LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG
Ministère de l'Agriculture, de la Viticulture
et du Développement rural

UCLouvain

RECENT EVOLUTION OF SOIL ORGANIC CARBON IN THE GRAND-DUCHY OF LUXEMBOURG

November 2020

C. Chartin, B. van Wesemael
Georges Lemaître Centre for Earth and Climate Research-Earth and Life Institute
Université Catholique de Louvain
Place Louis Pasteur, 3 - 1348 Louvain-la-Neuve, Belgium
Email : caroline.chartin@uclouvain.be

S. Marx, M. Steffen, L. Leydet
Ministère de l'Agriculture, de la Viticulture et du Développement rural
Administration des Services Techniques de l'Agriculture – ASTA
Service de pédologie
72, avenue L. Salentiny - 9080 Ettelbruck, Grand Duchy of Luxembourg
Email: simone.marx@asta.etat.lu

TABLE OF CONTENTS

1 Introduction 2
2 The Grand-Duchy of Luxembourg 3
3 Current state of SOC in GDL and its short-term evolution 8
3.1 Context 8
3.2 Methodology 8
3.2.1 Recent SOC data in GDL (2012-2019) 8
3.2.2 Spatial covariates 10
3.2.3 Spatial SOC modelling 14
3.2.4 Comparing SOC at two periods 16
3.3 Results 16
3.3.1 Implications of the SOC data filtering 16
3.3.2 SOC by period - descriptive statistics and differences test 17
3.3.3 SOC relations to spatial covariates 25
3.3.4 Model results and performance 32
3.3.5 SOC maps - description and comparison 34
4 Implication of Good Agricultural Practices (GAP) on short-term SOC dynamics 40
4.1 Context 40
4.2 Methodology 40
4.2.1 Characteristics of the Good Agricultural Practices 40
4.2.2 Database preparation 41
4.2.3 Analysis of SOC differences between management practices 42
4.2.4 Conditional inference trees 42
4.3 Results 42
4.3.1 Implications of the SOC data filtering and merging 42
4.3.2 Impact of GAP on SOC 43
4.3.3 Relative importance of management practices vs environmental covariates on SOC variability51
5 Conclusions 54
6 References 56
7 Annexes 60
7.1 Methodology: map of the Minimum Depth of soil Hydromorphy 60
7.2 SOC summary statistics: reference tables for texture classes L, M, OM and S 62
7.3 Additional SOC maps for cropland and grassland 63

1 INTRODUCTION

In 2014, the UCLouvain produced the first versions of soil organic carbon (SOC) content and soil organic carbon stock maps for the Grand-Duchy of Luxembourg (GDL; Stevens et al., 2014a; Stevens et al., 2014b). The models fitted covered croplands, grasslands, vineyards and forests, based on analytical data from agriculture and viticulture (2012-2014) and from the National Forest Inventory (2008-2013).

The SOC content and stock maps are essential tools in the present discussion about climate change and the potential storage of carbon in cultivated and forested soils. Hence, these maps are used as central documents for reporting activities such as the National Inventory Report (NIR) required by the UNFCCC as part of the Kyoto protocol.

Soil organic matter, a key indicator of soil quality regarding its many positive benefits in terms of agronomy and environment, consists of 57% C, and plays an important role:

- for improving soil stability and reducing erosion (Chenu et al., 2000 ; Bronick et al., 2005) ;
- for providing plant nutrients (Clivot et al., 2017 ; Oldfield et al., 2019) enabling to reduce the use of nitrogenous fertilizers;
- for the degradation and adsorption of phytopharmaceuticals used for crop protection and pest control (Fenoll et al., 2011; 2014);
- for carbon storage (Buysse et al., 2013 ; D'Hose et al., 2014 ; Vanden Nest et al., 2014 ; Wiesmeier et al., 2019) in the fight against global warming.

Since 2014, the soil analytical database in GDL was enlarged with thousands of samples from croplands, grasslands and vineyards.

Objectives

This project aimed at exploring all the SOC data available in GDL for improving our knowledge on the recent evolution of soil organic matter, on the present SOC regional baseline and on the potential effects of management practices on the future SOC trends.

N.B. *this project (2019-2021) contains three work packages (WP):*

WP1: Historical trends of SOC (historical data are also available for GDL)

WP2: Updating current SOC content maps

WP3: Indicators assessing effects of management practices on SOC

The ASTA (Administration des Services Techniques de l'Agriculture) and UCLouvain agreed early in the project in prioritizing the WPs in this order: WP2, WP3 and WP1. The WP2 was first modified in order to optimize the research around the recent short-term trends in SOC contents, then the research was focused on the impact of recent management practices (WP2). Considering the new directions of the project, WP1 has been finally set aside. To finish, given the necessity for complementary SOC measurements during the research process and the COVID-19 crisis, this report was finalized later than foreseen (originally expected for early 2020).

2 THE GRAND-DUCHY OF LUXEMBOURG

- *Location and climate*

The Grand-Duchy of Luxembourg (GDL) covers ~2600 km² in northwestern Europe sharing borders with France, Germany and Belgium. The climate is temperate semi-oceanic with mean annual temperature ranging from 7.3 °C to 9.9 °C and annual precipitation of ca. 830mm (Fig. 1).

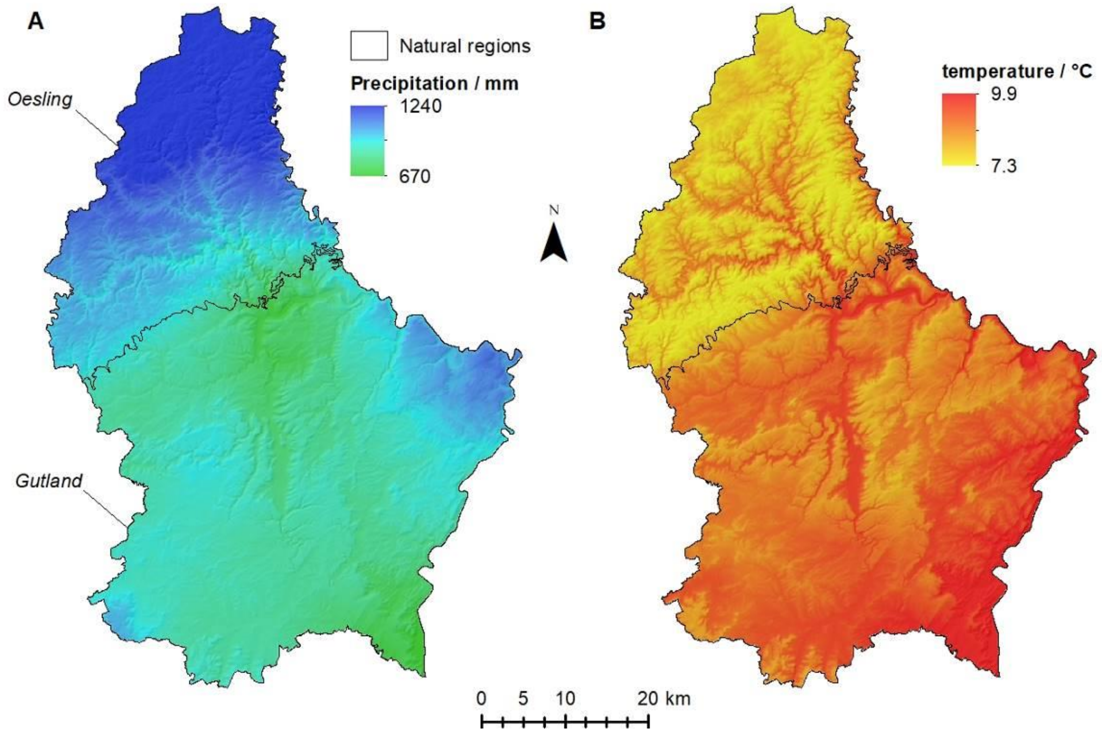

Figure 1: Maps of A/ mean annual precipitation (mm) and B/ mean annual temperature (°C) in the Grand-Duchy of Luxembourg over the period 1971-2000 (source: Stevens et al.; 2014)

- *Geology and relief*

The country consists of two main natural regions, the Oesling in the north (~830 km²) and the Gutland in the center and south (~1770km²). The Oesling, like the Ardennes in Belgium and Eifel region in Germany, is a massif of the Primary Period made of Lower Devonian slate, quartzite and sandstone. The Oesling is now a sub-horizontal peneplain with deeply incised valleys and a mean altitude of ca. 450m (Fig. 2). The Gutland is a more heterogeneous region characterized by a south-west-facing cuesta topography (mean altitude of ca. 245m; Fig. 2) which developed on monoclinal Triassic and Jurassic sediments. Rocks formed during the same period can be found in the Gaume region in Belgium, north of the Lorraine in France and Bitburger Gutland in Germany. Triassic deposits are made of marls, sandstone and dolomites, all containing mineral dolomite while Jurassic sediments are made of sandstone and marls with calcium carbonates.

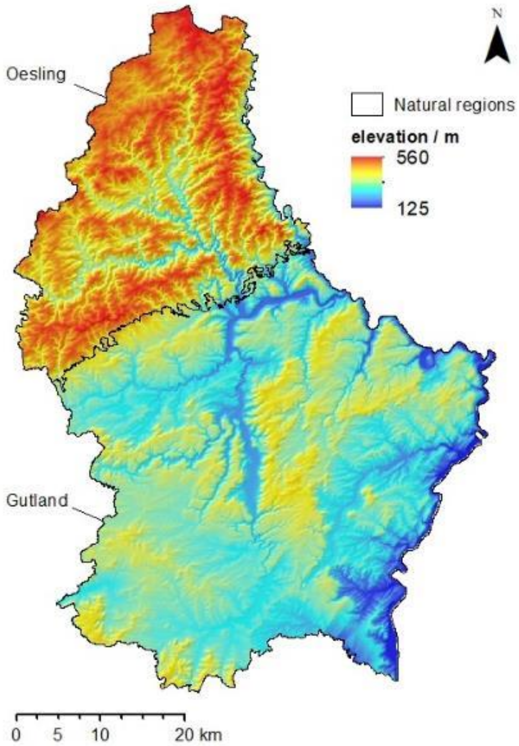

Figure 2: Natural regions and elevation (m) in Grand-Duchy of Luxembourg.

- *Soil types*

The Oesling is predominantly covered by shallow stony silt loam soils (cambic Umbrisol in the WRB classification) while soils of the Gutland are mainly Luvisols (IUSS Working Group WRB, 2006). There are 26 soil associations (SMU = soil mapping units) in GDL, according to the major geological units, that we further regrouped into 10 classes (Table 1; Figure 3A) representing variations based mainly on mineralogy (and texture).

Table 1: Soil associations and corresponding WRB classification (IUSS Working Group WRB, 2014)

Ref.	Soil association	Natural region	Geologic	WRB classification	Relative area (%)
1	Oesling	Oesling	Lower Devonian	skeletic dystric Cambisol (siltic)	29.1
2	Argiles lourdes des schistes bitumineux	Gutland	Triassic	vertic calcaric Cambisol (clayic)	4.6
3	Argiles lourdes du Keuper	Gutland	Triassic	vertic dolomitic Cambisol (clayic)	5.1
4	Argiles du Lias Inf. et moyen	Gutland	Jurassic	gleyic/stagnic endocalcaric Luvisol (loamic)	1.3
5	Dépôts limoneus sur grès	Gutland	Jurassic	haplic Luvisols (loamic)	10.6
6	Grès du Luxembourg	Gutland	Jurassic	haplic Luvisols (arenic)	9.6
7	Calcaire du Bajocien	Gutland	Jurassic	leptic calcaric Cambisol (loamic)	12.5
8	Dolomies du Muschelkalk	Gutland	Triassic	leptic dolomitic Cambisol (loamic)	10.1
9	Bundsandstein	Gutland	Triassic	endolomitic Luvisol (loamic)	3.2
10	Autres	Gutland and Oesling	Alluvium, colluvium	Fluvisol, Cambisol, Regosol	13.9

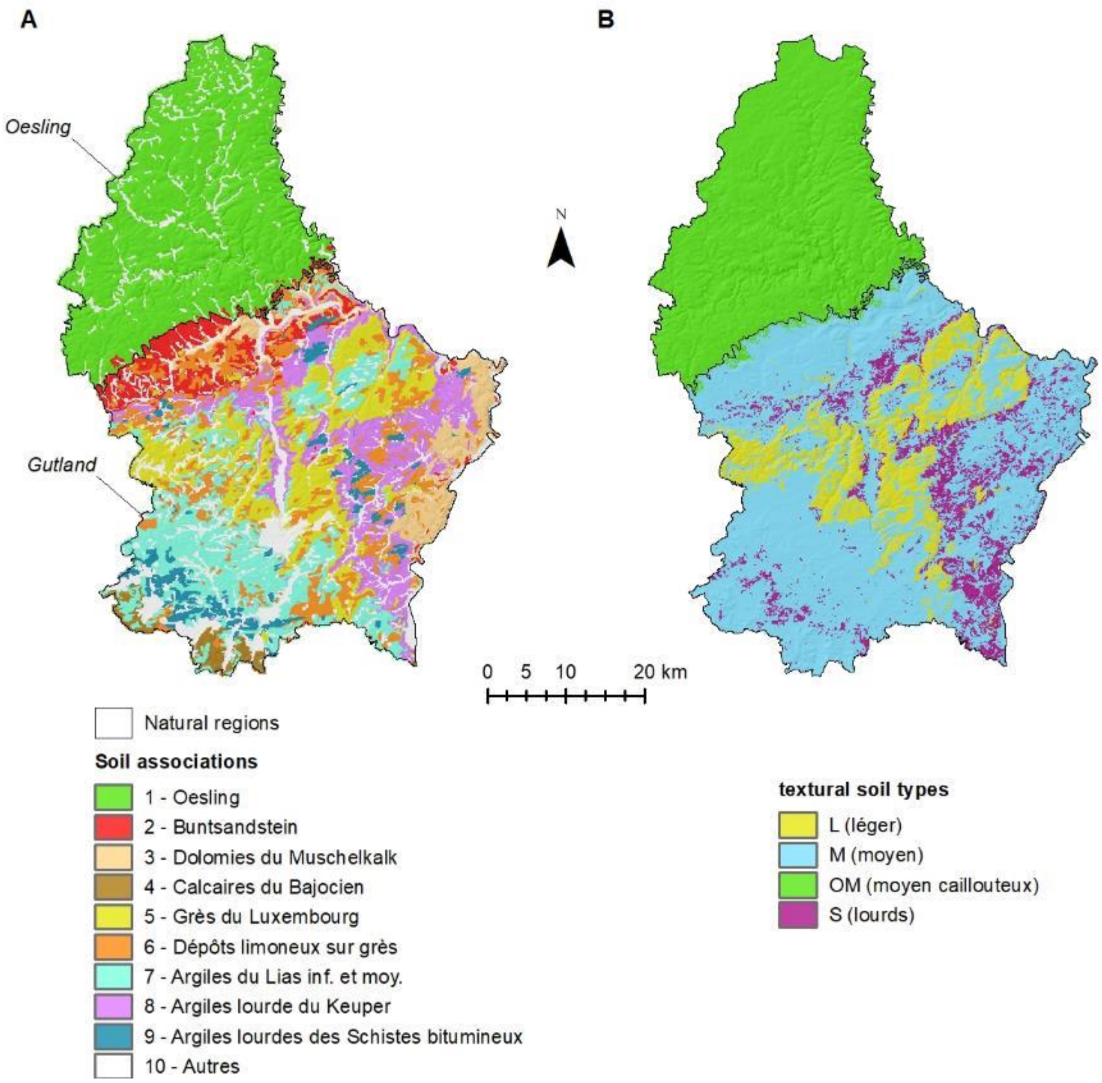

Figure 3: Maps of soils of GDL classified by A/ regrouped soil associations (mineralogy / geology) and B/ textural properties.

In addition, based on the texture identified by finger testing on soil samples entering the ASTA - soil laboratory, GDL has defined four main textural soil types (Fig. 3B). Each of them cover different classes of the texture triangle designed for soils of Belgium and GDL (Tab. 2 and Fig. 4). The shallow stony silt loam[1] soils (OM) cover the northern natural region - Oesling. The southern natural region – Gutland - is mainly covered by clay loam - loam - silt loam soils (M), loamy sand (L) and clay (S) soils.

[1] *According to FAO texture classes*

Table 2: textural soil types and corresponding groups from the texture triangle designed for Belgium and GDL.

Type de sol par test tactile (sigle d'abréviation)	Région naturelle	Classe texturale (triangle textural LU)	Surface relative (%)
Sol léger (L)	Gutland	Z (sable), S (sable limoneux)	11
Sol moyen (M)	Gutland	L (limon sableux), P (limon sableux léger), A (limon), E (argile)	48
Sol lourd (S)	Gutland	U (argile lourde)	8
Sol moyen caillouteux (OM)	Oesling	G… (Sols argilo-limono-caillouteux)	33

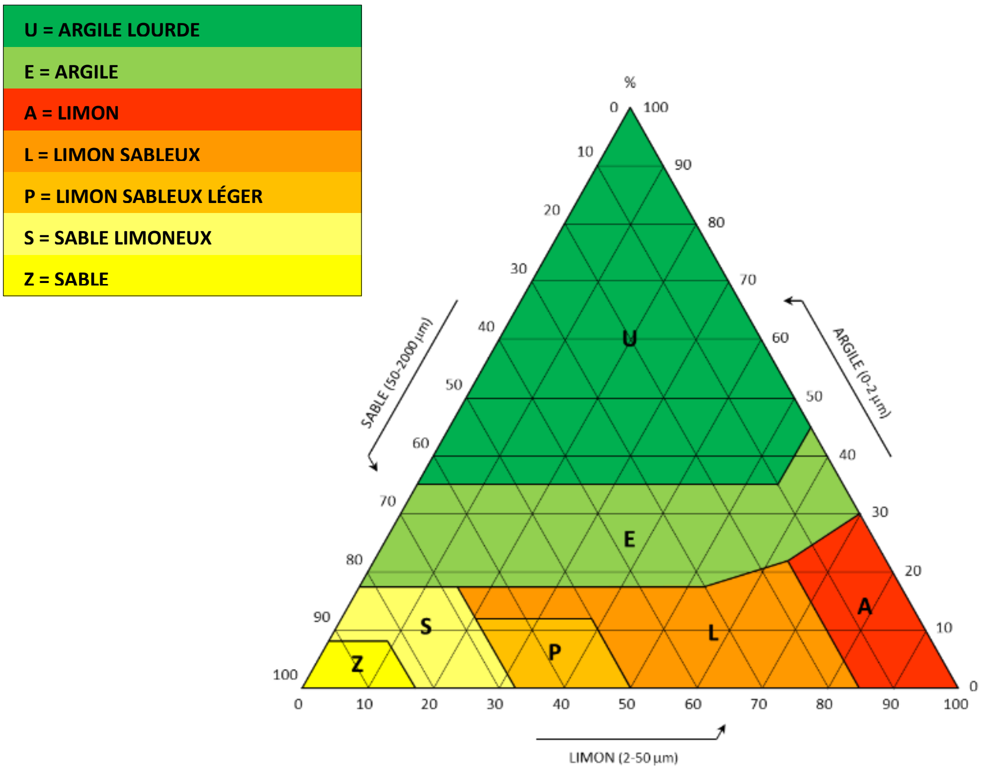

Figure 4: Texture triangle designed for soils of Belgium and Grand-Duchy of Luxembourg.

- *Land use*

According to the simplified LANDUSE2018 layer (https://data.public.lu/, Fig. 5), forest occupied about 35% of the GDL territory in 2018, followed by grassland with ~26% and cropland with ~22%, while vineyard covered only ~0.5% of the territory. The remaining percentage of ~16.5% represent mainly built-up areas, wetland, water and other agricultural land. The last three classes only cover a small area.

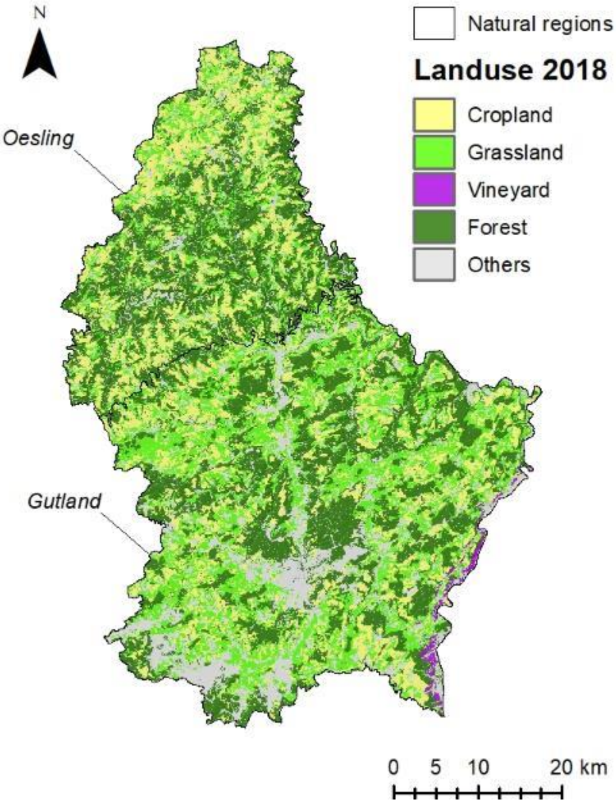

Figure 5: Map of landuses in GDL (after reclassification of 'Sub-type 1 LU classes'; https://download.data.public.lu/resources/landcover-landuse-2018/20200504-135337/lisl-landuse-2018-documentation.pdf).

- *Agriculture and management practices*

The most common crop rotation in the Oesling area is a 6 to 8 year rotation with cultivation of silage maize and cereals for 3 or 4 years, followed by temporary grassland for another 3 or 4 years. The most common crop rotation in the Gutland is a three years rotation with winter wheat, winter barley and silage maize.

Amongst the variety of management practices used in modern agriculture, some are recognized as not sustainable enough for soils and surrounding environment, e.g. intensive tillage, heavy spreading of pesticides, bare soils exposed to rainfall, etc. These last years, national and European policies (as Common Agricultural Policy – CAP) have been created to encourage farmers to apply environmentally-friendly farming techniques (we would call them 'Good Agricultural Practices' or 'GAP' here) that go beyond legal obligations, as the Agri-Environment Measures[2] and the Greening initiative[3] from EU. So far, the most common GAP applied and promoted in cropland over GDL are cover crops, reduced or no-till strategies, and temporary grassland.

[2] *https://ec.europa.eu/info/food-farming-fisheries/sustainability-and-natural-resources/agriculture-and-environment/cap-and-environment/agri-environment-measures?2nd-language=fr*

[3] *https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/income-support/greening_en*

3 CURRENT STATE OF SOC IN GDL AND ITS SHORT-TERM EVOLUTION

3.1 CONTEXT

In the context of the second pillar of the Common Agricultural Policy, the European commission approved the RDP (Rural Development Program) of GDL for the period 2014-2020. Based on SOC data from 2012 to 2019, the aim of this study is to assess how the SOC content has evolved in general during this period, and if, and how, the RDP could have affected the soil quality. To this aim, SOC content baselines were computed for the period right before the beginning of the PDR (T1: 2012-2015; considering one year delay for a complete implementation of the RDP all over the GDL) and for the period right after (T2: 2016-2019).

Various data (in raster or vector formats) were used to understand the impact of natural environment vs anthropogenic activities on spatial and temporal SOC variability. It is now well-known that human activities, especially the ways humans occupy and manage soils, have important consequences on soil properties such as SOC contents and stocks (Post and Kwon, 2000; Guo and Gifford, 2002). Hence, the data analysis and the spatial modeling were performed separately for each landuse, i.e. cropland, grassland and vineyard. Consequently, the results are also reported by land use.

N.B.: during the previous study dedicated to SOC mapping in GDL (Stevens et al., 2014), data for forest from the National Forest Inventory (NFI; period 2008-2013) were used. As the database from the NFI was not updated at the time of this research, forest soils were not included.

3.2 METHODOLOGY

3.2.1 Recent SOC data in GDL (2012-2019)

SOC data for the period 2012-2019 were first extracted from the ASTA database compiling all the results of routine analyses performed for farmers in GDL. Then, more than 300 SOC analyses were performed on soil samples from 2019 and early 2020[4] in order to improve the spatial cover and the number of paired observations in the database. Only soil data for cropland, grassland and vineyard were considered, i.e 14,972 observations originally (Table 3).

Table 3: Number of SOC analyses extracted from the ASTA database by landuse and year of sampling.

Landuse	2012	2013	2014	2015	2016	2017	2018	2019	2020
Cropland	314	812	777	811	698	706	956	1348	68
Grassland	139	136	474	149	228	318	411	817	37
Vineyard	823	658	754	675	903	708	458	794	0
Total	1276	1606	2005	1635	1829	1732	1825	2959	105

- **SOC analysis**

SOC content is analyzed on composite soil samples. The sampling unit being the field, one soil sample corresponds to a mean representation of soils within the field (with an average surface of 1.85 ha in 2019). Samples were taken at 0-25cm in cropland, 0-15 cm in grassland and 0-30 cm in vineyard.

[4] *Despite the use of some samples from early 2020, period T2 is referred as 2016-2019.*

Sampling is done by farmers as part of Agri-Environment Measures (AEM) of RDP, with an obligation of one soil analysis per field every five years.

Soil samples are dried at < 40°C, first sieved at 2 mm according to ISO 11464, and then gently ground mechanically. Total Organic Carbon (TOC = SOC) is measured according to ISO 10694 (TOC = TC-TIC; with TC: Total Carbon and TIC: Total Inorganic Carbon). TC is measured by dry combustion, i.e. burning the sample at 1200°C in an O_2 atmosphere and analyzing the CO_2 produced with an infrared detector. Inorganic carbon (TIC) is measured by an automatic acidification of the sample with H_3PO_4 (20%) and measuring the CO_2 produced with an infrared detector. The analysis of the SOC is under accreditation according to ISO 17025.

Between November 2012 and June 2020, two successive TC analyzers were used: the Multi EA 4000 analyzer (Analytik Jena AG, Germany) and the Skalar Primacs SNC-100 Carbon / Nitrogen Analyzer (Skalar, The Netherlands). Earlier analyses made with the TruSpec CN (LECO Corporation, Michigan, USA) until October 2012 were not used in this study as the replicate errors between the TruSpec CN and the Multi EA 4000 analyzer were too large (0.3-0.37%C; Stevens et al., 2014).

Between October 2012 and August 2018, SOC content analysis was performed on the Multi EA 4000 analyzer. Since September 2018, samples are analyzed using the Skalar Primacs SNC-100. No significant bias was detected between 52 SOC measurements on both machines (38 on non-carbonated samples, 14 on carbonated samples). However, 7 out of 52 samples presented a difference higher than the enlarged incertitude *u* (Table 4; u in relative percentage %): 4 non-carbonated samples (*u* = 15%) and 3 carbonated samples (*u* = 20%). The enlarged analytical uncertainty *u* is estimated on results obtained from measuring samples of inter-laboratory test. A coverage factor k of 2 is used. Based on the comparison of the 52 samples, the mean errors (ME) were estimated and conclued that the Skalar Primacs SNC-100 (used since September 2018) produced slightly lower SOC content estimates than the Multi EA 4000 analyzer:

➔ For non-carbonated samples: ME = - 0.05 %C with 50% within [-0.15 , 0.10] %C;
➔ For carbonated samples: ME = - 0.06 %C with 50% within [-0.24 , 0.00] %C.

Table 4: Enlarged analytical uncertainties u associated to the devices used for the SOC measures used in this study.

