Geoderma Regional xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Geoderma Regional

journal homepage: www.elsevier.com/locate/geodrs

Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southern Belgium, Wallonia)

Caroline Chartin [a,*], Antoine Stevens [a], Esther Goidts [b], Inken Krüger [c], Monique Carnol [c], Bas van Wesemael [a]

[a] George Lemaître Centre for Earth and Climate Research, Earth & Life Institute, Université catholique de Louvain, Place Louis Pasteur 3, 1348 Louvain-la-Neuve, Belgium
[b] Soil Protection Direction, Direction générale Agriculture, Natural Resources and Environment, Public Administration of Wallonia, B-5100 Jambes, Belgium
[c] Laboratoire d'Ecologie Végétale et Microbienne, Université de Liège, Botanique B22, Quartier Vallée 1, Chemin de la Vallée 4, 4000 Liège, Belgium

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 1 December 2016
Accepted 21 December 2016
Available online xxxx

Keywords:
Soil Organic Carbon stocks
Digital Soil Mapping
Soil Monitoring Network
Luvisol
Cambisol

ABSTRACT

The quantification and the spatialisation of reliable SOC stocks (Mg C ha^{-1}) and total stock (Tg C) baselines and associated uncertainties are fundamental to detect the gains or losses in SOC, and to locate sensitive areas with low SOC levels. Here, we aim to both quantify and spatialize SOC stocks at regional scale (southern Belgium) based on data from one non-design-based or model-based sampling scheme. To this end, we developed a computation procedure based on Digital Soil Mapping techniques and stochastic simulations (Monte-Carlo) allowing the estimation of multiple (here, 10,000) independent spatialized datasets. The computation of the prediction uncertainty accounts for the errors associated to both the estimations of i) SOC stocks and ii) parameters of the spatial model. Based on these 10,000 individuals, median SOC stocks and 90% prediction intervals were computed for each pixel, as well as total SOC stocks and their 90% prediction intervals for selected sub-areas and for the entire study area. Hence, a Generalised Additive Model (GAM) explaining 69.3% of the SOC stock variance was calibrated and then validated ($R^2 = 0.64$). The model overestimated low SOC stock (below 50 Mg C ha^{-1}) and underestimated high SOC stock (especially those above 100 Mg C kg^{-1}). A positive gradient of SOC stock occurred from the northwest to the center of Wallonia with a slight decrease on the southernmost part, correlating to the evolution of precipitation and temperature (along with elevation) and dominant land use. At the catchment scale higher SOC stocks were predicted on valley bottoms, especially for poorly drained soils under grassland. Mean predicted SOC stocks for cropland and grassland in Wallonia were of 26.58 Tg C (SD 1.52) and 43.30 Tg C (2.93), respectively. The procedure developed here allowed to predict realistic spatial patterns of SOC stocks all over agricultural lands of southern Belgium and to produce reliable statistics of total SOC stocks for each of the 20 combinations of land use/agricultural regions of Wallonia. This procedure appears useful to produce soil maps as policy tools in conducting sustainable management at regional and national scales, and to compute statistics which comply with specific requirements of reporting activities.

© 2016 Published by Elsevier B.V.

1. Introduction

Soil Organic Carbon (SOC) is a key element for soil quality and fertility, as it improves aggregate stability, nutrient availability and water retention capacity (Hallett et al., 2012; Robinson et al., 2012). Thus, SOC contributes to many ecosystem services, such as food production, maintenance of soil structure and the prevention of soil erosion (FAO, 2015). The sequestration of CO_2 in the form of SOC in particular through agricultural practices that increase organic matter inputs and reduce their decomposition is considered to be a potential offset to current anthropogenic CO_2 emissions (Paustian et al., 2016).

The quantification and the spatialisation of reliable SOC stocks (Mg C ha^{-1}) and total stocks (Tg C) and associated uncertainties appear therefore fundamental to: i/ detect the trends of gains or losses in SOC linked to recent and future national, European or global policy decisions related to soil quality and carbon sequestration, and, ii/ locate sensitive areas characterized with low SOC levels for prioritizing conservation practices.

In order to monitor SOC stocks at regional or national scales, a Soil Monitoring Network (SMN) has to be developed as a set of sites or areas where changes in soil characteristics can be observed through regular assessment of different soil properties (Morvan et al., 2008). An important decision related to the design of a SMN is the choice of the sampling locations that can be executed by probability sampling (random sampling with known inclusion probabilities) or by non-

* Corresponding author at: TECLIM, Université catholique de Louvain, Place Louis Pasteur 3, 1348 Louvain-la-Neuve, Belgium.
E-mail addresses: caroline.chartin@uclouvain.be (C. Chartin), esther.goidts@spw.wallonie.be (E. Goidts), inken.kruger@ulg.ac.be (I. Krüger), m.carnol@ulg.ac.be (M. Carnol), bas.vanwesemael@uclouvain.be (B. van Wesemael).

http://dx.doi.org/10.1016/j.geodrs.2016.12.006
2352-0094/© 2016 Published by Elsevier B.V.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

ARTICLE IN PRESS

probability sampling. Hence, two main types of soil sampling strategies exist, i.e. the design-based strategy and the model-based strategy (Brus and de Gruijter, 1993; de Gruijter et al., 2006). The design-based strategy is a classical sampling theory based on probability sampling: i.e. random sampling within the area of interest. The population is then assumed independent and identically distributed. For the model-based strategy, also called geostatistical modelling, randomness is introduced via the model, which is a stochastic model and therefore contains a random error term. The inference is based on assuming the validity of the stochastic model.

Hence, design-based sampling is a combination of probability sampling and design-based inference, whereas model-based sampling is a combination of purposive sampling and interpolation (de Gruijter et al., 2006). The choice between these two sampling strategies relies finally on the purpose of the SMN itself, i.e., to quantify or to spatialize the target variable. Design-based sampling is more suitable to deal with calculation of global mean, quantiles, and standard deviation, i.e. the frequency distribution of the target variable. The model-based sampling is more suitable to predict values at unsampled locations, and so to map the target variable in the whole study area along with its confidence interval through the associated stochastic model parameters (Wang et al., 2012).

In practice, the means (e.g., funding, time, and labour) needed to develop SMNs are far from being ideal for the proper development of these types of soil sampling networks, i.e. design-based and model-based. A random sampling (design-based sampling) is difficult to achieve because of the impossibility to access some sites, and/or because the environment (forest, buildings…) can occasionally hamper a precise localization of random selected sites by GPS devices (Kidd et al., 2015). Concerning the realization of a model-based sampling, it requires the prior knowledge of some information about the spatial variability of the target parameter, which often does not exist or is too scarce. Simpler approaches for a spatial sampling to be used in a model-based strategy can be designed, such as a regular grid or spatial coverage samples (de Gruijter et al., 2006). These are often more time-consuming and costly, as their density are higher due to the lack of a priori knowledge of the spatial variability or due to the existence of conflicting trends in the spatial variability of different soil parameters.

Many western European countries have been developing their own SMN within the last decades such as Netherlands, France and Denmark (Bloem et al., 2005; Jolivet et al., 2006; Østergaard, 1989), mostly choosing one particular type of SMN amongst design-based or model-based ones, or capitalizing on a legacy database such as in southern Belgium (Goidts et al., 2009). The Soil Monitoring Network 'CARBOSOL' was developed to survey the quality of agricultural soil in Wallonia by quantifying SOC stocks and their evolution (Goidts and van Wesemael, 2007). A reasonable number of locations from a previous national sampling campaign from 1949 to 1972 (used to map soils in Belgium) were chosen according to their belonging to 15 spatial LandScape Units (LSUs) and resampled from 2005 to 2009. The LSUs are characterized by homogeneous climate conditions, soil type (texture, drainage and stoniness), land use (cropland or grassland) and main management practices (inherent to the agricultural regions). Two main limitations arise for the quantification of the total SOC stock: the 15 LSUs considered here all together covered only 63% of the Walloon agricultural areas, and the legacy database on which the SMN is based was created as a convenience design sampling. Hence, SOC stocks were only computed on a limited area in Wallonia based on a non-independent sampling data invalidating the statistical inference.

The first attempts to map SOC stock in Belgium consisted of assigning the mean and associated uncertainties observed in each LSU to the corresponding spatial polygons (Goidts et al., 2009; Lettens et al., 2005b). Thus, this process of mapping lacked statistical inference and hampered a proper understanding of SOC stock spatial variability. To produce more realistic results, a Digital Soil Mapping (DSM) technique was first applied to Flanders (northern Belgium; Meersmans et al., 2008) and later, to the entire Belgian territory for two different periods (Meersmans et al., 2011). The DSM technique consisted in fitting a multiple regression to observed data using a set of spatialized layers of covariates, i.e. environmental factors controlling soil formation and explaining the spatial variability of the target variable (Lagacherie, 2008; McBratney et al., 2003). Although Meersmans et al. (2011) used a multiple linear regression including some higher order interaction terms, this particular model structure may prevent the representation of potentially more complex non-linear relations existing between SOC and environmental covariates. These authors also applied a PTF (PedoTransfer Function) to estimate soil bulk density, increasing considerably the uncertainties (2.6%–29.0%) for SOC stock calculation (Meersmans et al., 2011).

Here, we aim to both quantify and spatialize SOC stocks at regional scale (southern Belgium) based on data from a non-design-based or model-based SMN. To this end, we developed a computation procedure based on Digital Soil Mapping techniques and stochastic statistic approaches allowing the estimation of multiple independent spatialized datasets and their uncertainties associated to both SOC stock calculation and spatial modelling. Hence, reliable statistics, mean or median SOC stock and confidence interval, can be computed for each pixel, as well as total SOC stocks and their confidence intervals for selected subareas or the entire study area.

2. Materials and methods

2.1. Study area

The study area covers southern Belgium, also called Wallonia, which covers about 16,900 km^2. From the northwest to the southeast, there is an increase in precipitation (from 800 to 1200 mm; Fig. 2A) along with elevation (from 180 to 690 m; Fig. 1A) and a decrease of mean annual temperature (from 10 to 8 °C; Fig. 2B). In the same direction, a shift occurs from deep sandy loam and silty soils to shallow silt loam and stony soils, from Haplic Luvisol to Distric Cambisol mainly (with Fluvisol on valley bottoms), along with a shift from intensive arable agriculture to more extensive cattle breeding (Fig. 1B; for further details, see Goidts and van Wesemael, 2007).

Ten agricultural regions (Fig. 1) are defined on homogeneous characteristics in terms of agricultural practices (types of crops and/or livestock) and crop yields, both directly linked to soil types (e.g., texture, fertility, stoniness, and drainage) and climatic characteristics (Lettens et al., 2005a). Agricultural regions are now used as references in most of the recent reporting activities linked to agricultural activities and diagnosis of soil quality evolution (such as National Inventory Report – NIR – commanded by EU or the state of the environment report published by the Public Administration of Wallonia).

2.2. Soil data

2.2.1. SOC monitoring network in Wallonia

Between 1949 and 1965, the National Soil Survey (NSS) resulted in the description and analysis of about 13,000 soil profiles to create the Belgian soil map (De Leenheer et al., 1968). The resulting digitized legacy database ("Aardewerk", Van Orshoven et al., 1988) contains precious information per horizon (up to a depth of about 1.5 m) for each soil profile, e.g., texture, SOC content and rock fragment content (RF). Other information such as "soil series", land use, and coordinates of the profile, is also available.

