Water Resources Research: 5, 1969, 967-988 **H25**

Robustness of *R/S* in measuring noncyclic global statistical dependence (M & Wallis 1969c)

• ***Chapter foreword: two warnings.*** This heavily illustrated paper documents that the ratio *R/S* gives rise to an extremely useful statistical procedure. However, it is necessary to repeat two warnings already spelled out in Chapter H5.

A) When the sample variance $S^2(t, \delta)$ converges rapidly to a finite population variance, dividing $R(t, \delta)$ by $S(t, \delta)$ brings no value while introducing biases.

B) A chastening special example is provided by the PFSP processes, abbreviation for the "partly random fractal sums of pulses." The PFSP show by example that the "standard" relation $R/S \sim \delta^{1/2}$ is *not* always a signature of local dependence; to the contrary, it is compatible with an important new form of global statistical dependence. Therefore, this paper was recently obsoleted on a vital point. The exponent of *R/S* does *not* suffice to discriminate between local and global dependence. However, once again, *R/S* remains of wide utility. *Editorial changes. Strong*, as applied to statistical dependence, was viewed as too vague and replaced by *global*. The word *bridge* as applied to range was inserted where appropriate. The original denoted the lag by a lower case letter s and the standard deviation by an upper case letter S. The lag is now denoted by δ instead of s. Figure 1.

A matter of layout. To better accommodate the illustrations of this paper, many do not follow the first reference to them, but precede it. •

✦ **Abstract.** This paper investigates the rescaled range $R(t, \delta)/S(t, \delta)$ by extensive computer simulation and shows it to be a very robust statistic for testing the presence of noncyclic global statistical dependence. This robustness extends to processes that are extraordinarily far from from being Gaussian, for example, have huge values for skewness and/or kurtosis (that is, third and/or fourth moments). In cases where long dependence is present, we show how to estimate its intensity. ✦

THIS PAPER ADDRESSES THE ANALYSIS of empirical records in which very long statistical dependence (excluding seasonals or other cycles) may be present. Problems relative to global dependences are increasingly recognized as being on the forefront of both theoretical and practical statistics. Until now, however, the main statistical technique to treat the very global was spectral analysis, which performs poorly on processes that are far from being Gaussian. Figure 2. Figure 3. Figure 4. Figure 5. •

We reinterpreted Hurst's graphs as suggesting a new technique for data analysis, to be called "R/S analysis." We found this technique to be very effective in its context, and this paper examines the principal reasons for our enthusiasm. Actual records are examined in M & Wallis 1969b{H27}, but this paper comments on Hurst's early empirical results concerning R/S. In this ratio, $R(t, S)$ is the bridge range defined by

$$R(t, \delta) = \max_{0 \le u \le \delta} \{X_\Sigma(t+u) - X_\Sigma(t) - (u/\delta)[X_\Sigma(t+\delta) - X_\Sigma(t)]\}$$
$$- \min_{0 \le u \le \delta} \{X_\Sigma(t+u) - X_\Sigma(t) - (u/\delta)[X_\Sigma(t+\delta) - X_\Sigma(t)]\}.$$

As to $S(t, \delta)$, it is the sample standard deviation of $X(s)$. That is, $S^2(t, \delta)$ Figure 6. Figure 7. Figure 8. Figure 9. Figure 10. •

$$S^2(t, \delta) = \delta^{-1} \sum_{u=1}^{\delta} \{X(t+u) - \delta^{-1}[X_\Sigma(t+\delta) - X_\Sigma(t)]\}^2$$
$$= \delta^{-1} \sum_{u=1}^{\delta} X^2(t+u) - [\delta^{-1} \sum_{u=1}^{\delta} X(t+u)]^2.$$

FIGURE C25-1. Construction of the sample range $R(t, \delta)$ {P.S. 1999. This figure in the original reproduces Figure 1 of M & Wallis 1969a{H13}. To reprint it here would be redundant hence it was deleted, but this caption was preserved to avoid renumbering the other figures.}

This paper shows, that in many cases one has $R(t, \delta)/S(t, \delta) \sim \delta^{J+1/2}$. When the exponent J is defined, it satisfies $-0.5 < J < 0.5$ and measures what may be called the R/S intensity of statistical dependence. This is a precise and useful measure for at least one aspect of the more general concept of the intensity of noncyclic global statistical dependence. The special value $J = 0$ corresponds to the absence of very global dependence. Consequently, the dependence on δ of the average of the sample values of $R(t, \delta)/S(t, \delta)$, carried over all admissible starting points t within the sample, can be used to test whether the R/S intensity is nonvanishing and to estimate the value of this intensity. The execution of these statistical

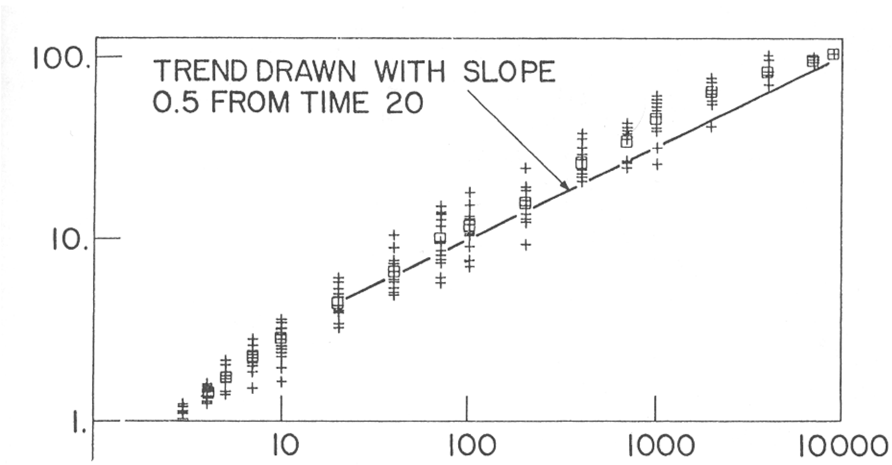

FIGURE C25-2. Pox diagram of $\log R/S$ versus $\log \delta$ for a sample of 9000 values of a discrete white noise $G(t)$, that is, of independent, identically distributed Gaussian random variables, as plotted in M & Wallis 1969a{H12}. Skewness (0.1) and kurtosis (2.96) are as expected.

The boxes correspond to estimates of $\log \mathscr{E}[R/S]$. Their disposition is evidence that the $\delta^{0.5}$ law in the mean applies to $G(t)$ after a very short transient, shorter than the transient of $R(t, \delta)$ derived by Anis & Lloyd 1953.

The crosses $(+)$ correspond to sample values of R/S for $\log \delta$ restricted to the sequence 3, 4, 5, 7, 10, 20, 40, 70, 100, 200, 400, 700, 1000, 2000, 4000, 7000 and 9000. For $\delta < 500$, 14 crosses $(+)$ are plotted, corresponding to values of t equal to 1, 100, ... , 1400. For $\delta > 500$, t was made successively equal to 1000, 2000, up to the smaller of 8000 or $T - \delta + 1$.

The dispersion of R/S around $\delta^{-0.5}$ (that is, the dispersion of $\delta^{-0.5}\ R/S$) depends little on δ. This gives evidence that the $\delta^{0.5}$ law in distribution applies to $G(t)$ after the same short transient, the relative dispersion of $\delta^{-0.5}R/S$ being small. Plotting $\log R(t, \delta)$ instead of $\log R/S$ would not significantly change this diagram, but the initial transient would lengthen.

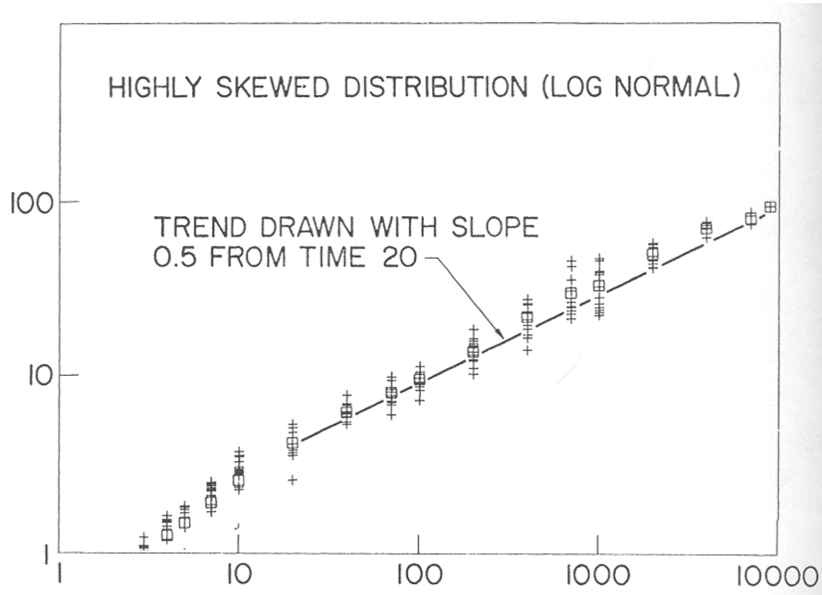

FIGURE C25-3. Pox diagram of log $R(t, \delta)/S(t, \delta)$ versus log δ for an independent log normal random function, i.e., $10^{G(t)}$, where $G(t)$ is a discrete white noise of mean 0 and variance 1. Skewness is 27.33 and kurtosis is 1050.16, both very high, which are symptoms of the Noah Effect. Nevertheless, the disposition of the boxes indicates that the $\delta^{0.5}$ law in the mean is satisfied by $10^{G(t)}$ after a short transient, whereas the dispersion of the sample values (+) indicates that the $\delta^{0.5}$ law in distribution is satisfied with a small relative dispersion.

A plot of $10^{G(t)}$, on linear coordinates could show either the peaks or the details of the small values, but not both. Logarithmic coordinates make the plot legible, but amount to plotting $G(t)$ itself.

The argument in Feller 1951 shows that the $\delta^{0.5}$ law holds asymptotically for $10^{G(t)}$. The study of the penultimate region of moderate values of δ is an entirely different matter. The proof that the asymptotic results remain applicable proceeds as follows. The basic fact is that in the range of moderately large values of δ the log normal density can be approximated by an appropriate hyperbolic, as defined in the text. (This assertion is proved in an unpublished paper by BBM {P.S. 1999, whose contents is incorporated in Chapter E9 of M 1997E}.) As a result, the penultimate behavior of R/S is essentially the same for the log normal process and an appropriate hyperbolic process.

