Estimating soil bulk density with information metrics of soil texture

Miguel Ángel Martín^{*1}, Miguel Reyes² and F. Javier Taguas¹

* Corresponding author, Phone: +34 91 336 58 13, Fax: +34 91 336 5817 , E-mail address:

miguelangel.martin@upm.es

¹ Dpt. of Applied Mathematics, E.T.S.I. Agrónomos, Universidad Politécnica de Madrid, Avda.

Complutense s/n, 28040 Madrid, Spain.

² Dpt. of Applied Mathematics, E.T.S.I. Informáticos, Universidad Politécnica de Madrid,

Campus de Montegancedo s/n, 28660 Boadilla del Monte, Madrid, Spain.

Abstract

An information metrics of soil texture -the Shannon Information Entropy- is proposed to analyze the contribution of the particle size distribution (PSD) in soil bulk density values.

A database including 6239 soil samples from Florida is used. For each soil the Information Entropy is computed using mass proportions of the seven texture fractions that the database provides. The range interval of Information Entropy values is divided into an increasing number of subintervals of equal length, to study how differences in the soil texture metrics corresponded with differences in soil bulk density values. Grouping soil samples according to their information entropy, the average information entropy value in any of the resulting subintervals is plotted versus the corresponding average soil bulk density values. It was found that variations of less than 0.04 g/cm^3 in the mean bulk density values are explained with variations of mean information entropy values with an R^2 coefficient of determination equal to 0.99, being lower variations explained with no significant decrease in the fitting goodness. Predictions based in that linear regression give a mean predicted error (MPE) equal to 0.0015 g/cm^3 over the total number of soils, and a normal distribution of errors with standard deviation (SDPE) equal to 0.16 g/cm^3 . These results strongly support that Information Entropy serves as an indicator of the typical bulk density for a soil with a given PSD and average structure features.

Information Entropy (IE) was also computed for all samples (6239) using clay, silt and sand content. Simple linear regression, now implemented using values of each soil sample, was used to predict bulk density values using the corresponding Information Entropy value as input. Additionally, different published bulk density pedotransfer functions (PTFs), including organic carbon (OC) content and texture inputs, are applied to the same data bank. Results show similar mean square predicting error (RSMPE) and standard deviation of predicted error (SDPE) when Information Entropy is used as unique input respect to those obtained using PTFs. The results become worse for soils in horizon A and better in horizon E, respectively, possibly due to the influence of the different OC content in those horizons. Notably, the MPE is, by average, about

3 orders of magnitude lower when IE is used respect to the MPE obtained by the PTFs, reflecting the potential of Information Entropy of texture in capturing mean soil bulk density values.

Results show that Information Entropy metrics of soil texture provide a useful input for estimating bulk density, which also might be used together with other inputs as depth or OC content.

Keywords: bulk density, soil texture heterogeneity, soil particle size distribution, information entropy.

1. Introduction

Bulk density (ρ_b) is an important soil physical property needed for estimating soil water characteristic and used as input parameter for water and nutrient transport models (Boucneau et al., 1998). Accurate evaluation of bulk density is needed to get precise estimates of soil organic carbon (OC) stocks by means of weight-to-volume conversions (Howard et al., 1995).

Different factors as depth, organic matter content or compaction have influence in the bulk density values. Overall, differences in bulk density value among soils are primarily attributed to differences in particle size distribution (PSD) (Manrique and Jones, 1991). However the way that PSD data are included in the pedotransfer functions (PTFs) used to estimate bulk density values have enormous variation depending of authors (Rawls, 1983; Manrique and Jones, 1991; Tomasella and Hodnett, 1998; Bernoux et al., 1998; Kaur et al., 2002; De Vos et al., 2005; Benites et al., 2007; Nanko et al., 2014). There exist great differences in the weight given to any textural component in formulas generally obtained by multiple regressions on particular data sets. In spite of large number of existing PTFs, the search for additional properties such as inputs is an important direction for improving PTF accuracy and

reliability (Wösten et al. 2001). As it has been recently noticed, developing physics-based PTFs should be explored (Pachepsky and Romano, 2015).