Enlarged analytical uncertainty *u* *(relative percentage %)*			
Analyzer	**Period of use**	**Samples without carbonate minerals**	**Samples with carbonate minerals** (e.g. Calcite and Dolomite)
Multi EA 4000	Nov 2012 – Aug 2018	15%	25%
Skalar Primacs SNC-100	Sep 2018 – June 2020	15%	20%

- **Database preparation**

Each field has an identifier (named FLIK) corresponding to a unique agricultural field in the official land field information system (LPIS) of the GDL. The LPIS is the spatial register within the Integrated Administration and Control System (IACS), which ensures that payments of the EU Common Agricultural Policy (CAP) to the farmers are correctly made. Hence, LPIS identifies and quantifies agriculture land for targeting CAP payments. Each soil sample is identified by his FLIK. The ASTA soil data were merged with the LPIS considering the period 2008-2019. This allowed retrieving the location of each sample (the position of the samples were defined as the centroid of the fields[5]) as well as the crop grown and what has been effectively cultivated on the related field for each year of the period 2008-2019[6]. Then, the data were cleaned and filtered to minimize the presence of miscoded information or errors that could hamper the detection of trends in SOC data analysis by inducing biases or noises. The different steps of cleaning and filtering were:

1. Keeping the SOC analyses obtained by the devices Analytic Jena EA 4000 and Skalar Primacs SNC100 only;
2. Removing the observations from 2016 coming from soils sampled by the operator 'LAKU' (the operator sampled at 0-30cm depth instead of 0-25cm in cropland);
3. Removing the observations for which the FLIK polygons were not available;
4. Removing the duplicates by FLIK and year, and replacing them by their mean SOC value;
5. Removing all observations related to fields submitted to a recent land use change (within the 5 years preceding the sampling) considering RPG data from years of interest;
6. Removing the duplicates by FLIK and period (T1: 2012-2015 and T2: 2016-2019), and replacing them by their mean SOC value;
7. Removing FLIK not in RPG layers, i.e. miscoded or referred to former FLIK (before 2008);
8. By landuse and soil association, removing right-skewed data[7] (filtering greatest outliers).

The filtered dataset obtained was called LU-SOC-map.

3.2.2 Spatial covariates

The Digital Soil Mapping (DSM) approach used in this study fitted a statistical regression model between the soil property to predict (SOC content in %C) and independent environmental covariates at the same location. The environmental covariates were chosen considering their known influence on topsoil SOC content, i.e. their potential implication in the balance between organic matter (OM) inputs into the soil and its decomposition (or mineralization) by micro-organisms. Hence, SOC values can be predicted at unsampled locations by applying the fitted model to the spatial continuous layers of covariates.

A set of spatial layers in raster format were then prepared with a resolution of 90 meters and with the same grid topology. While some of the soil covariates were initially available at a higher resolution (e.g. the digital elevation model has a 5 m pixel resolution), we resampled all rasters to the resolution corresponding to the one of the raster with the lowest resolution (90 x 90m) by bilinear interpolation. These operations and generally all the manipulations related to spatial data were realized with the raster and sp R packages (Hijmans and van Etten, 2012; Bivand et al., 2013).

A large part of the layers of covariates used here was already present in the mapping procedure of Stevens et al. in 2014: all the layers related to the '*relief*', '*climate*' and '*land use*' sections of Table 1. Most of the layers representing the *'soil'* covariates were recently updated/modified or created thanks to the recent works of Steffen et al. (2019) (e.g., Figs. 6 and 7).

---

[5] *Converting areas (fields) to points using polygon centroids is a great simplification and is not strictly appropriate as it assumes that the spatial support is constant in shape and size (Kerry et al., 2012) but this greatly facilitate the spatial modeling procedure.*

[6] *Considering the code culture from RPG layers allowed to detect any land use change during the years preceding each sampling, which could induce outliers or a bias in the sub-datasets (the mapping procedure is applied separately to each considered land use, i.e. cropland, grassland and vineyard).*

[7] *All data superior to Q3 + 3*SE (with Q3 = 3rd quartile and Se = standard error).*

- *Relief*

We used the DEM with a resolution of 5 m from the Base de Données TOPO/CARTO (BD-L-TC) - altimetric product of the Administration du Cadastre et de la Topographie – (Fig. 2). We derived from the DEM a series of morphometric and hydrologic variables using the SAGA-GIS software (Olaya, 2004): slope, topographic position index (TPI; Jenness, 2006), upstream flow length of the RUSLE equation, eastness and northness. Eastness and northness represent the degree to which aspect is close to the East or to the North and take values in the range [-1 , 1]. Combined, these parameters are more convenient to use in spatial geometry than aspect. Hence, following Zar (1999), we converted the aspect (in degrees) into two separate continuous variables according to Eq. 1 and 2:

$$eastness = \sin(aspect * \frac{\pi}{180}) \quad (1)$$

$$northness = cos(aspect * \frac{\pi}{180}) \quad (2)$$

- *Climate*

Spatial layers representing climatic variations in GDL were created by spatial interpolation of temperature and precipitation averages for the period 1971-2000 data from weather stations in GDL and neighboring countries. Aggregated meteorological data of Luxembourg weather stations were obtained from the Observatory for Climate and Environment, Department of Environmental Research and Technology of the Luxembourg Institute of Science and Technology (LIST). This dataset includes precipitation records of 25 stations and temperature records for 7 stations. We combined this dataset with weather data of Belgium (Koninklijk Meteorologisch Instituut, KMI), France (Meteo France) and Germany (Deutscher Wetterdienst, DWD) obtained from Dr Jeroen Meersmans and that he gathered for creating precipitation and temperature maps of Belgium (Meersmans et al., 2011). Climatic maps of GDL were created by modeling temperature and precipitation with altitude using thin-plate splines regression (Stevens et al., 2014). Using altitude as covariate for mapping climatic variables can improve predictions dramatically (Boer at al., 2001). The smoothing parameter is chosen automatically by generalized cross-validation. Elevation data were derived from the Shuttle Radar Topography Mission (SRTM) mission of the NASA (Jarvis et al., 2008). The resulting temperature and precipitation maps are given in Figure 1.

- *Soil*

Three maps of textural classes (sand 50 µm-2mm, silt 2-50µm, clay < 2 µm; Fig. 6) were created based on a historical and recent multi-sources dataset (Steffen et al., 2019). Random Forest algorithms (package Rborist) was selected and tuned for the modelling of soil texture. Only the topsoil texture map (0-30 cm) was used in this study.

Maps for pH $CaCl_2$, Mg, K_2O (Fig. 7) were computed based on the standard analysis (n=43,000) performed for farmers between 2014 and 2018 (Steffen et al., 2019) by co-kriging technics.

TIC map (Total Inorganic Carbon) was obtained from analysis needed in TOC measurement protocol (ISO 10694; see section 3.2.1) on which a co-kriging with pH $CaCl_2$ was applied. The number of sample was more limited here because TIC analysis is only performed when TOC analysis is required.

The layer detailing the minimum depth of soil hydromorphy was developed within this project (Fig. 8). This parameter corresponds to the minimum depth in soil profile where physical indicators of temporary or continuous surface water saturation were observed by soil types (for further details concerning the methodology, please see ANNEX 7.1).

Table 5: Description of the environmental covariates used in the SOC mapping procedure.

Covariate	Definition	Discrete (D) or Continuous (C)	Unit	Source
Relief				
elevation	elevation	C	m	Derived from a 5m DEM provided by the 'Administration du cadastre et de la topographie' (https://data.public.lu/)
slope	slope gradient	C	°	
tpi500	Topographic Position Index (Jeness, 2006)	C	(-)	
flow length	Flow length as used in RUSLE	C	m	
eastness	aspect, orientation towards east (Zar, 1999)	C	(-)	
northness	aspect, orientation towards north (Zar, 1999)	C	(-)	
Climate				
precip	Precipitation	C	mm	Annual mean data (1971-2000) from meteo stations in GDL (AGE - Division de l'hydrologie - service hydrométrie, ASTA - Service météorologique, LIST - environmental sensing and modeling) and neighboring countries (KMI - Belgium, Meteo France - France, DWD - Germany); maps modelled using elevation from NASA SRTM DEM (Stevens et al., 2014)
temp	Temperatures	C	°C	
Soil				
clay	clay content	C	%	Maps based on a historical multi-sources dataset (ASTA-Soil Department; Steffen, 2019)
silt	silt content	C	%	
sand	sand content	C	%	
ph	pH CaCl2	C	(-)	Maps based on standard analysis performed for farmers between 2014 and 2018 (ASTA-Soil Department; Steffen, 2019)
TIC_90	Total Inorganic carbon	C	%	
Mg_90	Available Magnesium	C	mg/100g (d.s.)	
K2O_90	Available Potassium	C	mg/100g (d.s.)	
Hydromorphy min. depth	Minimum depth of hydromorphy in the soil profile	C	cm	Derived from a fusion of two digital soil maps: ~75% the 1:25000 map and ~25% of the 1:10000 map (ASTA-Soil Department; Bah et al., 2015)
Land use				
Landuse	Main land use	D	(-)	Based on a reclassification of the landuse vector layer for 2018 provided by the'Ministère de l'Environnement, du Climat et du Développement durable & Ministère de l'Énergie et de l'Aménagement du territoire' (https://data.public.lu/)
C factor	Crop factor according to RUSLE (ref.)	D	(-)	Based on analysis of crop rotation for 2012-2015; ERRUISSOL3 project (ASTA-Service de pédologie; Bah et Marx, 2016)
UF_mean	Livestock intesity	D	UF / ha	Livestock intensity of 2018 in fertilizing units per ha (UF = unité fertilisante / 80 kg N ha-1) aggregated at the farms level provided by the 'Ministère de l'Agriculture'

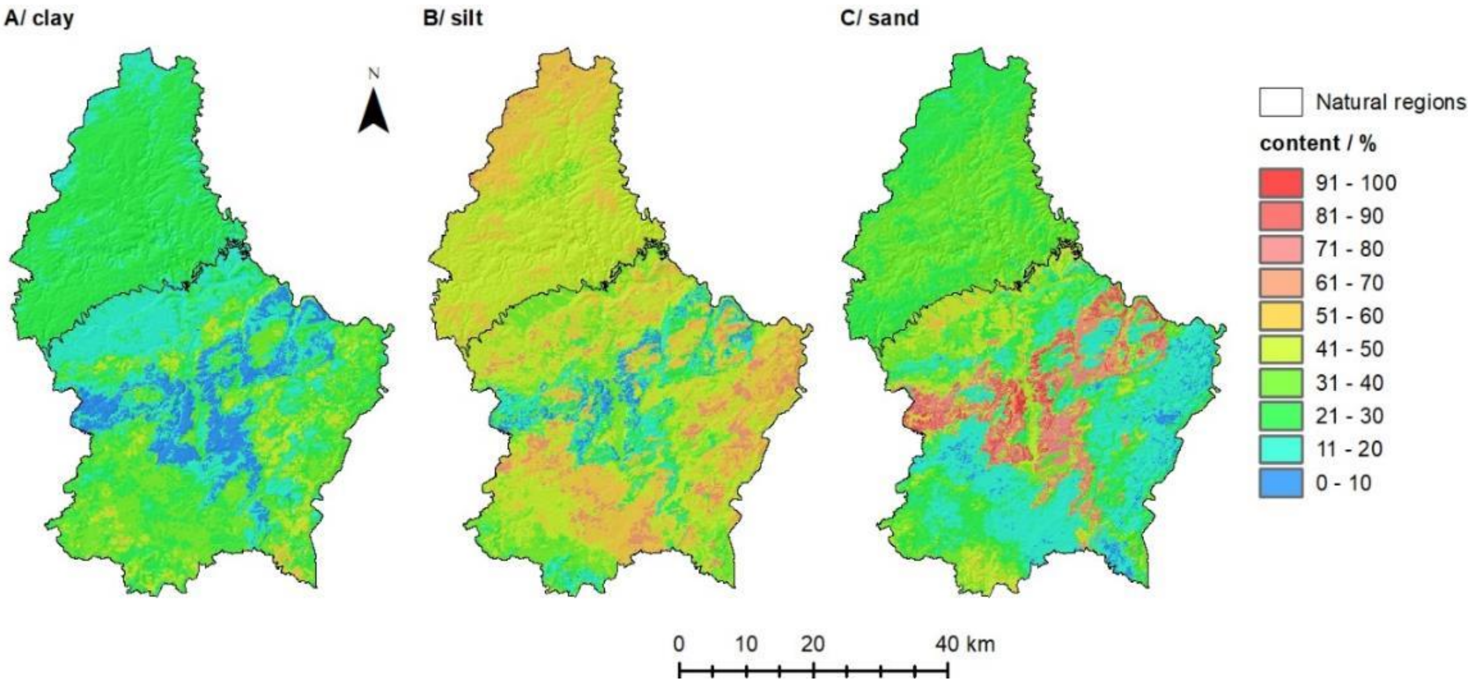

Figure 6: Map of A/ clay, B/ silt and C/sand content (%) in topsoil of Grand-Duchy of Luxembourg (source: Steffen et al., 2019).

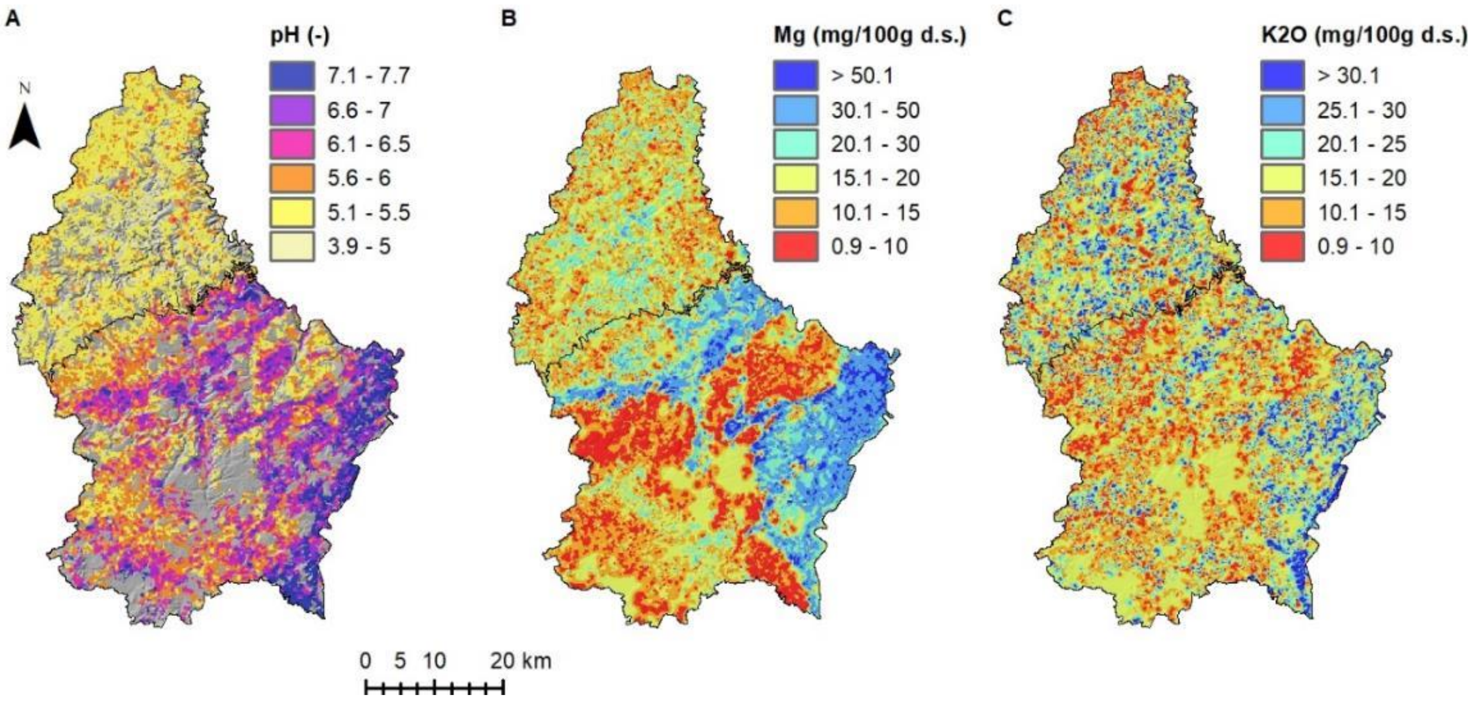

Figure 7: Map of A/ pH $CaCl_2$ (-), B/ available Mg (mg/100g of dry soil) and C/available K_2O (mg/100g of dry soil) in topsoil of Grand-Duchy of Luxembourg (source: Steffen et al., 2019) based on data for 2014-2018 period. The pH layer (A) covers agricultural soils only, i.e. croplands and grasslands.

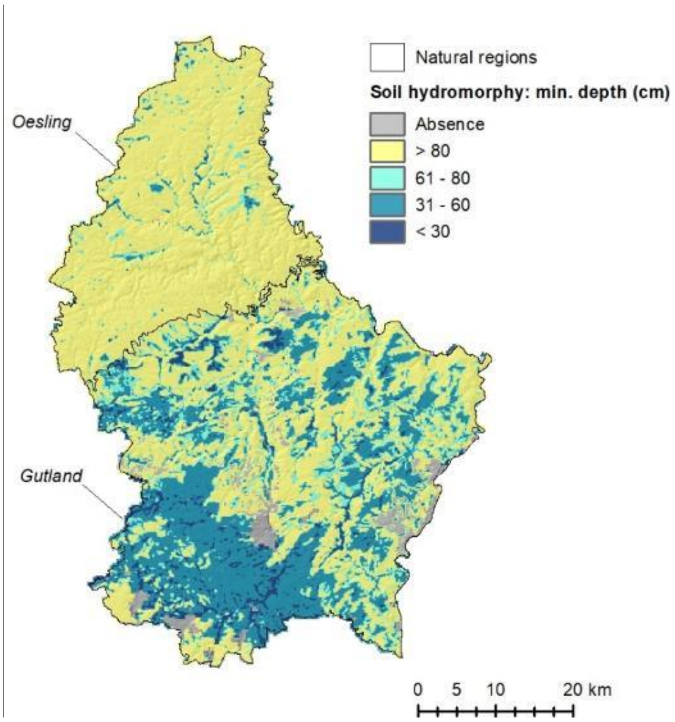

Figure 8: Map of the minimum depth of soil hydromorphy (cm) in Grand-Duchy of Luxembourg.

- *Land use & human influence*

The landuse layer used to finalize SOC maps, i.e. for applying the final models to the proper landuse areas, was derived from aerial images, LiDAR and ancillary GIS data for year 2018 (https://download.data.public.lu/resources/landcover-landuse-2018/20200504-135337/lisl-landuse-2018-documentation.pdf). The 45 'Sub-type 1 LU classes' depicted in the Landuse2018 layer provided by the Luxembourgish data platform were reclassified into five main land cover types: cropland, (permanent) grassland, vineyard, forest and other, as shown in Figure 5. The data is in polygon format and was therefore converted to raster with a 90 m resolution. We also incorporated a raster map of the C (crop) factor of the Universal Soil Loss Equation, computed from an analysis of the crop rotation systems 2013-2015. This dataset comes from the ERRUISSOL3 project aiming at mapping the risks of erosion and runoff in GDL and commissioned by ASTA. Finally, we included the livestock density of 2018 expressed in fertilizing units per ha (UF = unité fertilisante / 85 kg N ha-1) aggregated at the level of the farms.

3.2.3 Spatial SOC modelling

Attributing the values of the covariates (independent variables) to SOC observations (dependent variable) in the LU-SOC-map dataset is required to develop the spatial models with a sampling unit corresponding to a field and produce maps. The mean values of the covariates for each field were computed and attributed to the corresponding observations in LU-SOC-map through their FLIK number (see §3.2.1). Then, a model was fitted for each land use, i.e. a total of three models were fitted.

The dataset LU-SOC-map was first split according to land use and a Generalized Additive Model (GAM; Hastie and Tibshirani, 1986) was fitted on the totality of this subset (2012-2019). This regression technique is a generalization of linear regression models in which the coefficients can be a set of smoothing functions, then accounting for the non-linearity that could exist between the dependent variable and the covariates Eq. (3):

$$g\,[\![\mu\,(Y)]\!] = \alpha + f_1 x_1 + f_2 x_2 + \cdots + f_p x_p \qquad (3)$$

where Y is the dependent variable, X1,X2,…Xp represent the covariates and the fi's are the smooth (non-parametric) functions. As for generalized linear models, the GAM approach specifies a distribution for the conditional mean μ(Y) along with a link function g relating the latter to an addictive function of the covariates. The LU-SOC-map dataset being continuous and strictly positive, we applied a Gamma distribution in the GAM model. The log-link function was chosen for the model fitting considering the positively-skewed unimodal characteristic of the SOC content distribution. The GAM model was built using regression splines, and the smoothing parameters were estimated by penalized Maximum Likelihood to avoid an over-fitting (Wood, 2001). An extra penalty added to each smoothing term allowed each of them to be set to zero during the fitting process in case of multi-collinearity or concurvity[8].

The aim of the modeling procedure being to spatialize the SOC content all over GDL, the geographical coordinates (x, y) were integrated in each model (as a two-dimensional spline on latitude and longitude) to account for the spatial dependence and main trends of the target variable at the regional scale. Then, a first model with all the covariates was developed followed by a backward selection of the terms using their approximate p-values. This was done by sequentially dropping the single term with the highest non-significant p-value from the model and re-fitting until all terms are significant as indicated in Wood (2001). The level of significance was set at $p < 0.05$.

After the model calibration, the landuse subset was split in two periods (T1 and T2). Then, we estimated the goodness-of-fit of the model for each period by computing a stratified 10-fold cross-validation on each landuse-period subset. The stratification of the cross-validation was performed considering the soil associations in order to keep a balance on their representation at each fold. Model accuracy was evaluated with the Mean Error (ME; Eq. 4):

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \qquad (4)$$

where $\hat{y}_i$ is the predicted value of observation i in the validation set, y_i is the observed value and n the total number of observations in the validation set. We also computed the Root Mean Square Error (RMSE, Eq. 5):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((\hat{y}_i - y_i)^2)} \qquad (5)$$

After this validation phase, a final model was built with all the samples (i.e. in both the calibration and the validation sets) using the covariates selected by the stepwise procedure, in order to improve model accuracy and representativity over the GDL territory. The model for each landuse was then applied to the stack of spatial layers (covariates) to map topsoil SOC content and associated model uncertainties (the mgcv package provides a Bayesian approach to compute standard errors and confidence interval for the model predictions). We should note that RMSE computed in Eq. 5 does not give a correct

[8] *Concurvity defines the non-linear form of collinearity, i.e. the non-linear dependencies among the predictor variables here.*

measure of the true map accuracy, which should preferably be assessed using a set of samples taken from randomly selected locations (Loveland and Webb, 2003).

3.2.4 Comparing SOC at two periods

- *Observed data*

For each landuse /soil association class, the significance of difference between the distribution of SOC_{T1} and SOC_{T2} was tested with a non-parametric Wilcoxson test. The tests were applied on subsets of LU-SOC-map (non-paired tests) and on paired observed data (paired tests).

NB: Here, the difference of statistical distribution between two datasets was tested. The analytical uncertainties were not considered (see 3.2.1).'

Indeed, each site (i.e., FLIK) with one observation for both periods (one pair of observations) was identified. The difference in SOC content dSOC (%C; Eq. 6) between periods and associated statistics were computed:

$$dSOC = SOC_{T2} - SOC_{T1} \qquad (6)$$

- *Predicted data*

Two SOC maps were produced by applying the methodology proposed in section 3.2.3: the first map SOC_{T1} for the period 2012-2015, and the second map SOC_{T2} for the period 2016-2019 (both in %C). The relative differences (%) between the two predictions were computed (Eq. 7).

$$Relative\ dSOC = \frac{dSOC}{SOC_{T1}} * 100 \qquad (7)$$

A visual analysis was made on the spatial distribution of the relative differences obtained[9], and basic statistics on absolute differences were computed for each landuse.

N.B.: As explained in section 3.2.1, the enlarged SOC estimation of analytical uncertainties (in relative percentage) are 15% for non-carbonated samples and 20% for carbonated samples.

3.3 Results

3.3.1 Implications of the SOC data filtering

After the cleaning and the filtering of the extracted raw data, the LU-SOC-map dataset contained 11,819 observations, including 4820 for the period T1 (2012-2015) and 6999 for the period T2 (2016-2019). As we did not keep the SOC data analyzed by the LECO device, the effective total period covered was from October 2012 to June 2020, which can explain why the number of observations is smaller for T1. Not considering the observations obtained with the Tru Spec CN analyzer induced a loss of 1101 observations (Table 6). Replacing the duplicates, by year first and then by period, by their mean values diminished the dataset by 1750 observations.

[9] *Due to the complex distribution of the different landuses in GDL and the autocorrelation of the variable Corg, no valid statistical test could have been applied to test the difference significance between the two rasters Corg T2 and Corg T1.*

Table 6: Filtering steps on the LU-SOC-map database preparation and associated numbers of observations.

Filtering step	Total Obs.	Obs. eliminated
None	**14972**	-
- Tru Spec CN analyzer	13871	1101
- LAKU 2016 cropland	13772	99
- FLIK NA	13390	382
- Duplicates by year	12676	714
- LU of no interest	12658	18
- Duplicates by period	12216	1047
- FLIK not in RPG	12016	200
- Soil association NA	11921	95
- outliers	**11819**	102

3.3.2 SOC by period - descriptive statistics and differences test

The landuse has a major role on the topsoil SOC (Figs 9 and 10). Croplands and vineyards showed close SOC, while grassland showed SOC almost two times higher. A small proportion of observations showed SOC content < [1.1 , 1.2%C] (~2% de MO), i.e. soils depleted in SOC with a poor potential for aggregation (van Camp et al., 2004). These observations corresponded mainly to sandy soil under cropland formed on the Grès du Luxembourg (haplic Luvisol (arenic), Arenosol; Fig. 10). Although both histograms for cropland showed a unimodal distribution dominated by observations from Gutland (mode around 1.2 - 1.5 %C; Fig. 9), the subdatasets for Oesling are visible around 2.5 – 3.0 %C (soil association 1 in Fig. 10). The comparison of those two histograms show that Oesling is proportionally more represented in T2 than in T1 subset. This is mainly due to a recent specific agricultural advisory service (LAKU) in the drinking water protection area around Esch-sur-Sûre.

The coverage of croplands in GDL was more homogeneous for T2 than for T1 (Figs. 11 and 12), especially in the natural region of Oesling. While grasslands were represented by almost twice as much observations for period T2 than for T1 (Fig. 10), the coverage of GDL appeared more homogeneous for T1. Some areas of grasslands within the eastern and southern parts of the natural region Gutland were covered by few or no samples for the period T2, and grasslands from northwest (west of Wiltz canton and north of Redange canton located in Oesling) showed a higher density of observations in T2.

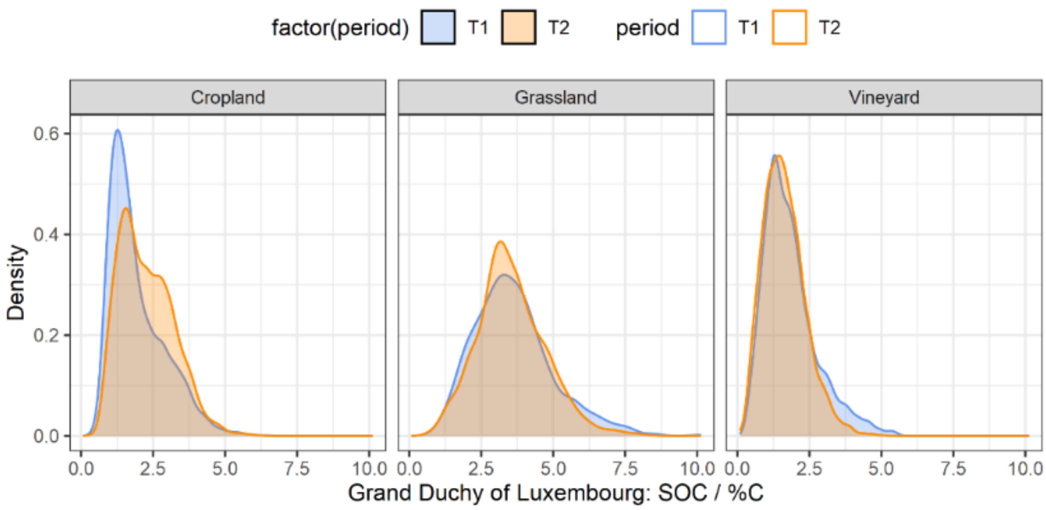

Figure 9: Histograms of topsoil SOC (%C) at T1 (2012-2015) and T2 (2016-2019) for croplands, grasslands and vineyards in Grand-Duchy of Luxembourg.

