Actuarial Applications of Multifractal Modeling
Part I: Introduction and Spatial Applications

John A. Major, ASA, MAAA and
Yakov Lantsman, Ph.D.

Actuarial Applications of Multifractal Modeling

Part I: Introduction and Spatial Applications

by John A. Major, ASA, MAAA and Yakov Lantsman, Ph.D.[1]

email:jmajor@guycarp.com, lantsman@netrisk.com

Abstract

Multifractals are mathematical generalizations of fractals, objects displaying "fractional dimension," "scale invariance," and "self-similarity." Many natural phenomena, including some of considerable interest to the casualty actuary (meteorological conditions, population distribution, financial time series), have been found to be well-represented by (random) multifractals. In this part I paper, we define and characterize multifractals and show how to fit and simulate multifractal models in the context of two-dimensional fields. In addition, we summarize original research we have published elsewhere concerning the multifractal distribution of insured property values, and discuss how we have used those findings in particular and multifractal modeling in general in a severe storm catastrophe model.

Introduction

In this section, we introduce the concepts of fractals and multifractals.

Fractals

Mathematicians have known of sets whose dimension is not a whole number for some time, but the term "fractal" emerged on the scientific and popular scenes with the work of Benoit Mandlebrot in the 1960s and 1970s [Mandlebrot 1982].

Mathematically, a fractal can be defined as a point set with possibly non-integer dimension. Examples of fractals include continuous random walks (Weiner processes), the Cantor set, and the Sierpinski triangle (the latter two discussed below). Phenomena in nature that resemble fractals include dust spills and coastlines.

Regular fractals possess the attribute of *self-similarity*. This means that parts of the set are similar (in the geometrical sense of equivalence under a transformation consisting of magnification, rotation, translation, and reflection) to the whole. This gives regular fractals an "infinite regress" look, as the same large-scale geometrical features are repeated at ever smaller and smaller scales. Self-similarity is also known as *scale*

[1] The authors would like to thank John Mangano for his contributions to this paper, Shaun Lovejoy and Daniel Schertzer for their helpful conversations, and Gary Venter for his review of an early draft. Errors, of course, are solely the responsibility of the authors.

symmetry or *scaling* – the fractal doesn't have a characteristic scale at which its features occur; they occur at all scales equally.

Irregular fractals do not possess strict self-similarity, but possess *statistical* self-similarity and scaling. This will be clarified below.

The key numerical index of a fractal, its *fractal dimension*, deserves further explanation. It is not immediately obvious how the concept of dimension from linear algebra, the maximum number of linearly independent vectors in a space, can be generalized to include the possibility of noninteger values. While there are several ways of doing so – and they often coincide – the so-called capacity dimension (sometimes misnamed the Hausdorff dimension[2]) is perhaps the easiest to understand.

Consider a closed and bounded subset S of N-dimensional Euclidean space R^N. We define a covering of S of size λ to be a set of hypercubes $\{H_i\}$ such that (1) each hypercube is of size λ on a side and (2) the set S is contained within the union of all hypercubes $\cup H_i$. For any λ, let $n(\lambda)$ be the minimum number of hypercubes needed to be a covering of S. The dimension of S can then be defined in terms of the scaling behavior of coverings of S, i.e., the behavior of $n(\lambda)$ as $\lambda \to 0$.

Examples:

- If S consists of a finite number of points, then for all λ less than the minimum distance between points, a covering needs to have as many hypercubes as there are points: n is constant for small λ.
- If S consists of a line segment of length L, then $n(\lambda)=L/\lambda$: n varies as the reciprocal of the first power of scale λ.
- If S consists of a (sub-) hypercube of dimension m and length L on a side, then $n(\lambda)$ is approximately $(L/\lambda)^m$: n varies as the reciprocal of the mth power of scale λ.

This exponential relation, $n(\lambda) \propto \lambda^{-m}$, motivates the definition of fractal dimension:

$$d = -\lim_{\lambda \to 0}\left(\frac{\log(n(\lambda))}{\log(\lambda)}\right) \qquad (1)$$

The previous examples show that by this definition, a set of isolated points has dimension zero, a line segment has dimension one, and an m-hypercube has dimension m, as we would expect.[3]

Subsets of the unit interval may have various dimensions less than or equal to one, and cardinality is no guarantee of dimension for infinite sets. Finite point sets have

---

[2] The definition of Hausdorff dimension is more technically complicated, involving an infimum rather than a limit, thereby handling cases where the limit (in equation 1 below) does not exist.

[3] Note that the dimension N of the embedding space is irrelevant. While it is true that a line segment of finite length can be made to fit in a hypercube of arbitrarily small side if the dimension of the hypercube is big enough, what really matters is the scaling behavior. That is, if the side of the hypercube is halved, then two of them are needed to cover the line segment – implying the line segment has dimension one.

dimension zero, of course, but there are countable subsets with dimension zero and those with dimension one. For example, the set of rational numbers (a countable set) is a dense subset of the real numbers, meaning that any open set around a real number contains a rational. Therefore, the fractal dimension of the rationals is the same as that of the reals (they need exactly the same covering sets), that is, one.

On the other hand, the countable set consisting of points $x_k = \alpha^k$, k=1,2,3,... where $0<\alpha<1$, has dimension 0. This can be seen by considering coverings by blocks of length $\lambda = \alpha^j$ for some arbitrary j. The first block covers all points x_j, x_{j+1}, and at most (j-1) blocks are needed to cover the other (j-1) points. Thus,

$$\log(n(\lambda))/\log(\lambda) \leq \log(j)/(j*\log(\alpha)) \rightarrow 0.$$

Nothing in the definition of fractal dimension precludes the possibility of a set S having a noninteger dimension d. We now present some examples to show how this can happen.

The *Cantor set* is a subset of a line segment and is defined recursively as follows. Start with the entire line segment. Remove the middle third, leaving two disconnected closed line segments. Repeat the process on each remaining line segment, ad infinitum. In the limit, we have the Cantor set. At stage k of the construction (the whole segment being stage 0), we have 2^k subsegments each of length 3^{-k}, for a total length of $(2/3)^k$. In the limit, the Cantor set has measure[4] zero (it consists of points with no net length) because in the limit, $(2/3)^k$ goes to zero. For any length $\lambda=3^{-k}$, we need 2^k segments H_i to cover the set. Therefore the fractal dimension of the Cantor set is log(2)/log(3) = 0.63093..., corresponding to something between a line and a set of isolated points.

The self-similarity of the Cantor set follows directly from its construction. Each sub-segment is treated in precisely the same way (up to a scale factor) as the original segment.

As an example of a noninteger fractal dimension in a 2-dimensional space, consider the *Sierpinski triangle* (also known as the Sierpinski *gasket*). This subset of the unit square is defined recursively as follows: Start with an equilateral triangle and its interior. Draw an inscribed triangle (point down) connecting the midpoints of each side. This divides the triangle into four similar and congruent sub-triangles. Remove the interior of the inscribed triangle. Repeat the process on each of the remaining three sub-triangles. Figure 1 shows an approximation to the result. As with the Cantor set, the Sierpinski triangle has zero measure (no area), because each stage of the construction takes up $(3/4)^k$ area of the outer triangle. Assuming the original triangle is inscribed in a unit square, at stage k of the construction, we need 3^k squares H_i of side $\lambda=2^{-k}$ to cover the set. Therefore, the Sierpinski triangle has fractal dimension log(3)/log(2) = 1.584963..., corresponding to something between a linear and a planar figure.

The self-similarity of the Sierpinski triangle again follows directly from its construction. Each sub-triangle is a miniature version of the original triangle and is similar to all other triangles appearing in the set.

---

[4] Measure theory is reviewed in the next section.

The analysis of fractal dimension by this method is generally termed *box-counting*. There are other approaches, but they will not be discussed here. Note that the method applies to arbitrary sets, not just self-similar ones. A non-self-similar set is called an *irregular* fractal if it has a noninteger fractal dimension.

Among natural phenomena, coastlines are frequently cited as good examples of irregular fractals. The measured length of a coastline depends on the scale of accuracy of the measuring tool. Comparing maps at various scales, one can see progressive deterioration of detail as larger scales are used. What appears as a wrinkled inlet on one map is abstracted to a simple polygon on the next and then obliterated completely on the next. [Barnsley] gives the fractal dimension of the coast of Great Britain as approximately 1.2. [Woo] discusses numerous areas where fractal laws relate to natural hazard processes.

This notion of scale-dependent measurements will play a central role in the practical application of fractal and multifractal theory to real-world problems.

Multifractals

Multifractals, also known as *fractal measures*, generalize the notion of fractals. Mandlebrot also worked on multifractals in the 1970s and 1980s [Mandlebrot 1988], but the first use of the term is credited to U. Frisch and G. Parisi [Mandlebrot 1989]. Rather than being point sets, multifractals are measures (distributions) exhibiting a spectrum of fractal dimensions.

A brief review of measure theory is in order. A measure μ on a space X is a function from a set of subsets of X (a σ-algebra of "measurable sets") to the real numbers R. In order to be a measure, the function μ must satisfy $\mu(\varnothing)=0$, $\mu(S)\geq 0$, and μ of any countable collection of disjoint sets must equal the sum of μ on each set. Actuaries typically encounter only *probability* measures, where, in addition, $\mu(X)=1$. The usual measure on R^N is Lebesgue measure $\nu(S)$, characterized by the fact that if S is a rectangular solid with sides of lengths λ_i, $i=1,\ldots,N$, then $\nu(S)=\Pi_i\lambda_i$.

If a measure μ on R^N is zero on every set for which ν is zero (i.e., it is absolutely continuous), then the ratio of measures $\mu(H)/\nu(H)$ where H is a neighborhood (with non-zero measure) around a point x is well-defined, and in the limit, as the neighborhood shrinks to measure zero, the ratio f(x), if it exists, is the *density* of μ, also known as the Radon-Nikodym derivative. Not all measures have densities; think of a probability function with a point mass at zero. As H shrinks around the point mass, $\mu(H)$ cannot become less than the point mass, but $\nu(H)$ goes to zero; the density becomes infinite.

Multifractals, as measures, tend to be extremely ill-behaved, not characterizable in terms of densities and point, line, plane, etc., masses.

The simplest way to create a multifractal is by a *multiplicative cascade*. Consider the "binomial multifractal," constructed on a half-open unit interval (0,1] with uniform density as follows: Divide the interval into two halves (open on the left) of equal length. Distribute $0<p<1$ of the mass uniformly on the left half and 1-p of the mass uniformly on the right half (here p is a constant throughout all stages of the construction). Repeat on each subinterval. Figure 2 shows several stages of construction with p=1/3. The horizontal axes show the unit interval and the vertical axes show density. The upper left

panel shows stage 1, where 1/3 of the mass is on the left half and 2/3 is on the right. Note that the average density is 1. The upper right panel shows stage 2, where the left and right halves have each been divided. The 2nd and 3rd quarters of the interval have the same density because they have masses of (1/3)*(2/3) and (2/3)*(1/3), respectively. The lower left panel shows stage 4 where the interval has been divided into $2^4 = 16$ subsegments. Some local maxima seem to be appearing, they are labeled. The lower right panel shows stage 7, and begins to give a sense of what the ultimate multifractal looks like. Note the similarity of left and right halves.

As you can see, at the local maxima, the density "blows up" as the scale resolution gets finer. Note how the maximum density increases from panel to panel. However, the rate of divergence is different at different points. The set of locations with particular (different) rates of divergence turn out to be fractals (with different fractal dimensions). Thus we have layers of fractals representing different "orders of singularities," with a relationship between the rate of divergence and the fractal dimension. See Appendix A for mathematical details.

This relationship is known as the *spectrum of singularities* - no single fractal dimension suffices to characterize the fractal measure, hence the name *multifractal.*

Having a spectrum of singularities means that the multifractal measure consists of infinitely spiky peaks sprinkled throughout predominant valleys, but that with proper mathematical technology, the peaks can be classified by the rate at which they diverge to infinity, and comparable peaks can be collected together into fractal "mountain ranges."

Figure 3 shows a real-world density field that approximates a multifractal. It is the population density of the northeastern USA. The big spike in the middle is New York City. Lesser spikes pick out other densely-populated cities.

In their analysis of turbulent meteorological phenomena, [Schertzer & Lovejoy] write the functional relationship between a chosen scale of resolution λ and the average densities φ_λ measured at that scale as:

$$\Pr\left\{\varphi_\lambda > \lambda^\gamma\right\} \propto \lambda^{-c(\gamma)} \qquad (2)$$

This is very much in the spirit of box-counting for fractals, except the equivalent formulation for fractals would have (1) the event inside Pr{ } being the probability of finding *any* point of the fractal in a λ-neighborhood, instead of points that satisfy a certain degree of singularity, and (2) the exponent on the right hand side being a constant, *the* fractal dimension of the set, instead of a function. In this formulation, the function $c(\gamma)$ carries all the information necessary to characterize, in a statistical sense, the multifractal.[5]

---

[5] It is tempting to read this equation as a statement about the probability of encountering a point with exponent γ or higher or the probability of fractal dimension. However, if the fractal dimension of points having exponent γ or higher is less than the dimension of the embedding space, then such points make up a set of (Lebesgue or probability) measure zero. In the typical multifractal, "almost" all the mass is concentrated in "almost" none of the region. The equation is really a statement about the *scaling relationship* between intensity and probability.

Compare Figure 3 with Figures 4 and 5. The former measures population density at the resolution of 8 miles. The latter two measure it at resolutions of 16 and 32 miles, respectively. Clearly, one's impression of this density field is largely driven by the scale of resolution used. A systematic investigation of the appearance of a field using various scales of resolution is at the heart of multifractal analysis.

A box-counting approach developed by [Lavallée et. al.] known as *Probability Distribution Multiple Scaling* (PDMS) can be used to estimate the probabilities of singularities with assorted rates of divergence. (See also [Lovejoy & Schertzer 1991].) It turns out that directly estimating $c(\gamma)$ in such a fashion is not a productive approach to analyzing real data sets for multifractality due to the severe demands that the procedure places on the sample data. In the next section, we will show how multifractals can be understood equally well through the behavior of their moments.

[Pecknold et. al.] give many examples of (apparent) multifractals in nature. See also [Ladoy et. al.] These include rain and cloud fields (measured from scales of a thousand kilometers and years down to millimeters and seconds – see [Lovejoy & Schertzer 1991]), human population density (as above, also see further discussion below), and foreign exchange rates. Part of the impetus for the development and practical application of multifractal analysis came from "the burgeoning mass of remotely sensed satellite and radar data" [Tessier et. al., 1993]. Depending on the scale of resolution used, measurements of cloud cover could be made to vary drastically; moreover, how this variation with scale behaved was also dependent on the level of intensity chosen as a threshold – just the sort of fractals-within-fractals behavior to be expected from multifractal fields.

Spatial Fields

In this section, we delve into the general theory of self-similar random fields, focusing on the two-dimensional case. (The extension to three or more dimensions is straightforward.) Examples are taken from our applications in property-liability insurance.

Analysis of Multifractal Fields

Analysis of random multifractals is an extension of the analysis of random fields. Recall that a random field $\varphi(\mathbf{r})$ is a collection of real-valued random variables φ, indexed by $\mathbf{r}$, where $\mathbf{r}$ may be

- an integer, for example, in the case where the random field is a (discrete) time series,
- a real number, for example, in the case where the random field is a (continuous) stochastic process,
- a vector in D-dimensional Euclidean space $\mathbf{R}^D$, in the case of a general random field.

Typically, we focus on $D = 1$ for financial/econometric time series and $D = 2$ for spatial distributions in geography or meteorology.

To analyze a random multifractal, we must first respect the fact that it is a measure, and strictly speaking may not (typically does not) possess real-valued densities. Therefore,

we cannot treat a random multifractal as a random field $\varphi(r)$.[6] However, as we have seen in previous sections, *when viewed at a finite scale of resolution L*, a multifractal does have a well-behaved density that we can treat as a random field $\varphi_L(r)$. Thus, the approach to studying random multifractals is to consider sequences of random fields that describe the density of the measure at various scales of resolution, and to study the scaling behavior of those sequences.

Appendix B outlines the mathematics. The box-counting approach appears to admit straightforward application (and becomes PDMS) as discussed above. For various reasons discussed below, it is more fruitful to deal with *moments* of the random fields. The key object of the analysis is the so-called $K(q)$ function, describing the scaling behavior of the q^{th}-moments of the sequence of random density fields as the scale of resolution λ varies. At finer resolutions, the density fields appear more "spiky" and average q-powers of the fields for q>1 (q<1) get arbitrarily large (small) according to the power law:

$$E(\varphi_\lambda^q) = (\lambda)^{-K(q)}. \tag{3}$$

The boundary conditions $K(0) = K(1) = 0$ further constrain the $K(q)$ curve.

Synthesis of Multifractals; 2-D Multiplicative Cascade

Above, we described how recursive application of multiplication of densities – a multiplicative cascade – generated the simple binomial multifractal on a line segment. A similar operation, in two dimensions, can be used to generate a multifractal akin to the Sierpinksi triangle. Consider the following two-by-two matrix:

$$a = \begin{bmatrix} 2.0 & 1.4 \\ 0.6 & 0 \end{bmatrix} \tag{4}$$

Take a unit square with uniform density. Divide it into four quadrants and multiply the density in each quadrant by the corresponding element of a. Note that the average of the four elements of a is 1.0, so the average density across the entire square is unchanged. Repeat the procedure on each quadrant, recursively. In the limit, we have a multifractal. At stage k, neighborhoods of the upper left corner have average density 2^k. That point has the highest degree of singularity.[7] The lower left corner has a different sort of singularity, with density 0.6^k approaching zero as the scale shrinks. The entire lower right half is empty (density zero). Like the Sierpinski triangle, in fact, the square is almost everywhere empty: at each stage, the area with nonzero density is $(3/4)^k$ which approaches zero as k increases without bound. Figure 6 depicts the result.

---

[6] It might be tempting to consider a random measure as a collection of random variables indexed by *subsets* of the underlying $\mathbf{R}^D$ space, but that quickly becomes awkward to work with.

[7] Countably many other points have the same degree of singularity. These are the "upper left corners" of nonzero subcells; at all stages k after some stage a, they have density $m2^{k-a}$.

