NBS TECHNICAL NOTE 1022

U.S. DEPARTMENT OF COMMERCE / National Bureau of Standards

Sunspot Cycle Simulation Using a Narrowband Gaussian Process

NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, and the Institute for Computer Sciences and Technology.

THE NATIONAL MEASUREMENT LABORATORY provides the national system of physical and chemical and materials measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; conducts materials research leading to improved methods of measurement, standards, and data on the properties of materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

Absolute Physical Quantities[2] — Radiation Research — Thermodynamics and Molecular Science — Analytical Chemistry — Materials Science.

THE NATIONAL ENGINEERING LABORATORY provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

Applied Mathematics — Electronics and Electrical Engineering[2] — Mechanical Engineering and Process Technology[2] — Building Technology — Fire Research — Consumer Product Technology — Field Methods.

THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

Programming Science and Technology — Computer Systems Engineering.

[1]Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted;
mailing address Washington, DC 20234.
[2]Some divisions within the center are located at Boulder, CO 80303.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

Sunspot Cycle Simulation Using a Narrowband Gaussian Process

J.A. Barnes

H.H. Sargent III

P.V. Tryon

Time and Frequency Division
National Measurement Laboratory
National Bureau of Standards
Boulder, Colorado 80303

U.S. DEPARTMENT OF COMMERCE, Philip M. Klutznick, Secretary

Luther H. Hodges, Jr., Deputy Secretary
Jordan J. Baruch, Assistant Secretary for Productivity, Technology and Innovation

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director

Issued September 1980

NATIONAL BUREAU OF STANDARDS TECHNICAL NOTE 1022

Nat. Bur. Stand. (U.S.), Tech. Note 1022, 24 pages (Sept. 1980)
CODEN: NBTNAE

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1980

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402

Price $1.75 (Add 25 percent additional for other than U.S. mailing)

SUNSPOT CYCLE SIMULATION USING
A NARROWBAND GAUSSIAN PROCESS

J. A. Barnes, H. H. Sargent III, and P. V. Tryon

The square of a narrowband Gaussian process is used to simulate sunspot cycles at computer speeds. The method is appealing because: (i) the model is extremely simple yet its physical basis, a simple resonance, is a widely occurring natural phenomenon, and (ii) the model recreates practically all of the features of the observed sunspot record. In particular, secular cycles and recurring extensive minima are characteristic of narrowband Gaussian processes. Additionally, the model lends itself to limited prediction of sunspot cycles.

Key words: ARMA models; forecasts; Maunder minimum, models; simulation; statistics; sunspots.

1. Introduction

Since the discovery of the cyclic behavior of sunspots by Schwabe in 1843, many authors have referred to the sunspot record as an example of naturally occurring periodic behavior - not easily explained by the dynamics of rotating systems. Yule [1] characterized the sunspot numbers as a "disturbed harmonic function," which he likened to the motion of a pendulum that boys are pelting with peas. Time series analysis texts [2] and statistical works [3] commonly cite the sunspot number series as a function that is more or less periodic. The noisy, but nearly periodic, character of the sunspot record suggests a very simple model of solar activity that simulates the observed sunspot numbers to a surprising degree. The observed annual mean sunspot numbers [4] and simulated annual mean sunspot numbers (produced using methods described in this paper) are shown in figure 1.

Our model, in its simplest form, is the squared output of a narrowband filter driven by white Gaussian noise. If elaborate filtering schemes are used, it is possible to mimic the stochastic properties of <u>any</u> existing signal. In other words, given sufficient resources, it is possible to mimic almost any existing signal. Suppose, on the other hand, a white noise signal (perhaps the simplest of signals) is filtered by a simple narrowband filter and the squared output signal resembles a complicated existing signal in all its gross characteristics - does this not compel one to give serious consideration to the physical implications of the stochastic process?

We see no reason to be uncomfortable with the suggestion that some Gaussian-noise process may be at work in the interior of the Sun. White Gaussian-noise is common in the Universe at all levels from the microscopic to the macroscopic; whether we consider the noise produced in thermionic emission in an electron tube or the noise received from some distant radio galaxy. By the same token, it is not discomforting to consider that the Sun might have resonant modes which act as filters. Many of the physical objects in our everyday lives exhibit the properties of filters. That is, they respond to certain modes of excitation and they are to a greater or lesser extent resonators. Is it not natural then to expect that the massive solar body with great mechanical, thermal and gravitational forces at work has its own natural modes of response that cause it to behave as a filter? Indeed, calculations of the solar thermal diffusion constant [5] indicate that where solar luminosity is concerned, the Sun does act like a low pass filter.

It is well known that if a narrowband resonant filter is driven by white Gaussian-noise, the resulting output signal has a nearly periodic behavior with slowly but randomly varying amplitude and

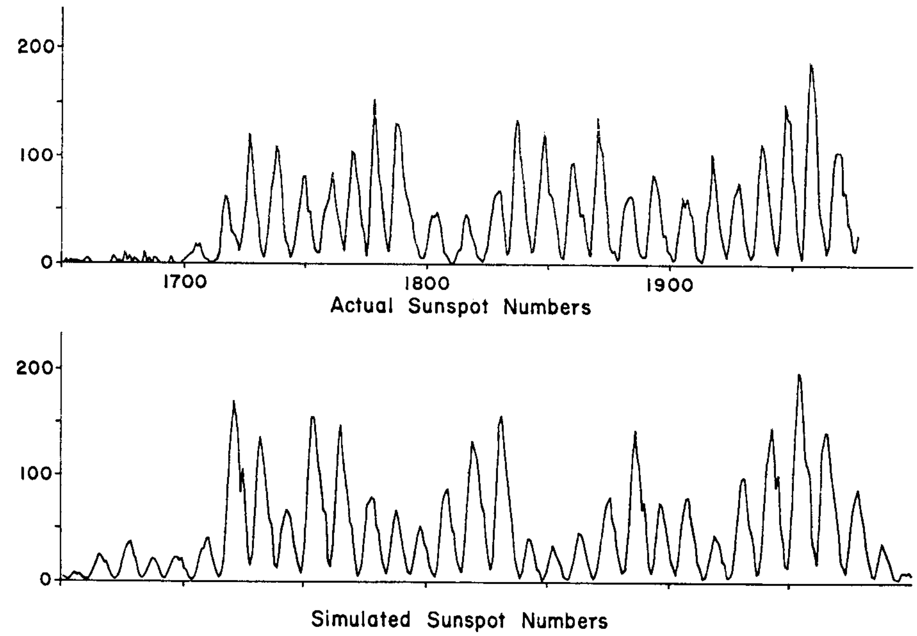

Figure 1.

Annual mean sunspot numbers from 1650 to 1977 compared with simulated annual mean sunspot numbers generated by the resonant model.

