Reprinted from Nature, Vol. 319, No. 6048, pp.43–44, 2 January 1986

Fractal characterization of inhomogeneous geophysical measuring networks

S. Lovejoy*, D. Schertzer* & P. Ladoy†

* Physics Department, McGill University, 3600 University Street, Montreal, Quebec, Canada H3A 2T8
† EERM/CLIM, Météorologie Nationale, 2 Ave Rapp, Paris, 75007, France

The measuring stations of most *in situ* geophysical networks are spatially distributed in a highly inhomogeneous manner, being mainly concentrated on continents and population centres. When inhomogeneity occurs over a wide range of scales in a space of dimension E, it can be characterized by a fractal dimension D_m. For measuring networks, there is no reason to assume *a priori* that D_m equals E; it will usually be less than E. The world meteorological network studied here is an example on a surface for which $E = 2$ (the surface of the Earth) whereas, the network has an empirical dimension $D_m = 1.75$. Whenever $D_m < E$, any sufficiently sparsely distributed phenomena (with dimension $D_p < E - D_m$, here 0.25), cannot be detected—even if the network is infinite. Because these rare phenomena are the most intense, this insufficient dimensional resolution is associated with biases in geophysical statistics, serious difficulties in interpolating measurements to a uniform grid, and problems in calibrating remotely-sensed information.

The intuitive notion of dimension D_m of a set (whether fractal or otherwise) is given by the variation of the number of points ($\langle n(L)\rangle$) with the size of the region (L):

$$\langle n(L)\rangle \propto L^{D_m}$$

When D_m is less than the dimension (E) of the space in which the set is embedded, when L increases, the volume of space available increases as L^E which is faster than the number of points of the set that are available to fill it ($\langle n(L)\rangle$). Hence, larger and larger 'holes' (empty regions) appear and the set is concentrated on a decreasing fraction of the total space. Furthermore, at a given scale, sets with lower D_m are more dominated by holes, and are hence more sparse.

Although there are many ways of estimating the dimension of sets of points, many of which have been developed for studying strange attractors[2–5] they generally yield similar results. The method used here, actually determines a particular dimension called the correlation dimension[6]. To apply this method to a geophysical network (considered here as a set of points), determine, for each station in the network, the number ($n(L)$) of other stations within various radii L of the point, and its average $\langle n(L)\rangle$ over all the stations. We then obtain D_m as the slope of $\log\langle n(L)\rangle$ against $\log L$. Scales over which $\langle n(L)\rangle$ varies non-algebraically, define the characteristic lengths of the network.

To apply this technique to a surface network (such as the highly clustered world meteorological network, Fig. 1), we must discuss a complication that takes into account the curvature of the Earth's surface. If the latter is covered uniformly with stations in a region of area S, then $\langle n(L)\rangle \propto S$. Taking $S(\theta)$ as the area of the spherical cap defined by two points subtending an angle θ at the Earth's centre (radius r) we may define the corresponding scale $L(\theta)$ by:

$$S(\theta) = (\pi/4)L^2(\theta) = 2\pi r^2(1 - \cos\theta/2)$$

Note that with this definition of L, for small θ, the formula reduces to the usual great circle distance ($= r\theta$), and we recover $D_m = 2$ when the stations are homogeneously distributed.

Figure 2 shows $\langle n(L)\rangle$ calculated using the above L, for the 9,563 stations in this network separated by at least 0.01° of arc

Fig. 1 The locations of the 9,563 stations in the global meteorological measuring network (defined as the stations which the World Meteorological Organisation lists as performing at least one meterological measurement per day), showing their high degree of non-uniformity. Most are clustered on the continents and major industrial areas. The dimension (D_m) is ~1.75.

