ARIMA Modeling and Forecasting:

An Interactive Program

Based on IMSL Subroutine Package - I

by

Prem K. Goel* and A. Gregory Rocco

Technical Report #82-11

Department of Statistics

Purdue University

May 1982

*This work was supported in part by National Science Foundation under grant #MCS-8002336 at Purdue University.

ARIMA Modeling and Forecasting: An Interactive Program Based on IMSL Subroutine Package - I

by

Prem K. Goel and A. Gregory Rocco

Introduction. The program "arimaf" was written as a project for the course 'Stat 520: Time Series Analysis and Applications'. Since then it has been revised several times.

"arimaf" is designed to implement Box-Jenkins Methodology for fitting non-seasonal ARIMA models to a set of time series data. One can then obtain short-term forecasts from the fitted model. The program also lets the user plot the raw data and update the forecasts. It is written in Fortran 77. "arimaf" is user friendly and completely interactive. The user can easily route any part of the output to a disk file.

"arimaf" requires the IMSL Subroutines Package for data analysis. The program listed in Goel and Rocco (1982) is currently being run on a VAX-11/780 under the UNIX operating system. Very slight modifications will make it portable to any system with a Fortran 77 Compiler. The listing is available on request from the Department of Statistics, Purdue University, West Lafayette, IN 47907, or from one of the authors.

We shall now list the help file containing the 'Command Summary'. Some examples will be given to illustrate the usage and to explain the output.

"arimaf" help file

"arimaf" is a program designed for fitting ARIMA models (Auto Regressive, Integrated, Moving Average) to a given set of data. The fitted model can then be used for forecasting. There are 3 columns and one array which the user is concerned with, namely:

1. column of original data
2. column of data to be analyzed
3. column of residuals from last ARIMA fit
4. array containing forecasts

When the program is ready for a command it prompts the user with "arm>". At this point the user can input the desired command. Only the first three letters of the command are looked at, the rest of the line is available for comments. The commands are interactive and will prompt the user with the necessary questions; so as you go through the command summary below don't worry about the specifics of how to use each command.

Command Summary

help - Prints this file

input - Input from a user specified file into columns 1 and 2 given above. The Input file should be in the form of a single column of integers or floating point numbers (scientific notation is OK).

Note: The maximum length for a filename including the path name is 60 characters.

output - Output column 2 to disk (see note under input).

open - Open a disk file for output, this will direct most output to a disk file as well as the terminal (see note under input).

close - Close above output file.

file - Allows output to the open disk file to be turned on and off.

term - Allows long output (graphs and tables) to the terminal to be turned on and off.

print - Output the column of data being analyzed.

plot - Plot column of data being analyzed.

pwidth- Allows width of the plot to be changed (the default is 66).

plength- Allows the length of the plot to be changed (default is 19).

acf - Finds the autocorrelation function of the data.

pacf - Finds the partial autocorrelation function of the data.

transform - Transform the data in column 2 and store the transformed data back into column 2. The transformations which can be performed are: data(I)=log(data(I)+shift) where shift is equal to minimum value of data minus one and log is base E; or data(I)=data(I) ** exponent. If you wish to analyze the original data, it can be copied back into column 2 using the original command.

difference - Perform non-seasonal and seasonal difference on data in column 2 and store the results back in column 2.

arima - Find the maximum likelihood estimates of the parameters in the user identified ARIMA model and place the residuals in column 3 given above.

again - Reenter the ARIMA fitting routine using the last fit as a starting point.

smodel - Save the current ARIMA model in a file (see note under input).

rmodel - Retrieve an ARIMA model from a file (see note under input).

swap - Swaps the data and residual columns so that the residuals can be analyzed.

original - Copies the data from column 1 to column 2 for further analysis.

forecast - Use the last fitted model to forecast.

update - Used to update a forecast by doing the following:
1. Retrieve the data file (same as input command).
2. Retrieve the desired ARIMA model (same as rmodel).
3. Add new point to data column 2.
4. Make new forecast (same as forecast command).
5. Output new data file.
Note: any of these file operations can be avoided by entering just a carriage return when asked for the file name.

stat - Lists the number of data points in each column.

concatenate - Concatinate the forecasted values onto the data column.

note - A command which dosn't do anything (for comments).

exit - Close output disk file if open and exit.

note: In case of any problems please contact prof. Prem K. Goel (Purdue University Stat Dept.) or A. Gregory Rocco (Purdue University EE Grad. student '82).

