



Pedodiversity in the United States of America

Yinyan Guo, Peng Gong, Ronald Amundson*

Division of Ecosystem Sciences and Center for Assessment and Monitoring of Forest and Environmental Resources (CAMFER), 151 Hilgard Hall, College of Natural Resources, University of California, Berkeley, CA 94720-3110, USA

Received 8 July 2002; accepted 17 March 2003

Abstract

Little attention has been paid to analyses of pedodiversity. In this study, quantitative aspects of pedodiversity were explored for the USA based on the State Soil Geographic database (STATSGO). First, pedodiversity indices for the conterminous USA were estimated. Second, taxa–area relationships were investigated in each soil taxonomic category. Thirdly, differences in pedodiversity between the USDA-NRCS geographical regions were compared. Fourth, the possible mechanisms underlying the observed relative abundance of soil taxa were explored. Results show that as the taxonomic category decreases from order to series, Shannon's diversity index increases because taxa richness increased dramatically. The relationship between the number of taxa (S) and area (A) is formulated as $S=cA^z$. The exponent z reflects the taxa-richness of soil 'communities' and increases constantly as taxonomic categories decrease from order to series. The "West" USDA-NRCS geographical region has the highest soil taxa richness, followed by the "Northern Plains" region. The "South Central" region has the highest taxa evenness, while taxa evenness in the "West" region is the lowest. The "West" or the "South Central" regions have the highest overall soil diversity in the four highest taxonomic categories, while the "West" or "Northern Plains" regions have the highest diversity in the two lowest taxonomic levels. The high diversity index in the "West" region results from high taxa richness while the high diversity index in the "South Central" region results from an evenness of taxa. As the taxonomic level decreases, the pattern of taxa abundance approaches a lognormal distribution. One of the key findings of this research is that at lower levels of soil taxonomic divisions (especially the series level), soil taxa increase continuously with increasing area, indicating considerable soil endemism in the USA (and likely around the world), a key consideration in conservation and preservation planning.

© 2003 Published by Elsevier Science B.V.

Keywords: Soil; Soil survey; Taxonomic diversity; Taxa–area relationships; Abundance models

1. Introduction

Biodiversity has become a major topic in ecological research in the past few decades (Gaston and Spice, 1998). Studies on biodiversity have focused on

the aboveground flora and fauna, though in recent years, underground biodiversity has been receiving attention (Copley, 2000). However, few studies have considered pedodiversity (i.e. soil diversity), which is a basis for beginning to characterize regional biodiversity given the strong relationship between soils and flora and fauna (Major, 1951; Jenny, 1980).

The soil mantle has the distinguishing characteristic of continuous change (variability) in both horizon-

* Corresponding author. Fax: +1-510-643-5098.

E-mail address: earthy@nature.berkeley.edu (R. Amundson).

tal and vertical dimensions (Mausbach and Wilding, 1991; MacBratney, 1992). Soil taxonomy is based on breaking the continuum into discrete geographical segments of restricted ranges in soil properties, and therefore is a potentially useful means of isolating the features relevant to supporting regional biodiversity. Studies on the relationship between pedodiversity and biodiversity have been hampered due to the limitation of geo-referenced GIS data for both pedodiversity and biodiversity across large geographic areas. Pedodiversity studies were first started by analyzing soil series–area relationships (Beckett and Bie, 1978). The term “pedodiversity” was developed by McBratney (1992) who discussed landscape preservation strategies based on pedodiversity. Recently, examinations of pedodiversity using indices commonly used to characterize bio-diversity have been made (Ibáñez et al., 1995, 1998). In these studies, the FAO soil map of the world was used as the soil database, a system that contained only 26 soil types for the world when the analysis was conducted. This analysis revealed that pedodiversity in the continents of the world was characterized by similarity, and that taxa richness did not increase as the area studied increased. More research on pedodiversity at finer scales of resolution needs to be carried out in order to better understand how soil resources are distributed geographically (Yaalon, 1998).

Generally, there are two ways to characterize biodiversity: (1) develop diversity indices and (2) study the patterns of the relative abundance of species. Diversity of a community depends on two components: the number of species (or species richness) and the evenness of the species—i.e. how the individuals are apportioned among the species present. Usually, three types of biodiversity indices (richness, evenness, and diversity taking both richness and evenness into account) are considered. While a biodiversity index characterizes the diversity of a sample or a community by a single number, only a detailed study on the pattern of quantitative abundance will reveal the mechanism on how a community is formed or characterized. Many species abundance models have been proposed (MacArthur, 1957, 1960; Williams, 1964; May, 1975, 1986; Hughes, 1986; Tokeshi, 1990, 1993; Bell, 2001), although only four models have repeatedly been highlighted in the ecological literature. These models are the MacArthur broken-stick model, geometric-series models, log-series model, and

the lognormal model (Magurran, 1988; Tokeshi, 1990; Ibáñez et al., 1998). The first two are niche-orientated models, i.e. based on ecological environments required by the species studied while the last two are statistically orientated ones, i.e. based on statistical methods.

Pedodiversity can be studied in the same way and with the same abundance models as biodiversity (Ibáñez et al., 1995). Previous research has outlined the nature of the biodiversity models and has applied them to the distribution of FAO soil taxa at continental scales (Ibáñez et al., 1998). The results of this work shows that of the four most used species abundance models, the soil taxa in North America fit the lognormal model while the soil taxa in other continents fit broken-stick or lognormal models (Ibáñez et al., 1998). However, a fundamental problem interpreting the analyses so far is the coarse nature of the database in the analysis (Yaalon, 1998), and the very small number of taxa used to represent total pedodiversity.

The objective of this study is to make a primary exploration of pedodiversity patterns at national and regional scales in the USA, and to understand the mechanisms underlying the patterns so that the results can provide a reference for soil resource assessment and soil conservation planning in USA. For this purpose, the most common diversity indices were estimated for the soils of the conterminous USA as a whole, and for soils within each of US Department of Agriculture (USDA)-Natural Resources Conservation Service (NRCS) geographical regions using State Soil Geographical Data Base (STATSGO). Taxa–area relationships were examined for each taxonomic category. Differences in pedodiversity between the regions were compared, and the four abundance models were tested for taxa in each taxonomic category to investigate the soil ‘community’ structure at various taxonomic levels.

2. Materials and methods

2.1. Data sources

The STATSGO (1997 version) soil database was utilized in this study (SCS, 1992). STATSGO is a Geographic Information System (GIS)-based relational database compiled by the National Resources

Conservation Service. It was generated by generalizing more detailed soil survey data, and is the only soil database available for evaluation of soil resources at a national scale with reasonably fine resolution. The mapping scale for the STATSGO map is 1:250,000, with a minimum mapping unit area of 6.25 km²—which is equivalent to square cells 2.5 by 2.5 km in size. The basic structure of STATSGO consists of map units and components. A map unit, composed of one or a few GIS polygons, may contain up to 21 components. Out of the 10,439 map units (74,577 polygons) in the STATSGO of the conterminous USA, there are 111,074 soil components having taxonomic information. For each soil component, the proportion (%) occupied in the map unit, as well as soil classification information for the six taxonomic categories (in descending rank: order, suborder, great group, subgroup, family, and series), are recorded in a relational database format (Soil Survey Staff, 1975, 1999). Data for the 50 states and the Puerto Rico territory are available, but in this study, data for 48 states in the conterminous US were used. The original projection of STATSGO was retained except that the datum was changed from Clarke 1866 to NAD83 using ARC/INFO software (Environmental Systems Research Institute, 1998).

2.2. Analysis methods

2.2.1. Calculation of pedodiversity indices

The pedodiversity indices were estimated with the taxonomic information stored in STATSGO. The area of a taxon in each map unit was calculated by multiplying the component percentage of the taxon by the area of the map unit (area of several GIS polygons). The total area of each taxon from the 48 states was summarized. The areas of all taxa in each of the six soil classification categories were calculated. The diversity indices were calculated based on the area abundance of the taxa.

