Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence

BENOIT B. MANDELBROT AND JAMES R. WALLIS

International Business Machines Research Center
Yorktown Heights, New York 10598

Abstract. The rescaled range $R(t, s)/S(t, s)$ is shown by extensive computer simulation to be a very robust statistic for testing the presence of noncyclic long run statistical dependence in records and, in cases where such dependence is present, for estimating its intensity. The processes examined in this paper extend to extraordinarily non-Gaussian processes with huge skewness and/or kurtosis (that is, third and/or fourth moments).

The present paper is addressed to scientists concerned with analyzing empiric records in which very long run statistical dependence (excluding seasonals or other cycles) may be present. Problems relative to the very long run are increasingly recognized as being on the forefront of both theoretic and practical statistics, but statistical techniques to treat the very long run were yesterday all but nonexistent. We have found a new technique of data analysis, to be called 'R/S analysis,' very effective in this context [*Mandelbrot and Wallis*, 1968; 1969*a*, *b*, *c*, *d*], and we shall try in this paper to examine the principal reasons for our enthusiasm. References to actual records will be limited to a comment (below) on Hurst's empiric results concerning R/S.

The letters 'R/S' stand for the rescaled range $R(t, s)/S(t, s)$, where $R(t, s)$ is the cumulated range of a process between times $t + 1$ and $t + s$ after removal of the sample average and $S^2(t, s)$ is the corresponding sample variance. That is, given any function $X(t)$ in discrete integer valued time with $X^*(t)$ defined by $X^*(t) = \Sigma_{u=1}^{t} X(u)$ and given any lag $s > 1$, $R(t, s)$ is defined by

$$\begin{aligned}R(t, s) = \max_{0\leq u\leq s} \{&X^*(t + u) - X^*(t) \\ &- (u/s)[X^*(t + s) - X^*(t)]\} \\ - \min_{0\leq u\leq s} \{&X^*(t + u) - X^*(t) \\ &- (u/s)[X^*(t + s) - X^*(t)]\}\end{aligned}$$

(Figure 1), and $S^2(t, s)$ is defined by

$$\begin{aligned}S^2(t, s) &= s^{-1} \sum_{u=1}^{s} \{X(t + u) - s^{-1} [X^*(t + s) - X^*(t)]\}^2 \\ &= s^{-1} \sum_{u=1}^{s} X^2(t + u) - [s^{-1} \sum_{u=1}^{s} X(t + u)]^2.\end{aligned}$$

We shall show here that the dependence of $\mathcal{E}[R(t, s)/S(t, s)]$ upon s can introduce the concept of R/S intensity, which is one precise and useful measure for the more general concept of intensity of noncyclic very long run statistical dependence. Consequently the dependence upon s of the average of the sample values of $R(t, s)/S(t, s)$, carried over all admissible starting points t within the sample, can be used to test whether the R/S intensity in a sample is nonvanishing and/or to estimate the value of this R/S intensity. The performance of these statistical tasks will be labeled R/S analysis, a term we have coined after spectral analysis. The measure of R/S intensity will be designated by $H - 0.5$, with $0 < H < 1$, the special value $H = 0.5$ corresponding to the absence of very long run dependence.

Without question the first discipline in which the presence of noncyclic very long run dependence has been reported is hydrology. We have therefore proposed that all fields exhibiting noncyclic very long run dependence be said to

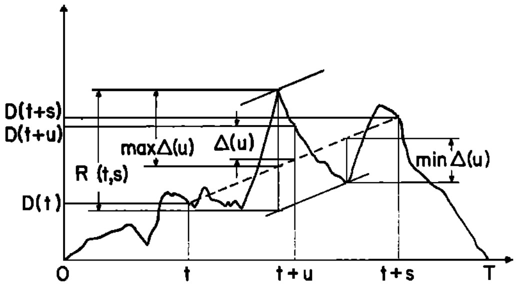

Fig. 1. Construction of the sample range $R(t, s)$ reproduced from *Mandelbrot and Wallis* [1969b]. To make this graph more legible the function $X^*(t)$ was measured arbitrarily from its sample average over the sample from $t = 1$ to $t = T$. That is, instead of $X^*(t)$ itself, we have plotted as a bold line the function $D(t) = X^*(t) - (t/T)X^*(T)$. The replacement of $X^*(t)$ by $D(t)$ does not affect the value either of $\Delta(u)$ as defined below or of $R(t, s)$. Moreover since empiric records are necessarily taken in discrete time, the function $D(u)$ should have been drawn as a series of points, but it was drawn as a line for the sake of clarity. The function $\Delta(u)$ as marked stands for $X^*(t + u) - X^*(t) - (u/s)[X^*(t + s) - X^*(t)]$, and the sample range is defined as $R(t, s) = \max_{0 \le u \le s} \Delta(u) - \min_{0 \le u \le s} \Delta(u)$.

exhibit the Joseph Effect. That is, the original Joseph Effect expresses the well-established fact that high or low levels in rivers tend to persist, perhaps over the Biblical 'seven fat and seven lean years,' but more often over decades, centuries, and millenia. Similar observations have been made in meteorology, geophysics, economics, physics, and other sciences. Using our convenient terminology, we shall characterize R/S intensity as one possible measure of how strongly the Joseph Effect is present in a given class of phenomena.

The question immediately arises of whether other measures of the strength of the Joseph Effect could have been used. The answer depends upon whether the record in question is nearly Gaussian. If it is, one can also use the method of variance-time curves introduced by G. I. Taylor in 1921 or perhaps one of a few alternative statistical techniques. However, natural records are very frequently extremely non-Gaussian. This finding was also first reported in hydrology, so that we have proposed to call it the Noah Effect. The original Noah Effect expresses the fact that the levels of rivers may be extraordinarily high and that intense rain may last over the Biblical forty days and nights. The Noah Effect is classically studied under the heading of 'theory of extreme values.' Because of the Noah Effect the question we raised at the beginning of this paragraph is transformed into the question of how the R/S intensity and other measures of the Joseph Effect are affected by the superposition of the Joseph and the Noah Effect.

Our investigations have led us to conclude that the unique virtue of R/S analysis lies in its being 'blind' to the Noah Effect, that is, in its being equally applicable to Gaussian records and to records with a strong Noah. When there is no Noah, other available techniques may be comparable to R/S analysis in effectiveness. When Noah is strong, however, all alternative techniques known to us are in some way inferior to R/S analysis. The best are simply less effective, their sampling distribution being less favorable. The worst alternative techniques are worthless because the results they yield confuse Noah and Joseph inextricably.

Because of the audience anticipated for this paper, the theorems to be stated will not be proved mathematically but rather demonstrated by computer simulation buttressed by some heuristic argument. The captions of the figures are unusually detailed and should be considered an integral part of the exposition.

Our use of the term 'law' will follow the custom established by the law of large numbers, which designates a statement to the effect that some sample average tends asymptotically towards its expectation. As is well known, many classical theorems of probability are of the form 'under such and such hypotheses, the law of large numbers holds.' Similarly behind every development that follows lurks a theorem of the form 'under such and such hypotheses, such and such 's^H law' applies to $R(t, s)/S(t, s)$.' That is, the phrase 's^H law' will not designate a theorem but only the conclusion common to a number of theorems.

For stationary Gaussian processes without long run dependence, division by $S(t, s)$ is a useful but not vital detail, so that R/S analysis is a small improvement over the R analysis of R itself, as carried out for white Gaussian noise by *Feller* [1951] and by *Anis and Lloyd* [1953]. The importance of the division of $R(t, s)$ by $S(t, s)$ increases as the process diverges from the Gaussian and/or as one introduces dependence of increasingly longer extent.

It has been shown in *Mandelbrot* [1965], *Mandelbrot and Van Ness* [1968], and *Mandelbrot and Wallis* [1969*b*] that Gaussian random processes with a fractional spectrum in which dependence can have as long an extent as one may wish are very good for modeling the Joseph Effect. It remained, however, to study R/S for non-Gaussian processes and for processes with strong cyclic components and to compare R/S analysis with other methods of analyzing long run statistical dependence (including R analysis and other techniques, many of which were apparently first considered in our work as possible alternatives to R/S analysis). We have attacked all these tasks with the help of computer simulation, but only our main results concerning R/S will be reported.

MATHEMATICAL PRELIMINARY

If $X(t)$ is a stationary random process, the ratio $R(t, s)/S(t, s)$ considered for fixed s as a function of t is another stationary random process, a transform of the original $X(t)$. Recall that a random process $X(t)$ is called stationary if identical rules generate the process $X(t)$ itself and all the processes deduced from $X(t)$ by a time shift, namely, all the processes of the form $X(t + s)$.

To appreciate some of our manipulations concerning R/S, it is useful to recall the corresponding features of the classical covariance analysis. One can consider that technique as based on the fact that when $X(t)$ is stationary, the transformed process $Y(t) = X(t)\ X(t + s)$ is also stationary for every s. Since the covariance of $X(t)$ may be written as $\mathcal{E}[X(t)X(t + s)] = \mathcal{E}[Y(t)]$, such covariance depends on s but is not a function of t and can be designated by $C(s)$. It is known that many features of a process are fully described by the functional dependence of $C(s)$ upon s.

R/S analysis also is based on the properties of a family of stationary random functions obtained by transforming $X(t)$, namely, the function $R(t, s)/S(t, s)$. Stationarity implies that $\mathcal{E}[R(t, s)/S(t, s)]$, like $C(s)$ above, depends on s but not on t, and we shall show that some important properties of a process are described by the functional dependence of this $\mathcal{E}[R(t, s)/S(t, s)]$ upon s.

The body of this paper will discuss first R/S testing, then R/S estimation. A short additional section on cyclic effects will follow. The last sections will comment briefly on R/S self similarity and on R/S analysis for nonstationary processes.

R/S TESTING

Abstract of this section. The behavior of $R(t, s)/S(t, s)$ as $s \to \infty$ can define the concept of R/S dependence, which is a form of noncyclic long run statistical dependence. The first application of R/S analysis thus occurs in testing whether R/S dependence is present in a record.

Preliminary. The concept of long run statistical dependence is obviously important, but it is complex and many faceted, and generally accepted definitions are lacking. Some random processes exist, however, for which long run dependence is unquestionably absent, and some other processes exist in which long run dependence is unquestionably present. Moreover, cyclic and noncyclic long run dependence must be distinguished [*Mandelbrot,* 1969, section 2.2]. Having R/S-analyzed many processes, we have observed a relation between noncyclic long run dependence and the following law.

Definition. A random process will be said to satisfy the $s^{0.5}$ law in the mean or, to be more precise, to satisfy the $R/S \sim s^{0.5}$ law in the mean if the expression

$$\lim_{s\to\infty} s^{-0.5}\,\mathcal{E}[R(t, s)/S(t, s)]$$

exists (that is, is well defined) and is positive and finite. In more intuitive terms this $s^{0.5}$ law means that the graph of the function log $\mathcal{E}[R(t, s)/S(t, s)]$ versus log s is asymptotically a straight line of slope 0.5.

The $s^{0.5}$ law in the mean fails to hold in two cases, (1) when

$$s^{-0.5}\mathcal{E}[R(t, s)/S(t, s)]$$

tends to no limit as $s \to \infty$ and (2) when this quantity tends to zero or to infinity. All these possibilities are expressed in more intuitive terms by saying that the graph of the function log $\mathcal{E}[R(t, s)/S(t, s)]$ versus log s does not possess a straight asymptote of slope 0.5.