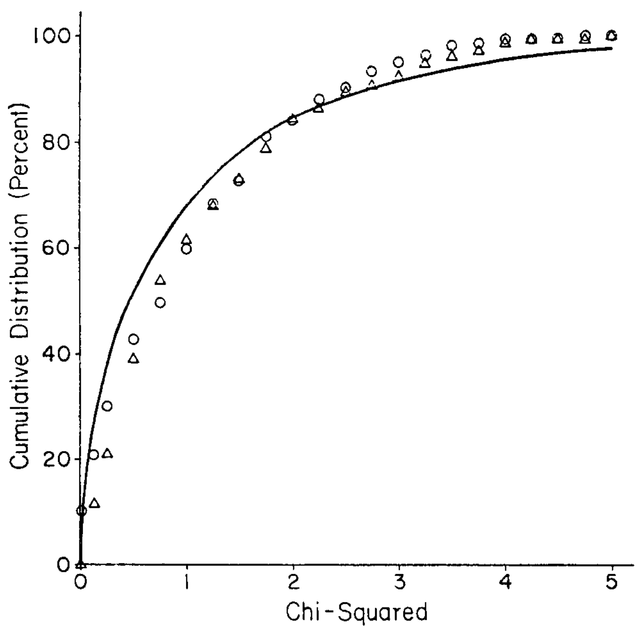

Figure 2.

Cumulative distributions for observed annual mean sunspot numbers (circles) and simulated annual mean sunspot numbers (triangles) divided by their means for 328 years. The solid curve is the Chi-square distribution with one degree of freedom.

phase. A filter output swings symmetrically in both positive and negative directions, but the Wolf sunspot numbers are a positive number set and are clearly not symmetric about their average value. Therefore, we chose the square of narrowband noise as the principal component of our model. This is consistent with the idea that the sunspot numbers are a measure of the magnitude of the 22-year solar magnetic cycle and with the use of square root transformation for improving the symmetry of sunspot cycles suggested by Bloomfield [6]. Since the narrowband filter is linear and is driven by Gaussian-noise, the output of the filter must also have a Gaussian distribution. The ratio of the squared filter output to its mean then must be distributed as Chi-square with one degree of freedom. If our decision to square the filter output permits realistic simulation of sunspot cycles, the cumulative distribution of the observed sunspot numbers divided by their average should be a reasonably good approximation to the Chi-square distribution. Figure 2 shows the empirical cumulative distributions of the observed and simulated sunspot numbers produced by our complete model (which contains refinements still to be discussed) together with the χ^2 distribution function. The simulated and actual sunspot distributions appear to approximate each other even better than either approximates the Chi-square distribution!

2. The Basic Model

The basic model is the square of a narrowband Gaussian process, which is generated by passing white noise through a single stage resonant filter. For convenience in numerical simulation, digital filters were used based on ARMA (AutoRegressive, Moving Average) models [7]. Figure 3 is a block diagram of the sunspot number simulation method. Equations used for simulation are as follows:

a_n are independent normal random deviates with zero mean and variance σ_a^2

$$z_n = a_n = 0 \quad \text{for } n < 1 \tag{1}$$

$$z_n = \phi_1 z_{n-1} + \phi_2 z_{n-2} + a_n - \theta_1 a_{n-1} - \theta_2 a_{n-2} \tag{2}$$

where $n = 1,2, \ldots$ counts the years. The ϕ_i's and θ_i's forming the AR and MA parts of the model respectively, and σ_a are constants. The simulated series is then

$$x_n = z_n^2. \tag{3}$$

The AR portion of the model alone provides the desired resonant filter. Simulations using the AR part of the model showed many similarities to the observed record. There were numerous Gleissburg cycles [8] and Eddy minima [9] present in long runs of simulated data. However, the cycle-to-cycle variation was much smoother than in the observed record. This is due to inadequate broadband noise levels. The MA portion of the model adds and shapes the spectrum of the broadband noise.

The broadband noise represents "measurement error" in its broadest sense. The Wolf sunspot number is itself only an indicator of solar activity. The formation of sunspots may well have a significant stochastic component that makes the Wolf number a noisy indicator. Abrupt large changes in the day-to-day numbers are possible (but not common). The changes result from the appearance, or disappearance, of spots and spot groups on the disk or at the limbs. While one 27-day span of daily sunspot numbers often bears some resemblance to the next (due to persistence of active longitudes on a rotating sun), the raw monthly mean sunspot numbers may vary wildly from one month to the next.

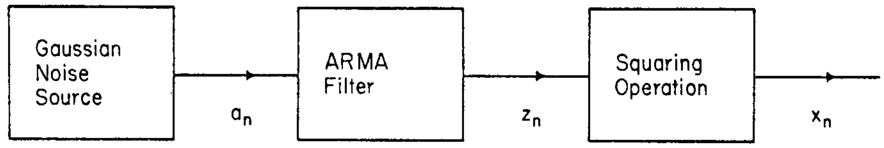

Figure 3.

Block diagram of the sunspot number simulator in its simplest form.

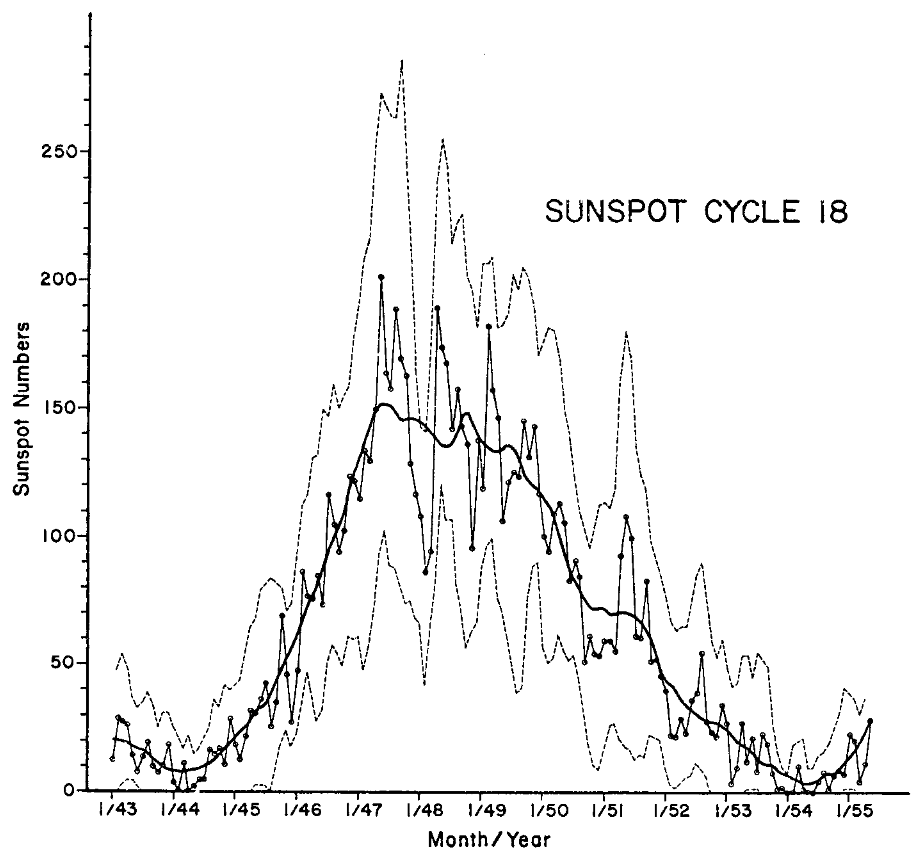

Figure 4.

Solar Cycle 18. The heavy line indicates the smoothed monthly mean sunspot numbers through the entire cycle. The connected points show the raw monthly mean sunspot numbers. The dashed lines are a slightly smoothed evelope of the maximum and minimum daily sunspot numbers observed each month during the cycle.