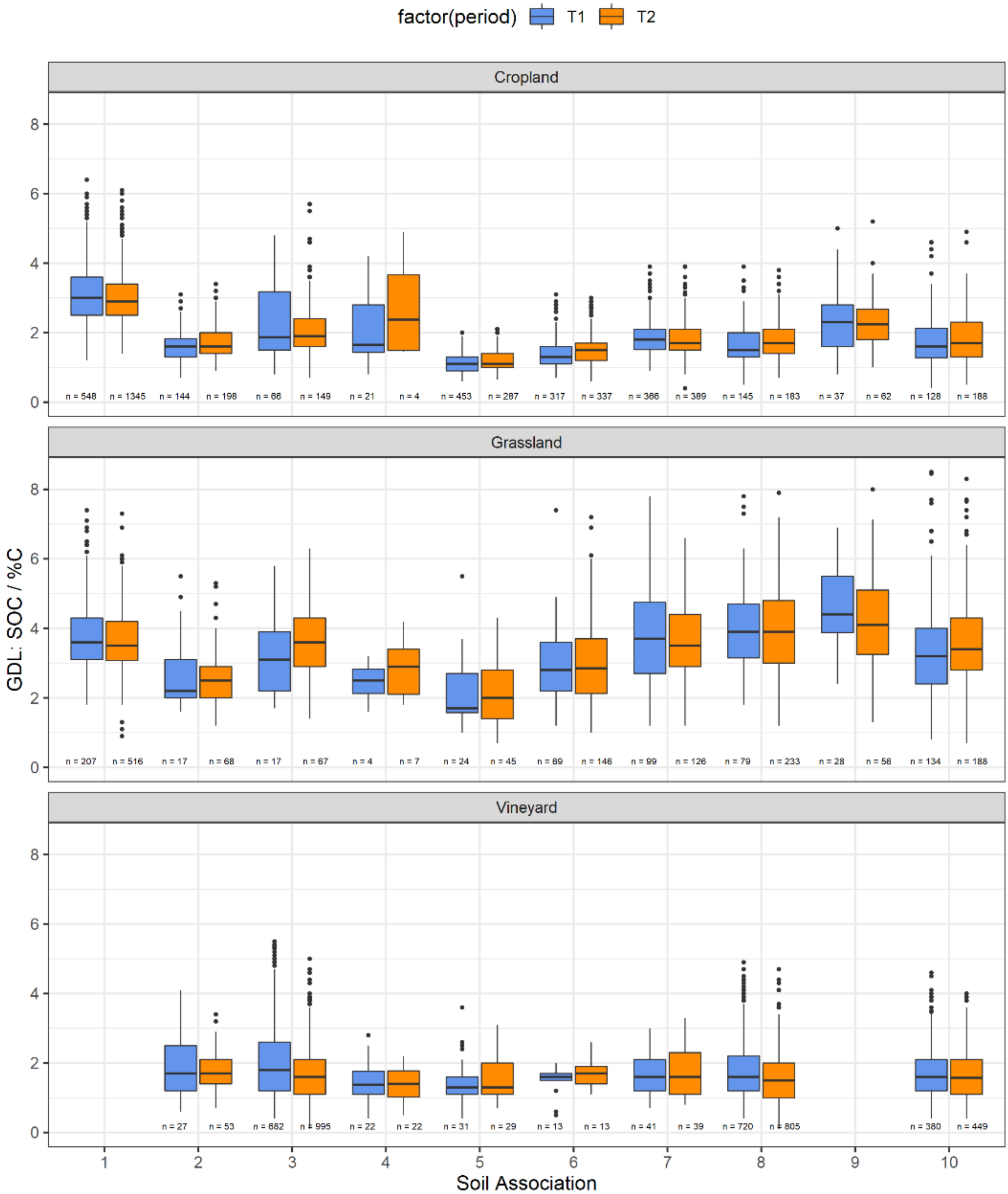

Figure 10: Box-plots of topsoil SOC (%C) in croplands, grasslands and vineyards per regrouped soil associations at periods T1 (2012-2015) and T2 (2016-2019) in Grand-Duchy of Luxembourg. (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)

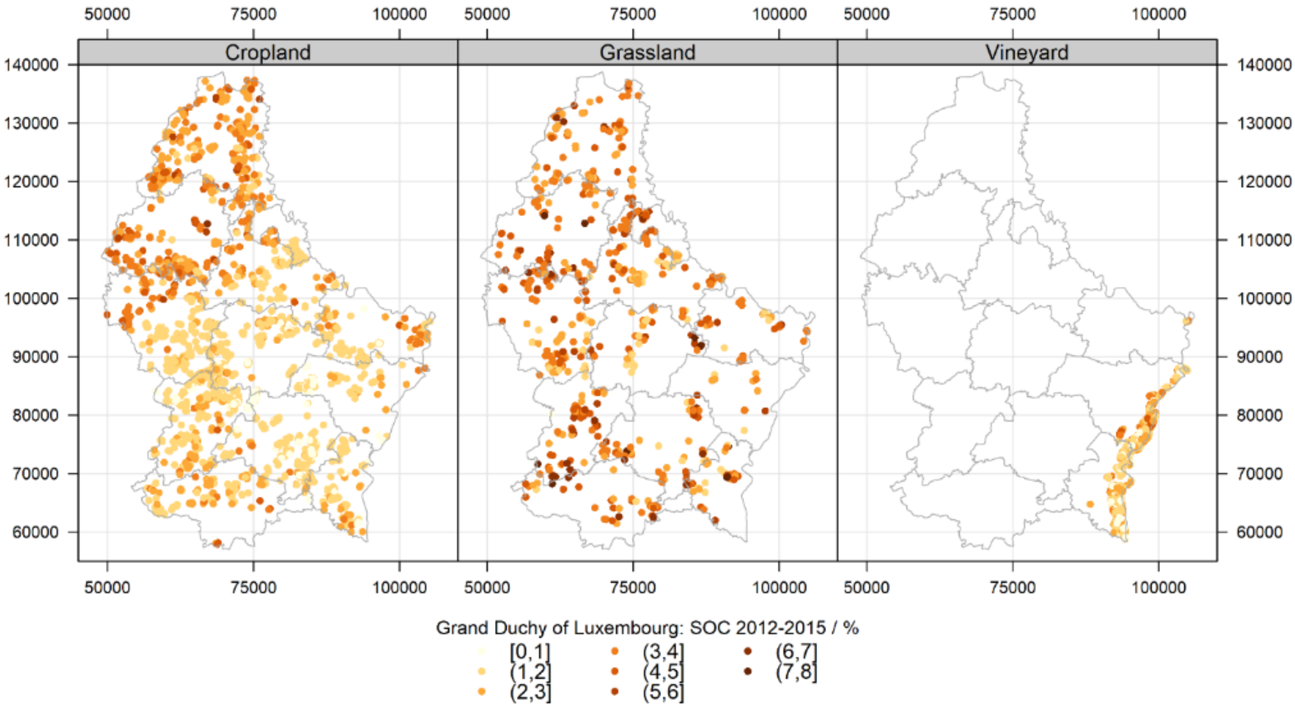

Figure 11: Observed SOC values (%C) of topsoil under cropland, grassland and vineyard for period T1 (2012-2015).

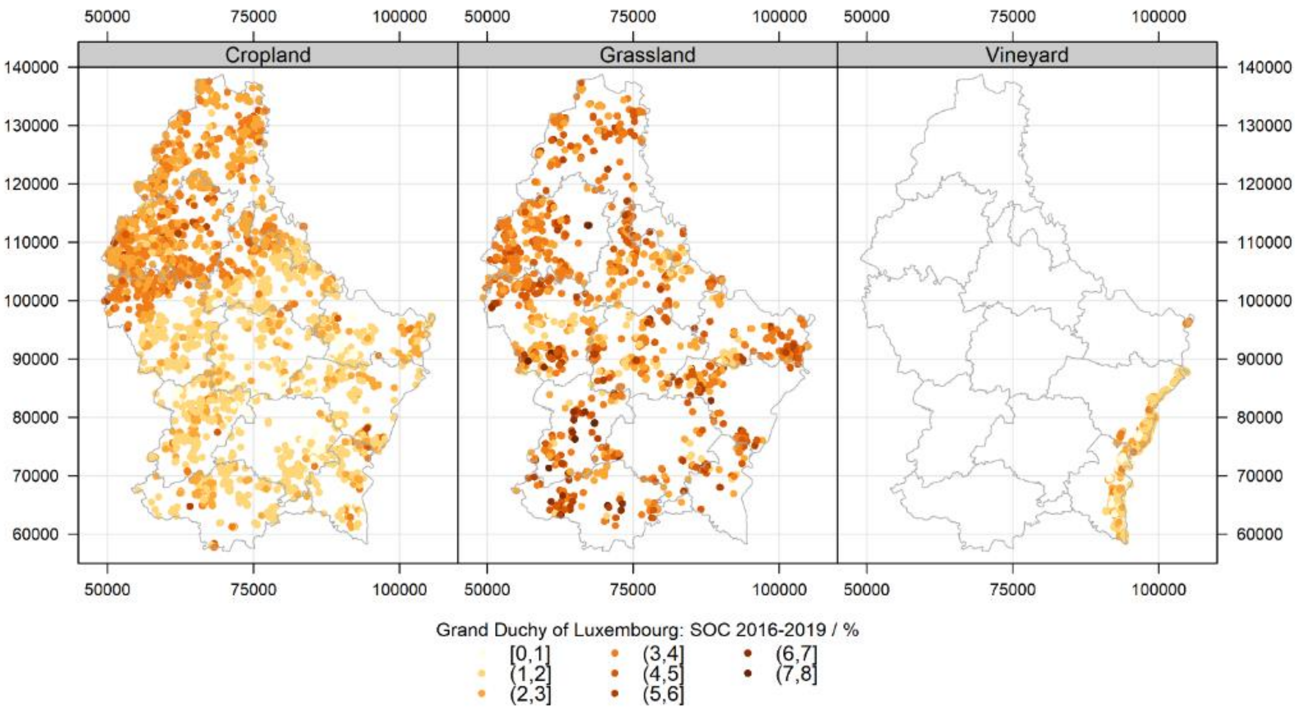

Figure 12: Observed SOC values (%C) of topsoil under cropland, grassland and vineyard for period T2 (2016-2019).

We propose to classify SOC observations based on soil associations (which also reflect climatic conditions) and use the median and quartiles (Q1 and Q3) of the SOC distributions (Tables 7 - 9) to define indicative SOC range in each soil association-landuse class. In each category, any value within the boundaries defined by the quartiles (i.e., [Q1 , Q3]) can be qualified as normal or standard values prevailing under environmental conditions related to the period considered. Values outside [Q1 , Q3] indicate soils either depleted or enriched in SOC in relation to the majority of observations in the same soil association.

N.B.1: In each land use-period class, statistics describing each subset (T1 and T2) without regard of the soil association (i.e., 'ALL' in Table 7 – 9) can NOT be compared directly as the soil associations are represented by very different proportions of samples in T1 and T2.

N.B.2: In this section 3.3.2, the differences of statistical distribution between different datasets were tested (Tables 7 to 10). The analytical uncertainties were not considered (see 3.2.1).

- **Cropland**

Table 7 compiles the descriptive statistics about SOC content in topsoil under cropland during the periods T1 and T2 showing large variations amongst the different soil associations. Large variations were also apparent when samples are grouped according to the four texture classes defined by ASTA soil laboratory (Figure 3B), giving the following sequence in terms of SOC content: L (leicht = light texture) < M (mittel = medium texture) < S (schwer = heavy texture) < MO (mittel Oesling = medium texture of the Oesling region, stony soils) (see ANNEX 7.2). With more details, the subset representing the period T1 had a median of 1.6%C with a range between quartiles of [1.2 , 2.5]%C. The Oesling region (1) reached a median of 3.0%C while in Gutland the median values ranged from 1.10%C ("Grès du Luxembourg" -5) to 2.3%C ("Argiles lourdes des schistes bitumineux" - 9). The subset representing the period T2 had a median of 2.1%C and a range between quartiles of [1.5 , 2.9]%C. As observed above, the subset T2 contains a greater proportion of samples from Oesling than the subset T1. And, as for T1, the soils under cropland in Oesling during T2 contain more SOC than the other soil associations with a median values of 2.9%C. However, SOC contents in "Oesling" slightly decreased from T1 to T2 (mean decrease of -0.09%C; $p<0.05$). Four soil associations showed significant increase of their SOC content: "Buntsandstein" (+0.12%C; $p<0.05$), "Grès du Luxembourg" (+0.07%C ; $p<0.05$), "Dépôts limoneux sur Grès" (+0.14%C; $p<0.01$) and "Argiles lourdes des schistes bitumineux" (+0.09%C ; $p<0.05$). Finally, five out of the ten soil associations showed no significant evolution in SOC contents.

These SOC observations in croplands are similar to values published in Belgium for similar environmental conditions for the period 2004-2014 (SPW - DGO3 - DEMNA - DEE, 2017). In the Belgian Ardennes, corresponding to the Oesling region, the mean is 3.15%C (Q1 = 2.90%C, Q3 = 3.38%C) and in the Belgian Jurassic region, corresponding roughly to the Gutland, the mean is 1.78%C (Q1 = 1.30%C, Q3 = 2.09%C).

*Table 7: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** at T1 (2012-2015) and T2 (2016-2019), and significance of the difference between these two periods (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	2225	0.40	1.20	1.60	1.94	2.50	6.40	3142	0.40	1.50	2.10	2.27	2.90	6.10		
1	548	1.20	2.50	3.00	3.11	3.60	6.40	1345	1.40	2.50	2.90	3.02	3.40	6.10	**-0.09**	**< 0.05**
2	144	0.70	1.30	1.60	1.62	1.83	3.10	198	0.90	1.40	1.60	1.74	2.00	3.40	**0.12**	**< 0.05**
3	66	0.80	1.50	1.87	2.26	3.18	4.80	149	0.70	1.60	1.90	2.16	2.40	5.70	-0.09	NS
4	21	0.80	1.43	1.65	2.09	2.80	4.20	4	1.45	1.49	2.38	2.78	3.66	4.90	0.68	NS
5	453	0.60	0.90	1.10	1.12	1.30	2.00	287	0.65	1.00	1.10	1.19	1.40	2.10	**0.07**	**< 0.05**
6	317	0.70	1.10	1.30	1.38	1.60	3.10	337	0.60	1.20	1.50	1.52	1.70	3.00	**0.14**	**< 0.01**
7	366	0.90	1.51	1.80	1.86	2.10	3.90	389	0.40	1.50	1.70	1.84	2.10	3.90	-0.02	NS
8	145	0.50	1.30	1.50	1.69	2.00	3.90	183	0.70	1.40	1.70	1.78	2.10	3.80	**0.09**	**< 0.05**
9	37	0.80	1.60	2.30	2.39	2.80	5.00	62	1.00	1.80	2.24	2.26	2.68	5.20	-0.13	NS
10	128	0.40	1.28	1.60	1.80	2.13	4.60	188	0.50	1.30	1.70	1.85	2.30	4.90	0.05	NS

- **Grassland**

Table 8 compiles the descriptive statistics about SOC content in topsoil under grassland during the period T1 and T2. For both periods, the SOC content in grassland of Oesling and Gutland are similar with median values of 3.4%C. There are however large variations between soil associations in Gutland. Median SOC values reach ca. 4.00%C for clay-rich soils ("Argiles du Lias Inf. et Moyen","Argiles lourdes du Keuper" and "Argiles lourdes des schistes bitumineux"), while loamy and sandy soils of ("Buntsandstein", "Dolomies du Muschelkalk", "Calcaires du Bajocien", "Grès de Luxembourg", "Dépôts limoneux sur Grès") have median SOC generally less than 3.0%C. As illustrated in Figure 8, the differences between Quartiles for grassland appeared larger than in cropland soils indicating larger variation of SOC in grassland than cropland. Only the soil association "others", regrouping mainly the "Alluvions et Colluvions", showed significant differences in SOC content between T1 and T2 (mean of +0.26%C; $p<0.05$).

*Table 8: Descriptive statistics of topsoil SOC (%C for the 0-15cm depth) in **grasslands** at T1 (2012-2015) and T2 (2016-2018), and significance of the difference between these two periods (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	679	0.80	2.70	3.40	3.62	4.35	10.10	1452	0.70	2.80	3.40	3.55	4.30	8.30		
1	207	1.80	3.10	3.60	3.82	4.30	7.40	516	0.90	3.08	3.50	3.65	4.20	7.30	-0.17	NS
2	17	1.60	2.00	2.20	2.80	3.10	5.50	68	1.20	2.00	2.50	2.56	2.90	5.30	-0.24	NS
3	17	1.70	2.20	3.10	3.14	3.90	5.80	67	1.40	2.90	3.60	3.60	4.30	6.30	0.45	NS
4	4	1.60	2.13	2.50	2.45	2.83	3.20	7	1.80	2.10	2.90	2.84	3.40	4.20	0.39	NS
5	24	1.00	1.58	1.70	2.15	2.70	5.50	45	0.70	1.40	2.00	2.18	2.80	4.30	0.03	NS
6	69	1.20	2.20	2.80	2.99	3.60	7.40	146	1.00	2.13	2.85	3.04	3.70	7.20	0.05	NS
7	99	1.20	2.70	3.70	3.87	4.75	7.80	126	1.20	2.90	3.50	3.64	4.40	6.60	-0.23	NS
8	79	1.80	3.15	3.90	3.96	4.70	7.80	233	1.20	3.00	3.90	3.93	4.80	7.90	-0.03	NS
9	29	2.40	3.90	4.40	4.83	5.80	10.10	56	1.30	3.25	4.10	4.15	5.10	8.00	-0.68	< 0.10
10	134	0.80	2.40	3.20	3.43	4.00	8.50	188	0.70	2.80	3.40	3.69	4.30	8.30	**0.26**	**< 0.05**

- **Vineyard**

Table 9 compiles the descriptive statistics about SOC content in topsoil under vineyard during period T1 and T2. ***Vineyards soils are predominantly located on 'Dolomies du Muschelkalk', 'Argiles Lourdes du Keuper' and 'Others', and to a lesser extent on 'Buntsandstein'. The location of dozens of TOC observations on other geological formations than these four latter could be an artefact induced by the map of soil associations (Fig. 3A).*** Soils under vineyard have about the same median content and interquartiles differences as cropland soils of GDL (Tab. 7; Fig. 10). In vineyard soils, soil associations showed a median SOC content oscillating around 1.60-1.80%C for T1, and 1.50-1.70%C for T2, indicating low interclass variation. From T1 to T2, two soil associations showed significant evolution of their SOC contents: "Dolomies du Muschelkalk" and "Argiles lourdes du Keuper" with respective mean decreases of -0.38%C and -0.16%C (both at p<0.01).

*Table 9: Descriptive statistics of topsoil SOC (%C for the 0-30cm depth) in **vineyards** at T1 (2012-2015) and T2 (2016-2018), and significance of the difference between these two periods (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	1916	0.40	1.20	1.65	1.83	2.21	5.50	2405	0.10	1.10	1.55	1.63	2.10	5.00		
2	27	0.60	1.20	1.70	1.85	2.50	4.10	53	0.70	1.40	1.70	1.78	2.10	3.40	-0.06	NS
3	682	0.40	1.20	1.80	2.05	2.60	5.50	995	0.10	1.10	1.60	1.66	2.10	5.00	**-0.38**	**< 0.01**
4	22	0.40	1.10	1.38	1.46	1.76	2.80	22	0.50	1.03	1.40	1.39	1.78	2.20	-0.07	NS
5	31	0.40	1.10	1.30	1.45	1.60	3.60	29	0.70	1.10	1.30	1.52	2.00	3.10	0.07	NS
6	13	0.50	1.50	1.60	1.49	1.70	2.00	13	1.10	1.40	1.70	1.68	1.90	2.60	0.19	NS
7	41	0.70	1.20	1.60	1.65	2.10	3.00	39	0.80	1.10	1.60	1.75	2.30	3.30	0.1	NS
8	720	0.40	1.20	1.60	1.73	2.20	4.90	805	0.10	1.00	1.50	1.57	2.00	4.70	**-0.16**	**< 0.01**
10	380	0.40	1.20	1.60	1.72	2.10	4.60	449	0.40	1.10	1.58	1.66	2.10	4.00	-0.06	NS

- **Comparison of paired observed data**

Within the LU-SOC-map dataset (n = 11,819), we identified 560 sites with paired observations (i.e., FLIKs with 1 observation for each period) for croplands, 149 for grasslands and 1027 for vineyards. Figure 13 presents the spatial location of these sites by landuse. Those in croplands are rather evenly spread over the GDL territory. Those in grasslands are concentrated in the Oesling and in the western part of Gutland. Finally, the sites in vineyards well cover the valley of Mosel where this landuse is concentrated.

Considering the sites classed by land use and soil association, it was difficult to produce proper statistics and test the difference (Fig. 14). Indeed, 3 out of 10 soil associations showed less than 30 pairs of observations for croplands, 8 out of 10 for grasslands and, 5 out of 8 for vineyards. No test or statistics computed on these groups (n < 30) should be considered robust (Tab. 10; results in gray italic correspond to groups under 30 sites).

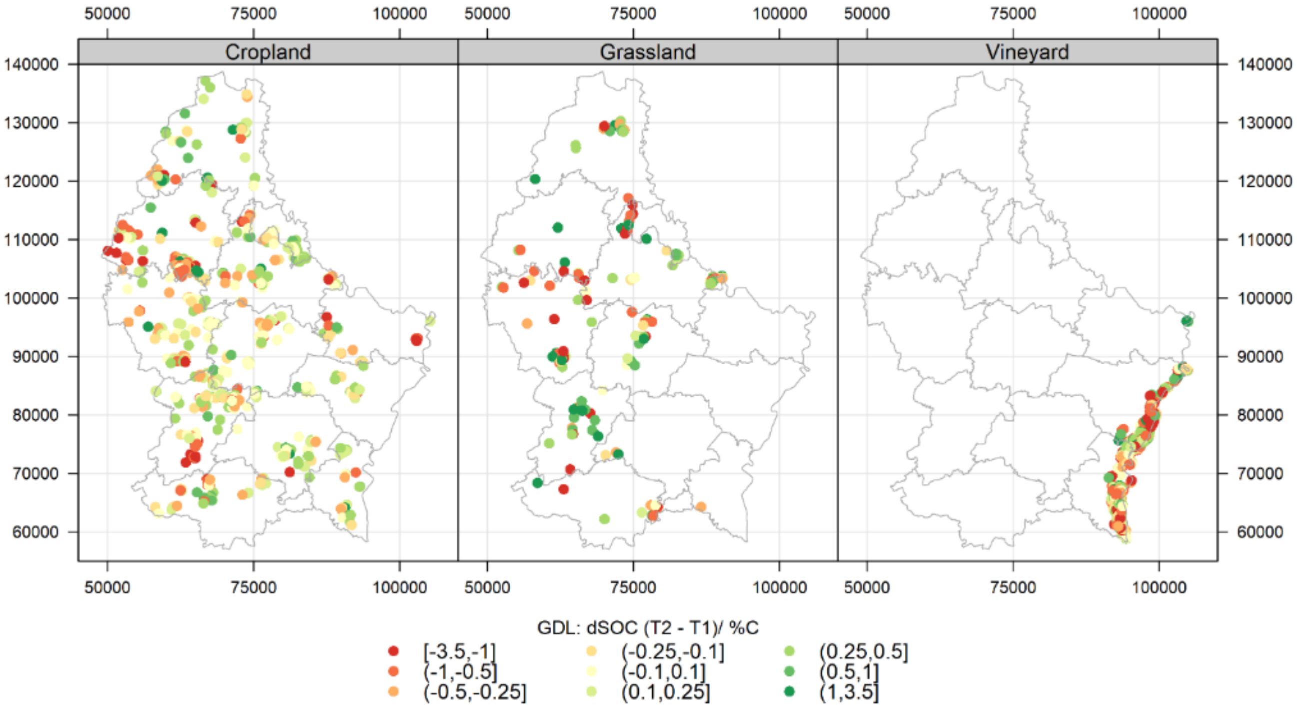

Figure 13: Location and differences in SOC content (dSOC in %C) between T1 (2012-2015) and T2 (2016-2019) computed on paired observations (same FLIKs) in Grand-Duchy of Luxembourg.

Tests on paired observations confirmed that a significant increase in SOC content occurred in croplands for soils of “Grès de Luxembourg” (mean of +0.07%C; $p<0.05$) between T1 and T2. While a non-significant decrease in SOC (-0.02%C; Tab. 7) was observed for soils on “Argiles du Lias inférieur et moyen” when considering all the data, a significant decrease of -0.16%C ($p<0.05$) was detected here considering 92 paired observations.

None of the soil associations with more than 30 sites in grasslands showed significant differences in SOC between T1 and T2. However, we noticed that a significant mean increase of +0.46%C ($p<0.05$) was detected for the 18 paired observations in soil association “others” (alluvium and colluvium).

Finally, significant decreases were confirmed for soils in vineyards developed on “Dolomies du Muschelkalk” (-0.39; $p<0.01$) and “Argiles lourdes du Keuper” (-0.20; $p<0.01$).

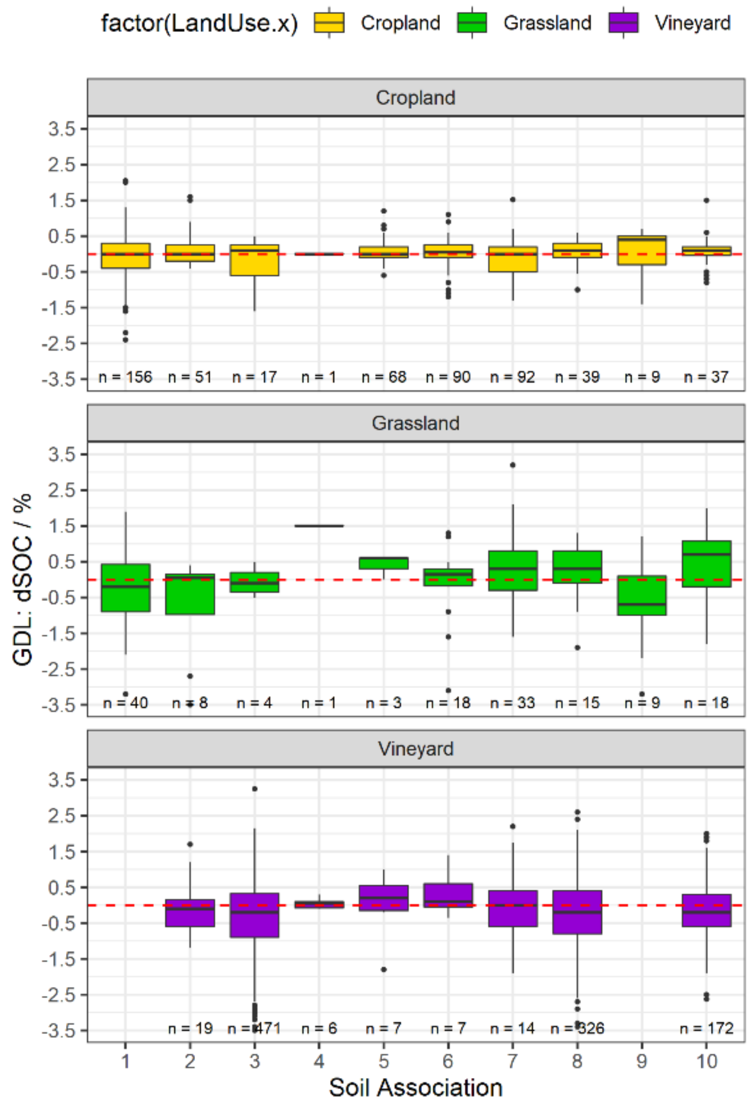

Figure 14: Box-plots of differences in SOC content (dSOC in %C) between T1 (2012-2015) and T2 (2016-2019) computed on paired observations (same FLIKs) in Grand-Duchy of Luxembourg.