Basic result. We have found that the $s^{0.5}$ law in the mean holds for every process for which long term dependence is unquestionably absent and does not hold for many processes exhibiting unquestionable noncyclic long term statistical dependence.

Examples. The stationary process of independent reduced Gaussian variables is unquestionably the simplest process with no dependence. The term 'reduced' means that the expectation vanishes and the variance is unity. For this process the law of large numbers shows that $\lim_{s\to\infty} S(t, s) = 1$. In addition *Feller* [1951] has computed the value of $\lim_{s\to\infty} s^{-0.5}$ $\mathcal{E}[R(t, s)]$ and showed it to be $\sqrt{\pi/2}$, which is about 1.25. Thus $C = \lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)]$ is also about 1.25 for independent Gaussian processes. The fact that this limit is both positive and finite establishes that these processes satisfy the $s^{0.5}$ law in the mean (Figure 2).

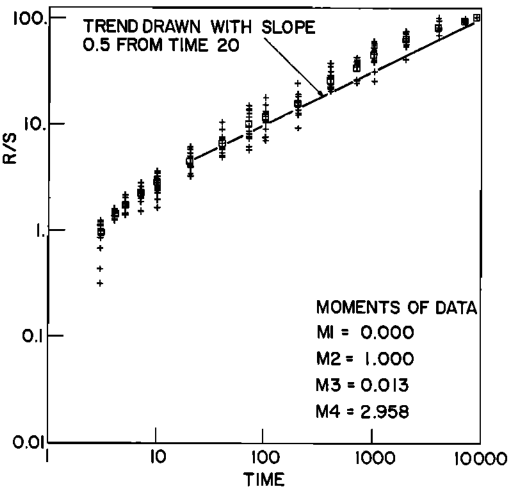

Fig. 2. Pox diagram of log R/S versus log s for a sample of 9000 values of a discrete white noise $G(t)$, i.e., a sample of independent identically distributed Gaussian random variables. The boxes correspond to estimates of log $\mathcal{E}[R/S]$. Their disposition is evidence that the $s^{0.5}$ law in the mean applies to $G(t)$ after a very short transient, shorter than the transient of $R(t, s)$ derived by *Anis and Lloyd* [1953]. The dots (+) correspond to sample values of R/S (see below). They show that the dispersion of R/S around $s^{-0.5}$, that is, the dispersion of $s^{-0.5}R/S$, depends little on s. This disposition is evidence that the $s^{0.5}$ law in distribution applies to $G(t)$ after the same short transient, the relative dispersion of $s^{-0.5}$ being small.

A sample function of white noise has been plotted in *Mandelbrot and Wallis* [1969*a*]. Had we plotted log R (t, s) instead of log R/S, this diagram would not look substantially different, though the initial transient would have become longer.

This pox diagram was constructed as follows. The dots (+) correspond to values of the lag s restricted to the sequence 3, 4, 5, 7, 10, 20, 40, 70, 100, 200, 400, 700, 1000, 2000, 4000, 7000, and 9000. For every s satisfying $s < 500$, 14 dots (+) are plotted, corresponding to values of t equal to 1, 100, ... , 1400. For every s satisfying $s > 500$, t was made successively equal to 1000, 2000, up to either 8000 or $T - s + 1$, whichever is the smaller.

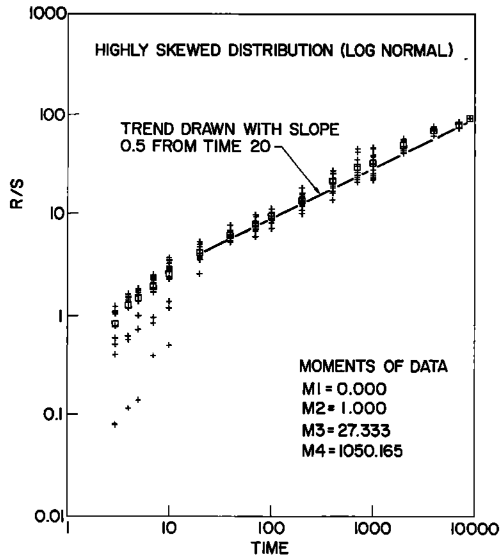

Fig. 3. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for an independent log normal random function. Such functions have been used often to fit hydrologic data. The function we used, namely, $10^{G(t)}$ (where $G(t)$ is a discrete white noise), has very high skewness (27) and kurtosis (1050), which are symptoms of the Noah Effect. Nevertheless, the disposition of the boxes indicates that the $s^{0.5}$ law in the mean is satisfied by $10^{G(t)}$ after a short transient, whereas the disposition of the sample values (+) indicates that the $s^{0.5}$ law in distribution is satisfied with a small relative dispersion.

It might have been illuminating to plot a sample function of $10^{G(t)}$, but a plot on linear coordinates would be illegible. It could show clearly either the peaks or the details of the small values but could not show both. Logarithmic coordinates would have to be used in order to make the plot legible, but this would have simply consisted in plotting $G(t)$ itself.

The fact that in the case of $10^{G(t)}$ the $s^{0.5}$ law holds asymptotically is a consequence of the argument in *Feller* [1951]. The study of the penultimate region of moderate values of s is an entirely different matter. To prove that the asymptotic results should apply there, one can proceed as follows. The basic fact, proved in *Mandelbrot* [in preparation], is that in the range of moderately large values of s the log normal density can be approximated by an appropriate hyperbolic, as defined in the legend of Figure 5. As a result the penultimate behavior of R/S for the log normal process is essentially like the asymptotic behavior of R/S for the appropriate hyperbolic process.

Figures 2, 3, 5, and 6 demonstrate that the $s^{0.5}$ law also applies in the mean to processes of independent values having a variety of other marginal distributions: truncated Gaussian, log normal, and hyperbolic, respectively. In every case the expression $s^{-0.5}\ \mathcal{E}[R(t, s)/S(t, s)]$ attains its limit value very rapidly, that is, after a very brief initial transient. Note, however, that the precise value of $\lim_{s\to\infty} s^{-0.5}\ \mathcal{E}[R(t, s)/S(t, s)]$ greatly depends on the process; this will be exemplified later in the paper.

It should be noted that the value of $R(t, s)/S(t, s)$ for small lag s as plotted in the figures of this paper have been incorrectly computed. They must be disregarded. The paper's conclusions, however, remain unaffected. The nature of the error will be explained in a paper by M. Taqqu, Note on Evaluations of R/S for

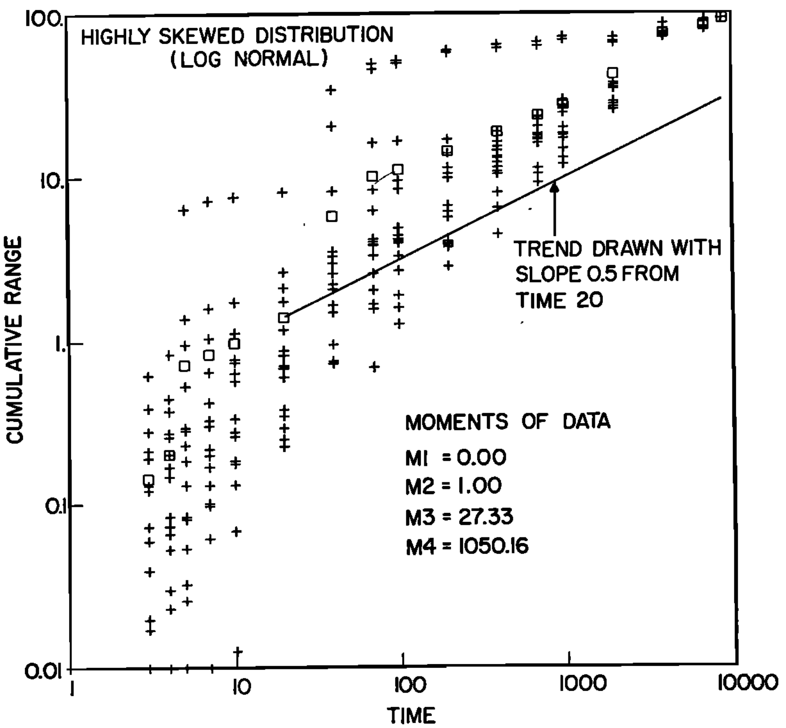

Fig. 4. Pox diagram of log $R(t, s)$ versus log s for the sample of the random function $10^{G(t)}$ used for Figure 3. The present diagram differs dramatically from the diagram of log R/S shown on Figure 3. The $s^{0.5}$ law in the mean (boxes) shows evidence of applying, but the transient goes up to $s \sim 70$, which means it is longer than the transient of log R/S for either $G(t)$ or $10^{G(t)}$. There is no visible stabilization, however, in the distribution of the sample values of $R(t, s)$ (+) around their average. Thus even a sample of 9000 values gives no evidence that the $s^{0.5}$ law in distribution is valid. The scatter of the dots (+) is so extreme that were the sample much shorter, testing whether the $R(t, s)$ function of $10^{G(t)}$ obeys the $s^{0.5}$ law would be hard at best and often hopeless. That is, the statistic $R(t, s)$ is much less robust than R/S.

The pox diagram of log $S(t, s)$ versus log s has also been plotted but need not be reproduced in this paper because it is so similar to the present diagram. In other words, the diagrams of both log $S(t, s)$ and $R(t, s)$ are widely scattered, but the values of the two functions are so precisely meshed together that, as seen on Figure 3, little scatter remains in the difference log $R(t, s)$ − log $S(t, s)$ = log R/S. Additional insights about meshing are yielded by the behavior of R/S for the very skew Gamma process discussed in the body of this paper. Still further insights are yielded by the legend of Figure 11.

The meshing of $R(t, s)$ and $S(t, s)$ teaches an important lesson: an excellently behaved statistic can sometimes be obtained by combining several expressions that behave badly when considered separately.

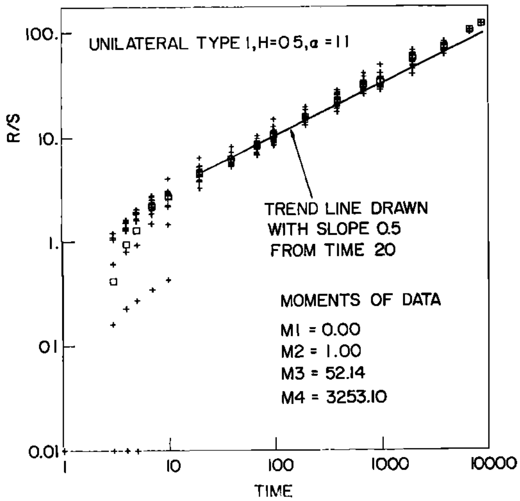

Fig. 5. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for a sequence of independent identically distributed hyperbolic random variables with exponent $\alpha = 1.1$, where hyperbolic random variables are defined below. We shall not try to plot a sample function of this process on linear coordinates because, as was the case for the function $10^{G(t)}$, such plot would be illegible. The disposition of the boxes is evidence that the $s^{0.5}$ law in the mean applies after a short transient, and the disposition of the sample values (+) is evidence that the $s^{0.5}$ law in distribution applies with a relative dispersion, perhaps even smaller than in the Gaussian case.

The random variable X is called 'hyperbolic' if it satisfies for large values of x the two relations

$$Pr(X > x) \sim (x/\sigma')^{-\alpha}$$

$$Pr(X < -x) \sim (x/\sigma'')^{-\alpha}$$

where α is a positive constant. If, moreover, either σ' or σ'' vanishes, X is called 'unilaterally hyperbolic' or 'Paretian.' If both σ' and σ'' are positive, X is called 'bilaterally hyperbolic.' (The possibility that $\sigma' = \sigma'' = 0$ must be excluded.) For a discussion of the special role of such random variables, see for example *Mandelbrot* [1963*a*, 1964*a*].

