Available online at www.sciencedirect.com
SCIENCE DIRECT
Geoderma 134 (2006) 415–427
GEODERMA
www.elsevier.com/locate/geoderma

Inter-scale behavior of balanced entropy for soil texture

José-Manuel Rey [a,*], Eloísa Montero [b], Miguel Angel Martín [c]

[a] Universidad Complutense, Madrid, Spain
[b] Universidad Alfonso X El Sabio, Madrid, Spain
[c] Universidad Politécnica, Madrid, Spain

Available online 24 April 2006

Abstract

The balanced-entropy index – recently proposed by Martín et al. [Martín, M.A., Rey, J.-M., and Taguas, F.J., 2005. An entropy-based heterogeneity index for mass–size distributions in Earth science. *Ecological Modelling*, 182, 221–228.] as a soil texture indicator – can be computed for any arbitrary non-uniform particle size partition provided that associated data are available. In this paper, properties of the index with respect to refinements of the partition are derived. In particular, we analyze how the index value responds when a partition is refined. Variations in the values of the index are shown to be related to the mass splitting in the finer partition. Fine textural data – representing soil volume–size distribution – obtained by laser diffractometry from 70 different types of soils in the Iberian Peninsula are used as a case study to illustrate the theory. The evenness of the underlying distributions is explored at different (finer) size resolutions by computing the balanced-entropy index for the associated partitions. It is observed that, in general, the index values increase when the partition is refined. This is shown to be consistent with the spreading of the mass becoming more uniform when refining the scale. Also, the relative orderings induced by the index computations at different scales do not differ qualitatively, which is significant for classification purposes. Further, for distributions with a continuous probability density function, the index values are shown to approach one when the partition size goes to zero. This is a theoretical property of the index that can be used to test continuity of soil particle size distributions. In our case study, the continuity of the processed distributions cannot be discarded from the analysis.

Keywords: Soil particle size distributions; Entropy

1. Introduction

Particle size distribution (PSD) measurement and associated textural classification is one of the most widely used physical analyses in soil science. Pedotransfer methods have received much attention lately (e.g. Wösten et al., 2001; Pachepsky and Rawls, 2004). They aim at devising functions to infer significant soil properties – notably difficult to measure directly – from simple soil measurements readily accessible in soil databases. The fact that obtaining textural data is relatively easy, together with the valuable information it provides in order to predict other physical properties, accounts for the use of PSD, or its associated parameters, in almost any pedotransfer function. Such functions have been used to predict pore distribution factors (Giménez et al., 2001), to estimate soil water retention (Clapp and Hornberger, 1978; Arya and Paris, 1981; Haverkamp and Parlange, 1986; Tyler and Wheatcraft, 1989; Kravchenko and Zhang, 1998), and also bulk density and permeability. These approaches

* Corresponding author. Postal address: Departamento de Análisis Económico, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain. Tel.: +34 91 394 24 07; fax: +34 91 394 2561.
E-mail address: j-man@ccee.ucm.es (J.-M. Rey).

0016-7061/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2006.03.015

require the PSD to be quantified by means of parameters such as mean particle size, geometric standard deviation or fractal scaling exponents of cumulative distributions.

The heterogeneity of soil particle sizes may play an important role in packing and soil compaction, and consequently it could be related to some of the soil physical properties mentioned above. Textural analysis based on heterogeneity parameters could be useful, not only to characterize soil textures, but also to produce pedotransfer indices to be used for modelling purposes.

Shannon's entropy (Shannon, 1948a,b) enjoys a well established reputation as a natural indicator of the heterogeneity (evenness) of a distribution, e.g. in ecology as an index of biodiversity (MacArthur, 1955; Margalef, 1958), and in economics as a measure of inequality in income distribution (Theil, 1967), or of the concentration of firms within an industry (Hart, 1971).

Soil texture can be thought of in terms of the relative contribution of particle sizes to soil mass, that is, in terms of the *evenness* of the PSD. Thus, entropy appears as a suitable candidate to report texture. However, a naive use of Shannon's formula $-\sum_{i=1}^{N} p_i \log(p_i)$ in the case of PSD – when reported in terms of the soil mass or the soil volume p_i carried by the class intervals ($i=1, 2, \ldots, N$) defined by a partition of the interval of particle sizes – may be misleading. This is because the lengths of the class intervals commonly used to report texture differ wildly, e.g. employing the standard clay, silt, and sand contents results in extremely unequal class sizes of 0.002 mm for clay, 0.048 mm for silt, and 1.95 mm for sand. In this situation, Shannon's entropy gives a distorted measure of evenness in the mass contributions of particle sizes, and in turn may not be trusted as a textural parameter. Moreover, a heterogeneity parameter that takes into account the particle size factor, beyond the corresponding mass contributions, is expected to have a closer correlation with soil physical properties — which in turn could be strongly influenced by such a factor.

In order to solve this problem, a generalization of the Shannon entropy has been recently proposed[1] as a meaningful index of soil texture heterogeneity (Martín et al., 2005). The so-called *balanced-entropy* (BE) index is a parameter with information-theoretic content that can be easily computed from standard textural data, reported using a partition of the considered particle size interval. The BE index has been recently shown to be significant as a pedotransfer input for soil water retention prediction (Martín et al., 2005).

The value of the BE index depends on the prescribed size partition. In this article the behavior of the index with respect to the partition is analyzed. In Section 2 the basic theoretical properties of the balanced-entropy index with regard to partitions are described. Section 2 relies on the theoretical analysis of the index developed in the Appendix[2] to this paper. Section 3 describes the soil data processing and the BE analysis developed for the practical analysis carried out in this study. In Section 4 the results of the analysis are presented and discussed. Conclusions are given Section 5.

2. Balanced entropy: theory

2.1. Basic theory

If a probability distribution P (indicating, say, physical mass or volume) is reported – quantized – by means of a histogram with N class intervals I_i so that $p_i=P(I_i)$ represents the probability mass to the i-th interval, the Shannon formula is defined as

$$H = -\sum_{i=1}^{N} p_i \log(p_i). \tag{1}$$

It is a standard convention that $0 \times \log 0 = 0$. In general, H takes values between 0 and $\log N$. The entropy index is a fair measure of the evenness of the quantized distribution if all class intervals are equal: the larger the value of H the more even is the distribution, the extreme values $H=0$ and $H=\log N$ corresponding, respectively, to the most uneven case (a Dirac delta: the whole mass is supported by only one class interval) and to the most even case (the distribution with the "uniform" property: every class interval carries the same mass fraction).

Thinking of soil texture in terms of the evenness of the soil particle size distribution (PSD) makes entropy an appealing candidate as textural indicator — H being computed from the values of soil mass or soil volume p_i carried by soil particles whose sizes belong to the class interval I_i. A main obstacle with this approach is the huge disparity of the size classes I_i usually employed in soil taxonomy (Soil Conservation Service, 1975). In fact, reporting PSD by using the standard USDA class intervals of clay (sizes below 0.002 mm), silt (sizes 0.002–0.05 mm), and sand (sizes 0.05–2 mm) produces extremely

[1] The index was first introduced by M.A. Martín and J.-M. Rey (2002) "Balancing entropy to evaluate biodiversity, soil texture and economic inequality", preprint.

[2] J.-M. Rey, "Basic quantization properties of balanced-entropy", Appendix.

unequal class sizes of 0.002 mm (clay), 0.048 mm (silt), and 1.95 mm (sand). In this situation, Shannon's entropy does not give an adequate measure of the evenness of the mass contributions, in terms of class particle size.

Martín and Rey proposed to balance Shannon's entropy with a multiplier $E=-\sum p_i \log r_i$ that takes into account differences in class sizes. Here the parameters r_i denote the relative lengths of the class intervals, that is, $r_i=\text{length}(I_i)/\text{length}(I)$, where I is the full interval of sizes considered, typically $I=[0, 2]$ (mm). For instance, in the case of the clay–silt–sand size partition, it is $r_1=0.001$, $r_2=0.024$, and $r_3=0.975$.

