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

UCLouvain

ORGANIC CARBON IN SOILS OF THE GRAND-DUCHY OF LUXEMBOURG: PAST, PRESENT AND FUTURE

April 2022

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

EXECUTIVE SUMMARY

SOC content and stock maps are essential tools in the discussion about soil health, climate change and the role of soils as sources or negative emissions of greenhouse gases. The first versions of soil organic carbon (SOC) content (required to assess the role for soil health) and soil organic carbon stock (required for the emission/sink of CO_2) maps for the Grand-Duchy of Luxembourg date back to 2014. Recently, the soil analytical database in the Grand Duchy of Luxembourg (GDL) was enlarged with thousands of samples from croplands, grasslands and vineyards. Moreover, soil databases from the beginning of the twentieth century and dating back to the 1960s have been recovered. The wealth of (historic) data on soil organic carbon, land use and land management is the main motivation for a more extensive investigation of the SOC status in Luxembourg including historical trends and effects of management options within a context of marked spatial trends in climate, relief and lithology across the GDL.

The SOC data for the beginning of the 20th century relate to humic acids and therefore only their spatial pattern can be analyzed. The Oesling region can clearly be distinguished by the higher humic acid content. This pattern is similar to the one for the modern SOC content. The higher humic acid contents in the region around Ettelbrück do not match modern SOC content patterns. The spatial patterns in SOC content in the 1960s are similar to the current ones with highest values in grasslands of the clayey soil associations in the Oesling and in the dolomites du Muschelkalk. The trends in cropland are the same but at slightly lower values compared to the grasslands. Overall, SOC contents in croplands have slightly decreased over time, while for grassland there was a slight increase in the clayey soil associations.

Spatial regression models were fitted for two recent periods - T1 (2012-2015) and T2 (2016-2019) - to predict SOC contents and its evolution during the last decennium. Overall SOC contents are highest in the Oesling for both grasslands and croplands, while in the Gutland land use and clay content are the main drivers for the spatial patterns. The spatial patterns of SOC content and its dynamics over the last decade can be seen in Figures 4.20-4.22.

Fields under 'good agricultural practices' (GAP) such as cover crops, reduced tillage and temporary grassland, showed higher SOC content than fields where GAP were not applied. The introduction of temporary grassland in the crop rotation seems the most effective practice for improving SOC content in croplands. Reduced tillage leads to higher SOC content but changes were not statistically significant.

There is a general decline of the cropland area in favor of grassland since the beginning of the 20th century. These historic conversions from cropland to grassland increased SOC stocks of the concerned areas in Oesling, Buntsandstein and Dolomies du Muschelkalk regions by 20-50%, while the SOC stocks have more than doubled in the converted areas of the rest of Gutland.

Overall patterns in SOC stocks for the topsoil (0-30 cm) are similar to the ones in SOC content (Figure 5.13). Statistically significant losses in SOC stocks in cropland soils over the last decade were observed in Oesling, while statistically significant gain occurred in 'Dépôts limoneux sur Grès'. While SOC stocks in grassland remained stable, losses were detected in vineyard soils of 'Dolomies du Muschelkalk' and 'Argiles Lourdes du Keuper' (an improvement of the Total Inorganic Carbon measurement could partly explain these trends). Overall, SOC stocks in GDL appeared to be stable at the national scale since 2012, while local statistically significant increase or decrease occurred. In the future, more attention should be payed to the regions submitted to these recent losses.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AEM: Agri-Environmental Measure
ASTA: Administration des Services Techniques de l'Agriculture (Administration of Agricultural Technical Services)
BSP: Backward Stepwise Procedure
CAP: Common Agricultural Policy
CC: Cover Crop
CI: Confidence Interval
CV: Cross-Validation
DSM: Digital Soil Mapping
ESM: Equivalent Soil Mass
EU: European Union
GAM: Generalized Additive Model
GAP: Good Agricultural Practices
GBM: Generalized Boosting Model
GDL: Grand Duchy of Luxembourg
IACS: Integrated Administration and Control System
IQR: Inter-Quartile Range
L: loamy sand and sandy soils in Gutland
LPIS: Land Parcel Information System
M: clay loam - loam - silt loam soils in Gutland
ME: Mean Error
OBS: Occupation Biophysique du Sol
OM: shallow stony silt loam soils in Oesling
PCN: Plan Cadastral Numérisé (Digitized cadastral plan)
PI: Prediction Interval
Rel. CI: relative Confidence Interval
Rel. PI: relative Prediction Interval
RF: Random Forest
RMSE: Root Mean Square Error
RT: Reduced Tillage
S: clay and heavy clay soils in Gutland
SE: Standard Error
SOC: Soil Organic Carbon
SOM: Soil Organic Matter
TIC: Total Inorganic Carbon
TG: Temporary Grassland

HOW TO CITE THIS REPORT

CHARTIN C., VAN WESEMAEL B., MARX S., STEFFEN M., LEYDET L., 2022. ORGANIC CARBON IN SOILS OF THE GRAND-DUCHY OF LUXEMBOURG: PAST, PRESENT AND FUTURE. EARTH AND LIFE INSTITUTE, UNIVERSITÉ CATHOLIQUE DE LOUVAIN - MINISTRY OF AGRICULTURE, VITICULTURE AND RURAL DEVELOPMENT, LUXEMBOURG. 110P.

Further thanks go to Cedric Ries, Joe Mulbach, Michèle Muller, Steve Nickels, Sandra Bousson, Frank Flammang and Ben Leiner from the soil laboratory and the soil survey of the Soil Department of ASTA.

1 INTRODUCTION

Soil organic matter (SOM) is a key indicator of soil quality regarding its many positive benefits in terms of agronomy and environment. SOM consists for 57% of Soil Organic Carbon (SOC) and plays an important role for:

- improving soil stability and reducing erosion (Chenu et al., 2000; Bronick et al., 2005);
- providing plant nutrients (Clivot et al., 2017; Oldfield et al., 2019) enabling to reduce the use of nitrogenous fertilizers;
- the degradation and adsorption of phytopharmaceuticals used for crop protection and pest control (Fenoll et al., 2011; 2014);
- 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.

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. 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. The first versions of SOC content and stock maps for the Grand-Duchy of Luxembourg (GDL) date back to 2014 and used a Digital Soil Mapping technique (DSM; Stevens et al., 2014a,b). The DSM consists in fitting a multiple regression to observed data (routine analyses for fertilizer advice in the case of agricultural soils) using a set of spatialized layers of covariates, i.e. environmental factors controlling soil formation and explaining the spatial variability of the target variable (Lagacherie, 2008; McBratney et al., 2003). 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 GDL is characterized by a strong variation in climate and relief along with a heterogeneity in lithology. This variation enhances the potential for DSM techniques as it extends the range in nearly all covariates, but it forms a challenge for soil monitoring. Unravelling the effects of natural variation in climate, relief and lithology from those of management on the dynamics of SOC require a comprehensive spatial and temporal analysis.

Since the last SOC inventory of 2014, the soil analytical database in the GDL was enlarged with thousands of samples from croplands, grasslands and vineyards. Moreover, soil datasets from the early 20th century and the 1960s have been georeferenced and the associated analytical protocols retrieved. Together with these soil datasets, accurate historical maps of land use and more recent databases on Agri-Environmental Measures (AEM) exist. The wealth of these data combined with the systematic variation in factors driving the SOC for agricultural soils is one of the main motivations for a more extensive investigation of the SOC status in GDL; including historical trends and effects of management options.

Objectives

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

This report depicts the works and results of a 3-years (march 2019 - march 2022) research collaboration between the Ministry of Agriculture, Viticulture and Rural Development (GDL) and the UCLouvain (Belgium).

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. 2.1).

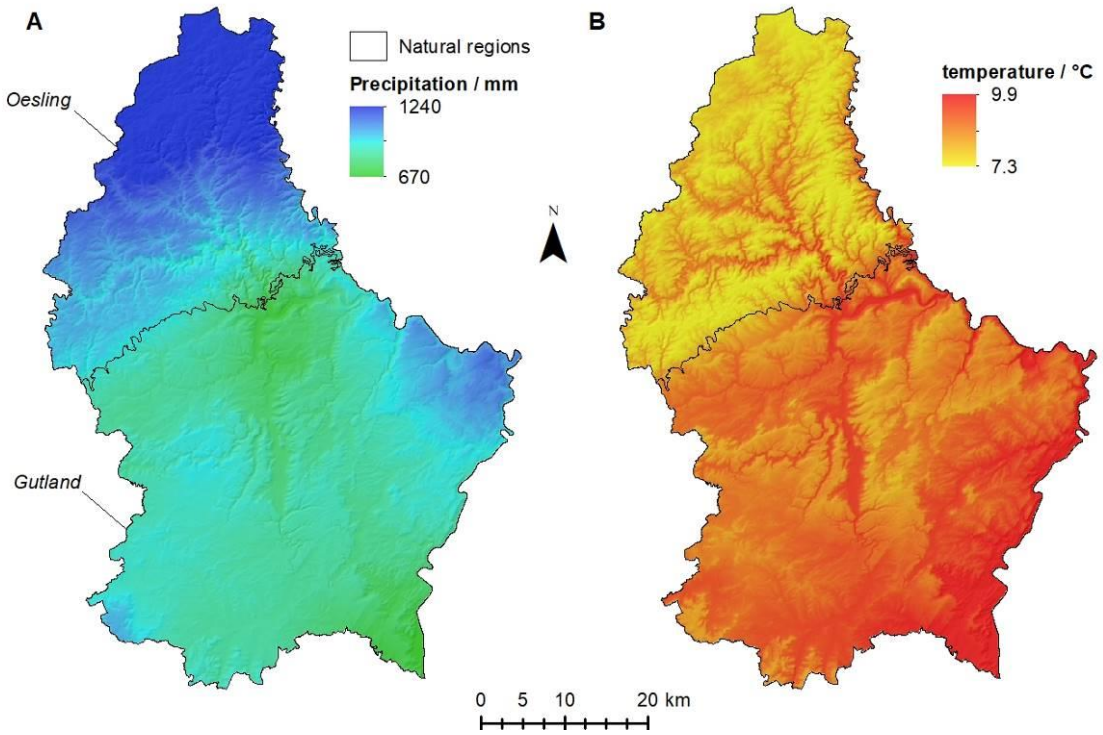


Figure 2.1: Maps of A/ mean annual precipitation (mm) and B/ mean annual temperature (°C) in the Grand-Duchy of Luxembourg over the period 1971-2000 (source: Stevens et al.; 2014a)

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 elevation of ca. 450m (Fig. 2.2). The Gutland is a more heterogeneous region characterized by a south-west-facing cuesta topography (mean elevation of ca. 245m) 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.

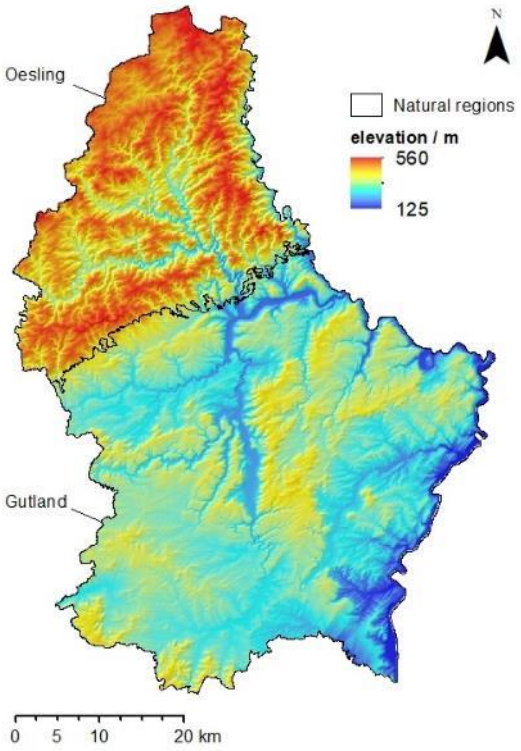


Figure 2.2: Natural regions and elevation (m) in Grand-Duchy of Luxembourg.

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 types (SMU = soil mapping units) in GDL (Carte des sols du Grand-Duché de Luxembourg, 1969, 1/100.000), according to the major geological units, that we further regrouped into 10 associations (Table 2.1; Figure 2.3A) representing variations based mainly on mineralogy (and texture). Soil association 'autres' is made up of alluvium and colluvium.

In addition, based on the texture identified by finger testing on soil samples entering the soil laboratory of the Administration des services techniques de l'agriculture (ASTA), GDL has defined four main textural soil types (Fig. 2.3B). Each of them covers different classes of the texture triangle designed for soils of Belgium and GDL (Table 2.2 and Fig. 2.4). The shallow stony silt loam soils (OM) covers the northern natural region - Oesling. The southern and central natural region – Gutland - is mainly covered by clay loam - loam - silt loam soils (M), loamy sand (L) and clay (S) soils.

Table 2.1: Soil associations and corresponding WRB classification (IUSS Working Group WRB, 2014)

Ref.	Soil association	Natural region	Geology	WRB classification	Relative area (%)
1	Oesling	Oesling	Lower devonian	skeletic dystric Cambisol (siltic)	29.1
2	Buntsandstein	Gutland	Triassic	endolomitic Luvisol (loamic)	3.2
3	Dolomies du Muschelkalk	Gutland	Triassic	leptic dolomitic Cambisol	10.1
4	Calcaires du Bajocien	Gutland	Jurassic	leptic calcaric Cambisol (loamic)	12.5
5	Gres du Luxembourg	Gutland	Jurassic	haplic Luvisol (arenic), Arenosol	9.6
6	Depots limoneux sur Gres	Gutland	Jurassic	haplic Luvisol (loamic)	10.6
7	Argiles du Lias inf. et moyen	Gutland	Jurassic	gleyic/stagnic endocalcaric Luvisol (loamic)	1.3
8	Argiles lourdes du Keuper	Gutland	Triassic	vertic dolomitic Cambisol	5.1
9	Argiles lourdes des schistes bitumineux	Gutland	Triassic	vertic calcaric Cambisol (clayic)	4.6
10	Autres	Gutland and Oesling	Alluvium, Colluvium	Fluvisol, Cambisol, Regosol	13.9

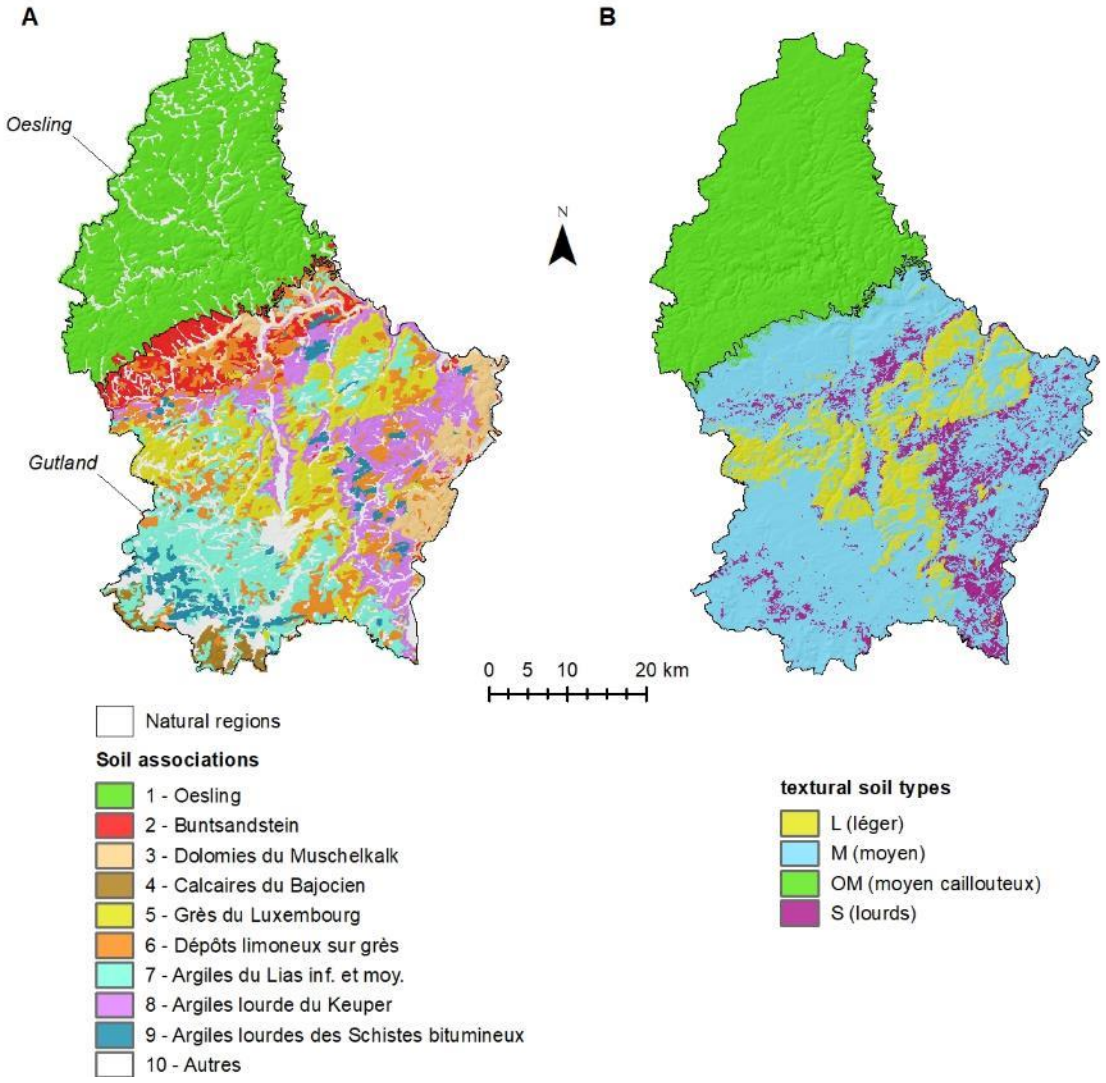


Figure 2.3: Maps of soils of GDL classified by A/ regrouped soil associations (mineralogy / geology) and B/ textural properties.

Table 2.2: 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

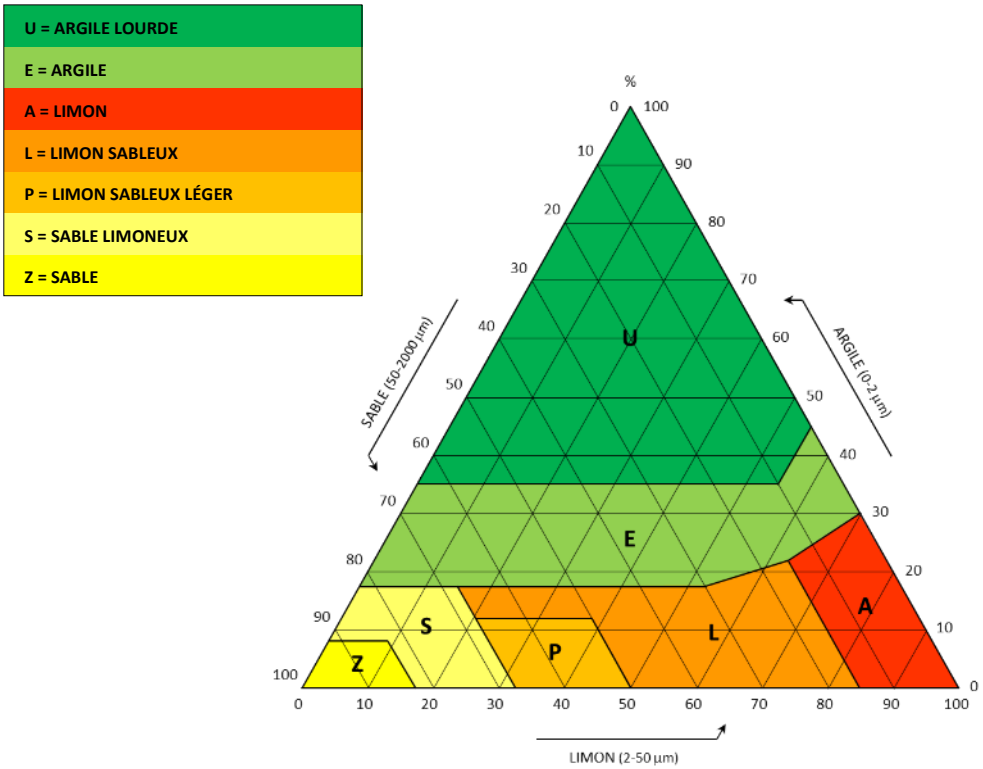


Figure 2.4: Texture triangle designed for soils of Belgium and Grand-Duchy of Luxembourg.

Land use

According to the simplified LANDUSE2018 layer (https://data.public.lu/; Fig. 2.5), 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.

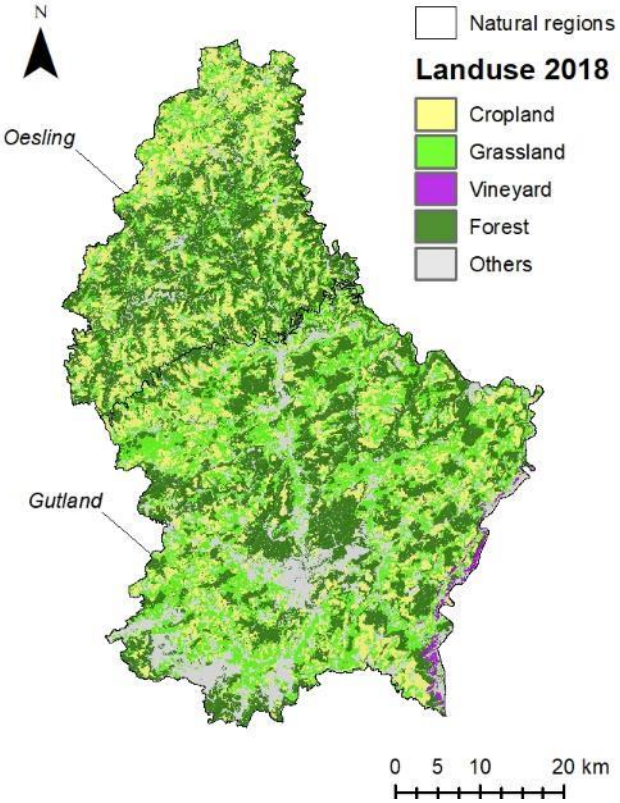


Figure 2.5: Map of land use types in GDL (after reclassification of 'Sub-type 1 LU classes'; https://download.data.public.lu/resources/landcover-landuse-2018/20200504-135337/lisl-landuse-2018-documentation.pdf).

Agriculture and management practices

The most common crop rotation in the Oesling area is a 6 to 8 years 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 (such as Common Agricultural Policy – CAP) have been created to encourage farmers to apply environmentally-friendly farming techniques (referred to as 'Good Agricultural Practices' or 'GAP') that go beyond legal obligations, as the Agri-Environment Measures[1] and the Greening initiative[2] from the EU. So far, the most common GAP's applied and promoted in cropland throughout the GDL are cover crops, reduced or no-till strategies, and temporary grassland.

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

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

3 SOIL ORGANIC CARBON AND ENVIRONMENTAL DATASETS

This section presents the main data on which this research project is based. The data covers observed SOC, land use, management practices, and environmental parameters influencing SOC content (covariates). Point observations of SOC (with their respective land use and coordinates) were gathered for three periods. Spatial layers of land use and management practices were mainly available for recent years only. Most of the covariates used in the exploratory analysis and mapping protocols were considered constant for the study period, i.e. ~110 years.

3.1 SOC DATA

Different sources of SOC or humus data were used, corresponding to three distinct periods: 1900-1914, 1964-1973 and 2012-2019 (Table 3.1).

*Table 3.1: description of the SOC databases available for recent and historical SOC evolution in GDL (*landuses and **SOC observations used in this study after dataset cleaning and filtering)*

Period	Name	Sources	Sampling	Landuses studied*	SOC observations used**	SOC analysis	Additionnal data
1900-1914	BDAT_1900	Bodenanalyse1900	Topsoil	Cropland	3177	Humic acid	fractions of stones and fine earth + lime (chaux) + phosphoric acid
1964-1973	BDSOL_1964-1973	BDSOL for 1964-1973 + Thesis data from Vermeire (1967)	Histoical profiles and topsoil survey	Cropland, Grassland	555 horizons (with 160 profils included)	Modified Walkley-Black	stone content + texture + pHs + C/N + CaCO3 + P2O5 + K2O + Mg (+ Fe2O3 + CEC)
2012-2019	BDAT_2012-2019	BDAT for 2012-2019	Topsoil	Cropland, Grassland, Vineyard	11,819	Dry Combustion	pH_CaCl2 + P2O5 + K2O + Mg + Na
2009-2019	BDSOL_2009-2019	BDSOL for 2009-2019	Recent profiles	Cropland, Grassland, Vineyard, Forest	43 profils (for total of 155 horizons)	Dry Combustion	stone content + texture + pHs + P2O5 + K2O + Mg +Na + Bulk Density + pFs…

N.B.: during the previous study (Stevens et al., 2014a), data for forest from the National Forest Inventory (NFI; period 1998 - 2000) were used. As the database from the NFI was not updated at the time of this research, forest soils were not included.

3.1.1 Early 20th century SOC data: 1900 – 1914

Source

Before the second world war, the GDL had only one type of soil map consisting of land registers established between 1900 and 1914, covering the majority of the communes, where the location of the sample points for soil analyses made for assessing needs on Thomas slag were georeferenced. The sampling sites were ***croplands*** and samples taken from the ***topsoil only*** (A. Thull, 1939). But it's not clear if it was a point sample or a composite sample of a certain area. Amongst the parameters analyzed back then we found the content of stones, fine earth, and what they called 'humus' and plant nutrients. At the time of writing, 4138 sampling sites were identified, georeferenced, and their analysis results reported in a numeric database (called 'Bodenanalyse1900'). Amongst them, 3225 were analyzed for 'humus' content.

SOC analysis

The method described by Aschman et Faber (1899; see Annex 8.1) used potassium permangate to oxidize the organic matter. More precisely, this method used the solubility of the humic acid(s) in a basic medium and their titration with potassium permanganate to determine the C content of the humic acids (Annex 8.5). No correction was applied in order to obtain the total SOC or SOM contents (SOM being composed of humic and non-humic substances). Therefore, it seems that back then they assessed the organic status of soil by measuring their content in humic acid which is the major component of humus.

Database preparation

Different steps of filtering were applied to the original 'Bodenanalyse1900' database:

1. Removing the samples with no humus values;
2. Removing the samples which coordinates where outside the PCN limits (Plan Cadastral Numérisé; https://data.public.lu/fr/datasets/plan-cadastral-numerise-pcn/);
3. Removing the samples with humus < 1 $g.kg^{-1}$ (minimum that could be weigh on 20-30g at this time);
4. Removing the samples from hydromorphic sites, after PCN (i.e., potential permanent grassland);
5. By soil association, removing right-skewed data (filtering extreme outliers)[3].

A total of 3177 samples were kept. This database was named BDAT_1900.

3.1.2 SOC data for 1964 – 1973

Sources

The SOC data for the period 1964-1973 came from two sources, both relating to soil profile analyses taken in many different types of land use: i/ the ASTA historical profiles database related to soil survey, and ii/ the profiles database from thesis work of Vermeire (1967). Both databases contained measures for profiles sampled by horizons including their geographic coordinates which have been verified and corrected when needed. The ASTA historical profile database contains ~2430 SOC observations. In addition, data on stone content and texture (clay, silt and sand fractions), pH (H_2O, $CaCl_2$, KCl), C/N ratio, and contents of $CaCo_3$, P_2O_5, K_2O, Mg (more rarely CEC and Fe_2O_3) are available. The Vermeire's database contains 176 observations relative to 33 profiles. Data on texture (clay, silt and sand fractions), pH (H_2O and KCl), CEC, and $CaCo_3$, and $FeSO_3$ are also available in this database.

SOC analysis

The 'Mémoire de méthodes d'analyses' by P. Gillen (year unknown), chief chemist in the soil laboratory from 1949 to 1973, described the analytical methods used in the 1960s in the laboratory from the soil service of the 'Station de Chimie Agricole' of Ettelbruck. According to Gillen, the SOC content was then measured by a modified Walkley-Black method, i.e., a classical Walkley-black protocol with an additional heating to complete the oxidation.

They applied the analysis on one gram of air-dried soil. The Organic Matter (OM) was oxidized by Potassium dichromate and sulfuric acid. The additional heating aimed to reach a temperature of 175 °C within 90 seconds. After cooling, the concentration in Cr^{3+} ions was determined by colorimetry. Finally, SOC content was computed based on remaining Cr^{3+} concentration. Gillen precised in his mémoire that

[3] *All data superior to Q3 + 3*IQR (with Q3 = 3rd quartile and IQR = InterQuartile Range Q3 – Q1).*

this modified version of Walkley-Black method oxidized 86.9 % of the OM. We supposed that the correction factor of 1.15 was applied to obtain the final data in the database.

Database preparation

Few steps of filtering were needed here. Both the ASTA historical profiles and Vermeire's database were filtered by:

1. Removing the observations with no geographic coordinates;
2. Removing the observations where SOC was not measured;
3. Removing the samples for which the coordinates were outside the PCN limits (Plan Cadastral Numérisé; https://data.public.lu/fr/datasets/plan-cadastral-numerise-pcn/);
4. Checking information about the types of land use, keeping only observations related to cropland and grassland.

A total of 555 observations were kept (including profiles and topsoil surveys). The database was named BDSOL_1964-1973.

3.1.3 Recent SOC data: 2012-2019

Sources

SOC data for the period 2012-2019 were first extracted from the ASTA database compiling all the results of analyses performed for farmers in GDL. Then, more than 300 SOC analyses were performed on soil samples from 2019 and early 2020 in order to improve the spatial cover and the number of paired observations in the database. Other soil parameters were available in the extraction: the pH ($CaCl_2$) and contents in P_2O_5, K_2O, Mg and Na. The last version of the extraction contained originally 14,837 observations (i.e., before filtering).

NB. Additional recent data - soil profiles data from BDSOL_2009-2019 (Tab. 3.1) - were used as calibration dataset for modeling parameters needed to compute SOC stocks at 0-30cm depth (§5.1.2). Indeed, this dataset contained SOC content by horizons, along with Bulk Density and stone content, for different landuses: cropland, grassland, vineyard and forest.

SOC analysis

In BDAT_2012-2019, 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 0-25 cm in cropland, 0-15 cm in grassland and 0-30 cm in vineyard. Sampling is done by farmers as part of Agri-Environment Measures (AEM) of the Rural Development Plan. Especially on grassland, sampling depth might vary between 10 and 15 cm.

Soil samples are dried at < 40°C, first sieved at 2 mm according to ISO 11464, and then gently crushed using 2 different mills from FRITSCH (Modell P8, Modell P13). 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 spectrometer. 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 CN 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., 2014a).

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 statistically 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 3.2; u in relative percentage %): four non-carbonated samples (*u* = 10%) and three carbonated samples (*u* = 15%). 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 concluded 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 3.2: 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	10 %	15%

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 was defined as the centroid of the fields[4]) as well as the crop grown on the related field for each year of the period 2008-2019[5]. 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:

[4] *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.*

[5] *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).*

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

After the cleaning and the filtering of the extracted raw data, the BDAT_2012-2019 dataset contained 11,819 observations, including 4820 for the period T1 (2012-2015) and 6999 for the period T2 (2016-2019; Tab. 3.3). 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. Replacing the duplicates, by year first and then by period, by their mean values diminished the dataset by 1761 observations.

Table 3.3: Filtering steps on the LU-SOC-map database preparation and associated numbers of observations.

Filtering step	Total Obs.	Obs. eliminated
None	**13871**	-
- 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

3.2 LAND USE AND MANAGEMENT DATA

Landuse

The datasets presented above (section 3.1) contained land use information for the period of sampling, along with geographic coordinates or FLIK code (present field identifier; see 3.1.3). Hence, we were able to compare the land use of the sampling period to the recent one (LPIS layers), and to identify the fields which have undergone conversion since then.

Fields geometry

We also paid attention to the potential evolution of the fields' geometry associated to the soil samples by using information from the PCN (Plan Cadastral Numérisé; https://data.public.lu/fr/datasets/plan-cadastral-numerise-pcn/) where parcels delimitations are based on historical plans from 1824.

[6] *All data superior to Q3 + 3*IQR (with Q3 = 3rd quartile and IQR = InterQuartile Range Q3 – Q1).*

Nowadays fields often consist of different neighboring parcels as different land consolidation schemes occurred since 1824.

Good Agricultural Practices

Different Good Agricultural Practices (GAP) consist of cover crops (CC), reduced tillage (RT) and temporary grassland (TG).

- *Cover crops and reduced tillage* both constitute GAPs introduced in the Agri-Environmental Measure (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 CC or RT separately, or combined. Operational since 2015, the Greening Initiative supports the sustainable use of farmland through CC 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).
- *Reduced tillage* aims to reduce intensity of tillage operations, and may over time evolve into to stopping tillage completely (no-tillage). These practices result in different 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). In Luxembourg, reduced tillage is limited to conservation tillage (Mulchsaat) by switching from plow to cultivator or disk harrow and locally strip till; but hardly any no-till (Direktsaat).
- *Temporary grasslands* induce positive residual effects on the next arable crops, increasing soil fertility and reducing crop diseases and weed infestation (Panattieri et al., 2017; Viaud et al., 2018). The European Commission makes a clear distinction between temporary grassland (fields under grassland for less than or equal to 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.

Different spatial layers where provided by the ASTA or were available in the website of the Luxembourg government (https://data.public.lu/) in order to identify the fields where the GAP presented above were applied these last years. For the AEM 262-362-462 which concerns the application of CC and/or RT, layers for years 2008 to 2020 were used. The Greening Initiative concerns CC and we used declaration for 2015-2019.

3.3 Spatial covariates

The Digital Soil Mapping (DSM) approach fits a multivariate 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 rasters 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 was already used by Stevens et al. in 2014a: all the layers related to the '*relief*', '*climate'* and '*land use'* sections of Table 3.4. Most of the layers representing the *'soil'* covariates were recently updated/modified or created by Steffen et al. (2019) (e.g., Figs. 3.1 and 3.2).

3.3.1 Relief

We used the digital elevation model (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. 2.2). 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:

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

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

3.3.2 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. Meteorological data of Luxembourg weather stations from ASTA were aggregated by 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 seven stations. We combined this dataset with weather data of Belgium (Koninklijk Meteorologisch Instituut, KMI), France (Meteo France) and Germany (Deutscher Wetterdienst, DWD; Meersmans et al., 2011). Climatic maps of GDL were created by modeling temperature and precipitation with elevation using thin-plate splines regression (Stevens et al., 2014a). Using elevation 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 2.1.

Table 3.4: Description of the environmental covariates used in the SOC mapping procedure.

Covariate	Definition	Discrete (D) or Continuous (C)	Unit	Source
Relief				
elevation	elevation (height above sea level)	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., 2014a)
temp	Temperatures	C	°C	
Soil				
clay	clay content	C	%	Maps based on historical and recent soil profiles (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-Soil Department; Bah et Marx, 2016)
UF_mean	Livestock intesity	D	UF / ha	Livestock intensity of 2018 in fertilizing units per ha (UF = unité fertilisante / 85 kg N ha-1) aggregated at the farms level provided by the 'Ministère de l'Agriculture'

3.3.3 Soil

Three maps of textural classes (sand 50 µm-2mm, silt 2-50µm and clay < 2 µm; Fig. 3.1) were used based on a historical and recent multi-sources dataset (Steffen et al., 2019). Random Forest algorithms (package Rborist) were 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. 3.2) 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 analyses in calcareous soils in order to convert total carbon to soil organic carbon (ISO 10694; see section 3.2.1) on which a co-kriging with pH $CaCl_2$ was applied. The number of samples was more limited here because TIC analysis is only performed when TOC analysis is required.

