Carpathian Journal of Earth and Environmental Sciences, May 2015, Vol. 10, No. 2, p. 225 - 236

VALIDATING SOIL BULK DENSITY PEDOTRANSFER FUNCTIONS USING A ROMANIAN DATASET

Ionuţ VASILINIUC[1] & Cristian Valeriu PATRICHE[2]

[1]*"Alexandru Ioan Cuza" University of Iaşi, Department of Geography, Blvd. Carol I nr. 20A, 700505, Iaşi, Romania, vasiliniucionut@yahoo.com*

[2]*Romanian Academy, Iaşi Branch, Blvd. Carol I nr. 8, 700505, Iaşi, Romania, pvcristi@yahoo.com*

Abstract: Bulk density is one of the most important soil parameters, being used in numerous hydrological, agricultural or environmental models. Nevertheless, this property is missing from most soil surveys in Romania and not only, its determinations being quite time-consuming. Because of this recent tendencies have been of deriving this property with the help of pedotransfer functions. Twenty-two pedotransfer functions have been tested using a database consisting of 430 soil profiles from Romania, with 2017 samples. The results show that many of the PTFs presented by different authors cannot be applied for the Romanian territory, being derived from small datasets, for small areas or for regions differing in their geographical conditions. The functions with better results have obtained R^2s of 0.403-0.436, MPEs of 0.003-0.04, RMSPEs of 0.157-0.171 and SDPEs of 0.156-0.159. In this situation, it is recommended that when estimating bulk density with the help of PTFs, one should take into account the mathematical expression and the characteristics of the dataset used in deriving the function.

Key words: Romania, pedotransfer functions, soil bulk density, validation, independent data set

1. INTRODUCTION

One of the frequent problems related to soil surveys from Romania and not only is the lack of measured values for some properties. Being difficult and time-consuming to determine, they do not enter the usual measurements conducted by the Romanian system of agricultural soil mapping, and hence the need for their estimation.

A recent tendency in soil science is the estimation of these parameters with the help of pedotransfer functions (PTFs), which have gained recognition during the last years as approaches to translate simple soil characteristics into more complex parameters. They can be defined as predictive functions of soil parameters from other more easily determined properties (Bouma, 1989; McBratney et al., 2002).

One of the most important parameters that are missing from soil surveys is bulk density. It is defined as the ratio between the undisturbed dry soil mass and the total soil volume. Knowing soil bulk density values is important in characterizing the soil state, total and air porosity and expressing different properties in volume percentage. Bulk density is a key parameter in many hydrological or environmental models, and is needed in estimating carbon or nutrients stocks or water retention characteristics (De Vos et al., 2005; Suuster et al., 2011).

Bulk density also determines several soil physical properties. High values imply a decrease in water retention capacity, in permeability, air capacity, as well as an increase in the mechanical resistance opposed by soil to tillage and plant rooting. The relation between bulk density and qualitative characteristics of the soil cannot be correctly interpreted but in relation to other soil properties such as grain-size distribution or organic matter content (Canarache, 1990).

Results of different studies have shown that bulk density depends most often on the organic matter (OM) or organic carbon (OC) contents and on the proportion of different grain-size fractions (Curtis & Post, 1964; Adams, 1973; Alexander, 1980; Harrison & Bocock, 1981; Huntington et al., 1989; Manrique & Jones, 1991; Federer et al., 1993; Bernoux et al., 1998; Tomasella & Hodnett, 1998

Prevost, 2004; Benites et al., 2007; Perie & Ouimet, 2008; Brahim et al., 2012; Han et al., 2012; Hollis et al., 2012), that it varies with depth (Huntington et al., 1989; Leonaviciute, 2000), soil type (Alexander, 1980; Manrique & Jones, 1991; Tranter et al., 2007; Suuster et al., 2011; Nanko et al., 2014), land use or vegetation (Harrison & Bocock, 1981; Suuster et al., 2011).

Although PTFs sometimes also use data on chemical properties (Benites et al., 2007; Brahim et al., 2012; Han et al., 2012), water content (Heuscher et al., 2005; Patil & Chaturvedi, 2012), management (Hollis et al., 2012), traffic, crop systems or even morphometric parameters (Wang et al., 2014), one of the most important factor remains the OM content (Ruehlmann & Korschens, 2009), which frequently explains over 30-50% of the bulk density variance.

Usually, soil bulk density values decrease with the clay content and vary inversely in relation to the OM content. Small values are determined for organo-mineral and especially organic soils. In the case of Romanian mineral soils which present an upper horizon with relative high humus content, bulk density varies between 0.8-1.2 g/cm^3 for the topsoil and 1.4-1.6 g/cm^3 for the subsoil.

Due to the absence of this parameter from most of the soil surveys, the alternative might be its estimation based on correlated properties. There are diverse estimation methods (PTFs) elaborated by different authors. The methods approached in deriving PTFs vary from basic statistical models to the most commonly used stepwise multiple regressions (Heuscher et al., 2005; Benites et al., 2007; Bernoux et al., 1998; Brahim et al., 2012; Kaur et al., 2002) or to more advanced techniques such as artificial neural networks (ANN) and regression trees. Some authors hold that the advanced methods produce better results (Martin et al., 2009; Patil & Chaturvedi, 2012; Al-Qinna & Jaber, 2013) while other studies (Tranter et al., 2007) have demonstrated that these techniques do not necessarily improve the performance of the models.

Many of these functions have been elaborated for small areas, from small datasets or for certain regions, soils or environments, which makes their application in other conditions uncertain.

In the case of Romanian soils, some attempts in this direction have been those of Chiriţă (1970) and Canarache (1995) for andic and spodic soils and using only the OM content.

In any case, Romania has a very diverse soil cover, so no function elaborated has analyzed or included all or at least most of the soil types. Taking into account these aspects, there even might be the possibility that a single PTF would not be appropriate for all soil classes.

The objectives of this paper have thus been of testing the performance of a number of published PTFs with the help of a Romanian database, and to evaluate the respective functions using validation indices in terms of accuracy, precision and operability.

2. MATERIALS AND METHODS

2.1. Study area and dataset

In order to validate the published PTFs, a database was constructed that has included diverse soil types from Romania, analyzed at national soil conferences, but also other legacy data with detailed analyses. Including measured bulk density data and many other parameters, the database can be used to compute the PTFs and to validate or invalidate the published functions for the Romanian territory.

