FRACTAL ANALYSIS OF YIELD MAPS

R. Ehsani, D. Karimi, K. H. Lee

ABSTRACT. *Agricultural crop yield data usually are highly nonlinear and complex. Basic mathematical and statistical techniques are sometimes insufficient to describe the nature, trend, or cause of variations in yield. This article investigates fractal analysis of agricultural yield maps and describes a method of applying fractal analysis to multiple years of yield data. It also shows how patterns of yield variation can be described by fractal geometry. Crop yield was measured and mapped for an agricultural field for five consecutive years. In order to obtain a sufficiently dense set of points necessary for valid fractal analysis, a method was proposed to transfer the data points from* $\Re^3$ *to* $\Re^2$. *Analysis of the resulting data set revealed multifractality of the yield variations. It was shown that multifractal measures such as the Rényi spectrum can be used to quantify and compare global and local yield variations.*

Keywords. *Fractal analysis, Multifractals, Precision agriculture, Spatial variability, Yield map.*

Analyzing and interpreting yield data is one of the main steps in successful application of precision agriculture. Assessment and interpretation of yield data can be very difficult because sometimes there are not enough data on the soil or plant factors that cause yield variability. Most of the common attempts to characterize in-field spatial yield variability have used basic statistical measures or geostatistical methods (Eghball et al., 1999; Blackmore et al., 2003). Because in-field variations are highly non-linear and complex, these methods cannot fully describe the variation patterns.

Maps of crop yield variability have been studied more than any other type of variability map because crop yield maps show the final outcome of management decisions and indicate overall field variability better than other variability maps. Spatial variability in crop yield is the result of interactions among many factors, including topography of the terrain, soil physical and chemical properties, and soil water content. These factors generate variations at nested scales that may result in self-similar patterns of variation (Green and Erskine, 2004). If so, these variation patterns can be described by fractal geometry (Peitgen et al., 2004) and quantified by fractal dimensions or fractal dimension spectra. Moreover, using methods such as joint multifractal analysis, the variability of crop yield can be related to other spatial variables of interest (Meneveau et al., 1990).Knowledge of scale-dependent variation can also be useful when generalizing analysis results or management decisions from a smaller to a larger scale or vice versa (Bekele et al., 2005).

A few studies have used fractal analysis to characterize temporal and spatial variability in an agricultural field. Perfect and Blevins (1997) used fractal analysis of soil structure to compare different tillage practices. Evaluation of the effect of variable-rate fertilizer application strategies has also been completed using fractal analysis of soil nitrate (Eghball et al., 1999). Fractal analysis of temporal and spatial variability in yield data has shown that temporal (year-to-year) variability in crop yield might be more significant than any spatial variability (Eghball and Varvel, 1997). Multifractal and joint multifractal analysis methods have been used to investigate the variability of crop yield with terrain slope (Kravchenko et al., 2000). Furthermore, fractal dimensions have been used to quantify the variability in crop yield and near-surface soil water (Green and Erskine, 2004) and to identify the scales of variation of soil electrical conductivity (Bekele et al., 2005). Fekete (2001) computed fractal dimension of crop yield by a time series analysis of yield monitor output and the threshing cylinder torque.

The objectives of this study were to introduce an appropriate methodology for fractal analysis of agricultural yield maps and to demonstrate how the developed method can be used to quantify and compare global and local variations in crop yield.

Submitted for review in December 2007 as manuscript number IET 7300; approved for publication by Information & Electrical Technologies Division of ASABE in August 2008.

The authors are **Reza Ehsani, ASABE Member,** Assistant Professor, Department of Agricultural and Biological Engineering, University of Florida, IFAS, Citrus Research and Education Center, Lake Alfred, Florida; **Davood Karimi,** Graduate Student, Department of Biosystems Engineering, University of Manitoba, Winnipeg, Manitoba, Canada; and **Kyeong-Hwan Lee, ASABE Member,** Postdoctoral Fellow, Department of Agricultural and Biological Engineering, University of Florida, IFAS, Lake Alfred, Florida. **Corresponding author:** Reza Ehsani, Department of Agricultural and Biological Engineering, University of Florida, IFAS, Citrus Research and Education Center, 700 Experiment Station Road, Lake Alfred, FL 33850; phone: 863-956-1151, ext 1228; fax: 863-956-4631; e-mail: ehsani@ufl.edu.