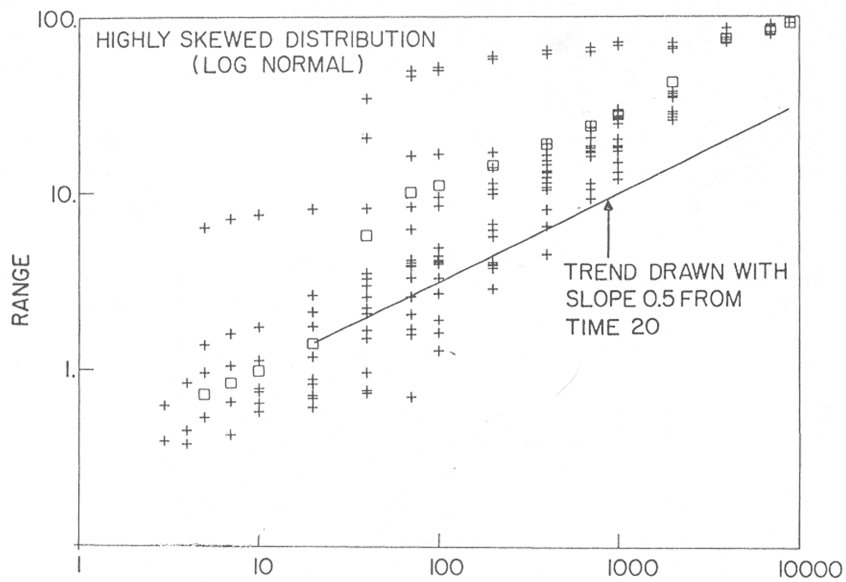

FIGURE C25-4. Pox diagram of $\log R(t, \delta)$ versus $\log \delta$ for the sample of the random function $10^{G(t)}$ used in Figure 3. Again, skewness is 27.33 and kurtosis is 1050.16. The diagrams of $\log R$ and $\log R/S$ (Figure 3) differ dramatically. There is evidence that the $\delta^{0.5}$ law in the mean (boxes) applies to, but the transient goes up to $\delta \sim 70$, making it longer than the transient of $\log R/S$ for either $G(t)$ or $10^{G(t)}$. However, the distribution of the sample values of $R(t, \delta)$ (indicated by +) around their average never stabilizes. Even a sample of 9000 values gives no evidence of the $\delta^{0.5}$ law in distribution. The scatter of the crosses (+) is so extreme that if the sample was much shorter, testing whether the $R(t, \delta)$ function of $10^G(t)$ obeys the $\delta^{0.5}$ law would be hard at best and often hopeless. That is, the statistic $R(t, \delta)$ is much less robust than R/S.

The pox diagram of $\log S(t, \delta)$ versus $\log \delta$ need not be reproduced because it is very similar to the present diagram. In other words, the two functions $\log S(t, \delta)$ and $R(t, \delta)$ are widely scattered but mesh so precisely together that, as seen in Figure 3, little scatter remains in the difference $\log R(t, \delta) - \log S(t, \delta) = \log R/S$. It is an important lesson: an excellently behaved statistic can sometimes be obtained by combining expressions that behave badly when considered separately.

Additional insights into meshing are yielded by the behavior of R/S for the very skew Gamma process discussed in the body of this paper. Further insights are yielded by the legend of Figure 11.

tasks is labeled *R/S* analysis, a term we coined after "spectral analysis." R/S intensity is measured by the exponents J or $H = J + 0.5$, with $0 < H < 1$.

Without question, the first discipline in which the presence of noncyclic very global dependence has been reported is hydrology. Therefore, we say that all fields exhibiting noncyclic very global dependence exhibit the *Joseph Effect.* As described in M & Wallis 1968 {H10}, the original Joseph Effect expresses a well-established fact that high or low levels in rivers tend to persist, such as over the Biblical "seven fat and seven lean years," but more often over decades, centuries and millennia. Similar observations have been made in meteorology, geophysics, economics, physics, and other sciences. Using our 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.

Could other measures of the strength of the Joseph Effect have been used? The answer depends on whether or not the record in question is nearly Gaussian. If it is, one can also use the method of variance-time

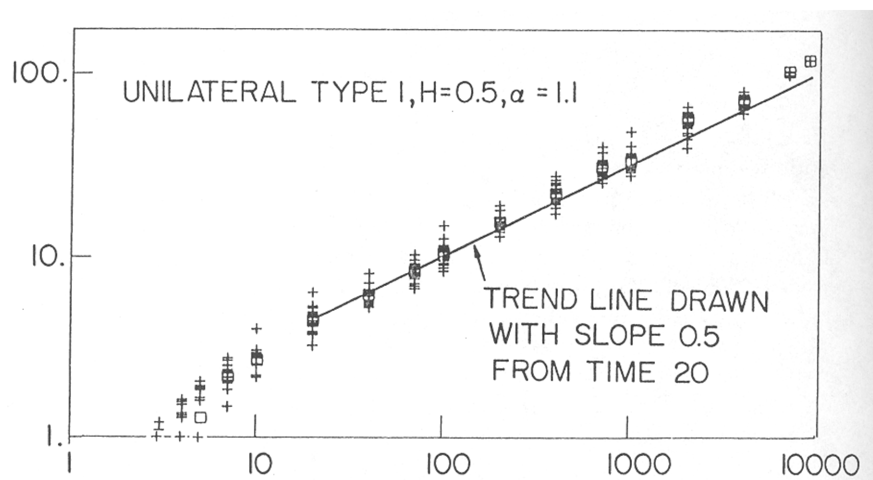

FIGURE C25-5. Pox diagram of $\log R(t, \delta)/S(t, \delta)$ versus $\log \delta$ for a sequence of independent, identically distributed, hyperbolic random variables with exponent $\alpha = 1.1$, the term "hyperbolic" being defined in the text. Skewness is 52.14 and kurtosis is 3253.10, both very high. A linear coordinate plot of this function, would, like that of the function $10^{G(t)}$ itself, be illegible. The disposition of the boxes is evidence that the $\delta^{0.5}$ law in the mean applies after a short transient. The disposition of the sample values (indicated by $+$) is evidence that the $\delta^{0.5}$ law in distribution applies with a relative dispersion, even smaller than in the Gaussian case.

curves introduced by G. I. Taylor in 1921 or perhaps one of a few alternative statistical techniques.

However, many natural records are extremely non-Gaussian. This finding was also first reported in Biblical hydrology, and 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. Standard books of statistics include the Noah Effect under the heading of the "theory of extreme values." Figure 14. Figure 15. Figure 16. Figure 17. •

Because of the Noah Effect, the question 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 Effects.

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

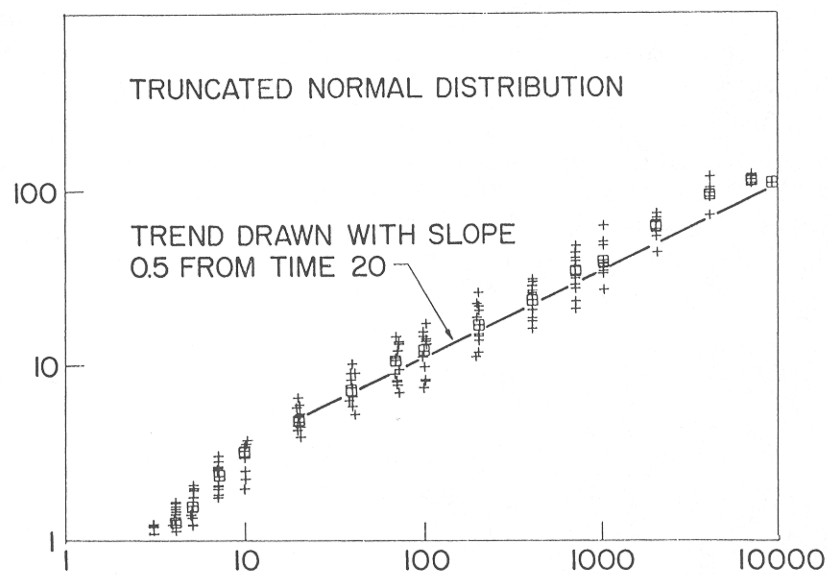

FIGURE C25-6. Pox diagram of $\log R(t,\delta)/S(t,\delta)$ versus $\log \delta$ for a sequence of independent random variables with the skewness 0.00 and the very low kurtosis 1.82. To achieve this result, a normal random variable were truncated so severely as to make its distribution almost uniform. The validity of the $\delta^{0.5}$ law in the mean is no more affected by low kurtosis than it was by the high kurtosis in Figures 3 and 5. The relative dispersion of $\delta^{-0.5}$ R/S is larger perhaps than in the Gaussian case.

Noah Effect. When there is no Noah Effect, other available techniques may be comparable in effectiveness to R/S analysis. However, when the Noah Effect is strong, the best alternative techniques known to us are simply less effective, their sampling distribution being less favorable. The worst are worthless because they confuse the Noah and Joseph Effects inextricably. Figure 18. Figure 19. •

The claims made in this paper, are not proved mathematically rather demonstrated by computer simulation buttressed by some heuristic arguments. Many of the figures have unusually detailed captions, which should be considered an integral part of the exposition. Figure 20. •

Our use of the term "law" follows the custom established by the law of large numbers, which designates the statement 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 in this paper lurks a theorem of the form "under such and such hypothesis, the function $\delta^{J+1/2}$ describes in such and such way the behavior of $R(t, \delta)/S(t, \delta)$." The phrase "$\delta^{J+1/2}$ law" will not designate a specific theorem, but rather the conclusion common to a number of theorems. Figure 21. Figure 22. Figure 23. Figure 24. •

For stationary Gaussian processes without global dependence, division by $S(t, \delta)$ is a useful, but not vital, detail. In this case, R/S analysis is a small improvement over the analysis of R itself, as performed on white Gaussian noise in Feller 1951 and Anis & Lloyd 1953. The importance of the division of $R(t, \delta)$ by $S(t, \delta)$ increases as the process diverges from the Gaussian and/or as one introduces dependence of an increasingly longer

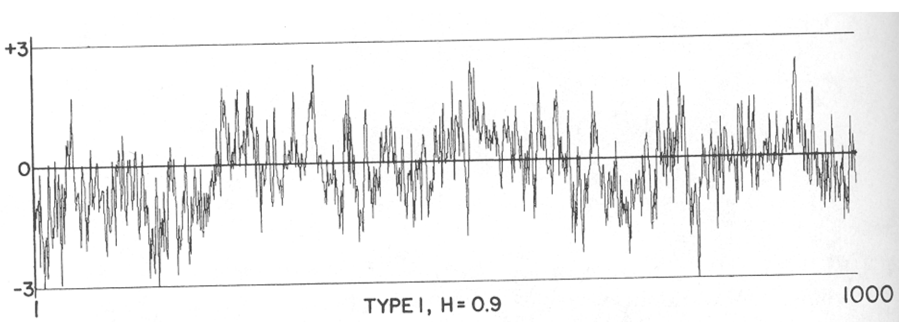

FIGURE C25-7. The first 1000 values from a 9000 value sample of Type 1 approximation to fractional noise with $H = 0.9$ and M equal to either 10 or 100. The whole sample of 9000 has been normalized to have zero mean and unit variance. This figure is from M & Wallis 1969a{H12}.

extent. M 1965h{H9} and M & Van Ness 1968{H11} argue, and M & Wallis 1969c{H25} shows that Gaussian random processes with a fractional spectrum, in which dependence is global, are very effective for modeling the Joseph Effect. However, it still remained to study the behavior of R/S for non-Gaussian processes and for those with strong cyclic components, and to compare R/S analysis with other methods of analyzing global statistical dependence (including R analysis and many other techniques that were apparently first considered in our work as possible alternatives to R/S analysis). We have pursued all these tasks with the help of computer simulation, but only the main results concerning R/S are reported.

