Vegetation and Topographic Control on Spatial Variability of Soil Organic Carbon

Vegetation and Topographic Control on the Spatial Variability of Soil Organic Carbon*1

I. OUESLATI[1], P. ALLAMANO[1,*2], E. BONIFACIO[2] and P. CLAPS[1]

[1] *DITIC, Politecnico di Torino, 10129 Torino (Italy)*
[2] *DIVAPRA, Università di Torino, 10095 Grugliasco (Italy)*

(Received , 2011; revised , 2011)

ABSTRACT

SOC (Soil Organic Carbon) is one of the most important parameters affecting the hydraulic characteristics of natural soils. Despite being rather easy to measure, SOC is known to be highly variable in space. In this work we used vegetation, climate, and morphology factors to reproduce the spatial distribution of soil organic carbon (SOC) in the mineral horizons of forest and grassland areas (in North-Western Italy) and evaluated the feasibility of the approach. When the overall sample was analysed (114 samples), annual precipitation and elevation were significant descriptors of the SOC variability. However, a large part of the variability remains unexplained. Two stratification criteria were then adopted, based on vegetation and topographic properties. We obtained an improvement of the quality of the estimates, particularly for grasslands and forests in the absence of local curvatures. These results indicate that the spatial variability of soil organic matter is scarcely reproducible at the regional scale, unless an a-priori reduction of the heterogeneity is applied. A discussion on the feasibility of applying stratification criteria to deal with heterogeneous samples closes the paper.

Key Words: Natural soils, regression analysis, soil organic carbon, spatial variability

Citation Oueslati, I., Allamano, P., Bonifacio, E. and Claps, P. 201???. Vegetation and Topographic Control on the Spatial Variability of Soil Organic Carbon. Pedosphere. 2???(???): ???--???.

---

[1] Supported by the Italian Ministry of Education (grant 2007HBTS85)
[*2] Corresponding Author: E-mail: paola.allamano@polito.it

INTRODUCTION

Soils are among the most important global carbon cycle reservoirs. They represent 75% of the total terrestrial organic carbon pool, and twice the atmospheric carbon content (Prentice, 2001). Most soil carbon is stored as soil organic matter; inorganic carbonates are largely restricted to dry lands (Lal, 2008). Forest and grassland soils are therefore of utmost importance as their organic matter content is much higher than that of agricultural soils.

In addition to its well-known effect on soil fertility, soil organic carbon (SOC) is also one of the most important parameters affecting soil hydraulic properties: by enhancing soil aggregation and increasing pores continuity and stability (Tisdall and Oades, 1982), it ameliorates the water holding capacity and the saturated hydraulic conductivity (Nemes *et al*., 2005). As a consequence, it is used as an independent variable in several pedotransfer functions to estimate soil hydraulic properties (e.g. Guber *et al*., 2006).

Despite being rather easy to measure, SOC is highly variable in space. Evaluating its concentration in areas greater than an agricultural plot can be therefore rather difficult. A large effort has been made to predict its areal distribution through Digital Soil Mapping (DSM) approaches (e.g., Arrouays *et al*., 1995, 1998; Bell *et al*., 2000; Chaplot *et al*., 2001; Abnee *et al*., 2004a,b; Lagacherie and McBratney, 2004; Terra *et al*., 2004; Thompson and Kolka, 2005). Typically, the approach is implicitly based on Jenny (1941)'s factors of soil formation: it produces empirical quantitative relations between soil characteristics and spatially referenced factors, with the aim of using these relations as soil spatial prediction functions (McBratney *et al*., 2003).

The most relevant pedogenic factors controlling SOC are climate, vegetation, and topography (Baldock and Nelson, 2000). Climate impacts directly on SOC contents and rate of mineralization primarily through the effects of temperature and moisture. Temperature is a key factor controlling the rate of decomposition of plant residues. For example, reaction rates double for each increase of 8–9 °C in the mean annual air temperature (Bot and Benites, 2005). As a consequence, soils in cooler climates commonly contain more organic matter because of slower mineralization rates. In addition, SOC content increases as mean annual precipitation increases (e.g. Meier and Leutschner, 2010).

Changes in SOC concentration also correlate with changes in vegetation. For instance, herbaceous vegetation produces a big root mass compared with the above-ground component whereas, in forests, most organic matter is produced above ground and residues are more resistant to decomposition (Melillo *et al*., 1993).

The latter pedogenic factor controlling SOC is topography. Topography affects soil properties mainly through its effects on water movements. In terrain depressions, in fact, soils are moister because they receive runoff, sediments (including organic matter), and seepage from the surroundings, leading to a higher SOC concentration than in drier uphill soils (Yoo *et al*., 2006). On the contrary, soils on steep slopes tend to lose organic matter because the topsoil is constantly eroded. Therefore, the spatial distribution of topographic attributes that characterize the flow paths can help in capturing the soil variability and in predicting soil properties (Moore *et al*., 1991 and 1993; Florinsky and Kuryakova, 1996).

Landscape characteristics can be expressed by means of geomorphic and hydrological indices generated with digital terrain analysis (Speight, 1968; MacMillan and Pettapiece, 1997; Wilson and Gallant, 2000). These indices are widely used in hydrologic modelling, e.g. to characterize the spatial distribution and the extent of zones of saturation for runoff generation (e.g. Beven and Kirkby, 1979). Climatic data are often available through gridded datasets, and vegetation characteristics can be derived qualitatively from land cover maps such as CORINE (European Environment Agency, 2005), or quantitatively by making use of remote sensing. Quantitative environmental variables can be then associated to SOC, resulting in long sets of potential descriptors of SOC variability.

In this work we evaluated the spatial variability of SOC and the feasibility of the prediction of organic carbon concentration by means of quantitative descriptors of vegetation, climate and topography, in forest and grassland soils of alpine and sub-alpine areas of North-western Italy. No specific survey was performed for this study and the data were all derived (and homogenised) from existing databases.

STUDY AREA

The Piedmont region is located in North-Western Italy (see Fig. 1). Its area, of about 25 000 km^2, is characterized by marked heterogeneities determined by the presence of the Western Alps chain.

Fig. 1

Fig. 1 Fig. 1 Shown is a map of the study region with the locations of the sampled sites. The gray tones represent the elevation variations within the region.

In few hundreds kilometres the climate changes from a typical Mediterranean regime to an Alpine-continental regime. The average monthly temperatures range between a minimum of -10 °C in the mountain areas during the cold season and a maximum of +35 °C in the Po Plain during the summer. The mean annual precipitation ranges between 500 mm and more than 2 000 mm.

Forests cover the mountain areas of the region up to an elevation of about 1800 m. Coniferous and broadleaved species, either as monocultures or in admixture, are present in the whole area. Grasslands are found either above the timberline, or scattered at lower elevation and on hilly areas.