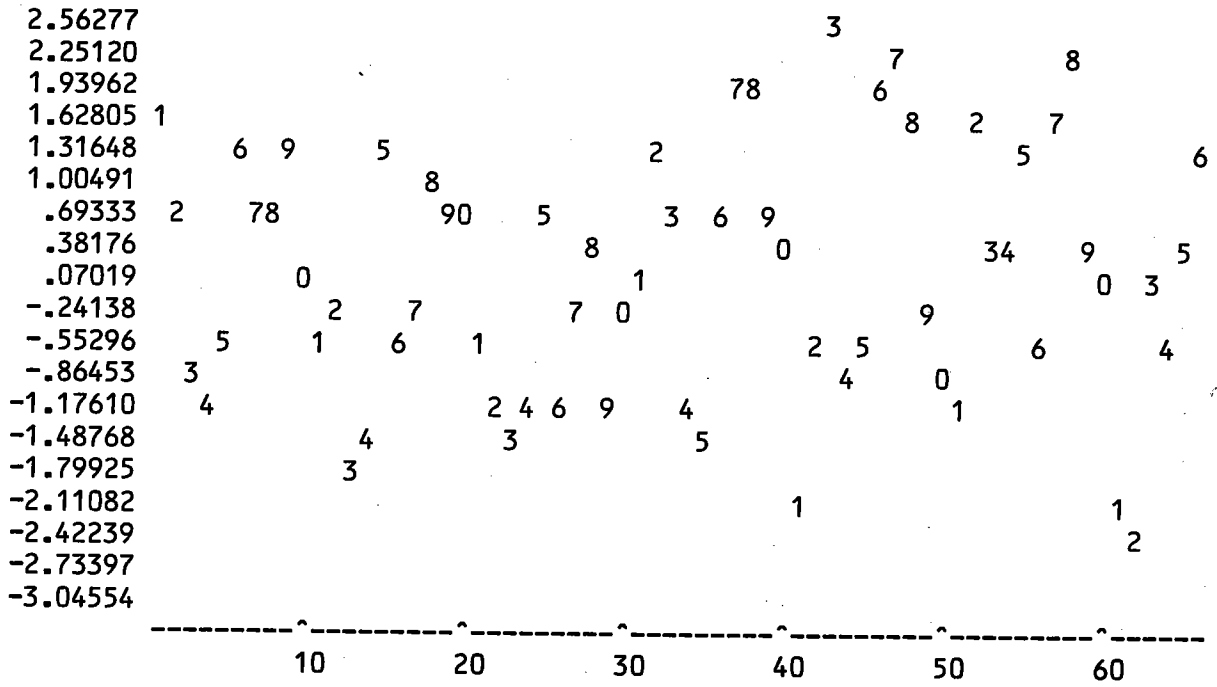
2. Some Illustrative Examples.