MATHEMATICAL PRELIMINARY

If $X(t)$ is a stationary random process, the ratio $R(t,\delta)/S(t,\delta)$, considered for fixed δ 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

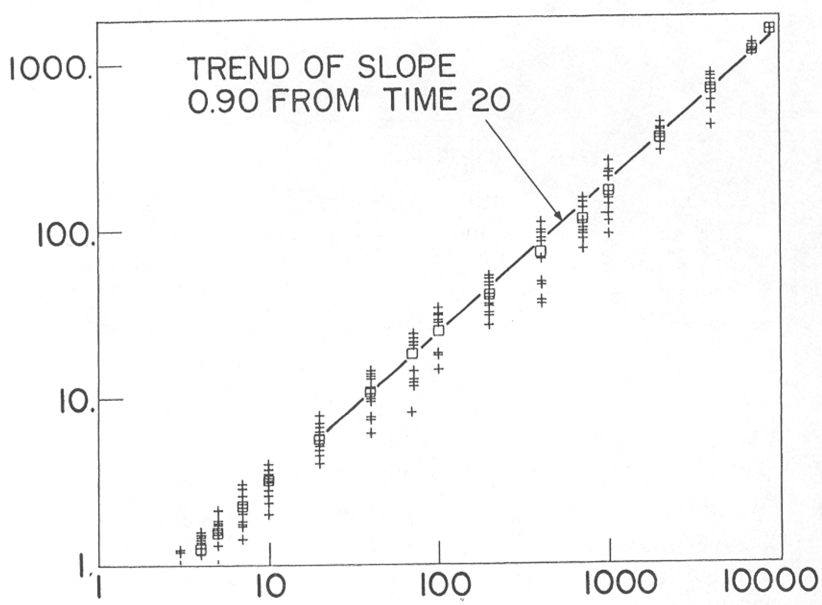

FIGURE C25-8. Pox diagram of $\log R(t,\delta)/S(t,\delta)$ versus $\log \delta$ for 9000 values of a Type 1 approximation to the fractional noise with $H = 0.9$ and $M = 10,000$. This figure is reproduced from M & Wallis 1969a. The R/S intensity of statistical dependence is clearly equal to $H = 0.9$.

stationary if identical rules generate the process $X(t)$ and all processes of the form $X(t+\delta)$, which are deduced from $X(t)$ by a time shift.

To appreciate fully the manipulations concerning R/S in this paper, it is useful to view classical covariance analysis as being based on the fact that, when $X(t)$ is stationary, the transformed process $Y(t)=X(t)X(t+\delta)$ is also stationary for every δ. Since the covariance of $X(t)$ may be written as $\mathscr{E}[X(t)X(t+\delta)]=\mathscr{E}[Y(t)]$, such covariance depends on δ but is not a function of t and, therefore, can be designated by $C(\delta)$. It is known that many features of a process are fully described by the functional dependence of $C(\delta)$ on δ.

R/S analysis is based also on the properties of a family of stationary random functions obtained by transforming $X(t)$, namely, the function $R(t,\delta)/S(t,\delta)$. Stationarity implies that $\mathscr{E}[R(t,\delta)/S(t,\delta)]$, like $C(\delta)$ above,

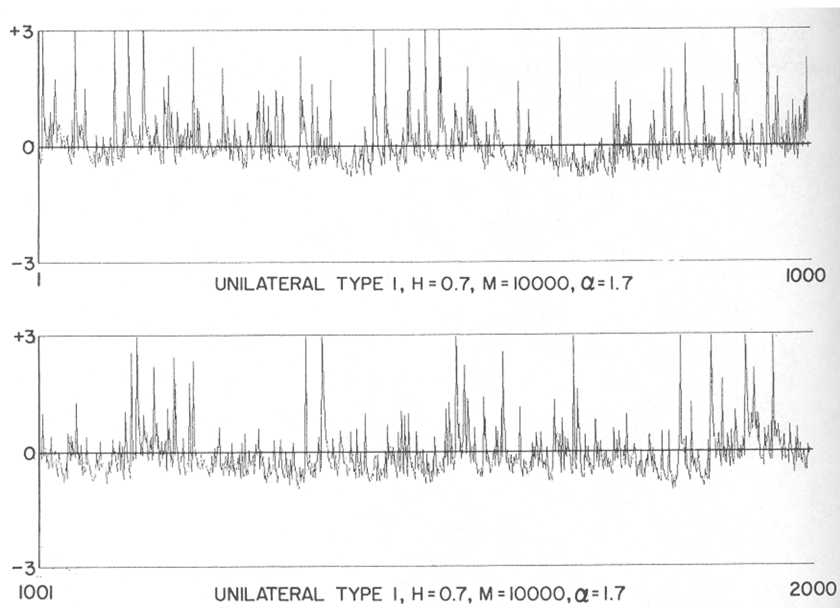

FIGURE C25-9. The first 2000 values from a 9000 value sample of a hyperbolic non-Gaussian fractional noise. To construct this sample, we preserved the same moving average kernel already used to construct a Type 1 fractional Gaussian noise, M & Wallis 1969a{H12}. However, the variables to be averaged were hyperbolic, a concept defined in the text and a symptom of a very strong Noah Effect. The largest values of this sample of hyperbolic fractional noise exceeded the bounds of the graph and were truncated. Thus values that seem to equal the maximum plottable value are in fact larger. Each large value has strong and long-lived after effects.

depends on δ but not t. We shall show that some important properties of a process are described by the functional dependence of $\mathscr{E}[R(t,\delta)/S(t,\delta)]$ on δ.

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

R/S TESTING

The behavior of $R(t,\delta)/S(t,\delta)$ as $\delta \to \infty$ defines the concept of R/S dependence, which is a form of noncyclic global statistical dependence. Thus, the first application of R/S analysis occurs in testing for R/S dependence in a record.

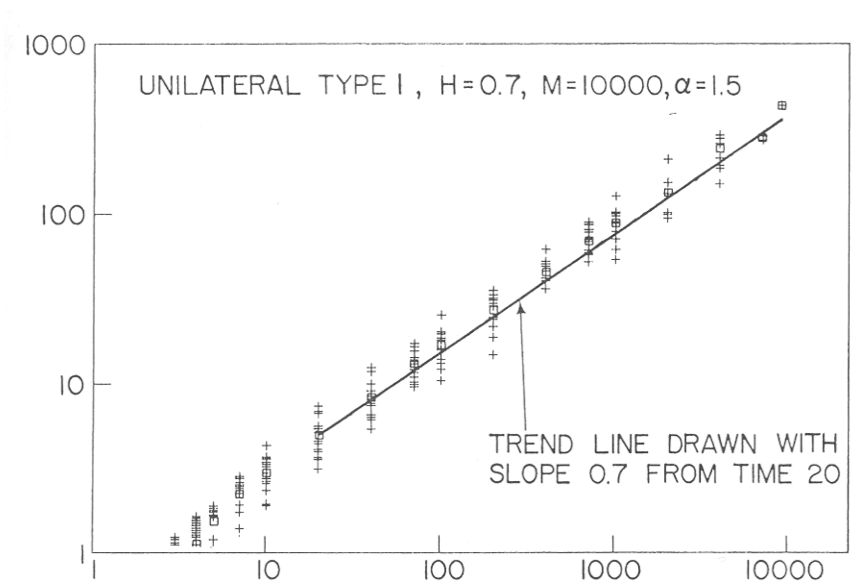

FIGURE C25-10. Pox diagram of $\log R(t,\delta)/S(t,\delta)$ versus $\log \delta$ for 9000 values of the hyperbolic non-Gaussian fractional noise, plotted in part in Figure 9. The skewness is 18.46 and the kurtosis is 602.39, both very high. The $\delta^{0.7}$ law clearly holds both in the mean and the distribution, indicating that R/S analysis is blind to the extremely non-Gaussian character of the marginal distribution (strong Noah Effect) even when very global dependence (strong Joseph Effect) has been built into the process in question.

Preliminary. The concept of global statistical dependence is obviously important, but it is complex and many faceted; generally accepted definitions are lacking. However, some random processes exist for which global dependence is unquestionably present. Moreover, cyclic and noncyclic global dependence must be distinguished (M 1969e, Section 2.2). Having analyzed many processes, we have observed a relation between noncyclic global dependence and have defined the following law.