Three types of indices were considered in this study: richness, evenness, and diversity. Richness expresses how many soil taxa exist in a region studied; evenness addresses area equitability of the soil taxa; diversity is an index taking both richness and evenness into account, i.e. the higher the richness and evenness, the higher the diversity. There are several indices in each of the types, but most of them

are highly correlated (Magurran, 1988; Ibáñez et al., 1998). The number of taxa in each taxonomic category was considered as taxa richness (S). For evenness, Smith's evenness index (E) was adopted because it has been recommended for general use after comparison with 13 other evenness indices (Smith and Wilson, 1996) and is considered to be the best-discriminating index for the random function like lognormal distribution (Mouilliot and Wilson, 2002). For diversity, Shannon's diversity index (H') was used, which is the most widely used index to describe both species diversity (Pielou, 1975; Magurran, 1988; Ibáñez et al., 1998) and patch diversity in landscape ecology (O'Neill et al., 1988; Hulshoff, 1995; Ritters et al., 1995). To consider integrated diversity, Shannon's hierarchical diversity index (Hierarchy H') was applied since there are six hierarchical taxonomic categories in the current Taxonomy System (Pielou, 1975; Magurran, 1988; Soil Survey Staff, 1999). The O'Neill dominant index was used to examine the deviation of the calculated Shannon's diversity index from the maximum diversity (O'Neill et al., 1988).

Smith's evenness index was calculated with the following relationship (Smith and Wilson, 1996):

$$E = 1 - \frac{2}{\pi} \arctan \left\{ \frac{\sum_i^S \left[\ln(x_i) - \frac{\sum_j^S \ln(x_j)}{S} \right]^2}{S} \right\} \quad (1)$$

where S is taxa richness; x is area abundance of i th (or j th) taxa. E is taxa evenness and has a range of 0–1, with 0 representing the minimum evenness, and 1 the maximum (area of all taxa is equal to N/S , where N is total soil area studied and S is taxa richness).

Shannon's diversity index was calculated with the following equation (Shannon and Weaver, 1949; Magurran, 1988):

$$H' = - \sum_i^S p_i \ln(p_i) \quad (2)$$

where S is taxa richness; p_i is the proportion of i th taxa. p_i is estimated by n_i/N , where n_i is the area covered by the i th taxa and N is the total area studied. Shannon's hierarchical diversity index was calculated

using the following equation (Pielou, 1975; Magurran, 1988):

$$H'(i) = H'(1) + H'_1(2) + H'_{12}(3) + \dots + H'_{12\dots(i-1)}(i) \quad (3)$$

where $H'(i)$ is hierarchical diversity index at i th taxonomic levels. $H'(1)$ is diversity index at first taxonomic level (Order in Soil Taxonomy System). $H'_1(2)$ is the mean of second taxonomic levels within first category, which is averaged over all taxa in the second taxonomic level (average suborder diversity within each order category in Soil Taxonomy System). The same applies to the definitions of the other hierarchical components.

Shannon's maximum diversity index was calculated:

$$H_{\max} = \ln(S) \quad (4)$$

where S is taxa richness. When each taxon has equal area abundance, i.e. $n_i = (N/S)$ where N is the total area studied, Shannon's diversity index (H') reaches its maximum H_{\max} .

Shannon's maximum relative diversity index was calculated (Ibáñez et al., 1998):

$$H_{\max r} = \frac{\ln(S)}{\ln(S_T)} \quad (5)$$

where S is taxa richness in the region studied. S_T is the total taxa richness in the conterminous USA.

O'Neill dominant index was calculated with the following equation (O'Neill et al., 1988):

$$D = \ln(S) + \sum_i^S p_i \ln(p_i) \quad (6)$$

where S and P_i are the same as the those in the Eq. (2).

There are six USDA-NRCS geographical regions in the conterminous USA: West, Northern Plains, Midwest, South Central, Southeast, and East (Fig. 1). The basic characteristics of topography, mean annual temperature, mean annual precipitation, main soil parent materials for each region were presented in Table 1. Taxa richness, Smith's evenness index, Shannon's diversity index, Shannon's hierarchical diversity index, and O'Neill dominant index were calculated for each region.

Calculations of abundance of each taxon (area) and the diversity indices described above were processed

using programs compiled by the senior author with Visual Basic in Microsoft Access (Microsoft Corporation, 2000) and Avenue in ArcView GIS (ESRI, 1999).

2.2.2. Comparison of soil diversity among USDA-NRCS geographical regions

To test whether taxa richness is influenced by the size of the study area (which would bias a region by region analysis of diversity), a GIS moving window technique was used to develop area–soil “species” relationships. First, STATSGO data was converted from polygon to point coverage to obtain the coordinates of the geometric centroids for each polygon (excluding those polygons with pure water) in Arc/info software (ESRI, 1998). The conversion resulted in 74,633 pairs of coordinates for the conterminous USA. To avoid an edge effect caused by encountering the US boundary during the moving window operation, three rectangles (each 2400 by 1300 km in length/height) were selected in the upper left, middle and low right parts of the conterminous USA as study areas (the window size was determined based on the size of rectangles that could be fitted into the conterminous USA). A window with a dimension of 50 km (in length/height) was moved from upper left corner, rightwards, then left to right in the next line, etc., across each rectangle to capture the points within the windows. The taxa richness in each taxonomic category inside each window was tabulated, and the average of taxa richness was calculated. This step was repeated using window sizes from 100 km (in square length) to 1300 km in increments of 50 km. The relationship between taxa richness and the size of the window sampled was analyzed.

One hundred random samples were selected from each USDA-NRCS region to compare differences in pedodiversity between the regions. First, each of the polygons in STATSGO was numbered with a range of 1 to n_i (n_i is total number of polygons in the i th region). A random polygon was then selected by generating a random number (within a range of 1 to n_i) in uniform distribution (equal probability) within SAS software (SAS Institute, 1996). The selection process was repeated until total area of polygons selected reached the required range of $400,000 \pm 4$ km² (the area of the smallest region studied is 504,651 km², excluding the polygons of pure water). The taxa richness within each taxonomic category was