To carry out the simulations reported in the figures of this paper, we considered the case where $\sigma'' = 0$ and $\sigma' = 1$ and the case where $\sigma' = \sigma''$, $2^{-1/\alpha}$. We began with a sequence $F(t)$ of independent random variables distributed between 0 and 1 with uniform probability density. Next a bilateral hyperbolic function $Z(t)$ was constructed using the formulas

$$\text{If } 0 < F(t) < \tfrac{1}{2}, \qquad Z(t) = [2F(t)]^{-1/\alpha}$$

$$\text{If } \tfrac{1}{2} < F(t) < 1, \qquad Z(t) = [2 - 2F(t)]^{-1/\alpha}$$

Finally, to simplify subsequent calculations, $Z(t)$ was rounded off by dropping away all the digits after the decimal point. As to our unilateral hyperbolic function, it was defined as $|Z(t)|$.

Fractional Noises and Geophysical Records, submitted for publication to *Water Resources Research.*

When the values of the process $X(t)$ are statistically dependent but the dependence is limited to the short run, the transient is much longer, but the $s^{0.5}$ law in the mean holds asymptotically. We shall return later to a discussion of the practical meaning of such asymptotic results.

On the other hand Figures 7 to 17 and additionally many figures in *Mandelbrot and Wallis* [1969*b*] show the $s^{0.5}$ law in the mean to fail for a variety of processes for which the dependence between $X(t)$ and $X(t + T)$ decreases to zero as $T \to \infty$ but does so extremely slowly. A more detailed discussion of these figures is best postponed to a later section devoted to R/S estimation.

Remark concerning the explanation of Hurst's empiric s^H law. In empiric records the values

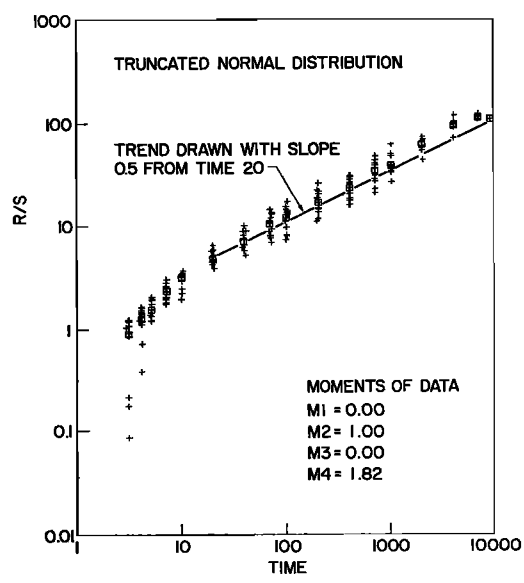

Fig. 6. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for a sequence of independent random variables with low kurtosis. These variables were obtained by truncating a normal random variable so severely as to make its distribution almost uniform. The validity of the $s^{0.5}$ law in the mean is no more affected by the low kurtosis of the process used in this figure than it was affected by the high kurtosis of the processes used in Figures 3 and 5. The relative dispersion of $s^{-0.5}$ R/S is perhaps larger than in the Gaussian case.

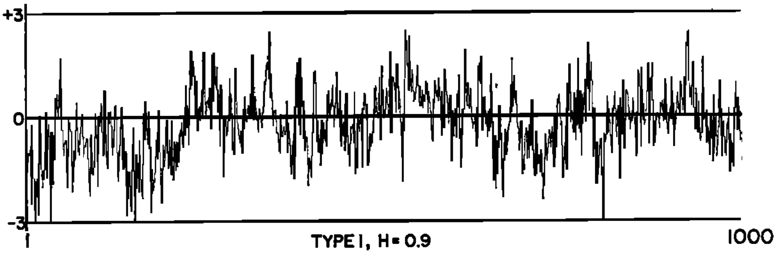

Fig. 7. The first 1000 values from a 9000 value sample of a Type 1 approximation to the fractional noise with $H = 0.9$ and $M = 10{,}000$. The whole sample of 9000 has been normalized to have zero mean and unit variance. This figure is reproduced from *Mandelbrot and Wallis* [1969*b*].

of R/S were found to cluster closely along a function of the form s^H with $H > 0.5$. The finding that $H > 0.5$ was originally made by *Hurst* [1951], although Hurst's estimates of H involve a far-reaching conceptual error discussed in *Mandelbrot and Wallis* [1969*d*]. After *Feller* [1951] proved that the empiric s^H law is incompatible with the idea that the records in question are generated by an independent Gaussian process, several authors, including *Moran* [1964; 1968], have conjectured that the empiric s^H law could be explained by postulating that the records are generated by a random process with independent values and a very skew marginal distribution. These authors thus postulated that the empiric s^H law relates to the Noah Effect. The results in the present paper show that this Noah explanation is insufficient. Earlier [*Mandelbrot* 1965; *Mandelbrot and Van Ness,* 1968; *Mandelbrot and Wallis,* 1969*b*] the Noah explanation was shown to be also unnecessary, since it was shown to be possible to explain the s^H law by the Joseph Effect.

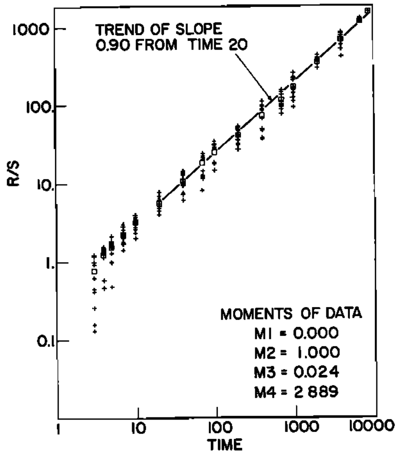

Fig. 8. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for 9000 values of a Type 1 approximation to the fractional noise with $H = 0.9$ and $M = 10{,}000$ (including and continuing the sample of Figure 7). This figure is reproduced from *Mandelbrot and Wallis* [1969*b*]. The R/S intensity of statistical dependence is clearly equal to $H = 0.9$.

More specifically, *Moran* [1968, p. 495] has attempted to illustrate his proposed explanation of Hurst's empiric law by considering Gamma distributed random variables of density $[\Gamma(\gamma)]^{-1} x^{\gamma-1} e^{-x}$, where the parameter γ is assumed very small. Moran's illustration is fallacious, as we shall now proceed to demonstrate.

The key fact is that a very skew Gamma process $X(t)$ exhibits a Noah Effect so extreme in its intensity that unless t is made extremely large, of the order of $1/\gamma$, there is a very high probability for $X^*(t) = \Sigma_{u=1}^{t} X(u)$ to be almost indistinguishable from $\max_{0\le u\le t} X(u)$. As a result $R(t, s)$ is nearly equal to $\max_{0\le u\le s} X(t+u)$. In addition, $\Sigma_{u=1}^{s} X^2(t+u)$ is nearly equal to $[\max_{0\le u\le s} X(t+u)]^2$ and $S^2(t, s)$ is nearly equal to

$$s^{-1}[\max_{0\le u\le s} X(t+u)]^2 - [s^{-1} \max_{0\le u< s} X(t+u)]^2$$
$$= s^{-1}(1 - s^{-1})[\max_{0\le u\le s} X(t+u)]^2$$

Finally we find $R(t, s)/S(t, s)$ to be nearly equal to $s^{0.5}(1 - s^{-1})^{-0.5}$, independently of t. After an initial transient until s^{-1} becomes $\ll 1$, say up to $s = 10$, one has $R(t, s)/S(t, s) = s^{0.5}$ with negligibly small statistical scatter. This argument ceases to apply when s exceeds $1/\gamma$, but it suffices to show that Moran's claims were unfounded.

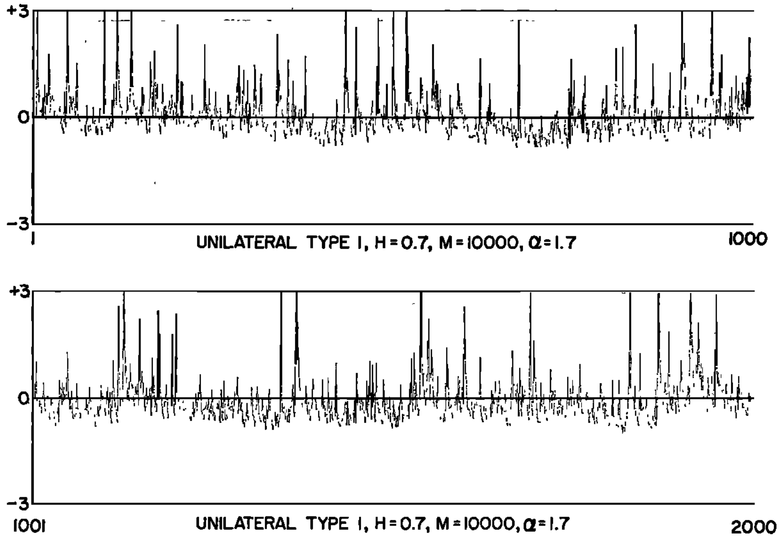

Fig. 9. The first 1000 values from a 9000 value sample of a hyperbolic non-Gaussian fractional noise. To construct this sample, we preserved the same moving average kernel already used to construct a Type 1 fractional Gaussian noise [*Mandelbrot and Wallis*, 1969*c*]. However, the variables to be averaged were hyperbolic, a concept defined in the legend of Figure 5. The largest values of this sample of hyperbolic fractional noise are very large indeed, a symptom of a very strong Noah Effect. In fact such extrema exceeded the bounds of the graph and had to be sharply truncated. This means that at the instants where $X(t)$ apparently equals the maximum plottable value of $X(t)$ the actual value is much larger. Each large value has strong and long-lived aftereffects.

Incidentally, we do not question Moran's mathematics. His error lies elsewhere, in believing that Hurst's empiric findings applied to the range itself and not to the ratio R/S. The behavior of $R(t, s)$ will be examined below in a subsection devoted to nonrobust variants of R/S.

Formal definition of R/S independence. The examples we considered introduce a distinction between two kinds of random process: those for which the limit $\lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)]$ is nontrivial (positive and finite) and those for which the limit in question is either trivial or nonexistent. This alternative has purposefully been stated in such terms that every random process falls on one or the other side, and this alternative can therefore be used as the basis of the following formal definition of dependence: Every process such that $\lim s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)]$ is finite and positive will be said to be R/S independent. All other processes will be said to be R/S dependent.

Definition of R/S testing and the relativity of the concepts of short run and long run. The above definition suggests that having computed the values of $R(t, s)/S(t, s)$ corresponding to some available finite sample of $X(t)$'s one could try to determine the category to which the process that generated $X(t)$ is likely to belong from the behavior of the sample values of R/S. However, this proposed statistical technique immediately raises a major conceptual difficulty: the concept of R/S dependence was defined by the asymptotic behavior of $R(t, s)/S(t, s)$. It remains to interpret R/S for finite samples of ordinarily available size.

Given a sample of size T so that the values of $R(t, s)/S(t, s)$ are known up to s equal to a finite T, the ideal case occurs when the variations of the sample average of $s^{-0.5}\, R(t, s)/S(t, s)$ die out for s much less than T. Two conclusions can then be drawn: (1) the value near which this sample average stabilizes can be taken as an estimate of the asymptotic limit $\lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)]$, and (2) one can say not only that there is no long run R/S dependence in $X(t)$ but also that the R/S dependence of $X(t)$ has a span much shorter than T.

