PEDO-TRANSFER FUNCTIONS FOR ESTIMATING SOIL BULK DENSITY IN CENTRAL AMAZONIA

Henrique Seixas Barros[(1)]* and Philip Martin Fearnside[(1)]

[(1)] Instituto Nacional de Pesquisas da Amazônia, Laboratório de Capacidade de Suporte Humano e Agroecossistemas, Manaus, Amazonas, Brasil.

* Corresponding author.
E-mail: henriqueseixasbarros@yahoo.com

ABSTRACT

Under field conditions in the Amazon forest, soil bulk density is difficult to measure. Rigorous methodological criteria must be applied to obtain reliable inventories of C stocks and soil nutrients, making this process expensive and sometimes unfeasible. This study aimed to generate models to estimate soil bulk density based on parameters that can be easily and reliably measured in the field and that are available in many soil-related inventories. Stepwise regression models to predict bulk density were developed using data on soil C content, clay content and pH in water from 140 permanent plots in *terra firme* (upland) forests near Manaus, Amazonas State, Brazil. The model results were interpreted according to the coefficient of determination (R^2) and Akaike information criterion (AIC) and were validated with a dataset consisting of 125 plots different from those used to generate the models. The model with best performance in estimating soil bulk density under the conditions of this study included clay content and pH in water as independent variables and had $R^2 = 0.73$ and AIC = -250.29. The performance of this model for predicting soil density was compared with that of models from the literature. The results showed that the locally calibrated equation was the most accurate for estimating soil bulk density for upland forests in the Manaus region.

Keywords: Amazon forest, *terra firme* (upland), Manaus, Brazil, soil carbon stocks.

RESUMO: *ESTIMATIVA DA DENSIDADE DO SOLO POR EQUAÇÕES DE PEDOTRANSFERÊNCIA NA AMAZÔNIA CENTRAL*

A densidade do solo é um parâmetro de difícil aferição no campo da floresta amazônica e exige rigor metodológico para realizar inventários dos estoques de C e nutrientes nos solos, fazendo com que esse

Received for publication on June 24, 2014 and approved on December 23, 2014.
DOI: 10.1590/01000683rbcs20140358

processo se torne oneroso e, certas vezes, inviável, dependendo das condições de trabalho. Objetivou-se com este estudo gerar modelos para estimar a densidade do solo a partir de parâmetros facilmente medidos em campo com confiabilidade, presentes em muitos inventários pedológicos. A construção dos modelos de regressão para estimar a densidade do solo usou dados sobre o teor de C, conteúdo de argila e pH em água medidos em 140 parcelas permanentes em florestas de terra-firme da região de Manaus, Estado do Amazonas, Brasil. Os resultados foram interpretados de acordo com o coeficiente de determinação (R^2) e o critério Akaike de informação (AIC) dos modelos. Os modelos foram validados por meio de uma série de dados composta por 125 parcelas diferentes das usadas na geração dos modelos. O modelo que melhor estimou a densidade nas condições desse estudo foi o que incluiu o conteúdo de argila e o pH em H_2O como variáveis independentes e apresentou $R^2 = 0,73$ e AIC = -250,29. O modelo desta pesquisa foi comparado a equações existentes na literatura quanto ao desempenho da predição dos valores de densidade do solo. Os resultados apontaram que o uso da equação calibrada localmente evidenciou melhor desempenho para estimar a densidade do solo para as florestas de terra-firme da região de Manaus.

Palavras-chave: densidade do solo, floresta amazônica, terra-firme, Manaus, Brasil, floresta tropical, solos tropicais.

INTRODUCTION

Soil bulk density is used in the quantification of soil C stocks (Veldkamp, 1994), and is therefore an important parameter for national inventories of greenhouse gas (GHG) emissions under the United Nations Framework Convention on Climate Change (UNFCCC). However, this parameter is a major source of uncertainty for estimates of soil C stocks (Fearnside and Barbosa, 1998; Taalab et al., 2012). Data are scarce because sampling undisturbed cores is laborious and data are only reliable if the methodological rigor is high. The result is that soil carbon stocks are commonly estimated from mean bulk density (Bd) values from the literature and from values for C concentration measured in the field (Bernoux et al., 1998).

Although Bd is calculated as the ratio of soil mass to volume, both of which are easily measured variables, reliable information on soil Bd is difficult. This has stimulated predictions of soil density that exploit the relationship between this parameter and other variables commonly found in soil-related inventories in order to ensure the reliability of carbon stock assessments and to reduce evaluation costs (Federer et al., 1993; Bernoux et al., 1998, 2002; Tomasella and Hodnett, 1998; Calhoun et al., 2001; Heuscher et al., 2005; Benites et al., 2007; Tranter et al., 2007; Steller et al., 2008; Gharahi-Ghehi et al., 2012; Chaudhari et al., 2013).

Pedo-transfer functions (PTFs) have been widely used in soil studies to estimate values that are difficult to measure in the field (Minasny and Hartemink 2011). The PTFs constructed from commonly available parameters in soil inventories, such as organic C content and clay amount, are highly promising to represent direct Bd measurements where these are difficult to access or unavailable (Benites et al., 2007).

Thus, the trend is to increasingly generate estimates locally, to reduce the uncertainties of equations and calculations derived from this parameter. It should be noted that measuring the actual soil density values in the field is always more reliable than estimates based on variables that are equally or even more complex and with high spatial variability.

Bernoux et al. (1998) generated equations to estimate density from data series of the RADAMBRASIL inventories, published between 1973 and 1982 by the Ministry of Mines and Energy. Despite the vast amount of data on geology and geomorphology, vegetation, soils and land use, these surveys lack information on soil density. This led to the creation of equations to estimate this parameter.

