Landscape Ecology vol. 9 no. 4 pp 271-278 (1994)
SPB Academic Publishing bv, The Hague

A fractal model of vegetation complexity in Alaska

Willem **W.S.** van Hees
U.S. *Department of Agriculture, Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3301 C Street, Suite 200, Anchorage, Alaska* ***99503***

Keywords: Alaska, complexity, fractal, vegetation

Abstract

A methodology using fractals to measure vegetation complexity in three regions of Alaska is presented. Subjective, binomial (0 = simple, 1 = complex) classifications of the complexity of mapped vegetation polygon patterns within continuous forest inventory plots measured in the regions were made by interpreters of aerial photographs. The fractal dimensions of the vegetation patterns within the plots then were estimated. Subsequently, the subjective classifications of the photo-interpreted plots were regressed against fractal dimension by using logistic regression.

Assessment of interobserver agreement among the aerial photo interpreters, by using estimated unweighted Kappa coefficients, indicated substantial classification agreement among observers.

Examination of general versus regional applicability of the logistic models provided strong support for applicability of a single model to all three regions. The logistic model provides numerical identification of the division between simple and complex patterns. Possible applications beyond the needs of the study are discussed.

1. Introduction

Measurement of the irregularity of natural objects, such as shorelines and vegetation edges, has involved the efforts of many researchers. Approaches to this problem have included development of diversity indices, whereby the perimeter of an area is related to the circumference of a circle of equal area (Patton **1975).** Czaplewski and Cost (1985) used a measure of edge dependent on the number of times a circular plot border was intersected by forest and nonforest borders, forest type changes, or stand size changes. These approaches do not provide a direct measure of the complexity, or ''convolutedness,'' of the lines. Rather, they produce indices reflecting either the amount of edge of the object@) under study relative to the amount of edge of some 'standard' object or the abundance of vegetative edges.

Fractal geometry (Mandelbrot **1983)** has provided an alternative framework for describing natural objects. This geometry is based on the observation that the lengths, areas, and volumes of natural objects increase when the precision of measurement increases. The parameter governing this change is the fractal dimension. In Euclidean geometry, dimensions are integers. Fractal dimensions exceed Euclidean dimensions by fractional amounts.

Correspondence to: Willem W.S. van Hees, FAX: **907-271-2898**

Thus, the fractal dimension of a line on a plane can be some value between 1 (Euclidean dimension of a line) and 2 (Euclidean dimension of an area).

Fractal geometry has been used to describe a variety of natural objects and phenomena. Two examples are descriptions of tree crowns by Zeide and Gresham (1991) and land terrain modeling by Polidori, Chorowicz and Guillande (1991) who used fractional Brownian motion as a surface model.

In a separate study, vegetation complexity was hypothesized to be a factor influencing area estimation error arising from imprecise collocation of aerial-photo plots with ground plots. This separate study required that a large number of plots (200–300) be separated into one of two complexity categories: simple or complex. It was felt that aerial-photo interpreters consistently could visually separate complex vegetation patterns from simple ones. An objective measure of vegetation pattern complexity was needed to verify this belief and to provide an objective measure of the point where a simple pattern becomes a complex pattern.

Vegetation complexity, in the construct presented here, is the degree to which the borders making up the vegetation pattern of the plot are convoluted. A large number of small polygons of mapped vegetation classes would not necessarily constitute a complex pattern if the edges of those polygons were not convoluted. Smooth-bordered polygons are, in this framework, 'simpler' than convoluted polygons. In the context of fractal geometry, the fractal dimension of the borders between vegetation polygons would range between 1 and 2. Complex borders would have a fractal dimension closer to 2 than would simple borders. Because of the sensitivity of fractal dimension to the convolutedness of the borders, fractal dimension was chosen as an appropriate measure of complexity.

Three goals were developed to satisfy the needs that prompted this study. The first was to establish whether there was substantial classification agreement among observers. The second was to model the fractal dimension value (classification threshold) below which plots would be classified as simple and above which plots would be classified as complex. The third was to examine whether or not the classifiation threshold was specific to a particular region or was generally applicable to all regions considered.