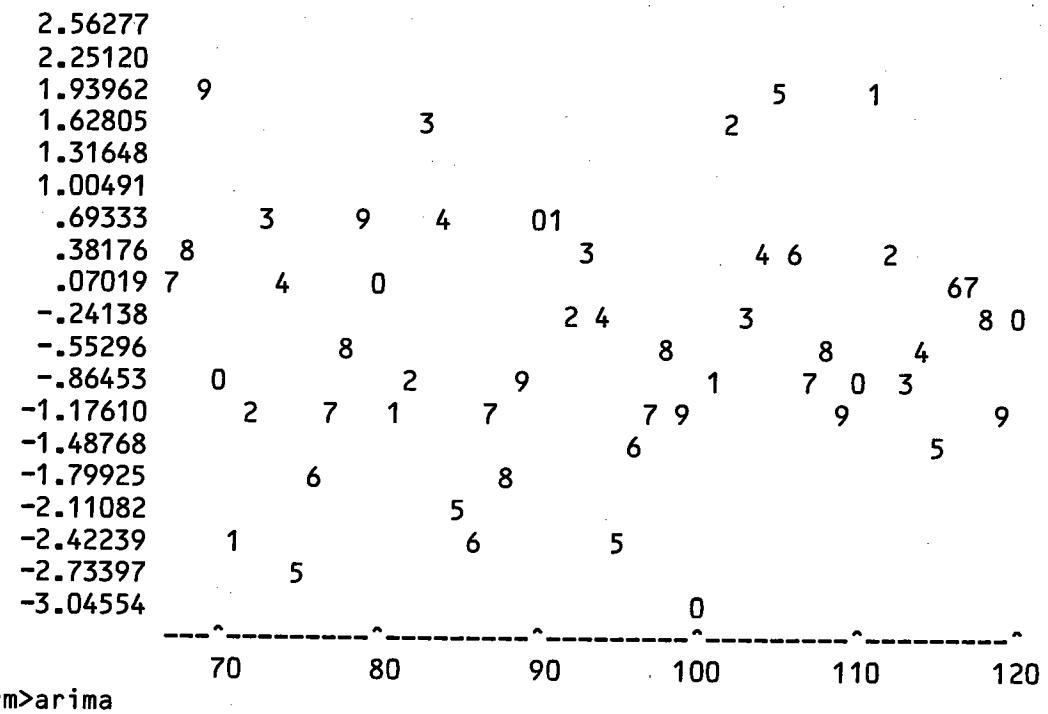
Example 1. The data used for this example is based on a simulation run from an MA(2) model, namely

$$Z_t = a_t - 0.7a_{t-1} + 0.2a_{t-2}$$

in the Box-Jenkins notation. Therefore we only illustrate the usage of the commands 'input', 'plot', 'arima' and 'exit'.

```
   Ec> arimaf
arimaf version 2.03

arm>input
input data from file? st31.dat
    120 data points read from file: st31.dat
        1.72204          .58956          -.85666          -1.16608 . . .

arm>plot
plot period?
```






```
arm>arima
input desired arima model? 0 0 2
would you like to input preliminary estimates? no

initial ma parameters:
       -.53388        .19593
```

```
 max likelihood est.of arima model ( 0, 0, 2)
 for    120 data points after differencing
     est. of moving average constant =          -.13741
     est. of white noise variance =            .99076

    MA terms:
        -.65530         .30809


 arm>exit
Ec>
```

Remarks. The fitted MA(2) model from this simulated data is

$$Z_t = -.137 + a_t - .655a_{t-1} + .308a_{t-2} ,$$

with an estimated variance $(a_t) = .991$.

Example 2. The time series analysed in this example is the 'Series A: Chemical Process Concentration Readings' from Box and Jenkins (1976). This run illustrates the usage of the commands 'input', 'acf', 'pacf', 'difference', 'original', 'arima' and 'again'.

```
Ec> arimaf
 arimaf version 2.03

 arm>input
 input data from file? bjsera.dat
     197 data points read from file: bjsera.dat
         17.00000          16.60000          16.30000          16.10000 . . .

 arm>acf
 number of terms? 20
```

```
time series  mean=          17.06244  and  variance=           .15859
acf
            -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1    .570                          xxxxxxxxxxxxxx
   2    .495                          xxxxxxxxxxxxx
   3    .398                          xxxxxxxxxx
   4    .356                          xxxxxxxxx
   5    .327                          xxxxxxxx
   6    .350                          xxxxxxxxx
   7    .392                          xxxxxxxxxx
   8    .322                          xxxxxxxx
   9    .304                          xxxxxxxx
  10    .255                          xxxxxxx
  11    .188                          xxxxxx
  12    .163                          xxxxx
  13    .195                          xxxxxx
  14    .236                          xxxxxxx
  15    .140                          xxxxx
  16    .180                          xxxxxx
  17    .197                          xxxxxx
  18    .202                          xxxxxx
  19    .140                          xxxxx
  20    .183                          xxxxxx

arm>pacf
number of terms? 20


time series  mean=          17.06244  and  variance=           .15859
pacf
            -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1    .570                          xxxxxxxxxxxxxx
   2    .252                          xxxxxxx
   3    .068                          xxx
   4    .069                          xxx
   5    .066                          xxx
   6    .124                          xxxx
   7    .156                          xxxxx
   8   -.032                         xx
   9    .010                          x
  10   -.016                          x
  11   -.073                        xxx
  12   -.021                         xx
  13    .061                           xxx
  14    .087                           xxx
  15   -.123                       xxxx
  16    .047                          xx
  17    .097                          xxx
  18    .068                          xxx
  19   -.071                        xxx
  20    .052                          xx
```

```
arm>difference
warning this routine will destroy the contents of the data column.
order of non-seasonal difference? 1
order of seasonal difference? 0

finding non-seasonal difference of order    1

arm>acf
number of terms? 20
```

