Published September, 1998

DIVISION S-5—PEDOLOGY (NOTES)

USE OF STATISTICAL TESTS TO DESCRIBE THE BASIC DISTRIBUTION PATTERN OF IRON OXIDE NODULES IN SOIL THIN SECTIONS

J. M. AROCENA* AND J. D. ACKERMAN

Abstract

Soil formation is often associated with the spatial reorganization of soil components, particularly Fe oxides. However, the methods available to describe the spatial distribution of Fe oxides generally lack measurable parameters that allow comparison between different types of soils. In this work, we demonstrate the use of statistical methods (e.g., χ^2 goodness-of-fit test, mean, variance, and cluster coefficient) to describe the basic distribution pattern of Fe oxide nodules in soil thin sections. Using a computer, we mapped the spatial distribution (i.e., x and y coordinates) of Fe oxide nodules in six soil thin sections from the B horizons of three sandy soils from Alberta, Canada. We calculated the statistical tests at a range of grid sizes (2–10 mm) with a Fortran program to (i) examine the effect of grid size on the statistical description of the distribution and (ii) determine the "appropriate scale of observation" for each thin section. We defined this scale as the grid size at which the mean number of observation is at least five, which was found to vary in size from 5 to 9 mm in our samples. We recommend the determination of the appropriate scale of observation before any statistical description of the Fe nodule distribution. In all but one of our six samples, the statistical distribution of nodules was clustered; the remaining one was uniform in distribution. The methods described in this study provide statistical methods whereby results from different studies can be compared.

THE SPATIAL PATTERNS of soil features are commonly described in terms of their (i) basic, (ii) related, and (iii) referred distributions (Brewer, 1976; Bullock et al., 1985). The *basic distribution* refers to the arrangement of soil features with respect to each other (e.g., among Fe oxide nodules), which can be classified as uniform, random, and clustered (Brewer, 1976; Bullock et al., 1985). The *referred distribution* is the arrangement of soil features with respect to specific reference features such as planar voids or mineral grains. The *related distribution* pattern is a referred distribution of the coarse and fine fractions where the size limit for the size fractions varies from one thin section to another (Bullock et al., 1985). Of particular interest here are the ubiquitous Fe oxides concentrated in coatings, mottles, and nodules. The purpose of this study was to demonstrate the use of statistical methods (i.e., χ^2 goodness-of-fit tests, variance, mean, and cluster coefficient) to describe the basic distribution pattern of Fe nodules in soil thin sections, including a recommendation on the determination of the optimum area of observation.

Faculty of Natural Resources and Environmental Studies, Univ. of Northern British Columbia, Prince George, BC, V2N 4Z9, Canada. Received 27 May 1997. *Corresponding author (arocenaj@unbc.ca).

Published in Soil Sci. Soc. Am. J. 62:1346–1350 (1998).

Materials and Methods

Soil Thin Sections, Identification, and Mapping of Iron Oxide Features

The location, field morphological descriptions, and properties of the three sandy soils in the study are described in Arocena et al. (1992). The soils are transitional between Alfisol (Luvisolic) and Spodosol (Podzolic) and have Fe oxide nodule concentrations in the B horizon (Table 1). Two thin sections of ≈30 μm thick and 8 by 5 cm from each of the B horizons of the three soils were used in this study. Mapping the spatial distribution of Fe nodules involved displaying the digitized image on a 50-cm video monitor (0.31-mm dot pitch resolution) using graphics software (Corel Draw 4.0, Ottawa, Canada) and manually recording the location for each nodule. The spatial coordinates for each nodule were noted but no attempt was made to record the size of the nodules.

For the statistical analysis, the number of Fe oxide nodules in each frame (the area delineated between grids imposed on the thin section), was tabulated following the "forbidden line rule", which prevents overlap and gaps in enumeration of objects along two adjacent edges of each frame (Ringrose-Voase, 1994). Nodules that fell completely or partially within a frame were counted unless they touched the "forbidden line" (Fig. 1).