The data of soil profiles were taken from a pre-existing database belonging to the DIVAPRA of the University of Torino (67 samples) and to the Regional Soil Survey Service (IPLA, 55 samples). The chemical characteristics which we considered were the soil organic carbon (SOC) and the total nitrogen concentrations in the mineral horizons. The data were available for all genetic horizons (A, B, or C types) and also aggregated on the 0--30 cm depth (SOC_{30}). The chemical characteristics of the dataset are reported in Table I.

TABLE I

Soil organic carbon (SOC), Nitrogen concentration and C/N ratio in soil horizons according to vegetation types. Different (small) letters indicate significant differences in the chemical property among vegetation types ($P < 0.05$)

Horizon or layer		n	SOC (g kg^{-1}) Mean	Std. Dev.	N (g kg^{-1}) Mean	Std. Dev.	C/N Mean	Std. Dev.
A	Broadleaves	72	33.5	23.7	2.2	1.5	15.4^b	4.8
	Mixed Forests	24	34.3	30.2	2.4	1.4	13.4^b	4.4

	Conifers	3	27.3	15.7	1.5	1.0	21.7^{a}	17.1
	Grasslands	21	38.4	24.4	2.9	1.4	12.1^{b}	5.4
B	Broadleaves	38	11.9	11.6	0.9	0.7	12.6	3.6
	Mixed Forests	15	11.6	10.4	1.2	0.7	13.9	7.4
	Conifers	0	.	.	.	.	.	.
	Grasslands	10	11.5	5.7	1.4	0.2	11.3	.
C	Broadleaves	39	9.1	11.6	0.8	0.8	13.3	6.3
	Mixed Forests	8	3.6	1.7	0.6	0.3	6.6	3.0
	Conifers	1	14.4	.	1.2	.	12.2	.
	Grasslands	7	14.4	8.2	1.6	0.7	8.8	2.1
0--30 cm	Broadleaves	69	26.4	18.9	2.2	3.3	15.3^{ab}	6.2
	Mixed Forests	22	23.8	12.1	1.8	0.6	13.1^{ab}	3.6
	Conifers	3	25.2	12.8	1.5	0.6	17.7^{a}	9.8
	Grasslands	20	32.0	15.8	2.7	1.1	10.1^{b}	2.9

No significant differences were found in SOC and Nitrogen concentrations between forests and grasslands, while the C to N ratio clearly reflected the differences in organic matter decomposition, (with the lowest values in grasslands and the highest in conifers, both in the A horizon and in the 0--30 cm layer). The concentrations of SOC and N were generally more variable in the genetic soil horizons than in the 0--30 cm layer, reflecting the different thicknesses of the SOC enriched A horizons. The same behaviour was observed for the C to N ratio, with lower standard deviations in the 0--30 cm layer, particularly in grassland samples.

The sample variogram (see e.g., Burgess and Webster, 1980) was fitted by an exponential model with nugget, where the sample and spatial variance were nearly the same (namely 0.196 and 0.195, respectively) with a 10 km range (Fig. 2) The relative weights of the two components indicated that the organic carbon concentration was controlled by two kinds of variables, having very short range and long range variability, respectively.

Fig. 2

Fig. 2 Semivariogram of the natural logarithm of the soil organic carbon concentration for the 0--30 cm layer.

As an attempt to capture the relative roles of the long and short range mechanisms controlling SOC variability at the regional scale we described each sampled point according to its climatic, vegetation and topographic characteristics (also called “morphoclimatic” factors in the following of the paper). The data guaranteed a spatial coverage of the entire region, in view of possible mapping applications.

In details:

- To obtain a characterization of the soil water balance in the profiles locations, the average annual rainfall (P, mm) produced by Bartolini *et al*. (2008) for the study region was used. In addition, we calculated the potential evapotranspiration (PET) with the Thornthwaite relation according to the monthly temperature estimates proposed by Claps *et al*. (2008). Precipitation and PET values were also combined by means of the Thornthwaite (1948) Wetness Index (WI) as WI = (P - PET)/PET. The WI allows a synthetic characterization of the soil water balance in a certain location and provides a relative measure of potential biological productivity of the site.
- NDVI maps, derived from SPOT-VEGETATION data (http://free.vgt.vito.be/) for the years 1998-2005, were used to characterize vegetation. For each month the composite NDVI values were

averaged over 1 km^2 cells. For each cell containing at least one soil sampling we considered the minimum NDVI (identified with the subscript “Min”), the maximum NDVI (identified with the subscript “Max”), the difference between maximum and minimum NDVI (identified with the subscript “Diff”) and the mean NDVI (identified with the subscript “Mean”). The CORINE land cover map (European Environment Agency, 2005) and the forests map of the Piedmont Regional Administration were also taken into account for further evaluations of vegetation types in correspondence of the sampling points.

- The selection of the terrain descriptors was based on the prior experience by a) Chai *et al*. (2008), who considered elevation, slope and CTI (see Table II) as external drift variables for predicting soil organic matter; b) Mueller and Pierce (2003) and Simbahan *et al*. (2006), who used terrain attributes as secondary spatial variables to increase the accuracy of predictions at different scales. To determine the primary and secondary terrain attributes (Table II), we used a digital elevation model (derived from contour line interpolation) with a spatial resolution of 50 m. The primary attributes in Table II include elevation (EL), slope (SL), profile (kv) and plan curvature (kh). In kv and kh, the sign ensures that positive curvatures refer to convex forms and negative curvatures refer to concave forms. The expected values for a hilly area (moderate relief) can vary from -0.5 to 0.5, while for steep and rugged mountains (extreme relief), they can vary between -4 and 4. A value of zero indicates that the surface is flat. As secondary terrain attributes in Table II, we considered the compound topographic index (CTI), also known as topographic wetness index (Wilson and Gallant, 2000), and the stream power index (SPI). The CTI, expressed as reported in Table II, is related to zones of surface saturation (Moore *et al*., 1993) and quantifies a point in the landscape in terms of water and sediment accumulation: the higher the value of this index in a cell, the higher the soil moisture that can be found in it. SPI instead, expressed as reported in Table II, is related to erosion processes, being an indicator of the capability of a flow to generate net erosion.

