Regional scale characterization of the geomorphic control on the spatial distribution of soil organic carbon in cropland

F. STEVENS*, P. BOGAERT, K. VAN OOST, S. DOETTERL & B. VAN WESEMAEL

Uiversité catholique de Louvain -

Earth and Life Institute – Environment. Croix du Sud, 2. L7.05.16

Louvain-la-Neuve 1348

Belgium

***Corresponding author** email francois.stevens@uclouvain.be

Running title: Soil organic carbon prediction at regional scale

Summary

The heterogeneity of the spatial distribution of Soil Organic Carbon (SOC) at the landscape scale is generally not considered in regional or national SOC dynamics models. In cropland, this heterogeneity is largely controlled by topography, which influences the distribution of water, energy and sediments, and thus the SOC dynamics. Sediment redistribution rates have strongly increased since the mechanization of agriculture. The oversimplification of landscape processes in regional models of C dynamics may add to the uncertainty in C balances. Therefore, a better characterization of the importance of landscape scale effects on the SOC distribution throughout a region is needed. This study proposes to characterize the relative importance of geomorphology on the SOC horizontal and vertical variability across the croplands in the Belgian loess belt region. A large legacy dataset of soil horizons was exploited together with 147 recently sampled profiles. Mean SOC depth profiles for different soil types were compared. Various topographic attributes were computed from a digital elevation model, and their influence on SOC was quantified through simple linear models. Finally, SOC content was mapped at three depth layers through multiple linear models, and results were cross-validated. The legacy dataset allowed identifying significant differences in the mean SOC profile according to texture, drainage or profile development classes. A clear relationship between SOC content and topographic attributes was highlighted, but only for the recently sampled profiles. This may be explained by a substantial error on the location of the profiles of the legacy dataset. This study thus shows evidence that the major control on the vertical distribution of SOC is related to topography in a region where observed heterogeneities for other commonly involved factors are limited. However, the large amount of unexplained variability still limits the usefulness of the spatial prediction of SOC content,

and suggests the importance of additional influencing factors.

Keywords: erosion, arable soil, soil organic carbon.

Introduction

The soil organic carbon (SOC) pool is of great importance for the global carbon (C) cycle. Soil represents the largest terrestrial C pool, containing more organic C than biosphere and atmosphere together (Grace, 2004). However, it is the largest source of uncertainty in regional and continental C balances of terrestrial ecosystems (van Wesemael et al., 2011). For example, predictions about the response of the soil carbon store to global warming are diverging (e.g. Trumbore & Czimczik, 2008). As a result, soils are increasingly receiving attention for the potential role they can play in CO_2 mitigation (Milne et al., 2007). One of the sources of uncertainty in C flux modeling between soils and atmosphere arises from the lack of consideration of the landscape processes influencing SOC (Ciais et al., 2010). Models predicting the temporal change of SOC for a region typically represent the soil system as a collection of large spatially homogeneous units (Easter et al., 2007). However, horizontal and vertical variability of SOC within landscapes is large (Stevens et al., 2006; VandenBygaart et al., 2007; Hbirkou et al., 2012). For example, Goidts & van Wesemael (2007) showed that SOC variability at the field scale is of the same order of magnitude as the variability inside a large map unit of cropland. The implications of this large variability for model parameterization, predictions and soil monitoring schemes are still poorly understood.

The topography is typically related to the spatial patterns of SOC in the landscape , as geomorphic landscape features control hydrologic and erosional processes, and soil temperature (Moore et al., 1993; Florinsky et al., 2002). Transport of sediments influences C fluxes between soil and atmosphere through the dynamic replacement of eroded C at eroding sites and reduced decomposition of buried C at depositional sites (Van Oost et al., 2005; Quinton et al., 2010; Vandenbygaart et al., 2012). Conceptual models have been developed that combine geomorphic models with models of carbon dynamics. They have been applied to eroding micro-catchments and were able to closely reproduce the observed SOC depth profiles for a wide range of erosional and depositional settings (Liu, 2003; Rosenbloom et al., 2006; Dlugoss et al., 2010). Application of these models suggests that lateral fluxes of SOC, sediments and water will further enhance the spatial heterogeneity in SOC storage within agricultural landscapes. However, these studies typically focus on small areas with high erosion rates and pronounced topography. So far, the importance of erosion-induced and topography related variability in C stocks at the regional scale remains unclear.

However, the spatial prediction of SOC content accounting for landscape-scale variability is still a challenging task, given the large number of additional processes occuring at that scale (Viaud et al., 2010). Even if the main controls are identified, the true landscape condition differs from the ideal conditions of field and hillslope scale studies (e. g. Van Hemelryck et al., 2011). A complex lateral and vertical distribution of SOC may result from the historic and current interactions between many processes incuding sediment transfers or agricultural management (Sleutel et al., 2007a; Goidts et al., 2009). Doetterl et al.(2012) showed that even if the subsoil SOC content is generally larger at depositional than at eroding positions along a slope, differences between similar geomorphological positions of different slopes may be important. The difficulty to predict soil horizon thickness at a high spatial resolution (Vanwalleghem et al., 2010) also suggests that the landscape processes result in important point or plot scale variability in soil properties. A spatially explicit description of the soil processes and properties at landscape scale is, however, of prime importance to simulate the evolution of SOC stocks in the coming decades (Viaud et al., 2010).

To adress this issue, the main objective of this paper is to assess the importance of the topographic control on SOC spatial distribution, over a large area of cropland, accounting for landscape scale variability. The effect of topography will be characterised on both the vertical SOC variability directly, and the lateral variability at different depths. The second objective is to compare the potential of a a legacy dataset and a more recent dataset to fulfill the first objective.

Material and methods

Study area

The study area was chosen to maximize variability in topographic features while limiting variability in other environmental factors influencing SOC. To this end, the cropland of the Belgian Loess Belt, in central Belgium was selected (Fig. 1). The Belgian Loess Belt is an area of 9921 km² of which 43% is occupied by cropland. It is characterized by a rolling topography with plateaus, slopes and some incised rivers with generally well drained, dry valley bottoms. The climate of the region is a temperate oceanic climate with mild winters and cool summers. The geological substrate is a several meters thick Pleistocene aeolian deposit of calcareous loess in which luvisols have developed (Gullentops, 1954). Loess deposits are typically thicker on south facing slopes than on north facing slopes and are overlaying tertiary sands (Vanwalleghem et al., 2010). In some locations, these sandy layers are already apparent at the surface, as soil erosion has removed several meters of the loess. At present, the main crops in the region are cereals (46%), sugar beet (20%), silage maize (10%) and potato (7%) (van Wesemael et al., 2010) with typical C inputs to the soil by plants estimated at 1.93 Mg C ha^{-1} $year^{-1}$ and by manure at 1.1 Mg C ha^{-1} $year^{-1}$. Forests, grassland and urban areas are typically located in the floodplains.

Carbon datasets

Characterizing the SOC spatial distribution in both topsoil and subsoil required the use of spatial datasets describing the SOC profile. In our analysis, both a legacy and a recently-sampled dataset of soil profiles were used. The legacy dataset is composed of soil profiles described during the Belgian National Soil Survey (1947-1962) (De Leenheer et al., 1968). Soils were sampled by horizon from observation pits, and physical and chemical analyses were performed in the laboratory. SOC was estimated by dichromate wet combustion (Walkley & Black, 1934). Soil profile location, horizon limits, physical and chemical properties and classification were recorded on paper.

In 1993, the legacy dataset was digitized creating the digital soil database ‘Aardewerk’ (Van Orshoven et al., 1988). However, due to storage constraints, data description was simplified and standardized and, as a result, part of the information was lost. A critical simplification was the coding of the horizon boundaries as intervals instead of exact values. Some horizons were also omitted during the transcription. The methodology we developed to reconstruct detailed SOC profiles from incomplete and uncertain information of Aardewerk is described later. From all the profiles under cropland, 543 were removed because they had missing horizons or inconsistent values for some variables, leaving 2449 profiles available. In the following, the dataset containing the reconstructed SOC profiles located in the study area will be called AW93.

Recently, original descriptions of soil profiles located in Flanders were again digitized, but this time with the exact limits of the horizons (Van De Vreken et al., 2011). This dataset was used to validate the SOC profile reconstruction method (see below). We selected 944 profiles for which SOC values were coherent with the corresponding AW93 profiles. This dataset is referred to as AW10. A common factor of 1.33 (Sleutel et al., 2007b) was used to correct SOC concentration for the incomplete oxidation of Walkley and Black method, in AW93 and AW10.

Finally, we also used 139 profiles from a recently sampled dataset (Doetterl et al., 2013). These profiles were randomly selected from existing AW93 profile locations within a 40 x 40 km² sub-area (Fig. 1). It is not possible to affirm that their position exactly matches the position of the corresponding AW93 profile, since no marker was left in the soil during the original sampling of AW93 profiles. Besides, it was checked that the current land use had not changed since the Belgian National Soil Survey. Between 2010 and 2012, soil cores were extracted and analysed for carbon with a spectrometer by 3 cm depth intervals up to one meter soil depth. Reflectance was measured using a Fieldspec-Pro spectroradiometer (ASD, Boulder, CO) in the Vis-NIR range (350 - 2500 nm) under laboratory conditions. SOC concentrations were predicted from spectral information using a multiple tree algorithm. The root mean square error associated with these estimates was low (1.22 g C kg^{-1}), and similar to the analytical error of the reference technique of dry combustion. In the following, this dataset will be referred as R_AW (for Resamples Aardewerk).

