Qüestiió - V. 8, n.º 1 (març 1984) pp. 9-19

PROBLEMS IN SCIENTIFIC TIME SERIES ANALYSIS

TUNNICLIFFE WILSON
UNIVERSITY OF LANCASTER, U. K.

The paper reviews the statistical methods of time series analysis used in a selection of papers from respected scientific journals. In particular, problems are considered in the search for cycles, the use of regression to establish causal links between variables, transfer function modelling and the use of filtering to extract components of time series.

An attempt is made to assess how useful the ideas of ARMA and Transfer Function modelling might be in improving the efficiency of statistical inference in these contexts.

Keywords: PARAMETRIC MODELLING, ARMA MODELS, TRANSFER FUNCTION MODELS, LINEAR FILTERING, CLIMATOLOGICAL TIME SERIES.

1. INTRODUCTION.

Time series analysis is an activity with an extremely broad range of applications, including engineering, and the physical and social sciences. Within the subject a selection of techniques for 'signal processing' have been developed which are designed to reveal the structures within series and the relationships between them, so that they are evident to the analyst.

In so far as it is concerned with the manipulation of observational data, time series analysis may be formally considered a branch of statistics. The applications of methods of statistical inference to time series analysis has developed over the years, leading to model building procedures with a sound statistical foundation.

Gwilym M. Jenkins, through his major books with D.G. Watts /18/ and G.E.P.Box /3/, his extensive consulting and Instructional activities, and his recent books describing practical experiences and case studies (1979-1983) contributed to and promoted such an approach to time series analysis. This has had a major impact on the application of the subject in many fields, particularly business and economic forecasting. Engineering, and some of the physical and environmental sciences such as hydrology have also readily accepted this approach.

There are however, several areas of scientific time series analysis which have not been fully exposed to this approach. The aim of this paper is to look at a selection of published scientific analyses, and to assess whether and how they might benefit from the application of the ideas of ARMA and Transfer Function modelling. To leave it at that would be rather arrogant. An attempt has therefore been made to understand some of the techniques used in these papers, particularly the widespread use of data filtering. In seeking the statistical justification for such techniques, I believe we can uncover questions of real interest to the statistician concerning procedures of statistical inference for time series.

The selection of papers for consideration, was made mostly from well known and widely real scientific journals. Many are associated with climatology a subject of considerable popular interest with a reputation for attracting explanations in terms of cycles of various origins. I shall consider the papers under four headings: the search for cycles, the use of regression, transfer function models and lastly, the use of filtering. Under each heading I shall briefly describe some scientific data

- G. Tunnicliffe Wilson - University of Lancaster - Dep. of Mathematics - Cartmel College Bailrigg, Lancaster, LA1 4YL
- Article rebut el novembre del 1983.
- Aquest treball es una ponencia invitada al "First Catalan International Symposium on Statistics" a Barcelona del 23 al 24 de setembre de 1983.

analyses, comment on what has been done, and describe some of the problems which may be worthy of further consideration. I hope others will be encouraged to examine these problems, since I believe the statistician could do much more to assist the scientist in the interpretation of his data.

2. THE SEARCH FOR CYCLES.

A distinction should be made here between the search for evidence of cycles with fairly well precribed frequencies, and the more suspect search for cycles of indeterminate frequency.

Consider first an excellent example in Hays, Imbrie and Shackleton /16/. They examine a set of three indices obtained from measurements on deep sea sediment cores. Each series is of approximately 160 points representing a geological time span of 468KY (thousands of years). They seek evidence to support the Milankovitch theory of the ice age. This theory implies that glaciation of the earth follows cycles in the obliquity of the earth's axis (with a period of ∿41KY), precession of the equinoxes (period ∿21KY), and possibly also eccentricity of the earth's orbit (period ∿93KY). Berger /1/ presents tables of the periods of the main cycles predicted by the theory. Their indices are expected to reflect the variation in climate, so that significant cyclical components corresponding to these periods will be taken as supporting evidence for the theory. In fact the postulated cycles are only quasi-periodic, and over the time span of interest the cycle of period 21KY is better represented as being split into components of period 19 and 23 KY. There is also some slight uncertainty in the dating of the core, due to variations in sedimentation rate, so the time scale attached to the data points may be in error by about 5%. The authors therefore attach greater importance to finding a pair of cycles with approximate periods 41 and 23K years, but with a relatively precise ratio of 1.8.