Table 10: Results of the paired Wilcoxson test computed on paired observations by landuse and soil association, comparing SOC evolution from T1 to T2. (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)

	Cropland			**Grassland**			**Vineyard**		
Soil Assoc.	**n**	**mean dif.**	**p-value**	**n**	**mean dif.**	**p-value**	**n**	**mean dif.**	**p-value**
1	156	-0.07	NS	40	-0.20	NS			
2	51	0.08	NS	8	-0.71	NS	19	-0.09	0.3936
3	17	-0.21	NS	4	-0.05	NS	475	**-0.39**	**< 0.01**
4	1	0.00	-	1	1.50	-	6	0.05	0.4982
5	68	**0.07**	**< 0.05**	3	0.40	NS	7	0.03	0.5781
6	90	0.03	NS	18	-0.02	NS	7	0.32	0.4017
7	92	**-0.16**	**< 0.05**	33	0.29	NS	14	0.06	0.9249
8	39	0.03	NS	15	0.22	NS	327	**-0.20**	**< 0.01**
9	9	0.03	NS	9	-0.73	NS			
10	37	0.08	NS	18	**0.46**	**< 0.05**	172	**-0.15**	**< 0.1**

3.3.3 SOC relations to spatial covariates

The covariates that potentially contribute to the modeling procedure are presented here. **We paid more attention to the form and dispersion of the plots than the Pearson correlation coefficient, as the latter considers only linear relations while the GAM are designed to handle non-linear relations.** The main drivers of SOC spatial distribution in GDL, other than the landuse, are coordinates (x, y), soil type (texture), climatic condition and elevation. In addition, we observed conspicuous relations between topsoil SOC and new *'soil'* covariates implemented in this project (section 3.2.2; Tab. 5). The latter are both influenced by natural processes and agricultural activities.

- **Cropland**

We observed in §3.3.2 a clear dichotomy in SOC level between the Oesling and Gutland (Figs 11 and 12). The Oesling showed higher SOC level than the Gutland describing a negative trend from NW to SE in GDL (see the shapes and the strengths of the relations between SOC and the geographic coordinates in Figure 15A). As observed in many other studies, soil texture, especially clay content, has an important role in SOC spatial distribution in croplands (Fig. 16). This is especially true for the Gutland region. Indeed the Oesling has soils showing a smaller textural variability (clay content mainly between ~ 15 to 25%) than soils of Gutland (clay content mainly between 5 and 45%). Consequently, the relation between clay and SOC is of ρ=0.57 in Gutland and ρ=0.15 in Oesling. Clay fraction has an important role on SOC through chemical stabilization (Six et al., 2002) and higher soil moisture content (due to poor drainage status) leading to lower SOC mineralization rates (Skopp et al., 1990; Davidson and Janssens, 2006). Also, fine-textured soils with their greater nutrient and water-holding capacity favor plant production and thereby the amount of fresh OM returning to the soil.

The relation of SOC with elevation was strong (ρ=0.57). The same is true for precipitation (ρ=0.53) and temperature (ρ=-0.54; Fig. 15B). The two natural regions present very different geomorphologies, especially their ranges of altitude (Fig. 1) associated to different climatic contexts. The high plateaus of Oesling experience rainier and colder conditions than the lower cuestas of Gutland. Higher SOC content in areas with higher precipitation and lower temperature is often observed due the effects of precipitation on Net Primary Productivity, lower level of oxygen concentration in wetter soils (anaerobic conditions), and decreased microbial activity or decomposition of organic matter in colder climate (Kirschbaum, 1995; Post et al., 1982; Trumbore et al., 1998). However, soils of Oesling present a good infiltration capacity due to their texture, stoniness and SOC content. SOC content and slope showed a weak positive relation here (ρ=0.18) due to a regional effect as Oesling present higher slope gradients and SOC contents than Gutland.

A clear relation was observed between SOC and the C-factor (ρ=-0.35). The values of C-factor ranges between 0 and 1, increasing when crop cover decreases (i.e. 1 is for bare soils). The crop cover influences the rate of OM incorporation into the soils through plant residues and OM decomposition through variations in runoff and erosion rates. Livestock intensity showed no significant relation with SOC in croplands (ρ=0.05) maybe due to the data aggregation at the farm level (not the field), and/or that the regional natural variations (induced by climate, elevation and texture) could hide the local effect of farming practices.

Amongst the environmental covariates included in the procedure here, we observed clear relations of SOC content with available Mg (ρ=0.22) and K_2O (ρ=0. 27). The available Mg is linked to the presence of dolomite or dolomitic marls in the soils, coming mainly from geology but also from amendments.

The available K_2O is linked to the amount of some minerals in the soils; e.g., mica, feldspath and illite (Steffen et al., 2019). K_2O also depends on the (historical and actual) type and amount of fertilizer and organic amendments applied. This is less the case for Mg. Mg and K_2O are both important elements needed for plant growth and development: the first as primary macronutrient and the second as secondary macronutrient (Parikh and James, 2012). Minimum depth of hydromorphy between 0 and 80cm showed a clear negative relation with SOC. Indeed, SOC content increased while drainage deteriorated, i.e. while hydromorphy features appeared closer to the surface. The decomposition of Organic Matter is slower under anaerobic conditions (Gale and Gilmour, 1988).

- **Grassland**

No clear regional trend was observed for topsoil SOC in grassland of GDL (Figs. 11 and 12). Consequently, coordinates and climatic covariates showed no clear relation with SOC (Fig. 17A). Amongst covariates linked to management, the plots showed a positive relation with available Mg and K_2O, although these appear weak when considering linear relations (ρ=0.20 for both; Fig. 17B). Clay and sand contents have a strong influence on SOC level in grasslands of Gutland ($\rho = 0.55$ and $\rho = -0.50$, respectively; Fig. 16). Similar to the croplands, minimum depth of hydromorphy < 80cm depth showed a clear negative relation with SOC.

- **Vineyard**

Vineyards in GDL are concentrated in the Mosel river valley characterized by rather homogeneous environmental conditions, which can explain the poor correlations between SOC and environmental covariates (Fig. 18). Moreover, vineyards are located where land consolidation and terracing have occurred in the past. However, as for grassland and cropland, we observed a positive but weak relation with available Mg, ($\rho = 0.20$).

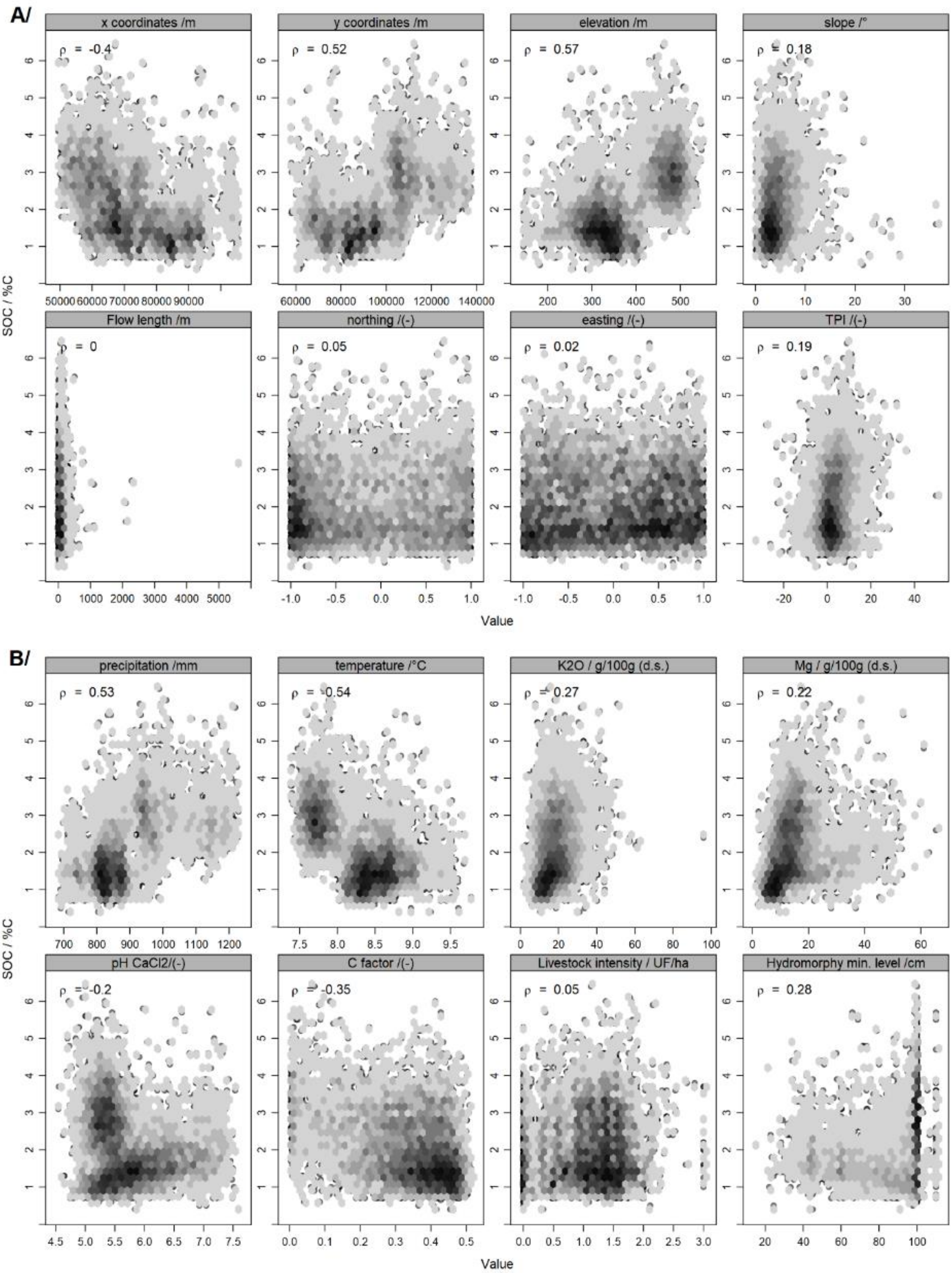

Figure 15: Scatter plots of topsoil SOC (%C) ***in croplands*** *as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and landuse parameters.*

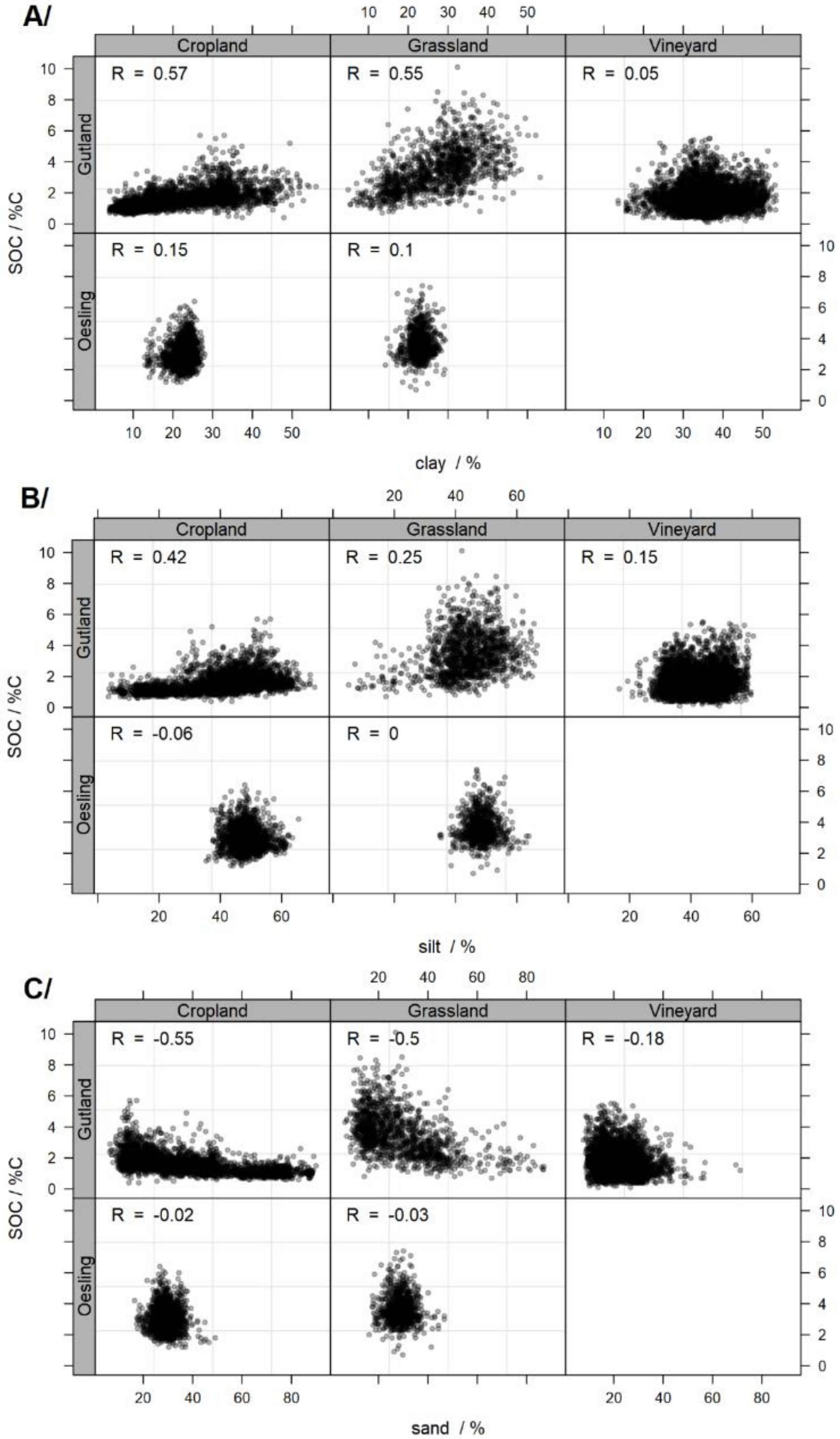

Figure 16: Scatter plots of topsoil SOC (%C) by land use and natural region as function of A/ clay content (%), B/ silt content (%) and C/ sand content (%).

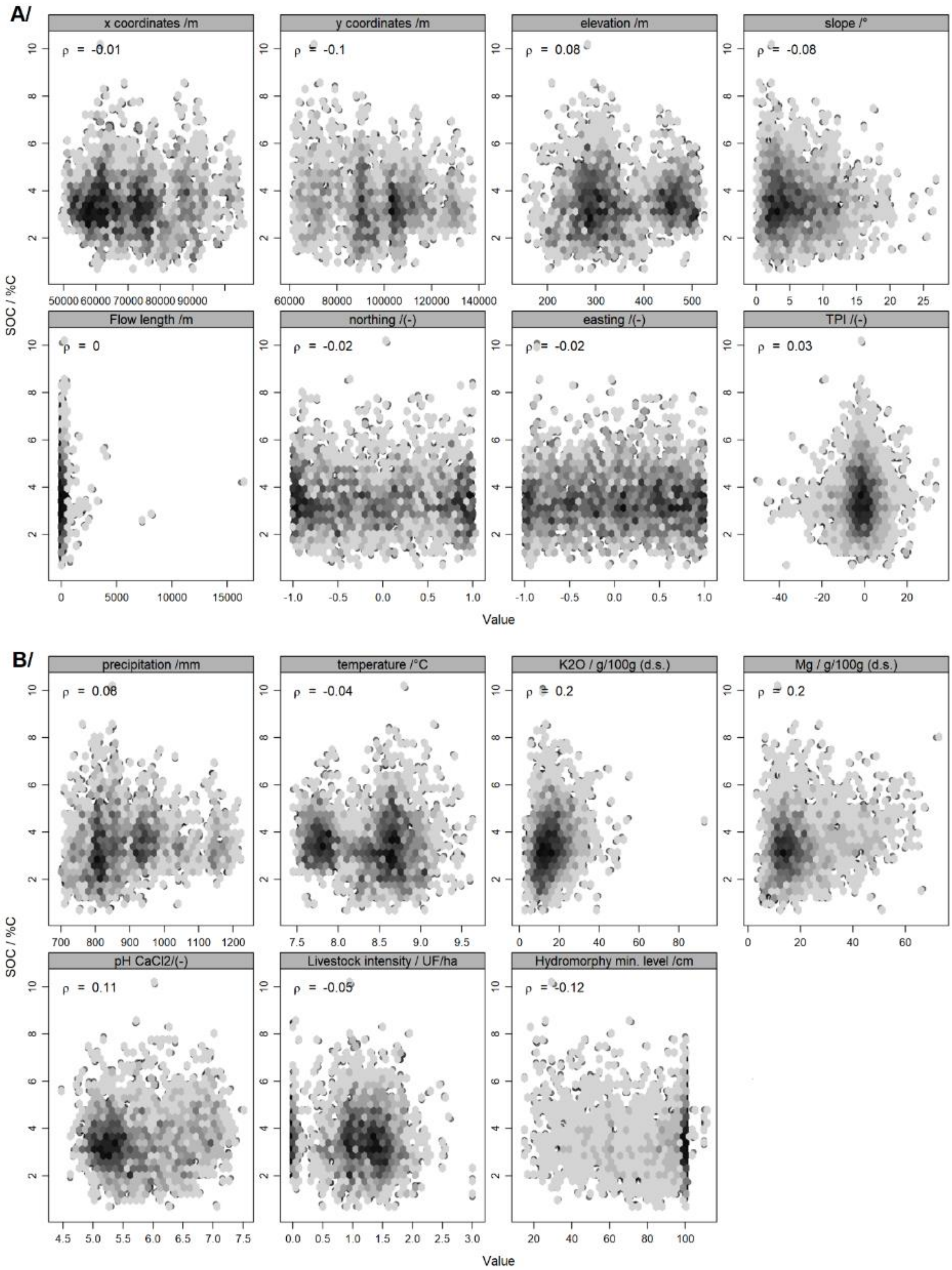

Figure 17: Scatter plots of topsoil SOC (%C) ***in grasslands*** *as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and landuse parameters.*

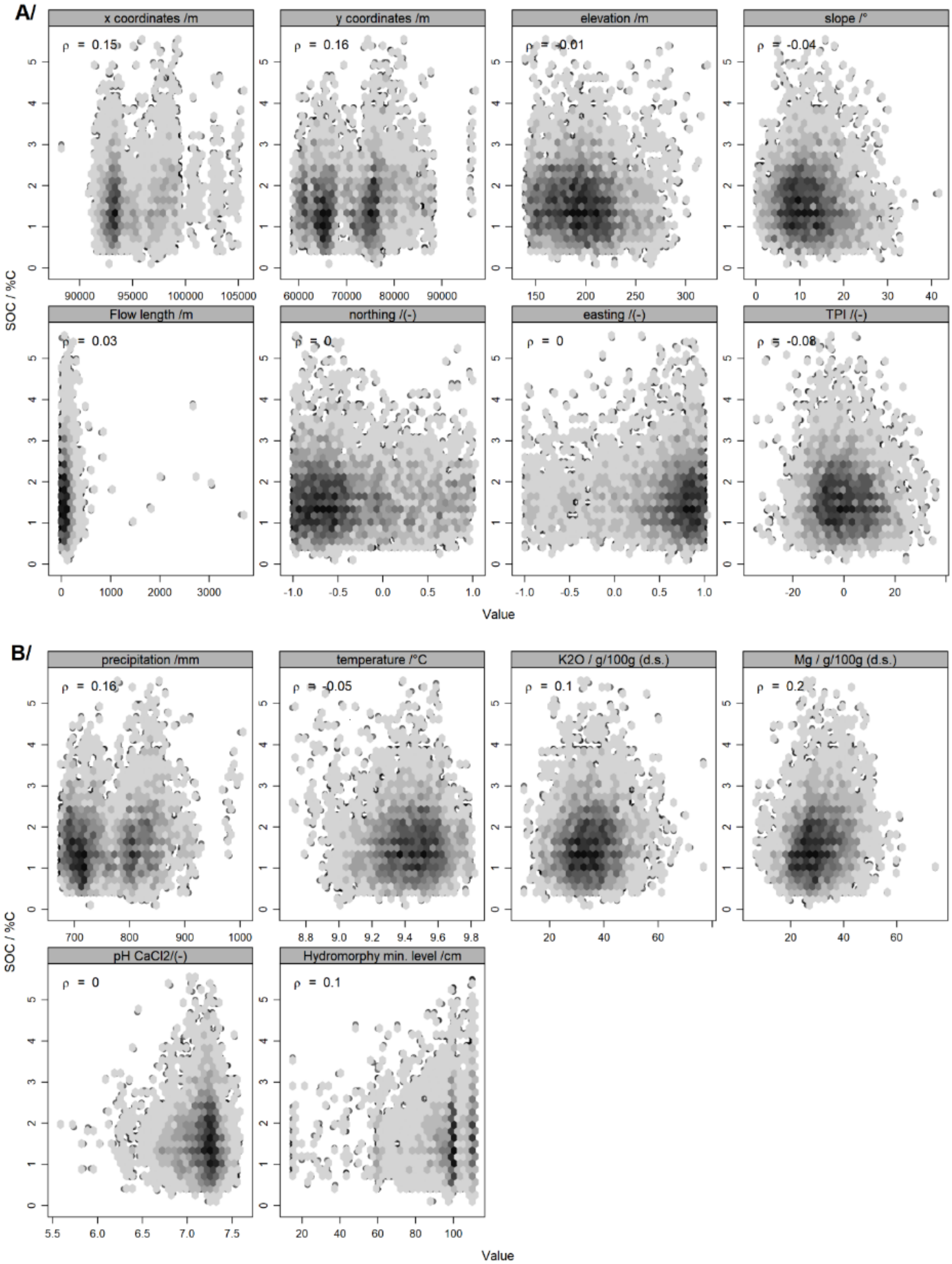

Figure 18: Scatter plots of topsoil SOC (%C) ***in vineyards*** *as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and landuse parameters. (Minimum level of soil hydromorphy was homogeneous all over the vineyard dataset)*

The semivariograms of the SOC observations show very different spatial structures between the different landuses (Figure 19). Cropland soils have a small nugget-to-sill ratio, indicating a high degree of spatial dependence, and the large range indicates that SOC in croplands is mostly determined by long-range factors (e.g. climate variables). The semivariogram for croplands is also possibly unbounded, which can be related to the presence of a trend in the data (probably due to the differences between the Oesling and Gutland regions). In grasslands, SOC content has also a low nugget-to-sill ratio but with a much smaller range (ca. 2 km). The spatial dependence of SOC in grasslands occurs at a much lower distance than croplands and can be due to the above-mentioned role of the clay content (which can vary on short distances). Vineyard soils show very little spatial structure suggesting that SOC almost varies randomly in space. Spatial variation of SOC in vineyards should be very diffcult to model.

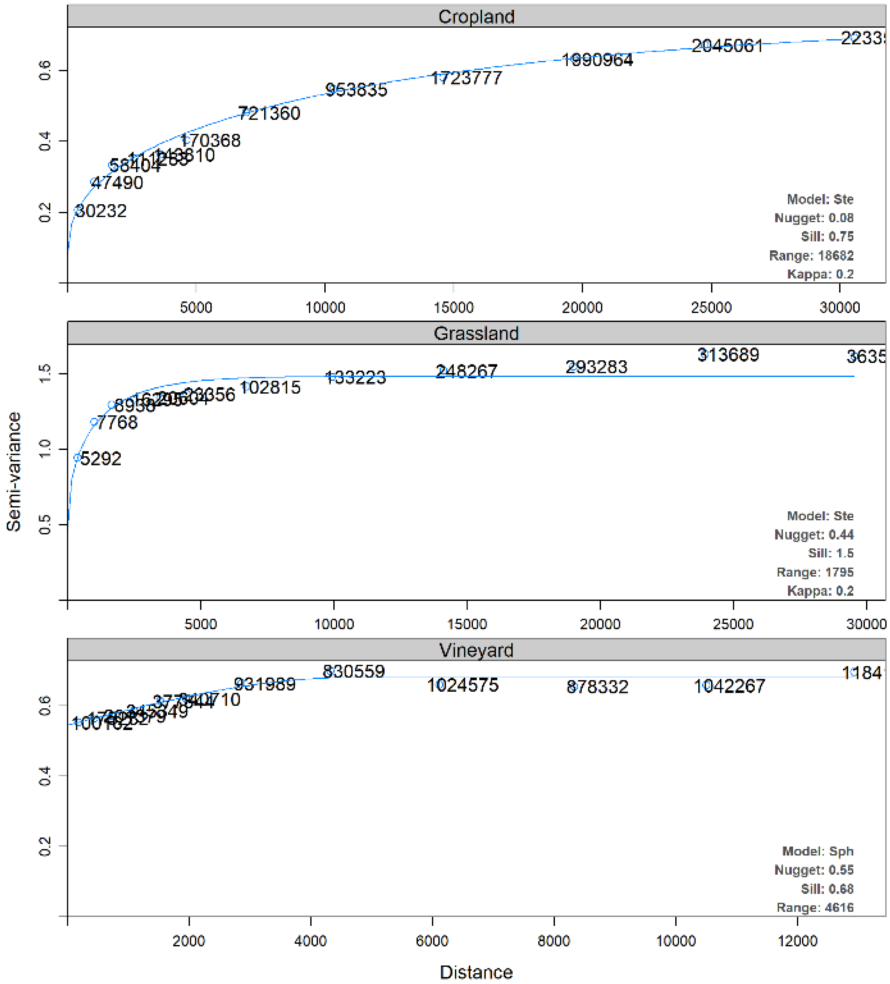

Figure 19: Semi-variograms of SOC under cropland, grassland and vineyard in Grand-Duchy of Luxembourg.

3.3.4 Model results and performance

The exploratory analysis showed some clear relations between topsoil SOC content and different natural and anthropogenic covariates, especially for croplands. For each landuse, a GAM model was calibrated on the whole subset containing T1 and T2 observations, and then the performance of the model was tested separately on each landuse-period subset. Before fitting, a pre-selection of covariates considering their relations, especially collinearity (concurvity is managed in the GAM procedure; section 3.2.3) was done. The performance of the final models supported the observations during the exploratory analysis about the relations between SOC and the different covariates in the different landuse classes. Figure 20 presents the point plots between observed and predicted data produced during the stratified 10-fold cross-validation for each landuse and period. Hence, ranked by descending order, the final models performed best for croplands, then grasslands and finally vineyards. According to the maps elaborated from the models fitted on T1 and T2, the national mean (and standard deviation) is of 2.25(0.74)%C for cropland, 3.57(0.76)%C for grassland and 1.74(0.31)%C for vineyards.

- **Cropland**

To model the spatial variation of topsoil SOC in cropland, the backward stepwise procedure selected, in addition to the geographical coordinate couple (x,y - supporting the main regional trends), the clay content, the C-factor, the Mg content, the K_2O content, the minimum depth of hydromorphy, the slope and the elevation, (here in decreasing order of importance; Fig. 21). The model explains 74% of the variance in the SOC content. Considering T1 and T2 separately, the model achieved a deviance explained of 77% and 73%, respectively. The remaining non-explained deviance could be related to factors not included in this study, especially management factors as crop rotation or good agricultural practices application. The predicted-observed point plots in Figures 20 showed that the model fitted well for both periods with R^2 of 0.70 and 0.66, RMSE of 0.52 %C for T1 and 0.55%C for T2. Although the predictions seemed unbiased, the observations with SOC > 4%C were underestimated for both period, even if croplands with such high SOC content are scarce.

