A novel soil organic C model using climate, soil type and management data at the national scale in France

Meersmans, Manuel Martin, de Ridder, Lacarce, Wetterlind, de Baets, Bas, Benjamin Louis, Thomas Orton, Bispo, et al.

Agron. Sustain. Dev. (2012) 32:873–888
DOI 10.1007/s13593-012-0085-x

RESEARCH ARTICLE

A novel soil organic C model using climate, soil type and management data at the national scale in France

Jeroen Meersmans · Manuel Pascal Martin · Fjo De Ridder · Eva Lacarce · Johanna Wetterlind · Sarah De Baets · Christine Le Bas · Benjamin P. Louis · Thomas G. Orton · Antonio Bispo · Dominique Arrouays

Accepted: 20 February 2012 / Published online: 23 March 2012

Abstract This report evidences factors controlling soil organic carbon at the national scale by modelling data of 2,158 soil samples from France. The global soil carbon amount, of about 1,500 Gt C, is approximately twice the amount of atmosphere C. Therefore, soil has major impact on atmospheric CO_2, and, in turn, climate change. Soil organic carbon further controls many soil properties such as fertility, water retention and aggregate stability. Nonetheless, precise mechanisms ruling interactions between soil organic carbon and environmental factors are not well known at the large scale. Indeed, most soil investigations have been conducted at the plot scale using a limited number of factors. Therefore, a national soil survey of 2,158 soil samples from France was carried out to measure soil properties such as texture, organic carbon, nitrogen and heavy metal content. Here, we studied factors controlling soil organic carbon at the national scale using a model based on stepwise linear regression. Factors analysed were land use, soil texture, rock fragment content, climate and land management. We used several criteria for model selection, such as the Akaike information criterion (AIC), the corrected AIC rule and the Bayesian information criterion. Results show that organic carbon concentrations in fine earth increase with increasing rock fragment content, depending on land use and texture. Spreading farmyard manure and slurry induces higher carbon concentrations mostly in wet and stony grasslands. Nonetheless, a negative correlation has been found between carbon and direct C input from animal excrements on grasslands. Our findings will therefore help to define better land management practices to sequester soil carbon.

Keywords Land use · Manure · SOC · Rock fragment content · Climate · Soil texture · AIC · AICc · BIC · Stepwise regression

Abbreviation list

AIC	Akaike information criterion
AICc	Corrected Akaike information criterion
BIC	Bayesian information criterion
dg	Geometric mean particle size
Gt	Giga tonne
man	Slurry and farmyard related C production
man_df	Direct on field C input from animal excrements
pet	Potential evapotranspiration
prec	Precipitation
RMSE	Root mean square error
RMQS	French National Soil Survey (Réseau de Mesures de la Qualité des Sols)

J. Meersmans (✉) · M. P. Martin · E. Lacarce · J. Wetterlind · C. Le Bas · B. P. Louis · T. G. Orton · D. Arrouays
INRA Orléans, InfoSol Unit,
Avenue de la pomme de pin 2163, CS 40001, Ardon,
45075 Orléans, Cedex 2, France
e-mail: J.Meersmans@exeter.ac.uk

F. De Ridder
Earth System Science and Department of Geography,
Vrije Universiteit Brussel,
Pleinlaan 2,
1050 Brussels, Belgium

S. De Baets
Earth and Life Institute (ELI), Georges Lemaître Centre for Earth and Climate Research (TECLIM),
Université catholique de Louvain,
Place Louis Pasteur 3,
B-1348 Louvain-la-Neuve, Belgium

A. Bispo
Agriculture and Forestry Department, ADEME,
49004 Angers, Cedex 01, France

RPD	Ratio of performance to deviation
R^2_{adj}	Adjusted coefficient of determination
temp	Temperature

1 Introduction

Worldwide soils store more carbon (ca. 1,500 Gt C) than the biosphere (ca. 500 Gt C) and atmosphere (ca. 700 Gt C) combined (e.g. Grace 2004). Given the short-term dynamic behaviour of soil organic carbon (SOC), in essence determined by the balance between soil C input (mainly coming from plant residues) and the mineralization of organic material, the soil is recognised as an important reservoir of the global C cycle that can act as a significant source or sink of atmospheric CO_2 (e.g. Schulze et al. 2010). Furthermore, organic carbon is considered a key element for soil quality because it controls several functions and environmental factors, such as water-holding capacity, aggregate stability and fertility. Because organic matter is composed for 50% of carbon, soil organic carbon is commonly used as main indicator for soil fertility (Dawson et al. 2007). Moreover, its capacity to adsorb pesticides and breakdown excess nitrogen in the soils underlines its crucial role in combating current soil and groundwater contamination threats (Poissant et al. 2008). Consequently, the importance of studying soil organic carbon in a multidimensional context is reflected in different international frameworks and treaties such as the Kyoto Protocol (climate change) and European Union Soil Thematic Strategy (soil protection).

Different attempts were made to catch the SOC variability at the regional/national scale, starting with the calculation of average or median carbon values by land use and/or soil type (e.g. Lettens et al. 2005). Given the rather complex interactions between many determining factors, more enhanced statistical methods were recently developed to capture the heterogeneity of SOC and predict it as a function of a wide set of environmental variables such as land use, soil type, climate and agricultural management. Martin et al. (2011) constructed for example a boosted regression tree model and Jones et al. (2005) developed a rule-based system using pedotransfer rules to predict topsoil carbon in France and Europe, respectively. Meersmans et al. (2011) developed a multiple linear regression model to predict soil organic carbon in Belgium, where the model terms (i.e. variables and their interaction terms) were added following a trial-and-error procedure in order to maximise the adjusted coefficient of determination (R^2_{adj}) under the condition that all parameters are significant ($p<0.05$). Because a very high number of potential linear combinations of the different terms exits, this manual model construction procedure is time consuming and the ultimate best model may not be found. Therefore, we aim to further develop this part of the data processing by applying an automatized model selection procedure.

Selecting the most complex model will result in the best fit with the data used to calibrate the model, but will (most probably) end up with too much variation in the model. This will result in worse predictions when the model is extrapolated to novel positions. So the precision of the model is not optimal. On the other hand, when the simplest model is used, much of the variation in the measurements cannot be explained. This obviously results in inaccurate extrapolations. The aim of model selection criteria used here is to tune the model complexity so that precision and accuracy are balanced. In the history of model selection criteria, in general two thoughts are present. Both groups start from a measurement series and a set of models and assume that the measurement series can be described by a mathematical model disturbed by a stochastic noise source. Akaike (1974) wonders which model to select from this set so that the selected model will predict the outcome if the measurements were to be redone. Schwarz (1978), on the other hand, assume that one of the models in this model set is the ‘real’ model, which can explain the variation in the data. They came up with rules to identify that model, which has the largest chance of being that ‘real’ model. Both selection criteria are based on assumptions regarding the ‘real’ model. In practice, we rarely know if the ‘real’ model meets these assumptions, so the selection criteria favour certain classes of models. For that reason, we compared several criteria and looked if any general pictures emerged. Later, these selection criteria have further been refined. For example, when the measurement series becomes too short, both types of criteria select too complex models. This is remedied in, e.g. De Ridder et al. (2005). If the number of observations would tend to infinity, both the classical and refined model selection criteria would converge to the same results.

In this study, we aim to model topsoil SOC (0.3 m) as a function of land use, soil type, climate and agricultural management variables. To do this, a stepwise linear regression procedure in combination with the use of model selection criteria is programmed in order to automate the model construction process. The question, which can be answered by so-called model selection criteria is which model is the most appropriate to describe the variations observed in the measurements. Hence, the output obtained by using different model selection criteria will be compared and discussed in detail.

2 Materials and methods

2.1 Soil data

A total of 2,158 sites from the Réseau de Mesures de la Qualité des Sols (RMQS) soil survey, gathered by the French National Institute for Agriculture Research between 2000 and 2009, are used in this study. In the framework of this large-scale soil sampling campaign, many topsoil properties (e.g. texture, plant nutrients and heavy metals) have been measured all over France following a systematic 16×16 km grid (Fig. 1). Composite samples for each site were made from 25 subsamples taken within a 20×20 m grid (Arrouays et al. 2002).

In this study, SOC in the top 0.3 m of the soil is considered. Under cropland, average soil organic carbon, clay, silt, sand and rock fragment concentrations until this reference depth were calculated as a weighted average of their concentrations measured in the uppermost layer (i.e. plough layer, Ap) and the layers below, whereby the thickness of the horizon (with a maximal depth equal to 0.3 m) were used as weights. For all other land uses, the uppermost horizon corresponds to the 0–0.3 m depth interval. Organic carbon concentrations (g C kg^{-1}) of the fine earth were measured using the dry combustion method (e.g. Meersmans et al. 2009a). Here we refer, according to the definition given by Poesen and Lavee (1994), with fine earth to soil particles with a diameter smaller than 2 mm and with rock fragments to particles with a diameter of 2 mm or larger.

Given the rather large diversity in soil types covering France, i.e. vast areas of sandy soils (Podsols) in the southwest, fertile loess soils (Luvisols) in the north, various shallow soils (Leptosols) developed from calcareous rocks and large areas of dystric Cambisols in West and Central parts of the country (Fig. 2), in this study, Food and Agriculture Organization of the United Nations texture classification (CEC 1985) was used to analyse model output. Moreover, clay and silt contents were attributed to the corresponding texture class by calculating their median values using the entire French national soil inventory database (i.e. DoneSol 2.0, N=17,484 horizons; Table 1).