Romania is characterized by a varied geology, with diverse rock types lying over platform or orogen regions and also including volcanic areas. The altitude of landforms covering these deposits ranges from zero to 2544 m altitude, including plains, plateaus, hills and mountains. The climate is temperate continental, with variations imposed by latitude and longitude, but also by the presence of the Carpathian belt (Apostol & Sfîcă, 2013) or of large river valleys. Mean annual temperatures vary from <0ºC, in high mountain areas, to >11ºC, in the southern part of the country, annual rainfall quantities range from <350 mm to >1200 mm, with differences determined mainly by altitude. The vegetation areas vary accordingly from steppe in the south-eastern part to silvo-steppe, forests (mainly beech, oak and spruce) and alpine pastures at the highest altitudes (Popovici et al., 2013). In relation to this environmental variety, in Romania are found, according to the national soil classification system (Florea & Munteanu, 2012), 12 classes with 29 soil types, which develop in altitude from Cernisols and Luvisols to Cambisols, Spodisols and Umbrisols. The soil cover is completed by diverse types, from Pelisols, Andisols, Hidrisols, Histisols to Anthrisols and Salsodisols.

The database has included 430 soil profiles with 2017 horizon samples, from all over the country. Most of the samples have been taken from altitudes lower than 500 m, about 200 samples from 500-1000 m and 100 samples from altitudes over 1000 m. From the total, 658 samples are from Cernisols, 578 from Luvisols, 151 from Cambisols, 209 from Protisols, 80 from saline soils, 63 from vertic soils, the rest being represented by other mountainous soils (Spodisols, Umbrisols). Very few samples have been included from Andosols (6) and Histisols (5). About 1500 samples are from topsoils (0-50 cm) and the rest from lower horizons.

A first step in the analysis has been bringing to a common denominator the data, because Romania uses a different grain-size distribution system than those of the databases used for deriving the published PTFs. In order to do this, the Atterberg textural data from our database (with the upper limit of silt fraction at 0.02 mm) have been converted in the USDA system (with the limit at 0.05 mm) with the use of the log-linear transformation suggested by Wösten et al., (1999).

2.2. Sampling and analytical methods

The soil sampling and analysis has been conducted according to the standard methods used in Romania (Florea et al., 1987). Disturbed and undisturbed samples have been taken from the genetic horizons. The disturbed sampling has been conducted for the physical and chemical characterization, while the undisturbed samples for physical determinations have been taken in metallic cores of known volume (100 cm^3), at the actual soil humidity, with four repetitions for each horizon.

The preparations for laboratory analyses have included the separation of organic materials and skeleton, followed by grounding and sieving. Grain-size analysis involves a three variant pre-treatment according to the sample composition:

- for samples including carbonates a treatment with 2 n hydrochloric acid and dispersion with 1 n sodium hydroxide, Kacinski method;

- for samples without carbonates and OM > 5%, the oxidation of OM with hydrogen peroxide 6% and dispersion with potassium hexametaphosphate solution 10% or 1 n sodium hydroxide, Kacinski method;

- for samples without carbonates and OM < 5%, dispersion with potassium hexametaphosphate solution 10%.

The determination of grain-size fractions is conducted by pipetting for the fractions < 0.002 mm, by wet sieving for the 0.002-0.2 mm fractions and dry sieving for the fractions > 0.2 mm. The results are expressed as percentages reported to the material left after the pretreatment.

Bulk density (Bd) is determined by the metallic cylinders with known volume (100 cm^3) method at the momentary soil humidity.

Total organic carbon is determined by humid oxidation according to the method of Walkley-Black, modified by Gogoaşă (1959). Organic matter content is computed using the van Bemmelen factor (OM = OC x 1.724).

2.3. Published PTFs

Numerous PTFs have been derived for estimating or indirectly determining soil bulk density. The problem with these functions is that they are frequently created departing from a dataset specific for a certain area or soil type, so their application in other regions first needs a validation.

From the many published functions, for validation using the Romanian dataset have been chosen those of Curtis & Post (1964), Adams (1973), Alexander (1980), Harrison & Bocock (1981), Federer et al., (1983), Huntington et al., (1989), Manrique & Jones (1991), Tomasella & Hodnett (1998), Post & Kwon (2000), Kaur et al., (2002), Tremblay et al., (2002), Prevost (2004), Perie & Ouimet (2008), Han et al., (2012), Hollis et al., (2012) (Table 1).

These functions have been classified in three categories, according to the soil parameters that have been used by authors in estimating bulk density (organic carbon, organic matter, organic carbon and grain-size data).

One of the first attempts at validating PTFs has been that of Boucneau et al. (1998), who used 40 soils (182 samples) from northern Belgium. Their results indicated that locally developed functions had very small mean prediction errors but weak correlation coefficients, while the PTFs of Manrique & Jones (1991) had inverse performances, with determination coefficients of 0.5-0.6 and mean errors of 0.11-0.17 g/cm^3.

Kaur et al., (2002) have evaluated PTFs with the help of 224 samples with different land use from India, discovering a very weak predictive potential due to the development of the functions for specific soils or ecosystems. In order to obtain a visible accuracy and a better precision in estimating bulk density, they recommend using a function for each soil class. Their results have shown that the PTFs of Curtis & Post (1964), Adams (1973), Federer et al., (1993) and Huntington et al., (1989) underestimated bulk density values, while Alexander (1980) and Manrique & Jones (1991) overestimated. Among the methods derived using OM, the best to perform seemed to be those of Alexander (1980) and Manrique & Jones (1991). The functions of Curtis & Post (1964), Adams (1973), Federer et al., (1993), Huntington et al., (1989) and Tomasella & Hodnett (1998) have been developed for forest soils with high contents of OM, and also with specific grain-size distribution. The validation set included data with smaller OM contents and different textures, which might explain their weak performance. It seems that the models of Manrique & Jones (1991) and Alexander (1973), based on larger datasets with diverse soils, had a better performance.

De Vos et al., (2005) have evaluated 12 PTFs with the help of a 1614 samples dataset. All the functions have produced an underestimation of bulk density, with errors of 0.01-0.51 g/cm^3. Their evaluation has demonstrated the weak performance of some PTFs and raised problems related to their prediction ability. Validating several functions, Perie & Ouimet (2008) obtained correlation coefficients of 0.42-0.82, the polynomial models predicting better in the case of soils with OM <1.

Han et al., (2012) evaluated 19 PTFs, showing that the models developed by Alexander (1980), Manrique & Jones (1991) and Perie & Ouimet (2008) give relatively good predictions, although the first two models are not so good for soils with high OM contents. Nanko et al., (2014) have also compared a series of functions, most of them explaining 63-68% of the variance of bulk density. Most of the functions have overestimated, but this is due to the use only of soils developed on volcanic materials. The nature of the mathematical relations that define PTFs is reflected in the estimation of bulk density. Non-linear (logarithmic, exponential) functions are apparently more realistic, indicating an attenuated decrease of bulk density values with the increase in the OM content.