The method used is classical spectral analysis, and a precise description is presented. Both Blackman and Tukey /2/ and Jenkins and Watts /18/ are referenced, and acknowledgement is given of advice from J.W. Tukey. Because the aim is to detect discrete components in the spectrum, a high resolution spectrum (hrs) of handwidth ∿ .01 cycles/KY is calculated. Peaks in the spectrum are found close to the prescribed frequencies. Their statistical evaluation uses a low resolution spectrum calculated for the same data. Then: 'Based on the null hypothesis that the data are a sample of a random signal having a general distribution of variance like that in the observed low-resolution spectrum, confidence intervals are calculated as a guide to the statistical significance of spectral estimates in the high resolution spectrum. A particular peak in the spectrum is judged significant if it extends above the lrs by an amount that exceeds the appropriate one-sided confidence interval.'

This is an extremely rigorous test, as might be expected given the advise received by the authors . Of six peaks examined, two are found significant at $P = .05$, one at $P = .02$.

COMMENT:

(a) The choice of bandwidth = .01, means that between the postulated peaks at frequencies $f = .01$, .024, .043, .052 and the origin $f = 0$, there are only one of two bandwidths. To properly assess a peak in a spectrum one needs to compare it with the general level of the spectrum at neighbouring frequencies. With the bandwidth chosen in this analysis, the four frequencies of interest are so close, and close to $f = 0$, that there is very little information from neighbouring frequencies that can be used for comparison. The chosen course of computing a lrs for comparison, incorporates the variance from the peaks in the level of the lrs. Even if the peaks were real, unless one of them were particularly dominant, their deviation from the lrs would be small. Unfortunately therefore, the method used for assessing the peaks would seem to be lacking in sensitivity.

(b) One might question why an even smaller bandwidth was not used, because the presence of discrete components is best detected in the absence of smoothing, using the raw sample spectrum. Even the small amount of smoothing used in the hrs could cut a discrete component down by a factor of about four. The reason may be that the expected cycles are not in fact so well defined in frequency, due e.g. to modulation of amplitude or time scale, so that a narrow peak is a better description than a discrete component.

SUGGESTIONS:

(a) A time domain model for a mixed spectrum might usefully be considered for this data, e.g.

$$x_t = A \cos wt + B \sin wt + e_t \qquad (2.1)$$

where x_t is the observed series and e_t the error. The sinusoidal terms could be repeated for each frequency w expected in the data. A low order ARMA model would be used for the error process. The coefficients A and B would be estimated, together possibly with the frequency w if there was some uncertainty in the value of this, arising from the time scale calibration limitations. The range of possible frequencies would however have to be narrowly defined for each component.

The selection of the ARMA model would provide a problem. One possibility is to fit an ARMA model to x_t without the sinusoidal components, and to use this with the parameter values held fixed, in the mixed spectrum model (2.1). This would be equivalent to the use of the lrs for assessment of the hrs peaks. It would also suffer a similar criticism, in that if substantial discrete components were present the ARMA model would be biased away from that appropriate in the true model (2.1). The significance of the discrete component would tend to be diminished. A more appealing procedure is to estimate the ARMA model parameters in (2.1) along with the coefficients A, B and frequency w. This follows the treatment given by Campbell and Walker /5/ to mixed spectrum estimation. The possible danger here however is that removal of the discrete components from x_t would deplete the error spectrum over the range of frequencies of interest. This could bias the ARMA model in the opposite direction so that the significance of the discrete components would tend to be misleadingly enhanced. Provided the discrete components were neither too many in number, not too close in frequency, the bias in the ARMA model should not be of great concern, and for the data set considered here, should be tolerable. Preliminary analysis of one of the data sets reveals that the mixed spectrum model (2.1) fits very well, with significant coefficients A, B and an ARMA (2,1) error structure.

Another example which merits a much more skeptical approach to its conclusions, is given by Luo et al /22/ whose work is reported by Gribbin /14/ . They seek cycles in a series of 150 annual measurements of the change in daylength. Their method is to select peaks in the periodogram, and then to refine the frequencies and amplitudes by fitting a model such as (2.1) using ordinary least squares applied to the errors e_t. They find twelve different frequency components in the series.