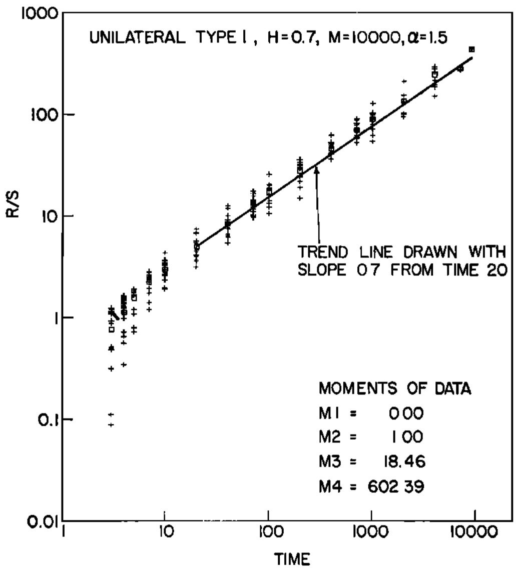

Fig. 10. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for 9000 values of the hyperbolic non-Gaussian fractional noise, part of which was plotted on Figure 9. The $s^{0.7}$ law clearly holds both in the mean and in distribution, indicating that R/S analysis is blind to the extremely non-Gaussian character of the marginal distribution (strong Noah Effect) even when very long run dependence (strong Joseph Effect) has been built into the process in question.

The alternative case is when there is a tendency for the observed averages of $s^{-0.5}$ $R(t, s)/S(t, s)$ to continue to vary greatly while s approaches its upper bound T. This has two possible causes: either (1) $X(t)$ is R/S independent, but the transient zone of $R(t, s)/S(t, s)$ is longer than T, or (2) $X(t)$ is R/S dependent. From a sample of finite duration T, one cannot conceivably distinguish between the above two possibilities.

In summary, given a sample of duration T, R/S testing consists in deciding which is more likely among the following possibilities: (1) the span of R/S dependence is much less than T, or (2) the span of R/S dependence is either of the order of magnitude of T or greater or even infinite.

Relation between R/S dependence and other forms of long run dependence. The idea of forming the ratio R/S first arose in hydrology, $R(t, s)$ being related to Rippl's ideal minimum capacity of reservoirs for long term storage [*Hurst*, 1951]. The distinction between R/S dependence and R/S independence is therefore likely in one field at least to be practically useful. The examples we studied show, moreover, that the concept of R/S independence catches some of the aspects of the intuitive idea of long run statistical independence, but it does not catch all aspects. We may add that it is unlikely that any single definition of the long run independence will ever catch all aspects of this concept and that alternative definitions for long run independence will always exist. For instance, random processes may be R/S dependent but long run independent according to other criteria [*Mandelbrot*, 1969], but this is not the proper place to discuss this feature. Also some processes are R/S dependent but independent according to other criteria. The principal example of this last possibility is provided by processes with both a sinusoidal cyclic component and a noise component. The presence of the sine expresses that there is a very long run statistical dependence, but it will be shown that such processes are R/S independent.

Effect of strong cyclic components on R/S analysis. The best known kind of long run dependence is not R/S dependence but is exemplified by the pure sine wave $A \sin (2\pi t/L + \phi)$. The wavelength L is to be thought of as prescribed, and the amplitude A and phase ϕ are both chosen at random in advance according to any specified probability distribution. For this process the covariance between $X(t)$ and $X(t + s)$ is itself a sine function that oscillates up and down without limit. Now consider the ratio R/S of the pure sine. Clearly, $\lim_{s\to\infty} R(t, s) = AL/\pi$ and $\lim_{s\to\infty} S(t, s) = A/2$, so that $\lim_{s\to\infty} [R(t, s)/S(t, s)] = 2\, L/\pi = .636\, L$ (see Figure 18). Division by $s^{0.5}$ yields $\lim_{s\to\infty} s^{-0.5}$ $\mathcal{E}[R(t, s)/S(t, s)] = 0$, so that pure sine waves are R/S dependent. In other words, R/S dependence is not in conflict with pure sine dependence.

When $G(t)$, a white Gaussian noise of zero mean and unit variance, is added to the sine wave to obtain

$$X(t) = A \sin (2\pi\, t/L+\phi) + G(t),$$

the situation changes radically. One can check that $X^*(t)$ satisfies the double inequality

$$G^*(t) - AL/2\pi \le X^*(t) \le G^*(t) + AL/2\pi$$

Since for $t \to \infty$, $AL/2\pi$ becomes negligible in relative value, the ranges of the two processes

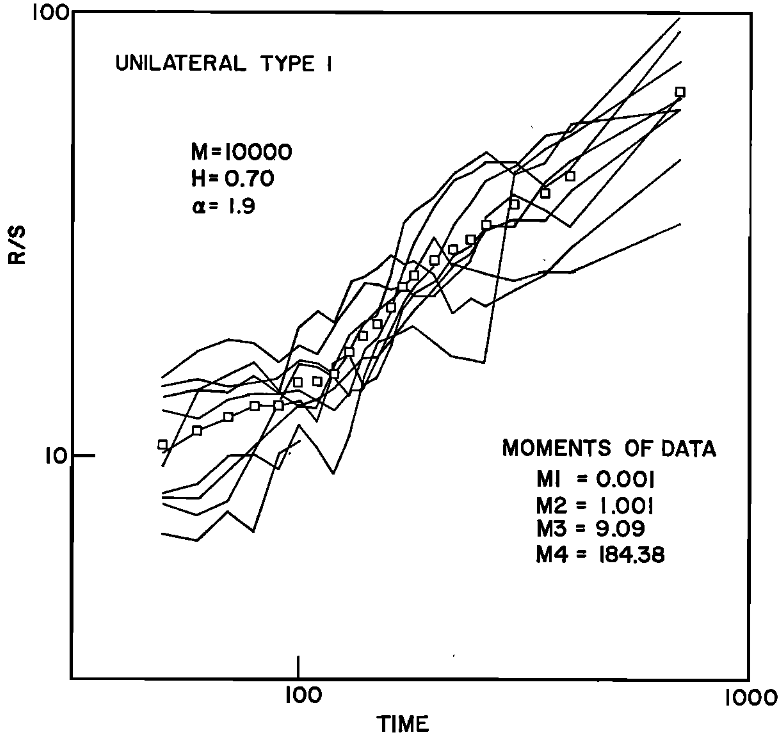

Fig. 11. Alternative plot of a greatly enlarged detail of the variation of $R(t, s)/S(t, s)$ for a hyperbolic process different from the process used in Figure 10. This figure is not a pox diagram. Several starting points t having been selected, the sample path of $R(t, s)/S(t, s)$ as function of s was plotted. It is visible that this sample path rarely stays at one side of the pox diagram. It rather tends to flip up and down. This makes R/S analysis more reliable than it would have been if sample paths did not constantly cross the trend line.

$X(t)$ and $G(t)$ are asymptotically identical, and $\lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)] = 1.25$. On the other hand it is readily seen that $\mathcal{E}S^2(t, s) = 1 + A/2$. Consequently for the process $X(t)$ one has $\lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)] = 1.25\,(1 + A/2)^{-0.5}$. That is, in the case of a sine wave plus a white noise of arbitrary amplitude, the $s^{0.5}$ law in the mean is valid and there is no R/S dependence.

The values of $\lim_{s\to\infty} s^{-0.5}\, \mathcal{E}[R(t, s)/S(t, s)]$ as well as the speed with which this limit is attained are, however, highly dependent on A. Two things happen as A increases: the limit of $s^{-0.5}\, \mathcal{E}(R/S)$ tends to zero, and $s^{-0.5}\, \mathcal{E}(R/S)$ takes an ever longer time to reach its limit. For example in the case where $L = 100$ we found that the point where the asymptotic $s^{0.5}$ behavior prevails is beyond 9000 when $A = 3$ but is about 200 when $A = 0.5$.

Sharp cyclic components rarely occur in natural records. One is more likely to find mixtures of waves that have slightly differing wavelengths but greatly differing high subharmonics. As a result a number of cycles covering a whole band of frequencies will perturb R/S analysis less than would a single sharp sine of comparable total energy.

Statistical robustness of the mean variance $s^{0.5}$ law. The relative deviation of R/S is defined as $\sqrt{\mathrm{Var}}\, [R(t, s)/S(t, s)]/\mathcal{E}[R(t, s)/S(t, s)]$. For the stationary process of independent Gaussian variables this relative deviation tends as $s \to \infty$ towards $\lim_{s\to\infty} \sqrt{\mathrm{Var}}\, [R(t, s)]/\mathcal{E}[R(t, s)]$. *Feller* [1951] showed that quantity to be $\sqrt{\pi/3 - 1}$, which is about 0.217 and which we consider small. For other processes we studied, independently of whether R/S dependence is strong or absent, we again found the relative deviation of R/S to be small. More precisely, the relative deviation is smaller for R/S than for any alternative expression we thought might be used to study long run dependence.

It will be convenient in the sequel to use the

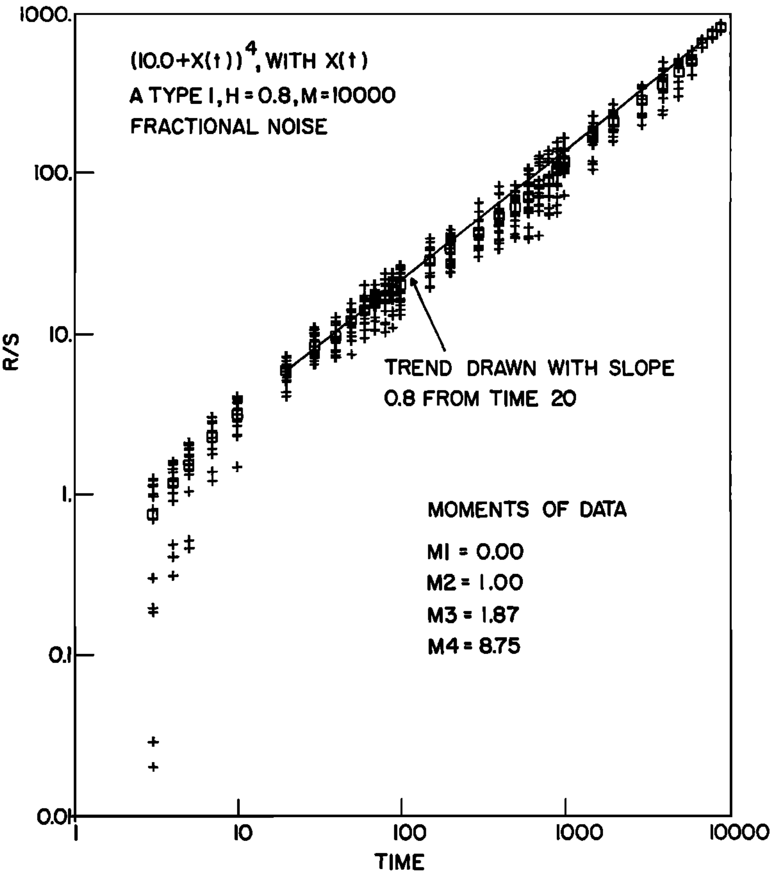

Fig. 12. Effect on the value of the R/S intensity of a transformation with fairly strong nonlinearity applied on an approximately fractional Gaussian noise. The initial process exhibited a strong Joseph Effect but no Noah Effect. The transformed process exhibits a moderate Noah Effect, but the R/S intensity of dependence is unaffected by the transformation. The practical importance of such nonlinear transformations is exemplified by the cases of tree rings and river levels. The thickness of a tree's rings is an increasing but probably nonlinear function of the total yearly precipitation in the site of the tree. The yearly maximum and minimum of a river's levels are presumably increasing but nonlinear functions of the yearly discharge. We studied a fourth power merely to illustrate such nonlinearity. The result shown on this figure strongly suggests that the R/S intensity that is estimated from tree ring thickness (resp., from river levels) can reasonably be expected to apply also to yearly precipitation (resp., to yearly flows).

term 'mean variance $s^{0.5}$ law' as an abbreviation for the combination of the two statements that the limit $\lim_{s\to\infty} s^{-0.5}\mathcal{E}[R(t, s)/S(t, s)]$ is nontrivial and that the limit of the relative deviation $\sqrt{\text{Var}}\ [R(t, s)/S(t, s)/\mathcal{E}[R(t, s)/S(t, s)]$ is small.