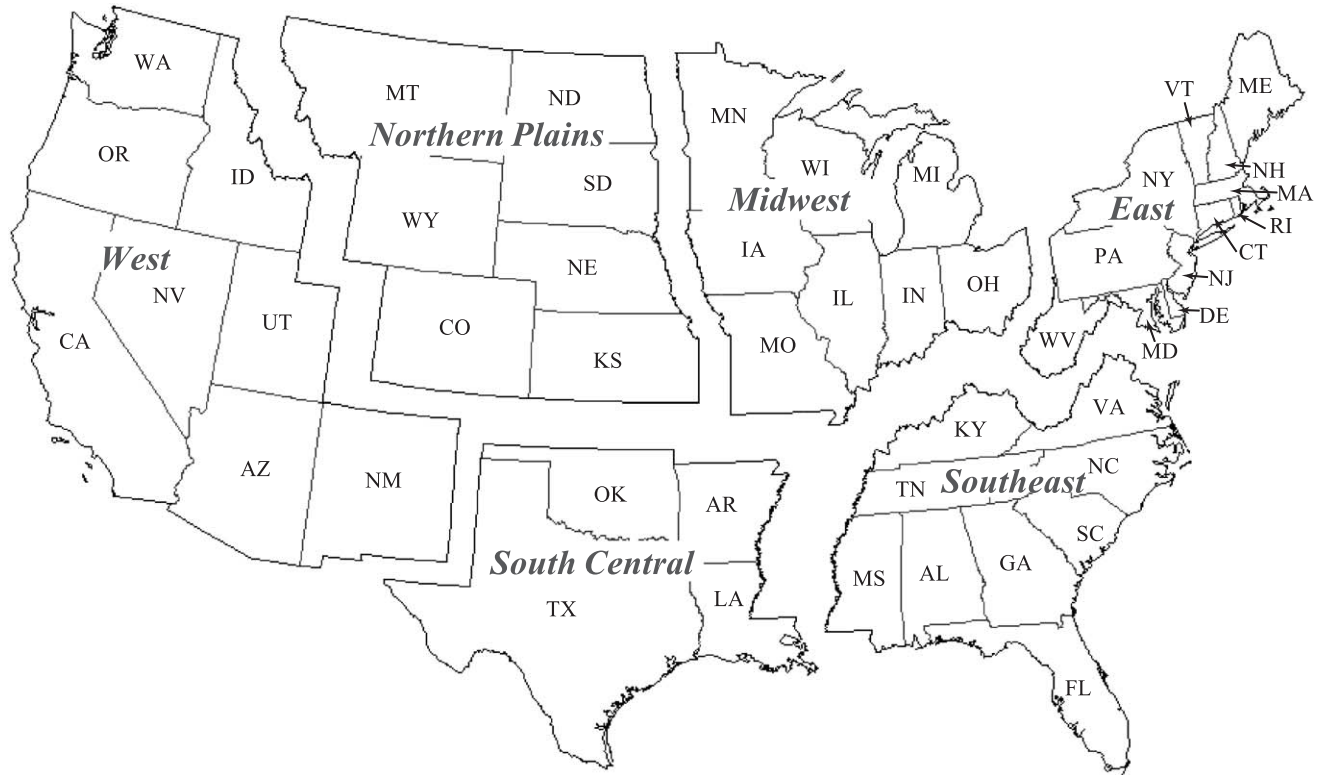


Fig. 1. USDA-NRCS geographical regions and state boundaries in the conterminous USA.

Table 1
Basic characteristics of elevation, climate, and soil parent materials in each geographical region

Region	Elevation (m) ^a			MAT (°C) ^b			MAP (mm) ^b			Main soil parent materials ^c
	Mean ^d	Variation		Mean	Variation		Mean	Variation		
		Range ^e	STD ^f		Range	STD		Range	STD	
West	1341	<0–4314	685	10.9	–4.1–24.6	0.46	529	48–7108	517	Alluvial fan and desert accumulations, Complex mountainous areas and outwash, etc.
Northern plains	1211	213–4328	743	7.2	–4.2–15.1	0.31	505	141–2959	215	Soft sedimentary rock, Glacial drift, Complex mountainous areas and outwash, etc.
Midwest	289	32–670	89	8.7	2.0–15.4	0.30	878	452–1287	147	Glacial drift, Loess, etc.
South central	395	<0–2558	385	17.5	11.3–23.2	0.22	884	219–1826	366	Marine, Soft sedimentary rock, Sandstone and shale, etc.
Southeast	168	<0–1981	196	16.3	7.6–25.2	0.29	1316	850–3029	155	Marine, Crystalline rocks, etc.
East	316	<0–1873	222	8.2	1.5–14.3	0.26	1088	732–2674	125	Glacial drift, Sandstone and shale, etc.

Elevation is meters above sea level. MAT: mean annual temperature. MAP: mean annual precipitation.

^a Calculated based on data sources (Gesch and Larson, 1996) with resolution of 1 km.

^b Calculated based on data sources (Daly et al., 2001) with resolution of 2 km.

^c Data sources (Troech and Thompson, 1993) p. 19.

^d Mean in the region studied.

^e Range of minimum to maximum in the region studied.

^f Standard deviation in the region studied.

counted and the area abundance of each taxon was calculated within each sample.

The “Jack-knifing” method, a technique that allows the estimate of virtually any statistic to be improved, was used for the estimation of the four types of indices (Zahl, 1977; Magurran, 1988). First, each of the four diversity indices was estimated by counting or by applying (Eqs. (1), (2) and (6) using all 100 samples in the region studied to get the overall index V ; then, the overall index was recalculated with one of the 100 samples excluded, which resulted in VJ_i ($i=1, 2, \dots, 100$) “jack-knife” estimates. Each of these Jack-knife estimates was converted to a pseudo-value, VP_i , using the following equation (Magurran, 1988):

$$VP_i = (nV) - [(n-1)(VJ_i)] \quad (7)$$

where n =the number of samples (e.g. 100 in this study).

Multivariate analysis of variance (MANOVA) was used to analyze the pseudo-values (taxa richness, Smith’s evenness index, Shannon’s diversity index, and O’Neill dominant index) of Jack-knifing with

SAS software to compare diversity differences between the USDA-NRCS regions (SAS/STAT, 1994).

The indices described above show pedodiversity, however, they do not illustrate the similarity of the taxa between the regions. Two regions when compared can have the same pedodiversity, but totally different soil taxa. To study the taxa similarity between region A and region B, each of the 100 samples from region A was compared with each of the 100 samples from region B to calculate Sorenson’s similarity measure (index) using the following equation (Magurran, 1988):

$$C_N = \frac{2 \sum_j N_j}{N_a + N_b} \quad (8)$$

where C_N is Sorenson’s similarity measure; N_a and N_b are the sample area from A region and B region, respectively. N_j is the minimum value of the area abundance of j th taxa in two samples (a or b) being compared. When $C_N=0$, taxa in the two samples are completely different, whereas taxa are exactly the

same if $C_N=1$. The mean of the 10,000 (100×100) Sorenson's similarity indices was used as a representative of the taxa similarity between the two regions studied.

Sorenson's similarity indices between each pair of regions were calculated. Similarity indices were converted into dissimilarity indices by $C_D=1-C_N$. The dissimilarity indices were used to conduct cluster analysis by the method of average linkage in SAS software (SAS/STAT, 1994). The dissimilarity indices were converted back to similarity indices by $C_N=1-C_D$ when a cluster tree was made following cluster analysis.

2.2.3. Relative abundance model fitting

The fitness of data to a predicated model is usually assessed by a goodness-of-fit test (especially χ^2). The χ^2 test has a notoriously low discriminating power (Engen, 1979; Routledge, 1980; Hughes, 1986), which is especially true in abundance model testing where the data are grouped to a small number of points (by being grouped into octaves or groups of octaves). As a result, any model used to describe data will commonly be rejected by the goodness-of-fit tests (Hughes, 1986). Thus, a method of plotting expected values against the observed value, as well as residual variance, was used to test the model fitness in this study.

The four most commonly used biodiversity models were tested for the taxa in the conterminous USA (Magurran, 1988). In testing the lognormal model, all taxa abundance data of the conterminous USA were transformed on a \log_2 scale to obtain "octaves" of taxa abundance. If taxa abundance fits a lognormal model, "octaves" of abundance will display a normal distribution, and standardized octaves will fit a standardized normal distribution by the equation of standardization:

$$u_i = \frac{x_i - \mu}{\sigma} \quad (9)$$

where u_i is standardized octaves of abundance; x_i is octaves of abundance; μ and σ are the mean and standard deviation of octaves abundance in each taxonomic category, respectively.

The probability of a standardized normal distribution within an octave was obtained using SAS software. The expected frequency per octave was estimated by multiplying the total number of taxa

with the probability within the corresponding octaves. The observed frequency per octave was plotted against expected frequency to test the fitness of the model in each taxonomic category. The residual variance between expected frequency and observed frequency was calculated by the following equation:

$$S_e^2 = \frac{\sum_i^n \left(\frac{y_i}{S} \times 100 - \frac{\hat{y}_i}{S} \times 100 \right)^2}{n-1} \quad (10)$$

where S_e^2 is residual variance between expected frequency and observed frequency; n is the number of octaves; and y_i and \hat{y}_i are the observed and expected frequency in each octave. S is the total number of taxa.

3. Results and discussion

3.1. Pedodiversity in the conterminous USA and its geographical regions

Pedodiversity indices for each taxonomic category in the conterminous USA are presented in Table 2. As taxonomic category decreases from order to series, Shannon's diversity index increases because taxa richness increases dramatically. In terms of taxa evenness, taxa in all taxonomic categories are not of equal area (the maximum $E=1$ occurs when all taxa in a certain category have equal area). Lower evenness of taxa indicates some taxa are relatively rare while some others have a large area abundance. When diversity indices are considered, a comparison of similarity between the Shannon's diversity index and the hierarchical Shannon's diversity index in each taxonomic level showed that there is no obvious difference, suggesting that either one can be used in pedodiversity studies.