The layer detailing the minimum depth of soil hydromorphy (Fig. 3.3) was developed here, based on drainage classes from the detailed soil map (1 : 25,000). This parameter corresponds to the minimum depth where physical indicators of temporary or continuous surface water saturation were observed in the soil profile (for further details concerning the methodology, please see Annex 8.2).

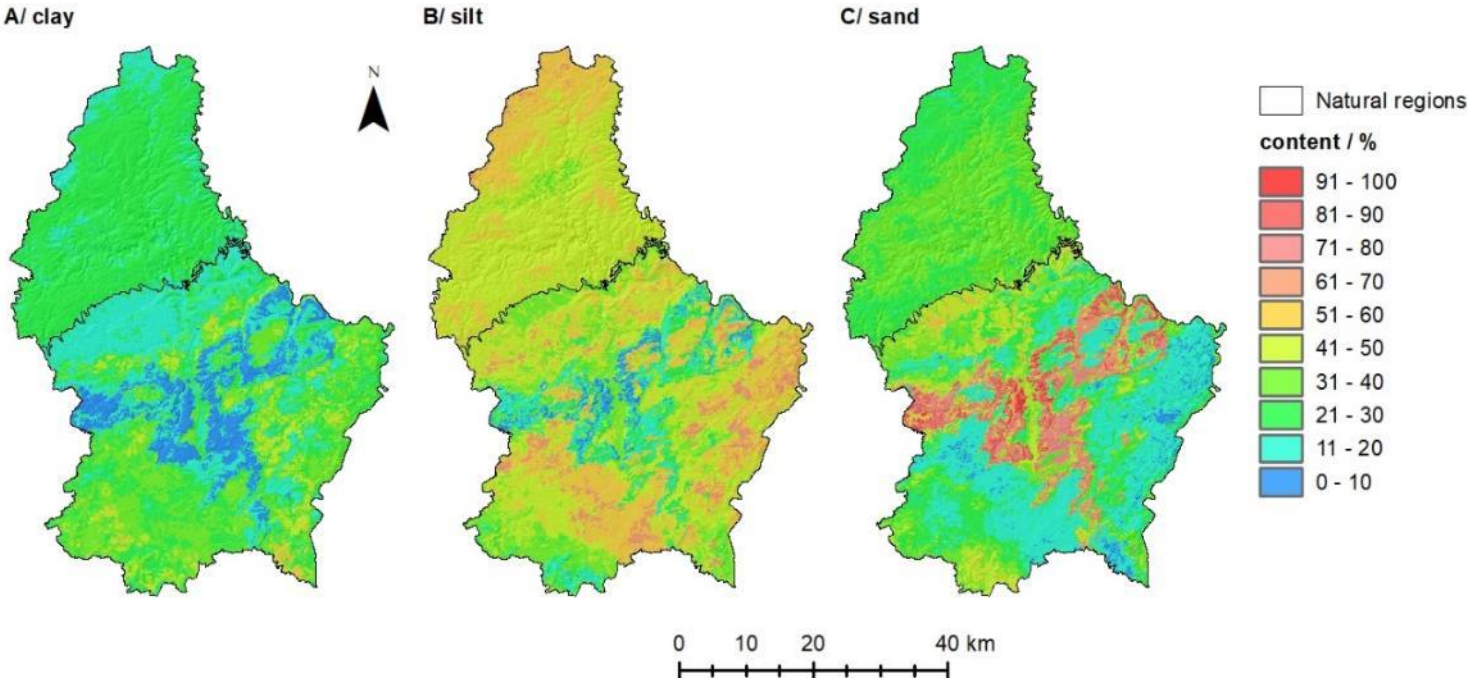


Figure 3.1: Map of A/ clay, B/ silt and C/sand content (%) in topsoil of Grand-Duchy of Luxembourg (source: Steffen et al., 2019).

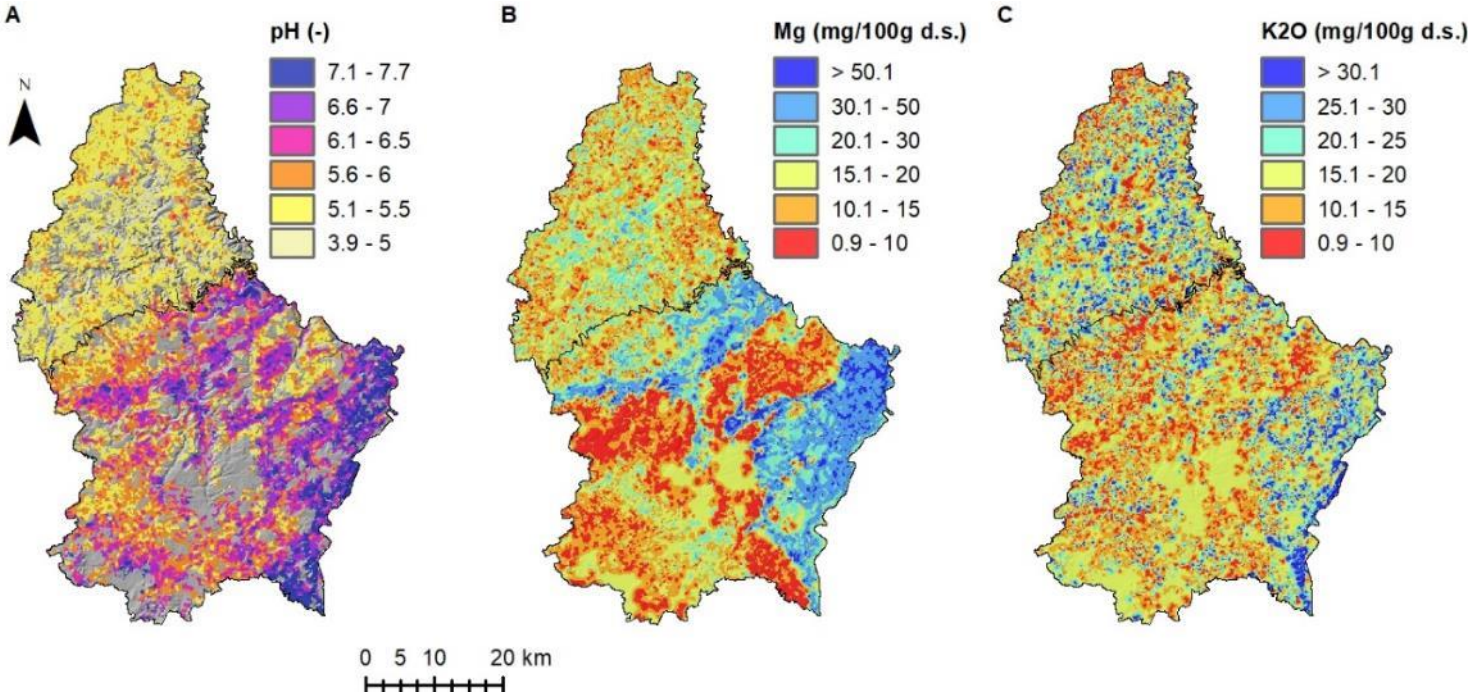


Figure 3.2: Map of A/ pH $CaCl_2$ (-), B/ available Mg (mg/100g of dry soil) and C/available K_2O (mg/100g of dry soil) in topsoil of Grand-Duchy of Luxembourg (source: Steffen et al., 2019) based on data for 2014-2018 period. The pH layer (A) covers agricultural soils only, i.e. croplands and grasslands.

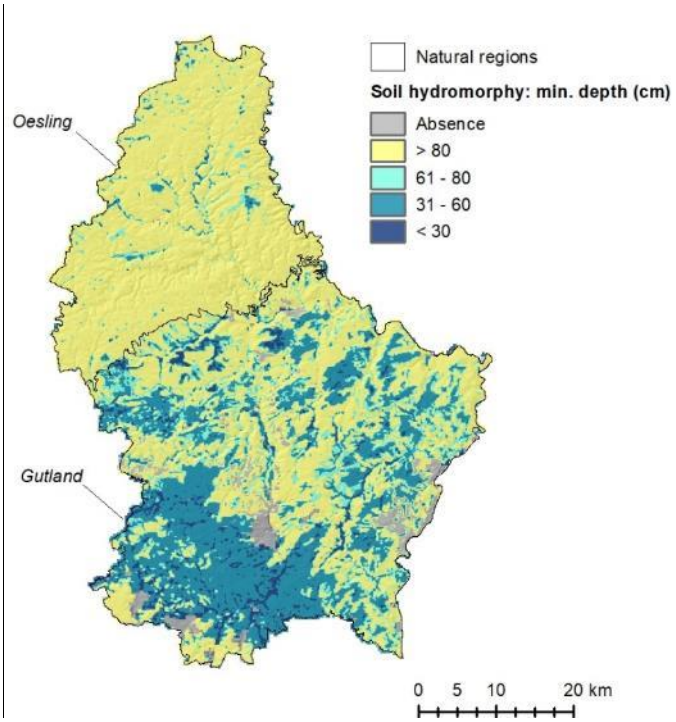


Figure 3.3: Map of the minimum depth of soil hydromorphy (cm) in Grand-Duchy of Luxembourg.

3.3.4 Land use and human influence

The land use layer used to finalize SOC maps, i.e. for applying the final models to the proper land use areas, was derived from aerial images, LiDAR and ancillary GIS data for the year 2018 (https://data.public.lu/fr/datasets/r/40288ffc-6f97-4d35-afaf-32415e1337b7). 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 2.5. 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 (Bah et al., 2016). Finally, we included the livestock density of 2018 expressed in fertilizing units per ha (UF = 85 kg N ha-1) aggregated at the level of the farms.

4 ORGANIC CARBON CONTENT IN AGRICULTURAL SOILS

4.1 HISTORICAL TRENDS IN SOC CONTENTS

The data from the three different periods (1900 - 1914, 1964 – 1973 and 2012 – 2019) were not directly comparable as the analysis protocols evolved over time (§ 3).

4.1.1 Methodology

The ASTA has compiled data describing soil organic status for three periods: 1900 – 1914, 1963 – 1974, and 2012 – 2019 (§ 3.1). These last decades, the main parameter measured to assess soil organic status is the SOC content, from which the SOM can be derived by applying the Van Bemmelen conversion factor (1.724). Today, the methodology estimated as the most effective and precise for estimating SOC is the dry combustion method, i.e. this method has the best recovery factor of ~ 100 % and follows the ISO 17025. Hence, the dry combustion is the reference methodology for SOC analysis.

Please, note that the three datasets had their specific characteristics:

- Data from 1900 – 1914 were exclusively sampled in croplands, then any comparison for grassland and vineyards were possible. In addition, the parameter describing the organic status appeared to be the humic acid content (a major component of humus), which is different from SOC or SOM content. To our knowledge, there is no specific correction factor in the literature converting acid humic content to SOC (or SOM). To finish, some areas of tenths km² were not prospected during this campaign;
- Data for the period 1963 – 1974 were associated to hundreds of soil profiles for different types of land use as cropland, grassland, vineyard and forest. After filtering the database (section 3.1.1), only 67 profiles for cropland and 93 for grassland were kept. Unfortunately, these low effectives and their sparse spatial repartition hamper to perform proper statistics test of comparison and spatial analysis;
- Data for 2012 – 2019 was the reference, as containing thousands of observations by landuse all over the GDL and SOC contents measured by dry combustion.

As the data for 1900 - 1914, 1964 - 1973 and 2012 - 2019 were not directly comparable (see above), we proposed (when possible) to: i/ analyse the influence of environmental and soil/agricultural parameters on the organic status through a Conditional Inference Forest analysis[7] and, ii/ map/project the parameter describing the organic status and proceed to a qualitative study of the evolution of the spatial patterns.

4.1.2 Results

1900 – 1914

The spatial repartition of humic acid data was very uneven, hampering a proper mapping of the parameter (Fig. 4.1). Overall the humic acid contents were quite similar at a low level of ca. 6-7 $g.kg^{-1}$ with the exception of the Oesling where the content is ca. 15 $g.kg^{-1}$ (Fig. 4.2). There was a positive correlation between humic acid content and precipitation (r = 0.49) and a negative correlation with temperature (r = -0.43). A Conditional Inference Forest analysis provided a reasonable estimate of the humic acid content using covariates available in the original archives such as KALK: $CaCO_3$ content;

[7] *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).*

SAB: lime deficit or excess and P: phosphoric acid content. Moreover, modern covariates related to climate and topography were used under the assumption that their spatial trends have remained stable. The model performed reasonably well with $R^2 = 0.55$ and elevation, precipitation and temperature as the main explaining variables (Fig. 4.3).

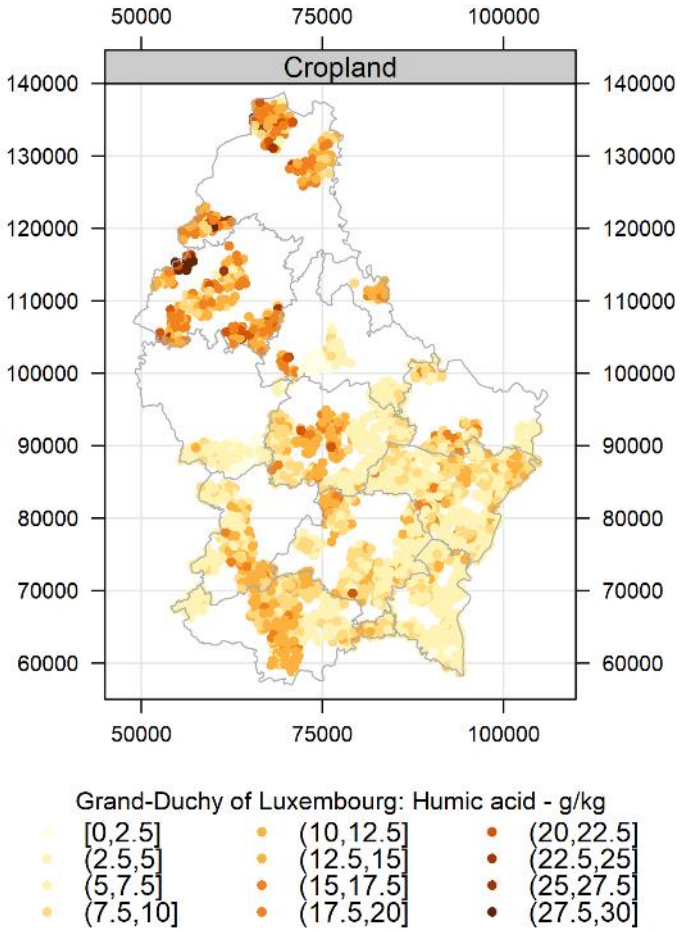


Figure 4.1: Humic acids in croplands for the period 1900-1914 (BDAT_1900).

The Oesling region can clearly be distinguished by the higher humic acid content. This pattern is similar to the one for the modern SOC content (Fig. 4.20). The higher humic acid contents in the region around Ettelbrück do not match modern SOC content patterns.

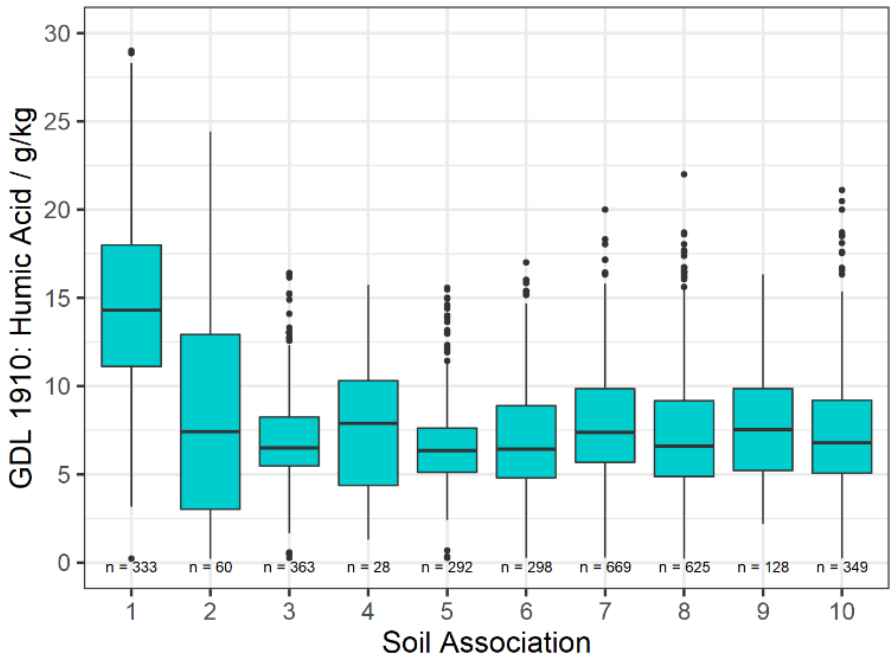


Figure 4.2: Humic acid concentration in croplands (1910-1914). The soil associations :1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)

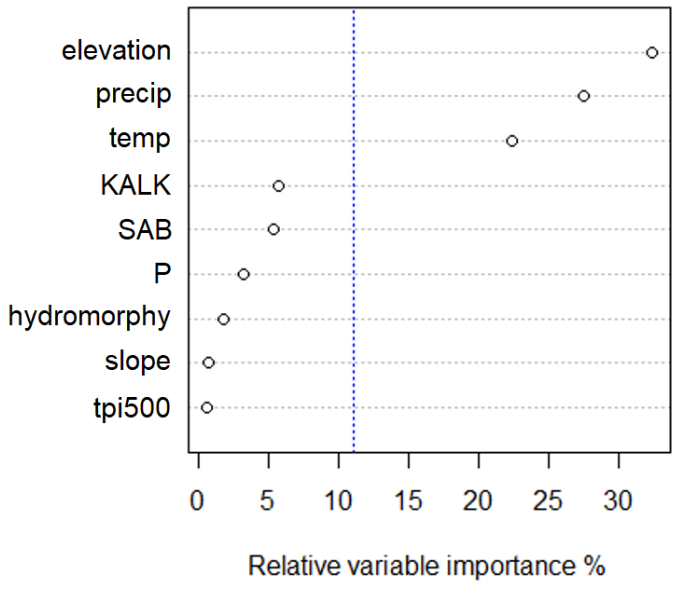


Figure 4.3: Conditional inference tree as a spatial model for humic acid content in croplands (1900-1914). The vertical blue dashed line indicates the average relative importance. KALK: $CaCO_3$ content, SAB: lime deficit or excess, P: phosphoric acid.

1963 – 1974

The dataset for SOC content in croplands and grasslands from the period 1963-1974 was much less extensive (n = 160; Fig. 4.4). A Conditional Inference Forest model gave reasonable predictions for the SOC content (R^2 = 0.58) with clay, precipitation, land use and pH as statistically significant variables (Fig. 4.5).

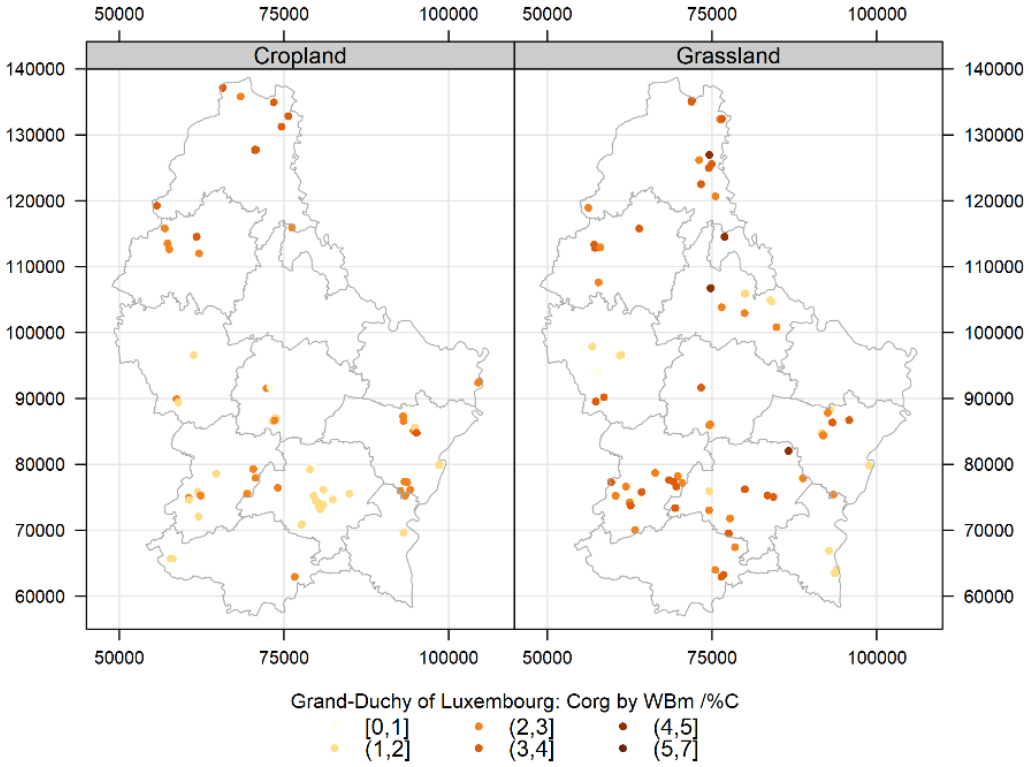


Figure 4.4: SOC content data in cropland and grassland for the period 1963-1974.

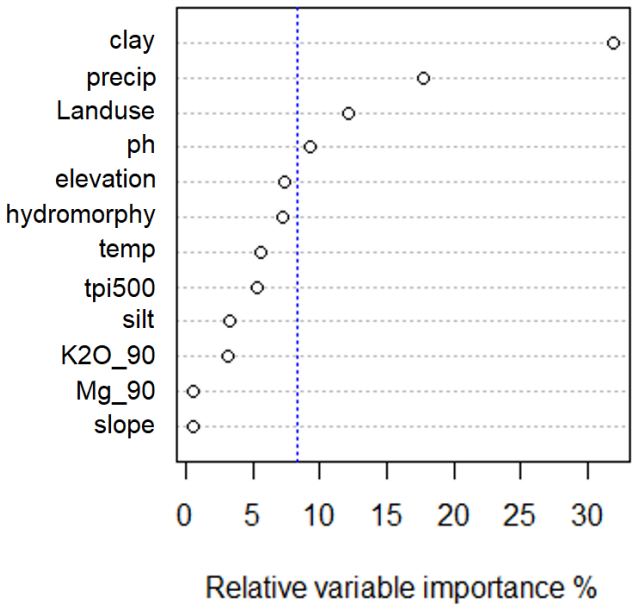


Figure 4.5: Conditional inference model for the SOC contents in 1963-1974. The vertical blue dashed line indicates the average relative importance.

Although the SOC analysis protocol was different (modified Walkley and Black) compared to the modern one (dry combustion), the trends were to a large extent comparable (Fig. 4.6). Overall, the spatial patterns in SOC did not changed over time with highest values in grasslands of the clayey soil associations (argiles du Lias, argiles lourdes du Keuper and argiles lourdes des schistes bitumineux), in the Oesling and in the Dolomites of Muschelkalk. The trends between soil associations in cropland were the same but at slightly lower values compared to the grasslands. Overall, SOC contents in croplands had slightly decreased over time, while for grassland there was a slight increase in the clayey soil associations.

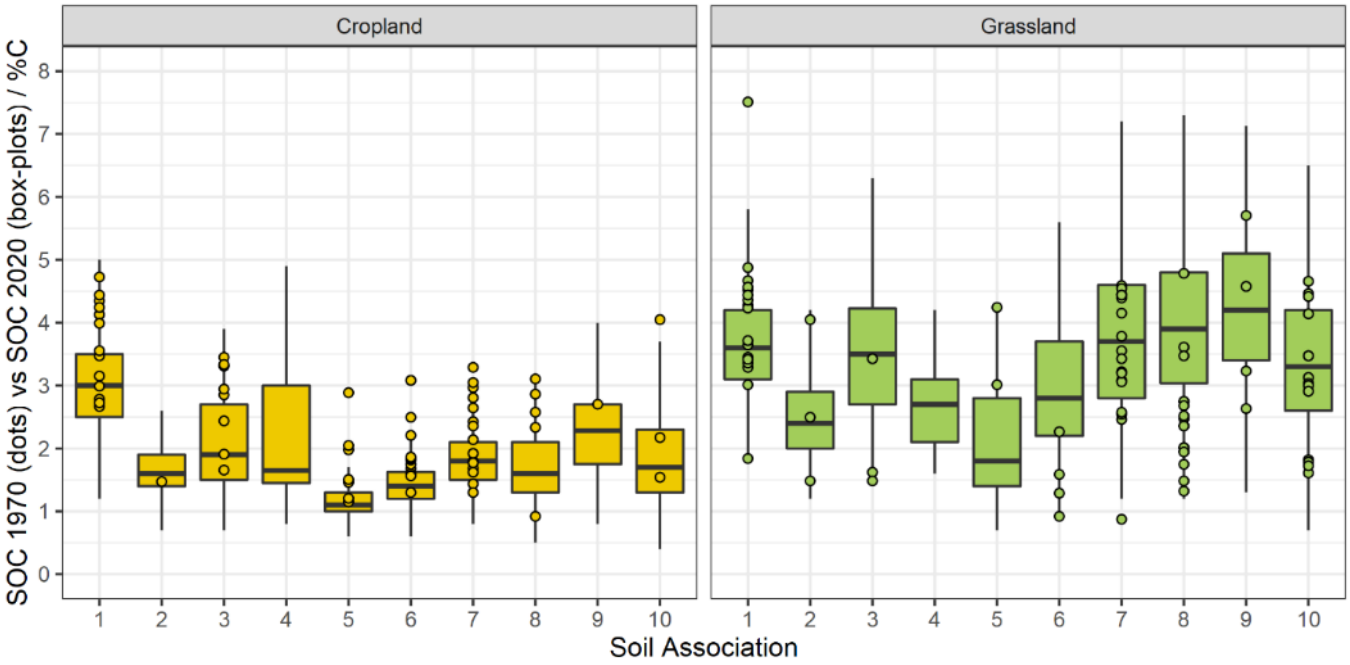


Figure 4.6: Individual points refer to SOC contents for 1963-1974, while box plots refer to SOC contents for 2012-2019. (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)

4.2 Current state of SOC content in GDL and its short-term evolution

4.2.1 Context

In the context of the second pillar of the Common Agricultural Policy (CAP), 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 here was to assess how the SOC content has evolved 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 RDP (T1: 2012-2015; considering a 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 (§ 3.2 and 3.1). 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 land use, i.e. cropland, grassland and vineyard. Consequently, the results are also reported by land use.

N.B.: during the previous study (Stevens et al., 2014a), 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.

4.2.2 Methods

Spatial modelling

Attributing the values of the covariates (independent variables; § 3.2 and 3.3) with a sampling unit corresponding to a field to SOC observations (dependent variable; § 3.1) of the BDAT_2012-2019 dataset was first required to develop the spatial models and produce maps. The mean values of the covariates for each field were computed and attributed to the corresponding observations through their FLIK number. Then, a model was calibrated for each land use, i.e. a total of three models were fitted, and then validated by periods (T1 and T2). The models fitted here were GAM (Generalysed Additive Models). The detailed protocol is depicted in Annex 8.3.

Comparing SOC at two periods

- *Observed data*

Here, we proceeded to test the differences of statistical distribution between different subsets of BDAT_2012-2019. The enlarged analytical uncertainties (Tab. 3.2.; § 3.1.3) did not influence the test results because these uncertainties apply to each SOC observations of BDAT_2012-2019, and therefore has no effect on the respective spatial distribution of the subsets.

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 BDAT_2012-2019 (non-paired tests) and on paired observed data (paired tests).

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:

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

- *Predicted data*

Two SOC maps, and their respective Standard Error (SE) of prediction maps, were computed by applying the spatial modelling methodology proposed in Annex 8.3: the first map SOC_{T1} for the period 2012-2015, and the second map SOC_{T2} for the period 2016-2019 (both in %C). We assessed the significance of the predicted differences between the two periods for croplands and grasslands of GDL. For both landuses, we computed dSOC (i.e., SOC_{T2} – SOC_{T1}). Then, at each pixel, the p-value was associated to dSOC, potential significance was estimated based on a comparison between dSOC itself and the SE of prediction for both SOC_{T1} and SOC_{T2}.

4.2.3 Results

SOC by period - descriptive statistics and differences test

The land use has a major role on the topsoil SOC (Figs 4.7 and 4.8). Croplands and vineyards showed similar SOC contents, while SOC contents in grasslands were 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. 4.8). Although both histograms for cropland showed a unimodal distribution dominated by observations from Gutland (mode around 1.2 - 1.5 %C; Fig. 4.8), the sub-datasets for Oesling are visible around 2.5 – 3.0 %C (soil association 1 in Fig. 4.8). 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. 4.9 and 4.10), especially in the natural region of Oesling. While grasslands were represented by almost twice as much observations for period T2 than for T1 (Fig. 4.9), 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.

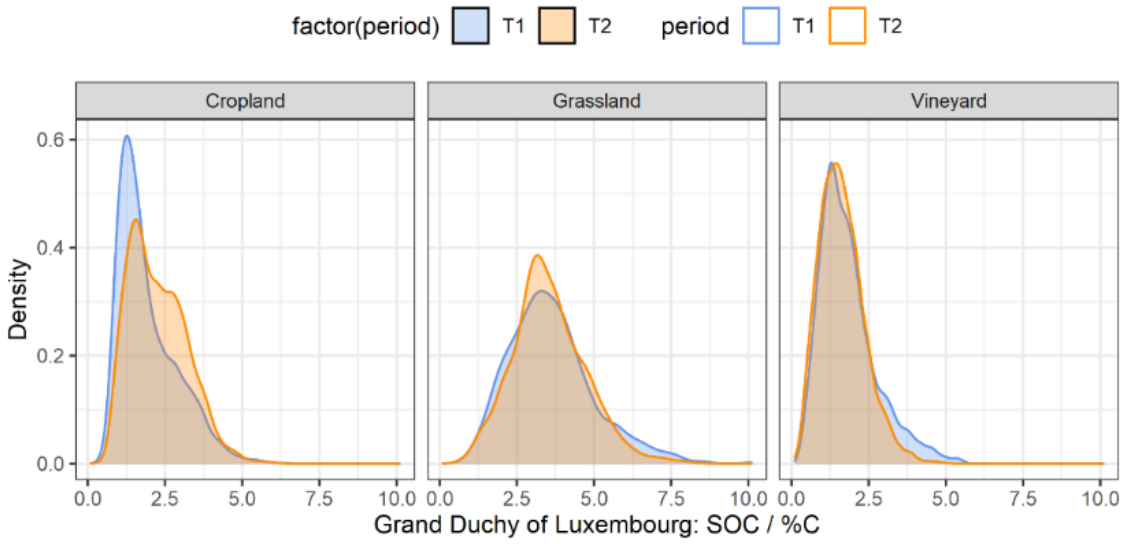


Figure 4.7: Histograms of topsoil SOC (%C) at T1 (2012-2015) and T2 (2016-2019) for croplands, grasslands and vineyards in Grand-Duchy of Luxembourg.

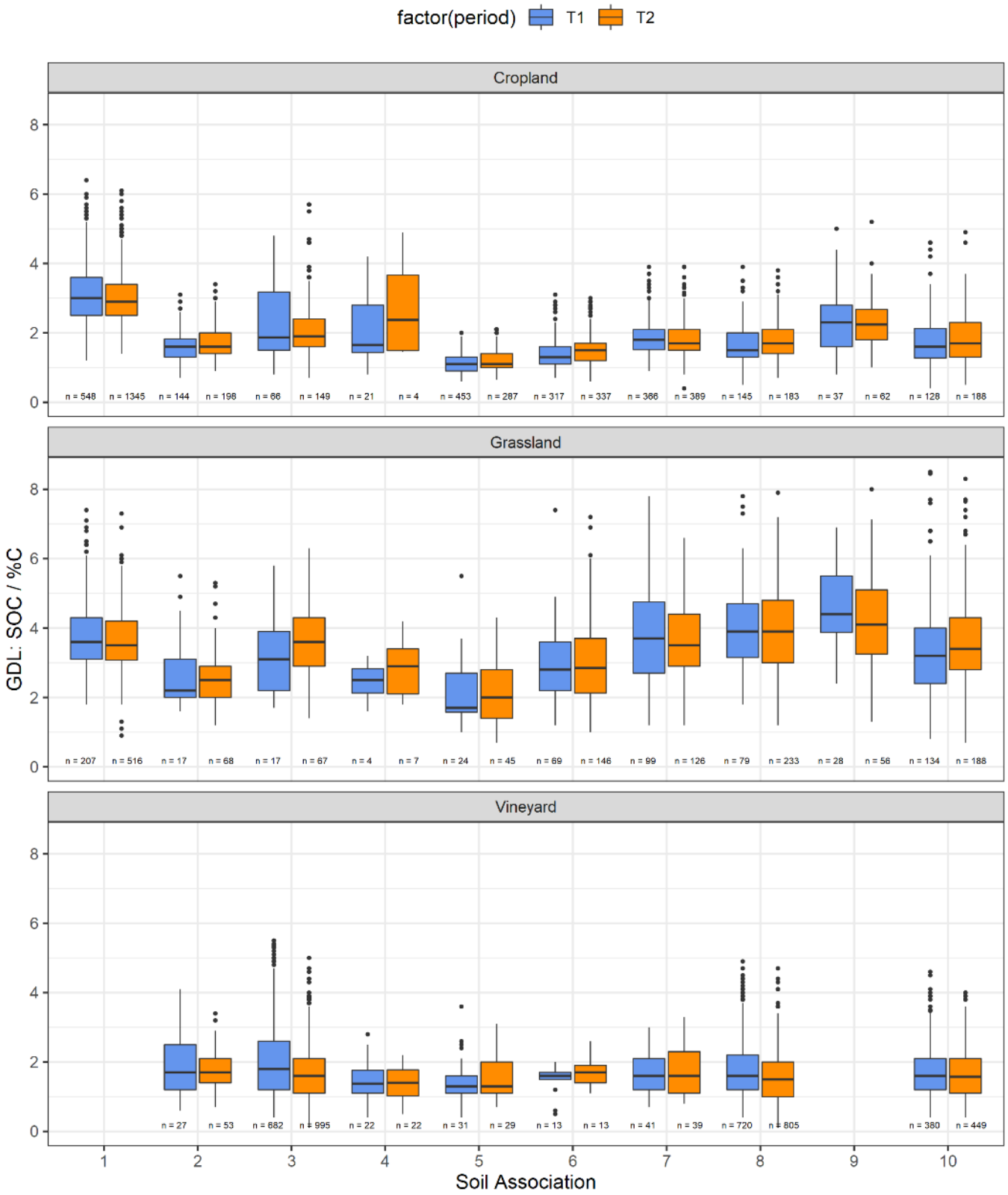


Figure 4.8: Box-plots of topsoil SOC (%C) in croplands, grasslands and vineyards per regrouped soil associations at periods T1 (2012-2015) and T2 (2016-2019) in Grand-Duchy of Luxembourg. (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 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).