time series mean= .00204 and variance= .13642

acf

```
            -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1   -.413                 xxxxxxxxxx
   2    .019                       x
   3   -.066                    xxx
   4   -.011                      x
   5   -.069                    xxx
   6   -.020                     xx
   7    .146                      xxxxx
   8   -.068                    xxx
   9    .037                      xx
  10    .021                      xx
  11   -.046                     xx
  12   -.064                    xxx
  13   -.011                      x
  14    .160                      xxxxx
  15   -.171                  xxxxx
  16    .033                      xx
  17    .014                      x
  18    .082                      xxx
  19   -.123                   xxxx
  20    .151                      xxxxx
```

```
arm>pacf
number of terms? 20
```

```
time series  mean=          .00204  and  variance=          .13642
pacf

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+----+
   1   -.413                xxxxxxxxxx
   2   -.183                     xxxxxx
   3   -.165                      xxxxx
   4   -.139                      xxxx
   5   -.193                     xxxxxx
   6   -.214                     xxxxxx
   7   -.002                          x
   8   -.047                        xx
   9   -.018                         x
  10    .041                         xx
  11   -.005                         x
  12   -.077                       xxx
  13   -.102                      xxxx
  14    .103                         xxxx
  15   -.085                       xxx
  16   -.127                      xxxx
  17   -.096                       xxx
  18    .045                         xx
  19   -.074                       xxx
  20    .091                         xxx
```

Remark. Since the 'difference' command destroyed the data to be analysed column, we use the 'original' command to copy the original data into column 2 again. One of the models suggested by Box and Jenkins for this series is ARIMA (0,1,1). By using the 'arima' command in different modes, we now illustrate the inputting of initial parameter values for this model.

```
arm>original

arm>arima
input desired arima model? 0 1 1
would you like to input preliminary estimates? yes
preliminary estimate for ma model? .2

initial ma parameters:
        .20000

max likelihood est.of arima model ( 0, 1, 1)
for    196 data points after differencing
    est. of moving average constant =          .00204
    est. of white noise variance =           .09981

   MA terms:
        .69450
```

```
arm>arima
input desired arima model? 0 1 1
would you like to input preliminary estimates? no
*** terminal error          (ier = 130) from imsl routine zsystm
*** terminal error          (ier = 130) from imsl routine ftmps
error= 130 while finding preliminary est. for MA model .1(s) assumed.

initial ma parameters:
        .10000

max likelihood est.of arima model ( 0, 1, 1)
for    196 data points after differencing
    est. of moving average constant =          .00204
    est. of white noise variance =           .09981

   MA terms:
        .69450
```

Remark. For this series, the subroutine for finding initial values of the MA coefficient bombs out, but the main program assumes the value of 0.1. The model fitting with an intial value of zero also works. Some other B-J methodology programs require that the initial values be non-zero.

```
arm>arima
input desired arima model? 0 1 1
would you like to input preliminary estimates? yes
preliminary estimate for ma model? 0

initial ma parameters:
        .00000

max likelihood est.of arima model ( 0, 1, 1)
for    196 data points after differencing
    est. of moving average constant =          .00204
    est. of white noise variance =           .09981

   MA terms:
        .69449
```

