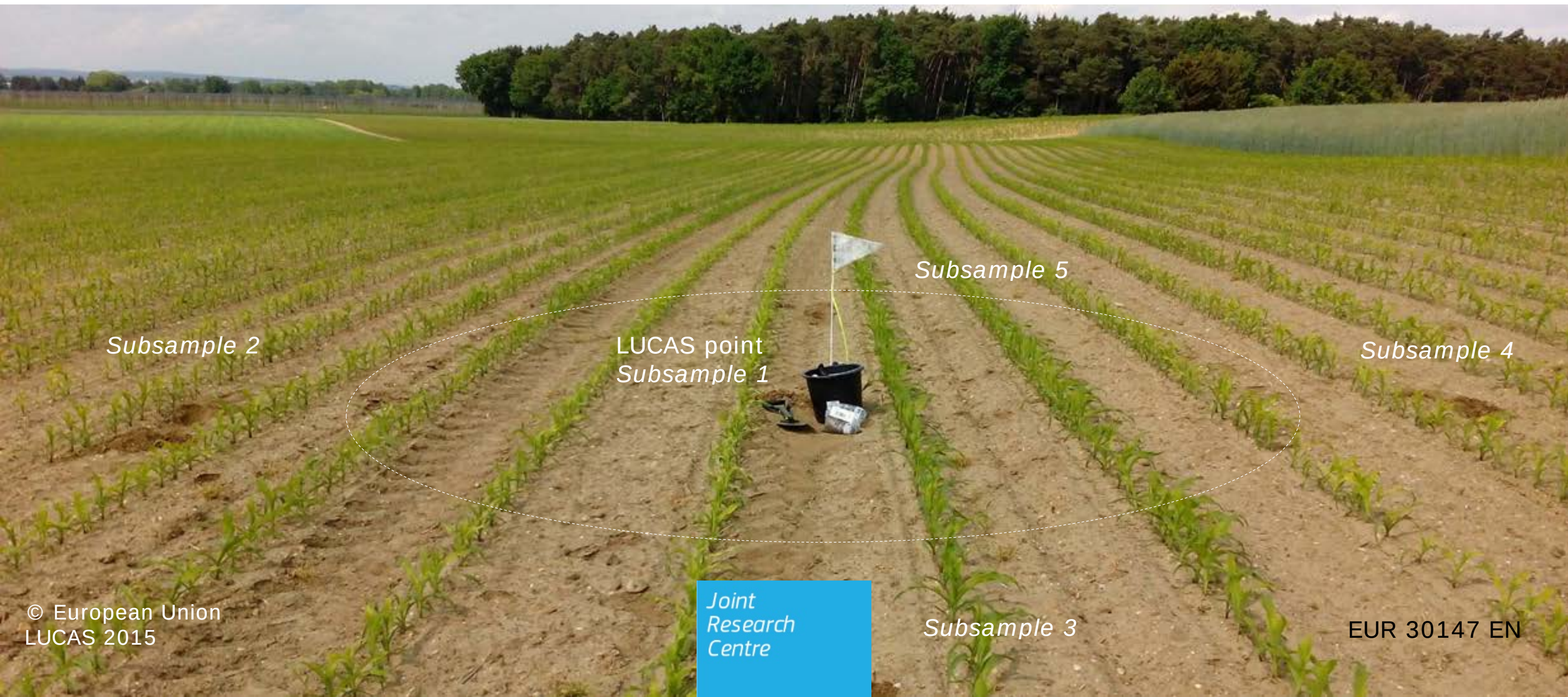
European Commission

JRC TECHNICAL REPORTS

Assessment of changes in topsoil properties in LUCAS samples between 2009/2012 and 2015 surveys

Fernández-Ugalde, O., Ballabio, C., Lugato, E., Scarpa, S., Jones, A.

2020



Joint Research Centre

EUR 30147 EN

Contact information
Name: Arwyn Jones
Address: European Commission, Joint Research Centre (JRC), Sustainable Resource Directorate, Land Resources Unit, Via E. Fermi 2749, I-21027 Ispra (VA), Italy
Email: arwyn.jones@ec.europa.eu

EU Science Hub
https://ec.europa.eu/jrc

JRC120138

EUR 30147 EN

PDF ISBN 978-92-76-17430-1 ISSN 1831-9424 doi:10.2760/5503

Publications Office of the European Union, Luxembourg, 2020

How to cite this report: Fernández-Ugalde, O., Ballabio, C., Lugato, E., Scarpa, S., Jones, A., *Assessment of changes in topsoil properties in LUCAS samples between 2009/2012 and 2015 surveys*, EUR 30147 EN, Publications Office of the European Union: Luxembourg 2020, ISBN 978-92-76-17430-1, doi:10.2760/5503, JRC120138.

Contents

Acknowledgements

Special thanks are due to our colleagues in EUROSTAT, in particular, Alessandra Palmieri, Beatrice Eiselt, Ruxandra Roman Enescu, Paolo Dominici and Jane Schofield that contribute to the planning and implementation of the LUCAS Soil Survey.

We note the funding provided by DG AGRI, DG CLIMA and DG ENV to support the field programme and subsequent laboratory analysis. In addition we are grateful to colleagues from DG ENV, DG AGRI and DG CLIMA who provided an invaluable support for the planning of the survey and the review of this report.

We would also like to thanks to our colleagues at JRC D3 Land Resources Unit for their support to the 2015 LUCAS Soil Survey, in particular, for their willingness to help with logistical issues and for providing comments during the preparation of this report.

Sincere thanks also to SGS Hungary for the effort in analysing all the soil samples.

Abstract

Soil delivers fundamental ecosystem services that support human well-being. These include the provision of food, feed, fuel, fibre and genetic resources, the regulation of storage, filtration and cycling of nutrients and water, cultural (aesthetic, spiritual and recreational) values and supporting the provision of all other services. Policies for sustainable land and soil management should be based on monitoring systems that are able to provide evidence of the impact of land use/land cover changes and climate change in soil condition, both in space and in time. In this context, the topsoil assessment module of the Land Use and Cover Area Frame Survey (LUCAS) is the first harmonised soil monitoring network at European Union (EU) level that uses a common sampling procedure and standard analysis methods.

Eurostat has carried out the LUCAS survey every 3 years since 2006. The surveys are based on the visual assessment of environmental and structural elements of the landscape in georeferenced control points, a subsample of which is selected to be visited to collect field-based information. In 2009, a soil assessment module was added within the LUCAS survey with the scope to create a harmonised and comparable dataset of physical and chemical properties of topsoil across the EU to support policymaking. About 20,000 soil points were selected across 27 member states (except Bulgaria and Romania) based on a stratified sampling scheme with land use and terrain information as attributes. At each point, samples were collected from a depth of 20 cm using a common sampling procedure. Subsequently, the samples were analysed for several properties in a single laboratory using standard analytical methods. The same point selection procedure, sampling method and analysis methods were extended in 2012 to Bulgaria and Romania, where samples were collected for about 2,000 soil points.

The LUCAS Topsoil Survey was repeated in 2015, year in which 17,613 soil points sampled in the LUCAS 2009 and 2012 surveys were revisited. Furthermore, new soil points at an altitude of 1,000 - 2,000 m were added to the survey (the altitude limit was 1,000 m in LUCAS 2009 and 2012 surveys). The soil module was also extended by the JRC to Albania, Bosnia and Herzegovina, Croatia, Montenegro, Republic of North Macedonia and Serbia. In total, 27,069 points were selected for the topsoil survey in 2015, of which 25,947 were located in the EU-28 MS.

In this report, we provide a detailed evaluation of the LUCAS topsoil sampling and the laboratory analysis. We also assess changes in topsoil properties between LUCAS 2009/2012 and 2015 surveys based on data of paired samples (i.e. samples collected in revisited LUCAS soil points in 2009/2012 and in 2015). The ultimate goal of this report is to assess the efficacy of the LUCAS Topsoil Module for the early detection of changes in soil conditions, since this is a primary objective for scientific and policy organizations to improve their policies for a sustainable land use and management.

The LUCAS spade sampling is an efficient and cost-effective method for topsoil monitoring at regional/continental scale, although a better control of litter removal in woodland and sampling depth in all LC classes is needed. When comparing sampling locations of revisited points, almost 97 % of the samples taken in 2015 were taken at a distance <100 m from their baseline locations in 2009/2012 as indicated in the sampling protocol. Three percent of the samples were taken at a distance between 100 and 400 m from one survey to the other. As a result, changes in soil properties were not significantly affected by the distance between sampling locations in the 2009/2012 and 2015 surveys.

Regarding laboratory analysis, the data of the properties analysed showed a coherence from the soil point of view. Organic carbon and N levels showed a positive correlation, $CaCO_3$ content was lower in samples were pH was below 7, and the sum of sand, silt and clay percentages was between 99 and 101 in the fraction <2 mm of all samples.

Overall, OC and N levels were highest in woodland, followed by grassland and cropland in both the 2009/2012 and 2015 surveys. On the contrary, P and K levels were higher in cropland and grassland than in woodland in the surveys. Carbonate content was lowest in

woodland from northern member states and highest in cropland from southern member states in both surveys. In agreement with these results, pH was lower in woodland than in cropland in both surveys.

Soil properties showed large standard deviations within surveys and between surveys due to uncertainties arising from the sampling. Unfortunately, some LC classes were under sampled. Consideration should be given to increase the number of sampling sites in future surveys to ensure representative data.

Overall, most of the soil properties showed limited changes between 2009 and 2015 (over the six-year period) in the 27 member states. Changes in Bulgaria and Romania were even less evident over the three-year period (from 2012 to 2015). Thus, the survey confirms that soil properties change very slowly over time. From a policy perspective, a time lapse longer than six years is necessary in order to observe small variations in soil conditions, unless a marked change has occurred due to erosion processes, extreme meteorological events or land use/cover changes.

Despite uncertainties arising from the sampling, it has been possible to draw some conclusions when assessing changes in soil properties between 2009/2012 and 2015 surveys <u>in mineral soils</u> (i.e. where OC <120 g kg^{-1}).

— Taking the revisited points, a statistically significant increase in OC content of 3.74 % was observed in grassland over six years in the 27 member states. This is in line with the annual 0.4 % increase in the topsoil (30-40 cm) targeted by the ‘4 per 1000’ initiative. This would contribute to climate change mitigation.

— Similarly, for the revisited points in cropland, a statistically significant decrease in OC content of 2.5 % was observed while points that changed from grassland to cropland over six years decreased by 11 %. This suggests that cropland soils are not working as carbon sinks.

— In other land cover categories, the number of repeated points was insufficient to assess statistical significance.

— No tangible changes were observed in Bulgaria and Romania over three years.

— Nitrogen content increased in cropland, grassland, woodland points, and in points that changed from cropland to grassland over six years in the 27 member states. In Bulgaria and Romania, N content increased in cropland points and in points that changed from cropland to grassland and vice-versa over three years. In non-agricultural conditions, this may reflect airborne deposition of nitrogen.

— Phosphorus content increased in cropland, grassland and woodland points over six years in the 27 member states. On the contrary, K content decreased in cropland points in the 27 member states. In Bulgaria and Romania, no tangible changes were observed over three years.

— pH in $CaCl_2$ was a more consistent measurement and was less affected by seasonal fluctuations of electrolyte concentration in soil solution.

— pH in $CaCl_2$ increased in cropland and woodland points, and in points that changed from woodland to shrubland over six years in the 27 member states. On the contrary, pH in $CaCl_2$ decreased in grassland points. In Bulgaria and Romania, pH in $CaCl_2$ decreased in grassland points over three years.

1 Introduction

Soil is a key component of the biosphere that delivers fundamental ecosystem services to support human well-being. Among these services are provisioning (food, feed, fuel, fibre and genetic resources), regulating (storage, filtration and cycling of nutrients and water), cultural (aesthetic, spiritual and recreational values) and supporting (essential for the provision of all other services). In order to ensure that soil delivers these ecosystem services, it is necessary to develop pan-European policies for a sustainable land and soil management while preventing degradation. The development of such policies should be based on land and soil monitoring networks that are able to provide evidence of the impact of land use and land cover (LC) changes in soil physical and chemical properties, both in space and in time. In this context, the topsoil assessment module of the LUCAS (Land Use and Cover Area Frame Survey) programme is the first harmonised topsoil-monitoring network that uses a common sampling procedure and standard analysis methods at the European Union (EU) level.

The LUCAS Programme started as an area frame statistical survey organised and managed by Eurostat (the Statistical Office of the EU to monitor land use and LC, and their changes, over time across the EU. Since 2006, Eurostat has carried out LUCAS surveys every three years. The surveys are based on the visual assessment of environmental and structural elements of the landscape in georeferenced control points. The points belong to the intersections of a 2 x 2 km regular grid covering the territory of the EU. This results in around 1,000,000 georeferenced points. Each point has been classified by LC class using orthophotos or satellite images. In every survey, a subsample of these points is selected to be visited for collecting field-based information. In LUCAS 2009, about 200,000 points were visited across 27 member states (EU-28 except Bulgaria and Romania). In LUCAS 2012 and 2015, about 270,000 points were visited in the EU-28 member states.

Eurostat, together with DG-ENV and DG-JRC, implemented a topsoil assessment module within the LUCAS programme (LUCAS-Topsoil) in 2009. The scope was to create a harmonised and comparable dataset of physical and chemical properties of topsoil in the EU to support policymaking. About 20,000 soil points (i.e. 10 % of the 200,000 points of the LUCAS grid visited in the field) were selected across the participating member states for the topsoil sampling based on a stratified sampling scheme with land use and terrain information as attributes (Tóth et al., 2013). At each point, soil samples were collected from a depth of 20 cm following a common sampling procedure. These were analysed for several physical and chemical properties in a single laboratory following standard analytical methods: coarse fragments, particle-size distribution, organic carbon (OC), carbonates ($CaCO_3$), nitrogen (N), phosphorus (P), potassium (K), cation exchange capacity (CEC), pH and multispectral spectroscopy. The same point selection procedure, sampling method and analysis standards were extended in 2012 to Bulgaria and Romania, where topsoil samples were collected from about 2,000 soil points.

In 2015, it was planned to repeat the LUCAS Topsoil Survey in 90 % of the soil points sampled in the LUCAS 2009 and 2012 surveys. The other 10 % soil points of the 2009 and 2012 surveys were substituted by new points in the revisited member states. Part of the new points were assigned at altitudes above 1,000 m, which were out of scope of the LUCAS 2009 and LUCAS 2012 surveys. The LUCAS Topsoil Survey was also extended by the JRC to Albania, Bosnia and Herzegovina, Croatia, Montenegro, Republic of North Macedonia, and Serbia. Switzerland also participated following standard LUCAS protocols. Overall, 27,069 points were selected for the topsoil survey of LUCAS 2015. Soil samples were finally collected in 23,902 points, from which 22,631 were collected in the EU-28 member states, following the common LUCAS sampling procedure (see chapter 2). As in LUCAS 2009 and 2012 surveys, all samples were analysed for physical and chemical properties in a single laboratory using the same analytical methods (see chapter 2). In addition, electrical conductivity (EC) and clay mineralogy were analysed for the first time.

The aims of this report are to (1) provide a detailed insight into the LUCAS topsoil sampling and into the analysis of soil properties in the laboratory and (2) assess changes in topsoil properties between LUCAS 2009/2012 and 2015 surveys based on data of paired samples (i.e. samples collected in revisited LUCAS soil points in 2009/2012 and in 2015). The ultimate goal of this report is to assess the efficacy of the LUCAS Topsoil Module for the early detection of changes in soil conditions, since this is a primary objective for scientific and policy organizations to improve their policies for a sustainable land use and management.

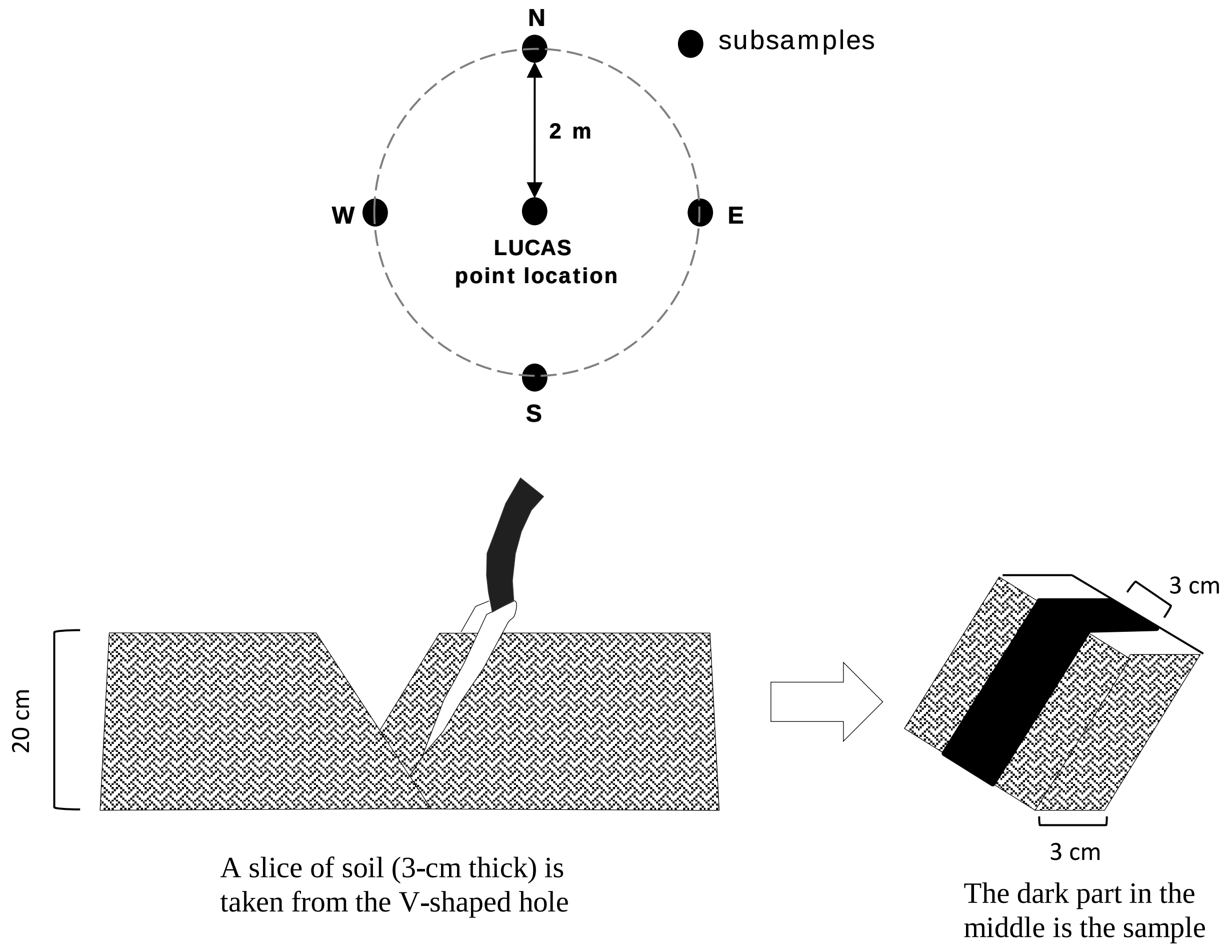
The report is divided is two parts. The first part provides a description of the sampling and laboratory analysis techniques (chapter 2), an assessment of the efficiency of the LUCAS soil sampling protocol (chapter 3) and an evaluation of the soil data produced in the laboratory in the context of LUCAS 2015 (chapter 3). The second part regards the assessment of changes in soil properties between 2009/2012 and 2015 surveys (chapter 6). Changes are described by NUTS 2 units and main LC classes. Changes in soil properties are presented at NUTS 2 level to match how DG ESTAT presents LUCAS statistics in their website ([1]).