- **Grassland**

The final GAM model fitted for grasslands explained 40% of the variance in the whole subset (T1+T2). The backward stepwise procedure selected, in addition to the geographical coordinate couple (x,y): the clay content, the Mg content, the minimum depth of hydromophy, the K_2O content , the elevation and the pH (by decreasing order of importance). The model showed poor results in validation procedure for both period with R^2 of 0.29 and 0.31, RMSE of 1.08 %C for T1 and 0.97 %C for T2 (Fig. 20). The results for T1 appeared a bit poorer than for T2 certainly due to the smaller number of observations (n=679 for T1, n=1452 for T2). The model tended to overestimate observations < 4 %C whereas, as for cropland, observations with SOC > 4%C were underestimated. The difference in RMSE between the T1 and T2 models can also be explained by the fact that the subset T1 contained much more observations > 6 %C than T2 creating a more pronounced bias induced by the underestimation of high SOC contents (i.e., > 4 %C) in this model.

Three possible explanations for the poor model fit in grassland are:

- i/ soils samples from grasslands contain more or less vegetal debris like roots that influence greatly the OC analysis inducing higher variability in SOC measurements than in cropland or vineyard soils;

- ii/ a non-negligible part of present grassland fields were converted from cropland after Second World War till the 1980s. Hence, soils of numerous grassland fields have probably not yet reached their SOC equilibrium phase, i.e. they did not yet reach their maximum SOC storing capacity;
- iii/ the vertical SOC gradient in the topsoil is much stronger in grasslands and a slight variation in sampling depth can have an effect in SOC content.

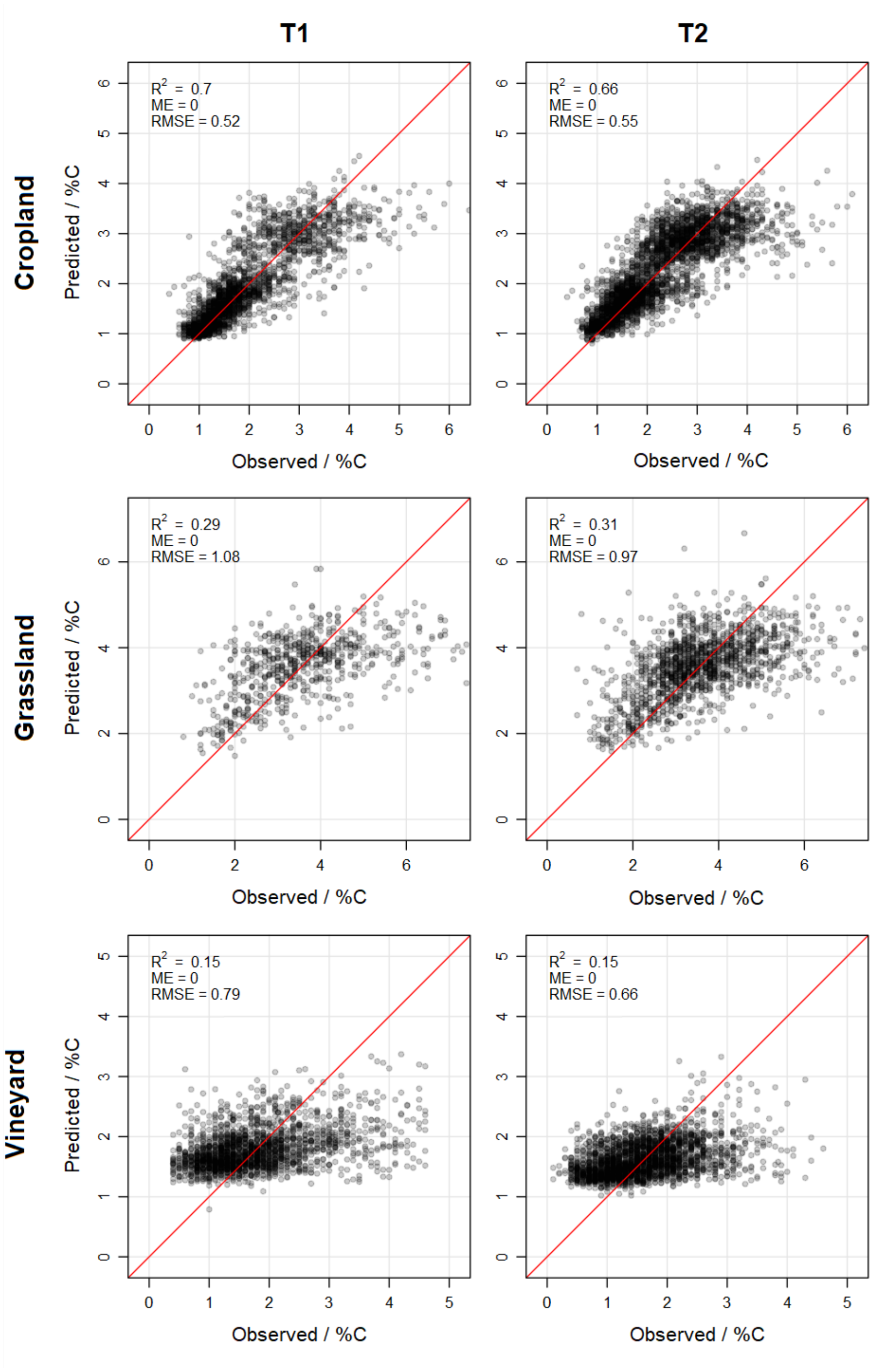

Figure 20: Observed vs predicted SOC (%C) as obtained by the models fitted for cropland, grassland, and vineyard soils at T1 (2012-2015) and T2 (2016-2019).

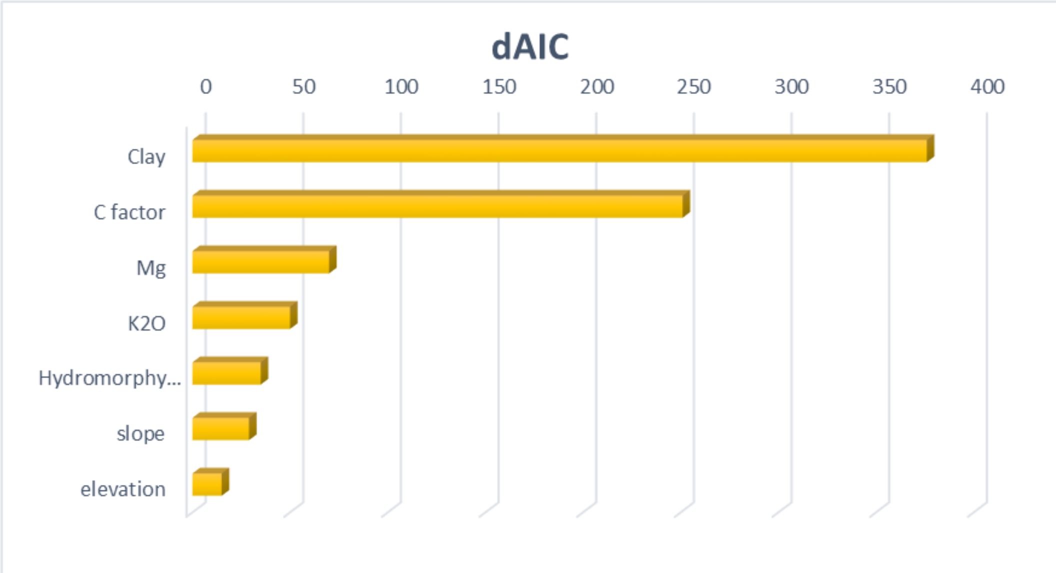

*Figure 21: Implication of covariates, complementary to the coordinates (x,y), in the GAM fitted on topsoil SOC (%C for the 0-25cm depth) in **croplands** of Grand-Duchy of Luxembourg according. dAIC represents the difference of AIC to the final model (Akaike Information Critrerion; Akaike, 1974). Only the covariates showing a p-value < 0.05 in the final GAM model were kept in this Figure.*

- **Vineyard**

The deviance explained by the model fitted for vineyards was very poor (14% on the whole subset). This poor fit was expected since the exploratory analysis demonstrated that the covariates were very poorly correlated with SOC observations. In the past, 84 % of the vineyards have been reallocated and undergone terracing. This emphasizes that vineyard soils have been disturbed so that topsoil SOC content and environmental covariates are not spatially linked anymore. Large variation of SOC on very small distance can be related to land reallocation and terracing operations that are very common in vineyards.

N.B.: In addition to GAM, ordinary kriging and regression kriging were tested as mapping technics for SOC in vineyards. Unfortunately, the results did not improve compared to the GAM procedure.

3.3.5 SOC maps - description and comparison

The models developed were applied to the layers of the selected covariates (90m x 90m resolution) to produce two maps of topsoil SOC contents under cropland, grassland and vineyard (one for each period T1 and T2). These maps are presented in Fig. 22 and 23 along with their respective map of standard error of prediction. Additional maps produced for croplands and grasslands separately are available in ANNEX 7.3.

Both SOC maps show the same general trends. Oesling and Gutland have very distinctive patterns. Oesling shows a smaller range of topsoil SOC content and a more homogeneous spatial variability than Gutland. The patterns in Oesling are mostly determined by the spatial location of the land uses of interest: croplands are mainly located on the plateaus (summits and shoulders of hillslopes) whereas grasslands are located on the shoulders and hillslopes. In addition, Oesling is mainly represented by only one soil association and textural soil type (OM) with relative homogenous clay content and the climatic conditions are quite homogeneous within this natural region. Topsoil SOC patterns in Gutland are mainly controlled by the clay content (the three textural soil types L, M and S can be clearly distinguished here; Fig. 3 and 5) and by land use repartition. The error of prediction map for T1 shows no clear regional trend (Fig. 22 – right part). The prediction errors are higher in areas with scarce observations for both SOC maps. Consequently, wider areas with higher prediction errors can be observed in southern and eastern parts of GDL for SOC_{T2} map (Fig. 23 – right part).

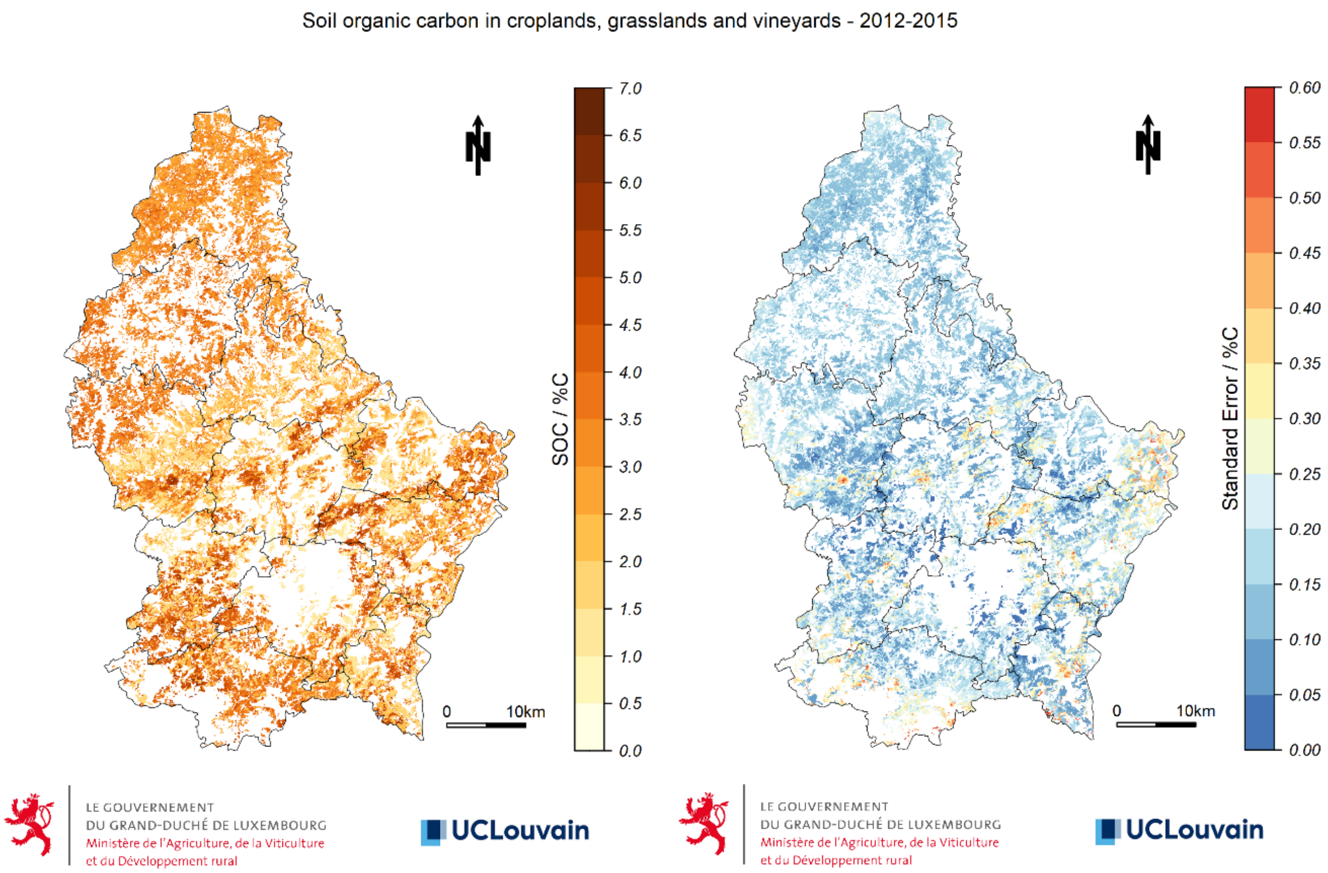

Figure 22: Maps of predicted SOC content (%C; on the left) and standard error of prediction (%C; on the right) for topsoil of Grand-Duchy of Luxembourg under Croplands, Grasslands and Vineyards for period T1 (2012-2015).

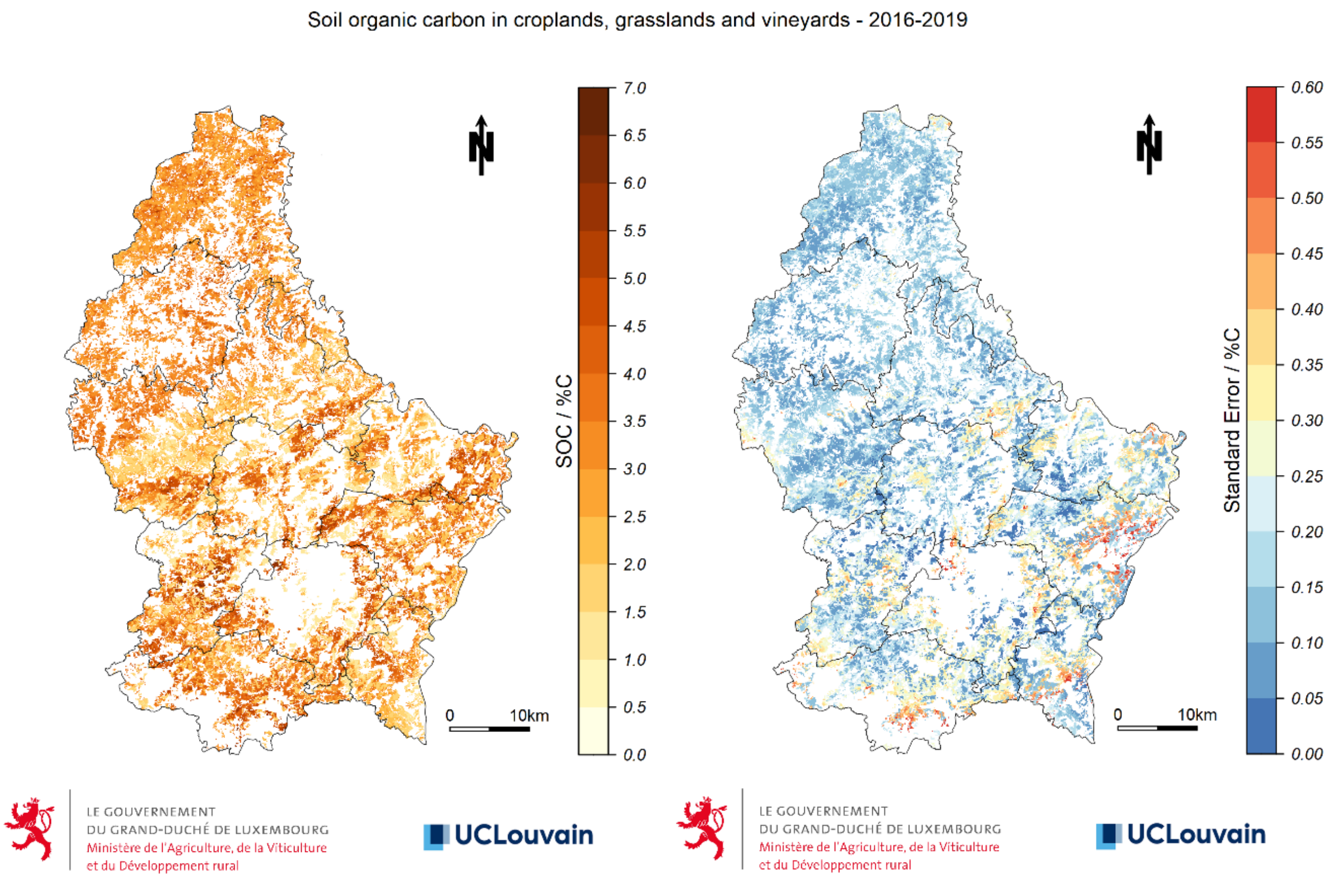

Figure 23: Maps of predicted SOC content (%C; on the left) and standard error of prediction (%C; on the right) for topsoil of Grand-Duchy of Luxembourg under Croplands, Grasslands and Vineyards for period T2 (2016-2019).

Figure 24 presents the significance of predicted SOC differences between T1 (2012-2015) and T2 (2016-2019) for cropland (left part) and grassland (right part). The p-values were estimated based on a comparison between predicted dSOC (i.e., $SOC_{T2} - SOC_{T1}$) and the standard errors (SE) of prediction for both SOC_{T1} and SOC_{T2} (depicted in right parts of Figures 22 and 23, respectively). For cropland and grassland, we estimated that ca. 40% of their respective areas had Non-Significant (NS) evolution of their topsoil SOC content. The predicted gain of SOC was estimated significant for ~25% of the cropland areas and for ~30% of the grassland areas (of which ~13% were significant at $p < 0.05$ for each landuse). Regarding the predicted loss of SOC, it was estimated significant for ~35% of the cropland areas and ~30% of the grassland areas (of which ~17% were significant at $p < 0.05$ for each landuse). However, it is important to note that, considering the weak goodness-of-fit of the models for grasslands, the estimations for grassland are less reliable than those for croplands.

The **cropland** areas subject to a significant predicted **loss** of SOC (Fig. 24 – left part) are mainly located in:

- The west and north-east parts of Oesling (Fig. 3A) where we observed a significant mean decrease of -0.99%C ($p<0.05$; Tab. 7);

- the easternmost part of Gutland which corresponds to the northern area of "Dolomies du Muschelkalk" soil association where we observed a mean decrease of -0.09%C (NS; Tab. 7);
- southernmost parts of Gutland corresponding mostly to the "Argiles du Lias inf. et moyen" and the "Argiles Lourdes des Schistes bitumineux" where we observed a decrease of -0.02%C and -0.13%C, respectively (NS; Tab.7).

The **cropland** areas subject to a significant predicted **gain** of SOC (Fig. 24 – left part) are located in:

- the southernmost part of "Oesling" (although the trend is negative at the entire region scale; Tab. 7);
- the northernmost part of Gutland, i.e in the "Buntsandstein" which showed a mean SOC increase of +0.12%C ($p<0.05$; Tab. 7);
- the center and eastern parts of Gutland, i.e. in the "Grès du Luxembourg", the "Dépôts limoneux sur Grès" and the "Argiles lourdes du Keuper" which had a respective mean SOC increase of +0.07%C, +0.14%C and +0.09%C ($p<0.05$; Tab. 7).

The **grassland** areas subject to a significant predicted **loss** of SOC (Fig. 24 –right part) are mainly concentrated in Gutland:

- in its northwestern most part which corresponds to the "Buntsandstein" soil association (fig. 3A) where we observed a mean non-significant decrease of -0.24%C (Tab. 8);
- in its southwestern part, mainly on soil associations "Argiles du Lias inf. et moyen" and "Argiles Lourdes des Schistes bitumineux" where we observed respective mean decrease of -0.23%C and – 0.68%C (NS; Tab. 8).

The **grassland** areas subject to a significant **gain** of SOC are also mostly concentrated in Gutland (Fig. 24 – right part):

- in most of the "Alluvions et Colluvions" located in valley bottoms where we observed a significant increase of 0.26%C ($p<0.05$; Tab. 8) and;
- locally on soils developed on "Grès du Luxembourg" and "Dépôts limoneux sur Grès" where we observed respective non-significant increase of +0.03%C and +0.05%C, (Tab. 8).

Grassland of Oesling tend to gain SOC in the northern part of this natural region, whereas they tend to loss SOC in the southern part. Overall, the mean absolute differences between predicted SOC at T1 and T2 are of -0.05%C for croplands and -0.02%C for grasslands (Fig. 25).

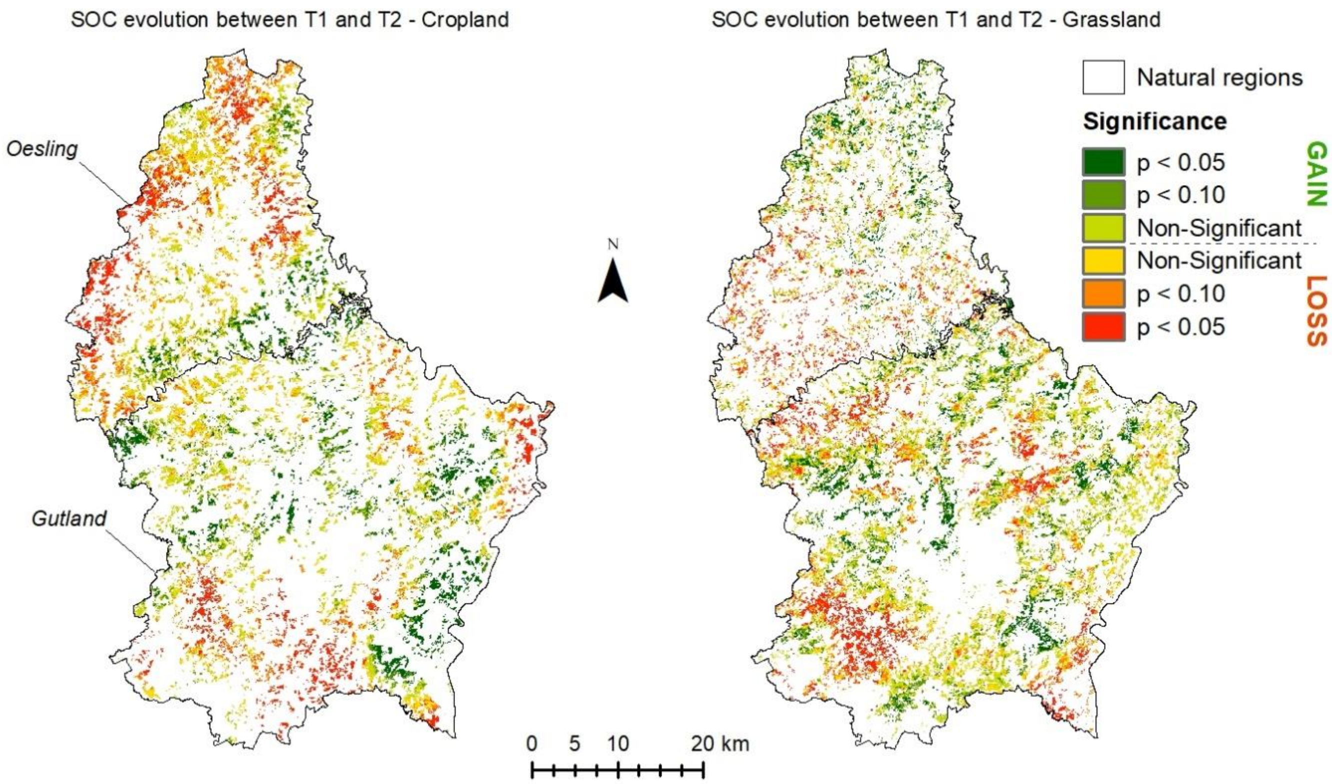

*Figure 24: **Significance of predicted SOC differences (p-value)** between T1 (2012-2015) and T2 (2016-2019) for soils under croplands (left) and grasslands (right).*

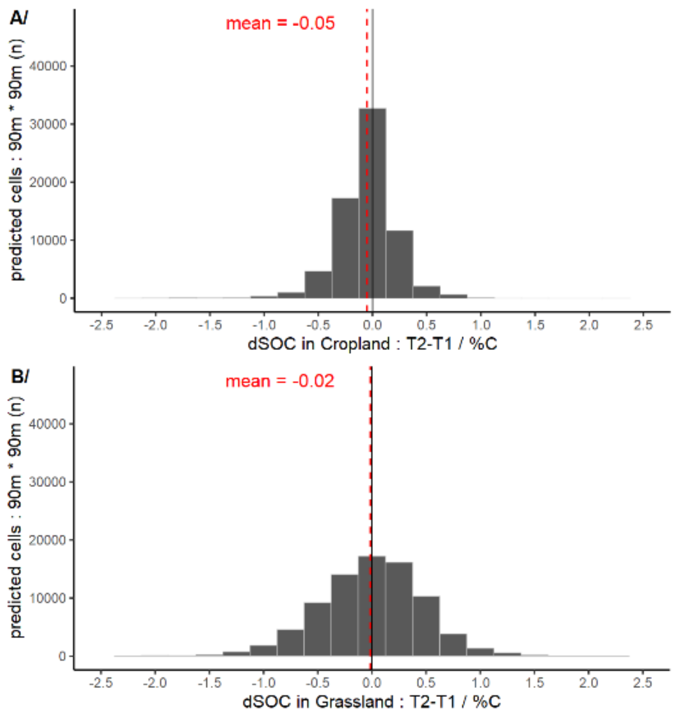

Figure 25: Histograms of the predicted difference in SOC (%C) between T1 (2012-2015) and T2 (2016-2019) for soils under cropland (top) and grasslands (bottom).

Considering the expertise of ASTA and the results of chapters 3 and 4, hypotheses are put forward on the processes and practices involved in the recent SOC trends. The recent temperature increase is likely to have had a positive impact on C mineralization. Also, these last years were characterized by drier summers that could have a negative impact on biomass production, and by more frequent extreme events enhancing topsoil erosion, especially in cropland. However, soils of wet areas (mainly occupied by grasslands) could have benefitted from a better productivity in these drier and warmer conditions, i.e. increased productivity in a warmer soil and less stress from asphyxiation. Finally, in the framework of Good Agricultural Practices (see chapter 4), changes in management practices could have induced more C inputs and/or less C outputs. Unfortunately, more data and additional research are needed to confirm or not these hypotheses, and to identify which of them is/are involved in the SOC dynamics depicted above.

4 Implication of Good Agricultural Practices (GAP) on short-term SOC dynamics

4.1 Context

In the context of the RDP (Rural Development Program) for the period 2014-2020, the European commission asked the Member States to monitor the effects of the environmentally-friendly farming techniques, as applied in the context of Agri-Environment Measures (AEM) for example, and develop indicators in order to highlight their benefits on different environmental compartments including soils. Here, we studied the effects of three Good Agricultural Practices (GAP) applied under cropland - cover crops (CC), reduced tillage (RT), and temporary grassland (TG) – on the evolution of topsoil SOC content at short-term (since 2012).

4.2 Methodology

4.2.1 Characteristics of the Good Agricultural Practices

Cover crops and reduced tillage both constitute the GAP introduced in the AEM 262-362-462. This AEM exists since 2000 and was designed to prevent soil erosion and to limit nitrate leaching from cropland soils. Farmers can apply cover crop or reduced tillage separately, or combined. Operational since 2015, the Greening Initiative supports the sustainable use of farmland through cover crop cultivation as GAP.

