THE STATISTICAL CHARACTERIZATION OF RAIN AREAS IN TERMS OF FRACTALS

S. Lovejoy
McGill Radar Weather Observatory
Macdonald Campus of McGill University
Ste. Anne de Bellevue, PQ., Canada H9X 1CO

1. INTRODUCTION

Ever since the publication of Mandelbrot’s celebrated book “Fractals” in 1977, this new chapter of classical geometry has been invading physics. The fractal geometry of nature has been discerned in a bewildering variety of places. Mandelbrot himself has examined the fractal structure of language, stock-market price series, the organization of matter in the universe, the geometry of rivers and coastlines, hydrological time series, biological structures such as trees and bronchioli, as well as the structure of turbulence.

What unites these diverse aspects of nature is the existence of structure at almost every space or time scale. This is perhaps the most fundamental property of fractals. Since one of the most striking features of satellite cloud pictures or of radar rain maps is the wealth of spatial and temporal structure, the idea that rain and cloud areas are fractal, is intuitively appealing. The aim of the present paper is to show as simply and non-mathematically as possible, that this intuition is indeed well founded. Coincidentally, the theoretical groundwork for the introduction of fractals into meteorology has recently been laid. Chorin (1981) has proved Mandelbrot’s conjecture that the solution to the Euler equations is a fractal set.

2. WHAT IS A FRACTAL?

The best explanation of fractals is to be found in Mandelbrot’s highly readable book “Fractals”, although briefer and more intuitive accounts are Gardner (1978) and Mandelbrot (1981). Undoubtedly, the best way to understand fractals is to start by examining what they look like. Figure 1 is a reproduction of one of Mandelbrot’s (1981) beautiful illustrations: a fractal cloud. (Although meteorologists may doubt whether some real clouds are fractals, there is no doubt about this one.) To use Mandelbrot’s expression, it is a 100% geometric “fake”, drawn by a computer after specifying three parameters. The fact that only three parameters were needed to unambiguously specify this rather contorted, complicated shape suggests that the figure has a basic geometric simplicity, although it is clearly unlike anything in classical geometry. Indeed, the conceptual distance from Euclid may be judged by the fact that one way of mathematically defining the dimension (strictly speaking, the “Hausdorf-Besicovitch” dimension), is a non-integer fraction. This should be compared to another type of dimension, namely the topological dimension. The topological dimension clearly has the value 1 for a parameter.

Comparing the fractal cloud in Fig. 1 with the simplest Euclidean shapes such as a circle is illuminating. Whereas a circle is uniquely determined by one size parameter – which is not length - by a dimension (D) which controls the “disconnectedness” or degree of “contorsion”, and by a random seed which selects a particular cloud from the family of all clouds. If real clouds are fractals, one would expect them to differ only in size, and random seed, not in dimension. This is because raising or lowering D has the effect of making more or less “disconnected” clouds, and more or less jagged edges respectively. Intuitively, D ought to be determined by the physical processes which shape clouds. In particular, one might expect D to be determined by the properties of turbulence. It is therefore significant that Komolgorov’s theory of isotropic homogeneous turbulence predicts D = 4/3, 5/3 for isobars and isotherms respectively, in a two-dimensional cross-section of three dimensional turbulence. These isobars and isotherms are fractals because they satisfy Mandelbrot’s definition: a fractal set is a set which has a Hausdorff-Besicovitch dimension greater than the topological dimension.

3. REAL CLOUDS AND RAIN AREAS

Mandelbrot produced Fig. 1 by varying the dimension of the perimeter until a realistic-looking cloud was produced, choosing D = 3/2. No actual data went into it, indeed, meteorologists may feel that a lower value should have been chosen because real clouds look ”smoother” than this fractal cloud. We shall presently show that real clouds have perimeters with D ~ 4/3.