The two datasets which are used – separately – for describing the spatial distribution of SOC, AW93 and R_AW, are very different but complementary. The legacy dataset AW93 contains a large number of data, with associated observations on soil properties, but a low precision in both the coordinates and the retrieved SOC content at a given depth. The recent dataset R_AW contains a one order of magnitude-smaller number of profiles and no associated soil observations, but a very high precision in both coordinates and depth-explicit SOC content. In this study, the large number of profiles in AW93 will be an advantage to highlight statistical differences in SOC depth distribution between soil classes, when the good precision in the coordinates of R_AW profiles will be an advantage for 3D spatially explicit predictions.

Terrain attributes

Terrain attributes derived from a DEM are used in our analyses. The DEM was produced by merging two DEMs of 10 and 5 meters resolutions that partially cover Belgium (Demarcin et al., 2009; AGIV, 2006). A new DEM was then created by interpolating at a 10 meters grid resolution and by smoothing the surface using a 3x3 mean filter. Terrain attributes were computed using Matlab and the TopoToolbox package (Schwanghart & Kuhn, 2010). They are elevation (ELEV), slope gradient (GRAD), total curvature (CURV), south orientation (SOUTH), topographic wetness index (TWI, Beven & Kirkby, 1979), stream power index (STR) and topographic position index for different sizes of neighbourhood (TPI) (Weiss, 2001). The flow accumulation, implied in the calculation of hydrologic attributes, was computed using multiple flow direction algorithms. TPI is the relative difference between the elevation of a cell and the average elevation of the cells within a given radius. Three radiuses were used (32 m, 128 m and 512 m), in order to represent different scales. The south orientation was taken as the cosine between the slope orientation vector and a vector pointing south. It was selected to reflect the anisotropy of the geomorphic variables observed in the region, with loess deposits thickness and texture depending on slope orientation (Goossens, 1997).

Horizons distribution reconstruction

Due to the loss of information while digitizing the legacy profile dataset, the vertical position of each horizon in AW93 is not given as an exact value but as two intervals: one interval indicates the depth of the upper boundary of the horizon; the other indicates its thickness class. These intervals belong to a set of predefined intervals (Table 1). In order to address this issue, a Monte Carlo simulation based method was developed (Fig. 2). For a given profile, the thickness of each horizon was simulated from a uniform distribution within the given thickness interval. Then the corresponding horizon depths were derived assuming, naturally, no gap or overlap between horizons. The simulated set of horizon thickness values was retained only if the derived horizon depths were also enclosed in the depth intervals given in AW93, for the corresponding horizons. This means that simulated sequences of horizons that were not coherent with the original information given by AW93 were discarded. A smooth SOC concentration profile was then obtained from the simulated distribution of the horizons and the SOC horizon values, by equal area quadratic spline (EAQS, see below). The whole process was repeated until 10,000 valid horizon sequences were simulated. The estimated profile was then taken as the mean of these 10,000 profiles. This method allowed us to use all the information on horizon distribution contained in AW93 simultaneously without adding arbitrary information.

The AW10 dataset was used to validate the horizon distribution reconstruction method, because it contains the same profiles than AW93, but provided the exact horizon limits. The profiles of AW10 were also smoothed using a spline method (see below). Then, all the AW10 profiles were compared with the corresponding AW93 profiles. The mean profile of both datasets and the profile of mean difference (bias) and root mean square difference (RMSD) were computed. This resulted in an estimate of the magnitude of the error caused by the coding of the horizon limits used in the legacy dataset including its depth distribution.

Smoothing the profiles

SOC concentration was given by horizon in AW93 and AW10. Nevertheless, estimating depth explicit profiles of SOC was needed for further analysis. Traditionally, it is assumed that the horizon value of a particular attribute represents its average value over that horizon (Malone et al., 2009). In luvisols that developed in thick loess deposits, soil properties including SOC content are expected to vary continuously with depth, except at the transition zone below a mixed plough layer (Minasny et al., 2006; Kempen et al., 2011; Myers et al., 2011). This is confirmed by the re-sampled cores that do not display SOC vertical patterns of abrupt transitions between horizons (Doetterl et al., 2013). Thus, using a step function with mean SOC value over each horizon could lead to bias and cumulative errors may produce under- or overestimation for a given soil layer (Ponce-Hernandez et al., 1986). Although a SOC profile is often described using a negative exponential depth function, fitting this function does not ensure that the mean SOC content measured for each horizon is respected. In this study we therefore chose to apply the equal area quadratic spline method of Bishop et al. (1999). This method produces smoothly varying profiles that are consistent with the observed mean SOC values for each horizon. However, we assume that the first horizon , the soil surface horizon, is a tillage horizon and, given the continuous mixing by tillage operations, the SOC content is constant. Therefore, inside the depth interval corresponding to the first horizon, interpolated values from the spline method were not used, and the mean horizon value given by the database was used instead. The SOC concentration was then calculated for three depth layers (0-30 cm, 30-60 cm, 60-90 cm) from the interpolated profiles.

Computing class-representative SOC profiles

Even though the legacy dataset did not permit to spatially predict SOC content (see Results and discussion), it allowed characterizing the influence of topography on the vertical distribution of SOC. AW93 profile descriptions contain information about soil texture, drainage and profile development classes. The profile development classification was simplified to the presence or absence of a colluvial layer above 120 cm depth ("colluvial" and "not colluvial" classes). This variable may indeed be considered as a proxy for the topography since the occurrence of a colluvium largely depends on the position in the hillslope. To compare the influence of these properties on SOC profile inside the study area, profiles were grouped based on combinations of soil properties. Then, for each group, a representative profile was computed by taking the mean of all the profiles in this group, and the error on the mean was computed for all depths assuming normal distribution of the errors. Finally, representative profiles of each group were compared. Representative profiles were computed for the interaction between drainage and development, and between texture and development. In order to assure a sufficient number of elements in each group to permit meaningful statistical analyses, the classifications of AW93 drainage and texture were also simplified. Drainage was split into two classes: well drained and poorly drained. Poorly drained profiles display traces of temporary water saturation (pseudo gley) above 80 cm when well drained profiles do not. 24 profiles (0.9 %) showing traces of permanent water saturation (gley) were discarded. Texture was split into three classes: “silt loam”, “(heavy) sandy loam” and “light sandy loam”. 103 profiles (4.2 %) belonging to other texture classes (mainly clay) were not used.. Only groups containing more than 30 profiles are shown in the results.

To evaluate the direct influence of the topography on the SOC profile, a method similar to the computation of soil class representative profiles was applied, but this time using terrain attributes and R_AW dataset. Indeed, the use of terrain attributes with AW93 or AW10 profiles was not possible due to the error in the coordinates (see below). For each terrain attribute, the profiles were sorted by increasing value of this terrain attribute, and divided in three groups of equal size. Mean profiles were then computed for the three groups, and compared between them.

For both the soil property and terrain attributes classifications, the within class variability of SOC concentration for our three successive 30 cm thick layers was computed and displayed as a boxplot using the ggplot2 package (Wickham, 2009). For each layer, an ANOVA F-test was used to test the hypothesis that all the means are equal. p-values are given in Tables 2-5. When the mean of the SOC values in one or many groups showed a relevant similarity or difference with the mean of the SOC values of another group, a t-test was performed to question the hypothesis of equality of the means. The AW93 profiles are distributed all over the study area, compared to the ones of R_AW which are restricted to a smaller sub-area. Thus, the large number of profiles increases the power of statistical tests for AW93.

Spatial prediction of SOC

In order to estimate the importance of topography on the spatial distribution of SOC, statistical analyses were performed with AW93, R_AW and terrain attributes. In a first step, the Pearson coefficients of correlation between the observed SOC content in the three depth layers and the terrain attributes were computed. The correlations were analysed for both AW93 and R_AW, in order to assess their ability to spatially predict SOC. Because of the very weak values observed with the legacy dataset AW93 (see Table 6), a procedure was developed to check if these weak values could originate from a lack of precision in profile geolocations. Simulated random errors of given magnitude and random direction were added to the coordinates of the profiles in R_AW. They were chosen because they are characterised by a standard deviation of the error in their (unbiased) position of only approx. 5 meters, due to precision of the GPS. The magnitude of the simulated errors was increased gradually from 0 to 300 meters by 10 meters increment (Fig. 6). Then, correlations between SOC layer contents and terrain attributes extracted at modified positions were computed. Because of the issue of the precision of the AW93 samples, the next steps of the SOC spatial prediction were only performed using the R_AW dataset.

In a second step, a multiple linear regression model was used to predict SOC from the terrain attributes. A model with an arbitrary number of predictors was selected as the one which minimizes the residual sum of squares during a 10-fold cross-validation procedure. This lead to the best predictive model among tested ones. Terrain attributes, terrain attributes at second power, and interaction between all pairs of terrain attributes were all included as predictors. However, the number of predictors in the selected model was limited to three because of computational constraints.

Finally, predicted SOC spatial stock from the best model was mapped in a 5 km × 5 km plotting area (see Fig. 1), randomly selected inside the sub-area containing the R_AW profiles. The stock was calculated by multiplying on a pixel and layer basis the predicted SOC concentration with the estimated bulk density (BD). BD was itself estimated from SOC concentration using the general model of Manrique & Jones (1991). The choice of the model is of prime importance given that different estimation models applied over a whole region may lead to differences in mean BD up to 7.5%, and in mean SOC stock up to 6% (Liebens & VanMolle, 2003). The model was chosen for its simplicity and, relatively to other models, its good ability to predict soil bulk density at surface as well as along soil profile in Belgium (Boucneau et al., 1998).

