast is wanted, that is a forecast starting at observation 110. 90% confidence limits are wanted. We assume that KA=5 i.e. all 5 variables in Y are non-deterministic. Generally this will be the case if the array IN(..) is not used in the call to MARIMA.
The various indicators may be written directly in the CALL statement or they may be given via, for example

```
    DATA KA,IGET,NFY,IPRY,NFV,IPRV,LPR,INIT,PROB
   *     / 5,  0, 10,   3, 10,   3,  1,  1, 0.90/
    NSTART=110
    CALL DIFEQU(K,NSTART,Y,A,KA,IGET,YFOR,NFY,IPRY,VFOR,NFV,
   *            IPRV,MDIF,NPAT,LPR,INIT,PROB)
```

INIT=0

In subsequent calls INIT=0 can be used. If for example a forecast for the last 10 observations of the series is wanted one may now use:

```
     CALL DIFEQU(5,100,Y,A,5,0,YFOR,10,3,VFOR,10,0,
    *            MDIF,2,0,INIT,0.90)
```

giving a 10 step ahead forecast starting at observation 100 with 90% confidence bands and without repetition of the forecast covariance matrices and the aggregated MARIMA model.

If in the above example the user is interested in forecasting the data in differenced form the calling of DIFEQU should be with NPAT=0 and the recreation of the series via the second call to GENDIF should be left out. The calling sequence would then be :

```
         CALL GENDIF(K,N,Y,MDIFS,NPAT,-1,MAXDS)
         CALL MARIMA(K,N-MAXDS,Y(1,1+MAXDS),A(1,1+MAXDS),
        *            ... ,IN)
         CALL DIFEQU(5,110,Y,A,5,0,YFOR,10,3,VFOR,10,3,
        *            MDIF,0,1,1,0.90)
```

which will produce the required forecast for the differenced series.

Programming note:

DIFEQU makes use of all four common's RESULT, SPEC CLAGS and EXTRA. The three first ones are produced in MARIMA and SARMAX, but can of course be remembered from some previous call to MARIMA, as described elsewhere (in FOREC3 and the separate COMMON description).

The COMMON/EXTRA/ is a large work space used by MARIMA also. DIFEQU uses it for manipulation with the MARIMA model. In the first (K*K)(P+D+Q+2) places in EXTRA the aggregated MARIMA model is put in blocks of (K*K) by DIFEQU. Both the AR and the MA part have leading -1's.

The COMMON/RESULT/ is used by DIFEQU and it must be at hand when DIFEQU is called.

The common CLAGS is used by the K-step-ahead forecasting routine KSTEP. It must also be at hand when DIFEQU is called.

If, for some reason, MARIMA is called between two calls to DIFEQU, the initiation indicator INIT=1 must be used again. In the original version of DIFEQU the COMMON /EXTRA/ contains 7385 places. This number also limits the maximum forecasting length for the VFOR (covariances). If the length of EXTRA is shorter than 7385 the calculation of the variable 'MAXF' in DIFEQU must be changed accordingly.

The recursive one step ahead forecasts can be computed by using DIFEQU repeatedly from the start of the time series and finding the residuals and filling them into the A-series as the one step ahead forecasting progresses. This method is used by FOREC3.

Seasonal models using subroutine SARMAX.

MARIMA is able to analyse seasonal models via a calling routine SARMAX. Additive seasonal models and models including differenced variables can be treated quite easily.

Seasonal lagging using SARMAX.

Consider the following univariate seasonal model

$$y_t - a_1 y_{t-1} - a_{12} y_{t-12} = u_t - b_1 u_{t-1} - b_{12} u_{t-12}$$

where the 12-period seasonality is modelled via the two additive terms y_{t-12} and u_{t-12}.

Introducing the usual Box-Jenkins notation we may write:

$$a(B)y_t = b(B)u_t$$

where $a(\cdot)$ and $b(\cdot)$ are the two operatorpolynomials:

$$a(B)= 1 - a_1 B - a_{12} B^{12}$$
$$b(B)= 1 - b_1 B - b_{12} B^{12}$$

In SARMAX an artificial variable can be generated which makes it possible to analyse the above model as an ordinary ARMA-model. This can be done by introducing a new lagged variable as for example:

$$z_t = y_{t-11} \quad \text{with residual} \quad v_t = u_{t-11}$$

Using this new variabel we may write the original model for y_t as follows:

$$y_t - \{a_1 \; a_{12}\} \begin{Bmatrix} y_{t-1} \\ z_{t-1} \end{Bmatrix} = u_t - \{b_1 \; b_{12}\} \begin{Bmatrix} u_{t-1} \\ v_{t-1} \end{Bmatrix}$$

where as described

$z_t = y_{t-11}$ with residual $v_t = u_{t-11}$.

When MARIMA is called via SARMAX the series z_t can be generated automatically and the seasonally lagged residuals are taken into account in the proper way.

The definition of z_t is accomplished by an array MLAG(3,NV) where NV is 1 in this example

The new variable z_t must be No.2. The old variable y_t is No.1. The lagging period is 11 (and not 12). This is defined in MLAG by using

```
MLAG(1,1) = 2    (No. of new variable)
MLAG(2,1) = 1    (No. of old variable)
MLAG(3,1) = 11   (No. of lags = season -1)
```

The array MLAG is given as input to SARMAX and the original time series is treated as a two-variate time series for which the user must provide the necessary space when SARMAX is called. It is required that MLAG(1,1) > MLAG(2,1) and MLAG(3,1) $\geq$ 1.

The above description refers to a univariate series. If the original series is multivariate and/or several seasonally lagged new artificial series are needed this is done in the same way as already described. If for instance both a 12 hour seasonality and a 24 hour seasonality are needed one may use:

```
DIMENSION MLAG (3,2)
DATA MLAG / 2,1,11,3,1,23 /
```

which defines two new series called for example

$z_t = y_{t-11}$ with residual $v_t = u_{t-11}$

$x_t = y_{t-23}$ with residual $w_t = u_{t-23}$

If the original series y_t is two-dimensional

$$y_t = (y_{1,t}\ ,\ y_{2,t})^T$$

and a 12 period season is wanted for both variables two new variables $y_{3,t}$ and $y_{4,t}$ are defined by MLAG. An example is:

```
DIMENSION MLAG (3,2)
DATA MLAG /3,1,11,4,2,11/
```

which then defines:

$$y_{3,t} = y_{1,t-11} \quad , \quad u_{3,t} = u_{1,t-11}$$
$$y_{4,t} = y_{2,t-11} \quad , \quad u_{4,t} = u_{2,t-11}$$

thus the original two-variate problem is transformed into a four-variate MARIMA problem.

<u>NOTE</u> that the two new variables $y_{3,t}$ and $y_{4,t}$ must be no.3 and no.4, that is the last variables in the time series. For instance it would be illegal to read $y_{1,t}$ and $y_{3,t}$ into the users calling program and generate $y_{2,t}$ by MLAG from $y_{3,t}$. But it is legal to read $y_{1,t}$ and $y_{3,t}$ and then generate $y_{2,t}$ from $y_{1,t}$ and $y_{4,t}$ from $y_{3,t}$.

<u>Seasonal differencing via SARMAX.</u>

In the describtion of the routine GENDIF a general method of ordinary or seasonal differencing a K-variate time series is presented.

The differencing is defined by a pattern MDIF(K,NP) where K is the dimension of the time series and NP is the No. of differencing patterns to be used.

Suppose we have a two-variate time series

$$y_t = (y_{1,t} \; , \; y_{2,t})^T$$

where it is wanted to use ordinary first differencing for $y_{1,t}$

and 12 period differencing for both variables afterwards, then we can use

```
DIMENSION MDIF (2,2)
DATA MDIF /1,0,12,12/
```

defining two patterns, namely (1,0) and (12,12) to be applied on the time series.

When calling SARMAX the array MDIF is used as calling argument which causes SARMAX to do the necessary differencing before MARIMA is called and to do the opposite integration (of the series) after returning from MARIMA.

When both MLAG (for seasonal lagging) and MDIF are used it is noted that the MLAG definitions are operating on the time series firstly and MDIF afterwards.

Thus if for example a univariate time series $y_{1,t}$ has to be analyzed in differenced form and with a 12 hour seasonal lagging

$$\nabla_{12}y_t - a_1\nabla_{12}y_{t-1} - a_{12}\nabla_{12}y_{t-12} = u_t - b_1u_{t-1} - b_{12}u_{t-12}$$

where ∇_{12} is the 12-period seasonal differencing operator $\nabla_{12} = 1-B^{12}$, then we would use

```
DIMENSION MLAG (3,1),MDIF(2,1)
DATA MLAG /2,1,11/, MV/1/
DATA MDIF /12,12/, NP/1/
```

This will result in the following steps taken in SARMAX

input:	$y_{1,t}$
define:	$y_{2,t} = y_{1,t-11}$
difference:	$y_{1,t} \rightarrow y_{1,t} - y_{1,t-12}$
difference:	$y_{2,t} \rightarrow y_{2,t} - y_{2,t-12}$

At return from SARMAX the array Y(2,N) will contain the original series as well as the new series (as the second variable), both in non-differenced form.