Table 1. Published PTFs, sampling area and land use*

Pedotransfer functions (PTFs)		Code
Functions that use only OC		
Manrique & Jones, 1991	Bd = 1.510 - 0.113 x OC	**A1**
Manrique & Jones, 1991	Bd = 1.660 - 0.318 x OC$^{0.5}$	**A2**
Alexander, 1980	Bd = 1.66 - 0.308 x OC$^{0.5}$	**A3**
Alexander, 1980	Bd = 1.72 - 0.294 x OC$^{0.5}$	**A4**
Huntington et al., 1989	lnBd = 0.263 - 0.147 x ln OC - 0.103 (ln OC)^2	**A5**
Harrison & Bocock, 1981	Bd = 1.558-0.728 x log (OC)	**A6**
Functions that use only OM		
Curtis & Post, 1964	Log(Bd x 100) = 2.09963 - 0.00064 x (log OM) - 0.22302 x (log OM)^2	**B1**
Federer et al., 1983	lnBd = -2.31 - 1.079 x ln (OM/100) - 0.113 x (ln OM/100)^2	**B2**
Prevost, 2004	lnBd =-1.81-0.892 x ln(OM/100)-0.092 x ln(OM/100)^2	**B3**
Perie & Ouimet, 2008	Bd=-1.977 + 4.105 x (OM/100) - 1.229 x ln(OM/100) - 0.103 x ln (OM/100)^2	**B4**
Han et al., 2012	lnBd = 0.5379 - 0.0653 x (OM x 10)$^{0.5}$	**B5**
Adams, 1973	Bd = 100 / [(OM / 0.224 + (100 - OM) / 1.27)]	**B6**
Post & Kwon, 2000	Bd = 0.244 x 1.640 / [1.640 x OM + 0.244(1 - OM)]	**B7**
Tremblay et al., 2002	Bd = 0.120 x 1.400 / [1.400 x OM + 0.120(1 - OM)]	**B8**
Prevost, 2004	Bd = 0.159 x 1.561 / [1.561 x OM + 0.159(1 - OM)]	**B9**
Perie & Ouimet, 2008	Bd = 0.111 x 1.767 / [1.767 x OM + 0.111(1 - OM)]	**B10**
Han et al., 2012	Bd = 0.167 x 1.526 / (1.526 x OM + 0.167(1 - OM)]	**B11**
Functions that use OC and grain-size data		
Tomasella & Hodnett, 1989	Bd = 1.578 - 0.054 x OC-0.006 x S-0.004 x C	**C1**
Kaur et al., 2002	ln(Bd) = 0.313 - 0.191 x OC + 0.02102 x C - 0.000476 x C^2 - 0.00432 x S	**C2**
Hollis et al., 2012	Bd = 0.80806 + 0.823844 x exp(-0.27993 x OC) + (0.0014065 x N) - (0.0010299 x C)	**C3** (cultivated topsoils)
Hollis et al., 2012	Bd = 0.697941 + (0.750636 x exp(-0.230355*OC) + (0.0008687 x N) - (0.0005164 x C)	**C4** (all other mineral horizons)
Hollis et al., 2012	Bd = 0.38502 + (1.04817 x exp(-0.070638 x OC) + (0.00090 x N) – (0.000715 x C)	**C5** (all other horizons)

*Bd = soil bulk density (g/cm3); OC = organic carbon (%); C = clay (%), S = silt (%), N = sand (%); OM = organic matter (%). In the case of OM, units are frequently expressed as %, or g/g for the functions B7, B8, B9, B10, B11. Codes are used for graphical representations and for separating functions derived by the same authors.

A frequent problem arises in the case of logarithmic functions, namely that below a certain value of OM or OC the estimated bulk densities register a sudden, unrealistic decrease, so that the functions cannot be applied in these intervals. For example, in the case of the B1 relation, the decrease is manifested at values of OM under 1%, for B2 and B3 the effect is manifested under 0.8% OM while in the case of A5 at values of OM lower than 0.5%. This problem is reflected in the upper flattening of the correlation graphs between real and estimated values of bulk density (Fig. 1).

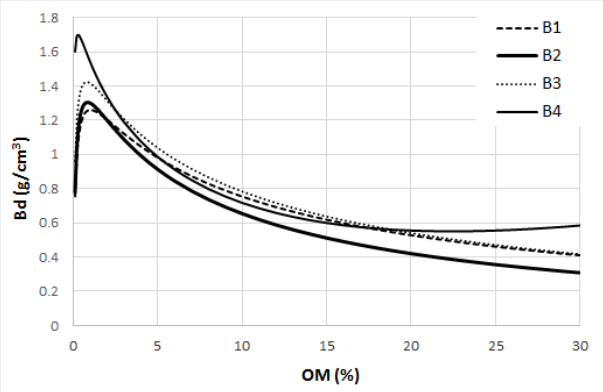


Figure 1. Theoretical distribution of predicted bulk density values for some logarithmic functions

In the case of the B1, B2 and B3 functions this flattening occurs at values of 1.257, 1.304 and 1.422 g/cm^3 respectively, which represent the maximum values that can be estimated by these functions. Similar problems occur in the case of linear functions (A1) or of those expressed as fractions (B6 … B11). In the case of A1 it can be seen that it also generates negative bulk densities at OC values higher than 13%.

All these problematic aspects impose the careful verification of the validity domains of the PTFs before they can be applied. The mentioned problems do not affect the exponential relations (A2, A3, A4, B5), which theoretically (mathematically) makes them superior to other functions.

Among the general findings that have been drawn in the papers analyzing, deriving or validating PTFs are the following:

- there is an inability of the models to explain more than 50-60% of the bulk density variance (Calhoun et al., 2001; Käterer et al., 2006), although some authors have obtained for their areas or datasets values of up to 80%;

- it seems that soil texture, OC and sample depth can only explain a part of the variation in bulk density. Terrain management, parent materials, particle density, some processes such as swelling or morphometric factors can affect bulk density (Calhoun et al., 2001; Wang et al., 2013);

- the prediction capacity of PTFs also depends on the mathematical concepts used in generating the models and on the variability of predictors (Al-Qinna & Jaber, 2013);

- with all the improvements brought in the last years, many of the published PTFs have limitations when applied to a larger area. This is explained by the fact that most models are based on small datasets, approach a single land use (e.g. forest) or include data that are not usually available (water contents). PTFs will give different results according to the environment / soil type for which they have been derived (De Vos et al., 2005; Hollis et al., 2012, Patil & Chaturvedi, 2012; Martin et al., 2007).

- the introduction of soil structure data might improve the prediction (Tranter et al., 2007; Hollis et al., 2012).

