Broken Line Process Derived as an Approximation to Fractional Noise

BENOIT B. MANDELBROT

Mathematical Sciences Department, IBM Thomas J. Watson Research Center Yorktown Heights, New York 10598

Abstract. The broken line process has been advanced recently as a possible replacement for fractional noise as a model representing long-run persistence in hydrology and turbulence. It is shown that in fact one can rederive the broken line process as an approximation to fractional noise and that efforts to preserve the value of $\rho''(0)$ may well be misplaced.

The 'broken line process' (BL) was first introduced by *Ditlevsen* [1969] and is being developed by *Mejia* [1971], *Rodriguez-Iturbe et al.* [1972], *Mejia et al.* [1972*a, b*], and *Garcia* [1972]. Their purpose has been to show that hydrology and turbulence theory can take account of long-run effects without using the fractional noises (FN). The FN were introduced by *Mandelbrot* [1965], explored initially by Mandelbrot and J. R. Wallis, and explored more recently by N. C. Matalas and others.

The purpose of the present note is to show that, irrespective of its actual historical motivation, the BL process can be obtained as an approximation of the fast Gaussian FN proposed in *Mandelbrot* [1971]. The possibility of such a derivation will, on the one hand, confirm that BL can indeed perform in conformity with its promoters' claims and, on the other hand, stress that BL is not an alternative to FN itself but rather an alternative to the fast FN and to any other computer-oriented approximations of FN. Also, I shall emphasize that BL is not a Gaussian process and that this characteristic may be a very serious drawback. In particular, the extensive literature on, among others, the problem of level crossings is not applicable to BL except as an approximation of uncertain status. Efforts to 'preserve $\rho''(0)$' may well be misplaced.

A PROPERTY OF UNICITY POSSESSED BY FN

The aim of a model of long-run statistical dependence in fields like hydrology is to account for the Hurst phenomenon relative to the behavior of the rescaled range R/S [e.g., *Mandelbrot and Wallis,* 1968]. When one limits oneself to Gaussian processes, one may just as well work with the nonrescaled range R. In this case, I observed in 1965 that a stationary Gaussian process whose range exhibits the Hurst behavior exactly and for all lags is unique. This statement is not a subjective judgment that one could dispute; it is related to the theorem that if a stationary process is Gaussian its covariance describes it fully. In the present case, this process is the continuous time Gaussian FN and can be obtained as the derivative $B_H'(t)$ of the fractional Brownian motion of *Mandelbrot and Van Ness* [1968].

THEORETICAL APPROXIMATIONS TO FN: DIFFERENTIABILITY

The high-frequency properties of the process $B_H'(t)$ are unmanageable; $B_H'(t)$ is not even an ordinary function. Therefore in practice $B_H'(t)$ must be replaced by approximate FN. In theoretical work the best approximations are those obtained as moving averages of $B_H'(t)$ [*Mandelbrot and Van Ness,* 1968, section 4.1]. For example, when $B_H'(t)$ is averaged once (by using a uniform weight) over the interval $(t, t+1)$, one obtains the process for which the covariance $\rho(s)$ is equal to

$$\tfrac{1}{2}[|s+1|^{2H} - 2s^{2H} + |s-1|^{2H}]$$

This process can be defined at will in either discrete or continuous time. In continuous time its sample function is (almost surely) continuous, but because $\rho''(0) = \infty$ the sample function is very irregular in its local behavior and as a result is nondifferentiable. When $B_H'(t)$ is averaged twice, or alternatively is averaged once by using a triangular weighting factor, one obtains an approximation for which $\rho''(0)$ is finite and the behavior of ρ near the origin is otherwise sufficiently regular for the standard theory of smooth Gaussian processes to apply to its sample function. For example, it has (almost surely) a continuous derivative.

The authors of BL attribute a great importance to properties of smoothness. By contrast, *Mandelbrot and Wallis* [1969, section starting on p. 262, column 2] have argued that, since practical work is always limited to processes in discrete time, smoothness is not important. I continue to hold this last opinion; nevertheless, we shall return to $\rho''(0)$ again.

PRACTICAL APPROXIMATION: FFN

The above smoothed out forms of B_H' are, unfortunately, themselves impractical. Some examples of practical approximations are the type 1 and type 2 functions of *Mandelbrot and Wallis* [1969], which are cumbersome to write and expensive to compute, as is pointed out by the unanimity of critics. Another example is fast fractional noise (FFN), proposed generator of *Mandelbrot* [1971]. A sample of T values of FFN is obtained as the sum of a number $N(T)$ of independent Markov-Gauss processes, where $N(T)$ increases with T roughly proportionately to log T. Consider one of these terms $X(t)$, defined as having the covariance $C(t) = a^2 \exp(-s/s_1)$, where a and s_1 are two constants. By definition,

$$X(t + s) = \exp(-s/s_1)X(t) + a[1 - \exp(-s/s_1)]^{1/2}G$$

where G is a reduced Gaussian term independent of $X(t)$.

PRACTICAL APPROXIMATION: PROGRESSION TOWARD BL

When the $X(t)$ above has been defined for continuous t, $X(t)$ is known to be continuous but nondifferentiable. In practice, of course, $X(t)$ is only computed when t is an integer, and the question arises whether it would be possible to evaluate $X(t)$ even more economically. One proposal is to compute $X(t)$ over some rough grid of t values looser than integers and then to interpolate. When s_1 is much larger than 1, the change between $X(t)$ and $X(t + s_1)$ is very small, and it may indeed be sensible to select an appropriate second constant s_2, to construct $X(t)$ for values of t distant by s_2, and finally to define $X_1(t)$ by interpolating $X(t)$ between the above successive values. If, for example, the interpolation is linear, the function $X_1(t)$ so obtained not only is continuous but also has a derivative at nearly all points except those belonging to the basic grid. If, moreover, the starting point t_0 (namely, the first value of t for which $X(t)$ is computed) is chosen at random (with uniform probability) between $-s_2$ and 0, then the random function $X_1(t)$ can be shown to be stationary.