The balanced entropy BE is defined by

$$\mathrm{BE}=\frac{H}{E}=\frac{\sum p_i \log(p_i)}{\sum p_i \log(r_i)} \tag{2}$$

This index preserves the main features of Shannon's entropy while being adapted to deal with non-uniform partitions. In general, BE is a normalized index that takes values between 0 and 1 regardless of the number and size of class intervals. The boundary values BE = 0 and BE = 1 correspond, respectively, to the most uneven case (a Dirac delta) and to the most even case (the distribution with the uniform property: every class interval carries a mass equal to its relative size). In the case that $N=2$ and for a given partition of the normalized size interval [0, 1], the distribution can be described by the probability vector $(p_1, 1-p_1)$ so that the values of BE can be represented as a function of p_1. In Fig. 1, where curves for BE vs. p_1 for different partitions have been plotted, it becomes apparent how BE incorporates size-variability together with the heterogeneity in mass assignment as measured by H.

Values of BE for soil particle size distributions defined in terms of clay, silt, and sand USDA contents are represented within the textural triangle in Fig. 2.

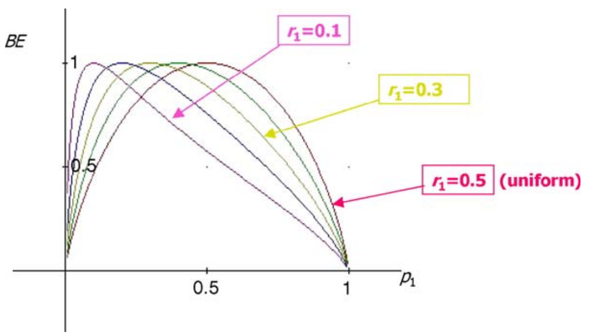

Fig. 1. BE curves for different partitions in the case of $N=2$ class intervals: the y-value gives the BE index computed for the mass distribution $(p_1, 1-p_1)$ and for the size partition $I_1=[0, r_1]$, $I_2=[r_1, 1]$ (sizes are considered normalized in the unit interval).

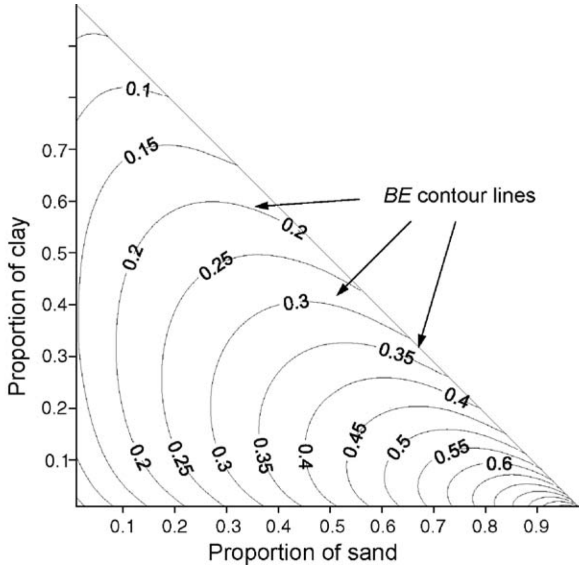

Fig. 2. Contour lines of balanced entropy within the USDA textural triangle.

The values of BE range from zero to one, the latter being approached at the rightmost corner of the textural triangle. Indeed, BE = 1 corresponds to the USDA fractions 97.5% sand, 2.4% silt and 0.1% clay. Of course, for uniform structures (sandy, silty or clayey) BE = 0, which properly indicates their textural homogeneity — the full mass being carried by one class only. Inside the textural triangle, BE typically increases with the increase in sand content. This is so because the sand interval supports – by two orders of magnitude – the widest class of particle sizes. In the practical comparison of textures, BE clearly distinguishes between predominantly sandy and predominantly clayey structures. These and other basic properties of the index BE are discussed in Martín et al. (2005).

2.2. Further properties: inter-scale features of BE

This article is generally concerned with the behavior of BE in respect to the partition, and specifically when the partition is refined. A partition $\prod'$ is finer than (or *refines*) another partition $\prod$ if its class intervals are either class intervals of $\prod$ or subintervals of some class interval of $\prod$. The BE analysis may render significant information about how the mass is globally distributed, e.g. when a class interval is further divided to generate a new (finer) partition.

To interpret properly the variations of the BE index when a partition is refined, some general working principles are to be taken into account. These principles are carefully derived from theory in the Appendix to this paper. Intuition about them, however, may be obtained from the information-theoretical interpretation of the BE formula considered below (details can be found in Martín et al., 2005).

Given a mass distribution P on the (normalized) interval of sizes [0, 1], a size partition $\prod=\{I_i\}$ of [0, 1] induces a discrete distribution (p_i) – the $\prod$-*quantization* of P – defined by $p_i=(P(I_i))$. Let BE($\prod$) denote the value of the balanced-entropy index for the quantizing partition $\prod$. Since $\sum r_i=1$, (r_i) always defines a probability distribution — the (normalized) *size distribution*. The notations $H(\prod)$ or $E(\prod)$ will be used when necessary. It turns out that the balanced-entropy index can be written as

$$\mathrm{BE}(\prod) = \frac{H(\prod)}{H(\prod) + d_\Pi(p_i||r_i)}, \tag{3}$$

where $d_\prod(p_i||r_i)$ is the Kullback–Leibler distance between the mass distribution (p_i) and the size distribution (r_i), that obviously depends on the partition $\prod$. Notice that BE($\prod$)$\approx$0 if and only if $H(\prod)\approx 0$, so that the $\prod$-*quantization* of P is a Dirac delta on $\prod$. Furthermore, BE($\prod$)$\approx$1 if, and only if, $d_\prod(p_i\|r_i)\approx 0$, which occurs if and only if $p_i\approx r_i$ (see Cover and Thomas, 1991), that is, the $\prod$-*quantization* of P is nearly uniform on $\prod$, because the mass of each interval of the partition is given by its length. It is thus convenient to think of the index BE as a sort of distance from the quantized distribution (p_i) to the distribution with the uniform property (r_i) — with respect to the quantizing partition $\prod$. As a particular case, if $\prod$ is the standard clay–silt–sand size partition, the higher the value of BE inside the textural triangle the more uniformly spread is the mass across the partition, i.e. the closer each mass fraction p_i is to its size supporting value r_i for every i.

Building on the interpretation above and on the fact that BE depends continuously on the p_i's, key inter-scale principles for BE can be formulated as follows:

#1. Small values of the index BE are consistent with P being *nearly discrete*, i.e. concentrated at a finite number of sizes.

#2. A lowering of the value of BE when the partition is refined is consistent with the measure spread inside the new class intervals being far from uniform (e.g., some size interval getting no mass in the mass splitting).

#3. Near-to-one values of the index BE are consistent with P being *nearly uniform*, every class interval supporting a mass approximately proportional to its size.

#4. An increase in the value of BE when the partition is refined is consistent with the measure spread for the new partition being close to uniform (every class interval getting nearly the mass share proportional to its size).

#5. BE-values approaching one when refining the size partitions is consistent with underlying distributions having continuous probability density functions (e.g. the normal or lognormal distributions).

#6. BE-values approaching a constant value below one when refining the partition is consistent with a fractal underlying distribution.

Rules #1 to #5 are general principles for the balanced-entropy index. They are mathematically established from the theory of BE in the Appendix by Rey to this paper. Proposition #5 can be derived from #4 and the fact that, for a distribution with a continuous probability density, the spreading of the mass is nearly uniform inside intervals of sufficiently fine partitions. That is, every interval gets a probability mass approximately proportional to its length (see Appendix). Property #5 is an interesting feature of the index that may be used as a test for continuity of mass distributions. In particular, it may contribute to the discussion on the singular or continuous nature of PSD (Buchan et al., 1993; Caniego et al., 2001).

