



**HAL**  
open science

# Mapping soil organic carbon stock change by soil monitoring and digital soil mapping at the landscape scale

Yosra Ellili, Christian Walter, Didier Michot, Pascal Pichelin, Blandine Lemerrier

## ► To cite this version:

Yosra Ellili, Christian Walter, Didier Michot, Pascal Pichelin, Blandine Lemerrier. Mapping soil organic carbon stock change by soil monitoring and digital soil mapping at the landscape scale. *Geoderma*, 2019, 351, pp.1-8. 10.1016/j.geoderma.2019.03.005 . hal-02280194

HAL Id: hal-02280194

<https://institut-agro-rennes-angers.hal.science/hal-02280194v1>

Submitted on 25 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

**Manuscript Title:** Mapping soil organic carbon stock change by soil monitoring and digital soil mapping at the landscape scale

**The structure is:** author name, author affiliation and author address

Yosra Ellili, UMR SAS, INRA, AGROCAMPUS OUEST 35000 Rennes, France

INRA, UMR1069 SAS, 65 rue de Saint-Brieuc 35042 Rennes cedex, France

Christian Walter, UMR SAS, AGROCAMPUS OUEST, INRA35000 Rennes, France

AGROCAMPUS OUEST, UMR1069 SAS, 65 rue de Saint-Brieuc 35042 Rennes Cedex, France

Didier Michot, UMR SAS, AGROCAMPUS OUEST, INRA35000 Rennes, France

AGROCAMPUS OUEST, UMR1069 SAS, 65 rue de Saint-Brieuc 35042 Rennes Cedex, France

Pascal Pichelin, UMR SAS, AGROCAMPUS OUEST, INRA35000 Rennes, France

AGROCAMPUS OUEST, UMR1069 SAS, 65 rue de Saint-Brieuc 35042 Rennes Cedex, France

Blandine Lemercier, UMR SAS, AGROCAMPUS OUEST, INRA35000 Rennes, France

AGROCAMPUS OUEST, UMR1069 SAS, 65 rue de Saint-Brieuc 35042 Rennes Cedex, France

### Acknowledgements

The authors gratefully acknowledge all farmers at the Zone Atelier Armorique site involved in our research. We thank technical staff who actively participated in field sampling and laboratory analysis. This research was performed in the framework of the INRA “Ecoserv” metaprogram. This work was also funded by the Soilserv program funded by ANR (Agence Nationale de la Recherche) (ANR-16- CE32-0005-01).

# 1 Introduction

2 Understanding the multifunctional role of soil in ecosystem functioning is crucial, and soil  
3 scientists have recently give more importance to quantifying the contribution of soils to climate  
4 change mitigation (Paustian et al., 2016). The soil system is considered a significant terrestrial  
5 sink of carbon. Therefore, quantifying and mapping Soil Organic Carbon (SOC) change over  
6 time is crucial to manage agroecosystems sustainably and preserve soil resources at multiple  
7 scales, from local to global. The landscape scale appears to be a relevant scale for addressing this  
8 issue for the environment and agroecosystems (Viaud et al., 2010). This scale allows interactions  
9 between SOC dynamics and natural and anthropogenic processes to be considered.

10 Two approaches can be explored to assess change in stocks of SOC (SSOC): process-based  
11 models and inventory measurement systems (De Gruijter et al., 2016). Applying each method  
12 depends largely on data availability and cost-effectiveness. In general, process-based models are  
13 not expensive and are strongly recommended for predicting small-scale temporal change in SOC.  
14 This approach consists of applying a dynamic SOC model to each pixel under different scenarios  
15 of change in land use or land management practices that influence carbon inputs and outputs  
16 (Minasny et al., 2013). Spatially explicit information about erosion and deposition can also be  
17 included, as done by Walter et al. (2003), who performed field-scale simulations of  
18 spatiotemporal evolution in topsoil organic carbon at the landscape scale over a few decades and  
19 under different agricultural practices. In a recent study, Lacoste et al. (2016) assessed long-term  
20 impacts of soil redistribution on the evolution of SSOC in a hedgerow landscape using a  
21 spatiality explicit landscape model. These authors coupled the soil-redistribution model LandSoil  
22 with the SOC-dynamics model RothC to perform 90-year simulations. They used soil data  
23 collected in 2010 following Conditioned Latin Hypercube sampling and climate data predicted

24 from an available dataset (1980-2010) considering similar trends and using statistical methods.  
25 Overall, this study indicated that considering SOC content change over time should improve the  
26 modelling of SOC sequestration and consequently the accuracy of predicted sequestration rates.

27 Although process-based soil models are the approach used most to estimate SOC change  
28 (Grunwald, 2009), they have two main drawbacks. First, they do not cover the existing wide  
29 range combinations of climate, soil type, and management practices (De Guijter et al., 2016).  
30 This implies a lack of parametrization for several combinations that influence soil formation and  
31 the soil-landscape system. Second, such models use soil data collected over a given period and do  
32 not incorporate the temporal dimension of SOC content change in the modelling process. More  
33 importantly, SOC dynamics models should provide accurate predictions of SOC change in  
34 manageable spatial units, but the disconnect between mechanistic modelling of soil carbon  
35 dynamics and advances in soil carbon mapping remains a great challenge (Minasny et al., 2013).

36 The second approach that can be explored to assess SOC change is based on monitoring  
37 networks. They consist of performing direct measurements of SOC content over time by  
38 revisiting and resampling soils at the same location. To determine trends, SOC is usually directly  
39 measured repeatedly by revisiting intensively sampled sites, mainly to reduce short-range  
40 variability and thus maximize detection of change. Measuring SOC content for a reasonable  
41 number of sampling sites allows SSOC to be extrapolated between the sites for an entire area,  
42 using an appropriate method. For an area with a proper monitoring network, SOC change can be  
43 predicted and mapped using spatiotemporal models. For example, Bellamy et al. (2005) assessed  
44 SOC change across England and Wales using a dataset from national soil inventories for 1978-  
45 2003. Their results showed a mean decrease in SOC of  $0.6 \text{ \% year}^{-1}$  over the survey period.

46 Classical statistical approaches that can be applied to estimate temporal change from direct soil  
47 measurements are either design-based or model-based (Papritz et al., 1995). In the first case, the  
48 sampling locations are selected by probability sampling and the inference of the statistical  
49 parameters is based on the sampling design used to select the sampling locations (De Gruijter et  
50 al., 2006). Typical applications are to estimate global quantities in space-time, e.g. evolution of  
51 soil carbon stocks at farm scale in a context of carbon auditing (Wheler, 2014; De Gruijter et al.,  
52 2016) where the total difference of SSOC between two dates is the target attribute which  
53 uncertainty must be assessed. By contrast, model-based approaches make no assumption on the  
54 sampling design but the spatial distribution of the property is considered as a realization of a  
55 random process and inference is based on a stochastic model of the variation in space (Lark and  
56 Cullis, 2004). Typical applications are mapping spatial variations of a property by taking profit of  
57 the existing samples and the spatial autocorrelation of the target variable, e.g mapping SSOC  
58 over a domain (Eglin et al., 2008; Chaplot et al., 2009)

59 Assessing SOC change can take advantage from the development of digital soil mapping (DSM)  
60 approaches. Several DSM techniques have recently proven to be efficient to spatially predict soil  
61 attributes especially when direct soil measurements are scarce but environmental covariates, at  
62 high spatial resolution, are available (Hengl et al., 2018). This concept was initially developed by  
63 McBratney et al. in 2003 and relies on the construction of soil attributes prediction models as a  
64 function of a set of environmental covariates, which include inherent soil properties, landscape  
65 topography features and land use. The main advantage of this model is that both quantitative and  
66 categorical soil and environmental variables can be considered as explanatory covariables. In the  
67 DSM literature, many studies have attempted to predict and map SSOC a fixed period of time  
68 (Arrouays et al., 2002; Malone et al., 2011; Lacoste et al., 2014; Chartin et al., 2016; Hinge et al.,

69 2018) but very few studies have mobilized DSM techniques to map SOC change as done Wheeler  
70 in 2014. This researcher used a legacy soil data and environmental covariates to map SOC in  
71 1990 in the southern part of New South Wales, Australia, using rules based models. The SOC  
72 status in 2030 was then estimated under a likely climate change scenario and the map of SOC  
73 change was generated considering estimates of the bulk density.