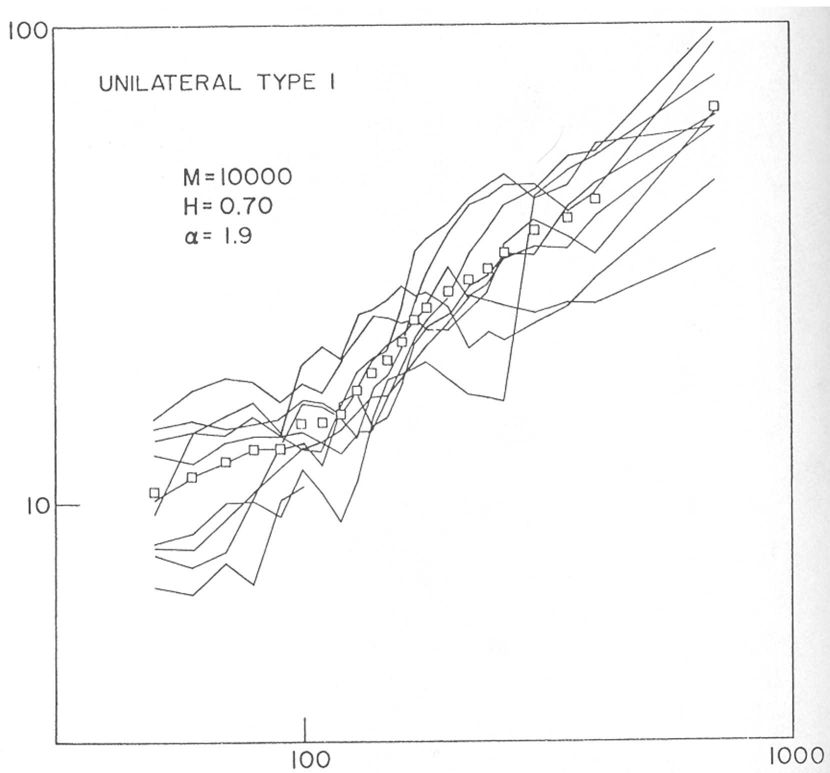

FIGURE C25-11. Alternative plot, not a pox diagram, of a greatly enlarged detail of the variation of $R(t, \delta)/S(t, \delta)$ for a hyperbolic process different from the process used in Figure 10. The skewness is 9.09 and the kurtosis is 184.38, both very high. The sample path of $R(t, \delta)/S(t, \delta)$ as a function of δ was plotted for several starting points. It is evident 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.

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

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

exists (that is, is well defined) and is positive and finite. In more intuitive terms, this means that the graph of $\log \mathscr{E}[R(t, \delta)/S(t, \delta)]$ versus $\log \delta$ is,

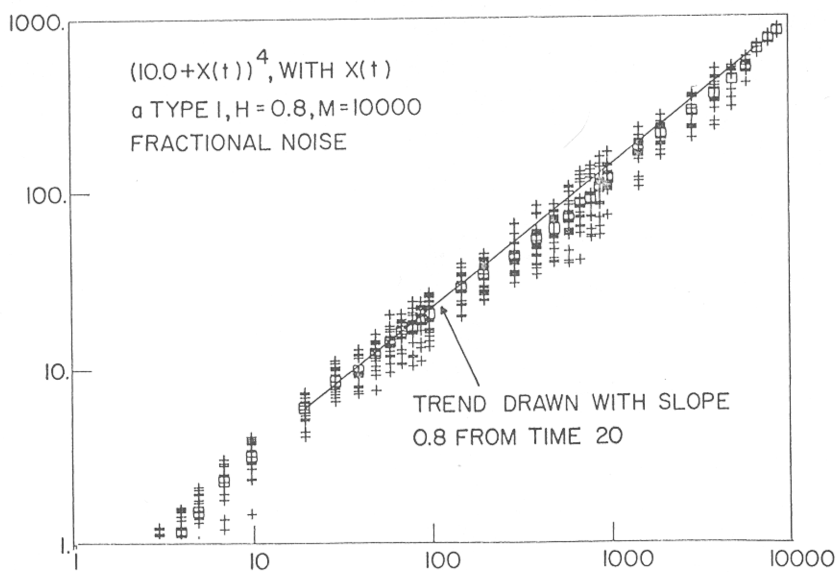

FIGURE C25-12. Effect of the nonlinear transformation $(10 + X)^4$ on an approximately fractional Gaussian noise. The skewness is 1.87 and the kurtosis is 8.75, both high. In the initial process the Joseph Effect is strong but the Noah Effect is absent. The transformed process exhibits a moderate Noah 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 function of the total yearly precipitation at the site of the tree but is probably nonlinear. The yearly maximum and minimum of a river's levels are increasing but presumably nonlinear functions of the yearly discharge. We studied a fourth power to illustrate such nonlinearity. This figure strongly suggests that the R/S intensity that is estimated from tree ring thickness (respectively, from river levels) can reasonably be expected to apply also to yearly precipitation (respectively, to yearly flows).

asymptotically, a straight line of slope 0.5. The $\delta^{0.5}$ law in the mean fails to hold in two cases:

(1) when $\delta^{-0.5}\mathscr{E}[R(t, \delta)/S(t, \delta)]$ oscillates with no limit as $\delta \to \infty$,

(2) when this quantity tends to either zero or infinity.

In either case, the graph of log $\mathscr{E}[R(t, \delta)/S(t, \delta)]$ versus log δ does not possess a straight asymptote of slope 0.5.

{P.S. 1999. Generality and mathematical exactitude would suggest that the existence of a limit be replaced by the two less demanding conditions lim sup $< \infty$ and lim inf > 0. The early treatment in this chapter forsook this higher level of generality as being premature.}

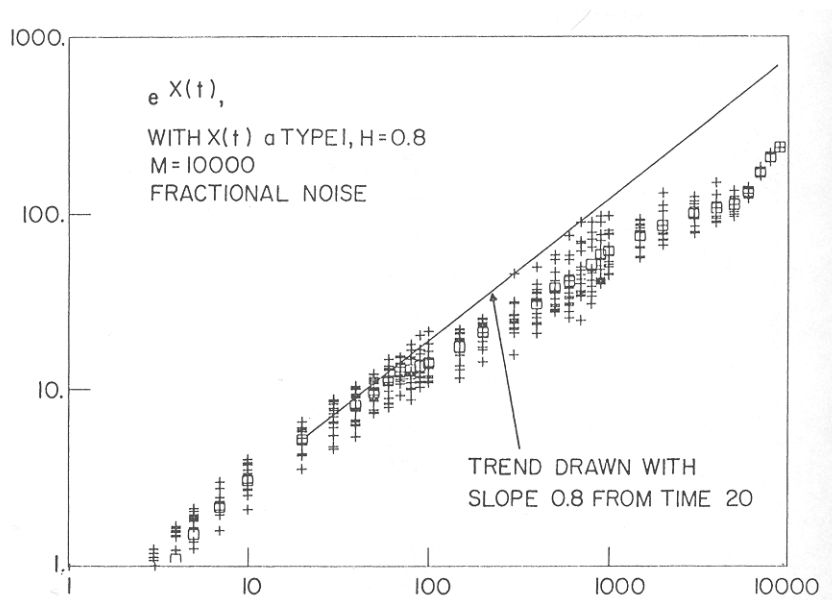

FIGURE C25-13. Effect of the extremely nonlinear transformation e^X, when applied to a fractional Gaussian noise. The skewness is 39.32 and the kurtosis is 2277.42, both very high. Over the span of values of δ that have been considered, the slope of the trend line of this diagram is much smaller than H. This shows that nonlinear transformations, if sufficiently extreme, may not preserve the R/S intensity of the original process. The resulting process need not even have a well-defined R/S density. {P.S. 1999. The effect of non-linear transformations on diverse "$1/f$ noises" are a topic of very great interest briefly discussed in M 1999N, Chapter N4.}

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

Examples. The stationary process of independent reduced Gaussian variables is unquestionably the simplest process with non existent, hence local, dependence. The term "reduced" means that the expectation has vanished and the variance is unity. For this process, the law of large numbers shows that $\lim_{\delta \to \infty} S(t, \delta) = 1$. In addition, Feller 1951 proved the existence of $\lim_{\delta \to \infty} \delta^{-0.5} \mathscr{E}[R(t, \delta)] = \sqrt{\pi/2}$, a number that is approximately 1.25. Thus $C = \lim_{\delta \to \infty} \delta^{-0.5} \mathscr{E}[R(t, \delta)/S(t, \delta)]$ is also about 1.25 for independent Gaussian processes that have not been reduced. The fact that this limit is both positive and finite establishes that these processes satisfy the $\delta^{0.5}$ law

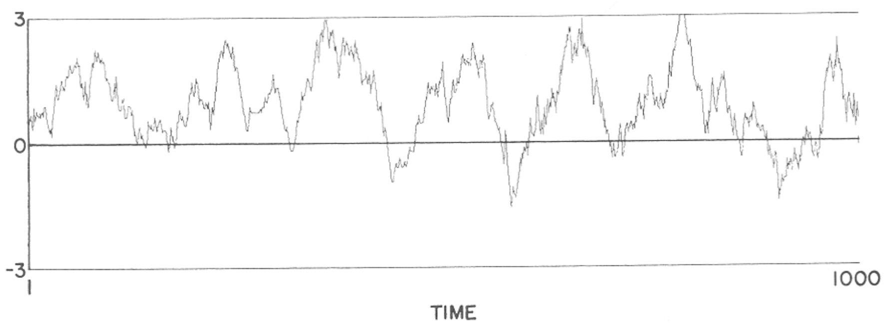

FIGURE C25-14. The first 1000 values from a 9000 value sample of a "locally Gaussian random process". This notion was introduced in M 1969a to resolve certain paradoxes encountered in attempts to model economic time series by strictly Gaussian processes. Because of seemingly contradictory properties, locally Gaussian processes are an especially tough challenge to data analysis.

This figure concerns the process $\Omega_N(t) = N^{-1/2}\sum_{m=1}^{N} W_m(t)$, where each $W_m(t)$ is a "coin" process constructed by the following three steps. The first step 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 satisfying $\Pr\{U_k > u\} = u^{-\beta}$. A second step selects for $W_m(T_k)$ a sequence of independent Gaussian random variables of zero mean and unit variance. A third step identifies the interval from T_k to T_{k+1} in which the instant t is located, and sets $W_m(t)$ 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 $N = 10$. {P.S. 1999. Coin processes, called "core processes" in the original, are investigated in M 1967i{N9}.}

of the mean; an experimental confirmation is shown in Figure 2. Figures 2, 3, 5 and 6 demonstrate that the $\delta^{0.5}$ law also applies to processes of independent values having a variety of other marginal distributions: truncated Gaussian, log normal and hyperbolic, respectively.

Here, the random variable X is called "hyperbolic" if, for large values of x, it satisfies the two relations $\Pr\{X > x\} \sim (x/\sigma')^{-\alpha}$ and $\Pr\{X < -x\} \sim (x/\sigma'')^{-\alpha}$, 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, M 1963e{E3}.