Ibáñez et al. (1998) studied pedodiversity for each continent of the world using the soil data from the FAO soil map and biodiversity indices. The soil classification system in the FAO soil map is different from the US Soil Taxonomy used in STATSGO. However, taxa in FAO might be considered equivalent (in terms of detail) to the order or suborder categories in US Soil Taxonomy. Map scales of FAO and STATSGO are different. Nevertheless, the scale effects are not strongly reflected in estimating diver-

Table 2
Pedodiversity indices in USDA-NRCS geographical regions and the USA

Taxonomy category	Regions	S	E	H_{\max}	$H_{\max r}$	H'	Hierarchy H'	D
Order	West	10	0.214	2.303	1.000	1.760	1.760	0.543
	Northern plains	10	0.055	2.303	1.000	1.349	1.349	0.954
	Midwest	8	0.094	2.079	0.903	1.524	1.524	0.555
	South central	8	0.442	2.079	0.903	1.791	1.791	0.288
	Southeast	8	0.290	2.079	0.903	1.402	1.402	0.677
	East	7	0.143	1.946	0.845	1.525	1.525	0.421
Suborders	Conterminous USA	10	0.428	2.303	1.000	1.961	1.961	0.342
	West	48	0.076	3.871	1.000	2.686	2.687	1.185
	Northern plains	33	0.067	3.497	0.903	2.317	2.317	1.180
	Midwest	25	0.086	3.219	0.831	2.404	2.404	0.815
	South central	26	0.107	3.258	0.842	2.491	2.491	0.767
	Southeast	24	0.094	3.178	0.821	1.891	1.891	1.287
Great-group	East	21	0.130	3.045	0.786	2.005	2.005	1.040
	Conterminous USA	48	0.140	3.871	1.000	3.100	3.100	0.771
	West	163	0.096	5.094	0.956	3.690	3.690	1.404
	Northern plains	103	0.068	4.635	0.870	3.300	3.300	1.335
	Midwest	66	0.088	4.190	0.786	3.168	3.168	1.022
	South central	82	0.101	4.407	0.827	3.512	3.513	0.895
Subgroup	Southeast	65	0.083	4.174	0.783	2.987	2.988	1.187
	East	56	0.083	4.025	0.756	2.571	2.572	1.454
	Conterminous USA	206	0.114	5.328	1.000	4.284	4.285	1.044
	West	696	0.133	6.545	0.940	5.255	5.267	1.290
	Northern plains	374	0.093	5.924	0.851	4.542	4.548	1.382
	Midwest	232	0.128	5.447	0.782	4.249	4.250	1.198
Family	South central	316	0.148	5.756	0.827	4.891	4.892	0.865
	Southeast	221	0.139	5.398	0.775	4.060	4.063	1.338
	East	149	0.097	5.004	0.719	3.382	3.403	1.622
	Conterminous USA	1057	0.134	6.963	1.000	5.746	5.751	1.217
	West	3317	0.179	8.107	0.933	6.954	6.976	1.153
	Northern plains	1351	0.133	7.209	0.829	5.803	5.810	1.406
Series	Midwest	916	0.145	6.820	0.785	5.554	5.555	1.266
	South central	1010	0.154	6.918	0.796	5.919	5.920	0.999
	Southeast	696	0.129	6.545	0.753	5.193	5.195	1.352
	East	452	0.130	6.114	0.703	4.398	4.419	1.716
	Conterminous USA	5959	0.161	8.693	1.000	7.361	7.369	1.332
	West	6433	0.241	8.769	0.927	7.767	7.812	1.002
Series	Northern plains	2581	0.169	7.856	0.831	6.641	6.666	1.215
	Midwest	1969	0.199	7.585	0.802	6.581	6.585	1.004
	South central	1542	0.179	7.341	0.776	6.420	6.422	0.921
	Southeast	1169	0.152	7.064	0.747	5.948	5.953	1.116
	East	850	0.149	6.745	0.713	5.573	5.595	1.172
	Conterminous USA	12,788	0.212	9.456	1.000	8.312	8.337	1.144

S =taxa richness in each taxonomy category. E =Smith's evenness index. H_{\max} =maximum diversity. $H_{\max r}$ =maximum relative diversity. H' =Shannon's diversity index. Hierarchy H' =Shannon's hierarchical diversity index. D =Oneill's dominant index.

sity indices of taxa area of higher taxonomic categories. The taxa richness and Shannon's diversity indices calculated by Ibáñez et al. (1998) for North America are 22 and 2.57, respectively, which are more or less similar to the values of these indices calculated for the order and suborder categories in the present study (Table 2).

Pedodiversity indices for each taxonomic category in the USDA-NRCS geographical regions, as well as maximum diversity and maximum relative diversity, are also presented in Table 2. The pedodiversity indices in each region follow the general trends of the entire conterminous USA as a whole. The sequence of taxa richness trends, by geographical re-

gion, is: West, Northern Plains, Midwest or South Central, Southeast, and East (in descending order). The diversity index would reach its maximum value if all taxa had an equal area abundance (complete evenness). The difference (D) between the calculated diversity index and its maximum value is another way to illustrate evenness of taxa or to detect the dominant taxa. The South Central region has the smallest dominant index in all taxonomic categories, suggesting that taxa abundance in this region is more evenly distributed. The diversity index implies both taxa richness and evenness, thus, the West and South Central regions have the highest diversity indexes among the regions studied, except at the series category because the number of series in the South Central region are relative small.

The diversity index (series category) versus area of the regions is presented in Fig. 2. There is an obvious positive linear trend between areas of the regions studied versus diversity, which is similar with the previous results of pedodiversity for the world (Ibáñez et al., 1998; MacBratney et al., 2000). The diversity index includes both taxa richness and taxa evenness. The previous studies showed that there is a strong species–area relationship either in biodiversity (Kilburn, 1966) or pedodiversity (Beckett and Bie, 1978),

and that the relationship can be commonly formulated as $S=cA^z$ (S , taxa richness; A , area; c and z are constants). The formula is well known in ecology as the power law (Borde-de-Agua et al., 2002). In the work by Ibáñez et al. (1998), there was no species–area relationship using the FAO small-scale soil map (1:5,000,000). However, it is important to test whether taxa richness is area-dependant when a map of more detailed scales is used. If taxa richness is area-dependant, the diversity index calculated in the Fig. 2 or Table 2 might be confounded by the area factor. A comparison of pedodiversity between the regions should be based on equal land areas in the same way as that commonly used in by biodiversity studies.

3.2. Pedodiversity by USDA-NRCS geographical region

Averaged taxa richness as a function of area was tested from GIS windows that were moved across the three study areas (Fig. 3). It is obvious that taxa richness increases dramatically as the study area increases. For orders and suborders, the richness values approach the maximal values of 10 orders and 48 suborders at an area of 1,690,000 km² (1300 by 1300 km window size) (Fig. 3 and Table 2). More

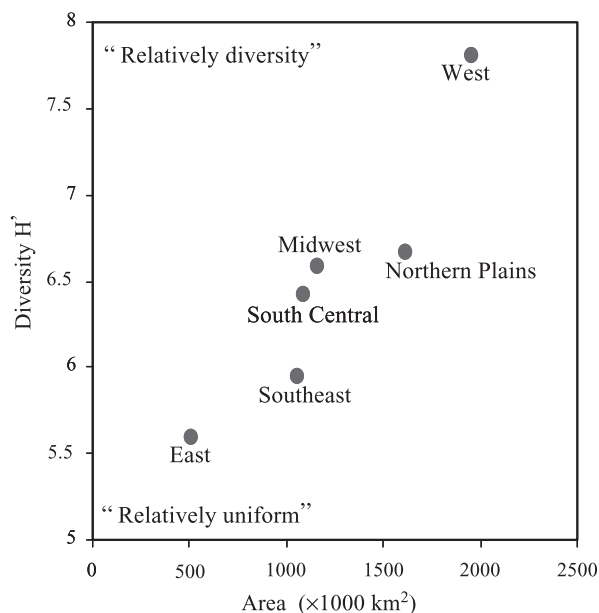


Fig. 2. Diversity of soil series versus area of USDA-NRCS geographical regions.