([1]) https://ec.europa.eu/eurostat/web/lucas/data

2 Sampling and laboratory analysis techniques

Soil sampling in the LUCAS surveys of 2009, 2012 and 2015 was carried out following the same standard protocol ([2]). Briefly, a composite sample of approximately 500 g was prepared from five subsamples collected with spade at each LUCAS point. The first subsample was collected in the geo-referenced point location; the other four subsamples were collected at a distance of 2 m following the cardinal directions (North, East, South and West) (Figure 1). Before collecting the subsamples, stones and boulders (>6 cm) (FAO, 2006), vegetation residues, grass and litter were removed from soil surface by raking with the spade. A V-shaped hole was dug to a depth of 20 cm using the spade and a slice of soil (approximately 3-cm thick) was taken from the hole with the spade. The slice was trimmed in the sides, which resulted in a 3-cm wide subsample. The subsample was placed in a bucket. The procedure was repeated to collect the other four subsamples. Finally, the five subsamples in the bucket were mixed with a trowel and extra vegetation residues and stones were removed. Approximately 500 g of the mixed soil were taken with a trowel from the bucket and placed in a plastic bag to derive the composite sample.

Figure 1. The LUCAS sampling schema and collection of a subsample



Source: Fernández-Ugalde et al. (2019).

([2]) http://ec.europa.eu/eurostat/documents/205002/6786255/LUCAS2015-C1-Instructions-20150227.pdf

The samples were then sent to a central laboratory to analyse physical and chemical properties with standard ISO methods, except for extractable K. Table 1 shows the soil properties measured, together with the method used. Soil samples from the three LUCAS surveys (2009, 2012 and 2015) were analysed following the same methods.

Table 1. Methods used for the analysis of physical and chemical properties in topsoil samples.

Soil properties	**Method**	**Description**
Coarse fragments	ISO 11464:2006	Sieving to separate coarse fragments (2-60 mm) from fine earth fraction
Clay, silt and sand contents	ISO 11277:1998 ISO 13320:2009	Sieving and sedimentation method (in 2009 and 2012) Laser diffraction (in 2015 only)
pH in $CaCl_2$ and in H_2O	ISO 10390:2005	Glass electrode in a 1:5 (V/V) suspension of soil in H_20 and $CaCl_2$
Electrical Conductivity	ISO 11265:1994	Metal electrodes in aqueous extract of soil
Organic carbon content	ISO 10694:1995	Dry combustion (elementary analysis)
Carbonates content	ISO 10693:1995	Volumetric method
Phosphorus content	ISO 11263:1194	Spectrometric determination of P soluble in sodium hydrogen $CaCO_3$ solution
Total nitrogen content	ISO 11261:1995	Modified Kjeldahl method
Extractable potassium content	USDA−NRCS, 2004	Atomic absorption spectrometry after extraction with NH_4OAc
Cation exchange capacity	ISO 11260:1994	Using barium chloride solution to saturate samples and extract cations
Multispectral spectroscopy	Soil Spectroscopy Group	Diffuse reflectance measurements
Clay mineralogy	X-ray diffraction	X-ray diffraction patterns of oriented aggregates (only in 2015)

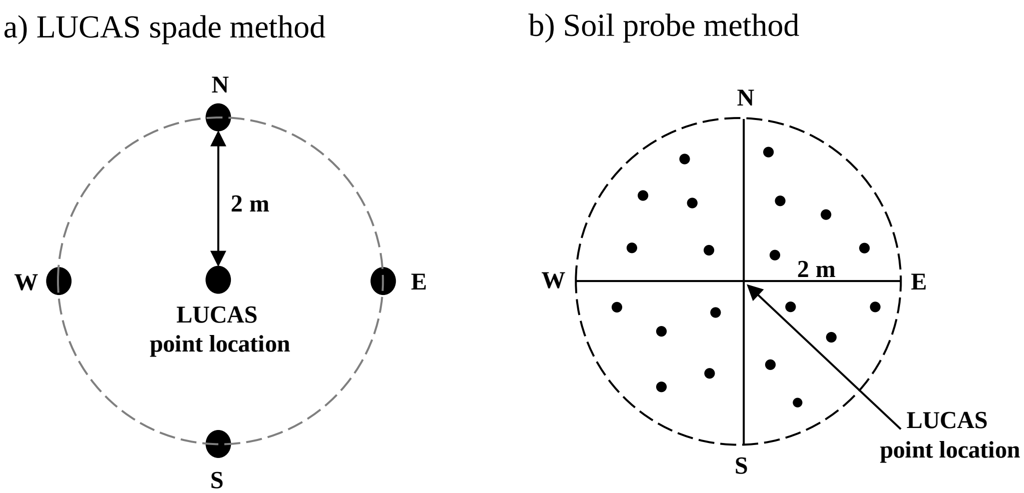
Samples from revisited LUCAS soil points were analysed for pH, EC, OC, $CaCO_3$, P, N, K, CEC and spectroscopy. Clay mineralogy was analysed only in samples taken in 400 revisited soil points. Particle-size distribution was not analysed in samples from revisited points because it is not expected to find significant changes in the short-term. In the samples taken in new soil points the whole set of properties in Table 1 was analysed.

3 Efficiency of the LUCAS sampling for topsoil monitoring

The election of the sampling strategy is a key aspect when planning a soil-monitoring network. Indeed, the determination of soil properties can be subject to uncertainties due to the sampling method. In the LUCAS 2015 survey, the spade sampling method used in LUCAS was compared to an alternative sampling with soil probe for 155 LUCAS points located in mineral soils in Switzerland. The 155 points selected covered the main LC classes in Switzerland: (i) agricultural areas included meadow (42 points), pasture (27 points) and cropland (28 points); (ii) woodland areas included coniferous forest (18 points), deciduous forest (21 points) and mixed forest (14 points). At each point, two topsoil samples were collected: one with the LUCAS spade method (see chapter 2 for more details) and other with the soil probe method (Figure 2).

In the method with the soil probe (internal diameter 2.5 cm), a composite sample of approximately 500 g of soil was collected at each point. The composite sample consisted of 20 soil cores collected following a stratified random sampling procedure. A circle was designed at a distance of 2 m around the geo-referenced point connecting the four cardinal directions where subsamples were collected in the LUCAS sampling method. The circle was divided in four quadrants (N–E, E–S, S–W and W–N) (Figure 2). At each quadrant, five soil cores were collected to a depth of 20 cm and placed in a bucket. Before collecting the soil cores, vegetation residues, grass and litter were removed from the soil surface. If a point was located in woodland, special care was taken to completely remove the litter layer by hand. The 20 soil cores in the bucket were mixed with a trowel and vegetation residues and stones were carefully removed by hand. The 20 mixed soil cores (i.e. the composite sample) were then transferred to a plastic bag.

Figure 2. The sampling schemas of the LUCAS spade method and the soil probe method



Source: Fernández-Ugalde et al. (2019).

The main differences between the two methods are the sub-sampling scheme around the main soil point (Figure 2), the equipment used to collect topsoil samples and the accuracy of the litter removal in woodland. These factors can affect the quality of the sample and the reliability of a soil analysis. The sampling scheme has to be able to show the spatial heterogeneity of topsoil in a field. Regarding the sampling equipment, a spade is a convenient tool to collect disturbed topsoil samples, whereas a soil probe is more appropriate to collect undisturbed soil samples with better depth control. An accurate litter removal, especially in woodland, is important for the analysis of properties such as OC and N content. Unlike in the LUCAS spade, vegetation residues in woodland are manually removed to ensure their complete elimination in the method with the soil probe. As a result, the litter removal was perceived as being more accurate in the method with the soil probe than in the LUCAS spade method in this study.

Both sets of samples were analysed with the standard methods detailed in chapter 2 for the following properties: clay, silt and sand, $CaCO_3$, OC, N, P, K, EC, pH in H_2O and pH in $CaCl_2$. The Bland and Altman diagram was used to detect potential outliers caused by errors of measurement. The diagram displays the difference between a pair of measurements (in our case, measurements made in paired samples collected with the LUCAS spade and the soil probe methods) plotted on the vertical 1:1 axis against the mean of the pair on the horizontal axis. Where normal distribution of the differences is met, approximately 95 % of the differences in the data are expected to lie between the limits of agreement, which are defined as the mean of the observed differences plus and minus 2 times the standard deviation of the differences (Bland & Altman, 1986). Where normal distribution of differences is not met, median, 2.5th and 97.5th percentiles are used as limits of agreement (Bland & Altman, 1999). From these limits, we decided whether the agreement between pairs of measurements is acceptable. The normal distribution of the differences between paired measurements was checked using normal Q-Q diagrams. The Lin's concordance correlation coefficient (LCCC) was used to compare the results of physical and chemical analyses between topsoil samples collected with the LUCAS spade and the soil probe methods. The LCCC measures the fit of the data along a one-to-one line passing through the origin (Lin, 1989). If samples collected with the two methods were able to reproduce the same results for the soil properties analysed, the relationship between them should fall on the line. Furthermore, the root mean square error (RMSE) was used to evaluate the average magnitude of the difference between the results of soil properties analysed in topsoil samples collected with the LUCAS spade and the soil probe method. The R statistical computing program (R ×64 3.0.3) was used for the statistical analyses.

Descriptive statistics of soil properties in mineral soils are detailed for the LUCAS spade and the soil probe methods in Table 2. The range of values, mean, median and standard deviation of analysed properties were similar for both methods. Mean and median were similar to each other for most of the soil properties, which means that data were evenly divided around the mean. The median content of $CaCO_3$ was 0 g kg^{-1} for the LUCAS spade and 1.0 g kg^{-1} for the method with the soil probe, while the mean content was 55.6 g kg^{-1} for the LUCAS spade and 51.2 g kg^{-1} for the method with the soil probe. This indicates that most of the samples contained no $CaCO_3$.

Table 2. Descriptive statistics of physical and chemical properties in topsoil (0–20 cm) of mineral soils collected with the LUCAS spade and the soil probe methods.

Soil properties	**LUCAS spade method (n = 155)**				**Soil probe method (n = 155)**			
	Range	**Mean**	**Median**	**SD**	**Range**	**Mean**	**Median**	**SD**
Clay / g kg^{-1}	40.0 – 400.0	198.8	200.0	7.1	40.0 – 420.0	204.4	210.0	7.2
Silt / g kg^{-1}	270.0 – 730.0	518.8	530.0	8.3	250.0 – 700.0	521.9	530.0	7.9
Sand / g kg^{-1}	60.0 – 690.0	282.1	260.0	12.6	60.0 – 710.0	274.3	260.0	11.6
OC / g kg^{-1}	8.9 – 151.5	43.6	37.4	27.3	5.6 – 147.2	41.8	33.8	26.4
$CaCO_3$ / g kg^{-1}	0.0 – 775.0	55.6	0.0	128.1	0.0 – 800.0	51.2	1.0	120.4
N / g kg^{-1}	0.5 – 10.1	4.0	3.5	1.8	0.4 – 10.9	4.0	3.5	1.9
P/ mg kg^{-1}	0.0 – 167.8	29.7	24.4	25.6	0.0 – 172.7	31.4	24.0	26.6
K / mg kg^{-1}	45.5 – 496.4	150.9	122.9	90.1	39.1 – 529.8	132.9	99.6	89.8
pH–H_2O	3.6 – 8.1	6.2	6.4	1.2	3.6 – 8.0	6.1	6.3	1.1
pH–$CaCl_2$	3.0 – 7.6	5.9	6.1	1.3	3.1 – 7.6	5.9	6.1	1.2
EC / mS m^{-1}	4.8 – 90.8	33.7	29.9	19.0	4.3 – 113.0	34.1	29.9	21.0

Source: Fernández-Ugalde et al. (2019).

This section focuses on the influence of the sampling method in key soil properties such as OC, N, P and K in the various LC classes (Table 3). In their publication, Fernández-Ugalde et al. (2019) present and discuss the results of the comparison between the two sampling methods for all physical and chemical properties. Overall, all properties showed an acceptable average magnitude of difference (measured by the RMSE). Regarding the concordance between measurements of samples collected with the LUCAS spade and the soil probe method, the greatest LCCC were observed in topsoil of croplands for OC, N and P. Differences in the control of sampling depth between the LUCAS spade and the soil probe method did not affect the accuracy of the analyses. This is explained by the homogeneous distribution of nutrients in topsoil of cropland because of ploughing (usually to 15 – 30 cm). In woodland classes, the LCCC were lower in topsoil of coniferous forest than deciduous and mixed forests. In forest soils, litter is an important source of nutrients that usually have shallow distributions (Jobbágy & Jackson, 2001). Differences in the accuracy on the removal of litter and the control of sampling depth can influence the results of the analyses of nutrients such as OC, N and K. The results suggested that K analysis was specially affected by the sampling method in coniferous forest (LCCC was 0.37). This is most likely due to the fact that litter is a key source of K in coniferous forest, where part of the soil K is unavailable because of soil acidity. In meadow and pasture classes, the LCCC was lower in pasture than in meadow. This can be explained by the heterogeneous distribution of livestock excreta, source of OC and N, on pasture surface (White et al., 2001). The spatial distribution of subsamples in the LUCAS spade and the soil probe methods (Figure 2) reflected this spatial variability of OC and N contents in topsoil.

Table 3. Lin's concordance correlation coefficient (LCCC) and root mean square error (RMSE) for physical and chemical properties between topsoil samples collected with the LUCAS spade and the soil probe methods in mineral soils.

Soil properties	**Coniferous forest (n = 18)**			**Deciduous forest (n = 21)**			**Mixed forest (n = 14)**		
	N	**Lin's CCC**	**RMSE**	**N**	**Lin's CCC**	**RMSE**	**N**	**Lin's CCC**	**RMSE**
OC / g kg^{-1}	17	0.89	15.73	21	0.97	8.17	14	0.92	12.76
N / g kg^{-1}	17	0.94	0.55	21	0.97	0.48	14	0.96	0.55
P/ mg kg^{-1}	17	0.86	6.60	20	0.89	6.88	14	0.95	3.40
K / mg kg^{-1}	17	0.37	25.80	21	0.91	33.41	13	0.95	32.05
Soil properties	**Cropland (n = 28)**			**Meadow (n = 42)**			**Pasture (n = 27)**		
	N	**Lin's CCC**	**RMSE**	**N**	**Lin's CCC**	**RMSE**	**N**	**Lin's CCC**	**I RMSE**
OC / g kg^{-1}	27	0.98	1.21	41	0.97	4.81	26	0.93	8.06
N / g kg^{-1}	27	0.97	0.17	42	0.94	0.63	26	0.89	0.77
P/ mg kg^{-1}	27	0.97	5.20	42	0.94	8.05	26	0.97	7.49
K / mg kg^{-1}	27	0.88	32.89	41	0.88	42.58	26	0.87	46.98

Source: Fernández-Ugalde et al. (2019).

The results demonstrate that the accuracy of the analyses of soil properties was similar in topsoil samples collected with the LUCAS spade and the soil probe methods. Thus, it can be concluded that the LUCAS spade method is as efficient as the method with the soil probe for topsoil monitoring at continental scale. Even so, the comparison of the LUCAS spade and the soil probe methods evidences that some improvements can be done to reduce uncertainty on the determination of soil parameter, when sampling topsoil with a spade. First, it is crucial to collect the samples accurately down to the target depth. Second, a conscientious removal of litter layer, especially in coniferous forest, should be ensured

4 Soil data evaluation in the LUCAS surveys

Data validation processes aims to provide certain guarantees of accuracy, completeness and consistency to data. Based on the methodology developed by Hiederer et al. (2008) for other pan-European soil databases, three aspects were assessed for the LUCAS soil data: compliance, conformity and uniformity. Compliance concerns the data format, conformity involves the data content, and uniformity is related to the comparability of data between different surveys. For each of these aspects, various tests were carried out:

— agreement of the data format with the specifications indicated in the call for tender of LUCAS 2015 (Compliance check, section 4.1),

— control of the identification and registration of samples in the LUCAS 2015 survey (Conformity check, section 4.2),

— evaluation of soil data and application of pedological criteria in the LUCAS 2015 survey (Conformity check, section 4.3),

— assessment of closeness of sampling locations in paired samples between the 2009/2012 and 2015 surveys (Uniformity check, section 4.4),

— assessment of the comparability of soil data between 2009/2012 and 2015 surveys (Uniformity check, section 4.5).

Furthermore, Hiederer (2018) performed a detailed validation of OC data, as it is a key property for all customer DGs due to its implications for climate change mitigation, assessing the impact of agricultural practices and the supply of ecosystem services.

4.1 Agreement of the data format with the specifications of the call for tender for laboratory analysis in LUCAS 2015

The technical specifications of the call for tender for the laboratory analysis in the LUCAS 2015 survey ([3]) includes the following conditions:

— Data generated in the laboratory for each soil sample shall be linked to the Soil ID in the dataset,

— Data of core soil properties ([4]) shall be delivered in an Excel (or 100 % Excel-compatible) workbook,

— Core soil properties shall be presented in columns following the order specified in the technical document,

— Units and number of decimals for each core soil property shall also follow the technical specifications,

— Additional soil properties (multispectral data and clay mineralogy) shall be delivered separately, following the specific indications for their data presentation.

The laboratory delivered the dataset of core soil properties in an Excel file with four sheets, one for each group of samples identified in the technical specifications. Group 1 includes potentially organic or organic-rich samples, Group 2 comprises mineral samples collected in georeferenced points previously sampled in 2009 or 2012 surveys, Group 3 contains mineral samples collected in new plots, and Group 4 contains soil samples in which clay mineralogy was analysed. Samples in Group 4 shall be included in Group 3, because they were collected in points sampled also in 2009 or 2012 surveys.