Remarks. As suggested by Box and Jenkins, we also fit ARMA (1,1) model to this series. Since the model parameters do not converge in 25 iterations of the maximum-likelihood estimation procedure during the first call to 'arima', one can use the command 'again' to restart the estimation with the current values of the parameters as the starting point. In the remainder of this example, the results of the model fitting for different initial values of the parameters are displayed. Clearly, the IMSL Subroutine seems to be working very well as far as the iterative procedure is concerned.

```
arm>arima
input desired arima model? 1 0 1
would you like to input preliminary estimates? yes
preliminary estimates for ar model? .2
preliminary estimate for ma model? .2

initial ar parameters:
        .20000

initial ma parameters:
        .20000
*** warning with fix error  (ier =  68) from imsl routine ftmxl
warning after    25 iterations the model parameters did not converge

max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.67318
    est. of white noise variance =           .09786

   AR terms:
        .90194

   MA terms:
        .55885


arm>again

initial ar parameters:
        .90194

initial ma parameters:
        .55885

max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.63077
    est. of white noise variance =           .09786

   AR terms:
        .90442

   MA terms:
        .56341


arm>arima
input desired arima model? 1 0 1
would you like to input preliminary estimates? no
*** terminal error          (ier = 130) from imsl routine zsystm
*** terminal error          (ier = 130) from imsl routine ftmps
error= 130 while finding preliminary est. for MA model .1(s) assumed.
```

```
initial ar parameters:
        .86828

initial ma parameters:
        .10000
*** warning with fix error  (ier =  68) from imsl routine ftmxl
warning after    25 iterations the model parameters did not converge

max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.73034
    est. of white noise variance =            .09786

   AR terms:
        .89859

   MA terms:
        .55302


arm>again

initial ar parameters:
        .89859

initial ma parameters:
        .55302

max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.62679
    est. of white noise variance =            .09786

   AR terms:
        .90466

   MA terms:
        .56403


arm>arima
input desired arima model? 1 0 1
would you like to input preliminary estimates? yes
preliminary estimates for ar model? 0
preliminary estimate for ma model? 0

initial ar parameters:
        .00000

initial ma parameters:
        .00000
*** warning with fix error  (ier =  68) from imsl routine ftmxl
warning after    25 iterations the model parameters did not converge
```

```
max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.74992
    est. of white noise variance =            .09787

   AR terms:
        .89744

   MA terms:
        .55195


arm>again

initial ar parameters:
        .89744

initial ma parameters:
        .55195

max likelihood est.of arima model ( 1, 0, 1)
for    197 data points after differencing
    est. of moving average constant =          1.62667
    est. of white noise variance =            .09786

   AR terms:
        .90466

   MA terms:
        .56405


arm>exit
Ec>
```

Example 3. The data set for this example is the updated version of the 'Unemployment Rate Data' given in Fuller (1976). This run presents the usage of the commands 'acf', 'pacf', 'arima', 'again', 'swap', 'smodel', 'original', 'forecast', 'stat', 'concatenate', 'plot', 'pwidth' and 'update'.

This example also illustrates the implementation of various steps in the Box-Jenkins Methodology for ARIMA model fitting, residual analysis, forecasting and updating of forecasts.

```
   Ec> arimaf
arimaf version 2.03

arm>input
input data from file? rate.dat
    132 data points read from file: rate.dat
        3.73000          3.67000          3.77000          3.83000 . . .

arm>acf
number of terms?



time series  mean=          5.21364  and  variance=          1.97228
acf
                -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
                  +----+----+----+----+----+----+----+----+----+----+
   1    .930                              xxxxxxxxxxxxxxxxxxxxxxx
   2    .796                              xxxxxxxxxxxxxxxxxxxx
   3    .645                              xxxxxxxxxxxxxxxx
   4    .515                              xxxxxxxxxxxxx
   5    .417                              xxxxxxxxxx
   6    .354                              xxxxxxxxx
   7    .314                              xxxxxxxx
   8    .283                              xxxxxxx
   9    .266                              xxxxxxx
  10    .255                              xxxxxx
  11    .243                              xxxxxx
  12    .230                              xxxxxx
  13    .224                              xxxxxx
  14    .223                              xxxxxx
  15    .233                              xxxxxx
  16    .241                              xxxxxx
  17    .231                              xxxxxx
  18    .202                              xxxxx
  19    .154                              xxxx
  20    .086                              xx
  21    .003                              x

arm>pacf
number of terms?
```

```
time series  mean=          5.21364  and  variance=          1.97228
pacf

           -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
             +----+----+----+----+----+----+----+----+----+----+
   1    .930                          xxxxxxxxxxxxxxxxxxxxxx
   2   -.502               xxxxxxxxxxxxx
   3    .006                          x
   4    .127                          xxxx
   5    .026                          xx
   6    .074                          xxx
   7   -.040                         xx
   8   -.007                          x
   9    .129                          xxxx
  10   -.039                         xx
  11   -.021                         xx
  12    .045                          xx
  13    .061                          xxx
  14    .034                          xx
  15    .070                          xxx
  16   -.077                        xxx
  17   -.098                        xxx
  18    .004                          x
  19   -.083                        xxx
  20   -.140                       xxxxx
  21   -.101                        xxxx
```