A random version of the Sierpinski multifractal can be seen in Figures 7 and 8. Here, the positions of the elements of a are randomly shuffled at each downward step in the cascade. Statistically, the random and regular versions are identical, but visually, the random version suggests phenomena taken from biology or geography.

Figure 9 shows the empirically fit and theoretical ("universal") *K(q)* curves for the Sierpinski multifractal. The latter will be explained in the next section.

Universality Classes; Form of *K(q)*

By making certain plausible assumptions about the mechanisms generating a multifractal, we can arrive at a "universal" theory, akin to a central limit theorem, for multifractals. The critical assumption is that the underlying generator (analogous to the multiplicative factors in the matrix of the previous example) is a random variable with a specific type of distribution: the *exponentiated extremal Lévy* distribution. This is plausible because Lévy distributions generalize the Gaussian distribution in the central limit theorem.

This leads to a two-parameter family of *K(q)* curves:

$$K(q) = \begin{cases} \dfrac{C_1}{\alpha - 1}(q^{\alpha} - q) & \alpha \neq 1 \\ C_1 \cdot q \log(q) & \alpha = 1 \end{cases} \tag{5}$$

where C_1 acts as a magnification factor and α, related to the tail index of the Lévy generator, determines curvature. These parameters in turn can be related to position and scale parameters μ and σ to be applied to a "standard" Lévy variable $\Lambda_\alpha(-1)$.

The derivation, and an introduction to Lévy variables, is presented in Appendix C.

Synthesis of Multifractals: Extremal Lévy Generators

In creating multifractals for liability applications, we adopt this still somewhat controversial theory of universality.[8] That is, each step of a simulated multiplicative cascade is a multiplication by the random factor a given by equation 32 (Appendix C) for appropriately chosen parameters. A cdf of random step factors corresponding to the best universal fit to the Sierpinski cascade example above is shown in Figure 11 (thick curve). A multiplicative cascade with these random step factors could be used instead of the four-element array used above (shown as thin line step function) to construct a multifractal with roughly the same properties as the Sierpinski multifractal.

The Laplace transform of the logarithm of these factors take on the particularly simple forms described in Appendix C. This fact is exploited in data analysis, as will be explained later in the discussion of Trace Moments.

[8] The scope and relevance of the necessary conditions to real-world phenomena are hotly debated.

Example Spectrum Analysis: Insured Property Portfolios

A preliminary step, to be taken before fitting a *K(q)* curve to suspected multifractal data, is spectrum (Fourier) analysis. The key point is that a multifractal must possess a spectral density having a certain shape: a straight line in a log-log plot. Furthermore, the slope of that line has additional implications. Therefore, spectrum analysis is used as a screening step before applying multifractal analysis. The mathematics relating *K(q)* to spatial spectral density is presented in Appendix D.

The spatial distribution of the human environment has been studied in geography and human ecology. [Major] analyzed homeowners insurance property as a two-stage Poisson process. Multifractal approaches include the analysis by [Tessier et al. 1994] of the global meteorological network (i.e., locations of weather stations) and [Appleby]'s study of population in the USA and Great Britain. Until [Lantsman et. al.], no one had studied the spatial distribution of insured property values (Total Insured Value, or "TIV").

[Lantsman et. al.] show that some portfolios of insured homeowners properties display a spatial distribution consistent with multifractal behavior (over appropriate scales). Figure 12 shows the isotropic power spectra of the insured value density of five geographically distinct regions of an insured property portfolio.

The preparation of such graphs starts with a grid of insured values at a sufficiently small scale of resolution. First, accumulate insured value totals over a 2^{Tm}-squared grid over the UxU area. In practice, we have found Tm=7 or 8 to be comfortable for Pentium-III class machines. If the data originates as individual observations (e.g., geocoded lat-lon locations) then each observation must be assigned to the appropriate grid cell. If the data originates as areal data (e.g., accumulated values for polygons) then the data must be *allocated* to the grid. In any case, make sure that $L=U/2^{Tm}$ is larger than the resolution of the data. For analysis of large portfolios with ZIP-level data, we typically use U = 512 or 1024 miles[9] with Tm = 6 or 7, resulting in a resolution of L=8 miles, which is a bit bigger than the square root of the average area of a ZIP code.[10]

The second step is to compute the 2-dimensional discrete Fourier transform (DFT) of the array. The third is to convert to an isotropic power spectrum. Appendix D has details. Roughly speaking, the isotropic power spectrum reveals the strength (vertical axis) of various periodicities (horizontal axis) in the spatial data, averaged over all directions.

The horizontal axis of Figure 12 represents the wavenumber (spatial frequency) r where, e.g., wavenumber r =10 corresponds to a periodicity of 512/10 = 51.2 miles. The plots stop at the finest resolution of 8 miles, corresponding to wavenumber r = 512/8 = 64. The vertical axis represents the power (spectral density – i.e. Fourier component amplitude – squared) $P(r)$, with arbitrary constant factors used to separate the five curves.

---

[9] A 1024-mile square covers about one-sixth of the USA.

[10] Since most of the population resides in smaller, more densely populated ZIP codes, we feel that an 8-mile resolution is appropriate.

All but one curve show the smooth, loglinear relationship between power and wave number that is to be expected from a self-similar random field. The exception displays higher than expected spectral amplitude at wavenumbers 45-50 (~11 miles) and less than expected at wave numbers 30-35 (~16 miles). This anomaly was traced to unique factors in this insurer's distribution channel. They had a strong affinity marketing program for military personnel. In Washington DC proper, the portfolio's spatial density of insured value was nearly zero. However, in two suburban enclaves adjacent to nearby military bases, the value density was among the highest observed in the region. The two groups were about 11 miles apart and 16 miles away from the center of DC. If not for this unusual geographic structure to the market, the power spectrum would have been similar to that of the other regions.

Fitting *K(q)*; Trace Moment Analysis

In this section, we discuss how to fit a universal *K(q)* curve to spatial data and use the US population density in the northeast (Figure 3 discussed previously) as an example. Conceptually, the idea is very simple: construct an empirical *K(q)* curve by measuring the moment scaling behavior as expressed in equation 14 (Appendix B), then find parameters C_1 and α that produce a best-fitting theoretical *K(q)* curve (equation 5, equation 31 of Appendix C). In practice, a few wrinkles emerge.

Data preparation starts with the gridding process discussed above in the context of spectral analysis. Most of the square grid should contain meaningful data; too many "structural zeroes" (e.g., representing water or other area that cannot by definition support positive values) will distort the analysis. In the case of Figure 3, each grid entry is an approximate[11] count of persons living in that 8x8 mile geographic square.

Having represented the field as a 2^{Tm}-square matrix, normalize the entries by dividing each by the average value of all the entries; this makes the average entry equal one.

The next step is to prepare a series of locally averaged ("dressed") versions of the data matrix, each 2^T on a side for T=0,2,...,Tm-1. Specifically, the four elements indexed by (2*r+i, 2*c+j) (where i=0,1 and j=0,1) of the 2^{T+1} grid are averaged to become the value of the (r,c) element of the 2^T grid. These represent the same field, but at progressively coarser scales of resolution.[12] See Figures 3 through 5, mentioned previously. Note that for each grid, the average cell value is one. The coarsest grid, corresponding to T=0 and scale U, consists of the single entry, one.

The fourth step is to compute q^{th} powers of the dressed fields and look for a loglinear relationship between them and the scale. If multifractal scaling is present, we should see, for each fixed q, a linear relationship between T (the label identifying the coarseness of a

---

[11] Recall the original data was at the ZIP code level of resolution, so entire ZIP codes were allocated to particular grid squares, introducing a bit of distortion at the smallest scales.

[12] As a refinement of this process, we start with two grids, the 2^{Tm}-sided grid as described, as well as a slightly coarser $3*2^{Tm-2}$-sided grid, and operate on them in parallel. This way, we get a factor of 1.5 or 1.33 (ideally it would be the square root of two) between adjacent scale ratios instead of a factor of two. This doubles the sample of scale ratios in the analysis.

grid, equal to $\log_2$ of the number of rows or columns in the grid) and the logarithm of the average of the q^{th} power of the grid entries.

Figure 13 shows this relationship for q = 0.6, 0.9, and 1.4. These so-called *trace moments* are close enough to linear to make the multifractal model appropriate.

Having satisfied ourselves that scaling is present, the fifth step is to estimate *K(q)* values as coefficients in a linear regression version of equation 16 (Appendix B), for each of a range of values for q. A certain amount of judgment is called for, however, in choosing the range over which the regression should be carried out. [Essex] and [Lavallée et. al.] discuss "symmetry breaking" that results from the limitations of sample data. The selected range of scaling must avoid these extremes in order to deliver unbiased estimates of moments, and hence undistorted *K(q)* estimates. Linear regression in this case suggests that K(0.6) = -0.2, K(0.9) = -0.1, and K(1.4) = 0.3. An example of the resulting *empirical K(q) curve* based on slopes estimated from regressions of trace moments corresponding to q values ranging from 0.16 to 4.5 is shown in Figure 14.

Before considering how to best fit a universal *K(q)* to the empirical curve, we must address additional limitations of the methodology. The relation between *K(q)* and *c(γ)* (the latter "box counting" exponent expressing the scaling behavior of probability of extreme values) is given by a Legendre transform; there is a one-to-one correspondence between moments and orders of singularities [Tessier et. al. 1993]. Realistic limitations to data (rounding low values to zero, finite sample size, bounded sample) can limit the range of observable singularities and consequently introduce distortions in the measured *K(q)*. In addition, estimating the universal parameters C_1 and α by nonlinear least squares may run afoul of a substantial degree of collinearity between the parameters.

For such instances, [Tessier et. al. 1993, 1994] developed the *double trace moments* technique. This is based on the observation that if a universal field is exponentiated first by η, then averaged to scale λ, then exponentiated to q, we have the relation

$$K(q,\eta) = \eta^{\alpha} \cdot K(q,1) \tag{6}$$

where the second arguments refer to the exponent of the original field from which the *K(q)* estimate is made. This allows an estimation of the field's α by fixing q and varying η. Figure 15 shows a graph of log K(q,η) vs. log η for various q. Due to the limitations cited above, this equation as applied to sample data tends to break down except for a limited range of η; thus we estimate α as the *maximum* slope observed in the graph. With a good estimate of α in hand, an ordinary least-squares estimate of C_1 is easy to obtain.

In this case, a standard two-parameter nonlinear regression does fine, with α =0.66 and C_1 = 0.72 obtained. The resulting theoretical *K(q)* curve is compared to the empirical version in Figure 16.

Simulating Universal Multifractal Data; Synthetic Geocoding

The utility of a model of insured value emerges when detailed geographical information about a portfolio of risks is lacking. Often the information fed into catastrophe models in the US is based on aggregations at the county or ZIP code level. While this may suffice

for hurricane risk analysis, it does not for thunderstorm wind, tornado or hail perils. On the one hand, the average size of a ZIP code is 8 by 8 miles, and the distribution of properties over the area is typically very sparse, irregular and non-uniform. A damage potential (expected damage rate) field representing a hail or tornado event is of a comparable scale (scattered patches less than a mile wide by a few miles long for hail; narrower and longer for tornadoes), and it, too, is highly non-uniform (e.g., 90% of the damage potential from a tornado occurs in less than 5% of its area). Given that the details of the hazard and exposure fields must be superimposed to obtain a reasonable estimate of losses sustained, one can appreciate the difficulty of working with aggregate data.

Previous solutions to the problem were simplistic and reflected a characterization of TIV over the area either as regularly or randomly uniformly distributed, or, at the other extreme, concentrated at a single point, (i.e., the area's centroid). The result of this kind of misrepresentation is a critical misestimation of the variability inherent in the process of loss generation. Figures 17 and 18 illustrate this.

Figure 17 is a map of a portion of a real homeowner property portfolio. The scale is 8 miles on a side, the average size of a ZIP code. Figure 18 shows a realization of the same number of homes assuming a uniform spatial point process. The true portfolio shows more "clumps and gaps" than the relatively smoother uniform random version. Figure 19 shows the results of applying the multifractal model. While it does not reproduce the original portfolio (no random model would be expected to), it does appear to exhibit the same spatial statistics. When intersected with a number of simulated damage footprints from hail or tornadoes, it will clearly do a better job of estimating the damage probability distribution than will either the uniform random version or a version that puts all the properties at the center of the figure. The uniform distribution will result in too many small loss events and not enough large loss events, and vice-versa for the centroid.

The construction of a synthetic geocoding proceeds as follows:

1. Create a multifractal field over the area in question. Typically, we use a five- to seven-stage process, depending on the outer scale. A seven-stage process divides a square into $2^7x2^7=16{,}384$ grid cells; this is sufficient to carve an 8-mile square into 2.5 acre parcels. At each stage i = 0 to T_m, instantiate a $2^i x 2^i$ array of independent and identically distributed exponentiated extremal Lévy random variables (see equation 32 of Appendix C).[13] In the example of Figure 19 we used the parameters $\alpha = 0.8$, $C_1 = 0.6$. In [Lantsman et. al.], we reported different parameters for industry and selected client portfolios.[14] Combine factors via multiplicative cascade as described for the Sierpinski multifractal.

[13] [Samorodnitsky & Takku] has an efficient algorithm for simulating Lévy variables.

[14] Specifically, $\alpha = 1.024$ and $C_1 = 0.560$ for industry TIV measured at the ZIP code level, and $\alpha = 0.552$, $C_1 = 0.926$ for a geocoded client portfolio. The implications of this difference are discussed in that paper.

2. Normalize the field and use it as a probability distribution to drive a multinomial point process. If the area is a polygon other than a square, then grid cells must be identified as to being inside or outside the polygon. Outside grid cells are zeroed out; inside cell intensity values are divided by the total of all inside values to renormalize. Say the grid probability in cell i is p_i. The desired number of homes, N, is then allocated to each cell N_i, by a multinomial(N, p_1, p_2, ..., $p_{4\sim rm}$) joint random variable draw. In practice, this is implemented by a sequence of conditional binomial r.v. realizations. The first r.v. is N_1~binomial(N, p_1). Subsequent cells' realizations are conditional on all that precede, viz., N_3~binomial(N-N_1-N_2, p_3/(1-p_1-p_2)), etc.

Project APOTH: Thunderstorm Simulation

Occurrence rates for small scale thunderstorm-related perils (wind, hail, tornadoes) have traditionally been computed as an annual rate averaged over a fairly wide region. This is done by counting the number of occurrences of some peril of interest – say, hail two inches or more in diameter, tornadoes F3 or more, etc. – in a given area (frequently, a one- or two-degree longitude/latitude box). This count is normalized to the area of the box and the period of record. When this process is continued for more boxes (usually overlapping), contour maps can be plotted showing the smoothed variation in the rate. These types of maps are often developed for differing severity levels, such as hail >1", >2", etc. or tornadoes >F2, >F3, etc. To this extent, both frequency and severity are incorporated into them.

Maps such as these are often used to estimate the probability of occurrence for an event of at least a certain severity at a single location. Such an application might be estimating the chance that a nuclear power plant will be hit by an F4 or F5 tornado. These maps can also be used to estimate probabilities of an event hit to a town or subdivision.

Unfortunately, point-frequency maps are not very useful for modeling the typical insurance catastrophe loss event. While there are cases where a single violent tornado or a single storm of large diameter hail hits a highly populated area and produces a large loss, there are also other cat events whose losses are aggregates of many moderate losses which occurred in different locations, possibly over several states and over several days.

The goal of Project APOTH (Atmospheric Perils Other Than Hurricane) is to develop the capability to credibly estimate probabilistic losses from the thunderstorm perils of hail, tornado, and straight-line winds (non-tornadic high wind gusts). The APOTH project presently has models that can realistically simulate both the geographical and seasonal characteristics of severe storms, as well as model their small-scale ground damage patterns as they affect homeowners anywhere in the lower 48 states of the USA. The model can be easily extended to include other lines of business once vulnerability functions become available from further research.

One objective of natural hazard simulation is to produce a "future history" of meteorological events, in sufficiently rich detail to be able to explore the range of damage effects on a subject portfolio of insured properties, yet maintain a statistically stationary relationship to the available historical data upon which the simulation is based.

The production of tornadoes and hail involves meteorological processes exhibiting complex behavior over a wide range of scales, from synoptic weather patterns (thousands of km) down to the size of the hailstone (millimeters or centimeters). We have made use of multifractal modeling, not only to distribute property values in statistically appropriate patterns, but directly in the simulation of the hazards themselves.

Multifractal modeling is not appropriate to all scales, however. Thunderstorms exhibit a strong seasonality during the year, nonhomogeneity of occurrence frequencies over distances of thousands of km, and anisotropy in terms of preferred directions of movement. At the smaller scales, the structure of tornado tracks and hail streaks (continuous bands of hailfall) are also highly idiosyncratic. In between, however, we have found that the scale of the *swath* (tens to hundreds of km) on a single day is amenable to multifractal modeling.

Figure 20 shows a set of reported hail occurrences for 3/30/98. Unfortunately, while swaths may make conceptual and meteorological sense, data are not reported in swath groupings. Before we can analyze swaths, we must identify them, using various tools including Bayesian classification, modal clustering, and nearest-neighbor methods. Figure 21 shows the same set of reports, now grouped into meaningful swaths.

In order to expand the data into a meaningful set of possible alternative scenarios, we have followed the practice of other modelers in using the historical data as a template for a synthetic "probabilistic database" of possible events. Figure 22 exemplifies the typical practice of equally displacing all reported events by some random X-Y translation vector.[15] One of our innovations is to use multifractal modeling to create and recreate alternative detailed patterns within a given swath.

Our procedure is as follows:

1. Historical reports are grouped into swaths as mentioned above.
2. Swaths are characterized by a small number of key parameters: the location, size, orientation, and eccentricity of the bounding ellipse; the prevailing storm motion direction within, and parameters describing the overall intensity level of the activity. In the case of hail, intensity is defined by a categorical type label and the total volume of hail (number of hailstones). In the case of tornado, intensity is defined by Fujita class-specific Poisson parameters for the number of touchdowns and two principal component scores defining the conditional distribution of tornado path lengths. In the case of non-tornado wind, intensity is defined as total wind power (watts).
3. When an historical swath is drawn from the database as a template for a simulated swath, the ellipse is gridded at the 1-km scale and a multifractal field (with parameters appropriate to the peril and type) is laid down over the grid. As described above for simulated geocoding, this field is "condensed" to a schedule of report (hail, tornado, or wind event) locations.