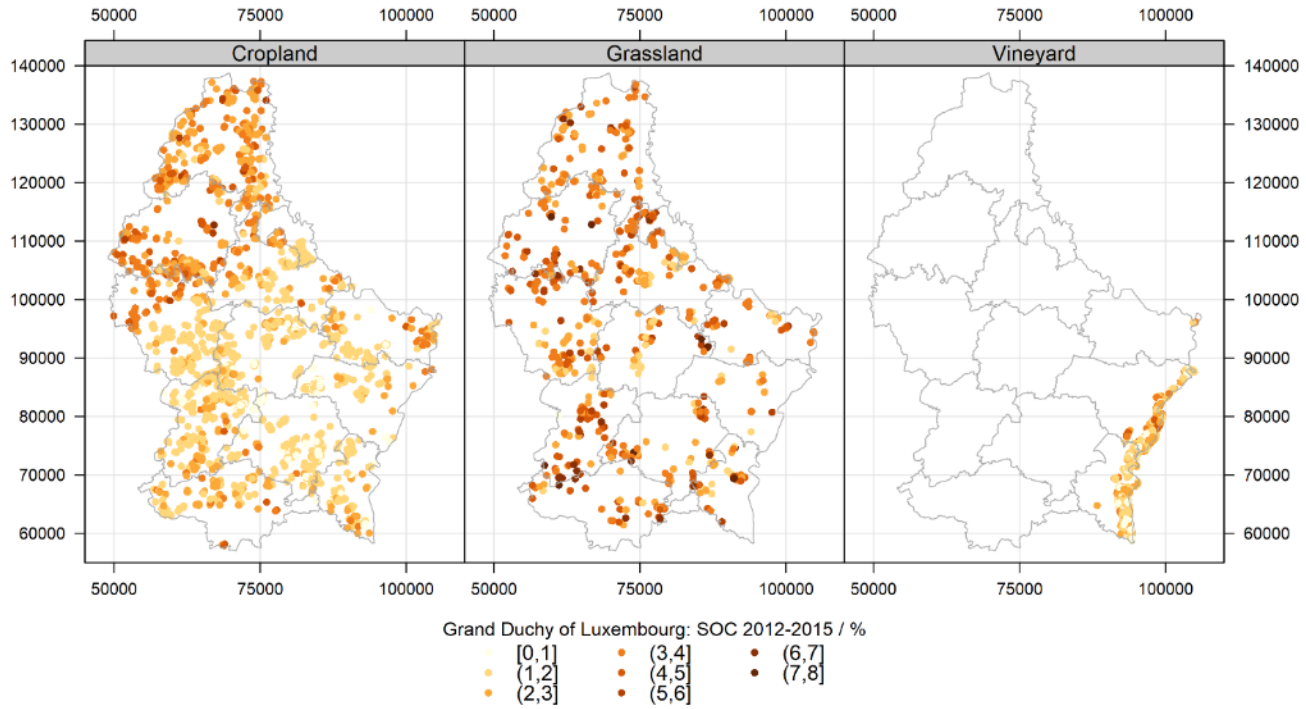


Figure 4.9: Observed SOC values (%C) of topsoil under cropland, grassland and vineyard for period T1 (2012-2015).

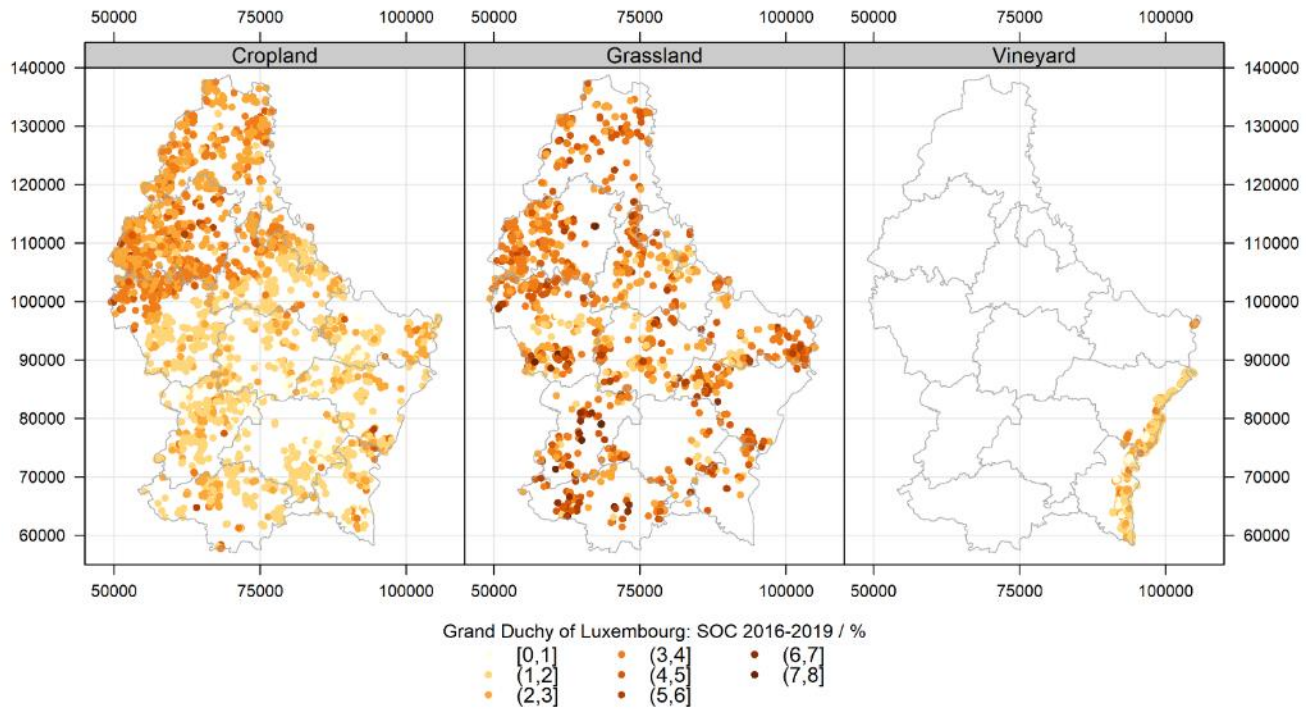


Figure 4.10: Observed SOC values (%C) of topsoil under cropland, grassland and vineyard for period T2 (2016-2019).

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 4.2 – 4.4) to define indicative SOC range in each soil association-land use 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.

NB1: In this section, the differences of statistical distribution between different datasets were tested (Tables 4.2 to 4.4). The analytical uncertainties did not influence the test results (see § 4.2.2).

NB2: Some subsets contained less than 30 individuals preventing the determination of proper statistics; i.e. description of the distribution. The statistics for these subsets are depicted in gray italics in the next tables.

- *Cropland*

Table 4.2 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 2.4), 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 8.4). 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 value 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 statistically statistically 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 statistically 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 4.2: 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
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 4.3 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 4.8, 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 statistically significant differences in SOC content between T1 and T2 (mean of +0.26%C; p<0.05; Tab. 4.3).

*Table 4.3: Descriptive statistics of topsoil SOC (%C for the 0-15cm 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). (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
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 4.4 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 1/100.000 scale map of soil associations (Fig. 2.3).*** Soils under vineyard have about the same median content and interquartiles differences as cropland soils of GDL (Tab. 4.2). 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 statistically 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). This negative trend is partly explained by a better assessment of the TIC (Total Inorganic Carbon) since 2018, revealing that the TIC was previously underestimated by ~0.2%C.

*Table 4.4: 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
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 BDAT_2012-2019 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 4.11 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. 4.12). 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 those groups (n < 30) should be considered robust (Tab. 4.5; results in gray italics correspond to groups under 30 sites). Moreover, we observed absolute differences >2%C for few sites, which were not likely to occurred because of climate or field management changes in less than a decade. These outliers could be the results of sampling or location errors (FLIK), or significant SOC variations within the affected fields.

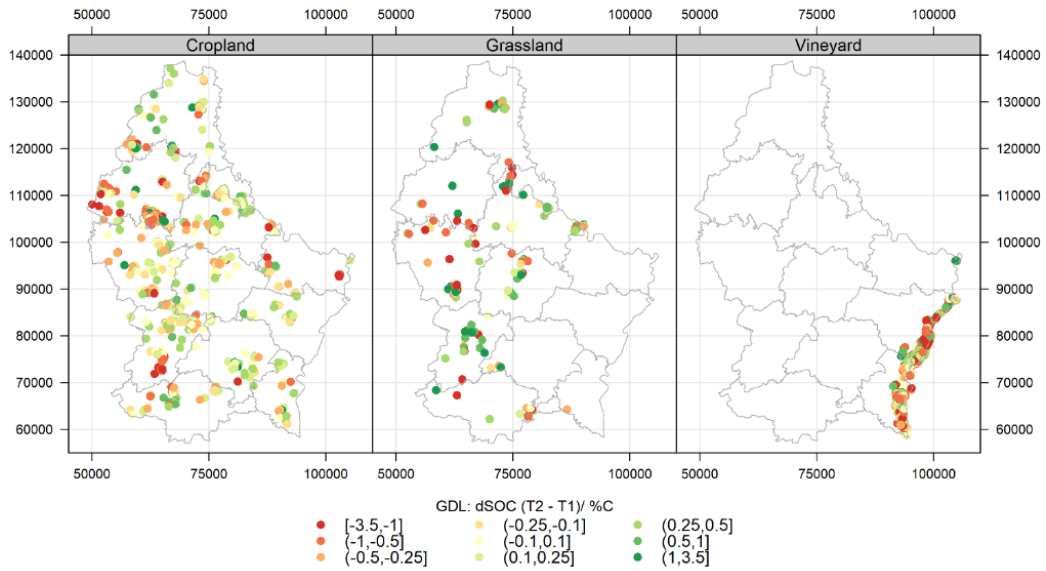


Figure 4.11: Location and differences in SOC content (dSOC in %C) between T1 (2012-2015) and T2 (2016-2019) computed on paired observations (same FLIKs) in Grand-Duchy of Luxembourg.

Tests on paired observations confirmed that a statistically 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 (Tab. 4.5). While a non-statistically significant decrease in SOC (-0.02%C; Tab. 4.2) was observed for soils on “Argiles du Lias inférieur et moyen” when considering all the data, a statistically significant decrease of -0.16%C ($p<0.05$) was detected here considering 92 paired observations (Tab. 4.5).

None of the soil associations with more than 30 sites in grasslands showed statistically significant differences in SOC between T1 and T2. However, we noticed that a statistically 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, statistically 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; $p<0.01$).

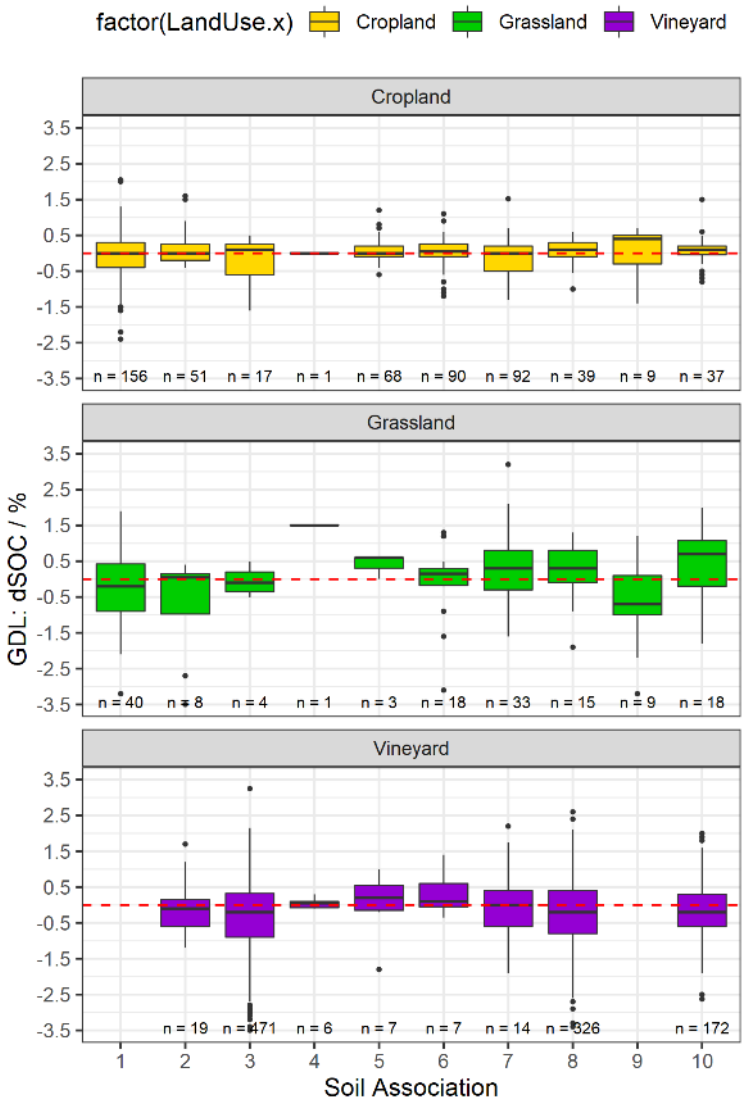


Figure 4.12: Box-plots of differences in SOC content (dSOC in %C) between T1 (2012-2015) and T2 (2016-2019) computed on paired observations (same FLIKs) in Grand-Duchy of Luxembourg.

Table 4.5: 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**

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 land use, are geographic coordinates (x: longitude, y: latitude), soil type (texture), climatic condition and elevation (z). While the geographic coordinates (x, y) support the main regional trend of relief (z) and climate, which characterize the GDL in the North-South direction, the relief and climate also support the local variations observed (e.g., induced by valley or plateaus; see § 2). In addition, we observed conspicuous relations between topsoil SOC and new *'soil'* covariates implemented in this project (Table 3.4). The latter are both influenced by natural processes and agricultural activities.

- *Cropland*

We observed a clear dichotomy in SOC level between the Oesling and Gutland (Figs 4.9 and 4.10). 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 4.13A). As observed in many other studies, soil texture, especially clay content, has an important role in SOC spatial distribution in croplands (Fig. 4.14). 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 (z) was strong (ρ=0.57). The same is true for precipitation (ρ=0.53) and temperature (ρ=-0.54; Fig. 4.13B). The two natural regions present very different geomorphologies, especially their ranges of elevation (Fig. 2.1) 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 range 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 statistically 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 hydromorphic features between 0 and 80cm showed a clear negative relation with SOC. Indeed, SOC content increased while drainage deteriorated, i.e. while hydromorphic 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. 4.9 and 4.10). Consequently, coordinates and climatic covariates showed no clear relation with SOC (Fig. 4.15A). 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. 4.15B). Clay and sand contents have a strong influence on SOC level in grasslands of Gutland (ρ = 0.55 and ρ = -0.50, respectively; Fig. 16). Similar to the croplands, minimum depth of hydromorphic features < 80cm depth showed a clear negative relation with SOC.

- *Vineyard*

Vineyards in GDL are concentrated in the Mosel river valley characterised by rather homogeneous environmental conditions, which can explain the poor correlations between SOC and environmental covariates (Fig. 4.16). 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, (ρ = 0.20).

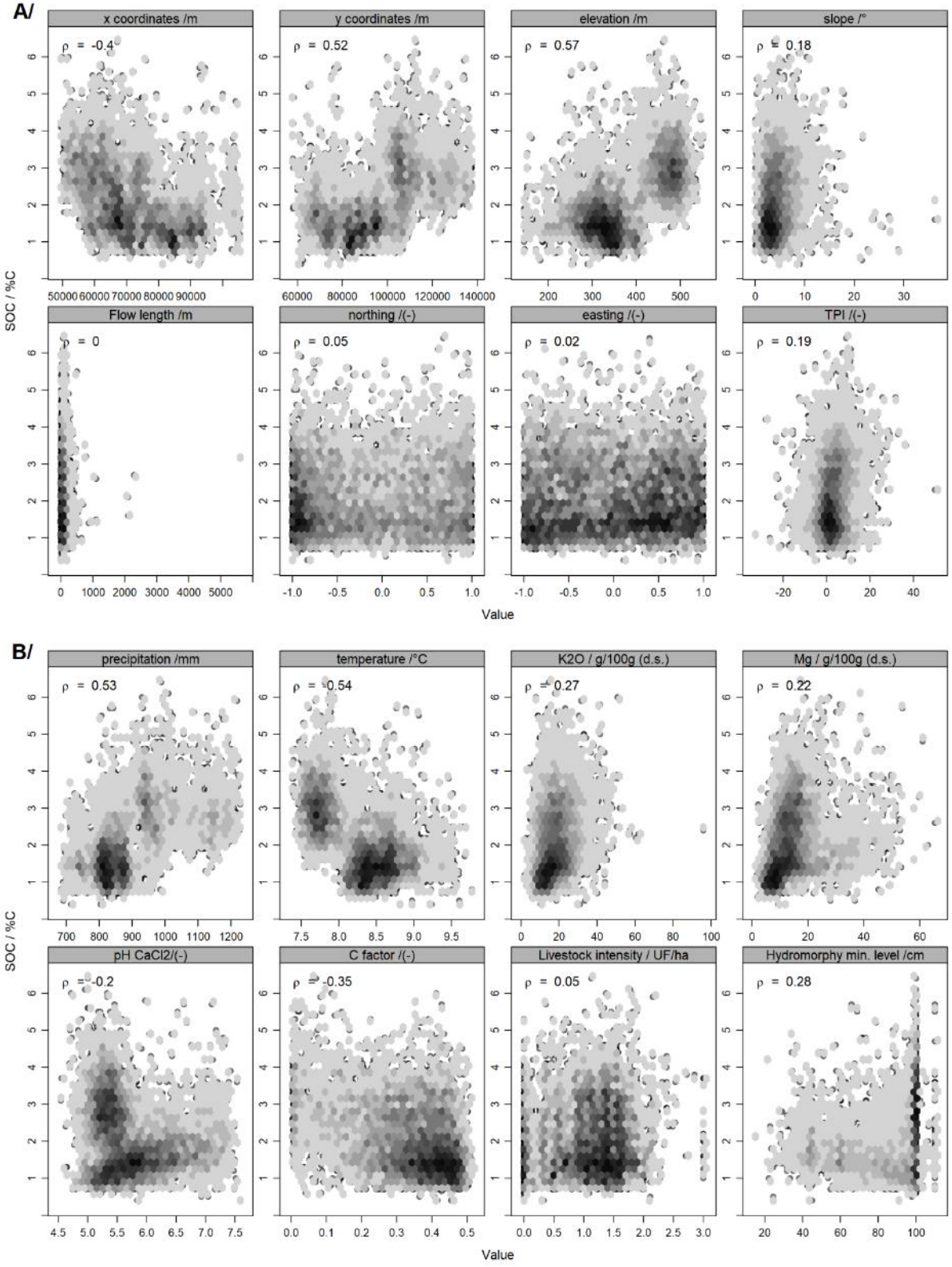


*Figure 4.13: Scatter plots of topsoil SOC (%C) **in croplands** as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and landuse parameters.*

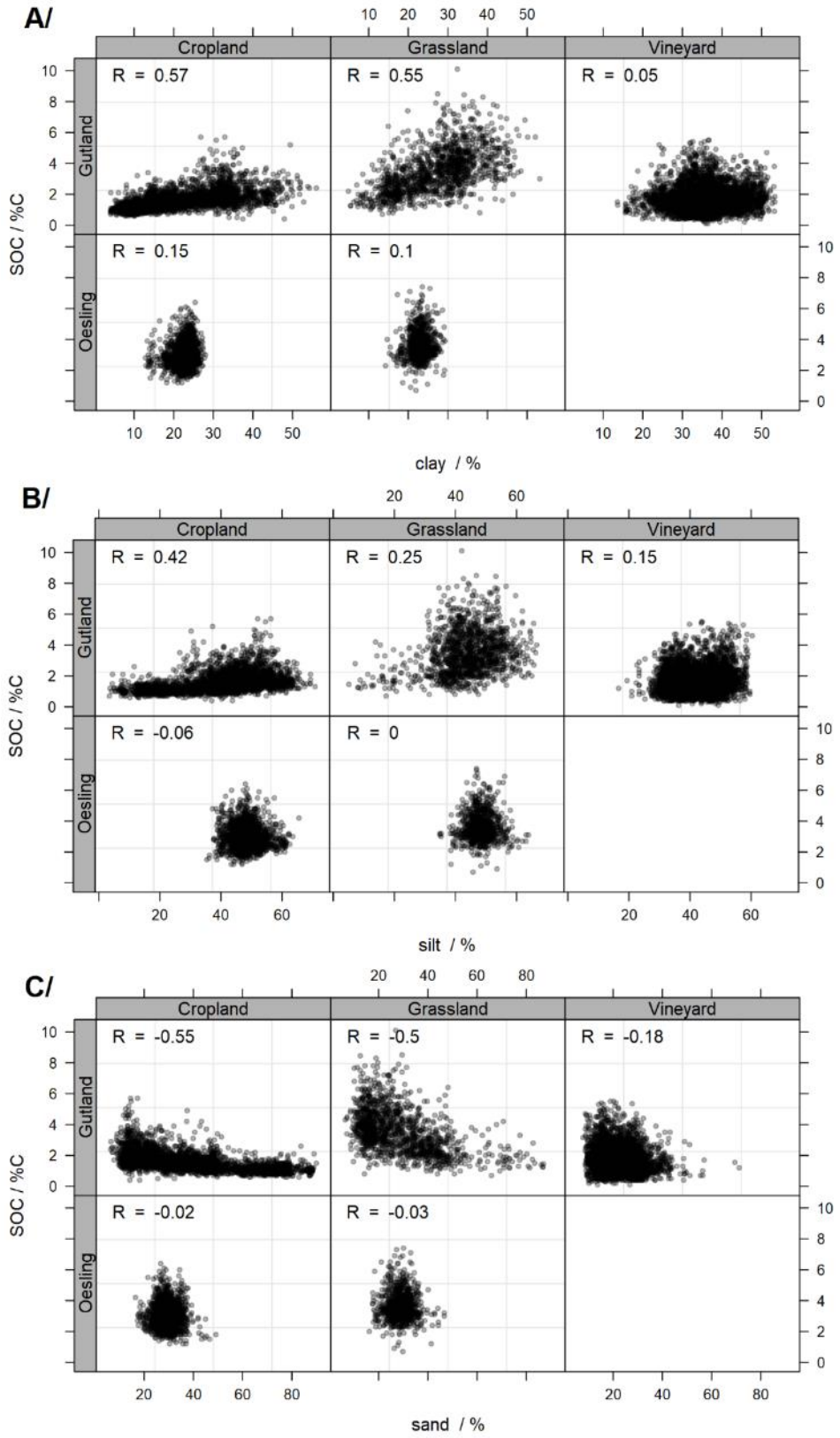


Figure 4.14: Scatter plots of topsoil SOC (%C) by land use and natural region as function of A/ clay content (%), B/ silt content (%) and C/ sand content (%).

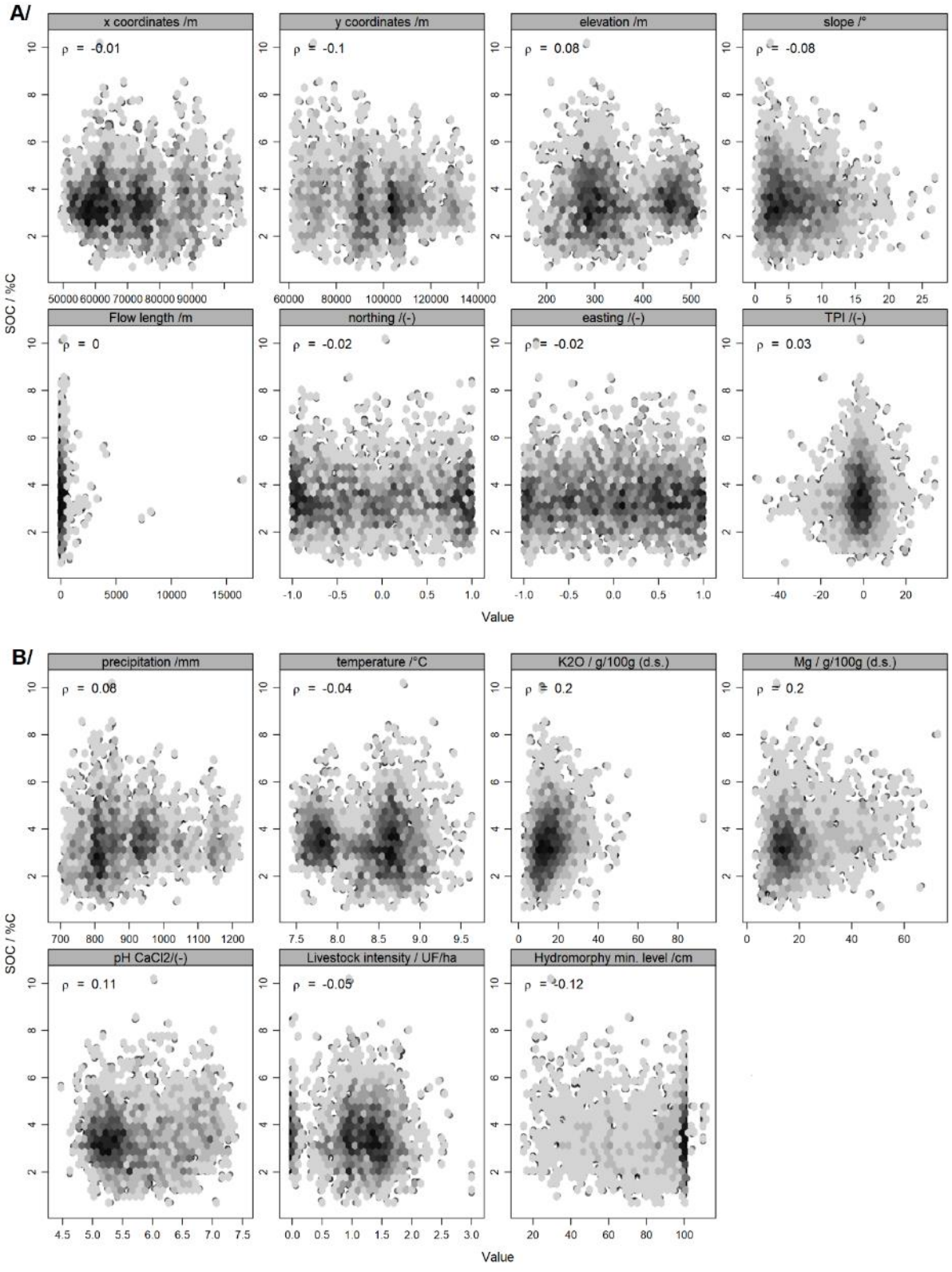


Figure 4.15: Scatter plots of topsoil SOC (%C) ***in grasslands*** *as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and land use parameters.*

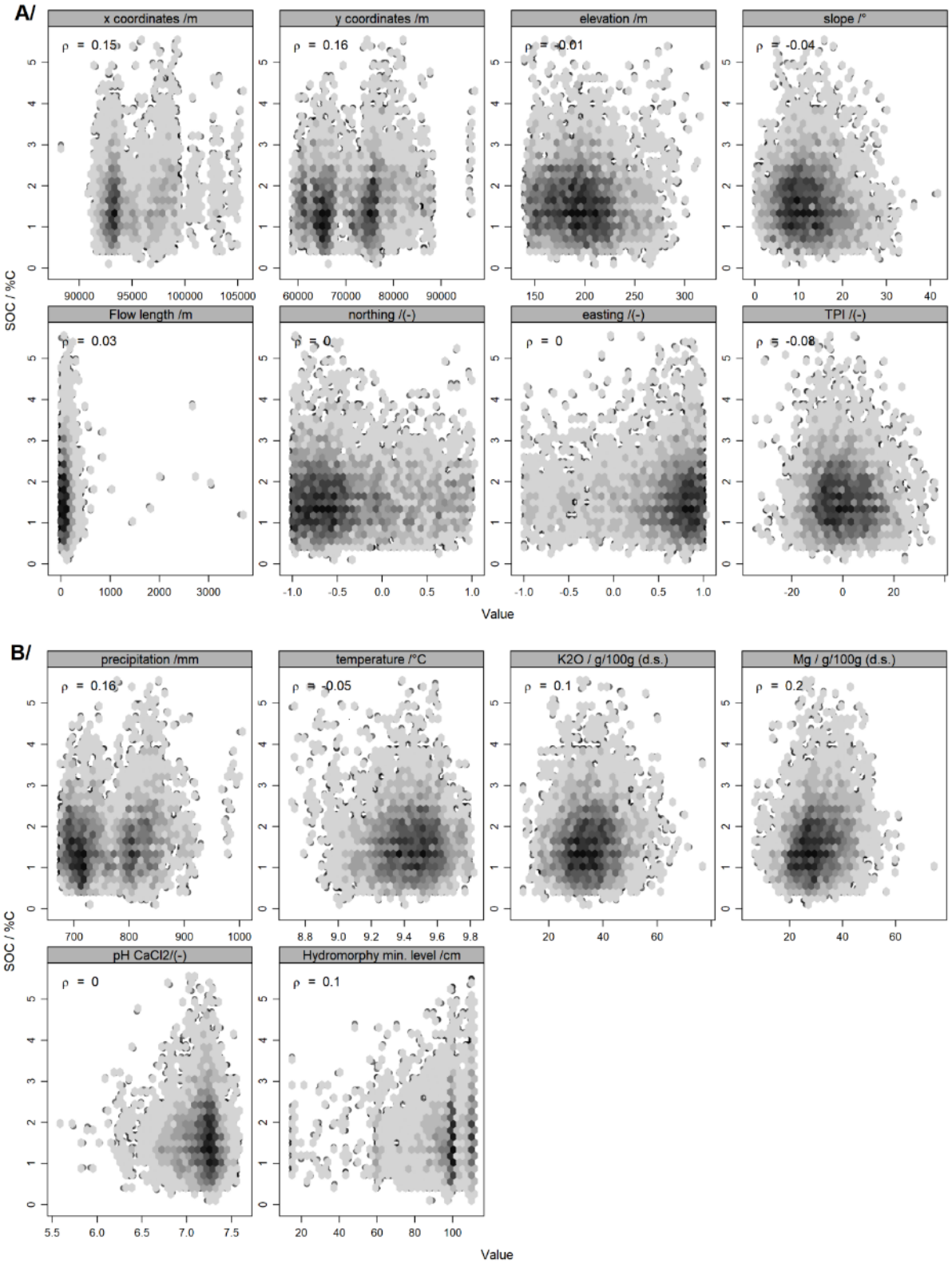


*Figure 4.16: Scatter plots of topsoil SOC (%C) **in vineyards** as function of A/ geographical coordinates (x and y in m) and relief parameters and, B/ climate, soil and land use parameters. (Minimum level of soil hydromorphic features was homogeneous all over the vineyard dataset)*

The semivariograms of the SOC observations show very different spatial structures between the different landuses (Figure 4.17). 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 difficult to model.

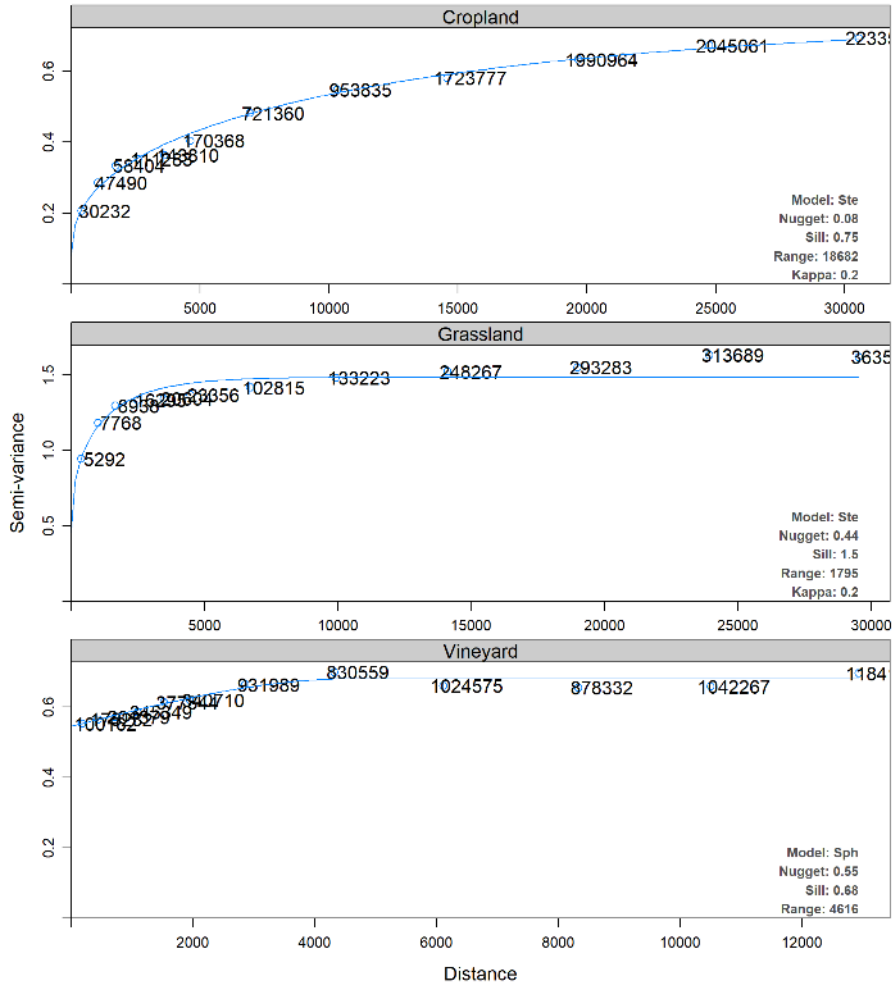


Figure 4.17: Semi-variograms of SOC under cropland, grassland and vineyard in Grand-Duchy of Luxembourg.

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 land use, 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 land use-period subset. Before fitting, a pre-selection of

covariates considering their relations, especially collinearity (concurvity[8] is managed in the GAM procedure; see §8.2) 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 land use classes. Figure 4.18 presents the point plots between observed and predicted data produced during the stratified 10-fold cross-validation for each land use 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), the clay content, the C-factor, the Mg content, the K_2O content, the minimum depth of hydromorphic features, the slope and the elevation (here in decreasing order of importance; Fig. 4.19). 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 Figure 4.18 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 periods, 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 hydromorphic features, 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. 4.18). 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;

[8] *Concurvity occurs when some smooth term in a model could be approximated by one or more of the other smooth terms in the model.*

ii/ a non-negligible part of present grassland fields were converted from cropland during the first half of the 20th century. 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.

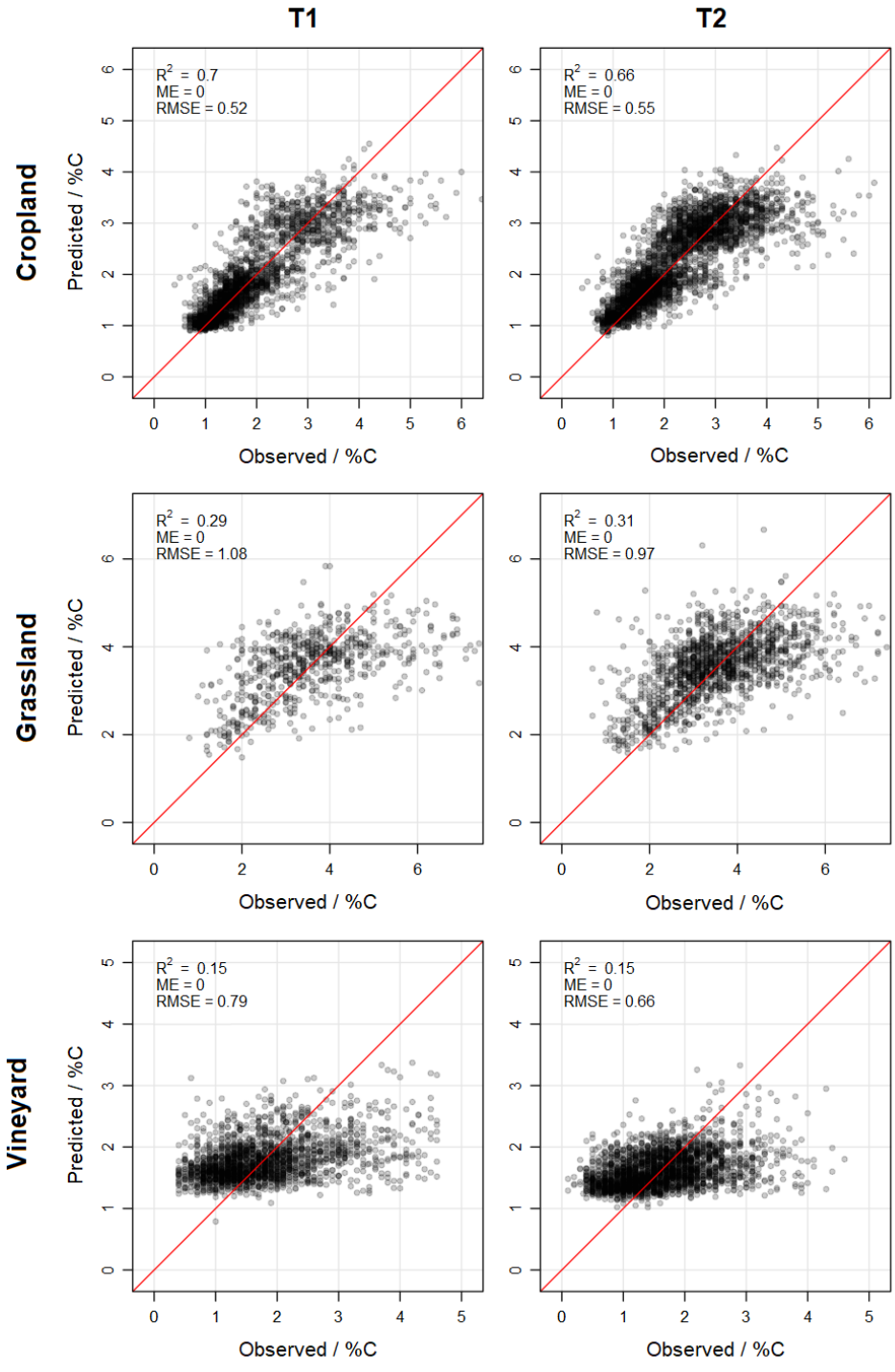


Figure 4.18: Observed vs predicted SOC (%C) as obtained by the models fitted for cropland, grassland, and vineyard soils at T1 (2012-2015) and T2 (2016-2019).