2.2 Manure data

Manure application and animal excrement production statistics (ton per hectare per year) at departmental level (ADEME 2007) were combined with dry matter C concentration values, i.e. 37.7% for farmyard manure and 36.6% for slurry (Lashermes et al. 2009). Land use area statistics (AGRESTE 2009) were then used to calculate average yearly C input related to farmyard manure and slurry production on agricultural soils as well as direct C input from animal excrements on grassland by department.

2.3 Climate data

Average yearly temperature (degree Celcius) and total annual precipitation and potential evapotranspiration (millimetre) are abstracted for each RMQS site from a 0.125×0.125° climatic grid distributed by Meteo-France, which has been obtained by interpolating observational data from the period 1993–2004. In this study, we defined seven climate zones for this study (Table 2): cold dry (mountainous dry, T<9°C; P<1,000 mm), cold wet (mountainous wet, T<9°C; P>1,000 mm), moderate–cold wet (continental, T>9°C and <10.5°C; P>850 mm), moderate wet (oceanic, T>10.5°C and <13.5°C; P>850 mm), moderate dry (continental–oceanic, T>9°C and <13.5°C; P< 850 mm), warm wet (oceanic warm T>13.5°C; P>850 mm), warm dry (Mediterranean, T>13.5°C; T<850 mm). The spatial distribution of these climate zones with regards to the soil sampling locations is given in Fig. 1. A more detailed description of the wide range of climatological, pedological and land use settings present in France can be found in Meersmans et al. (2012).

2.4 Soil organic carbon modelling

French soil survey (RMQS), climate and manure data have been used for calibration and validation of the model. A multicollinearity analysis has been performed to identify potential input variables by considering those that show strong correlations with organic carbon and avoiding input variables that are characterised by a strong internal correlation (r>0.70, e.g. Moore and McCabe 2001). A multiple linear regression model, predicting topsoil (0.3 m) carbon concentration in fine earth as a function of land use, soil type, management and climate (Eq. 1), was created using a combined forward and backward stepwise regression method. Therefore, a term could be added or removed from the expression based on the model selection criteria that was used. First- and second-order interactions between the input variables were also taken into consideration during model construction. The effect of land use was incorporated in the model by estimating terms land use independent or specific for 1, 2 or 3 land uses.

$$\mathrm{SOC} = \underbrace{\sum_{i=1}^{n} \alpha_i f_i}_{\text{linear}} + \underbrace{\sum_{j=1}^{n}\sum_{k>j}^{n} \beta_{j,k} f_j f_k}_{\text{first-order interaction}} + \underbrace{\sum_{l=1}^{n}\sum_{m\neq l}^{n} \gamma_{l,m} f_l f_m^2}_{\text{second-order interaction}} + \varepsilon \quad (1)$$

where, SOC is soil organic carbon concentration (percentage); n is the number of linear terms; f_i is variable i (e.g. clay, silt, rock fragment content, temperature, precipitation, manure, …); α, β and γ are model parameters to be estimated with a least squares estimator and ε is the error term, assumed here to follow a normal distribution with zero mean and constant variance.

The following three different model selection criteria were compared: the Akaike information criterion (AIC), the corrected Akaike information criterion (AICc) and the Bayesian information criterion (BIC) rule (Akaike 1974; Schwarz 1978; De Ridder et al. 2005). All these criteria take into account the distance between the predicted and observed values, but penalise the complexity of the model in order to avoid overfitting (i.e. when the model is describing stochastic random variations instead of significant deterministic variations; Eq. 2).

The only two factors that contribute are the number of observations and the number of model parameters. One aims at minimising the following expression:

$$\left(\ln\left(\frac{\mathrm{RSS}}{N}\right)+P\right) \tag{2}$$

where, RSS is the residual sum of squares, N is the number of observations, and P is the model complexity penalisation term, which differs depending on the selection criteria (Eq. 3).

$$\begin{array}{l} \mathrm{AIC}: P = 2n_\theta \\ \mathrm{AICc}: P = \frac{2N(n_\theta+1)}{(N-n_\theta-2)} \\ \mathrm{BIC}: P = \ln(N)n_\theta \end{array} \tag{3}$$

where, n_θ is the number of free parameters in the model.

2.5 Model validation

R^2 adjusted was computed for each model to assess the quality of the fit. Additionally, in order to assess model predictive performance, repeated tenfold cross-validation of the models was performed (1,000 replicates) where 90% of the data were used for calibration and 10% were used for validation as it is commonly recommended (Hastie et al. 2001). Root mean square error (RMSE, Eq. 4) and ratio of performance to deviation (RPD, Eq. 5) were calculated. These validation measures, in combination with the coefficient of determination (R^2), were used to compare the different models obtained by using the aforementioned model selection criteria.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{SOC}_{\mathrm{obs}(i)}-\mathrm{SOC}_{\mathrm{pred}(i)}\right)^2} \tag{4}$$

where, RMSE is the root mean square error (percentage), n is the number of samples used for validation, $\mathrm{SOC}_{\mathrm{obs}(i)}$ is the observed organic carbon concentration of sample i (percentage), and $\mathrm{SOC}_{\mathrm{pred}(i)}$ is the predicted organic carbon concentration of sample i (percentage).

$$\mathrm{RPD} = \frac{\mathrm{STD}}{\mathrm{RMSE}} \tag{5}$$

where, RPD is the ratio of performance to deviation, and STD is the standard deviation of organic carbon measurements (percentage).

2.6 Software

The R-software (version 2.9.0) was used for the modelling and error calculation part of this research. The pre-programmed function 'step' allowed the automation of the stepwise regression procedure.

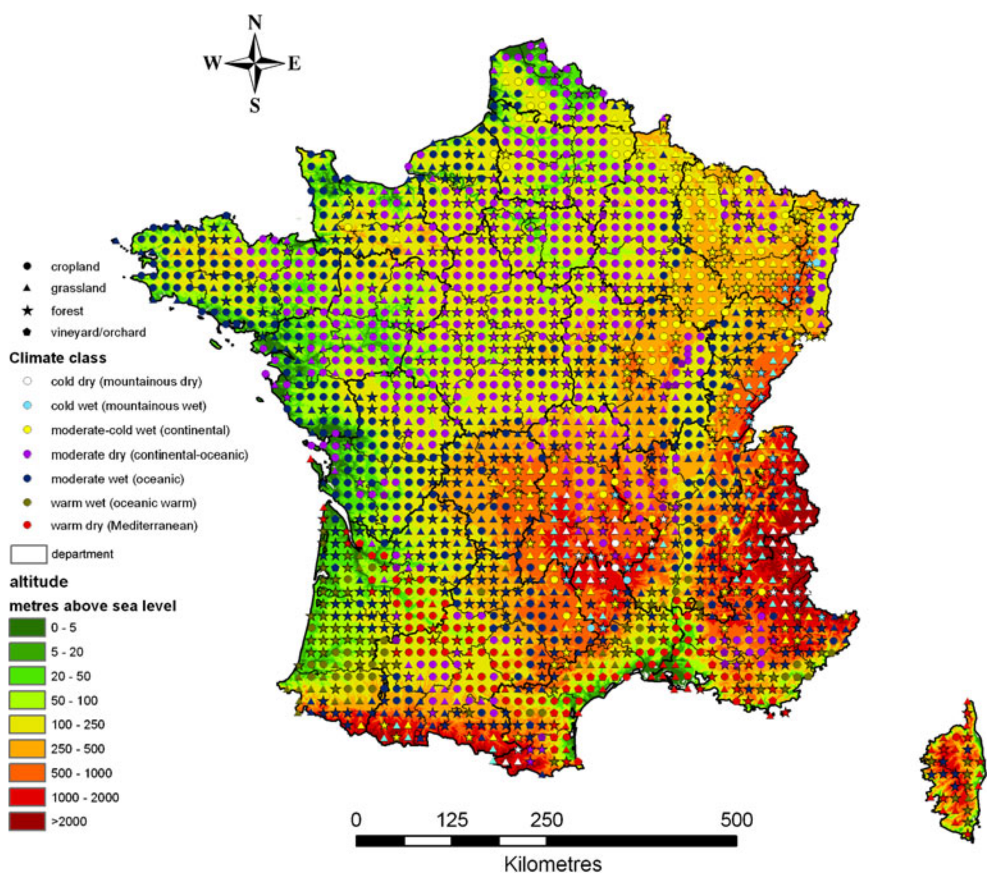


Fig. 1 Spatial distribution of sample locations with annotation of climate class (Table 1) and land use class in combination with digital elevation model, administrative regions (*thick line*) and departments (*thin line*)