As stated in rule #6, it is also plausible that a sequence of computed indexes BE($\prod_k$) accumulates around a certain value other than zero or one. According to the heuristics above, that would mean that the BE-"distance" to the uniform splitting remains approximately constant, which in turn would indicate that the mass splitting follows a scale-invariance rule — with respect to the resolutions defined by the partition sequence. This may happen, for example, when P is a fractal distribution. In fact, for a self-similar distribution a suitable sequence of partitions $\prod_k$ can be chosen so that BE($\prod_k$) remains constant for every k. Moreover, this constant value coincides with the fractal dimension of the distribution (Martín et al., 2001). See Appendix for further details.

3. Material and methods

3.1. Soils

Seventy mineral soils from Sierra del Segura and Sierra de Cazorla (Jaén, Spain) were chosen corresponding to different textural classes and with low organic matter content. Our samples belong to ten textural classes, following the USDA classification of soils. The most common texture was Clay (36 samples) followed by Clay loam and Sandy loam (7), and Sandy

Table 1
Sample output of the Longbench Mastersizer S, with the percentages of volume quantized according to the partition of the apparatus (the so-called Malvern partition in this paper)

Size interval I_i (μm)	Volume in %	Size interval I_i (μm)	Volume in %	Size interval I_i (μm)	Volume in %	Size interval I_i (μm)	Volume in %
0.05–0.06	0.00	0.81–0.97	0.80	13.18–15.69	2.14	213.95–254.66	1.93
0.06–0.07	0.00	0.97–1.15	0.91	15.69–18.67	1.82	254.66–303.12	0.82
0.07–0.08	0.00	1.15–1.37	1.12	18.67–22.22	1.62	303.12–360.81	0.12
0.08–0.10	0.01	1.37–1.63	1.39	22.22–26.45	1.65	360.81–429.46	0.02
0.10–0.12	0.02	1.63–1.94	1.67	26.45–31.49	1.74	429.46–511.19	0.04
0.12–0.14	0.04	1.94–2.31	1.99	31.49–37.48	1.97	511.19–608.46	0.09
0.14–0.17	0.10	2.31–2.75	2.31	37.48–44.61	2.32	608.46–724.24	0.14
0.17–0.20	0.22	2.75–3.27	2.60	44.61–53.10	2.79	724.24–862.06	0.13
0.20–0.24	0.44	3.27–3.89	2.84	53.10–63.20	3.37	862.06–1026.10	0.07
0.24–0.29	0.71	3.89–4.63	3.02	63.20–75.23	4.05	1026.10–1221.36	0.02
0.29–0.34	0.92	4.63–5.52	3.12	75.23–89.55	4.83	1221.36–1453.77	0.00
0.34–0.40	0.94	5.52–6.57	3.14	89.55–106.59	5.58	1453.77–1730.41	0.00
0.40–0.48	0.91	6.57–7.81	3.14	106.59–126.87	5.96	1730.41–2059.69	0.00
0.48–0.57	0.86	7.81–9.30	3.06	126.87–151.01	5.66	2059.69–2451.63	0.00
0.57–0.68	0.76	9.30–11.07	2.84	151.01–179.75	4.71	2451.63–2918.16	0.00
0.68–0.81	0.73	11.07–13.18	2.51	179.75–213.95	3.33	2918.16–3473.45	0.00

Figures in each even column represent soil volume percentages supported by the size interval $I_i=[\phi_{i-1},\phi_i]$ defined by the two neighbor figures in the column on its left side.

clay loam (6). Other textures were Loam (4), Silt loam (4), Silty clay loam (2), Sand (2), Loamy sand (1) and Sandy clay (1). Soils were under different management systems at the time the samples were collected, i.e., cultivation, rotary grazing, grazing and forestry.

Soils were collected and transported in plastic bags to the laboratory where they were dried at room temperature, large clods in the air-dry state were broken, and sieved to separate coarse materials. Soil samples were analyzed by laser diffraction after the aggregates were dispersed by stirring and ultrasonics lasting 5 min.

3.2. Laser diffraction processing

Soil samples were analyzed with the laser diffraction technique using a Longbench Mastersizer S (Malvern

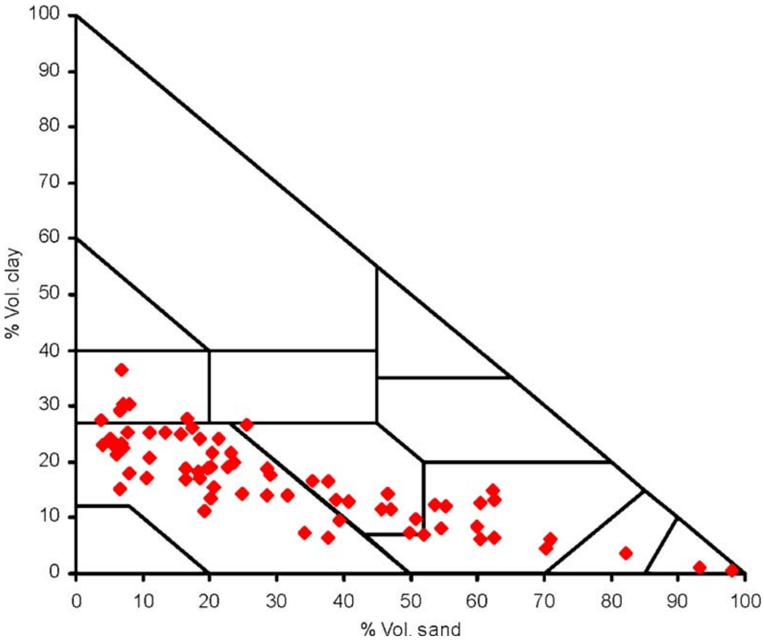

Fig. 3. The 70 soil samples considered in the paper represented inside a textural triangle (in percentages of soil volume by clay–silt–sand size fractions).

Table 2
Malvern equivalent intervals, normalized relative size (r_i) and volume (p_i) distributions for some USDA classes of particle sizes

Textural fraction	Interval of sizes (μm)	Malvern equivalent interval (μm)	Normalized length of interval (r_i)	Normalized relative volume (p_i)
Clay	<2	0.05–2.3	0.001	$p=\sum_{1}^{22} p_i$
Silt	2–50	2.3–53.1	0.025	$p=\sum_{23}^{40} p_i$
Sand	50–2000	53.1–2059.7	0.974	$p=\sum_{41}^{61} p_i$
Very fine sand	50–100	53.1–106.6	0.026	$p=\sum_{41}^{44} p_i$
Fine sand	100–250	106.6–254.7	0.072	$p=\sum_{45}^{49} p_i$
Medium sand	250–500	254.7–511.2	0.124	$p=\sum_{50}^{53} p_i$
Coarse sand	500–1000	511.2–1026.1	0.250	$p=\sum_{54}^{57} p_i$
Very coarse sand	1000–2000	1026.1–2059.7	0.502	$p=\sum_{58}^{61} p_i$

Instruments, Malvern, England) that employs a 5-mW He–Ne laser with a wavelength of 632.8 nm as a light source. Measurements were taken covering the interval of sizes 0.05–3474 μm. The laser beam is diffracted by soil particles in different angles according to their size. Light detectors receive the diffracted light and data are collected as shown in Table 1. Histogram data are produced from each analysis that represent relative volume (%) vs. soil particle diameter (μm) across 64 size subintervals. For instance, for the sample processed in Table 1, 0.01% of the total volume of measured soil particles corresponds to particles with diameters between 0.08 and 0.10 μm.

Fig. 3 displays the 70 samples inside a textural triangle in which each sample is represented by the percentages of soil volume carried by the clay–silt–sand class sizes.

3.3. Computing the BE-indices

The laser diffraction technique supplies a fixed non-uniform partition of the size interval $I=[0.05, 3473.5]$ (μm) into 64 sub-intervals $I_i=[\phi_{i-1}, \phi_i]$, $i=1, 2, \ldots, 64$, whose endpoints satisfy $\log\phi_i - \log\phi_{i-1} \simeq$ constant (see Table 1 for the entire sequence ϕ_i). While the first interval is $I_1=[0.05, 0.06]$, with length 0.01 μm, the last interval is $I_{64}=[2918.2, 3473.5]$, with length 555.3 μm. The raw data produced by the apparatus are relative volume values $V_1, V_2, \ldots, V_{64}$, expressed as percentages of the total volume, i.e., $\sum_{i=1}^{64} V_i = 100$, corresponding to the 64 subintervals of sizes. The value V_i corresponding to subinterval I_i is the percentage of total soil volume contributed by particles with diameters within I_i (see Table 1).