<u>Calling SARMAX.</u>

```
CALL SARMAX  (K,N,Y,U,IP,IQ,MU,MLAG,MV,MQ,MDIF,NP,
              KDET,MODEL,IA,IT,F)
```

where	K	= No. of variables in time series Y after use of array MLAG.
	N	= No. of observations in time series Y.
	Y	= time series , with dimension Y(K,N).
	U	= series of residuals with dimension U(K,N).
	IP	= autoregressive order of MARIMA-model formulated by SARMAX.
	IQ	= moving average order of MARIMA-model formulated by SARMAX. IQ $\geq$ MQ below.
	MU	= 1 if average has to be subtracted before calling MARIMA (this is the normal procedure) = 0 if no average subtraction is wanted
	MLAG	= array giving new lagged variables definitions , dimensioned MLAG(3,MV).
	MV	= No. of new variables defined by MLAG.
	MQ	= moving average order for new variables defined by MLAG. If no moving average terms are wanted for lagged variables use MQ=0. MQ $\leq$ IQ.
	MDIF	= differencing pattern to be applied for the time series including new lagged variables defined by MLAG, dimensioned MDIFF(K,NP).
	NP	= number of differencing patterns to be used. 0 $\leq$ NP $\leq$ 4.

KDET = No. of non-random variables in the input series. These variables must be the last variables in the time series. No model is defined for these variables. They are treated as x-variables.

MODEL = array of model indicators as described in the call of MARIMA.The organisation of MODEL is generally:
MODEL((IP+IQ)*K,K) giving indicators for the models for the variables in Y(K,N) one by one.
(Note that variables generated by MLAG will always be given zeroes in the MODEL array generated by SARMAX when it calls MARIMA.)

IA = 1 the array MODEL is used to define a MARIMA model (see MARIMA description).
0 a default MODEL array is generated with 1's for all variables not generated by MLAG, and with zeroes for deterministic variables.

IT = maximum number of iterations allowed in MARIMA. Generally around 25 is enough. If that is not the case the model may be overparametrized (see MARIMA description).

F = F-test value for including or excluding parameters in the MARIMA model (see MARIMA description).

NOTE: at page 56 and onwards a large example using SARMAX is given. The example includes seasonal lagging and differencing and a deterministic x-variable.

Common areas in MARIMA.

MARIMA uses four labelled commons.

COMMON/CLAGS/MMV,MMLAG(3,10),MAXLAG,MMQ,MPAT,MMDIF,(10,4),
MAXSEA,KAA,IGET

Routines : SARMAX (Computes the content of CLAGS)
MARIMA
FOREC2
KSTEP
FOREC3

MMV	= MV in call to SARMAX = No. of lagged variables generated by array MLAG in SARMAX.
MMLAG	= pattern for generation of lagged variables by SARMAX. The (3xMMV) places are in use if MARIMA was called via SARMAX. Otherwise MMLAG not in use (MMV=0).
MAXLAG	= maximum length of lagging used in call via SAR-MAX. Is found in SARMAX.
MMQ	= order of moving average part for the lagged variables defined by MLAG in call of SARMAX. Is equal to MQ in call of SARMAX.
MPAT	= No. of differencing patterns defined in MMDIF.
MMDIF	= Differencing patterns for at most 10 variables and 4 patterns.
MAXSEA	= Maximum value of differencing period as defined by MMDIF.
KAA	= No.of non-deterministic variables in timeseries.
IGET	= 0 if the routine DIFEQU uses future data from the forecast-array or 1 if DIFEQU uses data from the time series itself for future values.

The values of MPAT,MMDIF,MAXSEA,KAA and IGET are used by the routine FOREC3 when it calls DIFEQU.

All values in CLAGS are integer*4.

COMMON/RESULT/AVERA(10),COEF(70,10),UCOV(10,10)

Routines MARIMA (computes the content of RESULT)
FOREC2
DIFEQU
KSTEP
AVER
FOREC3

AVERA = array containing averages of the K variables in the time series analyzed by MARIMA. REAL*4 array

COEF = array with the coefficients of the MARIMA model. Space is provided for 10 variables. For each variable 70 places are available. The first K places are always zero while the last 70-K places are used for the parameters of the MARIMA model. Thus if IP and IQ are the autoregressive and moving average orders of the MARIMA model and there are $K \leq 10$ variables in the series (1+IP+IQ)*K must be at most 70. REAL*4 array.

UCOV = double precision array for storing the residual covariance matrix. It is organized (10,10). REAL*8 array.

COMMON/SPEC/INI,LAGY,NROT,KVARY,LAGU,MEANX,NDIF

```
Routines  MARIMA   (creates the content of SPEC)
          FOREC2
          DIFEQU
          KSTEP
          TRIN
```

INI = indicator determined in MARIMA. Used when the routine ESTIX is called K times in a K-variate problem in each iteration. INI = 1 tells ESTIX that a new regression problem is to be solved. INI = 0 tells ESTIX that the regression problem is not new, such that the solution can be found quicker.

LAGY = autoregressive order of MARIMA model

NROT = number of observations used in MARIMA

KVARY = dimension of time series

LAGU = moving average order of MARIMA model

MEANX = indicator for average subtraction in MARIMA analyses. 1: average subtracted, 0: no average subtraction.

NDIF = not used anymore.

All values in SPEC are integer*4

COMMON/EXTRA/COV(4900),STORE(2485)
DOUBLE PRECISION COV, STORE

COMMON/EXTRA/PAR(7385),EXT(7385)

Routines MARIMA(COV, STORE organization).
AWAY

DIFEQU (PAR, EXT organization).

COV = array used for storing a maximum of (70,70) array containing up till 70 variables linear regression analysis normal equations. A maximum of $K \leq 10$ dependent and 70-K independent variable can be treated, corresponding to the array COEF(70,10) in common /RESULT. REAL*8 array.

STORE array used by MARIMA to store the purely autoregressive part of normal equations from iteration. The use of STORE speeds up computations REAL*8 array.

Both COV and STORE are generated and used by MARIMA. COV is passed to routine ESTIX which does the regression analyses.

PAR = array used in DIFEQU for storing MARIMA model in integrated autoregressive moving average form and as matrix polynomials. REAL*4.

EXT = left over array not in use in DIFEQU. REAL*4 array.

SEASONAL MODELING EXAMPLE - USE OF S A R M A X

A time series with seasonal variation and a regression input variable is simulated. The input series is generated as a 12-period integrated white noise series but is treated as a deterministic input series. The input innovations are generated as white noise with added 12-period deterministic variation:

$$x_t = x_{t-12} + \text{white noise}$$

$$u_t = \text{sinus(12-period)} + \text{white noise}$$

as is described in the program below.

The time series is then generated according to the model:

$$y_t = 0.80y_{t-1} - 0.20y_{t-2} + 0.50y_{t-12} - 0.25y_{t-24}$$

$$- 0.75u_{t-1} + 0.25u_{t-12} + u_t + 0.50x_{t-1} + 0.25x_{t-2}$$

which then will be a seasonally varying timeseries with a long autocorrelation.

In the example SARMAX is asked to analyze the time series $y_{1,t}$ which is the above described series.