COMMENT:

(c) The dangers associated with the previous example occur here in an extreme manner. The frequencies are not prescribed, and those discovered are separated approximately by the fundamental frequency $w = 2\pi/150$. The fitted components therefore absorb almost all the low frequency variance in the series, and leave very little information in the residual spectrum for use in assessing their statistical significance. I have fitted the first 8 discrete components to this series using model (2.1) with ARMA error structure and the following difficulties arise.

- An ARMA (2,2) model with strongly autoregressive behaviour fits well the series x_t without discrete components.

- If this model with the ARMA coefficients fixed, is used as the error structure in the mixed spectrum model, the discrete components are not significant.

- If the same model is fitted but with the ARMA coefficients freed, the autoregressive parameters are much reduced and the discrete components become highly significant. Exact Gaussian Likelihood was used, and the likelihood ratio test statistic for the discrete components (+trend) was 101.2 which would be referred to chi-squared on 17 d.f.

To obtain some feel for where the truth lies a resort was made to simulation. This is very much to be recommended to the scientist who is concerned to know whether the structure he has discovered might be spurious. It requires only the acceptance that the data might arise by a random (or random walk) process with spectrum features such as can be produced by a low order ARMA (or ARIMA) model - although some might reject this proposal. In the example under consideration a series was simulated from the ARMA (2,2) model as fitted to x_t without discrete components. When subjected to the same modelling procedures as above, very similar results were obtained, the likelihood ratio test statistic being 31.8. One might conclude that the discrete components are spurious, but be uneasy about the use and reliability of likelihood inference in such examples. Having looked further at this problem, I believe that marginal likelihood inference can be used to improve the statistical assessment, as described in Tunnicliffe-Wilson /24/. In essence this makes inferences about the ARMA structure of the error, on information in the data which is not corrupted by the fitting of discrete components.

The search for cycles is therefore fraught with problems, and the subject has become notorious for spurious results. This is not because scientists are unaware of the problems, see for example the discerning article by Burroughs /4/. There are some very strongly cyclic series, such as the sunspot record, and it is natural to look for the reflection of such cycles in terrestial climatic variables. Thus in another example presented by Dicke /9/, the deuterium/hydrogen (D/H) ratio is measured in a set of 10 yr samples of material obtained from two bristle cone pines that cover 1000 yr of time without gaps. This is supposed to reflect climatic variation. His spectral analysis of this data reveals a high peak in the raw spectrum corresponding to a period of 22.36 years, coinciding remarkably well with the period of the solar magnetic field (twice the sunspot period). Besides the question of statistical significance of spectral peaks, there is also the question of whether the coincidence of peak frequencies in two series can be taken as evidence of a relationship. For a precisely defined frequency no statistical comment can be made, because the coherency cannot be estimated except over a frequency band. In another version of his article, Dicke /10/ concludes in a sceptical vein, but asks 'would nature be so malicious as to produce the close correspondence shown in the two frequency peaks as a statistical fluke?'

Success in prediction does of course add considerably to the credibility of cyclical models. Although this is not very relevant to series sampled at 3KY intervals, it is interesting to note for example that the 76 year period detected by Gilliland /12/ in low quality historical measurements of the solar radius, appears to be confirmed by recent, higher precision measurements of Parkinson /23/. The recent volcanic activity of Mt.St. Helens and Chichon might be expected to confirm relationships between the volcanic dust veils and the climate, such as are discussed by Kelly /21/, who also engages in the search for cycles.

3. THE USE OF REGRESION.

We have mentioned regression as applied to the fitting of sinusoids by least squares. It is known in that context that the fitted coefficients are unaffected by the assumed error model, although incorrect standard errors are given if the wrong error model is used. In other contexts it becomes more important for inference to allow for any autocorrelation in the error process. In the earlier part of his paper, Dicke /9/ is concerned with estimating the period of the sunspot cycle and its stability. He selects the times T_i of peak sunspot activity and carries out a linear regression

