USING MICROCOMPUTERS IN TEACHING TIME SERIES ANALYSIS

D. Alan Jones
University College of Wales
Aberystwyth, Wales. U.K.

1. Time Series Analysis

Time Series Analysis is a particular area of Statistics that has seen remarkable progress in the last 20 years. There has been increased activity and interest in both the theory and practice of the subject that has led in some sense to a unification of methodologies that existed previously. Early work in the field can be regarded as development associated with two main schools broadly described as the exact and the non-exact sciences.

There is a distinction between these two approaches, clearly linked to the area of application. The physicist wanted to interpret variation over time in terms of periodic components; the economist sought to first iron out what was regarded as "stray" variability and then project into the future.

A modern approach to many areas has the two components:

Theory: Postulate a class of models; investigate their properties ensuring the class is wide enough to cover and allow (most) observable phenomena!

Data analysis: Apply known and possibly new descriptive and summary techniques to adequately describe the data.

and the two components meet in the guise of "model fitting" whereby the observed summary statistics and characteristics are in some way matched to a set of theoretical characteristics exhibited by a member of the class of models. This model then lies at the heart of any inferences that need to be drawn, be they in the "explanation" or "forecasting" objective areas. Such a structured approach now exists in most areas of Statistics and in the particular sphere of Time Series Analysis underlies the 1960's approach of Box and Jenkins as exemplified in their book (Box & Jenkins,1970).

2. The role of computers

It is clearly not pure coincidence that the growth in Time Series Analysis has occurred at the same time as the growth of computing availability. Software of suitable quality was a little slow in appearing at first. but recent years have seen the introduction of many new and re-vamped statistical packages, and most of these nowadays contain quite extensive Time Series routines. This is true of established packages such as BMDP, SPSS, SAS, MINITAB. etc. and the growth of availability and usage continues with the gradual introduction of microcomputer versions of these packages. The earlier comments concerning the two areas of origin also apply when one looks at some of these packages; MINITAB for example is

clearly modern in its approach in that it its inbuilt routines almost all belong to the model-data-fit philosophy. Some of the newer and more specialised packages however clearly show parentage in the "smooth and forecast" camp with Box-Jenkins routines appearing alongside the ad hoc forecasting routines of old (examples being RATS and TSP).

The favoured package for teaching at Aberystwyth is the very powerful STATGRAPHICS. Available for IBM microcomputers and their clones, this menu driven package covers an extremely broad spectrum. Time Series is just one of its 22 sub-menus (another being Smoothing) and its user friendly interactive mode of operation allows students to perform standard analyses on data at their pace and in their order. Box-Jenkins Modelling is one of the 15 options offered in the Time Series Analysis sub-menu, and Fig.1 shows the display at the beginning of this routine. The various parameter, model specification and accuracy options can be modified by simple cursor movements, and the key at the bottom of the display shows the choice of next step – thus a single key press will display the autocorrelation function, while another will fit the model currently specified. Ease of operation is a boon.

Figure 1

```
-------------------------------------
GENERAL UNIVARIATE ARIMA MODEL FITTING
-------------------------------------

                Output time series: illegit

Order of nonseasonal diff.:  0          Order of seasonal diff.: 0
Constant contained in model: YES        Length of seasonality:   12
Order of nonseasonal AR factor: 2       Order of seasonal AR factor: 0
Order of nonseasonal MA factor: 0       Order of seasonal MA factor: 0

Maximum lag for acf plots: 24           Maximum lag for PACF plots:    16
Lags for chi-square test:  20           Number of forecasts desired:   10

Backforecasting:       YES
Maximum iterations:    25
Stopping criterion 1:  .00010
Stopping criterion 2:  .00100


PRESS ENTER TO UPDATE PANEL, THEN DESIRED PF KEY.

*1HELP    2SERIES 3ACF     4PACF    5ESTIM  6RACF    7RPACF  8INTPER 9FORCST 10QUIT
PRINT     WED AUG 6  1986  06:54:00 PM  VERSION 1.1                        REC:OFF
```

3. The Aberystwyth Microcomputer Laboratory Project

The combined Mathematics departments at Aberystwyth are currently engaged in a collaborative project geared towards the use of microcomputers in teaching. As part of this project, software is being developed for use in teaching a variety of mathematical and statistical topics. It is recognised that any such software should possess the following virtues:

(i) it should be easy to use;

(ii) it should be versatile, and any collection of software routines should have a common appearance;

(iii) the whole system should be adaptable and flexible to new and local requirements;

(iv) the whole system should not be too dependent on the particular hardware used, to allow for use in a wider mathematical community which can only lead to better testing and ultimately to improvement.

Even though Statistics teachers are well accustomed to using computational aids, experiences with STATGRAPHICS have impressed. Students readily take to the menu-driven system and staff welcome the facility of adding their own routines, a very unusual feature.