Models of Basic Distribution Patterns

Uniformity. If the Fe nodules are uniformly distributed, equal-sized subareas (frames) in the soil thin section should contain an equal number of nodules. This hypothesis can be evaluated by comparing the observed number of nodules per frame (O) with the expected number of nodules (E):

$$E = m/T \qquad [1]$$

(where m is the total number of Fe nodules and T is the number of frames) using a χ^2 goodness-of-fit test,

$$\chi^2 = \sum \frac{(O - E)^2}{E} \qquad [2]$$

The test has $T - 2$ degrees of freedom. For Sample 316B presented in Fig. 1, there are 230 nodules and 40 frames, each with an area of 49 mm^2 (i.e., grid size of 7 mm), and therefore an E of 5.75 (= 230/40). Using Eq. [2], the calculated χ^2 value is 146, and $v = 38$ is significant (i.e., $P < 0.05$) (Table 1). We must reject the null hypothesis that the values are uniform in their distribution. This result does not, however, provide information on the nature of the nonuniformity, which requires comparison with other spatial models (e.g., random and clustered).

Randomness. A random pattern is created when there is an equal probability of finding nodules in similar-sized subareas and their placement is statistically independent (Davis,

Table 1. Description, χ^2 values (and degrees of freedom), variance (s^2), mean (m/T), and cluster coefficient (k) for Fe oxide nodules in soil thin sections of B horizons of sandy soils from Alberta at grid size where the minimum number of nodules is five or more. The variance mean ratio (s^2:m/T) should be less than one for uniform, equal for random, and greater than one for clustered distribution patterns.

Sample	Description	Grid size	Uniform	χ^2 Random	Clustered	s^2	m/T	k
113	Bw, 14–21 cm, Loamy sand, single grain	5	186***(75)	204***(10)	19ns(14)	16	6.5	4
115	Bt, 28–35 cm, Loamy sand, single grain	6	117***(58)	52***(9)	9ns(11)	12	6.0	6
213A	Bt, 14–21 cm, Sandy loam, weak granular to single grain	9	53***(22)	26***(7)	7ns(8)	12	5.3	4
215	BC, 35–42 cm, Sand, single grain	5	79ns(75)	†	†	6	5.5	128
316A	Bt, 14–21 cm, Sand, weak granular to single grain	5	122***(58)	131***(9)	21ns(12)	13	6.3	6
316B	Bw, 21–28 cm, Sand, single grain	7	146***(38)	21**(8)	10ns(11)	22	5.8	2

*,**,*** Significant at the 0.05, 0.01, and 0.001 probability levels, respectively; ns is $P > 0.05$.
† Not applicable (i.e., uniformly distributed).

1986). The Poisson distribution for a random pattern (Davis, 1986) assumes that the probability (P) of finding nodules in a thin section is given by

$$P_r = \exp - E(E^r/r!) \qquad [3]$$

where r is the number of nodules in a frame. The expected number, E_r, of frames with r nodules is

$$E_r = TP_r \qquad [4]$$

The number of frames with r nodules (0, 1, 2, 3... nodules observed per frame) is then counted and compared with the expected number of frames that contain that r number of nodules calculated using Eq. [3]. So for Sample 316B (Fig. 1), the probability of finding zero number of nodules in a frame is calculated as

$$P_0 = \exp - (230/40)\left[\frac{(230/40)^0}{0!}\right] = 0.0031 \qquad [5]$$

while the expected number of frames with zero nodules is

$$E_0 = 40(0.0031) = 0.1273 \qquad [6]$$

The probability of finding a subarea with r number of nodules for Sample 316B (see Fig. 1) is provided in Table 2.

A statistical evaluation of the data is undertaken using χ^2 with $c - 2$ degrees of freedom, where c is the number of density class (e.g., 0, 1, 2, 3.... nodules per frame). The expected number of frames for each density classes should be at least one for the proper evaluation of the χ^2 test (Snedecor and Cochran, 1989) and if not, the E_r should be combined. If the χ^2-test statistic exceeds the critical value (i.e., $P < 0.05$), we reject the null hypothesis of a Poisson distribution (i.e., random) and conclude that the nodules are not distributed randomly in the soil thin section. It should be reiterated that this test cannot be applied if the sample data are determined to be uniform in distribution.

