THE METHODS OF TIME-SERIES ANALYSIS

Interdisciplinary Science Reviews 12, No.2, 1987

The methods which are to be discussed in this review are designed for the purpose of analysing series of statistical observations taken at regular intervals in time. The methods have a wide range of applications. We can cite astronomy [18], meteorology [9], seismology [21], oceanography [10], [12], communications engineering and signal processing [17], the control of continuous process plants [21], neurology and electroencephalography [1], [8], [26], and economics [11]; and this list is by no means complete.

The Frequency Domain and the Time Domain

The methods apply, in the main, to what are described as stationary or non-evolutionary time series. Such series manifest statistical properties which are invariant throughout time, so that the behaviour during one epoch is the same as it would be during any other.

When we speak of a weakly stationary or covariance-stationary process, we have in mind a sequence of random variables $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$, representing the potential observations of the process, which have a common finite expected value $E(y_t) = \mu$ and a set of autocovariances $C(y_t, y_s) = E\{(y_t - \mu)(y_s - \mu)\} = \gamma_{|t-s|}$ which depend only on the temporal separation $\tau = |t-s|$ of the dates t and s and not on their absolute values. We also commonly require of such a process that $\lim(\tau \to \infty)\gamma_\tau = 0$ which is to say that the correlation between increasingly remote elements of the sequence tends to zero. This is a way of expressing the notion that the events of the past have a diminishing effect upon the present as they recede in time. In an appendix to the chapter, we review the definitions of mathematical expectations and covariances.

There are two distinct yet broadly equivalent modes of time-series analysis which may be pursued. On the one hand are the time-domain methods which have their origin in the classical theory of correlation. Such methods deal preponderantly with the autocovariance functions and the cross-covariance functions of the series, and they lead inevitably towards the construction of structural or parametric models of the autoregressive moving-average type for single series and of the transfer-function type for two or more causally related series. Many of the methods which are used to estimate the parameters of these models can be viewed as sophisticated variants of the method of linear regression.

On the other hand are the frequency-domain methods of spectral analysis. These are based on an extension of the methods of Fourier analysis which originate in the idea that, over a finite interval, any analytic function can be approximated, to whatever degree of accuracy is desired, by taking a weighted sum of sine and cosine functions of harmonically increasing frequencies.

Harmonic Analysis

The astronomers are usually given credit for being the first to apply the methods of Fourier analysis to time series. Their endeavours could be described as the search for hidden periodicities within astronomical data. Typical examples were the attempts to uncover periodicities within the activities recorded by the Wolfer sunspot index and in the indices of luminosity of variable stars.

The relevant methods were developed over a long period of time. Lagrange [14] suggested methods for detecting hidden periodicities in 1772 and 1778. The Dutchman Buijs-Ballot [6] propounded effective computational procedures for the statistical analysis of astronomical data in 1847. However, we should probably credit Sir Arthur Schuster [18], who in 1889 propounded the technique of periodogram analysis, with being the progenitor of the modern methods for analysing time series in the frequency domain.

In essence, these frequency-domain methods envisaged a model underlying the observations which takes the form of

$$\begin{aligned} y(t) &= \sum_j \rho_j \cos(\omega_j t - \theta_j) + \varepsilon(t) \\ &= \sum_j \left\{ \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t) \right\} + \varepsilon(t), \end{aligned} \tag{1}$$

where $\alpha_j = \rho_j \cos\theta_j$ and $\beta_j = \rho_j \sin\theta_j$, and where $\varepsilon(t)$ is a sequence of independently and identically distributed random variables which we call a white-noise process. Thus the model depicts the series $y(t)$ as a weighted sum of perfectly regular periodic components upon which is superimposed a random component.

The factor $\rho_j = \sqrt{(\alpha_j^2 + \beta_j^2)}$ is called the amplitude of the jth periodic component, and it indicates the importance of that component within the sum. Since the variance of a cosine function, which is also called its mean-square deviation, is just one half, and since cosine functions at different frequencies are uncorrelated, it follows that the variance of $y(t)$ is expressible as $V\{y(t)\} = \frac{1}{2}\sum_j \rho_j^2 + \sigma_\varepsilon^2$ where $\sigma_\varepsilon^2 = V\{\varepsilon(t)\}$ is the variance of the noise.

The periodogram is simply a device for determining how much of the variance of $y(t)$ is attributable to any given harmonic component. Its value at

$\omega_j = 2\pi j/T$, calculated from a sample $y_0, \ldots, y_{T-1}$ comprising T observations on $y(t)$, is given by

$$\begin{aligned} I(\omega_j) &= \frac{2}{T}\left[\left\{\sum_t y_t \cos(\omega_j)\right\}^2 + \left\{\sum_t y_t \sin(\omega_j)\right\}^2\right] \\ &= \frac{T}{2}\big\{a^2(\omega_j) + b^2(\omega_j)\big\}. \end{aligned} \tag{2}$$

If $y(t)$ does indeed comprise only a finite number of well-defined harmonic components, then it can be shown that $2I(\omega_j)/T$ is a consistent estimator of ρ_j^2 in the sense that it converges to the latter in probability as the size T of the sample of the observations on $y(t)$ increases.

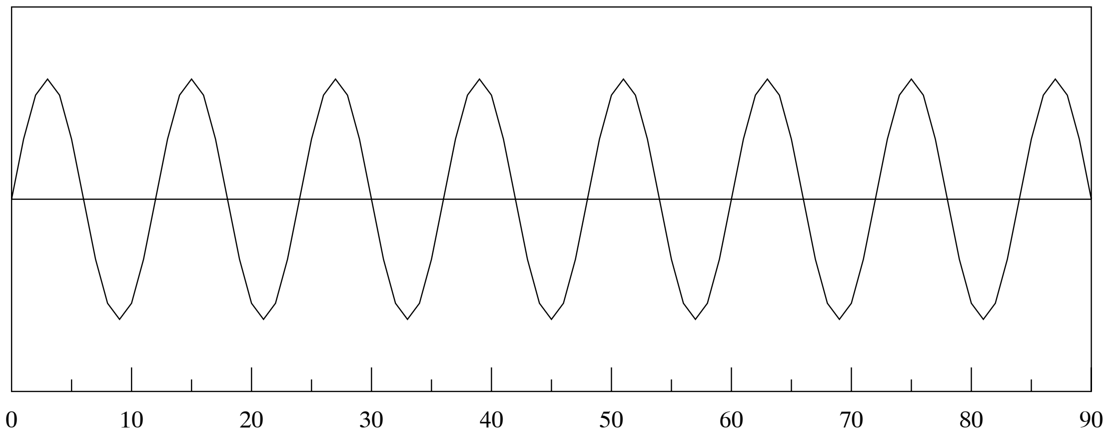

Figure 1. The graph of a sine function.