In view of the need to create equations that are locally better adjusted, the purpose of this study was to generate regression models to estimate density from soil parameters available in the soil inventories of the Project Biological Dynamics of Forest Fragments (PDBFF) and to compare their performance in predicting bulk density of upland forest soil of the region of Manaus between the locally generated equations and equations available in the literature.

MATERIAL AND METHODS

The study was conducted in the reserves of the PDBFF (Figure 1) located approximately 70 km north of Manaus (2° 30' S, 60° W). The altitude at the site varies between 50 and 100 m above sea level, average annual temperature was 26.7 °C in Manaus, and average annual rainfall 2,200 mm, with a pronounced dry season from July to September (Fearnside and Leal Filho, 2001).

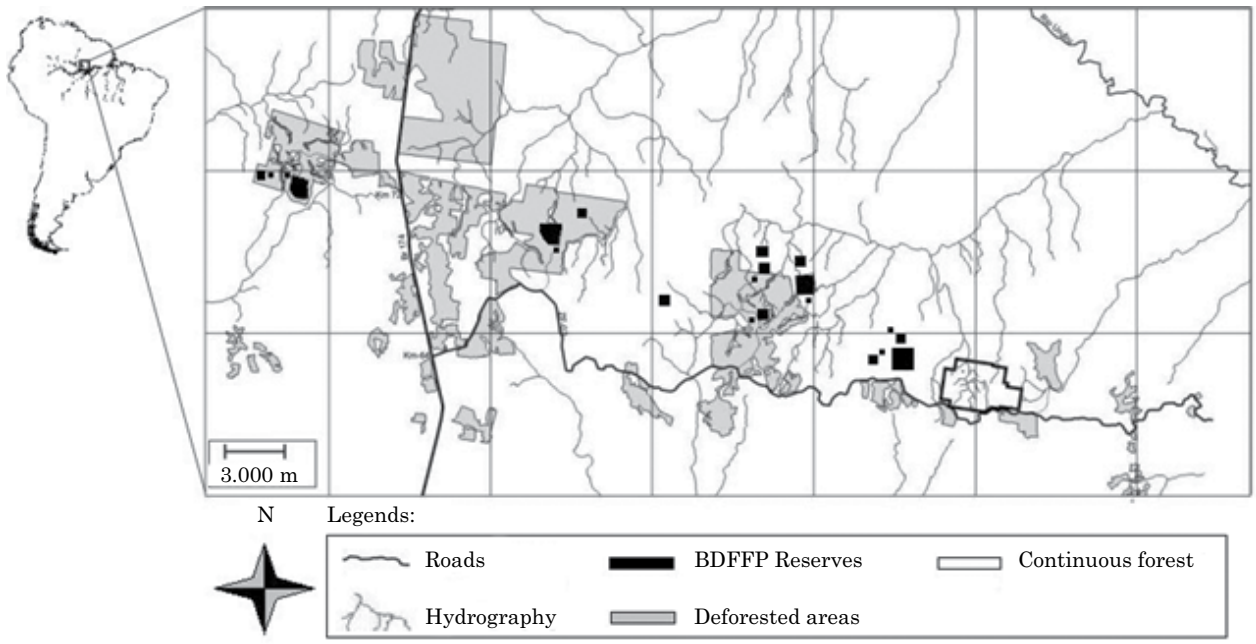

Figure 1. Study area. Source: Project Biological Dynamics of Forest Fragments [PDBFF].

The soils in the region were classified as Latossolos Amarelos álicos (Oxisol), according to RADAMBRASIL maps (Fearnside and Leal Filho, 2001); they are highly leached, acidic and nutrient-poor (Chauvel, 1982; Chauvel et al., 1987). According to the current Brazilian system of soil classification (SiBCS), these soils are classified in the category of Latossolos Amarelos alumínicos (Oxisol) (Embrapa, 2013).

The typical vegetation in this part of central Amazonia is characterized as upland dense forest. The diversity of trees can be considered high, with an average of over 280 species (diameter at breast height [DBH] > 10 cm) per hectare in continuous forest (Oliveira and Mori, 1999).

In each of the PDBFF reserves, one or more 1-ha grid plots were permanently installed and divided into 25 quadrats (20 × 20 m, here called "plots"). Each vertex of these plots was assigned a letter and a number, creating a system of coordinates for the exact localization of the sampling points in the field

Soil sampling

Samples for C concentration determination were obtained using of a screw auger. Each individual sample for C content was composed of five subsamples taken on all four sides and in the center of each plot. To determine soil density, a sample was collected in the center of each plot using a stainless steel cylindrical auger specifically designed to collect undisturbed samples, similar to kopeck volumetric rings (length 20 cm, diameter 5 cm, volume 0.3925 dm^3). All samples were collected from the 0-20 cm layer and stored in plastic bags to be transported and properly processed in the laboratory. In total, soil samples were collected in 265 plots distributed in 22 1-ha grids in the PDBFF forest inventories in Central Amazonia.

Preparation and processing of samples

In the laboratory, samples for C analysis were dried in a solar oven and then sieved first through 20 mm mesh and then through 2 mm mesh. The plant roots and other visible fractions were removed and a fraction of each specimen was ground and reduced to particles with maximum diameter of 50 microns before automatic chemical analysis.

Samples for determination of bulk density (Bd) were placed to dry in aluminum containers in an electric oven at 105 ° C for approximately 72 h.