Clustering. If the distribution is neither uniform nor random, a negative binomial probability distribution can be used to evaluate whether the distribution of Fe oxide nodules is clustered in soil thin sections. The model adapted from Davis (1986) is a combination of the Poisson and logarithmic distributions, which assume that the nodules are randomly clustered and that the individual nodules within a cluster follow a logarithmic distribution. The probabilities of finding a frame with zero or r nodules are given, respectively, by

$$P_0 = \frac{1}{(1 + p)^k} \qquad [7]$$

$$P_r = \left[\frac{(k + r - 1)\left(\frac{p}{1 + p}\right)}{r}\right][P(r - 1)] \qquad [8]$$

where, $P = E/k$ is the probability that a frame contains nodules at density of r, $k = E^2/(s^2 - E)$ is a measure of degree of clustering of the nodules, and $s^2 = [\Sigma_{i=1}^{T}(r_i - E)^2]/(T - 1)$ is the variance in the number of nodules per frame. The distribution of nodules approaches a random pattern if k is large, and is clustered if k approaches zero (Davis, 1986; Krebs, 1989). Estimates of k, p, and s^2 for Sample 316B (Fig. 1) are,

$$k = \frac{5.75^2}{21.55 - 5.75} = 2.09 \qquad [9]$$

$$p = \frac{5.75}{2.09} = 2.75 \qquad [10]$$

$$s^2 = \frac{840}{39} = 21.55 \qquad [11]$$

The probability of finding a frame with zero and one nodule in these data using Eq. [7] and [8] are, respectively

$$P_0 = \frac{1}{(1 + 2.75)^{2.09}} = 0.063 \qquad [12]$$

$$P_1 = \left[\frac{(2.09 + 1 - 1)\left(\frac{2.75}{1 + 2.75}\right)}{1}\right]0.063 = 0.097 \qquad [13]$$

The expected number E_r, of frames with r nodules is determined as per Eq. [4].

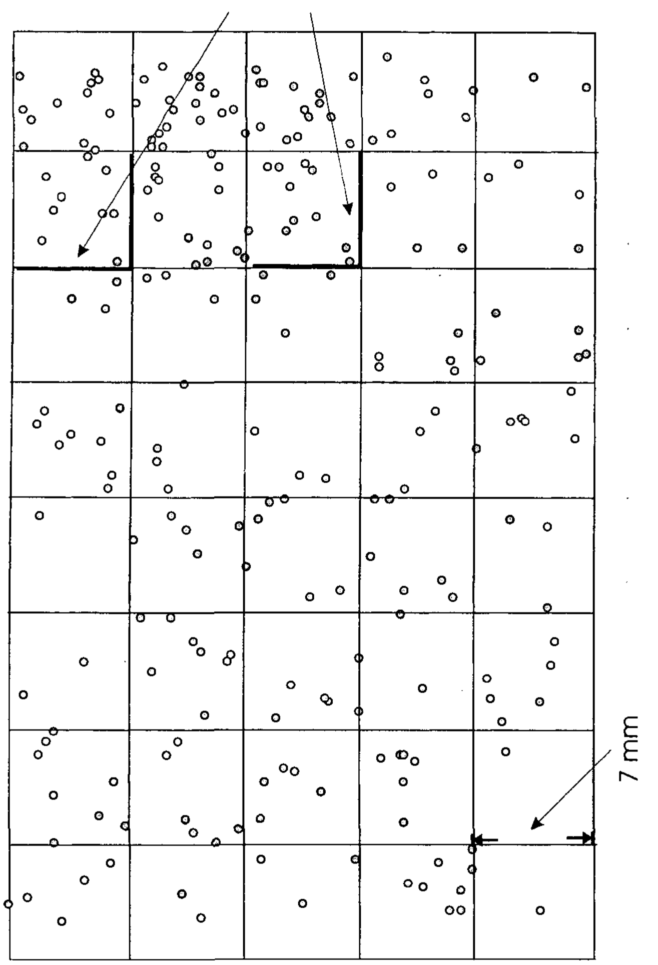

Fig. 1. Spatial distribution of Fe oxide nodules in soil thin section (4.0 by 5.6 cm) of Sample 316B on an 8 by 5 grid system (grid size = 7.0 mm) showing the forbidden lines. The forbidden line-rule states that any nodules touching the forbidden lines are not included in the tabulation within that frame (frame area = 49 mm²).