The simulations reported in this paper concern the case where $\sigma'' = 0$ and $\sigma' = 1$, and the case where $\sigma' = 2^{-1/\alpha}\sigma''$. We began with a sequence $F(t)$ of independent random variables, uniformly distributed between 0 and 1. Next, a bilateral hyperbolic function $Z(t)$ was constructed using the formulas:

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

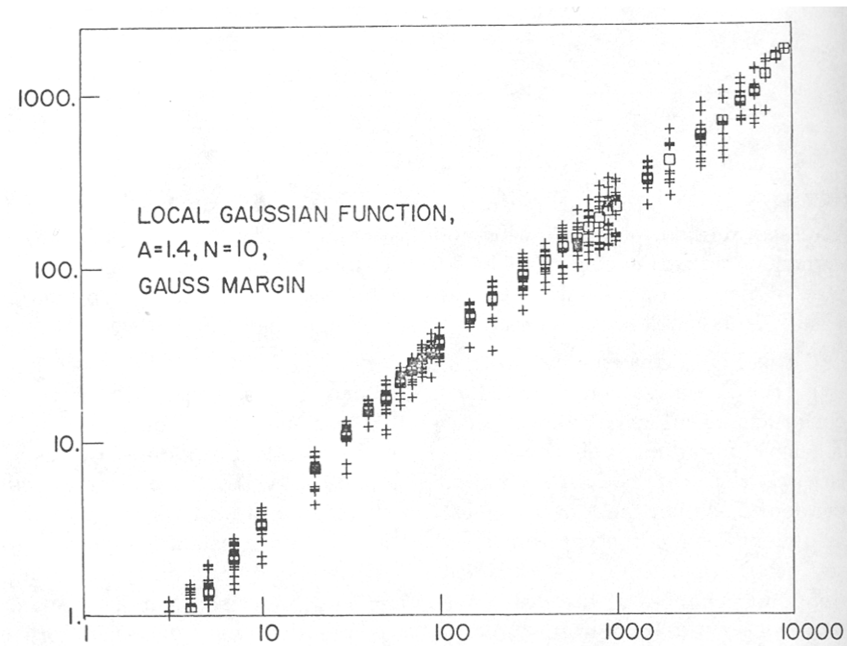

FIGURE C25-15. Pox diagram of $R(t,\delta)/S(t,\delta)$ for 9000 values of a locally Gaussian noise, including and continuing the sample of Figure 14.

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

To simplify subsequent calculations, $Z(t)$ was rounded to its integer part. The unilateral hyperbolic function was defined as $|Z(t)|$. In Figures 2 to 6, the expression $\delta^{-0.5}\mathscr{E}[R(t,\delta)/S(t,\delta)]$ attains its limit value very rapidly, that is, after a brief initial transient. Note, however, that the precise value of $\lim_{\delta\to\infty}\delta^{-0.5}\,\mathscr{E}[R(t,\delta)/S(t,\delta)]$ greatly depends on the process; this will be exemplified later in the paper.

It should be noted that the values of $R(t,\delta)/S(t,\delta)$ for small log δ, as plotted in the figures of this paper, have been computed incorrectly. They must be disregarded. However, the paper's conclusions remain unaffected. The nature of the error will be explained in Taqqu 1970. {P.S.

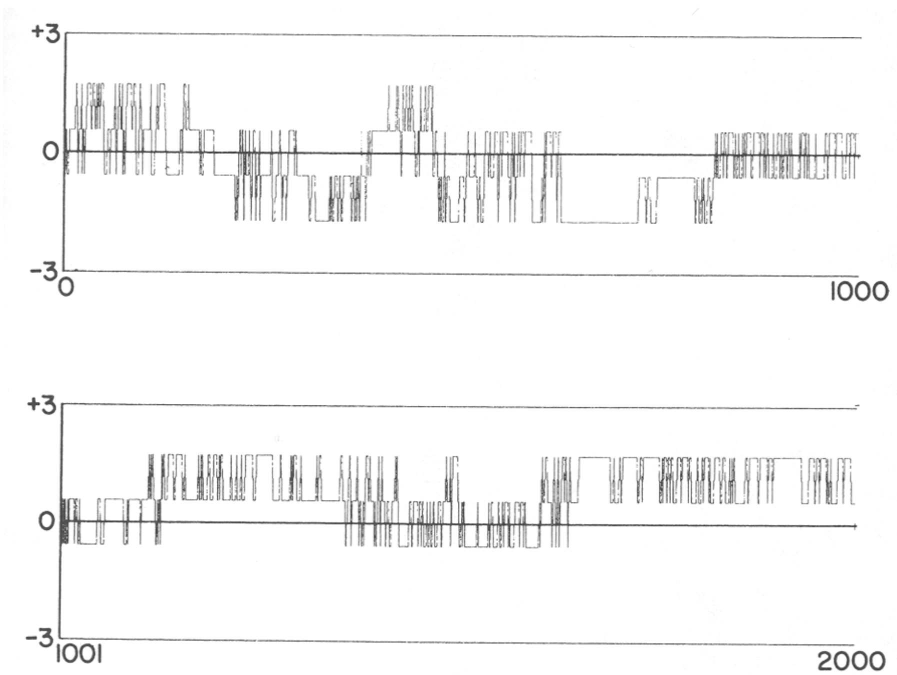

FIGURE C25-16. The first 2000 values from a 9000 value sample of a different locally Gaussian random process. The constructions proceeded as for the function plotted in Figure 14, except that we selected for $W_m(T_k)$ a sequence of independent binomial random variables of zero mean and unit variance equal to $+1$ or -1 with probabilities 0.5. In this figure, $N = 3$.

1999. In this reprint, the bottoms of the figures, which were incorrect, were not reproduced.}

When the values of the process $X(t)$ are statistically dependent, the dependence is limited to the local. The transient is much longer, but the $\delta^{0.5}$ 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 M & Wallis 1969a {H13} show that the $\delta^{0.5}$ law in the mean fails for a variety of processes for which the dependence between $X(t)$ and $X(t+T)$ decreases extremely slowly to zero as $T \rightarrow \infty$. A more detailed discussion of these figures is best postponed to a later section devoted to R/S estimation.

How to account for Hurst's empirical δ^H law. In empirical records, the values of R/S were found to cluster closely along a function of the form $\delta^{J+1/2}$ with $H > 0.5$. The finding that $H > 0.5$ was made originally by Hurst 1951, although Hurst's estimates of H involved a far-reaching conceptual

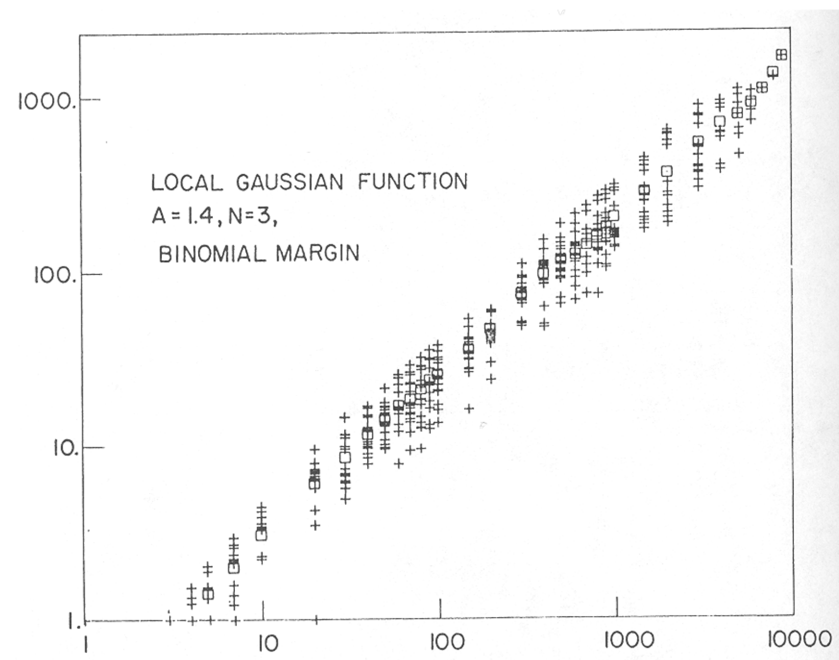

FIGURE C25-17. Pox diagram of $R(t,\delta)/S(t,\delta)$ for 9000 values of a locally Gaussian noise, including and continuing the sample of Figure 16.

error discussed in M & Wallis 1969b{H27}. Feller 1951 proved that the empirical $\delta^{J+1/2}$ law is incompatible with the idea that the records in question were generated by an independent Gaussian process. Then several authors, including Moran 1964, 1968, argued that the empirical $\delta^{J+1/2}$ law could be accounted for by postulating that the records were generated by

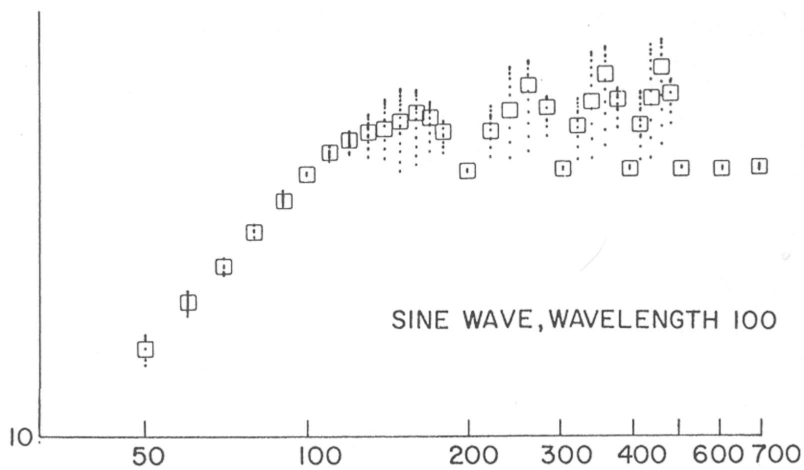

FIGURE C25-18. Pox diagram of $\log[R(t,\delta)/S(t,\delta)]$ versus $\log\delta$ for a pure sine wave with $L=100$. Values of the form $\delta=kL$ correspond to subharmonics of the sine wave. For them, $R(t,\delta)/S(t,\delta)$ is independent of t and k, as can be seen from the theory. Big lobes are, however, visible for other values of δ. If $R(t,\delta)/S(t,\delta)$ had been computed for δ in a grid that eventually merges with the grid of the subharmonics of the pure sine, $R(t,\delta)/S(t,\delta)$ would rapidly attain its asymptotic limit. However, if the grid is selected independently of the value of L, then R/S is more likely to fall within the lobes. The result is a pox diagram of $\log[R(t,\delta)/S(t,\delta)]$ versus $\log\delta$ having a positively sloped trend line. Thus, a small sample of a pure sine wave could be declared by R/S to have a small 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 the data to eliminate the cycles, one should compute $R(t,\delta)/S(t,\delta)$ or its average for as many values of δ as one can manage. M & Wallis 1969a{H12} show that, contrary to spectral analysis, one need not smooth out the behavior of $R(t,\delta)/S(t,\delta)$ by averaging its values over neighboring values of δ. We now add the observation that such smoothing would also mix the cyclic effects with noncyclic global dependence to produce an apparent value of H devoid of significance.