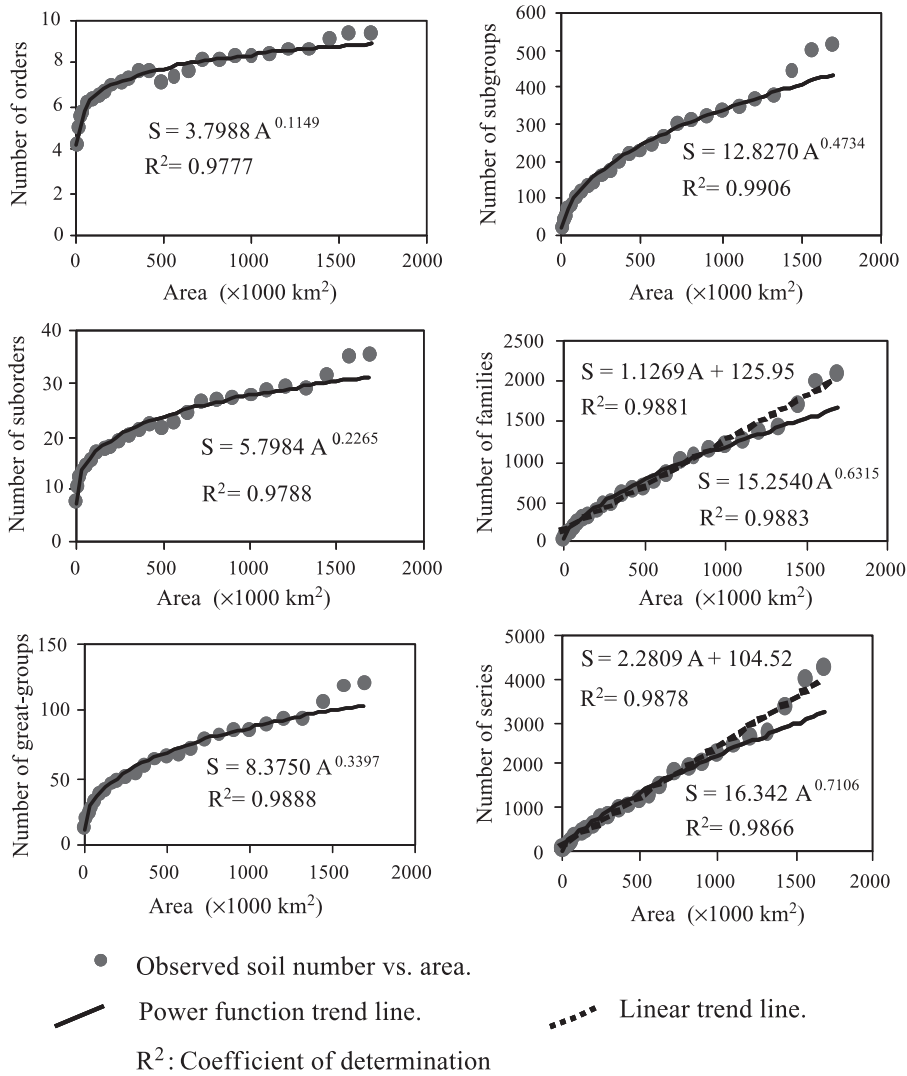


Fig. 3. Taxa richness (S) versus area of moving windows (A).

than half of the great groups and subgroups have also been captured in this window size. At the family and series levels, however, there is an obvious tendency for taxa to continue to increase with increasing area since less than half of the families and less than one third of the total series have been observed in a 1300 by 1300 window. The relationship between the number of taxa (S) and area (A, 1000 km²) are captured by power function ($S=cA^z$). The exponent z reflects the taxa-richness of soil ‘communities’ and increases constantly as taxonomic categories decrease from order to series. For higher taxonomic categories, taxa

richness increase very fast at small window sizes, and the increasing speed gradually decreases as the window area increases, until reaches its maximum. However, as taxonomic category declines, taxa richness increases constantly as the window size increases and the species–area relationship changes gradually from a power to linear function. At the series level, the species–area relationship is $\hat{S}=2.2809A+104.52$ with $R^2=0.9878$. Theoretically, $S(0)=0$. However, small-scale variability may often cause the predicting curve not to originate in the origin of the coordinate system, an artifact called the nugget effect. Extrapolating the

area–series relationship to the USA (area of the conterminous USA, Alaska, Hawaii and Puerto Rico, is about 9,243,474 km² (based on STASTGO GIS coverage excluding polygons of pure water in projection of Albers conic equal-area)) results in an estimate of 21,084 soil series, which agrees with the present listed number of 21,000 series (<http://soils.usda.gov>). These results strongly suggest that soil taxa richness is area-dependant. Because diversity implies both taxa richness and taxa evenness, a comparison of diversity between regions should be based on equal area size.

One hundred random samples, each equal to an area of 400,000±4 km², were selected from each region. *F* tests on the four diversity indices among the regions are significantly different at 0.05 probability level in a multivariate analysis of variance (MANOVA). A multiple-comparison on the means of the indices between the regions was further conducted (Table 3). Comparison of the taxa richness indices showed that the West region has the highest taxa richness, followed by the Northern Plains region;

the East region has the lowest taxa richness in all taxonomic levels except in the great group category, where the Southeast region has the lowest taxa richness in that category (Table 3). With respect to the evenness indices, the West region has the lowest evenness in all taxonomic categories (except in the great group category), indicating that the taxa in the region have large variations of area abundance. It is understandable that the West region has the highest taxa richness and lowest evenness because of its complicated variations in lithology, topography, and climate (Table 1). Similar to the results of the nation as whole, evenness indices in each region all are small and far away from the maximum value of 1, indicating that soil resources in each region are unevenly distributed and some taxa may be very rare. Rarity is one of the important concepts for conservation planning of all natural resources. Evenness is a useful index to depict a community structure. Developing new evenness indices and studying evenness performance still is a focused research field in ecology, and none of the

Table 3
The comparison of mean taxonomic diversity indices among the geographical regions

Indexes	Region	Order	Suborder	Great group	Subgroup	Family	Series
S	West	10 A	38 A	126 A	431 A	1463 A	2216 A
	Northern plains	10 A	30 B	94 B	305 B	939 B	1585 B
	Midwest	8 B	23 D	60 D	173 E	570 C	1171 C
	South central	8 B	25 C	76 C	238 C	572 C	829 D
	Southeast	8 B	23 D	55 E	194 D	493 D	772 E
	East	7 C	21 E	58 D	141 F	412 E	759 E
E	West	0.060 D	0.080 D	0.064 C	0.083 D	0.092 C	0.087 D
	Northern plains	0.062 D	0.097 B	0.087 B	0.116 AB	0.132 A	0.158 A
	Midwest	0.106 C	0.088 C	0.073 BC	0.086 CD	0.115 B	0.153 A
	South central	0.273 A	0.093 B	0.105 A	0.104 B	0.092 C	0.105 C
	Southeast	0.234 B	0.095 B	0.107 A	0.128 A	0.129 AB	0.129 B
	East	0.107 C	0.110 A	0.041 D	0.072 D	0.090 C	0.114 BC
H'	West	1.665 A	2.570 A	3.375 A	4.471 A	5.759 A	6.162 B
	Northern plains	1.525 C	2.306 C	3.180 C	4.262 C	5.429 B	6.271 A
	Midwest	1.209 E	1.871 F	2.604 E	3.551 E	4.782 E	5.799 C
	South central	1.662 A	2.390 B	3.356 B	4.362 B	5.181 C	5.662 D
	Southeast	1.560 B	2.018 D	2.994 D	4.122 D	4.983 D	5.797 E
	East	1.456 D	1.937 E	2.510 F	3.327 F	4.354 F	5.550 F
D	West	0.638 C	1.067 D	1.459 C	1.592 B	1.523 C	1.531 A
	Northern plains	0.778 B	1.095 C	1.363 D	1.458 C	1.414 D	1.096 D
	Midwest	0.870 A	1.263 A	1.490 B	1.602 AB	1.563 B	1.266 B
	South central	0.417 E	0.829 E	0.974 F	1.111 E	1.162 F	1.053 E
	Southeast	0.520 C	1.117 B	1.012 E	1.144 D	1.216 E	1.051 E
	East	0.417 F	1.107 B	1.559 A	1.620 A	1.666 A	1.131 C