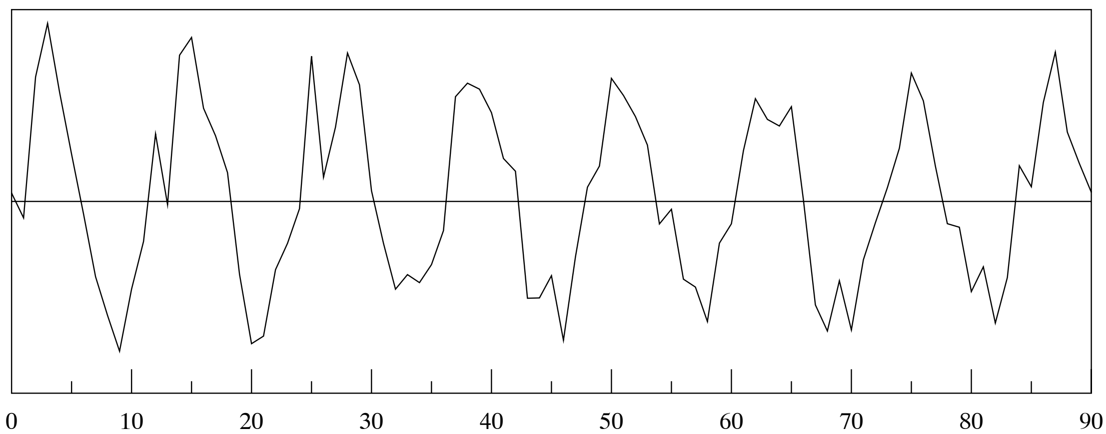

Figure 2. Graph of a sine function with small random fluctuations superimposed.

The process by which the ordinates of the periodogram converge upon the squared values of the harmonic amplitudes was well expressed by Yule [25] in a seminal article of 1927:

> If we take a curve representing a simple harmonic function of time, and superpose on the ordinates small random errors, the only effect is to make the graph somewhat irregular, leaving the suggestion of periodicity still clear to the eye. If the errors are increased in magnitude, the graph becomes more irregular, the suggestion of periodicity more obscure, and we have only sufficiently to increase the errors to mask completely any appearance of periodicity. But, however large the errors, periodogram analysis is applicable to such a curve, and, given a sufficient number of periods, should yield a close approximation to the period and amplitude of the underlying harmonic function.

We should not quote this passage without mentioning that Yule proceeded to question whether the hypothesis underlying periodogram analysis, which postulates the equation under (1), was an appropriate hypothesis for all cases.

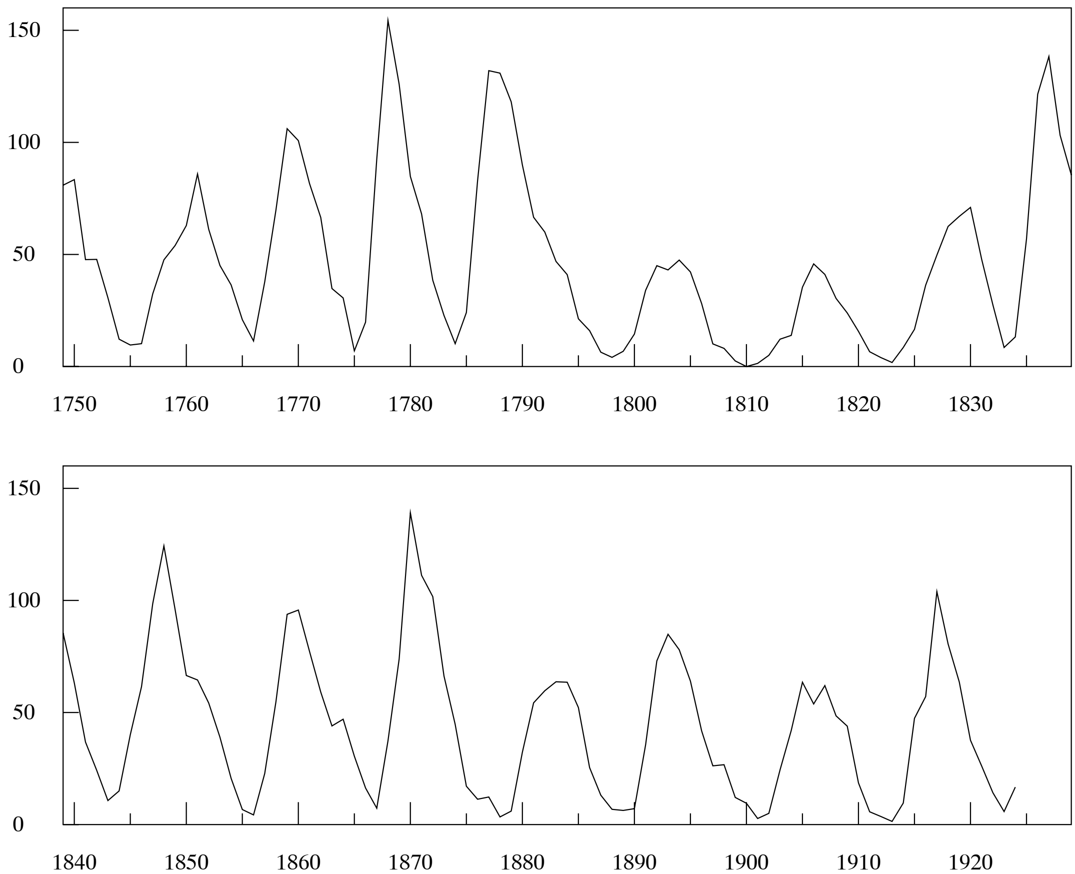

Figure 3. Wolfer's Sunspot Numbers 1749–1924.