Between 2005 and 2014, 592 locations from the original NSS sampling network were resampled to form the CARBOSOL SMN dedicated to agricultural soils (croplands and grasslands). Firstly, 434 locations were sampled from 2005 to 2009 within 15 homogeneous LSUs (Goidts et al., 2009) to study SOC stock changes between 1955 and 2005, and then completed in 2012–2014 by 158 locations sampled

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

ARTICLE IN PRESS

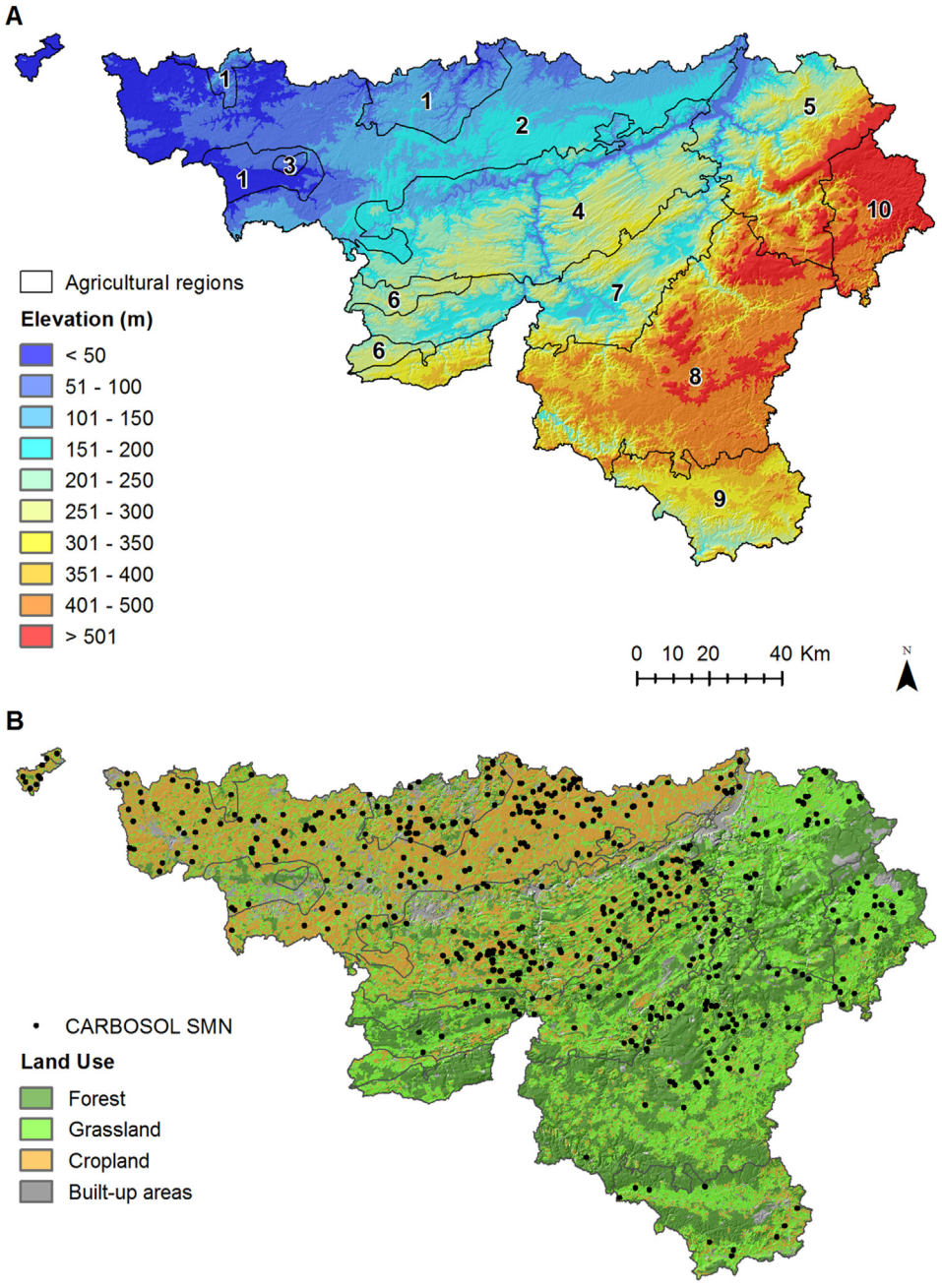

Fig. 1. Elevation (A) and main land uses (http://cartopro3.wallonie.be/) with location of sample sites (B) in southern Belgium (Wallonia) with delineation of the 10 official agricultural regions (1: Sandy-loam, 2: Loam, 3: Campine Hennuyère, 4: Condroz, 5: Herbagère Liège, 6: Herbagère Fagnes, 7: Famenne, 8: Ardenne, 9: Jurassic, and 10: Haute Ardenne).

within 30 additional LSUs to improve the spatial coverage of the network. As the CARBOSOL SMN has not been originally designed to assess SOC stocks by agricultural regions and main agricultural land use (i.e., cropland and grassland), the spread of the samples through the territory is not optimal to produce significant and reliable statistics of SOC stocks for the 20 existing agricultural region/land use combinations (Table 1).

2.2.2. Sampling method

At each location, composite samples of the 0–30 cm topsoil layer were taken from 5 points sampled by auger and located within a circle of 4 m radius centered on the known geographical coordinates. When the first 30 cm were composed of different horizons, the sampling was done by horizon. The samples were air-dried, gently crushed with a mortar and pestle, sieved at 2 mm and stored in plastic bags at room temperature. Fine earth (Ø < 2 mm) and rock fragments (Ø > 2 mm) were both weighed. Three intact soil cores of 100 cm^3 (Øcore = 5.3 cm) were taken in the middle of each sampled layer to measure the corresponding bulk density.

2.2.3. Laboratory analyses and SOC stock calculation

SOC content of fine earth (Ø < 2 mm) was measured using the classic dichromate oxidation method of Walkley and Black (1934) (C_{wb}) for soils sampled between 2005 and 2009. This method is known as requiring the application of a correction factor to raw measures in order to correct for incomplete oxidation and estimate total organic carbon concentration. The total C content (TC) was measured in soils sampled since 2009 using a VarioMax CN dry combustion Analyzer (Elementar GmbH, Germany) and was taken as SOC content (C_{dc}) unless a HCl test showed the presence of carbonates. In which case, Inorganic Carbon (IC) content was then determined separately by a pressure method (Sherrod et al., 2002) and C_{dc} was taken as the difference between TC and IC.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

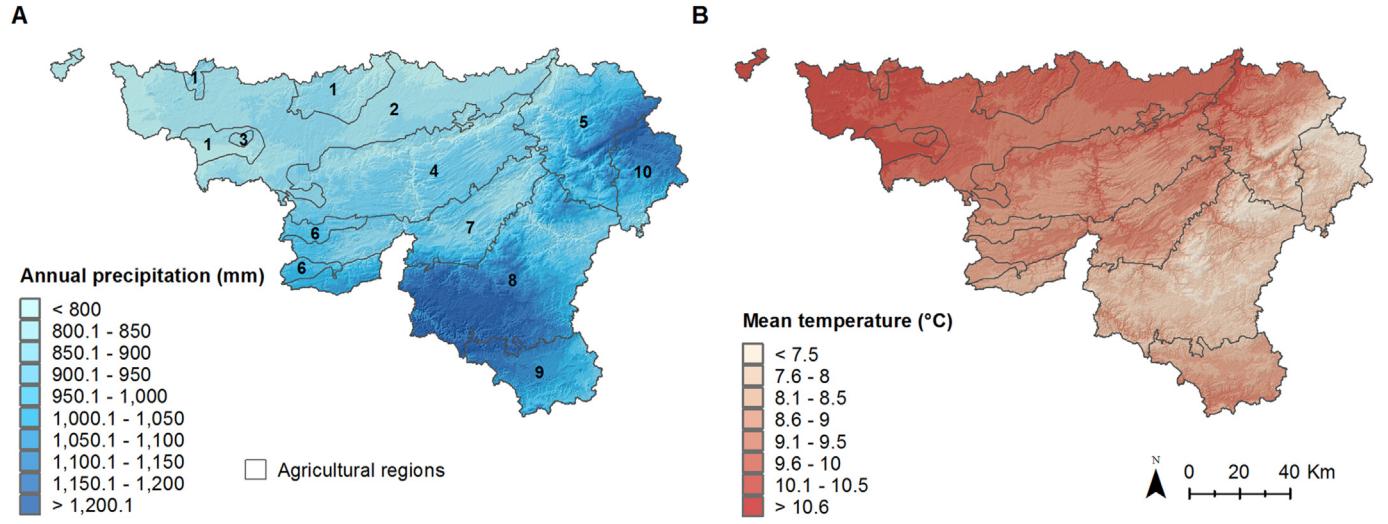

Fig. 2. Annual precipitation (mm) and mean temperature (°C) of southern Belgium (Wallonia) with delineation of the 10 official agricultural regions (see Fig. 1).

Amongst the 434 locations originally analysed for SOC content by the Walkley & Black method between 2005 and 2008, 100 were randomly selected and analysed by dry combustion and if required corrected for carbonates. These 100 measurements were used to determine a specific correction to apply to Walkey & Black raw data. To this end, a linear model was fitted between raw data obtained by the Walkley and Black method (C_{wb}) and by dry combustion analysis (C_{dc}), using lm function from the *stats* package in R (Eq. (1)):

$$C_{dc} = a * C_{wb} + b \quad (1)$$

where a and b are the constant parameters of the equation.

The bulk density of the fine earth (*BD*) was determined by drying and weighing undisturbed soil samples of 100 cm^3. The bulk density of the total soil was then corrected by using the weight of the rock fragments and their density (CRC, 1996). The SOC stock (SOC in Mg C ha^{-1}) has been calculated for the 0–30 cm soil depth according to Eq. (2):

$$SOC = \frac{d * C * [1 - Rv] * BD}{100} \quad (2)$$

where d is the depth (m), C is the content in organic carbon ($gC \cdot kg^{-1}$), Rv is the fraction of rock fragments by volume (−), and BD is the bulk density of the fine earth ($kg \cdot m^{-3}$).

Table 1
Number of soil samples from SMN CARBOSOL per landuse/agricultural region combinations in Wallonia. (Areas were computed using spatial layers of landuses and agricultural regions; see Section 2.3.)

Landuse	Agricultural region		% of total area per landuse	Number of samples
Cropland	Sandy-loam	(1)	9.3	59
	Loam	(2)	53.2	133
	Camipne Hennuyère	(3)	0.1	–
	Condroz	(4)	19.8	127
	Herbagère Liège	(5)	2.1	4
	Herbagère Fagnes	(6)	1.0	–
	Famenne	(7)	4.9	18
	Ardenne	(8)	5.6	23
	Jurassic	(9)	2.7	6
	Haute Ardenne	(10)	1.2	1
Cropland	All (~4745 km^2)		100	371
Grassland	Sandy-loam	(1)	4.6	–
	Loam	(2)	15.4	49
	Camipne Hennuyère	(3)	0.1	–
	Condroz	(4)	14.3	22
	Herbagère Liège	(5)	14.5	40
	Herbagère Fagnes	(6)	3.1	–
	Famenne	(7)	11.8	41
	Ardenne	(8)	22.4	42
	Jurassic	(9)	6.9	4
	Haute Ardenne	(10)	6.9	23
Grassland	All (~5710 km^2)		100	221

2.3. Environmental covariates considered in the DSM approach

The Digital Soil Mapping (DSM) approach used in this study fitted a statistical regression model between the soil property to predict (SOC stock in Mg C ha^{-1}) and independent environmental covariates at the same location. SOC stock values can be predicted at unsampled locations by applying the fitted model to the spatial continuous layers of covariates. A spatial resolution of 40 m was chosen as a compromise between the different original resolutions of covariates already continuous in space (varying from 20 to 90 m), the scale of the sampling sites (i.e., a circle of 4 m radius) and the limitation of the computation time. So, all the continuous variables used as environmental covariates in our mapping procedure were first rescaled via ArcGIS10 (ESRI, 2011) by bilinear interpolation, which determined the output value of a cell based on a weighted distance average of the four nearest input cell centers. Once the stack of layers was constituted, the values of these covariates were extracted at each of the 592 locations constituting the CARBOSOL SMN, then creating the dataset used in the computation procedure of SOC stocks and uncertainties. The environmental covariates considered in this study are discussed below.