S=Taxa richness in each taxonomy category. E=Smith's evenness index. H'=Shannon's diversity index. D=Oneill's dominant index. Means with the same letter are not significantly different at 0.05 probability level.

evenness indices studied has been found to be perfect. The J evenness index ($J=H/H_{\max}$; Pielou, 1975) was found to strongly overestimate evenness, especially when its real value is low (Molinari, 1989; Bulla, 1994). The Smith evenness index (E) is recommended for general use after comparison with 13 other evenness indices (Smith and Wilson, 1996), and is considered the best-discriminating index that is suitable for detecting low evenness in random functions like a lognormal distribution (Mouilliot and Wilson, 2002). However, it was also found not to be stable among replicate samples (Mouilliot and Wilson, 2002), it seems that it may not be sensitive to the high levels of evenness in this study. The dominant index (D) clearly showed that taxa in the South Central region have the highest evenness among the regions, although it cannot be compared directly with the evenness index E since D is not restricted to values in the range of 0–1. For the overall pedodiversity of the regions, taking both taxa richness and evenness into account (Shannon's diversity index), the West and South Central regions have the highest and second highest diversity indices at the order, suborder, great group, and subgroup categories. However, in the family and series categories, the West or Northern Plains regions have the highest diversity indices among the regions. The East and Midwest regions are the least or second least diverse regions in all taxonomic levels except at the series level. The high diversity index of the West region mainly comes from higher taxa richness while the higher diversity index in South Central region is due mainly to the evenness of taxa.

In Ibáñez et al.'s (1998) study of pedodiversity for the world, taxa richness, evenness and diversity indices were very similar between the continents. In some ways, this is expected because of the restricted range of taxa that were allowed on a map with a coarse resolution and the fact that at high taxonomic levels, most taxa would be expected to be present given a reasonable land area (Fig. 3). In our study, with a much more detailed database, taxa richness between USDA-NRCS geographical regions are very different, which, as expected, is more obvious in the lower taxonomic categories. In term of evenness, the indices are significantly different between the regions. However, the absolute value of differences between regions is small because the evenness indices of all

the regions are skewed toward a value of zero, the minimum evenness. The overall diversity index is of concern in planning issues, and we note that the indices are significantly different between the regions. In summary, our results indicate that there is far more geographical variation in soil diversity for the nation than expected based on the previous work.

Using indices to study landscape structure is a growing research area in landscape ecology. Indices of measuring contagion and fractal geometry of the landscape are commonly used to depict the aggregated or clumped landscape patches and patch shape in ways similar with indices used in our study: patch richness, patch evenness and patch diversity index (O'Neill et al., 1988; Hulshoff, 1995). However, we can not explore taxa contagion and fractal geometry in this study because the location of each soil taxon is not specified in STATSGO.

Results of a similarity analysis of taxa between the regions are presented in Fig. 4. The distance at the top of the figure indicates the level of similarity: 1.0 means that the taxa in two (or more) regions studied are exactly the same; and 0.0 indicates that the taxa in the regions are completely different. The general pattern of taxa similarity between the regions is that taxa in the South Central and Southeast regions are most similar while taxa in the Northern Plains most resemble to the taxa in the West region, especially at the higher taxonomic levels. Taxa similarity between the regions decreased dramatically as taxonomic levels descended. There was a 72% similarity between the West and the Northern Plains regions at the order category while 66% taxa similarity was found between the South Central and the Southeast regions. At the suborder level, the similarity was 60% between the South Central and the Southeast regions. At the great group level, 46% similarity existed between the South Central and the Southeast regions. At the subgroup and family levels, the similarity between the South Central and the Southeast was only 34% and 16%, respectively. At the series level, taxa similarity between each pair of regions was less than 7%.

The taxa similarity analysis illustrated in Fig. 4 indicates that taxonomic similarity at the higher categories is driven by geographic/climatic proximity. However, at the lower taxonomic levels, soil dissimilarity is the norm, and even geographical proximity exerts little effect on similarity indices. The trend

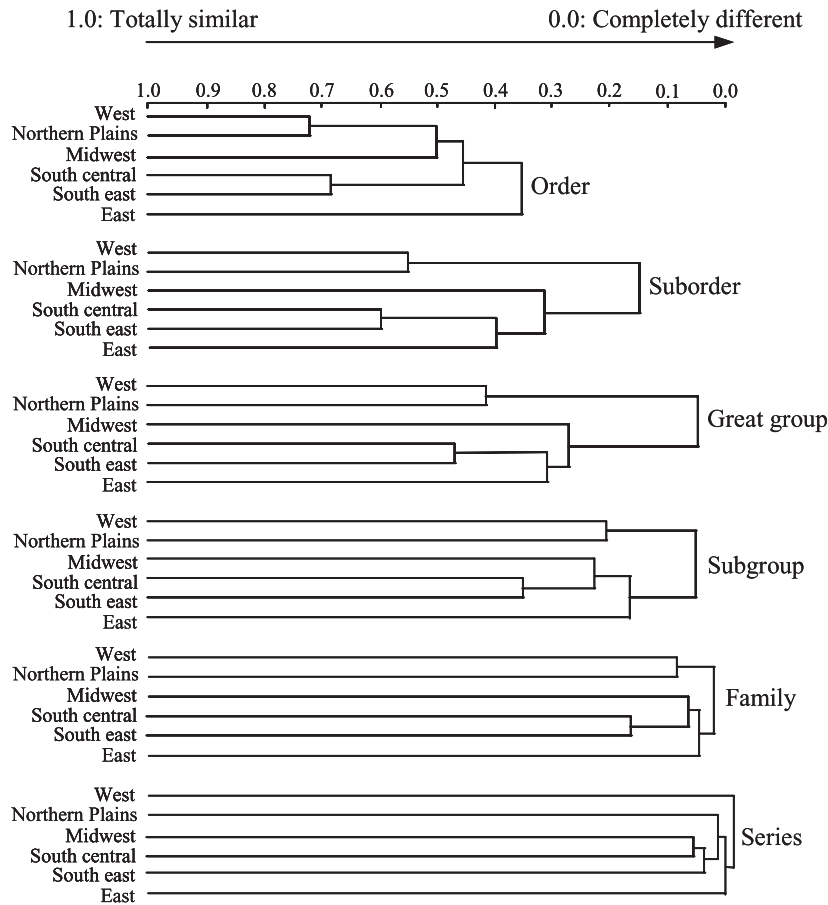


Fig. 4. Taxa similarity between the geographical regions for each taxonomy category.

towards “endemism” at the series and family levels is of considerable interest in conservation or preservation efforts, and it likely reflects, controls, or is in some ways related to the regional distribution in plant species (Amundson et al., in press).

3.3. Taxa abundance model and its implications

The structure of the quantitative abundance of taxa will help reveal the mechanisms or factors responsible for the distribution of the soil “community”. Out of the four abundance models tested, the lognormal model most fits the taxa abundance of the USA, which is consistent with results of Ibáñez et al.’s (1998) study. The typical characteristic of a lognormal distribution is asymmetry: i.e. the data are positively skewed, concentrated around the left end, close to the

taxon with the smallest abundance. A lognormal model will display a normal distribution after log transformation in any base (\log_e , \log_2 , and \log_{10} , etc.) However, the \log_2 scale usually works best since it captured most natural phenomena (Preston, 1948), and did so in our study. Transformed abundance in the \log_2 scale has been referred to a ‘octaves’ of abundance (Hughes, 1986). Our analysis shows that the taxa abundance gradually reaches a lognormal distribution as the taxonomic category decreases (Fig. 5). It is interesting to note that the fitness of taxa abundance to the model varies with taxonomic category.