A quantitative analysis of the contribution of the PSD to bulk density values seems to demand key metrics of PSD which enables to efficiently give accounts of the variation in bulk density values corresponding to variation in the PSD metrics. Similarly as was done in (Tranter et al., 2007), bulk density might be conceptually represented as the sum

$$\rho_b = \rho_{PSD} + \Delta\rho \quad (1)$$

where the component ρ_{PSD} is seen as the typical bulk density for a soil with a given PSD and average structure features and $\Delta\rho$ the variation in the bulk density attributed to other structural factors such as organic matter and depth. The objectives of this study are: (i) to propose unified information metrics of the soil texture on a rational basis, (ii) to study how differences in the soil texture metrics corresponded with differences in the typical soil bulk density value of related soils and to analyze the basic structure of the correspondence, (iii) to evaluate the accuracy of predictions of soil bulk density using information metrics as a single input, and (iv) to compare predictions made by using information metrics as single input with those made using traditional PTFs, which also include OC content together with variable textural inputs.

The soil texture information metrics proposed here -the Shannon Information Entropy-, is related with PSD heterogeneity. Indeed heterogeneity should influence bulk density values. Briefly, in a sandy soil solely composed by sand particles only pores greater than certain size might be expected to be present. On the contrary, in the case of a more heterogeneous texture with additional clay fraction, some of the parts of such pore space should be occupied by smaller particles, with no significant increase in volume. Similar arguments might be held true across the whole range of size particles: the richer the heterogeneity is, the higher chance of denser packing would be expected. A related entropy-like parameter, the Balance Entropy Index (BEI), agreeing with the Shannon Information Entropy except for a normalizing multiplier, has already been used as predictor of soil water retention estimation (Martín et al., 2005), which seemed to suggest that the index had a potential to reflect probable packing of soil particles in soils with heterogeneous PSD's. This is consistent with the fact that the BEI provides additional

information at smaller scales when a scaling in the PSD takes place (Martín and Taguas, 1998; Martín et al., 2001). This might be the rationale behind the ability of entropy indexes based on soil coarse texture data, to reflect further soil functions.

2. PSD heterogeneity metrics

Heterogeneity is an ambiguous term that needs to be quantified to be introduced as input parameter for bulk density prediction. This issue is addressed with concepts from the Information Theory introduced in the pioneering work of Shannon C.E. (1948).

Mathematically speaking, the PSD of granular media may be considered as a continuous mass particle-size distribution μ supported in the interval of grain sizes. Limited information on PSD is usually determined over a list of size ranges that covers all the sizes present in the sample. Grains sorted according to the size thus appear distributed in size classes $J_1, J_2, \dots, J_k$ defined by those size ranges. Different methods of analysis may provide the mass proportions $p_1, p_2, \dots, p_k$, where $\mu(J_i) = p_i$, $1 \leq i \leq k$, and $\sum_{i=1}^k p_i = 1$. The Shannon Information Entropy (IE) (Shannon, 1948) of the partition is defined by

$$H = - \sum_{i=1}^k p_i \log p_i \quad (2)$$

where $\log$ represented to the binary logarithm, and provided $p_i \log p_i = 0$ if $p_i = 0$.

The number $H = H(p_1, p_2, \dots, p_k)$ is expressed in information units (bits) and its extreme values are $\log k$, which corresponds to the most even (homogeneous) case — where all the intervals have the same cumulative mass — and 0, which corresponds to the most uneven (heterogeneous) case — where the whole mass is concentrated in a single interval. The Shannon Information Entropy $H = H(p_1, p_2, \dots, p_k)$ is a widely accepted measure of heterogeneity in the mass distribution μ at the scale level given by each partition $P = (J_1, J_1, \dots, J_1)$. It can be shown that any measure of heterogeneity with the natural properties for such goal must be a multiple of $H(p_1, p_2, \dots, p_k)$ (Khinchin 1957). In fact, entropy has long been used in the life sciences as a plausible metric of biodiversity (Margalef, 1958) in the sense of evenness or heterogeneity of

the diversity of species in an ecosystem. The same use of entropy to measure pedodiversity has been discussed in (Ibañez et al., 1998).