After selecting the PTFs to be tested according to data availability, the next step in determining the most suited function for bulk density is computing it for the samples in the database and comparing estimated and measured values.

2.4. Validation methodology

In general it is recommended that several statistical indices should be used in model validation. The validation parameters used by different authors (Donatelli et al., 2004) include the determination coefficient R^2 (1), the mean prediction error (MPE) (2), root mean square prediction error (RMSPE) (3), standard deviation of prediction error (SDPE) (4), maximum absolute error (ME) (5) and mean absolute error (MAE) (6).

$$R^2 = \frac{[\mathrm{cov}(E_i, M_i)]^2}{\mathrm{var}(E_i)\cdot \mathrm{var}(M_i)}\cdot 100 \quad \text{.........} (1)$$

$$MPE = \frac{1}{n}\sum_{i=1}^{n}(E_i - M_i) \quad \text{.........}(2)$$

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_i - M_i)^2} \quad \text{.........}(3)$$

$$SDPE = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[(E_i - M_i) - MPE]^2} \quad \text{.....} (4)$$

$$ME = \max|E_i - M_i| \quad \text{.........}(5)$$

$$MAE = \sum_{i=1}^{n}\frac{|E_i - M_i|}{n} \quad \text{.........}(6)$$

where n is the number of observations, Ei are the estimated values of bulk density, Mi are is the measured values of bulk density and var, cov stand for variance and covariance respectively.

The coefficient of determination indicates how well the data fits a statistical model, providing a measure of how well measured values are estimated by the model. The mean prediction error evaluates accuracy errors or the positive or negative systematic bias of the model, indicating a mean tendency of over- or underestimation. The RMSPE is a measure of the differences between predicted and observed values, being a good measure of model accuracy and of total prediction error. SDPE is often used for measuring precision errors, as it shows the random variation of the predictions after the application of a correction for global bias. SDPE is usually affected by the limited precision of the models, the local deviations of the models from real situations, the inherent variability of the measured property and by determination errors (De Vos et al., 2005).

While R^2 should have higher values, MPE, RMSPE and SDPE have to be as small as possible (Benites et al., 2007).

Alexander (1980) states that MPEs of 0.15-0.30 are acceptable, as bulk densities might increase through compaction. De Vos et al., (2005) consider that prediction errors should be of 0.14-0.29 for all the samples and of 0.12-0.25 for the topsoil ones.

3. RESULTS

3.1. Statistics of the database

The descriptive statistical indices calculated for the validation database (Table 2) show that the grain-size fractions are characterized by important dispersions in relation to the mean (variation coefficients between 39 and 47%). The reduced values of asymmetry indicate a relatively normal distribution of the values. On the other hand, values of organic matter and organic carbon are less dispersed around the mean (variation coefficients of 12.5-12.7%), yet their distributions present pronounced left asymmetries. This is indicated by the high skewness values and is explained by the presence of a few very large values associated to organic soils.

Bulk density values are characterized by a mean of 1.33 g/cm^3 and a maximum variation range between 0.46 and 1.85 g/cm^3. The distribution of the values is slightly right asymmetric due to the presence of some smaller values characteristic to the same organic / histic soils or of horizons with andic properties.

The values of bulk density are significantly differentiated according to soil classes (Fig. 2). The highest mean values are specific to Antrisols (1.48 g/cm^3) and Luvisols (1.41 g/cm^3). In the first case these values are explained by the inclusion of compacted soils, while in the later responsible is the high clay content of the Bt horizons. The lowest mean values of bulk density are those of Andisols (0.77 g/cm^3), explained by the influence of the volcanic parent material with high porosity, but also Histisols (0.84 g/cm^3), Spodisols (0.88 g/cm^3) and Umbrisols (1.006 g/cm^3) where the responsible factor is the abundance of weakly humificated organic materials. The most important variations in bulk density values are characteristic for Cambisols, due to varied parent materials these soils are formed on. Antrisols have the lowest variations, most probably due to generalized compaction that determines high bulk density values. The relation between the mean and the median indicates in the case Andisols, Antrisols, Cambisols, Histisols, Luvisols and Salsodisols positive asymmetries of the distributions, with a domination of low bulk density values. Umbrisols are characterized by a negative asymmetry, the mean being higher than the median and indicating a dominance of bulk density values higher than the mean. The other soil classes (Cernisols, Hidrisols, Pelisols, Protisols, Spodisols) have relatively equilibrated distributions.

Table 2. Descriptive statistics of the database used

	Sand	Silt	Clay	OM	OC	Bd
Min	3.36	0.5	0.17	0.03	0.017	0.46
Max	98.6	54.8	88.9	33.05	19.17	1.85
Mean	46.5	21.2	32.3	2.49	1.43	1.33
Median	43.3	21.9	31.7	1.68	0.96	1.36
St dev	19.4	8.29	15.3	3.16	1.78	0.197
Skewness	0.65	-0.21	0.33	4.32	4.22	-1.07
Kurtosis	0.14	0.55	0.21	25.94	24.80	2.11
1st quartile	34	16.9	22.3	0.8	0.46	1.24
3rd quartile	55.8	26.4	41.5	2.9	1.67	1.47
CV%	41.8	39.1	47.5	12.68	12.5	14.7

These differences in the statistics of the database indicate similar behaviors of some soil classes (Cernisols, Luvisols, Pelisols, Salsodisols, Protisols, characterized by high values of bulk density and moderate variations in relation to central values; Andisols, Histisols, Spodisols, Umbrisols with lower bulk density values and important variation limits), suggesting that a differentiated analysis in the elaboration of PTFs would be better.

3.2. Comparing PTFs performances

The determination coefficient values (Table 3, Fig. 4) demonstrate that a large part of the more simple functions, derived using only organic matter or carbon, achieve better results even if they do not exceed 0.44.