The basis of the R/S tests for noncyclic long run independence can now be rephrased in terms of the statistical concept of robustness. The extent to which Feller's results hold if $X(t)$ is not independent Gaussian is the extent to which statistics based upon $R(t, s)/S(t, s)$ are robust. Before we tackle this issue, it may be good to remind the reader of the definitions of the classical terms, 'statistic' and 'robust.'

Definition of the term 'statistic.' Given either T values of a random process $X(t)$ or T recorded observations thought to have been generated by a random process, the term 'statistic' is an awkward but entrenched synonym of one-dimensional or multidimensional function of the T arguments $X(t)$. The best known one-dimensional statistics are the sample moment for

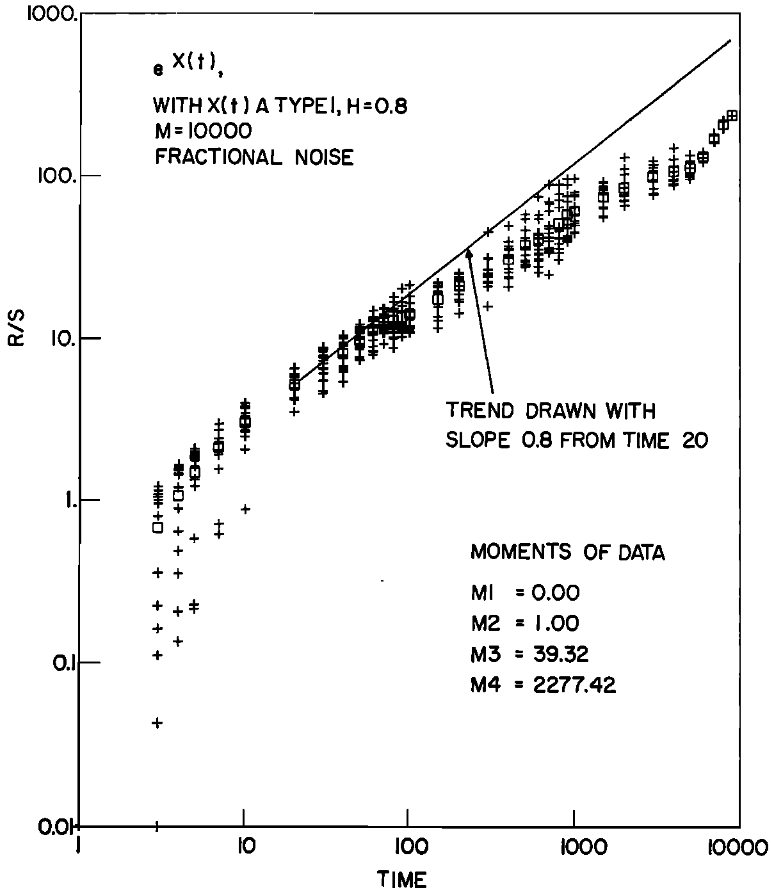

Fig. 13. Effect of an extremely nonlinear transformation on a fractional Gaussian noise. Note that the values of $X(t)$ are near the origin. This is why e^x is called more extremely nonlinear than $(10 + X)^4$. The apparent R/S intensity of dependence, which is the slope of the trend line of this diagram over the span of values of s that has been considered, is much smaller than H. This shows that extremely nonlinear transformation need not preserve the R/S intensity of an original process and in fact may lead to resulting processes that have no well defined R/S intensity.

given k, namely $T^{-1} \Sigma_{t=1}^{T} X^k(t)$, the sample covariance for given lag s, namely either $T^{-1} \Sigma_{t=1}^{T-1} X(t) X(t+s)$ or $(T-s)^{-1} \Sigma_{t=1}^{T-1} X(t) X(t+s)$, the sample lag correlation between $X(t)$ and $X(t+s)$ for given lag s, and the Fourier coefficient of $X(t)$ at given wave number k, for example

$$T^{-1} \sum_{t=1}^{T} X(t) \quad \sin \quad (2\pi kt/T)$$

Corresponding multidimensional statistics are the sets of all sample moments, correlations, or Fourier coefficients. Similarly the present work is concerned with statistics involving the rescaled range exemplified by

$$(T-s)^{-1} \sum_{t=1}^{T-s} R(t, s)/S(t, s)$$

Definition of the term robustness. A statistic is called robust if its distribution or the conclusions to which it leads are not drastically dependent upon specific assumptions about the process generating $X(t)$. The usual assumption against which robustness is assessed is that the process is Gaussian. But even then robustness is not a uniquely defined concept, since one can consider many different aspects for every statistic and since each of these aspects can be

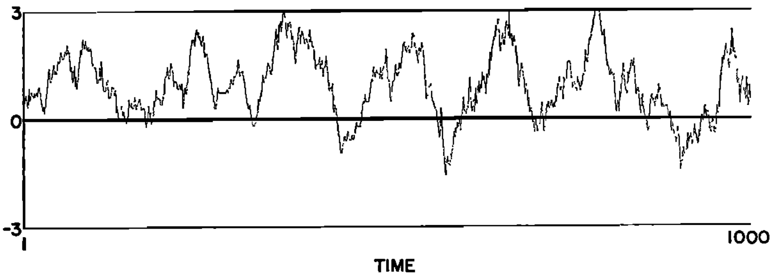

Fig. 14. The first 1000 values from a 9000 value sample of a locally Gaussian random process. Having been introduced in *Mandelbrot* [1969a] in order to resolve certain paradoxes encountered in attempts to model economic time series by strictly Gaussian processes, locally Gaussian processes are an especially tough challenge to data analysis.

The process plotted here is constructed as the expression $\Omega_N(t) = N^{-1/2} \Sigma_{m=1}{}^N W_m(t)$, where each $W_m(t)$ is a core process constructed by the following three steps. In a first step, one constructs a stationary renewal process, that is, a stationary sequence of points T_k such that the intervals $U_k = T_{k+1} - T_k$ are independent random variables all satisfying $Pr\ (U_k > u) = u^{-\beta}$. In a second step, one selects for $W_m(T_k)$ a sequence of independent Gaussian random variables of zero mean and unit variance. In a third step, one identifies the interval from T_k to T_{k+1} in which the instant t is located, and one makes $W_m(t)$ equal to $W_m\ (T_k)$. Thus each $W_m(t)$ is a step function representing a trend that changes at the instants T_k. Over any prescribed sample size from $t = 1$ to $t = T$, the random function $\Omega_M(t)$ tends to a fractional Gaussian noise as $N \to \infty$. When N is finite, however, $\Omega_m(t)$ is merely locally Gaussian. In this figure, $\beta = 1, 4$ and $N = 10$.

studied with respect to many different kinds of deviation from the independent Gaussian.

Nonrobustness of the precise value of the limit $lim_{s\to\infty}\ s^{-0.5}\ \mathcal{E}[R(t,\ s)/S(t,\ s)]$. *Feller* [1951] has proved, as we already noted, that for the process of independent Gaussian random variables, the value of this limit is equal to approximately 1.25. The same limit is also attained for every process with a finite variance. When the variance is infinite, however, the limit is different, typically between 1.25 and 1.

In addition the value of this limit can be arbitrarily modified by introducing short run statistical dependence, so that the property $\lim_{s\to\infty} s^{-0.5}\ \mathcal{E}[R(t,\ s)/S(t,\ s)] = 1.25$ is not robust with respect to short run deviations of $X(t)$ from the independent Gaussian process.

Consider for example the stuttering white Gaussian noise which is the stationary process such that its values for even instants of time $t = 2k$ are independent Gaussians $G(k)$, and its value at odd instants of time equals its value at the following even instant. When the value of s is large, the range $R_0(t, s)$ of the stuttering white Gaussian noise can be shown nearly to equal $\sqrt{2}\ R(t, s)$ with $R(t, s)$ the range of the independent Gaussian white noise $G(k)$, while the standard deviation $S_0(t, s)$ nearly equals the standard deviation $S(t, s)$ of $G(k)$. Thus

$$\lim_{s\to\infty} \mathcal{E}[s^{-0.5}R_0(t, s)/S_0(t, s)]$$

$$= \sqrt{2} \lim_{s\to\infty} \mathcal{E}[s^{-0.5}R(t, s)/S(t, s)] = 1.25\ \sqrt{2}$$

If the very short run dependence due to stuttering is made stronger, the limit of $\mathcal{E}[s^{-0.5}\,R(t,\,s)/S(t,\,s)]$ is further modified.

Extreme robustness of the mean variance $s^{0.5}$ law. As we have already said, for every process of independent values we examined, including extremely skew log normal processes (Figure 3) and processes with an infinite population variance (Figure 5), $\mathcal{E}[R/S]$ is asymptotically proportional to $s^{0.5}$, and the reduced variable $s^{-0.5}\ R/S$ has small variance. If anything, the variance is smaller in cases when $X(t)$ is a very long tailed random variable than in cases when $X(t)$ is Gaussian. We can now rephrase this result as saying that the mean variance $s^{0.5}$ law is extremely robust with respect to changes in the marginal distribution of $X(t)$.