TABLE II

Definitions, formulas and physical interpretations for the topographic variables considered in the study

Variable	Definitions and formula	Interpretation
Slope gradient (SL), °	Angle between a tangent plane and a horizontal one at a given point of the land surface	Velocity of substance flows
Profile curvature (kv), m^{-1}	Profile curvature is parallel to the direction of the maximum slope. A negative value indicates that the surface is upwardly convex at that cell. A positive profile curvature indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is linear.	Relative acceleration/deceleration of flows across the surface
Plan curvature (kh), m^{-1}	Planform curvature (commonly called plan curvature) is perpendicular to the direction of the maximum slope. A positive value indicates the surface is sidewardly convex at that cell. A negative value indicates the surface is sidewardly concave at that cell. A value of zero indicates the surface is linear.	Convergence/divergence of flows across a surface
Compound Topographic Index (CTI)	$CTI = \ln \frac{C_A}{SL}$ a)	Extent of potential flow accumulation
Stream power Index (SPI)	$SPI = C_A \cdot SL$ a)	Extent of potential flow erosion

a) C_A: specific catchment area, *i.e.* upslope area per unit width of contour (Wilson and Gallant, 2000), m^2 m^{-1}.

METHODS

Linear multiple regression is the most widely used method to capture soil carbon patterns, as demonstrated by numerous applications (e.g., Walker *et al*., 1968; Pennock *et al*., 1987; Moore *et al*., 1993; Gessler *et al*., 1995, 2000; Chaplot *et al*., 2000). This preference stems from its ease of use, computational efficiency, and straightforward interpretation (Hastie *et al*., 2001).

In order to select the possible relationships between the soil carbon concentration and the environmental factors, we implemented simple correlation and multiple regression schemes. The simple correlations between SOC and environmental variables were evaluated on all genetic horizons and on the fixed 0--30 depth. The multiple regressions were calculated only on the values related to 30 cm depth, to avoid considering horizons of variable thickness.

The models were tested against the assumptions of the linear regression analysis (see e.g., Kottegoda and Rosso, 1998), namely lack of multicollinearity, equal error variance (homoscedasticity), and normal and random residuals. In some cases we used the logarithmic transformation of the variable for complying to these assumptions, particularly as regards normality. The goodness of the relationships was verified using the standard error of the estimate (SEE) and the adjusted determination coefficient (Adj R^2). The variogram analysis was also applied to the residuals of the regressions in order to assess their residual spatial variability.

APPLICATION AND DISCUSSION

Simple correlations between morphoclimatic factors and SOC concentration were tested at first. The results are shown in Table III.

Among the topographic variables (Table III, lines 1 to 6), only the elevation was significantly correlated to SOC in all horizons, in agreement with the results obtained in other studies. For example, Powers and Schlesinger (2002) found that elevation explained much of the variability in soil carbon, likely because of the lower decomposition rate of organic matter at lower temperatures. This hypothesis was in agreement with the C to N ratio which, in our case, tended to increase with elevation, particularly in the top horizon (r = 0.24, P < 0.05, not shown).

TABLE III

Pearson correlation coefficients (r) between SOC and morphoclimatic variables associated to different horizons (identified by subscript capital letters)

	$\ln(SOC_A)$	$\ln(SOC_B)$	$\ln(SOC_C)$	$\ln(SOC_{30})$
1. number of samples	120	63	55	114
2. EL (elevation)	$0.369^{a)}$	$0.408^{a)}$	$0.399^{a)}$	$0.402^{a)}$
3. SL (slope)	0.148	$0.265^{b)}$	0.212	0.114
4. kh (profile curvature)	-0.075	-0.157	0.009	-0.001
5. kv (plan curvature)	0.073	0.014	-0.095	0.092
6. CTI (Compound Topographic Index)	-0.111	-0.171	0.009	-0.141
7. ln SPI (Stream Power Index)	0.012	0.152	0.194	-0.065
8. $NDVI_{Max}$ (maximum NDVI)	0.027	0.216	0.017	-0.061
9. $NDVI_{Min}$ (maximum NDVI)	$-0.292^{a)}$	$-0.278^{b)}$	$-0.340^{b)}$	$-0.355^{a)}$
10. $NDVI_{Mean}$ (maximum NDVI)	$-0.220^{b)}$	-0.088	$-0.266^{b)}$	$-0.281^{a)}$
11. $NDVI_{Diff}$ (maximum NDVI)	$0.328^{a)}$	$0.448^{a)}$	$0.322^{b)}$	$0.329^{a)}$
12. P (average annual rainfall)	$0.477^{a)}$	$0.547^{a)}$	-0.015	$0.464^{a)}$

13. WI (Thornthwaite wetness index)	0.556[a)]	0.599[a)]	0.197	0.553[a)]

[a)]Correlation is significant at the 0.01 level (2-tailed).
[b)]Correlation is significant at the 0.05 level (2-tailed).

All NDVI parameters, except the maximum, were significantly correlated with the soil carbon variables (Table III, lines 7 to 10). The minimum and the mean NDVI decreased with increasing C content. A possible explanation for this behaviour is the high ratio between roots and shoots in grasslands (Jackson *et al*., 1996), which had a high SOC content compared with forests (Table I). When grasslands were excluded from the analysis, no more correlation was found between C, $NDVI_{Mean}$ and $NDVI_{Min}$, while a significant correlation appeared between $NDVI_{Diff}$ and SOC concentration in all horizons (r up to 0.54, $P < 0.05$, not shown).

The soil carbon variables were also positively correlated with the precipitation and the wetness index (Table III, lines 11 and 12) in all but the C horizons. This confirms the fact that the amounts of SOC increase when precipitation and soil moisture increase (*e.g.*, McKenzie and Ryan, 1999). The C horizons, instead, presented no correlations with the environmental variables. Probably in these deeper layers, the soil water balance dynamics cannot be fully represented by climatic variables, as they do not provide an estimate of percolation.

Several environmental descriptors were therefore related to SOC concentration, even though their correlation coefficients were always low. In the following, the descriptors were combined using multiple regressions only on the carbon concentration of the 0--30 cm layer (SOC_{30}).

Four linear regression models were significant, demonstrating to be effective in reproducing SOC within a stepwise procedure (Table IV). The adjusted determination coefficient (Adj R^2), the standard error of the estimate (SEE), and the standardized coefficients (β) for each model are reported, allowing for a thorough evaluation of the model efficiency.