Results and discussion

Horizon distribution reconstruction

The profiles calculated from incomplete horizon information (AW93) were compared to a subset of corresponding profiles that were derived from another database with exact horizon descriptions (AW10). The mean profile of both sets and the vertical profile of mean difference (bias) and root mean square difference (RMSD) are presented in Fig. 3. Bias is very low across depth, and reaches, between 15 and 30 cm, a minimum followed by a maximum, all having an absolute value less than 1 g C kg^{-1} . This may indicate a trend to underestimate the thickness of the surface horizon, since the transition between the surface horizon and the second horizon is generally also between 15 and 30 cm, and the surface horizon contains generally more SOC than the second one. The RMSD profile shows peaks in the same region, reaching up to 25% of the mean SOC content of AW10. This is due to the fact that the SOC content is high, and decreases fastly below the plough layer, and thus absolute error is high also. Above 10 cm and below 50cm, the RMSD is stable and varies between 0.4 and 0.5 g C kg^{-1}.

To interprete these results, one should keep in mind that SOC profiles in both AW93 and AW10 were estimated using the spline smoothing method mentioned before, and are thus not the true SOC profiles. The smoothing method also assumes that the SOC concentration associated with each horizon is the exact mean SOC concentration over this horizon, an assumption that could not be verified. Furthermore, more complex horizon reconstruction methods could be tested, for example by using the exact horizon distributions of AW10 to estimate an a priori horizon distribution for AW93. But the goal of the study was not to concentrate efforts on the development of complex methods for this task, since the current results are already satisfying for our purposes.

Legacy datasets are still valuable, also because they represent a past situation. Our reconstruction method deals with a specific problem in the data format, and could not be directly transposed to other legacy datasets. However, it could prove to be useful if it inspires other researchers when finding a strategy to exploit other legacy dataset. Besides, we had the chance to receive a large number of profiles without this specific problem in the format (AW10), permitting to validate our reconstruction methods, which is rarely the case.

Soil property influence on SOC profile

The soil property influence on the vertical distribution of SOC in AW93 is displayed by Fig. 4 and Tables 2 & 3. Generally, the representative profiles of each group display similar shapes, and are well sorted with few intersection of the curves. The intraclass variability for the SOC content inside the depth layers is large (Fig. 4, right side). However, thanks to the large number of profiles in each group, significant difference between mean layer SOC content of different classes can be observed, as indicated by the very small p-values of the ANOVA F-tests.

The upper part of Fig. 4, and Table 2, show the combined influence of drainage and profile development on SOC distribution. In the top layer, differences between groups are small compare to their mean values. The lowest SOC content was found in non-colluvial poorly drained profiles, which contain 12 % less SOC in this layer than other profiles ($p < 10^{-15}$). In the deeper layers, however, all the classes show relatively larger differences. In the bottom layer, the poorly drained not colluvial profiles, contain in average more than two times less SOC than the well drained colluvial profiles ($p < 10^{-15}$). A good drainage is positively correlated with SOC content, which has not always been observed (e.g. Tan et al., 2004). It can be explained by the fact that, outside redoxymorphic conditions, water may enhance SOC decomposition. In the study region, incidentally, redoxymorphic soils are mainly under grassland, and since all our profiles are under cropland, even the ones we classify as poorly drained undergo only temporary water saturation. Nevertheless, in these layers, we observe that the influence of profile development is still clearer than the influence of drainage.

The classification based on texture and development combinations (lower part of Fig. 4) also shows differences in SOC vertical distribution. For the topsoil layer, the mean SOC values are first sorted by texture, then by profile development. In particular, light sandy loam not colluvial profiles contain in average 25% SOC less than other not colluvial profiles (p <

10^{-15}), showing that even when all soils are loamy, texture can have a drastic influence. In contrast, a difference for example between the mean SOC content for colluvial silt loam profiles and non-colluvial silt loam profiles cannot be observed ($p = 0.26$). Texture has thus a larger influence than profile development on topsoil SOC content. In the two deepest layers, however, substantial differences may be observed between classes with similar texture and different profile development. In the 60-90 cm layer, silt loam colluvial profiles contain in average 41 % more SOC than silt loam non colluvial profiles ($p < 10^{-15}$). And the curves are first sorted by profile development, then by texture, showing that the effect of profile development is the most important in the deeper soil layers.

In summary, the influence of soil properties on SOC differs with depth and the presence of colluvium is the soil property which influences the most the deeper SOC content for our study area. Indeed, within the range of variability of texture and drainage of the cropland, the differences of deep SOC resulting from differences in profile development are larger than the differences resulting from differences in texture and drainage. The use of texture or drainage classes in combination with profile development classes permits to better highlight the influence of profile development alone. In this context, it should be pointed out that the soil property class-representative profiles are not intended to represent actual soil profiles, but only average, or type-profiles to be used for regional scale SOC distribution quantification.

Terrain influence on SOC profiles

Terrain attributes of the R_AW profiles were divided in 3 groups of equal size according to their order. For the sake of brevity, only the results for TWI and GRAD are shown and discussed (Fig. 5, Tables 4 & 5). For TWI, the mean SOC profile of the third class (the one with the highest TWI values) show the highest SOC content (Fig. 5). This third class contains in average 7.6 % ($p = 16 \times 10^{-4}$), 40 % ($p = 13 \times 10^{-4}$) and 33 % ($p = 16 \times 10^{-5}$) more SOC than the two other classes, for the 0-30 cm, 30-60 cm and 60-90 cm layers respectively. The positive relationship between SOC and TWI may not be explained by the effect of the wetness itself. This would contradict the trends previously observed for the mean profiles by soil class showing that, for a given soil development class, the drainage class containing most SOC is the less humid (Fig. 4). It can however be explained by the role of water erosion, since the accumulation of runoff in concave areas with a gentle slope induces sedimentation. And this results in the burial of former C-rich soil below the plough layer in a low mineralization context (Van Oost et al., 2012).

The effect of the slope classes (GRAD) is mainly visible for the third class (Fig. 5, Table 5). For the three layers successively, the mean content of the third class contains respectively 4 % ($p = 0.0037$), 21 % ($p = 0.0029$) and 24 % ($p = 32 \times 10^{-9}$) less SOC than the mean of the two other classes. This third class includes the profiles with a slope gradient larger than 3.4 %. This may be explained by the loss of SOC when the soil profile is truncated during erosion process. In the plough layer, the loss of carbon could be compensated through dynamic replacement, which would explain why the difference between classes is small in the 0-30 cm layer. However, in the deeper layer, inputs of SOC are reduced and replacement of SOC could be only partial.

The classification of the profiles by terrain attributes allowed to put in evidence a clear and significant influence on the SOC profile. The grouping by terrain attribute classes was also performed with the AW93 profiles, on individual texture and drainage classes. But the results did not show interesting relationships as with R_AW (not shown). This may be due to the positioning errors in AW93 (see below).

Analysis of the correlations for legacy and new data

The correlations between terrain attributes and SOC layer contents, and between SOC layers themselves, were computed for the AW93 and R_AW datasets (Table 6). First, our results show a strong dependence between the SOC content of the three layers for both datasets. The coefficient of correlation reaches 0.43 between the top and the bottom layer, and varies between 0.5 and 0.65 for the correlation between adjacent layers. There are small differences between the two datasets probably due to the characteristics of AW93 and the methods used to estimate its SOC profiles: the weaker precision in AW93 SOC content decreases the correlation between layers, while the Monte Carlo based method developed to deal with the uncertainty in the position of the horizons has a smoothing effect, and this likely increases the correlation between layers.

Then, we considered the correlations between terrain attributes and SOC layers, as preliminary information for the spatial prediction of SOC from topography. Results widely differ between AW93 and R_AW. In AW93, the correlations are generally weak, with the maximum coefficient of correlation of 0.28 reached by the relationship between elevetion and top layer SOC. The effect of elevation may be explained by the smooth variations of texture inside the Loess Belt, since the parts in the North and in the East, in average, have lower altitude, and contain in average less fine texture fractions. This effect is not visible for R_AW, since R_AW profiles are limited to a smaller sub-area, containing mainly silt loam. For R_AW, correlations are generally higher and all the attributes, except elevation and south aspect, show a highly significant correlation with the SOC content of at least one layer. The highest correlation observed are between wetness index and the deepest SOC layer, reaching a value of 0.55. Note that weaker correlations may be more significant for AW93 than for R_AW because of the larger number of observations.

The weak correlations for AW93 dataset could be explained, at least partially, by an error in the profile positions recorded in the 1950's. This is likely to result in an error in the terrain attributes extracted at these profile positions. The possible sources of error are positioning errors on paper maps during the original survey, and transformation between different geoids. While the last error component was estimated at approximately 60 m (Van De Vreken et al., 2011), the other cannot be quantified since the pit locations are not visible in the field anymore. The correlations between SOC and terrain attributes for R_AW profiles indicate the importance of precise geolocations (Fig. 6). For all terrain attributes, the absolute value of the correlation tends to decrease when the size of the errors increases, until it reaches a minimum plateau with some noise. However, the rate of decrease varies. With local attributes, like slope or curvature, it decreases faster than with attributes related to the position in the landscape, like wetness or large range TPI (not shown). This analysis confirms that the error associated with the coordinates of AW93 is likely to be the cause of the weak correlation between SOC layers and terrain attributes with this dataset. The weak correlations may be also partly explained by the precision of the SOC analysis in AW93, the fact that the profiles had to be reconstructed (Fig. 2), or the influence of past land use conversions prior to the National Soil Survey.