3. Materials and methods

3.1. Data set

The data set used to evaluate the influence of soil texture heterogeneity in soil bulk density values is USKSAT. This data base is comprised from journal publications and technical reports collected coupled data on laboratory Ksat, USDA textural class, and bulk density obtained in the United States. Detailed information can be found in (Pachepsky and Park, 2015). Soil samples not reporting both, texture and bulk density, and soil samples with incoherent textural data (the sum of mass texture fractions not agreeing with the total mass) were rejected. Soil samples with the same textural data but bulk density values differing less than 0.1 g/cm^3 were grouped and considered as a single soil. Finally, soil samples having same textural data and corresponding to the same depth, were grouped assigning them the average value of their bulk density values. After these selection criteria a number of 6239 soils were considered for the study. A number of 3956 are sand, 570 are loamy sand, 698 are sandy loam, 27 are loam, 27 are silty loam, 4 are silt, 666 are sandy clay loam, 26 are clay loam, 3 are silt clay loam, 118 are sandy clay and 144 are clay. For the first part of the study the total data base has been considered. In the comparative analysis with different PTFs, four different groups of soils have been considered: the total data base (6239 samples), soils from horizon A (1088 samples), soils from horizon E (1113 samples) and sandy soils (1264 samples). Some descriptive statistics of the basic soil properties are listed in table 1.

3.2. Methods

Firstly, Information Entropy of soil texture is computed, following Eq. (2), for all samples (6239) using the seven available fraction contents. In order to analyze how differences in the Information Entropy explain differences in the typical soil bulk density value of related soils,

the range interval of information entropy values was first divided into 10 subintervals of equal length. The mean bulk density value of soil samples with information entropy within any subinterval was computed and the mean bulk density value was plotted against the mean information entropy value. Linear regression analysis is made. The same analysis is repeated dividing the range interval of information entropy values into 15 subintervals of equal length with the objective of studying patterns of variation for mean bulk density increments between 0.02 and 0.04 g/cm³.

Also the linear regression is used to predict bulk density values of the total number of soils using the corresponding Information Entropy value as input. Here and later, the prediction quality was determined by complementary indices: Eq. (3), the mean predicted error (MPE); Eq. (4), the standard deviation of the prediction error (SDPE); and Eq. (5), the root mean square of the predicted error (RMSPE). These are defined as:

$$\text{MPE} = \frac{1}{N} \sum_{i=1}^N (\widehat{\rho}_{b,i} - \rho_{b,i}) \quad (3)$$

$$\text{SDPE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\widehat{\rho}_{b,i} - \rho_{b,i})^2 - \text{MPE}^2} \quad (4)$$

$$\text{RMSPE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\widehat{\rho}_{b,i} - \rho_{b,i})^2} \quad (5)$$

where $\rho_{b,i}$ and $\widehat{\rho}_{b,i}$ are the observed and predicted values of the bulk density, respectively. On the base of these values, the analysis of the distribution of errors is made in the above prediction by mean of a test of normality (Feller, 1968).

Secondly, Information Entropy of soil texture was also computed for all samples (6239) using clay, silt and sand contents. In both cases (i.e., using seven or three fraction contents), simple linear regression analysis of bulk density values versus Information Entropy values was implemented on all individual soils. The respective fitting lines were used to compute predicted bulk density values using the corresponding Information Entropy value as input. Results are discussed.

Additionally, different published PTFs to estimate bulk density are applied to the same data bank. These PTFs include different textural information and OC content as inputs (Table 2).

Bulk density predicted values were also compared with real ones. Mean predicted errors and standard deviation obtained were compared with those produced by using Information Entropy as single input.

4. Results and discussion

When Information Entropy and bulk density mean values are computed within 10 or 15 subintervals, as has been described, the mean bulk density values ranged between 1.40 and 1.65 (results are summarized in tables 2a and 2b).

The plot of bulk density values against Information Entropy values showed R^2 coefficient equal to 0.99 when the number of subintervals is 10 and diminish to 0.98 when the number of subintervals is 15 (figures 1a and 1b). This means that variations between 0.02 and 0.04 g/cm³ in the mean bulk density values are explained with variations of mean Information Entropy values with an R^2 fitting regression close to 1. It can be observed that together with close R^2 values, the fitting lines practically mirror each other. The use of one or the other in predictions did not produce significant different results. Results show a mean predicted error (MPE) equal to 0.0015 and a standard deviation (SDPE) equal to 0.16. The number of soils for which the predicted errors fall within the intervals (MPE-SDPE, MPE+SDPE) and (MPE-2·SDPE, MPE+2·SDPE) (see Table 4) confirmed that the distribution of errors fit a normal distribution (Feller, 1968) with a mean value close to 0. Also the variance analysis of bulk density values within the subintervals confirmed a normal distribution of ρ_b values within each subinterval.

The structure of the correspondence between bulk density and Information Entropy mean values obeying a strong linear pattern, together with the normal distributions of errors with a mean value practically equal to 0, seem to confirm that Information Entropy of soil

texture serves as an indicator of the typical bulk density for a soil with a given PSD and average structure features.