Remarks. The acf and the pacf of the rate data suggest an AR(2) model as the first choice at the identification stage. Therefore, we fit the AR(2) model. Fuller (1976) also fitted this model for a part of this series. The iterative procedure required more than 25 iterations for model parameters convergence.

```
arm>arima
input desired arima model? 2 0 0
would you like to input preliminary estimates? no

initial ar parameters:
      1.39586       -.50168
*** warning with fix error  (ier =  68) from imsl routine ftmxl
warning after    25 iterations the model parameters did not converge

max likelihood est.of arima model ( 2, 0, 0)
for    132 data points after differencing
    est. of moving average constant =          .47390
    est. of white noise variance =           .13746

   AR terms:
      1.50795       -.59885


arm>again
```

```
initial ar parameters:
       1.50795       -.59885

max likelihood est.of arima model ( 2, 0, 0)
for    132 data points after differencing
    est. of moving average constant =          .47593
    est. of white noise variance =           .13744

   AR terms:
       1.51587       -.60716
```

Remark. To perform an appropriate residual analysis, we 'swap' the original data with the residuals of the fitted AR(2) model.

```
arm>swap

arm>acf
number of terms? 20


time series  mean=          .01098  and  variance=          .13732
acf
            -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
              +----+----+----+----+----+----+----+----+----+----+
   1    .065                          xxx
   2   -.011                          x
   3   -.053                         xx
   4   -.027                         xx
   5   -.070                        xxx
   6    .087                          xxx
   7    .188                          xxxxxx
   8   -.116                       xxxx
   9    .033                          xx
  10    .074                          xxx
  11    .113                          xxxx
  12   -.091                        xxx
  13    .057                          xx
  14   -.086                        xxx
  15    .009                          x
  16    .084                          xxx
  17    .055                          xx
  18    .123                          xxxx
  19    .024                          xx
  20    .004                          x

arm>pacf
number of terms? 20
```

```
time series  mean=          .01098  and  variance=          .13732
pacf

           -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
             +----+----+----+----+----+----+----+----+----+----+
   1    .065                          xxx
   2   -.015                          x
   3   -.052                         xx
   4   -.020                         xx
   5   -.068                        xxx
   6    .094                          xxx
   7    .175                          xxxxx
   8   -.152                       xxxxx
   9    .063                          xxx
  10    .090                          xxx
  11    .113                          xxxx
  12   -.099                        xxx
  13    .033                          xx
  14   -.087                        xxx
  15    .081                          xxx
  16    .050                          xx
  17   -.016                          x
  18    .129                          xxxx
  19    .060                          xxx
  20   -.040                         xx
```

Remarks. Since the residuals of the fitted AR(2) model do not provide any evidence of departure from the white noise hypothesis, this model seems to fit quite well to the enlarged data set as it did for Fuller's data.