In addition, there are many observational errors that contribute noise. Some of these are the subjective definition of regions, the difficulties of counting spots as they form, fade out, or pass around the limbs (counting spots near the limbs is made difficult by foreshortening). Figure 4 shows the highly noisy character of the daily and monthly average sunspot number in comparison to the smoothed monthly numbers often displayed (a 13-month moving average was used for smoothing).

The following values for the parameters were determined by methods described below:

$$\phi_1 = 1.90693 \qquad \phi_2 = -0.98751 \tag{4}$$
$$\theta_1 = 1.20559 \qquad \theta_2 = -0.62432$$
$$\sigma_a = 0.633$$

Since the parameters ϕ_1, ϕ_2, θ_1, and θ_2 interact with each other, the number of significant digits given here is very large relative to their standard errors. Dropping digits can materially alter the model beyond what one might normally expect, because roots of the "operator" equation are changed significantly. This is often an annoying feature of digital filters and does not imply exactness in the overall model.

While an ARMA digital filter is used here for numerical simulation purposes [10], it may be more informative to express the model parameters in physical terms rather than in terms of the ARMA coefficients. There are four such "physical" parameters:

1. Resonant frequency of the narrowband filter
2. Bandwidth of the resonant filter
3. RMS level of the noise signal driving the filter
4. RMS level and spectral shape of the broadband noise.

The selection of numerical values for these physical parameters is described below. The details of the reparameterization to the ARMA coefficients will be documented elsewhere.

The first parameter, resonant frequency of the narrowband filter, is taken to be 1/22 cycles per year. Squaring the filter output, z_n, then doubles the frequency of the model output signal, x_n.

The filter bandwidth is a more difficult parameter to estimate. The results provide a relatively broad range of possible values. However, three different approaches to fitting this parameter were employed, which gave consistent results.

Narrowband Gaussian noise locally appears as a sinusoid; but between cycles, the peak amplitude (envelope) and phase (determining the individual period of cycles) will be slowly varying random functions. In general, as the bandwidth becomes narrower, the amplitude and phase functions become more slowly varying. Mathematically, a narrowband process may be written as:

$$f(t) = E(t)\cos(w_c t + \rho(t)) \tag{5}$$

where E(t) and $\rho(t)$, the envelope and phase functions, are stochastic processes whose statistical properties are completely determined by the filter parameters (see Middleton or Davenport and Root [11]). We wish to point out only that features of the sunspot record, such as Gleisburg cycles and Eddy minima, are features of the envelope process that result from the strong autocorrelations caused by the narrow bandwidth. Similarly, the variability of the periods of individual cycles is related to the filter bandwidth through the stability of the phase process. These features are well-known to communications engineers and are the features that originally suggested the model to

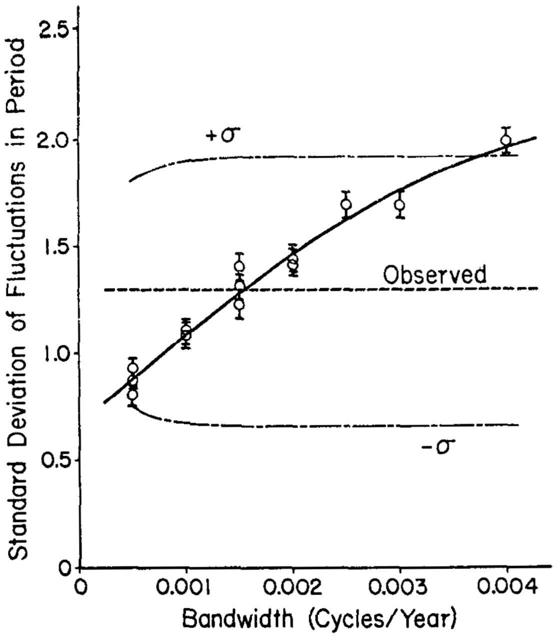

Figure 5(a).

Variability of sunspot cycle periods (rounded to integer years) for models based on various bandwidths. The circles are averages for 2500 cycles and the ±1σ lines indicate the confidence interval for averages based on samples of 25 cycles (275 years). The horizontal, dashed line is derived from observed "zero crossings" of actual sunspot cycles.

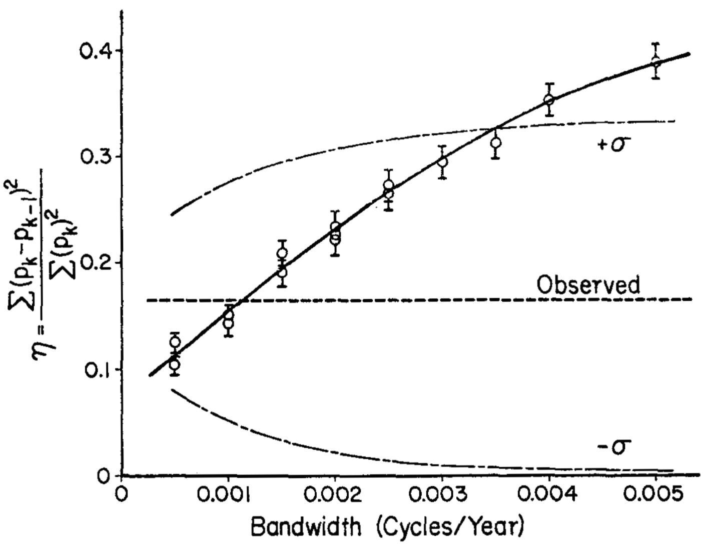

Figure 5(b).

Variability of differences of successive peak amplitudes as a function of bandwidth.

us. We have taken advantage of the relationship between bandwidth and variability in amplitude and period to empirically choose the filter bandwidth as described below.

For each of several bandwidths, a sequence of 2,500 cycles was generated and divided into 100 groups of 25 cycles (care must be used to exclude recurrent Eddy minimum periods). Integer values of the period were recorded for each cycle. Figure 5a shows the observed standard deviation of the 25 sunspot periods. The $\pm\sigma$ lines are $\pm$ one standard error of the standard deviation in a single group of 25 cycles as determined from the 100 simulated groups. The circles and error bars are the average standard deviation and standard error of the average of the 100 standard deviations within groups of 25 cycles. The observed standard deviation is consistent with a bandwidth of 0.001 to 0.002 cycles per year.

Figure 5b shows the results of a similar analysis based on normalized average square successive differences of peak values in groups of 25. Because of difficulties in choosing the peaks of the signal with the broadband noise included, the simulation was run with the AR portion of the model. The results were then checked with a full model simulation at 0.002 cycles/year. Reasonable agreement was obtained.

The third approach is to use standard Box and Jenkins [7] methods to fit the ARMA model to the actual sunspot data. The resulting fitted coefficients are:

$$\phi_1 = 1.90418 \qquad \phi_2 = -0.98642 \qquad (6)$$

$$\theta_1 = 0.63503 \qquad \theta_2 = 0.17255$$

$$\sigma_a = 1.04$$

These coefficients give a bandwidth of 0.002 cycles/year, to give (in engineering terminology) a "Q" of about 23. The center frequency is shifted to 0.046 cycles/year for a period of 21.5 years. The MA portion and the driving noise level are quite different from the empirical model given in eq (4). In the empirical model, the two noise levels were selected to produce a reasonable approximation to the power spectrum and realistic solar cycle simulators (neither too smooth nor too noisy). Standard statistical tests confirm that the fitted coefficients are different from both zero (and hence are necessary in the model) and the empirical model values.