Simple linear regression, now implemented using values of each soil sample, has been made using the Information Entropy computed with clay, silt and sand contents (IE3) and also using the Information Entropy computed with the seven available fraction contents (IE7). Also predictions by using PTF1 (Tomasella and Hodnett, 1998) and PTF2 (Kaur et al., 2002) to estimate bulk density have been made for the total number of soil samples. The results are summarized in Table 5.

It should be noted that predictions made by using the regression line fitting ($R^2 = 0.99$) obtained in the regression between mean values, as described in the beginning of this section, produced practically the same results. Results show very low predicted error (less than 10^{-4}) and standard deviation values close to 0.16 g/cm^3 . It can be observed that in spite of having rather different R^2 values of the corresponding linear fittings, depending on the use of three classical textural fractions or the seven textural fractions reported in the data base, the prediction quality is very similar. It can be observed that, in general, there are no great differences in the respective RMSPE and SDPE values while the MPE produced by using information metrics is much lower than the produced by PTF1 and PTF2.

Table 6a, table 6b and table 6c summarize results for PTF3, PTF4 and PTF5 for soil in horizon E, horizon A and sandy soils, respectively, compared in all cases with those obtained by using information metrics.

The results of prediction for soils in horizon A made by using the soil texture information metrics show a much lower MPE, although greater SDPE and RSMPE than those obtained by PTF3, suggesting that OC content present in soils of this horizon play a critical role for the better performance of PTF3, which includes OC content as input. For soils of horizon E, however, the results obtained by the soil texture information metrics are strictly better than those obtained by PTF4, specially proposed for that horizon. Finally the results for sandy soils are quite similar with those obtained by PTF5, also calibrated for these types of soils. Again, in these cases a remarkable result is that the MPE is notably lower in respect to the obtained by

above said PTFs, suggesting that information metrics have in all cases an excellent ability to catch the variation in bulk density values due to the PSD, and that it might be useful as single texture input in new PTFs for estimating bulk density.

5. Conclusions

There exist great differences in the way that soil texture data are included in the different pedotransfer functions used to estimate bulk density values. The soil texture Information Entropy metrics serves to make a quantitative analysis of the contribution of the PSD to bulk density values.

Linear regression analysis reveals a linear dependence pattern between mean information entropy and mean bulk density values. The analysis of predictions made on the basis of the above linear regression, strongly supports that Information Entropy has a great potential to serve as an indicator of the typical bulk density for a soil with a given PSD and average structure features.

Results obtained in this study show in general, similar predictive accuracy when Information Entropy is used as the only input respect to the PTFs which include OC content and different texture functions as input, that become worse for soils in horizon A and better in horizon E, respectively, possibly due to the OC content in those horizons.

Information Entropy metrics of soil texture provide a useful input for estimating bulk density, that anyway might be used together with other inputs such as as depth or OC content.

Acknowledgments

Authors thank to Yakov Pachepsky (USDA–ARS, Beltsville Agricultural Research Center, Environmental Microbial and Food, Safety Lab) for making available the USKSAT database and his valuable comments along the elaboration of this work.

This research work was funded by Spain's Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica (I+D+I), under ref. AGL2011-25175 and AGL2015-69697 P.

References

- Benites, V.M., Machado, P.L.O.A., Fidalgo, E.C.C., Coelho, M.R., Madari, B.E., 2007. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. *Geoderma* 139, 90–97.
- Bernoux, M., Cerri, C.C., Neill, C., De Moraes, J.F.L., 1998. The use of stable carbon isotopes for estimating soil organic matter turnover rates. *Geoderma* 82, 43-58.
- Boucneau, G., Meirvenne, M. Van, Hofman, G., 1998. Comparing pedotransfer functions to estimate soil bulk density in northern Belgium, *Pedologic Themata* 5, 67-70.
- De Vos, B., Meirvenne, M. Van, Quataert, P., Deckers, J., Muys, B., 2005. Predictive Quality of Pedotransfer Functions for Estimating Bulk Density of Forest Soils. *Soil Sci. Soc. Am. J.* 69, 500–510.
- Feller, W., 1968. An introduction to Probability Theory and its applications. John Wiley, New York.
- Howard, P.J.A., Loveland, P.J., Bradley, R.I., Dry, F.T., Howard, D.M., Howard, D.C., 1995. The carbon content of soil and its geographical distribution in Great Britain. *Soil Use Manage.* 11, 9-15.
- Ibáñez, J. J., de Alba, S., Lobo, A., Zucarello, V., 1998. Pedodiversity and global soil patterns at coarse scales (with Discussion). *Geoderma* 83, 71–192.
- Kaur, R., Kumar, S., Gurung, H., 2002. A pedo-transfer function (PTF) for estimating soil bulk density from basic soil data and its comparison with existing PTFs. *Aust. J. Soil Res.* 40, 847–857.