Analysis of total carbon content

Total C content (g/kg) of each sample was determined using an element analyzer (Varo Max model, CN Elemental Instruments, Hanau, Germany). This device employs the dry combustion technique, which is the conversion of elements in the samples to simple gases such as CO_2 and N_2. The resulting gases are mixed and maintained under standard conditions of pressure, temperature and volume, and are then depressurized in a column where they are detected, identified and separated based on their thermal conductivity (Pérez et al., 2001).

Determination of soil bulk density

The soil density in each plot was calculated by dividing the dry weight of the sample by the volume of the collection cylinder (0.3925).

Determination of clay content

Clay content (%) was determined by particle size analysis by the pipette method (Embrapa, 1979).

Determination of pH in water

Soil reaction (pH) in distilled H_2O [pH(H_2O)] was estimated by a pH-meter. The ratio between the amount of oven-dried soil and the amount of water used was 1:1 on a volume basis, 20 mL of soil to 20 mL of water (Fearnside and Leal Filho, 2001).

Model development

To estimate Bd, multiple-regression equations were generated by the stepwise routine in SYSTAT software (Wilkinson, 1990). The independent variables used in the models were chosen according to the information available in the PDBFF database, as well as their consistency with the soil variables most commonly used in developing PTFs to estimate density (C content amount of clay and pH measured in water). The model was constructed using data from 140 plots.

In order to assess the degree of correlation between the independent variables, the Pearson test was applied to this data series. The presence of co-linearity was observed according to the values of the parameters.

The performance of PTFs for predicting soil density was evaluated based on the Akaike information criterion (AIC) and the coefficient of determination (R^2), thereby validating the models. The AIC value allows a comparison and classification of multiple competing models and to estimate which is closest to the "real" process underlying the biological phenomenon under study (Akaike, 1973; Bozdongan, 1987; Burnham and Anderson, 2002; Burnham et al., 2011; Symonds and Moussalli, 2011). The coefficient of determination is the proportion of the variation in soil density that can be explained by the set of predictor variables.

The model was validated using a data set consisting of 125 samples different from those used to generate the equations. The predicted values were then plotted against the observed values to evaluate the performance of the estimates. We also plotted the residuals of the regressions against the estimated values to verify the premise of homoscedasticity. Normality of residuals was tested by the Kolmogorov-Smirnov method at the 95 % significance level.

Comparisons of pedo-transfer functions (PTFs)

This study's best model for predicting bulk density was compared with three models from the literature that were generated from samples collected throughout the Amazon Basin (Bernoux et al., 1998; Tomasella and Hodnett, 1998; Benites et al., 2007). These equations were used to predict density from the data series used to validate our models. For the model of Tomasella and Hodnett (1998), the dataset contained a limited number of plots due to the lack of information on soil texture in some locations. The result of these predictions were plotted in graphs of predicted values versus the values observed in the field. The comparison included analysis of the residual graphs to check homoscedasticity.

RESULTS

Model development and selection

Descriptive statistics for the data series used in the construction of density prediction models are presented in table 1. The Pearson correlation matrix indicates that there is co-linearity between soil carbon and clay (0.80), which creates an unwanted bias in composite models for these variables. The pH(H_2O) was not correlated neither with carbon (-0.21) nor clay (-0.40).

The regression model results show $R^2 = 0.74$ and AIC = -251.88 for Model 1 equation (Table 2). Figure 2 shows the observed density versus estimated values for each pedo-transfer equation. Model 1 fulfilled the homoscedasticity premise, as confirmed by analyzing the point distribution on the graph of regression residuals versus the estimated values (Figure 3). Furthermore, the result of the Kolmogorov-Smirnov test for normality at 95 % significance showed a normal distribution of the regression residuals (Table 3).

Table 1. Descriptive statistics for the model parameters

	Serie for model development			Serie for model validation		
Parameter	**Carbon**	**Clay**	**pH(H_2O)**	**Carbon**	**Clay**	**pH(H_2O)**
	g kg^{-1}	%		g kg^{-1}	%	
Mean	1.62	48.12	3.92	1.63	48.15	3.98
Median	1.59	49.68	4.02	1.62	50.71	4.06
Minimum	2.78	75.06	4.81	3.33	73.50	4.61
Maximum	0.81	15.23	3.20	0.73	5.81	2.20
Standard deviation	0.42	15.36	0.37	0.39	16.67	0.37

Table 2. Regression model selection results (ordinary least squares [OLS])

Model	Variable	Intercept	#1_Carbon	#2_Clay	#3_pH	Standard deviation	AICc	R^2
1	1, 2, 3	1.51	-0.06	-0.01	-0.07	0.09	-251.88	0.74
2	1, 2	1.19	-0.07	-0.01	...	0.10	-243.24	0.72
3	1	1.20	-0.32	...	...	0.12	-180.52	0.55
4	2, 3	1.49	...	-0.01	-0.08	0.09	-250.30	0.73
5	2	1.15	...	-0.01	...	0.10	-240.67	0.70
6	3	0.44	...	...	0.06	0.18	-73.07	0.01
7	1, 3	1.26	-0.33	...	-0.01	0.12	-178.73	0.55

Model 2, in which C and clay contents were explanatory variables, $R^2 = 0.72$ and AIC = -243.24. The residuals of this regression were shown to be homoscedastic (Figure 3) and followed the normal distribution (Table 3).

The third model generated was based solely on C content as predictor variable (Figure 2; Table 2). The model had a lower coefficient of determination compared to the previous models, with $R^2 = 0.55$ and AIC value = 180.52. Nevertheless, the regression residuals showed homoscedasticity and normal distribution, according to the Kolmogorov-Smirnov test (Figure 3; Table 3).

