APPENDIX B

FRACTAL MEASUREMENT METHODS

TruSoft-international (2004), the company produced Benoit 1.3, provided the following explanation for all five measurement methods including those discussed in Chapter Three (section 3.3.5) to show the mathematical principle behind the software:

B.1 Box Dimension Estimation Method interface

The box dimension is defined as the exponent Db in the relationship:

$$N(d) \approx \frac{1}{d^{Db}} \qquad \text{(equation 1)}$$

N(d) is the number of boxes of linear size d necessary to cover a data set of points distributed in a two-dimensional plane. The basis of this method is that, for objects that are Euclidean, equation (1) defines their dimension. One needs a number of boxes proportional to 1/d to cover a set of points lying on a smooth line, proportional to 1/d^2 to cover a set of points evenly distributed on a plane, and so on.

This dimension is sometime called grid dimension because for mathematical convenience the boxes are usually part of a grid. One could define a box dimension where boxes are placed at any position and orientation, to minimize the number of boxes needed to cover the set. It is obviously a very difficult computational problem to find among all the possible ways to cover the set with boxes of size d the configuration that minimizes N(d). Also, if the overestimation of N(d) in a grid dimension is not a function of scale (i.e., we overestimate N(d) by, say, 5% at all box sizes d), which is a plausible conjecture if the set is self-similar, then using boxes in a grid or minimizing N(d) by letting the boxes take any position is bound to give the same result. This is because a power law such as (1) is such that the exponent does not vary if we multiply N(d) or d by any constant.

In practice, to measure Db one counts the number of boxes of linear size d necessary to cover the set for a range of values of d; and plot the logarithm of N(d) on the vertical axis versus the logarithm of d on the horizontal axis. If the set is indeed fractal, this plot will follow a straight line with a negative slope that equals -Db. To obtain points that are evenly spaced in log-log space, it is best to choose box sizes d that follow a geometric progression (e.g. d = 1, 2, 4, 8,...), rather than use an arithmetic progression (e.g. d = 1, 2, 3, 4,...).

A choice to be made in this procedure is the range of values of d. Trivial results are expected for very small and very large values of d. A conservative choice may be to use as the smallest d ten times the smallest distance between points in the set, and as the largest d the maximum distance between points in the set divided by ten. Alternatively, one may exceed these limits and discard the extremes of the log-log plot where the slope tends to zero.

In theory, for each box size, the grid should be overlaid in such a way that the minimum number of boxes is occupied. This is accomplished in Benoit by rotating the grid for each box size through 90 degrees and plotting the minimum value of N(d). Benoit permits the user to select the angular increments of rotation.

B.2 Perimeter-Area Dimension Estimation method interface

Consider an object that is a closed loop in the two-dimensional plane, e.g., an island. Suppose that this island is a Euclidean object, i.e., a circle. Then the area A and the perimeter P of such an island are related as follows:

$$p = 2\pi r = r\sqrt{\pi A} \approx \sqrt{A}$$
$$A = \pi r^2 = \frac{p^2}{4\pi} \approx p^2 \qquad \text{(equation 2)}$$

"r" is the radius of the circle; note that the proportionality between A and P does not depend on r. If the island had a fractal perimeter, then the relationships (2) become

$$p \approx (\sqrt{a})^{Dp} = A^{Dp/2}$$
$$A \approx p^{2/Dp} \qquad \text{(equation 3)}$$

Dp is the perimeter-area dimension. Indeed, if Dp = 1, one obtains the Euclidean case, as in (2); if Dp = 2, then the figure is space-filling because P A. If Dp is between 1 and 2, equation (3) shows that the perimeter of the fractal figure is longer than the perimeter of a Euclidean figure with the same area, as expected.

In practice, to estimate Dp one measures perimeter P and area A with boxes of different side length d, and plots the logarithm of A on the vertical axis versus the logarithm of P on the horizontal axis. If the relationship is indeed fractal, this plot will follow a straight line with a positive slope that equals 2/Dp. Note that the estimation of perimeters and areas has to be done over a range of d.

B.3 Information Dimension Estimation Method

This fractal dimension is often encountered in the physics literature, and is generally different from the box dimension. In the definition of box dimension, a box is counted as occupied and enters the calculation of N(d) regardless of whether it contains one point or a relatively large number of points. The information dimension effectively assign weights to the boxes in such a way that

boxes containing a greater number of points count more than boxes with less number of points.
The information entropy I(d) for a set of N(d) boxes of linear size d is defined as

$$I(d) = -\sum_{i=1}^{N(d)} m_i \log(m_i) \qquad \text{(equation 4)}$$

where mi is:

$$m_i = \frac{M_i}{M} \qquad \text{(equation5)}$$

Mi is the number of points in the i-th box and M is the total number of points in the set.

