FIRST REPORT OF THE TIME SERIES PROJECT

Clive W. J. Granger*

*On leave from the University of Nottingham, England

# MEMBERS OF THE TIME SERIES PROJECT

The following are connected with the project to differing extents:

Professor Oskar Morgenstern

Professor John Tukey

* * * *

Dr. Michio Hatanaka

Dr. Clive Granger

* * * *

Dr. S. N. Afriat

Mrs. Dorothy Green

Mr. Herman Karreman

First Report of Time Series Project

## 1. Introduction.   Modern Tools

The methods of analysis for time series and the understanding of the philosophy underlying these methods by statisticians have made considerable progress in the last fifteen years; but the significance and power of these changes have been for the most part uncomprehended by casual observers of series in general, and the majority of economists in particular.   To indicate to non-experts in the field in what way the new approaches are better and more general than the traditional approach, can probably be best performed by first reviewing the now classical methods and then by showing both the limitations and possibilities of extension of these methods.

Even before the turn of the century, it had become obvious to statisticians that most time series occurring in the actual world were not merely strings of independent data.  The series were too "smooth" to be independent; that is, the fluctuations from one side of the mean to the other were not sufficiently violent.  The initial explanation of this smoothness was undoubtedly conditioned by the type of long time series then most readily available, i.e. astronomical data.  To the pioneer in the subject, Sir Arthur Schuster, looking at sun-spot data which has as its most obvious feature an apparent periodicity of eleven years, the idea of some inherent periodic function causing the smoothness was the only possible one.  That is, it was felt that if one could determine the amplitude and period of a sine curve sufficiently accurately and subtract this from the data then the remainder ought to be an

independent series. When in fact this was done and the remainder was still found to be somewhat too smooth, it was natural to re-use the current predominant idea of the cause of the smoothness and to look for yet further sine curves to fit to the data. It was to this end that Schuster proposed his ingenious periodogram which was subsequently used by other statisticians on all time series, whether or not they showed the kind of periodicity to be found in sun-spot data. In the periodogram the function $a_p \sin(pt + \theta)$ was fitted to the data and the estimate of $\frac{1}{2}a_p^2$ plotted against p. If sharp, high peaks were found at a certain value of p then this was taken to be the true frequency ($\propto 1/$period) of the "smoothing function," or at least one important component of it. In practice the diagram usually had many imposing looking peaks and so "tests of significance" were proposed to see whether these peaks were really due to the data containing a periodic function or if they could be explained merely as sampling fluctuations. All these tests, however, and in fact the whole philosophy of the era, worked under the assumption that if one could accurately estimate the amplitude, period (or frequency) and phase of a (small) finite number of sine terms and subtract these from the data then all that remained would be an independent series. The model upon which this assumption is based may be called the "linear cyclic model," "linear" here merely implying a "sum of" the sine terms.

The periodogram approach is closely related to the method of Fourier (or harmonic) analysis as each is attempting to decide which periods appear to be the most important, the main difference being that in the former one can choose the periods to be examined whereas in the

independent series.  When in fact this was done and the remainder was
still found to be somewhat too smooth, it was natural to re-use the
current predominant idea of the cause of the smoothness and to look
for yet further sine curves to fit to the data.  It was to this end
that Schuster proposed his ingenious periodogram which was subsequently
used by other statisticians on all time series, whether or not they
showed the kind of periodicity to be found in sun-spot data.  In the
periodogram the function $a_p \sin(pt + \theta)$ was fitted to the data and
the estimate of $\frac{1}{2}a_p^2$ plotted against p.  If sharp, high peaks were
found at a certain value of  p  then this was taken to be the true
frequency $(\propto 1/\text{period})$ of the "smoothing function," or at least one
important component of it.  In practice the diagram usually had many
imposing looking peaks and so "tests of significance" were proposed
to see whether these peaks were really due to the data containing a
periodic function or if they could be explained merely as sampling
fluctuations.  All these tests, however, and in fact the whole philo-
sophy of the era, worked under the assumption that if one could accu-
rately estimate the amplitude, period (or frequency) and phase of a
(small) finite number of sine terms and subtract these from the data
then all that remained would be an independent series.  The model upon
which this assumption is based may be called the "linear cyclic model,"
"linear" here merely implying a "sum of" the sine terms.