A highly successful application of periodogram analysis was that of Whittaker and Robinson [23] who, in 1924, showed that the series recording the brightness or magnitude of the star T. Ursa Major over 600 days could be fitted almost exactly by the sum of two harmonic functions with periods of 24 and 29 days. This led to the suggestion that what was being observed was actually a two-star system wherein the larger star periodically masked the smaller brighter star. Somewhat less successful were the attempts of Arthur Schuster himself [19] in 1906 to substantiate the claim that there is an eleven-year cycle in the activity recorded by the Wolfer sunspot index.

Other applications of the method of periodogram analysis were even less successful; and one application which was a significant failure was its use by William Beveridge [2, 3] in 1921 and 1922 to analyse a long series of European wheat prices. The periodogram of this data had so many peaks that at least twenty possible hidden periodicities could be picked out, and this seemed to be many more than could be accounted for by plausible explanations within the realm of economic history. Such experiences seemed to point to the inappropriateness to economic circumstances of a model containing perfectly regular cycles. A classic expression of disbelief was made by Slutsky [20] in another article of 1927:

> Suppose we are inclined to believe in the reality of the strict periodicity of the business cycle, such, for example, as the eight-year period postulated by Moore [15]. Then we should encounter another difficulty. Wherein lies the source of this regularity? What is the mechanism of causality which, decade after decade, reproduces the same sinusoidal wave which rises and falls on the surface of the social ocean with the regularity of day and night?

Autoregressive and Moving-Average Models

The next major episode in the history of the development of time-series analysis took place in the time domain, and it began with the two articles of 1927 by Yule [25] and Slutsky [20] from which we have already quoted. In both articles, we find a rejection of the model with deterministic harmonic components in favour of models more firmly rooted in the notion of random causes. In a wonderfully figurative exposition, Yule invited his readers to imagine a pendulum attached to a recording device and left to swing. Then any deviations from perfectly harmonic motion which might be recorded must be the result of errors of observation which could be all but eliminated if a long sequence of observations were subjected to a periodogram analysis. Next, Yule enjoined the reader to imagine that the regular swing of the pendulum is interrupted by small boys who get into the room and start pelting the pendulum with peas sometimes from one side and sometimes from the other. The motion is now affected not by superposed fluctuations but by true disturbances.

In this example, Yule contrives a perfect analogy for the autoregressive time-series model. To explain the analogy, let us begin by considering a homogeneous second-order difference equation of the form

$$(3) \qquad y(t) = \phi_1 y(t-1) + \phi_2 y(t-2).$$

Given the initial values y_{-1} and y_{-2}, this equation can be used recursively to generate an ensuing sequence $\{y_0, y_1, \ldots\}$. This sequence will show a regular pattern of behaviour whose nature depends on the parameters ϕ_1 and ϕ_2. If these parameters are such that the roots of the quadratic equation $z^2 - \phi_1 z - \phi_2 = 0$ are complex and less than unity in modulus, then the sequence of values will show a damped sinusoidal behaviour just as a clock pendulum will which is left to swing without the assistance of the falling weights. In fact, in such a case, the general solution to the difference equation will take the form of

$$(4) \qquad y(t) = \alpha\rho^t \cos(\omega t - \theta),$$

where the modulus ρ, which has a value between 0 and 1, is now the damping factor which is responsible for the attenuation of the swing as the time t elapses.

The autoregressive model which Yule was proposing takes the form of

$$(5) \qquad y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \varepsilon(t),$$

where $\varepsilon(t)$ is, once more, a white-noise sequence. Now, instead of masking the regular periodicity of the pendulum, the white noise has actually become the engine which drives the pendulum by striking it randomly in one direction and another. Its haphazard influence has replaced the steady force of the falling weights. Nevertheless, the pendulum will still manifest a deceptively regular motion which is liable, if the sequence of observations is short and contains insufficient contrary evidence, to be misinterpreted as the effect of an underlying mechanism.

In his article of 1927, Yule attempted to explain the Wolfer index in terms of the second-order autoregressive model of equation (5). From the empirical autocovariances of the sample represented in Figure 3, he estimated the values $\phi_1 = 1.343$ and $\phi_2 = -0.655$. The general solution of the corresponding homogeneous difference equation has a damping factor of $\rho = 0.809$ and an angular velocity of $\omega = 33.96^o$ The angular velocity indicates a period of 10.6 years which is a little shorter than the 11-year period obtained by Schuster in his periodogram analysis of the same data. In Figure 4, we show a series which has been generated artificially from the Yule's equation together with a series generated by the equation $y(t) = 1.576y(t-1) - 0.903y(t-2) + \varepsilon(t)$. The homogeneous difference equation which corresponds to the latter has the same value of ω as before. Its damping factor has the value $\rho = 0.95$, and this increase accounts for the greater regularity of the second series.

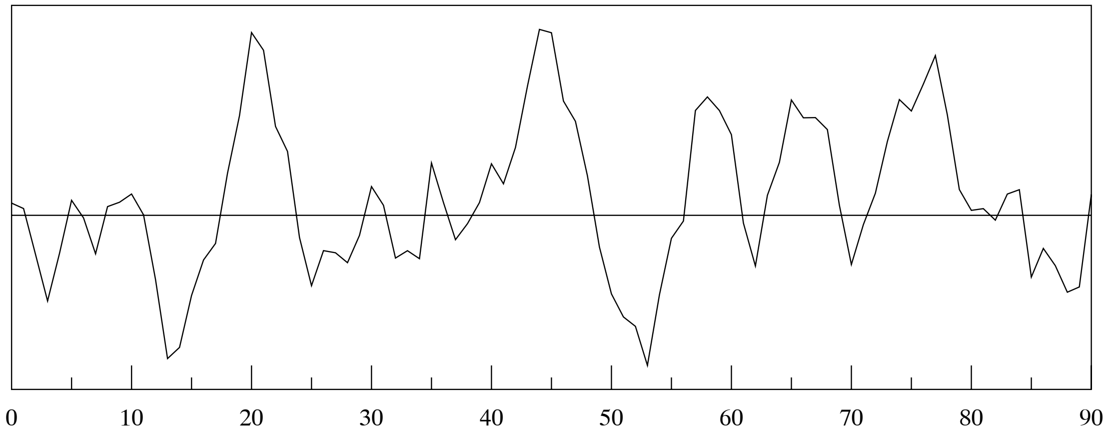

Figure 4. A series generated by Yule's equation
$y(t) = 1.343y(t-1) - 0.655y(t-2) + \varepsilon(t)$.