In the present analysis, we considered the interval of particle sizes $I=[0.05, 2059.7]$ (μm), corresponding to 61 – out of the 64 – size intervals, that we call "the Malvern partition". Notice that this produces a highly non-uniform size partition, that fits well into the BE theory. The associated relative volumes V_i, $i=1, 2, \ldots, 61$,

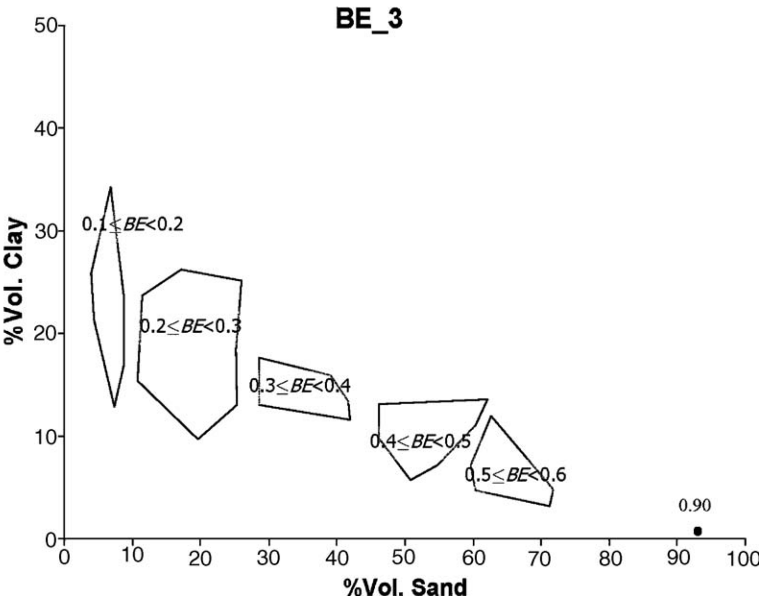

Fig. 4. Regions obtained by grouping the BE_3 values for the 70 samples.

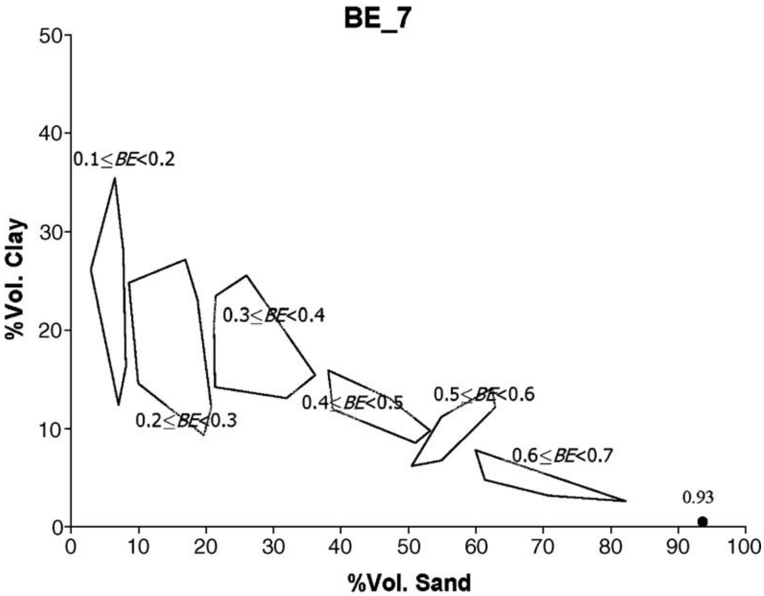

Fig. 5. Regions obtained by grouping the BE_7 values for the 70 samples.

were normalized to define the volume–size distribution $p_i = \frac{V_i}{\sum_{i=1}^{61} V_i}$. The lengths of intervals (ϕ_i–ϕ_{i-1}) were normalized to [0, 1] to represent sizes in relative terms. As above, r_i denotes the normalized length of the interval I_i, so that $\sum_{i=1}^{61} r_i = 1$.

Normalized relative volumes p_i corresponding to adjacent subintervals can be added to obtain the normalized relative volume of a larger interval of sizes. Because of the particular structure of the Malvern partition, standard particle size classes, say the clay fraction, can be obtained only approximately. Thus, using raw data from laser diffraction analysis, the interval of sizes equivalent to the clay fraction is $I_{\text{clay}} =$ [0.05, 2.3], which is obtained as the union of the first 22 subintervals given by the Longbench Mastersizer. Normalized relative volume corresponding to the clay

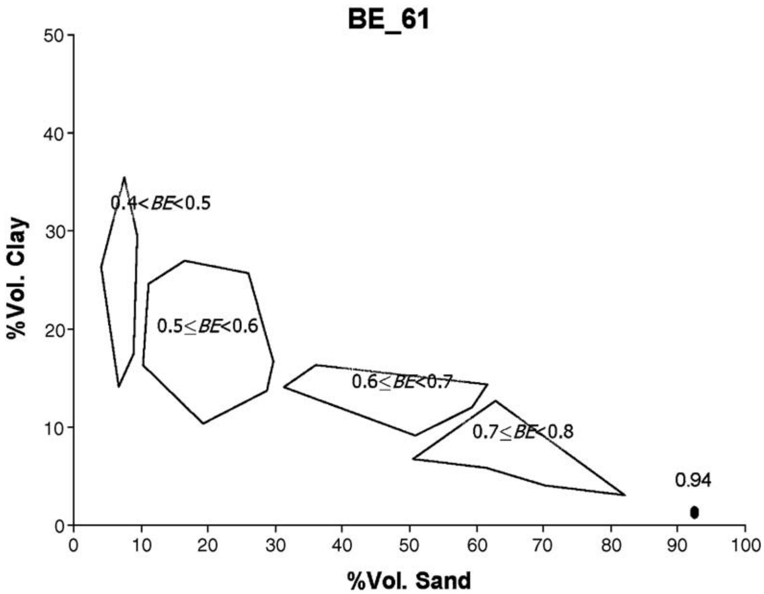

Fig. 6. Regions obtained by grouping the BE_61 values for the 70 samples.

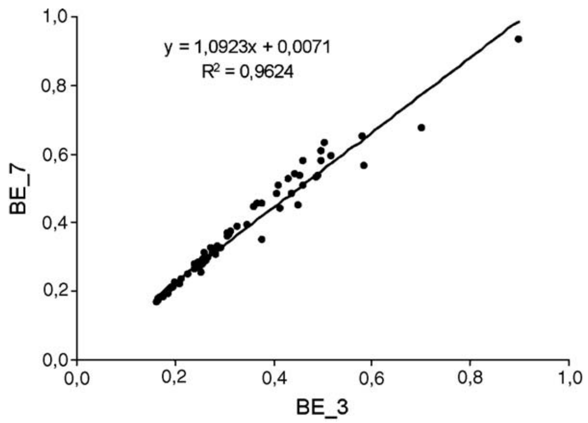

Fig. 7. Regression analysis for the indices BE_3 and BE_7.