The probability of finding a subarea with r number of nodules using a negative binomial model for Sample 316B (Fig. 1) is provided in Table 2. This probability can be evaluated by a χ^2-test with $c - 2$ degrees of freedom. If the χ^2 test statistic exceeds the critical value (i.e., $P < 0.05$), we reject the null hypothesis of a negative binomial (i.e., clustered) distribution and conclude that nodules are not distributed in clusters in the soil thin section. This test cannot be applied if the sample data are determined to be uniform or random in their distribution.

Grid Resolution and Spatial Distribution

The uniformity of an area in composition and distribution of any feature (e.g., Fe oxide nodule) will depend on the magnification or scale of observation (Moran, 1994; FitzPatrick, 1993). In order to compare the distribution of Fe oxides on various thin sections, an appropriate scale of observation (i.e., grid size) must be followed. Following the recommendation from Davis (1986), we propose an "appropriate scale of observation" as the grid size at which the mean number of observation is at least five. To facilitate the determination of this scale, we wrote a Fortran program to calculate the statistical parameters mentioned above at different scales of resolution. The program calculates the expected value (m/T), variance (s^2), Poisson and negative binomial probabilities, p and k, as well as χ^2 goodness-of-fit test for uniform, random, and clustered models. We calculated the statistical parameters for nine different grid sizes (2–10 mm at 1-mm increments) for the six thin sections described in this study.

Results and Discussion

Effect of Frame Area on the Description of the Distribution Pattern of Iron Nodules

Generally, the mean (m/T) and variance (s^2) increase from the smallest to largest grid sizes (Fig. 2). The increase in s^2 is expected as the degree of heterogeneity in soil increases with the scale of observation. The s^2 to m/T ratio increases with increasing grid size (frame area) and suggests that the clustered pattern is more pronounced for larger frame areas. With the exception of Sample 215, the clustering coefficient (k) values were around a value of 10 and show a decreasing trend with increasing grid size, further suggesting a clustered pattern (Fig. 2). For Sample 215, grid sizes of 4 and 5 mm may represent some spatial discontinuities where the imposed grid sizes caused the unusually low and high k values, respectively (Fig. 2), through a dramatic change in variance (see Eq. [9]). These discontinuities may lead one to interpret the low k value in Sample

Table 2. Probability (P_r) of finding r number of nodules in the frame and expected number (E_r) of frames with r number of nodules in Poisson (i.e., random) and negative binomial probability distributions (i.e., clustered) for Sample 316B with frame area of 49 mm² (data in Fig. 1 and equations in text).

R	P_r (Poisson)	E_r (Poisson)	P_r (Negative binomial)	E_r (Negative binomial)
0	0.0032	0.127	0.0630	2.52
1	0.0133	0.73	0.0967	3.87
2	0.0526	2.10	0.1096	4.38
3	0.1008	4.03	0.1096	4.38
4	0.1450	5.80	0.1023	4.09
5	0.1667	6.67	0.0914	3.66
6	0.1598	6.39	0.0792	3.17
7	0.1312	5.25	0.0671	2.69
8	0.0943	3.77	0.0560	2.24
9	0.0603	2.41	0.0460	1.84
10	0.0346	1.39	0.0374	1.50
11	0.0382	1.59	0.0302	1.21
12	–	–	0.1115	4.46
Total	1.00	40.3	1.00	40.0

215 as a clustered distribution for a grid size of 4 mm, or the high k value as nonclustered for a grid size of 5 mm. We do not have any explanation for these discontinuities, but their existence suggests that the distributions are not similar at different scales. The nature of the Fe oxide nodule distribution in the sample is one possible origin for this result. Nonetheless, this result supports the need for varying scales in the analyses of spatial data.