2.3.1. Digital Elevation Model and derivates

A Digital Elevation Model (DEM) with a resolution of 20 m provided by the National Geographic Institute (NGI) of Belgium was used, from which different morphologic and hydrologic derivatives were computed in ArcGIS10 (ESRI, 2011): slope, aspect, profile, plan and general curvatures, Topographic Position Index (TPI; Jenness, 2006), Wetness Index (WI; Phillips, 1990), flow accumulation, upstream flow length and LS-factor as used in the RUSLE equation (Renard et al., 1991; Fig. 1A). Finally, we converted aspect into easting and northing which represent the degree to which aspect is close to the east and the north, respectively (Zar, 1999).

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

2.3.2. Climate data

Annual mean temperature and precipitation data were extracted from weather stations in Belgium and neighboring countries for the period 1971–2000. These data were supplied by the national meteorological institutes of Belgium (Koninklijk Meteorologisch Instituut, KMI), the Netherlands (Koninklijk Nederlands Meteorologisch Instuut, KNMI), Germany (Deutscher Wetterdienst, DWD), France (Meteo France) and Grand Duchy of Luxemburg (Observatory for Climate and Environment). Based on these data, climatic maps were created by modeling temperature and precipitation with altitude using thin-plate splines regression (Fig. 2). Using altitude as covariate for mapping climatic variables can improve predictions dramatically (Boer et al., 2001). The smoothing parameter is chosen automatically by generalized cross-validation. Here, elevation data were derived from the Shuttle Radar Topography Mission (SRTM) mission of the NASA (http://srtm.csi.cgiar.org) to cover Belgium and neighboring countries.

2.3.3. Soil parameters

Numerous soil parameters were measured during the NSS and compiled in the digitized database AardeWerk and used to produce the Digital Soil Map of Wallonia (DSMW). All the polygons of the DSMW corresponding to the main mineral soil associations were selected and converted to a raster format with a resolution of 40 m. Minimum and maximum ground water levels, i.e. the position of the winter and summer ground water table, were mapped as proposed by Meersmans et al. (2008). Four maps of soil texture were created using the data of 6129 historical soil profiles from the AardeWerk database (Fig. 3). We computed three maps corresponding to the classical texture classes (sand 50 μm–2 mm, silt 2–50 μm, clay <2 μm), and one map corresponding to the clay and fine silt class (i.e., particles <20 μm) generally assumed to be linked to the stable SOC fraction (Hassink, 1997). Only the first soil horizon was kept in the AardeWerk database to compute these four maps. We spatialized the observations using regression kriging (Pebesma, 2006). First, we computed the mean sand, silt, clay, and clay and fine silt content per main soil associations and assigned the resulting values to the corresponding spatial polygons. Then, we computed the residuals between the observations and the mean and kriged the residuals in space by regression kriging. Finally, the obtained spatial layers of residuals were added to the corresponding map of mean values.

2.3.4. Anthropogenic influences

Land cover data were extracted from the Carte d'Occupation du Sol de Wallonie (COSW), a land cover map. Land cover classes were reclassified into cropland, (permanent) grassland, forest and others (Fig. 1). The data is in polygon format and was therefore converted to raster with a 40 m resolution. We computed the mean annual input of organic carbon by organic amendments (Mg C ha^{-1} yr^{-1}), e.g. farmyard manure and slurry, for each municipality of Wallonia based on data provided by the National Institute of Statistics for the period 2007–2012 and the method proposed by Dendoncker et al. (2004). Resulting calculations were then applied to the corresponding polygons of the Walloon

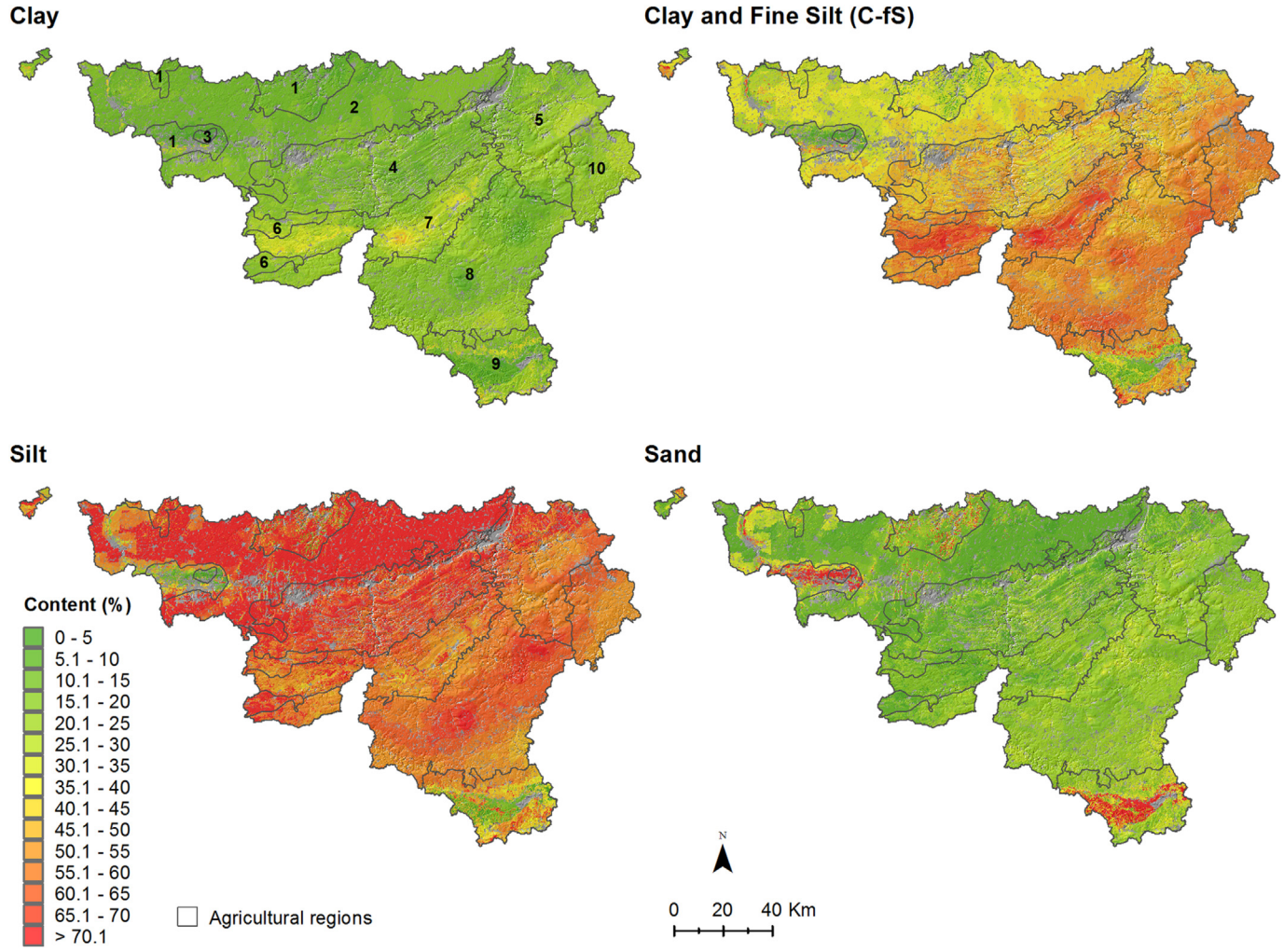

Fig. 3. Clay, clay and fine silt (C-fS), silt and sand contents (%) in topsoil of Southern Belgium (Wallonia) with delineation of the 10 official agricultural regions (see Fig. 1). Grey color corresponds to built up areas.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

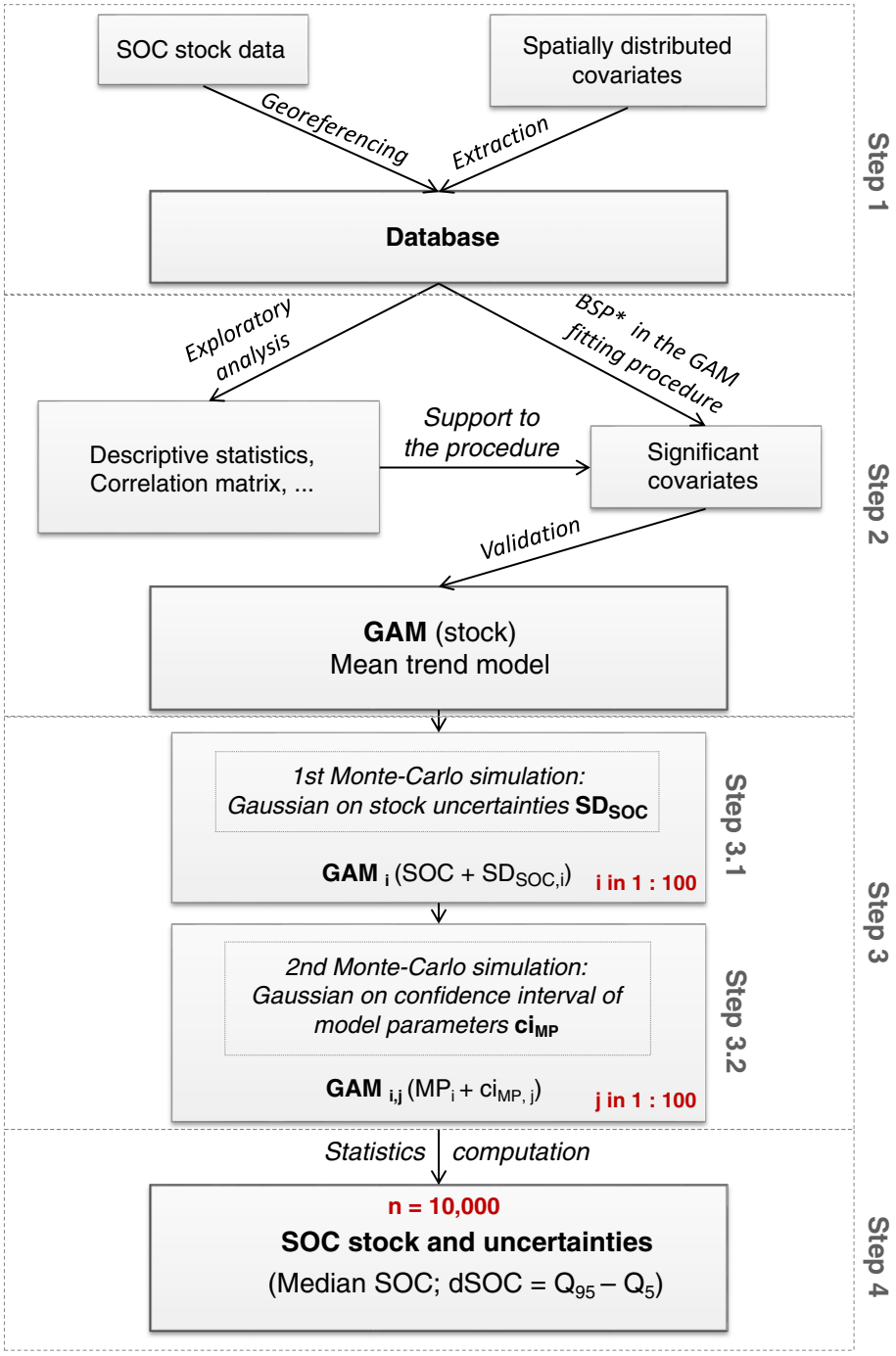

Fig. 4. Workflow of the consecutive procedures used to assess the SOC stocks and associated uncertainties for agricultural soils of southern Belgium.

municipalities and, transformed to raster format with a resolution of 40 m.

2.4. Simulation procedure of SOC stocks and confidence interval at the regional scale

After the creation of the database (step 1 in Fig. 4), two consecutive methods were applied to, first, fit the mean trend model linking SOC stock spatial variability to the most significant environmental covariates (step 2 in Fig. 4) and to, secondly, assess the frequency distribution of simulated SOC stocks (median and confidence interval) by taking into account uncertainties associated to SOC stock calculation and spatial modelling (step 3 in Fig. 4).