Model 4 included clay content and pH(H_2O) as independent variables. The coefficient of determination of this model was similar to that of the first equation with $R^2 = 0.73$ and AIC value = -250.29, this being the second lowest value of the models tested (Table 2). The validation of this model indicated approximately 60 % correspondence between the observed values and those predicted by the equation. The plot of the residuals versus model-estimated values (Figure 3) showed homoscedasticity of the points, and the Kolmogorov-Smirnov test at 95 % significance confirmed normal distribution (Table 3).

Clay content was the variable which, separately, had the highest predictive power, as indicated by the results of Model 5. This parameter was responsible for explaining about 70 % of the variation in soil density. The model had a random distribution (Figure 3), and the regression residuals were normally distributed (Table 3).

The pH(H_2O) was included as the only independent variable in Model 6 (Table 2). The extremely low coefficient of determination ($R^2 = 0.01$) and the high AIC value (-73.07) indicated that this variable is not significantly related to variation in soil Bd. The plot of residuals *versus* fitted values showed a pattern of clustering of points, and the Kolmogorov-Smirnov test at 95 % significance indicated a non-normal distribution, invalidating this model.

In the 7th and last model, C content and pH(H_2O) were used as predictor variables ($R^2 = 0.55$ and AIC = - 178.72). The model was validated by the homoscedasticity seen in figure 3 and the normal distribution of the residuals (Table 3).

Despite the similarity between the results of Models 4 and 1, the existence of co-linearity among the predictors of the latter model compromises its reliability. Therefore, Model 4 had a better performance in predicting Bd under the conditions of this study and was selected for comparison with models in the literature.

Model comparisons

Using the equation of Bernoux et al. (1998) to estimate soil density from the validation series of this study resulted in a $R^2 = 0.56$, which shows the degree of relationship between estimated density values for this model *versus* the values observed in the field. However, the graph of residuals versus estimated values showed clustering, suggesting that the residuals are not normally distributed, making the validity of this prediction questionable (Figure 4). Using this equation resulted in an overestimation of values, increasing the predicted average by around 72 % over the average of values obtained in the field (Table 4).

Results of the equation of Benites et al. (2007) also indicated a tendency for overestimation in the predicted values, with an average increase of about 130 %, compared to the observed values (Table 4). The yield of the regression between predicted and observed values was lower when the estimate was made from the validation series of this study, with $R^2 = 0.63$ in the original study *versus* $R^2 = 0.56$, when applied to the data set of this study (Figure 4).

When used to estimate soil density from the field data in this study, the equation of Tomasella and Hodnett (1998) showed the same trend observed earlier, resulting in overestimation of the predicted values of around 96 % above the average of the values for observed density (Table 4). The coefficient of determination of the regression was only 0.31, and the residuals were not randomized (Figure 4).