[15] Since this translation is by no more than a degree in either direction, it is a bit difficult to see at first.

4. Details of each report (hail streak size and intensity details; tornado F-class and track length, etc.) are drawn from conditional distributions, with correlation induced with the intensity of the underlying multifractal field at the point of condensation.

Figure 23 shows several realizations of the multifractal simulation of these particular swaths. Note how they respect the ellipse boundaries, yet vary dramatically in their inner detail. A much richer variety of possible outcomes is made possible, compared to simple location-shift models, but the statistics of event properties and their spatial colocation are still respected.

Conclusion

In this part I paper, we introduced the ideas of fractal point sets and multifractal fields. We showed that while those mathematical constructs are rather bizarre from a traditional point of view (e.g., theory of smooth, differentiable functions), they nonetheless have applicability to a wide range of natural phenomena, many of which are of considerable interest to the casualty actuary. We showed how to analyze sample data from multidimensional random fields, detect scaling through the use of the power spectrum, detect and measure multifractal behavior by the trace moments and double trace moments techniques, fit a "universal" model to the trace moments function $K(q)$, and use that model to simulate independent realizations from the underlying process by a multiplicative cascade. In particular, we discussed synthetic geocoding and the simulation of non-hurricane atmospheric perils.

In the companion part II paper, we focus on time series analysis and financial applications.

Appendix A: Binomial Multifractal

This appendix establishes a relationship between orders of singularities and fractal dimension in the binomial multifractal on the half-open unit interval (0,1]. We follow the presentation in [Mandlebrot 1989].

Divide the interval into two halves (each open on the left) of equal length. Distribute $0<p<1$ of the mass uniformly on the left half and 1-p of the mass uniformly on the right half (here p is a constant throughout all stages of the construction). Repeat on each subinterval.

At stage k of the construction, we have 2^k pieces of length 2^{-k}, of which $k!/(h!(k-h)!)$ of them have density $p^h(1-p)^{k-h}$.

Any point x in the interval can be expanded as a binary number $0.b_1b_2b_3\ldots$.[16] By considering the sequence of expansions truncated at b_k, we make meaningful statements about the behavior of the measure at x. For example, define

$$f(k) \equiv \frac{1}{k}\sum_{i=0}^{k} b_i \,. \tag{7}$$

Then, in a $\lambda=2^{-k}$ –wide neighborhood of x, the average density is $p^{1-f(k)}(1-p)^{f(k)} = \lambda^{-\alpha(k)}$, where $\alpha(k) = \log_2(p^{1-f(k)}(1-p)^{f(k)})$. If $f \equiv \lim_{k\to\infty} f(k)$, the proportion of 1's in the binary expansion, exists, then we can consider that the density behaves as $\lambda^{-\alpha}$ in the limit. Such a point is termed a *singularity of exponent* α. As λ gets smaller, the density may grow without limit or shrink to zero, but the rate of that growth is controlled by the exponent α, a quantification of spikiness at that point. (This is the classical Hölder exponent.)

What is the fractal dimension of the set of such points? (Here, the exposition becomes quite deliberately sketchy, as proper delta-epsilon arguments would take up an undue amount of space.) At stage k, we have a total of 2^k intervals, of which $n = k!/((kf)!(k(1-f))!)$ have density defined by $f(k) = f$. Recalling Stirling's approximation for factorial,

$$x! \approx \sqrt{2\pi}\cdot\exp(-x)\cdot x^{x+\frac{1}{2}} \tag{8}$$

we can write an approximation for n as

$$n(\lambda) = \frac{\left(f^f(1-f)^{1-f}\right)^{-k}}{\sqrt{2\pi k f(1-f)}} \tag{9}$$

This gives us a fractal dimension $d = f\log_2 f + (1-f)\log_2(1-f)$. Since the exponent $\alpha = f\log_2(1-p) + (1-f)\log_2(p)$ is also a function of f, we have a functional relationship between the order of the singularity α and the fractal dimension d of the set of points having that exponent.

---

[16] Since binary xyz0111... is the same as xyz1000..., let us agree to use only the 111... representation for such cases. (This is consistent with our closing the right side of intervals.)

Appendix B: Analysis of Multifractal Fields

A random field is called *stationary*[17] if the distribution of $\varphi(\boldsymbol{r}_1)$ is the same as that of $\varphi(\boldsymbol{r}_2)$ for any different $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. This does not imply the two random variables are independent, however. For example, a multivariate normal may have identical marginal distributions but nonetheless possess a nontrivial correlation structure. A nonstationary field is said to have *stationary increments* if the distribution of $\varphi(\boldsymbol{r}_1) - \varphi(\boldsymbol{r}_2)$ depends only on the difference vector $\boldsymbol{r}_1 - \boldsymbol{r}_2$. Furthermore, for $D > 1$, such a field is said to be *isotropic* if the distribution of $\varphi(\boldsymbol{r}_1) - \varphi(\boldsymbol{r}_2)$ depends only on the *magnitude* of that vector, $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$.

Our discussion follows [Novikov & Stewart], [Shertzer & Lovejoy], [Marshak et. al.] and [Menabde et. al.] in the general context of a D-dimensional space and for stationary fields. The generalization to non-stationary fields will be discussed in Appendix D.

Formally, consider a measure $\mu(X)$ whose domain consists of a σ-field of subsets X of $\mathbf{R}^D$. Define the scale-L average density as:

$$\varphi_L(r) = L^{-D}\mu(V) \tag{10}$$

where V is a D-dimensional hypercube of side length L centered at $\boldsymbol{r}$. Our first condition for μ to be a random multifractal is to assume $\varphi_L(\boldsymbol{r})$ is a random field. For a particular realization of μ, we can compare such field realizations at different scales of resolution L and U by considering their relative densities[18] defined as:

$$a_{L,U}(r) = \varphi_L(r)/\varphi_U(r) \tag{11}$$

where $L < U$, therefore $V_L \subset V_U$. This is only defined for nonzero values of φ_U, but note that when it is zero, so must be φ_L. We have the property that:

$$a_{L,U} = a_{L,\rho}a_{\rho,U} \tag{12}$$

where $L < \rho < U$ (therefore $V_L \subset V_\rho \subset V_U$), and we have suppressed mention of volume centers $\boldsymbol{r}$. This is true for any realization, and therefore can be considered a statement about the random variables as well.[19]

These relative densities are random fields in their own right. They characterize how the fluctuation (or *intermittency*) of the field varies as a function of scale. The assumption of

---

[17] Physicists would use the word *conservative*.

[18] These would be known as Radon-Nikodym derivatives to a statistician or *breakdown coefficients* to a physicist.

[19] It is helpful to think of the measure μ as a physical quantity, such as mass, rather than a probability measure. That way, probability statements about the random μ will not be confused with statements about the properties of particular realizations of μ.

stationarity implies that $a_{L,U}$ is a random variable whose distribution does not depend on the position of the volume-center $\boldsymbol{r}$. Furthermore, we assume it depends only on the *ratio* L/U and that the random variables $a_{L,\rho}$ and $a_{\rho,U}$ in equation (3) are independent. This last statement is the technical definition of μ(X) being a *statistically self-similar* (a.k.a. *scale-invariant*, or *scaling*) random measure.

Scaling of Moments, *K(q)* Function

It is possible to show that under these assumptions the statistical moments of $a_{L,U}$ have the property:

$$E(a_{L,U}^q) = E(a_{L,\rho}^q)E(a_{\rho,U}^q) \tag{13}$$

where $E(\cdot)$ is expected value operator. Since the moments of $a_{L,U}$ depend only on the ratio L/U, the most general expression for scaling behavior of statistical moments is:

$$E(a_{L,U}^q) = (L/U)^{-K(q)} \tag{14}$$

with the boundary conditions $K(0) = K(1) = 0$.

It is worth noting that for some processes $K(q) = \theta \cdot (q-1)$ (for $q>0$) which is usually referred to as *simple scaling* or the *(mono)fractal* case. However, in nature, most processes exhibit a more complex behavior and the $K(q)$ function evaluation requires a more elaborate approach. There is a wide class of random multiplicative cascade models that produce *multiscaling* behavior and *multifractal* fields [Parisi & Frisch].

For the special case where scale steps are factors of two, we can specialize equation 14.

From the definition of $a_{L,U}$ in equation 11, and noting that φ_U is equal to one, we can write:

$$E(a_{U\cdot 2^{-T},U}^q) = E(\varphi_{U\cdot 2^{-T}}^q) = 2^{T\cdot K(q)} \tag{15}$$

or

$$\log_2\left[E(\varphi_{U\cdot 2^{-T}}^q)\right] = T \cdot K(q) \tag{16}$$

This form reveals $K(q)$ as the coefficient in a log-linear regression between the scale index T and average q-power of the field, as used in empirical data analysis.

Appendix C: Universality Classes; Form of *K(q)*

To further explore the structure and behavior of the *K(q)* function we follow [Schertzer & Lovejoy], [Lovejoy & Schertzer 1990], and especially [Menabde et. al.] and formalize the idea of a multiplicative cascade generator (MCG):

$$G_{L,U} = -\ln(\varphi_L L^D / \varphi_U U^D) \tag{17}$$

We assume that the measure is not zero on any finite hypercube, therefore G is everywhere defined. By definition of φ_L (equation 10 of appendix B), the ratio is less than one and so $G_{L,U}$ is a non-negative random variable whose distribution depends only on the ratio L/U. For arbitrary n, we can introduce n hypercubes of side length ρ_n which nest between V_L and V_U so that:

$$L/\rho_1 = \rho_1/\rho_2 = ... = \rho_n/U = (L/U)^{1/n} \tag{18}$$

After a series of transformations the resulting expression for $G_{L,U}$ will be:

$$G_{L,U} = G_{L,\rho_1} + G_{\rho_1,\rho_2} + ... + G_{\rho_n,U} \tag{19}$$

The random variables on the right-hand side of equation (8) are assumed to be independent and *identically distributed* random variables with a pdf:

$$p(g;(L/U)^{1/n}) = p(g;\rho_i/\rho_{i+1}) \tag{20}$$

which depends solely on the scale ratio, $(L/U)^{1/n}$. The property expressed in equations 19 and 20 implies that the probability density for $G_{L,U}$ belongs to the class of *infinitely divisible* distributions [Feller]. The natural candidate for a MCG would therefore be a random variable with a stable Lévy distribution.

An aside on Lévy random variables is in order. Lévy random variables generalize Gaussian (normal) random variables in the Central Limit Theorem. The CLT states that the distribution of a sum of a set of N independent, identically distributed random variables *with finite variance* converges to a normal distribution as the number N increases without bound. More generally, if the restriction to finite variance is removed, we can say that the sum converges to a Lévy distribution.

Lévy distributions are characterized by four parameters: α, which must be in (0,2]; β, which must be in [-1,1]; and μ and $\sigma>0$, which are otherwise unrestricted. The latter two are location and scale parameters, respectively, allowing us to express a Lévy random variable as $\mu+\sigma\Lambda_\alpha(\beta)$ where Λ is "standardized" and depends on only two parameters. Note that σ is not the standard deviation because in general, variance is infinite for a Lévy random variable. The parameter α is the tail index: the case $\alpha=1$ gives the Cauchy distribution while the case $\alpha=2$ gives the Normal distribution. As x increases without bound, the probability that a Lévy random variable exceeds x is proportional to $x^{-\alpha}$. The second parameter, β, is a symmetry index: if $\beta=0$, then the distribution is symmetric; otherwise, the probability of the upper tail is proportional to $1+\beta$ and the probability of

the lower tail is proportional to 1-β (in the large-x limit). When α=1, the β parameter becomes irrelevant, and is conventionally set to 0. While there is no closed-form expression for the distribution function for Lévy variables, the characteristic function is analytically tractable.[20]

To develop a moment scaling relation for the random multifractal μ(X) we apply the Laplace transform to the density function *p(g; L/U)*:

$$\psi(s;L/U)=\int_0^\infty \exp(-s\cdot g)\,p(g;L/U)\,dg \tag{21}$$

where $s \geq 0$.

Because *p(g; L/U)* is the pdf of an infinitely divisible distribution, from equation 18 we can conclude:

$$\psi(s;L/U)=\psi^n(s;(L/U)^{1/n}) \tag{22}$$

Equation 22 has the solution:

$$\psi(s;L/U)=(L/U)^{\chi(s)} \tag{23}$$

where, according to the general properties of infinitely divisible distributions [Feller], $\chi(s)$ can in the most general case be represented by a Lebesgue integral:

$$\chi(s)=\int_0^\infty \frac{1-\exp(-s\cdot x)}{x} M(dx) \tag{24}$$

where M is a measure such that the integral:

$$\int_1^\infty x^{-1} M(dx) < \infty \tag{25}$$

For processes under consideration with some degree of rigor we can limit ourselves to considering only measures M having a density M^*. In such cases we can replace the Lebesgue integral with a Riemann integral, replacing $M(dx)$ with $M^* dx$. It is this density function M^* (or equivalently $\chi(s)$ or $p(g;L/U)$) that completely determines the properties of the MCG and therefore the (statistical) properties of the self-similar multifractal μ(X).

The expression in equation 21 could be considered as an expectation of $exp(-sG_{L,U})$ and can be rewritten as follows:

---

[20] Refer to [Samorodnitsky & Takku] for information on simulating and evaluating Lévy random variables.

$$\psi(s; L/U) = E\left[\exp(-sG_{L,U})\right]$$
$$= E\left[\exp(s\ln(\varphi_L L^D / \varphi_U U^D)\right] = E(\varphi_L / \varphi_U)^s (L/U)^{sD} \tag{26}$$

From equations 23 and 26 we can find the moment scaling relation:

$$E(\varphi_L / \varphi_U)^s = (L/U)^{-sD+\chi(s)} \tag{27}$$

From equations 14 (appendix B) and 27, after replacing s with q, we can get following expression:

$$K(q) = qD - \chi(q) \tag{28}$$

Since by definition in equation 14, $K(1)=0$, one has the normalization condition in equation 28 that $\chi(1)=D$. The asymptotic behavior of $K(q)$ could be deduced from equations 24 and 28 as:

$$K(q) = qD + O(1) \tag{29}$$

One can choose any form for the density measure M^* that satisfies the convergence and normalization conditions of equations 25 and 28. The most appealing measure is:

$$M^*(x) \propto x^{-\alpha} \tag{30}$$

(specifying only the limiting behavior for large x) which corresponds to a stable Lévy distribution [Feller]. With this choice of measure and proper renormalization we can express $K(q)$ as:

$$K(q) = \begin{cases} \dfrac{C_1}{\alpha - 1}(q^\alpha - q) & \alpha \neq 1 \\ C_1 \cdot q \log(q) & \alpha = 1 \end{cases} \tag{31}$$

This expression represents the classes of "universal generators" [Schertzer & Lovejoy]. The first remarkable thing to notice is that a universal generator is characterized by only two fundamental parameters (C_1, α). The idea behind the introduction of universality classes is that whatever generator actually underlies the multiplicative cascade giving rise to a random multifractal, it may "converge" (in some sense) to a well-defined universal generator.

With only two degrees of freedom, the $K(q)$ curves represented by universal multifractals are of a limited variety. As mentioned previously, $K(q)$ is constrained to go through the points (0,0) and (1,0) with negative values when $0<q<1$ and positive values for $q>1$. The parameter C_1 clearly behaves as a vertical scaling factor. The α parameter affects the curvature, as can be seen in Figure 10, with the extreme case of $\alpha \to 0$ converging to a straight line (with discontinuity at $q = 0$).

For this "universality" result to be useful, we must also investigate which classes of MCG are stable and attractive under addition and will at least converge for some positive moments (not necessarily integer order). The task to specify universality classes could be

accomplished by considering the Lévy distribution in a Fourier framework, i.e., its characteristic function. The restriction imposed by the Laplace transform (equation 21) is that we require a steeper than algebraic fall-off of the probability distribution for positive order moments, hence, with the exception of the Gaussian case ($\alpha = 2$), we have to employ strongly asymmetric, "extremal" Lévy laws ($\beta = -1$), as emphasized by [Schertzer & Lovejoy]. The Lévy location parameter μ is fixed by the normalization constraint and the scale parameter σ is derived from C_1 [Samorodnitsky & Takku]. Roughly speaking, the universality theory states that multifractals built from random multiplicative cascades are statistically equivalent to those built from a special class: the *exponentiated extremal Lévy variables*:

$$a = \exp\left(\mu + \sigma \cdot \Lambda_\alpha(-1)\right) \tag{32}$$

According to [Schertzer & Lovejoy], we can designate the following main universality classes by specifying the parameter α:

1. $\alpha = 2$: the Gaussian generator is almost everywhere (almost surely) continuous. The resulting field is a realization of the log-normal multiplicative cascade introduced by [Kolmogorov], [Obukhov], and [Mandelbrot 1972] to account for the effects of inhomogeneity in three-dimensional turbulent flows (turbulent cascades).
2. $2 > \alpha > 0$: the Lévy generator is almost everywhere (almost surely) discontinuous and is extremely asymmetric.
3. $\alpha = 0+$: this limiting case corresponds to divergence of every statistical moment of the generator and represents the so-called "β" model.

Appendix D: Spectrum Analysis; *K(q)* and Spectrum Slope

In this section, we explore the relation between the moment scaling function $K(q)$ and the power spectrum of the stationary field φ_L that represents a random multifractal at the (sufficiently small) scale of resolution L.[21] Recall that the power spectrum of a time series or one-dimensional stochastic process quantifies the magnitude (amplitude) of cycles of various lengths (periods). Spectral analysis generalizes to multidimensional fields by characterizing not only the amplitude and periodicity of such "waves" but their directions as well. An *isotropic* power spectrum averages the D-dimensional power spectrum over all directions, converting it to a one-dimensional spectrum.[22]

Because of Fourier duality between the correlation function of the field and its power spectrum [Feller] it is customary in analysis of empirical stochastic processes to examine the correlation structure of a process and then map it into Fourier space. But the correlation function is not well suited to analyzing non-stationary fields so we need to develop some guidance as to how to check for stationarity, and, if it exists, how to quantify the underlying field.