TABLE IV

Significant multiple regression models for SOC concentration in the 0--30 cm layer based on the whole dataset

Dependent variable	Independent variables	Adj R^2	SSE	β[a)]	P
$\ln(SOC_{30})$	EL, P	0.318	0.530	0.34, 0.41	< 0.001
	$NDVI_{Min}$, P	0.310	0.533	-0.33, 0.44	< 0.001
	$NDVI_{Mean}$, P	0.301	0.536	-0.31, 0.49	< 0.001
	$NDVI_{Diff}$, P	0.250	0.555	0.23, 0.41	< 0.01

[a)] Standardised coefficients

The best regression model was based on elevation (EL) and precipitation (P), but the percentage of SOC variance explained is low (R^2 = 0.32). As visible in Fig. 3a, the model lead to rather poor predictions, with remarkable underestimates when the actual SOC concentration was above 50--60 g kg^{-1} and visible overestimates in correspondence of low values of observed SOC. NDVI parameters contributed to SOC variability in the three other regression models, where the standardized coefficients for NDVI parameters assumed also negative values. This outcome could be related to the large heterogeneity of broadleaved forests which, in the study area, included both relatively highly productive beech stands, with poor understory vegetation, and degraded woodlands, where pioneer tree species, such as birch, shrubs and grasses, were rather abundant and could influence the NDVI parameters. Both the amount of carbon which is added annually to the soil and the recalcitrance to decomposition of organic residues are in fact tree-species dependent (Augusto *et al*., 2002).

Fig. 3

Fig. 3 Comparison between the observed and predicted values of SOC_{30} for: (a) the whole dataset (first model in Table IV); (b) the grasslands subset (best model for grasslands in Table V); (c) the LL subset (best model for the LL class in Table VI); (d) the forest-LL subset (best model for the forest-LL class in Table VII)

Fig. 4a shows the sample variogram of the residuals of the best regression model of Table IV (*i.e.*, EL, P). The variance related to the nugget was 0.197, the one for the exponential model was 0.089, the range was 6 km. The regression model therefore reduced the spatial variability but had no effect on the very short range variability represented by the nugget. This was confirmed by analysing separately the variograms of the two environmental factors (data not shown). We found that these variograms had rather large ranges (*i.e.*, 35 and 88 km, respectively, as calculated by fitting an exponential model to the observed variograms). This is due to the fact that these variables are strongly related to large scale phenomena, such as general topography and climate.

Fig. 4

Fig. 4 Semivariogram of the residuals of the best regression model for (a) the whole dataset (first model in Table IV); (b) the grasslands subset (best model for grasslands in Table V); (c) the LL subset (best model for the LL class in Table VI); (d) the forest-LL subset (best model for the forest-LL class in Table VII).

Although climate and vegetation appeared to be the main drivers of SOC variability, the relationships were still unsuitable to be used for estimation. This could be explained by the heterogeneity of vegetation and topographic characteristics of the samples. In fact, four land cover classes were present, namely broadleaved trees, coniferous, mixed forests, and grasslands, each of them giving rise to different SOC production, accumulation, and decomposition rates, which were only partially accounted for by NDVI parameters. Also the terrain morphology was highly variable: besides the elevation range of the samples, there was a large variety of topographic features, such as hilltops, slopes and plains. Evidences of the heterogeneity effects were found by Arrouays *et al.* (1995), Homan *et al.* (1995), Arrouays *et al.* (1998), Powers and Schlesinger (2002) and Perruchoud *et al.* (2000). From these studies it emerged that, at the regional and national scales, up to 90% of the SOC variance could be explained when major vegetation types and environmental gradients were taken into account. In addition, site-specific topographic information can also yield correlation to C content within regions of similar climate and land cover.

To reduce sample heterogeneity, especially the short range variability, the dataset was stratified. Samples were grouped according to vegetation types (see Table V, where conifers are not considered because of the low number of cases).

TABLE V

Significant multiple regression models for SOC concentration in the 0--30 cm layer after stratification based on vegetation type (the number of elements for each class is reported in parenthesis).

Vegetation type	Independent variables	Adj R^2	SSE	$\beta^{a)}$	p
Broadleaves (69)	a) $NDVI_{Max}$, EL, P	0.359	0.557	-0.32, 0.32, 0.60	<0.001
	b) $NDVI_{Mean}$, P	0.345	0.563	-0.26, 0.57	<0.001
	c) $NDVI_{Max}$, WI	0.344	0.563	-0.28, 0.71	<0.001
Mixed Forests (22)	d) $NDVI_{Min}$, EL, kv	0.309	0.446	0.72, 0.92, 0.55	<0.05
Grasslands (20)	e) $NDVI_{Mean}$, EL, CTI, ln(SPI)	0.695	0.293	1.11, 2.07, 0.86, -0.88	<0.001
	f) SL, CTI, ln(SPI)	0.594	0.338	1.15, 1.71, -1.33, 0.77	<0.001
	g) EL, CTI, ln(SPI)	0.564	0.350	0.94, 0.66, -0.65	<0.001

[a)] Standardised coefficients

The proportion of variance explained by the regression models increased up to $R^2 = 0.70$ in the case of grasslands. This result was obtained by coupling climate and vegetation with topographic variables, suggesting that in more homogeneous contexts also the topographic variables (CTI and SPI) play a role in controlling SOC amount (e.g., Homan *et al*., 1995). The standardized coefficients indicated a positive relation with $NDVI_{Mean}$, elevation and CTI, and a negative effect of SPI. The presence of the two topographic indexes suggested that, when vegetation was more homogeneous, the concentration of carbon in soils was largely determined by erosion and sedimentation. As visible in Fig. 3b, the model lead to an acceptable estimate of the concentrations of SOC in grasslands. The sample variogram of Fig. 4b was constructed on the residuals of the best regression model obtained for grasslands. This variogram had a nugget variance of 0.005 and the exponential model explained a variance of 0.074 in a range of 3 km, evidencing the positive effects of stratification on both the long and short range variability.

Scarce improvements were instead observed in the prediction of SOC in forest soils. In these cases, in fact, the SOC variance accounted for by the models remained rather low (around 30%) and far from the results obtained in the case of grasslands. The NDVI parameters still showed negative coefficients, as happened when the whole dataset was considered. In addition, no evidence of the role of topographic descriptors emerged, suggesting that a further classification based on local morphology could possibly enhance the model performance.

A quantitative subdivision of the territory into classes of curvature according to the combination of plan and profile curvature was used, based on a curvature classification model (Olaya, 2004). The original classes were regrouped into four classes according to the profile curvature. We classified as CV the points having a convex profile curvature (33 samples), as CC the points with concave profile curvature (40 samples), as LC the points having linear profile curvature (23 samples), and as LL the elements where both plan and profile curvatures were linear (15 samples). The aim was twofold: on the one hand to avoid classes with few elements, and on the other hand to introduce a hydrological criterion in the classification, by considering dispersion and accumulation characteristics. The approach is akin to the one adopted in Kyungsoo *et al*. (2006), who found that topographic curvature and soil thickness played a key role on SOC contents. A similar evidence was also pointed out by Johnson *et al*. (2009) who reported the largest amounts of soil organic matter in well drained soils on gently grading slopes.

The models obtained from the application of the stepwise multiple regression are presented in Table VI.