The periodogram approach is closely related to the method
of Fourier (or harmonic) analysis as each is attempting to decide which
periods appear to be the most important, the main difference being that
in the former one can choose the periods to be examined whereas in the

latter they are fully determined from the length of the data available.
If there is a period, then each is equally likely to give an indica-
tion of this fact but the periodogram is the more flexible and, as
will be shown later when the spectrum is discussed, the more natural.
Should the data contain almost periodic functions in the strict mathe-
matical sense, then again the periodogram is a better method as such
a function can always be represented by a finite sum of sine and co-
sine terms but cannot be represented by a finite number of harmonic
terms. Almost periodic data would produce a 'smearing' of the periodo-
gram peak about the average value for the period and thus would be more
difficult to detect, than strictly linear cyclic data.

The reason for the form of the periodogram, i.e. the plotting
of $\frac{1}{2}$ est.$a_p^2$ against p, is interesting and is also of importance in
explaining certain of the later developments. If the linear cyclic
model is actually the correct one for the data, then the total variance
of the data may be represented by

$$\sigma_y^2 = \frac{1}{2}a_b^2 + \frac{1}{2}a_c^2 + \dots + \frac{1}{2}a_d^2 + \sigma_\epsilon^2$$

where b,c, ... , d are the correct (and only) frequencies present and
$\sigma_\epsilon^2$ is the variance of the random term. Thus if a sine term,
$a_p\sin(pt + \theta)$ could be accurately fitted to the data $y_t$, (t=1,...,n)
$\frac{1}{2}$est.$a_p^2 \simeq 0$ if $p \neq a$, b,...,d (or at least is small), whereas if
p=a, b,..., d then this estimate is not insignificant in size. Another
way of looking at this, and later it will be seen to be a natural one,
is that $a_p\sin(pt + \theta)$ contributes a finite, significant part of the
overall variance if, and only if, p is one of the actual frequencies

A     B     C     G     H

$y_o$     $y_1$     $y_2$ ....     $y_{n-1}$     $y_n$
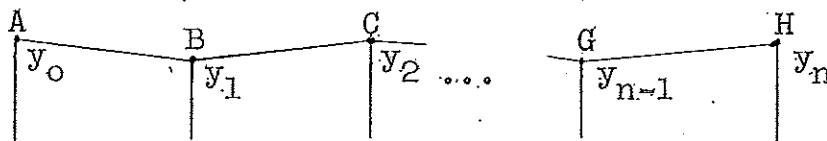
then obviously the length of the line ABC....GH will be a measure of the "smoothness" of the data. The square of this length is given by

$$K = \sum_{i=1}^{n} (y_i - y_{i-1})^2 + nH$$

where H is the distance between the readings.

Now the expected value of K is $2\sigma_y^2(1-\rho) + nH$ where $\rho$ is the correlation between $y_i$ and $y_{i-1}$ (assumed to be the same for all i) and so the nearer $\rho$ is to 1, the smaller the mean of K. With Yule's and Slutsky's models $\rho$ can be made positive and large by suitable choice of the parameters and so the smoothness of the series may be assured.

For several years the linear cyclic and the linear regressive models were treated as being alternative methods of explaining the underlying generating process of a time series and, in fact, were usually treated as being the only alternatives. It was not until after the second world war that it was realized that both models were specific cases of a very wide class of processes. Such processes are called stationary processes and have the properties (amongst others) that they are entirely trend-free and that the correlation between any two terms $y_i$ and $y_j$ is dependent only on the distance between the two terms ($|i - j|$) rather than on the actual positions of the terms.