The dependence of the distribution on grid size is consistent with the earlier report of Moran (1994) and FitzPatrick (1993). This is probably associated with the natural heterogeneity of the soil systems, which varies depending on the magnification and methods of observation (FitzPatrick, 1993). For practical purposes, Davis (1986) suggested that a grid size with a minimum of five observations should be used in this type of analysis. In this study, we refer to this grid size as the "appropriate scale of observation". The results of determining this scale to elucidate the Fe nodule distribution in various thin sections are given in the next section.

Description of Spatial Distribution of Iron Nodules Using χ^2 Test

Our results indicated that the "appropriate scale of observation" (i.e., at least five observations per frame) varied from 5 to 9 mm on different soil thin sections (Table 1). Our results indicate that the m/T ranged from five to seven nodules per frame (49-mm^2 area), while the s^2 ranged from 6 to 22 nodules per frame (Table 1). The results of χ^2-goodness-of-fit test for uniform distribution using Eq. [2], which compared the observed number of nodules with the expected m/T, indicated that the nodules in Samples 113, 115, 213A, 316A, and 316B were nonuniform in their distribution ($P < 0.05$) (Table 1). The nonuniform nodule distribution is also indicated by the consistently higher s^2 than m/T values

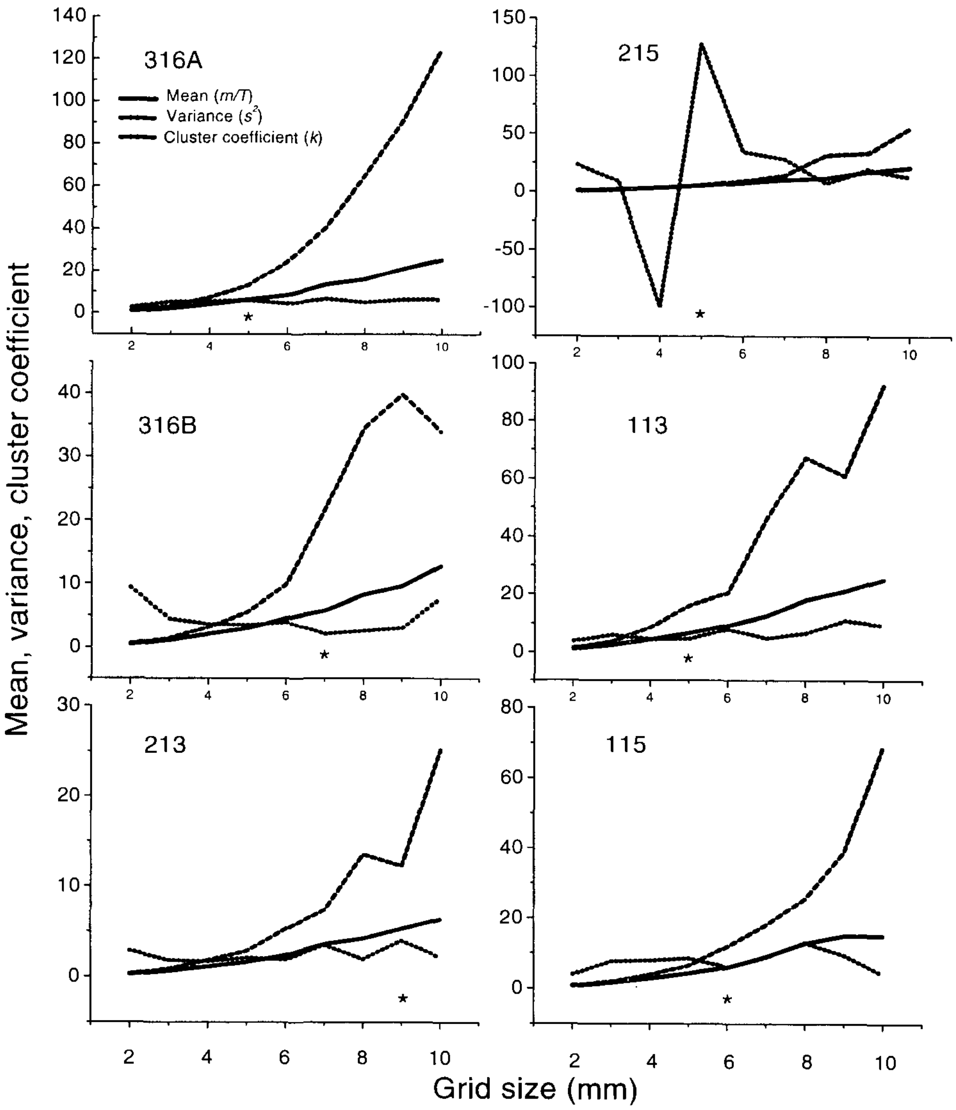

Fig. 2. Change in mean (m/T), variance (s^2), and cluster coefficient (k) with increasing grid size. Note the differences in scales in the ordinate among samples. The variation in k for Sample 215 represents a spatial discontinuity where the imposed grid system drastically affects the s^2 estimates. The asterisk (*) above the abscissa indicates the appropriate scale of observation (i.e., minimum of five observations per frame).

for these samples (Davis, 1986). The χ^2 values for Sample 215 were small ($P > 0.05$), and the null hypothesis of a uniform nodule distribution was accepted (Table 1). This was supported by a slightly higher s^2 than m/T values, which is indicative of a uniform nodule distribution pattern (Davis, 1986).