In order to do so the following variables are defined:

$$y_{1,t} = y_t$$
$$y_{2,t} = y_{1,t-11}$$
$$y_{3,t} = y_{1,t-23}$$
$$y_{4,t} = x_t$$

such that a 4-variate problem is treated. In MARIMA these 4 variables are all treated in 12-period differenced form. A simple ARMA(2,1) model is defined as is shown in the example.

```
$JOB WATFIV
      DIMENSION Y(4,1024),U(4,1024),V(4,1024)
      DIMENSION YY(1124),UU(1124),SEASON(12),MODEL(1)
      DIMENSION XX(1124)
      DIMENSION MLAG(3,2),MDIF(4,1),YFOR(4,24),VFOR(4,4,24)
      DOUBLE PRECISION DSEED,VSEED
C
      DATA YY,UU /2248*0.0/
      DATA MLAG/2,1,11 ,3,1,23/
      DATA MDIF/12,12,12,12/
      DATA
     1 K,N,IP,IQ,MU,MV,MQ,NP,KO,IA,IT,F/4,1000,2,1,1,2,1,1,0,0,25,2.0/
      DATA DSEED,VSEED/4711.D0,4911.D0/
C
C   SIMULATE SERIES WITH MAX. 1000 OBSERVATIONS. N IS THE ACTUAL NUMBER
C   ALSO SIMULATE INPUT X-SERIES TO ILLUSTRATE REGRESSION VARIABLE
C
      N=300
      N1=N+1
      N24=N+24
      N124=N+124
C
C GENERATE XX-SERIES AS 12-PERIOD INTEGRATED WHITE NOISE
C
      DO 33 I=1,N124
      CALL NORMVA(0.0, 10.,XX(I),VSEED)
      XX(I)=IFIX(XX(I))
      IF(I.GT.12)XX(I)=XX(I)+XX(I-12)
   33 CONTINUE
C
C   GENERATE A SEASONALLY VARYING UU-SERIES OF INNOVATIONS
C   BY SIMULATION.
C
      DO 2 I=1,12
      SEASON(I)=SIN(24.*3.141592/FLOAT(I))
    2 CONTINUE
C
C   THE ARRAY 'SEASON' IS A SINE WITH A 12 HOUR SEASONALITY
C
        DO 1 I=1,N124
C
C   NORMVA CREATES A NORMAL(0,1) NUMBER
C
      CALL NORMVA(0.0,1.0,UU(I),DSEED)
C
      J=MOD(I,12)+1
C
C   ADD THE 12 HOUR SEASONALITY TO THE NOISE INPUT
C
      UU(I)=UU(I)+SEASON(J)
    1 CONTINUE
C
C DEFINE A MODEL, WHICH CONTAINS SEASONALITY.
C
      AR1=0.80
      AR2=-0.20
      MA1=0.75
      SAR12=0.50
      SMA12=-0.25
      SAR24=-0.25
C
C   GENERATE TIMESERIES USING THE GENERATED XX-ARRAY AS INPUT
C   AND THE SEASONALLY VARYING UU-SERIES AS INNOVATIONS
```

```
C
      DO 4 I=25,N124
      YY(I)= AR1*YY(I-1)+AR2*YY(I-2)+SAR12*YY(I-12)+SAR24*YY(I-24)
     X +UU(I)-MA1*UU(I-1)-SMA12*UU(I-12)
     X +0.50*XX(I-1)+0.25*XX(I-2)
      IF(I.GT.100)Y(1,I-100)=YY(I)
      IF(I.GT.100)Y(4,I-100)=XX(I)
    4 CONTINUE
C
      DO 5 I=1,N24
      IF(Y(1,I).GE.0.)IY=Y(1,I)+.5
      IF(Y(1,I).LE.0.)IY=Y(1,I)-.5
C
C  ROUNDING OFF IN ORDER TO GET NICE NUMBERS
C
      Y(1,I)=IY
    5 CONTINUE
C
C
      WRITE(6,8)N
      WRITE(6,9)(Y(1,I),I=1,N)
C
    8 FORMAT('1TIME SERIES GENERATED BY SIMULATION ( N =',I4,' )'
     X //1X)
    9 FORMAT(1X,12F6.0)
C
      WRITE(6,3)(Y(1,I),I=N1,N24)
    3 FORMAT(////,' 24 NEXT VALUES OF TIME SERIES ARE :'//
     X 1X,12F6.0/1X,12F6.0//' WHICH WILL BE USED FOR TEST OF 24',
     X ' STEPS AHEAD FORECASTING'//1X)
C
      WRITE(6,28)N
      WRITE(6,9)(Y(4,I),I=1,N)
   28 FORMAT('1INPUT X-SERIES GENERATED BY SIMULATION ( N =',I4,' )'
     X //1X)
C
      WRITE(6,29)(Y(4,I),I=N1,N24)
   29 FORMAT(////,' 24 NEXT VALUES OF INPUT SERIES ARE :'//
     X 1X,12F6.0/1X,12F6.0//1X)
C
      KDET=1
C
C  USE  S A R M A X  IN ORDER TO EATIMATE A FORECASTING MODEL
C
      CALL SARMAX
     1(K,N,Y,U,IP,IQ,MU,MLAG,MV,MQ,MDIF,NP,KDET,MODEL,IA,IT,F)
C
C  DO FORECASTING FROM OBSERVATION NO.  N   AND ONWARDS
C  FORECAST LENGTH IS 24 STEPS INCLUDING 90% CONFIDENCE LIMITS
C
C  ROUTINE  D I F E Q U F  DOES FORECASTING USING THE ESTIMATED
C  MODEL WHICH IS PLACED IN COMMON BY  M A R I M A.
C
      CALL DIFEQU(K,N,Y,U,3,1,YFOR,24,3,VFOR,24,0,MDIF,NP,1,1,.90)
C
C  USE ROUTINE  F O R E C 3  FOR COMPUTATION OF RESIDUALS USING
C  MARIMA MODEL PREVIOUSLY PLACED IN COMMONS
C
      CALL FOREC3(K,N,Y,V)
C
C  PRINT OUT MARIMA RESIDUALS AND FOREC3 RESIDUALS.
C  THEY CAN DIFFER A LITTLE.
C
```

```
      DO 12 I=1,300
      WRITE(6,14)(Y(LL,I),LL=1,4),(U(LL,I),LL=1,4),(V(LL,I),LL=1,4)
   14 FORMAT(1X,4F9.4,2X,4F9.4,2X,4F9.4)
   12 CONTINUE
C
C
C     NOW CHECK ANALYSIS BY BI-VARIATE SPECTRAL ANALYSIS :
C
C     IN THIS EXAMPLE WHERE SEASONAL LAGGING AND DIFFERENCING IS
C     USED THE FIRST RESIDUALS ARE PUT EQUAL TO ZERO AND SHOULD
C     NOT BE INCORPORATED IN THE ANALYSIS.
C
      CALL TWOSPE(K,N-35,V(1,36),V,1,4,36,30,3,50,1,1,1,1)
C
C     FINALLY USE  A U T E S T  FOR CHECKING WHITENESS OF RESIDUALS
C
C     PUT 'RESIDUALS' FOR INPUT-X-SERIES AS NO. 2 FOR AUTEST
C     ANALYSIS. (THOSE RESIDUALS ARE THE VALUES OF THE X-SERIES USED
C     IN THE MARIMA-ANALYSIS. IN THIS EXAMPLE THEY ARE THE 12-PERIOD
C     DIFFERENCED SERIES CORRECTED FOR AVERAGE, AS DEFINED IN THE
C     CALLING OF SARMAX)
C
      DO 40 I=1,N
      V(2,I)=V(4,I)
   40 CONTINUE
C
      CALL AUTEST(K,N-35,V(1,36),2,12,4,TQ,NOF)
      STOP
      END
```

```
      SUBROUTINE NORMVA(AM,S,V,DSEED)
C
C     GENERATION OF APPROXIMATELY NORMAL(AM,S**2) DISTRIBUTED
C     RANDOM NUMBERS BY ADDING 12 UNIFORM(0,1) RANDOM NUMBERS
C
      DOUBLE PRECISION DSEED
      A=0.0
      DO 50 I=1,12
      Y=UNIFOR(DSEED)
   50 A=A+Y
      V=(A-6.0)*S+AM
      RETURN
      END
```

```
      REAL FUNCTION UNIFOR (DSEED)
      DOUBLE PRECISION   DSEED
      DOUBLE PRECISION   D2P31M,D2P31
      DATA               D2P31M/2147483647.D0/
      DATA               D2P31 /2147483648.D0/
      DSEED = DMOD(16807.D0*DSEED,D2P31M)
      UNIFOR = DSEED / D2P31
      RETURN
      END
```

TIME SERIES GENERATED BY SIMULATION (N = 300)

15.	-28.	-22.	5.	31.	54.	96.	93.	61.	35.	30.	21.
7.	-24.	5.	57.	87.	99.	129.	119.	83.	50.	43.	13.
-27.	-63.	-22.	57.	110.	144.	183.	170.	129.	85.	60.	6.
-53.	-106.	-72.	18.	87.	144.	212.	206.	166.	118.	91.	31.
-35.	-108.	-93.	-17.	46.	108.	202.	224.	192.	144.	115.	58.
-5.	-76.	-68.	-13.	33.	75.	179.	213.	181.	135.	112.	64.
23.	-21.	-4.	36.	56.	73.	166.	190.	146.	96.	83.	48.
12.	-5.	43.	92.	107.	105.	178.	183.	120.	62.	52.	23.
-9.	-24.	44.	111.	141.	140.	209.	206.	129.	56.	33.	0.
-34.	-51.	8.	69.	104.	117.	218.	239.	174.	96.	49.	-1.
-35.	-54.	-7.	29.	42.	51.	166.	222.	191.	136.	101.	42.
-5.	-38.	4.	25.	18.	10.	116.	180.	159.	135.	136.	92.
40.	0.	27.	45.	30.	14.	102.	144.	99.	79.	114.	104.
65.	35.	59.	78.	62.	49.	124.	144.	66.	18.	50.	63.
50.	37.	74.	104.	91.	90.	163.	177.	93.	22.	27.	23.
9.	1.	53.	107.	102.	101.	174.	196.	128.	54.	46.	24.
-12.	-43.	4.	79.	94.	96.	175.	203.	134.	62.	62.	55.
9.	-44.	-14.	52.	71.	77.	171.	215.	149.	79.	90.	93.
51.	-16.	-1.	51.	62.	65.	161.	211.	157.	102.	113.	111.
70.	8.	22.	74.	81.	78.	152.	189.	143.	99.	111.	111.
65.	3.	22.	82.	103.	103.	164.	180.	118.	75.	89.	97.
46.	-24.	-3.	65.	104.	113.	177.	191.	114.	61.	73.	75.
24.	-52.	-43.	31.	79.	99.	178.	201.	126.	65.	75.	68.
14.	-71.	-78.	-20.	25.	51.	146.	191.	136.	87.	94.	88.
30.	-72.	-89.	-45.	-22.	-13.	92.	153.	121.	90.	107.	117.