Cover crops are cultivated in order to protect the soils. They are ploughed in to increase soil organic matter and nutrients. In Northwestern Europe, they are usually cultivated 'off-season', sown after the harvest of the main crop (the commercial crop) in autumn and incorporated into the soil by plough or reduced tillage in early spring before seeding the next summer crop. Their cultivation helps preventing erosion and nitrate leaching, improving soil physical and biological properties, supplying nutrients to the following crop, improving soil water availability, and breaking pest cycles (Snapp et al., 2005).

Since the end of WWII, increased mechanization and intensive tillage, leading to an increase in erosion, have greatly degraded many agricultural soils (Lal, 1993; van Oost et la., 2005). Reduced tillage aims to reduce intensity of tillage operations, and may progress to stopping tillage completely (no-tillage). These practices result in some environmental benefits (e.g., reducing erosion, improving soil water availability, avoiding soil compaction), but also economic ones (as reducing fuel and labor costs) (Busari et al., 2015; Jacobsen and Ørum, 2010).

The third GAP is the application of temporary grassland. The EC (European Commission) makes a clear distinction between temporary grassland (fields under grassland less or equal than five consecutive crop years without ploughing) categorized under 'arable land' (i.e., cropland here), and permanent pasture (fields under grassland more than five consecutive crop years without tillage) categorized under grassland[10]. The temporary grasslands induce positive residual effects on the following arable crops, increasing soil fertility and reducing crop diseases and weed infestation (Panattieri et al., 2017; Viaud et al., 2018).

---

[10] *https://ec.europa.eu/eurostat/documents/2393397/8259002/Grassland_2014_Task+1.pdf/8b27c17b-b250-4692-9a58-f38a2ed59edb*

4.2.2 Database preparation

We used SOC data extracted from the ASTA database for the period 2012-2020 (as in section 3) merged to the Land Field Information System for the period 2008-2019[11]. Firstly, the data were filtered and cleaned, following the steps below:

1. Keeping the SOC observations obtained by the device Jena EA 4000 and the Skalar Primacs SNC100 only;
2. Removing the observations from 2016 from soils sampled by the operator 'LAKU' (the operator sampled at 0-30 cm depth instead of 0-25 cm in cropland);
3. Removing the observations for which the FLIK polygons were not available or the FLIK was miscoded;
4. Removing observations without date of sampling (mandatory to assess the cultural year);
5. Removing all observations related to fields submitted to a recent land use change (within the 5 crop years preceding the sampling) and related to a landuse of no interest[12];
6. Removing the duplicates by FLIK and crop year, and replacing them by their mean SOC value;
7. Removing those without GAP for more than 3 consecutive crop years (before the sampling year);
8. By land use and soil association, removing right-skewed data[13] (filtering greatest outliers).

The crop year of each observation was determined considering the date of sampling or entry at the laboratory: e.g. soils sampled between July XXXX and June XXXX+1 were related to the cultural year XXXX+1. As the information about what was grown in each field (as main crop) for cultural years 2008-2019 was collected, the fields concerned by a period of temporary grassland were already identified. Then, we merged the data with the spatial layers identifying fields where the AEM 362-462 (during cultural years 2008-2019 also) and the Greening Initiative (declared in 2015-2019[14]) were applied. These layers allowed us to compile the exact GAPs applied each cultural year in each field related to the SOC observations. Each observation was classified in a GAP or GAP combination by considering those applied from 2008 to the crop year corresponding to the sampling – even if no GAP was applied the year of sampling, the field was classified in the GAP category if a GAP was applied at least once from 2008 to less than 3 years before the crop year of sampling. Fields not submitted to any GAP between 2008 and the year of soil sampling were defined as 'Control' fields. The final dataset was called LU-SOC-GAP.

[11] *For thorough information about the methods of soil sampling, Corg analysis or merge between Corg data and LPIS, please refer to §3.2.1.*

[12] *To this aim, the cultural history of the FLIKs from cultural year 2008 has been reconstituted until 2019 to consider FLIK number changes over time. The methodology is, at the time of this reporting, still under improvement.*

[13] *All data superior to Q3 + 3*SE (with Q3 = 3rd quartile and SE = standard error).*

[14] *A cover crop sown in the context of the Greening Initiative during the calendar year XXXX is declared the same year. So, a cover crop from Greening initiative declared in XXXX is associated with cultural year XXXX+1.*

4.2.3 Analysis of SOC differences between management practices

- **LU-SOC-GAP**

For each soil association, the significance of differences between the distributions of SOC in fields under Control conditions and in fields under GAP (undifferentiated) were tested with a non-parametric Wilcoxson test (non-paired). Tests were also applied to fields submitted to a single GAP (i.e. cover crops - CC, reduced tillage – RT or temporary grassland – TG) against fields under Control conditions.

- **Paired observations**

We identified paired observations in the LU-SOC-GAP database, i.e. FLIKs with two or more SOC observations between 2012 and 2020. Then, we studied the relative SOC differences between paired observations considering the number of years separating them (relative annual difference in %/yr), and the application or not of GAP between both observation, and before the first observation.

4.2.4 Conditional inference trees

In order to assess the relative importance of farming practices and environmental covariates on SOC variability, we produced conditional inference trees. Each model ('a forest') was created based on 500 trees using the party package in R (Strobl et al., 2007). They are similar to a random forest and can be used to model non-linear interactions between the response variable (i.e. the SOC) and predictor variables without the requirements of normality and homoscedasticity (Hobley et al., 2016). Considering the analysis of relations between SOC and environmental covariates performed in section 3.3.3, we introduced the same covariates selected for introduction in the GAM models as predictor variables, i.e. elevation, slope, northness, eastness, precipitation, clay content, pH, available Mg, available K_2O, the C factor and the minimum depth of hydromorphy. To consider the impact of GAP on SOC variability, we added two specific variables:

- **GAP_app** informs about the type of GAP or combination of GAPs applied on the sites from 2008 to the crop year of sampling, i.e. Control, Cover Crops, reduced Tillage, Temporary Grassland and all their possible combinations.
- **GAP_app_years** informs about the number of years of GAP application (this variable is potentially biased as our database considered management practices form 2008 only).

Finally, the covariate **CROP_yr** was added as the crop year of sampling which could inform about the influence of weather, drought… of the concerned year.

The conditional inference forest was grown over 500 trees with the number of predictor variables randomly selected per split set to ~ square root of p (p being the number of covariates) and a significance relationship between predictor and response variable at $\alpha < 0.05$. The relative variable importance was expressed as $n=I/T*100$, where I is the covariate importance and T is the total variance explained by the model (Hobley et al., 2015). A full model was first fitted on the LU-SOC-GAP dataset and then sequentially the least important covariate was skipped until having the best goodness-of-fit. The overall performance of the models was evaluated on the RMSE and R^2 of the out-of-bag dataset (as a cross-validation).

4.3 RESULTS

4.3.1 Implications of the SOC data filtering and merging

After cleaning and filtering of the extracted raw data, the LU-SOC-GAP dataset contained 4016 observations, including 960 associated to Control fields and 3056 to fields under GAP (Table 11). Not considering the observations obtained with the Tru Spec CN analyzer induced a loss of 1101 observations. The elimination of fields submitted to recent landuse change or being under landuse of no interest here (i.e., not under cropland strictly speaking) induced a loss of more than 7850 observations (~5770 were vineyards). Replacing the duplicates by cultural year by their mean values led to diminish the set of 1082 observations.

Table 11: Filtering steps on the LU-SOC-GAP database preparation and associated numbers of observations

Filtering step	Total Obs.	Obs. eliminated
None	**14972**	-
- Tru Spec CN analyzer	13871	1101
- LAKU 2016 cropland	13772	99
- FLIK NA	13390	382
- Date NA	13344	46
- obs from 2020 (potential TG)*	13211	133
- LU of no interest	5352	7859
- duplicates by crop year	4270	1082
- FLIK not in RPG	4270	0
- potential miscoded GAP	4059	211
- Soil association NA	4051	8
- outliers	**4016**	35

4.3.2 Impact of GAP on SOC

- **LU-SOC-GAP**

Observations related to Control conditions and to GAP application in the LU-SOC-GAP dataset showed very different distributions (Fig. 26). ‘Control’ observations had a unimodal distribution with mode around 1.75%C, while ‘GAP’ observations had a bimodal distribution with a first mode around 1.70%C and a second around 2.80%C. Those distributions are mainly explained by the spatial distribution of the observations. The Control subset is dominated by observations from Gutland while the GAP subset is more evenly distributed between the Gutland and the Oesling. The latter is characterized by soils of higher SOC content compared to Gutland (Figs. 27-28; see also section 3.3.2.).

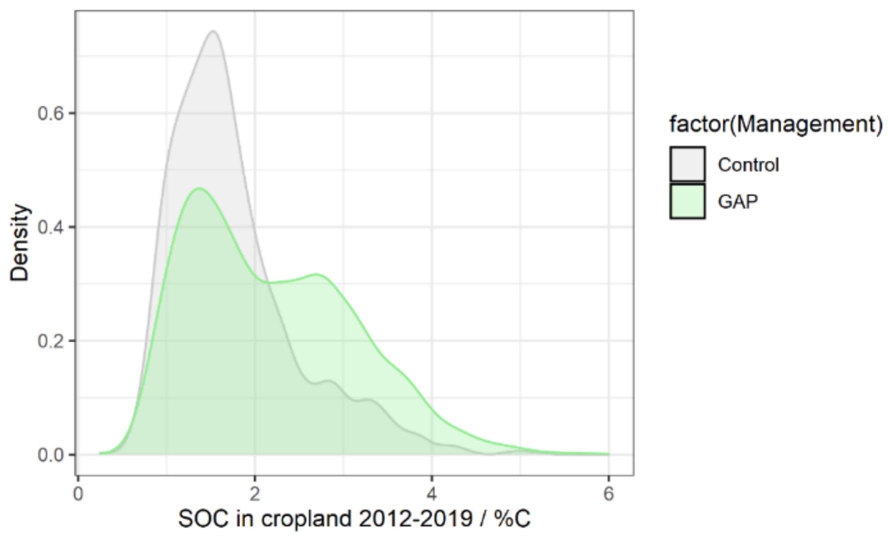

Figure 26: Histograms of topsoil SOC (%C) in croplands for fields under Control conditions and fields under GAP in Grand-Duchy of Luxembourg (2012-2019).

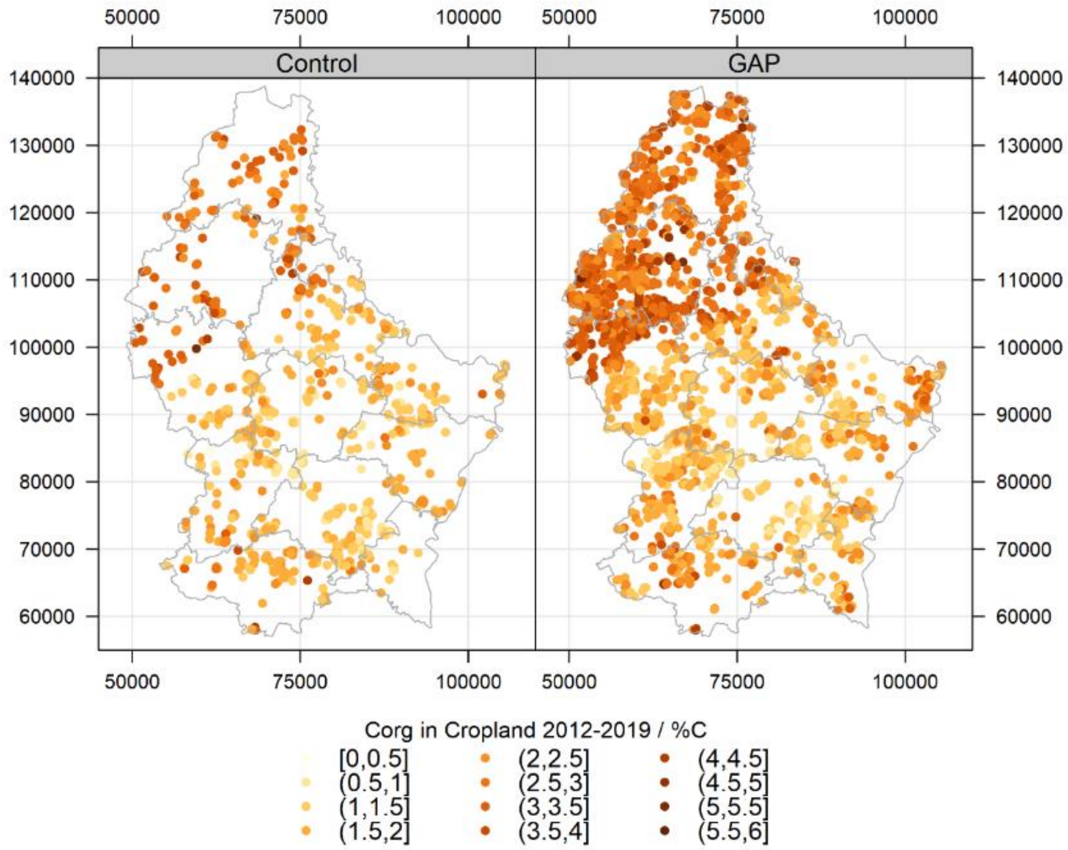

Figure 27: Observed SOC values (%C) of topsoil in croplands for fields under Control conditions and fields under GAP in Grand-Duchy of Luxembourg (2012-2019).

Considering the soil associations separately, GAP observations generally outnumbered Control observations (especially in 'Oesling'), and we observed a positive mean difference in SOC when comparing 'GAP' observations to 'Control' observations (Fig. 28; Table 12). However, this difference is statistically significant for only two soil associations: 'Oesling' (+0.16%C; $p<0.05$), 'Dolomies du Muschelkalk' (+0.29%C; $p<0.05$). The difference was also significant for 'Argiles lourdes des schistes bitumineux' (+0.65%C; $p<0.001$) but, **as the number of Control observations was < 30 (n=14), this result should not be considered as relevant**. As each GAP or combination of GAP are susceptible to impact soils differently, we will now consider them separately.

N.B.: In this section 4.3.2., the differences of statistical distribution between different datasets were tested (Table 12 to 15). The analytical uncertainties were not considered (see 3.2.1).

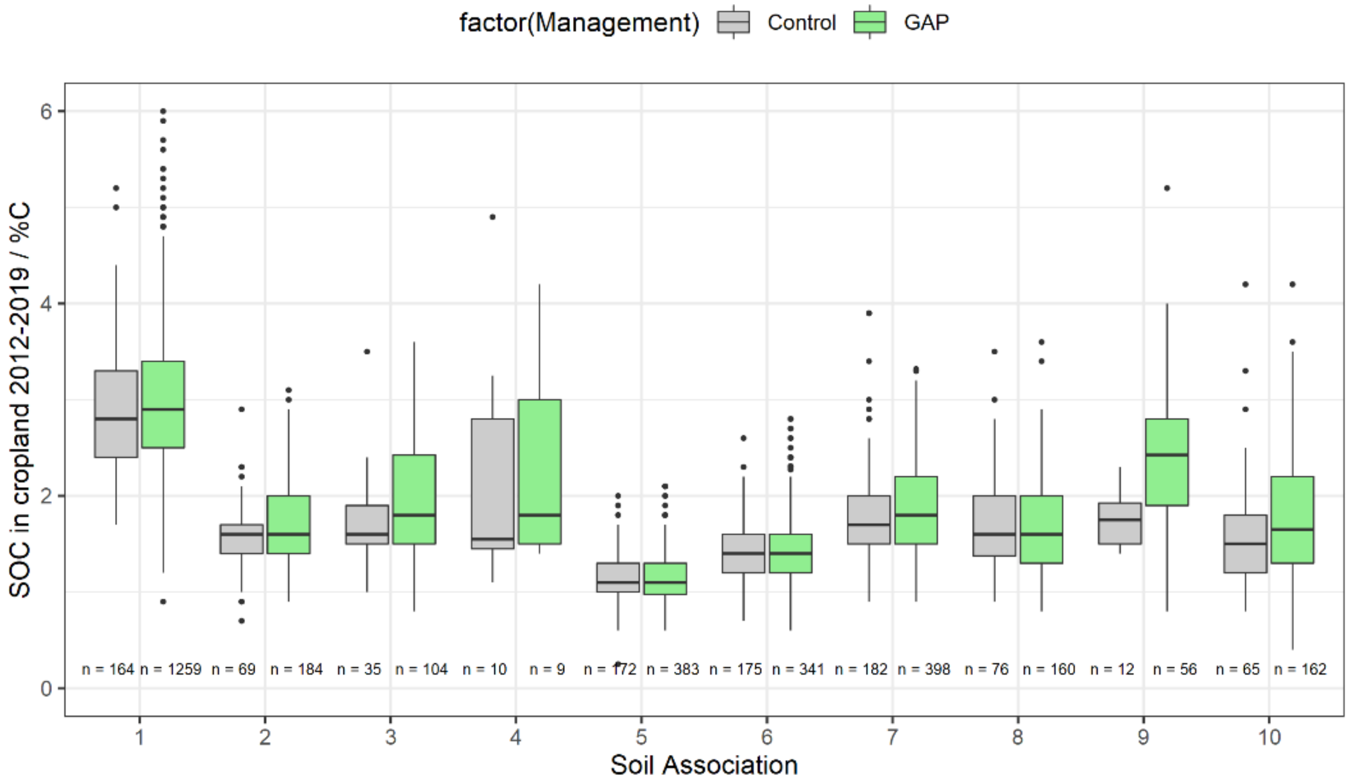

*Figure 28: Box-plots of topsoil SOC (%C) in **croplands** for fields under Control conditions and fields under **GAP** (2012-2019). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

*Table 12: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** for fields under **Control conditions** and fields under **GAP** (2012-2020)), and significance of the difference between these two types of management (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	Control							Good Agricultural Practices							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	960	0.3	1.2	1.6	1.8	2.0	5.2	3056	0.4	1.4	2.1	2.2	2.8	6.0		
1	164	1.70	2.40	2.80	2.86	3.30	5.20	1259	0.90	2.50	2.90	3.02	3.40	6.00	**0.16**	**< 0.01**
2	69	0.70	1.40	1.60	1.58	1.70	2.90	184	0.90	1.40	1.60	1.71	2.00	3.10	0.14	NS
3	35	1.00	1.50	1.60	1.74	1.90	3.50	104	0.80	1.50	1.80	2.03	2.43	3.60	**0.29**	**< 0.05**
4	10	1.10	1.45	1.55	2.16	2.80	4.90	9	1.40	1.50	1.80	2.44	3.00	4.20	0.29	NS
5	172	0.25	1.00	1.10	1.15	1.30	2.00	383	0.60	0.98	1.10	1.15	1.30	2.10	0	NS
6	175	0.70	1.20	1.40	1.39	1.60	2.60	341	0.60	1.20	1.40	1.43	1.60	2.80	0.04	NS
7	182	0.90	1.50	1.70	1.79	2.00	3.90	398	0.90	1.50	1.80	1.85	2.20	3.32	0.05	NS
8	76	0.90	1.38	1.60	1.70	2.00	3.50	160	0.80	1.30	1.60	1.69	2.00	3.60	-0.01	NS
9	12	1.40	1.50	1.75	1.77	1.93	2.30	56	0.80	1.90	2.43	2.41	2.80	5.20	**0.65**	< 0.01
10	65	0.80	1.20	1.50	1.61	1.80	4.20	162	0.40	1.30	1.65	1.80	2.20	4.20	**0.18**	**NS**

All GAPs (cover crops - CC, reduced tillage – RT or temporary grassland – TG) and possible combination of GAPs (CC RT, RT TG, CC TG and CC RT TG) were found in all soil associations except in 'Calcaires du Bajocien' (the latter has only 19 observations; Fig. 29). Most soil association subsets were composed of around 25-35% of Control observations, except Oesling with only ca. 12%. More than 50% of GAP observations in Oesling were related to fields where temporary grassland (TG, RT TG, CC TG and CC RT TG) has been or is currently applied. Soil associations from Gutland showed subsets dominated by observations related to cover crops and/or reduced tillage strategies application (CC, RT, CC RT), representing ca. 50% of each subset.

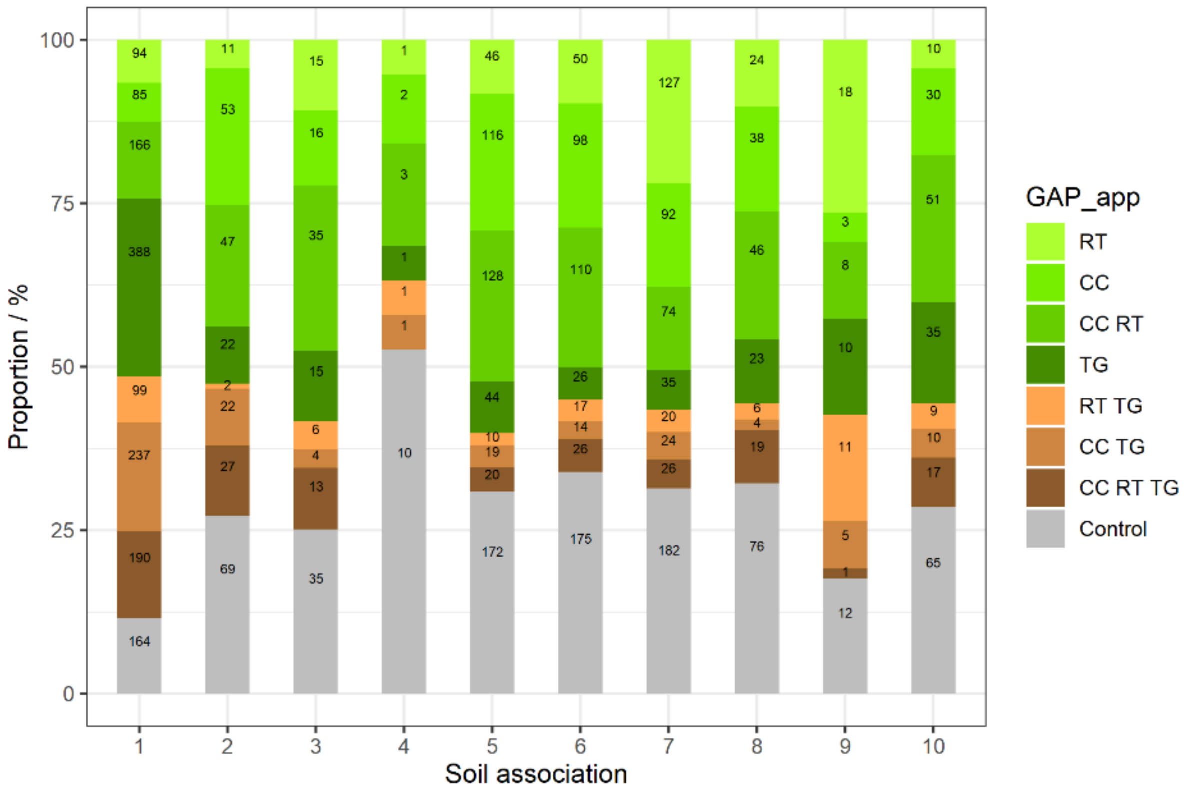

*Figure 29: Bar-plots of the proportion of observations in LU-SOC-GAP from fields under **Control conditions** and fields under **GAP or combination of GAPs** (2012-2020). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

It appears difficult to observe and confirm major trends in GAPs by soil association considering the number of box-plots and the fact that many subcategories have less than 30 individuals (Fig. 30). However, we can observe in soil associations the most represented in the dataset (1 – Oesling, 5 - Grès de Luxembourg, 6 - Dépôts limoneux sur Grès and 7 - Argiles du Lias inf. et moyen) that fields concerned by TG and/or combinations including TG could have higher SOC contents than fields under Control or CC and/or RT conditions.

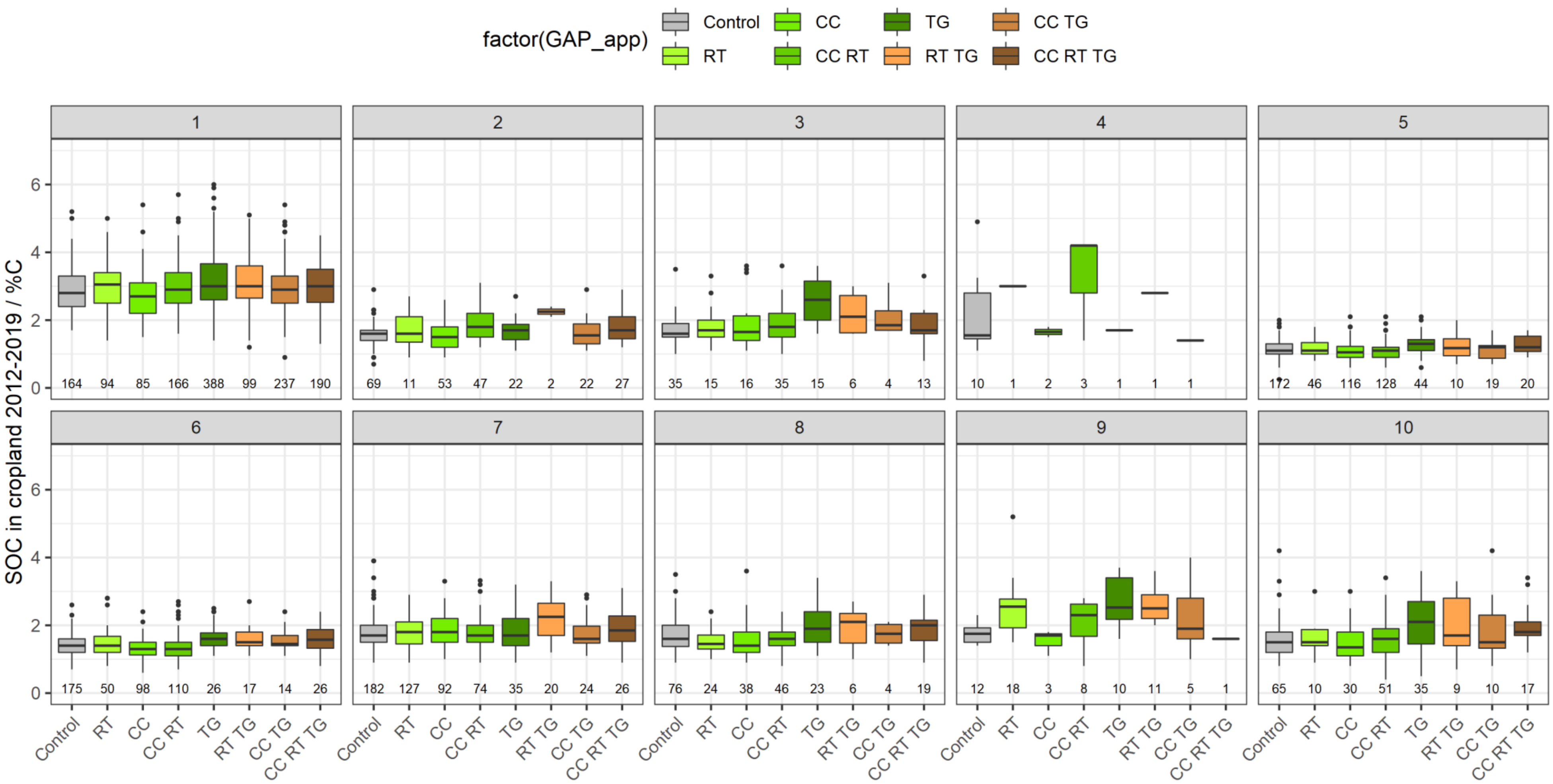

*Figure 30: Box-plots of topsoil SOC (%C) in **croplands** for fields under **Control conditions** and fields under **GAP or combination of GAPs** (2012-2020). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

Table 13 to 15 present summary statistics and results of statistical tests between observations under Control vs Reduced Tillage, Cover Crops and Temporary Grassland. In Table 13, only four out of ten soil associations (1 – Oesling, 5 - Grès de Luxembourg, 6 - Dépôts limoneux sur Grès, 7 - Argiles du Lias inf. et moyen) have more than 30 observations in each subcategory, i.e. Control and Reduced Tillage. The differences in SOC between Reduced Tillage and Control observations varied from -0.01%C to +0.15%C considering these 4 soil associations. None of their SOC differences was significant (p>0.05). In Table 14, seven out of the ten soil associations (1 – Oesling, 2 - Buntsandstein, 5 - Grès de Luxembourg, 6 - Dépôts limoneux sur Grès, 7 - Argiles du Lias inf. et moyen, 8 - Argiles lourds du Keuper, 10 - Others) present more than 30 individuals in each subcategory, i.e. Control and Cover Crops. The differences in SOC between CC and Control varies from -0.13% to +0.10%C for these 7 soil associations, most of those differences being negative. None of these differences were significant. Considering the GAP Temporary Grassland, four out of ten soil associations (1 - Oesling, 5 - Grès de Luxembourg, 7 - Argiles du Lias inf. et moyen, 10 - Others) have more than 30 observations in each subcategory. The differences in SOC between TG and Control ranged between +0.04% and +0.53%C and were all significant (p<0.05) except for 'Argiles du Lias inf. et moyen' (+0.04%C; p>0.05).