The χ^2 goodness-of-fit test for a random distribution (i.e., Poisson distribution) could only be applied to Samples 113, 115, 213A, 316A, and 316B, as Sample 215 had a uniform distribution of Fe oxides. Our results indicated that the distribution of the nodule was nonrandom as the large χ^2 values for these five samples lead to the rejection of the null hypothesis ($P < 0.05$). The ratio of s^2 and m/T in these samples was not unity, which further indicates that the distribution is nonrandom (Davis, 1986). Given that these distributions were not uniform nor random, we applied a χ^2 goodness-of-fit test for clustered distribution (i.e., negative binomial model). We could not reject the hypothesis of a clustered distribution as the χ^2 values were not significant ($P > 0.05$). This conclusion is also supported by the much higher s^2 to m/T ratio (i.e., >2:1) and low k values (i.e., <6), which approach zero when clustering is present (Davis, 1986).

The methods described in this study have important applications to soil science, particularly soil micromorphology. Statistical tests provide a measure to examine the spatial reorganization of soil components (e.g., Fe oxide nodules) during soil formation. These methods also provide comparable parameters for the comparison of the distribution of soil features between different types of soils.

References

Arocena, J.M., S. Pawluk, and M.J. Dudas. 1992. Genesis of selected sandy soils in Alberta, Canada as revealed by microfabric, leachate and soil composition. Geoderma 54:65–90.

Brewer, R. 1976. Fabric and mineral analysis of soils. Robert E. Krieger Publ. Co., Huntington, NY.

Bullock, P., N. Fedoroff, A. Jongerius, G. Stoops, T. Tursina, and U. Babel. 1985. Handbook for soil thin section description. Waine Research Publ., Wolverhampton, UK.

Davis, J.C. 1986. Statistics and data analysis in geology. 2nd ed. John Wiley & Sons, New York.

FitzPatrick, E.A. 1993. Soil microscopy and soil micromorphology. John Wiley & Sons, Toronto.

Krebs, J.C. 1989. Ecological methodology. Harper Collins Publ., New York.

Moran, C.J. 1994. Image processing and soil micromorphology. p. 459–482. *In* A.J. Ringrose-Voase and G.S. Humphreys (ed.) Soil micromorphology: Studies in management and genesis. Proc. IX Int. Working Meeting on Soil Micromorphology. Townsville, Australia. July 1992. Developments in Soil Science 22. Elsevier, New York.

Ringrose-Voase, A.J. 1994. Some principles to be observed in the quantitative analysis of sections of soil. p. 483–493. *In* A.J. Ringrose-Voase and G.S. Humphreys (ed.) Soil micromorphology: Studies in management and genesis. Proc. IX Int. Working Meeting on Soil Micromorphology, Townsville, Australia, July 1992. Developments in Soil Science 22. Elsevier, New York.

Snedecor, G.W., and W.G. Cochran. 1989. Statistical methods. Iowa State Univ. Press, Ames, IA.