24 NEXT VALUES OF TIME SERIES ARE :

59.	-60.	-77.	-26.	-28.	-61.	26.	95.	77.	55.	90.	125.
83.	-35.	-58.	-9.	-10.	-60.	11.	68.	45.	22.	58.	99.

WHICH WILL BE USED FOR TEST OF 24 STEPS AHEAD FORECASTING

INPUT X-SERIES GENERATED BY SIMULATION (N = 300)

-58.	51.	11.	35.	28.	61.	-18.	-7.	-16.	9.	-18.	-22.
-43.	64.	34.	29.	26.	68.	-10.	4.	-3.	25.	-46.	-33.
-49.	53.	42.	24.	49.	74.	-7.	10.	-4.	12.	-57.	-28.
-55.	56.	49.	18.	50.	85.	-24.	10.	-10.	25.	-49.	-20.
-58.	62.	45.	24.	47.	107.	-12.	10.	-11.	22.	-51.	-23.
-53.	68.	43.	34.	38.	136.	-28.	-6.	-19.	25.	-49.	2.
-29.	68.	41.	19.	38.	142.	-29.	-10.	-24.	29.	-49.	-20.
-14.	75.	30.	26.	24.	142.	-36.	-20.	-11.	28.	-44.	-15.
-17.	82.	28.	26.	14.	150.	-33.	-19.	-15.	23.	-39.	-18.
-15.	71.	5.	9.	1.	167.	-19.	-6.	-13.	18.	-37.	1.
-12.	78.	-8.	1.	-2.	153.	-14.	-10.	-6.	33.	-40.	1.
-21.	87.	-8.	6.	6.	165.	-7.	-23.	10.	35.	-40.	-4.
-11.	68.	4.	-3.	14.	154.	-15.	-45.	8.	52.	-44.	-6.
-6.	67.	9.	-5.	18.	143.	-19.	-57.	6.	51.	-33.	-3.
-9.	68.	9.	-9.	32.	128.	-7.	-44.	11.	61.	-31.	-10.
-9.	68.	29.	-23.	27.	115.	-5.	-44.	-8.	58.	-25.	-16.
-31.	72.	43.	-19.	26.	129.	-3.	-66.	-17.	56.	-8.	-25.
-23.	80.	43.	-14.	20.	150.	10.	-52.	-4.	73.	-6.	-23.
-37.	78.	45.	-6.	25.	148.	3.	-45.	10.	60.	-16.	-32.
-38.	77.	50.	0.	25.	124.	2.	-40.	3.	54.	-8.	-45.
-47.	66.	47.	6.	26.	120.	-3.	-52.	8.	49.	-1.	-55.
-52.	67.	47.	12.	16.	117.	-4.	-51.	14.	51.	-11.	-52.
-59.	55.	59.	9.	16.	125.	-4.	-55.	18.	56.	-13.	-47.
-74.	44.	34.	4.	5.	129.	-6.	-37.	18.	54.	-1.	-45.
-88.	50.	32.	-10.	-4.	142.	-10.	-37.	11.	55.	13.	-55.

24 NEXT VALUES OF INPUT SERIES ARE :

-100.	70.	47.	-34.	-35.	135.	1.	-46.	4.	58.	23.	-55.
-100.	55.	34.	-34.	-39.	151.	1.	-35.	7.	58.	20.	-48.

```
CALLING  S A R M A X  FOR MULTIVARIATE TIME SERIES ANALYSIS
INCLUDING GENERAL SEASONAL LAGGING AND/OR DIFFERENCING


NO. OF VARIABLES ADDED BY SEASONAL LAGGING =    2
NO. OF DETERMINISTIC VARIABLES SPECIFIED   =    1


*** BOTH SEASONAL LAGGING AND DETERMINISTIC VARI
*** ABLES SPECIFIED. MAKE SURE YOUR LAGGING STRUCTURE
*** IS THE ONE YOU WANT. FOR LAGGED VARIABLES NO
*** M A R I M A  MODEL IS FORMULATED BY SARMAX.
*** THEY ARE TREATED VIA THE LAGGING DEFINITIONS.

 Y( 2,T) = Y( 1,T- 11)   SEASON=  12

LEADING (MISSING) VALUES FOR NEW VARIABLE NO.   2
ARE PUT EQUAL TO AVERAGE FOR OLD VARIABLE NO.   1
THIS AVERAGE IS =       72.3267

 Y( 3,T) = Y( 1,T- 23)   SEASON=  24

LEADING (MISSING) VALUES FOR NEW VARIABLE NO.   3
ARE PUT EQUAL TO AVERAGE FOR OLD VARIABLE NO.   1
THIS AVERAGE IS =       72.3267


 STATUS OF VARIABLES FOR MARIMA ANALYSIS
  R=RANDOM  D=DETERMINISTIC  L=LAGGED

 1 = R ,  2 = L ,  3 = L ,  4 = D


 NO. OF VARIABLES IN MARIMA PROBLEM          =    4
 NO. OF DEPENDENT (RANDOM) VARIABLES         =    1
 NO. OF DETERMINISTIC OR LAGGED VARIABLES    =    3


 M A R I M A  MODEL INCLUDES (SEASONALLY) DIFFERENCING

 THE DIFFERENCING PATTERN(S) IS (ARE) AS FOLLOWS

      12  12  12  12
```

```
INITIALIZING    M A R I M A

  K        N      P      D      Q MEAN    IT         F      IN
  4      264      2      0      1    1    25     2.000      99

5 FIRST AND 5 LAST OBSERVATIONS IN INPUT ARE:

  OBS.NR.      DATA

     1  -26.00  -39.00    4.00   -6.00
     2  -43.00  -27.00   27.00    3.00
     3  -50.00    0.00   52.00    7.00
     4  -39.00   23.00   56.00   -6.00
     5  -23.00   45.00   45.00    1.00
   260  -38.00   10.00   12.00    0.00
   261  -15.00   22.00    4.00   -7.00
   262    3.00   19.00    2.00    1.00
   263   13.00   20.00   -7.00   14.00
   264   29.00   16.00  -10.00  -10.00

M A R I M A   HAS BEEN CALLED FROM   S A R M A X
DEFINING NEW SEASONALLY LAGGED VARIABLES

Y( 2,T) = Y( 1,T- 11)     SEASON=  12
Y( 3,T) = Y( 1,T- 23)     SEASON=  24
MAX. SEASONAL LAG VARIABLE DEFINITION IS   23



G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

A   M U L T I V A R I A T E   A R I M A - P R O C E S S
A N D   T R A N S F E R - F U N C T I O N   P R O G R A M
BY HENRIK SPLIID , I M S O R ,
THE INSTITUTE OF MATHEMATICAL STATISTICS AND OPERATIONS RESEARCH,
TECHNICAL UNIVERSITY OF DENMARK, DK 2800 LYNGBY, DENMARK.

DIMENSION OF TIMESERIES                              =    4  (K)
NUMBER OF OBSERVATIONS                               =  264  (N)
   -   -  LAGS IN AUTOREGRESSIVE PART OF MODEL       =    2  (P)
   -   -  DIFFERENCING OPERATIONS                    =    0  (D)
   -   -  LAGS IN MOVING AVERAGE PART OF MODEL       =    1  (Q)
INDICATOR FOR MEANS CORRECTION OF DATA               =    1  (M)
MAXIMUM NUMBER OF ITERATIONS                         =   25

FORMULATION OF THE AUTOREGRESSIVE MOVING AVERAGE  ( A R M A )  MODEL :
Y(T) = A(1)*Y(T-1)+...+A(P)*Y(T-P) + U(T)-B(1)*U(T-1)-...-B(Q)*U(T-Q),
A(1),....,A(P) AND B(1),....,B(Q) BEING THE UNKNOWN COEFFICIENT MATRICES,
Y(T) THE K-DIMENSIONAL TIMESERIES AND U(T) THE (UNKNOWN) INNOVATIONS.

MODEL WILL BE ESTIMATED FOR TIMESERIES CORRECTED FOR  M E A N S .
THOSE MEANS WERE FOUND (AFTER DIFFERENCING) TO BE :

      -1.3750          0.1705          1.7348          -0.0758

AN F-TEST MODEL  I D E N T I F I C A T I O N  WAS SPECIFIED.
PARAMETERS WITH F-VALUES LOWER THAN    2.00 ARE EXCLUDED.
```