One simple way to characterize the “smoothness” of the perimeter of a two-dimensional cloud picture is to measure how much perimeter is needed to enclose a given area of cloud. If a long perimeter encloses only a small area, D is nearly two times the dimension of a perimeter which literally fills the plane. If only a small amount of perimeter is required to enclose the same area, D is nearly 1, the dimension of a continuous line. In the latter case, we would expect an area (A) - perimeter (P) relationship of the form P ~ A since this is the formula for classical smooth shapes such as circles or squares. If the perimeter is so “contorted” that it tends to fill the plane (D~2), we obtain P ~ A. This suggests
Fig. 1. This figure is reproduced from Mandelbrot (1981), and represents the surface of a 0°C isotherm in homogeneous isotropic turbulence. Although Mandelbrot used this as a fractal cloud model, the dimension of a 2-D cross-section is 5/3 and is too high to be compatible with the data. This accounts for the fact that real clouds look "smoother". Any structure apparent in this cloud is purely random.

the general relationship $P \sim (A)^D$. This intuition is proved to be correct in Mandelbrot (1977). In order to estimate $D$, we therefore measure the slope of the log $P$ vs log $A$ graph indicated in Fig. 2. Here we have combined 1 x 1 km radar data from tropical Atlantic rain areas, with GOES stationary satellite infrared cloud area data, from over the Indian Ocean. As can be seen, a fractal model with $D \approx 4/3$ is strongly suggested by this data, ranging over 6 orders of magnitude in $A$. Later figures extend this at least another order of magnitude in the direction of small areas and it is possible that deviations do not occur until viscosity becomes important on a scale of meters or less. In the direction of increasing $A$, it may be hard to extend the curve much further simply because very few clouds attain these sizes. The largest cloud examined here extended over 3000 km from the centre of Africa to south of India.

It should be remarked that the apparent absence of a bend in Fig. 2 shows that no horizontal length scale is associated with these rain and cloud area perimeters. This strongly supports the position of Pinus (1968), Vinnechenko (1970), Gage (1979) and Gillet et al. (1980), who report $k^{-5/3}$ wind spectra (below $\sim 6$ km), out to distances of up to 1500 km (Vinnechenko 1970). A transition to the expected $k^{-2}$ (2-D turbulence) behaviour would yield an apparent lowering of $D$ (an increase in the slope of Fig. 2). This is because a $k^{-2}$ field is smoother than a $k^{-5/3}$ one. No such trend is apparent in Fig. 2.

4. DEVELOPMENT

4.1 Random Fractals and Non-Benign Chance

In the previous section we hinted at a basic aspect of fractal rain and cloud geometries: the stochastic element implied by the random "seed". Traditionally, meteorologists have regarded the "small scale" as a kind of random noise superimposed on a deterministic "large scale". The fractal nature of rain challenges this distinction unless the boundary between large and small is shifted off the graph in Fig. 2 to areas greater than $10^6$ km$^2$.
The reason we feel uncomfortable with the idea of random large scale structure is that until now most random processes studied in nature have been - to use Mandelbrot's (1973) term - "benign" or Laplacian. The following discussion closely follows that in Mandelbrot (1973). To illustrate what is meant by the term "benign", as well as to distinguish it from "non-benign" or "erratic" randomness, consider the random function $Y(t)$, e.g. $Y(T) - Y(0)$, as being composed of a sum of a large number of random increments $X(t)$:

$$Y(T) - Y(0) = \sum_{t=1}^{T} X(t)$$

Benign randomness is characterized by two basic properties:

(a) $\sum_{t=1}^{T} X(t)/T$ tends to a non-random limit as $T \to \infty$, denoted $\langle X \rangle$;

(b) the classical central limit theorem holds:

$$\sum_{t=1}^{T} (X(t) - \langle X \rangle)/\sqrt{T}$$

is distributed as a gaussian as $T \to \infty$. Taking the limit as $T \to \infty$ is equivalent to ignoring the fine structure of the process and studying the large scale, which can clearly be decomposed in the way indicated into the sum of a large number of small scale contributions. If properties (a), (b) hold, then fluctuations in $Y(t)$ must tend to cancel, and we are justified in regarding the change in $Y(t)$ as being due to a noise $X(t)$ superimposed on a signal $\langle X \rangle$. It is this separation between noise and signal that allows the noise to be removed and the signal to be studied independently. In meteorology, the hypothesis of benign randomness justifies the division between large and small scale processes.