One test of the validity of the model is to recursively solve the model equation for the sequence of a_n's (residuals) given z_n's and the parameters to see if the driving sequence is in fact white noise. The fitted model passes this test, but the empirical model does not. Its residuals show a significant, very broad, spectral peak in the region of its second harmonic. Simulations based on the fitted model, however, do not yield realistic sunspot records, being much too noisy, especially in peak amplitude variability. There is good reason for this.

Anyone who is familiar with the modern sunspot record will question the generally symmetrical appearance of the cycles produced by the ARMA model. The observed cycles (particularly the large cycles) exhibit a rapid ascent and a slower descent [12]. The ascent stage (minimum to maximum) takes four years on average, and the descent takes seven years. This is suggestive of a nonlinear phenomena which will introduce harmonics in the spectrum that will be phase-locked to the primary cycle. Second harmonics have been reported in previous spectral analysis and by Brillinger and Rosenblatt [13] using bispectral analysis, a method intended to study nonlinear effects in time series. Bloomfield [14] has discussed the phase relationship of the second harmonic.

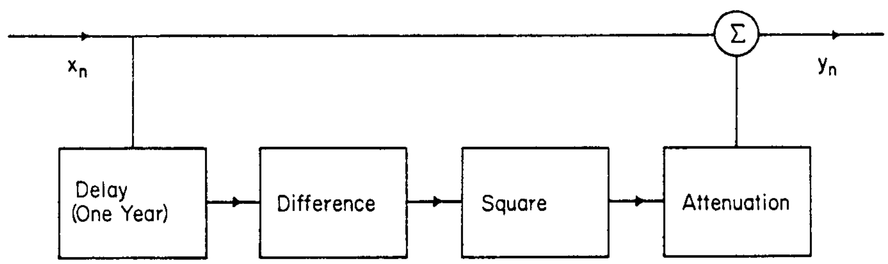

Figure 6.

Block diagram of the rise/fall correction element. The squaring circuit output is now modified by Eq. 7

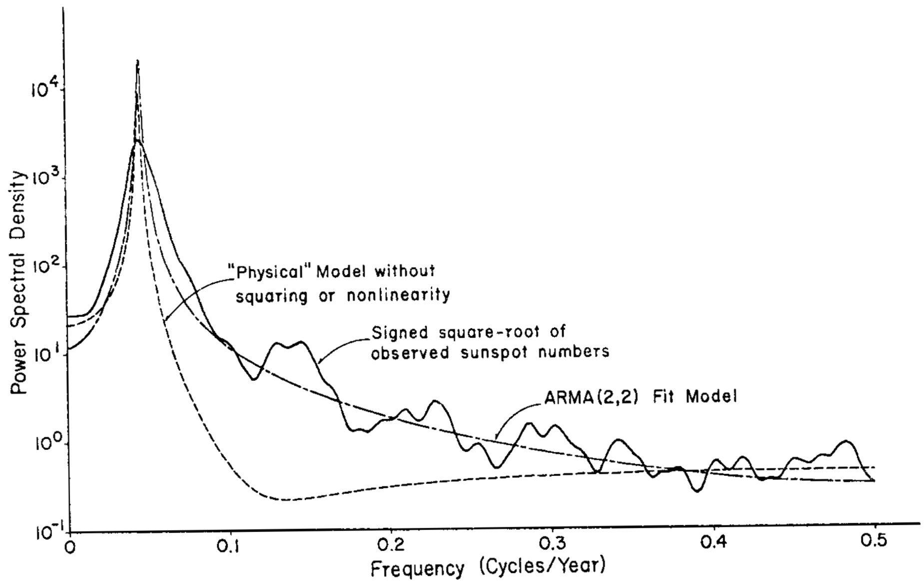

Figure 7(a).

Spectrum of the square root of the observed series with spectrums of the fitted ARMA model (Eq. 6) (broken line) and the empirical model without the nonlinear modification (Eq. 4) (dashed line) are shown.

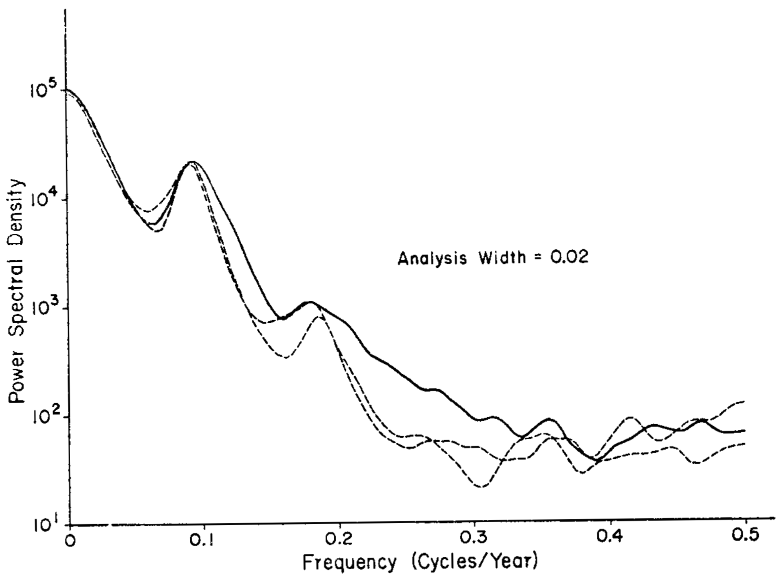

Figure 7(b)

Spectrum of the (untransformed) observed series (solid line) and spectra of two simulated series produced by the empirical (or "physical") model including the nonlinear modification (dashed lines).

In the least squares fitting process, the MA parameters are adjusted to shape the spectrum away from the resonant peak. This effectively replaces the phase-locked second harmonic with broadband noise. As will be seen in our spectral analysis, the fitted model spectrum is an excellent approximation to the data spectrum, yet for simulation purposes it is too noisy.

3. Addition of a Simple Refinement

Although the rise/fall time property of the sunspot record is perhaps a minor feature, we decided to try adding a nonlinear shaping function. This addition produced some unexpected improvements in the characteristics of the simulated sunspot data, and hence will be described here.

The shaping function, a lagged nonlinear term shown conceptually in figure 6, was added to shape the signal after the squaring operation. The squaring circuit output, x_n, is now modified by the following equation:

$$y_n = x_n + \alpha(x_{n-1} - x_{n-2})^2 \tag{7}$$

where y_n now simulates the sunspot numbers. The only new variable introduced was an amplitude parameter, α. It was also necessary to re-adjust the σ_a and the MA coefficients of the physical ARMA model eq (4). The new coefficients shown in eq (8) are those used to produce the simulated sunspot cycle records displayed in Fig. 1.

$$\begin{aligned} \phi_2 &= 1.90693 & \phi_2 &= -0.98751 \\ \theta_1 &= 0.78512 & \theta_2 &= -0.40662 \\ \sigma_a &= 0.4 & \alpha &= 0.03 \end{aligned} \tag{8}$$

This shaping function does create satisfactory rise and fall rates on simulated sunspot cycles. Figure 7(a) shows the spectrum of the square root of the observed series with the signs of successive cycles alternated, together with the theoretical spectra of the fitted ARMA model and the "physical" model before nonlinear modification. The latter model has considerably less power in the broad middle region and would appear to be unsatisfactory in comparison to the fitted version. However, figure 7(b) shows the spectrum of the (untransformed) observed series together with the spectra of two samples of simulated series using the empirical model including nonlinearity. The second harmonic is evident, the model spectrum is a good approximation to the observed spectrum, and excellent simulations of the observed sunspot cycle record are produced.