*Table 13: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** for fieldss under **Control conditions** and fields submitted to **reduced tillage strategies** (2012-2020), and significance of the difference between these two types of management (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	Control							Reduced Tillage							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	960	0.3	1.2	1.6	1.8	2.0	5.2	396	0.8	1.4	1.8	2.0	2.5	5.2		
1	164	1.70	2.40	2.80	2.86	3.30	5.20	94	1.40	2.50	3.05	3.01	3.40	5.00	0.15	NS
2	69	0.70	1.40	1.60	1.58	1.70	2.90	11	0.90	1.35	1.60	1.71	2.10	2.70	0.13	NS
3	35	1.00	1.50	1.60	1.74	1.90	3.50	15	1.10	1.50	1.70	1.87	2.00	3.30	0.13	NS
4	10	1.10	1.45	1.55	2.16	2.80	4.90	1	3.00	3.00	3.00	3.00	3.00	3.00	0.84	NS
5	172	0.25	1.00	1.10	1.15	1.30	2.00	46	0.80	1.00	1.10	1.14	1.34	1.80	-0.01	NS
6	175	0.70	1.20	1.40	1.39	1.60	2.60	50	0.80	1.20	1.40	1.46	1.68	2.80	0.07	NS
7	182	0.90	1.50	1.70	1.79	2.00	3.90	127	0.90	1.45	1.80	1.79	2.10	2.90	0	NS
8	76	0.90	1.38	1.60	1.70	2.00	3.50	24	1.00	1.30	1.45	1.55	1.71	2.40	-0.15	NS
9	12	1.40	1.50	1.75	1.77	1.93	2.30	18	1.50	1.93	2.55	2.52	2.78	5.20	**0.76**	**< 0.01**
10	65	0.80	1.20	1.50	1.61	1.80	4.20	10	0.90	1.40	1.50	1.63	1.88	3.00	0.02	NS

*Table 14: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** for fields under **Control conditions** and fields submitted to **cover crops cultivation** (2012-2020), and significance of the difference between these two types of management (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	Control							Cover Crops							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	960	0.3	1.2	1.6	1.8	2.0	5.2	533	0.6	1.2	1.5	1.7	2.0	5.4		
1	164	1.70	2.40	2.80	2.86	3.30	5.20	85	1.50	2.20	2.70	2.71	3.10	5.40	-0.15	NS
2	69	0.70	1.40	1.60	1.58	1.70	2.90	53	0.90	1.20	1.50	1.55	1.80	2.60	-0.02	NS
3	35	1.00	1.50	1.60	1.74	1.90	3.50	16	1.10	1.40	1.65	1.94	2.13	3.60	0.21	NS
4	10	1.10	1.45	1.55	2.16	2.80	4.90	2	1.50	1.58	1.65	1.65	1.73	1.80	-0.51	NS
5	172	0.25	1.00	1.10	1.15	1.30	2.00	116	0.60	0.90	1.05	1.10	1.23	2.10	-0.05	NS
6	175	0.70	1.20	1.40	1.39	1.60	2.60	98	0.60	1.13	1.30	1.34	1.50	2.40	-0.05	NS
7	182	0.90	1.50	1.70	1.79	2.00	3.90	92	1.00	1.50	1.80	1.90	2.20	3.30	0.1	NS
8	76	0.90	1.38	1.60	1.70	2.00	3.50	38	0.90	1.20	1.40	1.59	1.80	3.60	-0.1	NS
9	12	1.40	1.50	1.75	1.77	1.93	2.30	3	1.10	1.40	1.70	1.53	1.75	1.80	-0.23	NS
10	65	0.80	1.20	1.50	1.61	1.80	4.20	30	0.80	1.10	1.35	1.48	1.80	3.00	-0.13	NS

*Table 15: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** for fields under **Control conditions** and fields submitted to **temporary grassland** (2012-2020), and significance of the difference between these two types of management (non-paired Mann-Whitney test). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourds du Keuper, 9 = Argiles lourds des schistes bitumineux, 10 = Others)*

	Control							Temporary Grassland							Difference	
Assoc.	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	960	0.3	1.2	1.6	1.8	2.0	5.2	599	0.5	2.0	2.7	2.7	3.3	6.0		
1	164	1.70	2.40	2.80	2.86	3.30	5.20	388	1.40	2.60	3.00	3.14	3.66	6.00	0.28	**< 0.001**
2	69	0.70	1.40	1.60	1.58	1.70	2.90	22	1.10	1.43	1.70	1.69	1.88	2.70	0.11	NS
3	35	1.00	1.50	1.60	1.74	1.90	3.50	15	1.60	2.00	2.60	2.59	3.15	3.60	0.86	< 0.01
4	10	1.10	1.45	1.55	2.16	2.80	4.90	1	1.70	1.70	1.70	1.70	1.70	1.70	-0.46	NS
5	172	0.25	1.00	1.10	1.15	1.30	2.00	44	0.60	1.10	1.30	1.30	1.43	2.10	0.16	**< 0.01**
6	175	0.70	1.20	1.40	1.39	1.60	2.60	26	1.10	1.40	1.60	1.64	1.78	2.50	0.26	< 0.02
7	182	0.90	1.50	1.70	1.79	2.00	3.90	35	0.90	1.40	1.70	1.84	2.20	3.20	0.04	NS
8	76	0.90	1.38	1.60	1.70	2.00	3.50	23	1.10	1.50	1.90	1.93	2.40	3.40	0.23	NS
9	12	1.40	1.50	1.75	1.77	1.93	2.30	10	1.60	2.18	2.53	2.68	3.40	3.70	0.91	< 0.01
10	65	0.80	1.20	1.50	1.61	1.80	4.20	35	0.50	1.45	2.10	2.15	2.70	3.60	0.53	**< 0.01**

According to these preliminary results, Temporary Grassland seems to be the most effective GAP for increasing SOC in cropland soils from GDL. Reduced Tillage tends to help maintaining or slightly increasing original SOC content, while Cover Crops barely maintain SOC content when compared to Control fields. This latter fact is counter-intuitive as CC application enhances organic matter inputs in soils, and CC is recognized as one of the most effective GAP for improving SOC content and stock in cropland soils (Pellerin et al., 2019.). However, CC are mainly applied in GDL right before silage maize cultivation which is known as being a powerful humus consumer – the removal of straw/stover inducing a net reduction of the topsoil SOC stock (Xu et al., 2019). Here, 85.5% of the FLIK concerned by CC application only in the LU-SOC-GAP dataset are cultivated with maize silage at least once during their rotation. For comparison, 67.5% of the FLIK under Control condition, 54.5% of the FLIK with RT application, and 44.1% of the fields with TG are concerned by maize silage cultivation. Considering this information, the application of CC may appear as an effective way to counter-balance the negative effect of silage maize cultivation on SOC.

- **Paired observations**

Amongst the 500 fields concerned, from the first to the second observations (separated by min = 1yr, max = 6yrs, median = 4yrs), 53 remained under Control conditions, 86 went from Control to GAP, and 361 remained under GAP (Fig. 31). The 1st quartile of relative annual difference in SOC between the first and second sampling was -2.9%, the median 0.0% and the 3rd quartile 4.6%. Considering the enlarged analytical uncertainties of 15% for non-carbonated samples and 20-25% for carbonated sample (Table 4), plus the error propagation induced by the difference computation, **those results cannot be considered as significant and have to be considered very carefully. For information purposes only, we computed the relative annual difference in SOC and compiled the results in Table 16.**

Figure 31: Location and relative annual differences in SOC content between paired observations (same FLIKs) in the LU-SOC-GAP dataset.

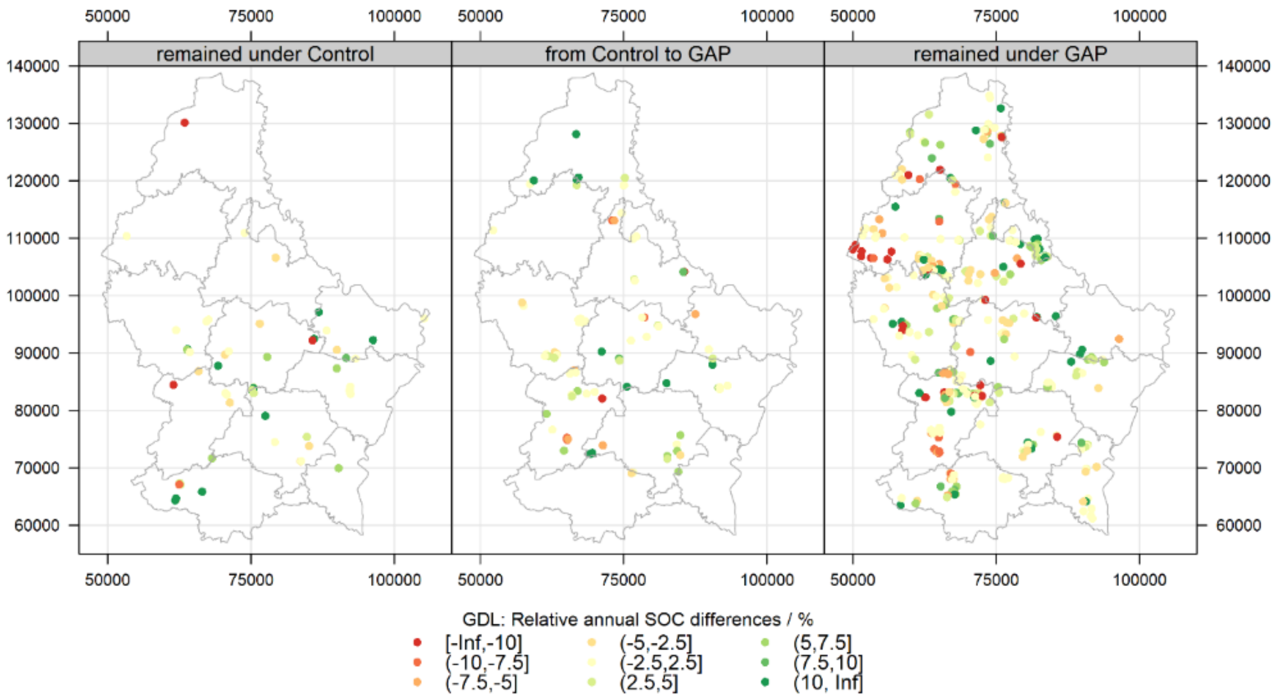

Table 16: Evolution of management practices and relative annual difference of SOC content (%/yr) between the first and the second sampling for the 518 fields with paired observations in ASTA-SOC-LU dataset. (CC = Cover Crops, RT = Reduced Tillage, TG = Temporary Grassland)

Management		**relative annual SOC difference (%/yr)**				
1st sampling	2nd sampling	n	Q1	median	mean	Q3
Control	Control	**53**	**-2.4**	**1.2**	**2.3**	**6.7**
Control	RT	14	-1.4	0.0	2.6	2.1
Control	CC	**47**	**-1.7**	**0.0**	**2.1**	**3.8**
Control	CC RT	11	-0.4	2.3	3.0	5.2
Control	TG	7	-5.4	0.0	-1.1	3.0
Control	RT TG	1	-2.4	-2.4	-2.4	-2.4
Control	CC TG	6	-3.0	4.3	1.4	5.8
RT	RT	**39**	**-1.9**	**1.3**	**2.0**	**7.2**
RT	CC RT	**31**	**-4.2**	**-1.9**	**-1.2**	**2.6**
RT	RT TG	7	5.1	6.0	6.2	7.8
RT	CC RT TG	1	3.0	3.0	3.0	3.0
CC	CC	**56**	**-5.4**	**0.0**	**-0.2**	**4.2**
CC	CC RT	3	-1.5	0.0	9.7	16.0
CC	CC TG	6	-8.1	-5.2	-4.6	-0.6
CC RT	CC RT	**90**	**-2.9**	**0.0**	**1.0**	**4.6**
CC RT	CC RT TG	4	-2.9	-0.2	0.3	3.0
TG	TG	**34**	**-4.1**	**0.0**	**1.9**	**5.7**
TG	RT TG	3	3.0	3.0	4.2	4.8
TG	CC TG	20	0.0	3.1	3.4	4.8
TG	CC RT TG	5	-1.8	1.2	-1.1	1.7
RT TG	RT TG	11	-1.6	2.1	5.1	7.5
RT TG	CC RT TG	10	1.6	4.1	4.9	5.5
CC TG	CC TG	15	-6.0	0.0	-0.3	3.4
CC TG	CC RT TG	2	0.8	2.4	2.4	3.9
CC RT TG	CC RT TG	24	-4.1	-0.8	-2.3	2.4

4.3.3 Relative importance of management practices vs environmental covariates on SOC variability

A first conditional inference tree was developed on the 4016 observations in croplands of the LU-SOC-GAP dataset covering the entire GDL territory (Fig. 32). The model was based on ten covariates and explained ~82% of SOC variance (R^2=0.82) with a RMSE of 0.39%C. The SOC variance was mainly explained by three environmental covariates (varying at the regional scale): the elevation, the clay content and the precipitation. Elevation had a relative importance of almost 35% in the model, whereas clay and precipitation were around 20%. Two others environmental covariates were selected in the model, the minimum depth of soil hydromorphy and their pH, each having a relative variable importance ~ 5%. Five covariates selected in the model were related (or in part related) to farming practices: Mg, K_2O, GAP_app, C factor and GAP_app_years. Each had also a relative importance < 5% in the final model. Finally, the crop year of the sampling was selected with a relative importance < 2%. *For thorough information about the relations between SOC and covariates, except the GAPs, please refer to section 3.3.3 (Fig. 15-16).*

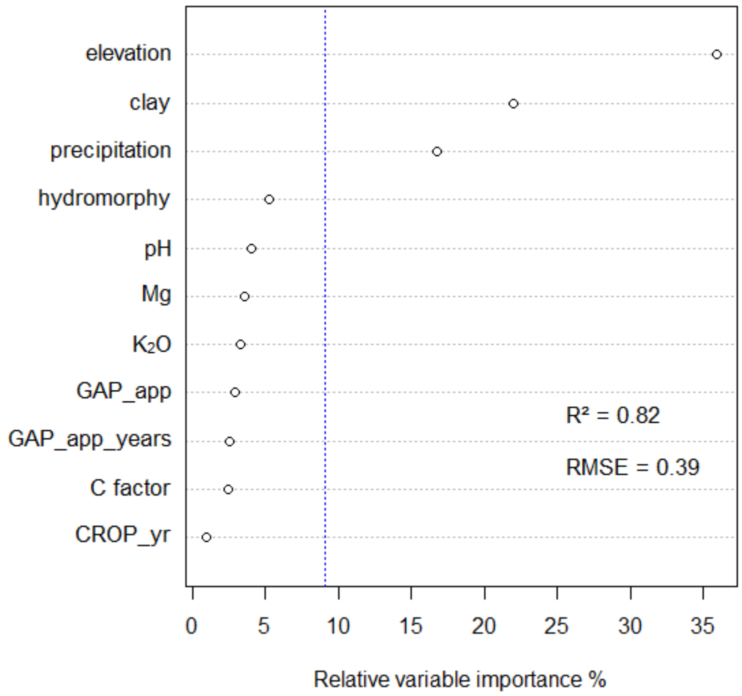

*Figure 32: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC in croplands **all over GDL** (n = 4016). The vertical blue dashed lien indicates the average relative importance.*

To get rid of the major trend induced by the differences between the Oesling and the Gutland, we fitted one additional inference forest for each (Fig. 33-34). The model for Gutland (R^2=0.67; RMSE=0.31) was stronger than that for Oesling (R^2=0.50; RMSE=0.50). SOC variance in Gutland was mainly explained by clay, for which the relative variable importance was > 45%, and then pH (relative variable importance ~ 10%; Fig. 34). The GAP applied and their duration (i.e., GAP_app and GAP_app_years) both had a relative importance ~ 5%. In Oesling, SOC variance is mainly explained by the minimum depth of hydromorphy, the dominant gradients (from N to S) of precipitation and elevation[15], the clay and the Mg contents, and the GAP application (Fig. 33). The GAP_app covariate showed a relative importance of ~ 9% while GAP_app_years accounted for 5%. These differences between Oesling and Gutland could have different origins as: i) Oesling has a smaller range of clay content than Gutland (section 3.3.3), ii) Oesling showed a higher proportion of GAP, especially under TG and combinations including TG, than Gutland, and iii) fields submitted to temporary grassland showed highest differences in SOC content than fields under cover crops or reduced tillage (section 4.3.2). The remaining unexplained SOC variance in each model could be induced by farming/management practices not reflected by the covariates and induced by sampling and SOC measurement.

[15] *The implication of the local positioning, e.g. in valley bottoms, on hillslopes or on plateaus, has a relative implication of ~4% (see TPI: Topographic Position Index in Fig. 33)*

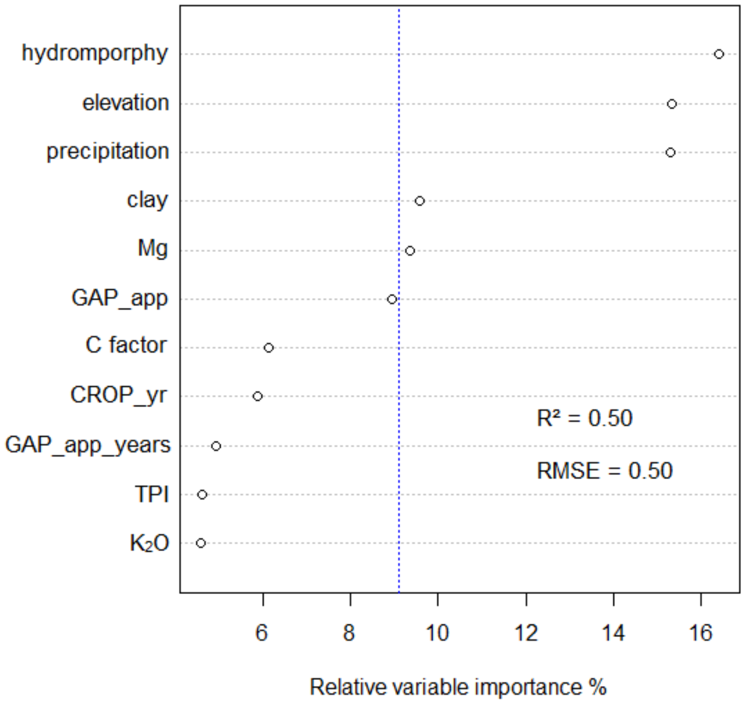

Figure 33: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC ***in croplands of Oesling*** *(n = 1464). The vertical blue dashed line indicates the average relative importance.*

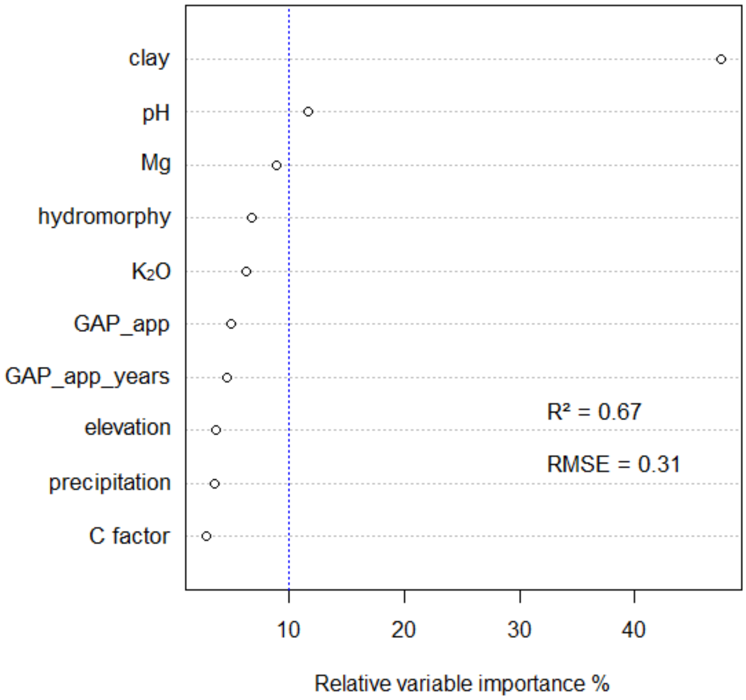

Figure 34: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC ***in croplands of Gutland*** *(n = 2552). The vertical blue dashed line indicates the average relative importance.*

5 CONCLUSIONS

1. Using observations from the ASTA database for the period 2012-2019, we analyzed the spatial variability of soil organic carbon (SOC) in croplands, grasslands and vineyards all over the Grand-Duchy of Luxembourg (GDL). In cropland, in addition to the geographical coordinate couple (x,y - supporting the main regional trends) the clay content, the C-factor, the Mg content, the K2O content, the minimum depth of hydromorphy, the slope and the elevation (by decreasing order of importance) explained SOC variability. For grasslands, in addition to the geographical coordinate couple (x,y), the SOC variability was explained by the clay content, the Mg content, the minimum depth of hydromophy, the K2O content , the elevation and the pH (by decreasing order of importance). Concerning vineyards, no explicit relations between SOC and environmental covariates was observed. Generalized Additive Models were fitted explaining 74% of SOC variance in cropland, 40% in grassland and 14% in vineyards.

2. We studied the recent evolution of soil organic carbon (SOC) in these three landuses considering the transition to the last Rural Development Program (RDP) for 2014-2020. To this aim, we split the data discussed in (1) to consider two distinct sub-periods - T1: 2012-2015 and T2: 2016-2019. In croplands, SOC increased significantly for the soil associations 'Buntsandstein', 'Grès du Luxembourg', 'Dépôts limoneux sur grès' and 'Argiles lourds du Keuper', and decreased significantly for soils of 'Oesling'. In grasslands, only soils from soil association 'Others' (mainly 'Alluvions etColluvions') showed a significant difference (an increase) in SOC content. In vineyards, we observed a significant decrease in SOC for soils on 'Dolomies du Muschelkalk' and 'Argiles lourdes du Keuper'. Few of the significant differences detected between T1 and T2 could have been confirmed by studying paired observations.

   N.B.: While we observed statistically significant differences/trends, those have to be considered very carefully considering the analytical errors of SOC.

3. The GAMs (see above; point 1) were applied to T1 and T2 subsets to map SOC at both period all over GDL. Both maps have the same main patterns. Oesling has significant higher SOC contents than Gutland. SOC patterns in Oesling seems mainly induced by landuse repartition (related to hillslope position), while in Gutland landuse repartition and clay content seem to dominate. By comparing both maps, we estimated that SOC in croplands of Oesling has decreased in its western and northeastern parts while increasing in the southernmost part. In croplands of Gutland, a significant decrease was estimated in the southwestern part and in the easternmost part of the region, though increasing in most of the eastern part. Considering the grasslands, we estimated that SOC decreased in the northwestern, center-eastern and southwestern areas of Gutland while increasing mainly in the valley bottoms of both natural regions. Grassland of Oesling tended to gain SOC in the northern part of this natural region, whereas they tended to loss SOC in the southern part.

   N.B.: Considering the goodness-of-fit of the GAM for grasslands, the differences between both maps have to be considered very carefully.

4. Hypotheses have been put forward on the processes and practices involved in the recent SOC trends. The recent temperature increase is likely to have had a positive impact on C mineralization. Also, these last years were characterized by drier summers that could have a negative impact on biomass production, and by more frequent extreme events enhancing topsoil erosion, especially in cropland. However, soils of wet areas (mainly occupied by

grasslands) could have benefitted from a better productivity in these drier and warmer conditions, i.e. increased productivity in a warmer soil and less stress from asphyxiation. Finally, in the framework of Good Agricultural Practices (see 5 below), changes in management practices could have induced more C inputs and/or less C outputs. Unfortunately, more data and additional research are needed to confirm or not these hypotheses, and to identify which of them is/are involved in the SOC dynamics.

5. By combining analyses from the ASTA for the period 2012-2019 and layers from the Land Parcel Information Service for crop years 2008 to 2019, we analyzed the impact of three Good Agricultural Practices (GAP) on SOC content: cover crops – CC, reduced tillage – RT, and temporary grassland – TG. For eight out of the ten soil associations, fields under GAP (undifferentiated) showed higher SOC content than 'Control' fields. Significant differences were detected in 'Oesling', 'Dolomies du Muschelkalk' and 'Others'. Considered separately, the introduction of temporary grassland in the crop rotation seems the most effective practice for improving SOC content in croplands. Indeed, fields submitted to temporary grassland had mainly higher SOC content than Control fields, with significant positive difference detected in 'Oesling', 'Grès du Luxembourg' and 'Others' (mainly 'Alluvions et Colluvions'). Fields managed with reduced tillage showed higher SOC content but no significance was detected. To finish, parcels submitted to CC only showed negative or equivalent SOC contents to Control fields. Only 'Oesling' showed a significant negative difference. It is worth noting that CC are mainly applied in GDL right before silage maize cultivation which is known as being a powerful humus consumer. So, the application of CC may appear as an effective way to counter-balance the negative effect of silage maize cultivation on SOC. However, more SOC observations from sites under CC and not associated to silage maize are needed to compare with these first results in order to properly test this hypothesis.

N.B.: Some sub-groups in this analysis (soil association x GAP) contained less than 30 observations inducing that statistics and tests of significance related to them could not be considered as relevant.

6. Using conditional inference trees, we studied the relative importance of environmental covariates vs management practices (GAP) on SOC variability in croplands. Considering data all over the GDL territory, the model was able to explain ~80% of the SOC variance predominantly by regional covariates as elevation, clay and precipitation. The application of GAP (considering the type of GAP or combination) and the duration of their application (number of years since the first application in the period 2008 – crop year of sampling) had both a relative importance < 5% in the model. When considering Oesling and Gutland separately, the application of GAP and their duration of application had a total relative importance in explaining SOC variance of ~14% in Oesling and of ~9% in Gutland. This difference between the two natural regions could be induced by a higher proportion of GAP application, especially temporary grassland (TG) and combinations including TG, in Oesling than Gutland.

6 References

Akaike H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723.

Bah B. et Marx S., 2016. Convention de recherche relative à la révision des cartes thématiques sur les zones à risque d'érosion et de ruissellement au Grand-Duché de Luxembourg, Rapport final. Axe Echanges Eau-Sol-Plante, Gembloux Agro-Bio Tech, Université de Liège. 81 p.

Bivand R.S., Pebesma E.J., Gomez-Rubio V., 2013.Applied spatial data analysis with R. Springer, New York, USA.

Boer E.P.J., de Beurs K.M., and Hartkamp A.D., 2001. Kriging and thin plate splines for mapping climate variables. International Journal of Applied Earth Observation and Geoinformation 3(2), 146-154

Bronick C.J., Lal, R. 2005. Soil structure and management: A review. Geoderma 124, 3-22

Busari M.A., Kukal S.S., Kaur A., Bhatt R., Dulazi A.A., 2015. Conservation tillage impacts on soil, crop and the environment. International Soil and Water Conservation Research 3(2), 119-219.

Buysse P., Roisin C., Aubinet M., 2013. Fifty years of contrasted residue management of an agricultural crop: Impacts on the soil carbon budget and on soil heterotrophic respiration. Agriculture, Ecosystems & Environment 167, 52-59.