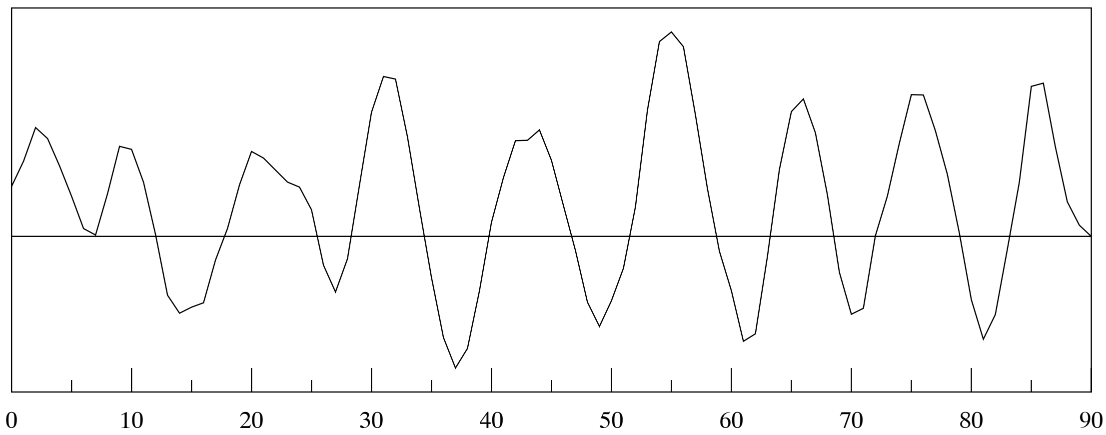

Figure 5. A series generated by the equation
$y(t) = 1.576y(t-1) - 0.903y(t-2) + \varepsilon(t)$.

Neither of our two series accurately mimics the sunspot index; although the second series seems closer to it than the series generated by Yule's equation. An obvious feature of the sunspot index which is not shared by the artificial series is the fact that the numbers are constrained to be nonnegative. To relieve this constraint, we might apply to Wolf's numbers y_t a transformation of the form $\log(y_t + \lambda)$ or of the more general form $(y_t + \lambda)^{\kappa - 1}$, such as has been advocated by Box and Cox [4]. A transformed series could be more closely mimicked.

The contributions to time-series analysis made by Yule [25] and Slutsky [20] in 1927 were complementary: in fact, the two authors grasped opposite ends of the same pole. For ten years, Slutsky's paper was available only in its original Russian version; but its contents became widely known within a much shorter period.

Slutsky posed the same question as did Yule, and in much the same manner. Was it possible, he asked, that a definite structure of a connection between chaotically random elements could form them into a system of more or less regular waves? Slutsky proceeded to demonstrate this possibility by methods which were partly analytic and partly inductive. He discriminated between coherent series whose elements were serially correlated and incoherent or purely random series of the sort which we have described as white noise. As to the coherent series, he declared that

> their origin may be extremely varied, but it seems probable that an especially prominent role is played in nature by the process of *moving summation* with weights of one kind or another; by this process coherent series are obtained from other coherent series or from incoherent series.

By taking, as his basis, a purely random series obtained by the People's Commissariat of Finance in drawing the numbers of a government lottery loan, and by repeatedly taking moving summations, Slutsky was able to generate a series which closely mimicked an index, of a distinctly undulatory nature, of the English business cycle from 1855 to 1877.

The general form of Slutsky's moving summation can be expressed by writing

$$y(t) = \mu_0\varepsilon(t) + \mu_1\varepsilon(t-1) + \cdots + \mu_q\varepsilon(t-q), \tag{6}$$

where $\varepsilon(t)$ is a white-noise process. This is nowadays called a qth-order moving-average process, and it is readily compared to an autoregressive process of the sort depicted under (5). The more general pth-order autoregressive process can be expressed by writing

$$\alpha_0 y(t) + \alpha_1 y(t-1) + \cdots + \alpha_p y(t-p) = \varepsilon(t). \tag{7}$$

Thus, whereas the autoregressive process depends upon a linear combination of the function $y(t)$ with its own lagged values, the moving-average process depends upon a similar combination of the function $\varepsilon(t)$ with its lagged values. The affinity of the two sorts of process is further confirmed when it is recognised that an autoregressive process of finite order is equivalent to a moving-average process of infinite order and that, conversely, a finite-order moving-average process is just an infinite-order autoregressive process.

Generalised Harmonic Analysis

The next step to be taken in the development of the theory of time series was to generalise the traditional method of periodogram analysis in such a way as to overcome the problems which arise when the model depicted under (1) is clearly inappropriate.

At first sight, it would not seem possible to describe a covariance-stationary process, whose only regularities are statistical ones, as a linear combination of perfectly regular periodic components. However any difficulties which we might envisage can be overcome if we are prepared to accept a description which is in terms of a nondenumerable infinity of periodic components. Thus, on replacing the so-called Fourier sum within equation (1) by a Fourier integral, and by deleting the term $\varepsilon(t)$, whose effect is now absorbed by the integrand, we obtain an expression in the form of

$$y(t) = \int_0^{\pi} \big\{ \cos(\omega t) dA(\omega) + \sin(\omega t) dB(\omega) \big\}. \tag{8}$$

Here we write $dA(\omega)$ and $dB(\omega)$ rather than $\alpha(\omega)d\omega$ and $\beta(\omega)d\omega$ because there can be no presumption that the functions $A(\omega)$ and $B(\omega)$ are continuous. As it stands, this expression is devoid of any statistical interpretation. Moreover, if we are talking of only a single realisation of the process $y(t)$, then the generalised functions $A(\omega)$ and $B(\omega)$ will reflect the unique peculiarities of that realisation and will not be amenable to any systematic description.

However, a fruitful interpretation can be given to these functions if we consider the observable sequence $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ to be a particular realisation which has been drawn from an infinite population representing all possible realisations of the process. For, if this population is subject to statistical regularities, then it is reasonable to regard $dA(\omega)$ and $dB(\omega)$ as mutually uncorrelated random variables with well-defined distributions which depend upon the parameters of the population.