2.4.1. Generalized Additive Model

The initial dataset was split into a calibration (2/3) and a validation (1/3) set by Latin hypercube sampling stratified by SOC stock values and land use (Minasny and McBratney, 2006). Then, a Generalized Additive Model (GAM; Hastie and Tibshirani, 1986) was fitted on the calibration dataset. This regression technique is a generalization of linear regression models in which the coefficients can be a set of smoothing functions, then accounting for the non-linearity that could exist between the dependent variable and the covariates Eq. (3):

$$E(Y \mid X_1, X_2, \ldots, X_p) = \alpha + f_1 X_1 + f_2 X_2 + \ldots + f_p X_p \quad (3)$$

where Y is the dependent variable, $X_1, X_2, \ldots X_p$ represent the covariates

and the f_i's are the smooth (non-parametric) functions. As for generalized linear models, the GAM approach specifies a distribution for the conditional mean $\mu(Y)$ along with a link function g relating the latter to an addictive function of the covariates Eq. (4):

$$g[\mu\,(Y)] = \alpha + f_1X_1 + f_2X_2 + \ldots + f_pX_p \quad (4)$$

The GAM model was built using regression splines, and the smoothing parameters were estimated by penalized Maximum Likelihood to avoid an over-fitting (Wood, 2001). An extra penalty added to each smoothing term allowed each of them to be set to zero during the fitting process in case of multi-collinearity or multi-concurvity. Finally, the couple of geographical coordinates (X, Y) was added to each model (as a two-dimensional spline on latitude and longitude) to account for the spatial dependence and trends of the target variable at the regional scale.

A model including all the covariates was first developed (the 'full model'). Then a backward stepwise procedure (BSP; Fig. 4) was applied to select the appropriate covariates where each of the covariates was sequentially removed from the 'full model' (de Brogniez et al., 2014; Poggio et al., 2013). The difference (dAIC) between AIC (Akaike's Information Criterion; Akaike, 1974) calculated for the 'full model' and each AIC obtained when removing sequentially one of the covariates from the model was computed to evaluate the influence of each covariate on the overall capabilities of model prediction. Indeed, the AIC allows measuring the relative quality of different models based on a given dataset: the lower the AIC, the higher the model quality. The 'full model' showed the lowest AIC. Only the covariates which sequential removal significantly increased the AIC of the 'full model' were included in the model.

For validation, predicted values of the independent validation set were compared to observed ones and mean error (ME) and root mean square error (RMSE) were calculated to appreciate the goodness of fit of the model, as well as the coefficient of determination. After this validation phase, a final model was built with all samples (i.e. in both the calibration and validation sets) using the covariates selected by the backward stepwise procedure in order to improve model accuracy and representativity over the entire territory.

2.4.2. Estimation of the uncertainties

In order to properly assess uncertainties of the estimated SOC stocks, the errors due to SOC stock measurement and those associated with the GAM model (i.e., uncertainties of the parameters) were both considered in the SOC stock estimation procedure (Fig. 4). For this purpose, two consecutive Monte Carlo simulations (Steps 3.1 and 3.2 in Fig. 4) were used (Hammersley and Handscomb, 1964; Rubinstein, 1981).

- Step 3.1: The errors associated with the SOC stock calculation at the sampling site scale, i.e. a 4-m radius circle, in the specific context of the CARBOSOL SMN were estimated by Goidts et al., 2009 considering the land use (cropland or grassland) and soil type (texture and stoniness). The coefficients of variation (CV = 100 * SD / μ; with SD the standard deviation and μ the mean) of SOC stock calculation derived from Goidts et al. (2009) were assumed to be representative for the different agricultural soils in southern Belgium, and were used as estimators of SOC stock calculation errors (Table 2). The SOC stock values were then randomly varied following a normal probability distribution with a mean equal to the parameter value and a standard deviation (SD_{SOC}; Fig. 4) based on the CV proposed in Table 2. Hundred MC simulations were performed and, based on each of these new 100 datasets, the parameters of the mean trend model were readjusted to create 100 new GAMs (i from 1 to 100; step 3.1 in Fig. 4).
- Step 3.2: To take into account errors associated with the modelling procedure, the prediction matrix of each of the 100 GAMs obtained above were used to obtain multiple realizations of the trend, according to the approach proposed by Wood (2006). For each of the 100 GAM models produced by the first Monte-Carlo simulation (step 3.1), confidence intervals of the model parameters were estimated. For each of these models, 100 replicate sets of the associated parameters were created through MC simulations considering the mean value of the parameters (MP), their confidence intervals (ci) and a normal distribution. To finish, the prediction matrix relating the model parameters to the predictor were used to obtain 100 replicates (j) of each GAM_i prediction on the 40 m grid (i.e. i and j from 1 to 100; step 3.2 in Fig. 4).

Hence, the simulation led to the computation of 10,000 (100 * 100) independent realizations of the trend model. SOC stock values were then predicted at unsampled locations of Walloon agricultural areas by applying separately each of the 10,000 models to the spatial continuous layers of concerned environmental covariates (step 4 in Fig. 4). So, the obtained 10,000 mapping products had a resolution of 40 m × 40 m and a support equivalent to a site sampling area, i.e. a 4-m radius circle area (~50 m^2).

2.4.3. Spatial estimation

Based on the 10,000 independent spatial estimations of SOC stock values produced above, we computed the median and upper (Q_{95}) and lower (Q_5) percentiles for each pixel. Resulting median values were used to map SOC stock (Mg C ha^{-1}) for the entire territory whereas the uncertainty (dSOC) was estimated as the difference between the 95th and the 5th percentiles, i.e. the 90% prediction interval Eq. (5), as proposed in the GlobalSoilMap project (Heuvelink, 2014):

$$\mathrm{dSOC} = Q_{95} - Q_5 \quad (5)$$

2.4.4. SOC stock quantification

For each of the 10,000 independent spatial estimations of SOC stock values, SOC stocks (Mt C) were computed by summing pixel values multiplied by the pixel surface (40 * 40 m) from the areas requested by GHG reporting, i.e. the 20 combinations of land use/agricultural regions (Table 1), and from the areas associated to WRB soil types according to European Soil Database v2.0 (ESDAC; Panagos et al., 2012). Finally, summary statistics of SOC stocks (mean and standard deviation; n = 10,000) were computed in order to be comparable with results from the literature.

The simulation presented in Fig. 4 and the summary statistics were run in R software.

3. Results and discussion

3.1. Descriptive statistics

Based on 100 randomly selected locations, a linear model Eq. (1) was fitted between raw data obtained by the Walkley and Black method (C_{wb}) and those obtained by the dry combustion method (C_{dc}) (a = 1.1692 and b = 2.9655; both significant at $p < 0.05$; Fig. 5). The correction regression in Meersmans et al. (2009) was also significant with similar parameter values for soils under cropland and dry silt loam grassland soils located in Northern Belgium. The positive significant intercept suggested the presence of recalcitrant organic matter

Table 2
Estimators used as uncertainties for SOC stock calculation considering land use and stoniness at the sampling site scale, i.e. a 4-m radius circle (after Goidts et al., 2009).

Landuse	Soil texture	CV[a] of SOC stock (%)
Cropland	Silt (loam)	8
	Stony loam	8
Grassland	Silt (loam)	11
	Stony loam	17

[a] CV = coefficient of variation

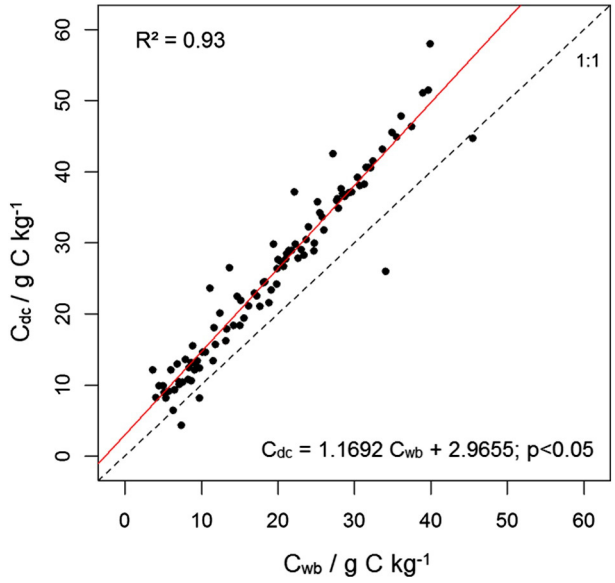

Fig. 5. Carbon content measured by dry combustion measurement (C_{dc}; of which inorganic carbon content has been subtracted) plotted against organic carbon content measured by the Walkley and Black method (C_{wb}).

resulting in an incomplete oxidation by the Walkley and Black procedure (Meersmans et al., 2009). The mean error between C_{dc} and $C_{wb}{}^{*}$ (i.e., C_{wb} corrected for incomplete oxidation by the linear model fitted above) was $6 * 10^{-5}$ g C kg^{-1} and non-significant (paired-sample *t*-test $P = 0.73$, $n = 100$). From these results, we assumed that the difference between SOC content values obtained from these two different methods were comparable after correction of the C_{wb} values and induced no bias in the SOC stock spatial trends and SOC stocks calculation (Figs. 6 and 11).

Descriptive statistics of the total dataset ($n = 592$) showed minimum and maximum SOC stock values of 28.81 and 163.8 Mg C ha^{-1} with a mean of 69.74 Mg C ha^{-1} (Fig. 6). The dataset showed a positively-skewed unimodal distribution ($p = 0.98$ at the Dip Test; Hartigan and Hartigan, 1985). The mode was located around 55 Mg C ha^{-1} and related to locations sampled under the predominant land use in the dataset, i.e. cropland ($n = 371$; median = 53.60 Mg C ha^{-1}; mean = 55.31 Mg C ha^{-1}; Fig. 6). Locations under grassland which represented ca. 37% of the total dataset ($n = 220$; median = 91.30 Mg C ha^{-1}; mean = 93.53 Mg C ha^{-1}) showed higher values of SOC stocks than those under cropland (mean difference of 38.22 Mg C ha^{-1} at $p < 0.01$).

A correlation matrix was first computed by including SOC stock and all the environmental covariates. From this first matrix (data not shown), we selected environmental covariates showing a minimum correlation coefficient of abs(0.10) with SOC stock to produce a simplified matrix which could be read in a paper format (Table 3). As the GAM fitting procedure used in this study already deals with collinearity and coconcurvity between covariates, the correlation matrix allows a first investigation on the relations between all the parameters and later to support the interpretation of the results from the backward stepwise procedure (BSP; Section 2.4). The highest correlations observed between SOC and environmental covariates were with the clay and fine silt content (C-fS; $\rho = 0.48$) and then the clay content (Clay; $\rho = 0.43$) (Table 3). The most significant correlations between SOC stock and morphometric parameters were 0.31 for elevation and 0.25 for ls-factor. Precipitation and mean temperature showed correlation coefficients of 0.35 and −0.27 with SOC stock, respectively. Amongst the independent covariates, strong correlations were observed between elevation and temperature (−0.99), rain and temperature (−0.86), and unsurprisingly between ls-factor and slope (0.88). Latitude and longitude also showed high correlation values with several covariates such as elevation and climatic covariates.

Minimum, mean and maximum SOC stock values were of 28.6, 68.7 and 163.8 Mg C ha^{-1} in the calibration dataset ($n = 396$), and of 31.2, 67.9 and 161.5 Mg C ha^{-1} in the validation dataset ($n = 196$; Table 4).

3.2. SOC stock model

3.2.1. Model calibration

The SOC stock dataset being continuous and strictly positive, we applied a Gamma distribution in the GAM model. The log-link function was chosen for the model fitting considering the positively-skewed unimodal characteristic of the SOC stock distribution. The *Q-Q* plot showed that the calibration residuals appeared to be mostly aligned along a Gamma distribution (Fig. 7).

The backward stepwise procedure selected the land use, the clay and fine Silt content (C-fS), the coordinates (*X*, *Y*), the LS-factor, the maximum ground water table depth (GWL max), and the temperature as the most influencing covariates (in decreasing order) on the prediction model (Fig. 8). The other covariates considered in this study were not retained in the final model as they did not influence the model performance, i.e. not significantly increase the AIC value when they were sequentially extracted from the model.