TABLE VI

Significant multiple regression models for SOC concentration in the 0--30 cm layer after stratification based on profile curvature (the number of elements for each class is reported in paranthesis).

Profile Curvature	Independent variables	Adj R^2	SSE	$\beta^{a)}$	p
CV (33)	a) EL, SL, CTI, ln(SPI), P	0.514	0.439	0.41, -1.07, -0.95, 1.38, 0.47	<0.001
	b) SL, CTI, ln(SPI), WI	0.482	0.453	-1.02, -1.02, 1.39, 0.61	<0.001
	c) $NDVI_{Mean}$, P	0.461	0.462	-0.43, 0.60	<0.001
	d) $NDVI_{Min}$, P	0.457	0.464	-0.42, 0.56	<0.001
	e) EL, P	0.429	0.475	0.39, 0.52	<0.001
CC (40)	f) EL, P	0.328	0.555	0.44, 0.35	<0.001

	g) $NDVI_{Min}$, P	0.313	0.561	-0.41, 0.38	<0.001
LC (23)	h) $NDVI_{Max}$, $NDVI_{Diff}$, WI	0.418	0.494	0.57, -0.55, 1.16	<0.01
	i) $NDVI_{Min}$, WI	0.443	0.483	0.74, 1.14	<0.001
LL (15)	l) $NDVI_{Diff}$, kh, kv, ln(SPI)	0.659	0.350	0.82, -0.37, -0.80, -0.82	<0.01
	m) $NDVI_{Diff}$, kv, ln(SPI)	0.523	0.414	0.74, -0.83, -0.79	<0.05
	n) $NDVI_{Min}$, kv, ln(SPI)	0.425	0.454	-0.59, -0.66, -0.72	<0.05

[a] Standardised coefficients

An improvement of the prediction with respect to the whole dataset was observed, but the accuracy was comparable with the one obtained by grouping according to vegetation types. The models for convex (CV) and concave (CC) classes explained 51% and 33% of the within-class variance, respectively. In the best model for the CV class, the highest standardised coefficients were related to topographic variables, indicating that carbon accumulation decreased with increasing slope, but the effect was partially compensated by the extension of the runoff contributing area (CA, see footnote in Table II). In the CC class, instead, the significant variables were the same as those found in the whole dataset (elevation, NDVI and precipitation). Climatic and vegetation variables were also significant in the models of the LC (linear profile curvature) class, with an explained variance up to 42%. A better regression model was obtained in the purely linear (LL) case ($R^2 = 0.66$). When both the profile and plan curvature were linear (*i.e.*, less than 0.1 in module) even small variations in concavity/convexity affected the carbon concentration, with higher amounts of SOC in relative concavities. The quality of the estimate (Fig. 3c) was comparable with that obtained for the grasslands when the dataset was stratified by vegetation type, suggesting that a further improvement of the quality of the estimation might be obtained by combining topography and vegetation classes.

To proceed in this direction and to have enough samples in each class, we merged mixed and broadleaved forest samples (Table VII). This further classification significantly improved the estimate and explained up to the 81% of the sample variance in the LL class. In all subsets the models included topographic attributes, NDVI parameters and climatic variables. In general, SOC increased when elevation and wetness (P and WI) increased. The heterogeneity of broadleaved forests, enhanced by the presence of mixed vegetation ones, was probably responsible for the lack of a general trend in the sign of the NDVI variables in the models. The effect of concavity and convexity on carbon accumulation was further enhanced by local slope whose steepness increased carbon concentrations when the curvatures were concave and decreased it in case of convexity. A similar result was found by Grimm *et al*. (2008), who concluded that the lowest SOC stocks can be found on mid-slopes while the highest were located on toe-slope positions. Also with this stratification the plan and profile curvature contributed to the carbon variance only in the LL class. The lowest determination coefficient was obtained for the LC class, the most heterogeneous one as the plan curvature could induce either flow dispersion or accumulation.

TABLE VII

Significant multiple regression analyses for the soil carbon concentration at 30 cm depth according to vegetation-curvature classification (the number of elements for each class is reported in parenthesis).

Veg.-Curv.	Independent variables	Adj R^2	SSE	$\beta^{¥}$	p
Forest - CV (29)	a) SL, CTI, ln(SPI), WI	0.501	0.452	-1.28, -1.16, 1.70, 0.59	<0.001
	b) SL, CTI, ln(SPI), P	0.475	0.464	-1.20, -1.11, 1.69, 0.56	<0.001
	c) $NDVI_{Mean}$, P	0.481	0.461	-0.37, 0.66	<0.001
	d) $NDVI_{Mean}$, WI	0.442	0.478	-0.31, 0.62	<0.001
	e) $NDVI_{Min}$, P	0.439	0.480	-0.31, 0.61	<0.001

Forest - CC (34)	f) $NDVI_{Min}$, SL, CTI, ln(SPI), P	0.337	0.561	-0.46, 1.09, 1.05, -1.31, 0.41	<0.01
	g) EL, P	0.275	0.586	0.37, 0.38	<0.01
	h) $NDVI_{Min}$, P	0.275	0.586	-0.36, 0.41	<0.01
Forest - LC (16)	i) $NDVI_{Min}$, WI	0.317	0.572	0.60, 0.66	<0.05
Forest - LL (13)	l) $NDVI_{Max}$, EL, kv, WI	0.807	0.248	-0.70, 0.76, -0.64, 0.60	<0.001
	m) $NDVI_{Max}$, EL, kv, P	0.800	0.252	-0.74, 1.02, -0.61, 0.46	<0.001
	n) $NDVI_{Min}$, kh, kv, CTI	0.728	0.294	-0.41, -0.56, -0.64, -0.76	<0.01
	o) kv, P, WI	0.644	0.336	-0.87, -1.26, 1.88	<0.01
	p) EL, kv	0.606	0.353	0.80, -0.78	<0.01
	q) kv, WI	0.433	0.424	-0.65, 0.64	<0.05
	r) kh, CTI	0.387	0.441	-0.68, -0.70	<0.05
	s) $NDVI_{Max}$, WI	0.340	0.458	-0.82, 1.02	<0.05

[a)] Standardised coefficients

The variogram presented in Fig. 4c was constructed on the residuals of the best regression model in Table VI (model l), where data were classified by morphology. It showed that in the LL class the short-range variability was removed almost completely, while a (spatial) variance of 0.164 remained unexplained. In this model both topographic and climatic-vegetation factors were present as independent variables, implying that all the ranges of spatial variability were considered. A better result emerged from the variogram of the residuals of the best regression model in Table VII (model l), associated to a small residual long-range variance (0.075, Fig. 4d). In Fig. 3d the estimates obtained with the best model of the forest-LL class were compared to the observed values. The most evident effect of the combined vegetation-curvature stratification was the improvement in the estimates of the low/high SOC samples, which were largely over/underestimated in the previous cases.