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

A SPECIAL MODEL HAS BEEN SPECIFIED AS FOLLOWS :

VARIABEL 1 INDICATORS	1 1 1 1	1 1 1 1	1 1 1 0
VARIABEL 2 INDICATORS	0 0 0 0	0 0 0 0	0 0 0 0
VARIABEL 3 INDICATORS	0 0 0 0	0 0 0 0	0 0 0 0
VARIABEL 4 INDICATORS	0 0 0 0	0 0 0 0	0 0 0 0

C O V A R I A N C E   M A T R I X   FOR VARIABLES IN ANALYSIS.

7.86060D 02	1.34954D 02	-4.49476D 02	1.33133D-02
1.34954D 02	7.93039D 02	4.29505D 02	9.89508D 00
-4.49476D 02	4.29505D 02	7.91344D 02	6.14808D-01
1.33133D-02	9.89508D 00	6.14808D-01	9.21995D 01

INNOVATION COVARIANCE MATRIX TRACE IN FIRST ITERATION   9.97205E 01

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

RESULTS FROM ESTIMATION   B E F O R E   T H E
M O D E L   I D E N T I F I C A T I O N .

NUMBER OF ITERATIONS USED = 9

MARIMA MODEL IS OF ORDER (K,P,D,Q) = (4, 2, 0, 1)

AUTOREGRESSIVE PART OF MODEL

A(1)=	0.7428	0.4947	-0.2528	0.5064
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
A(2)=	-0.1608	0.0320	0.0000	0.2580
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000

MOVING AVERAGE PART OF MODEL

B(1)=	-0.0380	0.4602	0.2348	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000

I N N O V A T I O N   C O V A R I A N C E   M A T R I X

THE MATRIX IS BASED ONLY ON THE LAST (N - MAX(P,Q) - D)
RESIDUALS. IT IS CALCULATED AS SUMS OF SQUARES AND
PRODUCTS DIVIDED BY (N - MAX(P,Q) - D - M) AND IT IS
NOT ADJUSTED FOR MODEL DEGREES OF FREEDOM OR AVERAGES.

A SEASONALLY LAGGING HAS BEEN DEFINED VIA   S A R M A X

THUS THE FIRST 23 RESIDUALS ARE ALSO TAKEN OUT
AND THE NO. OF RESIDUALS ANALYZED IN MARIMA IS REDUCED
FROM N = 264 TO N = 241

1.84790 00	1.37020-01	1.24600-02	-1.37240 00
1.37020-01	2.16060 00	7.96090-02	-3.67710-01
1.24600-02	7.96090-02	2.10850 00	-5.74710-01
-1.37240 00	-3.67710-01	-5.74710-01	9.27350 01

AVERAGES FOR RESIDUALS FROM ESTIMATION WAS

3.6385E-01   2.9312E-01   2.4649E-01   1.0861E-02

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S.

I N N O V A T I O N   C O R R E L A T I O N   M A T R I X

COMPUTED FROM THE ABOVE COVARIANCE MATRIX

1.0000	0.0686	0.0063	-0.1048
0.0686	1.0000	0.0373	-0.0260
0.0063	0.0373	1.0000	-0.0411
-0.1048	-0.0260	-0.0411	1.0000

M O D E L   S I G N I F I C A N C E   C O N T R O L

APPROXIMATE PARTIAL F-TEST-VALUES FOR COEFFICIENTS.

THE F-VALUES CAN BE COMPARED TO AN F(1, 251) DISTRIBUTION.

F-VALUES FOR AUTOREGRESSIVE PART OF MODEL

LAG= 1	352.33	413.01	316.40	754.27
	12.84	799.26	490.66	0.25
	249.86	6.58	913.43	0.22
	0.30	0.16	0.09	0.33
LAG= 2	31.35	1.43	0.00	83.12
	88.79	115.25	907.97	0.00
	234.29	13.61	115.94	0.14
	0.97	0.25	0.48	0.32

F-VALUES FOR MOVING AVERAGE PART OF MODEL

LAG= 1	0.08	13.60	3.35	0.00
	0.27	2.48	2.52	0.22
	0.43	0.03	3.09	0.22
	0.11	0.14	0.00	0.33

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

RESULTS FROM   F I N A L   E S T I M A T I O N

NUMBER OF ITERATIONS USED = 14

MARIMA MODEL IS OF ORDER (K,P,D,Q) = (4, 2, 0, 1)

AUTOREGRESSIVE PART OF MODEL

A(1)=	0.7664	0.5172	-0.2558	0.5047
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
A(2)=	-0.1740	0.0000	0.0000	0.2453
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000

MOVING AVERAGE PART OF MODEL

B(1)=	-0.0000	0.4628	0.2498	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000
	-0.0000	-0.0000	-0.0000	-0.0000

I N N O V A T I O N   C O V A R I A N C E   M A T R I X

THE MATRIX IS BASED ONLY ON THE LAST (N - MAX(P,Q) - D) RESIDUALS. IT IS CALCULATED AS SUMS OF SQUARES AND PRODUCTS DIVIDED BY (N - MAX(P,Q) - D - M) AND IT IS NOT ADJUSTED FOR MODEL DEGREES OF FREEDOM OR AVERAGES.

A SEASONALLY LAGGING HAS BEEN DEFINED VIA   S A R M A X

THUS THE FIRST 23 RESIDUALS ARE ALSO TAKEN OUT AND THE NO. OF RESIDUALS ANALYZED IN MARIMA IS REDUCED FROM N = 264 TO N = 241

1.8559D 00	6.2336D-02	-4.0414D-02	-1.2485D 00
6.2336D-02	2.1687D 00	9.4846D-02	-3.7126D-01
-4.0414D-02	9.4846D-02	2.1087D 00	-4.9637D-01
-1.2485D 00	-3.7126D-01	-4.9637D-01	9.2735D 01

AVERAGES FOR RESIDUALS FROM ESTIMATION WAS

3.9502E-01	3.2562E-01	2.7646E-01	1.0861E-02

G E N E R A L   M A R I M A   M O D E L   A N A L Y S I S .

I N N O V A T I O N   C O R R E L A T I O N   M A T R I X

COMPUTED FROM THE ABOVE COVARIANCE MATRIX

1.0000	0.0311	-0.0204	-0.0952
0.0311	1.0000	0.0444	-0.0262
-0.0204	0.0444	1.0000	-0.0355
-0.0952	-0.0262	-0.0355	1.0000

M O D E L   S I G N I F I C A N C E   C O N T R O L

APPROXIMATE PARTIAL F-TEST-VALUES FOR COEFFICIENTS.

THE F-VALUES CAN BE COMPARED TO AN F(1, 253) DISTRIBUTION.