There was one other effect of adding the shaping function which we consider fortuitous. The addition created an irregularity (a bump or stand-still) on the descending slope of many of the simulated cycles. Examination of sunspot cycles since 1818 (there is reasonable confidence in the fine structure of these cycles [15]) shows a bumplike feature on the descending slopes of a number of cycles. Recently ended Solar Cycle 20 produced a noteworthy example of this phenomenon. References to this type of phenomenon are rare in solar physics literature [16]. Stobie [17] remarked on such a bumplike feature in the descending portion of the light intensity curves of certain of the Cepheid family of variable stars. Incidently, these light intensity curves rise rapidly and decay slowly; but, of course, the periods are very short with respect to the eleven-year period of our star, the Sun. The bump in the variable star light curves may be due to some overtone harmonic mode of oscillation. This is presented as a curiosity and we do not consider it a major point.

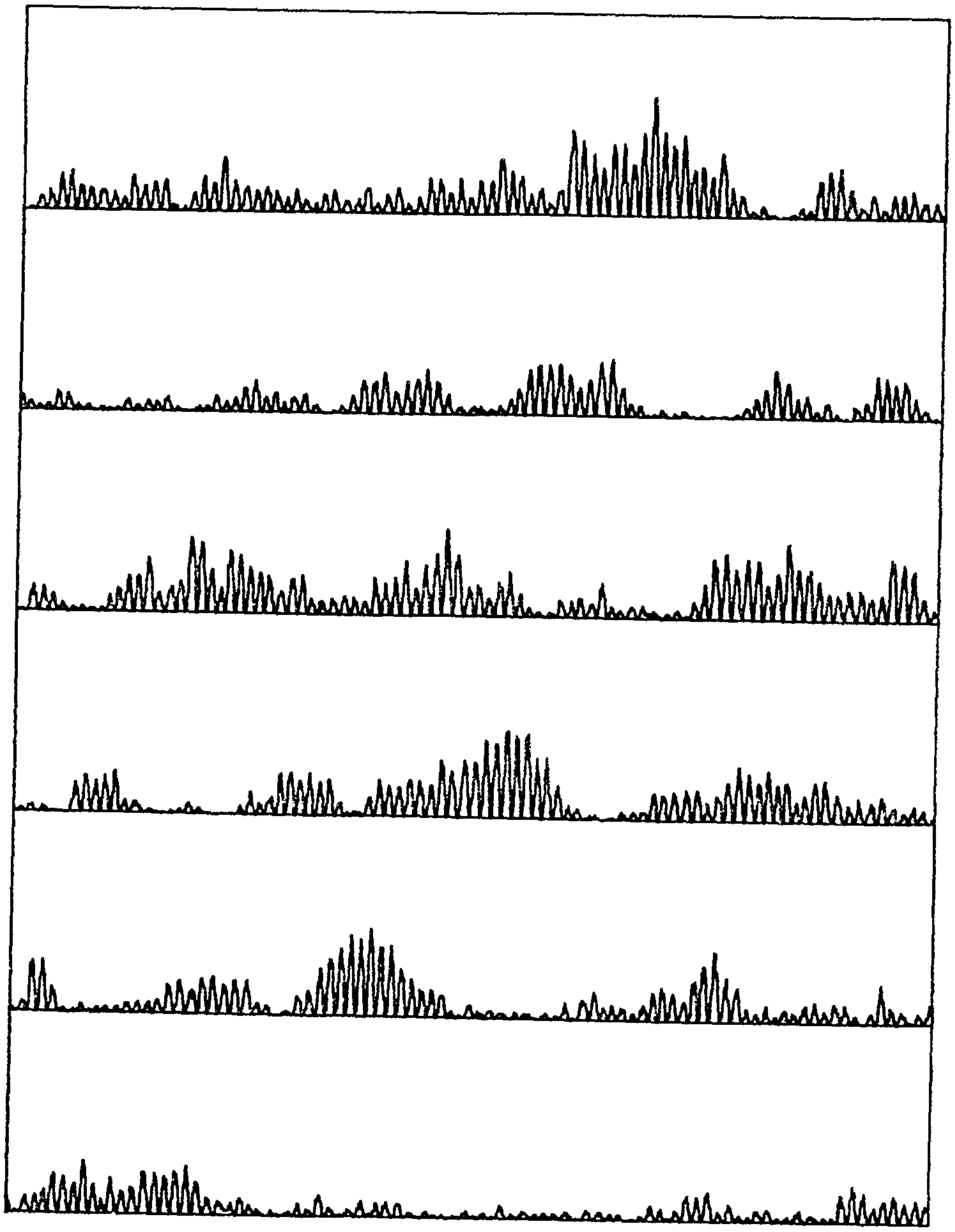

Figure 8.

6000 consecutive years of simulated annual mean sunspot numbers. Each line is 1000 years long and the lines are separated by a relative sunspot number of 400.

Recall that x_n, given by eq (3), divided by the average x_n will be distributed as Chi-square with one degree of freedom. However, if α is not equal to zero in eq (8), then $y_n/\bar{y}$ will not be distributed exactly as Chi-square. Recall also that the cumulative distribution of observed and simulated sunspot numbers approximate each other better than either approximates the Chi-square distribution.

4. Comparison of Observed and Simulated Records

Using the ARMA coefficients eq (8), the model was run to produce thousands of years of simulated sunspot numbers. The cycles shown in figure 8 are representative of this simulated data.

One has to admit that certain portions of figure 8 look very familiar and would rapidly be identified as sunspot cycles by an uncritical observer. Throughout hundreds of thousands of years of simulated cycles, there are frequent spans of data that are very reminiscent of the actually observed sunspot record. Indeed, one can find, without much trouble, patterns in cycle-to-cycle amplitude almost exactly like those described in the literature [18] and sometimes used for sunspot cycle forcasts.

A striking feature of the simulated sunspot data is the occasional (yet fairly regular) occurence of extensive sunspot minima (referred to in this paper as Eddy minima). If such a minimum is defined (arbitrarily) as a period of at least 50 years during which the annual mean sunspot number does not exceed 20, then these minima are observed to occur in the simulated series at the rate of twice per thousand years, on the average. Some extremely long minima show up in the simulations. For example, an Eddy minimum spanning more than 500 years is shown in the bottom row of figure 8.

Another feature of the simulated sunspot record is the presence, in practically every span of data, of Gleissburg cycles. If one connects successive cycle maximum values, an envelope is formed which itself appears to be cyclic in nature. Gleissburg pointed out that in the modern epoch (with maximum annual mean sunspot number seldom exceeding 150) the length of these cycles was about 80 to 90 years. Several such 80 to 90-year cycles are seen in the second row of figure 8. Gleissburg [19] used auroral data to extend his analysis and concluded that the length of his cycles was apt to vary between 55 and 121 years (5 to 11-year sunspot cycles). Rubashev [20] showed that the peak amplitude of the Gleissburg cycles appeared to be directly proportional to the duration of the cycles (the contrary appears to be true for eleven-year cycles). As discussed earlier, the parameters of the narrowband filter determine not only the size, shape and duration of the eleven-year cycles, but also the characteristics of the envelope of the cycles. This model produces very convincing Gleissburg cycles.