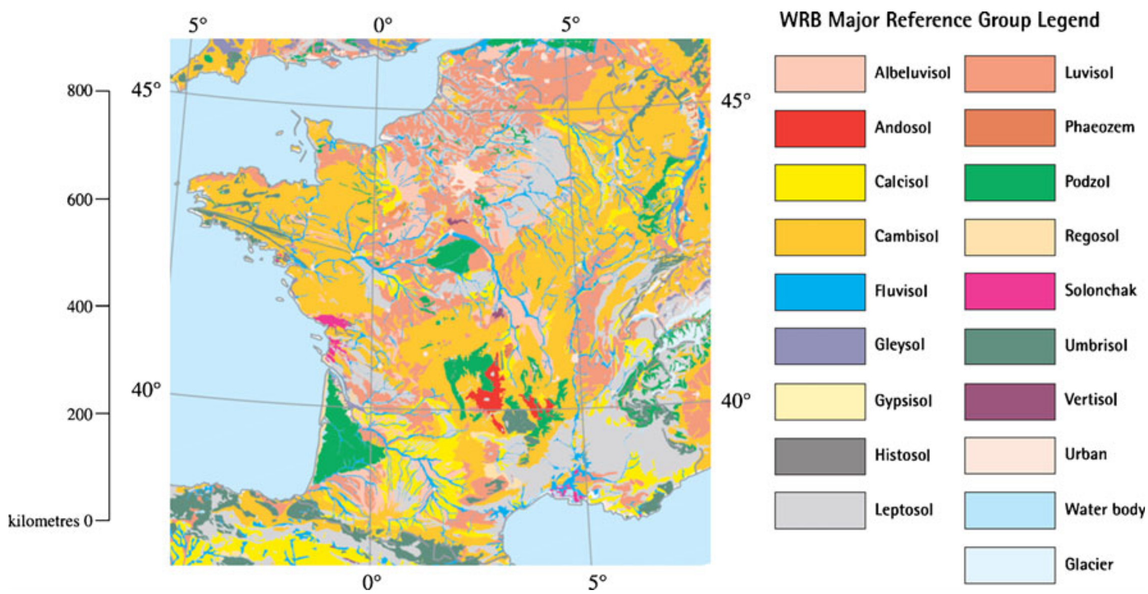


Fig. 2 Soil map of France showing major soil types following the World Reference Base for Soil Resources classification system, abstracted from the Soil Atlas of Europe (European Soil Bureau Network European Commission, 2005)

3 Results and discussion

3.1 Multicollinearity analysis

Regarding texture, the multicollinearity analysis shows that carbon is strongly correlated with clay content (i.e. r values ranging between 0.32 and 0.40) and weakly correlated with silt content under all land uses (Table 3). Moreover, SOC has a rather strong negative correlation with geometric mean particle size (dg, Eq. 8) and with sand content, especially under grassland and forest (i.e. r values ranging between −0.32 and −0.25 for the latter land uses).

$$\mathrm{dg} = \exp\left(\sum f_i \ln(M_i)\right) \tag{8}$$

where, dg is the geometric mean particle size, f_i is the relative proportion of particle diameter size class i and M_i is middle of particle diameter size class i (millimetre; i.e. sand=1.025 mm, silt=0.026 mm, clay=0.001 mm).

Table 1 Median clay and silt content (%) by texture class following Commission of the European Communities, revised Food and Agriculture Organization of the United Nations triangle (CEC 1985) and the entire France National Soil Inventory Database (i.e. DoneSol 2.0, N=17,484)

Texture class	N samples	Median clay (%)	Median silt (%)
Very fine	194	66.1	12.1
Fine	4,106	43.5	39.9
Medium fine	2,175	25.2	66.9
Medium	8,192	20.7	38.4
Coarse	2,817	7.5	12.1

Consequently, clay, sand and dg might all have the potential to serve as input variables for the model. Because the internal correlations between sand and silt, between sand and clay, and between dg and sand are above the threshold of Moore and McCabe (2001; i.e. r=0.7) for nearly all land uses, these variables may not be inserted together in the model. Only clay, silt and dg were selected as soil texture input variables for the model. In fact, a high correlation between the particle size fractions, sand, silt and clay is not surprising, due to the constraint of their summing to 100%. There are methods to deal with this constraint, e.g. the additive log-ratio transform (e.g. Lark and Bishop 2007), and we potentially lose some information by not accounting for the constraint prior to the multicollinearity analysis. However, we tested the additive log-ratio transform in this particular case study, and did not find it to give a better model (in terms of R^2 adjusted and RPD); therefore, we decided to proceed with analysis using the selected raw variables, clay and silt fractions, due to their direct interpretability. In addition, SOC shows a strong positive correlation with rock fragment content under cropland (r=0.37) and to a lesser extent under forest and vineyard/orchard (r=0.16 and 0.22, respectively). No significant correlation was found between carbon and rock fragment content for grassland. Nevertheless, because rock fragment content was not correlated to other potential input variables, it was also considered as an important model input variable.

By comparing the correlation coefficients between SOC and the climate-related potential input variables, one can conclude that the influence of precipitation is of primary importance under cropland, grassland and forest (i.e. r= 0.38, 0.35 and 0.38, respectively), whereas temperature

Table 2 Average standard deviation and median of average yearly temperature (degree Celcius) and total yearly precipitation amount (millimetre) by climate class

Climate class	N	Average yearly temperature (°C)			Total yearly precipitation amount (mm)		
		Average	SD	Median	Average	SD	Median
Cold dry	38	7.4	1.5	7.7	845	82	847
Cold wet	78	7.9	1.1	8.2	1,395	289	1,332
Moderate-cold wet	269	10.0	0.4	10.1	1,106	232	1,006
Moderate dry	769	11.5	0.8	11.4	757	58	764
Moderate wet	732	11.8	0.9	11.7	1,021	165	971
Warm wet	116	14.0	0.4	13.9	1,082	194	1,030
Warm dry	155	14.5	0.9	14.1	737	77	748

plays a key role under grassland, forest and vineyard/orchard (i.e. $r=-0.4$, -0.31 and -0.31, respectively). Because potential evapotranspiration shows a weaker correlation with carbon compared to precipitation or temperature and given the fact that this variable is too strongly correlated with temperature (i.e. for cropland, forest and vineyard/orchard, $r>0.7$), this variable was excluded from further analyses. Consequently, only temperature and precipitation were considered as climate-related input variables.

Slurry and farmyard manure-related C production shows a strong positive correlation with organic carbon under cropland and vineyard/orchard (i.e. $r=0.28$ and $r=0.41$, respectively). The results do not indicate such correlation under grassland. Nevertheless, the later land use shows a rather strong (negative) correlation between SOC and direct C input from animal excrements. Both slurry- and farmyard-related C production and direct C input from animal excrements were integrated in the model.

3.2 Models selected using different selection criteria (AIC, AICc and BIC)

The AIC model selection criterion method resulted in the most complex model, while the simplest model was obtained using the BIC model selection criterion. This is not surprising, since AIC tends to select models which are too complex, while BIC usually selects simpler model classes (De Ridder et al. 2005). This can be explained by the theory behind these selection criteria: BIC-type criteria are looking for that model which has the highest probability to be true. AIC-type models are developed to predict novel observations of the same underlying system. This is exactly what we do here: extrapolating the model to interpret regions with the same land use. So it is not at all surprisingly that AIC-type criteria outperform in this set-up. Both trends can be seen in these data. Due to the large number of predictors, it is hard to assess the origin of the differences in the selected models.

Notice that the AIC model and the AICc model have comparable R^2_{adj} (i.e. 0.492 and 0.491, respectively) and RPD (i.e. 1.40 and 1.39, respectively) values (Table 4). These model quality measures are remarkably lower for the BIC model (i.e. $R^2_{adj}=0.459$ and RPD=1.36). Furthermore, the AIC model has 36 parameters, of which only 30 are significant ($p<0.05$), whereas the AICc model has 30 parameters that are all significant. Only in the limited case of very large datasets, AIC and AICc identify the same model complexity. If the datasets are rather small, the AIC has the tendency to select models that are too complex, while AICc remains unbiased under these conditions. Although our dataset is not small ($N=$ 2,158), this is still the case in this particular study. Consequently, one might conclude that AIC risks overfitting the system.

3.3 Explaining factors

Table 5 presents the terms of the models with annotation of the selection criteria. Moreover, they were grouped by their land use independent or land use specific (i.e. estimated for 1, 2 or 3 land uses) character and into texture-, climate- and management-related variable groups and their interactions. The number of terms by variable/interaction type group can be considered as an indicator for their importance in predicting SOC.

The fact that the models have climate-related land use-independent terms (i.e. precipitation and interaction between precipitation and temperature) underlines the overall importance of climate on soil organic carbon (Table 5). The model expressions reveal, as well, land use-specific effects for texture, for the interaction between rock fragment content and texture and for the interaction between texture and climate. Moreover, these results show that the influence of manure on organic carbon depends on land use, texture and precipitation settings. Both slurry- and farmyard manure-related C production and direct C input from animal excrements are expressed in the model, in land use-specific interaction terms with texture, precipitation and rock fragment content. In addition, the model also contains interaction terms between farmyard manure and slurry production-related C input and texture.