The information of each soil sample is linked to its Soil ID in the dataset, so that soil information can be attributed to a monitoring point in the LUCAS 2015 database of Eurostat through the Soil ID. Soil properties are ordered in columns in the dataset as indicated in

([3]) Tender reference number: JRC/IPR/2016/H.5/0004/OC (https://etendering.ted.europa.eu/cft/cft-search.html).

([4]) Coarse elements, clay, silt, sand, pH in $CaCl_2$, pH in H_2O, organic carbon, carbonates, phosphorus, nitrogen, extractable potassium, cation exchange capacity.

the technical specifications, except for silt and sand that are interchanged. Data of soil properties are expressed in pertinent units and with the number of decimals requested in the technical specifications. The laboratory added two extra columns to the dataset with Client ID (i.e. internal identification of samples in the laboratory) and member state of origin of soil samples. As requested, the laboratory delivered the data of additional soil properties separately with the proper format.

The technical specifications do not include indications neither for the coding of missing data nor for data outside detection limits. From the dataset it can be concluded that empty fields indicate missing data. This is the case of contents of coarse elements, clay, silt and sand in samples from Groups 1, 2 and 3, in which these properties were not analysed. Regarding the detection limits, the laboratory provided the values for the methods used to analyse the soil properties in the Final Report.

4.2 Identification and registration of samples in LUCAS 2015

The LUCAS soil points in the EU-28 member states are identified by unique Point IDs. These Point IDs are used in every survey to record agro-environmental data relating to the points in the Data Management Tool (DMT) managed by Eurostat. Furthermore, topsoil samples collected in LUCAS points are identified by Soil IDs. The JRC creates these Soil IDs. In each LUCAS survey, surveyors randomly assign these Soil IDs to the samples when collected. Each sample is double-packed with twin labels that have the same Soil ID. At each LUCAS point, surveyors document agro-environmental observations by filling in a field form and by taking photographs. Surveyors have to indicate the Point ID and the Soil ID in the field form. All the data is then stored in the DMT. Thus, every topsoil sample has a double identification: the Soil ID and the Point ID. The Soil ID is used to identify the samples in the laboratory and provides the soil data, while the Point ID gives the field data and is used to link information from different LUCAS surveys.

Overall, 23,902 samples were taken in LUCAS 2015, from which 22,631 were taken in the EU-28 member states. In all, 241 samples in the EU-28 member states had repeated Soil IDs in the LUCAS 2015 survey. We were able to identify 58 out of these samples using the Point ID and member state in the DMT (Table 4). For the rest of the samples, it was not possible to find unique links between the soil data and the agro-environmental information.

During the laboratory analysis, 589 samples, which Soil IDs were not recorded in the DMT, were found through the EU-28 member states. Thus, it was not possible to relate these samples to any Point ID. On the opposite, there were 130 Soil IDs recorded in the DMT but the samples were not received (Table 4).

After the removal of samples that could not be identified, the LUCAS 2015 Soil dataset have 21,859 unique records with soil and agro-environmental data.

Table 4. Identification of samples taken in LUCAS 2015 in the EU-28 member states.

Identification of samples	N samples affected
Samples taken	22,631
Repeated Soil IDs	241
Recovered Soil IDs	58
Soil IDs not recorded in the DMT	589
Soil IDs recorded in the DMT but no physical samples available	130
Unique Soil ID / Point ID combinations	21,859

4.3 Evaluation of soil data and pedological criteria in LUCAS 2015

The limits of detection of the analytical methods were used to filter the data of soil properties and highlight the presence of values outside possible ranges in the dataset of the 2015 survey. Table 5 gives an overview of these outsider values in the dataset.

Table 5. Summary of outsider values per soil property in the dataset of the LUCAS 2015 survey.

Soil parameter	**LOD[1]**	**Range actual values**	***N* samples < LOD[1]**	**% of the data**
pH-$CaCl_2$	2–10	2.6–10	0	0
pH-H_2O	2–10	3.2–10.4	1	0
Electrical conductivity (mS m^{-1})	0.1	0.3–969	0	0
Organic carbon (g kg^{-1})	2.0	0.1–560.2	71	0.3
Carbonates (g kg^{-1})	1.0	0–976	11478	49
Phosphorous (mg kg^{-1})	10.0	0–1017.6	5464	23
Total nitrogen (g kg^{-1})	0.2	0–38.5	14	0
Extractable potassium (mg kg^{-1})	10.0	0–10030.9	136	0.6
Cation exchange capacity (cmol+ kg^{-1})	2.0	0–173.3	382	1.6

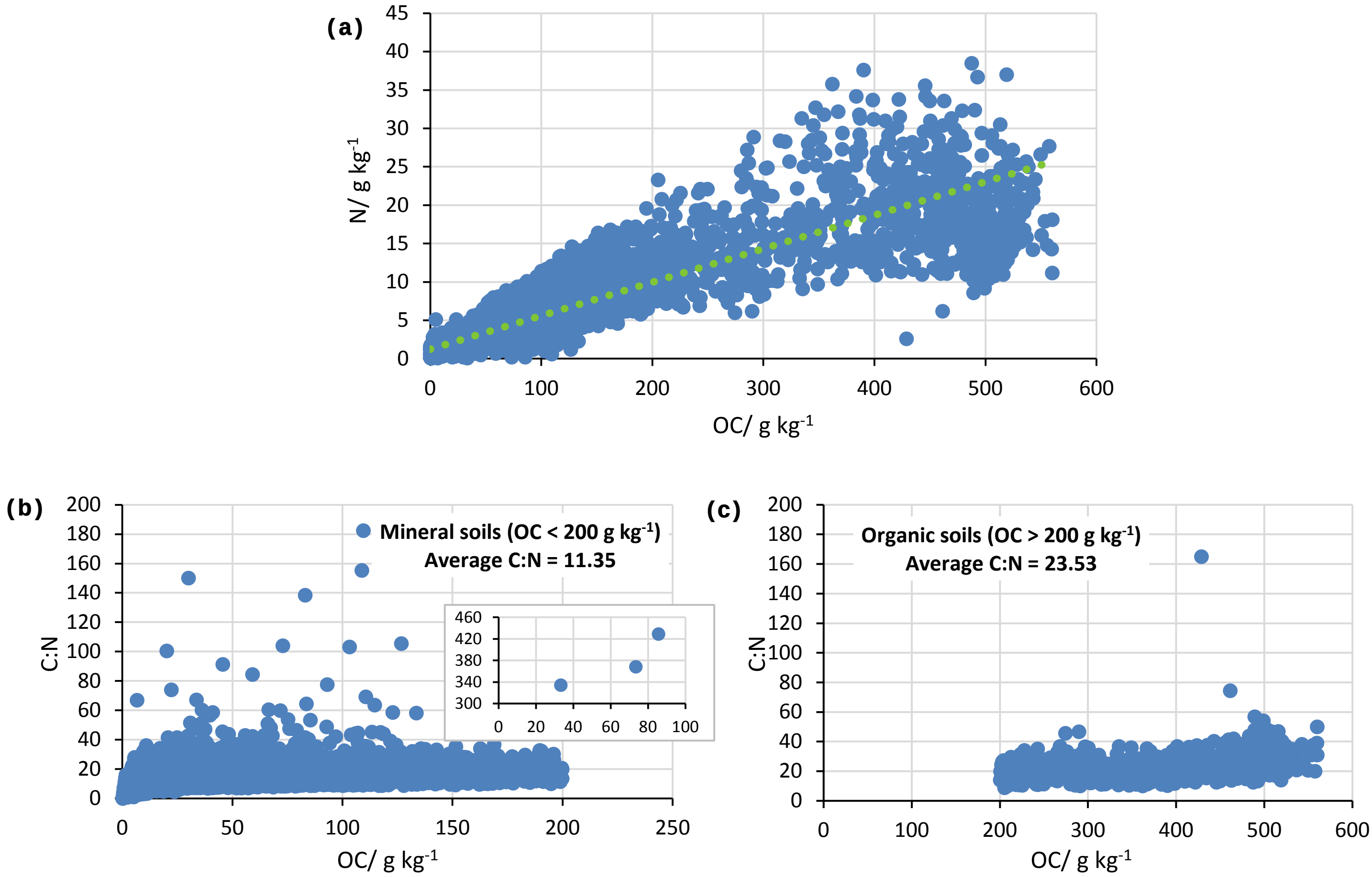
([1]) LOD : limit of detection.

The ranges of values reported for soil properties are within reasonable limits for soils in Europe. In fact, they are similar to those in the 2009 and 2012 surveys. The high number of outlier values for $CaCO_3$ is due to the use of the value "0" to indicate the absence of $CaCO_3$ in soil samples with low pH (pH<7). The value "0" has been substituted by "NA" in the dataset. Soil samples with the P content below the limit of detection are mainly located in woodland (36 %), and grassland and cropland most likely not subject to fertiliser applications (19 % and 21 %, respectively).

A range of correlations between soil properties were assessed for verifying coherence of the raw data from the soil point of view. These correlations include:

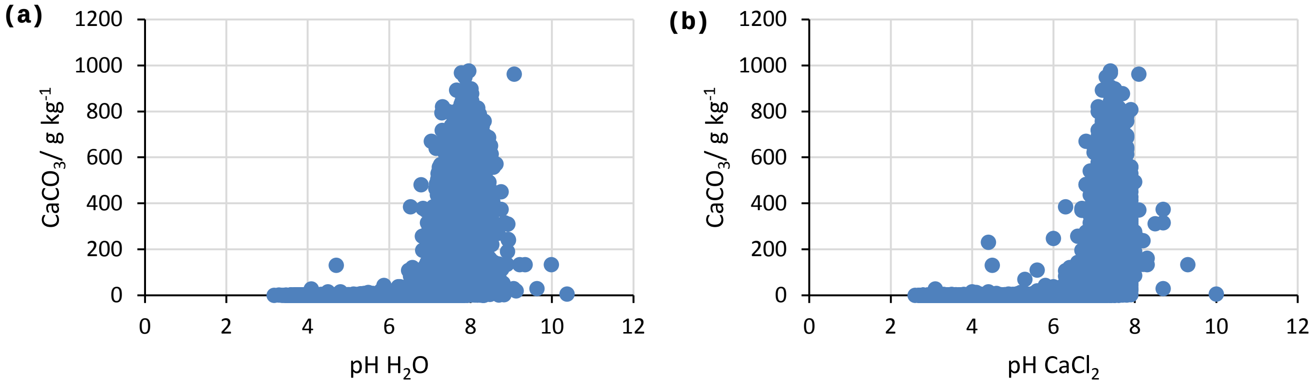
— *Correlation between OC and N*. A close relationship exists between OC and N levels in soil. The higher the OC concentration, the greater the N concentration (Figure 3a). Moreover, the C-to-N ratio is relatively stable across different soil types. Overall, mineral soils generally have a C-to-N ratio close to 12:1 (Figure 3b), while organic-rich soils shall have a C-to-N ratio close to 30:1 (Figure 3c). Soil samples with a C-to-N ratio greater than 40:1 need further consideration, since it is not usual for soil organic matter to have values higher than this.

Figure 3. (a) Relation between OC and N in the whole dataset, (b) relation between OC and C-to-N ratio in mineral soils and (c) relation between OC and C-to-N ratio in organic soils.



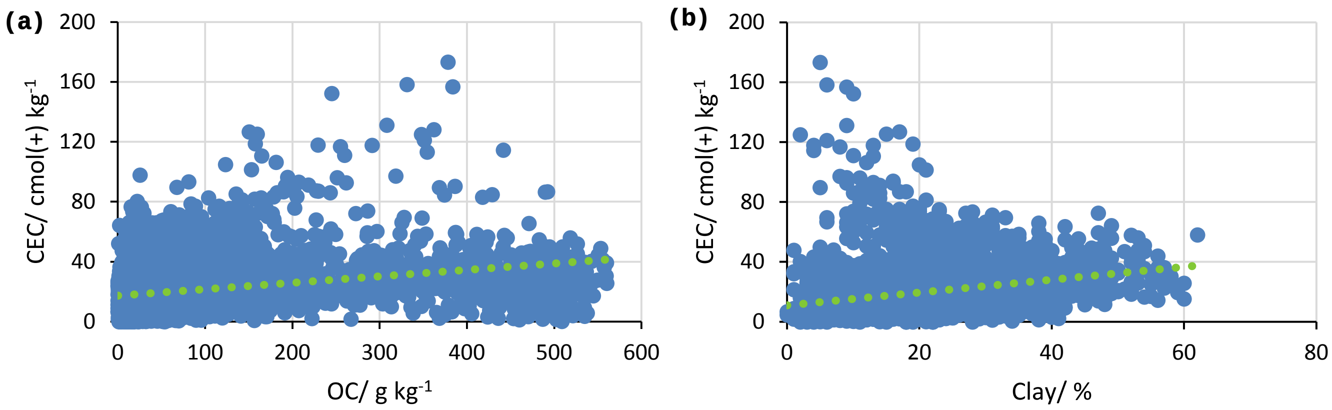
— *Correlation between pH and $CaCO_3$.* pH is a measure of the acidity or alkalinity in the soil. Soil pH can be measured in H_2O and in $CaCl_2$. The values of pH in $CaCl_2$ are normally lower than pH in H_2O by 0.5 to 0.9. Soils have commonly pH in H_2O values between 3.5 and 9.0. Calcium carbonate should not be present (or the concentrations should be very low) in soils where pH is below 7, as its solubility is pH dependent and it does not form under acidic conditions. In accordance with this criterion, Figure 4 shows (i) that pH in H_2O ranges between 3.4 and 10.4 while pH in $CaCl_2$ ranges between 2.6 and 10.0 in LUCAS 2015 samples, and (ii) that soil samples with pH around 7.0 – 8.5 have the greatest contents of $CaCO_3$.

Figure 4. Relation between pH and carbonates ($CaCO_3$) in the whole dataset: (a) pH-H_2O, (b) pH-$CaCl_2$



— *Correlations between CEC and clay and OC*. Cation exchange capacity is the total capacity of soils to hold exchangeable cations. Soils with higher organic matter content and/or clay tend to have greater CEC, because organic matter and clay minerals have negatively charged sites on their surfaces where cations are adsorbed by electrostatic force. Figure 5 shows a positive, though not strong, correlation between OC and CEC, and clay and CEC. This is because soil CEC not only depends on the OC and clay content, but also on the degree of decomposition of organic matter and mineral composition of clay fraction. Apart from OC and clay, soil pH also influences CEC.

Figure 5. (a) Relation between OC and CEC in the whole dataset and (b) relation between clay and CEC in new LUCAS sampling points (i.e. points that where sampled for the first time in 2015).



— *Coherence of particle size distribution data*: Mineral fraction < 2 mm in soil can be split in three size fractions: sand (2-0.05 mm), silt (0.05-0.002 mm) and clay (<0.002 mm). In the LUCAS 2015 survey, these three fractions were measured only in new LUCAS sampling points (i.e. points sampled for the first time in the LUCAS survey) because sand, silt and clay contents in soil are considered to be stable in the short to medium term. To validate the data, the sum of mass of the three fractions shall be equal to 100 %. In the LUCAS 2015 dataset, the sum of sand, silt and clay fractions ranged between 99 % and 101 % because the contents of the three fractions were rounded to the nearest whole number.

Overall, the raw soil data followed the expected trends for assessed correlations. Thus, it can be concluded that the data is coherent from the pedological point of view.

4.4 Closeness of sampling locations in paired samples between surveys

The LUCAS soil points for the topsoil module were selected from the LUCAS regular grid based on land use and terrain information, as indicated in the introduction chapter. Each of these soil points has its theoretical coordinates in the LUCAS grid. For the first topsoil survey in 2009, the triplet concept was used to design the survey (Tóth et al., 2013). Briefly, the surveyors received a list of three alternative sites from the LUCAS grid that have common characteristics of slope, aspect and LC within the surveyed area (this group of site is referred to as triplets). A soil sample was collected from only one of the sites of a triplet. As a general rule, the sample had to be taken in the first site of each triplet. The surveyor had to take the sample in the exact location. If this was not possible, the surveyor had to move to the next site in the triplet to collect the sample ([5]). This triplet concept was

([5]) LUCAS 2009. Instructions for Surveyors: http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS+2009+Instructions/8ffdb9d8-b911-40b6-8f9a-8788bf696aa3 also used in the LUCAS 2012 survey for Bulgaria and Romania. In the 2009 survey, 26 % of the sampling locations were at a distance less than a meter from the LUCAS soil points, 76 % of the locations were at a distance less than 5 m and 96 % of them were located less than 100 m from their LUCAS soil point. In the 2012 survey, 34 out of 1,454 sampling locations (2.3 %) were taken at a distance greater than 100 m from the LUCAS soil points in Bulgaria and Romania. According to the instructions, points should be monitored as close as possible (at a distance less than 100 m), always on the same field parcel.

In the 2015 survey, 80.6 % (17,613 out of 21,859) of the sampling locations of the 2009 and 2012 surveys were revisited. A maximum distance of 100 m, always from the same LC class as observed in the soil point, was allowed between the baseline samples collected in the 2009/2012 surveys and their paired samples collected in the 2015 survey ([6]). Altogether, 80 % of the sampling locations in 2015 were at a distance less than 10 m from their baseline sampling locations in 2009/2012. The percentage increased to 96.5 % when considering a distance less than 100 m between the sampling locations in 2009/2012 and in 2015. Among the 620 locations at a distance greater than 100 m between 2009/2012 and 2015, 362 were at a distance between 100 and 200 m and the rest at a distance between 200 m and 1 km from their baseline location. Surveyors gave different reasons to explain the inaccessibility to the LUCAS soil points for the topsoil sampling. The most common difficulties were the presence of high crops and dense vegetation, ground conditions (mainly waterlogged conditions, and stoniness), presence of fences, refusal of landowner and presence of roaming and dangerous animals.

4.5 Comparability of soil data between surveys

The LUCAS 2009/2012 and LUCAS 2015 samples have been analysed with the same methodologies, but under different conditions: in different laboratories, with different equipment, by different persons and at different times. These differences on analysis conditions can determine the comparability of soil data between LUCAS surveys. To study whether the soil data of the LUCAS surveys are comparable, a set of 214 samples collected in the context of the LUCAS 2009/2012 surveys were re-analysed by the laboratory selected for the analysis of the LUCAS 2015 samples.