2. Methods

2.1 Data

Analyses were conducted on plots randomly selected from three regions of Alaska (Fig. 1). In region 1, two independent sets of plots ultimately were used. The first set included 29 plots and the second set included 35 plots. For regions 2 and 3, respectively, 31 plots and 29 plots were used. There was no duplication of plots among the four sets of plots. The aerial-photo plots were circular with an approximate area of 8 hectares. The plots were drawn on color-infrared photographs with nominal scales ranging between 1:3,000 and 1:6,000 and then digitized. The digitized images were then adjusted to equalize scales.

All photo interpretation was accomplished before the beginning of this study. For the plots in region 1, the vegetation classes used for the photo interpretation were simple and broad. They were productive forest (land producing, or capable of producing $1.4 m^3 \cdot ha^{-1} \cdot yr^{-1}$ or more of industrial wood at culmination of mean annual increment [MAI]), nonproductive forest (producing less than $1.4 m^3 \cdot ha^{-1} \cdot yr^{-1}$ at culmination of MAI), nonforest, and water. The vegetation classes used for photo interpretation of the plots in the second and third regions were based on the vegetation classes of Viereck and Dyrness (1980). This latter vegetation classification system involved four vegetation formations (forest, scrub, herbaceous, and tundra) at classification level I and 415 discrete vegetation communities at classification level V (levels II, III, and IV contain 15, 39, and 126 units respectively).

Each of three interpreters of the aerial photographs was given black and white copies of transparent overlays of circular plot vegetation maps produced from aerial photos of the plots. The interpreters then made independent subjective separations of the plots into those they felt had complex vegetation patterns and those they felt had simple vegetation patterns.

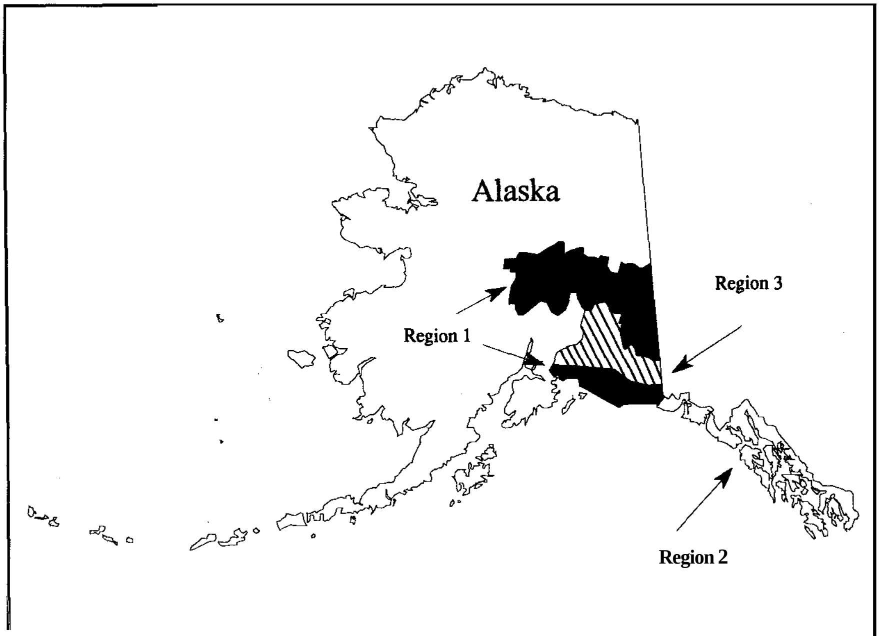

Fig. 1. Regions of Alaska from which aerial photo plots were selected for modeling of vegetation pattern complexity

The fractal dimension of the borders between vegetation classes within each plot was subsequently estimated according to the 'dividers' method of Sugihara and May (1990). Sugihara and May note that the lengths of many natural objects, such as coastlines (or, as applied in this study, the borders between vegetation classes), depend on the measurement scale, **6,** according to the following (over a range of **6** values):

$$L(\delta) = K\delta^{(1-D)}, \quad (1)$$

where: **6** = divider width (scale),
K = constant (scale reduction factor),
L(6) = apparent edge length as estimated at a given scale, and
D = fractal dimension ($2 > D \geq 1$).