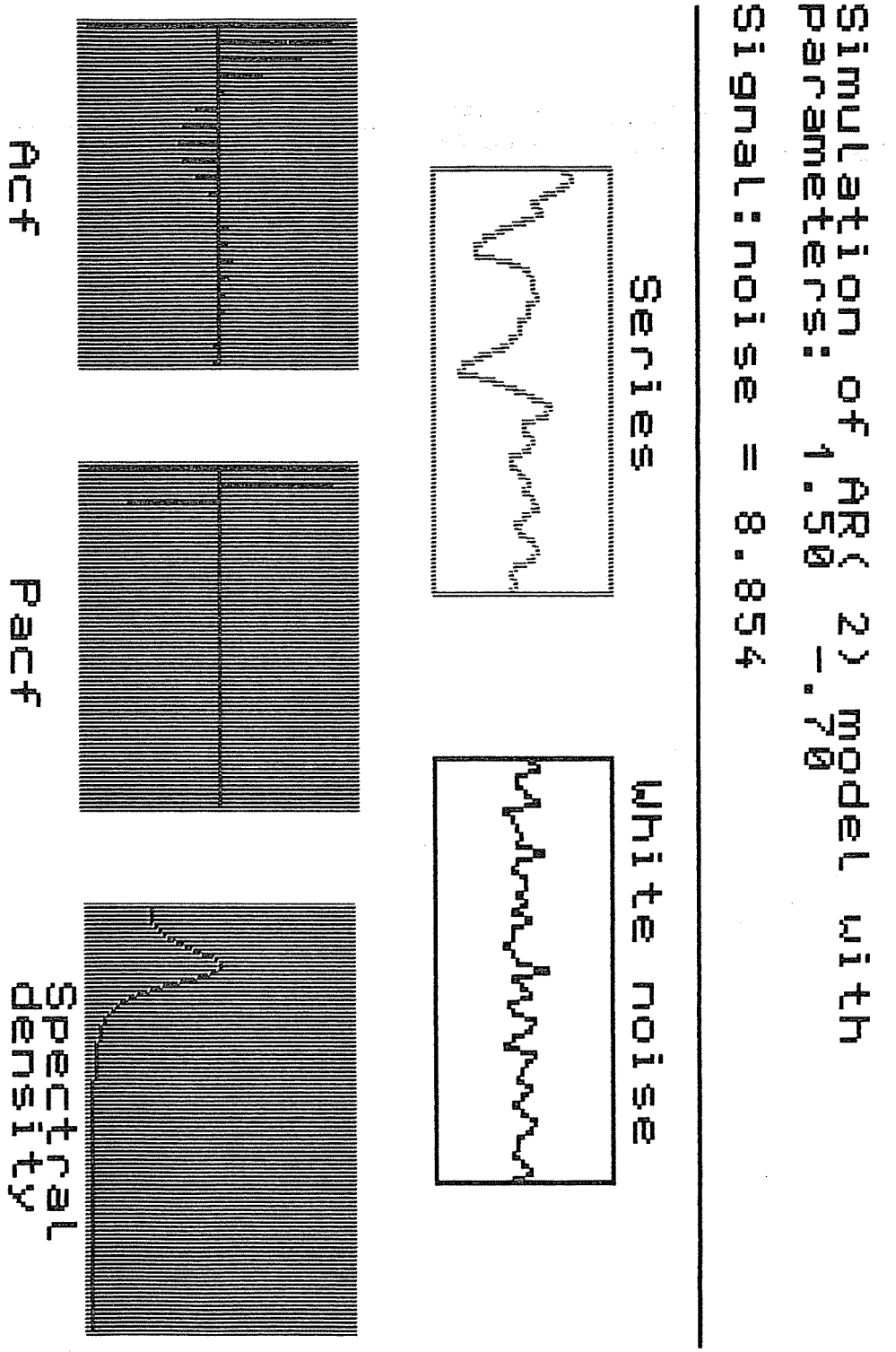
Packages invariably contain routines to perform certain analyses and carry out given operations on data. Thus the market provides reasonably well for the data analysis stage of the approach referred to at the end of section 1, as well as for the linking stage. That is, we can summarise data and, once we have decided upon a model, we can fit suitable parameter values. There is a void in the area of model teaching, the general process of familiarisation with models and what models mean. Time Series is perhaps an area where this problem is more acute than elsewhere. The rapid growth and development of the subject has led to a plethora of available models and presents to the newcomer a bewildering choice. There is therefore a need for systems that will allow a user to specify a given model and in return to receive:

(i) a description or plot of the corresponding summary measures;

(ii) a simulated realisation of the series.

In such a package in the course of development, summary functions are shown on the screen below a dynamic simulation of the process and the underlying white noise; these scroll across the screen allowing the user to observe a range of possible patterns generated by the ;model. This provides to be a useful device for the student, in particular demonstrating the difficulty of discriminating between many models of the same type but with slight differences in parameter values. The full version of this software allows for the simulation of any stationary ARMA model, and also permits investigation of the effects of non-stationarity as exhibited by the ARIMA class, and by the introduction of intervention and deterministic terms into the model.

Figure 2


Simulation of AR(2) model with
Parameters: 1.50 -.70
Signal:noise = 8.85%
Series
White noise
ACF
PACF
Spectral
Density

Whilst primarily aimed at aiding model familiarisation, there are already indications that some individuals use simulations to assist (or as a check) in the model identification stage of analysis, by simple visual comparison of the patterns generated.

Reference

Box, G.E.P., & Jenkins, G.M. (1970). Time series analysis: Forecasting and control. San Francisco: Holden-Day