74 The purpose of this study was to map SSOC change ( $\text{t C ha}^{-1} \text{ year}^{-1}$ ) between 2009 and 2016  
75 within a rural landscape by combining direct measurement of SOC change for a network of  
76 monitoring sites and DSM modelling. Uncertainties were assessed at two different scales, at the  
77 point scale and the landscape scale, to characterize the impact of the upscaling process.

## 78 2 Materials and Methods

### 79 2.1 Study area

80 The Zone Atelier Armorique (NW France,  $48^{\circ} 36' \text{ N}$ ,  $1^{\circ} 32' \text{ W}$ ), located in north-western France,  
81 covers  $10 \text{ km}^2$  (Fig. 1). This study area, part of the European Long-Term Ecosystem Research  
82 Network, has large soil heterogeneity over short distances (Lacoste et al., 2014). Soils are  
83 developed on shale and granite with a loamy cover. The main soil types are Cambisols and  
84 Luvisols, but the site also contains Leptosols and Fluvisols developed from alluvial and colluvial  
85 deposits (IUSS Working Group WRB, 2007). Its topography is strongly related to geological  
86 formations: a plateau on granite (southern end), a plain on soft schist (northern end), and a  
87 hillside on hard schist (transition between granite and soft schist) (Lacoste et al., 2014). The  
88 study site also has a marked micro-topography, related mainly to existing or former hedgerows  
89 and banks. Most of the farms are mixed crop-livestock farms and main land cover are annual  
90 crops (e.g. maize, wheat, barley) and temporary or permanent grasslands, but the study site also

91 includes woods and natural areas. A full inventory of crops and pastures is collected each year by  
92 a combination of field and remote sensing surveys: crop rotations differ depending on the  
93 farming system and include succession of annual crops (predominantly wheat and maize), annual  
94 crops following or followed by grassland, continuous grassland and finally natural areas (Fig. 1)

## 95 2.2 Quantifying changes in soil organic carbon

### 96 2.2.1 Data collection

97 Soil was sampled twice in 2009 and 2016 to quantify changes in SSOC over 7 years. The initial  
98 sampling strategy was based on 64 points selected using conditioned Latin Hypercube Sampling  
99 (cLHS). This method is a form of stratified random procedure that ensures an efficient way of  
100 sampling variables from their multivariate distributions (Minasny and McBratney, 2006). cLHS  
101 uses environmental covariates to select points that represent the study area well. In our study,  
102 four auxiliary variables were used: elevation (50 m of resolution), the topographic wetness index  
103 (Beven and Kirkby, 1979; Merot et al., 2003), the natural gamma emission of potassium  
104 (Bonijoly et al., 1999), and grassland frequency over 15 years (1993-2007).

### 105 2.2.2 Assessment of SSOC change

106 The 64 points were first sampled in May-June 2009 and resampled in April 2016 using a global  
107 positioning system with an accuracy of 3 m. Samples for SOC content analysis were taken with a  
108 7 cm diameter manual auger at four depths (0-7.5, 7.5-15, 15-30, and 30-45 cm): at each location,  
109 a composite sample was obtained by mixing four sub-samples collected within a radius of 1 m to  
110 attenuate local variability. A total of 256 composite soil samples were collected at each date, and  
111 total carbon content was measured by dry combustion with a CHN analyser (Thermo Finnigan  
112 EA 1112) according to the ISO 10694 certified method. As soil inorganic carbon may be  
113 considered negligible in the pedological context of the study area, SOC content was directly

114 derived from the total carbon measurement. Bulk density was measured at each site and for each  
115 depth using an 8 cm diameter root auger, which cored undisturbed samples of known volume  
116 (377 and 754 cm<sup>3</sup> for the first two and last two layers, respectively). All samples were oven dried  
117 at 105°C, weighted, and sieved at 2 mm. Samples were then weighted again to determine gravel  
118 content. In this study, we considered the bulk density of the fine earth fraction calculated from  
119 the dry mass and the core volume after being corrected for gravel content. At 6 randomly selected  
120 points, two additional replicates were taken to quantify the precision of local estimates of bulk  
121 density. Therefore, considering the maximum soil depth, 69 soil samples were used to assess bulk  
122 density uncertainty. As for bulk density, 30 replicates of SOC content analyses were performed to  
123 quantify the precision of SOC measurements.

124 Carbon stock was calculated considering a mass coordinate system, which attempts to correct for  
125 differences in bulk density over time (Minasny et al., 2013). The method consists of reporting  
126 carbon content for a fixed mass of soil mineral material (IPCC, 2006) by converting SOC content  
127 measured in soil layers to cumulative mineral mass and cumulative SOC content. This approach  
128 allows comparison of the change in SSOC between two dates for any given mineral mass,  
129 following three steps:

- 130 1. Calculate the mineral mass of each layer (0-7.5, 7.5 -15, 15-30, and 30-45 cm).

131 
$$m = Z \times \rho_b \times f_{\min} \quad [1]$$

132 where  $Z$  = thickness of the layer (m),  $\rho_b$  = volumetric mass density of the fine earth  
133 fraction (kg m<sup>-3</sup>), and  $f_{\min}$  = the mineral fraction (kg kg<sup>-1</sup>)

- 134 2. Calculate cumulative mineral mass at each point (i) by summing the mineral mass  $m$  of  
135 the  $n$  soil layers (j):

136 
$$M_i = \sum_{j=1}^n m_j \quad [2]$$



137 3. Calculate cumulative carbon stock. The amount of carbon for a given mineral mass (150,  
138 300, and 450 kg m<sup>-2</sup>) was estimated by linear interpolation:

$$139 \quad S_i = \sum_{j=1}^n S_j \quad [3]$$

140 Finally, the SOC stock change per year was calculated for a 300 kg soil mineral mass ( $\approx 30$  cm)  
141 at each sampling site using the interpolation function implemented in stats R package (R Core  
142 Team, 2017).

### 143 2.2.3 Estimation of SSOC uncertainty at point scale

144 According to Uusitalo et al. (2015), uncertainty stems from various source that can be divided  
145 into 6 categories: inherent randomness, measurement error, systematic error, natural variation,  
146 model uncertainty and subjective judgement. In our study, uncertainties were assessed for the 64  
147 points using Monte-Carlo simulations by varying bulk density, soil carbon content and thickness  
148 of soil layer. The Monte-Carlo approach used a normal probability distribution of all input  
149 variables used for SSOC calculation (Table 2) to generate a final probability distribution of the  
150 targeted variable. To achieve this, we used @ Risk software ver.4 (Palisade, 2005) with 1000  
151 iterations.

### 152 2.2.4 Environmental covariates for spatial modelling

153 To model the spatial distribution of SSOC change over time at the landscape scale, we used a set  
154 of 15 maps of predictor variables as shown in Table 1. Topographic and geomorphometric  
155 variables were derived from a 50 m resolution digital elevation model using ArcGIS 10.3  
156 software (ESRI, 2012) and the MNTsurf software developed by Squividant (1994). Soil parent  
157 material was derived from legacy geological maps and gamma radiometry using machine  
158 learning techniques (Lacoste et al., 2011). Predictor variables describing land use and crop

159 rotation stem from systematic field and remote sensing surveys carried out over the past eight  
160 years.