Since the AIC and AICc have comparable RPD and R^2_{adj} values but a different degree of model complexity (i.e. 36 versus 30 parameters, respectively) the predictions obtained

Table 3 Correlation coefficients between site variables under different land uses: soil organic carbon (SOC) concentration (percentage), clay, silt and sand content (percentage), geometric mean particle size (dg, millimetre), rock fragment content (rock, percentage), average total yearly precipitation amount (prec, millimetre), average year temperature (temp, degree Celsius), yearly total potential evapotranspiration (pet, millimetre), slurry- and farmyard-related C production (man, ton per hectare per year) and direct C input from animal excrements (man_df, ton per hectare per year)

	SOC	clay	silt	sand	dg	rock	temp	prec	pet	man	man_df
Cropland (*n*=882)											
SOC	1.00	0.37**	−0.05	−0.17**	−0.22**	0.37**	−0.19**	0.38**	−0.16**	0.28**	
clay		1.00	−0.05	−0.54	−0.52	0.04	0.02	−0.07	0.10	−0.23	
silt			1.00	−0.82[a]	−0.62	−0.32	−0.26	−0.04	−0.29	0.10	
sand				1.00	0.82[a]	0.25	0.21	0.07	0.19	0.05	
dg					1.00	0.08	0.09	0.04	0.06	0.01	
rock						1.00	0.03	0.13	0.12	0.03	
temp							1.00	0.00	0.75[a]	−0.15	
prec								1.00	−0.09	0.40	
pet									1.00	−0.36	
man										1.00	
Grassland (*n*=613)											
SOC	1.00	0.32**	0.06	−0.25**	−0.26**	0.07	−0.40**	0.35**	−0.12*	0.02	−0.18**
clay		1.00	0.12	−0.72[a]	−0.58	−0.16	−0.04	0.09	0.00	−0.06	−0.01
silt			1.00	−0.78[a]	−0.65	−0.20	−0.09	0.13	−0.23	0.21	0.31
sand				1.00	0.83[a]	0.24	0.09	−0.15	0.16	−0.11	−0.21
dg					1.00	0.06	0.16	−0.20	0.19	−0.12	−0.16
rock						1.00	−0.09	0.13	0.11	−0.12	−0.20
temp							1.00	−0.32	0.64	−0.14	0.14
prec								1.00	−0.28	0.22	−0.09
pet									1.00	−0.51	−0.42
man										1.00	0.63
man_df											1.00
Forest (*n*=597)											
SOC	1.00	0.40**	0.13*	−0.31**	−0.32**	0.16**	−0.31**	0.38**	−0.02		
clay		1.00	0.37	−0.79[a]	−0.65	0.09	−0.17	0.13	0.07		
silt			1.00	−0.86[a]	−0.74[a]	0.07	−0.20	−0.05	−0.11		
sand				1.00	0.84[a]	−0.10	0.23	−0.04	0.04		
dg					1.00	−0.22	0.29	−0.04	0.00		
rock						1.00	0.10	−0.02	0.31		
temp							1.00	−0.35	0.70		
prec								1.00	−0.19		
pet									1.00		
Vineyard/orchard (*n*=65)											
SOC	1.00	0.32*	−0.03	−0.19	−0.11	0.22	−0.31*	0.04	−0.22	0.41**	
clay		1.00	0.31	−0.83[a]	−0.64	−0.17	−0.28	0.05	−0.19	0.07	
silt			1.00	−0.79[a]	−0.71[a]	−0.25	0.00	−0.05	0.09	−0.17	
sand				1.00	0.83[a]	0.25	0.18	0.00	0.07	0.06	
dg					1.00	0.15	0.03	0.09	−0.11	0.04	
rock						1.00	0.08	−0.12	0.17	0.03	
temp							1.00	−0.33	0.83[a]	−0.55	
prec								1.00	−0.37	0.33	
pet									1.00	−0.55	
man										1.00	

*$P<0.01$, significant correlation between soil organic carbon (SOC) and site variables; **$P<0.001$, significant correlation between soil organic carbon (SOC) and site variables

[a] Too strong correlation between potential input site variables of the model (in order to insert them both in the model)

by using both models are compared and discussed in more detail in the following paragraphs.

3.3.1 Land use, climate and texture

Figure 3 shows the average organic carbon concentration by land use–climate–texture class combination predicted by the model (Table 5) with the AIC and AICc criteria. The error bars present the associated standard deviation due to variability in other factors (i.e. manure and rock fragment content). Average yearly temperature and precipitation values (Table 2) and median clay and silt contents (Table 1) were used as settings of climate and texture variables for each class in these simulations. Generally, carbon tends to increase towards more cold/wet climates and more fine-textured soils for all land uses. Very fine and fine-textured soils have remarkably high carbon concentrations compared to other texture classes and sand-textured soils have very low carbon concentrations. In most land use–climate–texture class combinations, the AIC and AICc model output are comparable. Nevertheless, the less complex model (AICc) shows an opposite trend under cropland for cold wet climates and under vineyard/orchard for cold wet, cold dry and moderate-cold wet climates. These classes are characterised by sparse data. Furthermore, De Ridder et al. (2005) pointed out that AIC tends to chose too complex models under these circumstances.

The importance of land use, climate and texture variables for predicting SOC contents (Fig. 3) is highlighted as well in other studies investigating soil carbon in relation to environmental factors at the regional, national or continental scale. Using a boosted regression tree model, Martin et al. (2011) identified land use, climate and clay content as the three most important factors explaining the variation of SOC in France. Organic carbon is physically protected against microbial mineralisation within soil (micro)aggregates or chemically stabilised through adsorption to clay and silt particles in fine textured soils (Six et al. 2002; Razafimbelo et al. 2008). Many other studies also found remarkably high C concentrations in the fine-textured soils compared to the coarse-textured soils and consequently used clay (or clay and silt) content as a predictor for carbon (Zinn et al. 2005). Rusco et al. (2001) illustrated that carbon is positively correlated with precipitation amount and negatively correlated with temperature at the continental scale. This corresponds with Martin et al. (2011) for France and Meersmans et al. (2011) for Belgium who found that carbon is more strongly correlated with precipitation than with temperature.

3.3.2 Rock fragment content

Figure 4 confronts predicted SOC contents with rock fragment content for AIC and AICc models by land use–texture class combination. The results obtained by AIC and AICc models are similar and organic carbon often (e.g. for forest and grassland) does not increase much for a rock fragment content less than 25%. Greater than 25%, the results between the methods can differ rather strongly and remarkable increases in carbon concentrations are predicted for many land use–texture class combinations; however, the number of samples characterised by a high rock fragment content is limited and so the associated uncertainty will be large. The results of the AIC and AICc are comparable for cropland on coarse-textured soils with constant carbon values (for different rock fragment contents). More explicit increases in SOC with increasing rock fragment content are reported for the more fine-textured soils. In addition, the increase in the medium- to fine-textured soils is larger using the AIC model than the AICc model.

Considering grassland, soil organic carbon increases strongly with increasing rock fragment content in silt-dominated texture classes (i.e. medium fine and to a lesser extent medium) in both models. This increase is less pronounced in clay-dominated texture classes (i.e. very fine and fine). Organic carbon tends to decrease slightly for rock fragment contents increasing from 0 to 25% in coarse textured soils. The models diverge for higher rock fragment contents (25–50%). The AIC model simulates a strong increase in carbon and the AICc predicts constant carbon values.

Under forest, all texture classes are characterised by an increase in organic carbon when the rock fragment content

Table 4 Model fit quality properties by model selection criterion (i.e. Akaike information criterion (AIC), the corrected AIC rule (AICc) and Bayesian information criterion (BIC)) for land use specific and overall cases

Method	N	n (p<0.05)	Max P	R^2adj	RMSE	RPD
AIC	2,158	36 (30)	0.1525	0.4921	1.45±0.24 (1.41)	1.40±0.10 (1.39)
AICc	2,158	30 (30)	0.0429	0.4908	1.46±0.24 (1.42)	1.39±0.09 (1.38)
BIC	2,158	12 (12)	0.0040	0.4586	1.48±0.26 (1.43)	1.36±0.09 (1.36)

Average and standard deviations as well as median values (in parentheses) of RMSE and RPD obtained by external cross-validation procedure are given

N number of samples, *n* number of parameters with annotation of number of significant parameters at p<0.05 within parentheses, max P is maximal P value of all estimates parameters, R_{adj} adjusted coefficient of determination, *RMSE* root mean squared error, *RPD* ratio of performance to deviation

Table 5 Land use dependent or independent estimated terms of overall multiple linear model

Land use	TXT	CLIM	MAN	TXT×TXT	CLIM×CLIM	TXT×CLIM	TXT×MAN	CLIM×MAN	Cte
Cr+Gr+Fo+Vi/Or		prec[a,b,c]			temp.prec[c] temp^2.prec[a,b,c]	clay^2.prec[a,b] clay^2.temp[c] rock.temp[a]			C_{all}[a,b,c]
Cr+Gr+Fo				silt^2.rock[a]		rock.prec^2[a]			
Cr+Gr+Vi/Or						rock^2.temp[a]		prec.man[a]	
Cr+Fo+Vi/Or						silt^2.prec[a] silt.prec^2[b]			
Gr+Fo+Vi/Or	rock[b]			clay^2.rock[a] clay.rock^2[a] dg.rock^2[a] dg^2.rock[a]					
Cr+Gr						dg.prec^2[a]			
Cr+Fo				clay.rock^2[a] clay^2.rock[b] silt.rock[a] silt^2.rock[b]		rock.temp^2[a,c]			
Cr+Vi/Or						clay^2.temp[a,b,c]		prec.man^2[a,b]	
Gr+Fo	clay[a,b,c]			clay.rock[a]		clay.temp[a,b] clay.temp^2[a,b] clay.prec[b] dg^2.prec[a]			
Gr+Vi/Or				silt.rock^2[a,b,c]		rock.prec[a]	rock.man[b] rock^2.man[a]		
Fo+Vi/Or				silt.rock^2[a]					
Cr				dg.rock[c]		clay^2.prec[a]	dg.man[b]		
Gr		prec[b]	man_df[b]	silt^2.rock[a]	temp.prec^2[a]	silt.temp^2[a]	clay^2.man[b] clay.man[b] rock.man_df[b] rock.man_df^2[b]	prec^2.man[a] prec.man^2[b,c] prec.man_df[b] prec^2.man_df[a,c] prec.man_df^2[a,b]	
Fo				dg.rock^2[b] dg^2.rock[b]		rock.temp[b]			
Vi/Or						rock.temp[b]			