$$T_i = c + pi + bN_i + e_i, \quad i = 1...25 \qquad (3.1)$$

where i is the index of the peak, p is the period between peaks, N_i is the sunspot number of peak i, which in part explains variation in the occurrence of peak times, and e_i is the error. The precision of estimation of the period p is of concern. Some theories suggest that the error e_i is a random walk, i.e. delays in the cycles are cumulative and passed on to subsequent cycles. This would imply differencing of (3.1) before carrying out regression and leads to a much larger estimated standard error for the period. Dicke attempts to discriminate between the two possibilities, that e_i is a random series, or is a random walk. Perhaps surprisingly, this is a challenging task, because in small samples much of the evidence for a random walk error is lost by fitting the regression. This is well appreciated by Dicke, who compares the variance of the residuals $\hat{e}_i$ with that of $\nabla\hat{e}_i$. The ratio of the second of these to the first is found to be 1.15. This increase is not as great as would be expected (i.e. 2.17), but appears to be in the right direction, especially since one might suppose that if the random walk error model was correct, then a massive variance reduction would be found. In fact the expected variance reduction in this latter case is only 0.58. The observed value is almost (geometrically) the mean of the two expected values! Using other evidence from the residuals, Dicke prefers to reject the random walk model.

COMMENT:

> The use of residual sum of squares as a criterion in selecting between different regression error structures can be misleading. Distributional properties should be investigated (possibly by simulation).

SUGGESTION:

> The use of marginal likelihood can again aid inference in such examples. However, in this case a more flexible error structure could be used. An AR(1) error model fits quite well, with parameter .625, and provides an estimate of the precision of the period (i.e. one S.E.=.o85) which is reasonable.

Another example which again reveal a cautious and thoughful statistical attitude, is to be found in a thesis by Gasquoine /11/. She re-examines models of transoceanic geomagnetic variations. Anomalies in the earth's magnetic fields are measured in an east-west direction in parts of the oceans, giving spatial series (magnetic anomaly profiles). Theories of ocean floor formation and spreading, together with knowledge that the earth's magnetic field undergoes reversals, suggest that the profiles are due to bands of rocks of alternating magnetic polarity across the ocean floor. Dates of magnetic reversals may be gleaned from continental volcanic rocks, so that the profiles mat be predicted for assumed values of the origin and time scale of ocean floor spreading. By matching the observed and predicted anomalies, this origin and time scale - particularly the ocean floor spreading rate - may be estimated. Both observed and predicted anomalies show a succession of somewhat irregular positive and negative peaks, so it is quite possible to find apparently good matches for a variety of parameter values. The criterion generally used to select the best parameters, was the sample correlation r between observed and predicted profile. The deficiency of this was recognised by Gascoigne as I summarise in the following comments: Plots of r against the spreading rate showed several peaks. Possibly more serious, the significance of r could not be assessed. The data were got by sampling a continuous trace. If the sampling interval is decreased, the number of data points n increases but it is evidently inapropriate to use this to extract from statistical tables an ever smaller critical value for r. The value of r itself stabilizes, and no new information is being gained by increasing n.

SUGGESTION:

> The problem arises from autocorrelation of the data. Regression of the observed profile on the predicted should be carried out with an ARMA error structure.
>
> This was in fact done, with an ARMA (2,1) error model. Reduction in the sampling interval below a certain threshold, although increasing the ARMA parameters, left the regression coefficient and its standard error almost unchanged.

> The spreading rate was determined as that which gave the highest 't' value for the regression coefficient, a quite reasonable procedure. Thus in one example a rate of 4.48 cm/yr gave t = 5.61, but this dropped to 2.75 at a rate of 4.60 cm/yr. In general the conclusions from this analysis were highly satisfactory.

The use of regression for matching curves is widespread. The error is commonly highly autocorrelated. This is probably because there are some unknown component curves which have been omitted from the regression and thus have to be treated as part of the error. Although these absent components may be of well determined form, our ignorance demands that we treat them stochastically - usually as a sample from an ARMA or ARIMA process. I can see no way to avoid this argument, although its application may bring into question some common scientific models. An area which is currently surrounded by controversy is the interpretation of observed infra-red spectra of galactic sources as discussed for example by Hoyle and Wickramasinghe /17/. To attach statistical significance to a match between an observed spectrum and a theoretical spectrum selected from a large emsemble of possible materials is a difficult challenge, but I suspect that the use of ARMA error models may be justified in such circunstances.