We may therefore assume that, for any value of ω,

$$\begin{aligned} &E\{dA(\omega)\} = E\{dB(\omega)\} = 0 \qquad \text{and} \\ &E\{dA(\omega)dB(\omega)\} = 0. \end{aligned} \tag{9}$$

Moreover, to express the discontinuous nature of the generalised functions, we assume that, for any two values ω and λ in their domain, we have

$$E\{dA(\omega)dA(\lambda)\} = E\{dB(\omega)dB(\lambda)\} = 0, \tag{10}$$

which means that $A(\omega)$ and $B(\omega)$ are stochastic processes—indexed on the frequency parameter ω rather than on time—which are uncorrelated in non-overlapping intervals. Finally, we assume that $dA(\omega)$ and $dB(\omega)$ have a common variance so that

$$V\{dA(\omega)\} = V\{dB(\omega)\} = dG(\omega). \tag{11}$$

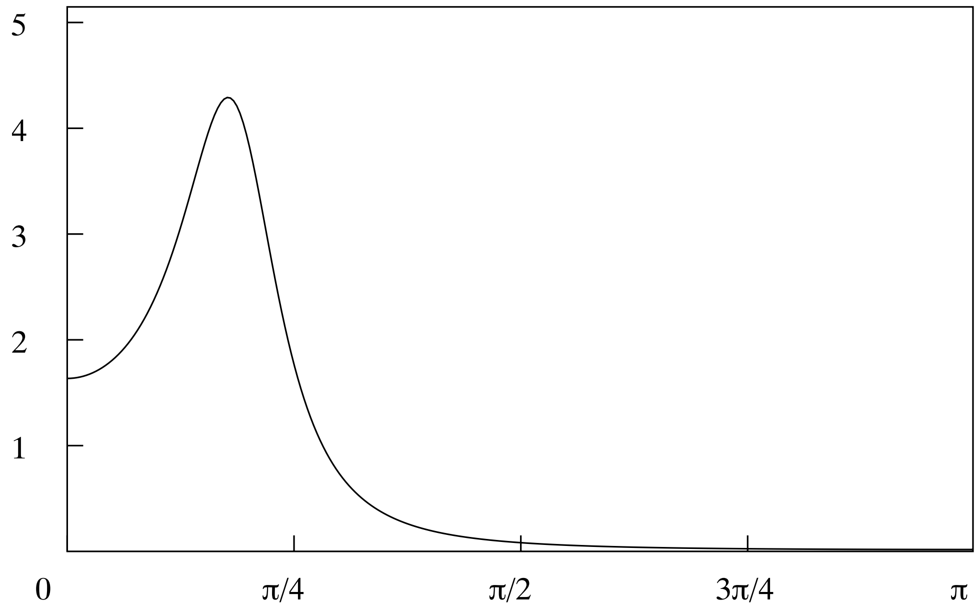

Figure 6. The spectrum of the process $y(t) = 1.343y(t-1) - 0.655y(t-2) + \varepsilon(t)$ which generated the series in Figure 4. A series of a more regular nature would be generated if the spectrum were more narrowly concentrated around its modal value.

Given the assumption of the mutual uncorrelatedness of $dA(\omega)$ and $dB(\omega)$, it therefore follows from (8) that the variance of $y(t)$ is expressible as

$$\begin{aligned} V\{y(t)\} &= \int_0^{\pi} \big[\cos^2(\omega t)V\{dA(\omega)\} + \sin^2(\omega t)V\{dB(\omega)\}\big] \\ &= \int_0^{\pi} dG(\omega). \end{aligned} \tag{12}$$

The function $G(\omega)$, which is called the spectral distribution, tells us how much of the variance is attributable to the periodic components whose frequencies range continuously from 0 to ω. If none of these components contributes more than an infinitesimal amount to the total variance, then the function $G(\omega)$ is absolutely continuous, and we can write $dG(\omega) = g(\omega)d\omega$ under the integral of equation (11). The new function $g(\omega)$, which is called the spectral density function or the spectrum, is directly analogous to the function expressing the squared amplitude which is associated with each component in the simple harmonic model discussed in our earlier sections.

Smoothing the Periodogram

It might be imagined that there is little hope of obtaining worthwhile estimates of the parameters of the population from which the single available realisation $y(t)$ has been drawn. However, provided that $y(t)$ is a stationary process, and provided that the statistical dependencies between widely separated elements are weak, the single realisation contains all the information which is necessary for the estimation of the spectral density function. In fact, a modified version of the traditional periodogram analysis is sufficient for the purpose of estimating the spectral density.

In some respects, the problems posed by the estimation of the spectral density are similar to those posed by the estimation of a continuous probability density function of unknown functional form. It is fruitless to attempt directly to estimate the ordinates of such a function. Instead, we might set about our task by constructing a histogram or bar chart to show the relative frequencies with which the observations that have been drawn from the distribution fall within broad intervals. Then, by passing a curve through the mid points of the tops of the bars, we could construct an envelope that might approximate to the sought-after density function. A more sophisticated estimation procedure would not group the observations into the fixed intervals of a histogram; instead it would record the number of observations falling within a moving interval. Moreover, a consistent method of estimation, which aims at converging upon the true function as the number of observations increases, would vary the width of the moving interval with the size of the sample, diminishing it sufficiently slowly as the sample size increases for the number of sample points falling within any interval to increase without bound.

A common method for estimating the spectral density is very similar to the one which we have described for estimating a probability density function. Instead of basing itself on raw sample observations as does the method of density-function estimation, it bases itself upon the ordinates of a periodogram which has been fitted to the observations on $y(t)$. This procedure for spectral estimation is therefore called smoothing the periodogram.

A disadvantage of the procedure, which for many years inhibited its widespread use, lies in the fact that calculating the periodogram by what would seem to be the obvious methods by can be vastly time-consuming. Indeed, it was not until the mid 1960's that wholly practical computational methods were developed.

The Equivalence of the Two Domains

It is remarkable that such a simple technique as smoothing the periodogram should provide a theoretical resolution to the problems encountered by Beveridge and others in their attempts to detect the hidden periodicities in economic and astronomical data. Even more remarkable is the way in which the generalised harmonic analysis that gave rise to the concept of the spectral density of a time series should prove itself to be wholly conformable with the alternative methods of time-series analysis in the time domain which arose largely as a consequence of the failure of the traditional methods of periodogram analysis.