Consider a set of points evenly distributed on the two-dimensional plane. In this case, we will have:

$$N(d) \approx \frac{1}{d^2} \qquad m_i \approx d^2 \qquad \text{(equation 6)}$$

so that (4) can be written as:

$$I(d) \approx -N(d)[d^2 \log(d^2)] \approx -\frac{1}{d^2}[2d^2 \log(d)] = -2\log(d)$$

(equation 7)

For a set of points composing a smooth line, we would find: $I(d) \approx -\log(d)$

Therefore, we can define the information dimension Di as in:

$$I(d) = -D_i \log(d) \qquad \text{(equation 8)}$$

In practice, to measure Di one covers the set with boxes of linear size d keeping track of the mass mi in each box, and calculates the information entropy I(d) from the summation in (4). If the set is fractal, a plot of I(d) versus the logarithm of d will follow a straight line with a negative slope equal to -Di.

At the beginning of this section, we noted that the information dimension differs from the box dimension in that it weighs more heavily boxes containing more points. To see this, let us write the number of occupied boxes N(d) and the information entropy I(d), in terms of the masses mi contained in each box:

$$N(d) = \sum_i m_i^0$$
$$I(d) = -\sum_i m_i \log(m_i) \qquad \text{(equation 9)}$$

The first expression in (9) is a somewhat elaborate way to write N(d), but it shows that each box counts for one, if mi > 0. The second expression is taken directly from the definition of the information entropy (4). The number of occupied boxes, N(d), and the information entropy I(d) enter on different ways into the calculation of the respective dimensions, it is clear from (9) that:

$$Db \leq Di \qquad \text{(equation 10)}$$

The condition of equality between the dimensions (10) is realized only if the data set is uniformly distributed on a plane.

B.4 Mass Dimension Estimation Method

Draw a circle of radius r on a data set of points distributed in a two-dimensional plane, and count the number of points in the set that are inside the circle as M(r). If there are M points in the whole set, one can define the "mass" m(r) in the circle of radius r as:

$$m(r) = \frac{M(r)}{M} \qquad \text{(equation 11)}$$

Consider a set of points lying on a smooth line, or uniformly distributed on a plane. In these two cases, the mass within the circle of radius r will be proportional to r and r^2 respectively. One can then define the mass dimension Dm as the exponent in the following relationship:

$$m(r) \approx r^{D_M} \qquad \text{(equation 12)}$$

In practice, one can measure the mass m(r) in circles of increasing radius starting from the centre of the set and plot the logarithm of m(r) versus the logarithm of r. If the set is fractal, the plot will follow a straight line with a positive slope equal to Dm. As the radius increases beyond the point in the set farthest from the centre of the circle, m(r) will remain constant and the dimension will trivially be zero. This approach is best suited to objects that follow some radial symmetry, such as diffusion-limited aggregates. In the case of points in the plane, it may be best to calculate m(r) as the average mass in a number of circles of radius r.

It can be shown that the mass dimension of a set equals the box dimension. This is true globally, i.e., for the whole set; locally, i.e., in portions of the set, the two dimensions may differ. Let us cover the set with N(d) boxes of size d, and let us define the mass, or probability, in the i-th box mi as:

$$m_i = \frac{M_i}{M} \qquad \text{(equation 13)}$$

Mi is the number of points in the i-th box and M is the total number of points in the set. We can now write the average mass, or probability, in boxes of size d as m(d), the average mi in the N(d) boxes:

$$m(d) = \frac{1}{N(d)} \sum_{i=1}^{N(d)} m_i = \frac{1}{N(d)} \qquad \text{(equation 14)}$$

(the sum of all the masses mi is obviously one). As the operation of calculating the mass contained in a box of size d is the same as calculating the mass in a circle of radius r, we can write our definition of mass dimension (12) in terms of d rather than r:

$$m(d) \approx d^{D_M} \qquad \text{(equation 15)}$$

By using (4) and re-arranging terms, we obtain:

$$N(d) \approx \frac{1}{d^{D_M}} \qquad \text{(equation 16)}$$

This is the definition of the box dimension; thus, the mass dimension equals the box dimension.

B.5 Ruler Dimension

Consider the problem of estimating the fractal dimension of a jagged, self-similar line, the typical example being a coastline. Define N(d) as the number of steps taken by walking a divider (ruler) of length d on the line, the ruler dimension Dr is defined as:

$$N(d) \approx d^{-Dr} \qquad \text{(equation 17)}$$

The basis of this method is as follows: if the line is Euclidean, Dr = 1, then the length of the line will be a constant independent of d. Note that this is bound to be true for values of d sufficiently small. For example, the perimeter of a circle measured by a ruler of length d will be constant when d is much less than the radius of the circle. At the other extreme, if the line completely fills space, Dr = 2, i.e., the length of the line is linearly related to the length of the ruler. This can be shown to be true by equating the measured length of the line N(d) with the number of boxes needed to cover the line N(d) times d: When Dr = 2, the number of filled boxes is proportional to 1/d^2, and the line fills the two-dimensional space. One can show the formal equivalence of the ruler and box dimension.

In practice, to obtain Dr one counts the number of steps N(d) taken by walking a divider (ruler) of length d on the line, and plot the logarithm of N(d) versus the logarithm of d. If the line is indeed fractal, this plot will follow a straight line with a negative slope that equals -Dr. It should be noted that in general, a ruler of length d will not cover exactly the line, but we will be left with a remainder. Benoit keeps this remainder and therefore has non-integer values of N(d).