Let us examine property (b) in greater detail; it can be decomposed into four parts:

1) there exists two functions $A(T)$, $B(T)$ such that

$$\sum_{t=a}^{T} X(t)/A(t) - B(t)$$

has a limit as $T \to \infty$.

2) this limit is a gaussian.

3) $A(T) = \sqrt{T}$

4) $X(t)$ and $X(t+t_0)$ are independent for sufficiently large $t_0$.

When any of these four conditions does not hold, we speak of "non-benign" or "erratic" fluctuations. Non-benign fluctuations manifest themselves either as an extreme fluctuation far from the norm or as a series of persistent departures from the norm. In reference to the biblical stories of the flood, and of the "seven lean, seven fat years", Mandelbrot has called these the Noah and Joseph effects respectively.

4.2 The Joseph Effect in Rainfall

An example of the Joseph effect is streamflow. In this case (see Mandelbrot and Wallis (1968)), $A(T) = T^H$ where $H \sim 0.6-0.9$ (depending on the river), and the future is by no means independent of the distant past (condition (iv) above does not hold). In the case of the Nile, it is possible to speak of "wet" millennia and "dry" millennia, indicating persistent departures from the mean over enormous time scales. A more familiar example is that of isotropic homogeneous turbulence. Although temperature fluctuations are distributed as gaussians, $A(T) = T^{1/3}$ and
Fig. 4a. The probability ($Pr(\Delta r > \Delta R)$) of a random difference $\Delta r$ (negative $\Delta r$ only), exceeding a fixed $\Delta R$ for spatial increments in .25 km x 1 km averaged rain rates. These radar data are from the tropical Atlantic, and curves are shown for .25 km, .5 km, 1 km separation respectively. The parallel, uniform spacing of these curves is evidence of a lack of length scale down to .25 km. This suggests that the P-A relationship plotted in Fig. 2 could be extended to .25 km x .25 km areas. Doubling the spatial separation, multiplies the distributions by a factor $2^n$ with $H \sim 1/2$. Also shown are the best fit symmetric Stable-Levy distributions for $\alpha = .75$, along with an asymptotic $\alpha = 2$ region (see Section 6).

Fig. 4b. The same as Fig. 4a, except for a different day. The only apparent difference between 4a, 4b is the scale parameter (width), and the location of the onset of the $\alpha = 2$ regime. The actual shapes of the rain areas appeared quite different, however.

thus the departures from the mean trend to cancel more quickly than they would in the benign $T^{1/2}$ case (which would prevail if the independence condition (iv) holds). For gaussian distributions, we therefore have the classification benign if $A(T) = TH = 1/2$, "antipersistent" for $H < 1/2$, "persistent" if $H > 1/2$. We shall later see that rainfall distributions require a generalization of this notion to non-gaussian cases.

In order to examine the Joseph effect in rain, we would like to evaluate $A(T)$ from the data. The simplest way of doing this is to regard each $X(t)$ value as a small increment in the rain rate $R(t)$ (just as for $Y(t)$ previously). Then $\sum_{t=1}^{T} X(t) = R(T) - R(0)$ would measure the change in rain rate over a time $T$. By plotting the distribution of $\Delta R$ for different $T$, $A(T)$ could be evaluated. In particular, for $A(T) = TH$, we expect:

$$Pr(\Delta R > \Delta R) = \frac{Pr(h^{-H}(R(t_0+hT)-R(t_0)) > \Delta R)}{Pr(h^{-H}(R(t_0+0)-R(t_0)) > \Delta R)}$$

for all $t_0$, $T$, $h$ where $Pr$ indicates "probability", $R(t)$ the rain rate as a function of time, $\Delta R$ is a given change in $R$. The validity of this relationship and the value of the parameter $H$ can be evaluated by plotting

$$\log Pr((R(t_0+hT)-R(t_0)) > \Delta R)$$

against $\log \Delta R$ for different values of $h$. On such a plot (Fig. 3), one expects to find parallel curves for different $h$, separated by a constant distance $H \log h$. As discussed in Lovejoy et al (1981), these distributions appear to be the same for rainfall in any diverse locations: e.g.,
Montreal, Spain, and the tropical Atlantic, and
yield \( H = 2/3 \) depending somewhat on the method
of estimation. An important point to note is the
asymptotic straight line behaviour indicating
\[ \Pr(\Delta R > \Delta R) \propto \Delta R^{-\alpha} \]
with \( \alpha \approx 1.65 \) for the probability
of a random difference \( \Delta R \) exceeding a fixed \( \Delta R \).