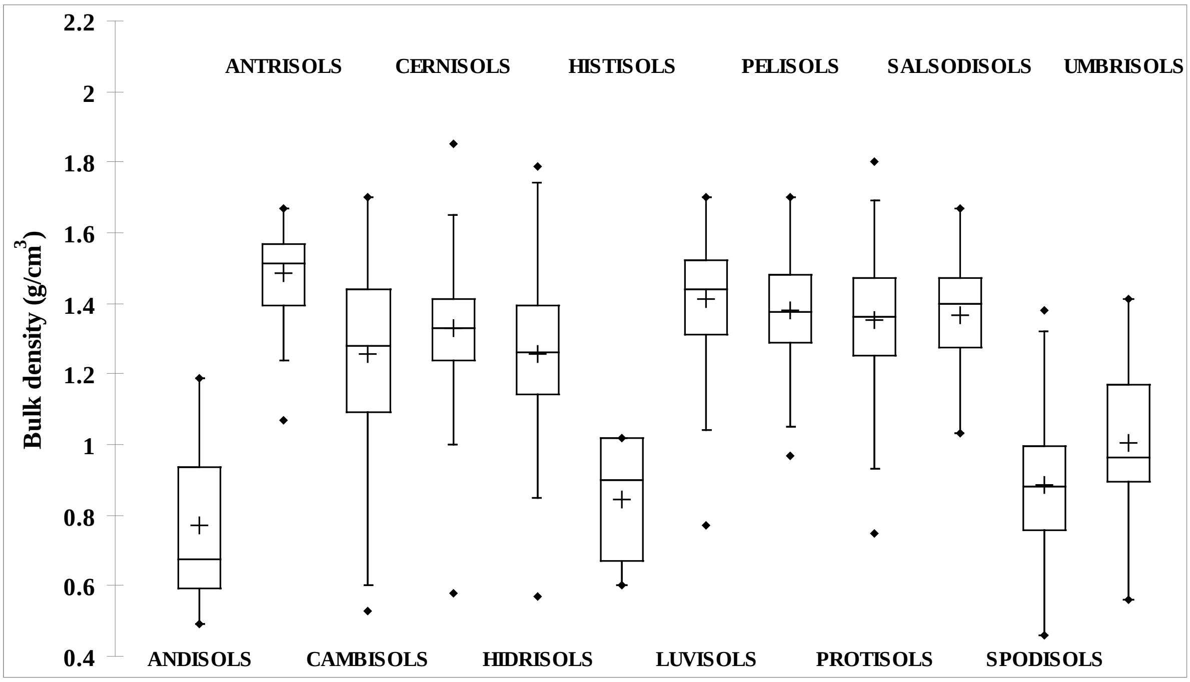


Figure 2. Bulk density variation according to soil class

The best results are obtained by the functions of Post & Kwon (2000), Han et al., (2012, B11), Tremblay et al., (2002), B10 of Perie & Ouimet (2008), B9 of Prevost (2004), Adams (1973), A2 of Manrique & Jones (1991) and Alexander (1980). The lowest values are obtained by the PTFs of Federer et al., (1993), Huntington et al., (1989), Curtis & Post (1964) and B3 of Prevost (2004) from the category of those using only organic matter, and by Tomasella & Hodnett (1998) or Kaur et al., (2002) among the PTFs including also textural data. The weak performance of most of the functions can be explained by the use of logarithms in PTFs, which determine a limitation at the upper limit of values, and in the case of Tomasella & Hodnett (1998) or Kaur et al., (2002) by the specific datasets used in PTF derivation. The reduced values of R^2 are also in strong relation to the database used, in which OM and texture data explain only 37-39%, respectively less than 1% of the bulk density variance. This indicates that the PTFs that include soil texture data will not necessarily give better results.

MPE values indicate the fact that a part of the functions underestimate bulk density (Kaur et al., 2002; Tremblay et al., 2002; Curtis & Post, 1964; Adams, 1973), while other overestimate (Harrison & Bocock, 1981). The PTFs that have the lowest MPEs are mostly those that use only organic matter or carbon: Manrique & Jones (1991), Alexander (1980), Prevost (2004), Han et al., (2012), but also some of the functions using soil texture, such as those of Hollis et al., (2012) (MPEs of 0.02-0.10).

In the case of RMSPE / SDPE, the functions giving better results are those of Manrique & Jones (1991), Alexander (1980), Han et al., (2012), B9 of Prevost (2004) and the PTFs of Hollis et al., (2012) for mineral horizons. The values of mean absolute errors indicate the same functions as performing better, with results of 0.123-0.132 g/cm^3. The ME are in most cases of 0.7-0.9 g/cm^3, although the functions of Harrison & Bocock (1981), A1 of Manrique & Jones (1991), Federer et al., (1993) and Kaur et al., (2002) stand out with values over 1.0 g/cm^3. The best results have been obtained by the B11, A3, B9, A2, B5 and B10 functions. With good overall results but with smaller R^2 values are the C4 and C5 functions of Hollis et al., (2012).

The same situation results from figure 3, which presents the MPE^2 vs SDPE^2 values of the tested models (suggested by De Vos et al., 2005) and shows the best PTFs located near the origin.

The performance of Han et al., (2012) and Hollis et al., (2012) functions might be influenced by the fact that they have been obtained from large datasets that have included diverse environmental conditions and soils from China or Europe. Manrique & Jones's (1991) PTFs have also been derived using a large dataset, even if for forest soils, many authors obtaining good results for it (including Hollis et al., 2012 for mineral soils). The use in PTF derivation only of forest soils might explain the weak performance for our database in the case of Curtis & Post (1964), Federer et al., (1993), Huntington et al., (1989), Adams (1973) or Tremblay et al., (2002).

Table 3. Validation statistics for the analyzed PTFs*

Models		**R^2**	**MPE**	**RMSPE**	**SDPE**	**MAE**	**ME**
Functions using OC							
Manrique & Jones, 1991	A1	0.376	**0.012**	**0.175**	0.175	**0.124**	1.283
Manrique & Jones, 1991	A2	**0.403**	**-0.013**	**0.160**	**0.159**	**0.123**	0.821
Alexander, 1980	A3	**0.403**	**-0.003**	**0.157**	**0.157**	**0.123**	0.831
Alexander, 1980	A4	**0.403**	**0.071**	**0.171**	**0.156**	**0.132**	0.904
Huntington et al., 1989	A5	0.334	-0.122	0.213	0.175	0.167	0.962
Harrison & Bocock, 1981	A6	0.302	0.257	0.367	0.261	0.287	1.638
Functions using OM							
Curtis & Post, 1964	B1	0.292	-0.199	0.263	0.171	0.217	1.044
Federer et al., 1983	B2	0.098	-0.151	0.261	0.213	0.196	1.216
Prevost, 2004	B3	0.131	**-0.025**	0.202	0.201	0.151	1.047
Perie & Ouimet, 2008	B4	**0.413**	0.151	0.230	0.173	0.189	1.028
Han et al., 2012	B5	**0.403**	**-0.040**	**0.164**	**0.159**	**0.129**	**0.741**
Adams, 1973	B6	**0.436**	-0.184	0.236	**0.149**	0.202	0.657
Post & Kwon, 2000	B7	**0.438**	**0.123**	0.194	**0.149**	0.156	0.939
Tremblay et al., 2002	B8	**0.432**	-0.215	0.265	**0.155**	0.226	0.829
Prevost, 2004	B9	**0.435**	**-0.016**	**0.159**	**0.158**	**0.123**	**0.787**
Perie & Ouimet, 2008	B10	**0.419**	**0.033**	0.206	0.204	0.162	0.834
Han et al., 2012	B11	**0.436**	**-0.031**	**0.157**	**0.154**	**0.122**	**0.775**
Functions using OC and grain-size data							
Tomasella & Hodnett, 1989	C1	0.187	**-0.090**	0.205	0.184	0.165	0.800
Kaur et al., 2002	C2	0.227	-0.246	0.371	0.278	0.288	1.166
Hollis et al., 2012	C3	0.340	**0.100**	0.198	0.170	0.155	0.998
Hollis et al., 2012	C4	0.372	**-0.043**	**0.163**	**0.157**	**0.128**	0.824
Hollis et al., 2012	C5	0.346	**0.022**	**0.160**	**0.159**	**0.125**	0.884

*Values in bold indicate functions performing better.