- Khinchin, A.I., 1957. *Mathematical Foundation of Information Theory*. Dover Publications, New York.
- Leonaviciute, N., 2000. Predicting soil bulk and particle densities by pedotransfer functions from existing soil data in Lithuania. *Geografijos Metrastis* 33, 317-330.
- Manrique, L.A., Jones, C.A., 1991. Bulk density of soils in relation to soil physical and chemical properties. *Soil Sci. Soc. Am. J.* 55, 476–481.
- Margalef, R., 1958. Information theory in ecology. *Gen. Syst.* 3, 36–71.
- Martín, M.A., Taguas, F.J., 1998. Fractal modelling, characterization and simulation of particle-size distributions in soil, *Proc. R. Soc. Lond. A* 454, 1457-1468.
- Martín, M.A., Rey, J.M., Taguas, F.J., 2001. An entropy-based parametrization of soil texture via fractal modelling of particle-size distribution. *Proc. R. Soc. Lond. A* 457, 937-947.
- Martín, M.A., Pachepsky, Y., Rey, J.M., Taguas, F.J., 2005. Balanced entropy index for soil water retention estimation. *Soil Sci.* 170, 759-766.
- Nanko, K., Ugawa, S., Hashimoto, S., Imai, A., Kobayashi, M., Sakai, H., Ishizuka, S., Miura, S., Tanaka, N., Takahashi, M., Kaneko, S., 2014. A pedotransfer function for estimating bulk density of forest soil in Japan affected by volcanic ash. *Geoderma*. 213, 36–45.
- Pachepsky, Y., Park, Y., 2015. Saturated Hydraulic Conductivity of US Soils Grouped According to Textural Class and Bulk Density. *Soil Sci. Soc. Am. J.* doi:10.2136/sssaj2015.02.006.
- Pachepsky, Y., Romano, N., 2015. Informing soil models using pedotransfer functions: challenges and perspectives. *Geophysical Research Abstracts* Vol. 17, EGU2015-6303, 2015 EGU General Assembly 2015.
- Rawls, W.J., 1983. Estimating soil bulk density from particle size analysis and organic matter content. *Soil Science* 135 (2), 123-125.
- Shannon, C.E., 1948. A mathematical theory of communication. *I. Bell Syst. Tech. J.* 27, 379-

423.

Tomasella, J., Hodnett, M.G., 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Sci.* 163, 190-202.

Tranter, G., Minasny, B., Mcbratney, A.B., Murphy, B., Mckenzie, N.J., Grundy, M., Brough, D., 2007. Building and testing conceptual and empirical models for predicting soil bulk density. *Soil Use and Management* 23, 437-443.

Wösten, J.H.M., Pachepsky, Y., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* 251, 123-150.

1 **Table 1**
2 Summary statistics of basic soil properties of USKSAT data base.

Dataset	Statistic	OC	Sand (%)	Silt (%)	Clay (%)
All soils (6239)	Min	0.01	0.20	0	0
	Max	15.61	99.90	94.50	93.40
	Mean	0.59	85.53	5.65	8.83
Horizon A (1088 soils)	Min	0.08	2.40	0	0
	Max	11.55	99.90	53.10	67.70
	Mean	1.60	89.92	6.10	3.98
Horizon E (1113 soils)	Min	0.01	36.40	0.10	0.10
	Max	3.61	99.70	42.60	21.00
	Mean	0.22	95.54	2.97	1.48

3

4

1
2
3
4
5

6
7

Table 2
Published pedotranfer functions used in this study (ρ_b , bulk density; ln, natural logarithm; OC, organic carbon %; clay, 0- to 2- μ m fraction %; silt, 2- to 50- μ m fraction %; sand, 50- to 2000- μ m fraction %).