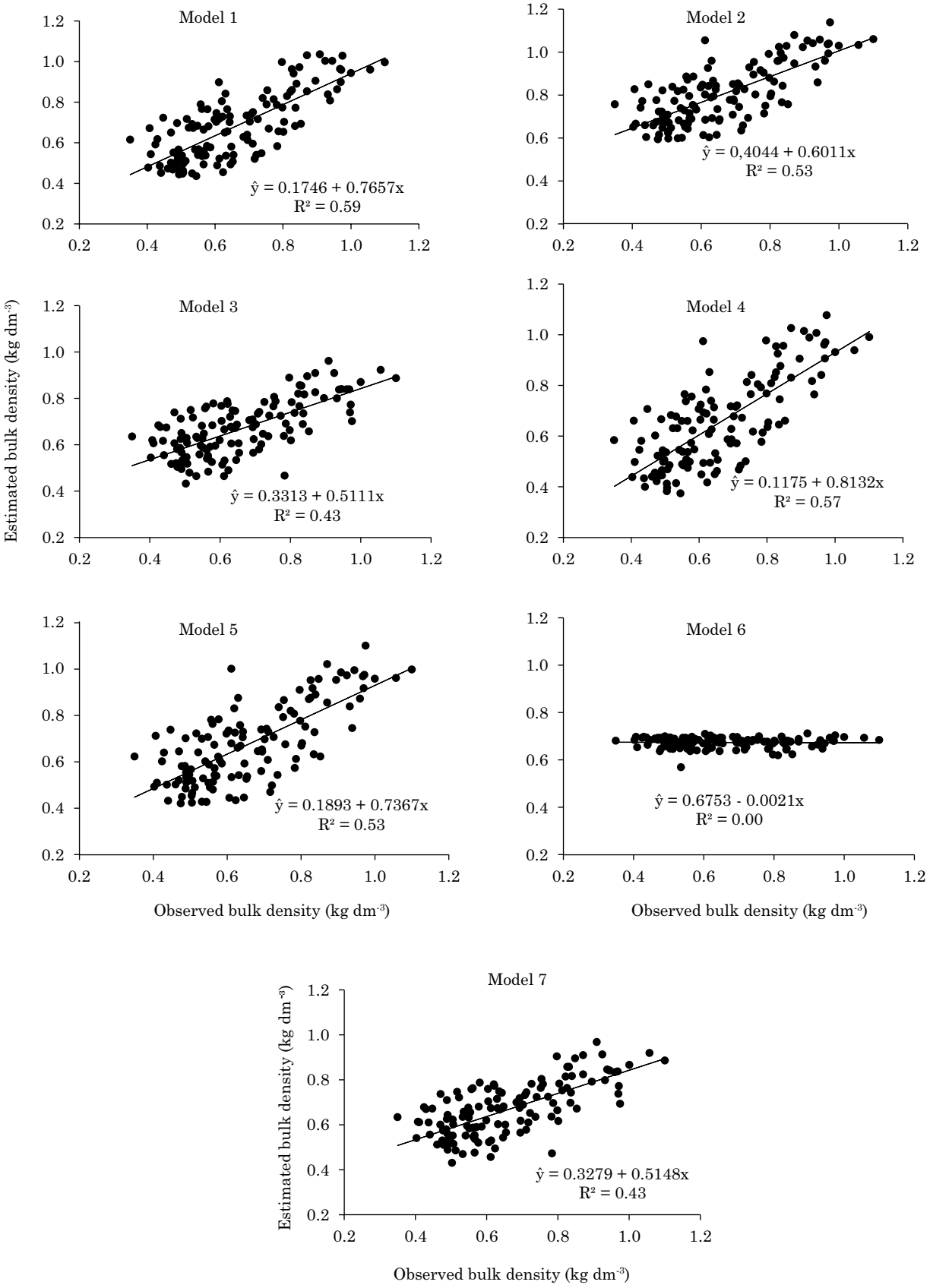

Figure 2. Validation of the regression models generated by the stepwise method for predicting soil density.

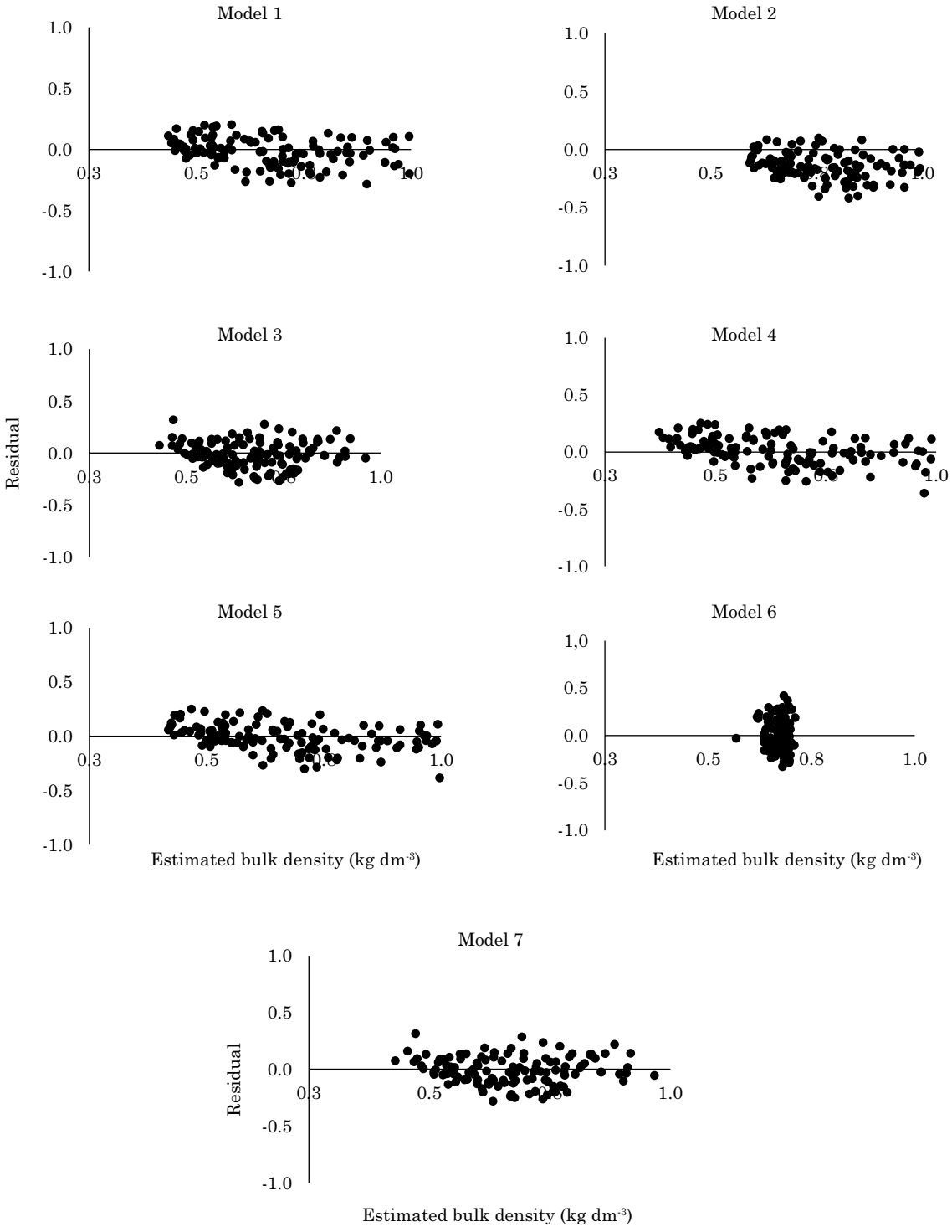

Figure 3. Regression residuals *versus* estimated values for soil density.

DISCUSSION

Assuming that there should be no co-linearity between independent variables that make up the equations in the regression models, the Models 1 and 2, despite having high R^2 values, are considered biased and unreliable for predicting soil density. Models that do not include both C and clay content are free of the bias generated by the co-linearity between these variables, as is the case for Models 3 to 7.