A stratified random selection was used to establish the set of samples. Stratification was based on a combination of HYPRES European soil texture classes ([7]) (5 classes: very fine, fine, medium fine, medium and coarse) and OC levels (3 levels: >20 % of OC, 10-20 % of OC, <10 % of OC). Then soil samples were randomly selected from each stratum according to the frequency of occurrence of the strata. The representativeness of the selection regarding the original LUCAS 2009/2012 dataset was verified in terms of OC, texture, $CaCO_3$, N, P, K, pH, and CEC. The selection was run using function `strata` with method `srswor` from the R statistical computing program (R ×64 3.0.3).

The samples were re-analysed with the standard methods detailed in chapter 2 for the following properties: OC, $CaCO_3$, N, P, K, and pH in H_2O and in $CaCl_2$. Clay, silt and sand contents were not re-analysed because different methodologies were used in the LUCAS 2009/2012 and the LUCAS 2015 surveys. The method of sieving and sedimentation (ISO 11277:1998) was used in the LUCAS 2009/2012 surveys, while the method of laser diffraction (ISO 13320:2009) was used in the LUCAS 2015 surveys ([8]). The Bland and Altman diagram was used to detect potential outliers caused by errors of measurement. The Lin's concordance correlation coefficient (LCCC) was calculated to compare the data of soil properties from the original analyses, carried out in the context of LUCAS 2009/2012 surveys, and from the re-analyses carried out in 2015. Furthermore, the root mean square error (RMSE) was calculated to evaluate the average magnitude of the difference between the original data and the data of the re-analyses. The R statistical computing program (R

([6]) LUCAS 2015. Instructions for surveyors: http://ec.europa.eu/eurostat/documents/205002/6786255/LUCAS2015-C1-Instructions-20150227.pdf
([7]) https://www.hutton.ac.uk/learning/natural-resource-datasets/hypres/european-soil-map-texture-classes
([8]) Pedotransfer functions can be used to convert clay, silt and sand contents measured by laser diffraction to the conventional (sieving and sedimentation) method as performed by Makó et al. (2017).

×64 3.0.3) was used for the statistical analyses (see section 3.1 for more details on statistical analysis).

Descriptive statistics of the original data (from analyses carried out in the LUCAS 2009/2012 surveys) and the data of re-analyses carried out in 2015 are detailed in Table 6. The range of values, mean, median and standard deviation of all soil properties were similar in the original data and the data of re-analyses. This suggests that the analytical variance might be very low. Mean and median values were similar for N, CEC, pH in $CaCl_2$ and pH in H_2O data. Thus, data of these properties were evenly divided around their mean values. However, the median was lower than the mean in OC, $CaCO_3$, P and K, which indicates that the distribution of the data is skewed to the right. This means that the mean value of these properties was pulled higher than the median because of few very high values.

Table 6. Descriptive statistics of the original data (LUCAS 2009/2012 surveys) and the data of re-analyses in 2015.

Soil properties	N samples analysed	Original data of LUCAS 2009/2012				Data of the re-analyses in 2015			
		Range	Mean	Median	SD	Range	Mean	Median	SD
OC / g kg^{-1}	214	3.6 – 352.4	34.3	19.6	46.9	2.8 – 323.6	32.8	19.7	42.7
$CaCO_3$ / g kg^{-1}	214	0.0 – 798.0	72.5	1.0	153.1	0.0 – 782.0	70.0	2.0	146.9
N / g kg^{-1}	211	0.0 – 17.0	2.5	2.0	2.4	0.0 – 18.0	2.7	2.0	2.5
P/ mg kg^{-1}	211	0.0 – 334.5	33.8	21.7	38.1	0.0 – 302.8	44.7	32.4	40.1
K / mg kg^{-1}	211	0.0 – 4793.4	231.5	141.9	376.9	0.0 – 4955.7	192.8	110.4	385.1
CEC/ cmol(+) kg^{-1}	213	1.0 – 74.7	15.9	13.6	11.8	2.4 – 46.4	25.5	25.5	8.7
pH−$CaCl_2$	206	2.9 – 8.1	5.8	5.9	1.4	2.9 – 8.1	5.9	6.1	1.3
pH−H_2O	206	3.6 – 8.4	6.4	6.4	1.3	3.4 – 8.1	6.2	6.2	1.3

Table 7 shows the results of the LCCC and RMSE for soil properties between the original data of the LUCAS 2009/2012 and the re-analyses of 2015. All properties had an acceptable average magnitude of difference, as indicated by the RMSE. The concordance between original data and the data of the re-analyses was good for OC, N, K, $CaCO_3$ and pH (the LCCC ranged from 0.97 to 0.99) and moderately good for P (LCCC = 0.85). The LCCC for CEC was very low (0.35), which indicated that the original data and the data of re-analyses did not fit well. When separating organic-rich (OC >100 g kg^{-1}) and predominately mineral samples (OC <100 g kg^{-1}), the LCCC was lower in organic-rich samples than in mineral samples for OC, N, P and K (Table 8). The difference of LCCC between organic-rich and mineral samples was the highest for P (0.67 in organic-rich samples vs 0.86 in predominately mineral samples). In organic-rich samples, P content was lower in the original data of LUCAS 2009/2012 than in the re-analysis of 2015. This is because the P was analysed in field-moist conditions in organic-rich samples in 2009/2012, but it was re-analysed when samples were air-dried in 2015. Drying has been found to substantially increase P solubility (Styles & Coxon, 2006). As a result, P measured in field-moist samples tends to be lower than in dried samples (Brogan et al., 1963) as it occurred in our case. In mineral samples, P was analysed in air-dried conditions both in 2009/2012 and in 2015, resulting in a better LCCC correlation between original and re-analysed results.

Table 7. Lin's concordance correlation coefficient (LCCC) and root mean square error (RMSE) for physical and chemical properties between original data of the LUCAS 2009/2012 surveys and the re-analyses in 2015.

Soil properties	N	Lin's CCC	RMSE
OC / g kg^{-1}	214	0.98	8.50
$CaCO_3$ / g kg^{-1}	214	0.99	21.34
N / g kg^{-1}	211	0.98	0.49
P/ mg kg^{-1}	211	0.85	21.78
K / mg kg^{-1}	211	0.98	81.64
CEC/ cmol(+) kg^{-1}	213	0.35	14.08
pH−$CaCl_2$	206	0.99	0.19
pH−H_2O	206	0.97	0.31

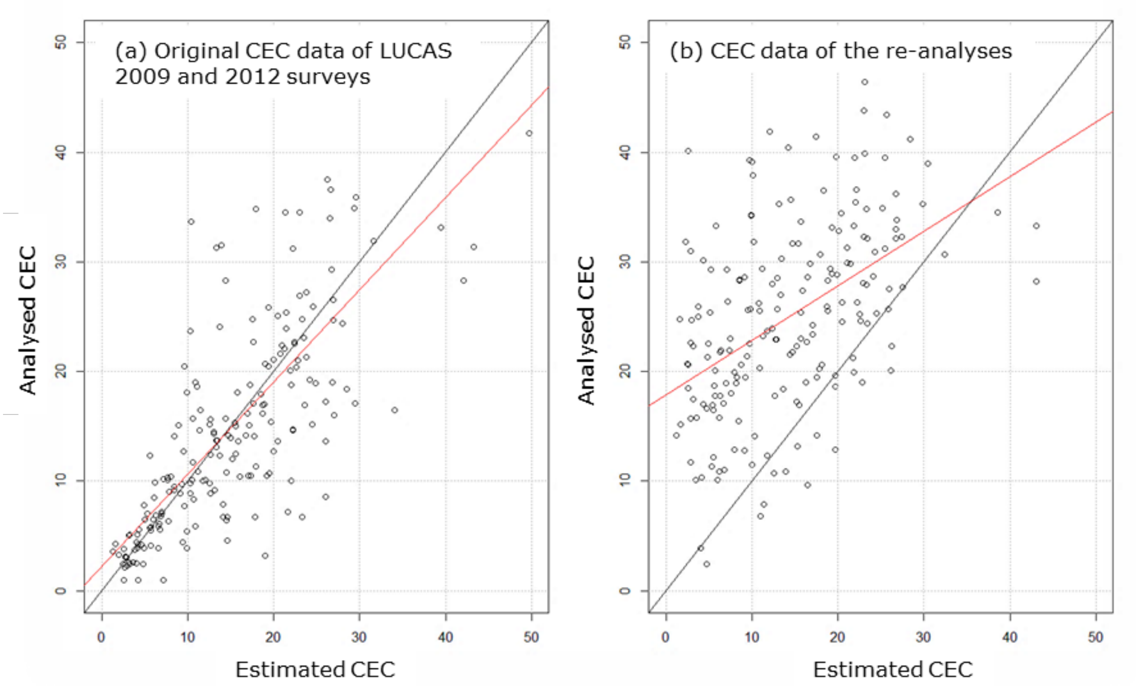
Table 8

Table 8. Lin's concordance correlation coefficient (LCCC) and root mean square error (RMSE) for soil properties between original data of the LUCAS 2009/2012 surveys and the re-analyses in 2015 in organic-rich (OC >120 g kg-1) and predominately mineral samples (

Soil properties	Organic-rich samples			Mineral samples		
	N	Lin's CCC	RMSE	N	Lin's CCC	RMSE
OC / g kg^{-1}	12	0.93	28.40	202	0.95	5.35
$CaCO_3$ / g kg^{-1}	12	0.99	1.96	202	0.99	21.96
N / g kg^{-1}	11	0.93	1.35	200	0.95	0.40
P/ mg kg^{-1}	11	0.67	25.80	200	0.86	21.54
K / mg kg^{-1}	11	0.92	84.48	200	0.98	81.48
pH−$CaCl_2$	6	0.99	0.09	200	0.99	0.20
pH−H_2O	6	0.98	0.18	200	0.97	0.31

Regarding the mismatch between the data of CEC, both the original data of CEC and the data of re-analysis were compared to the estimations of CEC calculated with a pedotransfer function developed by Horn et al. (2005) (Figure 6). The comparison showed that the original CEC data of the LUCAS 2009/2012 surveys fitted better than the CEC data from the re-analysis in 2015 with the CEC data estimated with the pedotransfer function.

Figure 6. Correlation between (a) CEC data estimated with a pedotranfer function and original CEC data (LUCAS 2009/2012 surveys) and (b) CEC data estimated with a pedotranfer function and the CEC data from the re-analysis in 2015.



The pedotranfer function used for the estimations has been developed by Horn et al. (2005): CEC = 0.05 + 0.026*OC*pH (CaCl2) + 0.055*CLAY

The laboratory was asked to control any factor that could affect to the analysis of CEC with the ISO 11260:1994 method. The laboratory checked the whole protocol of analysis, and carried out quality controls with local reference materials and certified reference materials. In addition, an inter-laboratory control was carried out, in which the CEC of a local reference material and various soil samples (with contrasting physical-chemical properties) was measured in two different laboratories. The repeatability and reproducibility of the results of CEC analysis carried out by the two laboratories was under the limit specified by the ISO 11260:1994 method (20 % and 40 % respectively). Thus, it was concluded that there was no methodological problems on the determination of CEC. However, the variability of CEC between original data of the LUCAS 2009/2012 and the data of re-analysis in 2015 suggested that CEC is difficult to determine accurately. Experimental factors related to clay minerals could explain this difficulty to determine CEC accurately (Bergaya & Lagaly, 2013). For example, drying duration, storage conditions, and grinding of soil samples can influence the properties of clay minerals (one of the major contributors of CEC in soils) and consequently affect to the determination of CEC.

This comparison showed that, except for CEC, the results of the analyses carried out in the LUCAS 2009/2012 surveys were reproducible, with a little unexplained variability, in the laboratory on charge of the analyses of the LUCAS 2015 samples. Thus, it can be concluded that the data from the different LUCAS surveys are directly comparable.

5 Dataset of revisited points for comparison of LUCAS 2009/2012 and 2015 surveys

In LUCAS 2015, samples were collected from 80 % of the soil points already visited in the surveys of 2009. We planned to repeat the sampling in 90 % of the soil points, however, the set was reduced by 10 % due to problematic weather conditions, such as freezing and flooding, which complicated the collection of samples together with issues on the identification and registration of samples taken (section 4.2). Finally, a set of 17,613 samples taken in revisited points (after excluding non-identified samples due to repetition of Soil IDs and missing registration in the DMT) through 27 member states were analysed to assess changes in key soil properties between 2009/2012 and 2015 surveys (Figure 7). All points are at altitudes below 1,000 m. Table 9 shows the distribution of revisited points in each member state. The percentage is greater than 60 % in all member states, except in Finland (59 %) and Ireland (53 %).

Some revisited points in Bulgaria and Romania had the value <6 g kg^{-1} for OC (153 points) and <5 mg kg^{-1} for P (459 points) and <0.5 g kg^{-1} for $CaCO_3$ (472 points) in the 2012 survey. These values indicated that OC, P and $CaCO_3$ contents in these points were below the limit of detection. Furthermore, the value -999 was used to indicate that no result is available for an analysis in the 2009 and 2012 surveys. The value -999 has been substituted by NA in the dataset. These revisited points with data below the limit of detection and missing information were excluded from the assessment of changes between surveys.

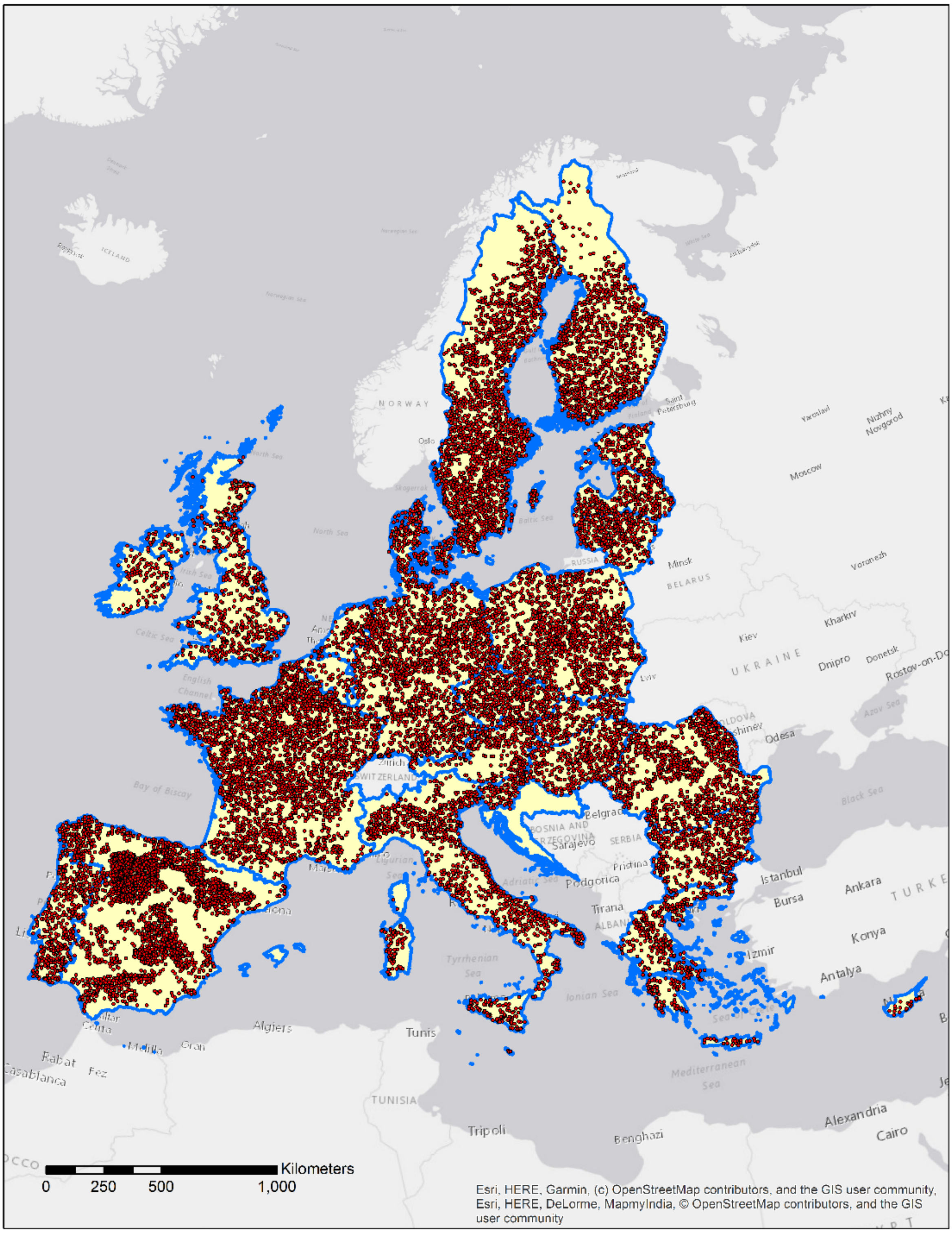
Table 9. Number of revisited points in each member states.

Member states	**N samples 2009/2012**	**Revisited points 2015**(1)	**% paired samples** (2)	**Member states**	**N samples 2009/2012**	**Revisited points 2015** (1)	**% paired samples** (2)
Austria	420	300	71	Latvia	349	298	85
Belgium	71	64	90	Lithuania	356	335	94
Bulgaria	661	467	71	Luxembourg	3	3	100
Cyprus	90	20	22	Malta	19	3	16
Czech Republic	431	420	97	The Netherlands	211	137	65
Denmark	232	218	94	Poland	1648	1,324	80
Estonia	220	182	83	Portugal	476	425	89
Finland	1716	1,021	59	Romania	1373	978	71
France	2952	2,640	89	Slovakia	268	206	77
Germany	1947	1,591	82	Slovenia	112	106	95
Greece	491	426	87	Spain	2696	2,484	92
Hungary	497	397	80	Sweden	2256	1,847	82
Ireland	233	123	53	UK	942	595	63
Italy	1333	1,003	75	**TOTAL**	**22,003**	**17,613**	**80**

(1) Points visited in 2009 and 2015 surveys in all member states, except in Bulgaria and Romania. In Bulgaria and Romania, points were visited in 2012 and 2015.

(2) Percentage of revisited points is calculated based on the points sampled in LUCAS 2009 survey (LUCAS 2012 survey for Bulgaria and Romania).

Figure 7. Spatial representability of revisited LUCAS.



Together with this report, the dataset of the LUCAS 2015 soil points can be accessed from the European Soil Data Centre (ESDAC) upon registration in the following URL: (https://esdac.jrc.ec.europa.eu/resource-type/datasets). The dataset contains analytical data of $$$$$$ soil points, including also the revisited points, from the EU-28 member states. Details of the dataset are described in the report of Fernández-Ugalde et al. (2020) that presents soil data and results of the 2015 survey.