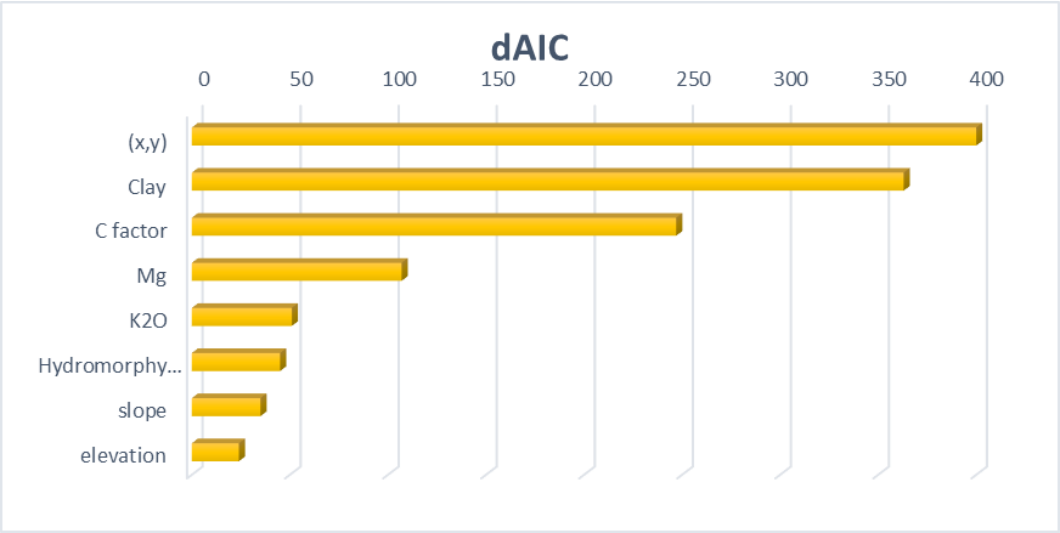


*Figure 4.19: Implication of covariates in the GAM fitted on topsoil SOC (%C for the 0-25cm depth) in **croplands** of Grand-Duchy of Luxembourg. dAIC represents the difference of AIC to the final model (Akaike Information Criterion; Akaike, 1974). Only the covariates showing a p-value < 0.05 in the final GAM model were kept in this Figure.*

- *Vineyard*

The variance 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.

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. 4.20 and 4.21 along with their respective map of standard error of prediction. Additional maps produced for croplands and grasslands separately are available in Annex 8.5.

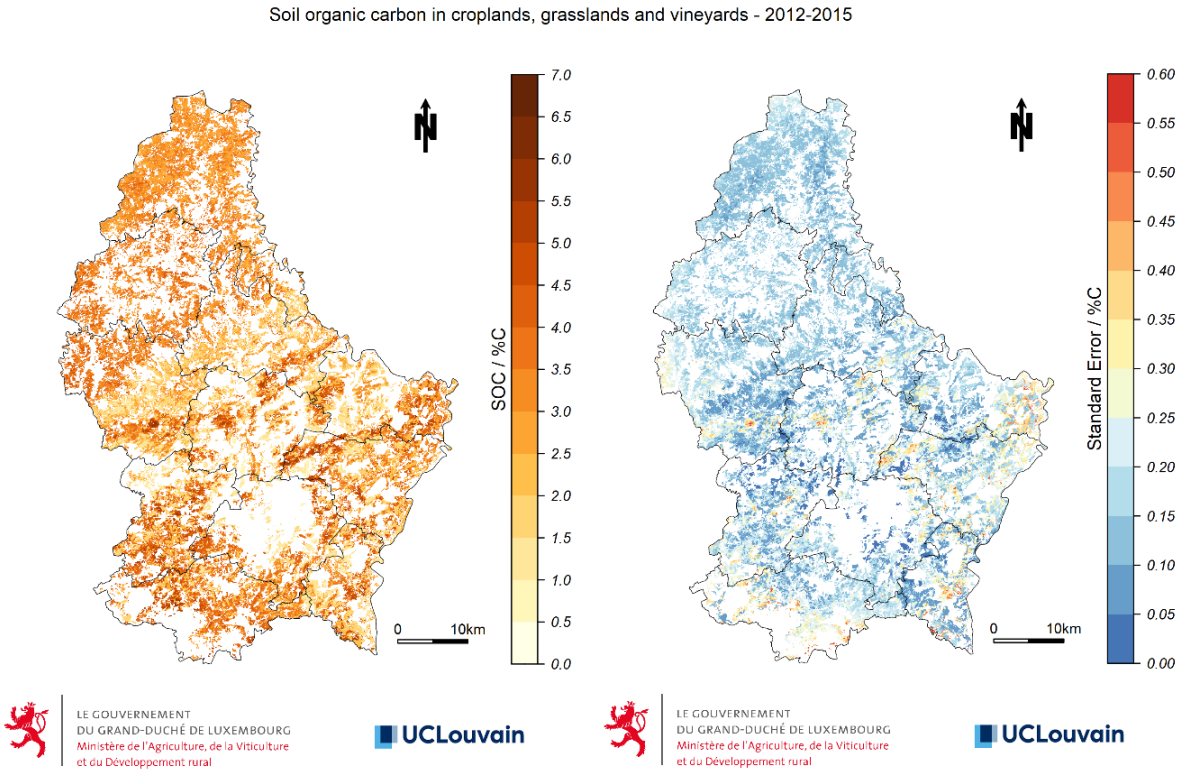


Figure 4.20: Maps of predicted SOC content (%C; on the left) and standard error of prediction (%C; on the right) for topsoil of Grand-Duchy of Luxembourg under Croplands, Grasslands and Vineyards for period T1 (2012-2015).

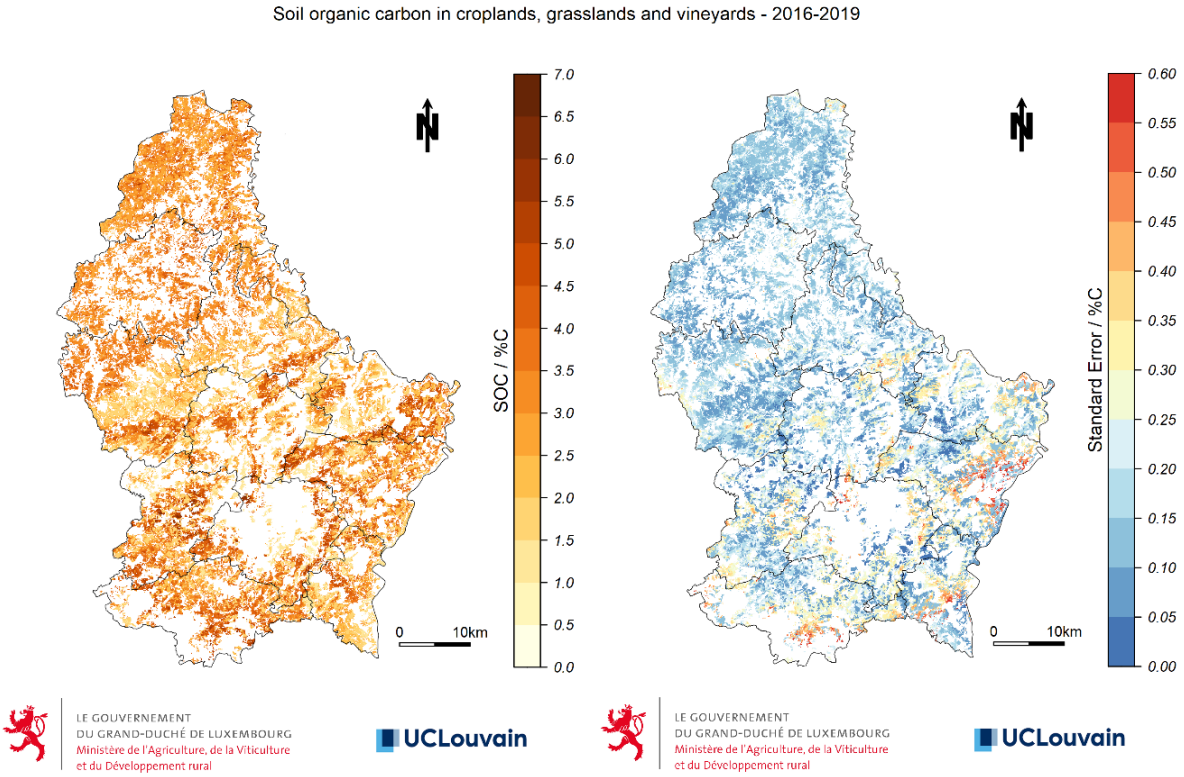


Figure 4.21: Maps of predicted SOC content (%C; on the left) and standard error of prediction (%C; on the right) for topsoil of Grand-Duchy of Luxembourg under Croplands, Grasslands and Vineyards for period T2 (2016-2019).

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. 2.3 and 2.5) and land use. The error of prediction map for T1 shows no clear regional trend (Fig. 4.20 – right-hand panel). 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. 4.21 – right-hand panel).

Figure 4.22 presents the significance of predicted SOC differences between T1 (2012-2015) and T2 (2016-2019) for cropland (left part) and grassland (right part). For cropland and grassland, we estimated that ca. 40% of their respective areas had Non-Statistically significant (NS) evolution of their topsoil SOC content. The predicted gain of SOC was estimated statistically significant for ~25% of the cropland areas and for ~30% of the grassland areas (of which ~13% were statistically significant at $p < 0.05$ for each land use). Regarding the predicted loss of SOC, it was estimated statistically significant for ~35% of the cropland areas and ~30% of the grassland areas (of which ~17% were statistically significant at $p < 0.05$ for each land use). 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 statistically significant predicted **loss** of SOC (Fig. 4.22 – lefthand panel) are mainly located in:

- The west and north-east parts of Oesling where we observed a statistically significant mean decrease of -0.09%C ($p<0.05$; Table 4.2);
- 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; Table 4.2);
- 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; Table 4.2).

The **cropland** areas subject to a statistically significant predicted **gain** of SOC (Fig. 4.22 – lefthand panel) are located in:

- the southernmost part of “Oesling” (although the trend is negative at the entire region scale; Table 4.2);
- the northernmost part of Gutland, i.e in the “Buntsandstein” which showed a mean SOC increase of +0.12%C ($p<0.05$; Table 4.2);
- 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 ($p<0.05$; Table 4.2).

The **grassland** areas subject to a statistically significant predicted **loss** of SOC (Fig. 4.22 – righthand panel) are mainly concentrated in Gutland:

- in its northwestern most part which corresponds to the “Buntsandstein” soil association (fig. 2.3A) where we observed a mean non-statistically significant decrease of -0.24%C (Table 4.3);
- 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; Table 4.3).

The **grassland** areas subject to a statistically significant **gain** of SOC are also mostly concentrated in Gutland (Fig. 4.22 – righthand panel):

- in most of the “Alluvions et Colluvions” located in valley bottoms where we observed a statistically significant increase of 0.26%C ($p<0.05$; Table 4.3) and;
- locally on soils developed on “Grès du Luxembourg” and “Dépôts limoneux sur Grès” where we observed respective non-statistically significant increase of +0.03%C and +0.05%C, (Table 4.3).

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. 4.23).

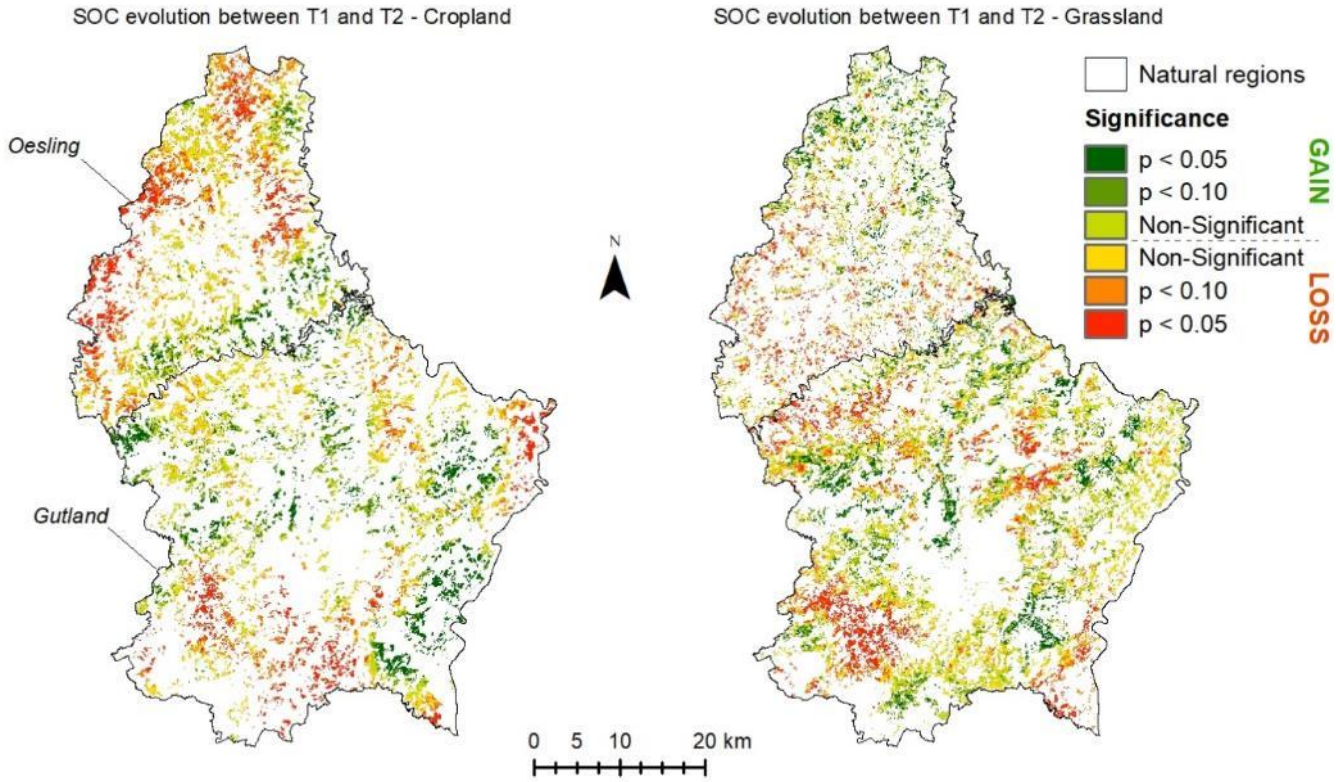


Figure 4.22: ***Significance of predicted SOC differences (p-value)*** *between T1 (2012-2015) and T2 (2016-2019) for soils under croplands (left) and grasslands (right).*

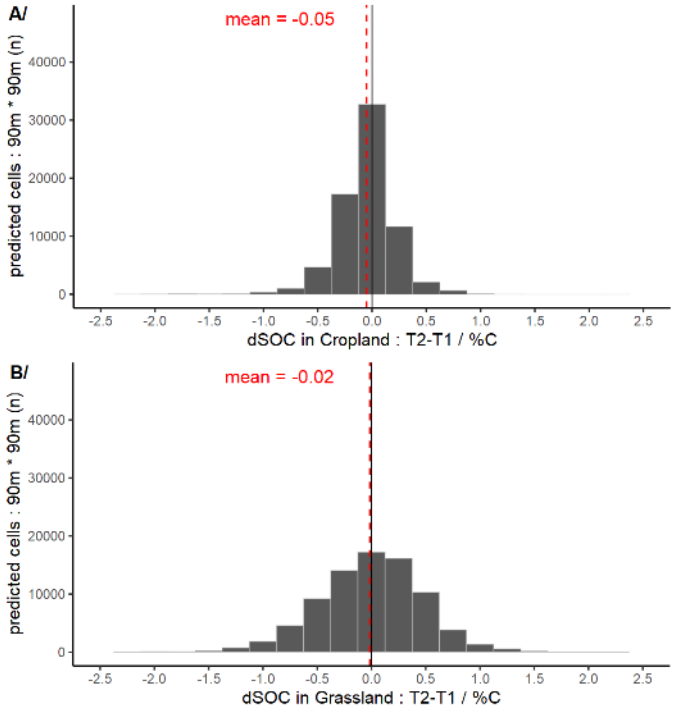


Figure 4.23: Histograms of the predicted difference in SOC (%C) between T1 (2012-2015) and T2 (2016-2019) for soils under cropland (top) and grasslands (bottom).

Considering the results of sections 4.1 and 4.2, 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 (Baveye et al., 2020; Wang et al., 2013; Wang et al., 2014). 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 (Keiluweit et al., 2017). Finally, in the framework of Good Agricultural Practices (see § 4.3), changes in management practices could have induced more C inputs and/or less C outputs (Bhupinder et al., 2018). 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.

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

4.3.1 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).

4.3.2 Methodology

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 additionally supports the sustainable use of farmland through cover crop cultivation as GAP.

Cover crops are cultivated in order to protect the soils. They are incorporated 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 al., 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[9]. 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).

Database preparation

We used SOC data extracted from the ASTA database BDAT of agricultural soil analysis for the period 2012-2019 (as in § 3) merged to the Land Field Information System for the period 2008-2019[10].
Firstly, the data were filtered and cleaned, following the steps below:

1. 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);
2. Removing the observations for which the FLIK polygons were not available or the FLIK was miscoded;
3. Removing observations without date of sampling (mandatory to assess the cultural year);

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

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

4. 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 land use of no interest[11];
5. Removing the duplicates by FLIK and crop year, and replacing them by their mean SOC value;
6. Removing those without GAP for more than 3 consecutive crop years (before the sampling year);
7. By land use and soil association, removing right-skewed data[12] (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[13]) 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.

Analysis of SOC differences between management practices

Here, we tested the differences of statistical distribution between different subsets of LU-SOC-GAP. The enlarged analytical uncertainties (Tab. 3.2.; § 3.1.3) did not influence the test results because these uncertainties apply to each SOC observations of the dataset, and therefore has no effect on the respective spatial distribution of the subsets.

- *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 observations, and before the first observation.

[11] 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.

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

[13] 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.

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 4.3.2, 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 hydromorphic features. 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).

4.3.3 Results

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 4.6). Not considering the observations obtained with the Tru Spec CN analyzer induced a loss of 1101 observations. The elimination of fields submitted to recent land use change or being under land use 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 4.6: Filtering steps on the LU-SOC-GAP database preparation and associated numbers of observations

Filtering step	Total Obs.	Obs. eliminated
None	**13871**	-
- 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

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. 4.24). '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. 4.25-4.26).

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. 4.26; Table 4.7). However, this difference is statistically 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 statistically significant for 'Argiles lourdes des schistes bitumineux' (+0.65%C; $p<0.001$) but, **as the number of Control observations was < 30 (n=12), 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.

NB1: In this section, the differences of statistical distribution between different datasets were tested (Tables 4.7 to 4.10). The analytical uncertainties did not influence the test results (see § 4.2.2).

NB2: Some subsets contained less than 30 individuals preventing the determination of proper statistics; i.e. description of the distribution. The statistics for these subsets are depicted in gray italics in the next tables.

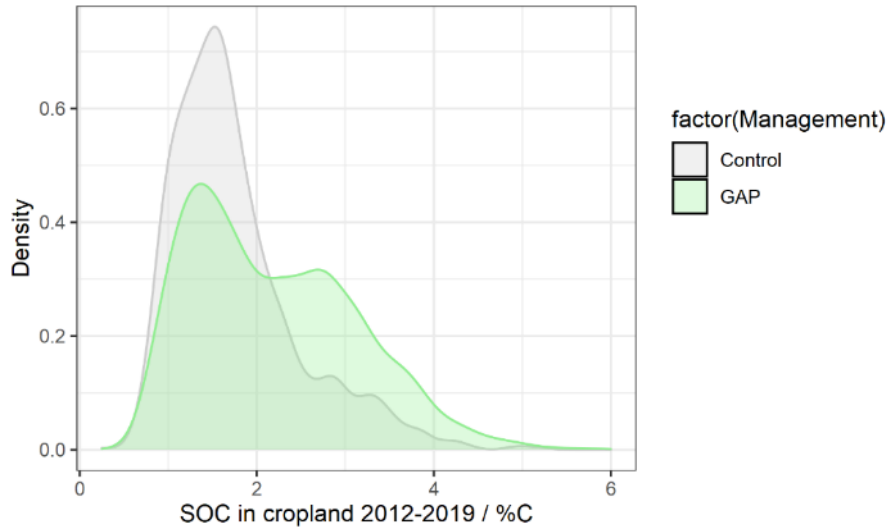


Figure 4.24: Histograms of topsoil SOC (%C) in croplands for fields under Control conditions and fields under GAP in Grand-Duchy of Luxembourg (2012-2019).

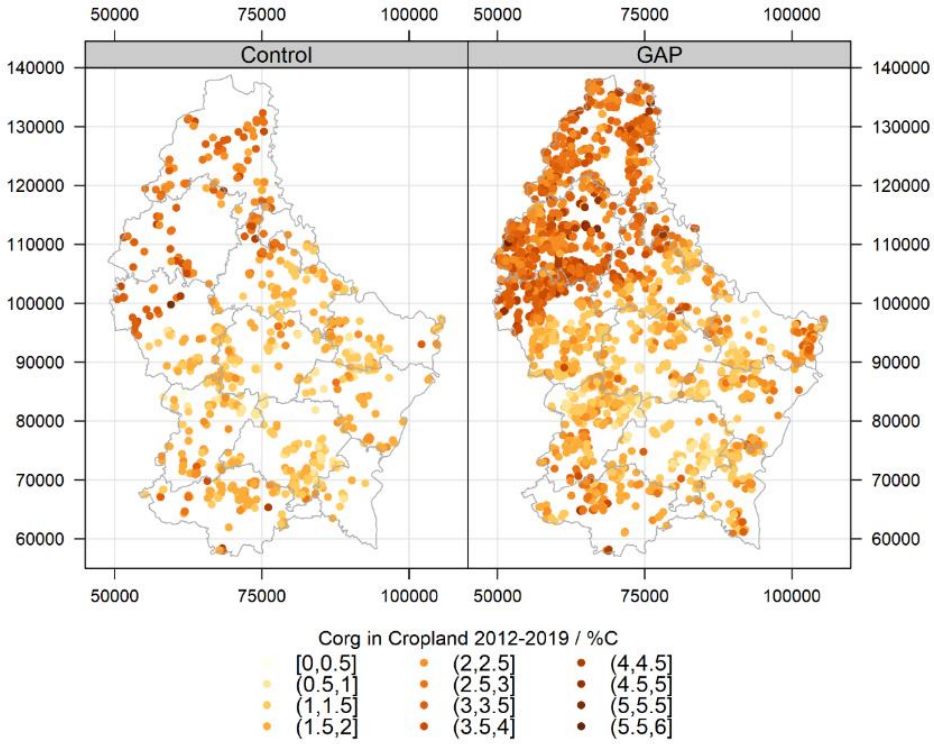


Figure 4.25: Observed SOC values (%C) of topsoil in croplands for fields under Control conditions and fields under GAP in Grand-Duchy of Luxembourg (2012-2019).

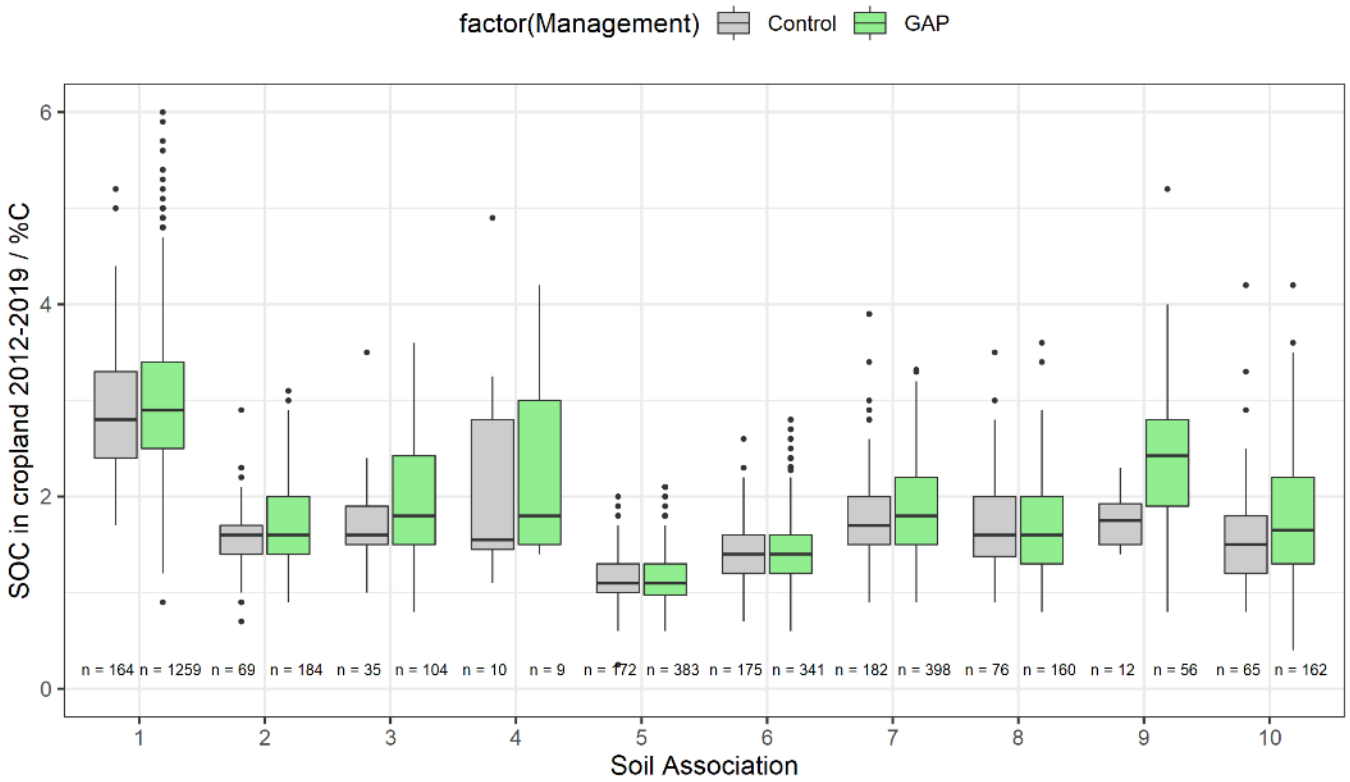


*Figure 4.26: Box-plots of topsoil SOC (%C) in **croplands** for fields under Control conditions and fields under **GAP** (2012-2019). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

*Table 4.7: 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
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. 4.27). 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.

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. 4.28). 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.

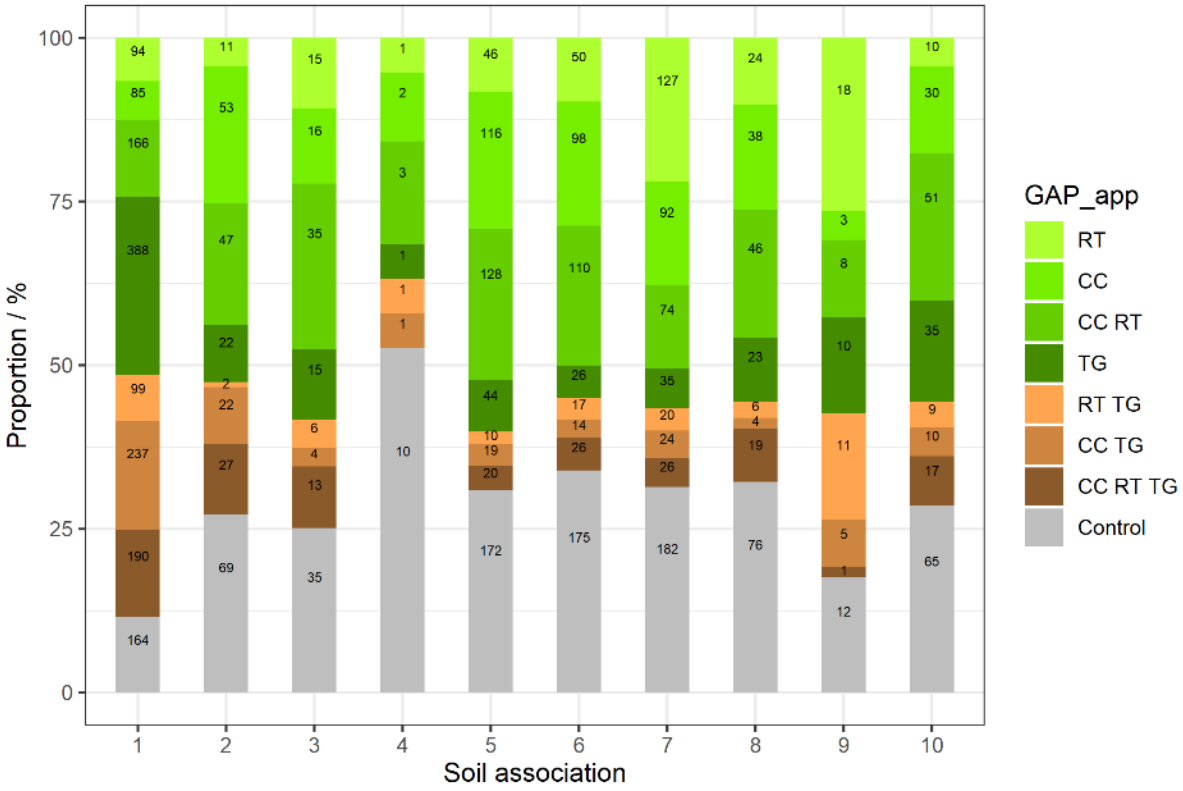


*Figure 4.27: Bar-plots of the proportion of observations in LU-SOC-GAP from fields under **Control conditions** and fields under **GAP or combination of GAPs** (2012-2020). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

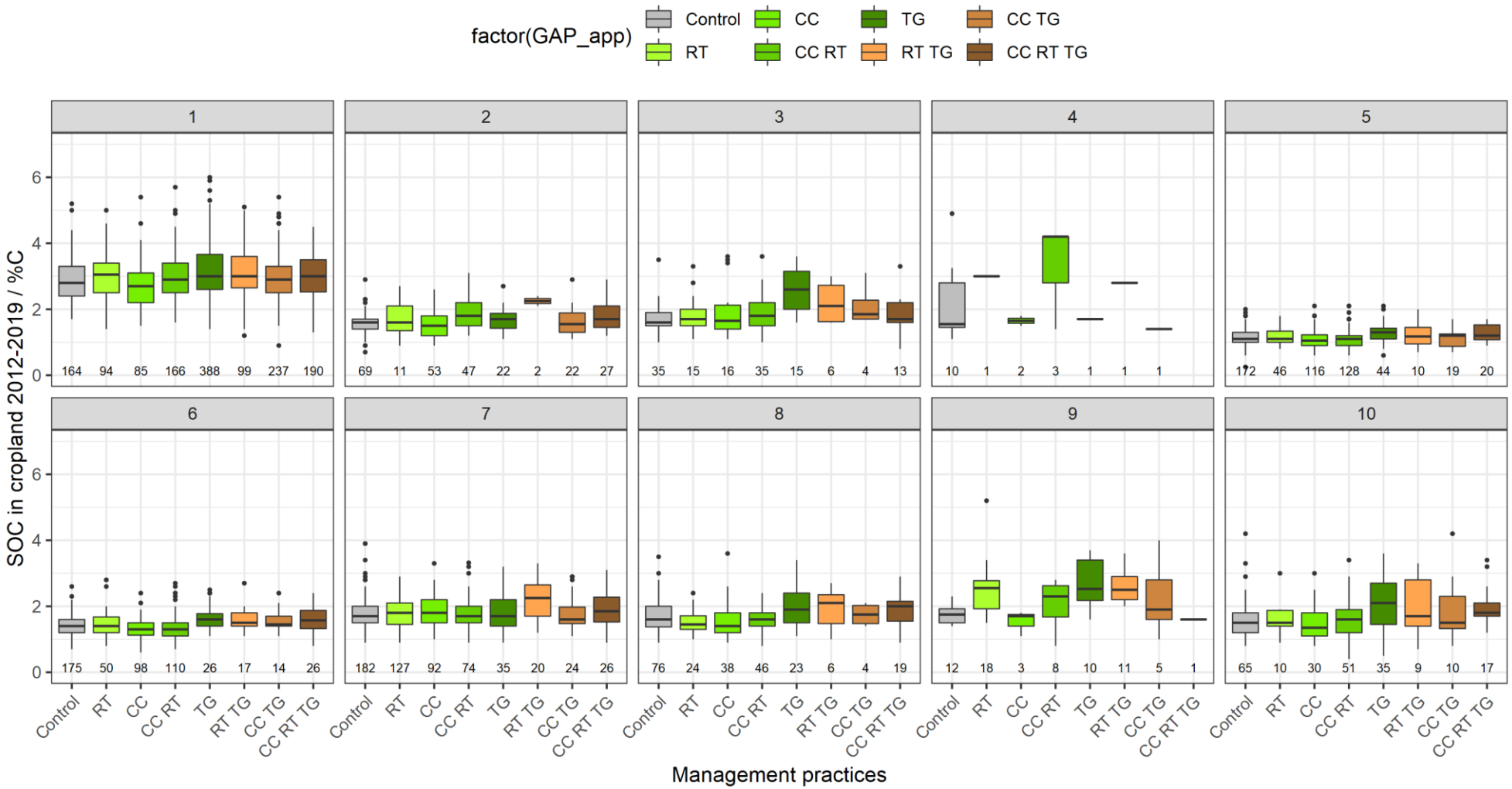


*Figure 4.28 Box-plots of topsoil SOC (%C) in **croplands** for fields under **Control conditions** and fields under **GAP or combination of GAPs** (2012-2020). (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires du Bajocien, 5 = Grès de Luxembourg, 6 = Dépôts limoneux sur Grès, 7 = Argiles du Lias inf. et moyen, 8 = Argiles lourdes du Keuper, 9 = Argiles lourdes des schistes bitumineux, 10 = Others)*

Tables 4.8 to 4.10 present summary statistics and results of statistical tests between observations under Control vs Reduced Tillage, Cover Crops and Temporary Grassland. In Table 4.8, 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 statistically significant ($p>0.05$). In Table 4.9, 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 statistically 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 statistically significant ($p<0.05$) except for 'Argiles du Lias inf. et moyen' (+0.04%C; $p>0.05$).

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.