Because in the case of stationarity the functional form of the correlation function closely relates to the $K(q)$ function, we can be reasonably confident in establishing a direct link between the power spectrum and $K(q)$ function of the field. Following [Menabde et al.], we demonstrate how it could be accomplished.

For a D-dimensional isotropic random field $\varphi_L(r)$:

$$E(\varphi_L(r_1)\varphi_L(r_2)) = C(|r_1 - r_2|) \tag{33}$$

where $C(r)$is the correlation function of the field.

The Fourier transform of $\varphi_L(r)$ field is defined as:

$$\psi(k) = \int \exp(-ikr)\varphi_L(r)d^D r \tag{34}$$

where $i = \sqrt{-1}$, the unit imaginary number. For an isotropic field (equation 33) one has that:

$$E(\psi(k)\psi(h)) \sim \delta(k-h)P(k) \tag{35}$$

---

[21] Historically, power spectrum analysis played a central role in identifying and characterizing the scaling properties of self-similar random fields. Recent advances [Marshak et al.] in understanding the limitations of applicability and sensitivity of power spectrum analysis leads one to realize that the issue of stationarity is critical in qualifying and quantifying intermittency of the field. The erroneous assumption that everything could be extracted from knowledge of the spectral exponent leads to a failure to discriminate between qualitatively different fields.

[22] This is explained more fully below.

where $\delta(\,)$ is a delta function (1 at 0, 0 elsewhere) and $P(k)$ is the isotropic power spectrum. On the other hand, from equations 33 and 34 we can get the expression:

$$E(\psi(k)\psi(h)) = \int\int \exp(-ikr_1 - ihr_2) C(|r_1 - r_2|) d^D r_1 d^D r_2 \tag{36}$$

After some mathematical manipulations with integrals involving change of variables, introduction of polar coordinates, and performing the integration over the angular variables, one can obtain the following elegant result for the power spectrum of a stationary isotropic random field:

$$P(k) \propto k^{-D+K(2)} \tag{37}$$

The practical implementation of the above on an NxN square grid V(m,n) of intensity values is as follows: Compute the array:

$$H(k,h) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} \exp\left\{-2\cdot\pi\cdot i\cdot\left(\frac{m\cdot k + n\cdot h}{N}\right)\cdot V(m,n)\right\}. \tag{38}$$

Convert this to the isotropic power spectrum by accumulating values $|H(k,h)|^2$ (i.e. complex magnitude squared) into one-dimensional array cells A(r) where

$$r = round\left(\left|\begin{matrix}(k+N/2)\bmod N - N/2\\ (h+N/2)\bmod N - N/2\end{matrix}\right|\right). \tag{39}$$

(Here, the vertical bars indicate vector magnitude, i.e., square root of sum of squares.) Then convert A values to averages P by dividing each accumulated A entry by the number of H cells contributing to the entry.

Equation 37 could be utilized in many ways: to check a D-dimensional stationary isotropic field for SS properties, to verify the validity of a numerical approximation of the $K(q)$ function at the point $q = 2$, or to examine a non-stationary field with stationary increments (Brownian motion and "fractional Brownian motion"). Note that $P(k)$ and $K(2)$ can be computed by independent methods from the same data, enabling one to verify the consistency of assumptions about stationary increments.

If we relax the assumption of stationarity, the problem of identification and characterization of SS fields develops some complications. We outline some important guidelines in handling non-stationary fields:

1. First of all, the power spectrum analysis still can indicate self-similarity of the field under investigation, revealing the following form:

$$P(k) \propto k^{-\beta} \tag{40}$$

2. For D-dimensional fields the condition $\beta > D$ will indicate lack of stationarity, but some transformations of the original field (like power-law filtering or taking the absolute value of small-scale gradients) could produce a stationary field.

3. The spectral exponent β contains information about the degree of stationarity of the field. The introduction of a new parameter H (sometimes called the Hurst exponent) related to β could aid in the task of characterizing the degree of persistence or long-term memory of the field. We will illustrate the importance of parameter H for time series in the part II paper.
4. The arguments that the correlation function is not well suited for non-stationary situations (because of its translation dependence) led to the development of new ideas about the statistical properties of non-stationary fields to be properly estimated by spatial averaging procedures. The Wiener-Khinchine relation applicable to fields with stationary increments [Monin & Yaglom] states that it is the second-order structure function – not the correlation function – that is in Fourier duality with the power spectrum. We will introduce the structure function in the context of time series analysis in the part II paper and illustrate how the structure function is the one-dimensional analog of the $K(q)$ function.

A further refinement of the multiplicative cascade is to pass from the *discrete* cascade, which is what has been described up to this point, to the *continuous* cascade. The idea behind a continuous cascade is that rather than proceeding in identifiable steps, the multiplicative transfer of intensity variation between scales happens continuously at all scales. [Schertzer & Lovejoy] describe a method of implementing continuous cascades by means of the Fourier transform.

The functional form for $K(q)$ (equation 31 in appendix C) could be extended to nonstationary fields, and fractional integration (power-law filtering in Fourier space) could be used to transform simulated stationary random fields to any desired degree of non-stationarity (in the sense of spectral exponent β). This is considered more fully in the part II paper.

References

Appleby, Stephen. (1996). "Multifractal Characterization of the Distribution Pattern of the Human Population," Geographical Analysis, vol. 28, no. 2, pp. 147-160.

Barnsley, Michael (1988). *Fractals Everywhere,* Boston: Academic Press.

Davis, Anthony, Alexander Marshak, Warren Wiscombe, and Robert Calahan. (1994). "Multifractal Characterizations of Nonstationarity and Intermittency in Geophysical Fields: Observed, Retrieved, or Simulated," *Journal of Geophysical Research* v. 99, no. D4, pp. 8055-8072.

Essex, Christopher. (1991). "Correlation Dimension and Data Sample Size," in *Non-Linear Variability in Geophysics: Scaling and Fractals,* Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications, volume 2,* New York: Wiley.

Kolmogorov, A. N. (1962). "A Refinement of Previous Hypothesis Concerning the Local Structure of Turbulence in Viscous Incompressible Fluid at High Reynolds Number," *Journal of Fluid Mechanics,* vol. 13, pp. 82-85.

Ladoy, P., S. Lovejoy, and D. Schertzer. (1991). "Extreme Variability of Climatological Data: Scaling and Intermittency," in *Non-Linear Variability in Geophysics: Scaling and Fractals,* Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Lantsman, Yakov, John A. Major, and John J. Mangano. (1999). "On the Multifractal Distribution of Insured Property," to appear in *Fractals.*

Lavallée, D., D. Schertzer, and S. Lovejoy. (1991). "On the Determination of the Codimension Function," in *Non-Linear Variability in Geophysics: Scaling and Fractals,* Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Lovejoy, Shaun, and Daniel Schertzer. (1990). "Multifractals, Universality Classes and Satellite and Radar Measurements of Cloud and Rain Fields," *Journal of Geophysical Research* v. 95, no. D3, pp. 2021-2034.

Lovejoy, Shaun, and Daniel Schertzer. (1991). "Multifractal Analysis Techniques and the Rain and Cloud Fields from 10^{-3} to 10^{6} m," in *Non-Linear Variability in Geophysics: Scaling and Fractals,* Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Major, John A. (1999). "Index Hedge Performance: Insurer Market Penetration and Basis Risk," in *The Financing of Catastrophe Risk,* K. Froot, ed., National Bureau of Economic Research, Chicago: University of Chicago Press, pp. 391-426.

Mandlebrot, B. (1972). "Statistical Models of Turbulence," in *Lecture Notes in Physics,* vol. 12, M. Rosenblatt and C. Van Atta, eds., Springer Verlag, p. 333.

Mandlebrot, B. (1982). *The Fractal Geometry of Nature,* Freeman.

Mandlebrot, B. (1988). "An Introduction to Multifractal Distribution Functions," in *Fluctuations and Pattern Formation,* H. E. Stanley and N. Ostrowsky, eds., Kluwer.

Mandlebrot, Benoit B. (1989). "The Principles of Multifractal Measures," in *The Fractal Approach to Heterogeneous Chemistry,* D. Avnir, ed., Chichester: John Wiley & Sons.

Marshak, Alexander, Anthony Davis, Warren Wiscombe, and Robert Calahan. (1997). "Scale Invariance in Liquid Water Distributions in Marine Stratocumulus. Part II: Multifractal Properties and Intermittency Issues," Journal of the Atmospheric Sciences, vol. 54, June, pp 1423-1444.

Menabde, Merab, Alan Seed, Daniel Harris, and Geoff Austin. (1997). "Self-Similar Random Fields and Rainfall Simulation," *Journal of Geophysical Research* v. 102, no. D12, pp. 13,509-13,515.

Monin, A. S., and A. M. Yaglom. (1975). *Statistical Fluid Mechanics, volume 2,* Boston: MIT Press.

Novikov, E. A., and R. Stewart. (1964). "Intermittency of Turbulence and Spectrum of Fluctuations in Energy-Dissipation," *Izv. Akad. Nauk SSSR, Ser. Geofiz,* vol. 3, pp. 408-412.

Obukhov, A. (1962). "Some Specific Features of Atmospheric Turbulence," *Journal of Geophysical Research* v. 67, pp. 3011-3014.

Parisi, G., and U. Frisch. (1985). "A Multifractal Model of Intermittency," in *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics,* Ghil, Benzi, and Parisi, eds., North-Holland, pp. 84-88.

Pecknold, S., S. Lovejoy, D. Schertzer, C. Hooge, and J. F. Malouin. (1998). "The Simulation of Universal Multifractals," in *Cellular Automata: Prospects in Astrophysical Applications,* J. M. Perdang and A. Lejeune, eds., World Scientific.

Samorodnitsky, Gennady, and Murad S. Taqqu, (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance,* New York: Chapman and Hall.

Schertzer, Daniel, and Shaun Lovejoy. (1991). "Nonlinear Geodynamical Variability: Multiple Singularities, Universality and Observables," in *Non-Linear Variability in Geophysics: Scaling and Fractals,* Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Tessier, Y., S. Lovejoy, and D. Schertzer. (1993). "Universal Multifractals: Theory and Observations for Rain and Clouds," *Journal of Applied Meteorology,* vol. 32, February, pp. 223-250.

Tessier, Y., S. Lovejoy, and D. Schertzer. (1994). "Multifractal Analysis and Simulation of the Global Meteorological Network," *Journal of Applied Meteorology,* vol. 33, December, pp. 1572-1586.

Wilson, J., D. Schertzer, and S. Lovejoy. (1991). "Continuous Multiplicative Cascade Models of Rain and Clouds," in *Non-Linear Variability in Geophysics: Scaling and*

Fractals, Daniel Schertzer and Shaun Lovejoy, eds., The Netherlands: Kluwer Academic Publishers.

Woo, Gordon (1999). *The Mathematics of Natural Catastrophes,* London: Imperial College Press.