Following the dividers method, the length of the vegetation class borders within each plot was stepped off with a pair of drafting dividers set at four different unit lengths: 2.54 cm, 1.27 cm, 0.64 cm, and 0.32 cm. The actual plot circle boundary was not included. **A** linear regression relation was then estimated for log L(δ) versus log δ. The estimated fractal dimension, D, of the vegetation class edges was then 1.0 minus the slope of the estimated linear regression.

2.2 Estimation of classification threshold

The LOGISTIC procedure **(SAS** 1990) was used to fit a linear logistic regression model to the binary classification data by the maximum likelihood method. In this study, the response, Y, could have one of two values, v *(0* = simple pattern or 1 = complex pattern), fractal dimension (D) was the explanatory variable, and $p = \Pr(Y = v|D)$ was the response probability modeled. The logistic model has the form:

$$E(Y) = p = e^{(\beta 0 + \beta 1 (D))}/[1 + e^{(\beta 0 + \beta 1 (D))}] \quad (2)$$

where:

β0 = intercept parameter,
β1 = slope parameter,
D = fractal dimension, and
$0 \leq E(Y) = p \leq 1$.

Applying the logit transformation to linearize (2) produces:

$$\text{logit}(p) = \log_e(p/(1-p)) = \beta0 + \beta1\ (D). \quad (3)$$

For this study, the estimated model was

$$\text{logit}(p) = b, + b_1\ (D), \quad (4)$$

where: b_0 = regression intercept,
$b,$ = regression slope, and
D = fractal dimension.

The fractal dimension value that is the estimated classification threshold is where $p = 0.5$, or where $\text{logit}(p) = 0$.

2.3 Analysis

As noted, the goals of this study were (1) to examine interobserver agreement on classification, (2) to construct models of regional fractal-dimension classification thresholds, and **(3)** to examine regional versus general applicability of the modeled classification thresholds.

To provide an objective measure of the agreement among observers, the estimated unweighted Kappa (κ) coefficients for frequency tables of observer 1 versus observer 2, observer 1 versus observer **3,** and observer 2 versus observer **3** (Agresti 1990) were calculated. The coefficient ranges between 0 and 1 (where 0 indicates any agreement between observers is totally by chance) and is defined as follows:

$$\kappa = (p_o - p_e)/(1 - p_e), \quad (5)$$

where $p_, = \sum_{ij} p_{ij}$,
$p_e = \sum_{ij} p_i.p.j$,
$p\cdot$ = frequency table row totals,
$p_{.j}$ = frequency table column totals, and
p_{ij} = proportion of N items classified into category *i* by the first observer and *j* by the second observer.

and

$$\text{var}(\kappa) = p_o(1 - p_o)/N(1 - p_e)^2.$$

To compare regional logistic models, one 'true' classification for each plot was needed. With three observers, the majority rule was chosen. Whichever classification was chosen by at least two observers was chosen as the correct classification for that plot.

The analysis of the logistic models for regional versus general applicability was conducted in two sequential steps. First, comparability of the slope coefficients of the estimated logistic regressions was investigated. The second step of the analysis depended on results of the first step. If first-step results indicated the slopes of the estimated models were similar, then second-step analysis would examine comparability of intercepts. This analytical approach ensured that if differences between regions did exist, they could be readily identified and incorporated into the model.

These analysis procedures required that an unrestricted model allowing different slopes and intercepts for each region be fit to a composite of the data for all three regions. The logit model formulation for this was:

$$\text{logit}(p) = \text{complexity} = b_0 + b_1 i_1 + b_2 i_2 + b_3 iD_1 + b_4 iD_2 + b_5 D, \quad (6)$$

where: complexity = 0 if vegetative pattern was simple, and
= **1** if vegetative pattern was complex,
D = fractal dimension for plot,
i_1 = 1 if observation was from region 1 and
= 0 otherwise,
i_2 = 1 if observation was from region 2 and
= 0 otherwise,
iD_1 = D if observation was from region 1 and
= 0 otherwise, and
iD_2 = D if observation was from region 2 and
= otherwise.

The first hypothesis addressed in this analysis concerned whether there were any slope differences among the estimated models for the three regions. The null hypothesis was that all three slopes were equal. The test of this hypothesis was accomplished by fitting a model restricted to having a common slope but having different intercepts and comparing chi-square (χ^2) values of the unrestricted and restricted models. The restricted model was:

$$\text{complexity} = b_0 + b_1 i_1 + b_2 i_2 + b_3 D. \quad (7)$$

The difference between χ^2 values for the unrestricted and restricted models was distributed χ^2 with 2 degrees of freedom. The appropriate χ^2 test value (a = 0.05) was 5.99.