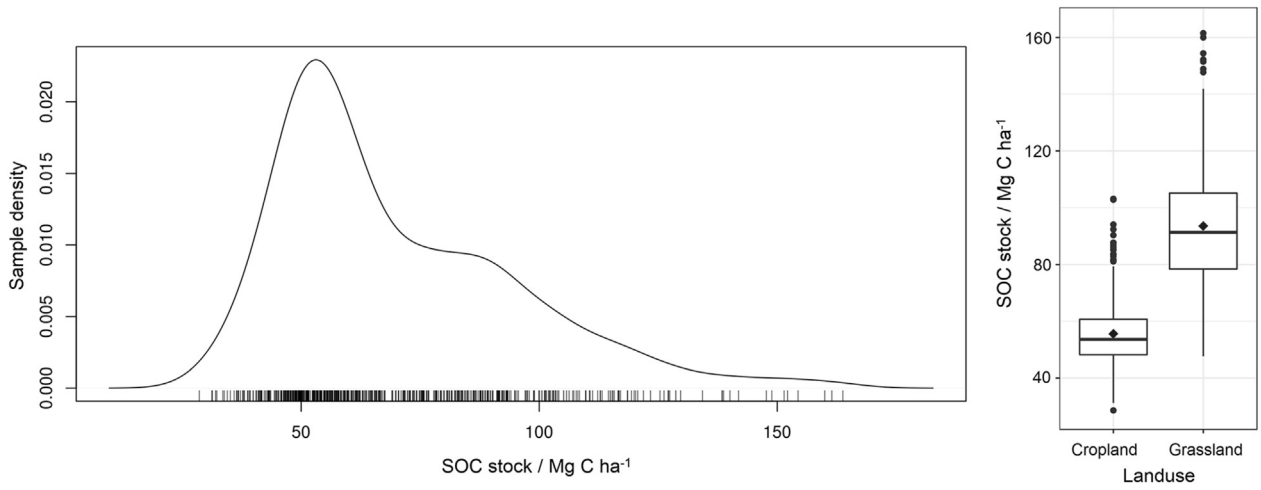

Fig. 6. Histogram of SOC stocks measured at the 592 investigated locations (A) and boxplot of SOC stocks by land use, i.e. cropland and grassland (B). In A, the small bars at the bottom of the graph indicate the presence or absence of samples. In B, the horizontal line in the middle of the boxes refers to the median, while their lower and upper limits are to the first and third quartiles, and the diamond represents the mean.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

Table 3
Correlation coefficients between SOC stock and different environmental covariates.

	SOC stock	X	Y	C-fS[a]	Clays	Silt	C-input	Elevation	Slope	Flow acc.[b]	Flow length	LS-factor	Rain	Temp.[c]	GWL min	GWL max
SOC density	1.00	0.39	−0.19	0.48	0.43	−0.19	0.30	0.33	0.24	0.16	0.19	0.26	0.36	−0.29	−0.14	−0.14
X		1.00	−0.39	0.57	0.28	−0.08	0.17	0.78	0.35	−0.02	0.34	0.33	0.65	−0.71	0.28	−0.05
Y			1.00	−0.60	−0.39	0.33	−0.28	−0.65	−0.28	−0.05	−0.67	−0.24	−0.62	0.71	−0.06	−0.04
C + fS[a]				1.00	0.77	−0.15	0.34	0.65	0.25	0.14	0.54	0.26	0.52	−0.64	0.03	0.04
Clays					1.00	−0.42	0.30	0.30	0.16	0.23	0.24	0.13	0.30	−0.28	−0.08	0.03
C-factor						1.00	−0.36	−0.45	−0.35	−0.09	−0.25	−0.36	−0.43	0.42	−0.09	−0.01
C-input							1.00	0.38	0.19	−0.02	0.17	0.19	0.35	−0.35	−0.04	−0.04
Elevation								1.00	0.31	−0.05	0.62	0.29	0.87	−0.99	0.22	0.00
Slope									1.00	−0.06	0.17	0.88	0.27	−0.29	0.21	−0.05
Flow acc.[b]										1.00	0.05	0.06	−0.02	0.04	−0.19	−0.07
Flow length											1.00	0.18	0.54	−0.69	0.03	0.05
LS-factor												1.00	0.22	−0.27	0.13	−0.10
Rain													1.00	−0.86	0.13	−0.05
Temp.[c]														1.00	−0.21	−0.01
GWL															1.00	0.10
																1.00

[a] Clay and fine silt content
[b] Flow accumulation
[c] Temperature

Several studies have already demonstrated that land use/land cover was by far the most influential environmental parameters on SOC content prediction (e.g., Jones et al., 2005; de Brogniez et al., 2014; Wiesmeier et al., 2014). Indeed, land management has a great impact on the balance between inputs and outputs of organic carbon in soils (Paustian et al., 1996; Rees et al., 2005). For example, fertilizing improved SOC storage while ploughing induces SOC loss through disaggregation and then a loss of protection of organic matter particles to oxidation (Kowalenko et al., 1978; Beare et al., 1993).

The clay and fine silt content (C-fS; Fig. 8) is the second parameter which influences the final model the most. Bulk SOC consists of heterogeneous pools with different turnover rates and stabilities depending on complex interactions between biological, chemical, and physical processes in the soil (Post & Kwon, 2000). The stable fraction of SOC, which has slower turnover (from years to decades), has been demonstrated to be associated to clay and fine silt content (Hassink, 1997). The stable fraction can represent around >60% of the bulk SOC in the cropland of northern Wallonia, i.e. the Loam agricultural region (Fig. 1 region 2) (Trigalet et al., 2014). This could explain the selection of the clay and fine silt content as the significant texture covariate selected to model the spatial variation of SOC stock instead of the classical clay content.

The LS-factor as used in the RUSLE equation combines flow length and slope gradient, explaining the removal of both of them from the model by the backward stepwise procedure. In the model, the LS-factor is the fourth parameter which influences the overall quality of the prediction the most, especially for cropland (Fig. 9). SOC stock increased with LS-factor in cropland, very slightly between 0 and ca. 10 (−) and then very rapidly. Fields located in Sandy-loam (n°1 in Fig. 1) and Loam (n°2) agricultural regions (intensive agriculture) are prone to severe water erosion processes which transport fine soil particles (rich in SOC) and deposit them temporarily downslope where LS-factor values are high due to flow length. Moreover, as highlighted in Table 3, LS-factor increases with elevation ($\rho = 0.29$). Some locations sampled in cropland with a high LS-factor are located in hilly areas where soils are managed very differently (extensive agriculture) than in flatter areas (intensive agriculture) because of the relief and the stoniness of the soils. Soils in hilly areas are mainly used as grassland especially in very stony areas, but it is not unusual that crop fields located in less stony areas include a grass ley in the rotation. Grasslands have higher SOC contents than croplands, as highlighted in this study and many others (see above), and croplands with grassland in their rotation have higher SOC content and stocks than those strictly used for crops. To finish, we have to consider that these specific rotations including grassland are located at higher altitudes and then submitted to climate conditions (higher rainfall and lower temperatures) which are more favorable to maintain higher SOC levels (i.e., lower mineralization rates and/or higher C-input rates).

The maximum ground water table depth (GWL max), which corresponds to the depth of the reduction horizon, was the fifth parameter influencing the SOC stocks model prediction. Indeed, SOC stocks increased while drainage deteriorated, i.e. while the reduced horizon appeared closer to the surface. A shallower reduced horizon indicates a higher ground water table. The decomposition of organic matter is slower under anaerobic conditions. Meersmans et al. (2011) also demonstrated that the ground water table depth, especially within grassland, greatly influenced SOC contents in Belgium.

Table 4
Summary descriptive statistics of measured SOC density (Mg C ha^{-1}) in calibration and validation datasets.

Dataset for	Calibration	Validation
n	396	196
min	28.6	31.3
Q_{25}	50.6	50.6
Q_{50}	60.6	60.5
mean	68.7	70.4
Q_{75}	84.9	84.8
Max	163.8	161.5
SD	24.3	25.6

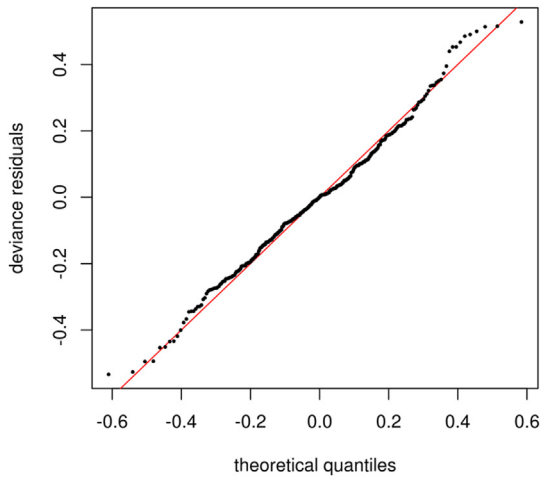

Fig. 7. Q-Q plot of the calibration residuals.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

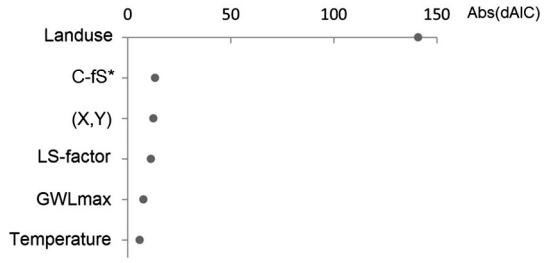

Fig. 8. Importance of covariates selected in the final GAM model for spatial interpolation of SOC density based on the difference in Akaike Information Criterion (abs(dAIC)). (* C-fS: clay and fine silt content)

To finish, the last parameters selected by the backward stepwise procedure is the mean annual temperature. SOC stock decreased with mean annual temperature increase which appeared in agreement with literature (e.g., Kirschbaum, 1995; Sleutel et al., 2007; Wang et al., 2013). Indeed, microbial decomposition of organic matter and Net Primary Production depend on temperature (Nemani et al., 2003; Davidson and Janssens, 2006).

Finally, the fitted GAM model explained 69.3% of the SOC stock variance in the calibration dataset.

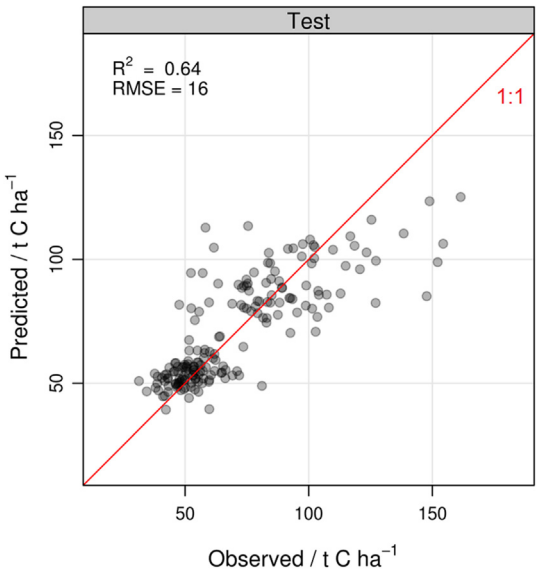

Fig. 10. Predicted SOC stock plotted against observed SOC stock for the validation dataset ($n = 196$) (red line is the 1:1 line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2.2. Model validation

The validation of the model gave satisfactory results with a R^2 value of 0.64 and a RMSE of 16 Mg C ha^{-1} (11.3 Mg C ha^{-1} for cropland and 21.6 Mg C ha^{-1} for grassland; Fig. 10). The model showed a trend to overestimate low SOC stock (below 50 Mg C ha^{-1}) and to underestimate high SOC stock (especially those above 100 Mg C kg^{-1}; Fig. 10).

Goidts et al. (2009) explained 12 to 29% of the SOC stock variance in southern Belgium by a digitizing approach, whereas Meersmans et al. (2008) explained 36% of the variance in northern Belgium by using a multiple linear regression approach, including some higher order interaction terms, based on land use, texture and soil drainage. For the entire Belgian territory, Meersmans et al. (2011) obtained a R^2 of 0.42 (for a legacy data set of the 1960s) and 0.65 for a data set from 2005.