PTF1 (Tomasella and Hodnett, 1988)	$\rho_b = 1.578 - 0.054 \cdot OC - 0.006 \cdot silt - 0.004 \cdot clay$
PTF2 (Kaur et al., 2002)	$\ln \rho_b = 0.313 - 0.191 \cdot OC - 0.000476 \cdot (clay)^2$ $- 0.00432 \cdot silt + 0.02102 \cdot clay$
PTF3 (Horizon E) (Leonaviciute, 2000)	$\rho_b = 0.99915 - 0.08415 \cdot \ln OC - 0.00592 \cdot \ln(silt)$ $+ 0.07712 \cdot \ln(clay) + 0.09371 \cdot \ln(sand)$
PTF4 (Horizon A) (Leonaviciute, 2000)	$\rho_b = 1.70398 - 0.00313 \cdot silt + 0.00261 \cdot clay$ $- 0.11245 \cdot OC$
PTF5 (Sandy soils) (Bernoux et al., 1998)	$\rho_b = 0.0181 \cdot sand - 0.08 \cdot OC$

Table 3
Information Entropy (IE) range, mean IE values and mean bulk density (ρ_b) values in: (a) 10 subintervals, and (b) 15 subintervals.

(a)

IE	Number of soils	Mean IE	Mean ρ_b
0.219-0.471	49	0.384	1.402
0.472-0.723	171	0.625	1.418
0.724-0.975	293	0.854	1.435
0.976-1.227	503	1.109	1.467
1.228-1.479	941	1.360	1.508
1.480-1.730	1203	1.604	1.513
1.731-1.982	1369	1.852	1.544
1.983-2.234	1032	2.100	1.581
2.235-2.486	548	2.344	1.598
2.487-2.739	130	2.557	1.639

(b)

IE	Number of soils	Mean IE	Mean ρ_b
0.219-0.387	22	0.326	1.409
0.388-0.555	59	0.482	1.414
0.556-0.723	139	0.648	1.416
0.724-0.891	188	0.810	1.435
0.892-1.059	250	0.981	1.437
1.060-1.227	358	1.146	1.478
1.228-1.395	599	1.315	1.498
1.396-1.563	759	1.484	1.520
1.564-1.731	792	1.649	1.511
1.732-1.899	969	1.816	1.536
1.900-2.067	777	1.982	1.575
2.068-2.235	653	2.147	1.577
2.236-2.403	396	2.309	1.597
2.404-2.571	234	2.471	1.617
2.572-2.739	44	2.623	1.646

1
2
3
4
5

6
7
8
9

Table 4
Results of predictions obtained by regressions on all individual soils, using Information Entropy computed with 3 textural data (IE3), 7 textural data (IE7), and with PTF1 (Tomasella and Hodnett, 1998) and PTF2 (Kaur et al., 2002)

	IE3	IE7	PTF1	PTF2
Number of soils	6239			
Mean predicted error (MPE)	-5.544E-05	2.609E-05	0.052	0.209
Standard deviation (SDPE)	0.167	0.161	0.158	0.171
Mean square predicted error (RMSPE)	0.167	0.161	0.167	0.270

Table 5

Results of predictions obtained by regressions on all individual soils, using Information Entropy computed with 3 textural data (IE3), 7 textural data (IE7), and with: (a) PTF3 (Leonaviciute, 2000), (b) PTF4(Leonaviciute, 2000), and (c) PTF5 (Bernoux et al.,1998).

(a)

	IE3	IE7	PTF3
Number of soils (Horizon E)	1113		
Mean predicted error (MPE)	3.412E-08	-1.037E-04	-0.031
Standard deviation (SDPE)	0.099	0.090	0.117
Mean square predicted error (RMSPE)	0.099	0.090	0.121

(b)

	IE3	IE7	PTF4
Number of soils (Horizon A)	1088		
Mean predicted error (MPE)	-1.626E-05	4.758E-06	-0.145
Standard deviation (SDPE)	0.180	0.182	0.151
Mean square predicted error (RMSPE)	0.180	0.182	0.209

(c)

	IE3	IE7	PTF5
Number of soils (Sandy soils)	1264		
Mean predicted error (MPE)	1.371E-05	3.185E-05	-0.143
Standard deviation (SDPE)	0.130	0.128	0.138
Mean square predicted error (RMSPE)	0.130	0.128	0.199

1

2

3

4

(a)

5

6

(b)

7

Figure 1

Mean bulk density value (mean ρ_b) plotted vs. mean information entropy value (mean IE) using: (a) 10 subintervals, and (b) 15 subintervals.

10

11