Nonrobustness of the statistic $R(t, s)$. None of the many variants of R/S that we studied is as robust as $R(t, s)/S(t, s)$. While some alternatives to the R/S ratio retain the property that their expected value is asymptotically propor-

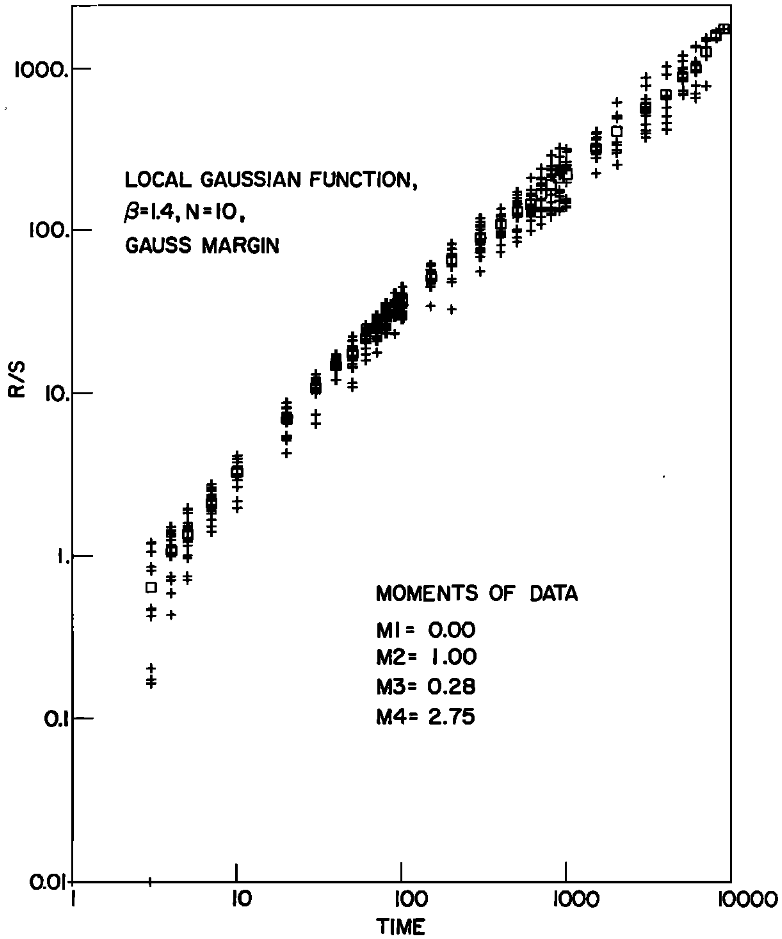

Fig. 15. Pox diagram of $R(t, s)/S(t, s)$ for 9000 values of a locally Gaussian noise, including and continuing the sample of Figure 14.

tional to $s^{0.5}$, none has as small a variance as R/S. In the present paper we shall be content to demonstrate the nonrobustness of $R(t, s)$ by examining two classes of non-Gaussian processes.

First class of examples. Consider the random process of independent hyperbolically distributed values for which R/S is studied on Figure 5. The marginal distribution is extremely skew and/or long tailed in this case, and it is possible to show [*Mandelbrot*, 1963*b*; *Moran*, 1964] that $\mathcal{E}R(t, s) \sim s^{1/\alpha}$ for this process, with α between 1 and 2 so that $1/\alpha$ is between 0.5 and 1. On the other hand the asymptotic population variance of $s^{-1/\alpha} R(t, s)$ is infinite for the process of Figure 5 which implies that sample values of $s^{-1/\alpha} R(t, s)$ are extremely erratically behaved, making it easy for sampling fluctuation to overwhelm and hide the functional dependence of $R(t, s)$ on s. As a consequence one may conjecture this: Had Hurst's rough graphic analysis been carried on $R(t, s)$ itself, given the highly non-Gaussian character of some of his records, Hurst might well have concluded that his records follow no simple law of general validity, and the topic might have been dropped. In other words, since on one hand sophisticated analysis is needed to make sure how $R(t, s)$ depends upon s and on the other hand sophisticated analysis is not ordinarily carried out unless there is evidence that it is worthwhile to do so, it is possible that ways to handle long run hydrologic effects would have been discovered

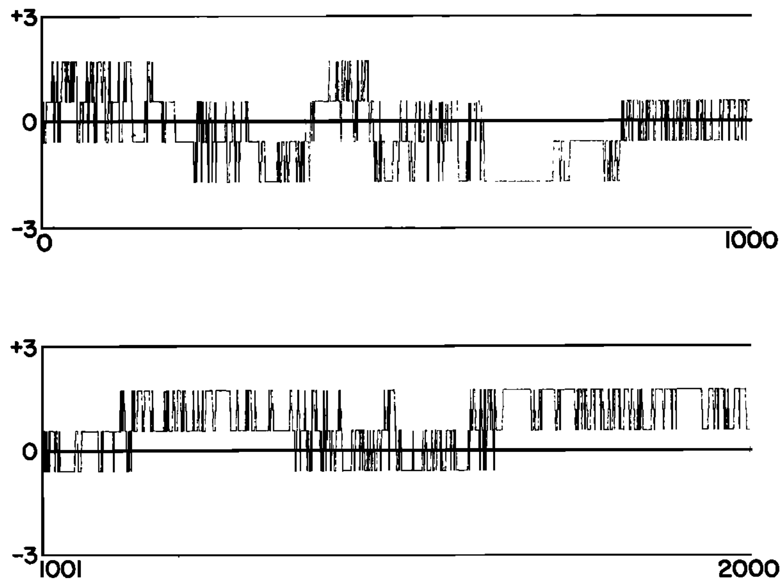

Fig. 16. The first 1000 values from a 9000 value sample of another locally Gaussian random process. The construction proceeded as for the function plotted on Figure 14, except that we selected for $W_m(T_k)$ a sequence of independent binomial random variables of zero mean and unit variance equal to $+1$ or -1 with probabilities 0.5. In this figure $N = 3$.

much later if Hurst had plotted R instead of R/S.

Second class of examples. Now consider the behavior of $R(t, s)$ for the process of independent log normal values (Figure 4). The corresponding behavior of R/S was reported on Figure 3. This example shows that random processes exist for which $\mathcal{E}R \sim s^{0.5}$ holds asymptotically, but the asymptotic behavior takes extraordinarily long to prevail. In the long transient that precedes this asymptote, the dispersion of R around $\mathcal{E}R$ may be enormous.

Similar remarks apply to Gamma distributed random processes, which (as we have noted already) were injected into this topic by Moran. For small values of s the range of such a process was found by Moran to satisfy $\mathcal{E}R \sim s$, but it is readily seen that the scatter of sample values around this expectation is enormous. Therefore the relation $\mathcal{E}R \sim s$ has no practical relevance.

Robustness of the $R/S \sim s^{0.5}$ law with respect to short run statistical dependence. Now consider random processes in which statistical dependence is present but intuitively felt to have a short range or, more accurately, to have a finite range. Examples are Markov random processes, finite autoregressive processes, and processes of finite moving averages of independent random variables. In such cases the value of $\lim_{s\to\infty} s^{-0.5} \mathcal{E}[R(t, s)/S(t, s)]$ is always positive and finite though much affected by details of the process. To eliminate this influence, one may consider the reduced random variable $[R(t, s)/S(t, s)]/\mathcal{E}[R(t, s/S(t, s)]$. It can be shown that the lim for $s \to \infty$ of that reduced variable is unaffected by the details of short run dependence.

Thus if it were legitimate to look at things from an asymptotic viewpoint, one could call the $R/S \sim s^{0.5}$ law robust with respect to the introduction of short run statistical dependence.

From a finite nonasymptotic viewpoint, however, things are always more complex, as we stressed earlier in this paper and in our preceding works.

R/S ESTIMATION

Abstract of this section. The behavior of $R(t, s)/S(t, s)$ as $s \to \infty$ can serve to define the concept of R/S intensity of dependence, which is a form of intensity of noncyclic long run statistical dependence. For that, one must divide the class of processes with long run

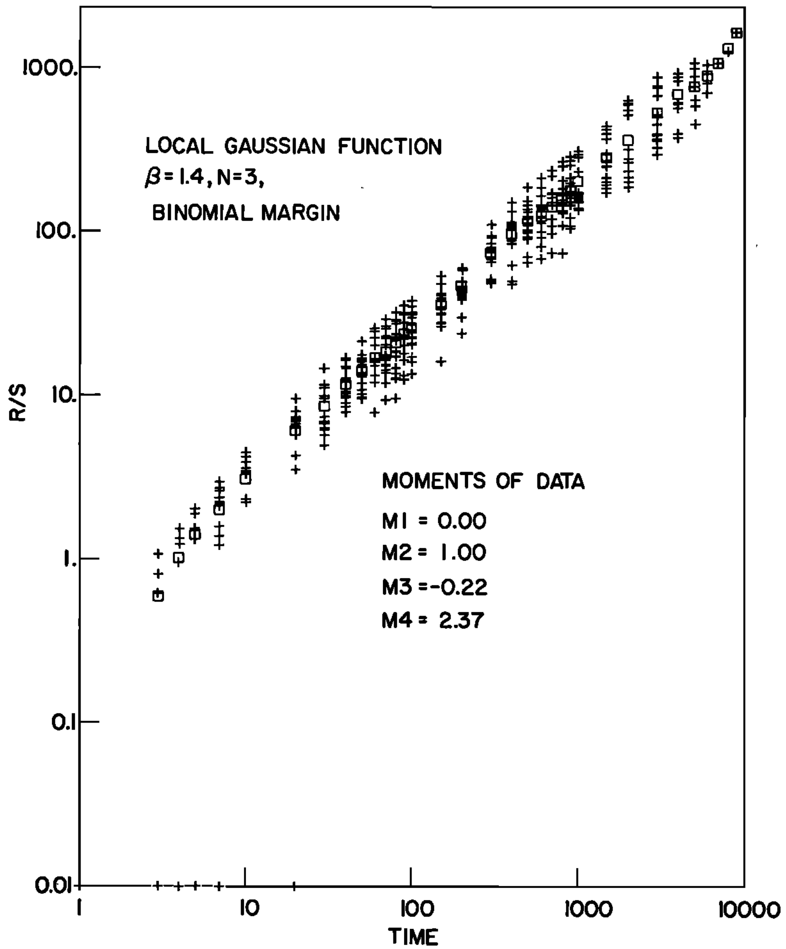

Fig. 17. Pox diagram of $R(t, s)/S(t, s)$ for 9000 values of a locally Gaussian noise, including and continuing the sample of Figure 16.

dependence more finely so that each subclass contains processes for which noncyclic long run dependence can be said to have the same intensity. With this finer subdivision we shall be able to proceed from the already discussed problem of testing for long run dependence to the problem of estimating the R/S intensity of a record precisely.

Definitions. We shall say that a random process satisfies the $R/S \sim s^H$ law in the mean if $\lim_{s\to\infty} s^{-H}\, \mathcal{E}[R(t, s)/S(t, s)]$ is defined and is positive and finite. We shall see that such processes exist for every H between 0 and 1. Following a common mathematical terminology, it is useful to say that all processes satisfying the $R/S \sim s^H$ law in the mean with identical H form a class of equivalence. The special class $H = 0.5$ corresponds to the absence of R/S dependence. If a process falls within the class $H \neq 0.5$, then $H - 0.5$ may be said to measure the R/S intensity of interdependence. Positive intensity expresses persistence. Negative intensity expresses a tendency of the values of $X(t)$ to compensate for each other to prevent $X^*(t)$ from blowing up too fast. Perfect compensation occurs in the pure sine wave, for which we saw that $H = 0$.