6 Assessment of changes in soil properties between 2015 and 2009/2012 surveys

Changes in soil properties in LUCAS topsoil samples between 2015 and 2009/2012 surveys are presented for main LC classes (cropland, woodland, grassland, bareland and shrubland) and at NUTS 2 levels. Cropland includes fields of cereal, root crops, industrial crops, dry pulses and vegetables, fodder crops, fruit trees, olive groves and vineyards. Woodland includes deciduous, coniferous and mixed forests, and grassland includes fields of grass with and without sparse trees below 1,000 m altitude. Shrubland refers to shrub cover with or without sparse trees, and bareland can be agricultural fallow lands (tilled or not), all kind of rock and stone litters, areas of sand, lichens and moss.

Table 10 shows the changes in the distribution of revisited points among LC classes between surveys. Almost 89 % of the cropland points in 2009 remained in the same LC class in 2015. Nearly 7 % of the cropland points changed to grassland and 3.9 % changed to bareland in 2015. Points that passed from cropland to grassland were homogeneously distributed in all agricultural zones of the EU. Although in the north of France, with a high density of agricultural points, almost no changes to grassland were observed (Figure 8). Regarding changes from cropland to bareland, many of the points were in the semiarid region of Spain (Figure 8). Among woodland points in 2009/2012, 95.1 % of the points remained in woodland in 2015, 2.3 % changed to grassland and 1.8 % changed to shrubland. Most of the changes from woodland to grassland or to shrubland occurred in the northern MS, such as Finland, Sweden and Latvia (Figure 8). Several changes to grassland or shrubland were also observed in Portugal. Among grassland points in 2009/2012, 75.7 % of the points remained in grassland in 2015, 15.9 % of the points changed to cropland, 4.1 % to woodland and 2.3 % to shrubland. Almost 54.4 % of the points in shurbland in 2009/2012 remained in the same LC class and 29.3 % changed to woodland in 2015. Changes from grassland to cropland were homogeneously distributed in all MS, although the number of points that changed in Finland and Sweden were very few (Figure 8). In these northern member states most of the changes observed were from grassland to woodland. Regarding bareland points, only 27.6 % of the points in 2009/2012 remained in bareland in 2015. Around 44 % of the bareland points changed to cropland and 17.6 % changed to grassland in 2015. Finally, 64.1 % of the wetland points in 2009/2012 remained in the same LC class in 2015, 16.9 % changed to woodland and 15.1 % changed to grassland or shrubland.

Table 10. Changes in the distribution of points among land cover (LC) classes between 2009/2012 and 2015 surveys.

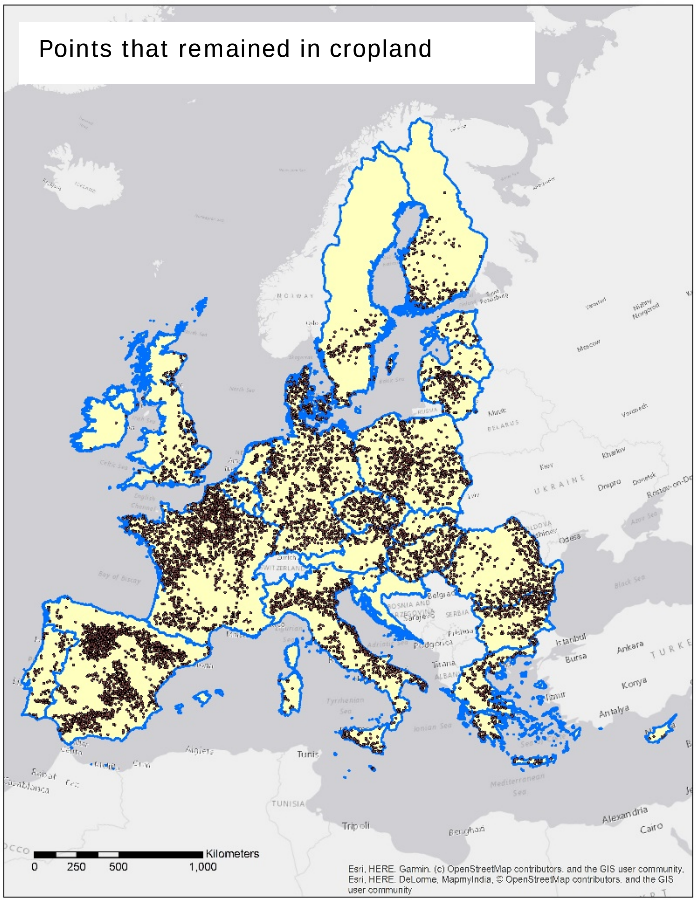
LC classes (from 2009/12 to 2015)	N points	Percentage (%)
Cropland to		
cropland	7374	88.7
grassland	567	6.8
bareland	327	3.9
other LC classes	41	0.5
Woodland to		
woodland	4424	95.1
grassland	106	2.3
shrubland	87	1.8
other LC classes	35	0.7
Grassland to		
grassland	2955	75.7
cropland	620	15.9
woodland	160	4.1
shrubland	88	2.3
other LC classes	72	1.8
Shrubland to		
shrubland	204	54.4
woodland	110	29.3
other LC classes	61	16.3
Bareland to		
bareland	86	27.6
cropland	138	44.2
grassland	55	17.6
other LC classes	33	10.6
Wetland to		
wetland	34	64.1
woodland	9	16.9
grassland & shrubland	8	15.1
other LC classes	2	3.8
Artificial land to		
artificial land	6	46.1
other LC classes	7	53.8

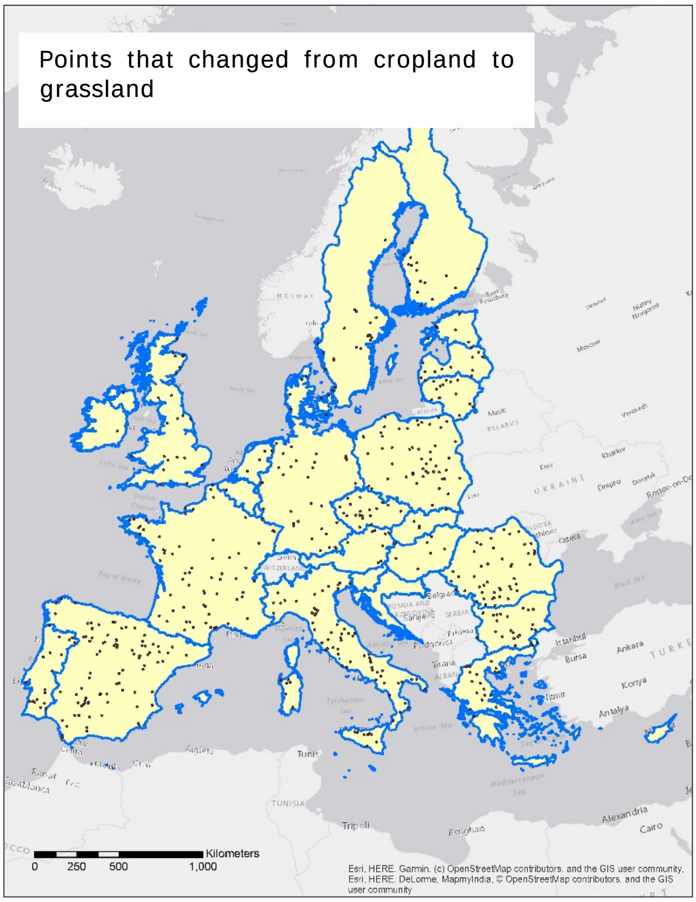
Overall, the distribution of revisited points in the main LC classes in 2015 was as follow: 46.3 % in cropland, 26.9 % in woodland and 21.2 % in grassland. The rest of LC classes (shrubland, bareland, wetland, artificial land and water) had less than 3 % of the revisited points each in 2015 (Table 11).

Table 11. Revisited points by land cover (LC) class in LUCAS 2015 survey.

LC classes	N points	Percentage (%)
Cropland	8154	46.3
Woodland	4746	26.9
Grassland	3741	21.2
Bareland	486	2.7
Shrubland	401	2.3
Wetland	42	0.2
Artificial land	40	0.2
Water areas	3	0.02

Figure 8. Distribution of revisited points by land cover between 2009/2012 and 2015 survey





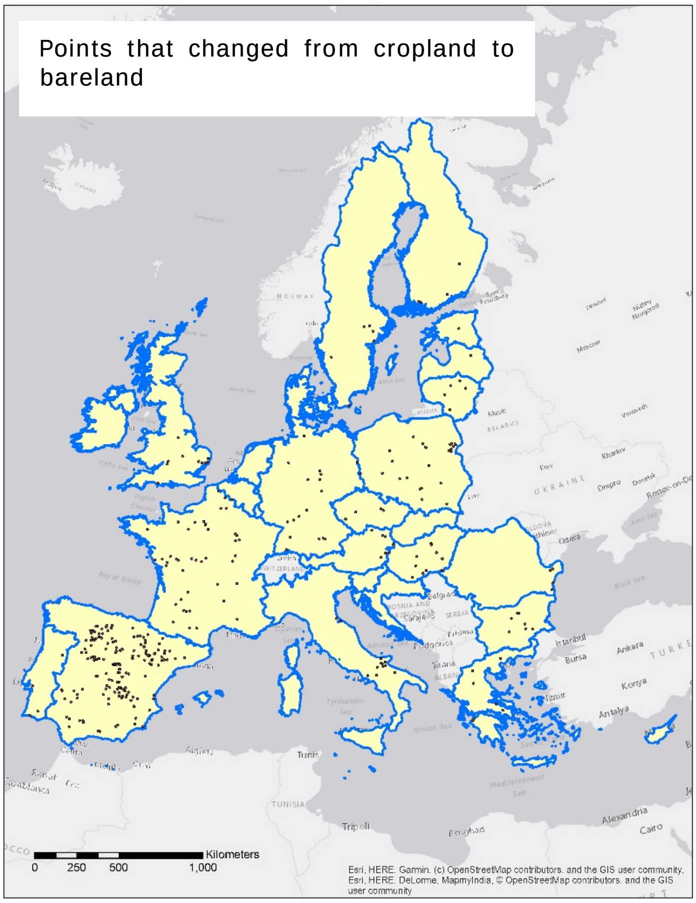


Figure 8 (cont). Distribution of revisited points by land cover between 2009/2012 and 2015 surveys

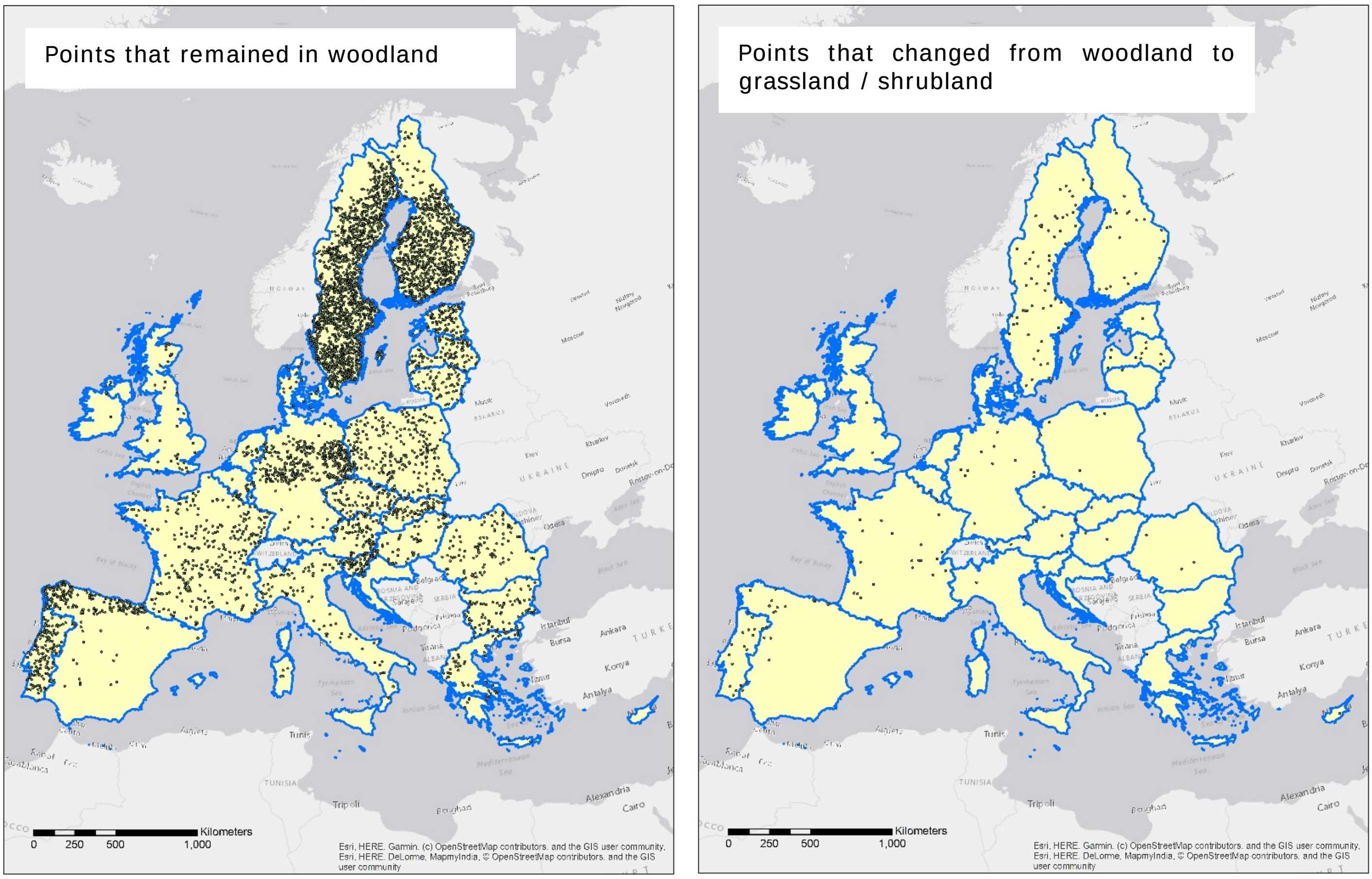
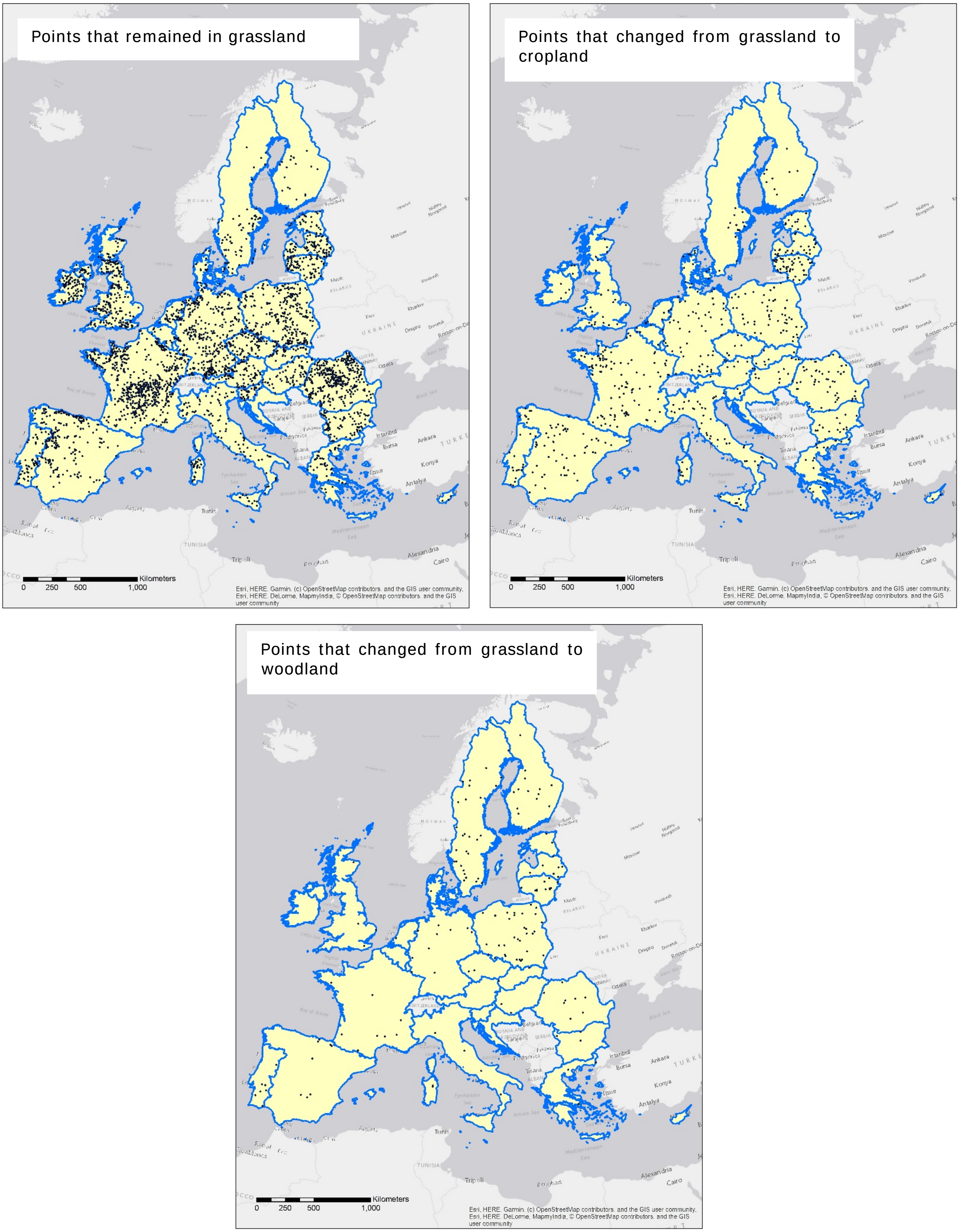


Figure 8 (cont). Distribution of revisited points by land cover between 2009/2012 and 2015 surveys



For each property, we used the linear least squares regression (LLSR) to study whether the changes in properties were correlated with the distance between sampling locations in revisited points between surveys. F-test was used to test whether this correlation was significant ($P < 0.05$). The coefficient of correlation (r) between distance and changes in soil properties was near 0, which indicated that there was very little relation between the two variables.

We performed the paired Student t-test at a confidence level of 0.95 to test the statistical significance of changes in soil properties between surveys. Variation in soil properties (both within a survey and between surveys) is often large, and this makes necessary a large number of samples for assessing significant changes in soil properties. We performed a power analysis for the Student t-test for determining how many soil samples we would need to detect a change of a specified size with a significant level of 0.05 and a power of 90 %. If the number of samples needed was larger than what we have, then we considered that changes in soil properties were tangible.