Models 4 and 5 had similar R^2 values (0.73 and 0.70, respectively). However, the results of the

Table 3. Test for normality of residuals (Kolmogorov-Smirnov at 95 % significance)

Model	Critical value at 95 % = 0.12	P value	Distribution of residual
1	0.06	0.43	normal
2	0.05	0.43	normal
3	0.05	0.38	normal
4	0.04	0.75	normal
5	0.06	0.45	normal
6	0.12	0.00	not normal
7	0.06	0.65	normal

validation of the models indicated better performance of Model 4, and much of the variation in the estimated data was explained by the independent variables (Figure 2). The predictive power of this model was similar to that of Model 1, with the advantage of having no bias of co-linearity according to the Pearson correlation test results.

It is worth highlighting that the higher the complexity of regression models, the more the principle of parsimony requires attention from the modeler. In the case of a series of models with adequate fits, the model with the fewest predictor variables is preferred. This is because models that include many explanatory variables

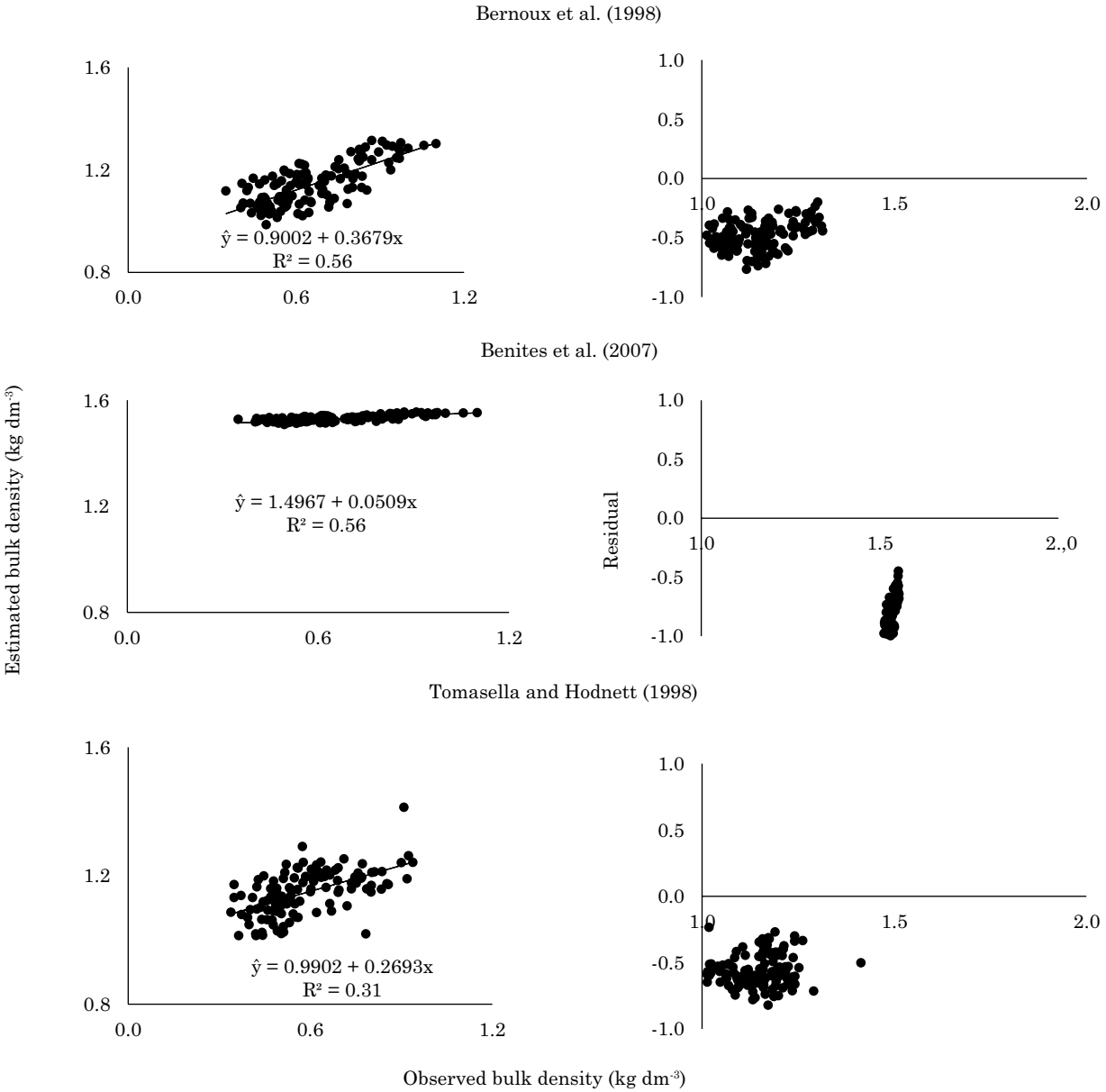

Figure 4. Comparison of our best model with models available in the literature.

Table 4. Descriptive statistics for observed densities and for densities estimated by the regression models

	Density				
	Observed in the field	Model 4 estimate	Bernoux et al. (1998)	Benites et al. (2007)	Tomasella and Hodnett (1998)
	kg/dm³				
Mean	0.66	0.65	1.14	1.53	1.15
Median	0.62	0.63	1.14	1.53	1.16
Minimum	0.35	0.37	0.99	1.51	1.01
Maximum	1.10	1.08	1.31	1.55	1.41

may seem to fit the data well, when, in fact, the fit is biased (Minasny and Hartemink, 2011). However, the difference between the values of the coefficients of determination for Models 4 and 5 should not be viewed in this light, since the difference in the number of explanatory variables is minimal between these models. Among the models generated, the performance of Model 4 was generally best for all assessment criteria. This model had the best coefficients of determination and AIC values, the residuals met the homoscedasticity assumptions and there was no co-linearity between the predictor variables.