Figures for Part I

John A. Major
Yakov Lantsman

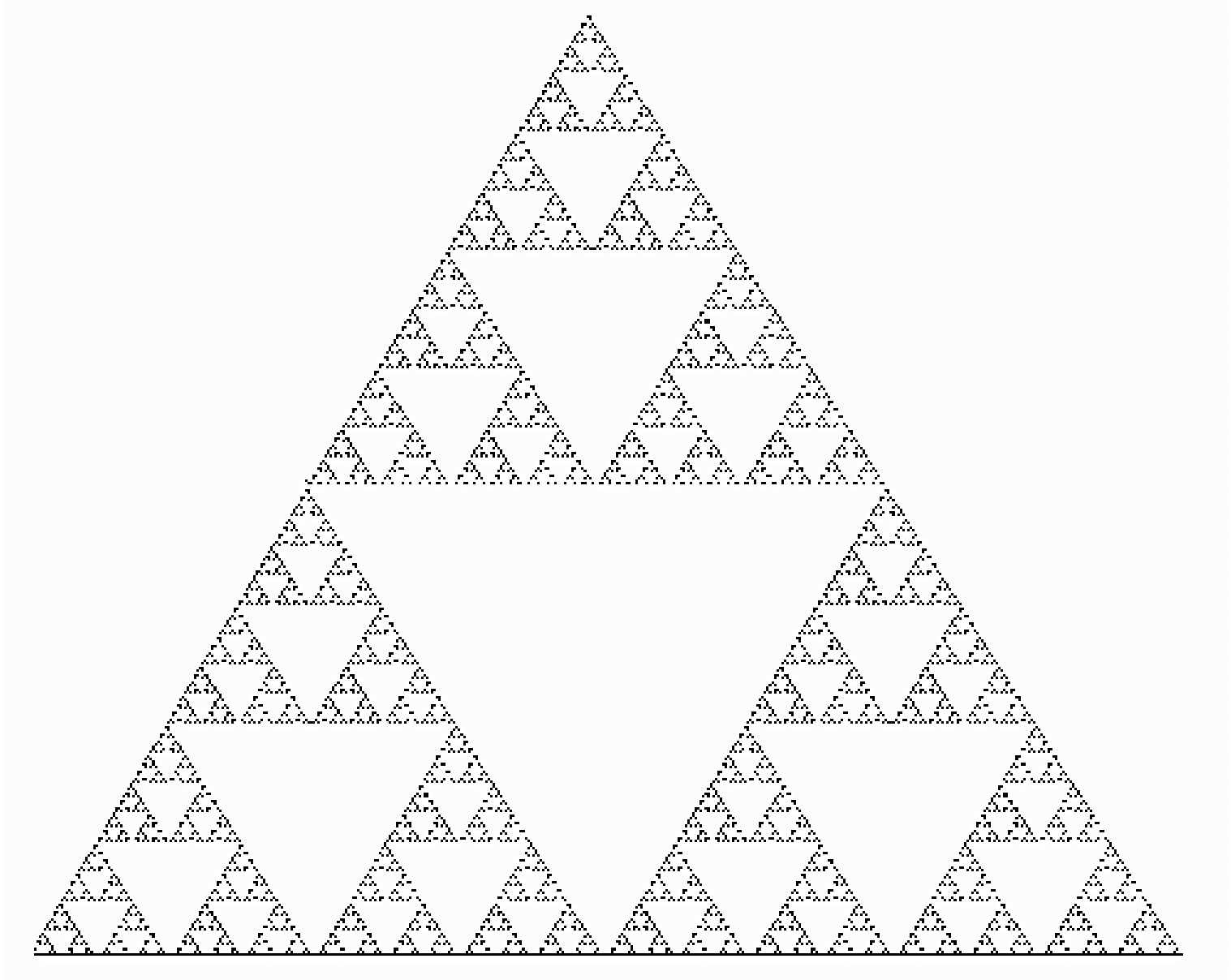

Figure 1: Sierpinski Triangle

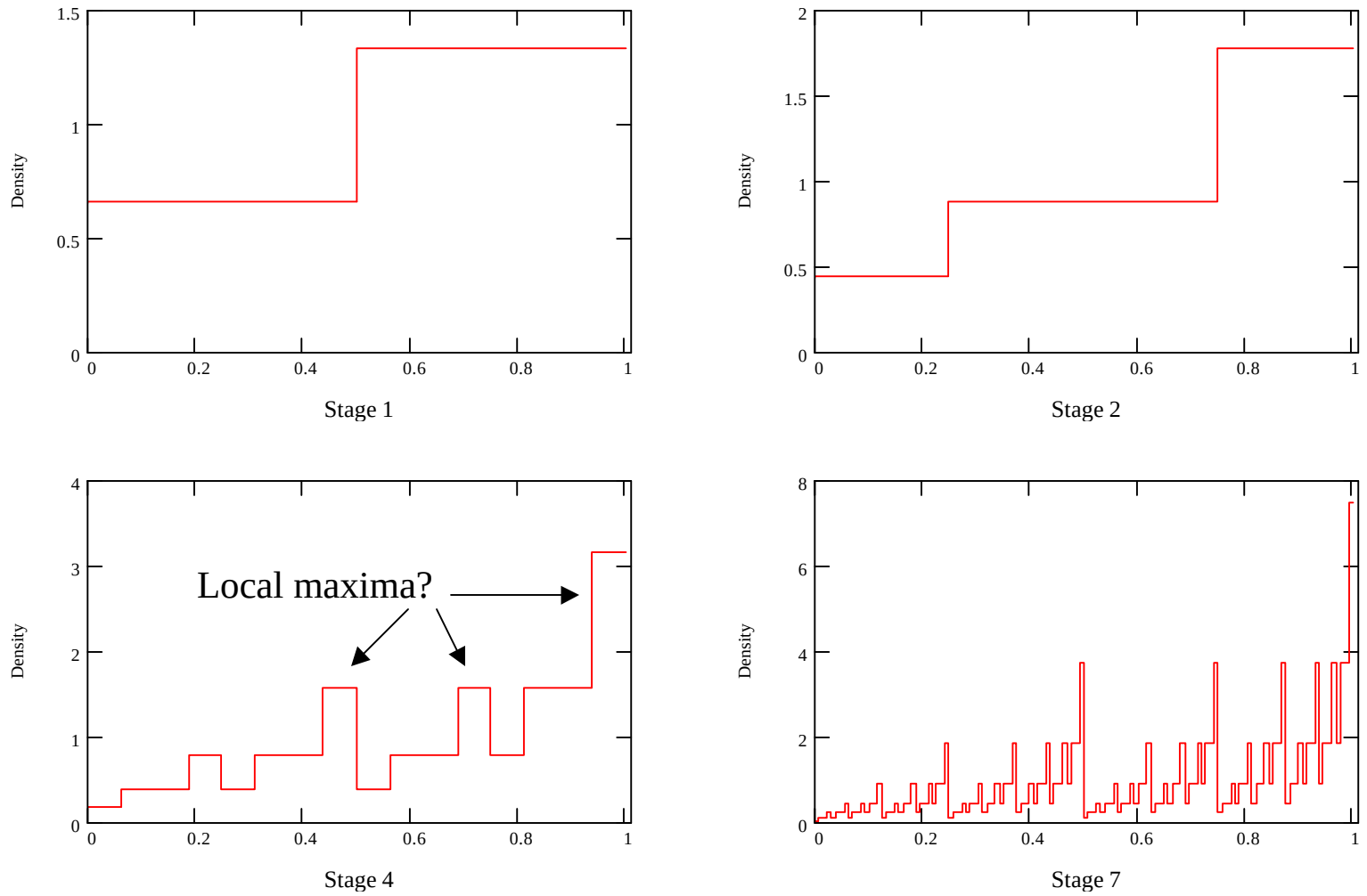

Figure 2: Stages of the Binomial Multifractal

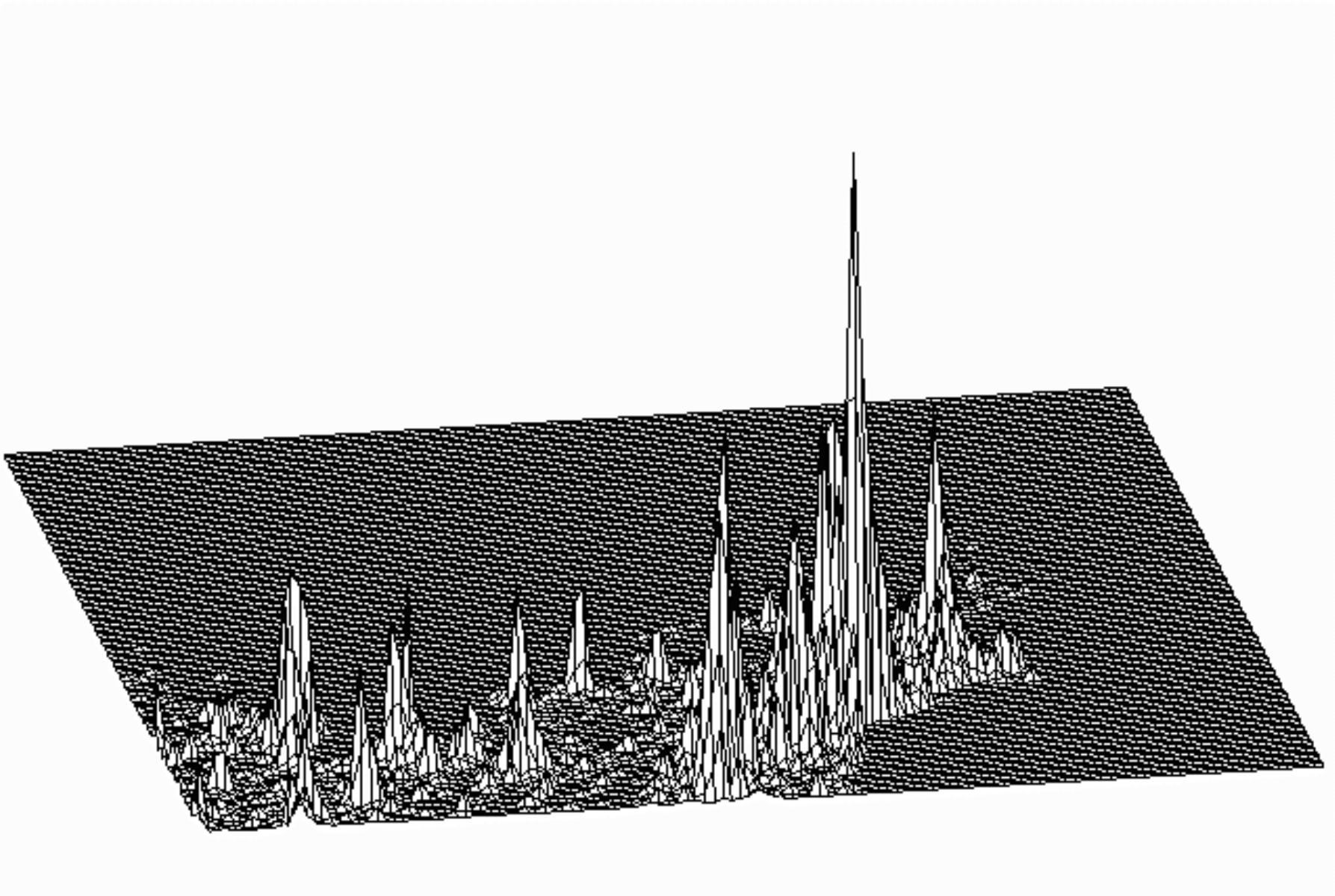

Figure 3: Northeastern USA Population Density

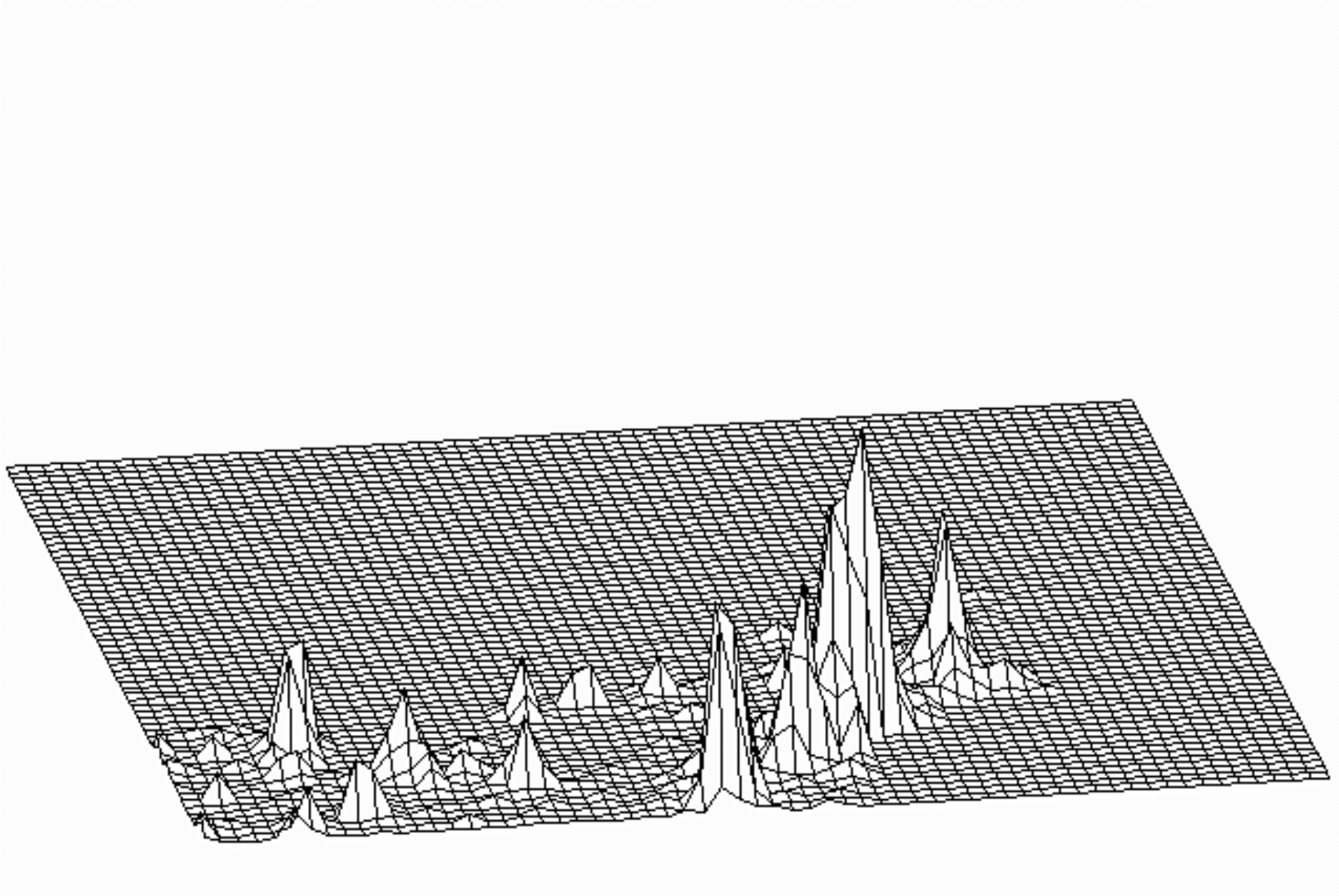

Figure 4: N.E. USA at 1/2 resolution

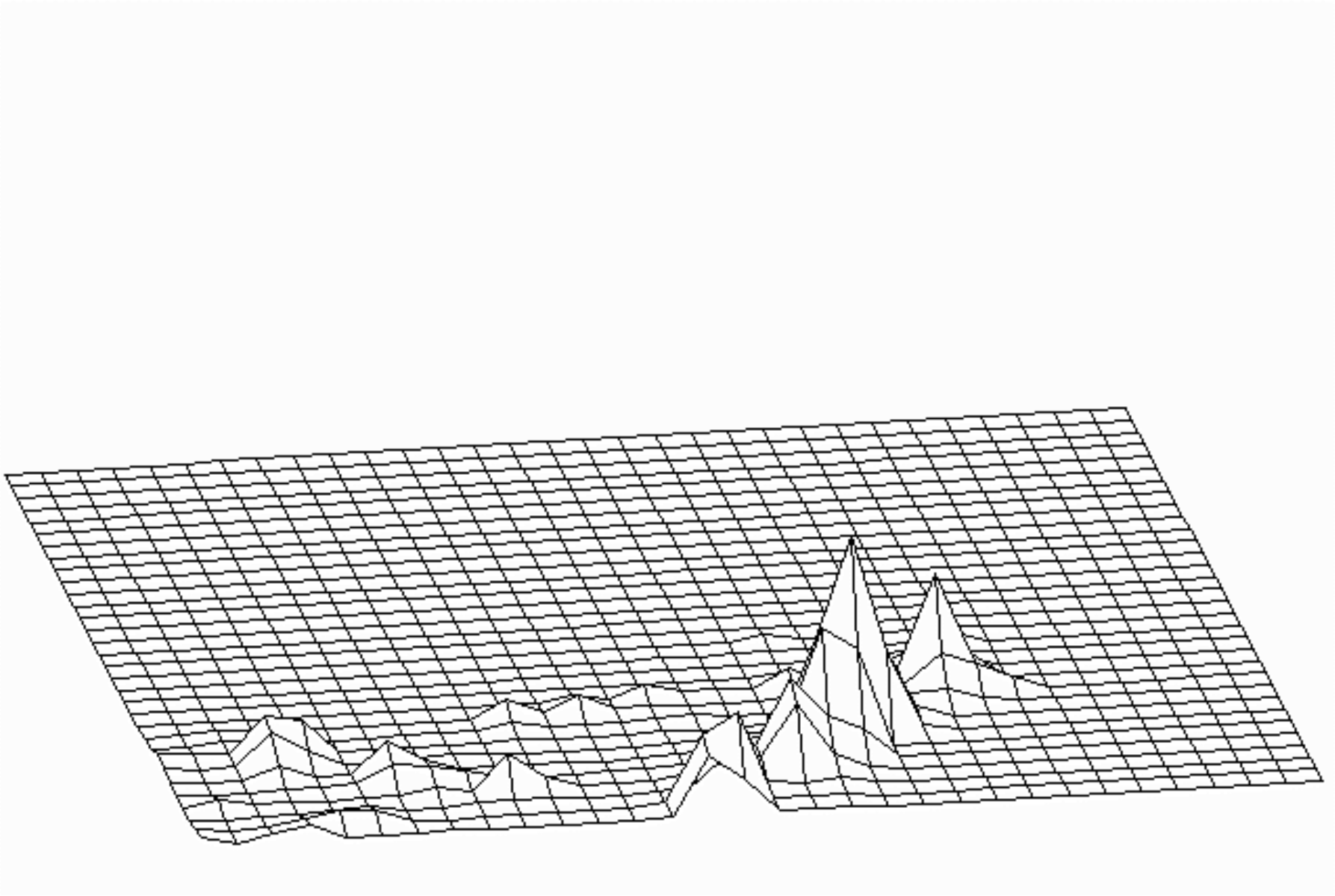

Figure 5: N.E. USA at 1/4 resolution

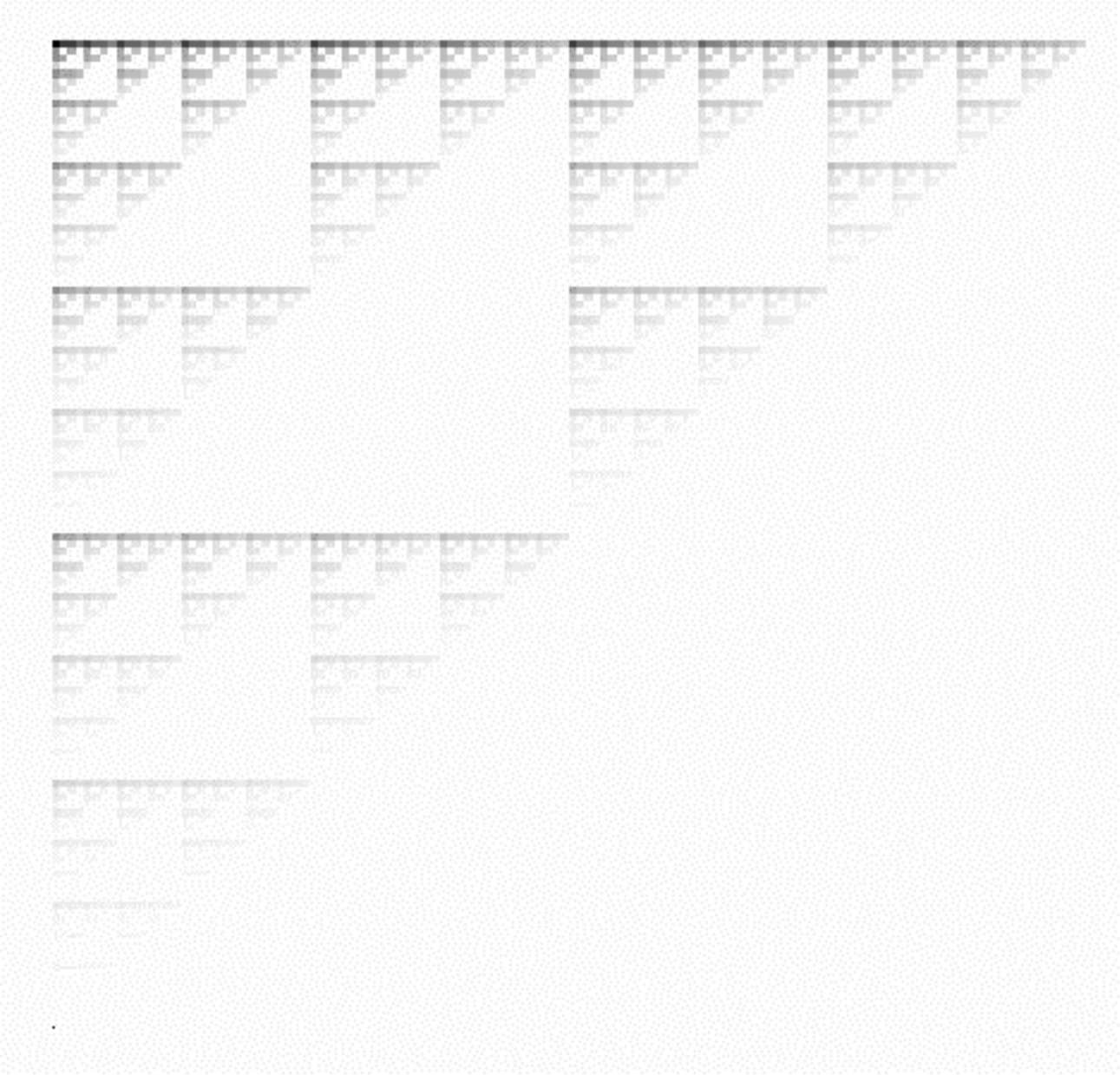

Figure 6: Sierpinski Multifractal

Figure 7: Random Sierpinski Multifractal

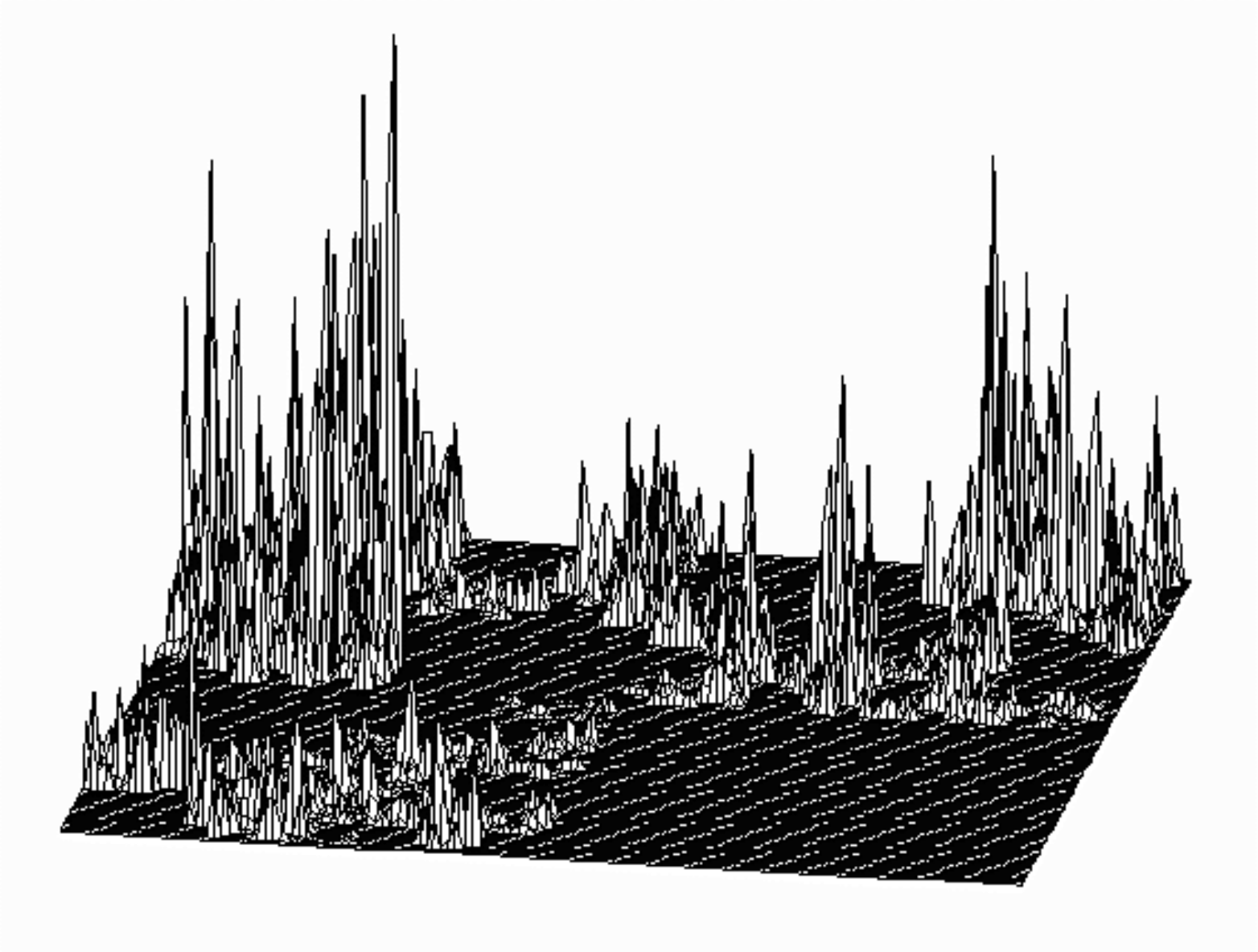

Figure 8: Random Sierpinski Multifractal (Perspective)