Spatially explicit prediction of SOC

A summary of the regression model used to predict SOC over the study area is given in Table

7. The table shows the terrain attributes selected by cross-validation, as well as the intercept, the coefficients and their standard errors, the mean error (ME) and root mean square error of prediction (RMSEP) , and the coefficient of determination R^2. Terrain attributes are sorted by decreasing importance (increasing p-values). All predictors are significant at $p < 0.05$. In all cases, the number of automatically selected attributes is equal to three, the maximum that was authorized. Interaction terms were selected for four predictors out of nine, terrain attributes at the power 2 were never selected.

For the top layer, TWI is the first predictor, with a positive coefficient. TWI was also the most correlated terrain attribute (Table 6). The selection of SOUTH as the next predictor in order of importance may be explained by small differences in texture related to the hillslope orientation (Goossens, 1997). The last predictor is the interaction between the elevation and the TPI with a range of 512 m. This may be due to small variations of texture correlated to the altitude. TWI, TPI with different radiuses, ELEV and GRAD are used, directly or inside interaction terms, in the models predicting SOC for the two deeper layers. The signs of the coefficients are coherent with the individual correlations (Table 6).

The models predicts respectively 21 % (0 – 30 cm) , 25 % (30 – 60 cm) and 33 % (60 – 90 cm) of the SOC variability (Table 7). The larger R^2 of the mid layer compared to the top layer, despite a similar RMSEP and a smaller mean SOC content, is explained by its larger variance. The mean error is negligible (~0.01 g C kg^{-1}) for all layers, confirming that the model is not biased. Points at the extreme of the range of SOC values have in general the larger errors in the prediction (Fig. 7, left column).

The fact that the prediction improves with deeper soil layers contradicts what has been observed in other studies (Minasny et al., 2006; Kempen et al., 2011). This may be due to both particularities in the study area and in the prediction method. In our study area, the role of topography is particularly important since the region combines slopes and intensive tillage, and the redistribution of sediments mainly influences the subsoil SOC dynamics through burial of SOC rich sediments. Besides, since we focus on the cropland of a pedologically homogeneous region, we do not exploit factors which mainly influence the topsoil layers, like soil type, vegetation or land-use, as it is often done in other studies through soil maps or remote images.

The predicted SOC layer contents were mapped into the 5 km × 5 km plotting area (Fig. 7, right column). For the three layers, the spatial patterns of SOC are very heterogeneous, even inside a field, which is coherent with the field or airborne observations (Stevens et al. 2010). Convergent positions and valley bottoms contain more SOC than other landscape positions, and this effect is stronger for deeper layers.

The large part of unexplained variability may be due in variable proportions to three factors. First, the quality of the data, i.e. the precision in the SOC content of R_AW and the terrain attributes. Secondly, the model structure, which does not permit to fit all the complexity and interaction of the processes controlling SOC. In their conceptual model, Rosenbloom et al. (2006) observe a sharp transition of SOC content between erosional and depositional surfaces, while Florinsky et al.(2002) observe a bimodal distribution for deep SOC. Finally, other factors influencing SOC were not considered. There could be small variations in the texture, differences in management practices or crop rotations between farms and influence of non cropped landscape features on adjacent fields (Viaud et al., 2010). However, an extensive knowledge of these variables is not easily available. For example, the information on the texture available in the soil map does not account for texture variability inside or between adjacent hillslopes.

More generally, comparing quantitavely our spatial prediction results with other studies should be made with caution. In our case, the study are was chosen to highlight the effect of topography in cropland but, usually, spatial predictions over regions are not restricted to one land use and soil type. Therefore, the range of variability of SOC is much larger, the relative effect of hardly observable landscape scale processes is reduced, and additional sources of information (vegetation images, land use/soil type maps, …) can be successfully exploited (e. g. Vasques et al., 2010).

Lastly, if in our case erosion-depostion processes are represented implicitly by terrain attributes, other models exist which aim at representing explicitly these processes, and the associated temporal evolution of the topography. These landscape evolution models (LEMs) use more flexible expressions for erosion, deposition and sediment transport processes, and run on a time step basis. Their calibration requires a large number of data, and for these reasons the models are generally applied to regions not larger than a catchment. However, some LEM have been calibrated an validated for millenial periods on a loess belt catchment (Temme et al., 2011). Since the Loess Belt is relatively homonegeneous in soil properties and agricultural management, the fact that the model parameters derived in this catchment are still partially valid for all the region, could be envisaged. This model could then give an estimate of the current sediment transport patterns, which could be related to SOC by empirical reationships, or integrated in a SOC dynamics model accounting for erosion (e.g. SPEROS-C, Van Oost et al, 2005).

Conclusions and perspectives

We conclude that topographical information, in relation to erosion and deposition processes, allows us to envisage the spatial prediction of SOC at regional scale but a large part of the variability still remains unexplained. The control of topography on SOC content increases with depth as we were able to explain up to 33% of the variability in SOC content in the 60-90 cm depth layer. This is substantially less than values reported for studies conducted in single fields or micro-catchments. Different strategies could be considered in order to improve prediction. In particular, for the topsoil layer, additional data could be exploited, like additionnal topsoil samples or remotely sensed images (Stevens et al., 2010). The dependency between topsoil and deeper SOC could also be used to improve the prediction of deeper SOC content from the additional data. Finally, the use of landscape evolution models to better estimate erosion-deposition pattens could be envisaged. Besides this, our study showed that precise geolocation of soil profiles is essential for spatial prediction in the landscape. The errors in the geolocations of the legacy dataset AW93 probably explains why high correlations between SOC content and terrain attributes are not observed for this dataset.

Acknowledgements

This research is part of an ARC project (Action de Recherche Concertée, contract number ARC 09/14-022) financed by the Communauté Wallonie-Bruxelles through the Académie universitaire ‘Louvain’. The support is gratefully acknowledged.

References

AGIV (Agentschap voor Geografische Informatie Vlaanderen). 2006. DHM-Vlaanderen, raster, 5m.

Beven, K.J., & Kirkby, M.J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* **24**, 43–69.

Bishop, T.F.A., McBratney, A.B., & Laslett, G.M. 1999. Modeling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma* **91**, 27–45.

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

Ciais, P., Wattenbach, M., Vuichard, N., Smith, P., Piao, S. L., Don, A., et al. 2010. The European Carbon Balance. Part 2: Croplands. *Global Change Biology* **16**(5), 1409–1428.

De Leenheer, Appelmans, F., & Vandamme, J. 1968. Cartes perforées et ordinateur comme instruments pour la caractérisation du sol et pédologie régionale. Le système des canes perforées de la section ”caractérisation du sol” de la cartographie des sols de Belgique. *Pédologie* **18**, 208–227.

Demarcin, P., Degré, A., Smoos, A. & Dautrebande, S. 2009. *Projet ERRUISSOL. Cartographie numérique des zones à risque de ruissellement et d'érosion des sols en région wallonne. Rapport final de convention DGO3-FUSAGx.*

Dlugoss, V., Fiener, P. & Schneider, K. 2010. Layer-specific analysis and spatial prediction of soil organic carbon using terrain attributes and erosion modeling. *Soil Sci. Soc. Am. J.* **74**, 922–935.

Doetterl, S., Six, J., van Wesemael, B., & Van Oost, K. 2012. Carbon cycling in eroding landscapes: Geomorphic controls on soil organic C pool composition and C stabilization. *Global Change Biol.* **18**, 2218–2232.

Doetterl, S., Stevens, A., Van Oost, K. & van Wesemael, B. 2013. Soil Organic Carbon Assessment at High Vertical Resolution using Closed-Tube Sampling and Vis-NIR Spectroscopy. *Soil Sci. Soc. Am. J.*

Easter, M., Paustian, K., Killian, K., Williams, S., Feng, T., Al-Adamat, R., et al.. 2007. The GEFSOC soil carbon modelling system: A tool for conducting regional-scale soil carbon inventories and assessing the impacts of land use change on soil carbon. *Agriculture, Ecosystems & Environment* **122**, 13–25.

Florinsky, I. V, Eilers, R.G., Manning, G.R. & Fuller, L.G. 2002. Prediction of soil properties by digital terrain modelling. *Environmental Modelling & Software* **17**, 295–311.

Goidts, E. & van Wesemael, B. 2007. Regional assessment of soil organic carbon changes under agriculture in Southern Belgium (1955-2005). *Geoderma* **141**, 341–354.

Goidts, E., van Wesemael, B., & Van Oost, K. 2009. Driving forces of soil organic carbon evolution at the landscape and regional scale using data from a stratified soil monitoring. *Global Change Biology* **15**, 2981–3000.

Goossens, D. 1997. Long-term aeolian loess accumulation modelled in the wind tunnel: The Molenberg case (central loess belt, Belgium). *Zeitschrift fur Geomorphologie*, ***41***(1), 115–129.

Grace, J. 2004. Understanding and managing the global carbon cycle. *Journal of Ecology*, **92**(2), 189–202.