CONCLUSIONS

In this work we exploited the information derived from vegetation, climate, and morphology for studying the spatial distribution of the soil organic carbon concentration in forest and grassland areas. The case study was formed by 114 samples belonging to a vast heterogeneous region of the North-West of Italy. The first attempts to statistically reproduce the SOC concentration by means of vegetation and climatic predictors were unsatisfactory. In fact, these explanatory variables allowed us to reduce the long-range spatial variability, but had no effect on the short-range dependence and were therefore unable to take into account all factors influencing organic matter concentration in a highly heterogeneous region. Better results were obtained after stratification of the samples, and by introducing in the regression models topographic variables that, by governing water fluxes, had a large influence on SOC inputs, losses and turnover. When classes of homogeneous vegetation and topography were considered, the quality of the estimates improved, particularly for grasslands and forests in the absence of local curvatures. The results of this work indicate that the spatial variability of soil organic matter can be overcome only when an a-priori reduction of the heterogeneity is applied. The only limitation to the feasibility of the approach is in the availability of the predictors, included the availability of finer scale digital elevation models. Further efforts in this field of investigation will be devoted to introduce more causative factors in the prediction models, provided that additional and more focused data will become available.

ACKNOWLEDGEMENTS

The IPLA Institute (Turin, Italy) is acknowledged for providing part of the soil data used in this study. G. Laguardia is gratefully acknowledged for assistance in data preparation and some computations.

REFERENCES

Abnee, A.C, Thompson, J.A., Kolka, R.K., 2004a. Landscape modeling of in situ soil respiration in a forested watershed of southeastern Kentucky. *Environ. Manage*. **33**(Suppl.): S168--S175.

Abnee, A.C., Thompson, J.A., Kolka, R.K., D'Angelo, E.M., Coyne, M.S., 2004b. Landscape influences on potential soil respiration in a forested watershed of southeastern Kentucky. *Environ. Manage*. **33** (Suppl.): S160--S167.

Arrouays, D., Daroussin, J., Kicin, J.L., Hassika, P., 1998. Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of France. *Soil Science.* **163**(2): 103--108.

Arrouays, D., Vion, I., Kicin, J.L., 1995. Spatial analysis and modeling of topsoil carbon storage in temperate forest humic loamy soils of France. *Soil Sci*. **159**: 191--198.

Augusto, L., Ranger, J., Binkley, D., Rothe, A. 2002. Impact of several common tree species of European temperate forests on soil fertility. *Ann. For. Sci*. **59**: 233-253.

Baldock, J.A., Nelson, P.N., 2000. Soil Organic Matter. *In* Sumner M.E (ed.), Handbook of soil science. CRC Press, Boca Raton. pp. B25--B84.

Bartolini, E., Claps, P., Laio, F., 2008. Analysis of the spatial variability of precipitation characteristics in Piemonte (in Italian). Working Paper 2008 – 01, DITIC, Politecnico di Torino (http://www.idrologia.polito.it/~bartolini/WP2008_01.pdf).

Bell, J.C., Grigal, D.F., Bates, P.C., 2000. A soil-terrain model for estimating spatial patterns of soil organic carbon. *In* Wilson, J.P., Gallant, J.C. (ed.), Terrain Analysis—Principles and Applications. Wiley, New York. pp. 295--310.

Beven, K.J., Kirkby, M.J., 1979. A physically-based variable contributing area model of basin hydrology. *Hydrological Science Bulletin*. **24**(1): 43--69.

Bot, A., and Benites, J., 2005. The importance of soil organic matter: Key to drought-resistant soil and sustained food and production. *FAO Soils Bullettin* **80**, Rome (http://www.fao.org/docrep/009/a0100e/a0100e00.htm).

Burgess, T.M., Webster R., 1980. Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. *J. Soil Sci*. **31**: 315--331.

Chai, X., Shen, C., Yuan, X., Huang, Y., 2008. Spatial prediction of soil organic matter in the presence of different external trends with REML-EBLUP. *Geoderma*. **148**: 159--166.

Chaplot, V., Bernoux, M.,Walter, C., Curmi, P., Herpin, U., 2001. Soil carbon storage prediction in temperate hydromorphic soils using a morphologic index and digital elevation model. *Soil Sci.* **166**: 48--60.

Chaplot, V., Walter, C., Curmi, P., 2000. Improving soil hydromorphy prediction according to DEM resolution and available pedological data. *Geoderma.* **97**: 405--422.

Claps, P., Giordano, P., Laguardia, G., 2008. Spatial distribution of the average air temperatures in Italy: quantitative analysis. *Journal of Hydrologic Engineering.* **13**(4): 242--249.

European Environment Agency, 2005. CORINE land cover 2000. Available at: http://dataservice.eea.europa.eu/dataservice/available2.asp?type=findkeyword&theme=clc2000 (last access: December 2008).

Florinsky, I.V., Kuryakova, G.A., 1996. Influence of topography on some vegetation cover properties. *Catena.* **27**(2): 123--141.

Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L.D., Holmes, K.W., 2000. Modeling soil–landscape and ecosystem properties using terrain attributes. *Soil Sci. Soc. Am. J.* **64**: 2046--2056.

Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J., 1995. Soil landscape modelling and spatial prediction of soil attributes. *International Journal of Geographical Information Systems.* **9** (4): 421--432.

Grimm, R., Behrens, T., Märker, M. and Elsenbeer, H., 2008. Soil organic carbon concentrations on Barro Colorado Island – Digital soil mapping using Random Forests analysis. *Geoderma.* **146**: 102--113.

Guber, A.K., Pachepsky, Y.A., van Genuchten, M.T., Rawls, W.J., Simunek, J., Jacques, D., Nicholson, T.J., Cady, R.E., 2006. Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. *Vadose Zone Journal.* **5**(1): 234--247.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer Series in Statistics. Springer-Verlag, New York, USA.

Homan, P.S., Sollins, P., Chappell, H.N., Stangenberger, A.G., 1995. Soil organic carbon in a mountainous, forested region: relation to site characteristics. *Soil Science Society of America Journal.* **59**: 1468--1475.

Jackson, R.B., Canadell, J., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D., 1996. A Global Analysis of Root Distributions for Terrestrial Biomes. *Oecologia*. **108**: 389--411

Jenny, H., 1941. Factors of Soil Formation, A System of Quantitative Pedology, McGraw-Hill, New York.

Johnson, K.D., Scatena, F.N., Johnson, A.H., Pan, Y., 2009. Controls on soil organic matter content within a northern hardwood forest. *Geoderma.* **148**: 346--356.