a random process with independent values and a very skew marginal distribution. In our vocabulary, these authors postulated that the empirical $\delta^{J+1/2}$ law relates to the Noah Effect. The results in the present paper show that this Noah Effect explanation is *insufficient.* Earlier, M 1965h, M & Van Ness 1968{H11} and M & Wallis 1969a{H12}, explained the $\delta^{J+1/2}$ as being one aspect of the Joseph Effect, hence made the Noah Effect explanation *unnecessary.*

Irrelevance of a specific example of Moran's argument. Moran 1968 (p. 495) attempted to illustrate his proposed account of Hurst's empirical law by considering Gamma distributed random variables of density $[\Gamma(\gamma)]^{-1}x^{\gamma-1}e^{-x}$, where the parameter γ is very small. Moran's illustration is fallacious, as we shall now 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

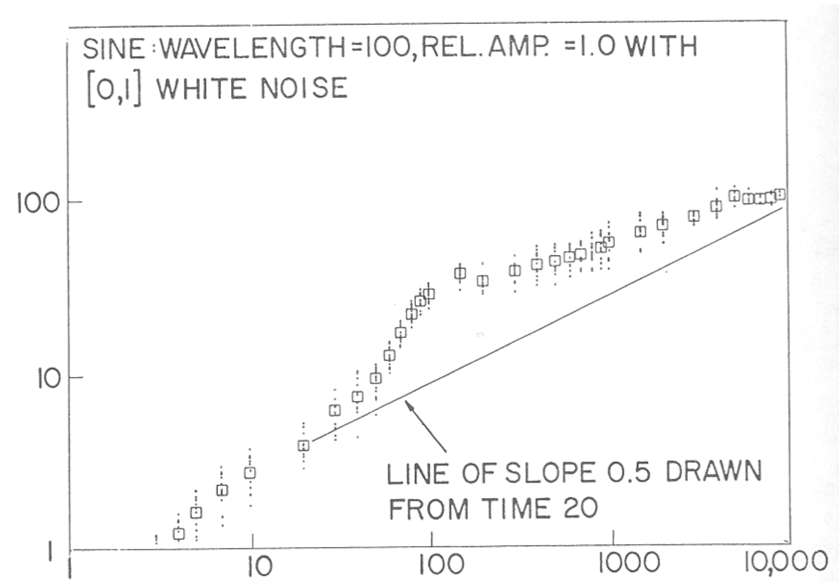

FIGURE C25-19. The pox diagram of log $R(t, \delta)/S(t, \delta)$ versus log δ for the sum of a pure sine wave and a white noise of comparable amplitudes. The behavior of this function is a clear hybrid of the behaviors of each of the functions plotted in Figures 2 and 18. If the sine amplitude was stronger, the asymptote of slope 0.5, characteristic of the noise component, would fail to prevail for the lags plotted on this figure. If the sine amplitude was smaller, the wiggles and lobes characteristic of the pure sine component would be less viable.

the order of $1/\gamma$, there is a very high probability that $X_\Sigma(t) = \Sigma_{u=1}^{t} X(u)$ is almost indistinguishable from $\max_{0 \le u < t} X(u)$. As a result,

$$R(t, \delta) \sim \max_{0 \le u \le \delta} X(t+u).$$

We can then write,

$$\sum_{u=1}^{\delta} X^2(t+u) \sim [\max_{0 \le u \le \delta} X(t+u)]^2,$$

and

$$\begin{aligned} S^2(t, \delta) &= \delta^{-1}[\max_{0 \le u \le \delta} \ X(t+u)]^2 - [\delta^{-1} \max_{0 \le u < \delta} \ X(t+u)]^2 \\ &= \delta^{-1}(1-\delta^{-1})[\max_{0 \le u \le \delta} \ X(t+u)]^2. \end{aligned}$$

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

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

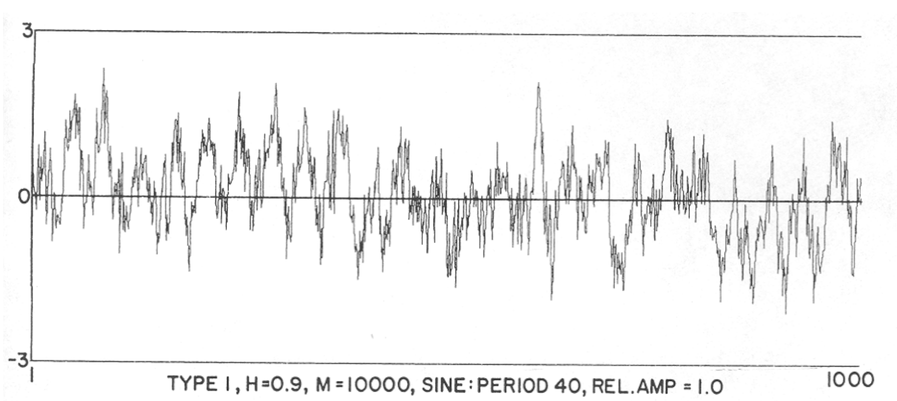

FIGURE C25-20. The first 1000 values from a 9000 value sample of the sums of a fractional noise and a moderately strong sine wave.

Formal definition of R/S independence. The examples we considered introduce a distinction between two kinds of random process: those for which $\lim_{\delta\to\infty} \delta^{-0.5}\mathscr{E}[R(t, \delta)/S(t, \delta)]$ exists and is positive and finite, and those for which the limit is either nonexistent, or 0, or infinite. This alternative has been stated purposefully in terms such that every random process falls on one or the other side and, therefore, can be used as a basis of the following formal definition of dependence. Every process with the quality that $\lim_{\delta\to\infty} \delta^{-0.5}\mathscr{E}[R(t, \delta)/S(t, \delta)]$ is finite and positive, will be said to be R/S independent. All other processes are R/S dependent.

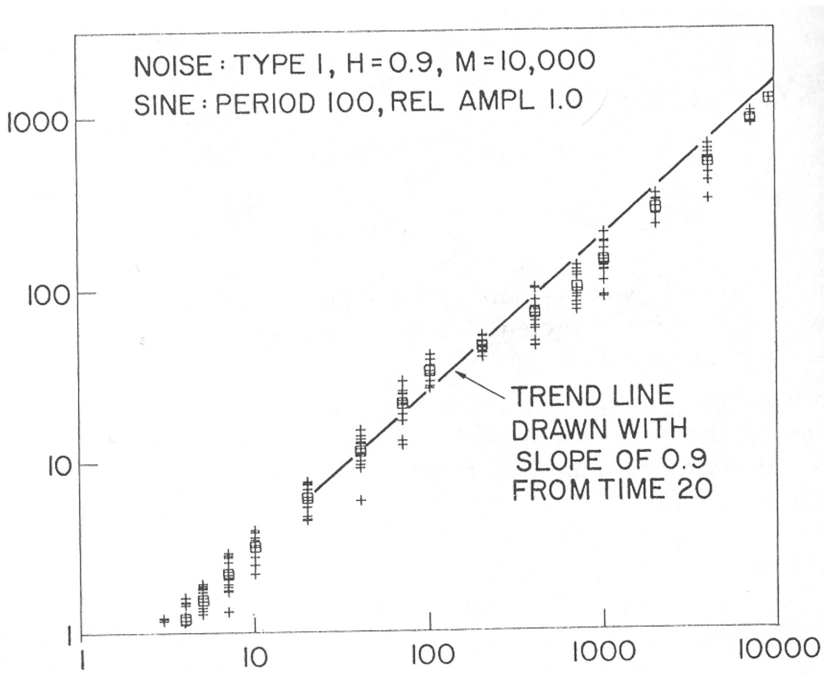

FIGURE C25-21. Pox diagram of $\log R(t, \delta)/S(t, \delta)$ versus $\log \delta$ for a sample of 9000 values of the sum of a fractional noise and a sine wave. For high values of H, such as $H = 0.9$, the presence of a comparatively strong sine component leaves the δ^H law in the mean valid. Thus, it does not greatly 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 $\delta = 200$. This means that the convergence towards the $\delta^{J+1/2}$ law in distribution is postponed to higher values of δ 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.

Definition of R/S testing and the relativity of the concepts of local and global. The above definition suggests that, having computed the values of $R(t, \delta)/S(t, \delta)$ 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 to, from the sample behavior 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 $R(t, \delta)/S(t, \delta)$. It remains to interpret R/S for finite samples of ordinary size.

Given a sample of size T such that the values of $R(t, \delta)/S(t, \delta)$ are known from $\delta = 1$ to $\delta = T$, the ideal case occurs when the variations of the sample average of $\delta^{-0.5}$ $R(t, \delta)/S(t, \delta)$ become negligible for $\delta \ll T$. Two conclusions can be drawn: (1) the value near the point where this sample average stabilizes can be taken as reasonable estimate of the limit $\lim_{\delta \to \infty} \delta^{-0.5}$ $\mathscr{E}[R(t, \delta)/S(t, \delta)]$ and (2) one can say not only that there is no global R/S dependence in $X(t)$ but also that the R/S dependence of $X(t)$ has a span much shorter than T.

However, the observed average of $\delta^{-0.5}R(t, \delta)/S(t, \delta)$ may continue to vary greatly while δ approaches its upper bound T. This behavior has two possible causes (1) $X(t)$ is R/S independent, with the transient zone of $R(t, \delta)/S(t, \delta)$ longer than T, or (2) $X(t)$ is R/S dependent. From a sample of finite duration T, one cannot conceivably distinguish between these two possibilities.

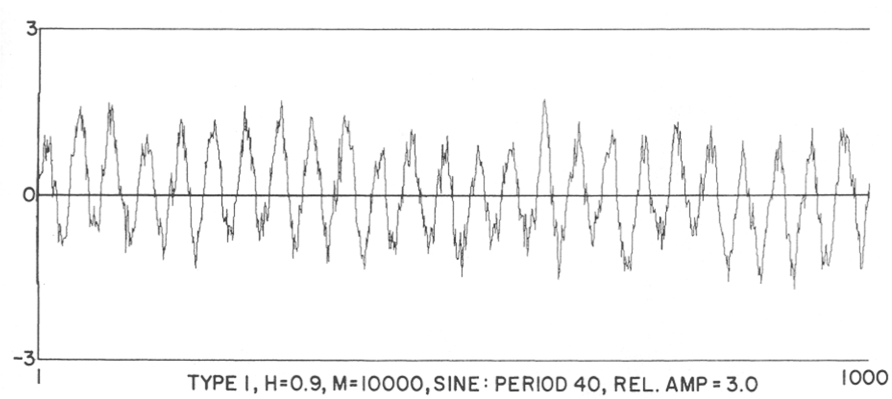

FIGURE C25-22. The first 1000 values from a 9000 value sample of the sum of a fractional noise and a sine wave of very large relative amplitude, comparable to that of meteorological records.