We also calculated the percent change of OC content between surveys and established threshold ranges to explain its policy relevance. Reflecting the goals of the '4 per 1000' initiative, we established a threshold range of ±2.4 % between 2009 and 2015 surveys and ±1.2 % between 2012 and 2015 surveys ([9]). If the percent change of OC was outside this range, we considered the change policy relevant. If the percent change was within the range, no policy-relevant change was assumed.

Changes in soil carbon content were assessed by fitting a boosted trees model on the measured SOC concentrations of the samples taken in the 2009/2012 and those taken in the 2015 LUCAS survey. The gradient boosting machines (GMB) aim to minimise the loss function (a measure of difference between the observed and predicted values) by combining a sequence of base-learner models. A common optimisation method to find a minimum is gradient decent which involves going down a gradient to reach a minimum. The key idea behind gradient boosting machines is to sequentially add a new base learner model to the ensemble sequence such that the new model is the model with the greatest correlation with the negative of the loss function's gradient calculated using the current ensemble sequence predictions.

The GMB algorithm is a boosting algorithm that sequentially combines decision trees such that each additional tree is trained with more weighting placed on correctly predicting data-points that the previous decision trees misclassified. In simple terms, each new tree aims to correct for the mistakes of the previous trees. Gradient boosting machines have been successfully implemented across a range of classification tasks but are known to have performance issues when there is noise present in the data.

The hyper-parameters of the gradient boosting machine control the complexity of the learning function. A GBM with a high max depth (maximum number of interactions between independent variables), high ntree (number of trees) and low observations per node (minimum number of data points required for each end node) is more complex and thus, more prone to overfitting. Therefore, limiting the max depth, ntree or increasing the observations per node can effectively perform regularisation and reduce the chance of overfitting.

The grid search for the hyper-parameters investigated in our models were: ntree = 10, 50, 125 and 200; max depth = 2, 4, 6, 8, and 10; the minimum observations per node was 2, 10, 50, and 100. The GMB model was chosen to have a Bernoulli distribution and the chosen model had the hyper parameters: ntree = 125, max depth = 4 (up to 4 variable

([9]) The '4 per 1000' initiative stands up for an increase of OC stock of 0.4% per year in the first 30-40 cm of soil. This would reduce significantly the annual increase of CO_2 in the atmosphere (https://www.4p1000.org/). Even if the initiative refers to stocks of OC, we decided to use the threshold value for OC concentration in this report. Overall, if OC concentration achieves the threshold of 0.4% in the 0-20 cm depth, the stock of OC in the same depth should also achieve it.

interactions were used by the model) and the minimum number of observations per node was 50.

Mapping methodology and data sources

To support the spatial predictions of soil properties, a series of datasets or covariates were selected according to their possible influence on soil chemical properties. The spatial resolution of the covariates was set to 250m, as a compromise between the resolution of the Moderate-resolution Imaging Spectro-radiometer (MODIS) data (500m), the finer resolution of the DEM (25m) and the coarser WorldClim climatic (1km) datasets. Overall 100 numeric and 99 dummy covariates were considered in the first steps of the analysis. The dummy covariates were obtained from the coding of the categorical variables classes (CORINE, parent material type) into dichotomous variables.

MODIS and derived data

A series of MODIS image products for 2009 was collected; in particular, the MODIS Global vegetation indices (Didan, 2005). These products are characterised by a spatial resolution between 250 and 500 m and a temporal resolution of 16 days. The products include blue, red and near- and mid-infrared reflectance, centered at 469 nm, 645 nm, and 858 nm respectively. The reflectance is used to determine the MODIS daily vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI).

NDVI is defined as $NDVI = (NIR - RED)/(NIR + RED)$,where NIR and RED stand for the spectral reflectance measurements acquired in the near-infrared and visible (red) regions, respectively. NDVI has been used to estimate a large number of vegetation properties from its value, such as biomass, chlorophyll concentration in leaves, plant productivity, fractional vegetation cover and accumulated rainfall.

The EVI index is defined as:

$$EVI = g \cdot \frac{NIR-RED}{NIR+c1 \cdot RED-c2 \cdot BLUE+L} \qquad (1)$$

where NIR, RED, and BLUE are the respective surface reflectance in the corresponding spectral bands, L is the canopy background adjustment, and c1 and c2 are coefficients for the aerosol resistance term, which uses the blue band to correct for aerosol influences on the red band. The coefficients adopted by the MODIS-EVI algorithm are; l = 1, c1 = 6, c2 = 7.5, and g (gain factor) = 2.5.

Phenological indices were derived from MODIS data using a first order harmonic model on the EVI and NDVI multi-temporal data. The harmonic uses a discrete Fourier processing that decomposes temporal curves in a linear trend plus amplitude, variance and phase metric terms. The harmonic model can be defined as:

$$\widehat{Y_t} = \alpha_0 + \sum_{j=1}^{m} \alpha_j \cos\left(\frac{j2\pi t}{l}\right) + \beta_j \sin\left(\frac{j2\pi t}{l}\right) \qquad (2)$$

where $\widehat{Y_t}$ is the vegetation index value, t is the time value for a given pixel, l is the cycle length (yearly) and m is the order of the trigonometric polynomial and coincides with the number of harmonics of the expansion (set as one in this study), α_j and β_j are the Fourier coefficients.

Harmonic analysis using Fourier series, has been used to model the temporal changes in the vegetation cover using satellite data for several decades and provides better spatial information on the different types of vegetation cover than using composite images alone. Additionally, a Principal Component Analysis (PCA) transformation of the full MODIS 16-day images time series was performed for each band in order to extract relevant features. The PCA projects the time correlated input images into uncorrelated PCA components ordered according to their variance. Thus, the first few components account for most of the time related variation in each MODIS band.

Terrain parameters

The EU-DEM digital elevation model was used to derive land features at a resolution of 25 m for all Europe. Both the DEM and the derived surface parameters were then rescaled to 250 m. The derivation of land surface parameters was made using the SAGA GIS software. Among the various parameters derived and tested, the most relevant were the Multi-resolution Valley Bottom Flatness (MRVBF) and the Multi-resolution Ridge Top Flatness (MRRTF) (Gallant and Dowling, 2003), slope, slope height and vertical distance to channel network (CNBL).

Land Cover

The CORINE (CORdinate INformation on the Environment) is a raster format land cover database comprising 44 classes. CORINE is derived from Earth observation satellites using computer-aided photointerpretation. The nominal scale of CORINE is 1:100,000 with a minimum mapping unit (MMU) of 25 ha and a change detection threshold of 5 ha. The CORINE dataset was used to represent the spatial distribution of land use/and land cover. The reliability of CORINE 2000 version at 95 % confidence level is 87.0 ± 0.7 %, according to the independent interpretation performed on the LUCAS data.

Climate data

Monthly temperature averages and extremes, and monthly average precipitation values were obtained from the WorldClim (http://www.worldclim.org/) dataset at a spatial resolution of 1 km^2. These data layers are the interpolated values of average monthly climate data collected from numerous weather stations. The approach uses a thin plate smoothing spline with latitude, longitude and elevation as independent variables to locally interpolate data. Climatic data was included explicitly in the model in the form of monthly values of minimum and maximum temperature and monthly rainfall rates. Also the bioclimatic variables (Temperature and precipitation indexes) of WorldClim were included in the analysis. Given the high collinearity of climate data, a careful feature selection procedure was applied in the model training stage.

Legacy soil data and parent material geochemistry

In the first stage of this study, the European Soil Database (ESDB) (was considered as a possible covariate to characterise soil properties. In this context, the ESDB was utilised as a multinomial variable by identifying and labelling soil types. However, the use of the ESDB soil data was found to provide little improvement to the model outcome and was then removed from the analysis. Nonetheless, the data within the ESDB was used to create a map of the parent material geochemistry that was included in the model.

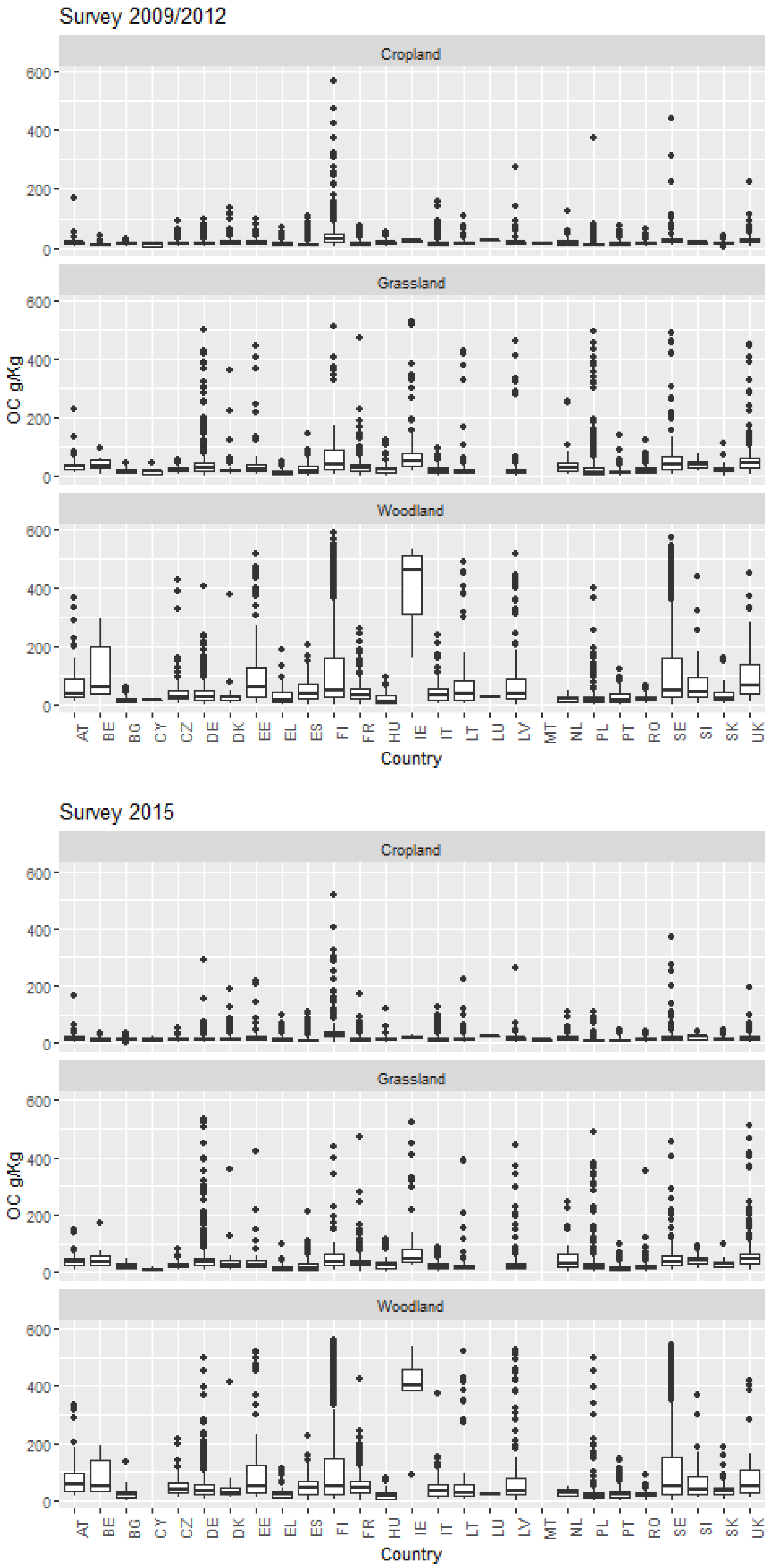
6.1 Changes in OC and N between surveys

Overall, OC and N contents were highest in woodland, followed by grassland and cropland in all member states in both the 2009/2012 and the 2015 surveys (Figures 9 and 10). Organic carbon and N contents in cropland were quite similar among member states in both surveys. On the contrary, OC content in woodland differed more among member states in the two surveys. This is due to the diversity of forest types and their litter layers across the different soil types and climatic conditions in the EU. For cropland, however, the regular tillage and similar management practices resulted in homogeneous contents of OC and N in soil despite pedological and climatic differences across member states/regions.

Both OC and N showed a positive linear relationship between surveys, when representing data of the 2009/2012 survey against the data of the 2015 survey (Figures 11 and 12). This means that OC and N contents moved in the same direction in the two surveys, despite the variability of the data. Sweden and Finland have the largest number of points with extreme changes of OC and N contents between surveys (Figures 11 and 12). Most of these points were located in woodland, as 81 % of the revisited points in Sweden and 76 % in Finland are in this LC class. In Sweden, OC content in woodland varied between -98 % and 8,031 % with respect to OC content in 2009 survey. In Finland, it varied between -98 % and 2,684 % with respect to 2009 survey. Similarly, N content varied between -98 % and 1,660 % in Sweden and between -96 % and 1,680 % in Finland with respect to N content in 2009 survey. These large variations on OC and N contents in woodland were probably due to an incorrect removal of the litter layer before the soil sampling in the surveys. The litter layer is a key source of OC and N for soil. If the litter is not completely removed or, on the contrary, it is excessively removed, OC and N content of soil would be measured inaccurately. This is a problem, especially in organic-rich woodland soils like those in Sweden and Finland. Difficulties of the soil sampling in woodland were already described in chapter 3. Briefly, surveyors need a better training to identify and accurately remove the litter layer, thus avoiding errors in the analyses due to the sampling.

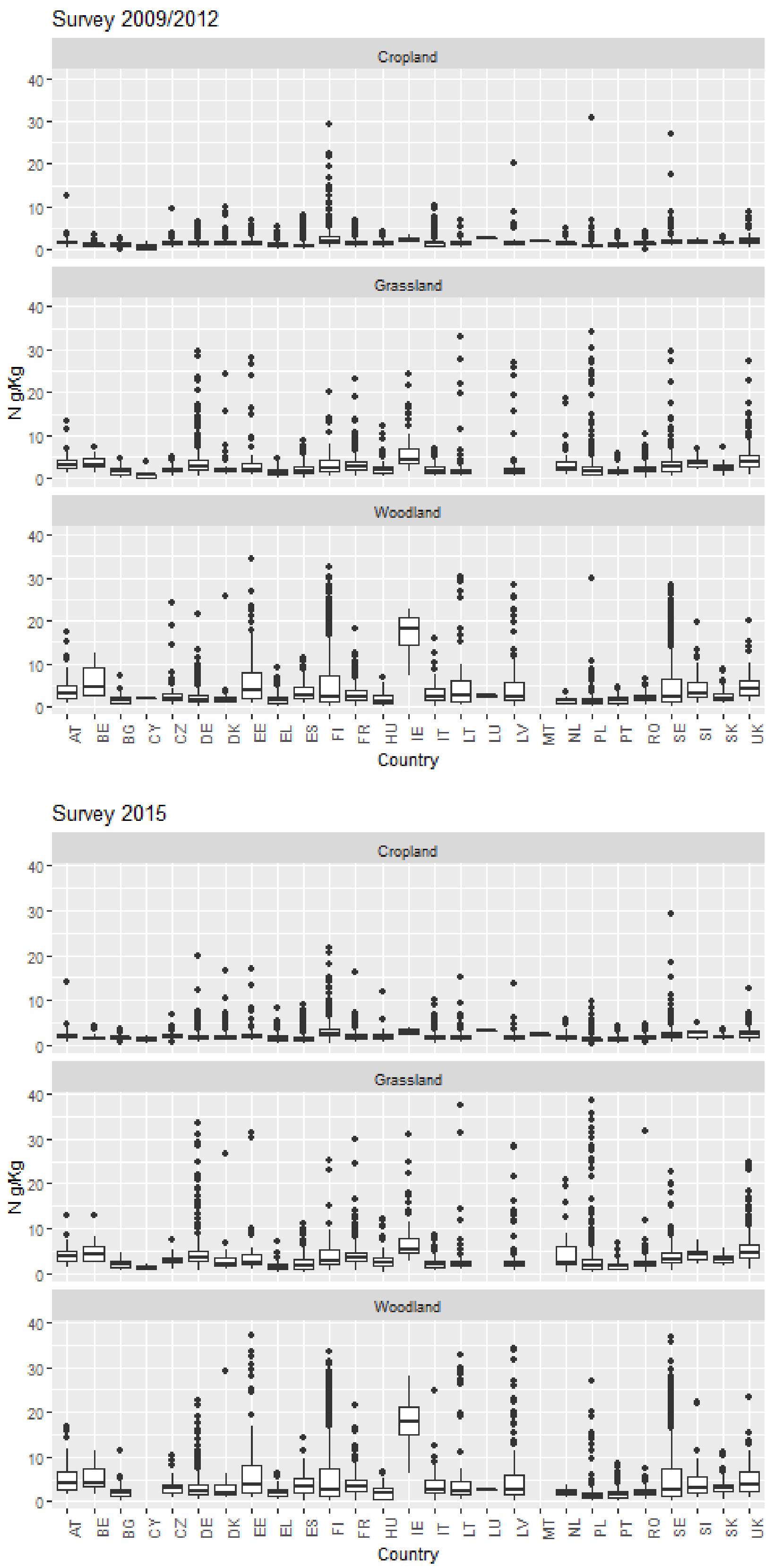
Figure 13 shows changes in OC content vs. distance between sampling locations (in 2009/12 and 2015) for each soil point in the EU member states. Similarly, Figure 14 shows changes in N content vs. distance between sampling locations. Both for OC and N contents, the correlation coefficient (*r*) for LLSR between the distance and the soil property was near 0, which indicated that variation in sampling locations very little explained variation in OC and N contents between surveys. Only in Ireland, changes in OC and N contents slightly increased while increasing the distance between sampling location in the two surveys. On the contrary in Belgium and United Kingdom, changes in OC and N contents slightly decreased while increasing the distance between sampling locations in the two surveys. In Sweden and Finland, despite the large changes in OC and N contents between surveys, those were not affected by the distance between sampling locations. As a result, we did not remove paired samples taken far away (>100 m, see chapter 4 section 4.4) from each other when performing the statistical analysis to assess changes in OC and N contents by LC and NUTS 2 region between surveys.

Figure 9. Box plots of organic carbon (OC) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys



Bold horizontal lines in the boxes are the median values of each member state

Figure 10. Box plots of nitrogen (N) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys. Box plots of nitrogen (N) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys.