(~1 km, the accuracy of the data). Avoiding double counting, each station defines 9,562/2 = 4,781 values of L, hence there are 9,563 × 4,781 = 45,720,703 independent values that go into the histogram for estimating $\langle n(L)\rangle$. Figure 2 shows that roughly between the minimum resolvable scale, (~1 km) and ~2,000 km,

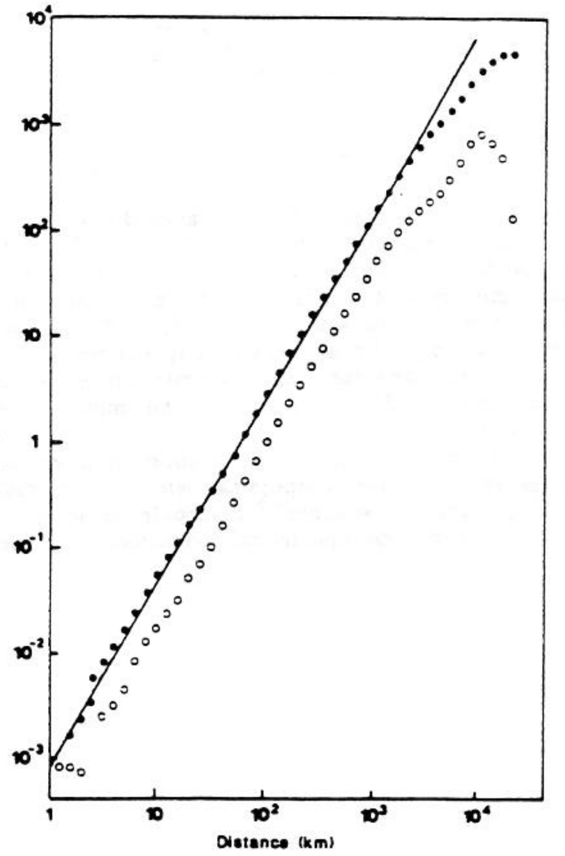


Fig. 2 ○, The average number of stations within annuli of geometrically increasing radii; ● (the integral of the previous function), the function $\langle n(L)\rangle$ described in the text. Over the scaling regime, both should be parallel and straight: the $L^{1.75}$ function is shown for reference. The scaling regime apparently continues down to scales comparable with the accuracy of the geographical locations used to determine L(~1 km) which is, therefore, the spatial resolution of the network. For comparison, the standard analysis method (which assumes $D_m = 2$, a surface area of ~10^8 km², and 10^4 stations), attributes an average area of $10^8/10^4 = 10^4$ km² per station hence a spatial resolution of ~10^2 km² which is about 100 times larger than its true resolution.

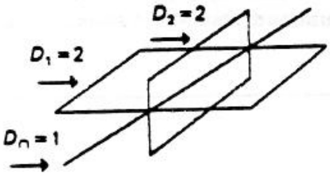
$\langle n(L)\rangle \propto L^{1.75}$, hence $D_m \sim 1.75$. The size of the scaling regime is limited as the number of points is finite: $\langle n(L)\rangle_{max} = 4{,}781$ and $\langle n(L)\rangle_{min} \sim 1/9{,}563 \sim 10^{-4}$. Assuming $L^{1.75}$ between these limits implies $L_{max} \sim 7{,}500$ km, $L_{min} \sim 0.3$ km, which shows that the scaling observed in Fig. 2 spans a range nearly as large as is possible. Similar analysis of the French climatological network (3,593) stations yielded $D_m \sim 1.8$, while the Canadian meteorological network (414 stations) yielded $D_m \sim 1.5$.

The fact that $D_m < E$, sets new detectability limits. A network, dimension D_m can detect a phenomenon with dimension D_p only if the two sets intersect. However, a theorem in geometry shows that this is certain to occur only when $D_p > E - D_m$. Two sets, dimension $= D_1$, D_2 embedded in a space dimension $= E$, intersect on a set dimension $= D_\cap$. (With the co-dimension $C = E - D$) according to the following rule:

$$C_\cap = \mathrm{Inf}((C_1 + C_2), E)$$

Example. Intersection in space ($E = 3$) of two planes

$$(D_1 = D_2 = 2 \Rightarrow C_1 = C_2 = 1)$$



$$C_\cap = C_1 + C_2 = 2 \Rightarrow D_\cap = E - C_1 = 1$$

Although the example is for two standard sets (planes) it holds for most fractal sets. Note that if $C_m + C_p > E$, the intersection set probably has dimension zero ($C_\cap = E$), hence, the sets typically miss each other. Hence, in the case studied here, sparse surface phenomena with $D_p < 2 - 1.75 = 0.25$, cannot be detected. In analogy with the network's spatial resolution, which is the minimum detectable scale, we may define, its dimensional resolution as $E - D_m$ which is the minimum resolvable dimension.