Remark. We could also have exhibited processes that do not satisfy in the mean any s^H law with $0 < H < 1$. Such processes when taken as a body constitute an additional class of equivalence, namely, a remainder class of proc-

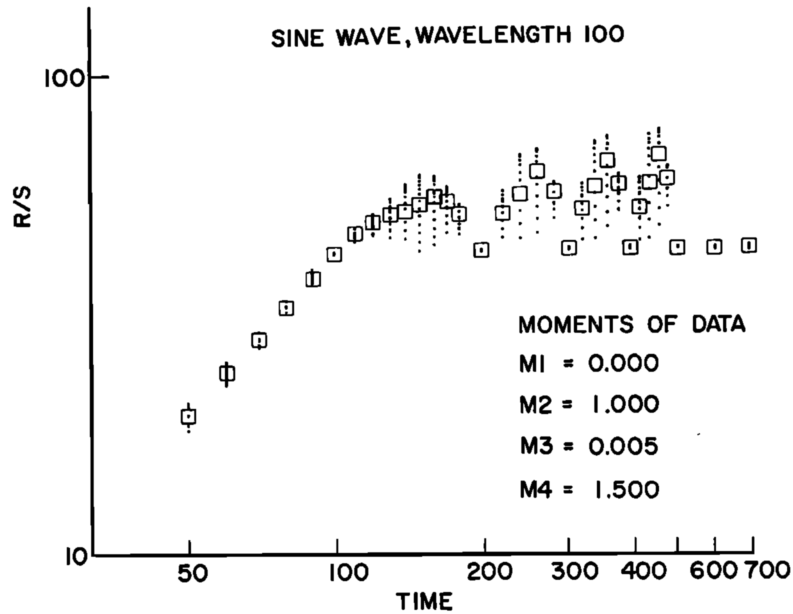

Fig. 18. Pox diagram of log $[R(t, s)/S(t, s)]$ versus log s for a pure sine wave with $L = 100$. First examine the values of time of the form $s = kL$ that correspond to subharmonics of the sine wave. There $R(t, s)/S(t, s)$ is independent of t and of k, as can be seen from the theory. Big lobes are, however, visible for other values of s. Had the values of $R(t, s)/S(t, s)$ been computed for s in a grid that eventually merges with the grid of the subharmonics of the pure sine, $R(t, s)/S(t, s)$ would rapidly attain its asymptotic limit. However, if the grid is selected independently of the value of L, it is more likely to fall within the lobes. This could yield a pox diagram of log $[R(t, s)/S(t, s)]$ versus log s having a positively sloped trend line. Thus a small sample of a pure sine wave could be R/S estimated to have a small but positive value of H. This conclusion would be incorrect.

This behavior of R/S is reflected in the remaining R/S pox diagrams and teaches important lessons. When cyclic effects are suspected but it is either undesirable or impossible to process data to eliminate cycles, one should compute $R(t, s)/S(t, s)$ or its average for as many values of s as one can manage. We have shown previously [*Mandelbrot and Wallis* [1969*b*] that, contrary to what happens in spectral analysis, it is quite unnecessary to smooth out the behavior of $R(t, s)/S(t, s)$ by averaging its values over neighboring values of s. We now see, in addition, that such smoothing would also mix the cyclic effects with noncyclic long run dependence to produce an apparent value of H devoid of significance.

esses to which no R/S intensity can be ascribed. But as of today, processes in this remainder class lack practical application.

Transformations with respect to which the R/S intensity of dependence is invariant. We must first return briefly to R/S testing, namely to the robustness of the $s^{0.5}$ law, because it will be useful to restate such robustness in an alternative fashion. It may be observed that every random process of independent non-Gaussian values $X(t)$ can be written as a nonlinear function of a process of independent Gaussian values $G(t)$. If for example $X(t)$ is log normal, one simply has $X(t) = c \exp [bG(t)]$ where c and b are arbitrary constants. The robustness of the mean variance $s^{-0.5}$ law discussed in the section on R/S testing can thus be rephrased by saying that this law is invariant with respect to nonlinear transformation of the white Gaussian noise. When discussing R/S testing, we also saw that the class of processes that exhibit no R/S dependence is left invariant by transformations that introduce short-term dependence.

We shall now address ourselves to the question of whether the equivalence classes with positive or negative R/S dependence are also left invariant by such transformations. It will be seen that fewer such transformations are admissible, so that the robustness under transformation is less when long run dependence is either positive or negative than when it is zero.

To describe the transformations we considered, let us take as a point of departure the fractional Gaussian noises of exponent H [*Mandelbrot and Van Ness,* 1968] and two approximations to fractional Gaussian noises [*Mandelbrot and Wallis,* 1969*b*]. Our Type 2 approximation is the grosser and less important of the two, but it is easier to define. It is given by the two parameter moving average

$$F_2(t|H, M) = (H - 0.5) \sum_{u=t-M}^{t-1} (t - u)^{H-1.5} \cdot G(u) + Q_H G(t)$$

$$= (H - 0.5) \sum_{u=1}^{M} u^{H-1.5} \cdot G(t - u) + Q_H G(t)$$

In this definition $G(u)$ is a sequence of independent Gauss random variables of zero mean and unit variance. The constant Q_H depends upon H, as follows:

$$Q_H = 0 \quad \text{if} \quad 0.5 < H < 1, \quad \text{and}$$

$$Q_H = (0.5 - H) \sum_{u=1}^{\infty} u^{H-1.5} \text{ if } 0 < H < 0.5$$

The final M parameter called the memory of the process is some very large quantity. Originally [*Mandelbrot,* 1965] M was set to $M = \infty$, but in *Mandelbrot and Wallis* [1969*b*] M was varied from 1 to 20,000.

The definitions of discrete fractional Gaussian noise itself, as well as of Type 1 approximation to it, are more cumbersome and they need not be repeated. It will suffice to recall that every variant considered in *Mandelbrot and Wallis* [1969*b*] is a linear function of independent Gaussian variables $G(u)$. Indeed, according to the definition in *Mandelbrot* [1965], fractional Gaussian noises are moving averages of the form $\int K(t - u)G(u)$, wherein the kernel $K(u)$ behaves for large values of u proportionately to $u^{H-1.5}$. The appearance of a typical fractional Gaussian noise is illustrated in Figure 7, and the corresponding R/S graph is plotted in Figure 8.

The words linear and Gaussian are crucial in answering the questions of whether, after various transformations have been applied to a fractional Gaussian noise of exponent H, the $R/S \sim s^H$ law continues to hold. We considered in detail two kinds of transformations:

(A) Replacement of the input variables $G(u)$ by extremely non-Gaussian variables, that is, nonlinear transformation of the input variables before they are combined linearly. We found that such transformations leave our classes of equivalence invariant (Figures 9, 10, and 11).

(B) Nonlinear transformation of intermediate variables obtained as linear forms of the input variables. We found that nonlinearity must be moderate if a class of equivalence is to stay invariant. For example, start from the function $F_2(t|H, \infty)$, whose R/S intensity of dependence is H. In the range of values of x between -8 and 8, the nonlinearity of the function $(10 + X)^4$ is sufficiently moderate for the R/S in-

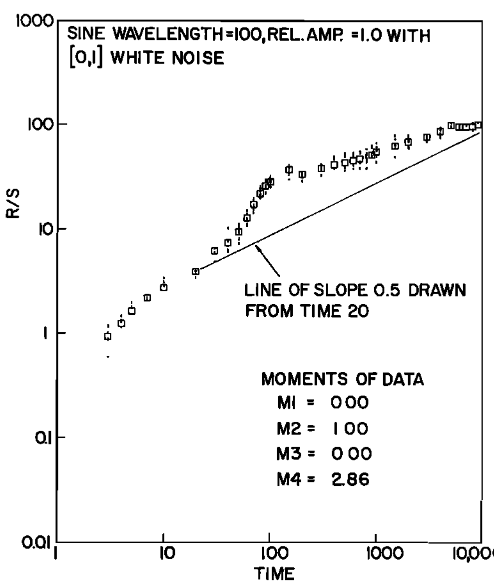

Fig. 19. The pox diagram of log $R(t, s)/S(t, s)$ versus log s for the sum of a pure sine wave and of a white noise of comparable amplitudes. The behavior of this function is a hybrid to which the behaviors of each of the functions plotted on Figures 2 and 18 contribute very clearly. Had the sine amplitude been stronger, the asymptotic of slope 0.5 characteristic of the noise component would have failed to prevail for the lags plotted on this figure. Had the sine amplitude been smaller, the wiggles and lobes characteristic of the pure sine component would have been less visible.

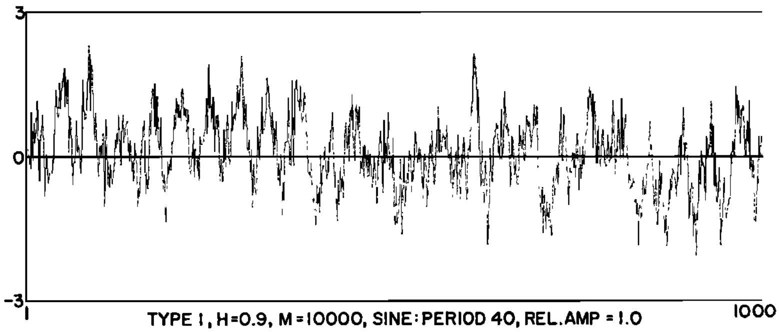

Fig. 20. The first 1000 values from a 9000 value sample of the sums of a fractional noise and a moderately strong sine wave.

tensity of $[10 + F_2(t|H, \infty)]^4$ to remain equal to H. But the function $\exp(X)$ is so nonlinear that the R/S intensity of $\exp[F_2(t|H, \infty)]$ is below H (Figures 12 and 13).

It will be interesting to combine the transformations (A) and (B) and to consider other transformations.

ADDITIONAL COMMENTS ON CYCLIC COMPONENTS

The effect of one cyclic component has already been studied under the assumption that $H = 0.5$ and that s is large. If more than one pure sine wave is added and H is made $\neq 0.5$, the asymptotic R/S intensity of dependence is unchanged, as might have been expected, but the nonasymptotic effects are not so obvious. The following unsystematic comments are meant as elaboration of what has been said earlier.

First examine in detail the graph of the $R(t, s)/S(t, s)$ function for the pure sine $A \sin(2\pi t/L + \phi)$ (Figure 18). The subharmonics of L, that is, the values of s multiple of L, stand out in two ways. First, when s is a subharmonic of L, the sample values of $R(t, s)/S(t, s)$ have no scatter, that is, are independent of t. Second, between those subharmonics, one finds lobes of decreasing amplitude with the scatter greatest halfway between subharmonics.

Next consider the function

$$X(t) = A \sin(2\pi t/L + \phi) + G(t)$$

where $G(t)$'s are independent Gaussian with zero mean and unit variance and L is large in comparison with the duration of the transient range before $R/S \sim s^{0.5}$ takes hold (Figure 19). When the lag s lies between the duration of the transient and T, the sine wave $A \sin(2\pi t/L)$ is practically a constant. Adding this constant to $G(t)$ leaves

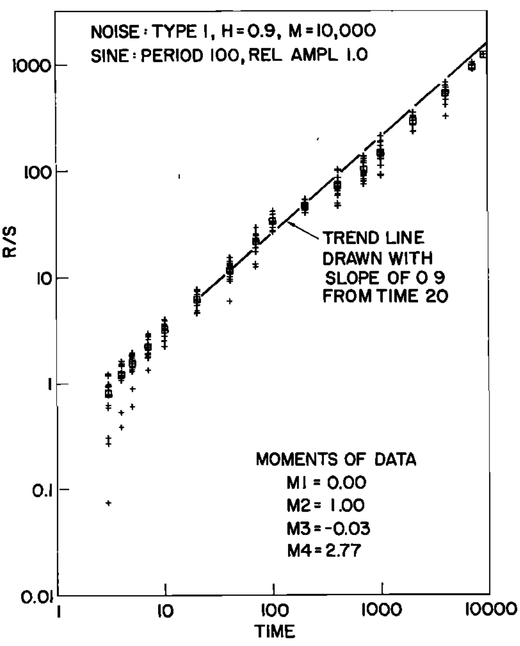

Fig. 21. Pox diagram of log $R(t, s)/S(t, s)$ versus log s for a sample of 9000 values of the sum of fractional noise and a sine wave. For high values of H, such as $H = 0.9$, the presence of a comparatively strong sine component leaves the s^H law in the mean valid. Thus it does not much affect R/S estimation. When the value of H is smaller, the effect of the cycle is more visible. Note also that the scatter of sample points around their trend line narrows near $s = 200$. This means that the convergence towards the s^H law in distribution is postponed to higher values of s when a sine wave is added. This tightening of the graph is even clearer on Figures 23 and 24 and will be discussed in the legend of Figure 24.