Bold horizontal lines in the boxes are the median values of each member state

Figure 11. Representation of organic carbon (OC) content in the 2009/2012 survey against the content in the 2015 survey

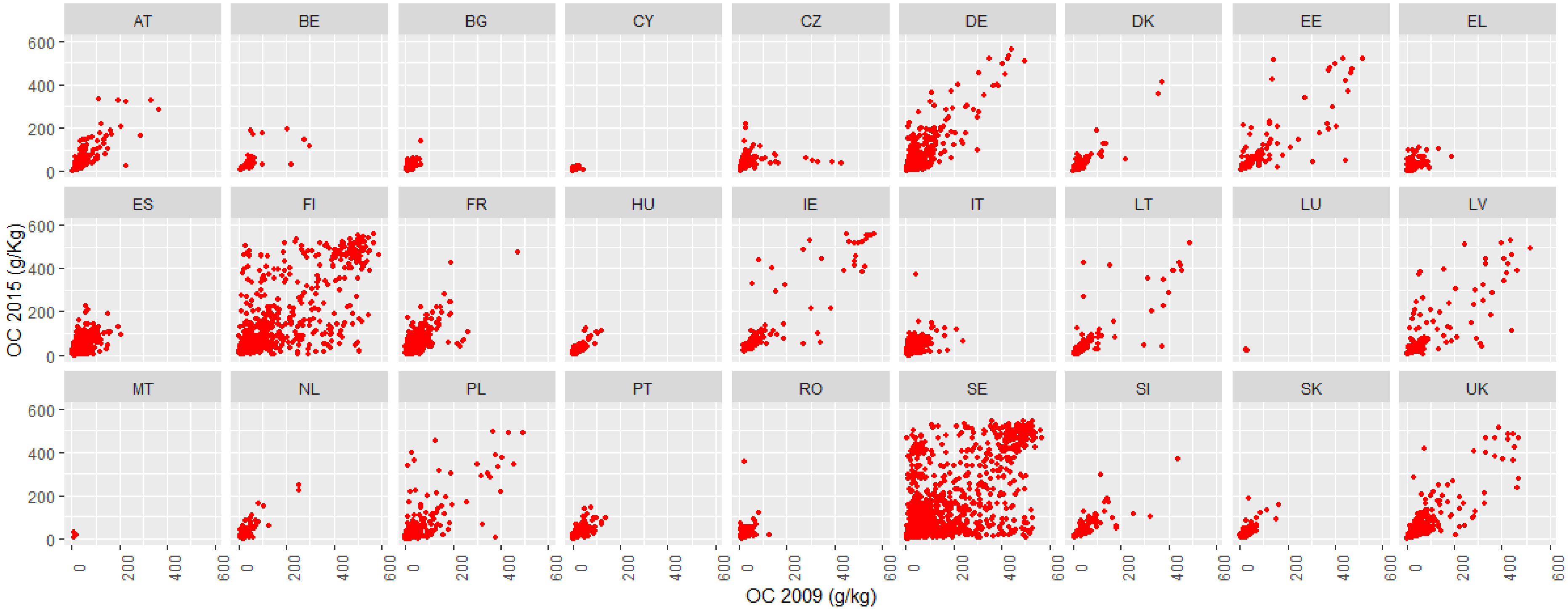


Figure 12. Representation of nitrogen (N) content in the 2009/2012 survey against the content in the 2015 survey

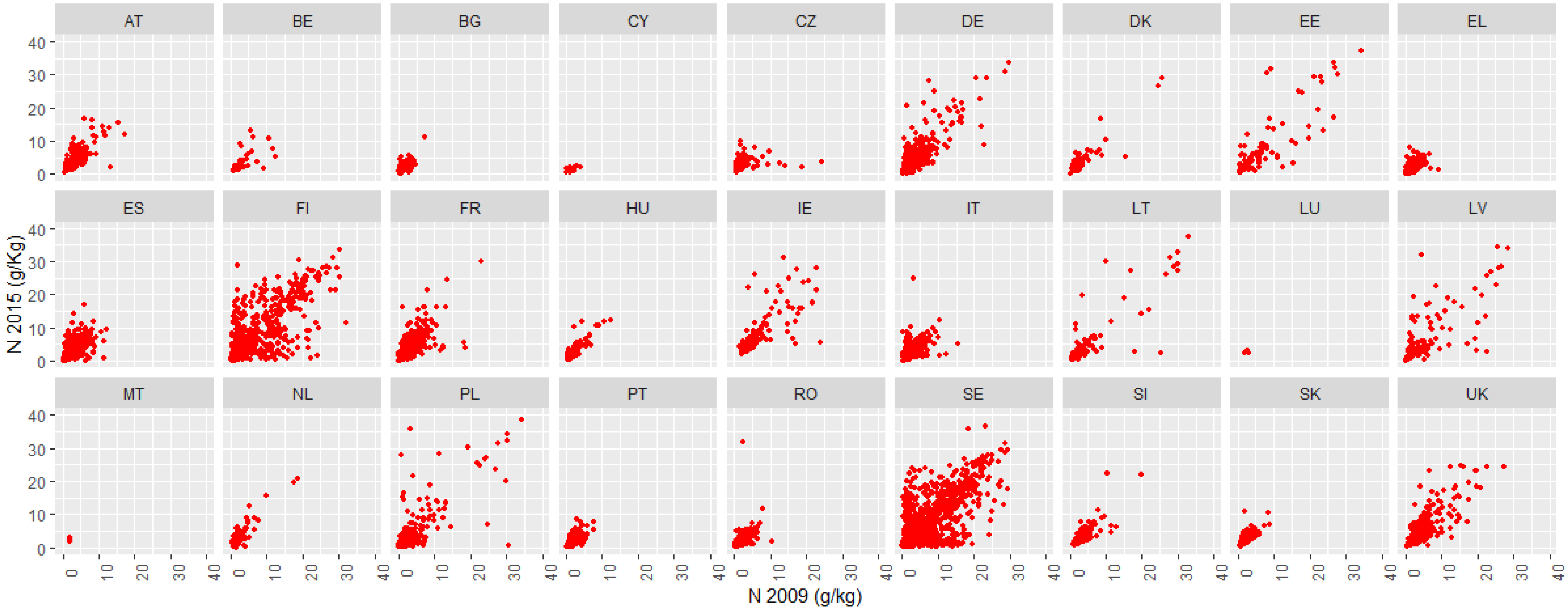


Figure 13. Changes in organic carbon (OC) content by member state (NUTS 0) against the distance of sampling locations between surveys

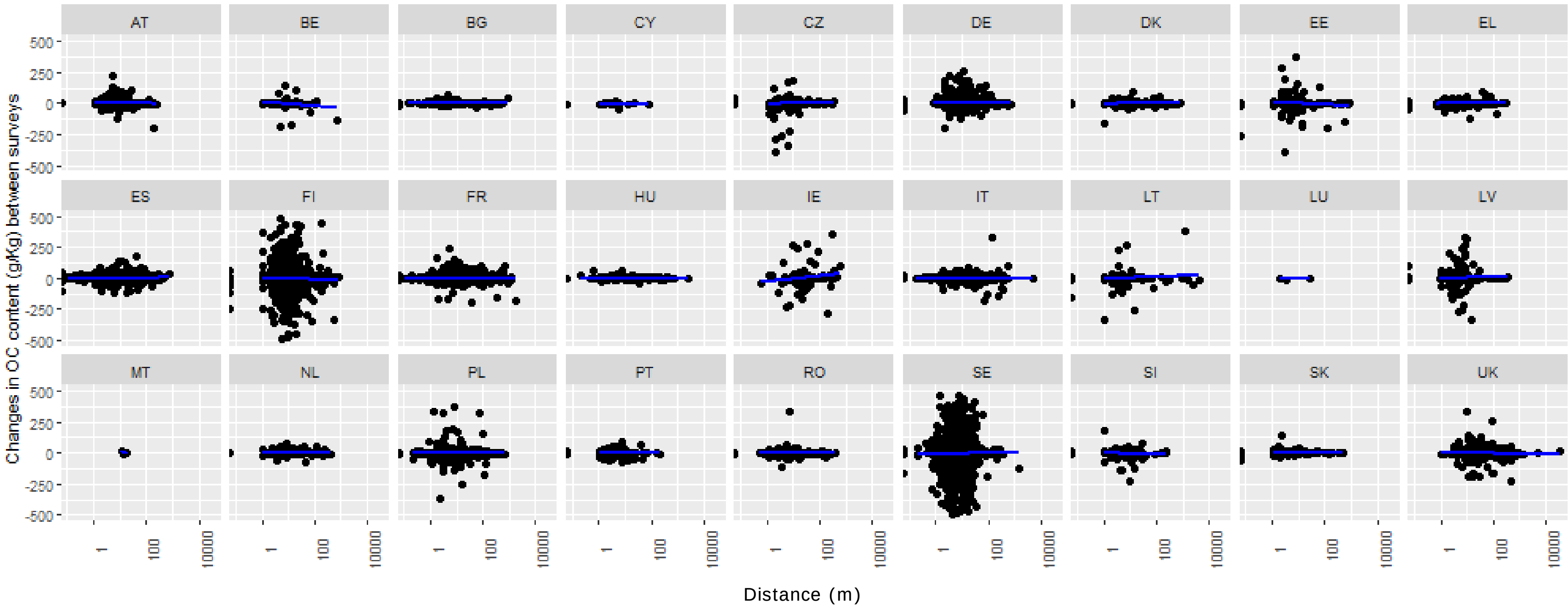
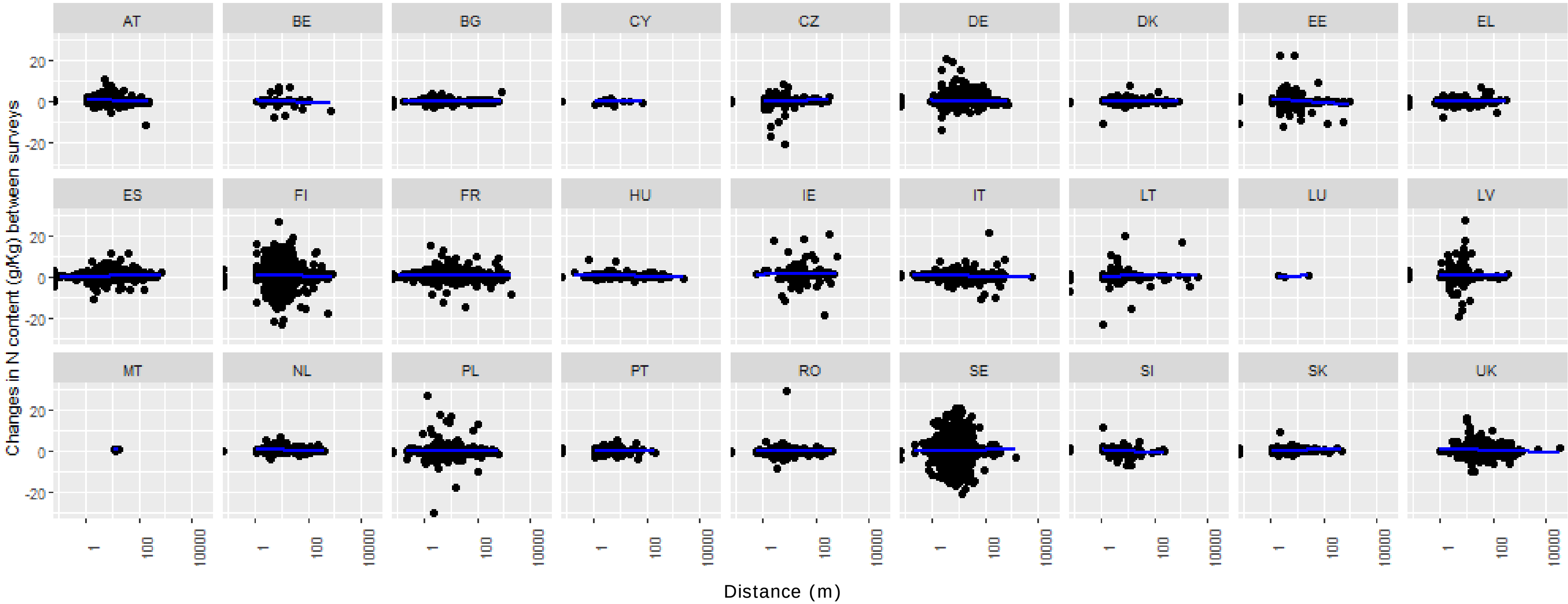


Figure 14. Changes in nitrogen (N) content by member state (NUTS 0) against the distance of sampling locations between surveys



When considering all revisited points, we observed that the standard deviation of changes in OC content was very large in all LC classes (Tables 12a and 13a). This was at most due to the presence of organic-rich points that showed a great variability of OC contents between surveys, as observed especially in woodland in Finland and Sweden (Figure 11). The large variability of OC content make necessary a large number of samples to detect little changes in OC content as those observed in many LC classes (Tables 12a and 13a). In some cases, we did not have enough number of samples/points to assume the changes in OC observed between surveys with a power of 90 % and a significant level of 0.05. In the next paragraphs, we only describe those changes that can be considered tangible/real with the number of samples we have in each LC class.

Despite the large standard deviation in OC changes of revisited points, OC content showed a significant decrease in points that remained in cropland and that changed from grassland to cropland in the 27 member states between 2009 and 2015 surveys (Table 12a). On the contrary, OC content showed a significant increase in points that remained in grassland between 2009 and 2015 surveys (Table 12a). In the three LC classes, the number of samples was enough to accept the significant changes in OC with a power of 90 % and a significant level of 0.05. In Bulgaria and Romania, the number of samples was not enough to assume a significant change of OC content, even if the Student t-test showed differences between surveys in some cases (Table 13a).

When focusing only in mineral points (OC content <120 g kg^{-1}), standard deviation of differences on OC content was lower in all LC classes (Tables 12b and 13b). Organic carbon content significantly decreased in points that remained in cropland and increased in points that remained in grassland between 2009 and 2015 surveys in the 27 member states (Table 12b). In both cases, the number of samples was enough to accept these significant changes with a power of 90 % and a significant level of 0.05. The percent decrease of OC in cropland points was close to the threshold value of -2.4 % (Table 12b), which suggests that these soils could be losing OC; and thus, are not carbon sinks. On the contrary, the percent increase of OC in grassland points was above the threshold of +2.4 % (Table 12b) suggesting that grassland points are gaining enough OC to assure that the increase of atmospheric CO_2 would be significantly reduced as proposed by the ‘4 per 1000’ initiative. These results are in line with the observations of Hiederer (2018) that concluded that there may have been real changes in OC content on agricultural and grassland points from 2009 to 2015. In Bulgaria and Romania, we did not have enough samples to assume that changes in OC showed by the Student t-test in points remaining in grassland are tangible with a power of 90 % and a significant level of 0.05 (Table 13b).

Regarding mineral points with changes in LC classes between surveys, OC content decreased significantly in points that changed from grassland to cropland in the 27 member states between 2009 and 2015 (Table 12b). The number of samples was enough to accept this change in OC with a power of 90 % and a significant level of 0.05. The percent decrease was greater than the threshold value of -2.4 %, suggesting that soils loss their carbon stocks when LC changes from grassland to cropland. In points that change from cropland to grassland in the 27 member states, the number of samples was not enough to accept an increase in OC content with a power of 90 % and a significant level of 0.05. Between 2012 and 2015 surveys in Bulgaria and Romania, the number of samples was not enough to assume significant changes in OC content with the standard deviation observed (Tables 13b).

Differences on OC content were not significant for points that changed from woodland to grassland and from grassland to woodland between surveys in the 27 member states (Tables 12b). Despite this lack of statistical significance, it has to be noted that the percent change was greater than ±2.4 % in both cases (Table 12b). These changes in OC between surveys can be explained by differences on the vegetation type, its biomass production and incorporation to the soils over the years due to LC changes.

Table 12. Results of paired Student t-test at a confidence level of 0.95 for changes in organic carbon (OC) content in relation to land cover (LC) between 2009 and 2015 surveys in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

(a) All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes from 2009 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value
Cropland to cropland *	6757	-0.65	12.52	-4.30	<0.0001
Woodland to woodland *	4244	-0.45	87.65	-0.33	0.74
Grassland to grassland *	2527	2.09	29.67	3.54	0.0004
Shrubland to shurbland	191	-1.95	35.17	-0.76	0.44
Bareland to bareland	86	1.33	8.67	1.42	0.16
Cropland to grassland *	498	0.40	14.95	0.59	0.55
Cropland to bareland	316	-0.07	6.56	-0.19	0.85
Woodland to grassland *	104	-8.52	74.43	-1.17	0.24
Woodland to shrubland	86	-17.83	96.59	-1.71	0.09
Grassland to cropland *	556	-3.00	19.22	-3.68	0.0002
Grassland to woodland *	149	2.54	64.95	0.48	0.63

(b) Mineral points (<120 g kg^{-1})

LC classes from 2009 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value	Percent change (%)
Cropland to cropland	6718	-0.44	7.85	-4.56	<0.0001	-2.49
Woodland to woodland	3170	-0.43	24.28	-0.99	0.32	-1.16
Grassland to grassland	2381	1.18	13.38	4.33	<0.0001	3.74
Cropland to grassland	492	0.92	8.55	2.38	0.018	4.76
Woodland to grassland	88	-2.83	20.09	-1.32	0.19	-7.90
Grassland to cropland	546	-2.54	10.27	-5.78	<0.0001	-11.21
Grassland to woodland	124	2.16	17.01	1.41	0.16	8.53

Table 13. Results of paired Student t-test at a confidence level of 0.95 for changes in organic carbon (OC) content in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes from 2012 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	554	0.08	3.62	0.54	0.58
Woodland to woodland *	160	1.81	15.15	1.51	0.13
Grassland to grassland *	375	-0.45	20.73	-0.42	0.67
Shrubland to shurbland	13	-3.21	11.51	-1.00	0.33
Cropland to grassland	58	-0.44	5.32	-0.63	0.53
Cropland to bareland	11	-1.27	3.98	-1.06	0.31
Grassland to cropland	55	-1.13	3.85	-2.18	0.03
Grassland to woodland	11	1.36	8.98	0.50	0.62

(a) Mineral points (<120 g kg^{-1})

LC classes from 2009 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**	**Percent change (%)**
Cropland to cropland	554	0.08	3.62	0.54	0.58	0.48
Woodland to woodland	159	1.35	14.07	1.21	0.23	5.82
Grassland to grassland	372	-1.15	9.99	-2.23	0.03	-5.02
Cropland to grassland	58	-0.44	5.32	-0.63	0.53	-2.42
Grassland to cropland	55	-1.13	3.85	-2.18	0.03	-6.18

The maps of changes in OC content at NUTS 2 level in cropland, woodland and grassland showed a regional variability (Figures 15, 16, 17). Still, the Student t-test (at a confidence level of 0.95) did not reveal significant differences on OC contents in most of the NUTS 2 regions between surveys for any of the LC classes. This can be linked to the large variability of the OC data and the low number of points to assess changes in many regions.