4. TRANSFER FUNCTION MODELING.

In models of climatic response, e.g. to variations in the amounts of solar radiation which reach the earth, the response takes place over widely varying time scales. The system of the earth's atmosphere, oceans and continents is dynamic, with time constants which may be measured in terms of days, weeks, years, centuries, or even longer. Equilibrium can only be considered relative to a given time span - in the very long run equilibrium is never attained. In their analysis, Hays et al /16/ avoid detailed modelling of the transfer functions of the responses to orbital variations, by noting that the sinusoidal inputs which are characteristic of these variations, will have sinusoidal responses of the same frequencies - provided only that the response is reasonably linear. In situations where the input is much more irregular, some attempt must be made at modelling the transfer functions involved.

The only serious example I have found of this, is given by Gilliland /13/, who models annual measurements of climatic variation over the last 100 years or so. He uses mean Northern Hemisphere Temperatures to measure the climate. His 'forcing functions' or inputs are:

(a) A cycle of period 76 years extracted from radius measurements and used as an indicator of solar radiation flux.

(b) A measure of atmospheric volvanic aerosols from Greenland ice sheet impurities as provided by Hammer /15/. This is used as an indicator for volcanism which is supposed to affect the earth's climate by veiling the sun. It is a fairly irregular series, though with some longer term trends.

(c) An exponential increase in atmospheric CO_2 due to human activity, at a rate which leads to a projected doubling in the level from 1925 to 2045. Via the greenhouse effect on solar radiation this will increase atmospheric temperatures.

Gilliland uses a two-compartment model of the ocean with time constants taken from physical considerations, to represent the most important part of the system controlling the earth's temperature on this time scale. The above inputs are used as forcing functions in his differential equations, and the output is matched to observed temperatures.

He also introduces a time lag for the solar cycle, to represent an expected delay between solar radius and radiation variation, and a negative time lag for the aerosols which are of course trapped in the ice sheed after they have circulated in the atmosphere.

COMMENT:

(a) The model used by Gilliland can be shown to follow the formulation of Box and Jen kins rather than the alternative ARMAX model, and in their notation becomes

$$y_t = \frac{(1-wB)}{(1-\delta_1 B-\delta_2 B^2)} (ax_{1,t}+bx_{2,t}+cx_{3,t}) + +e_t \qquad (4.1)$$

where y is the observed temperature and x_1, x_2, x_3 the inputs. The only fitted coefficients are a,b,c, ordinary least squares being applied to e_t. In particular w, δ_1 and δ_2 are not estimated, but follow from assumed physical constants and dimensions.

(b) Examination of the constants used by Gilliland lead me to believe that the transfer function would be well approximeted by a lower order form, $1/(1-\delta B)$, with $\delta \simeq .974$. It became apparent though that I must have misinterpreted his equations, because it was only through using a value $\delta \simeq .7$ that I was able to reproduce his graphical results with reasonable accuracy.

(c) The author examined the residuals from his model fit and concluded that they were far from random. His error analysis detected cycles with periods of 24 and 12 years, plausibly close to those associated with sunspots. The data had however been smoothed with a seven point moving average formula, which would in any case induce autocorrelation.

(d) The author is well aware that 'One could, with some validity, argue that fitting 11 least squares parameters to a climate record of 95 yr after smoothing with a low pass filter is a case of over-determination'. He says 'I am not attempting to argue the statistical significance of the above models, but merely that all of the included external forcing are plausible'. The plausibility derives very much from the fact that the estimated magnitudes of the effects of the forcings are consistent with theoretical expectation.

SUGGESTIONS:

The statistician feels obliged to make some attempt to assess significance, and insofar as I was able to duplicate the results of the analysis, I found that:

(e) Fitting the model (4.1) with ordinary least squares applied to the errors, but with the transfer function replaced by $1/(1-\delta B)$, gave apparently highly significant coefficients for a, b and c with t values of 9.8, 6.7 and 6.1. These were slightly improved when, following the author's example, sinusoidal terms of periods 22 and 12.4 years were added to the model. The validity and significance of these t values should be strongly questioned in the knowledge of the error autocorrelation still remaining.

(f) In view of the data smoothing which had been applied, a MA(7) error model was introduced into the previous regression. The MA coefficients were found to be large and reflected closely the weights used in the smoothing. The coefficients a, b, c of the inputs did not however change very much, though their t values were halved. The residuals had relatively little remaining autocorrelation, though what was left appeared to be due to low frequency components. The sinusoidal terms of periods 22 and 12.4 were clearly no longer significant. Halting the analysis at this point would leave one with appreciable confidence in the conclusion that the inputs adequately modelled the output.