The second hypothesis addressed in this analysis concerned whether there were significant differences among estimated regional model intercepts. The null hypothesis was that there were no differences. To test for differences, a procedure similar to that used to test slope differences was used. In this case, the unrestricted model is the restricted model of the first test (equation 7) and the restricted model for this instance is:

$$\text{complexity} = b_0 + b_1 D. \quad (8)$$

The difference between the χ^2 values for the unrestricted and restricted models was distributed χ^2 with **2** degrees of freedom and the appropriate χ^2 test value (α = 0.05) was 5.99.

Results

3.1 Interobserver agreement

Table 1 shows estimated unweighted Kappa coefficients. There is strong indication that significant agreement among observers occurred within each area. At worst, the difference between observed and chance agreement is about **60%** of the maximum possible difference (region 1, observer 2 vs. observer **3).**

Table 1. Estimated unweighted Kappa coefficients by observer pair and region (standard errors in parentheses).

Region	Observer 1 vs observer 2	Observer 1 vs observer 3	Observer 2 vs observer 3
1	0.86	0.71	0.60
	(0.09)	(0.13)	(0.15)
2	0.79	0.65	0.85
	(0.11)	(0.14)	(0.11)
3	0.68	0.81	0.74
	(0.11)	(0.11)	(0.12)

3.2 *Analysis of* estimated models

Initial analysis of the estimated models using the first set of data from region 1 (Table 2) did not provide definitive information to decide whether the estimated regressions were all similar. The difference in χ^2 values for the estimated unrestricted and partially restricted models was 5.289. The difference in χ^2 values for the estimated partially and fully restricted models was 6.156. These values are close to the test value of 5.99 (a = 0.05). The variation in region 1, data set 1 (sd = 0.10865) was higher than for all other data sets. It was felt that this variation contributed to observed ambiguity of results.

Three possible sources of increased variation were hypothesized for the results of region 1. The first possible source was observer lack of familiarity with the study procedure. Initially, there was some confusion regarding convolutedness as the index of complexity rather than mere numbers of polygons. The second possible source of variation was that the vegetation on the plots of region 1 was classified under a system different from that used for regions **2** and **3.** The third possible source was geographic distribution. Region 1 included plots from both interior and coastal Alaska whereas region 2 plots were from the interior and region **3** plots were from southeast coastal Alaska.

A second, independent sample was drawn from region 1 (Table 2). It was felt that if the second estimated model for region 1 proved to be similar to the first estimated model,-then observer learning could be ruled out as a source of variation. Also, there

Table 2. Number of plots and the means, standard deviations, and ranges of fractal dimension estimates by region.

Region	Number of plots	Mean	Standard deviation	Range
$1_{(1)}$*	29	1.20044	0.10865	1.00000 – 1.45490
$1_{(2)}$*	35	1.17382	0.08345	1.04755 – 1.37162
2	29	1.15677	0.08847	1.01155 – 1.41641
3	31	1.13176	0.07910	1.02588 – 1.33370

*Subscripts indicate data sets 1 and 2.

Table 3. Estimated coefficients (with standard errors in parentheses) for restricted and unrestricted forms of the logit model.

Coefficient Coefficients	Fully restricted model (common slopes, common intercepts)	Partially restricted model (common slopes, separate intercepts)	Unrestricted model (separate slopes, separate intercepts)
Intercept	26.1089*	30.4697*	46.3785*
	(5.6111)	(6.4243)	(17.0078)
Fractal dimension	–23.0104*	–27.4035*	–41.7252*
	(4.9408)	(5.7659)	(15.3044)
Region 1 indicator		1.6395*	–18.6481
		(0.7197)	(19.4132)
Region 2 indicator		0.1297	–20.5174
		(0.6747)	(20.0437)
Region 1*3 interaction			18.0778
			(17.2772)
Region 2*3 interaction			18.5409
			(17.9738)
df	1	3	5
Chi-squared (–2 log L)	40.070	47.191	48.731
	p = 0.0001	p = 0.0001	p = 0.0001

Hypothesis tests:

for common slopes:
- C_0: all slope coefficients equal
- C_1: at least two slope coefficients not equal
- χ^2 test value ($\alpha = 0.05$, 2 df) = 5.99
- (48.731 – 47.191) = 1.540; do not reject C_0

for common intercepts:
- C_0: all intercept Coefficients equal
- C_1: at least two intercept coefficients not equal
- χ^2 test value ($\alpha = 0.05$, 2 df) = 5.99
- (47.191 – 40.070) = 7.121; reject C_0

* significant at .05 level

would then be continued questions on the advisability of applying a general model versus one for each region.