Studies performed in other closed European countries showed similar results. Schulp and Verburg (2009) explained variances of 21 to 43% of SOC contents and stocks in different study areas in the Netherlands while extensive data on soil properties, topography, environmental conditions and recent historic land use were used as predictors. Wiesmeier et al. (2014) produced a spatial prediction using the regression tree technique explaining 39% of the spatial variability of SOC stock in Bavaria (Germany).

Hence, in this study, model validation often showed a higher R^2 value (0.64) than the other studies cited above (with R^2 ranging from 0.12 to 0.65). Except the fact that the modeling approaches were different between the different studies and this study, this could be also explained by the sampling technique. Here, we used composite samples from five sub-samples by site rather than classically one sample. By using a composite sampling technique in this study, the local SOC stocks variability was filtered out, getting rid of outliers which could have to a large extent affected R^2 and RMSE.

Despite the incorporation of numerous environmental covariates and the diversity of modelling techniques used in these different approaches, capturing the complexity of the processes controlling SOC storage and mineralization at regional scale remains challenging. Only low levels of explanation are expected for SOC predictions due to an inherent high spatial variability of SOC (Schulp and Verburg, 2009; Schulp et al., 2013).

The model fitted above has then been used to produce 10,000 independent spatial simulations of SOC stocks (Fig. 4), with a resolution of 40 × 40 m and a support corresponding to the surface of the sampling site (~50 m^2). In order to obtain raster layers with a support of the same size as the resolution, block kriging could have been applied to

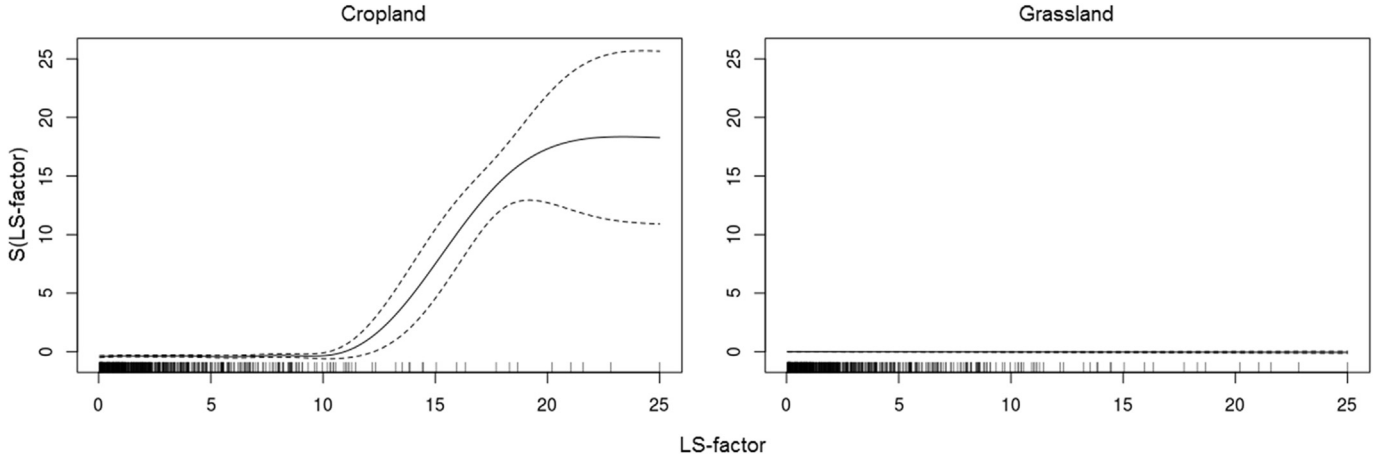

Fig. 9. Smooth functions of LS-factor as produced by the GAM model for cropland and grassland (land use is implemented as a categorical variable in the model). The dashed lines represent the 95% confidence interval. The bar at the bottom of the graph indicates the presence or absence of a sample.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

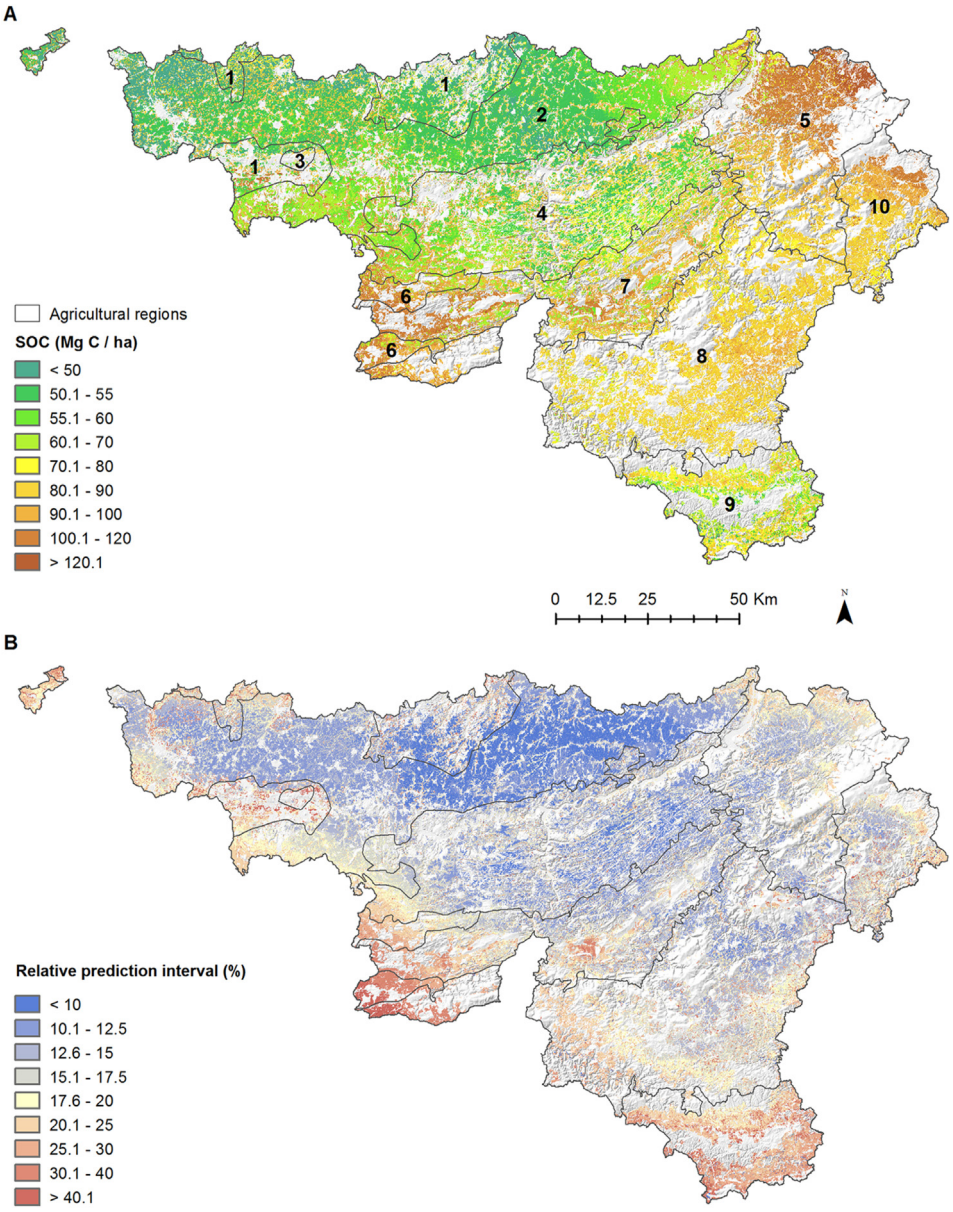

Fig. 11. Topsoil organic carbon stocks (SOC in Mg C ha^{-1}) (A) and relative prediction uncertainties (%; $((Q_{95} - Q_5) / SOC) * / 100$) (B) in Southern Belgium (Wallonia). Grey areas correspond to environments where no prediction of organic carbon stock could be made (e.g. built areas and forests). The agricultural regions are delineated by dark grey lines (1: Sandy-loam, 2: Loam, 3: Campine Hennuyère, 4: Condroz, 5: Herbagère Liège, 6: Herbagère Fagnes, 7: Famenne, 8: Ardenne, 9: Jurassic, and 10: Haute Ardenne).

each of the 10,000 simulations (Malone et al., 2013). Unfortunately, block kriging is computationally expensive and as yet not feasible for the 10,000 individual layers.

3.3. Spatial variability of SOC stocks

Median values of the 10,000 independent spatial simulations of SOC stock were applied to map the SOC stock (Mg C ha^{-1}; Figs. 4 and 10A). A positive gradient of SOC stock occurred from the northwest to the center of Wallonia with a slight decrease on the southernmost part, i.e. the Ardenne and locally in Jurassic agricultural regions (9 and 10 in Fig. 11A). This regional trend is mainly correlated with the evolution of elevation (Fig. 1A), precipitation and temperature (Fig. 2) and dominant land use (Fig. 1B). The visual comparison of the spatial distribution of SOC stock produced here and the SOC content produced by Meersmans et al., 2011 were in good agreement for the main regional trends described above. However, while the highest values appeared predominantly in Ardenne (8) and Haute Ardenne (10; Fig. 11A) in Meersmans et al. (2011), in this study the highest values were in Herbagère Fagnes (6), Famenne (7), Herbagère Liège (5) and Haute Ardenne (10; Fig 11A). These four agricultural regions are dominated by grassland (Fig. 1B), characterized by high clay and fine silt contents in topsoil (especially for regions 6 and 7; Fig. 3). The differences observed could probably come from the fact that Meersmans et al. (2011) predicted SOC content (g C $kg^{-}1$) whereas, here, we predicted SOC stock (Mg C ha^{-1}). Stone content is considerably higher in Ardenne (8) and Haute Ardenne (10) than in Herbagère Fagnes (6), Famenne (7) and Herbagère Liège (5; Fig. 10A). When calculating SOC stocks, the volume of soil occupied by stones is considered not containing any SOC. Thus, for two areas with similar mean SOC contents (g C kg^{-1}) and fine earth bulk density, the stony area will show lower SOC stocks than the non-stony one. Also, we used different methods for the spatial modelling, i.e. multiple linear regression including some higher order interaction terms for Meersmans et al. (2011) and GAM in this study, as well as different covariates layers (different in computation and resolution).

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

Based on the 10,000 independent spatial simulations produced in this study, the 5th and 95th percentiles were computed for each pixel to assess the uncertainties (the 90% prediction interval dSOC; Eq. (5)) which computation accounts for errors related to both SOC stock calculation and spatial modelling. The relative uncertainties – i.e., ((dSOC / 2) / SOC) * 100 (Fig. 10B) – exerted the same main regional spatial trends as those observed for SOC stock values. Indeed, the most reliable predictions were obtained for northwestern agricultural regions dominated by cropland, i.e. the Sandy-loam (n°1), Loam (n°2) and Condroz (n°4; Fig. 11A).

In 2009, Goidts et al. characterised the variability of SOC stock at different spatial scales in Wallonia and showed that SOC variation in a few meters could be locally on the same order of magnitude as the variation at the regional scale. This results from the multiple processes acting simultaneously on SOC balance (i.e. between SOC stabilisation and mineralisation) in topsoil but at different spatial scales. Indeed, these processes are controlled by environmental factors which show scale-dependent links to SOC content and stock. Stevens et al. (2015a) demonstrated that SOC content variance over central Belgium could be decomposed in three main signals acting at different spatial ranges, and potentially corresponding to different environmental factors controlling the entire SOC variability; i.e. long-range (~30 km, i.e. regional influences), medium-range (~5 km, i.e. linked to watershed systems) and short-range (~700 m, i.e. related to field/farm management). By separating the variation of the SOC at these three spatial scales, the authors disentangled the combined effect of the controlling factors and improve the qualitative and quantitative characterization of the controlling factors separately. In this study, we explained 64% of the spatial variability of SOC stock in agricultural topsoil through a GAM including five significant covariates, i.e. the land use, the clay and fine silt content, the LS-factor, the maximum Ground Water Level and precipitation in decreasing order of importance (Fig. 8). Besides the land use, these covariates hence act mainly at two different scales on SOC stock variation, the long-range scale for precipitation and the clay and fine silt content, and the medium-range scale for the maximum ground water level and the LS-factor. None of the covariates used initially in this study could have reflected field or farm management as this type of information is difficult to synthetize and moreover at the regional scale. However, a study from Luxembourg, and based on the use of mixed-effects models and airborne imagery showed that field effect explained locally from 23.3 ± 4.1% up to 68.8 ± 12.0% of the SOC variance (Stevens et al., 2015b).