Perhaps the most startling feature of the simulated sunspot record was the occasional appearance of "pathologically" large sunspot cycles. None of these is shown in figure 8. During some simulation runs, infrequent "supercyles" were observed with amplitudes approaching 800 (annual mean relative sunspot number). In comparing the monthly mean sunspot numbers for the first nineteen cycles to the general extreme value probability distribution function, Siscoe [21] indicated that 100 cycles of data would be required in order to find one equaling, or exceeding, 349. Analysis of lunar rock material [22] indicates that solar flare activity, averaged over 1000-year intervals, may have varied by a factor of 50 over the past 20,000 years. Eddy [23], in discussing the Maunder and Sporer minima, mentioned the intriguing evidence in Carbon-14 data for a possible Grand Maximum in the 12th century with higher maxima and higher minima than any seen in the modern era (the past 328 years). Such large maxima are consistent with our model. The reader is reminded that the nearly Chi-square distribution shown in figure 2 supports the small (but real) probability of observing sunspot cycles of exceptionally large amplitude. Of course, there may well be mechanisms which physically limit the maximum amplitudes of sunspot cycles.

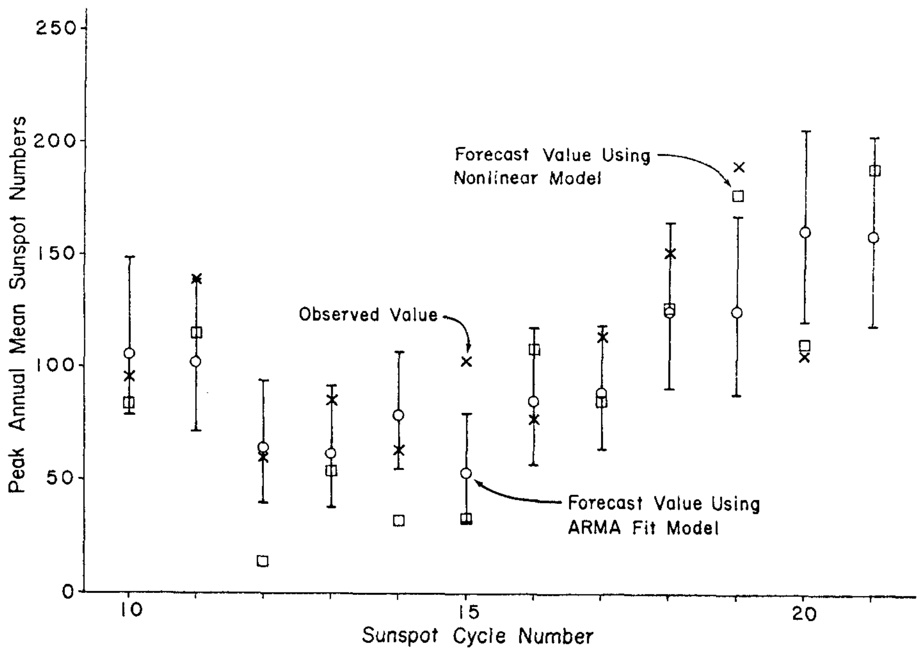

Figure 9.

Forecasts for subsequent maximum annual mean cycle values of sunspot numbers based on data to one year after previous sunspot cycle minima. Circles and error bars correspond to the ARMA forecasts using the fitted model (Eq. 6). Squares are the forecasts using the empirical model (Eq. 7 and 8). X indicates the actually observed values.

5. Forecasting Sunspot Cycles

In as much as ARMA models were developed for forecasting, the arguments presented in this paper must, inevitably, lead to a forecast for Solar Cycle 21. If one has a record of z_n, in eq (2), one can solve recursively for the values of $a_n = \hat{a}_n$ as the time series which historically has driven the model. Since the a_n are assumed to be random, independent numbers with zero mean, their optimum forecast (in a minimum mean square error sense) is just zero. Optimum forecasts for z_n, then, could be obtained by using eq (2) and:

$$a_n = \begin{cases} \hat{a}_n \text{ for } 1 < N \\ 0 \text{ for } n > N \end{cases} \tag{9}$$

where N is taken as "now" and $n > N$ implies a forecast.

We have done this using the fitted ARMA model (coefficients given by eq (6)) and approximately with the nonlinear model eq (8). Three problems arise, however. First, for the nonlinear model, the inverse of eq (7) shown here:

$$\hat{x}_n = y_n - \alpha(\hat{x}_{n-1} - \hat{x}_{n-2})^2 \tag{10}$$

is unstable and small deviations in y_n cause arbitrarlily large deviations in the computed x_n. Second, the inverse of eq (3) involves an ambiguity as to the sign of z_n. Third, even if one had an optimum forecast for z_n, this would not lead to an optimum forecast of y_n; since the square of an optimum forecast, in general, is <u>not</u> the optimum forecast of the square (for example, consider random numbers with zero mean).

These difficulties, however, do not totally prevent the use of these models in making forecasts. Although the resulting forecasts are not optimum (in the mean square error sense), the method is objective and can be tested (see fig. 9). The first problem (instability of eq (10)) can be avoided by ignoring the nonlinearity and recursively estimating the a_n's using the value $\alpha = 0.0$. If this is done, the spectrum of the a_n's is not white noise, but has a large broad peak near the region of the second harmonic. Forecasts are then made with α restored to 0.03. This approximation may be the cause of the bias noted below.

The ambiguity in the sign of z_n can be avoided by considering only that part of the historical record where clear cyclic behavior is apparent (i.e., all the data since 1700). Only the Maunder minimum (~1650 to 1700) cannot be used. An initial negative sign was used for z_n beginning in 1700 and subsequent signs were selected to make a reasonably smooth curve [24] with a period of 22 years.

To avoid the third problem, we simply forecast future z_n's, compute their confidence limits and square them.

Since data were available through 1977 (corresponding to one year after minimum, roughly), past cycles were computed using data up to, and including, one year past minimum for cycle 20 and for each of the past eleven sunspot cycles. Figure 9 shows the results of forecasting maximum annual mean sunspot numbers for these cycles. The maximum annual mean sunspot number forecast for Cycle 21 (using fitted model) is 159 ± 40. The confidence intervals given are 50 percent intervals and are based on forecast errors for the fitted ARMA models [6]. Due to the difficulties already elaborated on, these confidence intervals are probably optimistic. Confidence intervals for the empirical model (with nonlinearity) forecasts are probably similar. The RMS error for the fitted and empirical models are 34 and 27, respectively. The empirical model shows a significant bias of minus 23 (the standard error of the bias is 8).

In summary, to use ARMA models for forecasting, we estimated the original noise signal structure that produced the observed sunspot record. The narrowband resonant filter was then driven with this derived "original" noise signal until the beginning of the forecast period. The filter input was then set to zero, the most likely value, and the filter was allowed to "ring." The filter output, from that point on, was a damped harmonic response. For this reason, the confidence intervals became large rather rapidly and the forecasts are damped exponentially to zero. The model should not be used to forecast more than one cycle ahead.