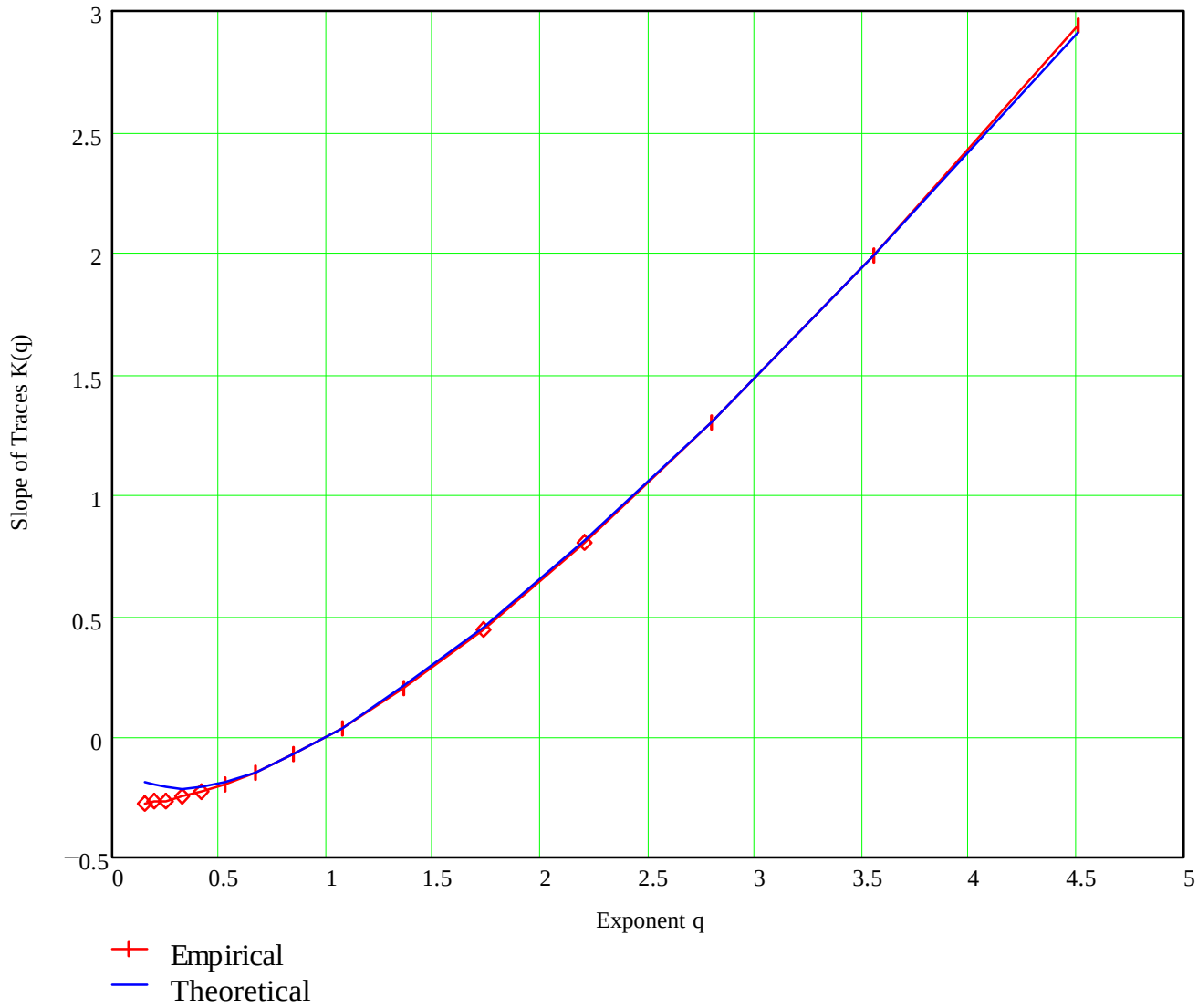

Figure 9: Sierpinski Multifractal K(q) Function

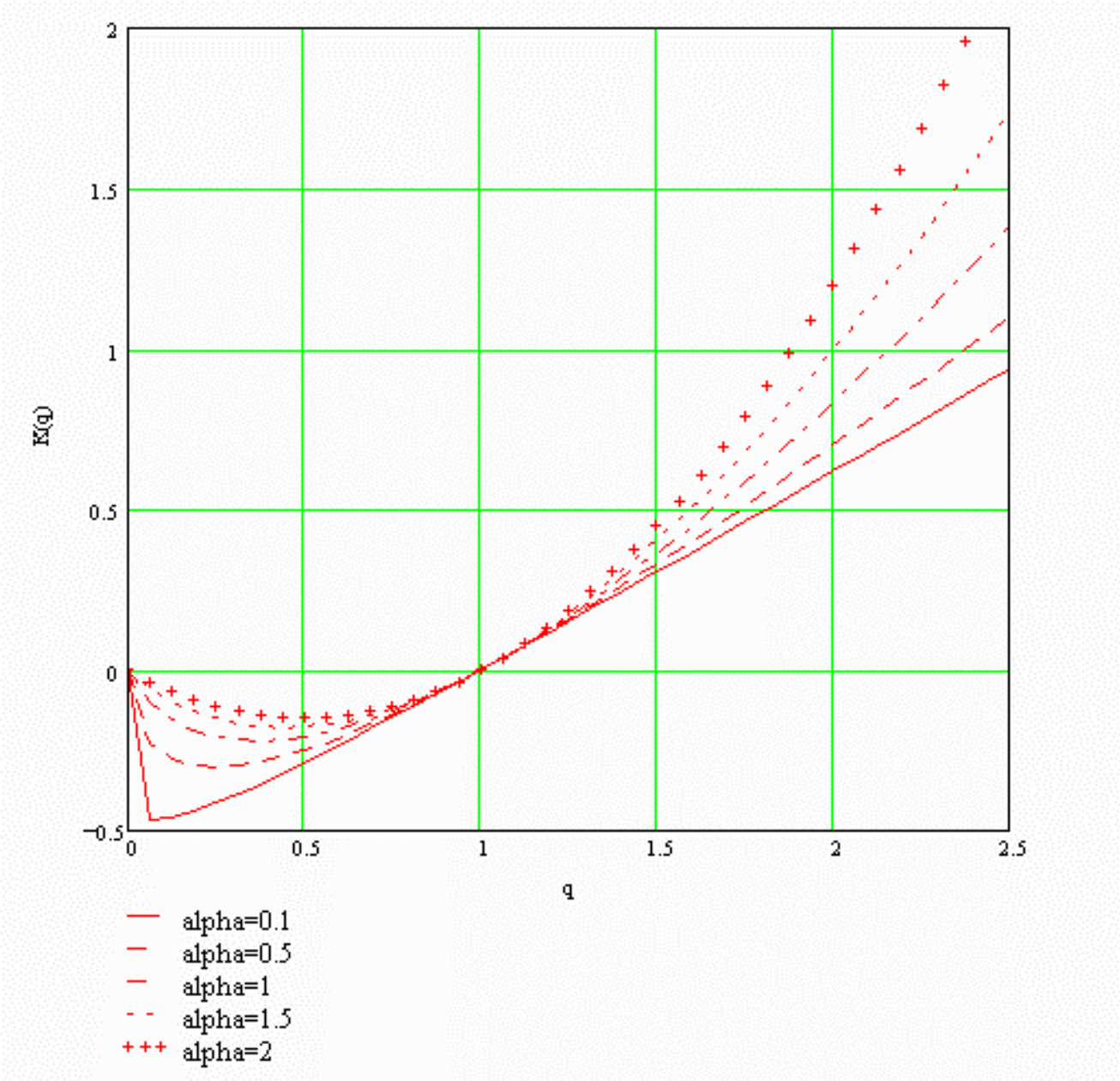

Figure 10: Effect of Alpha on K(q) Curve

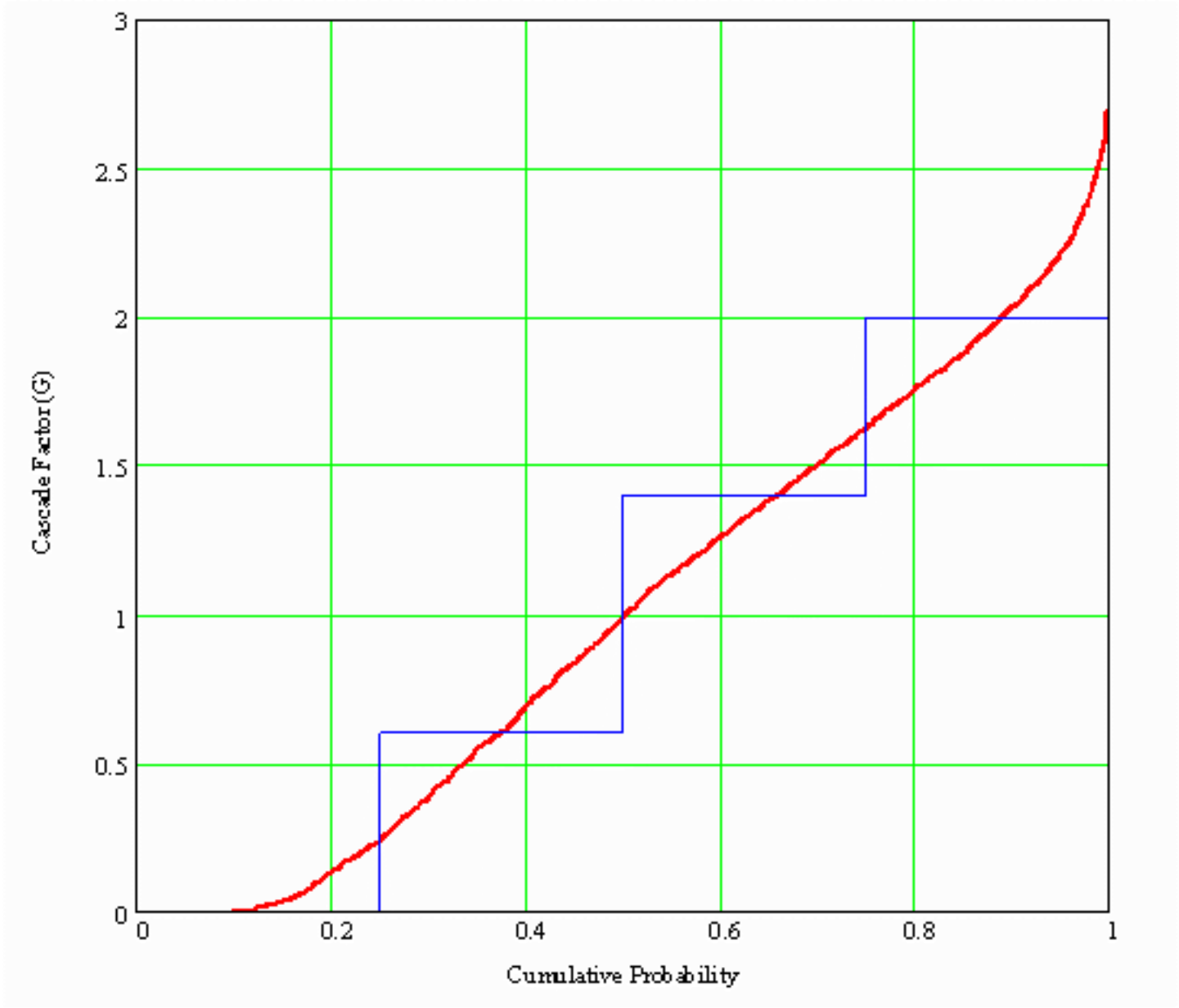

Figure 11: Universal Generator Distribution

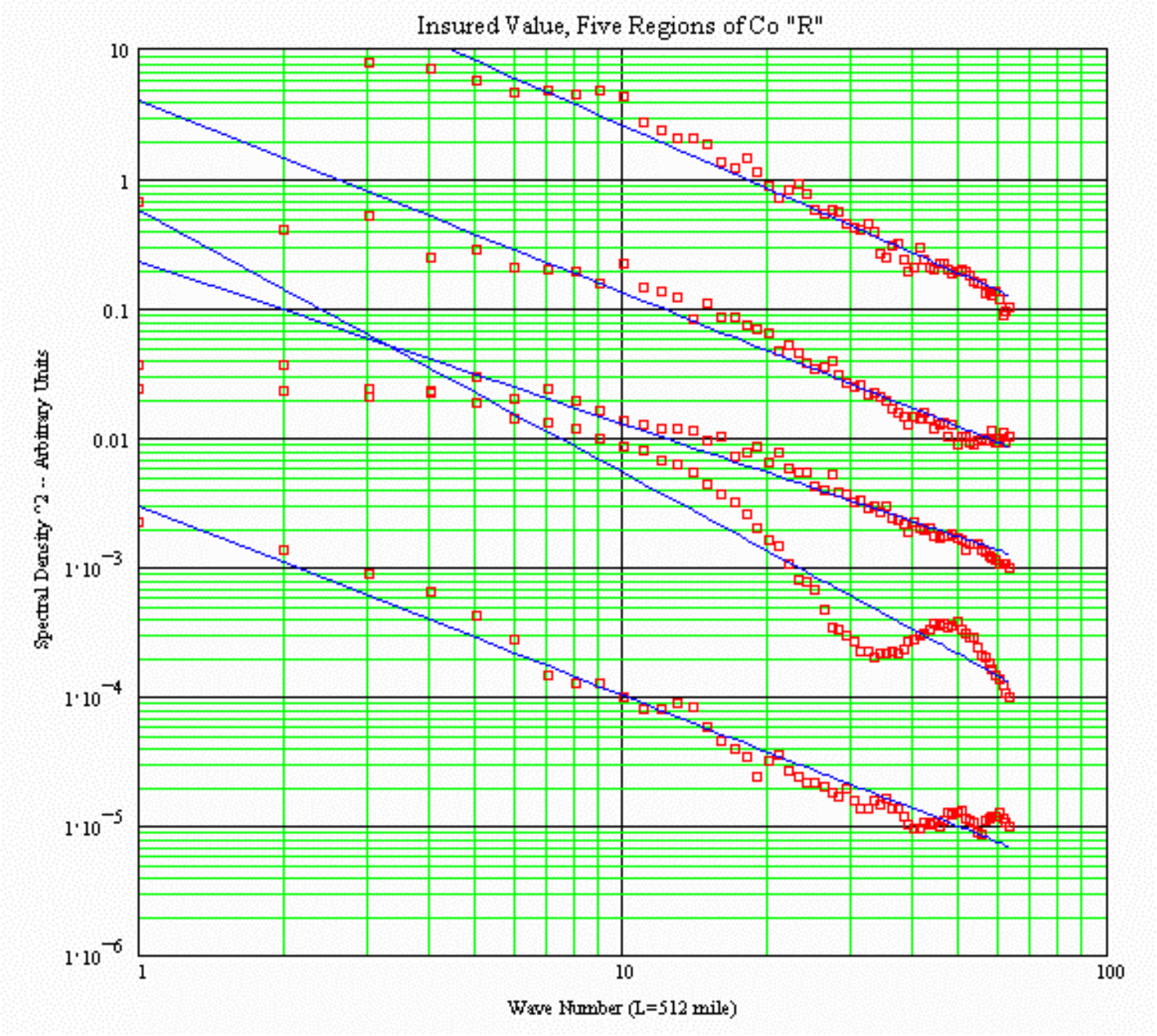

Figure 12: Comparative Portfolio Spectra

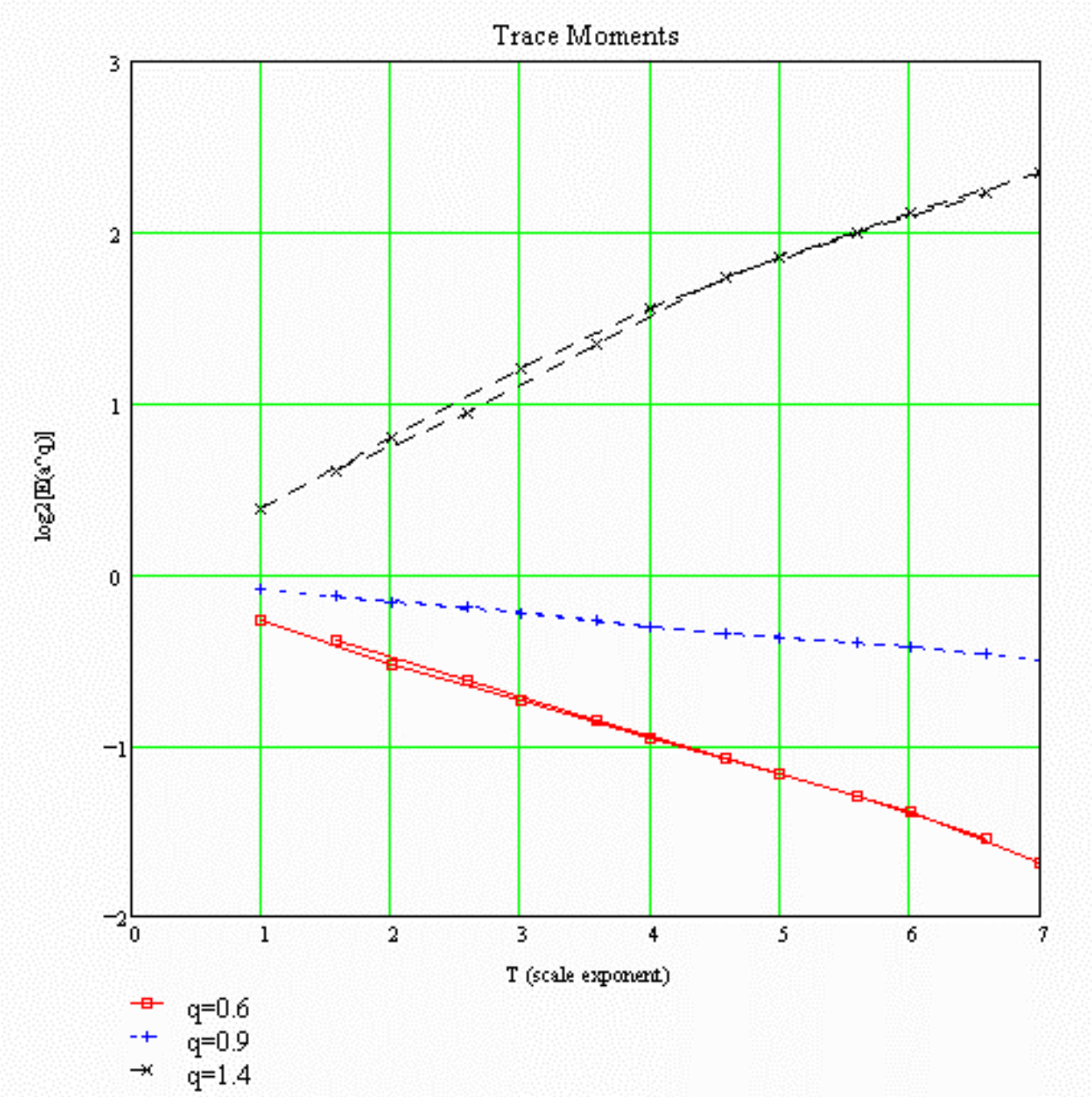

Figure 13: Selected Trace Moments of Population