Jones, R.J.A., Hiederer, R., Rusci, E. and Montanarella, L., 2005. Estimating organic carbon in the soils of Europe for policy support. *European Journal of Soil Science.* **56**: 655--671.

Kottegoda, N. and Rosso, R., (1998) Statistics, probability, and reliability for civil and environmental engineers, McGraw-Hill, International Edition.

Kyungsoo, Y., Amundson, R., Heimsath, A.M., Dietrich, W.E., 2006. Spatial patterns of soil organic carbon on hillslopes: Integrating geomorphic processes and the biological C cycle. *Geoderma*. **130**: 47--65.

Lagacherie, P., McBratney, A.B., 2004. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. *Developmnets in Soil Science*. **31**: 3-21.

Lal R. 2008. Sequestration of atmospheric CO_2 in global carbon pools. *Energy Environ. Sci*. **1**: 86--100.

MacMillan, R.A., Pettapiece, W.W., 1997. Soil Landscape Models: Automated Landscape Characterization and Generation of Soil- Landscape models. Technical Bulletin No. 1997-1E. Research Branch, Agriculture and Agri-Food Canada, Lethbridge.

McBratney, A.B., Mendonça Santo, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma* **117**(1-2): 3--52.

McKenzie, N.J, Ryan, P.J., 1999. Spatial prediction of soil attributes using terrain analyses. *Geoderma.* **89**: 67--94.

Meier, I.C., Leuschner, C., 2010, Variation of soil and biomass carbon pools in beech forests across a precipitation gradient, *Global Change Biology*. **16**: 1035--1045.

Melillo J.M. McGuire A.D., Kicklighter D.W. Moore B. III., Vorosmarty C.J., Schloss A.L. 1993. Global climate chage and terrestrial net primari production. *Nature.* **363**: 234--240.

Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal.* **57** (2): 443--452.

Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological and biological applications. *Hydrological Processes.* **5**(1): 3--30.

Mueller, T.G., Pierce, F.J., 2003. Soil carbon maps: enhancing spatial estimates with simple terrain attributes at multiple scales. *Soil Science Society of America Journal.* **67** (1): 258.

Nemes, A., Rawls, W.J., Pachepsky, Y.A., 2005. Influence of organic matter on the estimation of saturated hydraulic conductivity *Soil Science Society of America Journal*. **69**(4): 1330--1337.

Olaya, V., 2004. A gentle introduction to SAGA GIS. Edition 1.1.

Pennock, D.J., Zebarth, B.J., De Jong, E., 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan,Canada. *Geoderma.* **40** (3-4): 297--315.

Perruchoud, D., Walthert, L., Zimmerman, S., Luscher, P., 2000. Contemporary carbon stocks of mineral forest soils in the Swiss Alps. *Biogeochemistry.* **50** (2): 111--136.

Powers, J. S. and Schlesinger, W. H., 2002. Relationships among soil carbon distributions and biophysical factors at nested spatial scales in rain forests of northeastern Costa Rica. *Geoderma* **109**: 165--190.

Prentice, I.C., 2001. The Carbon Cycle and Atmospheric Carbon Dioxide, Climate Change 2001: The Scientific Basis IPCC. Cambridge University Press, Cambridge, UK. pp. 183–237.

Saby, N.P.A., Arrouays, D., Antoni, V., Lemercier, B., Follain, S., Walter, C. and Schvartz, C., 2008. Changes in soil organic carbon in a mountainous French region, 1990-2004. *Soil Use and Mangement.* **24**: 254--262.

Simbahan, G.C., Dobermann, A., Goovaerts, P., Ping, J., Haddix, M.L., 2006. Fine resolution mapping of soil organic carbon based on multivariate secondary data. *Geoderma.* **132** (3--4): 471--489.

Speight, J.G., 1968. Parametric description of landform. *In*: Stewart, G.A. (ed.), Land Evaluation. Macmillan, Melbourne.

Terra, J.A., Shaw, J.N., Reeves, D.W., Raper, R.L., van Santen, E., Mask, P.L., 2004. Soil carbon relationships with terrain attributes, electrical conductivity, and soil survey in a Coastal Plain landscape. *Soil Sci.* **169**: 819--831.

Thompson, J.A., Kolka, R.K., 2005. Soil carbon storage estimation in a central hardwood forest watershed using quantitative soil–landscape modeling. *Soil Sci. Soc. Am. J.* **69**: 1086--1093.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. The Geographical Review, 38(55), American Geographical Society: New York. pp.55-94.

Tisdall, J.M., and Oades, J.M., 1982. Organic matter and water stable aggregates in soils. *J. Soil Sci.* **33**: 141--163.

Walker, P.H., Hall, G.F., Protz, R., 1968. Relation between landform parameters and soil properties. *Soil Sci. Soc. Am. Proc*. **32**: 101–104.

Wilson, J.P., Gallant, J.C., 2000. Terrain Analysis—Principles and Applications. Wiley, New York. pp. 295–310.

Yoo, K., Amudson, R., Heimsath, A.M., Dietrich, D.E., 2006, Spatial patterns of soil organic carbon on hillslopes: Integrating geomorphic processes and the biological C cycle, *Geoderma*. **130**: 47--65.

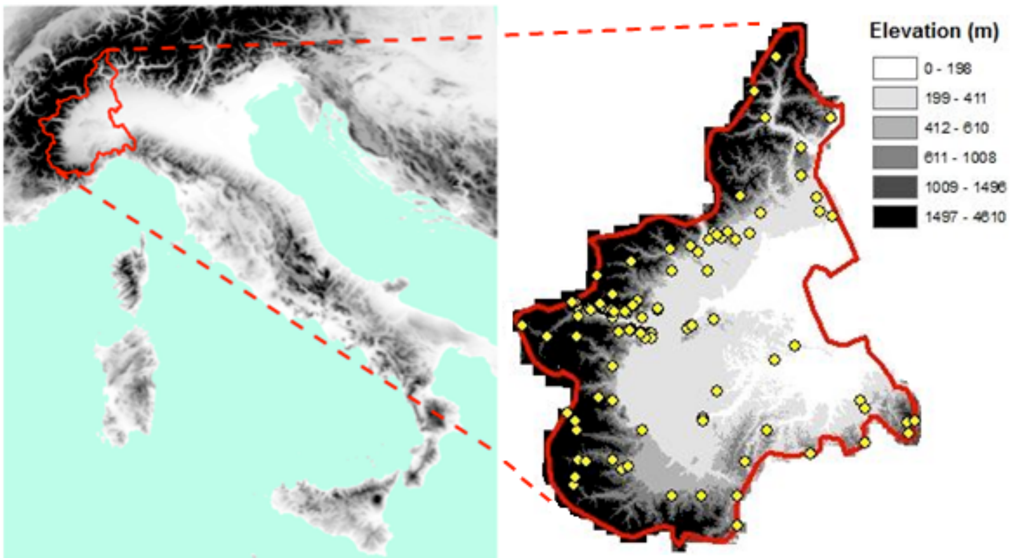

Fig. 1 Shown is a map of the study region with the locations of the sampled sites. The gray tones to the elevation variations within the region.