In summary, given a sample of duration T, R/S testing consists of deciding which is more likely between 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 in the order of magnitude of T or greater, or even infinite.

Relation between R/S dependence and other forms of global dependence. The idea of forming the ratio R/S first arose in hydrology, $R(t, \delta)$ being related to Rippl's ideal minimum capacity of reservoirs for global 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. Moreover, the examples we studied show that the concept of R/S independence quantifies some aspects of the intuitive idea of global statistical independence. We may add that it is unlikely that any single definition of global independence will ever adequately address all aspects of this concept, and that alternative definitions will always exist. For instance,

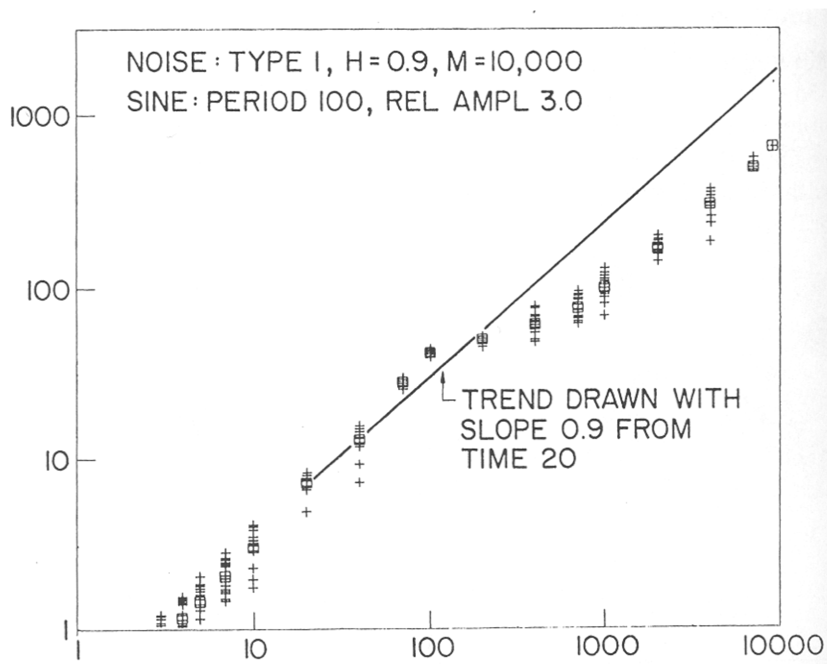

FIGURE C25-23. Pox diagram of $\log[R(t, \delta)/S(t, \delta)]$ for a sample of 9000 values that includes and continues the function in Figure 22. The effect of the sine wave is very strong. The critical bend starting at $\delta = 100$ is shown in detail on Figure 24.

random processes may be R/S dependent but global independent according to other criteria (see M 1969e), but this is not an appropriate place to discuss this feature. {P.S. 1999. Time vindicated the caution exerted in the late 1960s when drafting the preceding comments. In particular, many FSP processes are R/S independent yet globally dependent as seen in Chapter H5.}

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.

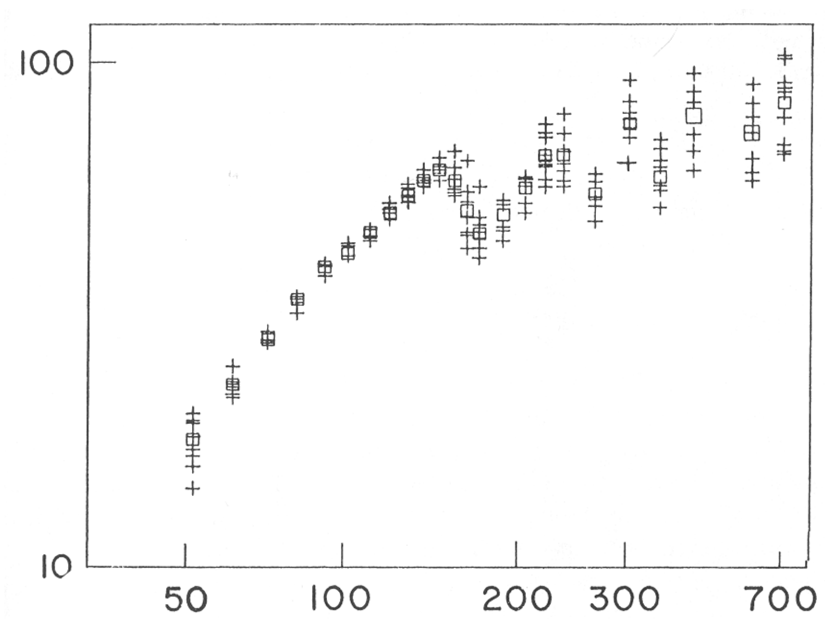

FIGURE C25-24. On this greatly enlarged detail, the single bend observed in Figure 23 splits into a richer structure of narrowing at the values of δ corresponding to the subharmonics of the sine wave, with broad lobes between these narrow points. Both features reflect the properties Figure 18 found for the pure sine wave. As $\delta \rightarrow \infty$, the lobes become negligible and the contribution of the noise again becomes determinant. However, unless the total available sample size is much larger than the wavelength of the pure sine, the addition of a strong sine wave makes the apparent R/S intensity decrease greatly.

The sine function creates global 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 type of global 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 prescribed, and the amplitude A and the phase ϕ are both chosen randomly in advance according to any specified probability distribution. For this process, the covariance between $X(t)$ and $X(t+\delta)$ is itself a sine function that oscillates up and down without limit. Now consider the ratio R/S of the pure sine wave. Clearly, $\lim_{\delta\to\infty} R(t,\delta) = AL/\pi$ and $\lim_{\delta\to\infty} S(t,\delta) = A/2$ so that (see Figure 18) $\lim_{\delta\to\infty} [R(t,\delta)/S(t,\delta)] = 2\ L/\pi = .636\ L$. Division by $\delta^{0.5}$ yields $\lim_{\delta\to\infty} \delta^{-0.5}\, \mathscr{E}[R(t,\delta)/S(t,\delta)] = 0$, leading to the fact that pure sine waves are R/S dependent. In other words, R/S dependence is not in conflict with pure sine dependence. When a white Gaussian noise of zero mean and unit variance $G(t)$ 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 $X_\Sigma(t)$ satisfies the double inequality

$$G_\Sigma(t) - AL/2\pi \le X_\Sigma(t) \le G_\Sigma(t) + AL/2\pi.$$

Since for $t \to \infty$, $AL/2\pi$ becomes negligible in relative value, the ranges of the two processes $X(t)$ and $G(t)$ are asymptotically identical and $\lim_{\delta\to\infty} \delta^{-0.5}\, \mathscr{E}[R(t,\delta)] = 1.25$. On the other hand, $\mathscr{E}S^2(t,\delta) = 1 + A/2$. Consequently, $\lim_{\delta\to\infty} \delta^{-0.5}\mathscr{E}[R(t,\delta)/S(t,\delta)] = 1.25(1+A/2)^{-0.5}$ for the process $X(t)$. That is, in the case of a sine wave plus a white noise of arbitrary amplitude, the $\delta^{0.5}$ law in the mean is valid and there is no R/S dependence.

However, the values of $\lim_{\delta\to\infty} \delta^{-0.5}\, \mathscr{E}[R(t,\delta)/S(t,\delta)]$, and the speed with which this limit is attained, are highly dependent on A. Two things happen as A increases: (1) the limit of $\delta^{-0.5}\mathscr{E}(R/S)$ tends to zero, and (2) $\delta^{-0.5}\mathscr{E}(R/S)$ takes even longer to reach its limit. For example, in the case where $L = 100$ we found that the point where the asymptotic $\delta^{0.5}$ 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 different lengths but differ greatly in high subharmonics. As a result, a number of cycles covering a whole band of frequencies will perturb R/S analysis less than a single sharp sine of comparable total energy.

Statistical robustness of the mean variance $\delta^{0.5}$ law. The relative deviation of R/S is defined as

$$\frac{\sqrt{\operatorname{Var}}\,[R(t,\delta)/S(t,\delta)]}{\mathscr{E}[R(t,\delta)/S(t,\delta)]}.$$

For the stationary process of independent Gaussian variables this relative deviation tends to $\lim_{\delta\to\infty}\sqrt{\operatorname{Var}}\,[R(t,\delta)]/\mathscr{E}[R(t,\delta)]$ as $\delta\to\infty$. Feller 1951 showed this limit to be $\sqrt{\pi}/3-1\sim 0.217$, 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 global dependence.

The term "mean variance $\delta^{0.5}$ law" conveniently combines two statements: (1) the limit $\lim_{\delta\to\infty}\delta^{-0.5}\mathscr{E}[R(t,\delta)/S(t,\delta)]$ is nontrivial and (2) the limit of the relative deviation $\sqrt{\operatorname{Var}}\,[R(t,\delta)/S(t,\delta)]/\mathscr{E}[R(t,\delta)/S(t,\delta)]$ is small.

The basis of the R/S tests for noncyclic global independence can 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 $R(t,\delta)/S(t,\delta)$ are robust. Before we tackle this issue, it may be beneficial to remind the reader of the definitions of the classical terms, "statistics" 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 the one-dimensional or multidimensional functions of the T arguments $X(t)$. The best known one-dimensional statistics are as follows: the sample moment for a given k, namely, $T^{-1}\sum_{t=1}^{T}X^k(t)$; the sample covariance for given lag δ, namely, either $T^{-1}\sum_{t=1}^{T-1}X(t)\;X(t+\delta)$ or $(T-\delta)^{-1}\sum_{t=1}^{T-1}X(t)\;X(t+\delta)$; the sample lag correlation between $X(t)$ and $X(t+\delta)$ for a given lag δ; and the Fourier coefficient of $X(t)$ at a given wave number k, for example,

$$\frac{1}{T}\sum_{i=1}^{T} X(t)\ \sin\,(2\pi kt/T).$$

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

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

Definition of the term "statistical 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 each of these aspects can be studied with respect to many different kinds of deviation from the independent Gaussian.