It was shown by Cramer in 1946[1] that all such processes could be represented as a complete generalization of the linear cyclic pro-

---

[1] Cramer, H., Ark. Mat. Astr. Fys. 28B (1942) No. 12.

cess but now with <u>all</u> frequencies being present rather than a finite number and with the amplitudes and phases being random variates[1], i.e.

$$x_t = \int_{-\pi}^{\pi} \sin(\omega t + \theta(\omega)) \cdot y(\omega) d\omega \quad .$$

Thus now <u>any</u> fitted sine or cosine curve is likely to make a non-zero contribution to the over-all variance although some bands of frequencies will be more important in this aspect than other bands. Concerning this, let $f(\omega)$ be the true contribution to the total variance attributable to frequency $\omega$, then the diagram of $f(\omega)$ plotted against $\omega$ is termed the <u>power spectral density function</u>. For the linear cyclic model this spectral function would be everywhere zero except at the frequencies present, i.e. $\omega$ = a, b, c, ... , d and at these frequencies $f(\omega)$ would be $\frac{1}{2}a_a^2$, $\frac{1}{2}a_b^2$, .... , $\frac{1}{2}a_d^2$ respectively, and the spectrum for an independent, random series is just a constant for all $\omega$. On the other hand, a simple autoregressive model may have a spectrum such as
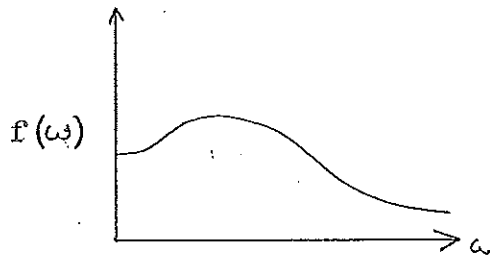


It has been shown that all stationary data will have spectra that are mixtures of these discrete and smooth cases. Thus, if the

---

[1] There are certain problems about how the integral is defined in this situation and the random variables $\theta(\omega)$, $y(\omega)$ have certain properties regarding their dependence.

spectrum one estimates is like:
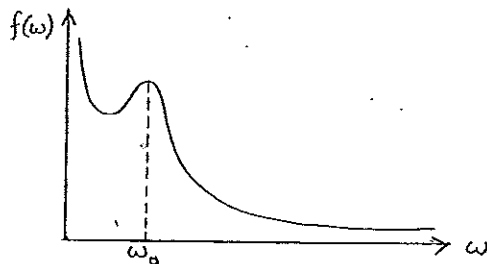


then one could guess that the main factors, at least, of the genera-
ting process of the series are a short auto-regressive, plus a
periodic term corresponding to the peak at $\omega_o$ with possibly also
a random term added. Such a model would be a sensible model in
that it would account for the majority of the variance found for
each frequency but this does not mean that it is necessarily the
correct generating process.

It is also known that if $\rho_j$ is the $j^{th}$ autocorrelation
coefficient, i.e. correlation between $x_t$ and $x_{t+j}$, then

$$\rho_j = \int_0^\infty \cos j\omega f(\omega) d\omega \quad .$$

If the spectrum were known perfectly, then any true periods
could be picked out and the relative importance of the various fre-
quencies is indicated. However, it is not possible to determine
exactly the true spectrum although many reasonably good estimates
are available.

It may here be noted that the periodogram is the only
available unbiased estimate of the spectrum but it is a bad esti-
mate as it has large variability and this variability does not
decrease as the amount of data available is increased. This par-

tially explains both the importance of the periodogram and also the reason for its rejection.

The idea of the spectrum becomes even more important and usually more useful when one is studying the interactions between two time series rather than the behavior of just one series. Suppose that both of the series have zero mean and denote the cross-variance with lag p (i.e., mean or expected value of $x_i y_{i+p}$) by $R_{xy}(p)$, then

$$R_{xy}(p) = R_{yx}(-p) = \int_0^\infty \cos p\omega \, c(\omega) d\omega + \int_0^\infty \sin p\omega \, q(\omega) d\omega$$

where $c(\omega)$ and $q(\omega)$ are termed the cospectral density function and quadrature spectral density function respectively. We may estimate $c(\omega)$ and $q(\omega)$ for a set of values of $\omega$ and, of course, may also estimate the power spectra of x and y, i.e. $f_x(\omega)$ , $f_y(\omega)$ , and thus may form what is called the coherency between the two series at the frequency $\omega$ , i.e.

$$C(\omega) = \frac{c^2(\omega) + q^2(\omega)}{f_x(\omega) \cdot f_y(\omega)} \ .$$

The coherency is a measure of the degree of dependence between the two series. In fact, a recent Russian paper[1] has shown that the average amount of information per unit of time concerning one series derivable from the other is given by

$$i(x,y) = -\tfrac{1}{2} \int_0^\infty \log(1-C(\omega)) d\omega \ .$$

---

[1] Gel'fand, I.M. and Yagham, A.M., American Math. Soc. Translations, Ser. 2, Vol. 12, p. 99.