Note that the approximation above has the asset that the number of random numbers to be generated is reduced considerably, roughly in the ratio of $1/s_2$.

In a further step away from the original $X(t)$, one may also change the process ruling the terms $X(t_0)$, $X(t_0 + s_2)$, $X(t_0 + 2s_2)$, $\cdots$. When $s_2/s_2 \gg 1$, those terms are nearly independent. Our final change will consist in making them strictly independent. The resulting process $X_2(t)$ is a broken line according to Ditlevsen's definition. Moreover, one can check that the relation between the constants a and s_1 for the addends of the FFN and the relation between the constants a and s_2 for the addends of the BL model are compatible. Thus BL is an approximation to FFN, as was announced.

NON-GAUSSIAN CHARACTER OF THE BL APPROXIMATION AND ITS CONSEQUENCES

The value of $X_2(t)$ taken at an instant t that is a breakpoint is a reduced Gaussian random variable. Its value at t, whose distances d and $1 - d$ from the nearest breakpoints are known, is an average of two independent reduced Gaussian random variables and therefore is itself a Gaussian random variable. Its variance is $d^2 + (1 - d)^2$, which depends on the position of t and attains a minimum at the midpoint between two breaks, where it equals ½. Consequently, the value of $X_2(t)$ at an arbitrary instant is a weighted mixture of all the above Gaussian averages and therefore is itself non-Gaussian. (For example, its kurtosis, though small, is positive and equal to 0.15.)

Though this non-Gaussian character is not extreme, its consequences must not be dismissed. For example, the authors of BL attribute great importance to $\rho''(0)$ and make sure that its values in the model and in reality are matched. Formally, such matching is possible, but its usefulness is not established. Indeed, the main point about $\rho''(0)$ is the central role that it plays in the theory of level crossings, maximums, and so on, and this role applies only to exactly Gaussian processes. Implicitly, the authors of BL work with the Gaussian process having the same covariance as BL. They imply that its formal properties also apply to BL itself as 'an approximation,' but they need not. As an example consider the matter of sample differentiability. The sample of the Gaussian approximation to BL is known to have a derivative, whereas BL itself has breakpoints where it has no derivative. For zero crossings and maximums a situation may well be worse. The existing mathematics and computer experiments about these problems only show them to be very difficult, and the main lesson that I personally draw from them is that the quality of the predictions based on the Gaussian approximation to the BL approximation is doubtful and unpredictable.

SUMMARY

On the one hand, I personally welcome BL because the fact that it has been advanced marks, despite continuing resistance elsewhere, a thickening of the ranks of those who believe low-frequency effects to be important in hydrology. Also, I tend to be prejudiced in favor of this innovation because it goes beyond an application to hydrology of techniques that have already made the textbooks in other fields. On the other hand, it should be recognized that the variety of genuinely different random processes is very limited. Economy in the generating mechanism must be paid for somewhere.

For some time now the selection of models has mobilized some of the best efforts of hydrologists and of friends of hydrology, but I submit that this phase should draw to an end.

Acknowledgment. In amplifying this note for publication, thoughts stimulated by a conversation with I. Rodriguez-Iturbe and J. C. Schaake, Jr. have been most useful indeed.

REFERENCES

Ditlevsen, O., Approximate extreme value distribution and first passage time probability in stationary and related stochastic processes, Ph.D. dissertation, Tech. Univ. of Denmark, Copenhagen, 1969. (*Extremes and First Passage Times with Applications in Civil Engineering,* Danmarks Ingeniorakademi, Copenhagen, Denmark, 1971.)

Garcia, L. E., D. R. Dawdy, and J. M. Mejia, Long memory monthly streamflow simulation by a broken line model, *Water Resour. Res., 8*(4), 1100–1105, 1972.

Mandelbrot, B., Une classe de processus stochastiques homothétiques à soi; Application à la loi climatologique de H. E. Hurst, *C. R. Acad. Sci. Paris, 260,* 3274–3276, 1965.

Mandelbrot, B., A fast fractional Gaussian noise generator, *Water Resour. Res., 7*(3), 543–553, 1971.

Mandelbrot, B., and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *SIAM Rev., 10*(4), 422–437, 1968.

Mandelbrot, B., and J. R. Wallis, Noah, Joseph, and operational hydrology, *Water Resour. Res., 4*(5), 909–918, 1968.

Mandelbrot, B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, 1, Averages and variances; 2, Rescaled ranges and spectra; 3, Mathematical appendix, *Water Resour. Res., 5*(1), 228–267, 1969.

Mejia, J. M., On the generation of multivariate sequences exhibiting the Hurst phenomenon and some flood frequency analysis, Ph.D. dissertation, Colo. State Univ., Fort Collins, Nov. 1971.

Mejia, J. M., I. Rodriguez-Iturbe, and D. R. Dawdy, Streamflow simulation, 2, The broken line process as a potential model for hydrologic simulation, *Water Resour. Res., 8*(4), 931–941, 1972*a*.

Mejia, J. M., D. R. Dawdy, and C. F. Nordin, Streamflow simulation, 3, The broken line process and operational hydrology, submitted to *Water Resour. Res.,* 1972*b*.

Rodriguez-Iturbe, I., J. M. Mejia, and D. R. Dawdy, Streamflow simulation, 1, A new look at Markovian models, fractional Gaussian noise, and crossing theory, *Water Resour. Res., 8*(4), 921–930, 1972.

(Manuscript received June 12, 1972.)