(g) The transfer function having been taken as fixed, I experimented with estimating its dynamic coefficient δ, and elaborating in other ways. When the OLS error model was used, I found a much better fit with a value of δ close to 1, but the estimated coefficients of the inputs changed markedly. In fact the volcanic effect became positive rather than negative.

(h) For the purpose of comparison, a univariate ARMA (1,8) model was fitted to the output series. An autoregressive term was included to allow representation of the low frequency behaviour which had otherwise been modelled by the inputs. The fit, as measured by the residual variance, was almost as good as that given by any of the previous models, casting serious doubt on their validity . My conclusion was that much more effort needed to be spent on this data set to establish useful relationships. It was the author's hope, I believe, that the solar cycle and CO_2 effect would account for the longer term (lower frequency) patterns in the data, whilst the volcanic index would mainly explain shorter term features. In practice, the solar cycle and volcanic index (especially after passing through the transfer function) show considerable correlation, which explains the instability I found in the coefficient of the volcanic index. This leads us to consider in what components of the data the required information might reside.

5. THE USE OF FILTERING.

In the examples referenced above, filtering is widely used. Thus Dicke /9/ uses a high pass filter to remove the trend from the (D/H) ratio series, before calculating its spectrum. Gilliand /13/ uses a low pass filter to remove noise from two of his series, the temperature record and the Volcanic Aerosol measurements. He also uses a very narrow handpass filter to extract the 76 year period cycle from solar radius measurements. Luo et al /22/ also smooth their data, with moving average filters of varying widths to allow for improvement in the quality of data over the 150 year record.

Hays et al /17/ use bandpass filters to extract components of their series associated with the peaks in their spectrum estimates. The same technique is used by Currie /7/to examine variations in the amplitude of the 11 year solar cycle which he had previously detected in the earth's rotation rate - see Currie /6/.

Apart from the use of a prewhitening filter by Hays et al /16/ and the transfer function filter in the model of Gilliland /13/, the use of filtering as reviewed above, is treated by statisticians with great suspicion in the context of ARMA and transfer function modelling. This is because filtering is viewed as imposing structure on the data which then complicates any modelling procedure. Seasonal adjustment has an Identical effect in economic time series analysis. The current view, I believe, is that it is far better to let the data dictate the filters, i.e. ARMA and Transfer Function operators which are necessary in the model. When the model has been fitted, then filters can be constructed in an optimal fashion for applications to signal extraction.

I would however like to argue that the use of filtering can be justified as a preprocessing procedure in parametric time series modelling, although it may be difficult to integrate the theory and practice.

Take for example the use of smoothing to reduce noise in a univariate analysis. The noise is essentially undesirable, particularly at high frequencies. A reasonably good picture of the underlying signal may be obtained even if the noise varies in amplitude over the data span as in the example of Luo et al. /22/ If however, an ARMA model is fitted to the filtered series in the (by now) classical manner of Box and Jenkins /3/ with the aim of achieving white residuals, the smoothing filter is undone and the noise recreated.The following example does I believe provide a context in which this classical modelling of the filtered series should not be carried out.

Suppose the observed data Z_t is the sum of unobserved components X_t and N_t, where the signal X_t is AR(1) and N_t is white noise,i.e.

$$Z_t = X_t + N_t$$

$$X_t = \phi X_{t-1} + \alpha_t \quad , \quad N_t = \beta_t \qquad (5.1)$$

where α_t, β_t are white noise with variances σ^2_α , σ^3_β . It is well known that Z_t follows an ARMA (1,1) model

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1}$$

where the parameter θ and variance σ^2_a are given by

$$(\phi-\theta)(1-\theta\phi)/\theta = \sigma^2_\alpha/\sigma^2_\beta = R \ ; \ \sigma^2_a = (\phi/\theta)\sigma^2_\beta$$

Moreover the optimal smoothing to extract an estimate $\hat{X}_t$ of X_t from Z_t is a two-sided exponentially weighted average,

$$\hat{X}_t = (\sigma_\alpha^2/\sigma_a^2)(1-\theta^2)^{-1}\{Z_t + \sum_1^\infty \theta^k (Z_{t+k}+Z_{t-k})\} \quad (5.2)$$

Given a finite sample $Z_1 \ldots Z_n$, forecasts and back forecasts of Z_t should be used in this formula for t outside the observed range.