Similarly, the argument of the cross-spectrum, given by

$$\phi(\omega) = \tan^{-1}\left(\frac{q(\omega)}{c(\omega)}\right)$$

is a measure of how much the two series are in phase at frequency $\omega$ .

The interpretation and intuitive meanings of the estimates of the spectra, coherency and angle are not easy to grasp, but, in an experienced user's hands, are powerful methods of investigating any "microscopic" relations and occurrences in the data. For example, if the angles appear to be lying about an increasing straight line (reduced modulo $360^{\circ}$), then this implies that one series is continually lagged behind the other and by a constant amount. On the other hand, if the angle seems to be moving periodically around a constant this may indicate that the longer the period (and shorter the frequency), the longer the lag between the series.


## 2. Non-Stationarity

As stated above, these new methods cannot strictly be used unless the data is stationary, and in particular there should be no trends in the mean, variance and covariances. However, it is visually clear that few, if any, economic series are stationary and that many have both a trend in mean and variance at the very least.

When the question of non-stationarity arises there are two main problems to be considered: (i) are we able to test whether the data is stationary and (ii) are we able to remove any non-stationarity

that is found. At the present stage of research, the answer to both of these questions is a slightly weak 'yes.'

A test for a continually increasing (or continually decreasing) trend in mean is available.[1] If $\bar{x}_1$ represents the mean for the first third of the series and $\bar{x}_2$ is the mean for the last third, then under unrestrictive conditions and with a sufficiently long series, $\bar{x}_1 - \bar{x}_2$ is found to be distributed normally under the hypothesis of stationarity with zero mean and variance $\frac{6K\sigma^2}{n}$ where n is the total length of the series and K is a constant which changes from series to series but which may be estimated. Using this information, confidence limits can be easily placed on $\bar{x}_1 - \bar{x}_2$. A similar test exists for testing against trend in variance and this could be extended to deal with autocovariances but the computations now become extremely laborious and the theory of the test less satisfactory.

If some non-stationary element is found to be present in the mean, then one of the new filtering techniques can be used to remove most of it. Suppose that this effect can be well represented by a polynomial, then this affects the spectrum only at zero frequency as the Fourier transformation of a polynomial is a constant. Thus, if a method could be designed that removed zero or very low frequencies whilst leaving the rest unaltered, we could remove the trend in mean. A weighted moving average with suitably chosen weights has such a property and a suitable filter has been designed for

---

[1] Tests for stationarity are fully discussed in "Tests of Discrimination and Stationarity for Time Series", C.W.J. Granger, Unpublished Thesis, University of Nottingham, England, 1959.

economic series. Using this point of view, the reason for the suggestion that a moving average introduces spurious periodic terms becomes easily explained. A badly designed filter will often remove the part required to be removed but at the cost of over-representing certain other frequencies. In many ways the ordinary moving average is a bad filter and such an over-representation or magnification of some frequencies will occur and so the 'spurious periods' arise.

When any trend in mean has been removed, a trend in variance can be removed by taking logarithms and again filtering out low frequencies.

One unfortunate aspect of this process is that the 'cycles' that have been 'found' by earlier researchers in economic time series are of low frequency and so special devices have had to be developed to investigate these cycles. The method suggested by Professor Tukey is essentially to 'stretch-out' the data, apply a filter and then use the spectral approach.

If, after removing any trends in mean and variance the series is still thought to be non-stationary in that the covariances may be changing with time, then we lack a theory of how to deal with such a situation. Any spectrum estimated in such circumstances is really a mean over time of a changing spectrum and it cannot be used unless specific assumptions are made about how the spectrum is changing.