*Table 4.8: 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
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 4.9: 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
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 4.10: 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
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**

- *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. 4.29). 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 small numbers of values by subsets, with n <30 often, **those results have to be considered very carefully. For information purposes only, we computed the relative annual difference in SOC and compiled the results in Table 4.11.**

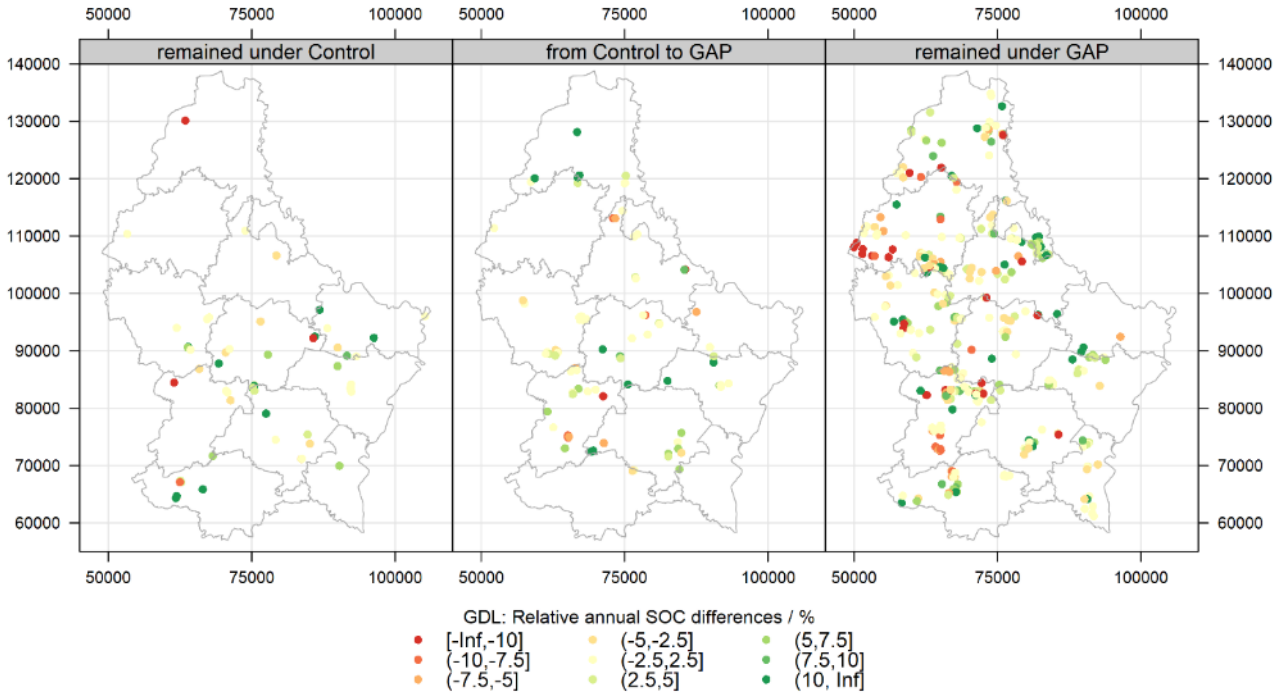


Figure 4.29: Location and relative annual differences in SOC content between paired observations (same FLIKs) in the LU-SOC-GAP dataset.

Table 4.11: 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

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. 4.30). 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%.

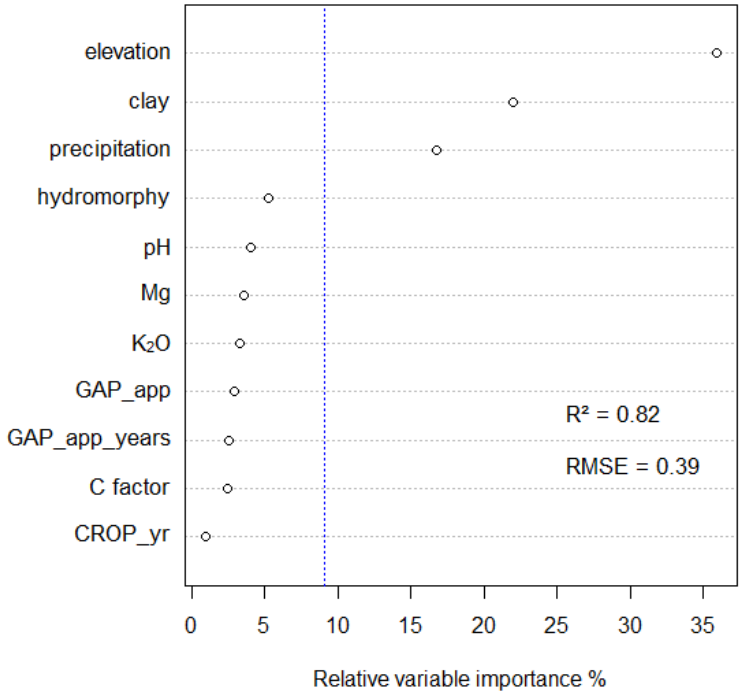


*Figure 4.30: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC in croplands **all over GDL** (n = 4016). The vertical blue dashed line indicates the average relative importance.*

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. 4.31-4.32). 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. 4.32). The GAP applied and their duration (i.e., GAP_app and GAP_app_years) together had a relative importance ~ 5%. In Oesling, SOC variance is mainly explained by the minimum depth of hydromorphic features, the dominant gradients (from N to S) of precipitation and elevation[14], the clay and the Mg contents, and the GAP application (Fig. 4.31). 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

[14] *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)*

of clay content than Gutland, 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. 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.

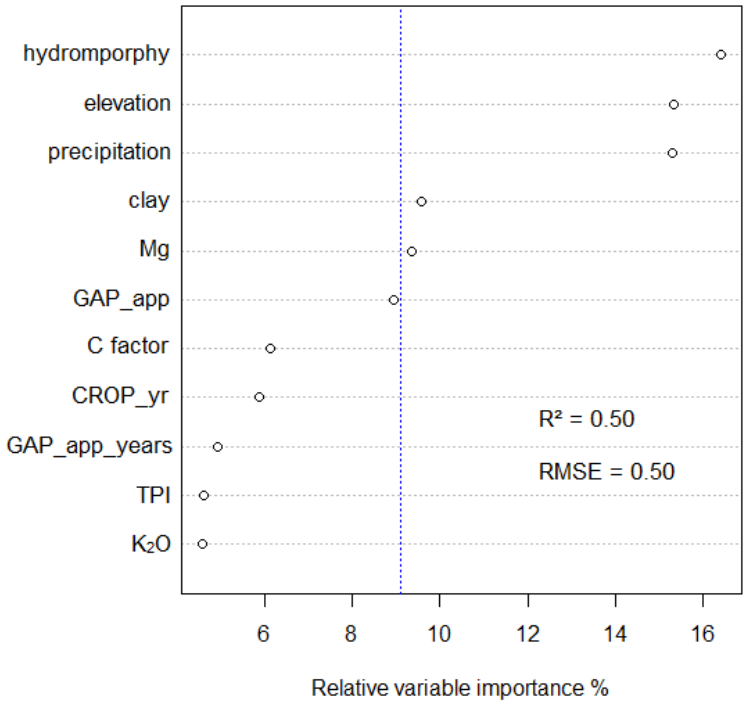


*Figure 4.31: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC **in croplands of Oesling** (n = 1464). The vertical blue dashed line indicates the average relative importance.*

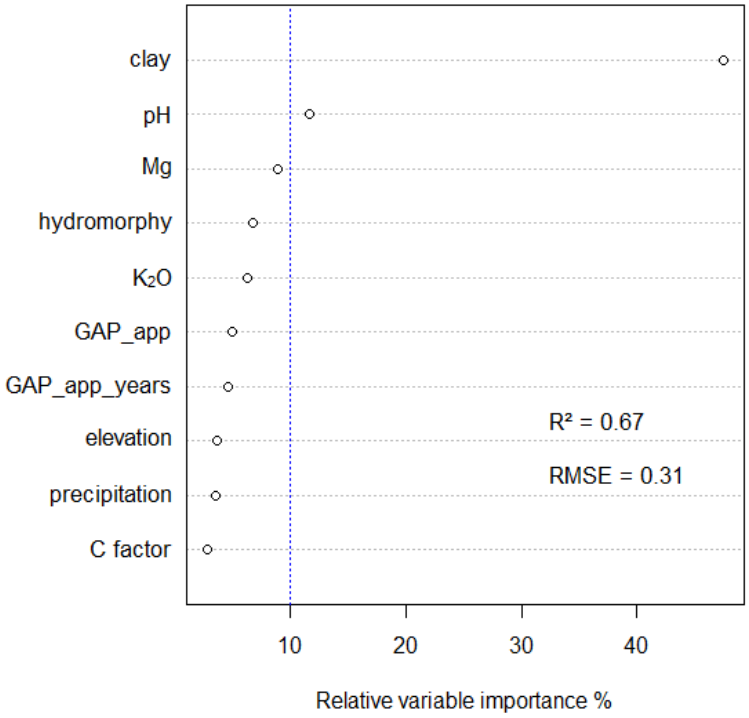


*Figure 4.32: Relative variable importance of covariates selected by the fitting procedure of the conditional inference forest model for SOC **in croplands of Gutland** (n = 2552). The vertical blue dashed line indicates the average relative importance*

5 ORGANIC CARBON STOCKS IN AGRICULTURAL SOILS

5.1 FROM SOC CONTENT TO SOC STOCK

5.1.1 Context

SOC content (or concentration; in %C) of topsoil was estimated all over GDL (§4.2). In order to assess the total amount (budget) of SOC stored in soils of a given area, we need to consider SOC stocks, i.e. to consider the soil density, depth and the proportion of rock fragments (stones; >2mm) to the total soil mass. SOC stocks were computed here according to equation 7:

$$SOCst = SOC * d * BD * (1 - RM) \quad (7)$$

With SOCst: the SOC stocks (MgC.ha^{-1}); SOC: the SOC content (%C) in the fine earth (<2mm); d: the soil depth (cm) over which stocks are computed; RM: the Rock fraction content by Mass (-) in the bulk sample; BD: the soil bulk density (g.cm^{-3}).

As required by the UNFCC, the SOC stocks were estimated and mapped for the 0-30cm layer (SOCst.30). A first mapping of SOC stocks for GDL was produced by Stevens et al. (2014b) for the period 2012-2014. This map was computed by combining continuous spatial layers representing SOC content, BD and RM. Here, we attempted to improve the methodology for SOC stocks assessment by estimating SOCst.30 for each observation where SOC content was provided.

The SOCst.30 were estimated based on the SOC observations compiled in the BDAT_2012-2019 (section 3.1.2). However, this database did not contain all the parameters required to apply Eq. 7.

5.1.2 Estimation of parameters required for SOCst.30 computation (Eq. 7)

The methodologies used to estimate SOC.30, BD and RM are developed below. Once the methods validated, they've been applied to BDAT_2012-2019. The related descriptive statistics (by landuse and soil association) were compiled in Annex 8.6.

SOC.30: SOC content in the 0-30cm layer

In the BDAT_2012-2019 database, samples were taken at 0-25 cm in cropland (SOC.25), 0-15 cm in grassland (SOC.15) and 0-30 cm in vineyard (SOC.30). We needed to estimate SOC in the 0-30cm layer (SOC.30) for cropland and grassland. We first gathered databases where SOC was available on soil profiles for GDL (section 3.1.2), i.e.:

1. the historical profiles of BDSOL_1964-1973;
2. and the recent profiles of BDSOL_2009-2019.

Then, we generated a continuous SOC distribution with depth using mass-preserving splines (using *'ithir'* R package) for each soil profile where SOC was originally measured by soil horizons. Mass-preserving splines allow fitting functions that preserve the mean value per horizon by ensuring that the area to the left of the fitted line above the horizon depth average is equal to the area to the right below the horizon depth average (Fig. 5.1). The lambda parameter controlling the degree of smoothness of the splines was set to 0.1 and chosen as to obtain a low root mean square error of prediction while keeping the functions smooth enough. Based on the fitted splines, we computed the average SOC content for the 0-30 cm layer (SOC.30), for the 0-15 cm layer for grasslands (SOC.15) and for the 0-25 cm layer for croplands (SOC.25). Then, we simply modeled SOC.30 as a function of SOC.15 for grasslands and SOC.25 for croplands.

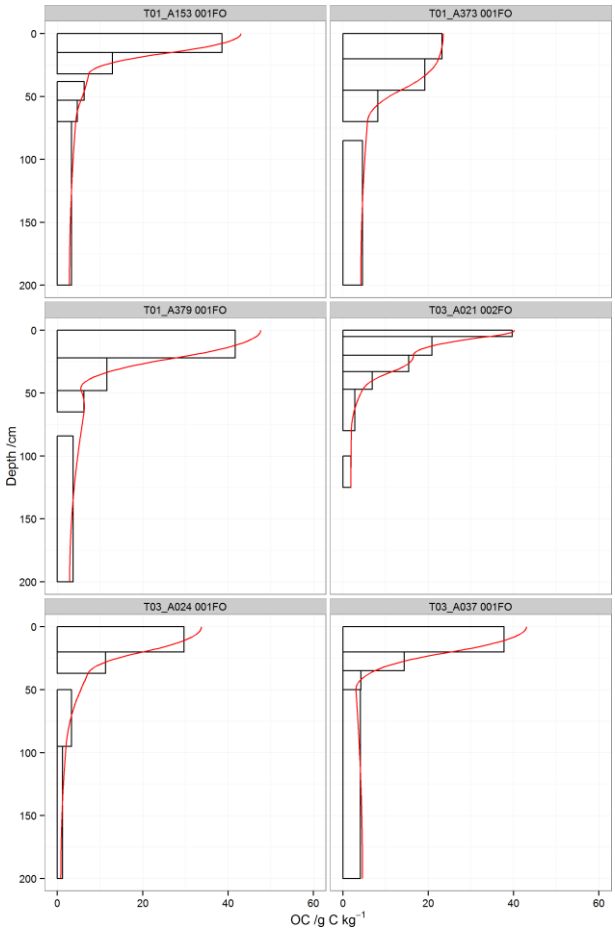


Figure 5.1: Examples of soil profiles showing horizon observed SOC values (gC.kg-1; horizontal bars) and interpolated SOC values using mass-preserving splines (red lines). Source: Stevens et al., 2014b

In croplands, the linear model lmSOC.30$_{crop}$ between SOC.30 and SOC.25 fits very well with a R^2 = 0.99 and a RMSE = 0.07 %C (Fig. 5.2; left panel) obtained in 10-fold cross-validation. The model showed that SOC.30 represents about 93 % of SOC.25 (Eq. 8):

lmSOC.30$_{crop}$: $$SOC.30 = 0.93 * SOC.25 + 0.01$$ (Eq. 8)

In grasslands, the linear model lmSOC.30$_{grass}$ between SOC.30 and SOC.15 fits well with a R^2 = 0.89 and a RMSE = 0.31 %C (Fig. 5.2; right panel) according to the 10-fold cross-validation. Here, SOC.30 was estimated representing around 75 % of SOC.15 (Eq.9):

lmSOC.30$_{grass}$: $$SOC.30 = 0.76 * SOC.15 + 0.15$$ (Eq. 9)

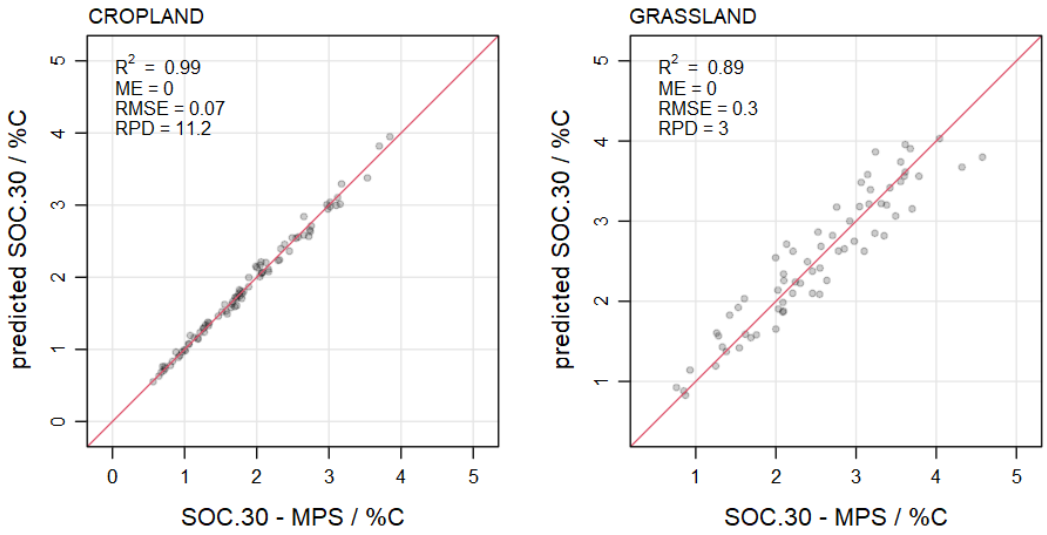


Figure 5.2: Results of the 10-fold cross-validations of the linear models lmSOC.30$_{crop}$ and lmSOC.30$_{grass}$ predicting, respectively, SOC.30 (%C) from SOC.25 in cropland (left panel) and SOC.30 from SOC.15 (right panel), against SOC.30 estimated by MPS (Mass Preserving Splines).

The validated models lmSOC.30$_{crop}$ and lmSOC.30$_{grass}$ were applied to BDAT_2012-2019. In Gutland (soils associations 2 to 9 mainly), soils under cropland had less SOC.30 than soils under grassland: in cropland, median SOC.30 values varied from 1.03%C in 'Grès du Luxembourg' to 2.09 %C in 'Argiles lourdes des schistes bitumineux' (Annex 8.6), whereas in grassland SOC.30 varied from 1.51 %C to 3.33 %C in the same soil associations. However, in Oesling, the difference in SOC.30 between cropland (median = 2.79 %C) and grassland (median = 2.88 %C) was slighter. Along with the SOC.30 values, the Standard Errors (SE) of the parameters were also computed to assess the 66% confidence interval of both models.

Bulk Density (BD)

When not measured, BD is generally estimated by a Pedo-Transfert Function (PTF) through its relation to SOC (or OM) and soil texture; i.e. PTF are empirical models. Several published PTFs could have been used, but as shown by Stevens et al. (2014b), those fitted not very well to the BD data available then for GDL (R^2 varying between 0.43 and 0.52 and sometimes large standard errors).

Since then, Bulk Density data have been added to the BDSOL_2009-2019 (recent soil profiles), especially in Gutland. More samples were required to cover a bigger range of soil types and climate conditions related to GDL. To this aim, we selected BD measures from Wallonia (in CARBOSOL dataset; Chartin et al., 2017) only related to soil types identified in both GDL and Wallonia. We restricted our research to Walloon natural regions having similar characteristics to Oesling (i.e., Ardenne and Haute-Ardenne regions in Wallonia) and Gutland (i.e., Jurassic region in Wallonia). After data cleaning, we counted n = 267 horizons with BD values (166 from GDL and 99 from Wallonia; Fig. 5.3).

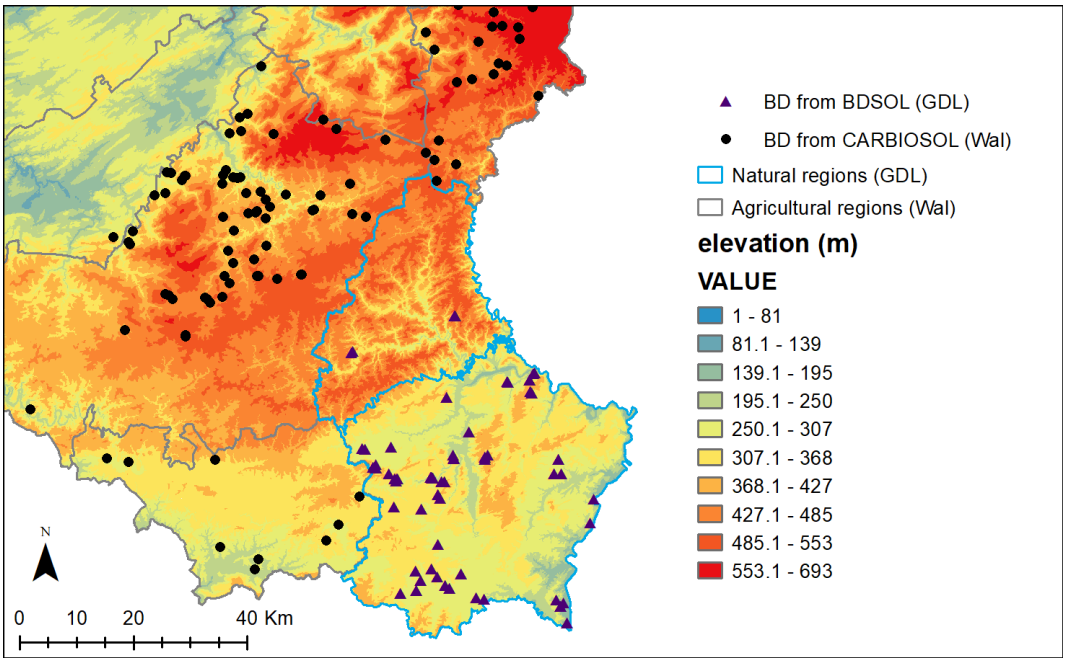


Figure 5.3: Location of Bulk Density (BD) data in Grand-Duchy of Luxembourg (GDL) from BDSOL_2009-2019 and Wallonia (Wal) from CARBIOSOL dataset (Chartin et al., 2017).

Based on this dataset, we tested different models in order to develop a specific PTF for estimating BD for soils of the GDL. We tested the most classical forms of model used as PTF for BD based on SOC and/or texture, and two machine learning techniques (Random Forest – RF, and Generalized Boosting Model - GBM). The most classical forms of model tested here to estimate BD ($g.cm^{-3}$) were:

- $BD = a * SOC + b$ (LM1)
- $BD = a * SOC + b * clay + c * silt + d * sand + e$ (LM2)
- $BD = a * SOC^2 + b$ (Sqrt)
- $BD = a * log_{10}(SOC) + b$ (Log10)
- $BD = a * \log(SOC) + b * (\log(SOC))^2 + c$ (Log)

With SOC: the SOC content (%C); clay, silt and sand: the respective clay, silt and sand fractions in soils; a, b, c, d and e: some constants.

Random Forest (RF) and Generalised Boosting Model (GBM) are both methods combining large number of classification and decision trees. RF combines the trees (using average or "majority rules"; Strobl et al., 2007) at the end of the process when GBM starts to combine trees at the beginning (Friedman, 2002). For RF and GBM, the covariates tested in the models (all available in the GDL and Wallonia subsets) were: the SOC content, the proportions of clay, silt and sand, the landuse, the mean depth of the horizon, and the texture classes (Mat_Text) from the triangle designed for soils of Belgium and Grand-Duchy of Luxembourg (Fig. 2.4). The texture class 'G' for stony loamy soils was also included (Tab. 2.2).

Seven models were tested to model BD and machine learning methods (RF and GBM) showed better results than the five classical models (LM1, LM2, Sqrt, Log10 and Log; Fig. 5.4). The GBM showed the best fit with a $R^2 = 0.91$ and a RMSE = 0.1 $g.cm^{-3}$ after a 10-fold cross-validation (Fig. 5.4) and was then

retained to model BD in GDL. SOC was the most influencing covariate with a relative influence of ~40 %, followed by the texture class (Mat_Txt) and a relative influence of ~15% (Fig. 5.5). Finally, the mean depth of the horizon with the clay, silt and sand fractions had a relative influence of ~10 %.

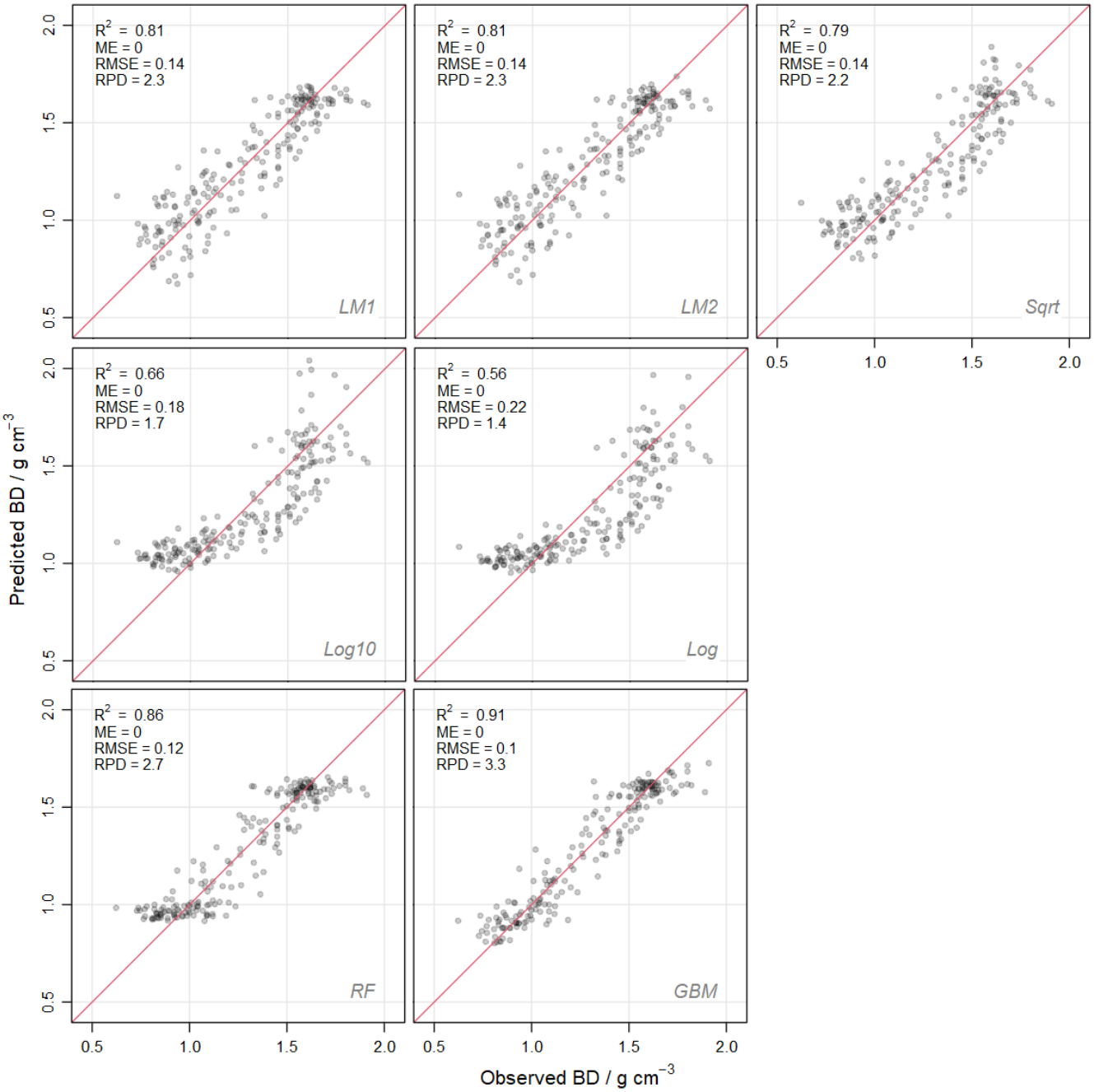


Figure 5.4: Results of the 10-fold cross-validations of the different forms of model tested to estimate soil Bulk Density (BD in g.cm³) in GDL.

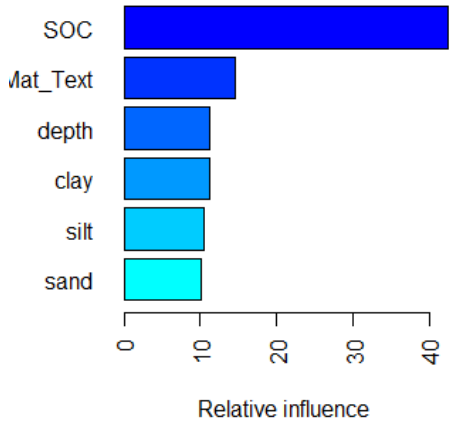


Figure 5.5: Relative influence of covariates selected in the GBM model to estimate soil Bulk Density (g.cm³) in GDL.

Once the GBM validated, it was applied to BDAT_2012-2019. As BD is known to decline with increasing SOC, it was logical to observe smaller BD values under grassland (median = 1.22 g.cm^{-3}) than in cropland (median = 1.38 g.cm^{-3}) and vineyard (median = 1.41 g.cm^{-3}; Annex 8.6). Also, cropland and grassland of Oesling (soil association 1) showed smaller BD than those of Gutland (soil associations 2 to 9). To finish, the 90% estimation interval of BD was assessed using two GBM quantile regressions (the 5% and 95% quantiles; § 5.3.2.).

Rock Fragment content by Mass (RM)

RM data were available in two different sources for GDL:

1. the historical profiles and topsoil survey of BDSOL for 1964-1973;
2. the BDSOL for 2009-2019 (topsoil only).

These two sources were gathered in one dataset containing RM data, the proportions of clay, silt and sand, the horizons depths (upper and lower), and the geographic coordinates. From the two original BDSOL datasets, we kept data from sites under cropland, grassland, vineyard and forest in order to optimize the number of observations (n = 1253) and the range of RM values ([0.00 ; 0.55] here). Then, the coordinates were used to extract complementary covariates, likely to influence RM in soils, from: i/ spatial covariates described in § 3.3. (elevation, slope, TPI and pH), and ii/ from the Numerical Soil Map (the texture classes - Mat_Text, and the profile development index – Dev_Profil). The geology could not be used here as a covariate because the observations did not cover the complete lithology of GDL. Random Forest and Generalized Boosting Model were then tested.

Similar to the estimation of BD, a GBM approach showed the best fit for RM with a R^2 = 0.77 and a RMSE = 0.05. As shown in Figure 5.6, the GBM model tended to underestimate RM in certain cases (not identified yet). However, the GBM fitting allowed us to identify the covariates influencing the most the RM content (Fig. 5.7). The three most influential covariates (relative influence > 10%) were 1/ the texture class (Mat_Txt), 2/ the elevation and 3/ the TPI (Topographical Position Index).

The GBM model was applied to the observations of BDAT_2012-2019 and descriptive statistics were computed by land use (cropland, grassland and vineyard; Annex 8.6). We noted that from one land use to another, there is very slight difference in RM by soil association, except for 'Calcaires du Bajocien' (4) which showed a median value of 0.05 in cropland and grassland when the median value is of 0.10

in vineyard. To finish, the 90% Confidence Interval of RM was assessed using two GBM quantile regressions (the 5% and 95% quantiles; section 5.3.2.).

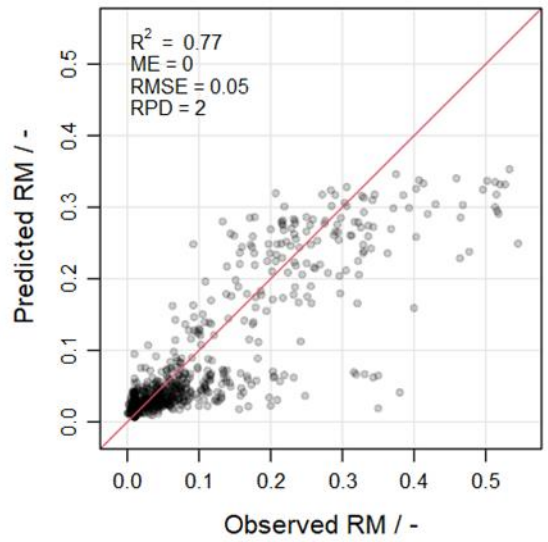


Fig. 5.6: Results of the 10-fold cross-validation of GBM model selected to estimate the Rock fraction content by Mass (RM; -) in soils of GDL.

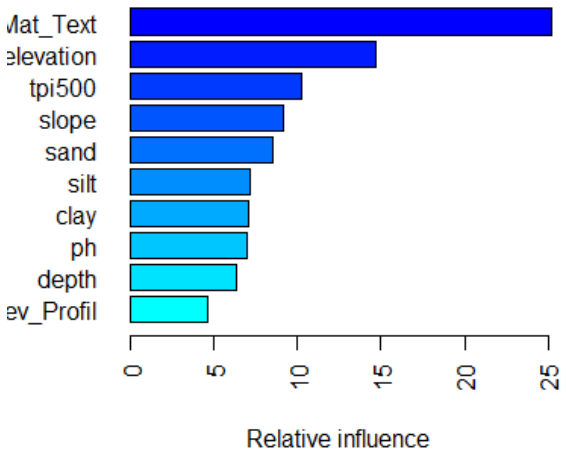


Figure 5.7: Relative influence of covariates selected in the GBM model to estimate Rock fraction content by Mass (RM) in soils of GDL.

5.1.3 Estimation of SOCst.30 and associated 90% confidence interval

SOCst.30 were computed for each of the 11,819 sites of the BDAT_2012-20169 dataset (see §3.1.3; data filtered and all covariates extracted at each site location). To this aim, the models developed in §5.1.2 were applied at each site of the dataset to estimate SOC.30, BD and RM. Then, mean estimation of SOCst.30 was calculated following Equation 8. For detail information on SOCst.30 values and statistics by agricultural region and period of interest, please refer to the section §5.3.2.

The 90% Confidence Interval (CI) for SOCst.30 was assessed using an Ensemble Uncertainty Propagation procedure:

1. The prediction intervals of SOC.30, BD and RM estimation were assessed once the models were validated (§5.1.2) and computed for each site of the BDAT_2012-2019;

2. Then, a Monte-Carlo (MC) simulation of 500 iterations was applied for each of these three parameters at each of the 11,819 sites (i.e., a total of 3 x 11,819 different MC simulations). The hypothesis of a normal distribution was applied;

3. For each site, 500 SOCst.30 values were computed based on the 500 iterations of SOC.30, BD and RM simulated by Monte-Carlo in step 2;

4. 90% CI of SOCst.30 (95% percentile – 5% percentile) was computed for each site.

In Tables 5.1 to 5.3, we averaged the 90% Estimation Interval (EI) and relative 90% Confidence Interval (Rel. CI = (CI / mean) * 100; in %) of SOC.30, BD, RM and SOCst.30 in each of the 10 soil associations for cropland, permanent grassland and vineyard, respectively. CI and Rel. CI. In cropland, the rel. CI for SOCst.30 varied from ~14% ('Argiles du Lias inf. et moyen') to ~22% ('Calcaires du Bajocien') in Gutland, and was of ~28% in Oesling. For grassland, SOCst.30 in Oesling had also a Rel. CI around 28% while in Gutland the range was between ca. 21 and 23%. Considering that vineyards soils are predominantly located on 'Dolomies du Muschelkalk', 'Argiles Lourdes du Keuper' and 'Others', and to a lesser extent on 'Buntsandstein', their Rel. CI for SOCst.30 varied mainly in a range of 12-15%.

Whatever the landuse, the most important source of variation in estimating SOCst.30 by the procedure developed here was mainly the Rock fragment content by Mass (RM; Rel. CI of ~67 to 218%), followed by the Bulk Density (BD; Rel. CI of ~11 to 26%) and, to finish, by the SOC content (SOC.30; rel. CI of ~0.7 to 7%). Goidts et al. in 2009 also observed the same parameters ranking for the sources of uncertainties identified when measuring SOC stocks 'in-field' through the Carbosol Soil Monitoring Network in Wallonia. Indeed, the RM is the most difficult parameter to estimate; it would necessitate an important amount of soil material to be accurately assessed. Here, the RM was determined by dry sieving at 2mm, based on a mean amount of soil material of ~1.75 - 2.00 kg (Nau, 2015; Annex 8.7). Also, the PTF developed here to predict RM showed a RMSE of 0.05 and an important bias of underestimation. Concerning the Bulk Density (BD), the Rel. CI appeared higher in Oesling than in Gutland (Tab. 5.1 and 5.2). The calibration data for BD PTF came mainly from Gutland while we had to complete with Walloon data from equivalent soils to Oesling ones.