Cr cropland, *Gr* grassland, *Fo* forest, *Vi/Or* vineyard and orchard, *TXT* texture variable, *CLIM* climate variable, *MAN* management variable

[a] Obtained using AIC selection criteria

[b] Obtained using AICc selection criteria

[c] Obtained using BIC selection criteria

increases. There is a remarkably strong increase in very fine (AIC model) and coarse (AICc) textured soils with a rock fragment content above 25%. The models give different results under vineyard/orchard. Carbon increases with increasing rock fragment contents mainly in clay-dominated texture classes following the AIC model, whereas the AICc model indicates increased SOC values foremost for silt-dominated texture classes. The AIC model predicts an early decline of organic carbon (between rock fragment content of 0 and 20%) up to 0% C and later (between rock fragment content of 20 and 50%) an increase, while the AICc model simulates a slight decrease in SOC between 0 and 25% and constant values above rock fragment contents of 25% in the coarse texture class.

The results indicate a positive correlation between rock fragment content and carbon concentration. One could hypothesise that when C input rate and mineralization conditions are constant, a higher rock fragment content would result in higher organic carbon concentrations in the fine earth proportional to the increase in volumetric rock fragment content in order to obtain similar carbon stocks. Our results indicate that higher rock fragment content often results in higher carbon contents, but the relationship is not as straightforward as stated in the hypothesis (Fig. 4).

One can state that SOC stock only increases when organic carbon concentration more than doubles when rock fragment content increases from 0 up to 50%. Figure 4 suggests that this is only (obvious) the case under medium-fine (i.e. silt) grassland soils and (very) fine-textured cropland soils. For the other soil/land use types, the organic carbon concentration increase (for rock fragment contents ranging between 0 and 50%) is smaller, indicating decreasing carbon stocks with increasing rock fragment contents. This is probably due to the direct or indirect influence of rock fragments on mineralisation and C input conditions. Poesen and Lavee (1994) stated that rock fragment contents have a first-order influence on plant productivity because high rock fragment contents limit the soil volume for nutrient supply and root development. Furthermore, rock fragments have a rather complex interaction with thermal properties in the topsoil, water-holding capacity, infiltration and evaporation rates, which all depend on the type (and porosity) of rock and depth of occurrence in the profile (Poesen and Lavee 1994; Cousin et al. 2003). The results from this study also indicate important differences according to land use and texture class. A higher rock fragment content only results in remarkably higher organic carbon contents in soils with a high clay and silt fraction under cropland and vineyard/orchard (Fig. 4). This can be explained by the fact that the associated additional C will only be protected against human-induced mineralisation due to soil disturbance (i.e. ploughing and rainfall effects on bare soils) when stored in micro-aggregates or when bound to clay minerals. The fact that SOC tends to increase with rock fragment content in the more silt-dominated texture class (i.e. medium fine) under grassland may be the result of a naturally high plant-available water capacity in the fine earth fraction of silt-textured soils. Despite the potential decline in water-holding capacity due to the presence of rock fragments, a high plant-available water capacity will be retained. Hence, plant growth and C input are less affected compared to other texture classes characterised by the same rock fragment content, but with lower fine earth plant available water capacity.

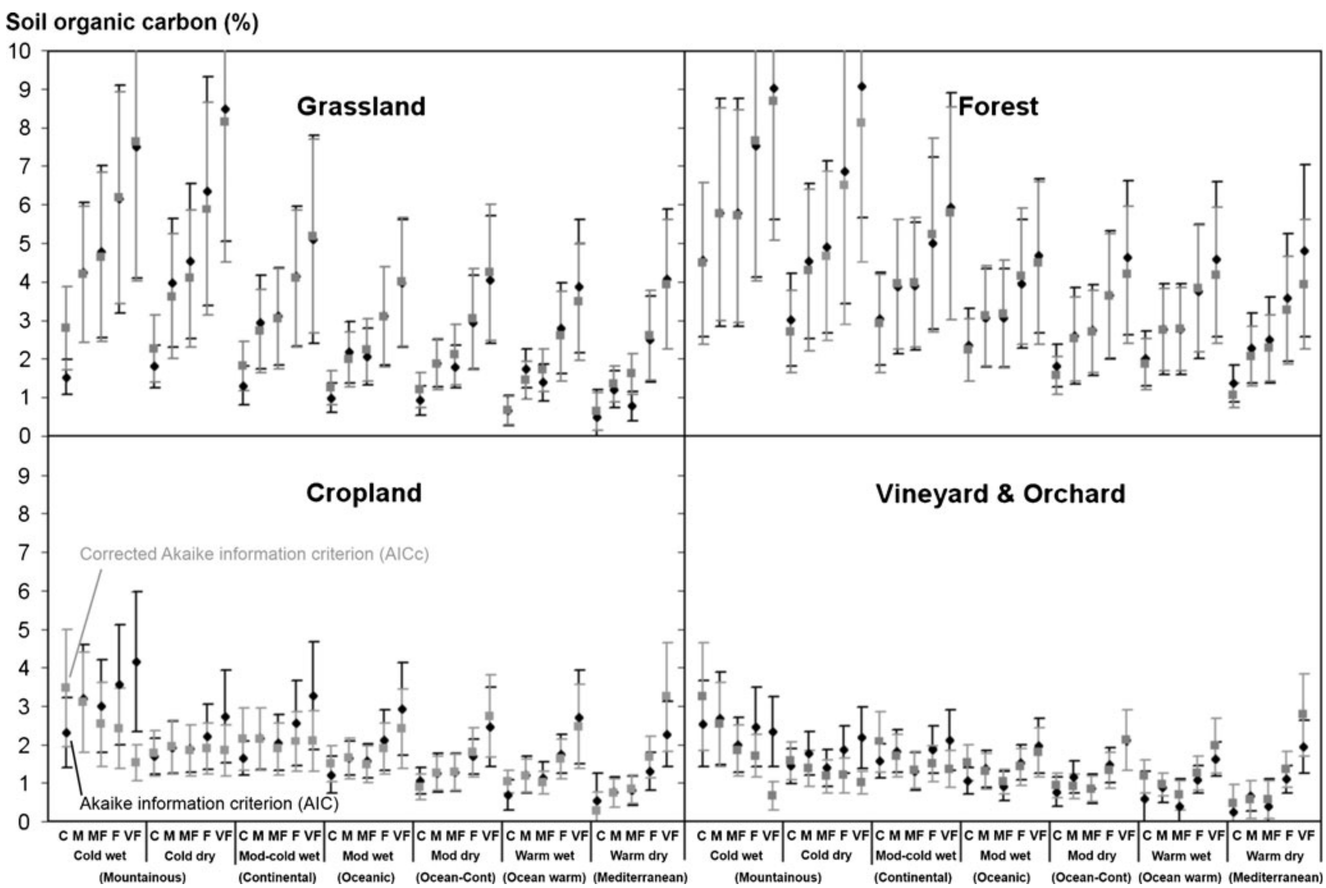
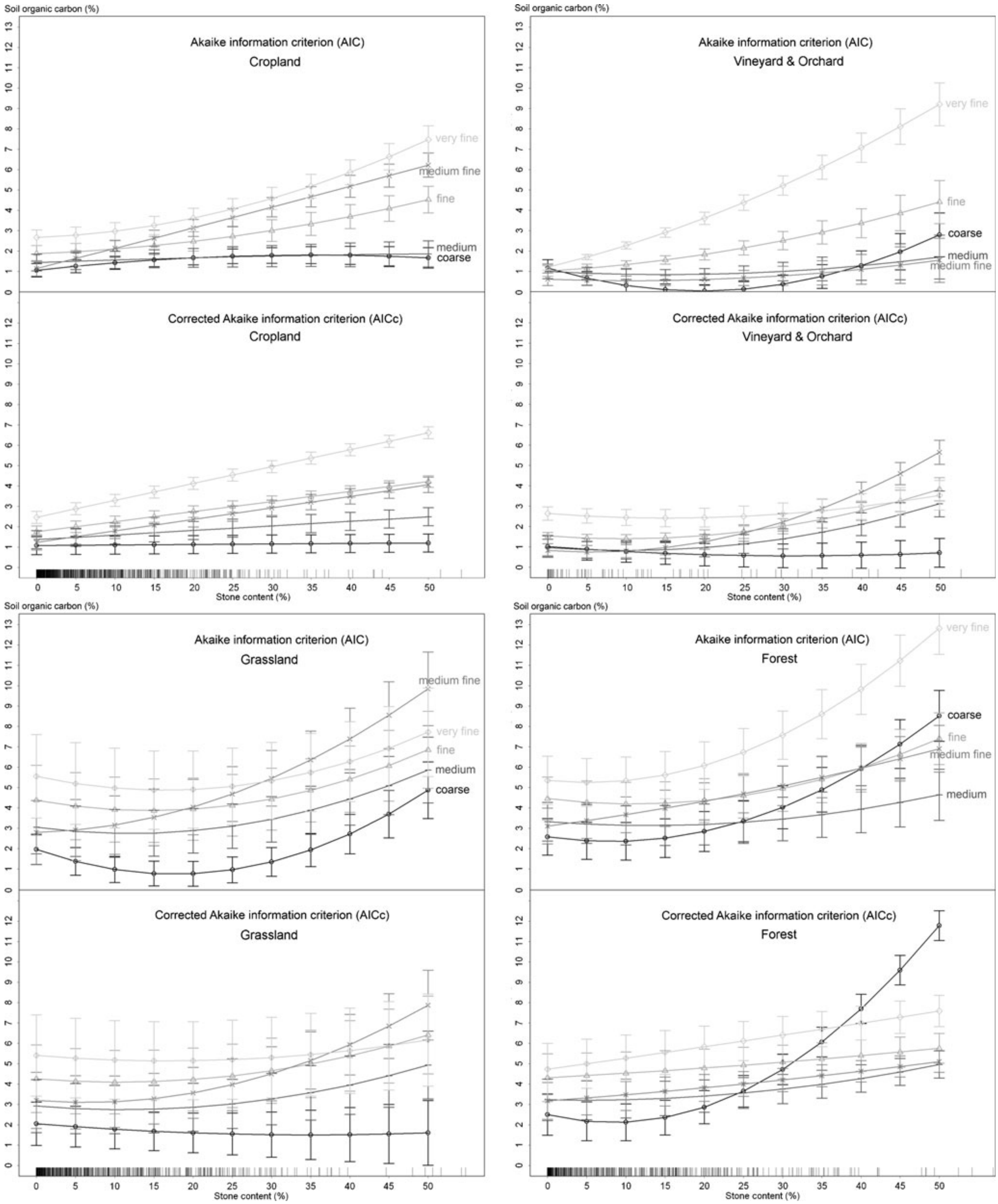