### 161 2.2.5 Spatial modelling of SSOC change at landscape scale.

162 Two approaches were explored for spatial modelling of SSOC change within our study area: Fig.  
163 2 details the different steps followed for each approach.

164 The first method consisted in performing stepwise multiple linear regression (MLR) (Hocking.,  
165 1976) implemented in MASS R package (R Core Team, 2017), in order to explain SSOC change  
166 by covariates linked to land cover between 2009 and 2016: annual land cover, crop rotation and  
167 grassland frequency between 2009-2016. The model was calibrated with SSOC change computed  
168 from the 64 soil samples collected in 2009 and 2016. Spatial autocorrelation of model residuals  
169 was checked to evaluate if a regression-kriging approach (Odeh et al., 1995) may be developed.

170 The second method consists in applying the *SCORPAN* model mobilizing a wider range of  
171 covariates characterizing both inherent soil properties and soil-landscape features using  
172 RandomForest model (Breiman, 2001; Liaw and Wiener, 2002). This model was calibrated using  
173 the same soil data expressing SSOC change and 15 environmental attributes covering the whole  
174 study area. The spatial autocorrelation of model residuals was also checked.

175 The performance of the two models: stepwise MLR- kriging model and Random Forest model  
176 was assessed using internal validation method and the best approach was chosen to spatially  
177 predict the target soil attribute within our study area using three error indexes: i) mean absolute  
178 error (MAE, Eq. 4), ii) root mean square error (RMSE, Eq. 5), and the coefficient of  
179 determination ( $R^2$ ), which measures the deviation between observed and predicted values (Eq. 6).

$$180 \quad MAE = \frac{1}{n} \sum_{i=1}^n |\hat{x}_i - x_i| \quad \text{where} \quad x_i = \text{measured delta SSOC} \quad [4]$$

$$181 \quad \hat{x}_i = \text{predicted delta SSOC}$$

182  $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i)^2}$   $x = \text{mean delta SSOC}$  [5]

183  $i = \text{point (1- 64)}$

184  $R^2 = \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$  [6]

185 Uncertainties of SSOC change predictions were estimated by constructing 95% confidence  
 186 intervals using K-means fuzzy clustering of the residuals from the Random Forest models  
 187 (Malone et al., 2011). This approach consists on dividing the conceptual space into clusters  
 188 sharing a similar distribution of error. Specifically, we i) performed fuzzy classification of the  
 189 conceptual space based on the training dataset, ii) constructed model error for each cluster, and  
 190 iii) determined confidence intervals of predictions for each cluster.

### 191 3 Results

#### 192 3.1 Point measurements of SSOC at two dates

193 SSOC at different soil mineral masses (150, 300, and 450 kg m<sup>-2</sup>) increased overall from 2009 to  
 194 2016 (Fig. 3a). At a mineral mass of 300 kg m<sup>-2</sup>, mean SSOC was 5.24 kg m<sup>-2</sup> (50.24 t C ha<sup>-1</sup>) in  
 195 2009 and 5.76 kg m<sup>-2</sup> (50.76 t C ha<sup>-1</sup>) in 2016, giving a mean increase in SSOC of 0.51 ± 0.12 kg  
 196 m<sup>-2</sup> (5.1 t C ha<sup>-1</sup>). Fig. 4 shows the evolution of SSOC change for the 64 sites, with their limits of  
 197 the 90% CI at 300 kg m<sup>-2</sup>. Overall, high variability among soil observations change was  
 198 highlighted. Added to this, associated limits of the 90 % CI were not constant and varied among  
 199 sites.

#### 200 3.2 Point assessment of SSOC change at measurement sites

201 SSOC varied greatly among the 64 points depending namely on the land use trajectory  
202 concerned. Mean ( $\pm$  90% confidence interval (CI)) annual change in SSOC was  $0.73 \pm 0.18$  t C  
203  $\text{ha}^{-1} \text{ year}^{-1}$ .

204 Fig. 4 shows the evolution of SSOC change for the 64 sites, with their limits of the 90% CI at  
205  $300 \text{ kg m}^{-2}$ . Overall, high variability among soil observations SSOC change was highlighted.  
206 Added to this, associated limits of the 90 % CI, was not constant and varied among sites.

207 Furthermore, crop rotations implemented between 2009 and 2016 led to differences in SOC  
208 sequestration. The levels of storage that the crop rotations entail are listed in Table 3. For  
209 instance, high increase in SOC storage characterises continuous grassland system ( $1.40 \pm 0.31$  t  
210  $\text{C ha}^{-1} \text{ year}^{-1}$ ). A similar trend was shown when croplands were converted to grasslands ( $0.94 \pm$   
211  $0.29$  t C  $\text{ha}^{-1} \text{ year}^{-1}$ ) and under continuous annual crops ( $0.82 \pm 0.26$  t C  $\text{ha}^{-1} \text{ year}^{-1}$ ). However,  
212 lower rate of SSOC change was obtained for annual crops to grassland rotation ( $0.42 \pm 0.18$  t C  
213  $\text{ha}^{-1} \text{ year}^{-1}$ ). As a result of the agriculture practices implemented, the accumulation of SOC was  
214 detected at almost all the sampling sites from 2009 to 2016 and a positive correlation  
215 characterised the relationship between SSOC at two dates ( $r^2=0.68$ ) (Fig. 3b).

### 216 3.3 Comparing SSOC change models

217 We investigated the effect of land cover, crop rotation and grassland frequency since 2009 on  
218 SSOC change. These covariates were used as explanatory variables and SSOC change as  
219 response variable. Using stepwise MLR model, only crop rotation covariate was selected to build  
220 up the regression model. Table 4 gives summary statistics of this MLR model where only crop  
221 rotation was selected as predictor variable. The results show that crop rotation accounts for only  
222 20 % SSOC change variability. Thus, SSOC change may be highly variable for a same modality  
223 of crop rotation.

224 The constant coefficient which represents continuous annual crops modality was significant at  $\alpha=$   
225 0.01. Permanent grasslands and grassland to crop rotation coefficients were significant at level of  
226  $\alpha=0.05$ . The F-statistic of the regression was 4.83 and the regression found to be significant at  
227  $\alpha=0.01$ .

228 As model residuals showed a spatial autocorrelation, they were kriged over the study area and  
229 were added to regression map to generate a final map of SSOC change at the extend of interest.

230 Internal validation gives  $R^2 = 0.23$ , concordance = 0.30, MEA= 1.3 t C ha<sup>-1</sup> year<sup>-1</sup>, RMSE= 1.1 t  
231 C ha<sup>-1</sup> year<sup>-1</sup> and bias=  $-7.77 \cdot 10^{-6}$  t C ha<sup>-1</sup> year<sup>-1</sup>.

232 A larger panel of predictors was used to construct the Random Forest regression tree for SSOC  
233 change (Fig. 5). Predictors with high influence on the predictions included inherent soil  
234 properties such as soil parent material, and certain terrain attributes describing soil hydromorphic  
235 conditions and the contrasting topography. Other variables also appeared important in the  
236 condition rules, namely crop rotation system. Model residuals showed no spatial autocorrelation  
237 and error indices of the predictive model derived from internal validation are  $R^2 = 0.92$ ,  
238 concordance = 0.84 MEA = 0.31 t C ha<sup>-1</sup> year<sup>-1</sup>, RMSE = 0.39 t C ha<sup>-1</sup> year<sup>-1</sup>, and bias= 0.012 t C  
239 ha<sup>-1</sup> year<sup>-1</sup>.

240 When comparing internal validation results of both models, MLR regression kriging model and  
241 Random Forest model, we note that the second model achieved highest  $R^2$  and concordance  
242 coefficients and lowest RMSE, MEA, and bias. Therefore, Random Forest model was used to  
243 spatially predict SSOC change over the study area.