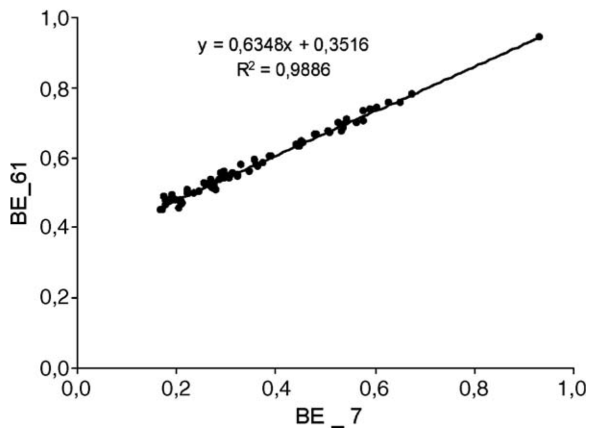

Fig. 9. Regression analysis for the indices BE_7 and BE_61.

fraction (p_{clay}) is in turn given by $p_{\text{clay}} = \sum_{i=1}^{22} p_i$. The Malvern equivalent size intervals, normalized relative volumes (p_i) and normalized length of intervals (r_i) corresponding to other USDA classes of particle sizes are shown in Table 2.

To perform an inter-scale BE-analysis of our textural data, we calculate the BE-index using Eq. (2) for three different partitions:

i) The partition $\prod_3$ defined by dividing the interval of sizes I into the 3 Malvern classes equivalent to clay–silt–sand in the USDA particle sizes classification.
ii) The partition $\prod_7$ obtained by partitioning I in 7 classes: the Malvern equivalent with clay–silt–very fine sand–fine sand–medium sand–coarse sand–very coarse sand following USDA particle sizes.

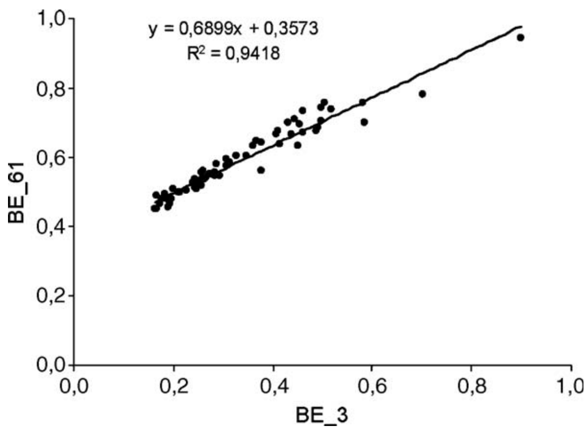

Fig. 8. Regression analysis for the indices BE_3 and BE_61.

iii) The Malvern 61-partition, i.e. the full partition into 61 size intervals obtained from Longbench Mastersizer S as described above.

Notice that the Malvern partition refines $\prod_7$ which in turns refines $\prod_3$. The BE values obtained from the three partitions defined above were denoted by BE_3, BE_7 and BE_61, respectively.

4. Results and discussion

The computed values of the BE_3 index for the 70 analyzed soil samples are represented in Fig. 4 within the textural triangle and grouped in regions.

Notice that the textural triangle is defined in terms of clay–silt–sand volume fractions rather than in terms of the usual mass fractions. Figs. 5 and 6 display similar regions classified according to the values of the indices BE_7 and BE_61. The high values of BE – around 0.93 – at the rightmost corner in each figure correspond to a soil sample with 94% sand content.

A shift to larger values of the BE_61 index for each sample is apparent, indicating that, at smaller scales, the soil volume is distributed more uniformly across class sizes.

It is significant that Figs. 4, 5 and 6 display the same qualitative representation. This implies that the three BE indices render the same ordered classification of the sample, and in turn that the fineness of the partition appears unimportant for comparative purposes. This enhances the role of the simpler BE_3 as a parameter to characterize textures. To establish this claim more consistently, we run regression analyses between the three indices. These are shown in Fig. 7 (BE_7 vs. BE_3), Fig. 8 (BE_61 vs. BE_3) and Fig. 9 (BE_61 vs. BE_7).

A remarkable positive correlation between indices is evident in all three plots, with the relationship

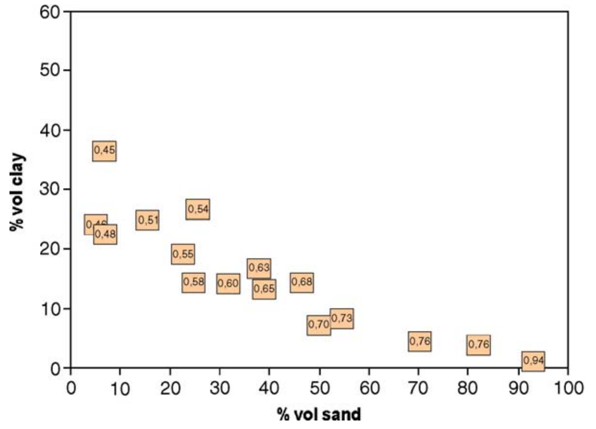

Fig. 10. A 16-subsample represented inside the textural triangle with their corresponding BE_3 attached.

BE_61 vs. BE_7 displayed in Fig. 9 being the most linear. The claimed increase in the index values when considering finer partitions is also evident from the figures.

These facts, namely (a) the observed systematic increase in index values when the partition is refined, and (b) the three computed indices comparing textures in the same manner, are by no means obvious. As explained in Section 2, when (a) occurs this is a symptom of a more uniform distribution of the soil volume across different class sizes when size resolution increases. Also, (b) may or may not be the case, as the following simple example proves. Consider $N=2$, a uniform partition $\prod_1$ composed of [0, 0.5] and [0.5, 1], and the finer partition $\prod_2$ defined by the class intervals [0, 0.5], [0.5, 0.75] and [0.75, 1]. Denote for the time being by $\mathrm{BE}(P(\prod))$ the BE index of a $\prod$-quantized distribution P. If P and P' are distributions such that their $\prod_1$-quantizations are (0.4, 0.6) and (0.5, 0.5), respectively, we have that the ordering $\mathrm{BE}(P(\prod_1))=0.97<1=\mathrm{BE}(P'(\prod_1))$. Besides, if their corresponding $\prod_2$-quantizations are (0.4, 0.3, 0.3) and (0.5, 0.0.5), it turns out that $\mathrm{BE}(P(\prod_2))=0.71>0.67=\mathrm{BE}(P'(\prod_2))$, giving the reversed order.

In order to illustrate the use of the BE index to test the continuity of the relative volume–size distributions, a subsample of 16 soils was selected. This subsample was chosen to cover a wide region inside the textural triangle so that the corresponding BE_3 index values ranged from 0.45 to 0.94 (see Fig. 10).

The BE index was computed for the three partitions defined above and for an intermediate partition with 9 subintervals, finer than $\prod_7$. The values of the BE indices for each sample increased with the refinement of each considered partition. Fig. 11 shows the mean values of the four BE indices together with their standard deviations.

The increase in the mean BE value is more pronounced when the refinement of the partition is finer. For instance, the average increases when passing from $\prod_3$ to $\prod_7$ in 0.037, whereas from $\prod_7$ to the 61-Malvern partition it goes up by 0.22. As explained in Section 2, this behavior is compatible with the continuity of the distribution — the BE index approaching 1 as the partition gets finer. The index values obtained for the 61-Malvern partition, however, might be thought not to be close enough to 1. This fact may be due to the huge disparity in interval lengths of the Malvern partition (the ratio length (I_{61})/length (I_1) being equal to 32928). In order to test further the continuity of the distribution, only the first 21 size classes of the Malvern partition were considered by analyzing the distribution within the interval of sizes [0.05, 194] (μm). The ratio length (I_{21})/length (I_1) is then only 31. The distribution of relative volumes inside the new smaller size interval is normalized and the associated BE index (BE_21) are

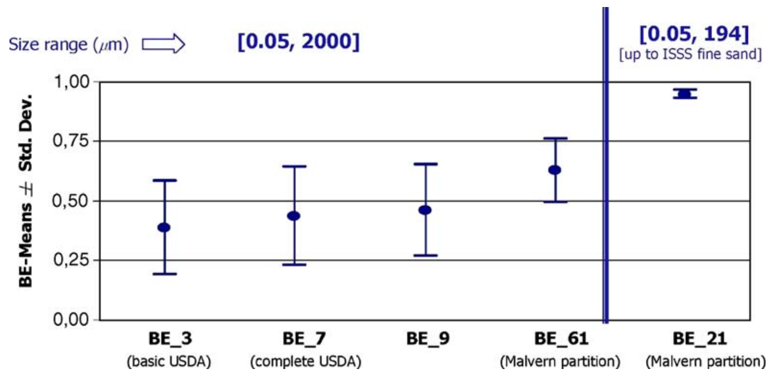

Fig. 11. Inter-scale analysis of BE for a sample: means and standard deviations of BE indices computed for different partition represented.

computed for the 16-subsample. The results are shown in Fig. 11. A substantial increase in the BE index is apparent for every sample, with the mean values of BE_21 approaching unity (indeed, the mean value is 0.948 and the standard deviation is 0.0184). These results are consistent with the volume–size distributions being continuous within the size interval [0.05, 194].