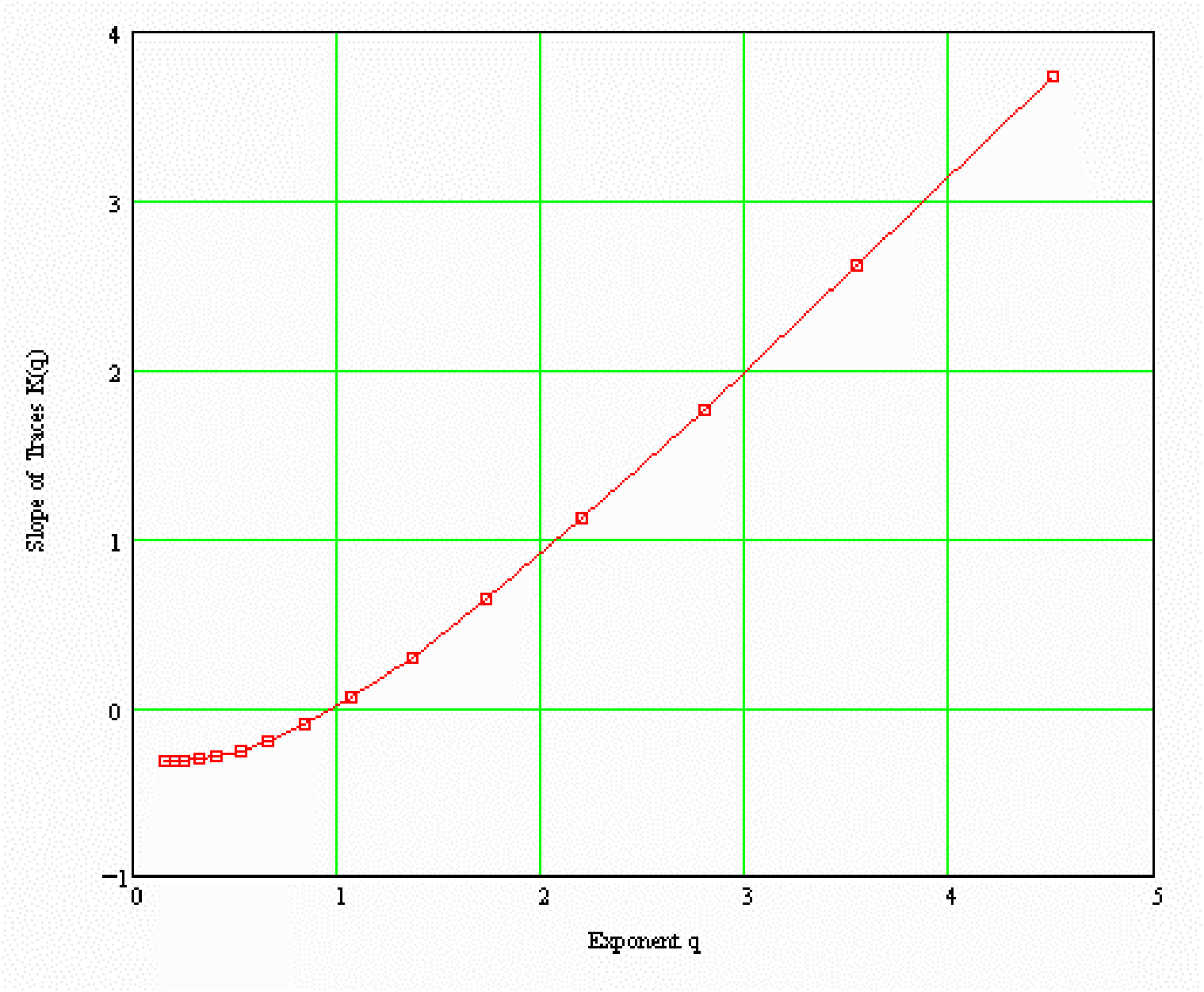

Figure 14: Empirical K(q) of Population

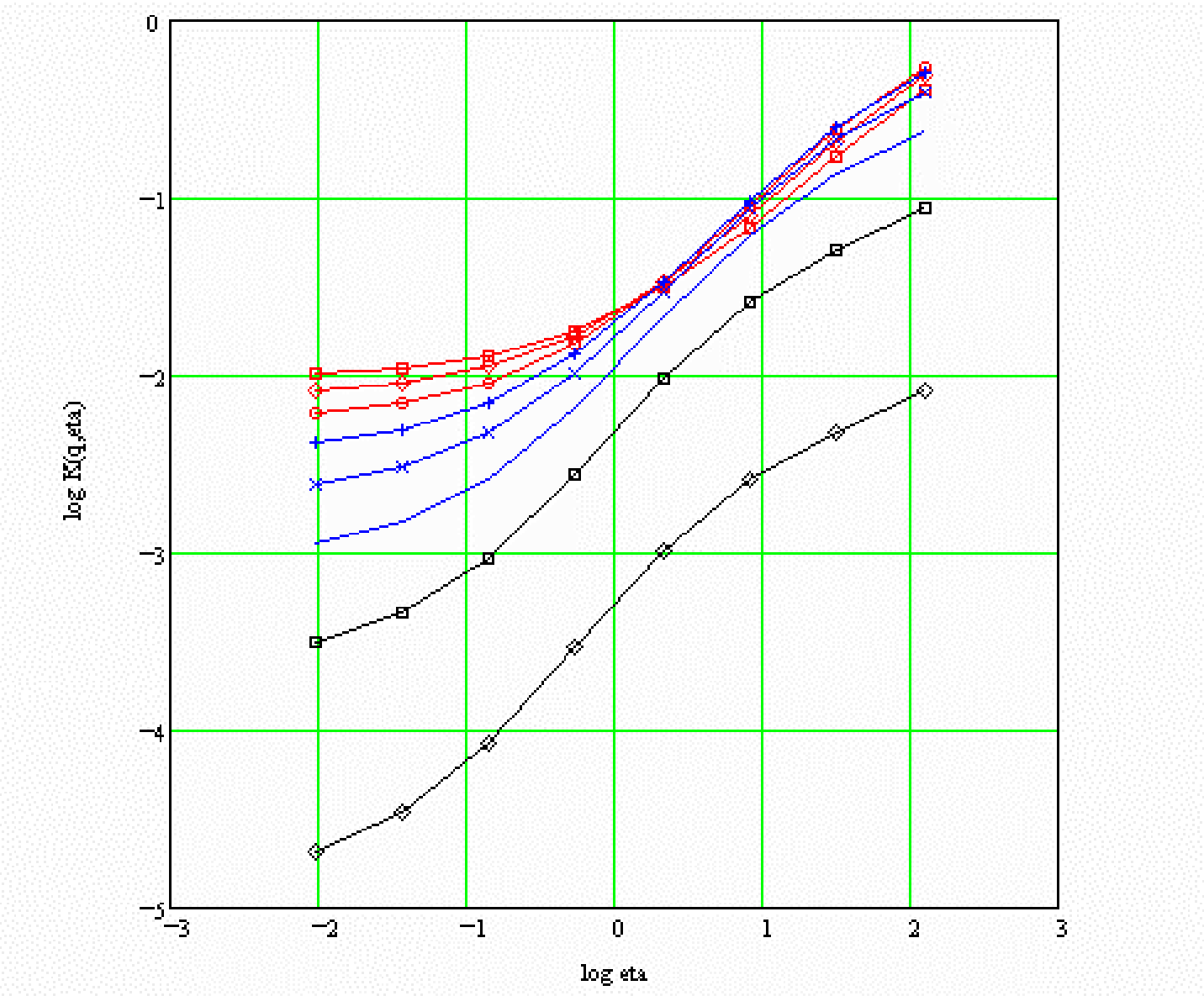

Figure 15: Double Trace Search for Alpha

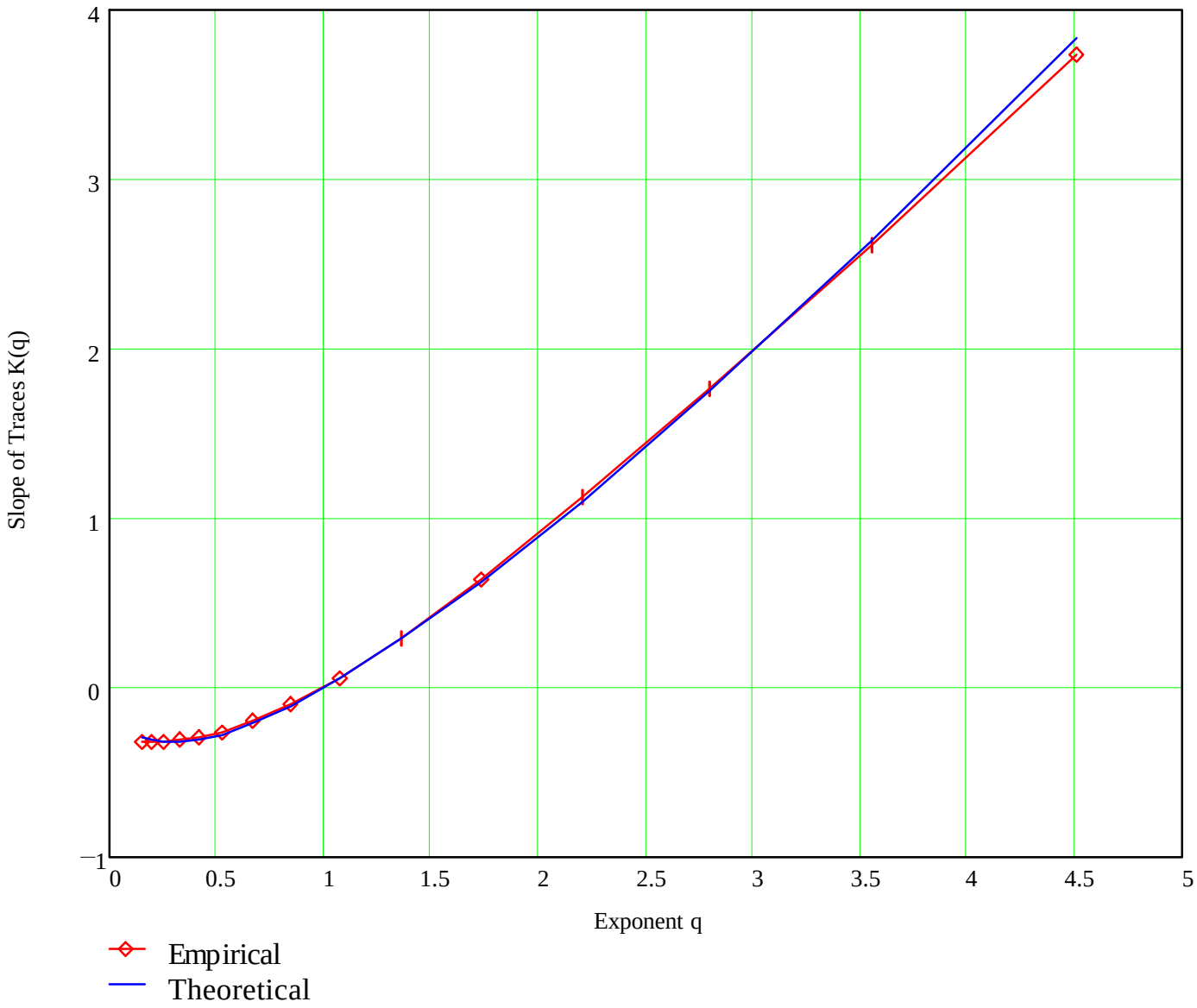

Figure 16: Empirical vs. Theoretical K(q) for Population

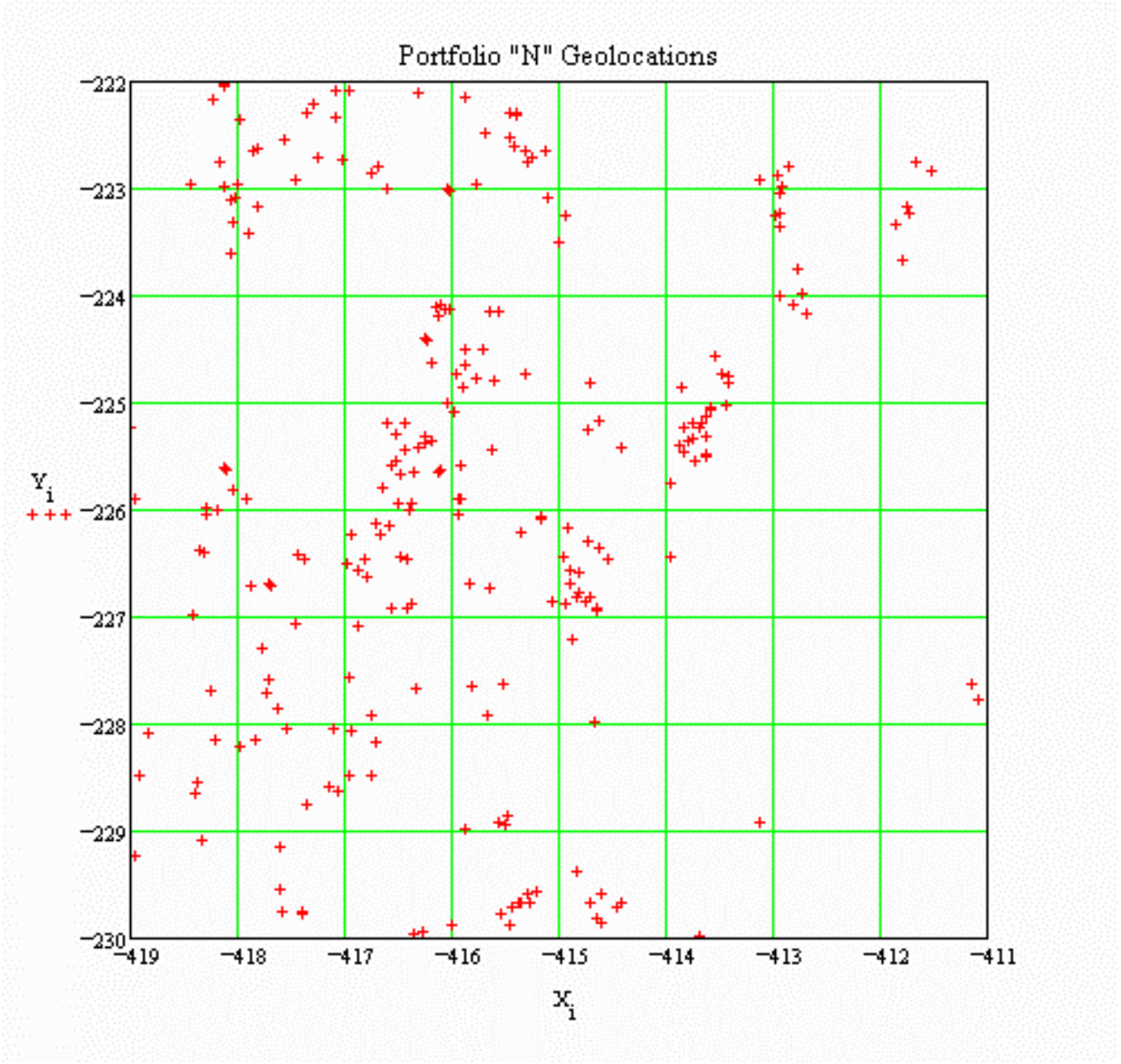

Figure 17: Actual Portfolio Map

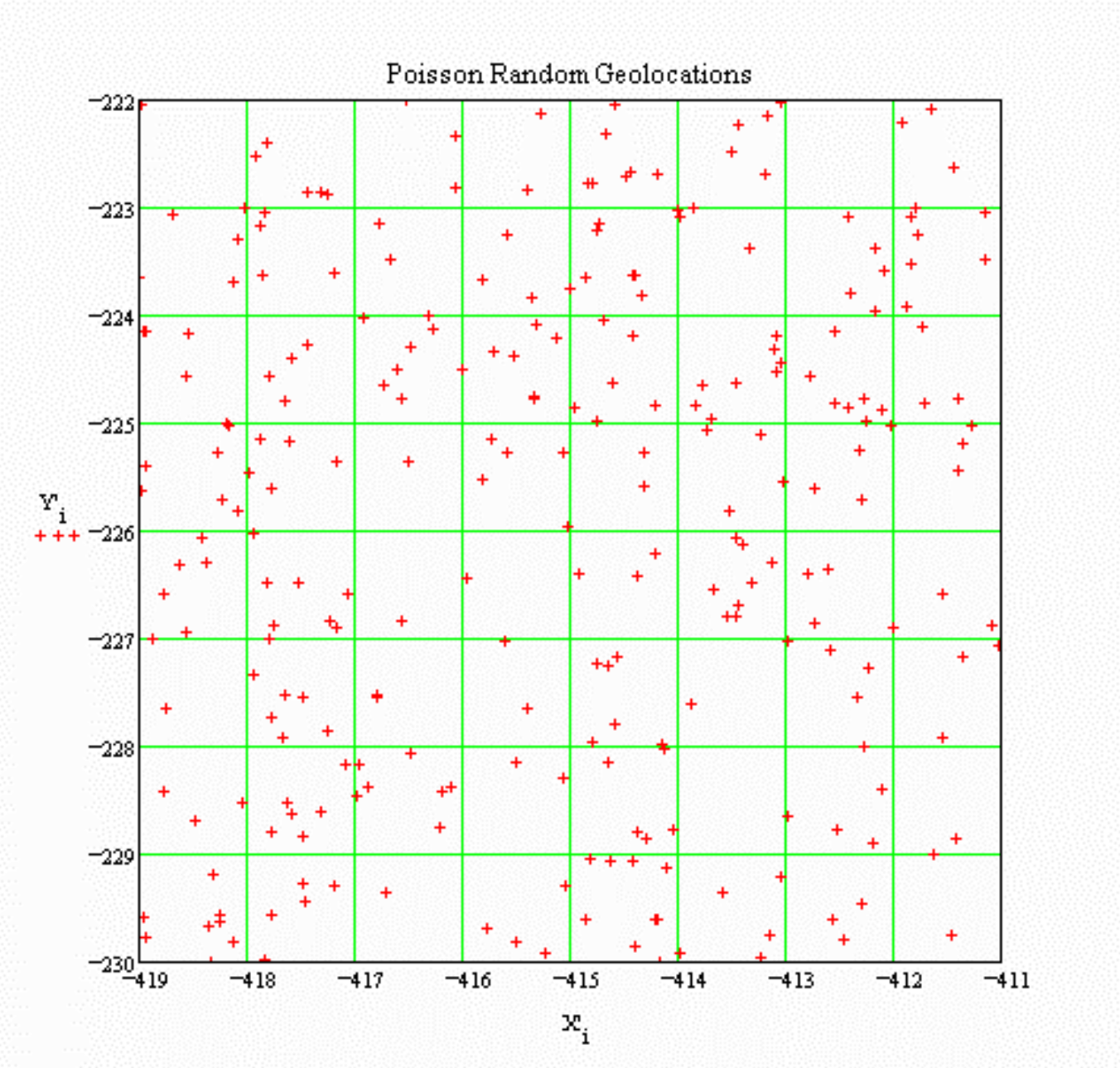