The relative influence of RM and BD in the SOCst.30 uncertainties could be reduced by new measurements. The RM could be measured throughout the GDL with a standard technique and compare to the calibration data used here. Then, decision could be taken to complete the already existing data or to create a new homogenized RM database. The BD should be investigated in Oesling in order to complete the data that already exists, using the same methodology as the latter.

Table 5.1: For cropland, average 90% Confidence Interval (CI) and relative CI (Rel. CI) by soil association of the soil organic carbon content in 0-30cm depth (SOC.30), bulk density (BD), rock fragment content by mass (RM) and soil organic carbon stock (SOCst.30). (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)

Soil Asso.*	SOC.30		BD		RM		SOCst.30	
	CI (%C)	Rel. CI (%)	CI (g/cm³)	Rel. CI (%)	CI (-)	Rel. CI (%)	CI (MgC/ha)	Rel. CI (%)
1	0.02	0.72	0.22	19.28	0.18	67.60	20.44	28.17
2	0.01	1.00	0.20	14.42	0.11	149.07	10.77	18.37
3	0.02	0.92	0.20	15.40	0.08	121.56	12.53	16.56
4	0.02	1.02	0.24	18.66	0.11	129.93	15.39	21.52
5	0.02	1.79	0.24	15.78	0.04	218.22	6.30	13.32
6	0.02	1.29	0.23	15.86	0.06	186.79	8.54	15.00
7	0.01	0.90	0.22	15.42	0.05	157.20	10.70	15.13
8	0.02	1.06	0.21	14.65	0.05	163.35	9.66	14.02
9	0.02	0.85	0.23	16.63	0.05	187.71	14.14	15.88
10	0.02	1.13	0.22	16.18	0.07	186.60	10.94	16.90

Table 5.2: For Grassland, average 90% Confidence Interval (CI) and relative CI (Rel. CI) by soil association of the soil organic carbon content in 0-30cm depth (SOC.30), bulk density (BD), rock fragment content by mass (RM) and soil organic carbon stock (SOCst.30). (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)

Soil Asso.*	SOC.30		BD		RM		SOCst.30	
	CI (%C)	Rel. CI (%)	CI (g/cm³)	Rel. CI (%)	CI (-)	Rel. CI (%)	CI (MgC/ha)	Rel. CI (%)
1	0.08	2.72	0.25	22.60	0.16	72.98	20.25	27.71
2	0.08	4.31	0.23	18.81	0.10	138.52	14.93	20.87
3	0.08	3.18	0.26	22.06	0.07	127.22	19.59	20.86
4	0.08	3.89	0.28	22.39	0.06	129.60	17.22	21.46
5	0.09	6.71	0.25	18.50	0.04	150.64	13.07	19.00
6	0.08	4.03	0.27	21.42	0.05	164.53	18.40	20.54
7	0.09	3.21	0.29	23.69	0.05	155.72	23.57	21.79
8	0.09	3.14	0.31	24.47	0.05	159.76	25.70	22.09
9	0.10	3.12	0.32	26.16	0.05	183.90	28.38	22.88
10	0.09	3.51	0.27	22.19	0.08	159.83	21.02	21.96

Table 5.3: For vineyard, average 90% Confidence Interval (CI) and relative CI (Rel. CI) by soil association of the soil organic carbon content in 0-30cm depth (SOC.30), bulk density (BD), rock fragment content by mass (RM) and soil organic carbon stock (SOCst.30). (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, 10 = Others)

Soil Asso.*	SOC.30		BD		RM		SOCst.30	
	CI (%C)	Rel. CI (%)	CI (g/cm³)	Rel. CI (%)	CI (-)	Rel. CI (%)	CI (MgC/ha)	Rel. CI (%)
2	-	-	0.15	11.18	0.09	133.99	10.04	14.74
3	-	-	0.16	11.88	0.09	116.32	9.98	14.87
4	-	-	0.15	10.89	0.10	112.75	8.03	14.76
5	-	-	0.14	10.08	0.05	154.50	6.74	10.40
6	-	-	0.14	10.12	0.09	188.22	8.05	13.10
7	-	-	0.15	10.90	0.04	143.03	8.37	11.18
8	-	-	0.16	11.45	0.05	147.78	8.47	12.01
10	-	-	0.16	11.60	0.06	157.34	8.79	12.54

5.2 HISTORICAL TRENDS IN SOC STOCKS: THE IMPACT OF CONVERSION FROM CROPLAND TO GRASSLAND

5.2.1 Context

Between 1910 and 2000, the total area of croplands in GDL was divided by 2 (Fig. 5.6) at the expense of permanent grasslands. The bulk of conversions occurred in the first half of the 20th century. We observed in section 4.2 that variability of SOC contents (and by extension of SOC stocks) in grasslands was more important and more difficult to explain by the environmental covariates and types of practices than in cropland. Here, we explored the influence of historical conversion from croplands to permanent grasslands on SOC stocks variability.

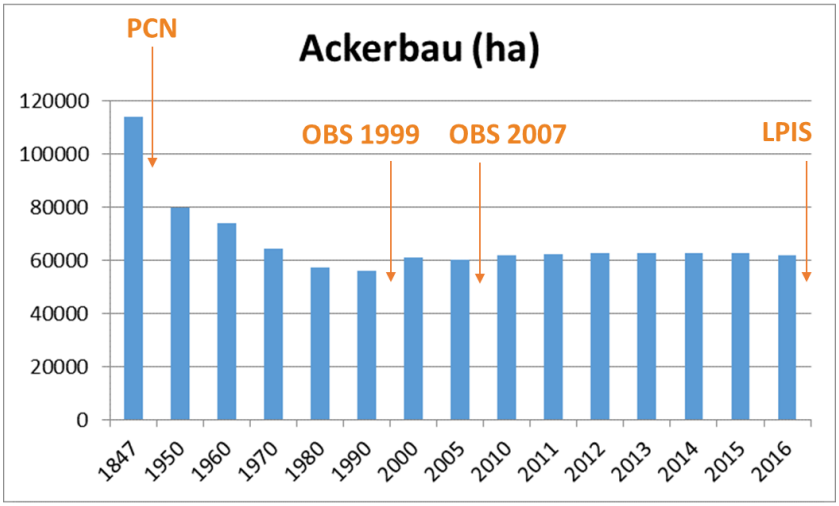


Fig 5.6: Evolution of cropland area over time with the dates for which land use maps are available.

5.2.2 Methodology

To examine the impact of historical conversion of croplands to grasslands on SOC stocks in GDL, we had two options:

1. Collect enough historical SOC data, with their location and land use, to produce proper statistics at different periods (from 1910 to 2020);
2. Or retrieve the land use history of current croplands and grasslands sampled in BDAT_2012-2019 and compare this history with the present SOC stock variability in a space for time replacement, or a so-called chronosequence.

As seen in 3.1.1, the historical observations for soil organic status (here, humic acid content), dating from 1900 – 1914, were not comparable to present SOC observations analyzed by dry combustion. SOC data for 1963 – 1974 were not numerous enough to produce proper statistics all over GDL. Hence, the option 1 was not feasible.

By choosing option 2, we shifted from a temporal to a spatial approach and chose to construct ***chronosequences.*** The spatial variability of a parameter (here, SOCst.30) is confronted to the

temporality of a specific covariate (here, date or time of conversion from cropland to grassland). Three steps were then followed:

- collecting land use information / layers at different times from 1910 to 2020 and extracting the information to the locations of present SOCst.30 observations of BDAT_2012-2019 (i.e., using FLIKcode, x and x). The areas were checked: a site of observation was kept when its current area differed less than 5% to its past area (i.e., in the historical land use layers collected above);

- categorizing the cropland sites from BDAT_2012-2019 considering if they were submitted to a conversion or not since 1910, and if yes, since when, i.e. defining the period of conversion;

- comparing SOCst.30 by periods of conversion and soil associations. To this aim, we compared SOC stocks at Equivalent Soil Mass (ESM; Ellert and Bettany, 1995), i.e. soil masses were normalized per unit area to get rid of the influence of the different soil management on soil masses (the BD). Hence, the mean SOC stocks estimated for the 0-30 cm depth in section 5.1.3 was divided by the related BD value at each site of BDAT_2012-2019. Then, we computed the relative differences to the 'control' group for a proper comparison between all the subsets (soil associations x periods of conversion).

5.2.3 Results

Landuse data were extracted to cropland and grassland sites of BDAT_2012-2019 from four different historical sources:

- the PCN, where parcels delimitations are based on historical plans from 1824. As the conversion from cropland to grassland begun mainly in early 20th century, we considered the landuse information from PCN correct for year 1910;
- the OBS (Occupation Biophysique du Sol) for 1999;
- the OBS (Occupation Biophysique du Sol) for 2007;
- the BDAT_2012-2019, where landuse at the time of sampling was given.

Hence, the period of conversion was determined for each site. Four periods of conversion were defined: before 1910, from 1910 to 1999, from 2000 to 2007, and from 2007 to present (2012-2019). All the cropland sites from BDAT_2012-2019 where landuse did not change since 1910 (i.e., no conversion occurred as they are managed as croplands since 1910) represented the 'control' category. Similarly, all the grassland sites from BDAT_2012-2019 where landuse did not change since 1910 represented the 'before 1910' category. For each soil association, the relative difference in SOC stocks was computed for each of the four period of conversion, i.e. the relative difference to the control group[15] (Tab. 5.4). The ten chronosequences (one by soil association) were grouped in Figure 5.7 - left panel – for comparison. The central year of the four period of conversion was used to define the time after conversion (in years), except for the period 'before 1910' where 1910 was kept as year of reference. The year of conversion for the control groups was set to 0 in Figure 5.7 as no conversion occurred in their case.

As observed in Figure 5.7 (left panel), SOC stocks in fields converted from cropland to grassland increased with time for the 10 soil associations. However, the increase in SOC stocks in Oesling (1), Buntsandstein (2) and Dolomies du Muschelkalk (3) soil associations was rather limited (between + 20 to 50% in 110 years; Tab. 5.4), while the majority of soil associations from Gutland have more than

[15] *The relative Hodges-Lehmann indicator was computed. It corresponds to the relative median of differences, computed based on all the possible paired combinations between the sub-group of interest and the related 'control' group.*

doubled their SOC stocks after more than 100 years of conversion (> + 100%; Tab. 5.4). Hence, soils of Oesling, Buntsandstein and Dolomies du Muschelkalk could have already been close to their maximum SOC storage capacity when converted, while the other soils associations have still storage capacity, especially in cropland soils. In 1900-1914, the Oesling showed the highest content of humic acids when compared to Gutland (§4.1.2). To finish, the SOC stocks increase seems to occur mainly within the first 20 years of conversion, and then slows down gradually. A meta-analysis from Peoplau et al. (2011) for fields converted from cropland to grassland in temperate region showed a similar trend (Fig. 5.7 – right panel). As most of the conversion occurred during the first half of the 20th century, we can expect that the spatial variability in SOCst.30 induced by the conversion has being less important with time.

Soil asso.	Relative change in SOC stocks (%) after...			
	7 years	17 years	65 years	110 years
1	10	20	11	21
2	57	-	16	55
3	-7	-	63	51
4	22	-	58	-
5	-2	36	34	140
6	37	104	62	180
7	-2	-	65	122
8	37	81	110	123
9	-14	-	86	113
10	51	-	71	102

Table 5.4: Relative median change in SOC stock (%; used for creating Fig. 5.7 above) after X years of conversion from cropland to grassland in GDL. Soil association: 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.

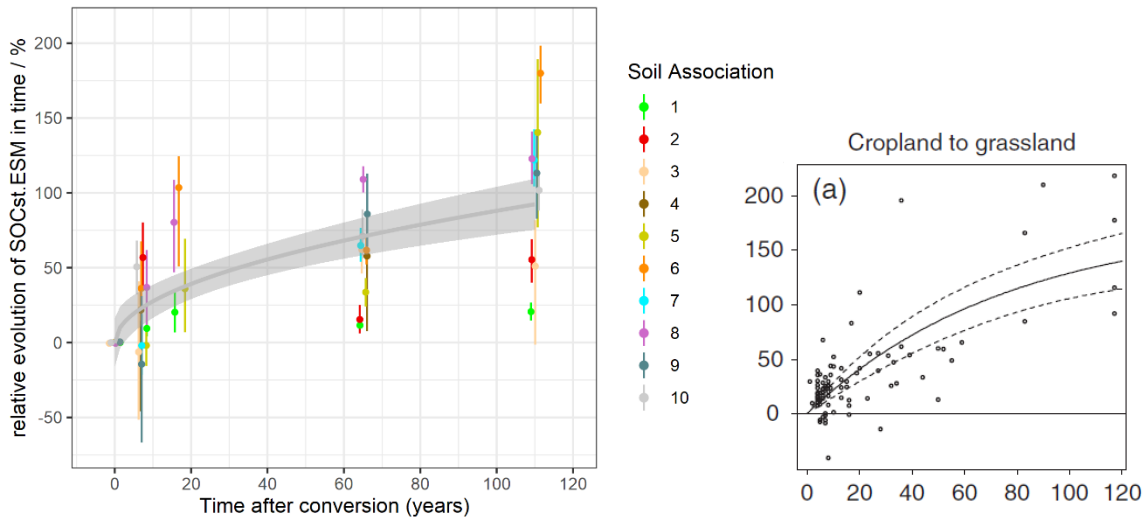


Figure 5.7 Relative change in SOC stock after conversion from cropland to grassland in GDL (left panel) and in temperate region (right panel; source: Poeplau et al., 2011). The colored points refer to the changes in each soil association (median with 90% confidence interval) while the grey line indicates the average model based on all soil associations together with the confidence limits (grey area). Soil association : 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)

5.3 Current state of SOC stock and its short-term evolution

5.3.1 Methodology

The SOC stocks for the 0-30 cm depth were estimated for all of the sites from BDAT_2012-2019 (see 5.1; step 1 in Fig. 5.8). These data were used in this section to study the recent spatio-temporal evolution of SOCst.30. Hence, a DSM approach combined to Monte-Carlo simulations was applied to predict SOCst.30 all over the GDL territory for the two periods T1 and T2 (Fig. 5.8).

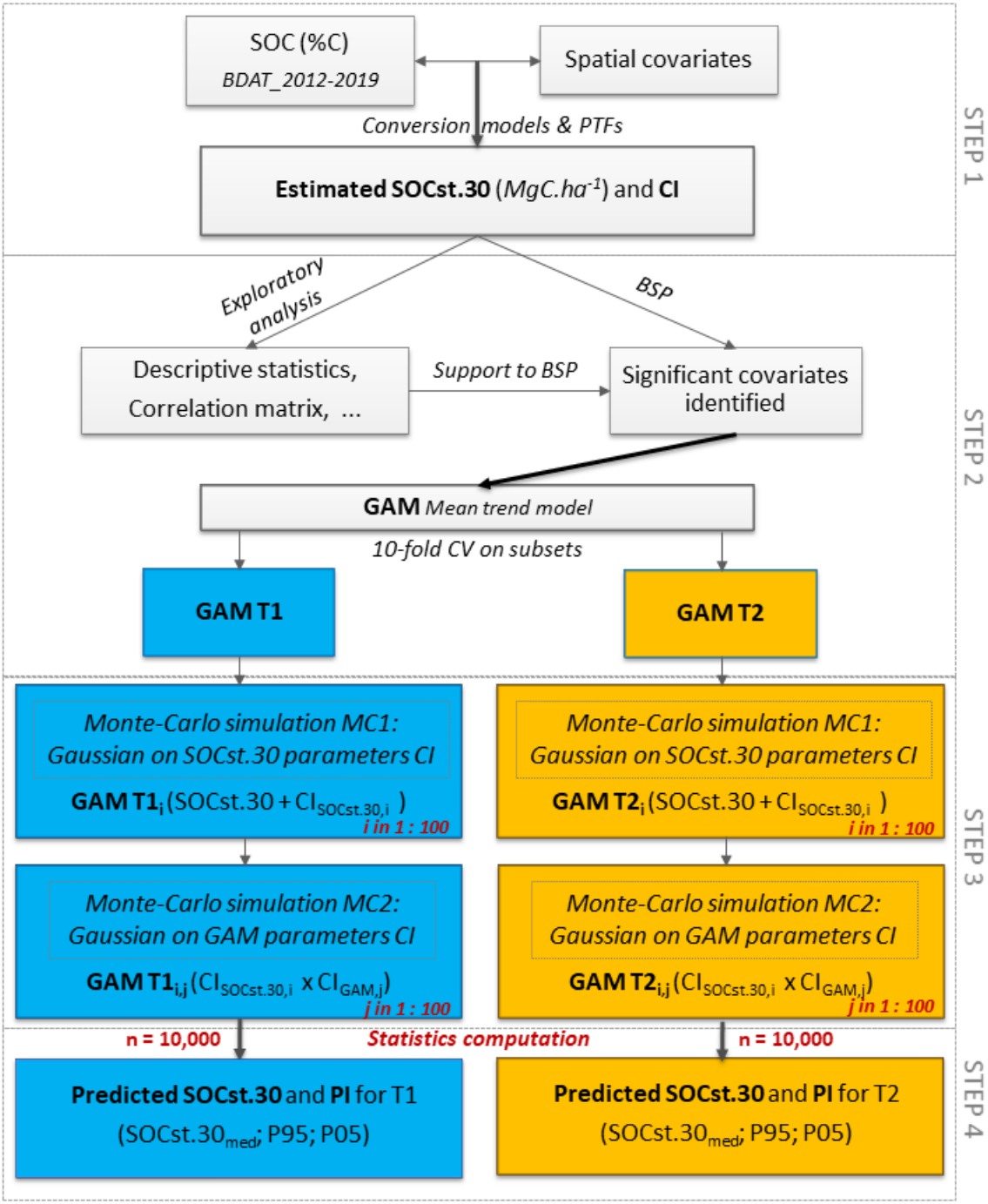


Figure 5.8: Workflow of the consecutive procedures used to predict SOC stocks for the 0-30 cm depth and associated Prediction Intervals on agricultural soils of GDL. (PTF: Pedo-Transfer Functions, BSP: Backward Stepwise Procedure, CI: Confidence Interval, CV: Cross-Validation, MC: Monte-Carlo, PI: Prediction Interval)

Predicting SOC stocks in agricultural soils all over the GDL

As for the SOC content, the mapping procedure was based on Generalized Additive Models (GAM; Hastie and Tibshirani, 1986). The GAM is a generalization of linear regression models in which the coefficients can be a set of smoothing functions accounting for the non-linearity that could exist between the dependent variable (here, SOCst.30) and the covariates (section 3.3.). A GAM model was fitted on all the estimations of SOCst.30 values computed for cropland and grassland in BDAT_2012-2019 (N = 7455; step 2 in Fig. 5.8). After the mean model calibration, this dataset was split in two periods, T1 (2012-2015) and T2 (2016-2019) for validation. We estimated the goodness-of-fit of the model for each period by computing a stratified 10-fold cross-validation (CV) on each period subset. The stratification of the CV was performed considering the 10 soil associations in order to keep a balance on their representation at each fold[16].

N.B. Unfortunately, the vineyard subset did not show significative relations between SOCst.30 and the covariates during the GAM fitting. Then, the median values of SOCst.30 by soil association in vineyard were applied.

Considering the mapping uncertainties

To properly assess the uncertainties implicated in the SOCst.30 prediction all over the GDL through the mapping procedure, the Estimation Interval (EI) associated to SOCst.30 computation (section 5.1.3) and the Confidence Interval (CI) associated to the GAMs parameters were combined here (step 3 in Fig. 5.8; as proposed by Chartin et al. in 2017 for Wallonia). For this purpose, two consecutive Monte-Carlo simulations were applied:

1. **MC1:** mean SOC.30, BD and RM values were randomly varied following a normal probability distribution with means and EI corresponding to those computed in 5.1.2 and 5.1.3. Hundred simulations were performed. SOCst.30 was computed for each simulation. Based on each of these new 100 SOCst.30 datasets, the parameters of the mean trend model were readjusted to create 100 new GAMs;

2. **MC2:** For each of the 100 GAM models produced by the first Monte-Carlo simulation, confidence intervals (CI) of the model parameters were estimated. For each of these models, 100 replicate sets of the associated parameters were created through new Monte-Carlo simulations considering the mean value of the parameters, their confidence intervals and a normal distribution. The prediction matrix relating the model parameters to the predictor were used to obtain 100 replicates of each GAM prediction on the 90m grid covering GDL.

Hence, this simulation led to the computation of 10,000 ($100 * 100$) independent realizations of the trend model. SOC stock values were then predicted at unsampled locations of Luxembourg agricultural areas by applying separately each of the 10,000 models to the spatial continuous layers of concerned covariates. So, we obtained 10,000 mapping products with a resolution of 90m × 90m.

N.B. No proper modeling procedure could be applied for SOCst.30 in vineyards; the median values by soil association were applied (Tab. 5.6). Hence, there were no uncertainties related to modeling procedure for vineyards.

[16] *For further consideration on the GAM calibration/validation procedure, please refer to Annex 8.3.*

SOCst.30 map finalization

Based on the 10,000 independent spatial estimations of SOCst.30, we computed the median ($SOCst.30_{med}$) and upper (P95) and lower (P5) percentiles for each pixel (step 4 in Fig. 5.8). The median values were used to map SOC stocks ($MgC.ha^{-1}$) in the 0-30 cm layer for all croplands and grasslands of GDL. The relative Prediction Interval ($SOCst.30_{rel.PI}$) was estimated as the ratio of the difference between the 95th and the 5th percentiles to the median, as proposed in the GlobalSoilMap project (Heuvelink, 2014; eq. 10):

$$SOCst.30_{rel.PI} = (P95 - P05)/SOCst.30_{med} \qquad \text{(Eq. 10)}$$

N.B. For vineyards, the relative uncertainties were computed considering P05, P95 and the SOCst.30 median computed by soil association in §5.1.3, as no uncertainties related to modeling procedure were available for this land use.

5.3.2 Results

Estimated SOCst.30 - descriptive statistics and differences test by periods

In BDAT_2012-2019, T2 2016-2019 (N = 4571) was represented by more observations than T1 2012-2015 (N = 2884). The distribution of data from croplands was more homogeneous for T2 than for T1 (Figs. 4.9 and 4.10), especially in the natural region of Oesling. While grasslands were represented by almost twice as much observations for period T2 than for T1, 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. These remarks could explain the differences observed in histograms (Fig. 5.9) between T1 and T2 for each land use.

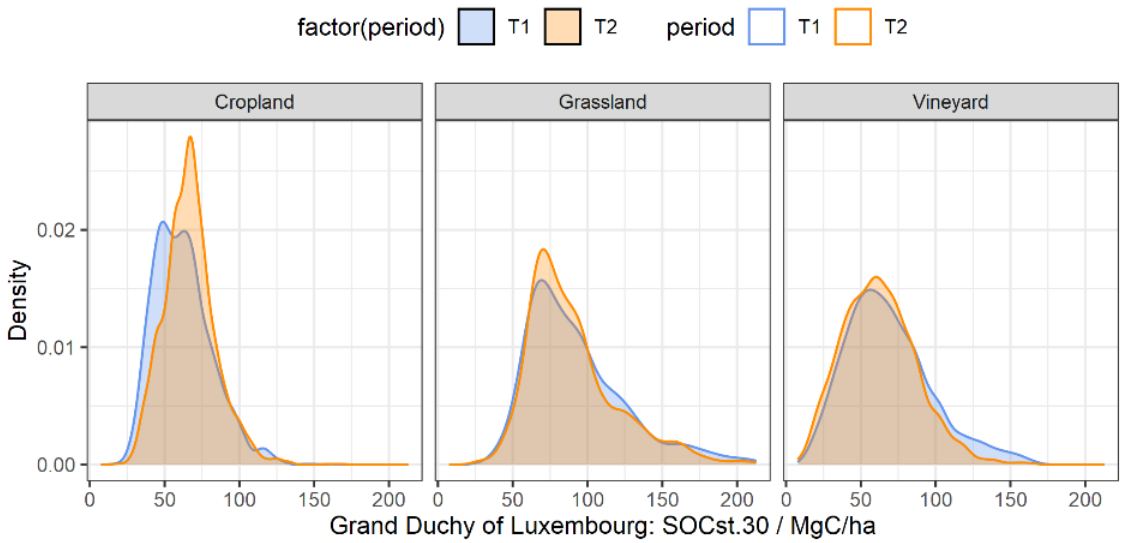


Figure 5.9: Histograms of SOC stocks in the 0-30cm layer (SOCst.30 in $MgC.ha^{-1}$) at T1 (2012-2015) and T2 (2016-2019) for croplands, grasslands and vineyards in GDL.

Overall, grassland have higher SOCst.30 (median = 82.6 MgC.ha-1) than cropland (64 MgC.ha-1) and vineyard (62.63 MgC.ha-1; Fig. 5.9). The SOCst.30 box-plots in Fig. 5.10 show the same relative trend from one soil association to another as the box plots for SOC content (Fig. 4.8).

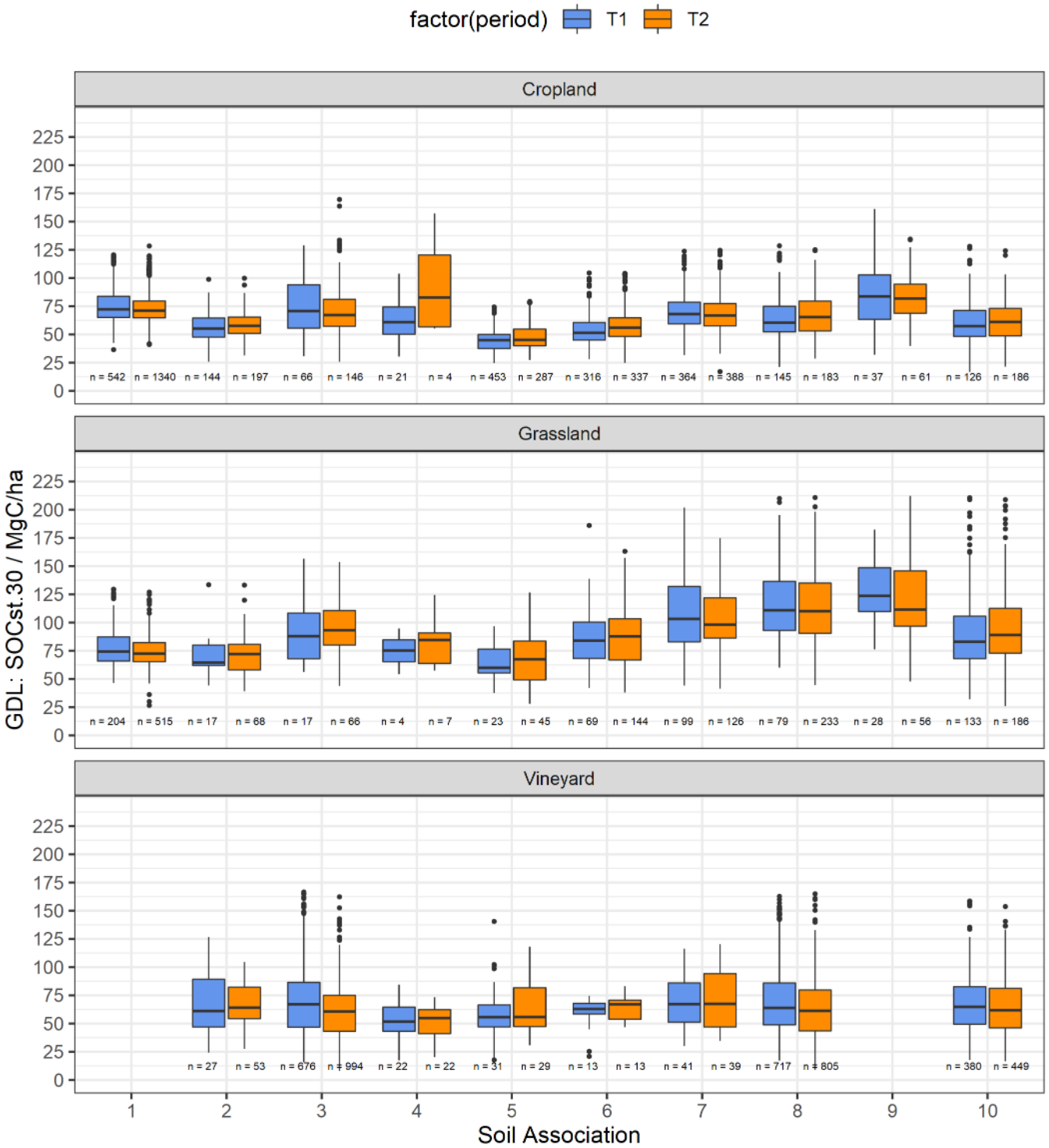


Figure 5.10: Box-plots of SOC stocks in the 0-30cm layer (SOCst.30 in MgC.ha^{-1}) in croplands, grasslands and vineyards per regrouped soil associations at periods T1 (2012-2015) and T2 (2016-2019) in Grand-Duchy of Luxembourg. (1 = Oesling, 2 = Buntsandstein, 3 = Dolomies du Muschelkalk, 4 = Calcaires 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)

Tables 5.5 to 5.7 summarize the descriptive statistics for T1 and T2, and the tested differences of distribution between both periods, of SOCst.30 (MgC.ha^{-1}) computed in 5.1.3, for cropland, grassland and vineyard, respectively.

NB: While significant statistical differences between T1 and T2 were detected for SOC content distributions in different landuse/soil association sub-groups (Tab. 4.2 to 4.4), they were not systematically detected again for SOC stocks. The procedure developed to estimate SOCst.30 (§ 5.1) implied the SOC content conversion from topsoil to 0-30 cm, and the addition of the Bulk Density and Rock fraction by Mass in the calculation (Eq. 7). Hence, the transition from contents to stocks may have reduced the differences previously observed for SOC contents.

- *Cropland*

The subset T1 showed a median SOCst.30 value of 71.4 MgC.ha^{-1} with an interquartile range (IQR) of [64.3 ; 82.1] MgC.ha^{-1} in 'Oesling' (1; Tab. 5.5). In Gutland, median values varied from 45.8 MgC.ha^{-1} (IQR = [38.2 ; 50.2] MgC.ha^{-1}) in 'Grès du Luxembourg' (5) to 84.2 MgC.ha^{-1} (IQR = [63.9 ; 103.5] MgC.ha^{-1}) in 'Argiles lourdes des schistes bitumineux' (9). For T2, the 'Oesling' (1) showed median SOCst.30 of 70.0 MgC.ha^{-1} (IQR = [63.8 ; 78.1 MgC.ha^{-1}), and Gutland had median ranging from 45.8 MgC.ha^{-1} (IQR = [40.6 ; 55.3] MgC.ha^{-1}) in 'Grès du Luxembourg' (5) to 82.8 MgC.ha^{-1} (IQR = [69.5 ; 94.8] MgC.ha^{-1}) in 'Argiles lourdes des schistes bitumineux' (9). From T1 to T2, only two soil associations in cropland went through a statistically significant evolution of SOC stock according to our data. SOCst.30 in the 'Oesling' decreased by 1.93 MgC.ha^{-1} ($p < 0.05$), while the stock increased by 4.45 MgC.ha^{-1} ($p<0.01$) in the 'Dépôts limoneux sur Grès' (Fig. 5.10).

*Table 5.5.: Descriptive statistics of estimated SOC stocks (MgC.ha^{-1} for the 0-30cm 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**
1	542	36.4	64.3	71.4	73.7	82.1	119.5	1340	40.0	63.8	70.0	71.8	78.1	120.6	**-1.93**	**<0.05**
2	144	26.1	47.7	55.4	56.9	64.8	99.3	197	31.4	51.2	58.3	58.9	65.7	100.5	2.00	N.S.
3	66	30.9	56.1	70.7	74.6	92.7	125.6	146	25.9	57.5	67.5	72.0	81.0	169.1	-2.59	N.S.
4	*21*	*31.1*	*50.2*	*61.2*	*64.4*	*72.9*	*101.7*	*4*	*55.6*	*57.0*	*82.1*	*94.3*	*119.4*	*157.4*	*29.92*	*N.S.*
5	453	25.9	38.2	45.8	45.8	50.2	75.4	287	28.1	40.6	45.8	48.0	55.3	79.7	2.18	N.S.
6	316	28.7	46.1	52.0	53.9	61.2	104.8	337	25.5	48.8	56.6	58.4	65.3	105.0	**4.45**	**<0.01**
7	364	31.8	60.0	68.8	69.6	79.0	123.5	388	17.4	57.9	67.1	68.9	77.8	124.1	-0.72	N.S.
8	145	21.8	53.0	60.9	65.7	75.7	129.2	183	28.8	53.6	65.6	68.0	80.2	125.9	2.31	N.S.
9	37	32.6	63.9	84.2	87.1	103.5	159.7	61	40.4	69.5	82.8	83.5	94.8	134.8	-3.59	N.S.
10	126	17.3	48.7	58.0	62.5	71.7	127.8	186	22.1	49.6	61.2	62.8	73.3	123.5	0.30	N.S.

- *Grassland*

For grasslands, median SOCst.30 in T2 was of 70.9 MgC.ha^{-1} (IQR = [63.8 ; 79.5] MgC.ha^{-1}; Tab. 5.6) in 'Oesling' (1), which is pretty similar to what we observed for croplands (Tab. 5.5). In Gutland, SOCst.30 in T2 ranged from 67.8 MgC.ha^{-1} (IQR = [49.4 ; 84.1] MgC.ha^{-1}) in 'Grès du Luxembourg' (5) to 110.2 MgC.ha^{-1} (IQR = [90.4 ; 134.5] MgC.ha^{-1}) in 'Argiles lourdes du Keuper' (8). No statistically significant changes between T1 and T2 could be detected in grasslands with the data of BDAT_2012-2019. Also, 5 out of the 10 soil associations presented less than 30 observations for T1 what hampered to compute proper statistics.

*Table 5.6: Descriptive statistics of estimated SOC stocks (MgC.ha^{-1} for the 0-30cm depth) in permanent **grasslands** 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**
1	204	46.2	63.3	70.6	75.5	85.4	123.2	515	26.9	63.8	70.9	72.6	79.5	122.1	-2.94	N.S
2	*17*	*44.5*	*62.1*	*64.4*	*70.3*	*77.4*	*129.5*	68	39.8	58.6	72.1	71.6	80.4	130.9	*1.34*	*N.S*
3	*17*	*56.8*	*68.2*	*86.7*	*91.4*	*106.9*	*154.8*	66	44.1	79.3	90.5	92.3	108.1	151.8	*0.90*	*N.S*
4	*4*	*53.9*	*65.3*	*75.6*	*74.9*	*85.2*	*94.3*	*7*	*57.9*	*63.9*	*84.4*	*81.6*	*88.6*	*123.8*	*6.72*	*N.S*
5	*23*	*38.2*	*55.7*	*60.4*	*64.8*	*76.7*	*96.6*	45	28.4	49.4	67.8	68.7	84.1	127.6	*3.85*	*N.S*
6	69	42.0	69.4	84.5	87.3	100.1	181.8	144	38.2	66.8	88.4	87.2	102.7	160.1	-0.12	N.S
7	99	44.5	83.2	103.2	109.0	131.2	205.8	126	42.0	86.9	97.7	102.8	120.0	173.6	-6.13	N.S
8	79	59.9	92.8	111.6	114.4	135.3	219.0	233	45.1	90.4	110.2	113.1	134.5	221.5	-1.26	N.S
9	*28*	*76.6*	*109.4*	*122.2*	*127.2*	*145.4*	*180.0*	56	48.7	96.4	109.7	118.6	146.3	224.6	*-8.63*	*N.S*
10	133	32.6	68.4	82.4	91.8	104.1	218.2	186	26.3	71.3	87.8	96.9	111.6	215.3	5.16	N.S

- *Vineyard*

Vineyards showed much less variability in SOCst.30 than croplands and grasslands (Tab. 5.7). Considering that vineyards soils are predominantly located on 'Dolomies du Muschelkalk' (3), 'Argiles lourdes du Keuper' (8) and 'Others' (10), and to a lesser extent on 'Buntsandstein' (2), their median SOCst.30 ranged around 62.5 MgC.ha^{-1} (IQR = [47.1 ; 89.3] MgC.ha^{-1}) in 'Buntsandstein' (3) to 67.4 MgC.ha^{-1} (IQR = [46.9 ; 86.4] MgC.ha^{-1}) in 'Dolomies du Muschelkalk' (3). From T1 to T2, the BDAT_2012-2019 allowed to detect that two soil associations went through a statistically significant evolution in SOCst.30: the SOC stocks in the 'Dolomies du Muschelkalk' (3) decreased by 9.70% and the SOC stocks in the 'Argiles lourdes du Keuper' (8) decreased by 5.3 %. This negative trend is partly explained by a better assessment of the TIC (Total Inorganic Carbon) since 2018, revealing that the TIC was previously underestimated by ~0.2%C.