The same plot may be made for the distribution
of \( \Delta R \) in space (Fig. 4). Here \( H \approx 1/2 \), and
the distributions appear to be hyperbolic with
\( \alpha \approx 0.75 \), with an extreme \( \alpha = 2 \) regime (see Section
6). These distributions were identical in shape
for all the cases examined in the tropical Atlan-
tic. The asymptotic regions of these plots are
quite important, and will be dealt with in Sec-
tion 4.3. For now, it suffices to note that
these distributions contain far more "extreme"
points than a gaussian, and thus the gaussian
classification persistence: \( H > 1/2 \), independence;
\( H = 1/2 \), antipersistence \( H < 1/2 \), does not hold.
In the next section, we shall see that the \( H = 1/2 \) case corresponds to antipersistence because the
increments in Fig. 4 must cancel considerably
more quickly than in the independent case to
yield such a small \( H \).

As can be seen from these figures, an accurate
estimate of \( H \) may be difficult. In fact, there
is a much better method of investigating persist-
ence known as R/S or "Re-scaled range" analysis
(Mandelbrot and Wallis (1969)). Since this
requires a fairly long uninterrupted series of \( R \)
values, it was only performed on the spatial dis-
tributions. R/S analysis also yields \( H \approx 0.50 \).

4.3 The Noah Effect in Rain

As mentioned earlier, the Noah effect occurs
when \( X(t) \) occasionally takes on such extreme
values that
\[ u = \left( \sum \Delta X(t) \right)/\Delta T = B(1) \]
cannot possibly be distributed as a gaussian.
Indeed, it is possible to show that if \( u \) has a
limiting distribution for \( T \to \infty \), that it must
belong to a four-parameter family of distributions
known as Stable-Levy distributions. This result
is a form of a generalized central limit theorem
which includes the gaussian limit as a special
case. If \( u \) has a finite variance, then the limit
is a gaussian, otherwise it is another member of
the Stable-Levy family. These distributions have
the following asymptotic form:
\[ \Pr(u > U) \approx (U/U^*)^{-\alpha} \quad \text{for} \quad \alpha < 2 \]
for the probability of a random \( u \) exceeding a
fixed \( U \). \( \alpha \) is the most important parameter since
it determines the total probability contained in
the extreme tail of the distribution; in other
words, it controls the frequency of occurrence of
"extreme" values. \( U^* \) is a parameter analogous to
the standard deviation and measures the width of
the distribution.

Before continuing, it is worth discussing in
detail the implications of the Noah effect in
meteorology. If changes in rainfall in time or in
space are hyperbolically distributed with \( \alpha < 2 \),
then occasionally such large values of \( \Delta R \) occur
that the variance (and mean, if \( \alpha < 1 \)) of \( \Delta R \) does
not converge. In other words, as we add more and
more \( \Delta R \)’s together, the variance (and mean for
\( \alpha < 1 \)) increase without limit. One immediate con-
sequence is that rainfall cannot be an RMS con-
tinuous random function; it is made of sharp dis-
continuities. Another way of looking at this is
to note that if we regard change in rain rate as
being made up of many small changes, that the
largest of these will be much larger than all the
others. This is illustrated in Fig. 5, which
shows a computer generated \( R(t) \) series with \( \alpha = 1.5 \).
The visual impression is one of a few large "jumps"
separated by noise. It is tempting to regard
these jumps as a signal, although clearly such a
distinction would have no basis. This behaviour
is known as "clustering" because the small jumps
appear to "cluster" around the large ones. In
rain fields, this effect would give rise to the
appearance of stochastic structure and organiza-
tion even if the Joseph effect (which can do the
same) was not present. Conversely, very different
looking rain areas yield similar probability dis-
tributions (see Figs. 4a,b). In this way, rain
areas differ by only two parameters (the width
and \( \alpha \), the mean) of which
at least one, and possibly both, are random.
(See Section 6).
Fig. 5. A random function $Y(t)$ produced with independent $\alpha = 1.5$ increments, shown with two different values of $Y(0)$ (at points 0, 0'). Also shown is the effect of "truncating" negative values (below axis marked "T"), a procedure required if $Y(t)$ is to model the rainrate $R$, since $R > 0$. (See Section 6).