Chenu C., Le Bissonnais Y., Arrouays D. 2000. Organic matter influence on clay wettability and soil aggregate stability. Soil Science Society of America Journal 64(4), 1479-1486.

Clivot H., Mary B., Valé M., Cohan J.-P., Champolivier L., Piraux F., Laurent F., Justes E. 2017. Quantifying in situ and modeling net nitrogen mineralization from soil organic matter in arable cropping systems. Soil Biology and Biochemistry 111, 44-59.

Davidson E.A. and Janssens I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. Nature 440(7081), 165-173.

D'Hose T., Cougnon M., De Vliegher A., Vandecasteele B., Viaene N., Cornelis W., Van Bockstaele E., Reheul D. 2014. The positive relationship between soil quality and crop production: A case study on the effect of farm compost application. Applied Soil Ecology 75, 189-198

Fenoll J., Ruiz E., Flores P., Vela N., Hellín P., Navarro S. 2011. Use of farming and agro-industrial wastes as versatile barriers in reducing pesticide leaching through soil columns. Journal of Hazardous Materials 187, 206-212

Fenoll J., Vela N., Navarro G., Pérez-Lucas G., Navarro S. 2014. Assessment of agro-industrial and composted organic wastes for reducing the potential leaching of triazine herbicide residues through the soil. Science of the Total Environment 493, 124-132

Gale P.M. and Gimour J.T., 1988. Net mineralization of carbon and nitrogen under aerobic and anaerobic conditions. Soil Sci. Am. Proc. 20, 218-224

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

Hijmans R.J. and van Etten, J., 2013. raster: Geographic analysis and modeling with raster data. R package version 1, 9-92.

Hobley E., Wilson B., Wilkie A., Gray J., Koen T., 2015. Drivers of soil organic carbon storage and vertical distribution in Eastern Australia. Plant Soil 390, 111–127.

Hobley E.U., Baldock J., Wilson B., 2016. Environmental and human influences on organic carbon fractions down the soil profile. Agric. Ecosyst. Environ. 223, 152–166.

ISO 10694:1995, Soil quality — Determination of organic and total carbon after dry combustion (elementary analysis)

ISO 11464:2006, Soil quality — Pretreatment of samples for physico-chemical analysis

Jacobsen B.H. and Ørum J.E., 2010. Farm economic and environmental effects of reduced tillage. Acta Agriculturae Scandincavica 6(2), 134-142.

Jarvis A., Reuter H.I., Nelson A. and Guevara E., 2008. Hole-_lled SRTM for the globe Version 4. available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org), 2008.

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

Kerry R., Goovaerts P., Rawlins B.G., and Marchant B.P., 2012 Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. Geoderma 170, 347-358.

Kirschbaum M.U. F., 1995. The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. Soil Biology and Biochemistry 27(6), 753-760.

Lal, R., 1993. Tillage effects on soil degradation, soil resilience, soil quality, and sustainability 27(1-4), 1-8.

Loveland P. and Webb J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil and Tillage Research 70(1), 1-18.

Marx S., Flammang, F., 2018. La cartographie des sols au Grand-Duché de Luxembourg - Légende de la carte des sols détaillée à l'échelle 1/25.000. Ministère de l'Agriculture, de la Viticulture et de la Protection des consommateurs. 39p.

Meersmans J. van Wesemael B., and van Molle M., 2009. Determining soil organic carbon for agricultural soils: a comparison between the Walkely & Black and the dry combustion methods (north Belgium). Soil Use and Management 25, 346-353.

Meersmans J., Van Wesemael B., Goidts E., Van Molle, M., De Baets S., and De Ridder F., 2011. Spatial analysis of soil organic carbon evolution in belgian croplands and grasslands, 1960-_2006. Global Change Biology 17(1), 466-479.

Olaya V., 2004. A gentle introduction to SAGA GIS. The SAGA User Group eV, Gottingen, Germany, 208, 2004.

Oldfield E.E., Bradford M.A., Wood S.A., 2019. Global meta-analysis of the relationship between soil organic matter and crop yields. SOIL 5, 15-32.

PANATTIERI M., RUMPEL C., DIGNAC M.-F., CHABBI A., 2017. DOES GRASSLAND INTRODUCTION INTO CROPPING CYCLES AFFECT CARBON DYNAMICS THROUGH CHANGES OF ALLOCATION OF SOIL ORGANIC MATTER WITHIN AGGREGATE FRACTIONS? SCIENCE OF THE TOTAL ENVIRONMENT 576, 251-263.

PARIKH S. J. AND JAMES B. R., 2012. SOIL: THE FOUNDATION OF AGRICULTURE. NATURE EDUCATION KNOWLEDGE 3(10):2

PELLERIN S., BAMIÈRE L ET AL., 2019. STOCKER DU CARBONE DANS LES SOLS FRANÇAIS, QUEL POTENTIEL AU REGARD DE L'OBJECTIF 4 POUR 1000 ET À QUEL COÛT ? SYNTHÈSE DU RAPPORT D'ÉTUDE, INRA (FRANCE), 114 P.

POST W.M., EMANUEL W.R., ZINKE P.J., STANGENBERGER A.G., 1982. SOIL CARBON POOLS AND WORLD LIFE ZONES. NATURE 298(5870), 156-159.

SIX J., CONANT R.T., PAUL E.A., PAUSTIAN K., 2002. STABILIZATION MECHANISMS OF SOIL ORGANIC MATTER: IMPLICATIONS FOR C-SATURATION OF SOILS. PLANT AND SOIL 241(2), 155-176.

SKOPP J., JAWSON M. D, DORAN J. W., 1990. STEADY-STATE AEROBIC MICROBIAL ACTIVITY AS A FUNCTION OF SOIL WATER CONTENT. SOIL SCIENCE SOCIETY OF AMERICA JOURNAL 54(6), 16-19.

SNAPP S.S., SWINTON S.M., LABARTA R., MUTCH D., BLACK J.R., LEEP R., NYIRANEZA J., O'NEIL K., 2005. EVALUATING COVER CROPSFOR BENEFITS, COSTS AND PERFORMANCE WITHIN CROPPING SYSTEM NICHES. AGRON. J. 97, 322-332.

SPW - DGO3 - DEMNA - DEE, 2017. RAPPORT SUR L'ÉTAT DE L'ENVIRONNEMENT WALLON 2017 (REEW 2017). SPW ÉDITIONS: JAMBES, BELGIQUE. EN LIGNE. HTTP://ETAT.ENVIRONNEMENT.WALLONIE.BE

STEFFEN M., MARX S., LEYDET L., 2019. ÉVALUATION ET COMPARAISON DES DIFFÉRENTES TECHNIQUES D'INTERPOLATION SPATIALE POUR L'ÉTABLISSEMENT DE LA CARTE DES TEXTURES POUR LE GRAND-DUCHÉ DE LUXEMBOURG. MINISTÈRE DE L'AGRICULTURE, DE LA VITICULTURE ET DU DÉVELOPPEMENT RURAL, ADMINISTRATION DES SERVICES TECHNIQUES DE L'AGRICULTURE - SERVICE DE PÉDOLOGIE, GRAND-DUCHÉ DE LUXEMBOURG, 106 P.

STEVENS A., VAN WESEMAEL B., MARX S., LEYDET L., 2014. MAPPING TOPSOIL ORGANIC CARBON IN GRAND-DUCHY OF LUXEMBOURG. TECHNICAL REPORT, MINISTÈRE DE L'AGRICULTURE, DE LA VITICULTURE ET DE LA PROTECTION DES CONSOMMATEURS, GRAND-DUCHÉ DE LUXEMBOURG, 59P.

STEVENS A., VAN WESEMAEL B., MARX S., LEYDET L., 2014. MAPPING TOPSOIL ORGANIC CARBON STOCKS IN GRAND-DUCHY OF LUXEMBOURG. TECHNICAL REPORT, MINISTÈRE DE L'AGRICULTURE, DE LA VITICULTURE ET DE LA PROTECTION DES CONSOMMATEURS, GRAND-DUCHÉ DE LUXEMBOURG, 16P.

STROBL C., BOULESTEIX A., ZEILEIS A., HOTHORN T., 2007. BIAS IN RANDOM FOREST VARIABLE IMPORTANCE MEASURES: ILLUSTRATIONS, SOURCES AND A SOLUTION. BMC BIOINF. 8, 25.

TRUMBORE S.E., CHADWICK O.A., AMUNDSON R., 1996. RAPID EXCHANGE BETWEEN SOIL CARBON AND ATMOSPHERIC CARBON DIOXIDE DRIVEN BY TEMPERATURE CHANGE. SCIENCE 272(5260), 393-396.

VAN-CAMP L, BUJARRABAL B, GENTILE AR, JONES RJA, MONTANARELLA L, OLAZABAL C, SELVARADJOU S-K, 2004. REPORTS OF THE TECHNICAL WORKING GROUPS ESTABLISHED UNDER THE THEMATIC STRATEGY FOR SOIL PROTECTION (EUR 21319 EN/3). OFFICE FOR OFFICIAL PUBLICATIONS OF THE EUROPEAN COMMUNITIES: LUXEMBOURG, GRAND-DUCHÉ DE LUXEMBOURG.

VANDEN NEST T., VANDECASTEELE B., RUYSSCHAERT G., COUGNON M., MERCKX R., REHEUL D., 2014. EFFECT OF ORGANIC AND MINERAL FERTILIZERS ON SOIL P AND C LEVELS, CROP YIELD AND P LEACHING IN A LONG TERM TRIAL ON A SILT LOAM SOIL. AGRICULTURE, ECOSYSTEMS AND ENVIRONMENT 197, 309-317

VAN OOST K., VAN MUYSEN W., GOVERS G., DECKERS J., QUINE T.A., 2005. FROM WATER TO TILLAGE DOMINATED LANDFORM EVOLUTION. GEOMORPHOLOGY 72, 193-203.

VAN WESEMAEL B., CHARTIN C., WIESMEIER M., VON LÜTZOW M., HOBLEY E., CARNOL M., KRÜGER I., CAMPION M., ROISIN C., HENNART S., KÖGEL-KNABNER I., 2019. AN INDICATOR FOR ORGANIC MATTER DYNAMICS IN TEMPERATE AGRICULTURAL SOILS. AGRICULTURE, ECOSYSTEMS AND ENVIRONMENT 274, 62-75

VIAUD V., SANTILLÀN-CARVANTES P., AKKAL-CORFINI N., LE GUILLOU C., CHEMIDLIN PRÉVOST-BOURRÉ N., RANJARD L., MENASSERI-AUBRY S., 2018. LANDSCAPE-SCALE ANALYSIS OR CROPPING SYSTEM EFFECTS ON SOIL QUALITY IN A CONTEXT OF CROP-LIVESTOCK FARMING. AGRICULTURE, ECOSYSTEMS & ENVIRONMENT 265, 166-177.

WIESMEIER M., URBANSKI L., HOBLEY E., LANG B., VON LÜTZOW M., MARIN-SPIOTTA E., VAN WESEMAEL B., RABOT E., LIEß M., GARCIA-FRANCO N., WOLLSCHLÄGER U., VOGEL H.-J., KÖGEL-KNABNER I., 2019. SOIL ORGANIC CARBON STORAGE AS A KEY FUNCTION OF SOILS - A REVIEW OF DRIVERS AND INDICATORS AT VARIOUS SCALES. GEODERMA 333, 149-162.

WOOD S.N., 2001. MGCV: GAMS AND GENERALIZED RIDGE REGRESSION FOR R. R NEWS 1(2), 20-25.

XU H., SIEVERDING H., KWON H., CLAY D., STEWART C., JOHNSON J.M.F., QIN Z., KARLEN D.L., WANG M., 2019. A GLOBAL META-ANALYSIS OF SOIL ORGANIC RESPONSE TO CORN STOVER REMOVAL. GCB BIONENERGY 11(10), 1215-1233.

ZAR J.H., 1999. BIOSTATISTICAL ANALYSIS. PRENTICE HALL, NEW JERSEY.

7 Annexes

7.1 Methodology: Map of the Minimum Depth of soil Hydromorphy

The minimum depth of soil hydromorphy corresponds to the minimum depth in soil profile where physical indicators of temporary or continuous surface water saturation were observed.

The table Anx1 compiled the range of depth for the presence of oxido-reduction and reduction features by drainage classes in each soil type of GDL (as defined on the texture triangle designed for soils of Belgium and GDL; Fig. 4).

Table Anx1: characteristics of the drainage classes for soils of GDL (source: Marx et Flammang, 2018).

Classe de drainage		Définition		Drainage naturel	Profondeur (cm) d'apparition des phénomènes d'oxydo-réduction (pseudogley) ou de réduction (gley)		Attribut SIG
		Texture Z, S, P	Texture L, A, E, U, G		Oxydo-réduction	Réduction	
.a.	.B.	sols très secs	-	excessif	-	-	DRAINAGE
.b.		sols secs	sols non gleyifiés	parfait	> 80	-	
.c.	.D.	sols modérément secs	sols faiblement gleyifés	modéré	60-80	-	
.d.		sols modérément humides	sols modérément gleyifiés	imparfait	30-60	-	
.h.	.I.	sols humides	sols fortement gleyifiés (à engorgement d'eau temporaire)	assez pauvre, sans horizon réduit	Taches légères entre 0-30	-	
.i.		sols très humides	sols très fortement gleyfiés (à engorgement d'eau temporaire)	pauvre, sans horizon réduit	Taches importantes entre 0-30	-	
.e.	.F.	sols humides	sols fortement gleyifiés à horizon réduit (à engorgement d'eau permanent... avec zone de battement)	assez pauvre, à horizon réduit	Taches légères entre 0-30	40-80	
.f.		sols très humides	sols très fortement gleyifiés à horizon réduit (à engorgement d'eau permanent... avec zone de battement)	pauvre, à horizon réduit	Taches importantes entre 0-30	40-80	
.g.		sols extrêmement humides	sols réduits (nappe phréatique permanente... sans zone de battement)	très pauvre	-	< 40	

In order to join the minimum depth of soil hydromorphy to the numerical soil map of GDL, we simplified the Table Anx1 to create the Table Anx2. For drainage classes c, d, D, e, f, F, h, i, I and g, we applied the central value of the observed depth ranges of oxido-reduction features compiled in the Table Anx1 as the minimum depth of soil hydromorphy. In GDL, soil augering investigations were performed between depths of 0 and 80cm. Hence, for drainage classes with potential minimum depth of soil hydromorphy > 80cm (a, b and B - referred as 'Absence' in figure 8), we used range of depths observed for equivalent soils in Belgium, where soil profiles were investigated till a depth of 120cm (Table 1 in Meersmans et al., 2009). In addition, when oxido-reduction features could not be observed because located deeper than 120cm in the profile or absent, a default value of 120cm was applied. ***These decisions were taken in order to create a covariate layer, continuous in space, covering the maximum territory of GDL, i.e. to optimize in the modeling procedure the inclusion of the known drainage influence on SOC content all over GDL.***

Table Anx2: Table joined to the attribute table of the 1:25000 numerical soil map (Bah and Marx, 2016) to create the map of minimum depth of soil hydromorphy (cm) in GDL.

Drainage class	Minimum depth of soil hydromorphy (cm)
a	120
b	100
B	110
c	70
d	45
D	60
e	15
f	15
F	15
h	15
i	15
I	15
g	0

Once the Table Anx2 was joined to the attribute table of the 1:25000 numerical soil map (Bah et Marx, 2016), we mapped the minimum depth of soil hydromorphy all over the GDL. Using a majority rule[16], this vector map (polygons) was then converted to a raster format (resolution of 90x90m) with the same exact characteristics as all the covariate layers used here in the SOC mapping procedure. To finish, a low-pass filter was applied to the raster layer in order to smooth the transitions between areas of different minimum depths of soil hydromorphy.

[16] *Majority rule: when converting polygons to a raster format, if a cell from the raster grid covers different polygons, this cell will take the value corresponding to the polygon covering the biggest proportion of the cell's area.*

7.2 SOC SUMMARY STATISTICS: REFERENCE TABLES FOR TEXTURE CLASSES L, M, OM AND S

This annex provides tables of summary statistics and results of difference test of SOC for the 4 texture classes defined by ASTA. Tables Anx3, Anx4 and Anx5 are related to cropland, grassland and vineyard, respectively.

*Table Anx3: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **croplands** at T1 (2012-2015) and T2 (2016-2019), and significance of the difference between these two periods (non-paired Mann-Whitney test). (L = léger, M = moyen, OM = moyen caillouteux, S = lourds)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Texture	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	2225	0.40	1.20	1.60	1.94	2.50	6.40	3142	0.40	1.50	2.10	2.27	2.90	6.10		
L	558	0.40	0.90	1.10	1.11	1.20	2.70	327	0.60	0.98	1.10	1.16	1.30	3.20	0.05	NS
M	928	0.70	1.30	1.60	1.74	2.00	5.00	1268	0.50	1.40	1.70	1.78	2.00	5.70	**0.03**	**< 0.05**
OM	555	1.30	2.50	3.00	3.10	3.60	6.40	1379	0.90	2.50	2.90	3.00	3.40	6.10	**-0.1**	**< 0.05**
S	184	0.50	1.50	1.80	1.96	2.30	3.90	168	0.40	1.68	2.10	2.08	2.40	5.20	**0.11**	**< 0.05**

*Table Anx4: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **grasslands** at T1 (2012-2015) and T2 (2016-2019), and significance of the difference between these two periods (non-paired Mann-Whitney test). (L = léger, M = moyen, OM = moyen caillouteux, S = lourds)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Texture	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	679	0.80	2.70	3.40	3.62	4.35	10.10	1452	0.70	2.80	3.40	3.55	4.30	8.30		
L	30	1.00	1.53	1.70	1.96	2.08	4.40	44	0.70	1.30	1.65	1.89	2.05	5.60	-0.08	NS
M	380	0.80	2.50	3.40	3.62	4.40	10.10	719	0.70	2.50	3.30	3.46	4.20	8.30	-0.15	NS
OM	232	1.70	3.08	3.60	3.75	4.20	7.40	581	0.90	3.00	3.50	3.63	4.20	7.30	-0.13	NS
S	37	1.90	3.30	3.90	4.07	4.90	7.30	108	0.80	3.50	4.35	4.44	5.23	7.20	0.36	NS

*Table Anx5: Descriptive statistics of topsoil SOC (%C for the 0-25cm depth) in **vineyards** at T1 (2012-2015) and T2 (2016-2019), and significance of the difference between these two periods (non-paired Mann-Whitney test). (L = léger, M = moyen, OM = moyen caillouteux, S = lourds)*

	T1: 2012-2015							T2: 2016-2019							Difference	
Texture	n	min	Q1	median	mean	Q3	max	n	min	Q1	median	mean	Q3	max	mean	p-value
ALL	1916	0.40	1.20	1.65	1.83	2.21	5.50	2405	0.10	1.10	1.55	1.63	2.10	5.00		
M	118	0.60	1.33	2.00	2.15	2.80	5.35	280	0.40	1.35	1.90	1.98	2.50	5.00	-0.17	NS
S	1798	0.40	1.20	1.60	1.81	2.20	5.50	2125	0.10	1.10	1.50	1.59	2.00	4.70	**-0.22**	**< 0.001**

7.3 Additional SOC maps for cropland and grassland

Figures Anx1 and Anx2 propose some additional SOC maps for cropland and grassland, respectively. Each figure contains 4 maps corresponding for each landuse to:

- A: the map resulting from the application all over the GDL territory of the GAM fitted for T1 (2012-2015);
- B: the map resulting from the application all over the GDL territory of the GAM fitted for T2 (2016-2019);
- C: the map resulting from the application all over the GDL territory of the GAM fitted for T1+T2 (2012-2019);
- D: the standard error of SOC estimation associated to the GAM pictured in C, i.e. the GAM fitted for T1+T2 (2012-2019).

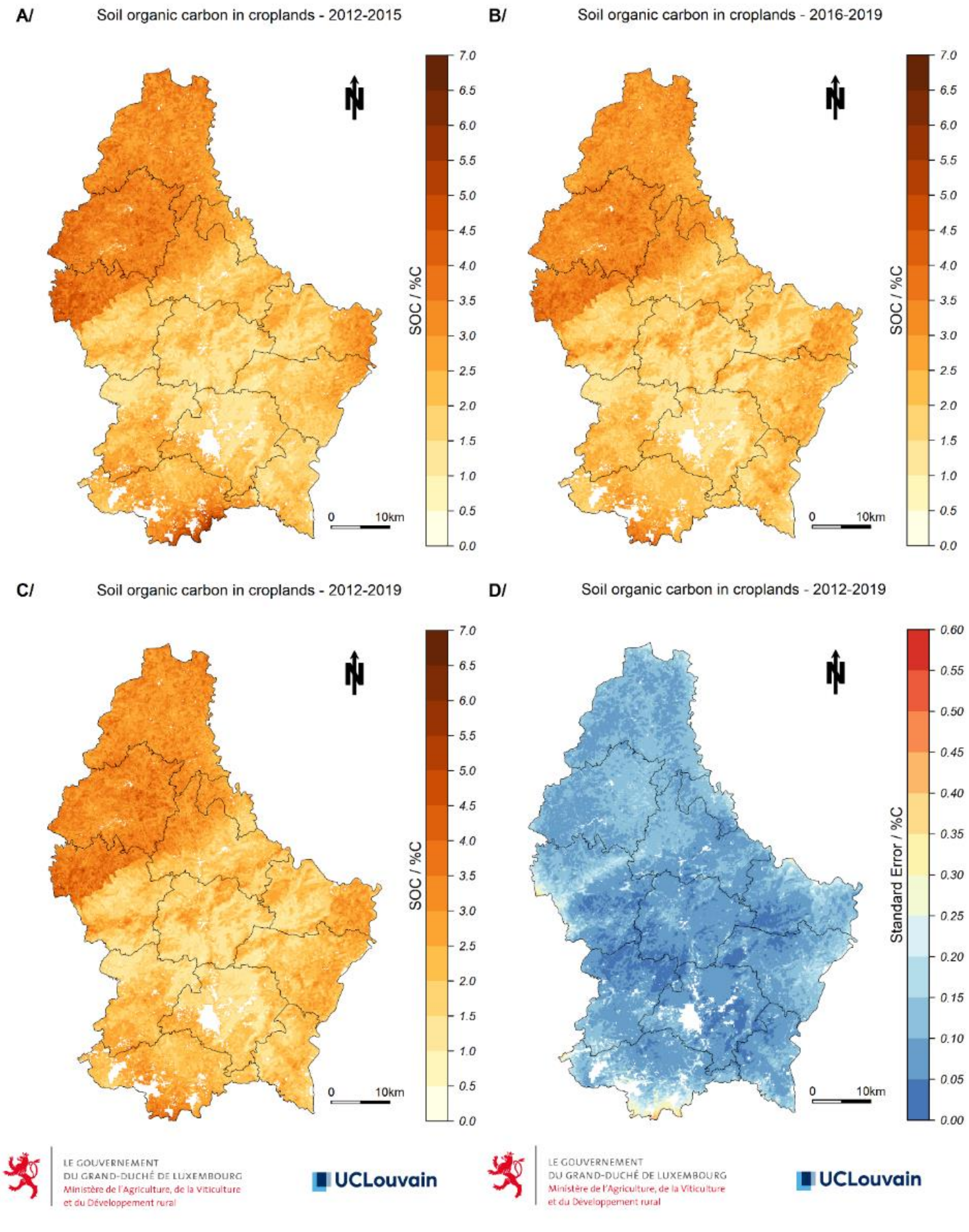

*Figure Anx1: additional maps of SOC considering the whole territory as **croplands** – A/ SOC (%C) for period 2012-2015, B/ SOC (%C) for period 2016-2019, C/ SOC (%C) for period 2012-2019 and D/ standard error of estimation (%C) for period 2012-2019.*

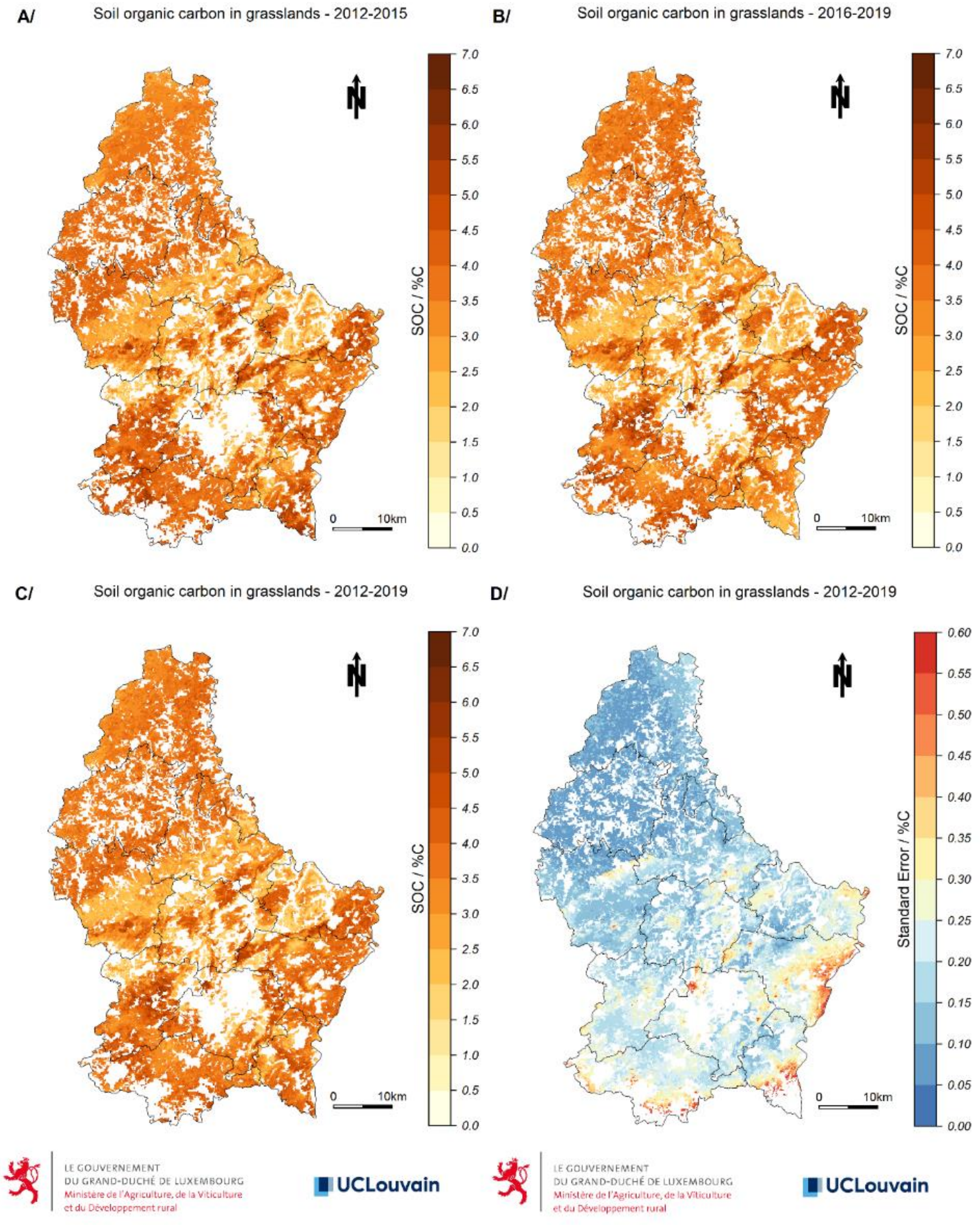

*Figure Anx2: additional maps of SOC considering the whole territory as **grasslands** – A/ SOC (%C) for period 2012-2015, B/ SOC (%C) for period 2016-2019, C/ SOC (%C) for period 2012-2019 and D/ standard error of estimation (%C) for period 2012-2019. (White areas are not covered by the pH map - Fig. 13A, a significant covariate in the SOC model for grasslands)*

LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG
Ministère de l'Agriculture, de la Viticulture
et du Développement rural

UCLouvain