Figure 18: Poisson Simulation of Portfolio

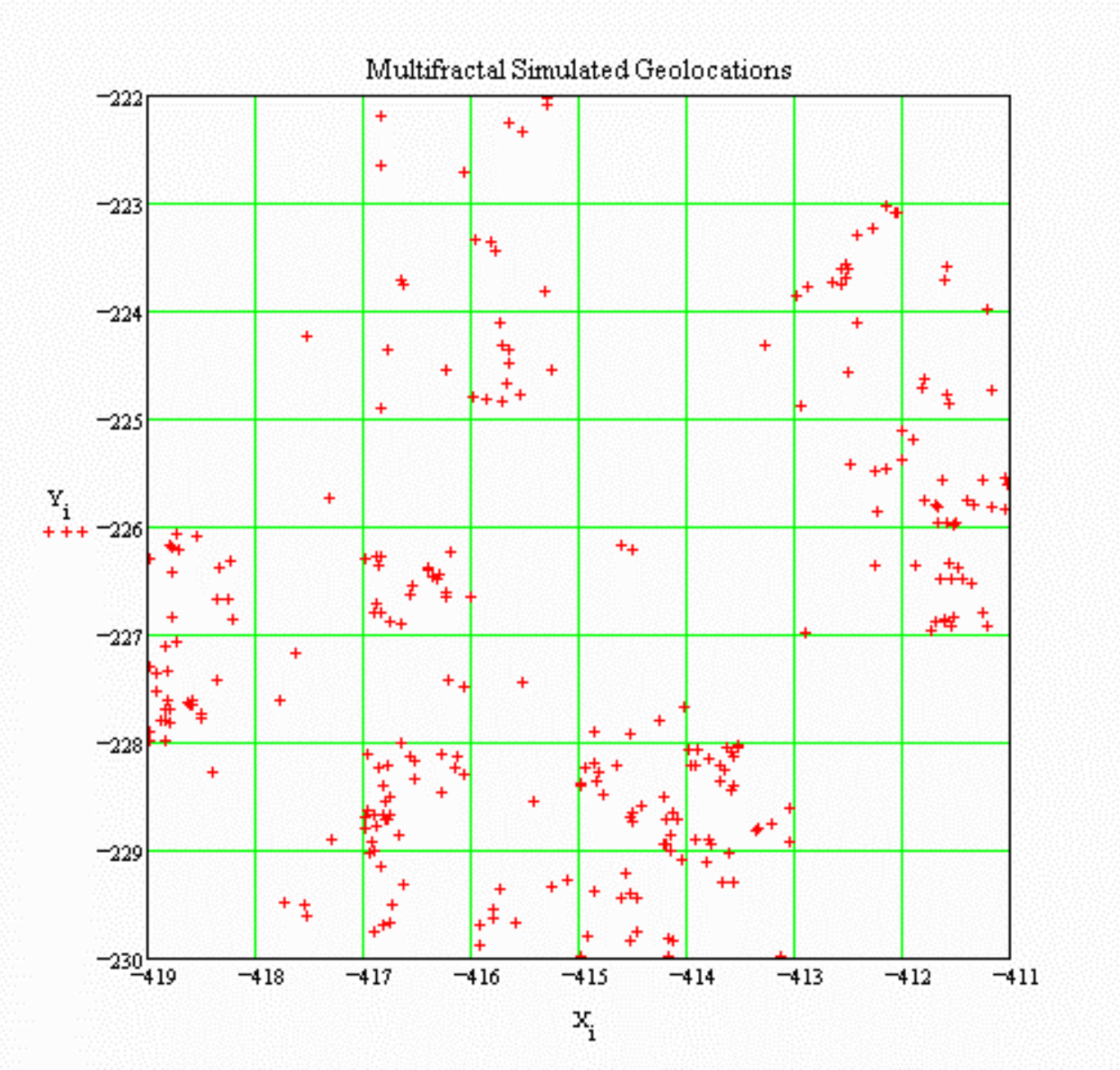

Figure 19: Multifractal Simulation of Portfolio

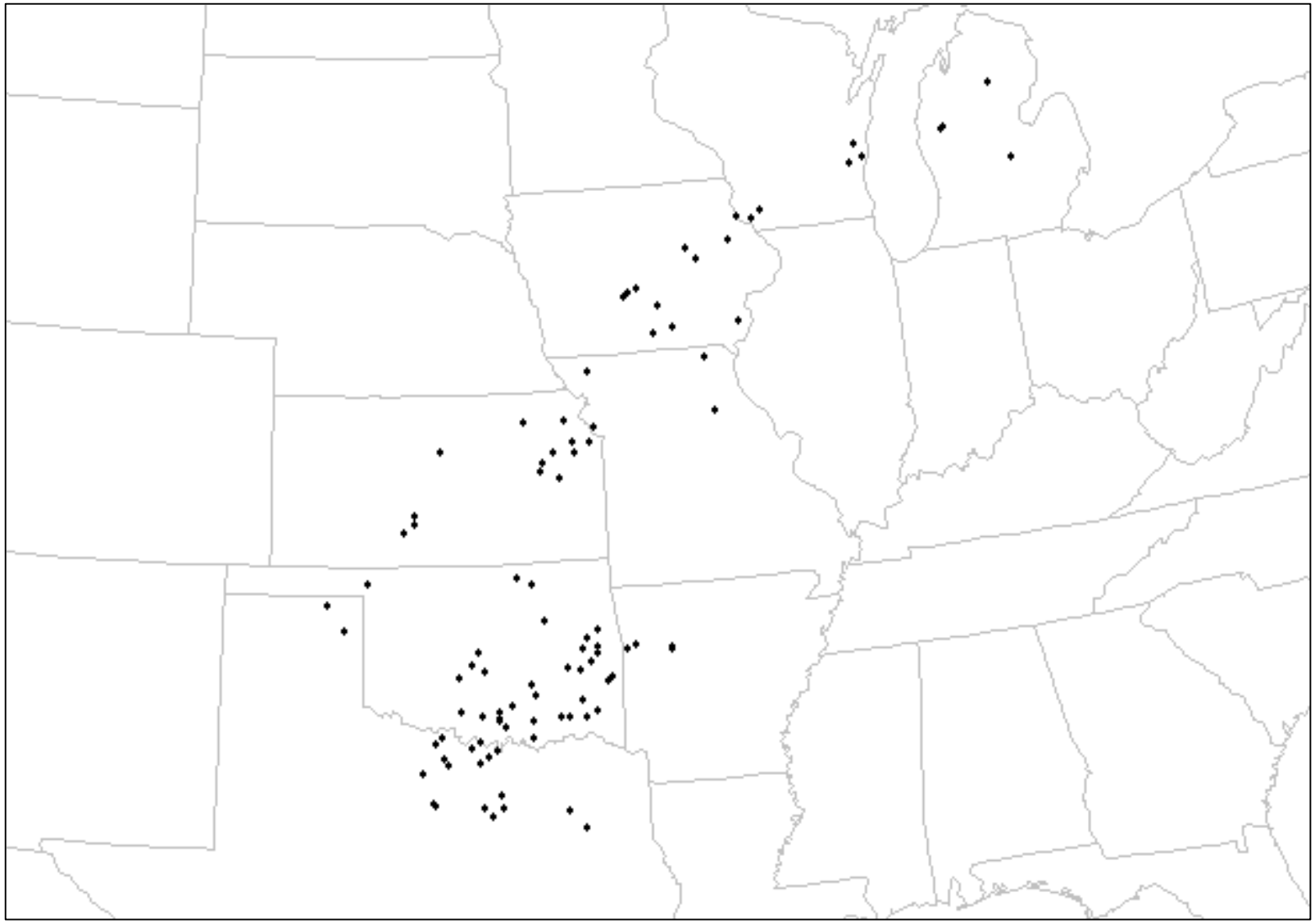

Figure 20: 3/30/98 Hail Reports As Given

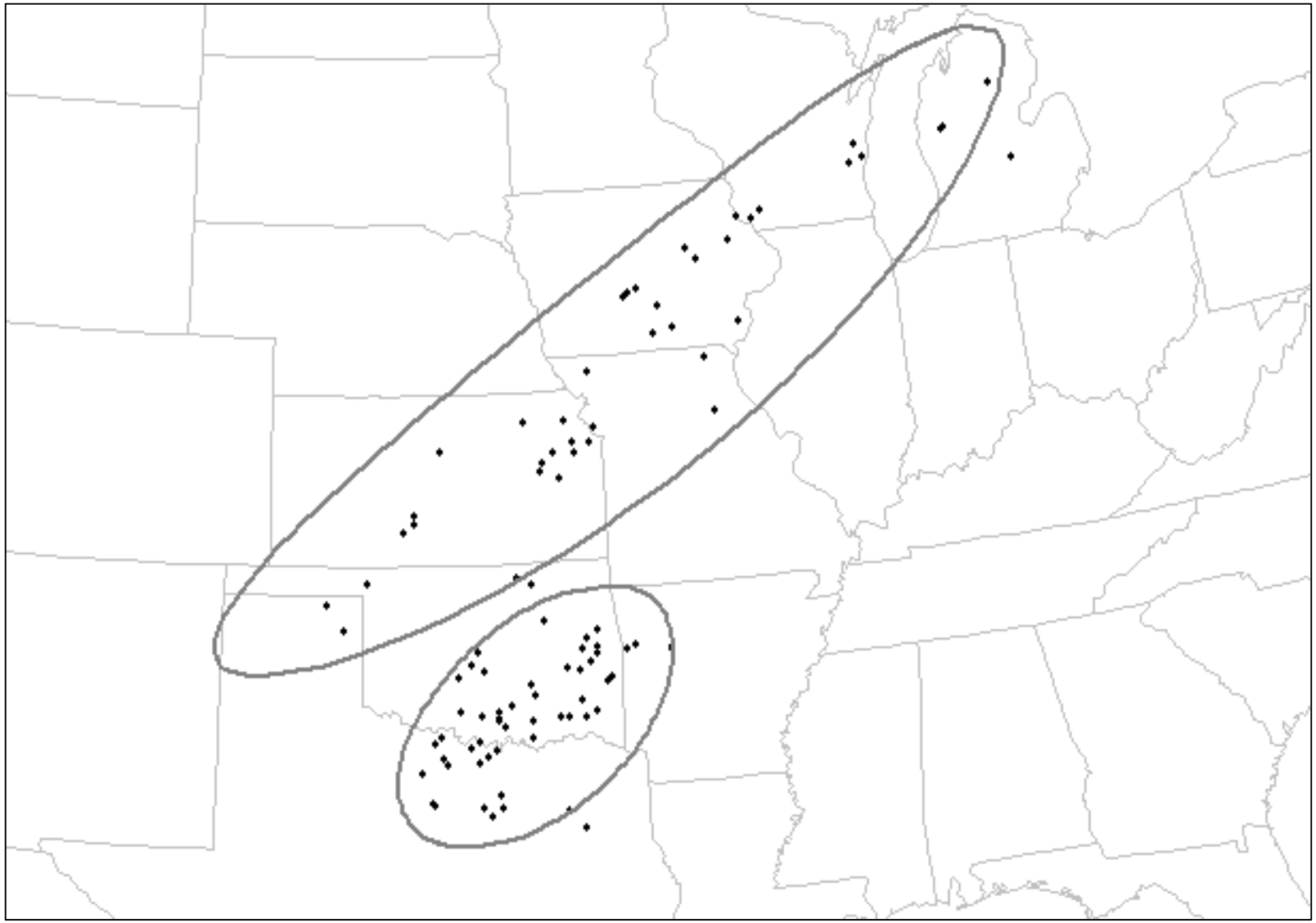

Figure 21: Hail Reports In Swaths

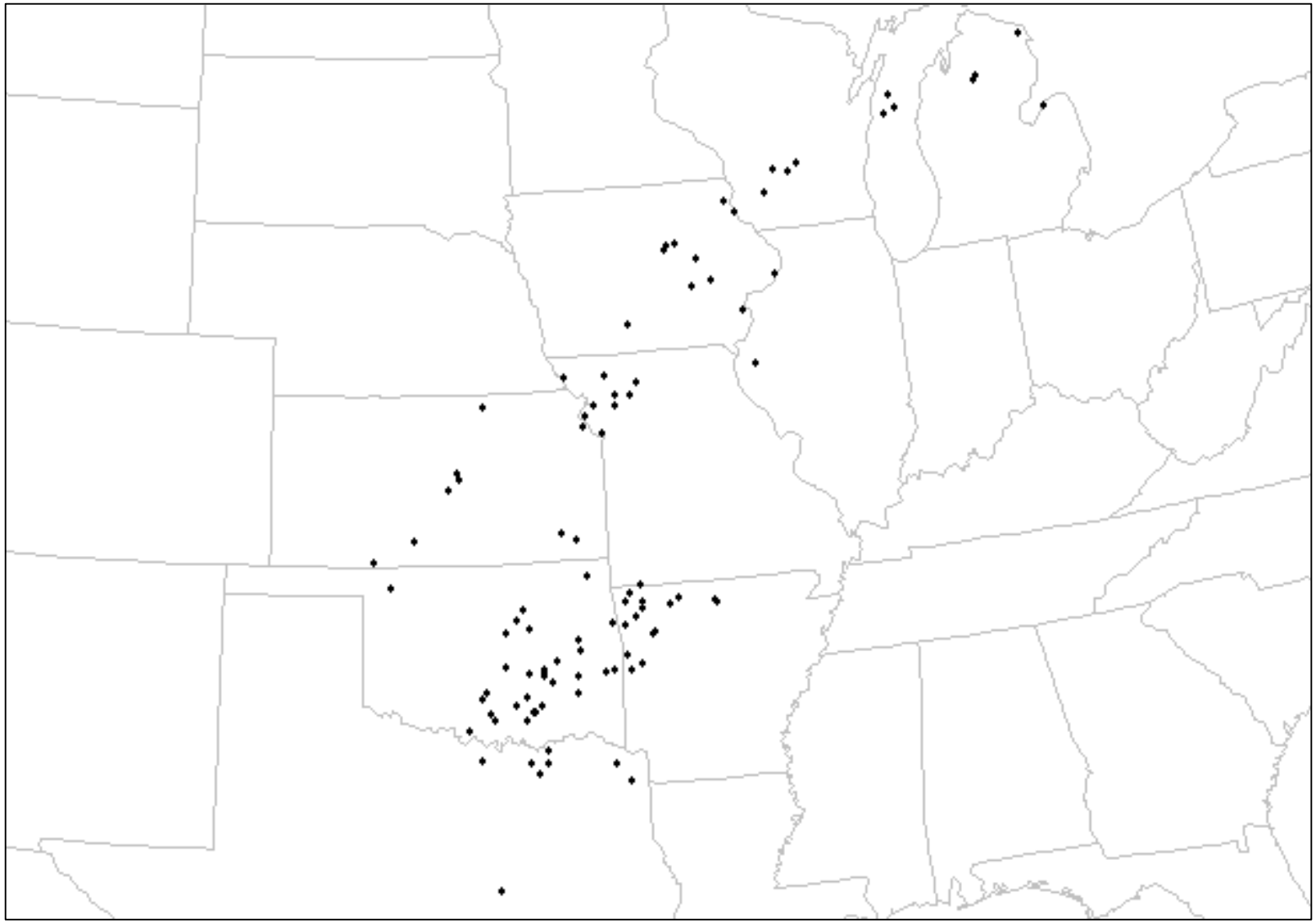

Figure 22: Hail Reports Shifted

Figure 23: Multifractal Simulations of Hail Reports