3. Model Fitting

Parallel with the spectral approach to the analysis of time series is that of model-fitting. Although the usefulness of this latter approach is arguable for time series in general, it is usually conceded that it is intuitively reasonable to expect that the important aspects of the behavior of an economic series can be well explained by a model based on the auto-regressive model. Various modifications of the basic auto-regressive model can be suggested to account for the non-stationarity; we may have: data $= A(t) x_t + B(t)$ where $A(t)$, $B(t)$ are functions changing with time and $x_t$ is a stationary series, or data generated by

$$a_0(t)x_t + a_1(t)x_{t-1} + \ldots + a_m(t)x_{t-m} = b(t)\epsilon_t \, ,$$

i.e., autoregressive with trending coefficients. Such models can be fitted to the data but more research is required before the use of such models becomes satisfactory.

4. Proposed Steps in the Analysis of a Group of Economic Time Series

(1) Select group of series. Business cycle theories are used to provide guiding principles for choosing the economic time series and for selecting the relationships among them as the subject of our studies. The series need to be all of the same length and this length needs to be at least 200 months. (Slightly shorter series can be considered but the possibility of getting spurious results is increased.)

(2) Test for stationarity of mean

    (2a) If non-stationary, filter out low frequencies

(3) Test for stationarity of variance

    (3a) If non-stationary, take logs, filter out low frequencies and then take exponential

(4) Data now suitable for spectral analysis. Form all spectra, cross-spectra, coherencies, etc. Investigate relations, degrees of dependence at different frequencies, possible lags, etc.

(5) Form spectra for two halves of some series. Investigate possibility of spectra changing with time.

(6) Stretch out original series to look for low frequency periods. Filter, form spectra etc. and investigate.

(7) Consider possibility of fitting a linear regression model and test goodness of fit.

(8) Consider fitting to original data some non-stationary model.

(9) Consider other methods of investigating interesting real or possible aspects of the data.


5. Progress of the Project

    After four months the following has been achieved:

(a) Compilation and preparation of some groups of data and in particular exchange and interest rates prior to the first world war.

(b) Part of the computation required under (2), (3) and (4) above, mostly using three series formed from

exchange rates. Initial investigations into the results obtained.

(c) The formation of a filter for use under sections (2a) and (3a).

(d) Initial considerations of sections (5) and (6) and some considerations of possible models for use with (8).

A great deal of time had to be spent on preparation of the data and preparing programs for carrying out computations on an IBM 650 computer. It is hoped soon to have sufficient data prepared so that spectra etc. may be computed on an IBM 704 computer.

To the best of our knowledge, none of the methods listed above have been previously used on economic data and several of the methods involve unpublished work, e.g. the use of cross-spectra and coherency etc., are methods evolved by Professor Tukey and not widely circulated. The test of stationarity evolved by Granger, and the methods of removing certain types of non-stationarity, Sections (2), (2a), (3), (3a), (5), (6) and (8) all involve, to varying extents, original research.

6. General Aims and Extent of Project

The general aim of the project might be said to apply modern time series methods to economic data and thereby possibly verify and extend previous findings and discover new relationships. In particular, some of the initial aims are to investigate the 'cycles' that previous investigators have found in economic series

and to consider the claims that some series 'lead' other series.

With respect to this, two early results by the project are the verification of the existence of a 40 month cycle in pre-first world war American interest rate data and the leading of the New York Commercial paper rate by the New York call money rate for the same period, although it was found that the extent of the lead varied over the frequency range.

The investigation involves considering both 'horizontal' cross-sections of the economy, i.e. considering together important series from various parts of the economy such as bank clearings, interest rates and production data, and 'vertical' sections, such as grouping together data for various aspects of production in a single industry, from the raw materials to the final product. By discovering the most important components of such sections it is hoped eventually to be able to compose an economy-wide model. An exhaustive investigation of this type may need to consider nearly a hundred time series and can be expected to extend over several years.

# APPENDIX

## Time Required to Analyze a Time Series

To give some idea of the time required to analyze a time series the following are the times in minutes to perform the operations listed for a series 400 months long and using an IBM 650:

|                                          | Minutes |
|------------------------------------------|---------|
| Test for stationarity of mean            | 15      |
| Filtering out trend in mean              | 15      |
| Test for stationarity of variance        | 17      |
| Filtering out trend in variance          | 17      |
| Forming spectra and cross-spectra        | 45      |
| Forming coherence etc.                   | 4       |

To fully prepare data for forming a spectrum and to calculate the spectrum will usually take 3 hours as several of the above operations need to be used more than once. This figure does not include subsequent analysis and plotting of the data.