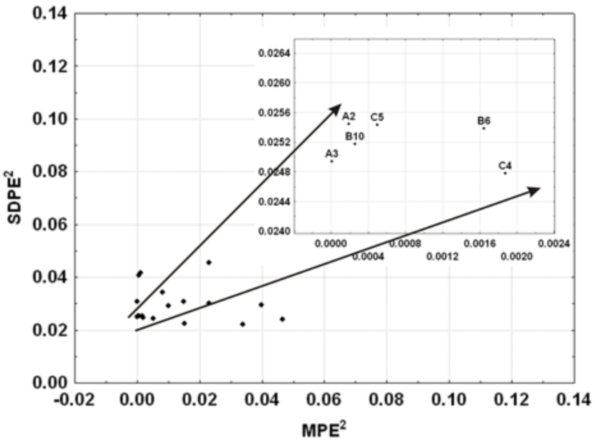


Figure 3. Standard deviation of prediction error (SDPE2) vs. mean prediction error (MPE2) of the tested PTFs

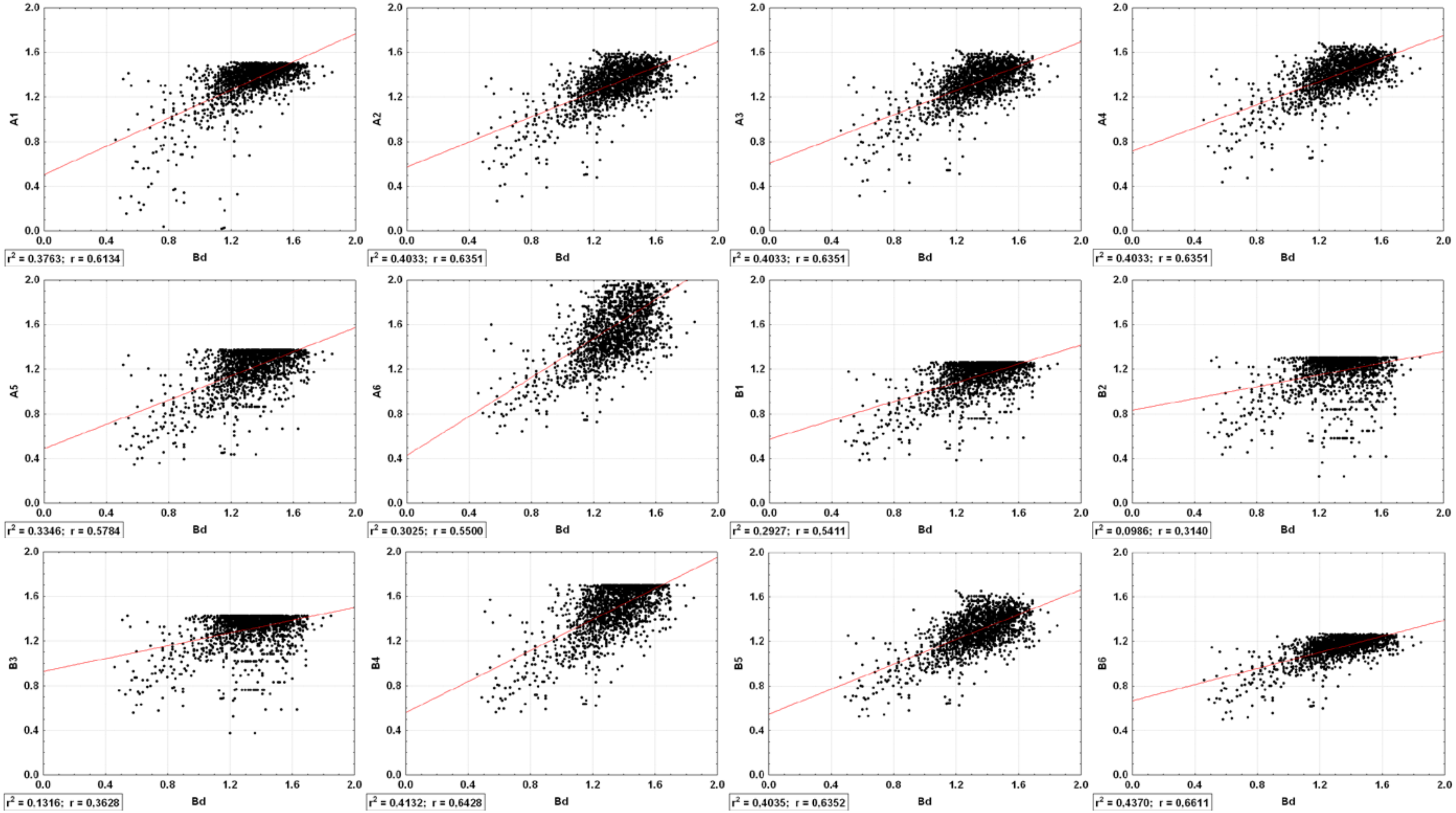
Besides the enumerated functions, other have also been tested (Bernoux et al., 1998; Leonaviciute, 2000; Benites et al., 2007; Brahim et al., 2012), but the results have been very weak, most probably due to the databases used in derivation (the functions of Benites et al., 2007 and Bernoux et al., 1998 have been elaborated for the Amazonian area, while the database used by Leonavicuite (2000) includes very high values of bulk density, up to 2.1).

Data stratification, recommended by Heuscher et al., (2005) or Hollis et al., (2012) does not help improving the results. For example in the case of the PTFs of Manrique & Jones (1991) and Alexander (1980), applying them only to forest soils led to a decrease in the MPE values (from 0.01 to 0.008, from -0.013 to 0.008, from 0.071 to 0.026) but also to a decrease in R^2 (from 0.376 to 0.346, from 0.403 to 0.399) and an increase in RMSE (0.17 to 0.26, 0.16 to 0.22, 0.15 to 0.21, 0.17 to 0.24).