#### 244 3.4 Mapping of SSOC change at the landscape scale

245 When mapped, SSOC change for a 300 kg mineral mass ranged from -1.00 to 2.8 t C ha<sup>-1</sup> year<sup>-1</sup>  
246 (mean = 0.98 t C ha<sup>-1</sup> year<sup>-1</sup>) and showed high spatial variability (Fig. 6b). Highest values were

247 found mainly in waterlogged soils under permanent grasslands. Similarly, zones with high  
248 vegetation cover, particularly woods, and less human development, located on steep slopes,  
249 showed high SSOC change. The spatial variability of SSOC change was driven mainly by soil  
250 redoximorphic conditions, which varied greatly within the study site.

## 251 4 Discussion

### 252 4.1 Estimates of SSOC change and associated uncertainties

253 Within a given area and a short time span, SSOC change is principally driven by evolutions in  
254 agriculture practices and crop rotations. Considering the direct measurements at 64 points, mean  
255 ( $\pm 90\%$  CI) annual change in SSOC (at  $300 \text{ kg m}^{-2}$  of soil mineral mass ( $\approx 30 \text{ cm}$ )) was  $0.94 \pm$   
256  $0.29 \text{ t C ha}^{-1} \text{ year}^{-1}$  when croplands were converted to grassland and  $1.40 \pm 0.31 \text{ t C ha}^{-1} \text{ year}^{-1}$   
257 under continuous grassland. By comparison, mean carbon sequestration of French soils in the 0-  
258 30 cm topsoil layer has been estimated at  $0.49 \pm 0.26 \text{ t C ha}^{-1} \text{ year}^{-1}$  on conversion of cropland to  
259 grassland (Arrouays et al., 2002) and can reach  $1 \text{ t C ha}^{-1} \text{ year}^{-1}$  for grassland systems involving  
260 high amounts of organic fertilisation (Gac et al., 2010). In our study, when grassland was  
261 converted into cropland, mean ( $\pm 90\%$  CI) carbon sequestration (at  $300 \text{ kg m}^{-2}$  of soil mineral  
262 mass) was  $0.42 \pm 0.18 \text{ t C ha}^{-1} \text{ year}^{-1}$ , which is a three times lower decrease than the French mean  
263 of  $-0.95 \pm 0.3 \text{ t C ha}^{-1} \text{ year}^{-1}$  in the 0-30 cm soil layer (Arrouays et al., 2002). The observed  
264 difference may be explained by the time step considered, as Arrouays et al.'s (2002) findings  
265 were based on a 20-year scenario while our results were based on a 7-year time step. Added to  
266 this, since 2009, farmers introduced more and more winter catch crops between two successive  
267 main crops and these catch crop are generally not harvested but introduced into the soil (Viaud et  
268 al., 2018). Another striking factor was the development of minimum tillage, which could also

269 explain positive evolution of SSOC within the study area through increased SOC inputs and  
270 slow-downing of SOC mineralization process.

271 Uncertainties of the SSOC change indicator were mainly related to the spatial variability of SOC  
272 contents, agriculture practices and measurement methods. We sought to minimise spatial  
273 variability in 2016 by returning as close as possible to the 2009 points and by mixing samples  
274 taken at 1m distance, but residual spatial variability remained over short distances. The accuracy  
275 of the global positioning system used (3 m) explains some of these uncertainties. The soil is  
276 defined as a phenomenon that varies at different scales in space and over time thereby affecting  
277 soil properties, especially SOC content, which is particularly affected by agriculture practices and  
278 landscape features like distance from hedgerows (Lacoste et al., 2014).

279 We estimated uncertainties at two different scales: (i) at the point scale by Monte-Carlo  
280 simulations considering the 64 sampling points and (ii) at the landscape scale by fuzzy clustering  
281 of the residuals from the RandomForest model. Comparing the CI of SSOC change at the 64  
282 points (Fig. 4) and the maps of the mean SSOC change (Figs. 6a and 6c) highlights an increase  
283 in uncertainty due to upscaling from the punctual scale to the landscape scale: the CI of the 64  
284 points was  $0.4 - 0.8 \text{ t C ha}^{-1} \text{ year}^{-1}$ , while that of spatial modelling was  $0.9 - 1.5 \text{ t C ha}^{-1} \text{ year}^{-1}$ .  
285 This issue was largely discussed by Dignac et al. (2017), who confirmed that upscaling from the  
286 fine spatial scale to landscape and global scale is a crucial issue that generates high uncertainties.  
287 Thus, our results demonstrate high carbon sequestration rates associated with high uncertainties  
288 amplified by spatial modelling. This challenge is mainly due to the complexity of carbon  
289 sequestration processes, which can be achieved through changes in land use and agricultural  
290 practices (Minasny et al., 2017). From a practical point of view, the SSOC greatly depends on the  
291 land use patterns and agriculture practices implemented such as the addition of exogenous  
292 organic matter, tillage, and fertilisation (Dignac et al., 2017). Actually, our results show that

293 SSOC change over time are highly variable for a same crop rotation and that land cover explains  
294 by itself only a relatively small fraction of the time evolution. This may be explained by the  
295 influence of other factors impacting the SOC dynamics namely local topographical and  
296 pedological conditions, but also by the practices adopted by the farmer to grow a given crop:  
297 tillage, crop residues management, exogenous organic matter inputs, grazing density, yield  
298 levels have a major effect on the quantity and quality of fresh organic matter inputs into the soil  
299 and considerably affect their dynamics. Inputs of organic matter to the soil, which mainly depend  
300 on the number of grazing animals or mowing frequency in grasslands, as well as on crop residues  
301 and exogenous organic matter in agriculture systems, promote SOC storage (Dignac et al., 2017),  
302 but may be variable according to the farming systems and the field. These practices markedly  
303 affect soil properties including soil physical structure (Dignac et al., 2017), soil microbial  
304 biomass and microbial activity (Tu et al., 2006) and consequently the distribution of SSOC in  
305 space and over time (Dignac et al., 2017).

306 Despite the importance of these factors, they could not be directly included in SSOC spatial  
307 modelling due to lack of appropriate spatialised data. In addition, our findings emphasise that  
308 using land use as a proxy of agriculture practices and crop rotations leads to very uncertain  
309 estimates of SSOC change. Thus, although direct measurement of SSOC change is costly and  
310 hampered by measurement uncertainties, they appear still essential to ensure unbiased assessment  
311 of soil carbon storage at landscape scale.

## 312 5 Conclusions

313 The SSOC change was assessed and mapped at the landscape scale by coupling direct soil  
314 monitoring and machine-learning regression modelling. Our study is innovative because in the  
315 soil science literature, to our best knowledge, no other study tried to apply machine-learning on



316 datasets of SSOC change measurements to map SSOC change in space and time. Predictive  
317 models were used to elucidate topographical, geological, pedological factors and human activities  
318 that caused patterns of spatial distribution. The main results showed that the spatial heterogeneity  
319 of SSOC change was driven by crop rotations, topographic attributes, and local soil conditions  
320 (waterlogging, soil parent material). These factors could be captured by the regression model and  
321 spatially represented. Other important factors such as agricultural practices may also influence  
322 the SOC dynamics and could not be informed at each point and therefore were not considered by  
323 the model. This study reveals complex interactions in storing SOC, which may have implications  
324 for future planning and monitoring of agro-ecosystems and can be integrated into land-use  
325 decision-making.

326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343

344  
345  
346  
347  
348  
349  
350  
351

352 **Figure captions**

Figure 1. Location of the study area in France and of the soil sampling points within the study area and for different crop rotations.

353 Figure 2. Diagram of the 4-step method used to map SSOC change across the landscape scale at  
354 10m of resolution (RF: Random Forest models, SMLRK: stepwise multiple linear regression  
355 kriging model)

356 Figure 3. a) Mean and 90% confidence intervals of soil organic carbon (SOC) stock in 2009 and  
357 2016 for different soil mineral masses. b) 2016 SSOC versus 2009 SSOC at the 64 sampling  
358 points for a mineral mass of 300 kg m<sup>-2</sup>. Blue line:  $y= 1.55+0.79x$  regression line ( $r^2=0.69$ );  
359 dashed red lines: 95 % predictive interval; grey zone: 95% confidence interval.