Suppose further that we are only supplied with $\hat{X}_t$, for $t = 1 \ldots n$ and wish to forecast X_t. The series $\hat{X}_t$, apart from end effects, does in fact follow an AR(2) model with operator $(1-\phi B)(1-\theta B)$. Treating $\hat{X}_t$ as the series of interest, would lead to the identification of this model and its application to forecasting. This does not however produce optimal forecasts of X_t. These are obtained by using the AR(1) model for X_t with parameter ϕ, and applying this to $\hat{X}_t$, an immediate consequence of least squares projection theory.

This lends substance to the idea that if smoothing is applied to extract a signal, and it is a forecast of this signal (not of its estimate) which is required, than the assumed model of the signal (not that of its estimate) should be applied to the estimate in forecasting.

What if the model for the signal requires estimating?. Can this be done from $\hat{X}_t$? . The E-M algorithm of Dempster et al /8/ may be applied to show that $\hat{X}_t$ can be used in place of X_t to estimate the AR(1) model for X_t. If X_t vere available the estimate $\phi = r_1$ (lag 1 acf) would be asymptotocally efficient. Using $\hat{X}_t$ in place of X_t one must however allow for bias, and should use instead:

$$\phi + \theta(1-\phi^2)(1-\theta\phi) = r_1 \; .$$

Thus knowledge of the smoothing parameter, and the smoothed series allows the estimation of the signal model which can be applied in prediction. Note that the residuals from this process of filtering the AR(1) model to $\hat{X}_t$. would not be white, but would themselves follow an AR(1) model with parameter θ!

The foregoing exploration of possible justification of smoothing practices, would hardly tempt the Box-Jenkins practioner from seeking first an ARMA model for the unsmoothed data, and pursuing any applications to forecasting from that point. When we consider the context of relating one time series to another, we may however have to concede some shortcomings in our transfer function model formulations unless we are prepared to admit smoothing. The simplest illustration arises when we have a direct regression between two variables X_t and Y_t with errors in the observation of both, e.g.

$$Y_t = aX_t + e_t$$

but we observe $V_t = Y_t + \alpha_t$ and $Z_t = X_t + \beta_t$. The error series e_t and noise series α_t, β_t are taken to be white noise. If our aim is to establish a predictive relationship for V_t given Z_t then the correlation structure of X_t becomes important. If for example X_t is AR(1) as in (5.1), then the best predictor of V_t given Z_t is just $a\hat{X}_t$, where $\hat{X}_t$ is given by filtering Z_t as in (5.2). In modelling V_t in terms of Z_t we would therefore require the two-sided transfer function which this filter defines. It does seem more appealing though, rather than estimating a two sided transfer function for V_t on Z_t, that we should instead smooth Z_t to get $\hat{X}_t$ and then estimate a simple regression of V_t on $\hat{X}_t$.

The next question is, having smoothed Z_t, why not smooth V_t to estimate $\hat{Y}_t$ and then regress $\hat{Y}_t$ upon $\hat{X}_t$. This is quite a common practice in science and engineering. There is some justification for this to be found in the E-M algorithm, but the situation is much more complicated than for univariate modelling. I would though, like to present an argument from spectral analysis.

The spectral estimation of transfer functions is just regression in the frequency domain. At a given frequency, or more precisely, over a given frequency band, this is equivalent to applying the corresponding band pass filter to the series and carrying out a simple lagged regression between the extracted components in the time domain, i.e.

$$\hat{V}_t(w) = a(w)\hat{Z}(w)_{t-b(w)} + e_t(w)$$

where $\hat{V}_t(w)$ etc. are the extracted frequency components. The error component will of course also belong to the designated frequency band. It will not be white, but will be a combination of independent frequency components within this band. The S.E. for the coefficient a(w) as supplied by this regression should however be scaled by 1/(bandwidth of the filter).

We might be persuaded that in circumstances where we have signals Y_t and X_t which are predominantly in a low frequency band, except for contamination by white noise α_t, β_t, then low pass filtering (i.e. smoothing) followed by OLS regression (possibly with a lag) is covered by the above argument. One important point is that it excuses us from using ARMA error models, which simply tends to undo the filtering.