5. Conclusions

The balanced entropy (BE) – obtained from the standard clay–silt–sand soil fraction content – was proposed in Martín et al. (2005) to characterize soil texture. Balanced entropy, however, is a flexible parameter that can be computed for an arbitrary partition of the interval of soil particle sizes. The behavior of BE with respect to the considered partition was addressed in this paper. In particular, the theoretical properties of the BE index were considered when the partition is refined, and the relationship between extreme values of this index and the nature of the underlying distribution was discussed. The variations of BE when refining the scale were explained in terms of the uniformity in the mass spreading. Also, it was argued that, for continuous distributions, the BE index values approaches unity as the partition gets finer.

The methodology was applied to a sample of 70 soil samples from the Iberian Peninsula. Significant conclusions that can be drawn from the analysis are as follows. First, for all samples the soil volume is more uniformly distributed across sizes when smaller scales are considered. Secondly, different BE indices – i.e. computed with respect to different partitions – play a role qualitatively similar as a parameter for comparing textures. Third, the continuous nature of the relative soil volume–size distribution cannot be discarded from the analysis.

As a general conclusion, balanced entropy is shown to be a useful tool to scrutinize the spreading of a given mass distribution within different scales — associated with size partitions. In turn, different BE indices may be used as textural indicators supplying inter-scale information when appropriately disaggregated soil data is available.

Acknowledgments

This research was partially supported by Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I) under ref. AGL2004-04079 AGR. Spain.

Appendix A. Basic quantization properties of balanced entropy

José-Manuel Rey[3,4]

Summary. Some basic quantization facts for the balanced-entropy index, introduced by Martín et al. (2005) are derived from theory. Specifically, mechanisms for mass partitioning are described that are consistent with the increase or decrease of index values when refining the size partition. Variations in index values are shown to respond to the uniformity of the mass spreading. A key result is that the index values approach one when the partition size goes to zero. Also, values of the index approaching a constant between 0 and 1 are shown to be consistent with an underlying fractal distribution.

Shannon's entropy has been successfully established in different fields as a useful heterogeneity index of a probability distribution. Balanced entropy is a natural generalization of Shannon's entropy introduced by Martín and Rey[5] as a measure of the evenness of a distribution with respect to a range of unevenly classified sizes. Consider the unit interval [0, 1] as the (normalized) interval of sizes and let $\prod=\{I_i\}$ be a (finite) size partition of [0, 1], that is, $\cup_i I_i=[0, 1]$ and $I_i \cap I_j=\varnothing$ for different i and j. Given a mass distribution P defined on the size interval, the partition $\prod$ induces a discrete distribution (p_i) defined by the probability vector $p_i=(P(I_i))$ that we call $\prod$-*quantization*[6] of P. Let $\mathrm{BE}(\prod)$ denote the value of the balanced-entropy index for the quantizing partition $\prod$:

$$\mathrm{BE}(\prod) = \frac{\sum P(I_i)\log P(I_i)}{\sum P(I_i)\log r_i}, \tag{A1}$$

where r_i is the length of the size interval I_i. Note that $\sum r_i=1$, so that (r_i) defines a probability distribution on the integer set $\{1, 2, \ldots, \#\prod\}$. The notations $H(\prod)$ or $E(\prod)$ will be used when necessary. The value $\mathrm{BE}(\prod)$ depends on the partition $\prod$. Privileged partitions do exist in some contexts. As an important example, for the classification of soil textures, the induced partition when

[3] Departamento de Análisis Económico, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain. E-mail: j-man@ccee.ucm.es, Tel.: +34 91 394 24 07; fax: +34 91 394 2561.

[4] Research partially supported by by Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I) under ref. BFM 2003-08204. Spain.

[5] Martín and Rey, submitted for publication.

[6] *Quantization* is the division of a quantity into a discrete number of small parts. The oldest application of quantization is estimating densities by histograms.

using the standard USDA system is defined by $r_1 = r_{clay} = 0.001$, $r_2 = r_{silt} = 0.024$, and $r_3 = r_{sand} = 0.975$ (see Soil Conservation Service, 1975). The value BE($\prod$) may be thought of as a measure of distance from P to the uniform distribution defined in $\prod$, for which each size interval gets a probability mass equal to its length (see Martín et al., 2005). In general, considering different partitions and computing the associated BE index gives significant information on the mass spreading of the distribution at different scales. As it is shown below (Claims 4 and 5), the index value may increase or decrease when the partition is refined. In this note basic mechanisms for the redistribution of the mass are formulated that are compatible with an observed increase or decrease of the index when the partition is refined.

For any distribution and partition, the index BE takes values in [0, 1]. First the occurrence of extreme values of BE is considered.

Claim 1. BE($\prod$)=0 for any partition $\prod$ $\Leftrightarrow$ The distribution P is a Dirac delta, i.e. the whole probability mass is concentrated at some size r_0.

It is clear that BE($\prod$)=0 if and only if $H(\prod)=0$ which occurs if and only if the $\prod$-quantized distribution satisfies $P(I_i)=1$ for some I_i. This fact occurs for every partition if and only if P is a Dirac delta, thus implying Claim 1.

The diameter of a partition $\prod = \{I_i\}$ is defined as diam$\prod$ = max {length(I_i)}.

Claim 2. If P is *discrete* – a (finite) sum of Dirac deltas – i.e. $P = \sum m_i \delta_{r_i}$ with $\sum m_i = 1$ and δ_{r_i} is a unit mass located at r_i, then BE($\prod$)→0 as diam$\prod$→0.

For any sufficiently fine partition $\prod$, we have $H(\prod) = -\sum m_i \log m_i$ = constant, whereas $E(\prod)$ tends to infinity as diam$\prod$ goes to zero. As a consequence, Claim 2 follows.

Claim 3. BE($\prod$)=1 for any partition $\prod \Leftrightarrow P$ is the uniform distribution.

The assertion BE($\prod$)=1 can be rewritten as $\sum p_i \log \frac{p_i}{r_i} = 0$. This means that the Kullback–Leibler distance between the $\prod$-quantized distribution (p_i) and the size distribution (r_i) is zero. This occurs if and only if $p_i = r_i$ (see Cover and Thomas, 1991). Since this is true for any arbitrary partition, P is the uniform distribution.

Next the response of the index BE is analyzed when the partition is refined. A partition $\prod'$ *refines* or is finer that another partition $\prod$ (denoted by $\prod \prec \prod'$) if its class intervals are either class intervals of $\prod$ or are subintervals of some class interval of $\prod$. It is well-known that the Shannon entropy index H does not decrease when the partition is refined. That is, $H(\prod) \le H(\prod')$ if $\prod \prec \prod'$ (see e.g. Gray, 1990). Also, the value of E increases when the partition is refined. To see this, consider a particular case when $\prod'$ is obtained from $\prod$ by partitioning just one class interval I of length r of $\prod$ into two subintervals, I_1, I_2, of lengths r_1 and $r_2 = r - r_1$. Without loss of generality, assume that $0 < r_1 \le \frac{r}{2}$. Assume also that the underlying distribution P splits the probability $p = P(I)$ into $p_1 = P(I_1)$ and $p_2 = P(I_2) = p - p_1$. Since $-(p_1 + p_2)\log r \le -p_1 \log r_1 - p_2 \log r_2$, we have that $E(\prod) \le E(\prod')$ in this case. This can be extended to prove that, in general, $E(\prod) \le E(\prod')$ if $\prod \prec \prod'$. Since both $H(\prod)$ and $E(\prod)$ increase, the value of BE($\prod$) may or may not increase when $\prod$ is refined. This depends on how the relative variations of both quantities compare when $\prod$ is refined: it is easy to check that

$$\mathrm{BE}(\prod) \le \mathrm{BE}(\prod') \Leftrightarrow \frac{H(\prod') - H(\prod)}{H(\prod)} \ge \frac{E(\prod') - E(\prod)}{E(\prod)}. \tag{A2}$$

As a consequence, the general mechanism in the mass spreading, producing an increase in the index BE, consists in redistributing the probability mass within finer partitions in such a way that the relative increase of the entropy exceeds that of E — the average of the logarithms of interval sizes. As shown next, however, a highly non-uniform spreading of the probability mass inside the finer partition $\prod'$ is compatible with a lowering of the value of BE($\prod$).