6. Conclusions

The model proposed is extremely simple, yet is based on simple resonant phenomena widely occurring in nature. The model accurately simulates <u>all</u> of the gross features of the observed sunspot record including:

(a) the approximately eleven-year period,

(b) the variability of period,

(c) the short-term fluctuations,

(d) the rapid rise and slow decay,

(e) the observed distribution of values and

(f) the general appearance of sunspot cycles.

In addition, the model provides forecasts for future sunspot cycle peaks and has given reasonable predictions when applied to old data. We acknowledge that the available sunspot data constitutes a very small sample upon which to base conclusions as to long-term solar behavior. Nevertheless, it is felt that the principal elements of this simple statistical model (i.e., the Gaussian-noise source and the narrowband resonant filter) do suggest possible physical processes in very general terms and beg for more specific physical explanation.

Dicke [25] has recently pointed out that the phase stability of the observed sunspot record suggests a periodic driving force or chronometer. He argues that the phase fluctuations are not the type of random walk that would be produced by an "eruption" mechanism. Bloomfield [26], using complex demodulation to estimate the envelope and phase processes, shows even more clearly that the phase is very stable during the period 1700 to the present. It is more difficult, however, to establish the existence of an aperiodic driving force against the alternative of a narrowband process. Middleton [11] derives the 4th order joint probability functions of the phase and envelope processes for arbitrary time delay. It is shown that, although the envelope and phase <u>variables</u> are independent at any given time, the envelope and phase <u>processes</u> are not independent. The distribution of the phase change occurring between times t and t + τ depends not only on the characteristics (primarily bandwidth) of the filter, but also locally on the envelope process. It is well known to communications engineers that the phase of a narrowband process is rather stable when the amplitude is large and unstable when it is small.

Our model thus implies that, while phase is a random walk in the sense that there is no "corrective" or "zeroing" force, the steps taken are likely to be small during periods of high amplitude, such as the period since 1700, thus giving the appearance of phase stability. Phase instability will occur during periods of low activity, such as the Eddy minima. Long lengths of data are needed to distinguish a process having a driving force from a simple narrowband process. An engineering rule-of-thumb is that many times the reciprocal of the bandwidth are required. In the case of the solar cycle, our model has a reciprocal bandwidth of 500 years (not accidentally the mean interval between Eddy minima), so many thousand years of data are required. The Epstein-Yapp Deuterium/ Hydrogen record (1,000 years) discussed by Dicke is not long enough to resolve this question. Perhaps some future solar sensitive, geochronological record will provide some answers.

We also point out that this is not necessarily an either/or situation. Suppose the "input" signal for our filter included a truly periodic but weak driving force, imbedded in noise. Such a model would still explain features such as Gleissburg cycles and Eddy minima, yet have greater long-term phase stability. Brier [27] describes a similar situation in his analysis of the Quasi-Biennial Oscillation. A statistical test for phase stability exceeding that explained by the bandwidth of a Gaussian process, is a challenging, unsolved problem.

The question of phase stability is interesting, but forecasting is of greater practical importance. The levels of solar activity we will face in the decades ahead are of increasing concern. If one is inclined to accept the model described in this paper as a valid simulator of sunspot activity, one must conclude that we cannot now forecast, and never will forecast with reasonable confidence, more than about one eleven-year cycle in advance. This, of course, has serious implications (especially for NASA); but, we think, Mankind may do well to know its limitations [28].

7. References

[1] Yule, G. U., Phil. Trans. Royal Soc. (Ser. A), 226, 267 (1927).

[2] Some examples are: Anderson, T. W., The Statistical Analysis of Time Series, 224, (Wiley & Sons, New York, 1971); Koopmans, L. H., The Spectral Analysis of Time Series, 210, (Academic Press, New York, 1974); and Bloomfield, P., Fourier Analysis of Time Series: An Introduction, 94, (Wiley & Sons, New York, 1976).

[3] For example, Smith, J. L. Spencer, J. Royal Statist. Soc., 107, 231, (1944); Moran, P. A. P., J. Royal Statist. Soc. (Ser. B), 16, 112 (1954) and Morris, M. J., J. Royal Statist. Soc. (Ser. 4), 140, 437, (1977).

[4] We used a listing of data from 1650 to present supplied by J. A. Eddy. This is essentially the listing that was published as part of Eddy, J. A., Science, 192, 1189, (1976).

[5] Gough, D., The Solar Output and its Variation, White, O. R., ed., 492, (Colorado Assoc. Univ. Press, Boulder, Co., 1977).

[6] Bloomfield, P., Fourier Analysis of Time Series: An Introduction, 101, (Wiley & Sons, New York, 1976).

[7] Box, G. E. P. and Jenkins, G. M., Time Series Analysis: Forecasting and Control (Holden-Day, San Francisco, 1970).

[8] Although first mention of the 80 to 90-year quasi-periodic changes in the heights of maxima of the eleven-year cycles is attributed to Wolf, Gleissburg's name is usually associated with these long cycles. Gleissburg, W., Naturwis., 42, 410, (1955); _______, J. Brit. Astron. Assoc., 75, 227, (1965); ______, J. Brit. Astron. Assoc., 76, 265, (1966) and ______, Solar Phys., 21, 240, (1971).

[9] Jack Eddy discussed the Maunder and Sporer minima and evidence for other similar periods of minimal solar activity in a previous article: Eddy op. cit., p. 1189. We believe that he should be credited with calling attention to this phenomenon and, hence, have named such periods in our simulated data Eddy minima.

[10] Our model can be implemented by a rather short program in BASIC. This is shown here for those who may wish to experiment with it. This program contains all the refinements discussed in the paper.

```
100 X=RND(-2):A=1.90693:B=-.98751
110 C=.78512:D=-.40662:E=.4:F=.03:G=0
130 FOR N=1 TO 300
140 IF G=1 THEN GOTO 180
150 X=RND(X):Y=SQR(-2*LOG(X))
160 X=RND(X):K=Y*E*COS(6.28318*X)
170 G=1:GOTO 190
180 G=0:K=Y*E*SIN(6.28318*X)
190 H=A*I+B*J+K-C*L-D*M
200 M=L:L=K:S=I*I-J*J
210 T=H*H+F*S*S
220 PRINT T
230 J=I:I=H
240 NEXT N
250 STOP
```

Note: The program assumes RND(-2) generates a "seed" number for a sequence of pseudo-random numbers rectangularly distributed between 0 and 1. Other negative arguments for RND will generate other sequences.

[11] Middleton, D., An Introduction to Statistical Communication Theory 397, Sec. 9.1 (McGraw-Hill, New York, 1960) Davenport, W. B. and Root, W. L., 158, An Introduction to the Theory of Random Signals and Noise (McGraw-Hill, New York, 1958).

[12] Vitinskii, Yu. I, Solar-Activity Forecasting, 12, (Israel Prog. for Sci. Translations, Jerusalem, 1965).

[13] Brillinger, D. R. and Rosenblatt, M., Spectral Analysis of Time Series Harris, B.,153, ed.(Wiley & Sons, New York), 1967.

[14] Bloomfield, op. cit., p. 98.

[15] Eddy, op. cit., p. 1189.