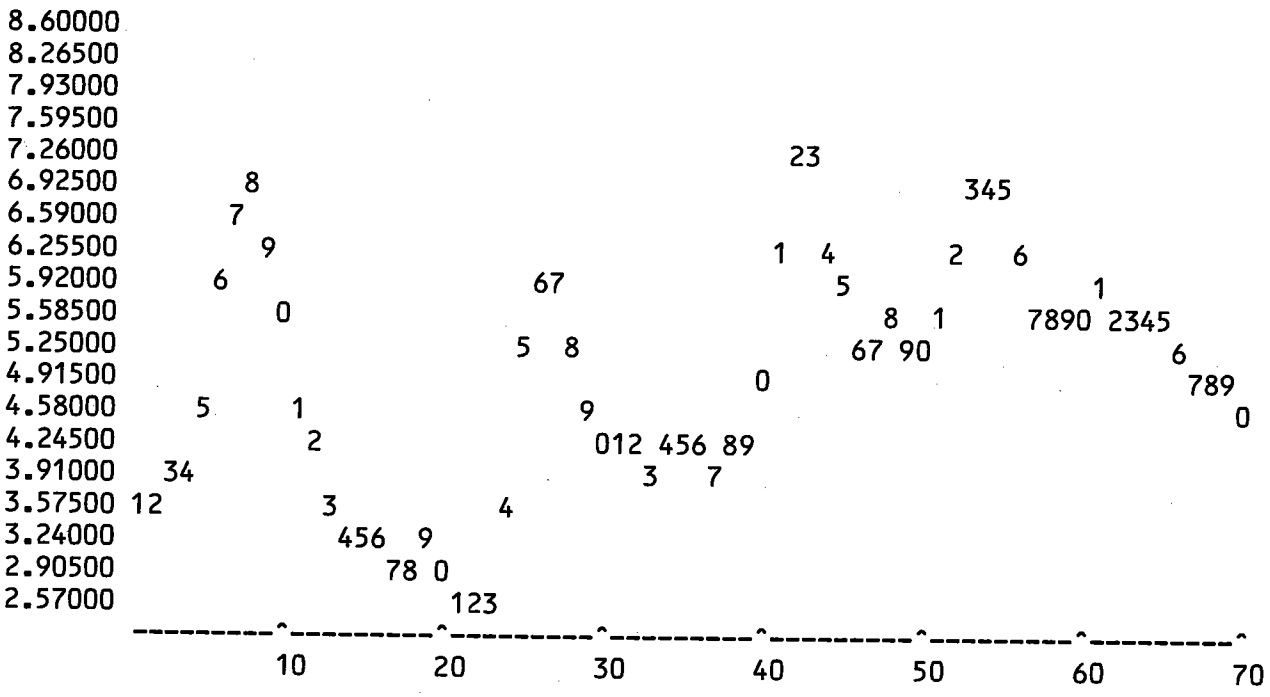
We can save the model information on a disk file, by using the command 'smodel', for future updating of forecasts.