The synthesis of the two branches of time-series analysis was achieved independently and almost simultaneously in the early 1930's by Norbert Wiener [24] in America and A. Khintchine [13] in Russia. The Wiener–Khintchine theorem indicates that there is a one-to-one relationship between the autocovariance function of a stationary process and its spectral density function. The relationship is expressed, in one direction, by writing,

$$g(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau \cos(\omega\tau) \quad ; \quad \gamma_\tau = \gamma_{-\tau}, \tag{13}$$

where $g(\omega)$ is the spectral density function and $\{\gamma_\tau; \tau = 0, 1, 2, \ldots\}$ is the sequence of the autocovariances of the series $y(t)$.

The relationship is invertible in the sense that it is equally possible to express each of the autocovariances as a function of the spectral density:

$$\gamma_\tau = \int_{\omega=0}^{\pi} \cos(\omega\tau) g(\omega) d\omega. \tag{14}$$

If we set $\tau = 0$, then $\cos(\omega\tau) = 1$, and we obtain, once more, the equation (12) which neatly expresses the way in which the variance $\gamma_0 = V\{y(t)\}$ of the series $y(t)$ is attributable to the constituent harmonic components; for $g(\omega)$ is simply the expected value of the squared amplitude of the component at frequency ω.

We have stated the relationships of the Wiener–Khintchine theorem in terms of the theoretical spectral density function $g(\omega)$ and the true autocovariance function $\{\gamma_\tau; \tau = 0, 1, 2, \ldots\}$. An analogous relationship holds between the periodogram $I(\omega_j)$ defined in (2) and the sample autocovariance function $\{c_\tau; \tau = 0, 1, \ldots, T-1\}$ where $c_\tau = \sum(y_t - \bar{y})(y_{t-\tau} - \bar{y})/T$. Thus, in the appendix, we demonstrate the identity

$$I(\omega_j) = 2 \sum_{t=1-T}^{T-1} c_\tau \cos(\omega_j \tau) \quad ; \quad c_\tau = c_{-\tau}. \tag{15}$$

The upshot of the Wiener–Khintchine theorem is that many of the techniques of time-series analysis can, in theory, be expressed in two mathematically equivalent ways which may differ markedly in their conceptual qualities.

Often, a problem which appears to be intractable from the point of view of one of the domains of time-series analysis becomes quite manageable when translated into the other domain. A good example is provided by the matter of spectral estimation. Given that there are difficulties in computing all T of the ordinates of the periodogram when the sample size is large, we are impelled to look for a method of spectral estimation which depends not upon smoothing the periodogram but upon performing some equivalent operation upon the sequence of autocovariances. The fact that there is a one-to-one correspondence between the spectrum and the sequence of autocovariances assures us that this equivalent operation must exist; though there is, of course, no guarantee that it will be easy to perform.

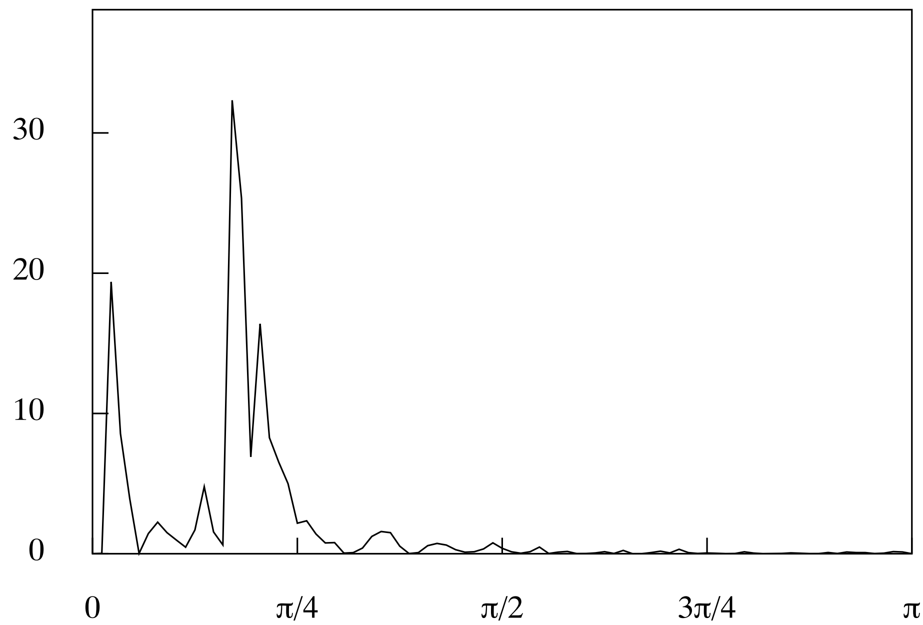

Figure 7. The periodogram of Wolfer's Sunspot Numbers 1749–1924.

In fact, the operation which we perform upon the sample autocovariances is simple. For, if the sequence of autocovariances $\{c_\tau; \tau = 0, 1, \ldots, T-1\}$ in (15) is replaced by a modified sequence $\{w_\tau c_\tau; \tau = 0, 1, \ldots, T-1\}$ incorporating a specially devised set of declining weights $\{w_\tau; \tau = 0, 1, \ldots, T-1\}$, then an effect which is much the same as that of smoothing the periodogram can be achieved. Moreover, it may be relatively straightforward to calculate the weighted autocovariance function.

The task of devising appropriate sets of weights provided a major research topic in time-series analysis in the 1950's and early 1960's. Together with the task of devising equivalent procedures for smoothing the periodogram, it came to be known as spectral carpentry.