Based on DSM, the modelling procedure applied here allowed to explain 64% of the SOC stock variability and to depict realistic spatial patterns. However, a better performance could not be achieved given the sampling density that is too coarse to detect the field scale variability. The produced map of SOC stocks and associated uncertainties confer tools in conducting sustainable management at regional scale, but would be also a support for designing a sampling scheme for carbon auditing at farm scale (de Gruijter et al., 2016).

3.4. SOC stocks in Wallonia

SOC stocks for the top 0–30 cm depth of soil were computed based on the 100,000 independent spatial simulations for each of the 20 land use/agricultural region combinations, including combinations which had no or very few sampling sites within the CARBOSOL SMN. Descriptive statistics of predicted SOC stocks per combination and the entire Wallonia, as well as observed and predicted SOC stocks, are presented in Table 5 or according to WRB classes in Table 6.

Mean predicted total SOC stocks for cropland and grassland in Wallonia were 26.58 Tg C in croplands and 43.30 Tg C in grasslands. This corresponds to a mean predicted SOC stock of 59.9 Mg C ha^{-1} in cropland, higher than the observed value of 55.3 Mg C ha^{-1}, and to a mean predicted SOC stock of 92.3 Mg C ha^{-1} in grassland, slightly lower than the observed value of 93.5 Mg C ha^{-1}. These differences were coherent as the model validation highlighted a trend to overestimate low SOC stock (below 50 Mg C ha^{-1}) and to underestimate high SOC stock (above 100 Mg C kg^{-1}; Fig. 10). Hence, the model induced an overestimation of SOC stocks for cropland and an underestimation of SOC stocks for grassland in Wallonia.

Lettens et al. (2005b) measured mean SOC stock values of 48 Mg C ha^{-1} in cropland and 75 Mg C ha^{-1} in grassland for Wallonia which appeared lower than the observed and predicted values calculated here (Table 5). Lettens et al. (2005b) used different sampling methods and estimated the soil bulk density through a PTF. At the whole Belgian territory scale, Meersmans et al. (2011) predicted SOC stock values of 49.5 Mg C ha^{-1} (SD 0.90) for cropland and 79.9 Mg C ha^{-1} (SD 1.0) for grassland. These results for Belgium are also based on a different sampling method than in the present study

Table 5
Mean observed and predicted SOC stocks (Mg C ha^{-1}) and predicted total stocks (Tg C) in cropland and grassland for the ten agricultural regions of Wallonia (see Fig. 1A).

Landuse	Agricultural region		Observed		Predicted	
			n	Mean (SD) SOC stocks — Mg C/ha	Mean (SD) SOC stocks — Mg C/ha	Mean (SD) total SOC stocks — Tg C
Cropland	Sandy-loam	(1)	59	51.0 (8.8)	53.5 (1.8)	2.19 (0.07)
	Loam	(2)	133	52.0 (12.5)	54.5 (1.6)	13.18 (0.39)
	Camipne Hennuyère	(3)	–	–	53.1 (4.8)	0.02 (0.00)
	Condroz	(4)	127	56.2 (9.4)	60.7 (2.2)	5.04 (0.18)
	Herbagère Liège	(5)	4	65.02 (8.2)	83.6 (5.1)	0.75 (0.05)
	Herbagère Fagnes	(6)	–	–	69.5 (11.7)	0.34 (0.06)
	Famenne	(7)	18	58.8 (13.7)	71.7 (7.8)	1.59 (0.17)
	Ardenne	(8)	23	74.5 (9.1)	88.1 (9.2)	2.16 (0.23)
	Jurassic	(9)	6	56.1 (7.5)	67.4 (6.6)	0.81 (0.08)
	Haute Ardenne	(10)	1	87.7 (–)	107.0 (16.5)	0.50 (0.08)
Cropland	All		371	55.3 (12.1)	59.9 (2.8)	26.58 (1.52)
Grassland	Sandy-loam	(1)	–	–	86.4 (7.6)	1.76 (0.15)
	Loam	(2)	49	90.3 (25.2)	90.5 (4.9)	6.26 (0.34)
	Camipne Hennuyère	(3)	–	–	94.0 (50.5)	0.04 (0.02)
	Condroz	(4)	22	84.4 (12.5)	87.1 (4.4)	5.28 (0.27)
	Herbagère Liège	(5)	40	109.0 (21.0)	103.4 (5.5)	7.32 (0.39)
	Herbagère Fagnes	(6)	–	–	103.9 (20.4)	1.66 (0.33)
	Famenne	(7)	41	99.4 (24.3)	94.9 (15.3)	5.63 (0.91)
	Ardenne	(8)	42	86.0 (15.1)	86.7 (5.1)	9.43 (0.55)
	Jurassic	(9)	4	68.3 (10.4)	79.1 (9.4)	2.76 (0.33)
	Haute Ardenne	(10)	23	90.0 (16.3)	91.5 (6.5)	3.16 (0.22)
Grassland	All		221	93.5 (22.2)	92.3 (6.2)	43.30 (2.93)

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006

Table 6
Mean predicted SOC stocks (Mg C ha^{-1}) and predicted total stocks (Tg C) in cropland and grassland for the different soil types (according to WRB) of Wallonia.

Landuse	Soil type	% of total area per landuse	Predicted	
			Mean (SD) SOC stocks — Mg C/ha	Mean (SD) total SOC stocks — Tg C
Cropland	Eutric Fluvisol	2.01	55.9 (2.1)	1.03 (0.04)
	Haplic Retisol	2.46	49.0 (2.7)	1.10 (0.06)
	Haplic Luvisol	38.66	57.7 (1.6)	20.45 (0.55)
	Dystric Cambisol	3.64	91.6 (9.0)	3.06 (0.30)
	Gleyic Luvisol	0.75	70.0 (5.7)	0.48 (0.04)
	Haplic Arenosol	0.31	51.7 (4.5)	0.15 (0.01)
	Dystric Histosol	<0.01	96.4 (38.8)	<0.01 (<0.001)
	Distric Gleysol	<0.01	87.4 (39.6)	<0.01 (<0.001)
	Calcaric Luvisol	<0.01	74.1 (11.0)	<0.01 (<0.001)
	Eutric Cambisol	0.07	84.3 (6.7)	0.06 (<0.01)
	Arenic Luvisol	0.41	61.3 (7.2)	0.23 (0.03)
	Calcaric Cambisol	0.01	71.7 (9.0)	<0.01 (<0.001)
Grassland	Eutric Fluvisol	1.54	91.1 (5.9)	1.28 (0.08)
	Haplic Retisol	0.83	86.9 (9.6)	0.66 (0.07)
	Haplic Luvisol	23.31	91.7 (3.5)	19.61 (0.76)
	Dystric Cambisol	20.02	89.4 (3.8)	16.43 (0.69)
	Gleyic Luvisol	4.54	100.8 (6.1)	4.20 (0.25)
	Haplic Arenosol	0.27	90.4 (36.2)	0.23 (0.09)
	Dystric Histosol	0.13	103.2 (13.0)	0.12 (0.01)
	Distric Gleysol	<0.01	95.5 (14.1)	<0.01 (<0.001)
	Calcaric Luvisol	0.01	77.7 (15.4)	<0.01 (0.001)
	Eutric Cambisol	0.24	80.2 (12.8)	0.18 (0.03)
	Arenic Luvisol	0.78	70.9 (20.7)	0.51 (0.15)
	Calcaric Cambisol	0.02	77.6 (14.6)	0.01 (<0.01)

and on soil bulk density estimated through a PTF: this could in part explain why the results in Meersmans et al. (2011) are significantly lower from the predictions described in this paper for the southern part of the country (59.9 Mg C ha^{-1} (SD 2.8) in cropland and 92.3 Mg C ha^{-1} (SD 6.2) for grassland; Table 5). In addition, northern Belgium showed lower SOC contents than in southern Belgium which could also explain that predicted mean national values were lower than mean regional values for Wallonia (Meersmans et al., 2011). The predictions for entire Belgium (Meersmans et al., 2011) also showed lower uncertainties than the predictions produced for Wallonia (Table 5). The uncertainties produced in this present study accounted for errors associated to the SOC stock calculation and to the model parameters, while only errors associated to the model parameters were considered in the uncertainties computation in the study consecrated to the whole Belgium.

In France, mean SOC stock values of 55.7 Mg C ha^{-1} and 85.8 Mg C ha^{-1} were predicted for cropland and grassland respectively (Meersmans et al., 2012), which are similar to our predictions for Wallonia. Similarly, total SOC stocks were estimated for France where the relative model error (i.e. model error divided by predicted value) on the total national SOC stock was 33.9% (Meersmans et al., 2012), whereas here the relative model error was ca. 6.7%. These differences could be explained by the fact that for France the model included not only cropland and grassland but also other land uses as vineyards and forests, and by the different methods used to compute the SOC stocks and associated uncertainties. Meersmans et al. (2012) also used a MC simulation but first perturbed the parameters of the model estimating SOC contents, and then the bulk density (as estimated by a PTF) and rock matter content were perturbed to finally obtain SOC stocks.

4. Conclusion

A methodology combining the Digital Soil Mapping technique and stochastic simulations (Monte-Carlo) was developed to both quantify and spatialize SOC stocks for southern Belgium based on a convenience sampling scheme. A Generalised Additive Model (GAM) explaining 69.3% of the SOC stock variance was calibrated and then validated ($R^2 = 0.64$). During the model fitting procedure, the landuse (cropland or grassland), the clay and fine silt content, the LS-factor, the maximum ground water level, and the temperature were selected as the most significant covariates explaining SOC stock variation at the regional scale. A positive gradient of SOC stock occurred from the northwest to the center of Wallonia with a slight decrease on the southernmost part. This regional trend is mainly correlated with the evolution precipitation and temperature (along with elevation) and dominant land use. At the catchment scale a higher SOC stock was predicted on poorly drained soils of valley bottoms, especially under grassland. A better performance for SOC stock spatial prediction could not have been achieved given the sampling density of the SMN CARBOSOL that is too coarse to detect the field scale variability. Two successive Monte-Carlo simulations produced 10,000 spatialized datasets for SOC stock all over Wallonia and allowed to consider uncertainties due to SOC stock measurement and those associated with the GAM model (i.e., uncertainties of the parameters). Based on these 10,000 independent datasets, reliable statistics for SOC stocks were computed. Mean predicted SOC stocks in Wallonia were 26.58 (SD 1.52) for cropland and 43.30 Tg C (2.93) for grassland. The model induced an overestimation of low SOC stock (below 50 Mg C ha^{-1} and predominantly under croplands), and an underestimation of high SOC stock (especially those above 100 Mg C ha^{-1}, predominantly under grasslands).

The procedure developed here allowed to predict realistic spatial patterns of SOC stock all over agricultural lands of southern Belgium and to produce reliable statistics of SOC stocks for each of the 20 combinations of land use/agricultural regions of Wallonia. This procedure appears useful to produce soil maps as policy tools in conducting sustainable management at regional and national scales, and to produce statistics which comply with specific requirements of reporting activities.