Separately, the pH(H_2O) was not significantly correlated with the target variable of estimation of the study. This indicates that changes in density must be more closely related to C content and especially clay content in the sampled soils, although the equation with the best performance included both clay and pH.

The study of Bernoux et al. (1998) provided a series of equations to estimate soil density, and the model with the best performance in predicting this parameter (R^2 = 0.56) was constructed from the quantities of clay, sand, organic matter and pH(H_2O) based on 323 observations. The study showed that the estimated performance improved when the input data in the model were separated according to soil type and horizon. In this case, the equation was based only on clay content and C restricted to samples collected in the A horizon. For the surface of Latosols (Oxisols) the performance was high (R^2 = 0.78), although the number of observations was small (only 26 samples). This model is represented by the following equation:

$$\text{Density (D, kg/dm}^3\text{)} = 1.419 - 0.0037 \times \text{clay (\%)} - 0.061 \times \text{carbon (\%)}$$

In the validation step of the PTFs, the application of this equation to our data resulted in a R^2 similar to that of Model 4 in this study (Figure 4). However, the graph of the residuals *versus* estimated values showed that this model is not adequately fit and that it does not satisfy the premise of homoscedasticity.

Estimates of Benites et al. (2007) showed improved performance when the models included Fe_2O_3 content instead of the sum of bases (SB), since this parameter was not correlated with soil density. In addition, these authors found increase in the predictive power of the models when these were generated from data collected at different depths. The same pattern was observed by Heuscher et al. (2005) and Sequeira et al. (2014), who observed that the depth was not an appropriate variable to predict soil density. However, density appears to increase with the depth of the soil horizon, suggesting an influence of the pressure generated by the soil load (Tranter et al., 2007).

However, in the model of Benites et al. (2007), according to the authors the most accurate, the sum of bases (SB) was included as follows:

$$\text{Density (D, kg/dm}^3\text{)} = 1.5600 - 0.0005 \times \text{clay (g/kg)} - 0.100 \times \text{C (g/kg)} + 0.0075 \times \text{SB (cmol}_c\text{/kg)}$$

In this particular case it was not possible to use this FTP to estimate Bd, since the data for this study contained no information on the parameter SB.

Benites et al. (2007) generated seven models for different soil parameters. For comparison, Model 6, which included the parameters clay and C, was selected due to its performance in predicting Bd. The equation of Model 6, of Benites et al. (2007), is as follows:

$$\text{Density (D, kg/dm}^3\text{)} = 1.5688 - 0.0005 \times \text{clay (g/kg)} - 0.009 \times \text{carbon (g/kg)}$$

The small difference between the R^2 of estimates made by this equation using the original data series and this validation series can lead to the impression that the PTF is suitable for predicting soil density under the conditions of this study, but the model was not considered valid since the residuals were not homoscedastic.

The best-performing model, presented by Tomasella and Hodnett (1998), included silt content, according to the equation:

$$\text{Density (D, kg/dm}^3\text{)} = 1.578 - 0.054 \times \text{carbon (\%)} - 0.006 \times \text{silt (\%)} - 0.004 \times \text{clay (\%)}$$

To use this PTF for predicting soil density under the conditions of this study, the validation series was reduced, since information on the silt fraction was available for only 118 of the plots. The

low performance of this PTF in estimating density from the data series in this study may reflect the high correlation between clay and silt, as these parameters are expressed as percentages of the total particle-size composition and are added to the sand fraction, with a total of 100 % of the three fractions.

Although previous studies provided important information on the relationship between density and other soil properties in the Amazon, the spatial scales evaluated are too large to be reliably applied to local estimates (Bernoux et al., 1998; 2002; Tomasella and Hodnett 1998; Benites et al., 2007). The superiority of locally fit models was also shown in China, where the predictive power of 19 pedo-transfer functions was tested for a range of soil data; the results showed that equations generated from local observations had greater accuracy in the prediction of soil density (Han et al., 2012).

Thus, it is understood that locally generated equations will perform better in estimating soil density. This is because the logic of spatial data confirms the idea that, in geographic space, everything is related. Observations at sites that are close together are more related to each other than to geographically more distant observations (Fotheringham et al., 2002; Charlton and Fotheringham, 2009).

For having been calibrated based on field information, Model 4 produced a better estimate of soil density in this study than the PTFs available in the literature. The equation generated in this study was able to reliably estimate the density of surface soil (0-20 cm) in upland forests of the region of Manaus region. However, the uncertainty attributed to this estimate is still relevant, making it necessary to generate a larger number of equations calibrated locally in order to increase the accuracy of predicted values of density, and thereby improve estimates of soil carbon stocks throughout the Amazon landscape.

CONCLUSIONS

The pedo-transfer function (PTF) with the best performance in predicting soil density under the conditions of this study was Model 4, which included clay content and pH in water as independent variables [$\text{Density} = 1.495 - 0.011 \times \text{clay (\%)} - 0.079 \times \text{pH}(H_2O)$].

Despite the bulk density estimates using the PTFs available in the literature showing coefficients of determination similar to that of Model 4, a tendency of overestimation at different percentages was found for all tested equations from the literature, showing that locally generated equations have lower levels of uncertainty.

ACKNOWLEDGEMENTS

The authors thank the entire staff of the Biological Dynamics of Forest Fragments Project (BDFFP) for logistical support in collecting soil samples in the field, the staff of the INPA Soil and Plant Laboratory (LTPS-INPA) and the Program of Graduate Studies in Tropical Forest Science (PPG-CFT/INPA). This is publication number 662 in the BDFFP technical series. The funding for this work was provided by CAPES, CNPq (Proc. 304020/2010-9, 573810/2008-7) and FAPEAM (Proc. 708 565, 363/2013).