Fig. 3 Predicted soil organic carbon concentrations by land use–climate–texture class combination using the model (Table 5) with Akaike information criterion (AIC) and corrected Akaike information criterion (AICc). *Error bars* RMSE of the corresponding subdataset. Texture classes: *C* coarse, *M* medium, *MF* medium fine, *F* fine, *VF* very fine. This figure clearly shows that high carbon concentrations are predicted in cold/wet climates and fine textured soils, which can be explained by lower carbon mineralization rates under these climate conditions and the physical protection of soil organic carbon within soil (micro)aggregates and chemical stabilisation through adsorption to clay and silt particles. In most land use–climate–texture class combinations the AIC and AICc model outputs are comparable


Soil organic carbon (%)
Akaike information criterion (AIC)
Cropland
very fine
medium fine
fine
medium
coarse
Corrected Akaike information criterion (AICc)
Cropland
Stone content (%)
Soil organic carbon (%)
Akaike information criterion (AIC)
Vineyard & Orchard
very fine
fine
coarse
medium
medium fine
Corrected Akaike information criterion (AICc)
Vineyard & Orchard
Stone content (%)
Soil organic carbon (%)
Akaike information criterion (AIC)
Grassland
medium fine
very fine
fine
medium
coarse
Corrected Akaike information criterion (AICc)
Grassland
Stone content (%)
Soil organic carbon (%)
Akaike information criterion (AIC)
Forest
very fine
coarse
fine
medium fine
medium
Corrected Akaike information criterion (AICc)
Forest
Stone content (%)

◀ **Fig. 4** Average predicted soil organic carbon concentration as a function of rock fragment content by land use–texture class combinations. *Error bars* standard deviation on predictions due to variability in other factors (i.e. climate and manure). Sample distributions along rock fragment content are indicated by *tick marks* on the *x*-axis. The general positive relationship between rock fragments and organic carbon concentration is related to the fact that a higher rock fragment content means a smaller volume of fine earth. But as rock fragments affect thermal properties, water-holding capacity, infiltration and evaporation rates, they potentially hamper plant growth, reduce C input and increase mineralization. More in detail, this figure shows that organic carbon concentration tends to increase with increasing rock fragment content in all texture classes under forest, in silt (i.e. medium fine and to a lesser extent medium) textured soils under grassland, and only in soils with high silt and clay fractions (i.e. very fine and fine texture classes) under cropland and vineyard and orchard. The latter seems to indicate that under cropland associated additional C will only be protected against human-induced mineralisation due to soil disturbance (i.e. ploughing and rainfall effects on bare soils) when stored in micro-aggregates or when bound to clay minerals. Under silt-dominated grassland, the naturally high plant-available water capacity in the fine earth fraction characteristic for this texture class counteracts the lowered water-holding capacity status due to the presence of rock fragments

3.3.3 Manure

Figure 5 presents predicted soil organic carbon by land use–precipitation–rock fragment content class combination as a function of departmental average farmyard manure and slurry or direct-on-field animal excrement-related C input. The AIC and AICc model simulations for cropland indicate that high farmyard manure- and slurry-related C input does not result in significantly higher organic carbon contents compared to low slurry- and farmyard manure-related C input. Under grassland and vineyard/orchard, SOC tends to increase with increasing slurry- and farmyard manure-related C input under wet and/or stony soils. Rock fragment content seems to influence this relation much more than precipitation. The AIC and AICc model outputs for grasslands that consider the influence of direct-on-field manure-related C input from animal excrements on soil carbon are different, especially for stony soils. The general declining trend of carbon with increasing direct C input from animal excrements identified by the model (Fig. 5) is in agreement with the negative correlation between these variables detected by the multi-collinearity analysis (i.e. −0.18, Table 3). However, the AIC model output under all conditions and AICc model output for non-stony soils suggest that for very high direct-on-field C input values carbon increases with C input.

Since a higher rock fragment content means a smaller volume of fine earth, farmyard manure and slurry-related C input will have a higher impact on organic carbon concentrations in stony soils (Fig. 5). The limited effect of farmyard manure- and slurry-related C input on soil carbon under cropland can be explained by the low physical protection of the added C due to intensive soil disturbance by tillage, resulting in higher mineralisation rates compared to other land uses or croplands characterised by no till or reduced tillage (West and Post 2002). A net flux of C from cropland to grassland can be created due to manure cycling at farm level. Manure applied under cropland might not always result in significantly higher SOC levels because the mineralization rate is higher for crops and part of the manure C inputs may come from the crop residues previously removed from the cropland to produce the manure. Crop residues generally account for a large portion of C in the manure (i.e. foremost as straw in farmyard manure) produced at the farm, which can then be spread both on cropland and grassland. Under grassland, the lower mineralization conditions and the external origin of part of the manure C inputs can cause stabilisation and accumulation of manure-derived C, which results in a net gain of carbon on the long term. A recent study (Bolinder et al. 2010) shows that when receiving the same inputs of manure, soils under long forage rotations accumulate more carbon than under short forage rotation or crop rotation. However, the results of this study show a clear trend of increasing SOC with increasing manure under stony vineyard/orchard soils. Large applications of farmyard manure and slurry under this land use are seldom (indicated by tick marks on the *x*-axis) because high nitrogen status can have a bad influence on wine quality (Bell and Henschke 2005).

While the results explained above indicate that manure applied on the field by the farmer will result, in most cases, in a direct net enrichment of organic carbon concentration in fine earth (Fig. 5), this picture is somewhat more complex for C added directly to the field by excrements of livestock, because grazing animals affect, as well other factors, determine SOC content (i.e. the balance between input and mineralization). More precisely, animals have a direct impact on the quality and quantity of vegetation growing on the field. Due to grazing, this source of C input diminishes considerably. In extreme situations, such as overgrazing, a loss of soil structure, which makes the soil more vulnerable to C mineralization and to soil and vegetation degradation, can be observed (e.g. Dai et al. 2011). In addition, a conversion of stable C in grass into labile C in manure as well as the direct transformation of carbon from plant material into CH_4 and CO_2 by enteritic processes and respiration takes place. All these factors will have most probably an important contribution in explaining the declining trend of carbon with increasing departmental average direct-on-field C input from animal excrements (Fig. 5). The slight increase toward very high values may be related to the fact that at high animal density, the livestock will need supplementary food sources (i.e. forage or fodder crop) other than those obtained directly by grazing. This causes C fluxes from cropland to grassland. Because the AIC and AICc model output are different, more research is needed to clarify and understand these trends.