Acknowledgement

This research was funded by the Public Administration of Wallonia (SPW-DGO3).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi: 10.1016/j.geodrs.2016.12.006. These data include the Google map of the most important areas described in this article.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723.
Beare, M.H., Hendrix, P.F., Cabrera, M.L., Coleman, D.C., 1993. Aggregate-protected and unprotected organic matter pools in conventional- and no-tillage soils. SSSAJ 58 (3), 787–795.
Bloem, J., Schouten, J.A., Sorensen, S.J., Rutgers, M., van der Werf, A., Breure, A.M., 2005. Monitoring and evaluating soil quality. In: Bloem, J., Hopkins, D.W., Benedetti, A. (Eds.), Microbiological Methods for Assessing Soil Quality. CABI Publishing, pp. 23–49.
Boer, E.P.J., de Beurs, K.M., Hartkamp, A.D., 2001. Kriging and thin plate splines for mapping climate variables. Int. J. Appl. Earth Obs. Geoinf. 3 (2), 146–154.
Brus, D.J., de Gruijter, J.J., 1993. Design-based Versus Model-based Estimates of Spatial Means. Theory and Application in Environmental Soil Science Environmetrics 4 pp. 123–152.
CRC, 1996. Handbook of chemistry & physics: a ready-reference book of chemical and physical data. CRC Handbook of Chemistry, 77th ed. CRC Press (boca raton edition).
Davidson, E.A., Janssens, I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature 440 (7081), 165–173.
de Brogniez, D., Ballabio, C., Stevens, A., Jones, R.J.A., Montanarella, L., van Wesemael, B., 2014. Eur. J. Soil Sci. 66 (1), 121–134.
de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotters, M., 2006. Sampling for Natural Resource Monitoring. Springer-Verlag, New York.
de Gruijter, J.J., McBratney, A.B., Minasny, B., Wheeler, I., Malone, B.P., Stockmann, U., 2016. Farm-scale soil carbon auditing. Geoderma 265, 120–130.
De Leenheer, L., Appelmans, F., Vandamme, J., 1968. Système des cartes perforées de la section caractérisation du sol de la cartographie des sols de Belgique. Pédologie 18, 208–227.

ARTICLE IN PRESS

Dendoncker, N., van Wesemael, B., Rounsevell, M., Roelandt, C., 2004. Belgium's CO_2 mitigation potential under improved cropland management. Agric. Ecosyst. Environ. 103, 101–116.

ESRI, 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

European Soil Data Centre (ESDAC), esdac.jrc.ec.europa.eu, European Commission, Joint Research Centre.

FAO, 2015. International Year of Soil. http://www.fao.org/soils-2015/.

Goidts, E., van Wesemael, B., 2007. Regional assessment of soil organic carbon changes under agriculture in Southern Belgium (1955–2005). Geoderma 141 (3), 341–354.

Goidts, E., van Wesemael, B., Crucifix, M., 2009. Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales. Eur. J. Soil Sci. 60 (5), 723–739.

Hallett, P.D., Loades, K.W., Krümmelbein, J., 2012. Soil physical degradation: threats and opportunities to food security. In: Hester, R.E., Harrison, R.M. (Eds.), Soils and Food Security. Royal Society of Chemistry, Cambridge, pp. 198–226.

Hammersley, J.M., Handscomb, D.C., 1964. Monte Carlo methods. Methuen.

Hartigan, J.A., Hartigan, P.M., 1985. The dip test of unimodality. Ann. Stat. 13 (1), 70–84.

Hassink, J., 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. Plant and Soil 191, 77–87.

Hastie, T., Tibshirani, R., 1986. Generalized additive models. Stat. Sci. 1 (3), 297–310.

Heuvelink, G.B.M., 2014. Uncertainty quantification of GlobalSoilMap products. In: Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A., McBratney, A. (Eds.), GlobalSoilMap: Basis of the Global Spatial Soil Information System. CRC Press, pp. 335–340.

Jenness, J., 2006. Topographic Position Index (tpi_jen. avx) Extension for ArcView 3. x, v. 1.3 a. Jenness Enterprises.

Jolivet, C., Arrouays, D., Boulonne, L., Ratié, C., Saby, N., 2006. Le réseau de mesures de la qualité des sols de France (RMQS). État d'avancement et premiers résultats. Etud. Gest. Sols 13 (3), 149–164.

Jones, R.J.A., Hiederer, R., Rusco, E., Montanarella, L., 2005. Estimating organic carbon in the soils of Europe for policy support. Eur. J. Soil Sci. 56, 655–671.

Kidd, D., Malone, B., McBratney, A., Minasny, B., Web, M., 2015. Operational sampling challenges to digital soil mapping in Tasmania, Australia. Geoderma Reg 4, 1–10.

Kirschbaum, M.U.F., 1995. The temperature dependence of soil organic matter decomposition and the effect of global warming on soil organic C storage. Soil Biol. Biochem. 27, 753–760.

Kowalenko, C.G., Ivarson, K.C., Cameron, D.R., 1978. Effect of moisture content, temperature and nitrogen fertilization on carbon dioxide evolution from field soils. Soil Bio. Chem. 10 (5), 417–423.

Lagacherie, P., 2008. Digital soil mapping: a state of the art. In: Hartemink, A.E., McBratney, A., Mindonça-Santos, M.L. (Eds.), Digital Soil Mapping With Limited Data. Springer, Netherlands, pp. 3–14.

Lettens, S., Van Orshoven, J., van Wesemael, B., De Vos, B., Muys, B., 2005a. Stocks and fluxes of soil organic carbon for landscape units in Belgium derived from heterogeneous data sets for 1990 and 2000. Geoderma 127 (1–2), 11–23.

Lettens, S., Van Orshoven, J., Van Wesemael, B., Muys, B., Perrin, D., 2005b. Soil organic carbon changes in landscape units of Belgium between 1960 and 2000 with reference to 1990. Glob. Chang. Biol. 11 (12), 2128–2140.

Malone, B.P., McBratney, A.B., Minasny, B., 2013. Spatial scaling for digital soil mapping. Soil Sci. Soc. Am. J. 77 (3), 890–902.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143 (1–2), 1–13.

Meersmans, J., van Wesemael, B., Van Molle, M., 2009. Determining soil organic carbon for agricultural soils in north Belgium: a comparison between the classic/modified Walkley & Black and the dry combustion method. Soil Use Manag. 25, 346–353.

Meersmans, J., van Wesemael, B., Goidts, E., van Molle, M., De Baets, S., De Ridder, F., 2011. Spatial analysis of soil organic carbon evolution in Belgian croplands and grasslands, 1960–2006. Glob. Chang. Biol. 17, 466–479.

Meersmans, J., Martin, P.M., Lacarce, E., De Baets, S., Jolivet, C., Boulonne, L., Lehmann, S., Saby, N.P.A., Bispo, A., Arrouays, D., 2012. A high resolution map of French soil organic carbon. Agron. Sutain. Dev. 32, 841–851.

Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32, 1378–1388.

Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A., et al., 2008. Soil monitoring in Europe: a review of existing systems and requirements for harmonisation. Sci. Total Environ. 391, 1–12.

Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J., Myneni, R.B., Running, S.W., 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. Science 300 (5625), 1560–1563.

Østergaard, H.S., 1989. Analytical methods for optimization of nitrogen fertilization in agriculture. In: Germon, J.C. (Ed.), Management Methods to Reduce Impact of Nitrates. Elsevier, London, pp. 224–234.

Panagos, P., Van Liedekerke, M., Jones, A., Montanarella, L., 2012. European Soil Data Centre: response to European policy support and public data requirements. Land Use Policy 29 (2):329–338. http://dx.doi.org/10.1016/j.landusepol.2011.07.003.

Paustian, K., Collins, H.P., Paul, E.P., 1996. Management controls on soil carbon. In: Paul, E.A., Paustian, K., Elliott, E.T., Vernon Cole, C. (Eds.), Soil Organic Matter in Temperate Agroecosystems: Long-term Experiments in North America. CRC Press, pp. 15–45.

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G.P., Smith, P., 2016. Climate-smart soils. Nature 532, 49–57.

Pebesma, E.J., 2006. The role of external variables and GIS databases in geostatistical analysis. Trans. GIS 10 (4), 615–632.

Phillips, J.D., 1990. A saturation-based model of relative wetness for wetland identification. Water Resour. Bull. 26, 333–342.

Poggio, L., Gimona, A., Brewer, M.J., 2013. Regional scale mapping of soil properties and their uncertainty with a large number of satellite-derived covariates. Geoderma 209–210, 1–14.

Post, W.M., Kwon, K.C., 2000. Soil carbon sequestration and land-use change: processes and potential. Glob. Chang. Biol. 6, 317–328.

Rees, R.M., Bingham, I.J., Baddeley, J.A., Watson, C.A., 2005. The role of plants and land management in sequestering soil carbon in temperate arable and grassland ecosystems. Geoderma 128 (1–2):130–154. http://dx.doi.org/10.1016/j.geoderma.2004.12.020.

Renard, K.G., Foster, G.R., Weesies, G.A., Porter, J.P., 1991. RUSLE: revised universal soil loss equation. J. Soil Water Conserv. 46, 30–33.

Robinson, D.A., Emmett, B.A., Reynolds, B., Rowe, E.C., Spurgeon, D., Keith, A.M., et al., 2012. Soil natural capital and ecosystem service delivery in a world of global soil change. In: Hester, R.E., Harrison, R.M. (Eds.), Soils and Food Security. Royal Society of Chemistry, Cambridge, pp. 41–68.

Rubinstein, R.Y., 1981. Simulation and the Monte Carlo method. John Wiley & Sons, New York.

Schulp, C.J.E., Verburg, P.H., 2009. Effect of land use history and site factors on spatial variation of soil organic carbon across a physiographic region. Agric. Ecosyst. Environ. 133 (1–2), 86–97.

Schulp, C.J.E., Verburg, P.H., Kuikman, P.J., Nabuurs, G.-J., Olivier, J.G.J., de Vries, W., Veldkamp, T., 2013. Improving national-scale carbon stock inventories using knowledge on land use history. Environ. Manag. 51, 709–723.

Sherrod, L.A., Dunn, G., Peterson, G.A., Kolberg, R.L., 2002. Inorganic carbon analysis by modified pressure-calcimeter method. Soil Sci. Soc. Am. J. 66 (1), 299–305.

Sleutel, S., De Neve, S., Hofman, G., 2007. Assessing causes of recent organic carbon losses from cropland soils by means of regional-scaled input balances for the case of Flanders (Belgium). Nutr. Cycl. Agroecosyst. 78 (3), 265–278.

Stevens, F., Bogaert, P., van Wesemael, B., 2015a. Spatial filtering of a legacy dataset to characterize relationships between soil organic carbon and soil texture. Geoderma 237-238, 224–236.

Stevens, F., Bogaert, P., van Wesemael, B., 2015b. Detecting and quantifying field-related spatial variation of soil organic carbon using mixed-effect models and airborne imagery. Geoderma 259-260, 93–103.

Trigalet, S., van Oost, K., Roisin, C., van Wesemael, B., 2014. Carbon associated with clay and fine silt as indicator for SOC decadal evolution under different residue management practices. Agric. Ecosyst. Environ. 196, 1–9.

Van Orshoven, J., Maes, J., Vereecken, H., Feyen, J., Didal, R., 1988. A structured database of Belgian soil prole data. Pédologie 38, 191–206.

Walkley, A., Black, I.A., 1934. An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. Soil Sci. 37, 29–37.

Wang, J.-F., Stein, A., Gao, B.-B., Ge, Y., 2012. A review of spatial sampling. Spat. Stat. 2, 1–14.

Wang, G., Zhou, Y., Xu, X., Ruan, H., Wang, J., 2013. Temperature sensitivity of soil organic carbon mineralization along an elevation gradient in the Wuyi Mountains, China. PLoS One 8 (1), e53914.

Wiesmeier, M., Barthold, F., Spörlein, P., Geuβ, U., Hangen, E., Reischl, A., Schilling, B., Angst, G., von Lützow, M., Kögel-Knabner, I., 2014. Estimation of total organic carbon storage and its driving factors in soils of Bavaria (southeast Germany). Geoderma Reg. 1, 67–78.

Wood, S.N., 2001. mgcv: GAMs and generalized ridge regression for r. R News 1 (2), 20–25.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC Press.

Zar, J.H., 1999. Biostatistical Analysis. Prentice Hall, New Jersey.

Please cite this article as: Chartin, C., et al., Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southe..., (2016), http://dx.doi.org/10.1016/j.geodrs.2016.12.006