Consider again the case when $\prod'$ is obtained from $\prod$ by partitioning a class interval I of length r into two subintervals, I_1, I_2, of lengths $0 < r_1 < r/2$ and $r_2 = r - r_1$. Let $p = P(I)$ and $p_1 = P(I_1)$ and $p_2 = P(I_2) = p - p_1$.

Claim 4. For any $\prod'$ refining $\prod$ as above so that $p_1 = 0$, BE($\prod'$)<BE($\prod$).

This follows from Eq. (A2) above since $H(\prod') = H(\prod)$ and $E(\prod') > E(\prod)$.

If the mass spreading across the subintervals I_1 and I_2 is uniform the value of the index BE goes up. This is the content of

Claim 5. For any $\prod'$ refining $\prod$ as above, in such a way that the mass splitting (p_1, p_2) is uniform, that is, $p_1 = p\frac{r_1}{r}$ and $p_2 = p\frac{r - r_1}{r} =$, we have BE($\prod'$)>BE($\prod$).

To check Claim 5, denote, $\delta = \frac{r_1}{r}$ for convenience. After some algebra:

$$\begin{aligned}\mathrm{BE}(\prod') &= \frac{H(\prod) + p\log p - \delta\, p\log(\delta\, p) - (1-\delta)p\log((1-\delta)p)}{E(\prod) + p\log r - \delta\, p\log(\delta r) - (1-\delta)p\log((1-\delta)r)} \\ &= \frac{H(\prod) + p\{-\delta\log\delta - (1-\delta)\log(1-\delta)\}}{E(\prod) + p\{-\delta\log\delta - (1-\delta)\log(1-\delta)\}} = \frac{H(\Pi) + pH_2(\delta)}{E(\Pi) + pH_2(\delta)},\end{aligned}$$

where $H_2(\delta)=-\delta\log\delta-(1-\delta)\log(1-\delta)$ is the Shannon entropy of the distribution $(\delta, 1-\delta)$. Since $H_2(\delta)>0$, it is always the case that[7]

$$\mathrm{BE}(\prod') = \frac{H(\prod)+pH_2(\delta)}{E(\prod)+pH_2(\delta)} > \frac{H(\prod)}{E(\prod)} = \mathrm{BE}(\prod),$$

which implies Claim 5.

An arbitrary refinement of a partition $\prod$ can be obtained in successive steps. At each step, some size interval of length r is partitioned into two subintervals of lengths r_1 and $r-r_1$, according to a choice of the parameter $\delta=\frac{r_1}{r}$ $(0<\delta<1/2)$. Refining $\prod$ in this way amounts to selecting a sequence of δ's, one for each step in which an original interval is split in two. For a sequence of refining partitions with decreasing diameters, a *partitioning rule* consists of selecting – at each partitioning step – some δ, $0<\delta<1/2$. The next one is a key result concerning the behavior of the index BE when partitions are refined.

Claim 6. If P is a distribution with a continuous probability density, then $\mathrm{BE}(\prod)\to 1$ as $\mathrm{diam}\prod\to 0$, provided that the partitioning rule satisfies $\delta>\delta_0>0$.

Let $\prod_1=\{I_i^1: i=1, 2, \ldots, N\}$ be an initial partition of [0, 1] and let $(p_i^1=P(I_i^1))$ be the induced $\prod_1$-quantized distribution. Consider some sequence of nested partitions $\{\prod_k\}$, where, $\prod_{k+1}$ refines $\prod_k$ for each k. Since the diameter of $\prod_k$ goes to zero as k increases, we may take $\prod_{k+1}$ as a refinement of $\prod_k$ obtained by dividing each class interval I_i^k of $\prod_k$ – using a partitioning rule δ_i^k – into two subintervals (that are themselves class intervals of $\prod_{k+1}$). Since P has a continuous probability density, it can be assumed that $\prod_1$ is fine enough so that the masses (p_i^1) will be split nearly uniformly inside the class intervals of the new partition $\prod_2$. This means that each new class interval gets a probability mass approximately proportional to its length, as stated in Claim 5.

In the first step – in which each class interval of $\prod_1$ is divided into two using partitioning rules δ_i^1 –, repeating the procedure used in the proof of Claim 5 gives:

$$\mathrm{BE}(\Pi_2)\approx\frac{H(\prod_1)+p_1^1H_2(\delta_1^1)+p_2^1H_2(\delta_2^1)+\ldots+p_N^1H_2(\delta_N^1)}{E(\prod_1)+p_1^1H_2(\delta_1^1)+p_2^1H_2(\delta_2^1)+\ldots+p_N^1H_2(\delta_N^1)}.$$

using the fact that BE depends continuously on the p_i so that the value of $\mathrm{BE}(\prod_2)$ is only approximately equal to the expression above. Since the partition rule is bounded from below by δ_0 and $H(\delta)$ is continuous and increasing for $0<\delta<1/2$, it holds that $H_2(\delta_i^1)>H_2(\delta_0)$ for any i. Therefore (see footnote 3),

$$\mathrm{BE}(\prod_2) > \frac{H(\prod_1)+H_2(\delta_0)\sum p_i^1}{E(\prod_1)+H_2(\delta_0)\sum p_i^1} = \frac{H(\prod_1)+H_2(\delta_0)}{E(\prod_1)+H_2(\delta_0)}$$

Repeating the argument for partition $\prod_2$ – using partitioning rules δ_i^2 and calling $P(I_i^2)=p_i^2$ – gives

$$\begin{aligned}\mathrm{BE}(\Pi_3)&\approx\frac{H(\prod_2)+\sum_{i=1}^{2N} p_i^2H_2(\delta_i^2)}{E(\prod_2)+\sum_{i=1}^{2N} p_i^2H_2(\delta_i^2)}\\ &=\frac{H(\prod_1)+\sum_{i=1}^{N} p_i^1H_2(\delta_i^1)+\sum_{i=1}^{2N} p_i^2H_2(\delta_i^2)}{E(\prod_1)+\sum_{i=1}^{N} p_i^1H_2(\delta_i^1)+\sum_{i=1}^{2N} p_i^2H_2(\delta_i^2)}\\ &>\frac{H(\Pi_1)+H_2(\delta_0)\Sigma_i p_i^1+H_2(\delta_0)\Sigma_i p_i^2}{E(\prod_1)+H_2(\delta_0)\Sigma_i p_i^1+H_2(\delta_0)\Sigma_i p_i^2}\\ &=\frac{H(\prod_1)+2H_2(\delta_0)}{E(\prod_1)+2H_2(\delta_0)}.\end{aligned}$$

Repeating the argument for partition $\prod_{k+1}$,

$$\mathrm{BE}(\prod_{k+1}) > \frac{H(\prod_1)+kH_2(\delta_0)}{E(\prod_1)+kH_2(\delta_0)}$$

The terms on the right hand side form an increasing sequence accumulating at unity. This argument justifies Claim 6. It also works when P has a density which is continuous only in a small interval I. Claim 6 is thus also valid for continuous densities with a high-peak and very small standard deviation. The rate of convergence of BE to 1 – when the diameter of the partitions goes to zero – may thus be used to differentiate between different continuous distributions. Moreover, if P has non-trivial, singular and continuous parts, it will also result in a BE index approaching 1 when the partition is fine enough within the support of the continuous part. It follows from Claim 6 that a necessary condition for BE to approach a value $d<1$ is that P be purely singular (e.g. fractal).