5. RAIN AREA DISTRIBUTIONS

Up till now, we have stressed two different aspects of fractal rain areas. On the one hand, a geometry with structure at all scales, and on the other, the non-benign stochastic element which permits this structure to arise in a non-deterministic fashion, so that no detailed information specifying the shapes is required. What finds these two elements together, and allows for a class of non-random fractals, is the concept of self-similarity. A self-similar curve, such as a rain area perimeter, has the property that the structure is in some sense the same at all scales, with the exception of a scale factor. This means that it is impossible to distinguish the perimeter of a large cloud from that of a small cloud enlarged by some factor. In both cases the relative sizes of the "bumps" and "wiggles" of the perimeter are the same. This invariance is expressed mathematically by Eq. 1 for the case of random fractals. For non-random fractals, such as the fractal "snowflake" in Fig. 6, self-similarity is determined by a fixed rule which builds perimeters (or other structures) with a hierarchy of identical shapes.

Self-similarity is a powerful concept because it can unite the various parameters we have estimated, such as $D \sim 4/3$, $H \sim 1.65$, .75 and $H \sim 2/3$, 1/2 (for the temporal and spatial structures respectively), into a coherent model which could permit the construction of "mock" or computer generated rain areas. The construction of such a model would be very important because it would allow, in principle, most of the statistical properties of rain to be derived theoretically. Unfortunately, it appears (Mandelbrot, private communication), that antipersistent, infinite variance fractals have not been adequately investigated, and thus it seems that the problem of rain structure is partly a mathematical one, requiring new techniques. If the problem is solved, mock radar scans and time histories could be produced which are completely indistinguishable from real ones, and the random structure of rain would be indisputable. Skeptics should look back at Fig. 1 or at any of the illustrations in Mandelbrot (1977). Naturally, such a fractal description of rain would pose the question not only of the origins of the particular parameters, but also of the consequences of stochastic rain structure.

Fig. 6. Reproduced from Mandelbrot (1981). This geometrical, non-random "snowflake" is shown in the first three stages of construction (which proceeds an infinite number of steps before becoming a true fractal). The result is known as the self-similar Koch curve, with $D = 1.26$. 
Returning to the notion of self-similarity, Mandelbrot (1977) has shown that if a 2-D field (such as \( R(x,y) \)), has self-similar relief (i.e., Eq. 1 holds), that the distribution of fractal "islands" (e.g., rain areas) must be of the form \( \Pr(a > A) \sim \left( \frac{A}{A^*} \right)^{-B} \) for the probability of a random area exceeding \( A \), whose \( A^* \) is the width parameter. In fact, in the case of gaussian distributions of \( \Delta R \), one would expect \( B = D/2 \). However, in the present case, we have two sources of hyperbolic behaviour of \( \Pr(a > A) \): the self-similar \( R \) field, and the hyperbolic \( \Delta R \) distribution. It is not known how these should combine to determine \( B \).