For the second analysis, the difference in χ^2 values for the estimated unrestricted and partially restricted models was 1.540 (Table 3). The difference in χ^2 values for the estimated partially and fully restricted models was 7.121. The former value (1.540) indicated that the model form likely is the same from one region to the next. Some question lingers on whether or not the model for region 1 has an intercept different from the others. The marginal nature of the significance of the difference between χ^2 values for the full and partially restricted models does not provide overwhelming evidence that this intercept is indeed different from the

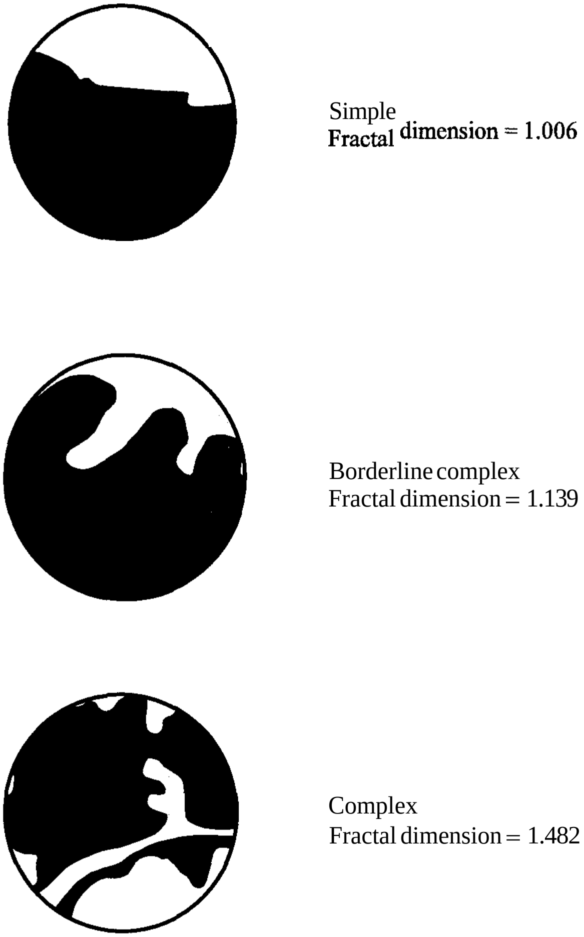

Fig. 2. Examples of aerial photo plot maps representing simple and complex vegetation patterns. Plot maps shown are not original size.

others. Also, it is likely that the observers did improve their understanding of the procedures, thereby reducing the variation in region 1. Questions of geographic range and vegetation classification differences as sources of variation still exist but are not addressed by this analysis.

3.3 Complex versus simple vegetation patterns

Examples of simple and complex vegetation patterns are shown in Fig. 2. A fractal dimension of 1.135 is the modelled value at which a simple pattern becomes **a** complex pattern.

4. Limitations

Application of the dividers method to estimate fractal dimension can produce inaccurate estimates of fractal dimension if the range of scales, 6, is too gross for plots with no vegetation polygons larger than the largest of the unit lengths. In cases where plots have a small polygon, the borders of which are shorter in length than the longest δ-values, fractal-dimension values could be artificially high. This limitation of the dividers method is noted by Sugihara and May (1990).

In this study, the effects of this limitation were reduced in that all plots with vegetation maps had at least one vegetation polygon whose edge could be measured by all four divider settings. Two plots in data set 1 for region 1 had no vegetation polygons (the whole plot was placed in one vegetation class) and therefore were not measured by any divider setting.