Since $\mathrm{BE}\approx 1$ for distributions with continuous densities whereas $\mathrm{BE}\approx 0$ for nearly discrete distributions, it may be thought that observing that BE approach intermediate values corresponds to more complex singular distributions. It may well happen that a certain sequence $\mathrm{BE}(\prod_k)$ stabilizes around some value fixed value d. This is actually the case for fractal (selfimilar) distributions. This claim can be illustrated with a standard Cantor distribution P, defined in the following way. Consider a partition $\prod_1$ of [0, 1], select its first and last subintervals, call them $I_1=[0, r_1]$, $I_2=[1-r_2, 1]$, and spread the probability mass by $P(I_1)=p_1$, and $P(I_2)=p_2$, $p_1+p_2=1$. Repeat the same procedure inside I_1, and I_2, i.e., define $P(I_{1,1})=p_1p_1$, $P(I_{1,2})=p_1p_2$, $P(I_{2,1})=p_2p_1$, $P(I_{2,2})=p_2p_2$, where, for $j=1, 2$, $I_{i,j}$ is a subinterval of I_i

[7] The fact that $\frac{a+x}{b+x}>\frac{a+y}{b+y}\Leftrightarrow x>y$ for non-negative a,b,x,y, $a<b$, $b\neq 0$, is often used. Take $y=0$ here.

with length $(I_{i,j})=r_ir_j$. Let $\prod_2$ be any partition of [0, 1] containing the subintervals $I_{i,j}$. At the k-th stage of the construction, there are 2^k subintervals $I_{i_1, i_2, \ldots, i_k}$, such that $P(I_{i_1, i_2, \ldots, i_k})=p_{i_1}p_{i_2}\ldots p_{i_k}$ and length $(I_{i_1, i_2, \ldots, i_k})=r_{i_1}r_{i_2}\ldots r_{i_k}$, $i_j=1, 2$. Let $\prod_k$ be any partition of [0, 1] containing $I_{i_1, i_2, \ldots, i_k}$ as class intervals. A computation gives

$$\begin{aligned}\mathrm{BE}\left(\prod_k\right) &= \frac{\sum_{i_1,i_2,\ldots,i_k} p_{i_1}p_{i_2}\cdots p_{i_k}\log(p_{i_1}p_{i_2}\cdots p_{i_k})}{\sum_{i_1,i_2,\ldots,i_k} p_{i_1}p_{i_2}\cdots p_{i_k}\log(r_{i_1}r_{i_2}\cdots r_{i_k})}\\ &= \frac{k(p_1\log p_1+p_2\log p_2)}{k(p_1\log r_1+p_2\log r_2)}\\ &= \frac{p_1\log p_1+p_2\log p_2}{p_1\log r_1+p_2\log r_2}\equiv d,\end{aligned}$$

so that $\mathrm{BE}(\prod_k)$ is the constant d for every k. It is remarkable that d also gives the entropy fractal dimension of the Cantor distribution. This is a consequence of a general result on the dimension of self-similar fractal constructions (Deliu et al., 1991). To illustrate this fact with a popular example, take $p_1=p_2=1/2$ and $r_1=r_2=1/3$. This implies $\mathrm{BE}(\prod_k)=d=\log 2/\log 3$ which is the well-known fractal dimension of the classical Cantor set and the natural Cantor distribution (see e.g. Falconer, 1990).

Invoking continuity of the index BE with respect to the probabilities p_i, the following general working principles are justified from the facts above:

#1. Small values of the index BE are consistent with P being *nearly discrete*.
#2. A lowering in the value of BE when the partition is refined is consistent with the measure spread being far from uniform (some size interval getting no mass in the splitting).
#3. Near to one values of the index BE are consistent with P *nearly uniform*.
#4. An increase in the value of BE when the partition is refined is consistent with the measure spread for the new partition being close to uniform (every class interval nearly getting the mass share proportional to its size),
#5. Computed BE values approaching one when partitions are refined is consistent with an underlying distribution with continuous density.
#6. Computed BE values approaching a certain positive value d below one when partitions are refined is consistent with an underlying fractal distribution.

References

Arya, L.M., Paris, J.F., 1981. A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. Soil Sci. Soc. Am. J. 45, 1023–1030.

Buchan, G.D., Grewal, K.S., Robson, A.B., 1993. Improved models of particle-size distribution: an illustration of model comparison techniques. Soil Sci. Soc. Am. J. 57, 901–908.

Caniego, J., Martín, M.A., San José, F., 2001. Singularity features of pore-size soil distribution: singularity strength analysis and entropy spectrum. Fractals 9, 305–316.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. Water Resour. Res. 14, 601–604.

Cover, T., Thomas, J.A., 1991. Elements of Information Theory. John Wiley.

Deliu, A., Geronimo, J.S., Shonkwiler, R., Hardin, D., 1991. Dimensions associated with recurrent self-similar sets. Math. Proc. Camb. Phil. Soc. 110, 327–336.

Falconer, K.J., 1990. Fractal Geometry — Mathematical Foundations and Applications. Wiley, New York.

Giménez, D., Rawls, W.J., Pachepsky, Y., Watt, J.P.C., 2001. Prediction of a pore distribution factor from soil textural and mechanical parameters. Soil Sci. 166 (2), 79–88.

Gray, R.M., 1990. Entropy and Information Theory. Springer-Verlag.

Hart, P.E., 1971. Entropy and other measures of concentration. J. R. Stat. Soc., Ser. A 134 (1), 73–85.

Haverkamp, R., Parlange, J.Y., 1986. Predicting the water retention curve from particle-size distribution: sandy soils without organic matter. Soil Sci. 142, 325–339.

Kravchenko, A., Zhang, R., 1998. Estimating the soil water retention from particle-size distributions: a fractal approach. Soil Sci. 163, 171–179.

MacArthur, R., 1955. Fluctuations of animal populations and measure of community stability. Ecology 36, 533–536.

Margalef, R., 1958. Information theory in ecology. Gen. Syst. 3, 36–71.

Martín, M.A., Rey, J.-M., Taguas, F.J., 2001. An entropy-based parameterization of soil texture via fractal modelling of particle-size distribution. Proc. R. Soc. Lond., Ser. A 457, 937–948.

Martín, M.A., Pachepsky, Ya. A., Rey, J.-M., Taguas, F.J., Rawls, W.J., 2005. The balanced entropy index to characterize soil texture for soil water retention estimation. Soil Sci. 170 (10), 759–766.

Martín, M.A., Rey, Balancing Shannon's entropy to indicate heterogeneity of mass-size distributions, submitted for publication.

Martín, M.A., Rey, J.-M., Taguas, F.J., 2005. An entropy-based heterogeneity index for mass-size distributions in Earth science. Ecol. Model. 182, 221–228.

Pachepsky, Ya. A., Rawls, W.J. (Eds.), 2004. Development of Pedotransfer Functions in Soil Hydrology. Elsevier, Amsterdam.

Shannon, C.E., 1948a. A mathematical theory of communication I. Bell Syst. Tech. J. 27, 379–423.

Shannon, C.E., 1948b. A mathematical theory of communication II. Bell Syst. Tech. J. 27, 623–656.

Soil Conservation Service, 1975. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. Agricultural Handbook, vol. 436. USDA-SCS, USA Department of Agriculture, pp. 486–742.

Theil, H., 1967. Economics and Information Theory. Amsterdam, North-Holland.

Tyler, S.W., Wheatcraft, S.W., 1989. Application of fractal mathematics to soil water retention estimation. Soil Sci. Soc. Am. J. 53, 987–996.

Wösten, J.H.M., Pachepsky, Ya. A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. (Amsterdam) 251, 123–150.