F-VALUES FOR AUTOREGRESSIVE PART OF MODEL

LAG= 1	660.00	964.02	343.44	762.83
	12.84	799.26	490.66	0.25
	249.86	6.58	813.43	0.22
	0.30	0.16	0.09	0.33
LAG= 2	49.82	1.64	1.25	103.86
	88.79	115.25	907.97	0.00
	234.29	13.61	115.94	0.14
	0.97	0.25	0.48	0.32

F-VALUES FOR MOVING AVERAGE PART OF MODEL

LAG= 1	0.00	13.76	3.81	0.00
	0.78	2.64	2.35	0.22
	1.31	0.00	3.30	0.22
	0.00	0.17	0.02	0.33

```
RELATIVE TRACE OF INNOVATION COVARIANCE MATRIX               CPU TIME USED
ITERATION      REL. TRACE    GRAPH OF TRACE                        SECONDS

     1        4.000000 00    I                               I       0.633
     2        3.552800 00    I   *                           I       1.249
     3        3.510030 00    I*                              I       1.881
     4        3.558400 00    I    *                          I       2.495
     5        3.599210 00    I      *                        I       3.110
     6        3.611910 00    I      *                        I       3.720
     7        3.612690 00    I      *                        I       4.354
     8        3.617350 00    I       *                       I       4.990
     9        3.624160 00    I       *                       I       5.628
    10        3.599260 00    I      *                        I       6.283
    11        3.600700 00    I      *                        I       6.854
    12        3.607060 00    I      *                        I       7.424
    13        3.622380 00    I       *                       I       7.996
    14        3.631210 00    I       *                       I       8.565
```

THE TRACE MUST REACH A STABLE VALUE - OTHERWISE CONVERGENCE CANNOT BE ASSU

TOTAL CPU TIME USED FOR ESTIMATION = 9.01

INITIALIZING   D I F E Q U

K	N	KA	IGET	NFY	IPRY	NFV	IPRV	NPAT	LPR	INIT	(INIT)
4	300	3	1	24	3	24	0	1	1	1	0

M A R I M A M O D E L IN DIFFERENCE EQUATION FORM

Y(T)-A(1)Y(T-1)-...-A(P)Y(T-P) = CST + U(T)-B(1)U(T-1)-...-B(Q)U(T-Q)

MARIMA MODEL INCLUDES (SEASONALLY) DIFFERENCING
THE FOLLOWING MODEL REFERS TO THE NOT DIFFERENCED SERIES
(SEASONAL) DIFFERENCING PATTERN(S) FOR THE VARIABLES :

12 12 12 12

AVERAGES FOR VARIABLES
ANALYZED IN MARIMA WERE :

-1.3750
0.1705
1.7348
-0.0758

VECTOR OF CONSTANTS (CST) :
IN THE (AGGREGATED) MODEL :

-0.1480
0.1705
1.7348
-0.0758

AUTOREGRESSIVE MATRIX COEFFICIENTS:

A(1) =	0.7664	0.5172	-0.2558	0.5047
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
A(2) =	-0.1740	0.0000	0.0000	0.2453
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
A(12) =	1.0000	0.0000	0.0000	0.0000
	0.0000	1.0000	0.0000	0.0000
	0.0000	0.0000	1.0000	0.0000
	0.0000	0.0000	0.0000	1.0000
A(13) =	-0.7664	-0.5172	0.2558	-0.5047
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
A(14) =	0.1740	0.0000	0.0000	-0.2453
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000

MOVING AVERAGE MATRIX COEFFICIENTS:

B(1) =	0.0000	0.4628	0.2498	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.0000

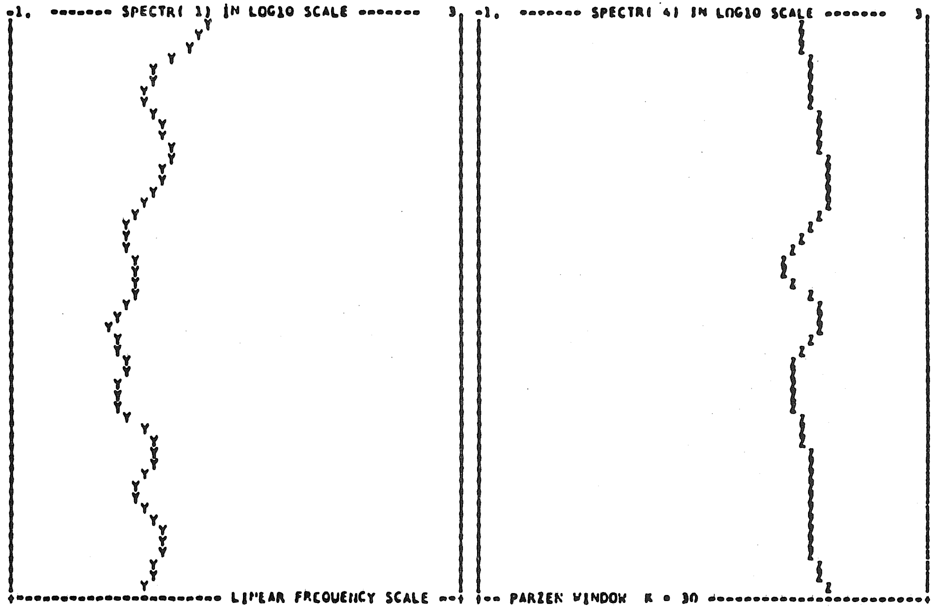
```
          MULTISTEP AHEAD FORECASTING ROUTINE
          -----------------------------------

          FORECASTING FROM OBSERVATION NO.  300
          FORECAST LENGTH IS                 24
          NO. OF VARIABLES IN SERIES          4
          NO. OF NONDETERMINISTIC VAR.S.      3
          NO. OF SEASONALLY LAGGED VAR.S      2


THE FOLLOWING SEASONAL LAGGING IS USED

Y( 2,T) = Y( 1,T- 11)
Y( 3,T) = Y( 1,T- 23)

  301 ( 1)        58.4134     -72.0000     -71.0000    -100.0000

LOW  5.0% =       56.1721     -72.0000     -71.0000
UPR 95.0% =       60.6546     -72.0000     -71.0000

  302 ( 2)       -60.4435     -89.0000     -78.0000      70.0000

LOW  5.0% =      -63.2672     -89.0000     -78.0000
UPR 95.0% =      -57.6197     -89.0000     -78.0000

  303 ( 3)       -75.3386     -45.0000     -20.0000      47.0000

LOW  5.0% =      -78.3105     -45.0000     -20.0000
UPR 95.0% =      -72.3667     -45.0000     -20.0000

  304 ( 4)       -24.6538     -22.0000      25.0000     -34.0000

LOW  5.0% =      -27.6540     -22.0000      25.0000
UPR 95.0% =      -21.6536     -22.0000      25.0000

  305 ( 5)       -27.8053     -13.0000      51.0000     -35.0000

LOW  5.0% =      -30.8094     -13.0000      51.0000
UPR 95.0% =      -24.8011     -13.0000      51.0000

  306 ( 6)       -63.1953      92.0000     146.0000     135.0000

LOW  5.0% =      -66.1999      92.0000     146.0000
UPR 95.0% =      -60.1908      92.0000     146.0000

  307 ( 7)        23.6824     153.0000     191.0000       1.0000

LOW  5.0% =       20.6778     153.0000     191.0000
UPR 95.0% =       26.6869     153.0000     191.0000

  308 ( 8)        94.7255     121.0000     136.0000     -46.0000

LOW  5.0% =       91.7210     121.0000     136.0000
UPR 95.0% =       97.7301     121.0000     136.0000

  309 ( 9)        75.0304      90.0000      87.0000       4.0000

LOW  5.0% =       72.0259      90.0000      87.0000
UPR 95.0% =       78.0349      90.0000      87.0000

  310 (10)        53.9899     107.0000      94.0000      58.0000

LOW  5.0% =       50.9854     107.0000      94.0000
UPR 95.0% =       56.9944     107.0000      94.0000

  311 (11)        88.1304     117.0000      88.0000      23.0000

LOW  5.0% =       85.1259     117.0000      88.0000
UPR 95.0% =       91.1349     117.0000      88.0000
```