Nonrobustness of the precise value of the limit $\boldsymbol{\lim_{\delta\to\infty}\delta^{-0.5}\,\mathscr{E}[R(t,\delta)/S(t,\delta)]}$**.** As we already noted, Feller 1951 has proved that, for the process of independent Gaussian random variables, the value of this limit is approximately equal to 1.25. The same limit is also attained for every process that has a finite variance. However, when the variance is infinite, the limit is different and typically between 1.25 and 1. In addition, the value of this limit can be modified arbitrarily by introducing local statistical dependence so that the property $\lim_{\delta\to\infty}\delta^{-0.5}\,\mathscr{E}[R(t,\delta)/S(t,\delta)]=1.25$ is not robust with respect to local deviations of $X(t)$ from the independent Gaussian process.

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

$$\lim_{\delta\to\infty}\mathscr{E}[\delta^{-0.5}R_0(t,\delta)/S_0(t,\delta)]=\sqrt{2}\lim_{\delta\to\infty}\mathscr{E}\left[\delta^{-0.5}R(t,\delta)/S(t,\delta)\right]=1.25\sqrt{2}\,.$$

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

Extreme robustness of the mean variance $\delta^{0.5}$ ***law.*** As we have stated previously, for every process of independent values we have examined, including extremely skew log normal processes (see Figure 3) and processes with an infinite population variance (see Figure 5), $\mathscr{E}[R/S]$ is asymptotically proportional to $\delta^{0.5}$ and the reduced variable $\delta^{-0.5}\,R/S$ has a 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 by saying that the mean variance $\delta^{0.5}$ law is extremely robust with respect to changes in the marginal distribution of $X(t)$.

Nonrobustness of the statistic $R(t,\delta)$ ***not divided by*** $S(t,\delta)$***.*** None of the many variants of R/S that we studied is as robust as $R(t,\delta)/S(t,\delta)$. While some alternatives to the R/S ratio retain the property that their expected value is asymptotically proportional to $\delta^{0.5}$, none has as small a variance as R/S. In the present paper we shall be content to demonstrate the nonrobustness of $R(t,\delta)$ 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 in Figure 5. In this case, the marginal distribution is extremely skew and/or long-tailed. M 1963b and Moran 1964 show that $\mathscr{E}R(t,\delta) \sim \delta^{1/\alpha}$ 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 $\delta^{-1/\alpha}R(t,\delta)$ is infinite, which implies that sample values of $\delta^{-1/\alpha}R(t,\delta)$ are extremely erratically behaved, making it easy for sampling fluctuation to overwhelm and to hide the functional dependence of $R(t,\delta)$ on δ. Consequently, one may conjecture that, given the highly non-Gaussian characters of some of his records, Hurst's rough graphic analysis had been carried on $R(t,\delta)$ itself. 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 verify how $R(t,\delta)$ depends on δ, and on the other hand, sophisticated analysis is not ordinarily attempted unless there is evidence that it is worthwhile to do so.

Had Hurst plotted R instead of R/S, it is possible that ways to handle global hydrologic effects would have been discovered much later.

Second class of examples. Now consider the behavior of $R(t,\delta)$ for the process of independent log normal values (see Figure 4). The corresponding behavior of R/S was reported in Figure 3. This example shows that random processes exist for which $\mathscr{E}R \sim \delta^{0.5}$ holds asymptotically, but the asymptotic behavior prevails only for extraordinarily large δ. In the long transient that precedes this asymptote, the dispersion of R around $\mathscr{E}R$ 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 δ, the range of such a process was found by Moran to satisfy $\mathscr{E}R \sim \delta$, but it is clear that the scatter of sample values around this expectation is enormous. Therefore, the relation $\mathscr{E}R \sim \delta$ has no practical relevance.

Robustness of the R/S $\sim \delta^{0.5}$ law with respect to local 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 $\lim_{\delta \to \infty} \delta^{-0.5}$ $\mathscr{E}[R(t, \delta)/S(t, \delta)]$ is always positive and finite, greatly though it is affected by the details of the process. To eliminate this influence, one may consider the reduced random variable $[R(t, \delta)/S(t, \delta)]/\mathscr{E}[R(t, \delta)/S(t, \delta)]$. It can be shown that the limit for $\delta \to \infty$ is unaffected by the details of local dependence. Thus, if an asymptotic viewpoint were legitimate, one could describe the $R/S \sim \delta^{0.5}$ law as robust with respect to the introduction of local 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 $R(t, \delta)/S(t, \delta)$ as $\delta \to \infty$ can serve to define the concept of the R/S intensity of dependence, which is a form of the intensity of noncyclic global statistical dependence. For this purpose, one must divide the class of processes with global dependence more finely so that each subclass contains processes for which noncyclic, global dependence can be said to have the same intensity. With this finer subdivision we shall be able to proceed from the previously discussed problem of testing from global dependence to the new problem of precisely estimating the R/S intensity of a record.

Definitions. We shall say that a random process satisfies the $R/S \sim \delta^{J+1/2}$ law in the mean if $\lim_{\delta \to \infty} \delta^{-H} \mathscr{E}[R(t, \delta)/S(t, \delta)]$ 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 $R/S \sim \delta^{J+1/2}$ 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 $X_{\Sigma}(t)$ from blowing up too fast. Perfect compensation occurs in the pure sine wave, for which we may say that $H = 0$.

Remark. We could also exhibit processes that do not satisfy in the mean any $\delta^{J+1/2}$ law with $0 < H < 1$. Such processes, when taken as a body, constitute an additional class of equivalence, namely, a remainder class of processes 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 the robustness of the $\delta^{0.5}$ law, because it will be useful to restate the robustness in an alternative fashion. Observe 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 $G(t)$. For example, if $X(t)$ is log normal, one simply writes $X(t) = c \exp[bG(t)]$, where c and b are arbitrary constants. Thus, the robustness of the mean variance $\delta^{-0.5}$ law discussed in the section on R/S testing can be rephrased by saying that this law is invariant with respect to a nonlinear transformation of the white Gaussian noise. When discussing R/S testing, we also saw that the class of processes which do not exhibit R/S dependence is left invariant by transformations that introduce short-term dependence. We shall now show that robustness under transformation is less when global dependence is either positive or negative than when it is zero.

To describe the transformations considered, take as a point of departure the fractional Gaussian noises of exponent H (M & Van Ness 1968{H11}) and two approximations to fractional Gaussian noises (M & Wallis 1969a{H12}.) 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:

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

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

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

The final parameter M, called the memory of the process, is some very large quantity. Early on, M 1965h, set $M = \infty$, but M & Wallis 1969a varied M from 1 to 20,000.

The definitions of discrete fractional Gaussian noise, as well as of Type 1 approximations, are more cumbersome. It suffices to recall that every variant considered in M & Wallis 1969a{H12} is a linear function of independent Gaussian variables $G(u)$. According to the definition in M 1965h, 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 plotted in Figure 8.

The words *linear* and *Gaussian* are crucial in answering the following question: after various transformations have been applied to a fractional Gaussian noise of exponent H, does the $R/S \sim \delta^{J+1/2}$ law continue to hold? We consider in detail two kinds of transformations:

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

(B) Nonlinear transformations 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, pick a value of $H = J + 0.5$ and a function $F_2(t \mid H, \infty)$, 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 intensity 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 (see 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 assumptions that $H=0.5$ and δ is large. If more than one pure sine wave is added and $H \neq 0.5$, 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 serve as an elaboration of earlier discussions.

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

Next, consider the function

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

where the $G(t)$'s are independent Gaussian variables with zero mean and unit variance and where L is large in comparison with the duration of the transient range before the asymptotic $R/S \sim \delta^{0.5}$ (see Figure 19). When the lag δ lies between the duration of the transient and T, the sine wave $A\sin(2\pi t/L)$ is nearly a constant. Adding this constant to $G(t)$ leaves $R(t,\delta)$ and $S(t,\delta)$ practically unaffected and leaves $\delta^{-0.5}\mathscr{E}[R(t,\delta)/S(t,\delta)]$ near Feller's asymptotic value of 1.25. Eventually, $\delta^{-0.5}\mathscr{E}[R(t,\delta)/S(t,\delta)]$ attains its asymptotic value, derived earlier in this 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 $R(t,\delta)/S(t,\delta)$ of a pure sine wave. For δ near L and also (but less markedly) for δ multiples of L, the scatter of $R(t,\delta)/S(t,\delta)$ is greatly reduced.

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

MATHEMATICAL DIGRESSION ON ASYMPTOTIC SELF-AFFINITY

In pursuing the study of R/S analysis, it becomes important to study the distribution of the ratio $R(t, \delta)/S(t, \delta)$. In the present paper, we have studied only its mean and variance. Simplest and most interesting are the processes where, as $\delta \to \infty$, the expression $\delta^{-H}R/S$ tends towards a nontrivial limit, that is, toward 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 extended readily to show that, in the case of an independent Gaussian process, $\delta^{-0.5}R/S$ has a nontrivial limit. This process and all others for which $\delta^{-J-1/2}R/S$ has a nontrivial limit can be said to satisfy the $R/S \sim \delta^{J+1/2}$ law in distribution, or to be asymptotically R/S self-affine. This last concept generalizes ordinary self-affinity, which is discussed in M 1967s.

MATHEMATICAL DIGRESSION ON THE SCOPE OF *R/S* ANALYSIS

As we noted when discussing the classical covariance analysis, $\mathscr{E}[X(t)X(t+\delta)]$ is independent of t, if $X(t)$ is a stationary random process. But the converse is not true; the property that $\mathscr{E}[X(t)X(t+\delta)]$ is independent of t does not require that $X(t)$ be stationary. When $\mathscr{E}[X(t)X(t+\delta)]$ is independent of t, the nonstationarity of $X(t)$ will remain unnoticed as long as the analysis does not proceed beyond the covariance. Such processes have been called covariance-stationary (or weakly stationary or second-order stationary) processes.

Similarly, nonstationary random processes may exist for which $\mathscr{E}[R(t, \delta)/S(t, \delta)]$ is independent of t. When such processes are R/S analyzed but not studied from other viewpoints, they will appear stationary; therefore, it might be useful to call them R/S stationary in the mean. Moreover, any such process may satisfy the $R/S \sim \delta^{H}$ law in the mean, the strict stationarity of $X(t)$ being unnecessary. If not only the expectation but the whole distribution of the random variable $R(t, \delta)/S(t, \delta)$ is independent of t, $X(t)$ deserves to be called strictly R/S stationary. Such processes may satisfy the $R/S \sim \delta^{J+1/2}$ 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.