*Table 5.7: Descriptive statistics of estimated SOC stocks (MgC.ha^{-1} for the 0-30cm 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). (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, 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**
2	*27*	*24.4*	*47.1*	*62.5*	*66.7*	*89.3*	*124.6*	53	27.6	54.7	63.9	66.1	82.6	104.9	*-0.63*	*N.S.*
3	676	15.8	46.9	67.4	70.2	86.4	163.1	994	4.3	43.2	61.0	60.5	75.0	159.7	**-9.70**	**<0.001**
4	*22*	*17.9*	*43.4*	*51.7*	*53.5*	*64.6*	*84.5*	*22*	*20.5*	*41.3*	*55.0*	*51.6*	*62.6*	*72.7*	*-1.87*	*N.S.*
5	31	18.1	47.1	56.2	60.7	66.7	141.4	*29*	*31.0*	*47.4*	*56.0*	*62.9*	*81.8*	*118.1*	*2.18*	*N.S.*
6	*13*	*21.2*	*58.8*	*63.8*	*57.7*	*68.0*	*74.4*	*13*	*46.9*	*54.2*	*67.4*	*63.8*	*70.8*	*82.8*	*6.08*	*N.S.*
7	41	30.3	51.4	67.3	68.0	86.0	116.9	39	34.7	47.1	67.3	71.4	94.6	120.3	*3.43*	*N.S.*
8	717	17.4	48.8	64.0	69.1	86.2	163.7	805	4.4	43.7	61.5	63.7	79.9	162.3	**-5.36**	**<0.001**
10	380	18.1	49.5	64.9	67.5	82.9	159.0	449	16.7	46.1	62.0	64.8	81.4	155.2	-4.00	N.S.

Model results and performance

A GAM model was calibrated on observations for cropland and grassland for both periods T1 and T2 (BDAT_2012-2019 dataset; N = 7455). The landuse was introduced in the model as a qualitative covariate. The backward stepwise procedure selected the most influencing covariates – here, by decreasing order of importance (Fig. 5.11): the land use, then the geographic coordinates and C-factor as almost ex-aequo, the pH and clay content also as almost ex-aequo, the Mg content, the minimum depth of hydromorphic features, and finally the K_2O content. This set of covariates is a fair mix of the ones selected for mapping SOC content in cropland and grassland separately (section 4.2.3 'model results and performance'). However, the pH was selected for grassland only during the SOC content mapping procedure, and was then less influent as here. As observed in section 5.1.2, the covariates Mat_Text (corresponding to the different classes of the texture triangle designed for soils of Belgium and GDL; Fig. 2.4) was selected as one of the main covariates influencing the BD and the RM (Figs. 5.3 and 5.5). As the pH, the Mat_Text class is greatly influenced by the nature of the parent material. Hence, the pH could fairly represent the spatial variability of parent material influencing SOCst.30 in GDL. The geographic coordinates were automatically implemented in the GAM model to spatialize and to avoid auto-correlation. Also, their strong implication in the model supported the main regional trend induced by the two natural regions (Oesling and Gutland) along with the elevation and the climatic conditions. The covariates selected for modeling SOCst.30 in the same type of land use in Wallonia were quite similar: the landuse, the clay and fine-silt content, the geographic coordinates, the LS-factor (combining slope gradient and length), the hydromorphic features and the temperature (Chartin et al., 2017).

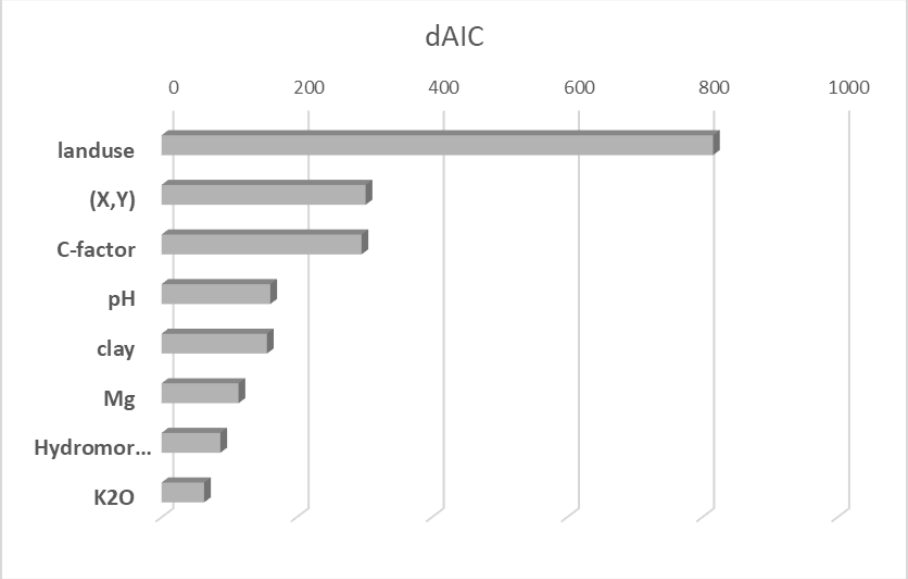


Figure 5.11: Implication of covariates in the GAM fitted on SOCst.30 (MgC.ha^{-1} for the 0-30cm depth) for cropland and grassland soils of Grand-Duchy of Luxembourg. dAIC represents the difference of AIC to the final model (Akaike Information Critrerion; Akaike, 1974). Only the covariates showing a p-value < 0.05 in the final GAM model were kept in this Figure.

The GAM model calibrated above explained 63 % of the SOCst.30 variance in croplands and grasslands of GDL. The parameters of this model were then readjusted to the T1 (N = 2884) and T2 (N = 4571) subsets, separately (Fig. 5.12). The validation procedure for T1 (2012 – 2015) gave a R^2 of 0.59, a RMSE of 16.8 MgC.ha^{-1} and a ME of +0.3 MgC.ha^{-1}. The validation results for T2 were almost similar with a

R^2 = 0.56, a RMSE = 16.4 MgC.ha^{-1} and a ME of +0.6 MgC.ha^{-1}. The SOCst.30 > 150 MgC.ha^{-1} were less accurately estimated by the model with a clear trend to underestimate (Fig. 5.12). The similar GAM approach applied to Wallonia obtained a R^2 of 0.64 and a RMSE of 16 MgC.ha^{-1} on data from a perennial monitoring network where BD and RM were measure in-field (Chartin et al., 2017). The results for GDL are very satisfactory compared to the ones obtained in neighboring countries with other approaches. For example, Goidts et al. (2009) explained 12 to 29% of the SOC stock variance in Wallonia by a digitizing approach. Meersmans et al. (2008) explained 36% of the variance in northern Belgium using a multiple linear regression approach. Later, Meersmans et al. (2011) obtained a R^2 of 0.42 (for a legacy data set of the 1960s) and 0.65 for a data set from 2005 for the entire Belgian territory. In Netherlands, Schulp and Verburg (2009) explained variances of 21 to 43% of SOC contents and stocks in different study areas while using extensive data on soil properties, topography, environmental conditions and recent historic land use as covariates. In Bavaria (Germany), Wiesmeier et al. (2014) produced a spatial prediction using a regression tree technique explaining 39% of the variance in SOC stock.

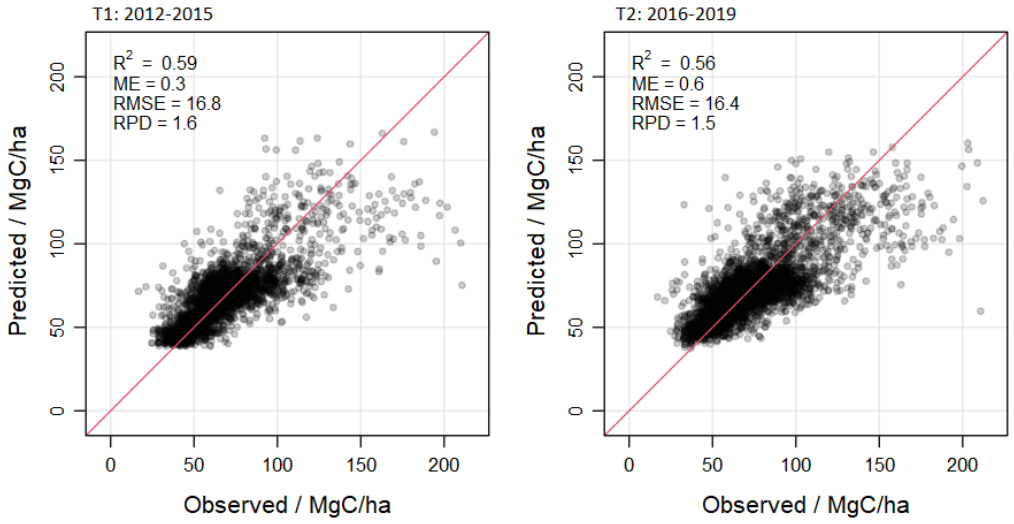


Figure 5.12: Observed vs predicted SOCst.30 (MgC.ha^{-1}) as obtained by the model fitted for cropland and grassland soils at T1 (2012-2015) and T2 (2016-2019).

Predicted SOCst.30 maps - description and comparison

The models validated above for T1 and T2 were applied to the layers of the selected covariates (90m x 90m resolution) to produce two maps of predicted SOC stocks in the 0-30cm layer under cropland, grassland and vineyard (one for each period; Fig. 5.13). The errors of estimation associated to both maps were also mapped (Fig. 5.14).

Both SOCst.30 maps have the general same patterns. Oesling and Gutland show very distinctive patterns. Oesling shows a smaller range of SOCst.30 and a more homogeneous spatial pattern than Gutland. Oesling is represented by only one soil association and a single soil texture class (OM), with relative homogenous clay content, pH and climatic conditions throughout the region. Moreover, cropland and grassland showed very similar SOCst.30 estimations with a median ~70 MgC.ha^{-1} for both types of land use (Fig. 5.10).

In the Gutland, we observe the same patterns as for SOC content (Figs. 4.20 and 4.21) mainly controlled by the land use, pH and clay content (the three textural soil types L, M and S can be clearly distinguished here; Fig. 2.3 and 2.5).

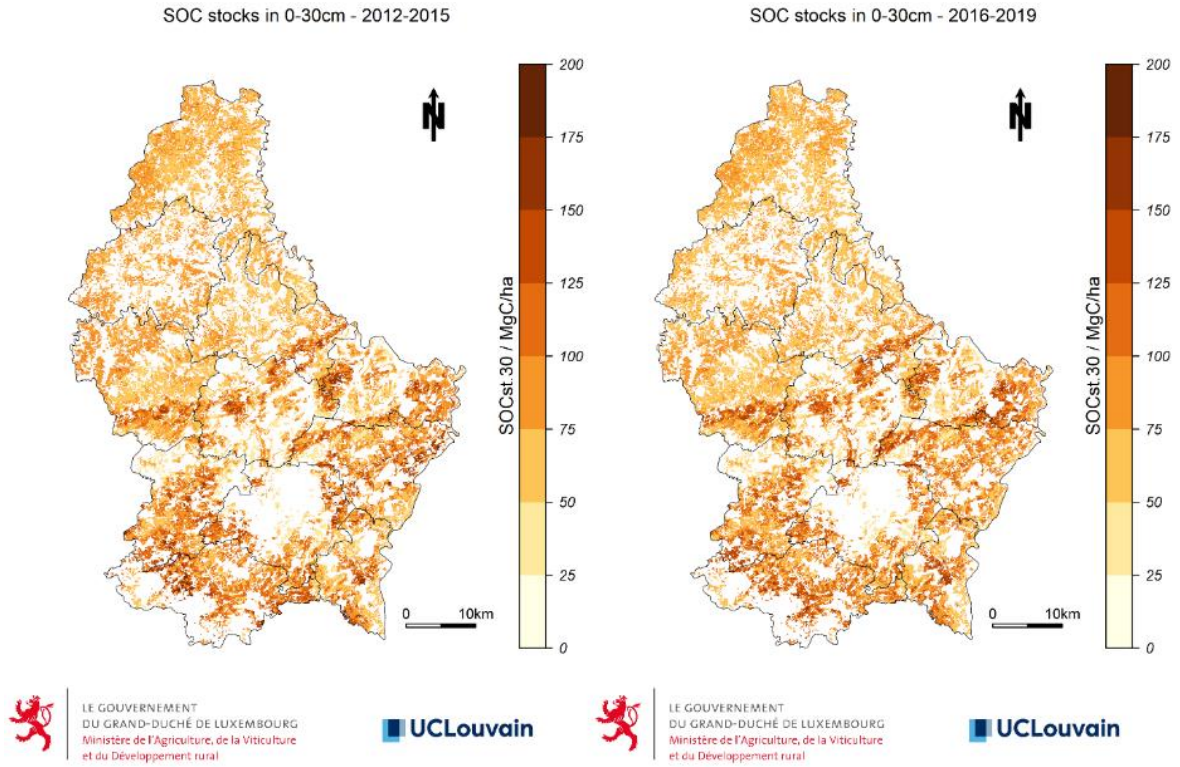


Figure 5.13: Maps of predicted SOC stocks (SOCst.30 in MgC.ha^{-1}) for the layer 0-30cm in croplands, grasslands and vineyards of GDL for periods 2012-2015 (left) and 2016-2019 (right).

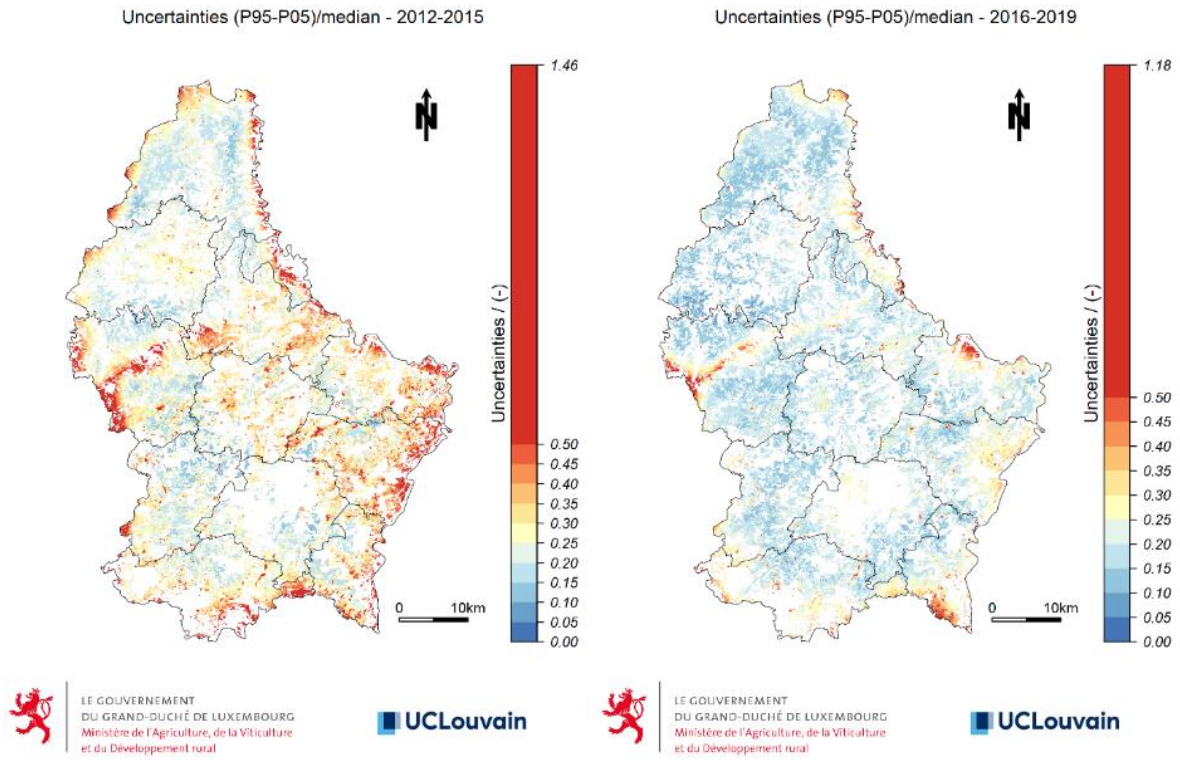


Figure 5.14: Maps of relative uncertainties (- ; relative Prediction Interval) associated to predictions of SOC stocks (SOCst.30 in MgC.ha^{-1}) for the layer 0-30cm in croplands, grasslands and vineyards of GDL for periods 2012-2015 (left) and 2016-2019 (right).

The relative uncertainties associated to SOCst.30 computing and spatial modelling were higher for T1 (median = 0.27) than for T2 (median = 0.19). The patterns in uncertainty are mainly related to the location and density of the original observations. Indeed, the uncertainties were higher in areas with a small number of observations (Figs. 5.14). As the T1 subset contained less observations than the T2 and more areas with few or no data, SOCst.30 map for T1 were related to higher uncertainties, especially at the Oesling – Gutland western limit and in the eastern and southern parts of Gutland (Fig. 5.14).

Figure 5.15 presents the significance of predicted SOCst.30 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 dSOCst.30 (i.e., $SOCst.30_{T2} – SOCst.30_{T1}$) and the standard errors (SE) of prediction for both $SOCst.30_{T1}$ and $SOCst.30_{T2}$ (deduced from Fig. 5.14). The change in SOC stock between T1 and T2 was estimated not statistically significant for 44 % of the croplands and 40 % of the grassland areas. For cropland, 11.3 % showed a statistically significant gain in SOC stock ($p < 0.05$), while more than 15 % showed a statistically significant loss. The SOC stock in grassland, statistically significantly increased for 12 % of the areas, while it decreased for 15 % of the area.

For both types of land use, loss of SOCst.30 mainly occurred in the western and eastern parts of Oesling, in the northwestern, southwestern and eastern parts of Gutland. The gains mainly occurred in Northern Oesling, and in the center, North-East and South-east of Gutland. More specifically, the patterns for each land use appeared almost similar to the changes in SOC content (see Fig. 4.22). So far, we have been able to make assumptions about these local trends but we still lack perspectives to properly explain them (see §4.2.3 p.49). Overall, despite these significant tendencies detected locally, the mean absolute differences between predicted SOCst.30 at T1 and T2 were close to 0 at the national scale with 0.41 $MgC.ha^{-1}$ for croplands and -0.15 $MgC.ha^{-1}$ for grasslands (Fig. 5.16).

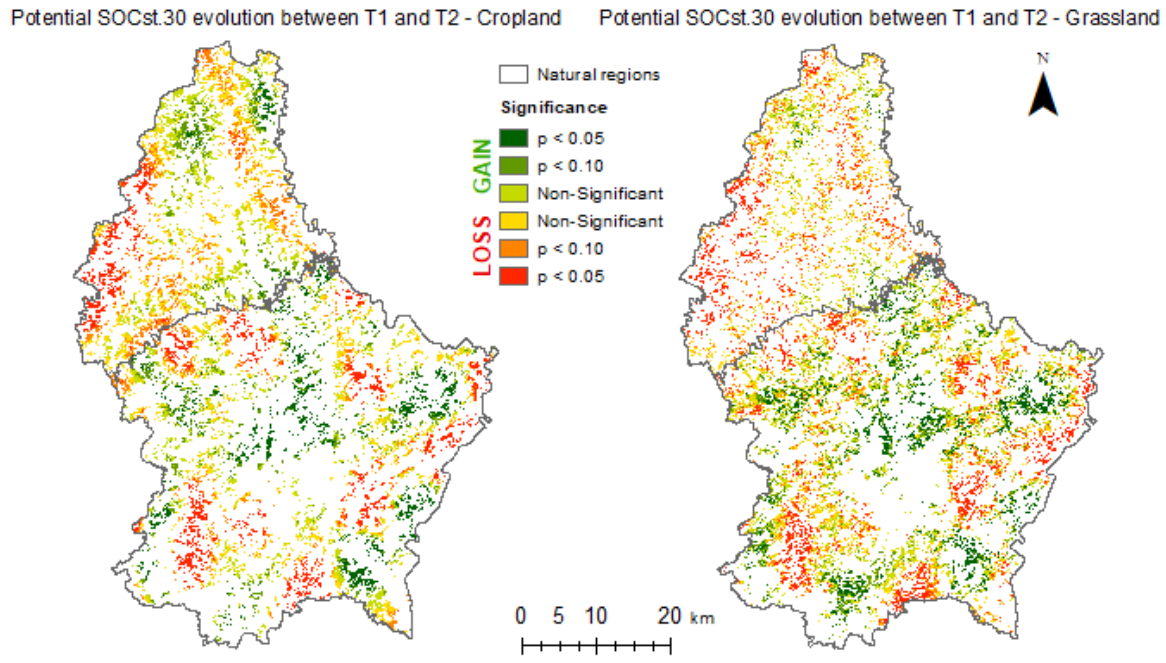


Figure 5.15: Significance of predicted SOCst.30 differences (p-value) between T1 (2012-2015) and T2 (2016-2019) for soils under croplands (left) and grasslands (right).

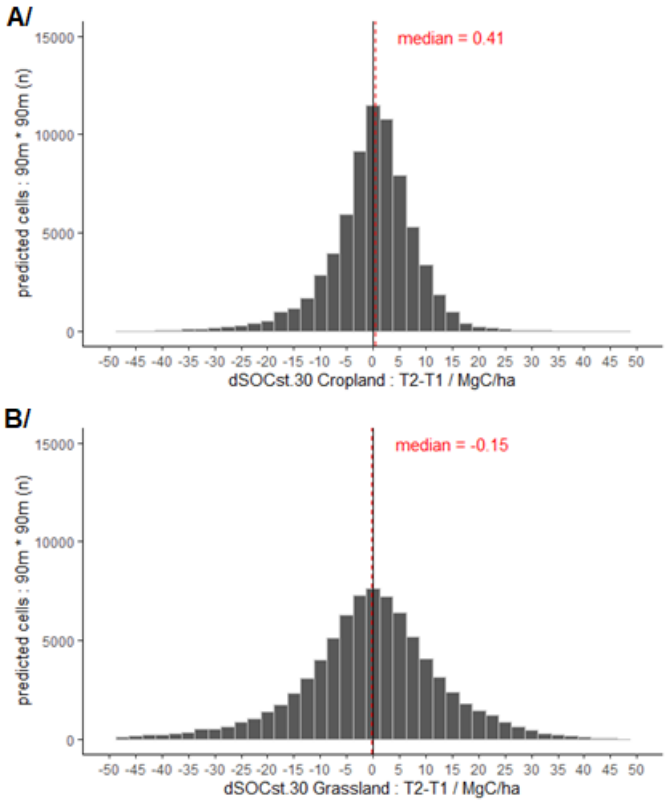


Figure 5.16: Histograms of the predicted difference in SOCst.30 (MgC.ha^{-1}) between T1 (2012-2015) and T2 (2016-2019) for soils under cropland (top) and grasslands (bottom).

Descriptive statistics of predicted SOCst.30

Based on predicted SOCst.30 maps produced above (Fig. 5.13), we used the pixel values to computed descriptive statistics of SOCst.30 predicted all over the GDL by the procedure depicted in Figure 5.8. Tables 5.8 to 5.10 summarize SOCst.30 for T1 and T2, for cropland, grassland and vineyards, respectively. As no proper modeling procedure could be applied for SOCst.30 in vineyards, the median values estimated by soil association were reported (Tab. 5.7). The relative trends between soil associations were equivalent to those observed for SOC contents (Tabs. 4.2 to 4.4) and for estimated SOCst.30 (Tabs. 5.5 to 5.7). The mean and median predicted SOCst.30 values appeared often superior to estimated SOCst.30, especially in soil associations where a low number of SOCst.30 values could be estimated. Also, we observed positive ME when validating GAM models used to predict SOCst.30 all over GDL (+0.3 and +0.6 MgC.ha^{-1} for predicted SOCst.30 for T1 and T2 respectively; Fig. 5.12), which have likely contributed to the differences observed here. From one period to another, trends were not systematically similar to those shown by estimated SOC stocks. The difference of representativity of SOC stocks (in the different soil associations and periods) between the original sampled sites (SOC stocks estimated) and the maps (SOC stocks predicted), along with the uncertainties induced the mapping procedure, could explained these differences.

Table 5.8: Descriptive statistics of predicted SOC stocks (MgC.ha^{-1}) for the layer 0-30cm in croplands for the periods T1: 2012-2015 and T2: 2016-2019, and related number of cells (90x90m) in the rasters. (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			
Assoc.	n cells	P05	median	mean	P95	P05	median	mean	P95
1	27603	63.1	73.0	74.2	89.1	63.3	72.0	72.6	83.9
2	4610	47.8	59.1	61.0	80.5	48.2	57.2	58.6	74.8
3	3784	52.3	72.1	73.4	99.1	54.9	71.5	71.4	86.0
4	330	46.8	66.9	68.2	95.6	50.9	73.4	73.7	108.1
5	5909	39.2	48.0	49.4	63.1	40.6	48.2	50.1	65.5
6	7796	43.4	57.7	59.3	79.8	45.0	57.7	59.3	77.5
7	10169	52.0	70.4	71.1	92.0	54.7	69.7	70.3	86.8
8	4898	50.1	71.6	71.1	90.0	52.2	75.1	75.8	100.0
9	1492	51.0	76.5	77.4	100.1	53.6	74.9	75.0	92.3
10	4299	46.7	67.5	66.7	87.5	47.9	67.2	66.5	85.8

Table 5.9: Descriptive statistics of predicted SOC stocks (MgC.ha^{-1}) for the layer 0-30cm in grasslands for the periods T1: 2012-2015 and T2: 2016-2019, and related number of cells (90x90m) in the rasters. (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			
Assoc.	n cells	P05	median	mean	P95	P05	median	mean	P95
1	16681	67.9	80.4	81.5	99.4	65.0	75.7	76.7	91.4
2	3530	57.6	75.5	78.6	109.5	57.2	73.4	75.6	102.7
3	5080	65.4	93.0	99.1	152.5	66.3	98.5	96.3	123.0
4	261	64.4	94.8	104.1	170.1	69.4	105.5	112.9	185.7
5	3515	46.3	67.0	70.6	108.0	52.5	68.2	72.6	113.6
6	7724	58.9	91.3	93.1	133.7	59.9	89.3	91.8	129.5
7	13694	72.7	108.4	109.0	148.1	72.7	109.2	107.0	137.8
8	13648	69.8	109.2	109.4	150.0	75.8	115.9	113.7	144.8
9	3411	82.9	119.2	122.4	167.9	88.5	122.8	121.1	146.5
10	14839	65.4	95.8	99.1	145.8	65.5	96.7	99.4	140.9

Table 5.10: Median estimated SOC stocks (MgC.ha^{-1}) for the layer 0-30cm in vineyards for the periods T1: 2012-2015 and T2: 2016-2019, and related number of cells (90x90m) in the rasters. (3 = Dolomies du Muschelkalk, 8 = Argiles lourdes du Keuper, 10 = Others)

		T1	T2
Assoc.	n cells	median	median
3	601	67.4	61.0
8	550	64.0	61.5
10	307	64.9	62.0

Based on the predicted SOCst.30 all over GDL, total budgets of SOC stocks in the 0-30cm (in TgC) were also computed as required by the UNFCCC. For cropland and grassland, their respective budgets were very similar during both periods. In cropland, the total budget of SOCst.30 was ~ 3.9 Tg C. In grassland, the budget was of ~ 6.3-6.5 Tg C.

- *Comparison with SOCst.30 predicted in the former project by Stevens et al. (2014b)*

Finally, we compared the results of SOCst.30 estimation in croplands and grasslands produced by Stevens et al. (2014b) for the period 2012-2014 and those produced here for 2012-2015 (Table 5.11). Overall, our predictions were 5 to 20 % lower than those of Stevens, except in three soil associations for croplands - 'Dépôts limoneux sur Grès' (6), 'Argiles du Lias inf. et moyen' (7) and Argiles lourdes du Keuper (8). As the period studied by Stevens et al. (2014) is pretty similar to the period T1 studied in this project ('Chartin'), the differences observed here in SOC stocks are likely to come from the differences in methodologies. Stevens et al. (2014) produced a map for each component of Equation 7 (i.e., for SOC30, BD and RM), with BD and RM estimated first with PTFs from literature, and used these maps to predict SOCst.30 at each pixel. Here, we first estimated SOC.30, BD and RM at each site physically sampled by developing conversion models and PTFs specific to GDL. Then, we produced maps of SOC stocks all over GDL. Also, we observed the PTF selected in 2014 for estimating the Bulk Density induced a bias which systematically overestimated this parameter (see Fig. 5 in Stevens et al., 2014b).

Table 5.11: median SOCst.30 (MgC.ha^{-1}) for 2012-2014 by Stevens et al. (2014b) and for 2012-2015 in this project (referred as Chartin) and their relative differences (Chartin relatively to Stevens) for cropland (in yellow) and grassland (in green). (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)

Assoc.	median Stevens	median Chartin	Relative difference
	MgC/ha		%
1	90.6	73.2	-19.2
2	63.8	59.5	-6.8
3	84.0	72.6	-13.5
4	77.7	68.2	-12.2
5	50.2	48.1	-4.2
6	57.2	57.9	1.3
7	67.7	70.5	4.2
8	65.4	71.8	9.8
9	89.2	76.8	-13.9
10	72.2	67.8	-6.1

Assoc.	median Stevens	median Chartin	Relative difference
	MgC/ha		%
1	89.8	80.8	-10.0
2	79.4	76.2	-4.1
3	116.0	94.0	-19.0
4	119.6	96.4	-19.4
5	80.0	67.4	-15.8
6	96.3	91.8	-4.7
7	122.9	108.9	-11.4
8	124.0	110.0	-11.3
9	147.2	119.8	-18.6
10	113.1	96.5	-14.7

6 CONCLUSIONS

1. **Historical geo-referenced soil datasets** for the periods 1900-1914 and 1963-1974 have been recovered together with detailed analytical protocols and historic land use maps. The SOC data for the beginning of the 20th century relate to humic acids measurements and therefore only their spatial pattern can be analyzed. The Oesling region can clearly be distinguished by the higher humic acid content. This pattern is similar to the one for the modern SOC content. The higher humic acid contents in the region around Ettelbrück do not match modern SOC content patterns. The SOC contents in the 1960s are similar to the current ones with highest values in grasslands of the clayey soil associations, in the Oesling and in the 'Dolomies du Muschelkalk'. The trends in cropland are the same but at slightly lower values compared to the grasslands. Overall, SOC contents in croplands have slightly decreased over time, while for grassland there was a slight increase in the clayey soil associations.

2. Using observations from the ASTA database BDAT for the period 2012-2019, we analyzed the **spatial variability of present soil organic carbon content (SOC)** in topsoil of croplands, grasslands and vineyards. 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 K_2O content, the minimum depth of hydromorphic features, 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 hydromorphic features, the K_2O 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 (GAM) were fitted explaining 74% of SOC variance in cropland, 40% in grassland and 14% in vineyards. The GAMs were applied to T1 (2012-2015) and T2 (2016-2019) subsets to map SOC content at both period all over GDL. Both maps have the same main patterns. Oesling has statistically significant higher SOC contents than Gutland. SOC patterns in Oesling seems mainly induced by land use repartition (related to hillslope position), while in Gutland land use repartition and clay content seem to dominate.

3. A modeling procedure has been developed for **estimating SOC stocks in the first 0-30cm (SOCst.30)**, and associated Prediction Interval (PI), based on relations between topsoil content of the BDAT_2012-2019 and environmental covariates. To this aim, models of conversion and Pedo-Transfer Functions were fitted especially for GDL, and an Ensemble Uncertainty Propagation was performed. SOCst.30 was mapped all over GDL for T1 and T2 through a specific GAM model fitted for cropland and grassland all together. Here, the covariates selected were - by decreasing order of importance: the land use, then the geographic coordinates and C-factor as almost ex-aequo, the pH and clay content also as almost ex-aequo, the Mg content, the minimum depth of hydromorphic features, and finally the K_2O content. This set of covariates is a fair mix of the ones selected for mapping SOC content in cropland and grassland separately (see point 2 above). Overall, spatial patterns in SOCst.30 were similar for both periods, but also to the ones observed for SOC content (see point 2 above).

4. However, the statistical analysis of SOC content from the BDAT_2012-2019 showed **statistically significant changes from 2012-2015 (T1) to 2016-2019 (T2)**. In cropland, topsoils of 'Oesling' lost

a mean of -0.09 %C, while those of 'Buntsandstein', 'Grès du Luxembourg', 'Dépôts limoneux sur grès', and 'Argiles lourdes du Keuper' gained +0.12, +0.07, +0.14 and +0.09 %C, respectively. In grassland, the topsoils located on alluviums and colluviums would have gained a mean of +0.26 %C. Topsoils of vineyards from the 'Dolomies du Muschelkalk' and the 'Argiles lourdes du Keuper' have lost about -0.38 and -0.16 %C, respectively - these negative trends are partly explained by a better assessment of the TIC (Total Inorganic Carbon) since 2018, revealing that the TIC was previously underestimated by ~0.2%C. Although the conversion protocol from topsoil SOC content to SOC stocks in the 0-30 cm depth induced new parameters (as BD and RM) and many uncertainties, some of these trends have been detected for SOCst.30. Cropland soils of 'Oesling' lost a mean of -1.93 $MgC.ha^{-1}$ while those of 'Dépôts Dépôts limoneux sur Grès' gained +4.45 $MgC.ha^{-1}$. Vineyards soils of 'Dolomies du Muschelkalk' and the 'Argiles lourdes du Keuper' have lost about -9.70 and -5.63 $MgC.ha^{-1}$, respectively. The statistically significant trends in SOC content were not detected in SOCst.30 for grasslands. To finish, total budgets of SOC stocks in the 0-30cm (in Tg C) for cropland and grassland have not changed over the last ten years. In cropland, the total budget of SOCst.30 was 3.9 Tg C. In grassland, the budget was ~6.3-6.5 Tg C. Hence, SOC stocks in GDL appear to be stable at the national scale since 2012, while local statistically significant increase or decrease occurred. In the future, more attention should be payed to the regions submitted to these recent statistically significant losses. In addition, the BD and the RM should be more prospected in-fields in a wider range of soil types and geology in GDL in order to improve the goodness-of-fit of the Pedo-Transfer Functions (i.e., decrease the uncertainties) developed here to estimate SOC stock in GDL.

5. By combining observations from the ASTA database BDAT 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 fields where GAPs were not applied. Statistically 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. Reduced tillage lead to higher SOC content but changes were not statistically significant. There was no effect of cover crops on SOC contents. 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.

6. Using conditional inference trees, we studied the **relative importance of environmental covariates vs management practices (GAP)** on present 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 in Gutland.