FRACTAL GEOMETRY

SELF-SIMILARITY AND FRACTALS

The word "fractal" was coined by Benoit B. Mandelbrot in 1975 (Mandelbrot, 1977) and refers to a rough or fragmented geometric shape or a quantity that displays self-similarity on several scales. In other words, a fractal is a geometric shape or a pattern of variation that can be divided into parts, each of which is, at least approximately or statistically, a scaled-down copy of the whole. Mandelbrot (1983) formally defines a fractal as "a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological di-

  ISSN 0001-2351

mension." Therefore, although self-similarity is the fundamental concept behind fractals, not every self-similar object is a fractal. Moreover, as was implied before, exact self-similarity is only seen in mathematical fractals such as the Sierpinski gasket and the Minkowski curve (fig. 1). Fractal patterns that are found in natural objects or phenomena show quasi-self-similarity or statistical self-similarity; in other words, they have numerical or statistical qualities that scale in a power law fashion with the scale of magnification. Simple mathematical fractals are monofractals, that is, they must be scaled by a constant number to achieve self-similarity. Most natural fractals, however, show self-similarity under arbitrary scaling (Chen, 1997). These objects are a mixture of many monofractals and are referred to as multifractals.

Fractal Dimensions

In the computation of the Hausdorff fractal dimension for the Minkowski curve shown in figure 1, the "vel" refers to the volumetric elements used. The fractal is covered by a set of volumetric elements of size r_k. In this study, we have adopted square vels, but any arbitrary shape works as well. The number (N_k) of the vels required to completely cover the object is counted. Then, the size of the vel is reduced, usually in a dyadic form, i.e., $r_{k+1} = r_k/2$, and the new number (N_{k+1}) of vels required to cover the fractal is counted. For mathematical fractals, this step can be iterated indefinitely. But for natural fractals, the power law relationships will saturate at some point because no further detail is seen after a finite number of magnifications (Vicsek, 1992). N_k is assumed to have a power law relationship with r_k:

$$N_k = \left(\frac{1}{r_k}\right)^{D_H} \tag{1}$$

where D_H is called the Hausdorff dimension of the object under consideration (Mandelbrot, 1977). The regression line on a plot of $\log(N_k)$ versus $\log(1/r_k)$ for several vel coverings is the simplest way to find D_H. In other words, in the limit:

$$D_H = \lim_{k\to\infty} \frac{\log(N_k)}{\log(1/r_k)} \tag{2}$$

For a non-fractal curve, such as a line or a differentiable curve, D_H equals one. For a fractal plane curve, however, it will be greater than one and less than or equal to two. For the Minkowski curve shown in figure 1, it equals 1.50.

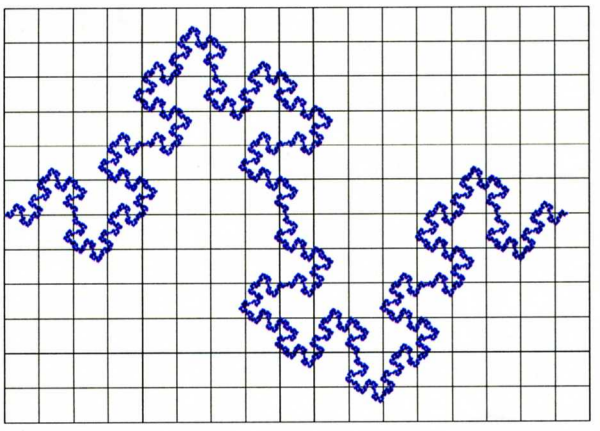

Figure 1. Minkowski curve: An example of mathematical fractals.