312 (12)	123.2082	58.4134	30.0000	-55.0000
LOW 5.0% =	120.2037	56.1721	30.0000	
UPR 95.0% =	126.2127	60.6546	30.0000	
313 (13)	79.1828	-60.4435	-72.0000	-100.0000
LOW 5.0% =	75.3603	-63.2672	-72.0000	
UPR 95.0% =	83.0053	-57.6197	-72.0000	
314 (14)	-39.7156	-75.3386	-89.0000	55.0000
LOW 5.0% =	-44.3952	-78.3105	-89.0000	
UPR 95.0% =	-35.0359	-72.3667	-89.0000	
315 (15)	-61.1113	-24.6538	-45.0000	34.0000
LOW 5.0% =	-66.2558	-27.6540	-45.0000	
UPR 95.0% =	-55.9667	-21.6536	-45.0000	
316 (16)	-11.0291	-27.8053	-22.0000	-34.0000
LOW 5.0% =	-16.3557	-30.8094	-22.0000	
UPR 95.0% =	-5.7024	-24.8011	-22.0000	
317 (17)	-14.0353	-63.1953	-13.0000	-39.0000
LOW 5.0% =	-19.4168	-66.1999	-13.0000	
UPR 95.0% =	-8.6539	-60.1908	-13.0000	
318 (18)	-66.3438	23.6824	92.0000	151.0000
LOW 5.0% =	-71.7379	20.6778	92.0000	
UPR 95.0% =	-60.9495	26.6869	92.0000	
319 (19)	4.3760	94.7255	153.0000	1.0000
LOW 5.0% =	-1.0204	91.7210	153.0000	
UPR 95.0% =	9.7725	97.7301	153.0000	
320 (20)	63.7131	75.0304	121.0000	-35.0000
LOW 5.0% =	58.3163	72.0259	121.0000	
UPR 95.0% =	69.1098	78.0349	121.0000	
321 (21)	39.2725	53.9899	90.0000	7.0000
LOW 5.0% =	33.8758	50.9854	90.0000	
UPR 95.0% =	44.6692	56.9944	90.0000	
322 (22)	16.3423	88.1304	107.0000	58.0000
LOW 5.0% =	10.9456	85.1259	107.0000	
UPR 95.0% =	21.7390	91.1349	107.0000	
323 (23)	52.7386	123.2082	117.0000	20.0000
LOW 5.0% =	47.3419	120.2037	117.0000	
UPR 95.0% =	58.1353	126.2127	117.0000	
324 (24)	96.4401	79.1828	58.4134	-48.0000
LOW 5.0% =	91.0434	75.3603	56.1721	
UPR 95.0% =	101.8368	83.0053	60.6546	

```
CALLING   F O R E C 3   FOR COMPUTATION OF RESIDUALS FROM MARIMA MODEL

NO. OF VARIABLES ADDED BY SEASONAL LAGGING =     2

Y( 2,T) = Y( 1,T- 11)

LEADING (MISSING) VALUES FOR NEW VARIABLE NO.    2
ARE PUT EQUAL TO AVERAGE FOR OLD VARIABLE NO.    1
THIS AVERAGE IS =         72.3267

Y( 3,T) = Y( 1,T- 23)

LEADING (MISSING) VALUES FOR NEW VARIABLE NO.    3
ARE PUT EQUAL TO AVERAGE FOR OLD VARIABLE NO.    1
THIS AVERAGE IS =         72.3267


NO. OF DIFFERENCING PATTERN(S) USED IS           1

PATTERN NO.  1  IS     12    12    12    12


LENGTH OF INPUT SERIES IS                 300
OBSERVATIONS LOST BY LAGGING               23
OBSERVATIONS LOST BY DIFFERENCING          12
LENGTH OF FULLY OBSERVED SERIES           265
FORECASTING STARTS AT OBSERVATION          36


INITIALIZING   D I F E Q U

  K    N   KA IGET  NFY IPRY  NFV IPRV NPAT  LPR INIT (INIT)
  4  301    3    1    0    0    1    3    1    1    1    0
```

At this point output from the routine FOREC3's call of DIFEQU is left out. Also the printout of the reconstructed residuals via the call of FOREC3 is left out.

FREQ.	SPECT-Y	SPECT-Z
0.0000	5.473E 00	7.942E 01
0.0100	4.954E 00	8.011E 01
0.0200	3.758E 00	8.164E 01
0.0300	2.618E 00	8.347E 01
0.0400	[illegible]	[illegible]
0.0500	[illegible]	[illegible]
0.0600	[illegible]	[illegible]
0.0700	[illegible]	[illegible]
0.0800	[illegible]	[illegible]
0.0900	[illegible]	[illegible]
0.1000	[illegible]	[illegible]
0.1100	[illegible]	[illegible]
0.1200	[illegible]	[illegible]
0.1300	[illegible]	[illegible]
0.1400	[illegible]	[illegible]
0.1500	[illegible]	[illegible]
0.1600	[illegible]	[illegible]
0.1700	[illegible]	[illegible]
0.1800	[illegible]	[illegible]
0.1900	[illegible]	[illegible]
0.2000	[illegible]	[illegible]
0.2100	[illegible]	[illegible]
0.2200	[illegible]	[illegible]
0.2300	[illegible]	[illegible]
0.2400	[illegible]	[illegible]
0.2500	[illegible]	[illegible]
0.2600	[illegible]	[illegible]
0.2700	[illegible]	[illegible]
0.2800	[illegible]	[illegible]
0.2900	[illegible]	[illegible]
0.3000	[illegible]	[illegible]
0.3100	[illegible]	[illegible]
0.3200	[illegible]	[illegible]
0.3300	[illegible]	[illegible]
0.3400	[illegible]	[illegible]
0.3500	[illegible]	[illegible]
0.3600	[illegible]	[illegible]
0.3700	[illegible]	[illegible]
0.3800	[illegible]	[illegible]
0.3900	[illegible]	[illegible]
0.4000	[illegible]	[illegible]
0.4100	[illegible]	[illegible]
0.4200	[illegible]	[illegible]
0.4300	[illegible]	[illegible]
0.4400	[illegible]	[illegible]
0.4500	[illegible]	[illegible]
0.4600	[illegible]	[illegible]
0.4700	[illegible]	[illegible]
0.4800	[illegible]	[illegible]
0.4900	[illegible]	[illegible]
0.5000	1.725E 00	1.241E 02

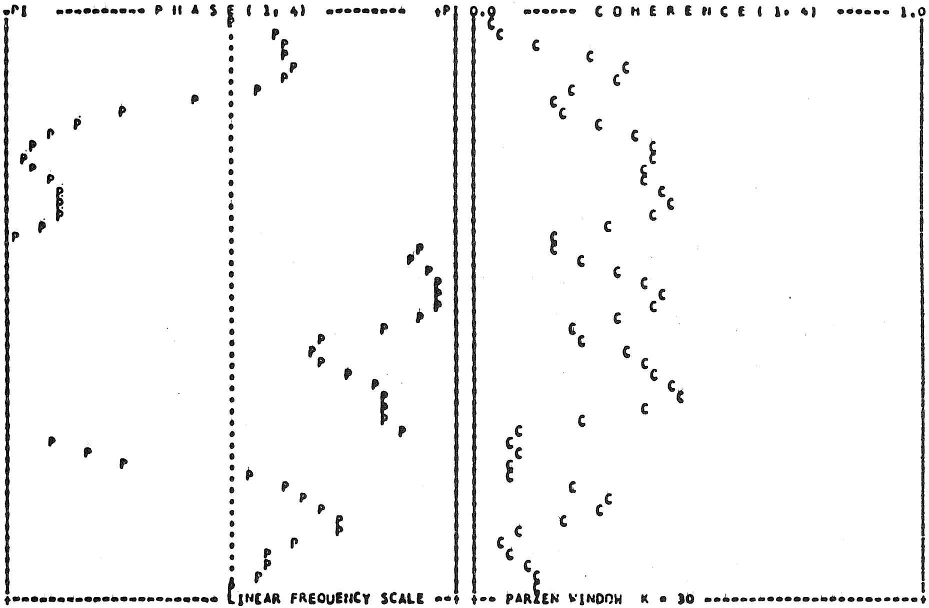


CROSS CORRELATION FUNCTION

RHO(1(T), 4(T-LAG))

LAG	
-36	-0.0821
-35	-0.0642
-34	-0.0029
-33	-0.0256
-32	0.0267
-31	0.0555
-30	0.0004
-29	-0.1101
-28	-0.1418
-27	-0.0613
-26	-0.0016
-25	0.0306
-24	0.0220
-23	0.0117
-22	-0.0644
-21	-0.0032
-20	-0.0463
-19	-0.0308
-18	-0.0820
-17	0.0448
-16	-0.1115
-15	-0.0023
-14	-0.0107
-13	-0.0361
-12	0.0061
-11	0.0055
-10	0.0995
-9	-0.0908
-8	-0.0689
-7	0.0362
-6	0.0156
-5	0.0879
-4	0.0926
-3	0.0414
-2	0.0012
-1	-0.0239
0	-0.0743
1	-0.0552
2	0.1593
3	0.0945
4	0.0317
5	-0.0441
6	-0.0159
7	-0.0747
8	-0.0831
9	-0.0642
10	-0.0417
11	-0.0199
12	-0.0535
13	0.0026
14	0.1549
15	0.0446
16	0.0176
17	-0.0171
18	-0.0467
19	-0.0520
20	-0.0241
21	-0.0111
22	0.1076
23	-0.0396
24	0.0314
25	-0.1698
26	-0.0395
27	0.1119
28	-0.0189
29	0.0621
30	0.0185
31	-0.0295
32	-0.0615
33	-0.0433
34	-0.0006
35	0.0124
36	0.0710