[16] One of the few references we know of is Cook, A. F., J. Geophys. Res., 54, 347 (1949). However, several authors have commented on an apparent double maximum in the solar cycle, among them Gnevyshev, M. N., Solar Phys., 1, 107 (1967) and Papagiannis, M. D., Zerefos, C. S. and Reparis, C. C., Solar Activity and Related Interplanetary and Terrestrial Phenomena, Xanthakis, J., ed.,75, (Springer-Verlag, Berlin, 1973).

[17] Stobie, R. S., Multiple Periodic Variable Stars, Fitch, W. S., ed., 93, (Reidel, D., Dordrecht, Holland, 1976). *See also*, Campbell, L. *and* Jacchia, L., The Story of Variable Stars, 63, (Bladiston Co., Philadelphia, 1946) and Rosseland, S., The Pulsation Theory of Variable Stars, 3, (Clarendon Press, Oxford, 1949). Pursuing this point further, Lockwood, G. W. and Thompson, D. T., Nature, 280, 43, (1979) have discussed measuring the Sun as a variable star by measuring the light reflected by the outer planets, and have found a convincing relation between planetary albedos of Neptune and Titan and the sunspot cycle.

[18] For example, Shapley, A. H., Terr. Mag. and Atmos. Elec., 49, 43, (1944)

[19] Gleissburg, W., Naturwis., 42, 410, (1955).

[20] Rubashev, B. M., Problems of Solar Activity, 34, (NASA, Washington, 1964),.

[21] Siscoe, G. L., J. Geophys. Res., 81, 6224, (1976).

[22] Zook, H. A., Hartung, J. B. and Storzer, D., Icarus, 32, 106, (1977).

[23] Eddy, op. cit., p. 1189.

[24] According to the method of Bracewell, R. N., Nature, 171, 649, (1953).

[25] Dicke, R. H., Nature, 276, 676, (1978).

[26] Bloomfield, P., Fourier Analysis of Time Series: An Introduction, 137, (Wiley & Sons, New York, 1976).

[27] Brier, G. W., Monthly Weather Review, 106, 938, (1978).

[28] The authors have discussed the concepts presented here with colleagues too numerous to cite individually. These discussions have helped develop and refine the ideas described here and some will recognize their contributions. The authors bear full responsibility for such errors as have crept into the work. Special mention is deserved by Jones, R. H. for his thoughtful review.

NBS TECHNICAL PUBLICATIONS

PERIODICALS

JOURNAL OF RESEARCH—The Journal of Research of the National Bureau of Standards reports NBS research and development in those disciplines of the physical and engineering sciences in which the Bureau is active. These include physics, chemistry, engineering, mathematics, and computer sciences. Papers cover a broad range of subjects, with major emphasis on measurement methodology and the basic technology underlying standardization. Also included from time to time are survey articles on topics closely related to the Bureau's technical and scientific programs. As a special service to subscribers each issue contains complete citations to all recent Bureau publications in both NBS and non-NBS media. Issued six times a year. Annual subscription: domestic $13; foreign $16.25. Single copy, $3 domestic; $3.75 foreign.

NOTE: The Journal was formerly published in two sections: Section A "Physics and Chemistry" and Section B "Mathematical Sciences."

DIMENSIONS/NBS—This monthly magazine is published to inform scientists, engineers, business and industry leaders, teachers, students, and consumers of the latest advances in science and technology, with primary emphasis on work at NBS. The magazine highlights and reviews such issues as energy research, fire protection, building technology, metric conversion, pollution abatement, health and safety, and consumer product performance. In addition, it reports the results of Bureau programs in measurement standards and techniques, properties of matter and materials, engineering standards and services, instrumentation, and automatic data processing. Annual subscription: domestic $11; foreign $13.75.

NONPERIODICALS

Monographs—Major contributions to the technical literature on various subjects related to the Bureau's scientific and technical activities.

Handbooks—Recommended codes of engineering and industrial practice (including safety codes) developed in cooperation with interested industries, professional organizations, and regulatory bodies.

Special Publications—Include proceedings of conferences sponsored by NBS, NBS annual reports, and other special publications appropriate to this grouping such as wall charts, pocket cards, and bibliographies.

Applied Mathematics Series—Mathematical tables, manuals, and studies of special interest to physicists, engineers, chemists, biologists, mathematicians, computer programmers, and others engaged in scientific and technical work.

National Standard Reference Data Series—Provides quantitative data on the physical and chemical properties of materials, compiled from the world's literature and critically evaluated. Developed under a worldwide program coordinated by NBS under the authority of the National Standard Data Act (Public Law 90-396).

NOTE: The principal publication outlet for the foregoing data is the Journal of Physical and Chemical Reference Data (JPCRD) published quarterly for NBS by the American Chemical Society (ACS) and the American Institute of Physics (AIP). Subscriptions, reprints, and supplements available from ACS, 1155 Sixteenth St., NW, Washington, DC 20056.

Building Science Series—*Disseminates technical information* developed at the Bureau on building materials, components, systems, and whole structures. The series presents research results, test methods, and performance criteria related to the structural and environmental functions and the durability and safety characteristics of building elements and systems.

Technical Notes—Studies or reports which are complete in themselves but restrictive in their treatment of a subject. Analogous to monographs but not so comprehensive in scope or definitive in treatment of the subject area. Often serve as a vehicle for final reports of work performed at NBS under the sponsorship of other government agencies.

Voluntary Product Standards—Developed under procedures published by the Department of Commerce in Part 10, Title 15, of the Code of Federal Regulations. The standards establish nationally recognized requirements for products, and provide all concerned interests with a basis for common understanding of the characteristics of the products. NBS administers this program as a supplement to the activities of the private sector standardizing organizations.

Consumer Information Series—Practical information, based on NBS research and experience, covering areas of interest to the consumer. Easily understandable language and illustrations provide useful background knowledge for shopping in today's technological marketplace.

Order the **above** *NBS publications from: Superintendent of Documents, Government Printing Office, Washington, DC 20402.*

Order the **following** *NBS publications—FIPS and NBSIR's—from the National Technical Information Services, Springfield, VA 22161.*

Federal Information Processing Standards Publications (FIPS PUB)—Publications in this series collectively constitute the Federal Information Processing Standards Register. The Register serves as the official source of information in the Federal Government regarding standards issued by NBS pursuant to the Federal Property and Administrative Services Act of 1949 as amended, Public Law 89-306 (79 Stat. 1127), and as implemented by Executive Order 11717 (38 FR 12315, dated May 11, 1973) and Part 6 of Title 15 CFR (Code of Federal Regulations).

NBS Interagency Reports (NBSIR)—A special series of interim or final reports on work performed by NBS for outside sponsors (both government and non-government). In general, initial distribution is handled by the sponsor; public distribution is by the National Technical Information Services, Springfield, VA 22161, in paper copy or microfiche form.

BIBLIOGRAPHIC SUBSCRIPTION SERVICES

The following current-awareness and literature-survey bibliographies are issued periodically by the Bureau:

Cryogenic Data Center Current Awareness Service. A literature survey issued biweekly. Annual subscription: domestic $35; foreign $45.

Liquefied Natural Gas. A literature survey issued quarterly. Annual subscription: $30.

Superconducting Devices and Materials. A literature survey issued quarterly. Annual subscription: $45. Please send subscription orders and remittances for the preceding bibliographic services to the National Bureau of Standards, Cryogenic Data Center (736) Boulder, CO 80303.