Gullentops, F. 1954. Contributions à la chronologie du Pléistocène et des formes du relief en Belgique. *Mémoires de l'Institut Géologique de l'Université de Louvain.* **18**. 125-252.

Hbirkou, C., Pätzold, S., Mahlein, A.-K. & Welp, G. 2012. Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale. *Geoderma* **175-176**, 21–28.

Kempen, B., Brus, D.J., & Stoorvogel, J.J. 2011. Three-dimensional mapping of soil organic matter content using soil type-specific depth functions. *Geoderma* **162**, 107–123.

Liebens, J., & VanMolle, M. 2003. Influence of estimation procedure on soil organic carbon stock assessment in Flanders, Belgium. *Soil Use and Management* **19**, 364–371.

Liu, S. 2003. Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition. *Global Biogeochemical Cycles* **17**, 1074.

Malone, B.P., McBratney, A.B., Minasny, B., & Laslett, G.M. 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* **154**, 138–152.

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.

Milne, E., Al-Adamat, R., Batjes, N.H., Bernoux, M., Bhattacharyya, T., Cerri, C.C., et al. 2007. National and sub-national assessments of soil organic carbon stocks and changes: The GEFSOC modelling system. *Agriculture, Ecosystems & Environment* **122**, 3–12.

Minasny, B., McBratney, A.B., Mendonça-Santos, M.L., Odeh, I.O.A. & Guyon, B. 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. *Aust. J. Soil Res.* **44**, 233–244.

Moore, I.D., Gessler, P.E., Nielsen, G.A., & Peterson, G.A. 1993. Soil Attribute Prediction Using Terrain Analysis. *Soil Sci. Soc. Am. J.* **57**, 443–452.

Myers, D.B., Kitchen, N.R., Sudduth, K.A., Miles, R.J., Sadler, E.J., & Grunwald, S. 2011. Peak functions for modeling high resolution soil profile data. *Geoderma* **166**, 74–83.

Ponce-Hernandez, R., Marriott, F.H.C., & Beckett, P.H.T. 1986. An improved method for reconstructing a soil profile from analyses of a small number of samples. *Journal of Soil Science* **37**, 455–467.

Quinton, J.N., Govers, G., Van Oost, K. & Bardgett, R.D. 2010. The impact of agricultural soil erosion on biogeochemical cycling. *Nature Geoscience* **3**, 311–314.

Rosenbloom, N.A., Harden, J.W., Neff, J.C., & Schimel, D.S. 2006. Geomorphic control of landscape carbon accumulation. *J. Geophys. Res.* **111**(G01004).

Schwanghart, W. & Kuhn, N.J. 2010. TopoToolbox: A set of Matlab functions for topographic analysis. *Environmental Modelling & Software* **25**, 770–781.

Sleutel, S., Neve, S., & Hofman, G. 2007a. Assessing causes of recent organic carbon losses from cropland soils by means of regional-scaled input balances for the case of Flanders (Belgium). *Nutrient Cycling in Agroecosystems* **78**, 265–278.

Sleutel, S., De Neve, S., Singier, B., & Hofman, G. 2007b. Quantification of organic carbon in soils: A comparison of methodologies and assessment of the carbon content of organic matter. *Communications in Soil Science & Plant Analysis*, ***38***(19-20), 2647–2657.

Stevens, A., Udelhoven, T., Denis, A., Tychon, B., Lioy, R., Hoffmann, L. & van Wesemael, B. 2010. Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma* **158**, 32–45.

Stevens, A., van Wesemael, B., Vandenschrick, G., Touré S. & Tychon, B. 2006. Detection of carbon stock change in agricultural soils using spectroscopic techniques. *Soil Sci. Soc. Am. J.* **70**, 844–850.

Tan, Z.X., Lal, R., Smeck, N.E., & Calhoun, F.G. 2004. Relationships between surface soil organic carbon pool and site variables. *Geoderma* **121**, 187–195.

Temme, A.J. a. M., Peeters, I., Buis, E., Veldkamp, A., & Govers, G. 2011. Comparing landscape evolution models with quantitative field data at the millennial time scale in the Belgian loess belt. *Earth Surface Processes and Landforms* **36**, 1300–1312.

Trumbore, S.E. & Czimczik, C.I. 2008. An Uncertain Future for Soil Carbon. *Science* **321**, 1455–1456.

Van De Vreken, P., Beckers, V., Jacxsens, P., Van Meirvenne, M. & Van Orshoven, J. 2011. *Historiek van de bodemdatabank AARDEWERK Vlaanderen-2010.* In opdracht van de Vlaamse Overheid, Departement Leefmilieu, Natuur en Energie, Afdeling Land en Bodembescherming, Ondergrond, Natuurlijke Rijkdommen (ALBON).

Van Hemelryck, H., Govers, G., Van Oost, K., & Merckx, R. 2011. Evaluating the impact of soil redistribution on the in situ mineralization of soil organic carbon. *Earth Surface Processes and Landforms* **36**, 427–438.

Van Oost, K., Govers, G., Quine, T.A., Heckrath, G., Olesen, J.E., De Gryze, S. et al. 2005. Landscape-scale modeling of carbon cycling under the impact of soil redistribution: the role of tillage erosion. *Global Biogeochem. Cycles* **19**(4), 1-13.

Van Oost, K. V., Verstraeten, G., Doetterl, S., Notebaert, B., Wiaux, F., Broothaerts, N., & Six, J. (2012). Legacy of human-induced C erosion and burial on soil–atmosphere C exchange. *Proceedings of the National Academy of Sciences*, **109**(47), 19492-19497.

Van Orshoven, J., Maes, J., Vereecken, H., Feyen, J., & Dudal, R. 1988. A structured database of Belgian soil profile data. *Pedologie: bulletin van de Belgische bodemkundige vereniging* **38**, 191–206.

van Wesemael, B., Paustian, K., Andrén, O., Cerri, C.E.P., Dodd, M., Etchevers, J. et al. 2011. How can soil monitoring networks be used to improve predictions of organic carbon pool dynamics and CO 2 fluxes in agricultural soils? *Plant Soil* **338**, 247–259.

van Wesemael, B., Paustian, K., Meersmans, J., Goidts, E., Barancikova, G.. & Easter, M. 2010. Agricultural management explains historic changes in regional soil carbon stocks. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14926–14930.

VandenBygaart, A.J., Gregorich, E.G., Angers, D.A. & McConkey, B.G. 2007. Assessment of the lateral and vertical variability of soil organic carbon. *Can. J. Soil Sci.* **87**, 433–444.

Vandenbygaart, A.J., Kroetsch, D., Gregorich, E.G. & Lobb, D. 2012. Soil C erosion and burial in cropland. *Global Change Biol.* **18**, 1441–1452.

Vanwalleghem, T., Poesen, J., McBratney, A., & Deckers, J. 2010. Spatial variability of soil horizon depth in natural loess-derived soils. *Geoderma* **157**, 37–45.

Vasques, G.M., Grunwald, S., Comerford, N.B., & Sickman, J.O. 2010. Regional modelling of soil carbon at multiple depths within a subtropical watershed. *Geoderma* **156**, 326–336.

Viaud, V., Angers, D.A., & Walter, C. 2010. Toward Landscape-Scale Modeling of Soil Organic Matter Dynamics in Agroecosystems. *Soil Sci. Soc. Am. J.* **74**, 1847–1860.

Walkley, A., Black., I., 1934. An examination of the Degtjareff method for determining organic carbon in soils: Effect of variations in digestion conditions and of inorganic soil constituents. Soil Science 63, 251–263.

Wickham, H. 2009. *ggplot2: elegant graphics for data analysis.* Springer. New York.

Weiss, A.D. 2001. *Topographic Position and Landforms Analysis*. Poster. The Nature Conservancy.

Captions for tables

Table 1: Predefined intervals for the horizon depth and thickness as coded in the legacy dataset AW93.

Table 2: Statistics for the grouping of the SOC layer contents by combined drainage and profile development classes, for AW93 dataset.

Table 3: Statistics for the grouping of the SOC layer contents by combined texture and profile development classes, for AW93 dataset.

Table 4 : Statistics for the grouping of the SOC layer contents by wetness index classes, for R_AW dataset.

Table 5 : Statistics for the grouping of the SOC layer contents by slope gradient classes, for R_AW dataset.

Table 6: Pearson correlation and level of significance between terrain attributes and SOC contents by layer, for AW93 and R_AW datasets. Levels of significance are *** : $p < 0.001$, ** : $p < 0.01$ and * : $p < 0.05$.

Table 7: Model formulae and performance indices of the regression models for the three layers. Please refer to text for complete names of terrain attributes. The sign before each predictor represent the sign of its cofficient in the model. Intercept are also present for each layer. The predictors are ordered by decreasing significance, i.e. by increasing p-value. Besides, all predictors are significant with $p < 0.05$.

Captions for figures

Fig 1: Left picture: the Belgian loess belt region (orange polygon) is largely occupied by cropland (orange fill). Center picture: AW93 (blue points) and R_AW (red triangles) profiles sites are presented across the resampled sub-area of the study region.

Fig 2: Reconstruction of the horizon distribution for AW93 data and validation of the method using AW10 data.

Fig 3: The inaccuracy and imprecision due to the use of incomplete horizon distribution information is limited thanks to the use of a Monte Carlo based simulation procedure. Error is maximal around the expected bottom limit of the plough layer (15-30 cm).