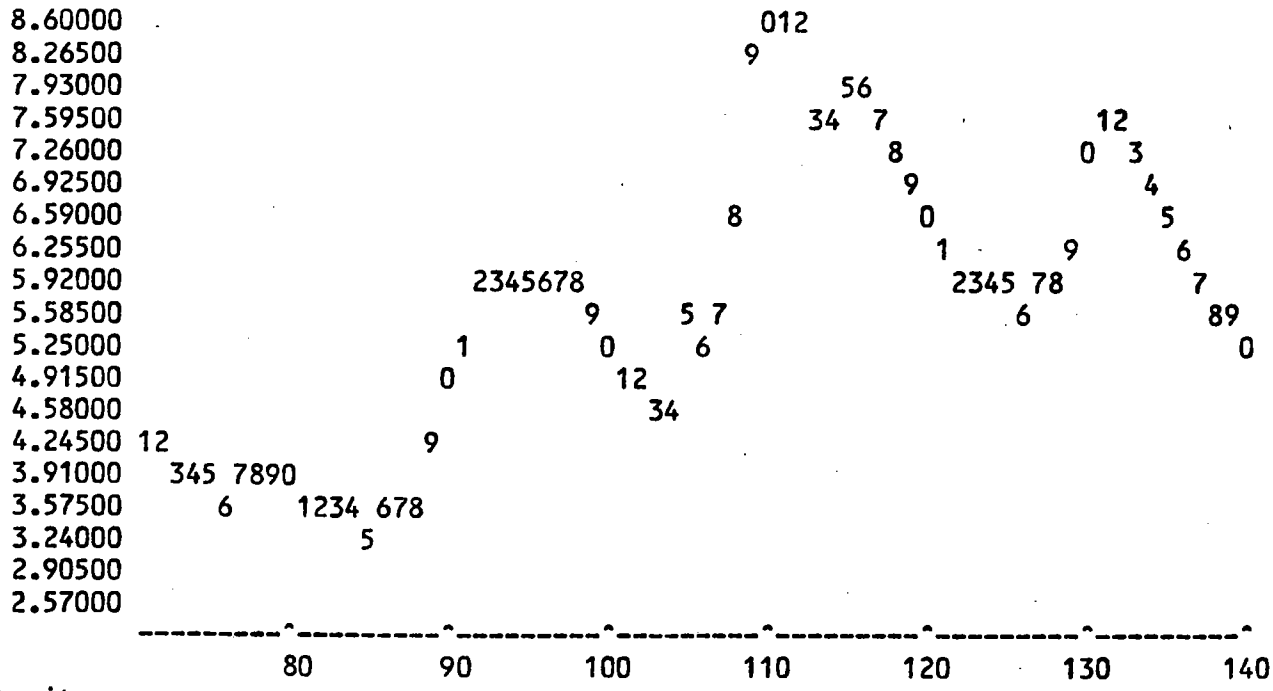
For forecasting from the fitted model, first we have to use the command 'original' to copy the data into column 2. The 'concatenate' command adds the forecasts to the column 2, which enables us to 'plot' the actual data along with the forecasts.

```
arm>smodel
write model to file? rate.arm
model written to file: rate.arm

arm>original

arm>forcast
forecast; alpha?
lead time? 8
```

lead time	psi weights	forecast	std. error 5.00 %
1	1.51587	7.27307	.72662
2	1.69071	6.94729	1.31954
3	1.64252	6.59124	1.80288
4	1.46333	6.24930	2.16213
5	1.22095	5.94715	2.40943
6	.96233	5.69674	2.56757
7	.71746	5.50060	2.66108
8	.50329	5.35531	2.71167

```
arm>concatenate

arm>pwidth
plot width? 70

arm>plot
plot period?
```






```
arm>exit
```

Example 3 (continued). To update the forecasts based on the last fitted model after one gets a new observation of the series, we need to use the command 'update'. The output below illustrates the usage of this command.

```
Ec> arimaf
 arimaf version 2.03

 arm>update
 input data from file? rate.dat
     132 data points read from file: rate.dat
          3.73000          3.67000          3.77000          3.83000 . . .

 retrieve model from file? rate.arm
 arima model retrieved from file: rate.arm

 max likelihood est.of arima model ( 2, 0, 0)
 for   132 data points after differencing
     est. of moving average constant =          .47593
     est. of white noise variance =           .13744

    AR terms:
        1.51587       -.60716


 new observation to be added to data? 7.35

 forecast; alpha?.05
 lead time? 8
```

lead time	psi weights	forecast	std. error 5.00 %
1	1.51587	7.06391	.72400
2	1.69071	6.72130	1.31478
3	1.64252	6.37566	1.79639
4	1.46333	6.05972	2.15434
5	1.22095	5.79066	2.40075
6	.96233	5.57463	2.55832
7	.71746	5.41051	2.65149
8	.50329	5.29289	2.70190

```
output data to file? rate2.dat
    133 data points output to file: rate2.dat


arm>exit
Ec>
```

References

1. Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis: Forcasting and Control, Revised Edition, San Francisco: Holden Day.

2. Fuller, Wayne A. (1976) Introduction to Statistical Time Series, New York: Wiley.

3. Goel, Prem K. and Rocco, A. G. (1982) ARIMA Modelling and Forecasting: An Interactive Program Based on IMSL Subroutines Package - II, Technical Report #82-12, Department of Statistics, Purdue University.

4. IMSL (1980) IMSL Subroutines Package, 8th edition, Houston: International Mathematical and Statistical Library.