Figure 7 shows the distribution of rain areas in the tropical Atlantic, clearly indicating hyperbolic behaviour with \( B = 0.82 \). Data from Spain, Florida and Montreal yield similar values of \( B \) (see Lovejoy et al (1981)). The fact that this distribution has neither finite variance nor mean indicates that enormous fluctuations in rain area are possible. Even if one argues that the width of the distribution (\( A^* \)), is determined by the traditional large scale meteorological parameters (the existence and nature of this relationship is the object of the cumulus parameterization project), fluctuations of actual rain areas (which are random samples from such a distribution), can still be very large. For example, computer simulations show that two sample rain areas taken from distributions with the same \( A^* \) (i.e., the same large scale conditions), differ by a factor 3, 40% of the time, and by a factor 10, 16% of the time (see Lovejoy et al (1981) for the details of this calculation). Because of the infinite mean property, the total area of rain covered by a large number of storms will behave in a similar way: relative cancellation of fluctuations does not occur, because the extreme rain area is always comparable in size to the sum of all the others. The existence of large fluctuations in rain area is important, especially in the tropics because in addition to latent heat and moisture transport the area of convection also controls the "friction" which balances tropical pressure gradients. Even in mid-latitudes where the dynamics are controlled by coriolis force/ pressure gradient balance, these stochastic effects may be very important. If \( A^* \) is not determined by the large scale but is itself a random variable, then the prospect of stochastic forecasting may be unavoidable.

6. A FURTHER NOTE ON "MCK" RAIN MAPS

Before concluding, a final note on "mock" rain maps and \( \Delta R \) distributions should be made. Although fractals with \( H = 1/2, \alpha = 3/4 \) cannot at present be computer generated, \( H = 4/3, \alpha = 3/4 \) fractals are quite easy to produce because \( H = 4/3 \) is the scaling one obtains for independent \( \alpha = 3/4 \) increments (see Section 4.3). Although such fractals show little resemblance to the strongly anti-persistent rain type fractal (\( H = 1/2, \alpha = 3/4 \)), they can explain the asymptotic \( \alpha = 2 \) region in Fig. 4, which we have previously ignored.

These computer generated fractals are produced in fashion similar to that shown in Fig. 5 by adding positive and negative random increments. The resultant \( R \) field will in general contain physically meaningless negative \( R \) values (those below the t axis in Fig. 5). These negative \( R \) values must be set to zero. When this procedure is performed, the extreme \( \Delta R \) are clearly more likely to be truncated than the others, and the result seems inevitably to be an \( \alpha < 2 \) asymptote at least over the range of \( H, \alpha \) values tested.

The abrupt transition point varying randomly for a fixed sequence of random increments. The \( \alpha = 2 \) region therefore has a natural and simple explanation as a truncation effect produced by the condition \( R > 0 \). As far as I know, no analytic derivation of this effect exists, although it seems to occur in many situations (see Lovejoy et al (1981)).
7. CONCLUSIONS

Convective rain and cloud area are shown to have fractal structure with dimension $D \approx 4/3$ for over more than 7 orders of magnitude in area. This suggests that convective structures up to $10^6$ km$^2$ are stochastic in origin and do not have a characteristic length scale. This view was supported by an analysis of the spatial and temporal structure of rain, which indicates that the rain field exhibits several unusual properties described under the rubric of "non-benign" or "erratic" fluctuations. Rainfall is shown to exhibit the Joseph effect, which means that rainfall at even distant points in time or space are not independent. Perhaps more significantly, it is also shown to exhibit the Noah effect, which means that rainfall varies in time and space in a much more erratic manner than in the usual models. The temporal and spatial scaling parameters ($H$) are estimated to be approximately 2/3, 1/2 respectively. The corresponding parameters characterizing the importance of fluctuations, are $\alpha \approx 1.65, 0.75$ respectively. The distribution of rain areas was shown to be hyperbolic, in agreement with the fractal model (exponent $\tau \approx -0.82$). These parameters suffice to specify the type of fractal involved. Unfortunately, there is no known method of combining these in a computer model that could generate "mock" rain area maps.

8. ACKNOWLEDGEMENTS

I would like to acknowledge Dr. B. Mandelbrot for showing a lively interest in this work, and for making many helpful suggestions. Also acknowledged is Dr. G. L. Austin, who encouraged me throughout, and Dr. K. Browning, for commenting on earlier versions of this paper.

9. REFERENCES

Vinnichenko, N.K., 1969: The kinetic energy spectrum in the free atmosphere - 1 second to 5 years. Tellus, 22, 158-166.