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 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 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, $H(\prod) \leq H(\prod')$ 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 $0 < r_1 \leq \frac{r}{2}$. 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 $-(p_1 + p_2) \log r \leq -p_1 \log r_1 - p_2 \log r_2$, we have that $E(\prod) \leq E(\prod')$ in this case. This can be extended to prove that, in general, $E(\prod) \leq E(\prod')$ 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

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

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

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

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

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

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.