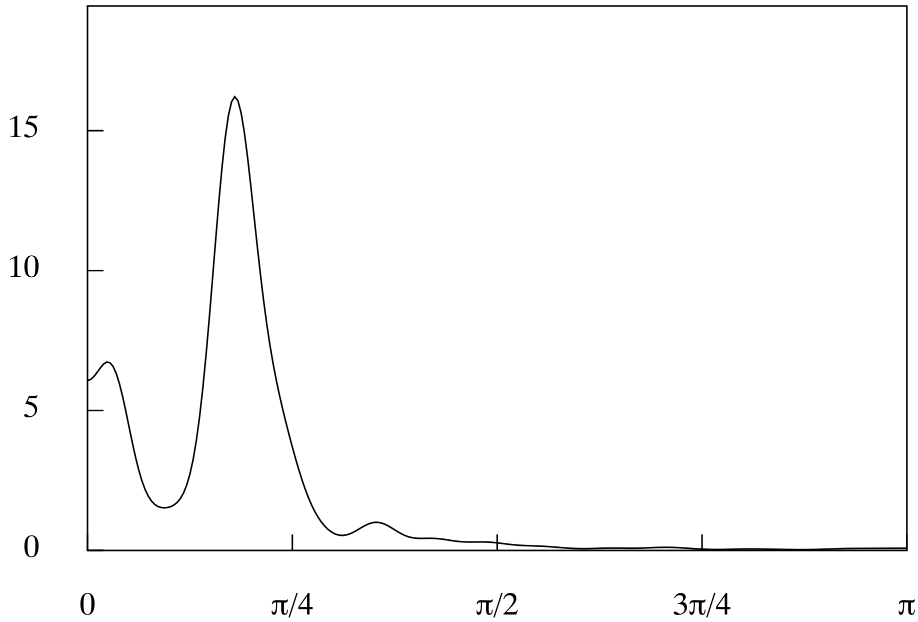

Figure 8. The spectrum of the sunspot numbers calculated from the autocovariances using Parzen's [16] system of weights.

The Maturing of Time-Series Analysis

In retrospect, it seems that time-series analysis reached its maturity in the 1970's when significant developments occurred in both of its domains.

A major development in the frequency domain occurred when Cooley and Tukey [7] described an algorithm which greatly reduces the effort involved in computing the periodogram. The Fast Fourier Transform, as this algorithm has come to be known, allied with advances in computer technology, has enabled the routine analysis of extensive sets of data; and it has transformed the procedure of smoothing the periodogram into a practical method of spectral estimation.

The contemporaneous developments in the time domain were influenced by an important book by Box and Jenkins [5]. These authors developed the time-domain methodology by collating some of its major themes and by applying it to such important functions as forecasting and control. They demonstrated how wide had become the scope of time-series analysis by applying it to problems as diverse as the forecasting of airline passenger numbers and the analysis of combustion processes in a gas furnace. They also adapted the methodology to the computer.

Many of the current practitioners of time-series analysis have learnt their skills in recent years during a time when the subject has been expanding rapidly. Lacking a longer perspective, it is difficult for them to gauge the significance of the recent practical advances. One might be surprised to hear, for example, that as late as 1971 Granger and Hughes [9] were capable of declaring that Beveridge's calculation of the Periodogram of the Wheat Price Index, comprising 300 ordinates, was the most extensive calculation of its type to date. Nowadays, computations of this order are performed on a routine basis using microcomputers containing specially designed chips which are dedicated to the purpose.

The rapidity of the recent developments also belies the fact that time-series analysis has had a long history. The frequency domain of time-series analysis, to which the idea of the harmonic decomposition of a function is central, is an inheritance from Euler (1707–1783), d'Alembert (1717–1783), Lagrange (1736–1813) and Fourier (1768–1830). The search for hidden periodicities was a dominant theme of 19th century science. It has been transmogrified through the refinements of Wiener's Generalised Harmonic Analysis which has enabled us to understand how cyclical phenomena can arise out of the aggregation of random causes. The parts of time-series analysis which bear a truly 20th-century stamp are the time-domain models which originate with Slutsky and Yule and the computational technology which renders the methods of both domains practical.

The effect of the revolution in digital electronic computing upon the practicability of time-series analysis can be gauged by inspecting the purely mechanical devices (such as the Henrici–Conradi and Michelson–Stratton harmonic analysers invented in the 1890's) which were once used, with very limited success, to grapple with problems which are nowadays almost routine. These devices, some of which are displayed in London's Science Museum, also serve to remind us that many of the developments of applied mathematics which startle us with their modernity were foreshadowed many years ago.

Mathematical Appendix

Mathematical Expectations

The mathematical expectation or the expected value of a random variable y is defined by

$$E(x) = \int_{x=-\infty}^{\infty} x dF(x), \tag{i}$$

where $F(x)$ is the probability distribution function of x. The probability distribution function is defined by the expression $F(x^*) = P\{x < x^*\}$ which denotes the probability that x assumes a value less than x^*. If $F(x)$ is a differentiable function, then we can write $dF(x) = f(x)dx$ in equation (i). The function $f(x) = dF(x)/dx$ is called the probability density function.

If $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ is a stationary stochastic process, then $E(y_t) = \mu$ is the same value for all t.

If $y_0, \ldots, y_{T-1}$ is a sample of T values generated by the process, then we may estimate μ from the sample mean

$$\bar{y} = \frac{1}{T}\sum_{t=0}^{T-1} y_t. \tag{ii}$$

Autocovariances

The autocovariance of lag τ of the a stationary stochastic process $y(t)$ is defined by

$$\gamma_\tau = E\{(y_t - \mu)(y_{t-\tau} - \mu)\}. \tag{iii}$$

The autocovariance of lag τ provides a measure of the relatedness of the elements of the sequence $y(t)$ which are separated by τ time periods.

The variance, which is denoted by $V\{y(t)\} = \gamma_0$ and defined by

$$\gamma_0 = E\left\{(y_t - \mu)^2\right\}, \tag{iv}$$

is a measure of the dispersion of the elements of $y(t)$. It is formally the autocovariance of lag zero.