4. CONCLUSIONS

In conclusion, the validation of published PTFs with the help of a Romanian dataset has shown that these functions give different results. Many of the PTFs derived by different authors cannot be applied to the Romanian territory.

These functions have obtained poor scores due most probably to their derivation for small areas or from small datasets. Also, some of the functions are limited through their mathematical expression at certain bulk density values, while others tend to underestimate at small OM contents. At the same time, OM only explains 37-39% of the bulk density variation, and soil texture has a small influence. With some exception, functions using OM or OC have had relatively similar performances.


A1
A2
A3
A4
A5
A6
B1
B2
B3
B4
B5
B6
Bd
r² = 0.3763; r = 0.6134
r² = 0.4033; r = 0.6351
r² = 0.4033; r = 0.6351
r² = 0.4033; r = 0.6351
r² = 0.3346; r = 0.5784
r² = 0.3025; r = 0.5500
r² = 0.2927; r = 0,5411
r² = 0,0986; r = 0,3140
r² = 0,1316; r = 0,3628
r² = 0,4132; r = 0,6428
r² = 0,4035; r = 0,6352
r² = 0,4370; r = 0,6611

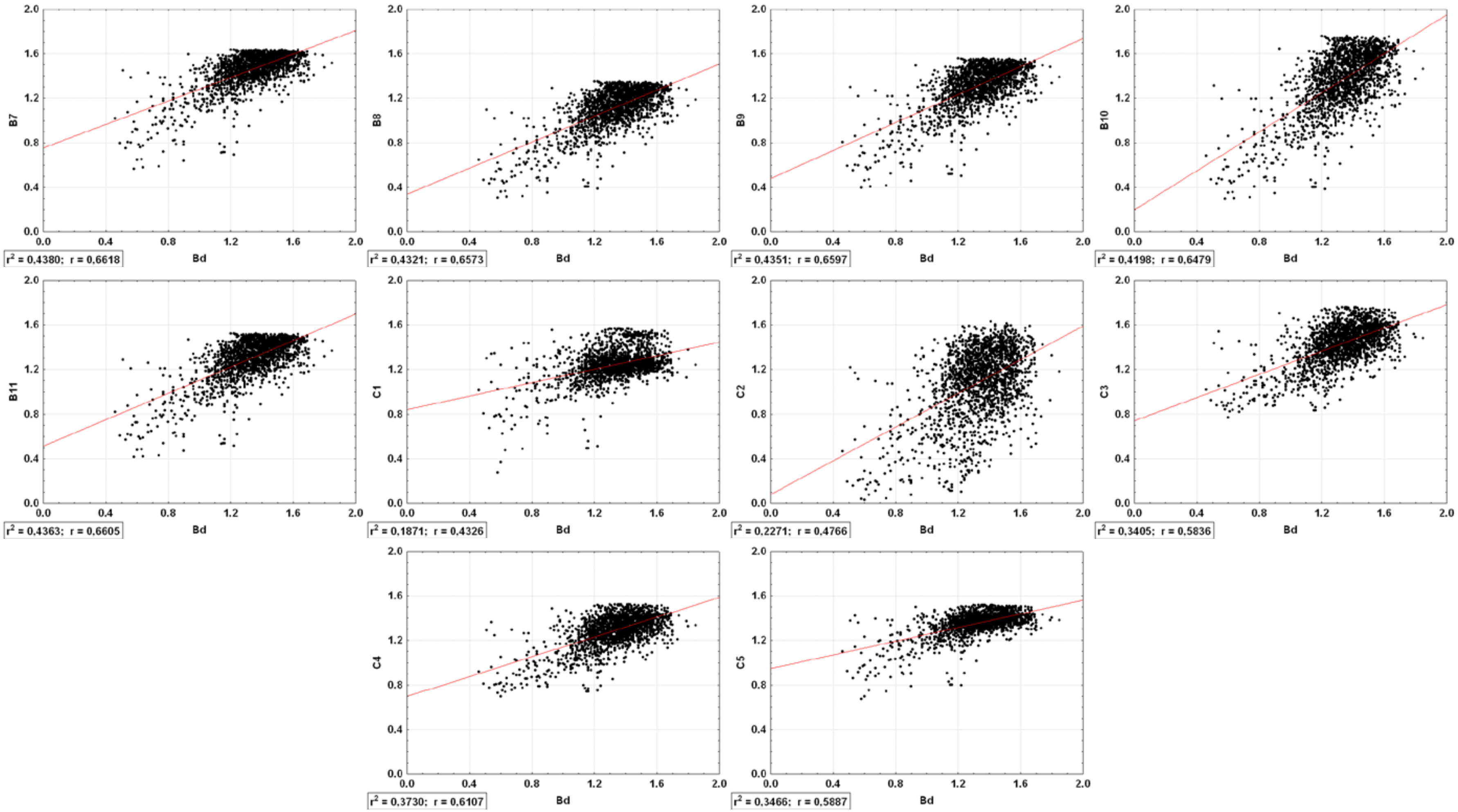


Figure 4. Measured vs. predicted values of bulk density for the analyzed functions

The functions performing better are A2 of Manrique & Jones (1991), Alexander (1980), Han et al., (2012) and B9 of Prevost (2004). With good statistics for SDPE, RMSE or ME, these functions still do not obtain R^2 values higher than 0.43. Although using besides OM grain-size data, the functions from group C did not have better performances.

It is clear in this situation that when estimating bulk density, one should take carefully into account the method of PTF derivation, the dataset used, and the performance of the mathematical expression used. Also, in the case of large datasets such as the one used, a further stratification based on soil classes or types might improve the results.

ACKOWLEDGMENTS

This work was supported by the strategic grant POSDRU/159/1.5/S/133652, co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 – 2013.

REFERENCES

Adams, W.A., 1973. *The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils*. J. Soil Sci. 24, 10-17.

Alexander, E.B., 1980. *Bulk densities of California soils in relation to other soil properties*. Soil Science Society of America Journal, 44, 689-692.

Al-Qinna, M.I. & Jaber, S.M., 2013. *Predicting soil bulk density using advanced pedotransfer functions in an arid environment*. Transactions of the ASABE, 56(3), 963-976.

Apostol, L. & Sfîcă, L., 2013, *Thermal differentiations induced by the Carpathian Mountains on the Romanian territory*. Carpathian Journal of Earth and Environmental Sciences, 8, 2, 215.221.

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., Arrouays, D., Cerri, C., Volkoff, B. & Jolivet, C., 1998. *Bulk densities of Brazilian Amazon soils related to other soil properties*. Soil Sci. Soc. Am. J., 62, 743-749.