360 Figure 4. Mean and 90% confidence intervals of SSOC change calculated for the 64 prospected  
361 sites using Monte Carlo simulations.

Figure 5. Importance of predictor variables in the RF modelling of SSOC change between 2009 and 2016.

362 Figure 6. Maps of SSOC change between 2009 and 2016 within the 10 km<sup>2</sup> study area: a) lower  
363 bound of the 90% confidence interval (CI) of SSOC change, b) mean SSOC change, c) upper  
364 bound of the 90% CI of SSOC change.

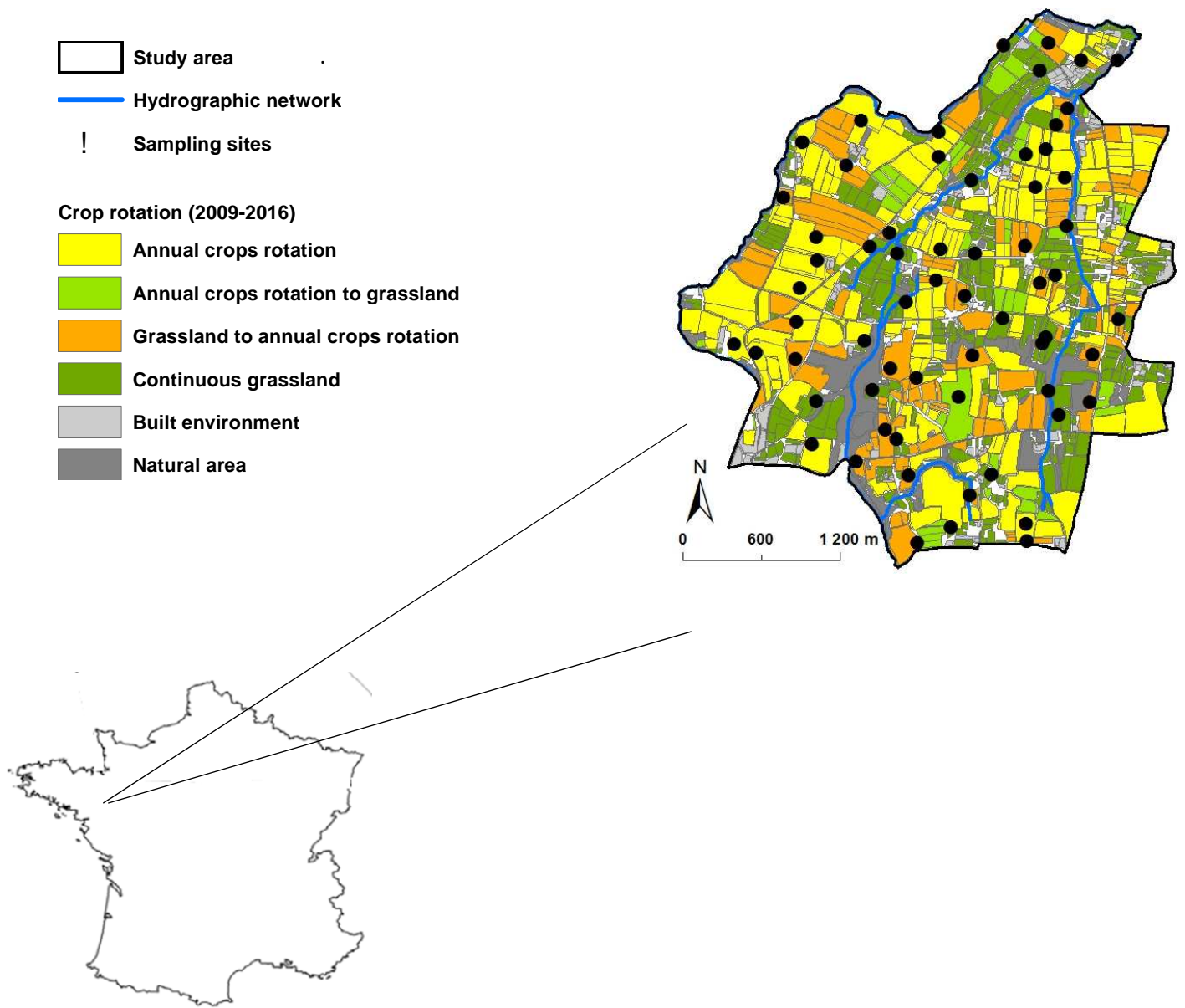
365 **Table headings**

Table 1. Description of the environmental covariates selected

Table 2. Probability distribution for input variables used in the Monte-Carlo procedure to estimate SSOC uncertainty at the 64 sampling points. Standard deviations for derived from replicates measurements for SOC content (n=30) and bulk density (n=69).

Table 3. Median values of SSOC change for different crop rotations between 2009 and 2016

Table 4. Summary statistics of stepwise MLR regression model



**Figure 1. Location of the study area in France and of the soil sampling points within the study area and for different crop rotations.**

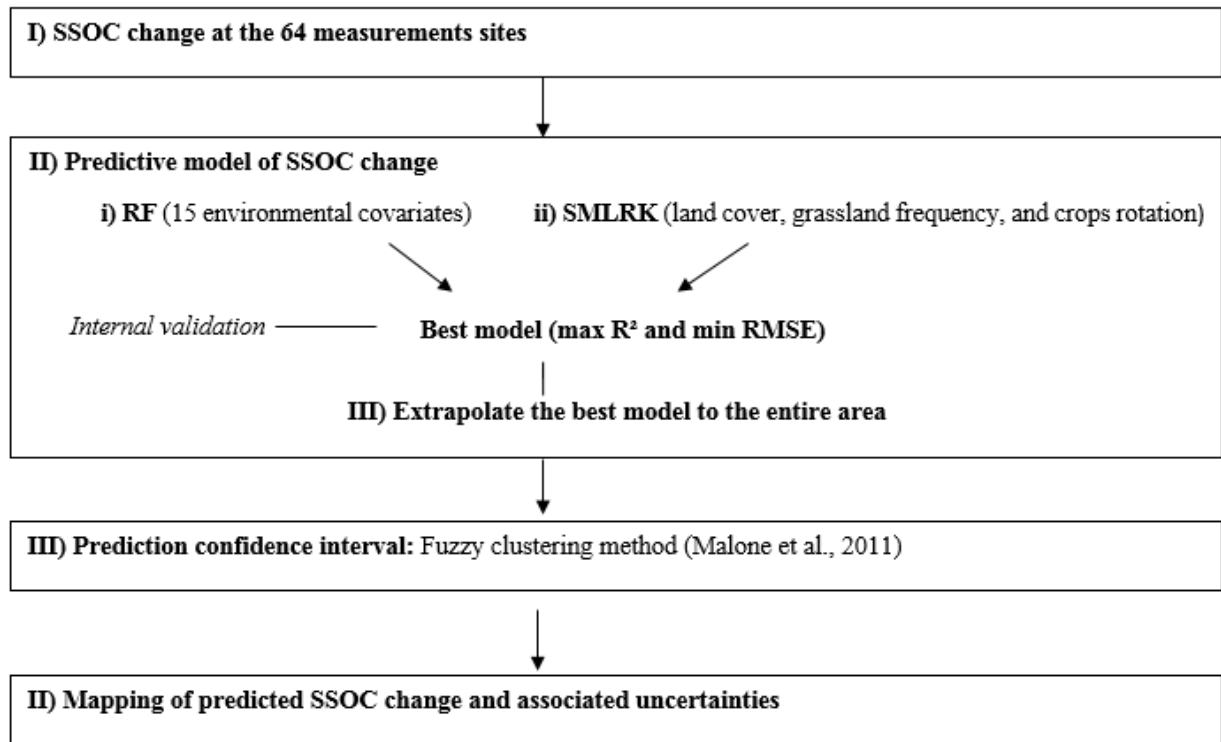


Figure 2. Diagram of the 4-step method used to map SSOC change across the landscape scale at 10m of resolution (RF: Random Forest models, SMLRK: stepwise multiple linear regression kriging model)