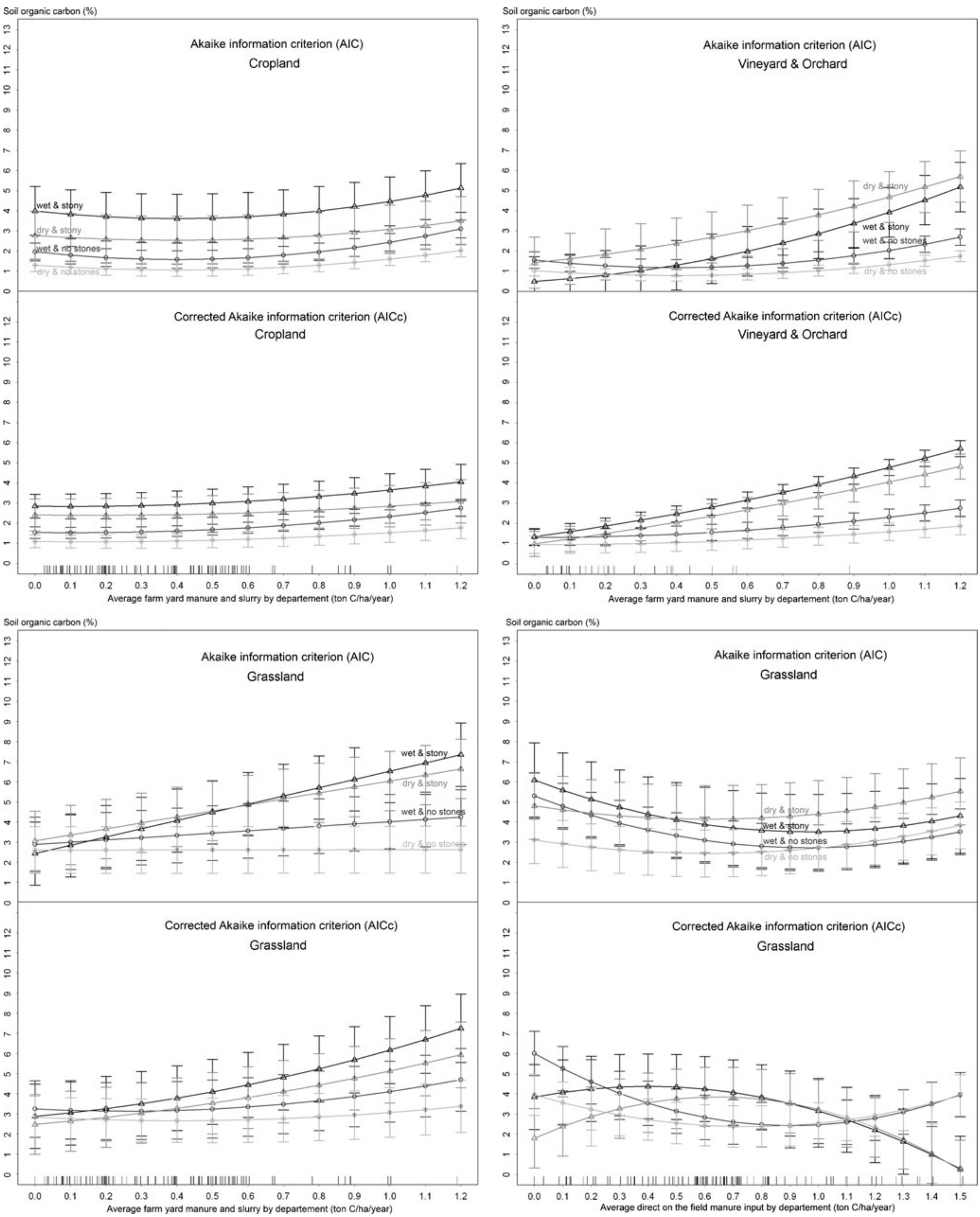

Soil organic carbon (%)
Akaike information criterion (AIC)
Cropland
wet & stony
dry & stony
wet & no stones
dry & no stones
Corrected Akaike information criterion (AICc)
Cropland
Average farm yard manure and slurry by departement (ton C/ha/year)
Akaike information criterion (AIC)
Vineyard & Orchard
Corrected Akaike information criterion (AICc)
Vineyard & Orchard
Akaike information criterion (AIC)
Grassland
Corrected Akaike information criterion (AICc)
Grassland
Average direct on the field manure input by departement (ton C/ha/year)

Fig. 5 Average predicted soil organic carbon concentration as function of departmental average slurry and farmyard manure or direct-on-field manure-related C input by land use–precipitation–rock fragment content class combination (no stones, rock fragment content=0%; stony, rock fragment content=30%; dry, yearly precipitation=700 mm; wet, yearly precipitation=1,100 mm). *Error bars* standard deviation on predictions due to variability in other factors (i.e. texture, temperature). Sample distributions along departmental average slurry and farmyard manure or direct-on-field manure-related C inputs are indicated by *tick marks* on the *x*-axis. Soil organic carbon concentrations increases with increasing farmyard manure and slurry production-related C input, mostly under wet and/or stony grassland soils. This is probably the consequence of reduced mineralization conditions and smaller fine earth volumes. The effect is much less pronounced under cropland, which can be explained by the low physical protection of the added C due to intensive soil disturbance by tillage. Moreover, high livestock densities (and direct on field manure C input rates) seems to lower the organic carbon level under grassland. The latter can be related to the fact that grazing reduces the quantity and quality of vegetation related C input, converses stable C in grass into labile C in manure, causes loss of soil structure and makes the soil more vulnerable to C mineralization

The results clearly illustrate that slurry- and farmyard manure-related C production are good predictors of organic carbon content, underlining their potential importance to be used to improve regional/national SOC estimates. Notice that manure application data were only available at departmental level. Probably, a more clear relation between SOC and manure could be identified if this variable was available at a finer spatial scale. This underlines the importance of collecting manure data at the field scale, e.g. detailed agro survey completed by the farmers at the sample sites.

3.4 Model error, validation and evaluation

Model error values for individual carbon predictions are rather high. The median relative error (i.e. model error divided by predicted value) is estimated at 46.6% for the AICc model and at 59.7% for the AIC model. The RPD value of the present study (i.e. 1.40 (AIC)–1.39 (AICc)) lies between the values obtained by Meersmans et al. (2011) for Belgium (i.e. 1.35 (1960) and 1.40 (2006)). Nevertheless, Meersmans et al. (2011) could integrate soil drainage (i.e. maximal and minimal depth of the ground water) in the model and show that this variable is one of the main factors explaining SOC variability at the regional scale. This variable was not available for France. We believe that the overall model uncertainty of the present study would be considerably lower if we would have been able to integrate soil drainage (e.g. depth of ground water) in the model. Other potentially interesting variables to increase the predictive power of the model can be topography (i.e. slope, curvature; e.g. Van Oost et al. 2007), net primary production (Martin et al. 2011), land use history (i.e. years since land use change or abandonment; e.g. Stevens and van Wesemael 2008), total soil depth and tillage depth (e.g. Meersmans et al. 2009b)

The adjusted coefficient of determination for the presented models (i.e. 0.49) corresponds with the predicted power obtained in other studies. Meersmans et al. (2011) applied a comparable method in order to map carbon in Belgium for 1960 and 2006, resulting in an R^2_{adj} of 0.42 for 1960 and 0.65 for 2006. Furthermore, Schulp and Verburg (2009) obtained R^2_{adj} values ranging between 0.21 and 0.42 for their predictions of organic carbon for different regions in the Netherlands.

Within the AIC-type selection criteria, one can expect that the classical AIC will be slightly worse than the AICc criteria. The AIC criterion will select the appropriate model if the measurement set will be infinitely large. Most real-world datasets are rather short. For the classical AIC criterion, this will results in the selection of too complex models (e.g. De Ridder et al. 2005). The corrected AIC criterion does not suffer from this shortcoming. This is exemplified in the remarkably higher median relative model error for the AIC model (i.e. 59%) compared to AICc (i.e. 47%), which is probably related to the existence of six insignificant parameters ($p>0.05$) in the AIC model. Furthermore, the unrealistic prediction of negative carbon values in very few cases (e.g. stony sandy grassland or vineyard/orchard soils, Figs. 3 and 4) by the AIC model is probably the consequence of overfitting. Given the small difference in RPD and R^2_{adj} values (Table 4), one might recommend using the AICc instead of the AIC model selection criteria method when constructing a similar multiple linear regression model predicting SOC as a function of environmental variables. However, Fig. 3 illustrates that the AICc model has some difficulties predicting organic carbon for land use–climate–texture class combinations characterised by sparse data. Hence, this model predicts a decrease of carbon toward more fine-textured soils under cropland and vineyard/orchard for relatively cold and/or wet climates. This is opposite to the AIC model output and to general literature findings (e.g. Zinn et al. 2005). Nevertheless, Leifeld et al. (2005) show for grasslands in Switzerland a significant relationships between SOC and clay (percentage) below 1,000 m asl (i.e. for more dry/warm climates) but not above this elevation (i.e. for more cold/wet climates). This result seems to be in accordance with the unclear organic carbon–texture relationship obtained for cold/wet climates in the present study, but more research is needed to unravel this interaction.

4 Conclusions

The results show that the influence of rock fragment content on soil organic carbon concentration depends on land use and texture settings. Organic carbon tends to increase with increasing rock fragment content in all texture classes under forest, in silt-textured soils under grassland, and only in soils with high silt and clay fractions under cropland, and vineyard and orchard. Moreover, farmyard manure- and slurry production-related C input is positively correlated to organic carbon concentrations mostly under wet and/or stony grasslands. This effect is much less pronounced under cropland, and orchard and vineyard. Furthermore, the results suggest that high livestock densities might decrease SOC values in grassland. This underlines the importance of integrating manure-related variables in studies investigating soil carbon stocks and dynamics. Model fit quality (R^2_{adj}) and validation (RMSE, RPD) measures of the AIC and the AICc model are almost the same, while the overall model performance of the BIC model is worse. However, not all parameters of the AIC model are significant and probably overfit the data. We therefore recommend using the less complex AICc model. Nevertheless, this model might have difficulties predicting reliable values for variable settings characterised by sparse data. The presented novel approach, unravelling the relation between soil organic carbon and a large range of site factors within a multidimensional context, can be a useful tool in optimising sustainable soil management at larger administrative levels in order to combat, in an appropriate way, soil fertility decline and climate change-related threats.