5. Conclusion

Complexity (convolutedness) of vegetation class patterns as mapped on aerial photos is a construct that is both readily identifiable subjectively and also objectively quantifiable numerically. The generally high Kappa values are good evidence that the process of subjectively separating plots into complexity categories is reproducible; that is, the separations made by one trained observer will agree well with those made by another trained observer.

Numerical identification of complexity classes was accomplished through logistic modeling of subjective classifications as a function of measured fractal dimension. Tests of estimated model slopes show model form likely was constant among regions. This may indicate that even though differing vegetation classification systems were used (the one being a generalization of the other), the self-sameness characteristic of fractal geometry was evident.

There remains a question as to whether the intercepts of regional models are equal. The intercept for region **1** may be different from the others. The evidence supporting this is not strong. On the basis of the evidence from this analysis, a model estimated from composite data likely would serve as well as models estimated independently for each region.

6. Possible applications

The use of fractal geometry to describe and analyze physical characteristics and processes has found numerous applications since its formulation by Mandelbrot **(1983).** Application of fractal dimension to characterize vegetation complexity could have a variety of uses in natural resources investigations. Some possibilities include:

6.1 Wildlife habitat description

For those species dependent on vegetation edges, it is conceivable there could be cover advantage in areas with many 'bays' or 'inlets' as opposed to areas with long, smooth, highly visible edges. **An** index combining length of edge with fractal dimension as an indicator of convolutedness may more accurately assess value of edge in a given area. To assess the likelihood that a given vegetated region is more or less attractive to wildlife than another region, aerial photo interpreters could classify plots into complexity categories. Ground sampling to determine actual use would follow. Resulting information may be useful for understanding of regional wildlife dynamics.

6.2 Aerial photo interpretation assessment

At certain combinations of vegetation complexity levels and aerial photo scales, it is possible that some information may become unavailable or may be significantly reduced in quality because of increasing inability to interpret well. In complex patterns on small scale photos, for example, a given vegetation classification system may cause photo interpreters to assign improper classification; particularly if the classification system is detailed. The interpreter may miss visual clues that aid classification, such as subtle changes in texture and shading, but are less obvious in complex vegetation patterns.

6.3 Landscape considerations

Complexity of vegetation patterns as seen from various viewing angles could be a determinant in the visual desirability of areas under consideration for recreational development. Users of recreational areas can be surveyed to determine scenic preferences. These preferences could then be compared with options for development.

6.4 Vegetation successional stage

Vegetation patterns change over time as one plant community gives way to another. It is possible that fractal dimension of the vegetation pattern changes predictably. If this were true, then current maturity of a vegetated community might be estimated.

References

Agresti, A. 1990. Categorical Data Analysis. John Wiley and Sons, Inc. New **York,** 558 p.

Czaplewski, R.L. and Cost, N.D. 1985. Photo interpretation of wildlife, recreation, and livestock-grazing variables on one-acre forest plots in South Carolina. *In* Proceedings: Use of Auxiliary Information in Natural Resource Inventories, pp. 24–37. Society of American Foresters Publication SAF 86-01, Blacksburg, Virginia.

Mandelbrot, B. 1983. The Fractal Geometry of Nature. W.H. Freeman and Co. San Francisco, 468 p.

Patton, D.R. 1975. A diversity index for quantifying habitat 'edge.' Wildlife Society Bull. 3(4): 171–173.

Polidori, L., Chorowicz, J. and Guillande, R. 1991. Description of terrain as a fractal surface, and application to digital elevation model quality assessment. Photogrammetric Engineering and Remote Sensing 57(10): 1329–1332.

SAS. 1990. SAS Technical Report P-200, SAS/STAT Software: CALIS and LOGISTIC Procedures, Release 6.04, SAS Institute, Inc., Cary, NC. 236 p.

Sugihara, G. and May, R.M. 1990. Applications of fractals in ecology. Trends in Ecology and Evolution, Elsevier Science Publishers Ltd., (UK), 5(3): 79–86.

Viereck, L.A. and Dyrness, C.T. 1980. A preliminary classification system for vegetation of Alaska. Gen. Tech. Rep. PNW-106, U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, OR. 38 p.

Zeide, B. and Gresham, C.A. 1991. Fractal dimensions of tree crowns in three loblolly pine plantations of coastal South Carolina. Canadian Journal of Forest Research 21: 1208–1212.