A lognormal model is a statistically orientated model, suggesting that the relative abundance of species in a community results from the combined effect of a large number of mutually independent causes that are multiplicative in their effect (Pielou,

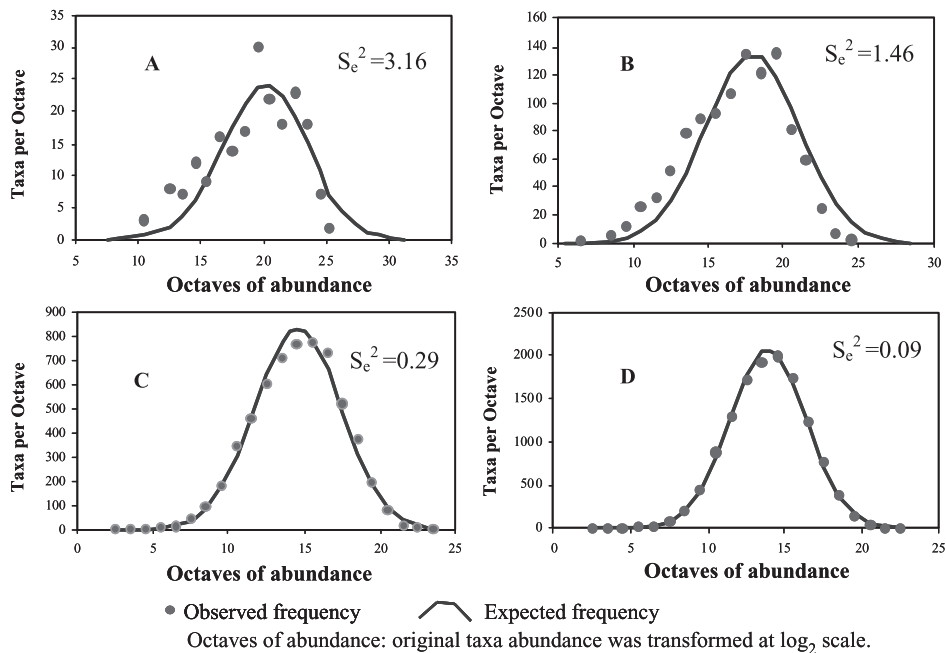


Fig. 5. Observed and expected frequency for a lognormal model in the great group (A), subgroup (B), family (C), and series (D) taxonomic categories.

1975). However, a lognormal model can also be considered as a niche-orientated (or resources-apportioning) model in that communal niche space is likened to a unit mass that has been sequentially split up into the component species, so that each fragment denotes relative species abundance. The successive subdivisions, in a hierarchical manner, may correspond to apportioning on different sets of niche axes (or resources availability), which could be driven by either ecological or evolutionary (or resource) forces. Such sequential fracturing can lead asymptotically to lognormal size-frequency distributions (Pielou, 1975; Sugihara, 1980).

Based on the above definition, we proposed a hypothesis on mechanisms underlying taxa abundance distributions in each taxonomic category. Taxa abundance gradually reaches a lognormal distribution as the taxonomic category decreases because of sequential breakage in taxonomic space. The total soil “mass unit” is broken into various soil orders by variations in a few soil formation factors. For instance, the formation of order A may be controlled mainly by soil parent materials (e.g. Vertisols) (Fig. 6). As the taxonomic category within that order (A) decreases

from the order level, taxa abundance is hierarchically and sequentially split in response to more soil formation factors, and the soil properties resulting from them, which vary across space. For example, the splitting of order A into suborder A_i may be caused by climate, the splitting of A_i into great group A_{ij} may be caused by vegetation, the splitting of A_{ij} into subgroup A_{ijk} by elevation, and the splitting of A_{ijk} into family A_{ijkp} by slope. As the taxonomic categories decrease further, the process of soil formation is driven by more and more factors (or factorial combinations), which results in a pattern of taxa abundance trending towards a lognormal distribution.

The Soil Taxonomy was deliberately designed to be a practical classification based on soil properties and climate regimes. The development of the system is in constant flux, subject to revision and expansion—particularly at the subgroup level. Yet, the soil properties (and series) themselves reflect variations in soil forming factors, and the system shows characteristics of theoretical distributions driven by variations in soil forming factors (Jenny, 1946). A number of soil properties have been found commonly following lognormal distribution, such as particle-size, contents

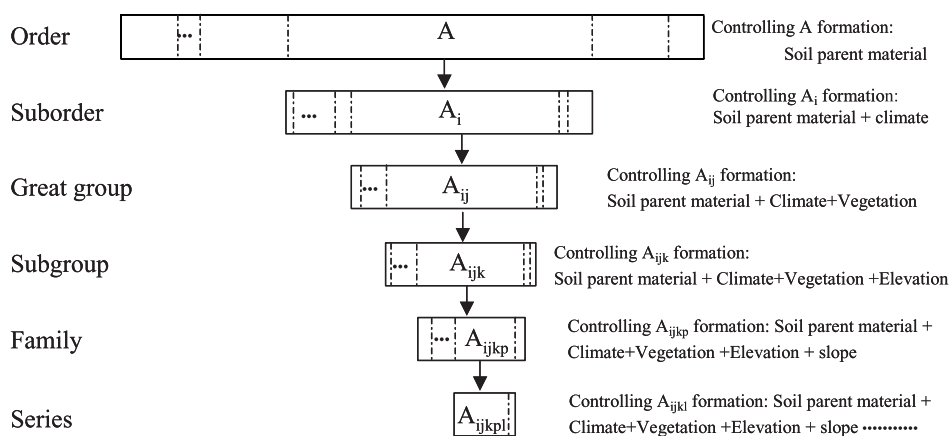


Fig. 6. A hypothetical sequential breakage model describing the manner in which new taxa are generated within various taxonomic categories.

of nutrient elements, etc. The mechanisms behind those phenomena may be explained in the same way as that of divisions of the taxonomic space.

As a concluding set of comments on our results, we emphasize that soil data bases, such as STATSGO, are the result of soil mapping efforts that vary in approach and intensity over time and space. Thus, the number and extent of any taxa may reflect artifacts of the soil survey process. All those may potentially contribute to the metrics calculated in this study. Additionally, the ecological indices used in this study are taken directly from species-oriented ecology, in which there are arguably more clear and objective criteria for distinguishing species. The suitability of treating soils as an individual “species” may be worth further discussion, although it can alternatively be argued that both soil taxa and biological species are to some degree arbitrary segments of their own respective continua. Darwin’s (1985) statement about species identification might apply as well to soils: “Certainly no clear line of demarcation has as yet been drawn between species and sub-species...or again between sub-species and well-marked varieties, or between lesser varieties and individual species. These differences blend into each other in an insensible series; and a series impresses the mind with the idea of an actual passage.”

4. Conclusions

Taxa abundance in the USA at any taxonomic category are unevenly distributed, i.e. some taxa are

rare while some taxa are very abundance. In addition, the relationship between the number of taxa (S) and area (A , 1000 km²) can be formulated by $S=cA^z$. The exponent z reflects the taxa-richness of soil ‘communities’ and increases constantly as taxonomic categories decrease from order to series. Previous research on soil taxa richness at higher levels of taxonomic abstraction (using the FAO soil map) showed no ‘taxa–area’ relationships. However, using the more detailed soil distributions found in STATSGO, we were able to show that as the taxonomic level of abstraction decreases (from order to series), there was an increasingly stronger trend of increasing taxa with area. This relationship demonstrates that at the soil series level, there is a large degree of local to regional soil endemism.

The quantitative determination of soil endemism is predictable from soil formation theory (Jenny, 1946). The unique combination of state factor combinations from region to region would be expected to result in unique soils, and the number of state factor combinations would be expected to increase with increasing land area. The fact that soil taxa are geographically restricted is important to planning and conservation efforts. Soil characteristics may have a role, in some cases, in determining plant community distributions and the combination of pedological and ecological expertise may be useful in preservation of plant biodiversity. Finally, soils themselves have characteristics of importance to the environment and society (Amundson, 1998; Amundson et al., in press). The clear regional distribution of soil types calls for

integrated planning in order to maintain undisturbed segments of soils for a variety of future uses and purposes.