7. We explored the **impact of the historical conversion of cropland to grassland** on the spatial and temporal variability of SOCst.30. To this aim, we constructed chronosequences. The historical conversion led to a rather limited increase of SOC stocks in the converted areas of Oesling, Buntsandstein and Dolomies du Muschelkalk regions, while the converted areas in clayey regions of Gutland have more than doubled their SOC stocks after conversion. Hence, soils of Oesling, Buntsandstein and Dolomies du Muschelkalk could have already been close to their maximum SOC storage capacity when converted, while other soils associations had still storage capacity. The SOC stocks increase occurs mainly within the first 20 years of conversion, and then slows down gradually. As most of the conversion occurred during the first half of the 20th century, we can expect that the spatial variability in SOCst.30 induced by the conversion has being less important with time.

8. **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, 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.

7 REFERENCES

AKAIKE H., 1974. A NEW LOOK AT THE STATISTICAL MODEL IDENTIFICATION. IEEE TRANS. AUTOM. CONTROL 19, 716–723.

ASCHMAN C. AND FABER H., 1899. ON THE ESTIMATION OF HUMUS. CHEM.ZTG.23, N0.7, P.61.

BAH B. AND 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 - MINISTÈRE DE L'AGRICULTURE, LUXEMBOURG. 81 P.

BAVEYE PC, SCHNEE LS, BOIVIN P, LABA M, RADULOVICH R, 2020 SOIL ORGANIC MATTER RESEARCH AND CLIMATE CHANGE: MERELY RE-STORING CARBON VERSUS RESTORING SOIL FUNCTIONS. FRONT. ENVIRON. SCI. 8:579904. DOI: 10.3389/FENVS.2020.579904

BIVAND R.S., PEBESMA E.J., GOMEZ-RUBIO V., 2013.APPLIED SPATIAL DATA ANALYSIS WITH R. SPRINGER, NEW YORK, USA.

BHUPINDER P. SINGH, RAJ SETIA, MARTIN WIESMEIER, ANITHA KUNHIKRISHNAN, 2018. CHAPTER 7 - AGRICULTURAL MANAGEMENT PRACTICES AND SOIL ORGANIC CARBON STORAGE, EDITOR(S): BRAJESH K. SINGH, SOIL CARBON STORAGE, ACADEMIC PRESS, 2018, PAGES 207-244, ISBN 9780128127667, HTTPS://DOI.ORG/10.1016/B978-0-12-812766-7.00007-X.

BOER E.P.J., DE BEURS K.M., 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.

CHARTIN C., STEVENS A., GOIDTS E., KRÜGER I., CARNOL M., VAN WESEMAEL B., 2017. MAPPING SOIL ORGANIC CARBON STOCKS AND ESTIMATING UNCERTAINTIES AT THE REGIONAL SCALE FOLLOWING A LEGACY SAMPLING STRATEGY (SOUTHERN BELGIUM, WALLONIA). GEODERMA REGIONAL 9, 73-86.

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.

Ellert B.H. and Bettany J.R., 1995. Calculation of Organic Matter and Nutrient Stored in Soil Contrasting Management regimes. Canadian Society of Soil Science 75(4), 529-538.

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.

Friedman J.H., 2002. Stochastic gradient boosting. Computational Statistics and Data Analysis 38(4), 367-378.

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.

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

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., 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.

Keiluweit, M., Wanzek, T., Kleber, M. et al., 2017. Anaerobic microsites have an unaccounted role in soil carbon stabilization. Nat Commun 8, 1771. https://doi.org/10.1038/s41467-017-01406-6

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., 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., 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.

NAU J., 2015. CONRTIBUTION À LA CARACTÉRISATION PHYSIQUE ET PHYSICO-CHIMIQUE DES SOLS CULTIVÉS DU GRAND-DUCHÉ DE LUXEMBOURG : APPLICATION À LA RÉGION DE L'OESLING. TRAVAIL DE FIN D'ETUDE, GEMBLOUX AGRO-BIO TECH – UNIVERSITÉ DE LIÈGE, 121P.

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.

POEPLAU C., DON A., VESTERDAL L, LEIFELD J., VAN WESEMAEL B., SCHUMACHER J., GENSIOR A., 2011. TEMPORAL DYNAMICS OF SOIL ORGANIC CARBON AFTER LAND-USE CHANGE IN THE TEMPERATE ZONE – CARBON RESPONSE FUNCTIONS AS A MODEL APPROACH. GLOBAL CHANGE BIOLOGY 17, 2415-2427. DOI: 10.1111/J.1365-2486.2011.02408.X

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., 2014a. 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., 2014b. 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.

Thull. a., 1939. Der Freie Luxemburger Bauer und seine Scholle. Buchdruckerei Ch.Hermann, Luxemburg, 152p.

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 A.R., Jones R.J.A., 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.

Vermiere, R., 1967. Oppervlaktegeologie en bodemgesteldheid van het westelijk Gutland (GrootHertogdom Luxemburg), Thèse de doctorat, Rijksuniversitéit Gent, Faculteit der Wetenschappen, Deel 1, 108 p.

VERMEIRE, R., 1967. OPPERVLAKTEGEOLOGIE EN BODEMGESTELDHEID VAN HET WESTELIJK GUTLAND (GROOTHERTOGDOM LUXEMBURG), THÈSE DE DOCTORAT, RIJKSUNIVERSITÉIT GENT, FACULTEIT DER WETENSCHAPPEN, DEEL 2, 267 P.

VERMEIRE, R., 1967. OPPERVLAKTEGEOLOGIE EN BODEMGESTELDHEID VAN HET WESTELIJK GUTLAND (GROOTHERTOGDOM LUXEMBURG), THÈSE DE DOCTORAT, RIJKSUNIVERSITÉIT GENT, FACULTEIT DER WETENSCHAPPEN, DEEL 3 (TABELLEN, FIGUREN, KAARTEN), 38 P.

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.

WANG G., ZHOU Y., XU X., RUAN H., WANG J., 2013 TEMPERATURE SENSITIVITY OF SOIL ORGANIC CARBON MINERALIZATION ALONG AN ELEVATION GRADIENT IN THE WUYI MOUNTAINS, CHINA. PLOS ONE 8(1): E53914. https://doi.org/10.1371/journal.pone.0053914

WANG, X., CAMMERAAT, E. L., ROMEIJN, P., KALBITZ, K., 2014. SOIL ORGANIC CARBON REDISTRIBUTION BY WATER EROSION--THE ROLE OF CO2 EMISSIONS FOR THE CARBON BUDGET. PLOS ONE, 9(5), E96299. HTTPS://DOI.ORG/10.1371/JOURNAL.PONE.0096299

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.

8 Annexes

8.1 HUMIC ACID ANALYSIS IN EARLY 20TH CENTURY

Source: Aschman C. et Faber H., 1899. On the estimation of Humus. Chem.Ztg.23, n0.7, p.61.

Noch bei 0,01 Proc. (Sättigungsconcentration) hindert es die Entwickelung von Schimmel auf guten Nährsubstraten.

Das für Pilze ziemlich schädliche Zimmtöl, in wässeriger sehr verdünnter Auflösung (die eigentlich 0,1-proc. sein sollte, aber wegen Ausscheidung eines Theiles des Zimmtöles beim Eingiessen der alkoholischen Lösung von 1 g Oel in ½ l Wasser viel schwächer war[1]), erwies sich auch als giftig für Mikroorganismen anderer Art. In einem Süsswasserschlamm hatte ich neben Oscillarien (Spaltalgen) als Hauptbestandtheil auch kleine Chlorophyceen wie Palmella, ferner Diatomeen, Infusorien, Spirillen und zahlreiche andere Spaltpilze. Sie alle starben bei 36-stündigem Liegen in jener Auflösung ab, bis auf wenige sehr kleine Bakterien, die noch etwas Beweglichkeit zeigten. Da das Zimmtöl den Zimmtaldehyd als Hauptbestandtheil enthält, so könnte man bei der Giftigkeit vieler Aldehyde zunächst an die Aldehydgruppe als die Ursache der Giftigkeit denken. Aber zweifellos ist sie nicht allein daran schuld. $C_6H_5.CH:CH.CHO$ Zimmtaldehyd. Sehr wahrscheinlich trägt auch die doppelte Bindung in der Gruppe CH:CH mit zur Giftigkeit bei; denn Körper mit doppelter Bindung erweisen sich allgemein giftiger als die entsprechenden gesättigten Substanzen. Ferner ist auch die Phenylgruppe C_6H_5, mit ihren doppelten Bindungen, als Ursache des Giftcharakters anzusehen. So ist ja auch die Benzoësäure ein starkes Gift, sie ist viel stärker giftig, als sie gemäss ihrer sauren Reaction sein würde; ersetzen wir die Phenylgruppe darin durch die Methylgruppe, so haben wir die Essigsäure, welche bekanntlich nur schwach giftig ist, d. h. nur gemäss ihrem Säurecharakter. Die mit Basen neutralisirte Essigsäure ist nicht giftig, sondern eine ziemlich gute Kohlenstoffnahrung für viele Pilze.

(Schluss folgt.)

Mittheilungen aus der analytischen Praxis.

Zur Bestimmung der Humussubstanz in der Ackererde.

(Mittheilung aus der Versuchsstation Ettelbrück, Luxemburg.)

Von **Dr. C. Aschman** und **H. Faber.**

Gegenüber der Umständlichkeit der üblichen Untersuchungsmethoden (Elementaranalyse, Oxydation durch Chromsäure) dürfte nachfolgendes, einfaches und zuverlässiges Verfahren zur Ermittelung der Humussubstanz im Boden von ziemlich allgemeinem Interesse sein. Dasselbe beruht auf der Löslichkeit der Humussäuren in alkalischen Flüssigkeiten und deren nachheriger Titration durch Kaliumpermanganat nach Art der Bestimmung der organischen Substanz im Wasser.

Hierzu sind folgende Lösungen erforderlich: 1. Eine Lösung von Natriumhydroxyd in Wasser: 50 g pro 1 l. 2. Chamäleonlösung: 0,32 g reines, übermangansaures Kali werden in Wasser gelöst und zum Liter aufgefüllt. 3. Oxalsäurelösung: 0,63 g pro 1 l. 4. Verdünnte Schwefelsäure 1:5.

Behufs Extraction der Humussubstanz werden 25 g lufttrockener, fein gesiebter Boden mit 100 ccm Natronlauge übergossen und eine Stunde lang auf dem kochenden Wasserbade digerirt. (Zweckmässig werden hierzu Porzellanschalen benutzt, da Glaskolben unter Einwirkung der caustischen Lösung sehr leicht springen.) Zur vollständigen Extraction der Humussäure ist ein mehrmaliges Decantiren und Auslaugen erforderlich. Dann füllt man den Gesammtinhalt auf 510 ccm (10 ccm werden durch den Bodensatz eingenommen, so dass die Flüssigkeit in Wirklichkeit auf 500 ccm vertheilt ist), mischt sorgfältig und hebt von dieser Lösung, die bis zur vollständigen Klärung einige Tage stehen muss, einen beliebigen Theil auf. Ein Filtriren derselben ist nicht zu empfehlen.

Zur Titerstellung des Chamäleons benutzt man reine Humussäure (Acid. huminic. pur., erhältlich bei E. Merck-Darmstadt). Man löst davon 0,125 g in einem Halbliterkolben vermittelst Natronlauge und füllt mit Wasser nach bis zur Marke. Da eine deutliche Endreaction nur in Gegenwart von Oxalsäure stattfindet, so ist zunächst durch einen Vorversuch festzustellen, wieviel Chamäleon durch 10 ccm Oxalsäure absorbirt wird. Es werden also 5 ccm Chamäleon nebst 10 ccm Schwefelsäure mit 100 ccm destillirtem Wasser gemengt, 5 Minuten lang gekocht, dann 10 ccm Oxalsäure hinzugefügt und titrirt. Verbrauch: 11,2 ccm Chamäleon. In gleicher Weise werden nun 5 ccm der reinen Humussäurelösung mit 100 ccm Wasser titrirt, indem man allmälich so viel Chamäleon zusetzt, dass bei längerem Kochen keine Entfärbung mehr stattfindet, nach 5 Minuten 10 ccm Oxalsäure beifügt und nun vorsichtig übermangansaures Kali zufliessen lässt bis zum Eintreten der Endreaction, welche stets sehr scharf ist. Es sei z. B. der Chamäleonverbrauch 21,9 ccm, so wird sich die Berechnung stellen, wie folgt:

21,9 — 11,2 = 10,7 ccm Chamäleon, entsprechend 5 ccm Humuslösung. 500 ccm erfordern 1070 ccm Chamäleon. 1070 : 0,125 g = 1 : x = 0,000116. 1 ccm Chamäleon entspricht also 0,000116 g fester Humussäure.

Bei Ausführung der Bestimmung ist es zum Erzielen genauer Resultate wichtig, darauf zu achten, dass die zu untersuchende Humuslösung keine erheblich höhere Concentration besitze, als die bei der Titerstellung benutzte, reine Humussäurelösung. Eine sonst unvermeidliche doppelte Titration ist indess hier nicht erforderlich, da die Intensität der Färbung den Concentrationsgrad leicht erkennen lässt. Man vergleicht einfach die beiden Lösungen in Reagenzröhrchen von gleicher Weite und verdünnt gegebenenfalls bis zum Eintritt einer übereinstimmenden Farbennüance. Von der extrahirten Humuslösung werden nun 5 ccm titrirt, wie vorhin angegeben. So wurden u. A. folgende Resultate gewonnen:

Guter Gartenboden enthielt 1,87 Proc. Humussäure. Sehr stark gedüngte Mistbeeterde 5,5—11,5 Proc. Waldboden: Obere Schicht 0,88 Proc., Untergrund 0,3 Proc. Boden aus Lohbeeten: Obere Schicht 1,25 Proc., Untergrund 0,44 Proc. Der Feldboden weist erhebliche Differenzen auf, zwischen 0,05 und 1,7 Proc., Mittelgehalt meist 0,5—0,7 Proc.

Es liegt klar vor Augen, dass an Hand dieser sehr leicht ausführbaren und praktischen Methode werthvolle Anhaltspunkte für die Beurtheilung eines Ackerbodens erhalten werden können.

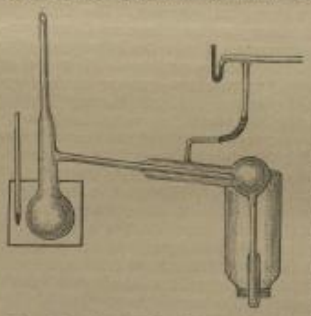
[1] Nach einer ungefähren Bestimmung nur 1 : 35 000.

Apparat zur Destillation unter stark vermindertem Druck mit einer Wasser-Quecksilberluftpumpe.

Von **Dr. L. T. C. Schey**-Leiden (Holland).

Mehr und mehr wird das Destilliren bei sehr geringem Drucke mit Hülfe der Wasser-Quecksilberluftpumpe eine gebräuchliche Operation bei organischen Synthesen. Nicht nur findet sie Anwendung beim Destilliren von solchen Stoffen, die nur im absoluten Vacuum zu destilliren sind, sondern auch in denjenigen Fällen, in denen eine Substanz kaum die Temperatur ertragen kann, bei welcher sie beim Vacuum der Wasserpumpe übergeht, ist sie empfehlenswerth. Mit der Quecksilberluftpumpe bekommt man ein grösseres Vacuum, und der Siedepunkt wird sehr herabgedrückt, wodurch eine Zersetzung der Substanz vermieden wird. Aber dann soll die Operation auch praktisch sein und nicht viel mehr Mühe kosten, als wenn man die Wasserpumpe verwendet.

Als ich bei einer Arbeit, die ich nächstens zu veröffentlichen hoffe, die synthetische Darstellung von Fetten, öfters mit der Quecksilberluftpumpe destilliren und fractioniren musste, stellte sich heraus, dass der nachstehend beschriebene Apparat gute Dienste leistet. Die alleinige Anfertigung habe ich der Firma Dr. H. Geissler Nachf. Franz Müller, Bonn a. Rh. übergeben, und der Apparat ist unter D. R. G. M. 104 144 gesetzlich geschützt. Das Princip ist die Vermeidung von Verbindungen, weil die Quecksilberluftpumpe, wenn der Apparat nicht absolut luftdicht schliesst, das grosse Vacuum nicht leisten kann. Die Anzahl der Verbindungen des ganzen Apparates ist auf eine reducirt. Das Destillirkölbchen hat einen langen Hals, und die Substanz wird mittelst eines Trichters mit langem Stiele eingeführt, und dann der Hals vor dem Gebläse capillar ausgezogen. Der Recipient ist an das Kölbchen festgeschmolzen und hat die Form eines Destillirkölbchens, dessen Seitenröhre durch einen Kautschukschlauch mit der Pumpe verbunden ist. An die Kugel ist eine zweite Röhre angesetzt, welche in einen weiteren Theil endet, der das Destillat aufnehmen soll. Der ganze Recipient ist umgeben von einem Kühlgefäss, welches zweckmässig mit Eis gefüllt wird, so dass die Kugel auch vollkommen von Eis umgeben ist. Man beobachtet nur die Temperatur des Heizbades (zweckmässig ein Luftbad). Ist die Destillation beendet, oder will man eine neue Fraction auffangen, so schiebt man über die Capillare des Destillirkölbchens ein Stückchen Kautschuk mit Quetschhahn, zertrümmert die Spitze und lässt trockne Luft ein. Dann schneidet man die Röhre, welche das Destillat aufgenommen hat, irgendwo am engeren Theil ab, nachdem man die Kühlvorrichtung entfernt hat. Will man den Apparat wieder benutzen, so schmilzt man an die abgeschnittene Stelle wieder einen neuen Recipienten an. Der lange Hals des Destillirkölbchens wird jedes Mal kürzer, ist er verbraucht bis an den breiteren Theil, so kann man wieder eine neue Röhre ansetzen. Sehr zweckmässig ist es beim Fractioniren, über zwei Apparate zu verfügen. Ist eine Fraction abdestillirt und aus dem Apparat in der beschriebenen Weise entfernt, so giesst man den Rückstand sofort in den zweiten und kann weiter fractioniren, während der erste gereinigt wird.



8.2 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 the detailed soil map of GDL (as defined on the texture triangle designed for soils of Belgium and GDL; Fig. 2.4).

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[17], 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.

[17] *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.*

8.3 DETAIL PROTOCOL FOR SOC CONTENT MODELING: GAM CALIBRATION AND VALIDATION

The BDAT_2012-2019 dataset 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 BDAT_2012-2019dataset 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[18].

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 stepwise procedure (BSP) for selecting the terms using their approximate p-values. This was done by sequentially dropping the single term with the highest non-statistically significant p-value from the model and re-fitting until all terms are statistically 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

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

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 the best measure of the map accuracy, ideally assessed using a set of external samples taken from randomly selected locations (Loveland and Webb, 2003).

8.4 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
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
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
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**

8.5 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).

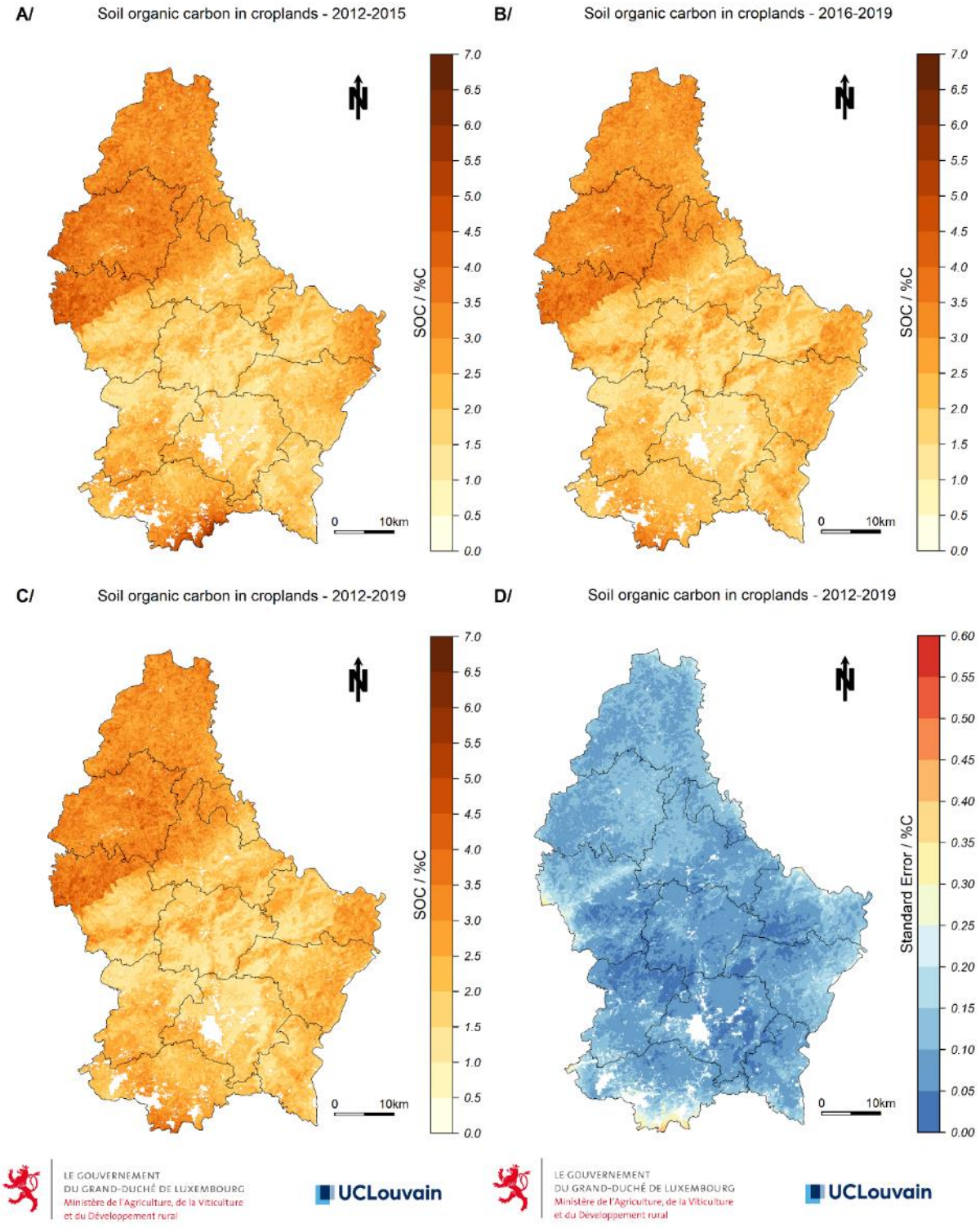


*Figure Anx1: additional maps of SOC considering the whole territory as **croplands** – A/ SOC (%C) for period 2012-2015, B/ SOC (%C) for period 2016-2019, C/ SOC (%C) for period 2012-2019 and D/ standard error of estimation (%C) for period 2012-2019.*

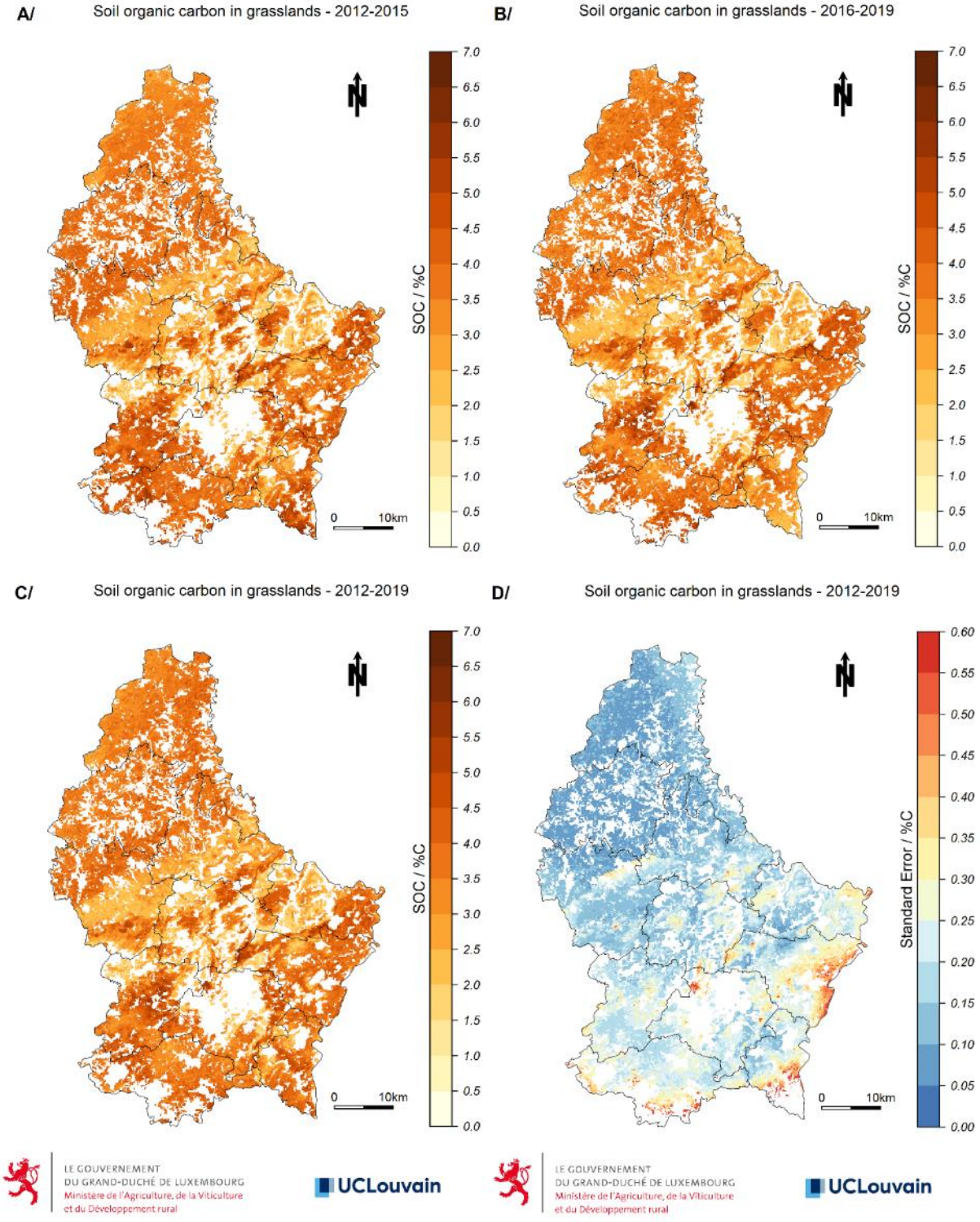


*Figure Anx2: additional maps of SOC considering the whole territory as **grasslands** – A/ SOC (%C) for period 2012-2015, B/ SOC (%C) for period 2016-2019, C/ SOC (%C) for period 2012-2019 and D/ standard error of estimation (%C) for period 2012-2019. (White areas are not covered by the pH map - Fig.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

8.6 DESCRIPTIVE STATISTICS FOR VARIABLES REQUIRED FOR CONVERSION SOC CONTENT TO SOC STOCK

8.6.1 SOC content in the 0-30cm layer

Table Anx6: Descriptive statistics of SOC.30 (SOC in %C for 0-30cm) for cropland estimated from the application of the linear model lmSOC.30$_{crop}$ to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	1882	1.12	2.33	2.79	2.82	3.26	5.58
2	341	0.66	1.31	1.49	1.57	1.77	3.17
3	212	0.66	1.40	1.77	2.00	2.45	4.47
4	25	0.75	1.36	1.54	2.05	2.79	4.56
5	740	0.57	0.93	1.03	1.07	1.22	1.96
6	653	0.57	1.12	1.31	1.36	1.49	2.79
7	752	0.38	1.40	1.68	1.72	1.96	3.63
8	328	0.47	1.22	1.49	1.62	1.96	3.63
9	98	0.75	1.61	2.09	2.13	2.52	4.65
10	312	0.38	1.22	1.54	1.68	2.14	4.28

Table Anx7: Descriptive statistics of SOC.30 (SOC in %C for 0-30cm) for grassland estimated from the application of the linear model lmSOC.30$_{grass}$ to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	719	0.83	2.50	2.88	2.94	3.33	5.76
2	85	1.06	1.66	1.97	2.13	2.35	4.32
3	83	1.21	2.19	2.80	2.78	3.33	4.55
4	11	1.36	1.74	2.19	2.19	2.50	3.33
5	68	0.68	1.21	1.51	1.75	2.21	3.41
6	213	0.90	1.81	2.27	2.41	2.95	5.76
7	225	1.06	2.27	2.95	2.98	3.63	6.06
8	312	1.06	2.45	3.10	3.13	3.79	6.14
9	84	1.13	2.72	3.33	3.42	4.01	6.21
10	319	0.68	2.12	2.65	2.85	3.33	6.59

8.6.2 Bulk Density

Table Anx8: Descriptive statistics of Bulk Density (BD in g.cm³) for cropland estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	1882	1.07	1.10	1.14	1.18	1.26	1.47
2	341	1.09	1.32	1.37	1.37	1.41	1.51
3	212	1.05	1.28	1.37	1.34	1.42	1.53
4	25	1.09	1.23	1.40	1.33	1.42	1.51
5	740	1.32	1.48	1.50	1.49	1.52	1.60
6	653	1.14	1.41	1.44	1.44	1.48	1.57
7	752	1.09	1.38	1.43	1.41	1.46	1.58
8	328	1.23	1.42	1.45	1.44	1.49	1.58
9	98	1.23	1.34	1.41	1.39	1.45	1.57
10	312	1.08	1.33	1.40	1.38	1.46	1.59

Table Anx9: Descriptive statistics of Bulk Density (BD in g.cm³) for grassland estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	719	0.97	1.02	1.07	1.11	1.21	1.45
2	85	0.99	1.21	1.27	1.25	1.33	1.40
3	83	1.01	1.08	1.22	1.20	1.28	1.46
4	11	1.07	1.27	1.29	1.28	1.30	1.37
5	68	1.06	1.30	1.41	1.36	1.44	1.51
6	213	1.03	1.23	1.30	1.28	1.36	1.51
7	225	1.03	1.18	1.23	1.24	1.31	1.46
8	312	0.99	1.23	1.26	1.26	1.30	1.45
9	84	1.08	1.16	1.24	1.23	1.27	1.47
10	319	0.98	1.14	1.22	1.22	1.30	1.50

TableAnx10 Descriptive statistics of Bulk Density (BD in g.cm³) for vineyard estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
2	80	1.11	1.30	1.37	1.35	1.43	1.50
3	1670	1.06	1.28	1.37	1.34	1.43	1.59
4	44	1.11	1.30	1.39	1.38	1.46	1.57
5	60	1.30	1.43	1.45	1.44	1.47	1.56
6	26	1.18	1.32	1.38	1.37	1.39	1.50
7	80	1.20	1.42	1.44	1.43	1.47	1.51
8	1522	1.10	1.40	1.43	1.42	1.47	1.56
10	829	1.09	1.35	1.42	1.39	1.46	1.57

8.6.3 Rock Fraction content by Mass

Table Anx11: Descriptive statistics of the Rock fraction content by Mass (RM; -) for cropland estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	1882	0.05	0.24	0.27	0.26	0.29	0.43
2	341	0.01	0.05	0.08	0.09	0.12	0.33
3	212	0.01	0.04	0.07	0.07	0.09	0.15
4	25	0.02	0.04	0.05	0.11	0.23	0.30
5	740	0.01	0.01	0.02	0.02	0.03	0.15
6	653	0.01	0.02	0.03	0.03	0.04	0.19
7	752	0.01	0.02	0.03	0.04	0.05	0.27
8	328	0.01	0.02	0.03	0.03	0.04	0.14
9	98	0.01	0.02	0.02	0.03	0.03	0.09
10	312	0.01	0.02	0.03	0.06	0.07	0.33

Table Anx12: Descriptive statistics of the Rock fraction content by Mass (RM; -) for grassland estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
1	719	0.04	0.20	0.25	0.24	0.28	0.37
2	85	0.02	0.04	0.07	0.08	0.11	0.37
3	83	0.02	0.03	0.06	0.06	0.08	0.16
4	11	0.03	0.04	0.05	0.05	0.06	0.06
5	68	0.01	0.02	0.03	0.03	0.04	0.13
6	213	0.01	0.02	0.03	0.04	0.04	0.15
7	225	0.01	0.02	0.03	0.04	0.04	0.14
8	312	0.01	0.02	0.03	0.03	0.03	0.15
9	84	0.01	0.02	0.03	0.03	0.03	0.06
10	319	0.01	0.02	0.04	0.08	0.10	0.33

Table Anx13: Descriptive statistics of the Rock fraction content by Mass (RM; -) for vineyard estimated from the application of the GBM model to BDAT_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)

Assoc.	n	min	Q1	median	mean	Q3	max
2	80	0.02	0.04	0.07	0.08	0.10	0.16
3	1670	0.02	0.07	0.08	0.08	0.09	0.17
4	44	0.03	0.05	0.10	0.09	0.12	0.13
5	60	0.02	0.02	0.03	0.03	0.03	0.05
6	26	0.03	0.04	0.05	0.05	0.06	0.10
7	80	0.02	0.02	0.03	0.03	0.04	0.07
8	1522	0.01	0.03	0.03	0.03	0.04	0.13
10	829	0.01	0.02	0.03	0.04	0.04	0.13

8.7 Determination of Stone Content in historical soil surveys

Source: 'Mémoire de méthodes d'analyses' by P. Gillen (year unknown), chief chemist (1949-1973), describing the analytical methods used in the 1960s in the laboratory from the soil service of the 'Station de Chimie Agricole' of Ettelbruck.

PRETRAITEMENT des échantillons de sols pour la Carte Pédologique (Analyses: Code 1, 2, 3 et 6)

Détermination du GRAVIER en %.

Matériel de laboratoire:

balance Bizerba: max. 2 kg à 5 gr.
cartons
dateur automatique (4x)
verres d'une contenance de 500 cc avec couvercle
mortier en fonte
machine de broyage et de tamisage 2 mm
tamis de 2 mm (rond) pour tamisage sous eau des fractions plus grandes que 2 mm
étuve à ventilation forcée.

Principe:

Faire sécher à l'air le sol qui a été prélevé (1-2 semaines) peser et tamiser.
Mise en stock du sol séché à l'air.

Procédure:

Les échantillons entrant au laboratoire sont munis d'étiquettes portant: la date de prise, le No. de la carte topographique, les coordinées de la prise de l'échantillon, le No. de profil avec indication du/des horizons, les profondeurs, les codes d'analyses à effectuer ainsi que le nom du prospecteur.

Au déballage les étiquettes sus-mentionnées sont munies d'un numéro courant qui est en même temps imprimé deux fois sur le carton où l'échantillon est versé.

Le No. des étiquettes est transcrit avec toutes les indications sur feuilles spéciales dont l'original reste dans le livre. Une copie (feuille jaune) est destinée à être mise dans le verre. Une autre (sur papier auto-collant) est destinée à être collée sur le verre.

Le sol séché à l'air est pesé (Poids total = X gr)
Ensuite le sol est prébroyé dans un mortier en fonte.
Après prébroyage le sol est concassé automatiquement et tamisé à 2 mm.- La fraction plus petite que 2 mm est recueillie séparément de la fraction plus grande que 2 mm (=gravier)
La fraction plus grande que 2 mm est pesée (= Z gr) (= X gr)
La fraction plus petite que 2 mm est pesée (= Y gr)

La fr. plus petite que 2 mm est introduite dans verres.
La fr. plus grande que 2 mm est traitée sur un tamis de 2 mm par l'eau courante. Les particules de terre adhérant au gravier sont lavées. La fraction est ensuite séché au four électrique à 60° Cels. en utilisant la ventilation forcée. Durée de séchage: 24 heures. Après refroidissement la fraction est pesée. (= Z' gr).

Calcul:

$$\frac{\text{Poids total de l'échant. X gr}}{\text{1 gr}} \quad \frac{\text{Z' gr x 100}}{\text{X gr}} = \text{\% gravier}$$

100 gr

Remarque: Une deuxième classification des échantillons est faite selon les numéros courants
Une troisième classification selon le numéro de profil est exécutée
Finalement nous utilisons une 4. classification suivant les planchettes (30 cartes topo couvrant tout le pays).