REFERENCES

Akaike H. Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. In: Proceeedings of the 2nd International Symposium on Information Theory; 1973; Budapest. Budapest: Akademiai Kiado, 1973. p.267-81.

Benites VM, Machado P, Fidalgo ECC, Coelho RM, Madari EB. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. Geoderma. 2007;139:90-7

Bernoux M, Arrouays D, Cerri C, Volkoff B, Jolivet C. Bulk densities of Brazilian Amazon soils related to other soil properties. Soil Sci Soc Am J. 1998;62:743-9.

Bernoux M, Carvalho MCS, Volkoff B, Cerri CC. Brazil's soil carbon stocks. Soil Sci Soc Am J. 2002;66:888-96.

Bozdongan H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52:345-70.

Burnham KP, Anderson DR. Model selection and multimodel inference: A practical information-theoretic approach. 2ª ed. New York: Springer; 2002.

Burnham KP, Anderson DR, Huyvaert PK. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behav Ecol Sociobiol. 2011;65:23-35.

Calhoun FG, Smeck NE, Slater BL, Bigham JM, Hall GF. Predicting bulk density of Ohio soils from morphology, genetic principles, and laboratory characterization data. Soil Sci Soc Am J. 2001;65:811-9.

Chaudhari PR, Ahire DV, Ahire VD, Chkravarty M, Maity S. Soil bulk density as related to soil texture, organic matter content and available total nutrients of Coimbatore Soil. Int J Sci Res Pub. 2013;3:1-8.

Chauvel A. Os Latossolos Amarelos, álicos, argilosos dentro dos ecossistemas das bacias experimentais do INPA e da região vizinha. Acta Amaz. 1982;12:38-47.

Chauvel A, Lucas Y, Boule R. On the genesis of the soil mantle of the region of Manaus, Central Amazonia, Brazil. Experientia. 1987;43:234-41.

Charlton M, Fotheringham AS. Geographically weighted regression white paper [internet]; 2009 [accessed Oct 10 2013]. Available: http://gwr.nuim.ie/downloads/GWR_WhitePaper.pdf.

Empresa Brasileira de Pesquisa Agropecuária - Embrapa. Sistema brasileiro de classificação de solos. 3ª ed. Rio de Janeiro; 2013.

Empresa Brasileira de Pesquisa Agropecuária - Embrapa. Manual de métodos de análise de solo. Rio de Janeiro: Serviço Nacional de Levantamento e Conservação de Solos; 1979.

Fearnside PM, Barbosa RI. Soil carbon changes from conversion of forest to pasture in Brazilian Amazonia. For Ecol Manage. 1998;108:14766.

Fearnside PM, Leal Filho N. Soil and development in Amazonia: Lessons from the Biological Dynamics of Forest Fragments Project. In: Bierregaard RO, Gascon C, Lovejoy TE, Mesquita R, editors. Lessons from Amazonia: The ecology and conservation of a fragmented forest. New Haven: Yale University Press; 2001. p.291-312.

Federer CA, Turcotte DE, Smith CT. The organic fraction-bulk density relationship and the expression of nutrient content in forest soils. Can J For Res. 1993;23:1026-32.

Fotheringham AS, Brunsdon C, Charlton ME. Geographically weighted regression: The analysis of spatially varying relationships. Chichester: Wiley; 2002.

Gharahi-Gheni N, Nemes A, Verdoodt A, van Ranst E, Cornelis WM, Boeckx P. Nonparametric techniques for predicting soil bulk density of tropical rainforest topsoils in Rwanda. Soil Sci Soc Am J. 2012;76:1172-83.

Han GZ, Zang GL, Gong ZT, Wang GF. Pedotransfer functions for estimating soil bulk density in China. Soil Sci. 2012;177:158-64.

Heuscher SA, Craig C, Brandt CC, Jardine PM. Using soil physical and chemical properties to estimate bulk density. Soil Sci Soc Am J. 2005;69:1-7.

Minasny B, Hartemink AE. Predicting soil properties in the tropics. Earth-Sci Rev. 2011;106:52-62.

Oliveira LAA, Mori SA. A central Amazonian terra-firme forest. I. High tree species richness on poor soils. Biodivers Conserv. 1999;8:1219-44.

Pérez DV, Alcântara S, Arruda RJ, Meneghelli NAC. Comparing two methods for soil carbon and nitrogen determination using selected Brazilian soils. Commun Soil Sci Plant Anal. 2001;32:295-309.

Sequeira CH, Willis SA, Seybold CA, West LT. Predicting soil bulk density for incomplete databases. Geoderma. 2014;213:64-73.

Steller RM, Jelinsk NA, Kucharik CJ. Developing models to predict soil bulk density in southern Wisconsin using soil chemical properties. J Int Biosci. 2008;6:53-63.

Symonds M, Moussalli A. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol. 2011;65:13-21.

Taalab KP, Corstanje R, Creamer R, Whelan MJ. Modeling soil bulk density at the landscape scale and its contributions to C stock uncertainty. Biogeosci Disc. 2012;9:18831-64.

Tomasella J, Hodnett MG. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. Soil Sci. 1998;163:190-202.

Tranter G, Minasny B, McBratney AB, Murphy B, McKenzie NJ, Grundy, M, Brough D. Building and testing conceptual and empirical models for predicting soil bulk density. Soil Use Manage. 2007;23:437-43.

Veldkamp E. Organic carbon turnover in three tropical soils under pasture after deforestation. Soil Sci Soc Am J. 1994;58:175-80.

Wilkinson L. Systat: The system for statistics. Evanston: Systat; 1990.