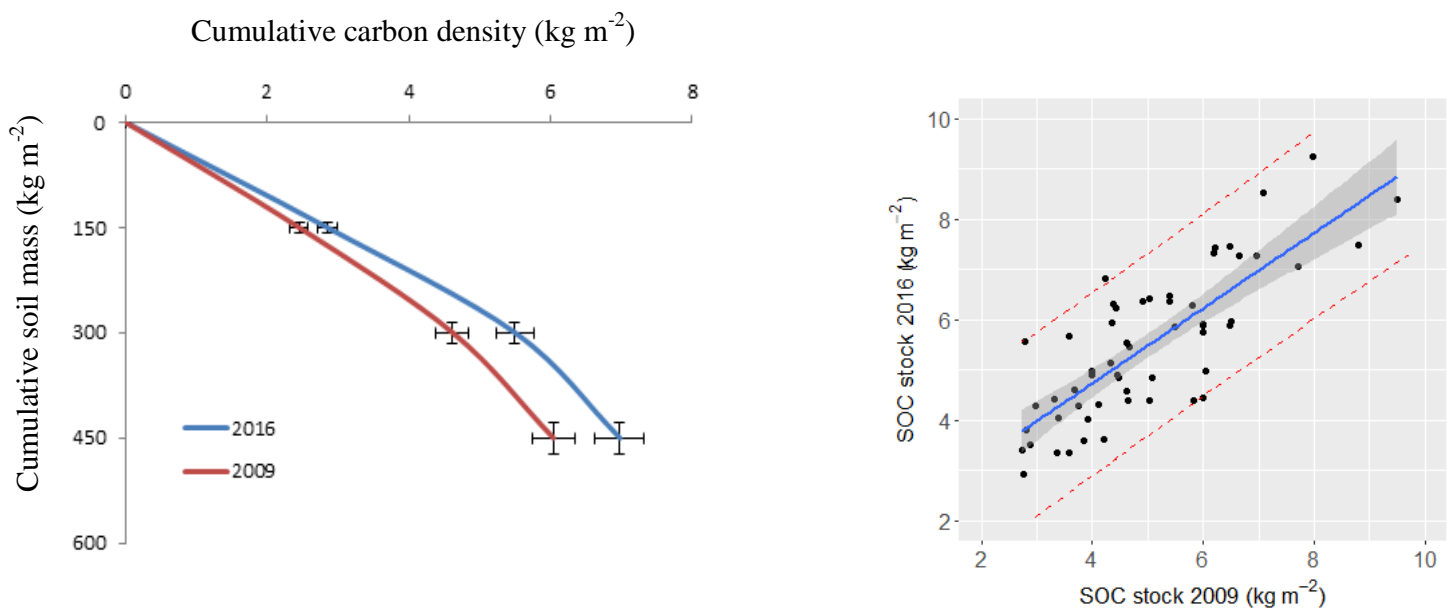
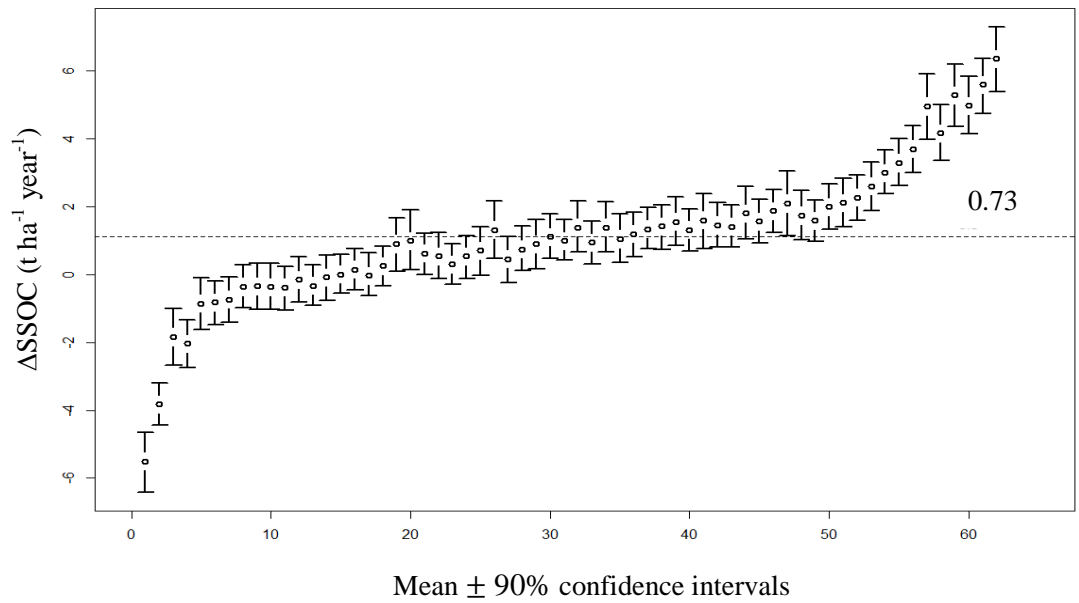
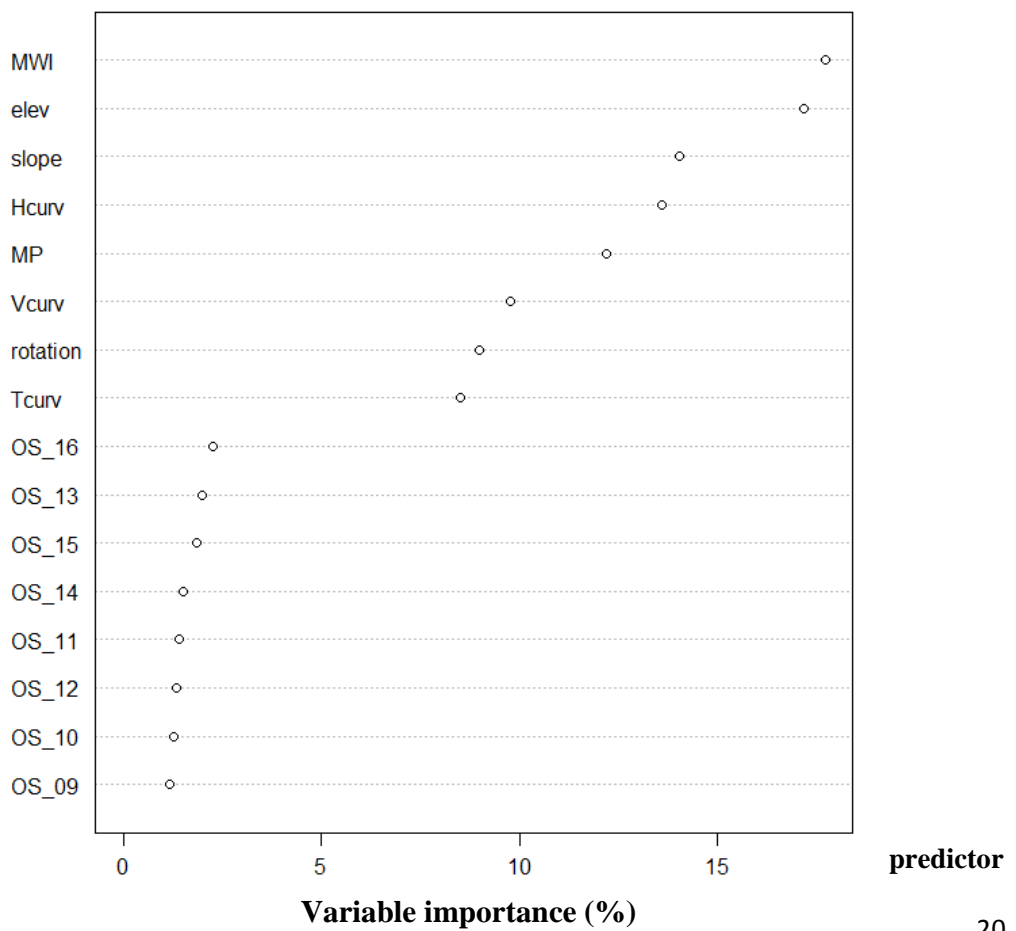