Fig 4: Left side: representative SOC profile for soil property classes using AW93 profiles. Standard error on the mean is represented by a ribbon along the curve. The number of profile of each class are inside parentheses. Right side: distribution of SOC for each class by 30 cm thick soil layers. Inside each layer, a common letter indicate not significantly different means, (cf Material and methods).

Fig 5: Left side: representative SOC profile for terrain attribute classes using R_AW profiles. Standard error on the mean is represented by a ribbon along the curve. The number of profile of each class are inside parentheses. Right side: distribution of SOC for each class by 30 cm thick soil layers. Inside each layer, a common letter indicate not significantly different means (cf Material and methods). (GRAD: slope gradient ; TWI : topographic wetness index).

Fig 6: Correlations between SOC layers and terrain attributes are sensitive to a perturbation of the originally precize geolocations of the profiles of the R_AW dataset. The weak correlations observed with the AW93 dataset could thus be explained by an insufficient precision in the coordinates. (GRAD: slope gradient ; TWI : topographic wetness index).

Fig 7: For each layers, result of the 10-fold cross validation of R_AW (left column) and mapping of the predicted SOC stocks inside the fields of the plotting area (right column)

Tables

Table 1

Horizon depth /cm	Horizon thickness /cm
0 - 5	0 - 2
6 - 10	3 - 5
11 - 20	6 - 10
21 - 40	11 - 20
41 - 60	21 - 30
61 - 80	31 - 50
81 - 100	51 - 100
101 - 150	> 100
>150	

Table 2 :

	well drained colluvial		poorly drained colluvial		well drained not colluvial		poorly drained not colluvial		Anova F-test
	mean	std	mean	std	mean	std	mean	std	p
$SOC_{0-30\ cm}$ /g C kg^{-1}	12.63 ± 0.18	3.33	11.92 ± 0.26	3.29	12.16 ± 0.11	3.50	10.70 ± 0.13	3.47	$< 10^{-15}$
$SOC_{30-60\ cm}$ /g C kg^{-1}	4.68 ± 0.13	2.47	4.08 ± 0.16	2.05	3.48 ± 0.06	1.91	3.02 ± 0.07	1.75	$< 10^{-15}$
$SOC_{60-90\ cm}$ /g C kg^{-1}	3.40 ± 0.10	1.9	2.68 ± 0.11	1.38	2.27 ± 0.05	1.46	1.65 ± 0.05	1.32	$< 10^{-15}$

Table 3:

	silt loam colluvial		(heavy) sandy loam colluvial		silt loam not colluvial		(heavy) sandy loam not colluvial		light sandy loam not colluvial		Anova F-test
	mean	std	mean	std	mean	std	mean	std	mean	std	p
$SOC_{0-30\ cm}$ /g C kg^{-1}	12.62 ± 0.16	3.28	11.7 ± 0.48	3.67	12.51 ± 0.11	3.38	11.09 ± 0.18	3.41	9.07 ± 0.17	2.92	$< 10^{-15}$
$SOC_{30-60\ cm}$ /g C kg^{-1}	4.77 ± 0.11	2.41	4.10 ± 0.28	2.12	3.52 ± 0.06	1.93	3.06 ± 0.08	1.49	2.77 ± 0.11	1.87	$< 10^{-15}$
$SOC_{60-90\ cm}$ /g C kg^{-1}	3.28 ± 0.08	1.78	2.70 ± 0.20	1.50	2.32 ± 0.05	1.45	1.85 ± 0.08	1.41	1.26 ± 0.07	1.22	$< 10^{-15}$

Table 4:

	wetness index I		wetness index II		wetness index III		Anova F-test
	mean	std	mean	std	mean	std	p
$SOC_{0-30\ cm}$ /g C kg^{-1}	9.58 ± 0.31	1.68	9.82 ± 0.18	0.94	10.44 ± 0.22	1.20	0.031
$SOC_{30-60\ cm}$ /g C kg^{-1}	2.96 ± 0.16	0.86	3.23 ± 0.19	1.02	4.32 ± 0.32	1.73	5.9×10^{-5}
$SOC_{60-90\ cm}$ /g C kg^{-1}	1.96 ± 0.07	0.37	2.14 ± 0.12	0.63	2.73 ± 0.17	0.93	2.2×10^{-5}

Table 5

	slope gradient I		slope gradient II		slope gradient III		Anova F-test
	mean	std	mean	std	mean	std	p
$SOC_{0-30\ cm}$ /g C kg^{-1}	10.25 ± 0.23	1.18	9.96 ± 0.23	1.26	9.65 ± 0.27	0.93	0.24
$SOC_{30-60\ cm}$ /g C kg^{-1}	3.76 ± 0.22	1.14	3.86 ± 0.33	1.75	2.99 ± 0.18	1.03	0.012
$SOC_{60-90\ cm}$ /g C kg^{-1}	2.56 ± 0.18	0.93	2.45 ± 0.14	0.74	1.90 ± 0.07	0.37	0.00021

Table 6:

	AW93			R_AW		
	SOC0-30	SOC30-60	SOC60-90	SOC0-30	SOC30-60	SOC60-90
SOC0-30	/	0.50***	0.43***	/	0.61***	0.43***
SOC30-60	/	/	0.45***	/	/	0.65***
ELEV	0.28***	0.11***	0.17***	-0.06	-0.05	-0.05
GRAD	0.05*	0.00	0.04	-0.34***	-0.26**	-0.37***
CURV	0.04	0.05*	0.05*	0.31***	0.45***	0.37***
SOUTH	-0.02	-0.03	-0.02	-0.27**	-0.10	-0.05
TPI32	-0.03	-0.06**	-0.07***	-0.28**	-0.45***	-0.38***
TPI128	-0.03	-0.14***	-0.14***	-0.19*	-0.42***	-0.38***
TPI512	-0.01	-0.14***	-0.13***	-0.02	-0.28**	-0.28**
TWI	0.00	0.06**	0.08***	0.40***	0.50***	0.55***
STR	0.09***	0.08***	0.14***	0.17	0.34***	0.27**

Table 7:

Model formula	RMSEP / g C kg^{-1}	ME / g C kg^{-1}	R^2
$SOC_{0-30\ cm} \sim TWI - SOUTH + (ELEV \times TPI_{512\ m})$	1.28	0.01	0.21
$SOC_{30-60\ cm} \sim -(TPI_{32\ m} \times TWI) - (ELEV \times GRAD) + TPI_{32\ m}$	1.28	<0.01	0.25
$SOC_{60-90\ cm} \sim -TPI_{128\ m} - GRAD + (GRAD \times TPI_{128\ m})$	0.67	0.01	0.33

Figures

Fig 1:

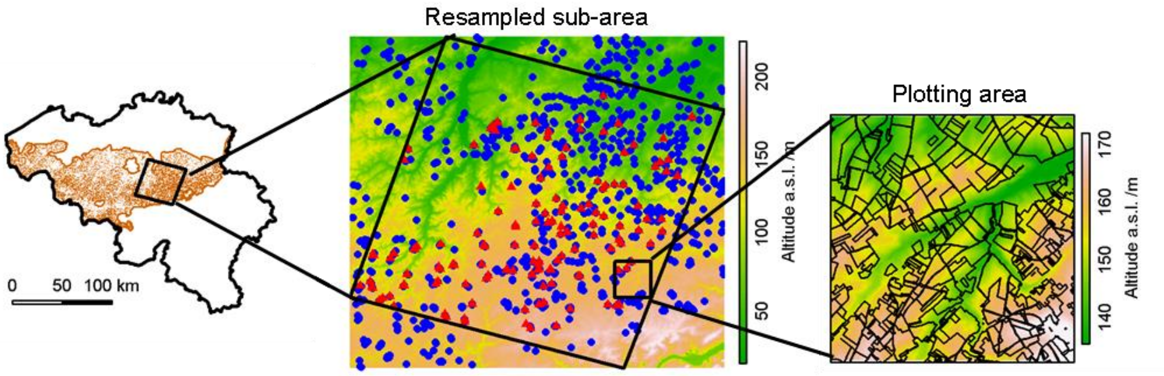


Fig 2:

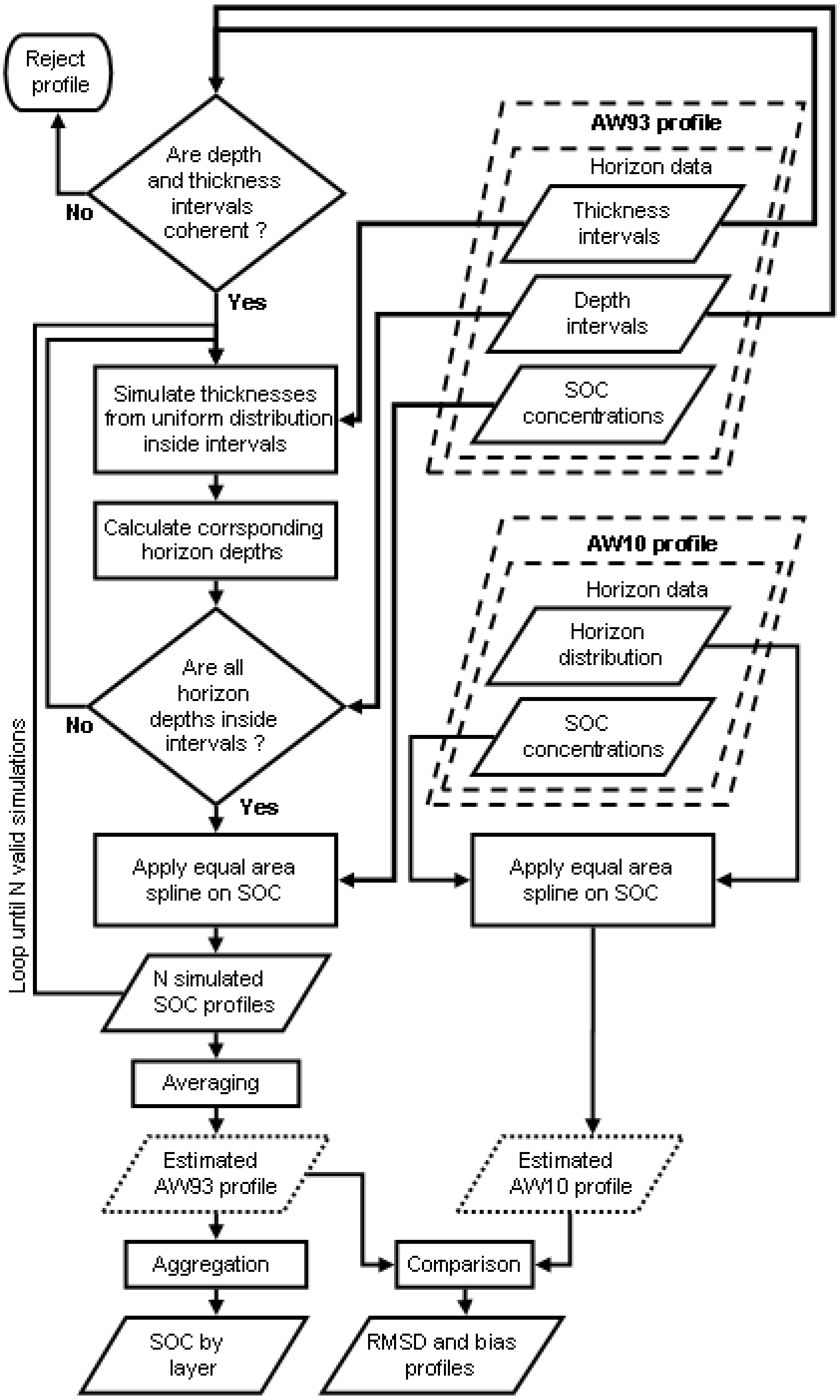
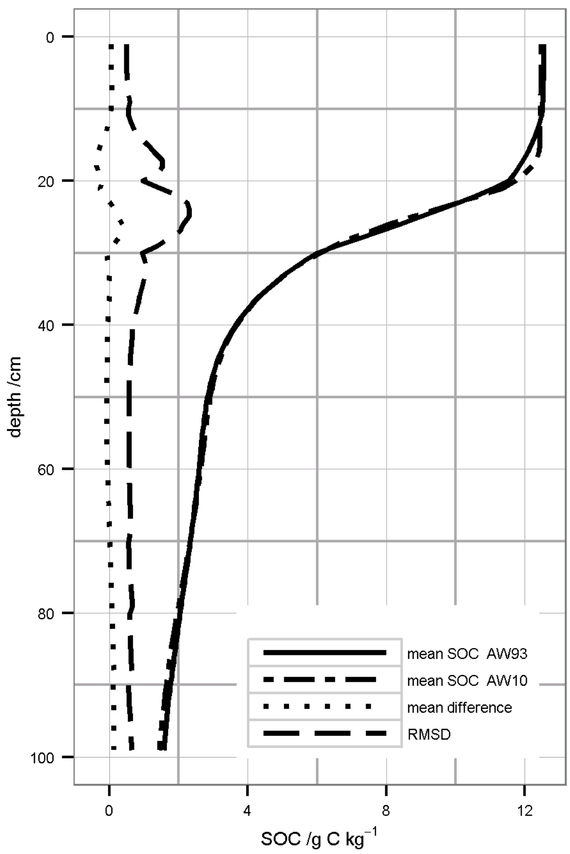

Reject profile
Are depth and thickness intervals coherent ?
No
Yes
AW93 profile
Horizon data
Thickness intervals
Depth intervals
SOC concentrations
Simulate thicknesses from uniform distribution inside intervals
Calculate corrsponding horizon depths
Are all horizon depths inside intervals ?
No
Yes
Loop until N valid simulations
AW10 profile
Horizon data
Horizon distribution
SOC concentrations
Apply equal area spline on SOC
Apply equal area spline on SOC
N simulated SOC profiles
Averaging
Estimated AW93 profile
Estimated AW10 profile
Aggregation
Comparison
SOC by layer
RMSD and bias profiles

Fig 3 :



depth /cm
0
20
40
60
80
100
0
4
8
12
SOC /g C kg^{-1}
mean SOC AW93
mean SOC AW10
mean difference
RMSD

Fig 4:

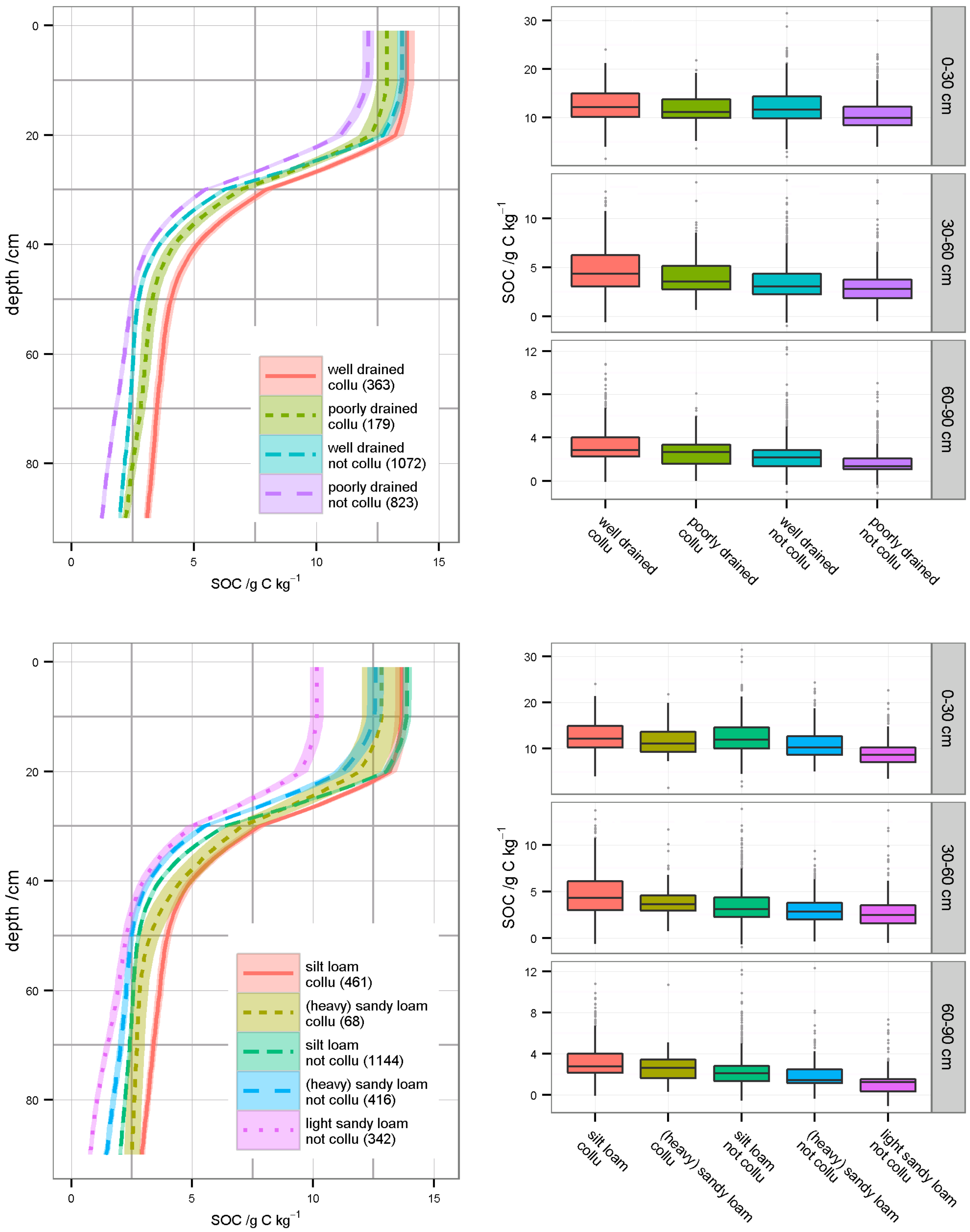

depth /cm
SOC /g C kg^{-1}
well drained collu (363)
poorly drained collu (179)
well drained not collu (1072)
poorly drained not collu (823)
0-30 cm
30-60 cm
60-90 cm
well drained collu
poorly drained collu
well drained not collu
poorly drained not collu
silt loam collu (461)
(heavy) sandy loam collu (68)
silt loam not collu (1144)
(heavy) sandy loam not collu (416)
light sandy loam not collu (342)
silt loam collu
(heavy) sandy loam collu
silt loam not collu
(heavy) sandy loam not collu
light sandy loam not collu

Fig 5:

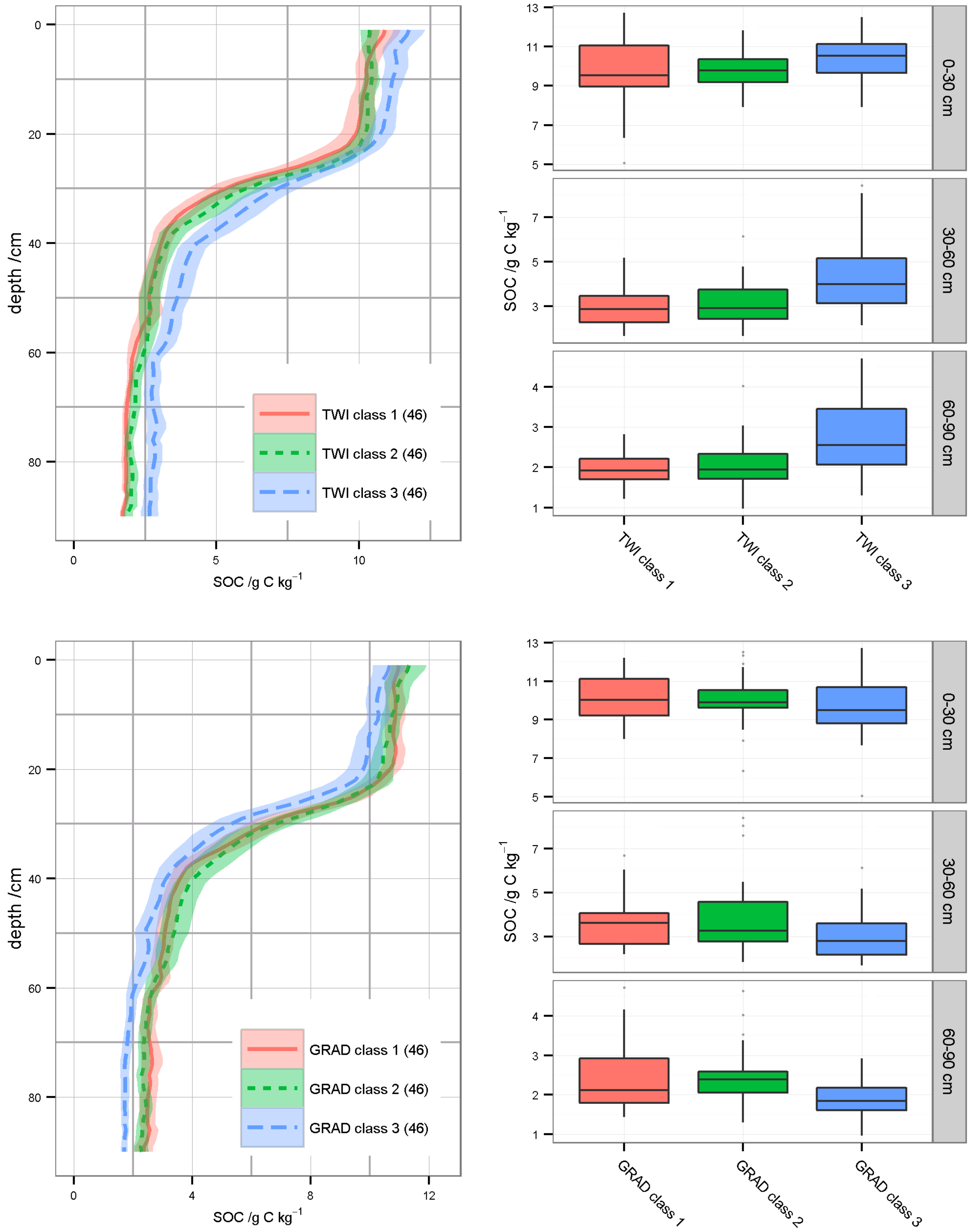

depth /cm
SOC /g C kg^{-1}
TWI class 1 (46)
TWI class 2 (46)
TWI class 3 (46)
0-30 cm
30-60 cm
60-90 cm
TWI class 1
TWI class 2
TWI class 3
GRAD class 1 (46)
GRAD class 2 (46)
GRAD class 3 (46)
GRAD class 1
GRAD class 2
GRAD class 3

Fig 6:

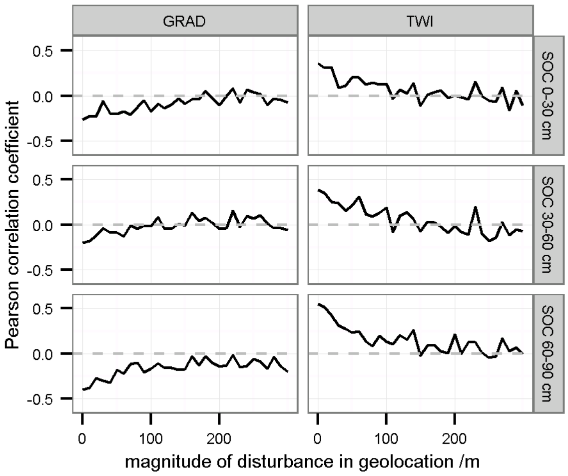
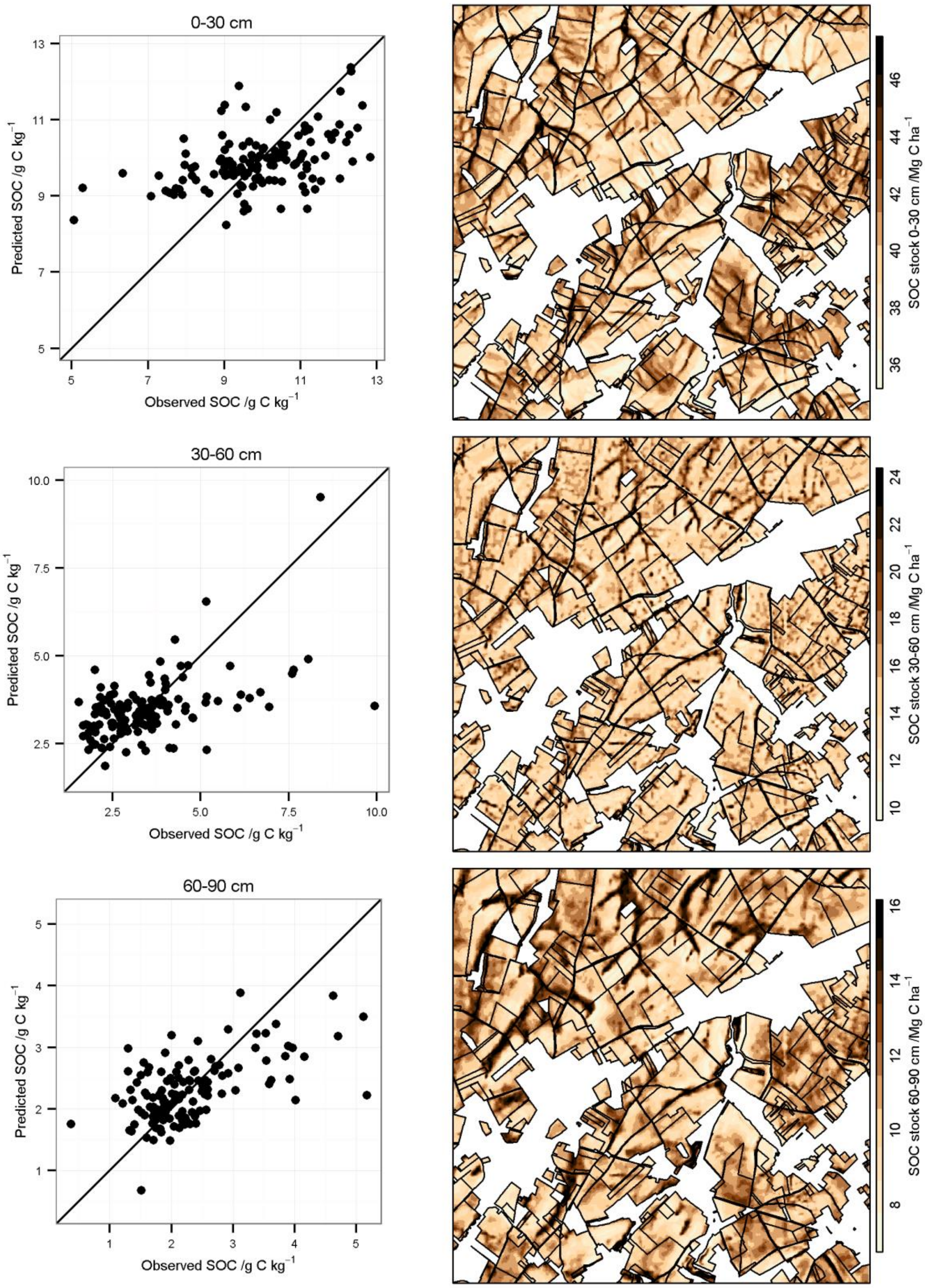

GRAD
TWI
SOC 0-30 cm
SOC 30-60 cm
SOC 60-90 cm
Pearson correlation coefficient
0.5
0.0
-0.5
0
100
200
magnitude of disturbance in geolocation /m

Fig 7.


0-30 cm
Predicted SOC /g C kg^{-1}
Observed SOC /g C kg^{-1}
SOC stock 0-30 cm /Mg C ha^{-1}
30-60 cm
Predicted SOC /g C kg^{-1}
Observed SOC /g C kg^{-1}
SOC stock 30-60 cm /Mg C ha^{-1}
60-90 cm
Predicted SOC /g C kg^{-1}
Observed SOC /g C kg^{-1}
SOC stock 60-90 cm /Mg C ha^{-1}