FREQ.	PHASE (P)	COHER. (C)
0.0000	0.0000	0.0336
0.0100	[illegible]	0.0658
0.0200	[illegible]	[illegible]
0.0300	[illegible]	[illegible]
0.0400	[illegible]	[illegible]
0.0500	[illegible]	[illegible]
0.0600	[illegible]	[illegible]
0.0700	[illegible]	[illegible]
0.0800	[illegible]	[illegible]
0.0900	[illegible]	[illegible]
0.1000	[illegible]	[illegible]
0.1100	[illegible]	[illegible]
0.1200	[illegible]	[illegible]
0.1300	[illegible]	[illegible]
0.1400	[illegible]	[illegible]
0.1500	[illegible]	[illegible]
0.1600	[illegible]	[illegible]
0.1700	[illegible]	[illegible]
0.1800	[illegible]	[illegible]
0.1900	[illegible]	[illegible]
0.2000	[illegible]	[illegible]
0.2100	[illegible]	[illegible]
0.2200	[illegible]	[illegible]
0.2300	[illegible]	[illegible]
0.2400	[illegible]	[illegible]
0.2500	[illegible]	[illegible]
0.2600	[illegible]	[illegible]
0.2700	[illegible]	[illegible]
0.2800	[illegible]	[illegible]
0.2900	[illegible]	[illegible]
0.3000	[illegible]	[illegible]
0.3100	[illegible]	[illegible]
0.3200	[illegible]	[illegible]
0.3300	[illegible]	[illegible]
0.3400	[illegible]	[illegible]
0.3500	[illegible]	[illegible]
0.3600	[illegible]	[illegible]
0.3700	[illegible]	[illegible]
0.3800	[illegible]	[illegible]
0.3900	[illegible]	[illegible]
0.4000	[illegible]	[illegible]
0.4100	[illegible]	[illegible]
0.4200	[illegible]	[illegible]
0.4300	[illegible]	[illegible]
0.4400	[illegible]	[illegible]
0.4500	[illegible]	[illegible]
0.4600	[illegible]	0.0509
0.4700	[illegible]	0.0819
0.4800	[illegible]	[illegible]
0.4900	[illegible]	[illegible]
0.5000	0.0000	0.1371



M U L T I V A R I A T E   A U T O C O R R E L A T I O N   T E S T .

```
NUMBER OF VARIABLES IN INPUT VECTOR   =     4
LENGTH OF TIME SERIES                 =   265
NUMBER OF VARIABLES TO BE TESTED      =     2
NUMBER OF LAGS INCORPORATED IN TEST   =    12
```

ORDINARY CORRELATION MATRIX (LAG ZERO)

```
 1.0000 -0.0743
-0.0743  1.0000
```

STANDARD DEVIATIONS OF VARIABLES IN TEST

```
1.3307  9.5657
```

INVERSE CORRELATION MATRIX

```
1.0055  0.0747
0.0747  1.0055
```

ORGANISATION OF CROSS-AUTO-CORRELATION MATRICES

R(I,J) = COR(A(I,T) , A(J,T-LAG)).

LAG = 1 CROSS-AUTO-CORRELATION-MATRIX

```
 0.1280 -0.0552
-0.0239  0.0355
```

CONTRIBUTION TO CHI-SQUARE-TEST = 5.20

LAG = 2 CROSS-AUTO-CORRELATION-MATRIX

```
0.2284  0.1593
0.0012  0.0352
```

CONTRIBUTION TO CHI-SQUARE-TEST = 22.82

LAG = 3 CROSS-AUTO-CORRELATION-MATRIX

```
0.0111  0.0945
0.0414 -0.0815
```

CONTRIBUTION TO CHI-SQUARE-TEST = 4.30

LAG = 4 CROSS-AUTO-CORRELATION-MATRIX

```
0.0207  0.0317
0.0926 -0.0391
```

CONTRIBUTION TO CHI-SQUARE-TEST = 3.01

LAG = 5 CROSS-AUTO-CORRELATION-MATRIX

```
0.0593 -0.0441
0.0879 -0.0642
```

CONTRIBUTION TO CHI-SQUARE-TEST = 4.61

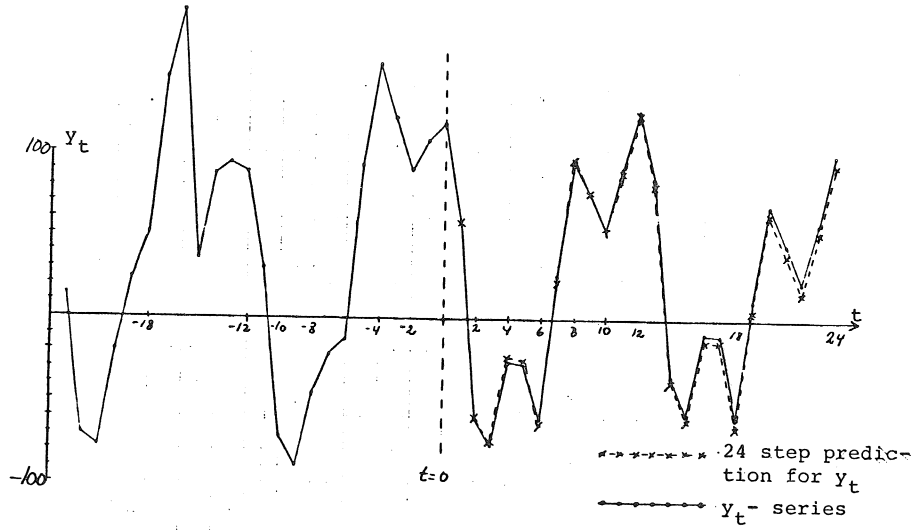
LAG = 6 CROSS-AUTO-CORRELATION-MATRIX

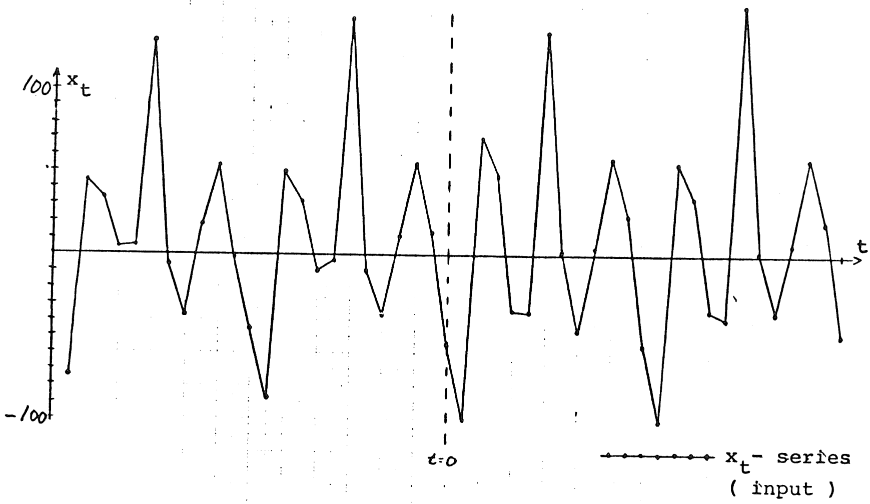
```
0.0686 -0.0159
0.0156 -0.0204
```

CONTRIBUTION TO CHI-SQUARE-TEST = 1.50

```
LAG =  7  CROSS-AUTO-CORRELATION-MATRIX

           0.0975 -0.0747
           0.0362  0.0728

              CONTRIBUTION TO CHI-SQUARE-TEST =        5.57

LAG =  8  CROSS-AUTO-CORRELATION-MATRIX

           0.1879 -0.0831
          -0.0689  0.0887

              CONTRIBUTION TO CHI-SQUARE-TEST =       13.08

LAG =  9  CROSS-AUTO-CORRELATION-MATRIX

           0.1505 -0.0642
          -0.0908 -0.0912

              CONTRIBUTION TO CHI-SQUARE-TEST =       11.22

LAG = 10  CROSS-AUTO-CORRELATION-MATRIX

           0.0884 -0.0417
           0.0995 -0.0628

              CONTRIBUTION TO CHI-SQUARE-TEST =        6.30

LAG = 11  CROSS-AUTO-CORRELATION-MATRIX

           0.1015 -0.0199
           0.0055 -0.0017

              CONTRIBUTION TO CHI-SQUARE-TEST =        2.82

LAG = 12  CROSS-AUTO-CORRELATION-MATRIX

          -0.0488 -0.0535
           0.0061  0.0891

              CONTRIBUTION TO CHI-SQUARE-TEST =        3.46


                 TOTAL CHI-SQUARE-TEST VALUE =        83.87

NOMINAL DEGREES OF FREEDOM IS    48    WHICH SHOULD BE CORRECTED :
SUBTRACT THE NUMBER OF PARAMETERS LINKING THE VARIABLES CONSIDERED
IN THE TEST.
FOR RESIDUALS FROM A FULL MARIMA(K,P,O,Q) MODEL SUBTRACT K*K*(P+Q);
IF THE MODEL IS REDUCED SUBTRACT THE NUMBER OF NONZERO PARAMETERS.
```

. o O o .





Seasonal modeling example - use of S A R M A X. The figure shows for the period until t=0 observed values for the series y_t. From then on the unknown future values are depicted. For that period, i.e. for t=1,2,..,24, a forecast is made. It is very close to the true values. Below the input x_t series is shown. It is assumed to be known also for the forecasting period.