Figure 3. a) Mean and 90% confidence intervals of soil organic carbon (SOC) stock in 2009 and 2016 for different soil mineral masses. b) 2016 SSOC versus 2009 SSOC at the 64 sampling points for a mineral mass of 300 kg m<sup>-2</sup>. Blue line:  $y = 1.55 + 0.79x$  regression line ( $r^2 = 0.69$ ); dashed red lines: 95 % predictive interval; grey zone: 95% confidence



**Figure 4. Mean and 90% confidence intervals of SSOC change calculated for the 64 prospected sites using Monte Carlo simulations.**



**Figure 5. Importance of**

**variables in the RF modelling of SSOC change between 2009 and 2016.**

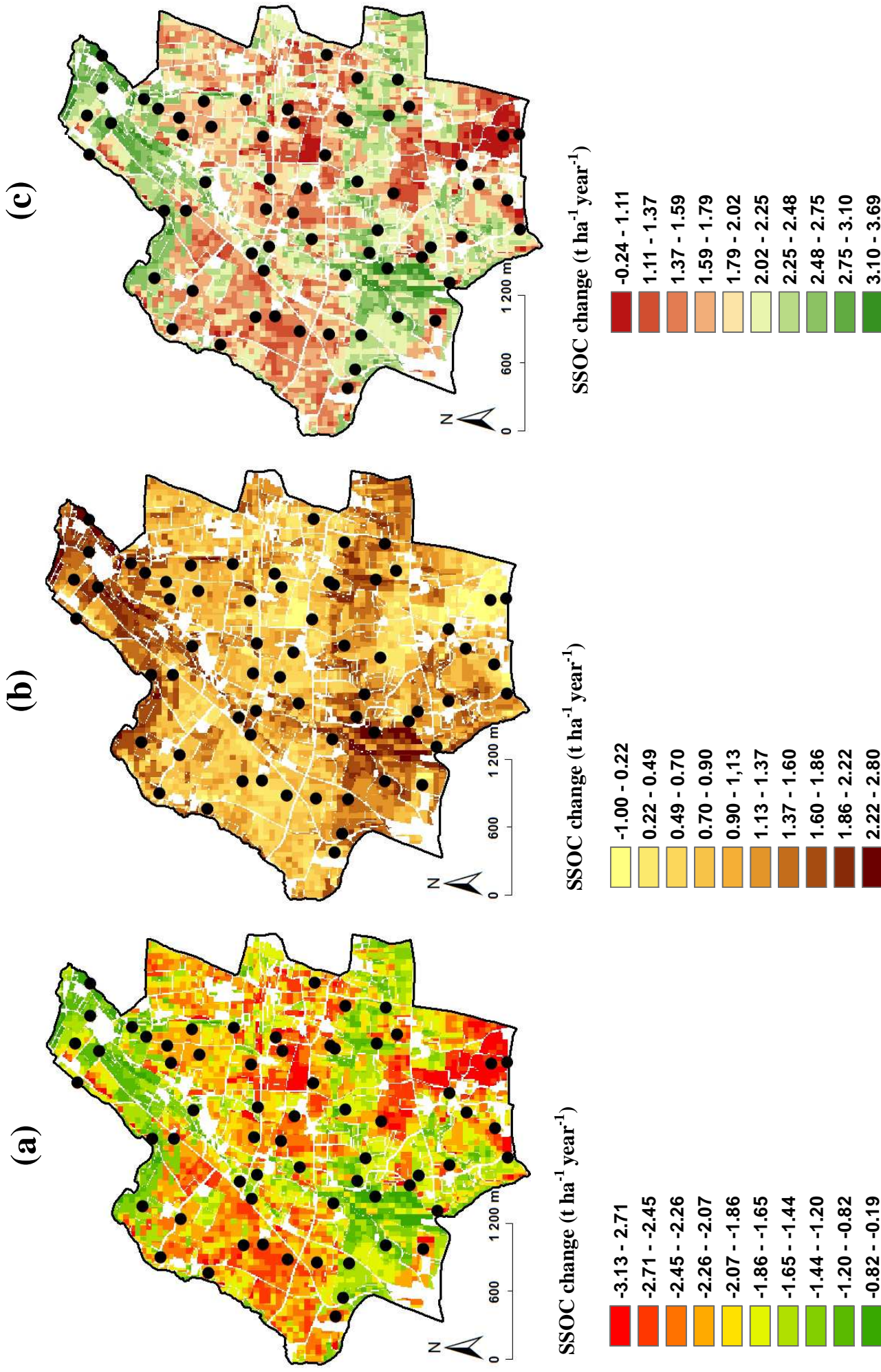


Figure 6. Maps of SSOC change between 2009 and 2016 within the 10 km<sup>2</sup> study area: a) lower bound of the 90% confidence interval (CI) of SSOC change, b) upper bound of the 90% CI of SSOC change, c) mean SSOC change.



**Table 1. Description of the environmental covariates selected**

Summary of environmental covariates. P: parent material; R: relief; O:organisms; C: categorical; Q: quantitative.

Label	Environmental covariate	SCORPAN factor	Type	Unit or number
<b>Terrain attributes derived from the digital elevation model</b>				
elev	Elevation	R	Q	m
Slope	Local hillslope gradient	R	Q	%
Vcurv	Profile curvature	R	Q	m.100m <sup>-1</sup>
Hcurv	Tangential curvature	R	Q	6 classes
Tcurv	Total curvature	R	Q	6 classes
MWi	Modified topographic wetness index	R	Q	5 classes
<b>Geology</b>				
MP	Soil parent material	P	C	6 classes
<b>Organism</b>				
OS-09	Land cover in 2009	O	C	5 classes
OS-10	Land cover in 2010	O	C	5 classes
OS-11	Land cover in 2011	O	C	5 classes
OS-12	Land cover in 2012	O	C	5 classes
OS-13	Land cover in 2013	O	C	5 classes
OS-14	Land cover in 2014	O	C	5 classes
OS-15	Land cover in 2015	O	C	5 classes
OS-16	Land cover in 2016	O	C	5 classes
Rotation	Crop rotation	O	C	5 classes

**Table 2. Probability distribution for input variables used in the Monte-Carlo procedure to estimate SSOC uncertainty at the 64 sampling points. Standard deviations for derived from replicates measurements for SOC content (n=30) and bulk density (n=69).**

Input variable	Distribution	Parameters
Thickness of soil layer: Z (0-7.5 cm)	Normal	Mean=Z Standard Dev.=0.5 cm
Thickness of soil layer: Z (7.5-15 cm)	Normal	Mean=Z Standard Dev.=1 cm
Soil organic carbon content: SOC (g kg <sup>-1</sup> )	Normal	Mean=SOC Standard Dev.=2 g kg <sup>-1</sup>
Bulk density: BD (g cm <sup>-3</sup> )	Normal	Mean=BD. Standard Dev.=0.03 g cm <sup>-3</sup>