Boucneau, G., Van Meirvenne, M. & Hofman G., 1998. *Comparing pedotransfer functions to estimate soil bulk density in northern Belgium*. Pedologie Themata, 5, 67-70.

Bouma, J., 1989. *Using soil survey data for quantitative land evaluation*. Advances in Soil Science, 9, 177–213.

Brahim, N., Bernoux, M. & Gallali, T., 2012. *Pedotransfer functions to estimate soil bulk density for Northern Africa: Tunisia case*. J Arid Environ, 81, 77-83.

Canarache, A., 1990. *Physics of agricultural soils* (in Romanian). Ed. Ceres, Bucureşti. 268 pp.

Canarache, A., 1995. *Observations regarding bulk density of organic and organo-mineral materials (special view on andic and spodic soils)*. Soil Forming Factors and Processes from the Temperate Zone, 3, 103-109

Calhoun, F.G., Smeck, N.E., Slater, B. L., Bigham, J.M. & Hall, G. F., 2001. *Predicting bulk density of Ohio soils from morphology, genetic principles, and laboratory characterization data*. Soil Sci. Soc. Am. J., 65, 811-819.

Chiriță, C.D., 1970. *Study on the acidity and base saturation of acid soils (with application to spruce soils from the Romanian Carpathians)* (in Romanian). Analele ISCP, 38, 45-76.

Curtis, R.O. & Post B.W., 1964. *Estimating bulk density from organic matter content in some Vermont forest soils*. Soil Sci. Soc. Am. Proc., 28, 285-286.

De Vos, B., Van Meirvenne, M., 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.

Donatelli, M., Wösten, H. & Belocchi, G., 2004. *Methods to evaluate pedotransfer functions*. In *Developments in Soil Science* (ed. Pachepsky, Y., Rawls, W.J.), 30, 357-411.

Federer, C.A., Turcotte, D.E. & Smith, C.T., 1993. *The organic fraction-bulk density relationship and the expression of nutrient content in forest soils*. Canadian Journal of Forest Research, 23, 1026-1033.

Florea, N., Bălăceanu, V., Răuță, C., Canarache, A., (eds.). 1987. *Methodology of soil surveys* (in Romanian). Center for Didactic Material and Agricultural Propaganda, Bucharest (vol. 1-3).

Florea, N., Munteanu, I., (eds.). 2012. *Romanian System of Soil Taxonomy* (in Romanian). Sitech, Craiova.

Gogoașă, T., 1959. *Titrimetric determination of soil humus* (in Romanian). D.S. Comit. Geol. Rom. 42, 117-128.

Han, G.Z., Zhang, G.L., Gong, Z.T. & Wang, G.F., 2012. *Pedotransfer functions for estimating soil bulk density in China*. Soil Sci., 177(3),158-164.

Harrison, A. F. & Bocock, K. L., 1981. *Estimation of soil bulk-density from loss-on-ignition values*. J. Appl. Ecol., 8, 919-927.

Heuscher, S.A., Brandt, C.C. & Jardine, P.M., 2005. *Using soil physical and chemical properties to estimate bulk density*. Soil Sci. Soc. Am. J., 69, 51-56.

Hollis, J.M., Hannam, J. & Bellamy, P.H., 2012. *Empirically-derived pedotransfer functions for predicting bulk density in European soils*. European Journal of Soil Science, 63, 96-109.

Huntington, T.G., Johnson C.E., Jonhson A.H., Sicama T.G. & Ryan D.F., 1989. *Carbon,* *organic matter, and bulk density relationships in a forested Spodosol*. Soil Sci., 148, 380-386.

Kätterer, T., Andrén, O. & Jansson, P.E., 2006. *Pedotransfer functions for estimating plant available water and bulk density in Swedish agricultural soils*. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science, 56 (4), 263-276.

Kaur, R., Kumar, S. & Gurung, H.P., 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.

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.

Martin, M.P., Lo Seen, D., Boulonne, L., Jolivet, C., Nair, K.M., Bourgeon, G. & Arrouays, D., 2009. *Optimizing pedotransfer functions for estimating soil bulk density using boosted regression trees*. Soil Sci Soc Am J., 73, 485-493.

McBratney, A.B., Minasny, B., Cattle, S.R. & Vervoort, R.W., 2002. *From pedotransfer function to soil inference system*. Geoderma, 109, 41-73.

Nanko, K., Ugawa, S., Hashimoto, S., Imaya, A., Kobayashi, M., Sakai, N., 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.

Patil, G.N. & Chaturvedi, A., 2012. *Estimation of bulk density of waterlogged soils from basic properties*. Archives of Agronomy and Soil Science, 58:5, 499-509.

Perie, C. & Ouimet R., 2008. *Organic carbon, organic matter and bulk density relationships in boreal forest soils*. Can. J. Soil Sci., 88, 315-325.

Popovici, E.A., Bălteanu, D. & Kucsicsa G., 2013. *Assessment of changes in land-use and land-cover pattern in Romania using Corine Land Cover database*. Carpathian Journal of Earth and Environmental Sciences, 8, 4, 195-208.

Post, W.M. & Kwon K.C., 2000. *Soil carbon sequestration and land-use change: Processes and potential*. Global Change Biol., 6, 317-327.

Prevost, M., 2004. *Predicting soil properties from organic matter content following mechanical site preparation of forest soils*. Soil. Sci. Soc. Am. J., 68, 943-949.

Ruehlmann, J. & Korschens M., 2009. *Calculating the effect of soil organic matter concentration on soil bulk density*. Soil. Sci. Soc. Am. J., 73, 876-885.

Suuster, E., Ritz, C., Roostalu, H., Reintam, E., Kolli, R. & Astover, A., 2011. *Soil bulk density pedotransfer functions of the humus horizon in arable soils*. Geoderma, 163(1-2), 74-82.

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 (4), 437-443.

Tremblay, S., Ouimet R. & Houle D., 2002. *Prediction of organic carbon content in upland forest soils of Quebec, Canada*. Can. J. For. Res., 32, 1-12.

Wang, Y., Shao, M., Liu, Z. & Zhang, C., 2014. *Prediction of bulk density of soils in the Loess Plateau region of China*. Surv. Geophyis., 35, 395-413.

Wösten, J.H.M., Lilly A., Nemes A. & Le Bas C., 1999 In the text is 1998. *Development and use of a database of hydraulic properties of European soils*. Geoderma, 90, 169-185.

Received at: 10. 01. 2015
Revised at: 16. 04. 2015
Accepted for publication at: 20. 04. 2015
Published online at: 22. 04. 2015