6. CONCLUSION.

I have considered mostly the field of climatology and the related area of geophysics. I find it interesting possibly because it seeks to relate so many variables such as the solar constant (sic), volcanism, orbital and rotational variations of the earth to our climate It is beset by problems of availability and quality of data. Spectral analysis is applied with some sophistication, and regression with somewhat less confidence. The sought for cycles and other effects are tantalisingly close to the borderlines of significance. I believe that there is a real challenge to the statistician in aplying the more recent techniques of parametric time series modelling to this subject. Hopefully this will confirm some of the conjectured relationship.

The subject raises problems of statistical inference which may stimulate further research into time series model building. I would encourage others to examine carefully the papers which I have referenced with a view to solving some of these problems.

7. REFERENCES.

/1/ BERGER, A.L.,: "Support for the astronomical theory of climatic change". Nature 269, 44-45. (1977).

/2/ BLACKMAN, R.B. AND TUKEY; J.W.: " The measurement of power spectra from the point of view of communication engineering". Dover, New York. (1958).

/3/ BOX,G.E.P. AND JENKINS,G.M.: "Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, 2nd Edition 1976. (1970).

/4/ BURROUGHS, W.: "Meteorological Cycles" , New Scientist, 87, 472-473 (1980).

/5/ CAMPBELL, M.J. AND WALKER, A.M.:"A survey of statistical work on the Mackenzie river series of annual Canadian lynx trappings for the years 1821-1934 and a new analysis. J.R. Statist. Soc. A, 140, 411-431 (1977).

/6/ CURRIE, R.G.: Geophys , J.R. Astrom. Soc. 61, 131. (1980).

/7/ CURRIE, R.G.: "Solar cycle signal in earth rotation: nonstationaty behaviour". Science 211, 386-389. (1981).

/8/ DEMPSTER, A.P., LAIRD, N.M. AND RUBIN,D.B. "Maximum likelihood from incomplete data via the EM algorithm". J.R. Statist.Soc. B. 39, 1-38. (1977).

/9/ DICKE, R.H.: "Is there a chronometer hidden deep in the sun?". Nature, 276, 676-680. (1978).

/10/ DICKE, R.H.: "The clock inside the sun" New Scientist, 83, 12-14. (1979).

/11/ GASQUOINE, S.L.: "The statistical analysis of magnetic anomaly profiles across mid-ocean ridges". B.A. Thesis, Dept. of Environmental Science, University of Lancaster, U.K. (1975).

/12/ GILLILAND, R.L. : "Solar radius variations over the past 265 years". The Astrophysical Journal, 248, 1144-1155. (1981).

/13/ GILLILAND, R.L. : "Solar volcanic and CO_2 forcing of recent climatic change". Climatic Change, 4, 111-131. (1982).

/14/ GRIBBIN, J.: "New Chinese results tie up sun cycles and earth weather". New Scientist, 76, 703. (1977).

/15/ HAMMER, C.U.: "Past volcanism revealed by Greenland ice sheet impurities". Nature, 270, 482-486. (1977).

/16/ HAYS, J.D., IMBRIE, J. AND SHACKLETON, N.J.: "Variations in the Earth's Orbit: pacemaker of the ice ages". Science, 184, 1121-1132.

/17/ HOYLE, F. AND WICKRAMASINGHE, N.C.:"Polysaccharides and infrared spectra of galactic sources", Nature 268, 610-612.

/18/ JENKINS, G.M. AND WATTS, D.G.: "Spectral analysis and its applications" Holden-Day, San Francisco. (1968).

/19/ JENKINS, G.M.: "Practical Experiences with modelling and forecasting time series", Gwilym Jenkins and Partners, Lancaster. (1979).

/20/ JENKINS, G.M. AND McLEOD,G.: "Case studies in time series analysis" Gwilym Jenkins and Partners, Lancaster (1983).

/21/ KELLEY, P.M.: "Volcanic dust veils and North Atlantic climate change". Nature 268, 616-617. (1977).

/22/ LUO, S., LIANG, S., YA, S, YAN, S., LI,Y "Analysis of periodicity in the irregular rotation of the earth". Chinese Astronomy 1, 221-227. (1977)

/23/ PARKINSON, J.: Nature, 304, 518.(1983).

/24/ TUNNICLIFFE WILSON, G.: "On the use of marginal likelihood in time series model estimation". Technical Report, Mathematics Research Centre. University of Madison. (1983).