**Table 3. Median values of SSOC change for different crop rotations between 2009 and 2016**

Crop rotations	( $\Delta$ SSOC) at 300 kg m <sup>-2</sup> soil mineral (t C ha <sup>-1</sup> year <sup>-1</sup> )	Number of sites
Continuous grassland	1.40 ± 0.31	14
Annual crops rotation (without grassland)	0.82 ± 0.26	26
Annual crops rotation to Grassland	0.94 ± 0.29	15
Grassland to annual crops rotation	0.42 ± 0.18	6

**Table 4. Summary statistics of stepwise MLR regression model**

Coefficients				
	Estimate	Standar error	t-value	Significance
Constant	0.9	0.31	2.86	0.005**
Annual crops rotation to Grassland	-0.1	0.72	-0.14	0.88
Continuous grassland	1.17	0.53	2.22	0.03*
Grassland to annual crops rotation	-1.08	0.51	-2.08	0.04*
Multiple R-squared: 0.20		Adjustes R-squared: 0.16		
F-statistics: 4.82		p-value: 0.004		
Signif. codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

## References

- Arrouays, D., Balesdent, J., Germon, J.C., Jayet, P.A., Soussana, J.F., Stengel, P., 2002. Contribution à la lutte contre l'effet de serre : Stocker du carbone dans les sols agricoles de France ? Expertise Scientifique Collective. INRA, Paris (332 pp.).
- Bellamy, P.H., Loveland, P.J., Bradley, R. I., Lark, R.M., Kirk, G.J.D., 2005. Carbon losses from all soils across England and Wales 1978-2003, *Nature* 437, 245-248.
- Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin* 24, 43–69.
- Breiman, L., 2001. Random forests: machine learning. Statistics department university of California, Berkeley, CV 94720. 5-32.
- Bonijoly, D., Perrin, J., Truffert, C., Asfirne, F., 1999. Couverture géophysique aéroportée du Massif armoricain. Report R 40471. BRGM, Paris.
- Chaplot, V., Bouahom, B., Valentin, C., 2010. Soil organic carbon stocks in Laos: spatial variations and controlling factors. *Global Change Biology* 16, 1380–1393.
- Chartin, C., Stevens, A., Goidts, E., Krüger, I., Carnol, M., van Wesemael, B., 2017. Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southern Belgium, Wallonia). *Geoderma Regional* 9, 73–86.
- 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.
- Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G.T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C., Basile-Doelsch, I., 2017. Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review. *Agronomy for Sustainable Development* 37,14.
- Eglin, T., Walter, C., Nys, C., Follain, S., Forgeard, F., Legout, A., Squividant, H. Influence of waterlogging on carbon stock variability at hillslope scale in a beech forest (Fougères forest–West France). *Annals of forest science*, 65(2), 1-10, 2008.
- ESRI, 2012. ArcMap 10.3. Environmental Systems Resource Institute, Redlands, California.
- Gac, Armelle., Dollé, J.B., Le Gall, A., Klumpp, K., Tallec, T., Mousset, J., Eglin, T., Bispo, A., Peyraud, J.L., Faverdin, P., 2010. Le stockage de carbone par les prairies une voie d'atténuation de l'impact de l'élevage herbivore sur l'effet de serre. Institut de l'élevage. pp12
- Gomes, L.C., Faria, R.M., de Souza, E., Veloso, G.V., Schaefer, C.E.G.R., Filho, E.I.F., 2019. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma* 340, 337–350.
- Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152, 195–207.
- Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* 6, e5518.
- Hocking, R. R., 1976. The Analysis and Selection of Variables in Linear Regression. *Biometrics*, 32.
- IUSS Working Group WRB, 2007. World Reference Base for Soil Resources 2006, first update 2007. World Soil Resources Reports No. 103. FAO, Rome. 116 pp.
- IPCC, 2006. Guidelines for National Greenhouse Gas Inventories. <http://www.ipcc-nggip.iges.or.jp/>

- Lacoste, M., Lemerrier, B., Walter, C. 2011. Regional mapping of soil parent material by machine learning based on punctual training data. *Geomorphology* 133 (1-2), 90-99.
- Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. *Geoderma* 213, 296–311.
- Lacoste, M., Viaud, V., Michot, D., Walter, C., 2016. Model-based evaluation of impact of soil redistribution on soil organic carbon stocks in a temperate hedgerow landscape: Soil Redistribution Impact on Soil Organic Carbon Stocks. *Earth Surface Processes and Landforms* 41, 1536–1549.
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R news* 2, 18–22.
- Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138–152.
- Malone, B.P., McBratney, A.B., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma* 160, 614–626.
- Lark, R.M., Cullis, B.R., 2004. Model-based analysis using REML for inference from systematically sampled data on soil. *European Journal of Soil Science* 55, 799–813.
- Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boullonne, L., Arrouays, D., 2011. Spatial distribution of soil organic carbon stocks in France . *Biogeosciences* 8, 1053–1065
- McBratney, A.B., Santos, M.L.M. and Minasny, B., 2003. On digital soil mapping. *Geoderma*, 117, 3-52.
- Merot, P., Squitiant, H., Auroousseau, P., Hefting, M., Burt, T., Maitre, V., Kruk, M., Butturini, A., Thenail, C., Viaud, V., 2003. Testing a climato-topographic index for predicting wetlands distribution along an European climate gradient. *Ecological Modelling* 163, 51–71.
- Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences* 32, 1378–1388.
- Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Digital Mapping of Soil Carbon, in: *Advances in Agronomy*. Elsevier, pp. 1–47.
- Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L., O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.-C., Vågen, T.-G., van Wesemael, B., Winowiecki, L., 2017. Soil carbon 4 per mille. *Geoderma* 292, 59–86.
- Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* 67, 215-226.
- Palisade, 2005. @Risk Software Ver.4. Palisade corporation.
- Papritz, A., Webster, R., 1995. Estimating temporal change in soil monitoring: I. Statistical theory. *European Journal of Soil Science* 46, 1–12.
- Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G.P., Smith, P., 2016. Climate-smart soils. *Nature* 532, 49–57.
- R Core Team, 2017. R: a language and environment for statistical computing. R. Foundation for

- Statistical Computing, Vienna, Austria 3-900051-07-0. <http://www.R-project.org>.
- Squidant H., 1994. MNTSURF: Logiciel de traitement des modèles numériques de terrain. ENSAR, Rennes, France, 36.
- Tu, C., Ristaino, J.B., Hu, S., 2006. Soil microbial biomass and activity in organic tomato farming systems: Effects of organic inputs and straw mulching. *Soil Biology and Biochemistry* 38, 247–255.
- Uusitalo, L., Lehtikoinen, A., Helle, I., Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software* 63, 24–31.
- Viaud, V., Angers, D.A., Walter, C., 2010. Toward Landscape-Scale Modeling of Soil Organic Matter Dynamics in Agroecosystems. *Soil Science Society of America Journal* 74, 1847-1860.
- Viaud, V., Santillán-Carvantes, P., Akkal-Corfini, N., Le Guillou, C., Prévost-Bouré, N.C., Ranjard, L., Menasseri-Aubry, S., 2018. Landscape-scale analysis of cropping system effects on soil quality in a context of crop-livestock farming. *Agriculture, Ecosystems & Environment* 265, 166–177.
- Walter, C., Viscarra Rossel, R.A. et McBratney, A.B., 2003. Spatio-temporal simulation of the field-scale evolution of organic carbon over the landscape. *Soil Science Society of America Journal* 67, 1477-1486.
- Walter, C., Bispo, A., Chenu, C., Langlais-Hesse, A., Schwartz, C., 2015. Les services écosystémiques des sols: du concept à sa valorisation. *Agriculture et Foncier- Concurrences entre Usages des sols et entre Usagers des sols Agricoles: La Question Foncière Renouvelée*, ed C. Demeter (Paris: Cahier Demeter), 51–68.
- Wendt, J.W., Hauser, S., 2013. An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers. *European Journal of Soil Science* 64, 58–65.
- Wheeler, I., 2014. Monitoring total soil organic carbon at farm scale. PhD Thesis. The university of Sydney.