5. Uncited references

Wilson et al., 1996
World Conservation Monitoring Centre, 1994

Acknowledgements

We thank Kit Paris, Eric Vinson, Sharon Waltman and Norman B. Bliss for detailed explanations of the STATSGO database; Ruiliang Pu, Ming Xu, and Yongqian Tian for commenting an earlier draft of this paper. This research was partially supported by the Kearney Foundation of Soil Science.

References

- Amundson, R., 1998. Do soils need our protection? *Geotimes*, 16–20 (March).
- Amundson, R., Guo, Y., Gong, P., 2003. Soil diversity and landuse in the United States. *Ecosystems* (in press).
- Beckett, P.H.T., Bie, S.W., 1978. Use of soil and land-system maps to provide soil information in Australia. Division of soils technical paper, vol. 33. Commonwealth Scientific and Industrial Research Organization, Australia.
- Bell, G., 2001. Neutral macroecology. *Science* 293, 2413–2418.
- Borde-de-Agua, L., Hubbell, S.P., McAllister, M., 2002. Species–area curves, diversity indices, and species abundance distribution: a multifractal analysis. *The American Naturalist* 159, 138–155.
- Bulla, L., 1994. An index of evenness and its associated diversity measure. *Oikos* 70, 167–171.
- Copley, J., 2000. Ecology goes underground (soil biodiversity research). *Nature* 406 (6795), 452–454.
- Daly, C., Taylor, G.H., Gibson, W.P., Parzybok, T.W., Johnson, G.L., Pasteris, P., 2001. High quality spatial climate data sets for the United States and beyond. *Transaction of the American Society of Agricultural Engineers* 43, 1957–1962 (Data available on line <http://www.climate-source.com>).
- Darwin, C., 1985. *The Origin of Species by Means of Natural Selection*. Penguin Books, London, England.
- Engen, S., 1979. *Stochastic Abundance Models, with Emphasis on Biological Communities and Species Diversity*. Chapman & Hall, London.
- Environmental Systems Research Institute (ESRI), 1998. Arc/info 7.2.1 (unix) software. <http://www.esri.com/software/arcgis/arcinfo>.
- Gaston, K.J., Spicer, J.I., 1998. *Biodiversity: an Introduction*. Blackwell Science, MA, USA.
- Gesch, D.B., Larson, K.S., 1996. Techniques for development of global 1-kilometer digital elevation models. Pecora Thirteen, Human interaction with Environment—Perspectives from Space, Sioux Falls, South Dakota, August 20–22, 1996. Data available on line <http://edcdaac.usgs.gov/gtopo30/gtopo30.html>.
- Hughes, R.G., 1986. Theories and models of species abundance. *The American Naturalist* 128, 879–899.
- Hulshoff, R.M., 1995. Landscape indices describing a Dutch landscape. *Landscape Ecology* 10 (2), 101–111.
- Ibáñez, J.J., De-Alba, S., Bermúdez, F.F., García-Álvarez, A., 1995. Pedodiversity: concepts and measures. *Catena* 24, 215–232.
- Ibáñez, J.J., De-Alba, S., Lobo, A., Zucarello, V., 1998. Pedodiversity and global soil patterns at coarse scales (with Discussion). *Geoderma* 83, 171–192.
- Jenny, H., 1946. Arrangement of soil series and types according to functions of soil forming factors. *Soil Science* 61, 375–391.
- Jenny, H., 1980. *The Soil Resource, Origin and Behavior*. Springer-Verlag, New York.
- Kilburn, P.D., 1966. Analysis of the species–area relation. *Ecology* 47, 831–843.
- MacArthur, R.H., 1957. On the relative abundance of bird species. *Proceedings of the National Academy of Sciences of the United States of America* 43, 293–295.
- MacArthur, R.H., 1960. On the relative abundance of species. *The American Naturalist* 94, 25–36.
- MacBratney, A.B., 1992. On variation, uncertainty and informatics in environmental soil management. *Australian Journal of Soil Research* 30, 913–935.
- MacBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. *Geoderma* 97, 293–327.
- Magurran, A.E., 1988. *Ecological Diversity and Its Measurement*. Princeton Univ. Press, New Jersey.
- Major, J., 1951. A functional, factorial approach to plant ecology. *Ecology* 32, 392–412.
- May, R.M., 1975. Patterns of species abundance and diversity. In: May, M.L. (Ed.), *Ecology and Evolution of Communities*. Blackwell, Oxford, pp. 197–227.
- May, R.M., 1986. The search for patterns in the balance of nature: advances and retreats. *Ecology* 67 (5), 1115–1126.
- Mausbach, M.J., Wilding, L.P., 1991. Spatial variabilities of soils and landforms. SSSA Special Publication 28 (Madison, WI, USA).
- Microsoft Corporation, 2000. Microsoft Access Software. <http://www.microsoft.com/office/access>.
- Molinari, J., 1989. A calibrated index for the measurement of evenness. *Oikos* 56, 319–323.
- Mouilliot, J.B., Wilson, J.B., 2002. Can we tell how a community was constructed? A comparison of five evenness indices for their ability to identify theoretical models of community construction. *Theoretical Population Biology* 61, 141–151.
- O'Neill, R.V., Krummel, J.R., Gardner, R.H., et al. 1988. Indices of landscape pattern. *Landscape Ecology* 3, 153–162.

- Preston, F.W., 1948. The commonness and rarity of species. *Ecology* 29, 254–283.
- Pielou, E.C., 1975. *Ecological Diversity*. Wiley, New York.
- Ritters, K.H., O'Neill, R.V., Hunsaker, C.T., Wichham, J.D., Yankee, D.H., Timmins, S.P., 1995. A factor analysis of landscape pattern and structure metrics. *Landscape Ecology* 10, 23–39.
- Routledge, R.D., 1980. The form of species abundance distributions. *Journal of Theoretical Biology* 82, 547–558.
- SAS/STAT, 1994. *User Guide, Version 6, 4th ed.* SAS Institute, Cary, NC, USA.
- SAS Institute, 1996. SAS (unix) software. <http://www.sas.com/software>.
- SCS, 1992. *State Soil Geographic Data Base (STATSGO) Data User Guide*. USDA-SCS. National Soil Survey Center, Lincoln, Nebraska. Available also online http://www.ftw.nrcs.usda.gov/stat_data.html.
- Shannon, C.E., Weaver, W., 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.
- Smith, B., Wilson, J.B., 1996. A consumer's guide to evenness indices. *Oikos* 76, 70–82.
- Soil Survey Staff, 1975. *Soil Taxonomy: a basic system of soil classification for making and interpreting soil surveys*. U.S. Dept. Agric. Handbook, vol. 436. U.S. Govt. Printing Office, Washington, DC.
- Soil survey staff, 1999. *Soil taxonomy: a basic system of soil classification for making and interpreting soil surveys*, 2nd ed. U.S. Dept. Agric. Handbook, vol. 436. U.S. Gov. Printing Office, Washington, DC.
- Sugihara, G., 1980. Minimal community structure: an explanation of species abundance patterns. *The American Naturalist* 116, 770–787.
- Tokeshi, M., 1990. Niche apportionment or random assortment: species abundance pattern revisited. *Journal of Animal Ecology* 59, 1129–1146.
- Tokeshi, M., 1993. Species abundance pattern and community structure. *Advances in Ecological Research* 24, 111–186.
- Troech, F.R., Thompson, L.M., 1993. *Soils and Soil Fertility*. Oxford Univ. Press, New York, Oxford, p. 19.
- Williams, C.B., 1964. *Patterns in the Balance of Nature*. Academic Press, London.
- Wilson, J.B., Wells, T.C.E., Trueman, I., Jones, G., Atkinson, M.D., Crawley, M.J., Dodd, M., Silverton, J., 1996. Are there assembly rules for plant species abundance? An investigation in relation to soil resources and successional trends. *Journal of Ecology* 84, 527–538.
- World Conservation Monitoring Centre, 1994. *Biodiversity Data Sourcebook* World Conservation Press, UK.
- Yaalon, D.H., 1998. Discussion of the paper. In: Ibáñez, J.J., et al. (Eds.), *Geoderma*, vol. 83, pp. 193–214.
- Zahl, S., 1977. Jackknifing an index of diversity. *Ecology* 58, 907–913.