The Hausdorff dimension belongs to morphology-based fractal dimensions. The value of these fractal dimensions depends only on the morphology of the fractal. More precisely, morphological fractal dimensions quantify the degree of irregularity of the fractal: the higher the degree of irregularity of the fractal, the larger the computed dimension. There are situations in which morphology-based fractal dimensions do not provide a sufficient description of the fractal pattern. For example, in fractal analysis of a strange attractor of a dynamical system, one is interested in time change, or evolution, of the fractal. In fractal analysis of a turbulent flow, one is interested in the distribution of a variable such as fluid pressure or velocity over a spatial fractal. Therefore, a series of entropy-, variance-, and transform-based dimensions have been defined to deal with these applications (Kinsner, 2005). One important variation of entropy-based fractal dimensions is the information dimension D_I (Peitgen et al., 2004):

$$D_I = \lim_{k\to\infty} \frac{H_{1k}}{\log(1/r_k)} \tag{3}$$

In this equation, H_1 is the Shannon entropy given by:

$$H_{1k} = -\sum_{j=1}^{N_k} p_{jk} \log p_{jk} \tag{4}$$

where p_j is the probability of the fractal entering the jth vel (for strange attractors) or the value of the measure of interest in the jth vel. The value of p_j has to be normalized before being put in equation 4, so that:

$$\sum_{j=1}^{N_k} p_{jk} = 1 \tag{5}$$

The Shannon entropy (H_{1k}) is in bits and quantifies information in a piece of data. Here, it is equivalent to the amount of information necessary to specify a point of the fractal with an accuracy of r_k (Peitgen et al., 2004).

Another important variation of entropy-based dimensions is the correlation dimension (D_c), defined as (Peitgen et al., 2004):

$$D_C = \lim_{k\to\infty} \frac{-\log \sum_{j=1}^{N_k} p_{jk}^{\,2}}{\log(1/r_k)} \tag{6}$$

The numerator in equation 6 is known as the correlation sum and can be interpreted as a correlation between pairs of adjacent points on the fractal (Kinsner, 2005). It is important to mention that equations 2, 3, and 6 assume that a power law relationship exists between the terms that appear in the numerator and the vel size; otherwise, the object or phenomenon under study is not a fractal.

Characterization by a single number, or dimension, is adequate only for those objects that show exact self-similarity (Chen, 1997). In fact, for such objects or phenomena, all of the fractal dimensions result in the same value. As mentioned before, natural fractals are only statistically self-similar and require more than one number to be characterized completely (Kinsner, 2005). The Rényi spectrum of dimension is based on the Rényi generalized entropy (Rényi, 1961), which is a

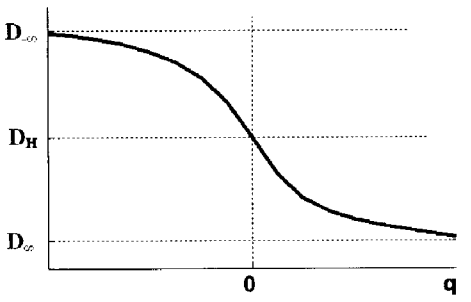

Figure 2. Typical shape of the Rényi spectrum of dimensions for a multifractal object.

generalization of the Shannon entropy (eq. 4) and can be used to characterize the distribution of multifractals. The Rényi spectrum of dimensions is defined as (Peitgen et al., 2004):

$$D_q = \lim_{k\to\infty} \frac{1}{q-1} \frac{\log \sum_{j=1}^{N_r} p_{jk}^{\,q}}{\log(1/r_k)} \tag{7}$$

where q is the moment order, and D_q is the dimension of the multifractal object for a given q. Theoretically, the spectrum can be computed for the whole range of $-\infty < q < \infty$, but in practice, due to limited accuracy in numerical calculations, the spectrum is computed on a limited range, usually $-10 \le q \le 10$. Therefore, instead of a single dimension, D_q provides a continuous range, or spectrum, of dimensions. It is easy to show that for $q = 0$, 1, and 2, the Rényi dimension reduces to Hausdorff, information, and correlation dimensions, respectively (Peitgen et al., 2004). In addition, some other single-valued fractal dimensions correspond to D_q for some value of the moment q; therefore, the Rényi spectrum can be thought of as the "unifying framework" of all single dimensions (Kinsner, 2005). For monofractals, the value of D_q is constant for all values of q, while for multifractals, it is a monotonically decreasing function of q (fig. 2). Therefore, the Rényi spectrum separates monofractals, for which the spectrum is a horizontal line, from multifractals, for which it is monotonically decreasing (Halsey et al., 1986). This curve can be interpreted as a "fingerprint" of the fractal. A very important feature of this signature is that it has limited bounds; that is, D_∞ and $D_{-\infty}$ are finite.