If y_t and $y_{t-\tau}$ are statistically independent, then their joint probability density function is the product of their individual probability density functions so that $f(y_t, y_{t-\tau}) = f(y_t)f(y_{t-\tau})$. It follows that

$$\gamma_\tau = E(y_t - \mu)E(y_{t-\tau} - \mu) = 0 \quad \text{for all} \quad \tau \neq 0. \tag{v}$$

If $y_0, \ldots, y_T$ is a sample from the process, and if $\tau < T$, then we may estimate γ_τ from the sample autocovariance or empirical autocovariance of lag τ:

$$c_\tau = \frac{1}{T}\sum_{t=\tau}^{T-1}(y_t - \bar{y})(y_{t-\tau} - \bar{y}). \tag{vi}$$

The periodogram and the autocovariance function

The periodogram is defined by

$$I(\omega_j) = \frac{2}{T}\left[\left\{\sum_{t=0}^{T-1}\cos(\omega_j t)(y_t - \bar{y})\right\}^2 + \left\{\sum_{t=0}^{T-1}\sin(\omega_j t)(y_t - \bar{y})\right\}^2\right]. \tag{vii}$$

The identity $\sum_t \cos(\omega_j t)(y_t - \bar{y}) = \sum_t \cos(\omega_j t)y_t$ follows from the fact that, by construction, $\sum_t \cos(\omega_j t) = 0$ for all j. Hence the above expression has the same value as the expression in (2). Expanding the expression in (vii) gives

$$
\begin{aligned}
I(\omega_j) =& \frac{2}{T}\left\{\sum_t\sum_s \cos(\omega_j t)\cos(\omega_j s)(y_t - \bar{y})(y_s - \bar{y})\right\}\\
&+ \frac{2}{T}\left\{\sum_t\sum_s \sin(\omega_j t)\sin(\omega_j s)(y_t - \bar{y})(y_s - \bar{y})\right\},
\end{aligned}
\tag{viii}
$$

and, by using the identity $\cos(A)\cos(B) + \sin(A)\sin(B) = \cos(A - B)$, we can rewrite this as

$$
I(\omega_j) = \frac{2}{T}\left\{\sum_t\sum_s \cos(\omega_j[t - s])(y_t - \bar{y})(y_s - \bar{y})\right\}.
\tag{ix}
$$

Next, on defining $\tau = t - s$ and writing $c_\tau = \sum_t (y_t - \bar{y})(y_{t-\tau} - \bar{y})/T$, we can reduce the latter expression to

$$
I(\omega_j) = 2\sum_{\tau=1-T}^{T-1} \cos(\omega_j\tau)c_\tau,
\tag{x}
$$

which appears in the text as equation (15).

Bibliography

[1] Alberts, W.W., L.E. Wright and B. Feinstein (1965), Physiological Mechanisms of Tremor and Rigidity in Parkinsonism, *Confinia Neurologica, 26,* 1965.

[2] Beveridge, Sir W.H., (1921), Weather and Harvest Cycles, *Economic Journal,* **31,** 429–452.

[3] Beveridge, Sir W.H., (1922), Wheat Prices and Rainfall in Western Europe, *Journal of the Royal Statistical Society,* 85, 412–478.

[4] Box, G.E.P., and D.R. Cox (1964), An Analysis of Transformations, *Journal of the Royal Statistical Society, Series B,* 26, 211–243.

[5] Box, G.E.P., and G.M. Jenkins (1970), *Time Series Analysis, Forecasting and Control,* Holden–Day, San Francisco.

[6] Buijs–Ballot, C.D.H., (1847), *Les Changements Périodiques de Température,* Utrecht.

[7] Cooley, J.W., and J.W. Tukey (1965), An Algorithm for the Machine Calculation of Complex Fourier Series, *Mathematics of Computation,* **19,** 297–301.

[8] Deistler, M., O. Prohaska, E. Reschenhofer and R. Volmer, (1986), Procedure for the Identification of Different Stages of EEG Background and its Application to the Detection of Drug Effects, *Electroencephalography and Clinical Neurophysiology,* **64**, 294–300.

[9] Granger, C.W.J., and A.O. Hughes (1971), A New Look at Some Old Data: The Beveridge Wheat Price Series, *Journal of the Royal Statistical Society, Series A,* **134,** 413–428.

[10] Groves, G.W., and E.J. Hannan, (1968), Time-Series Regression of Sea Level on Weather, *Review of Geophysics,***6,** 129–174.

[11] Gudmundsson, G., (1971), Time-series Analysis of Imports, Exports and other Economic Variables, *Journal of the Royal Statistical Society, Series A,* **134,** 383–412.

[12] Hassleman, K., W. Munk and G. MacDonald, (1963), Bispectrum of Ocean Waves, in *Time Series Analysis,*, M. Rosenblatt, (ed.) 125–139. John Wiley and Sons, New York.

[13] Khintchine, A., (1934), Korrelationstheorie der Stationären Stochastischen Prozessen, *Math. Ann.,* **109,** 604–615.

[14] Lagrange, Joseph Louis, (1772, 1778), *Oeuvres, 14 vols.,* Gauthier Villard, Paris, 1867–1892

[15] Moore, H.L., (1914), *Economic Cycles: Their Laws and Cause,* Macmillan, New York.

[16] Parzen, E., (1957), On Consistent Estimates of the Spectrum of a Stationary Time Series, *Annals of Mathematical Statistics,* **28,** 329–348.

[17] Rice, S.O., (1963), Noise in FM Receivers, in *Time Series Analysis,* M. Rosenblatt, (ed.) 395–422. John Wiley and Sons, New York.

[18] Schuster, Sir A., (1898), On the Investigation of Hidden Periodicities with Application to a Supposed Twenty-Six Day Period of Meteorological Phenomena, *Terrestrial Magnetism,* **3,** 13–41.

[19] Schuster, Sir A., (1906), On the Periodicities of Sunspots, *Philosophical Transactions of the Royal Society, Series A,* **206,** 69–100.

[20] Slutsky, E., (1937), The Summation of Random Causes as the Source of Cyclical Processes, *Econometrica,* **5,** 105–146.

[21] Tee, L.H., and S.U. Wu (1972), An Application of Stochastic and Dynamic Models for the Control of a Papermaking Process, *Technometrics,* **14,** 481–496.

[22] Tukey, J.W., (1965), Data Analysis and the Frontiers of Geophysics, *Science,* **148,** 1283–1289.

[23] Whittaker, E.T., and G. Robinson (1924), *The Calculus of Observations, A Treatise on Numerical Mathematics,* Blackie and Sons, London.

[24] Wiener, N., (1930), Generalised Harmonic Analysis, *Acta Mathematica,* **35,** 117–258.

[25] Yule, G.U., (1927), On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers, *Philosophical Transactions of the Royal Society, Series A* **226,** 267–298.

[26] Yuzuriha, T., (1960), The Autocorrelation Curves of Schizophrenic Brain Waves and the Power Spectrum, *Psych. Neurol. Jap.,* **26,** 911–924.