The inability to detect sparse phenomena is serious because, in general, we expect geophysical fields to be characterized by multiple fractal dimensions[6–9]. In turbulence, fields are generally characterized by multiple fractal dimensions[10–13]. In geophysics, there is empirical evidence for multidimensionality in the rainfield[14,15]. If a field is multidimensional, then regions exceeding a threshold T define fractal sets with dimension $D_p(T)$ decreasing with T. Thus, whenever averages are taken over sets with $D_m < E$, then, by the intersection theorem, the most intense fractals (with $D_p(T) < E - D_m$) are missed. While both mono- and multidimensional fields have statistical properties that are functions of scale, multidimensional fields are in addition dependent on the dimension (for example, line, plane, volume or fractal set) over which they are averaged.

We have argued that the characterization of network inhomogeneity by the fractal dimension (D_m) raises new problems concerning the detectability of sparse phenomena. Since phenomena (of any size) with $D_p < E - D_m$ will not intersect the network, we have obtained a new criterion for evaluating measuring networks: to detect phenomena, not only must a network have sufficient spatial resolution it must also have sufficient dimensional resolution. In general, we expect the fractal dimension of geophysical fields to be a decreasing function of intensity: thus, any lack of dimensional resolution leads to biases in the spatial averages. Finally, as information on the lowest dimensional fractals is lost, similar biases will arise when measurements are interpolated to uniform grids with dimension E.

We thank B. Luneau for help in preparing Fig. 1. We acknowledge the CNRS (France), ATP (RA) for partial financial support.

Received 2 June; accepted 21 October 1985.

1. Mandelbrot, B. *The Fractal Geometry of Nature* (Freeman, San Francisco).
2. Roux, J. C., Simoy, R. H. & Swinney, H. L. *Physica* 8D, 257–266 (1983).
3. Grassberger, P. & Procaccia, I. *Phys. Rev. Lett.* 50, 346–349 (1983).
4. Grassberger, P. & Procaccia, I. *Physica* 9D, 189–208 (1983).
5. Nicolis, C. & Nicolis, G. *Nature* 311, 529–532 (1984).
6. Hentschel, H. G. E. & Procaccia, I. *Physica*, 8D, 435–444 (1983).
7. Grassberger, P. *Phys. Lett.* 97, 227–230 (1983).
8. Schertzer, D. & Lovejoy, S. *Symp. on Turbulence and Chaotic Phenomena in Fluids*, 141–144 (*IUTAM*, Kyoto, Japan, 1983).
9. Mandelbrot, B. *J. Stat. Phys.* 34, 895–930 (1984).
10. Parisi, G. & Frisch, U. *Turbulence and predictability in geophysical fluid dynamics and climate dynamics* (eds Ghil, M. *et al.*) 84–88 (International School of Physics, Enrico Fermi, Varenna, Italy, 1985).
11. Shertzer, D. & Lovejoy, S. *Turbulent Shear Flow*, Vol. 4 (ed. Launder, B.) 7–33 (Springer, New York, 1985).
12. Schertzer, D. & Lovejoy, S. *PCH J.* (in the press).
13. Lovejoy, S., Schertzer, D. *AMS Bull.* (in the press).
14. Lovejoy, S. & Schertzer, D. *Digital Image Processing in Remote Sensing*, Ch. 14 (ed. Muller, P. J.) (Taylor and Francis, London, 1985).
15. Tsonis, A. A. Preprint (Univ. Wisconsin).