Perhaps the best way to understand the fractal dimension of a yield variation pattern is to visualize this pattern as a 3D surface. The Hausdorff fractal dimension shows the degree of unevenness of this surface. For a smooth, that is differentiable, surface, $D_H = 2$, while for an extremely rough surface, D_H will increase towards 3. The Rényi spectrum of dimensions can be thought of as a generalization of this concept. Using equation 7, the value of the Rényi dimension (D_q) for large values of q is mostly determined by those vels that have a high value of the variable under study, whereas for small values of q, D_q is determined mostly by those vels that have a low value of the variable. In other words, for small values of q, D_q indicates roughness of yield variation at low-yield areas, while for large values of q, it shows the same thing mainly for those spots with high yield. This insight is unique to fractal analysis and cannot be obtained using other methods.

METHODS

The study site was a 24 ha field in Ohio, and the period of the study was from 1997 to 2001. The site was under wheat cultivation in 1997 and 1999, under soybean cultivation in 1998 and 2001, and under corn in 2000. A 432 × 288 m rectangle-shaped central portion of the field was chosen for analysis. This choice was made based on the limitations in terms of the size and shape of the field and the requirement that the size of vel coverings should form a sequence of dyadic numbers.

A grain combine harvester was used for harvesting the corn, wheat, and soybean. During harvesting, the yield data were obtained by a yield monitoring system (AgLeader, Ames, Iowa). The yield monitoring systems were calibrated each season according to the manufacturer's instructions. The yield monitor software program was able to export the yield data in AgLeader advanced format (Shearer et al., 1999), which is an ASCII format. The spatial resolution of the yield data was determined by the combine head width and the yield log interval. As shown in table 1, the combine head and log interval were 5.9 to 7.6 m and 1.4 to 1.8 m, respectively, so that the spatial resolution was between 10.0 and 15.7 m^2. The number of data points in the chosen field of 432 × 288 m was between 7,900 and 12,500. The original data were interpolated to obtain the value of yield on a uniform grid of 2.25 m. Since the map size was 432 × 288 m, this provided 192 and 128 points in each direction, so that a total of 24,576 data points for the entire map was made. The triangle-based linear interpolation functionality in Matlab (ver. 7.1, The Math Works, Inc., Natick, Mass.) was used for interpolation. Table 1 also shows the mean and coefficient of variation (CV) of the yield data for each of the five years. Although the header width of the combines was larger than 2.25 m, this interpolation did not introduce an error because, as described later, the smallest vel size in this analysis was 4.5 m.

A dithering technique was used to convert the original data points into a set of points in $\Re^2$. As mentioned before, the yield was determined by interpolation on a grid of 192 × 128 points. These data points can be considered as graph of yield versus location in $\Re^3$. In the proposed approach, the value of yield at each interpolated point was first mapped to an integer in the range of 1 to 32. This range was chosen based on the reported accuracy for typical grain yield monitors

Table 1. Summary of the yield data.

Crop and Year	Combine Head Width (m)	Yield Log Interval (m)	Resolution of the Data (m^2)	Number of Data Points	Mean Yield (Mg ha^{-1})	CV (%)
Wheat in 1997	7.6	1.6	12.2	10,100	4.21	0.133
Soybean in 1998	8.5	1.7	14.5	8,600	3.62	0.144
Wheat in 1999	8.6	1.4	12.0	10,300	4.46	0.130
Corn in 2000	5.9	1.7	10.0	12,500	10.90	0.139
Soybean in 2001	8.7	1.8	15.7	7,900	3.60	0.133