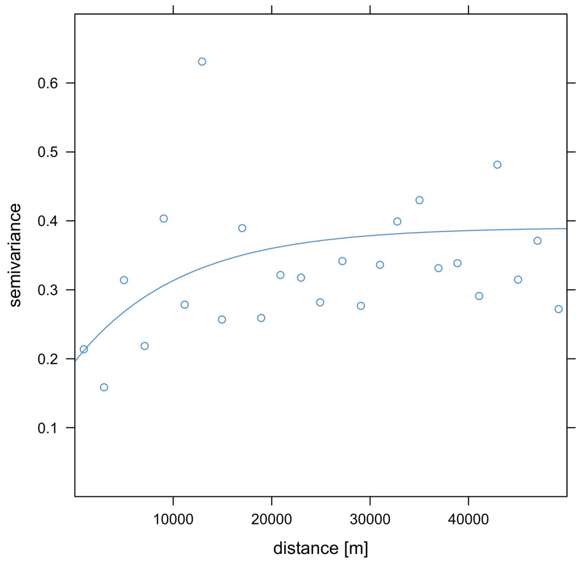

Fig. 2 Semivariogram of the natural logarithm of the soil organic carbon concentration for the 0--30 cm layer.

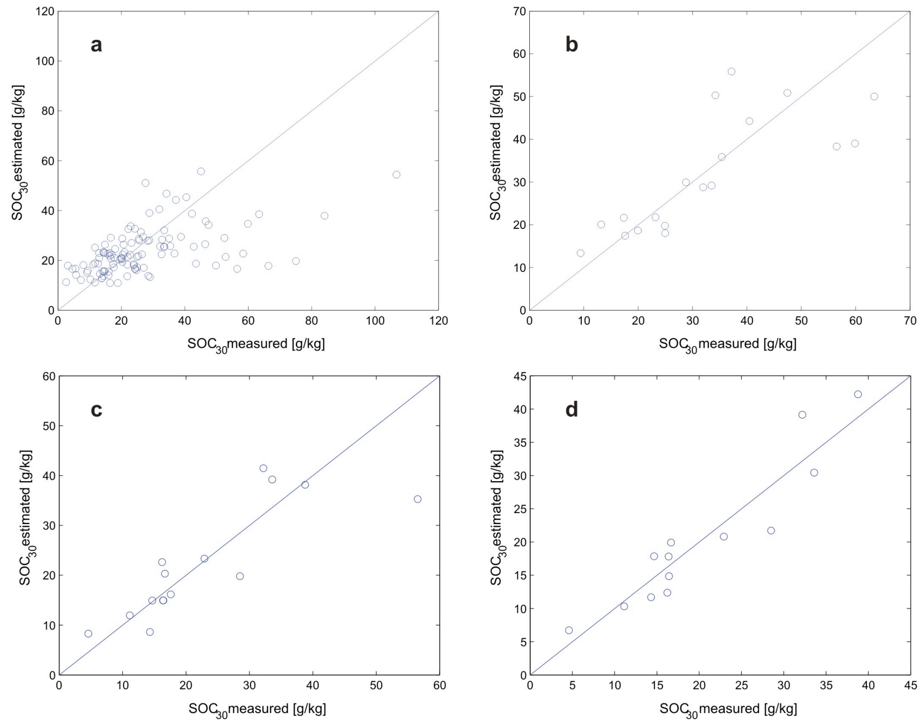

Fig. 3 Comparison between the observed and predicted values of SOC_{30} for: (a) the whole dataset (first model in Table IV); (b) the grasslands subset (best model for grasslands in Table V); (c) the LL subset (best model for the LL class in Table VI); (d) the forest-LL subset (best model for the forest-LL class in Table VII)

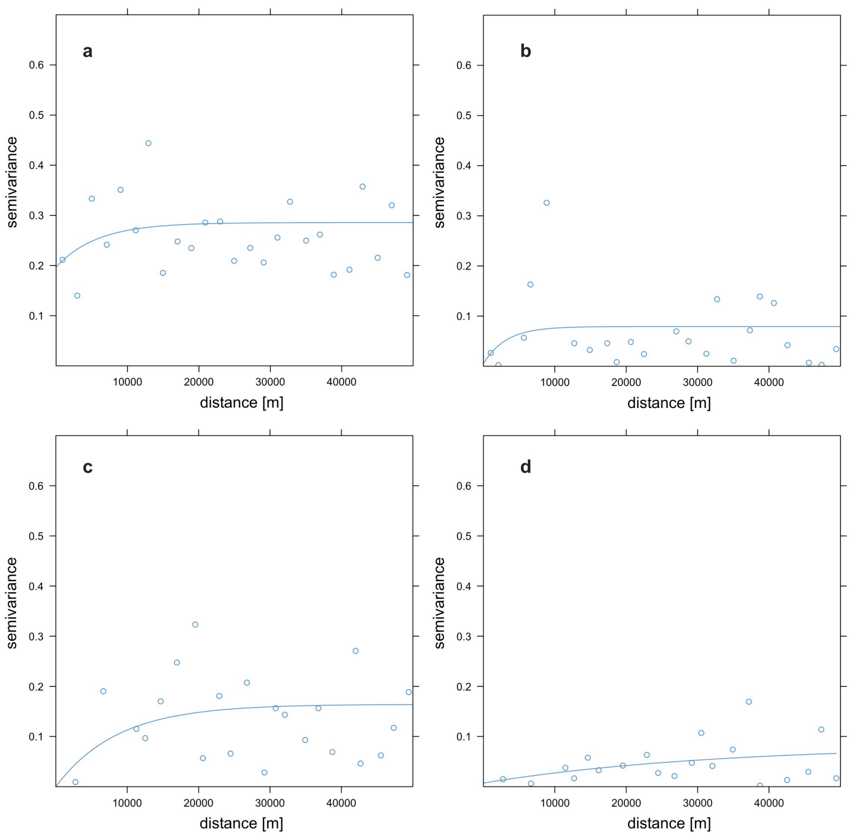

Fig. 4 Semivariogram of the residuals of the best regression model for (a) the whole dataset (first model in Table IV); (b) the grasslands subset (best model for grasslands in Table V); (c) the LL subset (best model for the LL class in Table VI); (d) the forest-LL subset (best model for the forest-LL class in Table VII).