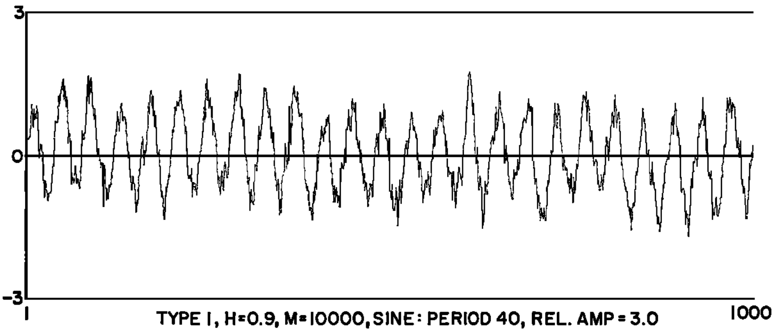

Fig. 22. The first 1000 values from a 9000 value sample of the sum of a fractional noise and a sine wave of very large relative amplitude, comparable to that of meteorologic records.

$R(t, s)$ and $S(t, s)$ practically unaffected and leaves $s^{-0.5}\,\mathcal{E}[R(t, s)/S(t, s)]$ near Feller's asymptotic value 1.25. Eventually $s^{-0.5}\,\mathcal{E}[R(t, s)/S(t, s)]$ attains its asymptotic value, derived earlier in the paper, of $1.25\,[1 + A/2]^{-0.5}$. But the transition from the initial value 1.25 to the final value $1.25\,[1 + A/2]^{-0.5}$ is not smooth and progressive; it proceeds in a series of wiggles that reflect the lobes of the function $R(t, s)/S(t, s)$ of a pure sine wave. For s near L and also (but less markedly) for s multiple of L the scatter of $R(t, s)/S(t, s)$ is greatly reduced.

Figures 20 to 24 refer to sums of a fractional Gaussian noise and various pure sines. The legends are self-explanatory.

MATHEMATICAL DIGRESSION CONCERNING ASYMPTOTIC SELF-SIMILARITY

In pursuing the study of R/S analysis it becomes important to study the distribution of the ratio $R(t, s)/S(t, s)$. In the present paper we have studied only its mean and variance. Simplest and most interesting are the processes such that as $s \to \infty$ the distribution of the expression $s^{-H}\,R/S$ tends towards a nontrivial limit, that is, tends towards the distribution of a random variable that does not reduce to either zero or infinity. Consider for example the independent Gaussian process. An argument due to *Feller* [1951] can be readily extended to show that in this case $s^{-0.5}\,R/S$ has a nontrivial limit. This process and all others for which $s^{-H}\,R/S$ has a nontrivial limit can be said to satisfy the $R/S \sim s^H$ law in distribution or to be asymptotically R/S self-similar. This last concept generalizes ordinary self-similarity, which is discussed in *Mandelbrot* [1967], *Mandelbrot and VanNess* [1968], and *Mandelbrot and Wallis* [1969*c*; 1969*d*].

MATHEMATICAL DIGRESSION CONCERNING THE SCOPE OF R/S ANALYSIS

As we noted when discussing the classical covariance analysis, $\mathcal{E}[X(t)X(t + s)]$ is independent of t if $X(t)$ is a stationary random process. But the converse is not true: the property that $\mathcal{E}[X(t)X(t + s)]$ is independent of t does

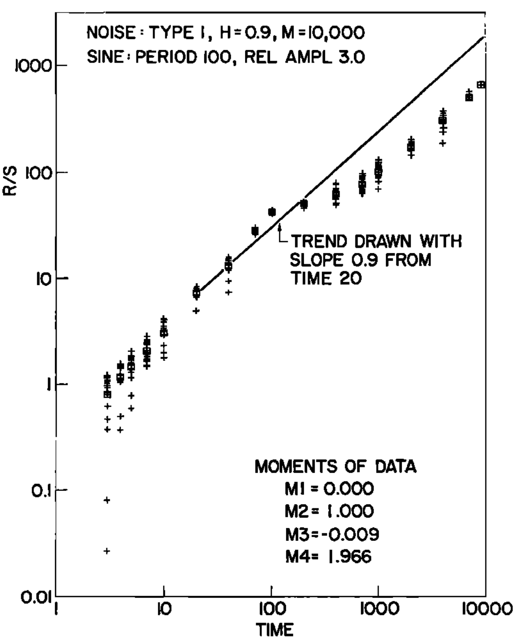

Fig. 23. Pox diagram of log $[R(t, s)/S(t, s)]$ for a sample of 9000 values that includes and continues the function of Figure 22. The effect of the sine wave is very strong. The critical bend starting at $s = 100$ is shown in detail on Figure 24.

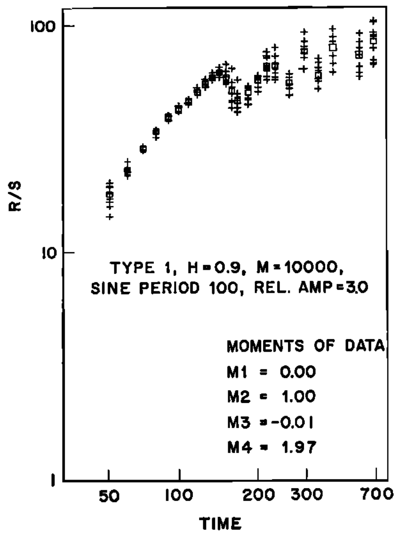

Fig. 24. Greatly enlarged detail of Figure 23. The single bend observed there is seen to divide into a richer structure of narrowings at the values of s corresponding to the subharmonics of the sine wave, with broad lobes between these narrow points. Both features reflect the properties noted on Figure 18 relative to the pure sine wave. As $s \to \infty$, the lobes die out and the contribution of the noise again becomes determinant. However, unless the total available sample size is much larger than the wave length of the pure sine, the apparent R/S intensity is greatly decreased by the addition of a strong sine wave.

not require that $X(t)$ be stationary. A nonstationary process such that $\mathcal{E}[X(t)X(t+s)]$ is independent of t will nevertheless appear stationary as long as examination is limited to the viewpoint of covariance analysis. Therefore it has been found useful to give a name to such processes: they have been called covariance-stationary (or weakly stationary or second order stationary) processes.

Similarly nonstationary random processes may exist for which $\mathcal{E}[R(t, s)/S(t, s)]$ is independent of t. When such processes are R/S analyzed but not studied from other viewpoints, they will appear stationary; it might therefore be useful to call them R/S stationary in the mean. Any such process may moreover satisfy the $R/S \sim s^H$ law in the mean, strict stationarity of $X(t)$ being unnecessary. If not only the expectation but the whole distribution of the random variable $R(t, s)/S(t, s)$ is independent of t, $X(t)$ would be called strict R/S stationary. Such processes may well satisfy the R/S s^H law in distribution, strict stationarity of $X(t)$ being again unnecessary. Thus R/S analysis may also apply to certain processes that are not stationary.

NOTATIONS

A,	(maximum) amplitude of a sine wave;
α,	The parameter of a hyperbolic distribution (also called Paretian);
β,	the basic parameter of a locally Gaussian process;
$C(s)$,	the covariance function of a stationary random function;
γ,	the parameter of a Gamma distribution;
$\Gamma(\gamma)$,	the Eulerian Gamma function of the parameter γ;
$\mathcal{E}$,	expectation, also called expected value, population mean or first moment;
$F_2(t \mid H, M)$,	Type 2 approximation of fractional Gaussian noise;
ϕ,	the phase of a sine wave;
$G(u)$,	sequence of independent Gaussian random variables;
H,	the principal parameter of a fractional noise. Also, the asymptotic slope of the plot of log $\mathcal{E}$ (R/S) versus log s if such plot is asymptotically straight;
k,	wave number in Fourier analysis;
$K(u)$,	a kernel function serving to compute a moving average;
L,	wavelength (period) of a sine wave;
M,	memory parameter of an approximate fractional noise, that is, point at which the sum defining such a noise is truncated;
N,	the number of core functions added to construct a locally Gaussian function;
Pr,	probability;
Q_H,	a constant;
$R(t, s)$,	sample sequential range for lag s;
s,	time lag;
$S^2(t, s)$,	the variance of s values of $X(t+1) \cdots X(t+s)$ around their sample average;
σ', σ'',	the scale parameters of the positive and negative tails of a hyperbolic distribution;
t,	time;
T,	total available sample size;
u,	'dummy' variable, that is, the index in a summation or integration;
$X(t)$,	a discrete time fractional noise;
$X^*(t)$,	$\sum_{u=1}^{t} X(u)$;
$Y(t)$,	the random process $X(t)$ $X(t+s)$;
$W_m(t)$,	a core function serving to define a locally Gaussian random function;
$Z(t)$,	a bilateral hyperbolic random function;
$\Omega_M(t)$,	a locally Gaussian random function.

REFERENCES

Anis, A. A., and E. H. Lloyd, On the range of partial sums of a finite number of independent normal variates, *Biometrika, 40,* 35, 1953.

Feller, W., The asymptotic distribution of the range of sums of independent random variables, *Ann. Math. Statis., 22,* 427, 1951.

Hurst, H. E., Long-term storage capacity of reservoirs, *Trans. Amer. Soc. Civil Engrs., 116,* 770, 1951.

Mandelbrot, B., New methods in statistical economics, *J. Political Econ., 71,* 421, 1963*a* (reprint in *Bull. Inst. Int. Statist., 40,* II-699, 1964*a*).

Mandelbrot, B., The variation of certain speculative prices, *J. Business Univ. Chicago, 36,* 394, 1963*b* (reprint in *The Random Character of Stock Market Prices,* edited by Paul H. Cootner, M.I.T. Press, Cambridge, Massachusetts, 1964*b*).

Mandelbrot, B., Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H. E. Hurst, *Compt. Rend. Acad. Sci. Paris, 260,* 3274, 1965.

Mandelbrot, B., How long is the coast of Britain? Statistical self-similarity and fractional dimension, *Science, 156,* 636, 1967.

Mandelbrot, B., Long-run linearity, locally Gaussian processes, H-spectra and infinite variances, *Int. Econ. Rev., 10,* 82, 1969.

Mandelbrot, B., Remarks on the lognormal and the Pareto distributions, in preparation.

Mandelbrot, B., and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Rev., 10,* 422, 1968.

Mandelbrot, B., and J. R. Wallis, Noah, Joseph and operational hydrology, *Water Resour. Res., 4,* 909, 1968.

Mandelbrot, B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, part 1, Averages and variances, *Water Resour. Res., 5,* 228–241, 1969*a*.

Mandelbrot, B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, part 2, Rescaled ranges and spectra, *Water Resour. Res., 5,* 242-259, 1969*b*.

Mandelbrot, B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, part 3, Mathematical appendix, *Water Resour. Res., 5,* 260–267, 1969*c*.

Mandelbrot, B., and J. R. Wallis, Some long run properties of geophysical records, *Water Resour. Res., 5,* 321-340, 1969*d*.

Moran, P. A. P., On the range of cumulative sums, *Ann. Inst. Statist. Math. (Tokyo), 16,* 109, 1964.

Moran, P. A. P., *An Introduction to Probability Theory,* Clarendon Press, Oxford, 1968.

(Manuscript received January 22, 1969.)