Acknowledgments The RMQS was financed by the "Groupement d'Intérêt Scientifique Sol". Jeroen Meersmans post-doctoral position was funded by the French Environment and Energy Management Agency (ADEME). We thank all the people involved in sampling and sample preparation and analysis.

References

ADEME (French Environment and Energy Management Agency) (2007) Bilan des flux de contaminants entrant sur les sols agricoles de France métropolitaine—final report: Bilan qualitatif de la contamination par les éléments tracés métalliques et les composés tracés organiques et application quantitative pour les éléments tracés métalliques

AGRESTE (Ministère de l'Alimentation, de l'Agriculture et de la Pêche) (2009) Chiffres et Données-Série Agriculture, 208, L'utilisation du territoire en 2008-Teruti-Lucas.

Akaike H (1974) New look at the statistical model identification. IEEE Trans Autom Control AC 19:716–723. doi:10.1109/TAC.1974.1100705

Arrouays D, Jolivet C, Boulonne L, Bodineau G, Saby N, Grolleau E (2002) A new initiative in France: a multi-institutional soil quality monitoring network. Comptes rendus de l'Academie d'Agriculture de France 88:93–105

Bell SJ, Henschke PA (2005) Implications of nitrogen nutrition for grapes, fermentation and wine. Aust J Grape Wine Res 11:242–295. doi:10.1111/j.1755-0238.2005.tb00028.x

Bolinder MA, Katterer T, Andren O, Ericson L, Parent LE, Kirchmann H (2010) Long-term soil organic carbon and nitrogen dynamics in forage-based crop rotations in Northern Sweden (63-64 degrees N). Agric Ecosyst Environ 138:335–342. doi:10.1016/j.agee.2010.06.009

CEC (1985) Soil Map of the European Communities at 1:1.000.000. CEC DG VI. Brussels-Luxembourg

Cousin I, Nicoullaud B, Coutadeur C (2003) Influence of rock fragments on the water retention and water percolation in a calcareous soil. Catena 53:97–114. doi:10.1016/S0341-8162(03)00037-7

Dai FQ, Su ZA, Liu SZ, Liu GC (2011) Temporal variation of soil organic matter content and potential determinants in Tibet, China. Catena 85:288–294. doi:10.1016/j.catena.2011.01.015

Dawson JJC, Godsiffe EJ, Thompson IP, Ralebitso-Senior TK, Killham KS, Paton GI (2007) Application of biological indicators to assess recovery of hydrocarbon impacted soils. Soil Biol Biochem 39:164–177. doi:10.1016/j.soilbio.2006.06.020

De Ridder F, Pintelon R, Schoukens J, Gillikin DP (2005) Modified AIC and MDL model selection criteria for short data records. IEEE Trans Instrum Meas 54:144–150. doi:10.1109/TIM.2004.838132

European Soil Bureau Network European Commission (2005) Soil Atlas of Europe, Office for Official Publications of the European Communities, Luxembourg

Grace J (2004) Understanding and managing the global carbon cycle. J Ecol 92:189–202. doi:10.1111/j.0022-0477.2004.00874.x

Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning, data mining, inference, and prediction, 2nd edn, Springer Series in Statistics. Springer, New York

Jones RJA, 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. doi:10.1111/j.1365-2389.2005.00728.x

Lark RM, Bishop TFA (2007) Cokriging particle size fractions of the soil. Eur J Soil Sci 58:763–774. doi:10.1111/j.1365-2389.2006.00866.x

Lashermes G, Nicolardot B, Parnaudeau V, Thuries L, Chaussod R, Guillotin ML, Lineres M, Mary B, Metzger L, Morvan T, Tricaud A, Villette C, Houot S (2009) Indicator of potential residual carbon in soils after exogenous organic matter application. Eur J Soil Sci 60:297–310. doi:10.1111/j.1365-2389.2008.01110.x

Leifeld J, Bassin S, Fuhrer J (2005) Carbon stocks in Swiss agricultural soils predicted by land-use, soil characteristics, and altitude. Agric Ecosyst Environ 105:255–266. doi:10.1016/j.agee.2004.03.006

Lettens S, Van Orshoven J, van Wesemael B, De Vos B, Muys B (2005) Stocks and fluxes of soil organic carbon for landscape units in Belgium derived from heterogeneous data sets for 1990 and 2000. Geoderma 127:11–23. doi:10.1016/j.geoderma.2004.11.001

Martin MP, Wattenbach M, Smith P, Meersmans J, Jolivet C, Boulonne L, Arrouays D (2011) Spatial distribution of soil organic carbon stocks in France. Biogeosciences 8:1053–1065. doi:10.5194/bg-8-1053-2011

Meersmans J, van Wesemael B, Van Molle M (2009a) Determining soil organic carbon for agricultural soils: a comparison between the Walkley & Black and the dry combustion methods (north Belgium). Soil Use Manage 25:346–353. doi:10.1111/j.1475-2743.2009.00242.x

Meersmans J, Van Wesemael B, De Ridder F, Dotti MF, De Baets S, Van Molle M (2009b) Changes in organic carbon distribution with depth in agricultural soils in northern Belgium, 1960–2006. Glob Change Biol 15:2739–2750. doi:10.1111/j.1365-2486.2009.01855.x

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 Change Biol 17:466–479. doi:10.1111/j.1365-2486.2010.02183.x

Meersmans J, Martin MP, Lacarce E, De Baets S, Jolivet C, Boulonne L, Lehmann S, Saby NPA, Bispo A, Arrouays D (2012) A high resolution map of French soil organic carbon. Agron Sust Devel. doi: 10.1007/s13593-012-0086-9

Moore DS, McCabe GP (2001) Introduction to the practice of statics. W.H. Freeman, New York, USA

Poesen J, Lavee H (1994) Rock fragments in top soils—significance and processes. Catena 23:1–28. doi:10.1016/0341-8162(94)90050-7

Poissant L, Beauvais C, Lafrance P, Deblois C (2008) Pesticides in fluvial wetlands catchments under intensive agricultural activities. Sci Total Environ 404:182–195. doi:10.1016/j.scitotenv.2008.05.030

Razafimbelo TM, Albrecht A, Oliver R, Chevallier T, Chapuis-Lardy L, Feller C (2008) Aggregate associated-C and physical protection in a tropical clayey soil under Malagasy conventional and no-tillage systems. Soil Tillage Res 98:140–149. doi:10.1016/j.still.2007.10.012

Rusco E, Jones R, Bidoglio G (2001) Organic matter in the soils of Europe: present status and future trends. European Soil Bureau, Soil and Waste Unit, Institute for Environment and Sustainability, JRC Ispra Institute, Joint Research Centre European Commission, Italy

Schulp CJA, Verburg PH (2009) Effect of land use history and site factors on spatial variation of soil organic carbon across a physiographic region. Agric Ecosyst Environ 133:86–97. doi:10.1016/j.agee.2009.05.005

Schulze ED, Ciais P, Luyssaert S, Schrumpf M, Janssens IA, Thiruchittampalam B, Theloke J, Saurat M, Bringezu S, Lelieveld J, Lohila A, Rebmann C, Jung M, Bastviken D, Abril G, Grassi G, Leip A, Freibauer A, Don A, Nieschulze J, Börner A, Gash JH, Dolman AJ (2010) Part4: integration of carbon and other trace-gas fluxes. Glob Change Biol 16:1451–1469. doi:10.1111/j.1365-2486.2010.02215

Schwarz G (1978) Estimating dimension of a model. Ann Stat 6:461–464. doi:10.1214/aos/1176344136

Six J, Conant RT, Paul EA, Paustian K (2002) Stabilization mechanisms of soil organic matter: Implications for C-saturation of soils. Plant Soil 241:155–176. doi:10.1023/A:1016125726789

Stevens A, van Wesemael B (2008) Soil organic carbon stock in the Belgian Ardennes as affected by afforestation and deforestation from 1868 to 2005. For Ecol Manage 256:1527–1539. doi:10.1016/j.foreco.2008.06.041

Van Oost K, Quine TA, Govers G, De Gryze S, Six J, Harden JW, Ritchie JC, McCarty GW, Heckrath G, Kosmas C, Giraldez JV, da Silva JRM, Merckx R (2007) The impact of agricultural soil erosion on the global carbon cycle. Science 318:626–629. doi:10.1126/science.1145724

West TO, Post WM (2002) Soil organic carbon sequestration rates by tillage and crop rotation: a global data analysis. Soil Sci Soc Am J 66:1930–1946. doi:10.3334/CDIAC/tcm.002

Zinn YL, Lal R, Resck DVS (2005) Texture and organic carbon relations described by a profile pedotransfer function for Brazilian Cerrado soils. Geoderma 127:168–173. doi:10.1016/j.geoderma.2005.02.010