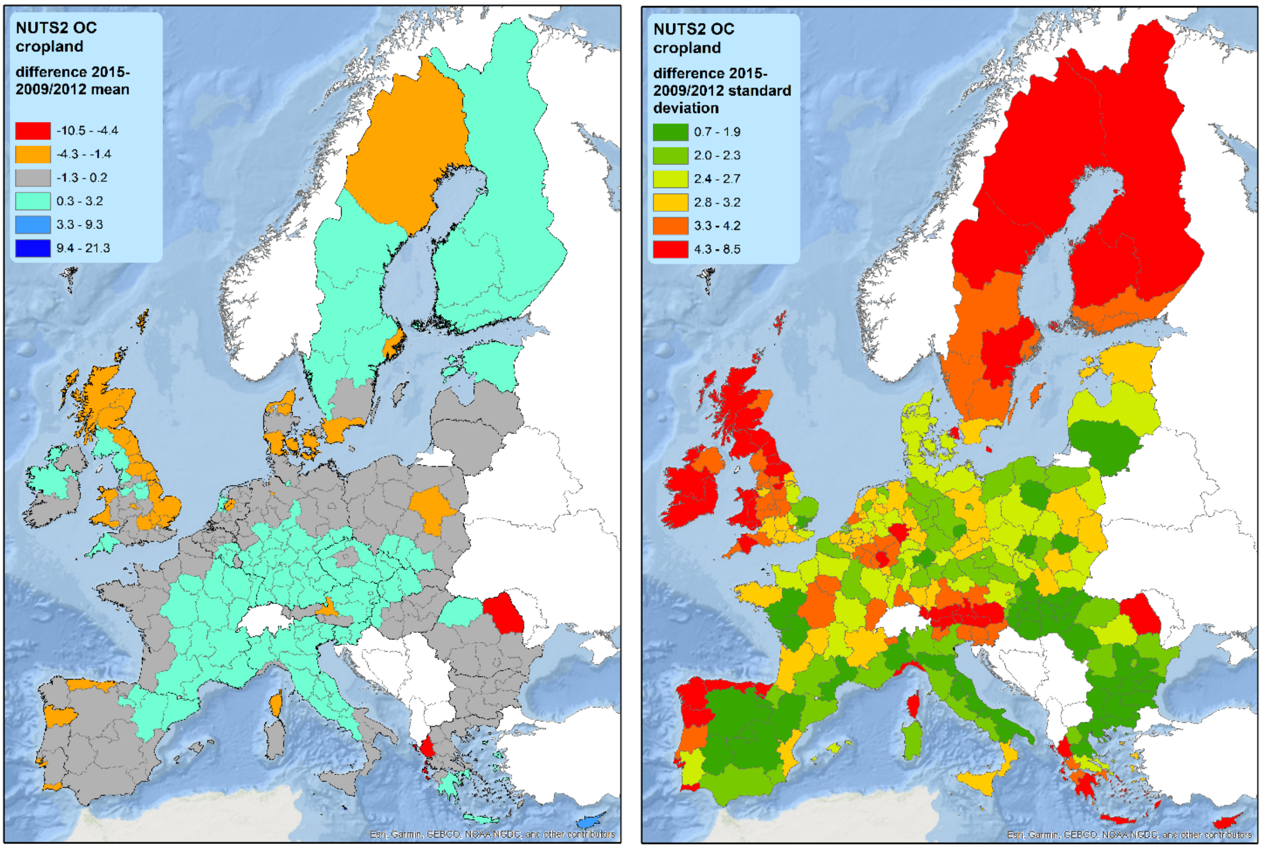
Although not significant differences, Figure 15 suggests that OC in cropland tended to increase in Mediterranean regions and central Europe, and to stay stable or decrease in many Atlantic and northern Europe regions. Organic carbon tended to stay stable or slightly decrease also in southern Italy, Greece and many regions in Bulgaria and Romania. A greater decrease of OC content occurred in regions of Ireland, United Kingdom, Denmark and Sweden, and in few regions of Portugal and Spain. It has to be noted that most of those regions have a low percentage of their surface under cropland, and thus, we have a low number of cropland points to be representative. The greatest decrease of OC content was observed in northern Romania.

Figure 16 suggests that OC content in woodland tended to increase in regions of central Europe, Romania and some northern regions, especially in Denmark, Sweden and Finland, although changes are not significant in most of the cases. Organic carbon content tended to remain more or less constant or to slightly decrease in the rest of the regions. In Ireland

and United Kingdom a greater decrease of OC content in woodland was observed. However, the results at NUTS 2 level in these two countries are not very meaningful due to the low surface woodland share of their surface and, consequently, low number of woodland points per region.

Grassland showed a greater regional variability than cropland and woodland (Figure 17). In many regions in central, southern and eastern Europe OC content tended to slightly decrease. On the contrary, OC tended to increase in Atlantic and most northern regions.

Figure 15. Changes in organic carbon (OC) content and standard deviation at NUTS 2 level in cropland



Between 2009 and 2015 surveys (27 member states) and between 2012 and 2015 in Bulgaria and Romania

Figure 16. Changes in organic carbon (OC) content and standard deviation at NUTS 2 level in woodland between 2009 and 2015 surveys and between 2012 and 2015 in Bulgaria and Romania

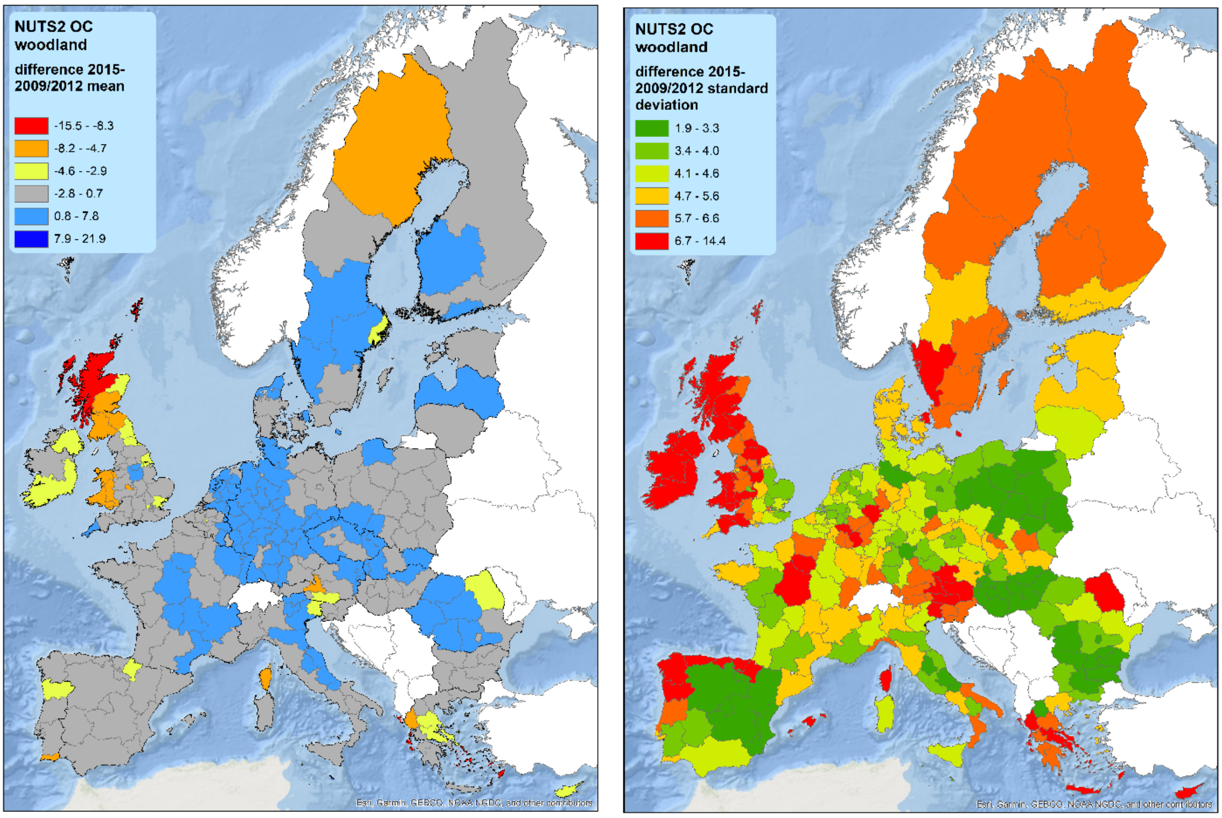
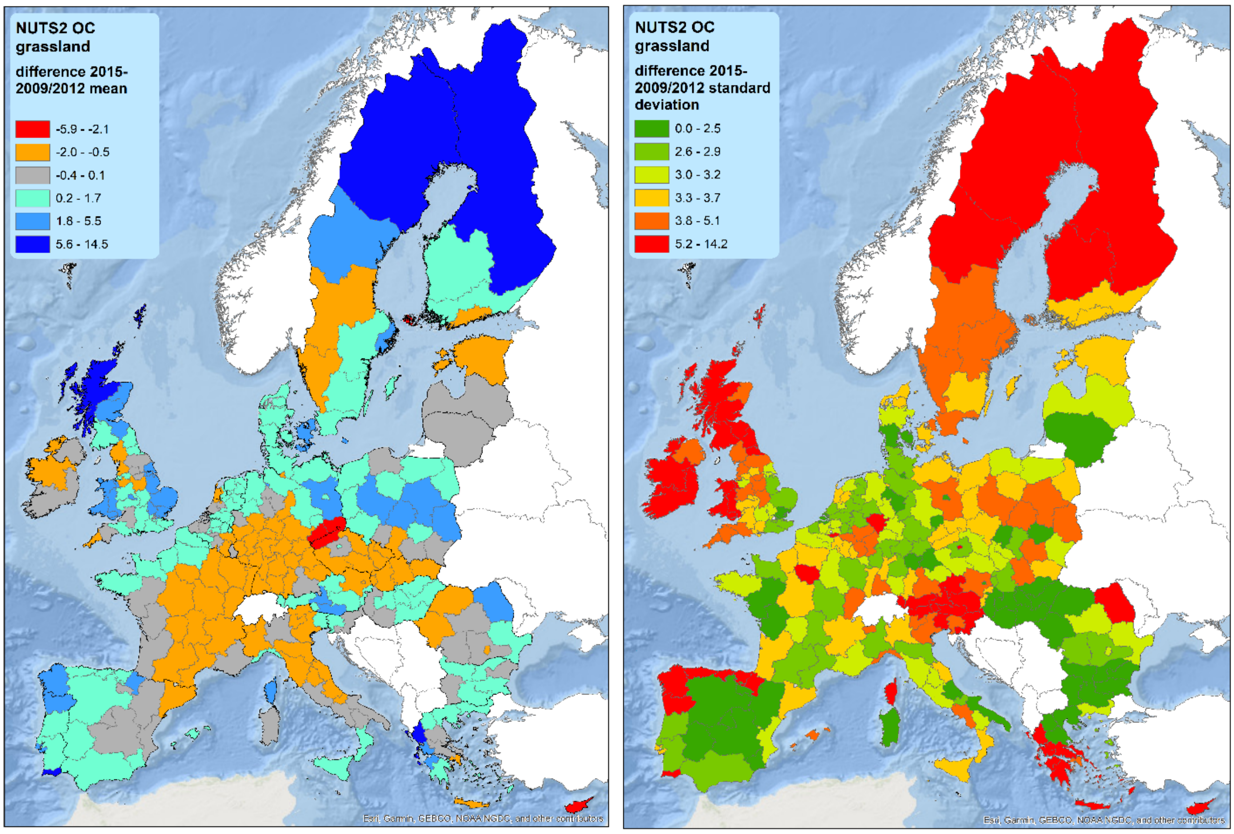


Figure 17. Changes in organic carbon content and standard deviation at NUTS 2 level in grassland between 2009 and 2015 surveys and between 2012 and 2015 in Bulgaria and Romania



In order to assess the reliability of OC changes across the two surveys (2009-2012 and 2015), a modelling exercise was undertaken using a consolidated modelling framework, which integrates LUCAS with the biogeochemistry DayCent model (Lugato, 2017 and 2018). The model was run on LUCAS points starting from the measured initial 2009 OC value (converted to stock) in points that remained in cropland in the 2009/2012 and 2015 surveys, since this has the largest number of paired samples and avoids other confounding factors (e.g. land use change, OC dilution).

The model predictions (red points) were narrowly ranging around the 1:1 line, suggesting small SOC changes in a short-time period (Figure 18). Even so, the preliminary results showed that these changes are not statistically significant in many cases. Although the simulations have different assumptions and source of uncertainty, the model is consistent in a mass balance and is, therefore, a useful indication of the expected changes. The preliminary results of the model are in line with the results of this report (Figure 15).

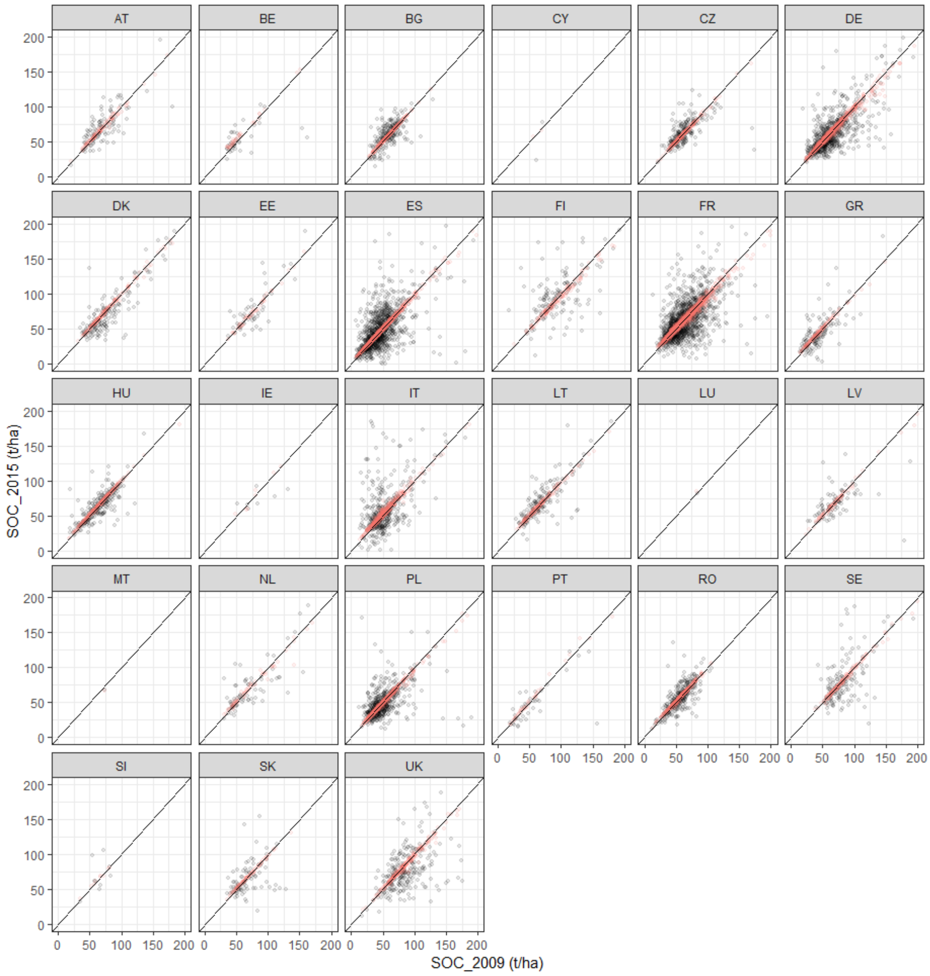
Both the results of OC changes in this report and the model predictions suggest that the LUCAS variability is likely overcoming small OC changes in the short-term in points that remain in cropland over time. As described also in the evaluation of LUCAS sampling methodology (chapter 3), a revision of the sampling protocol is necessary, especially for the assessment of LUCAS changes. Some measures have already be taken in the LUCAS 2018 survey, such as a more accurate removal of litter layer in woodland soils and a better control of sampling depth in all LC classes.

Nitrogen content showed significant differences between surveys for points that remained in cropland, grassland and woodland both when considering all revisited points and when taken into account only mineral points in the 27 member states (Table 14). In all cases, N content increased between the two surveys. Nitrogen content significantly increased also in points where LC changed from cropland to grassland and from grassland to woodland between 2009 and 2015 surveys in the 27 member states (Tables 14a and 14b). The number of samples was enough to accept the changes in OC with a power of 90 % and a significant level of 0.05 in all cases described. In Bulgaria and Romania, OC content significantly increased in points that remained in cropland and in points that changed from cropland to grassland between 2012 and 2015 surveys when considering all revisited points (Table 15a). The number of samples was enough to assume these changes in OC with a power of 90 % and a significant level of 0.05.However, we did not have enough points to accept changes on OC content between surveys when considering only mineral points (Table 15b). In cropland, this can be due to changes of crop type and fertilization over the years. In grassland, it can also be linked to the presence of livestock and the incorporation of its excreta to the soil over the years. These factors can also explain the changes in N content in points where LC changed from cropland to grassland. In woodland, the changes can be linked to the sampling. As explained before, large changes in OC and N contents in this LC class can be due to incorrect removal (excessive or insufficient) of litter layer while sampling.

At NUTS 2 level, the Student t-test did not show significant differences in most of the regions. A significant increase of N content was observed in some eastern NUTS 2 regions (in Austria, Bulgaria, Czech Republic, Poland, Romania and Slovakia) and northern regions (in Denmark, Finland, France, Germany, Ireland, United Kingdom and Sweden) in cropland, grassland and woodland. Most of these member states showed a growth in the consumption of N-based fertilisers between the years 2006-2016 ([10]). In southern member states, N content increased only in few NUTS 2 region of Spain in cropland and woodland. In fact, the consumption of N-based fertilisers in member states such as Spain, France, Germany, Italy and The Netherlands, with a great cropland share of their surface, decreased from 2006 to 2016 ([10]).

Figure 18. Comparison of organic carbon (OC) stocks in points that remained in cropland between 2009/2012 and 2015 surveys.

[10] https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/farming/documents/market-brief-fertilisers_june2019_en.pdf



Black and red points represent the laboratory data of LUCAS samples and model predictions, respectively

Table 14 Results of paired Student t-test at a confidence level of 0.95 for changes in nitrogen (N) content in relation to land cover (LC) between 2009 and 2015 surveys in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

(a) All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes 2009-2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value
Cropland to cropland *	6757	0.14	0.89	12.87	<0.0001
Woodland to woodland *	4244	0.55	3.70	9.71	<0.0001
Grassland to grassland *	2527	0.70	2.26	15.57	<0.0001
Shrubland to shrubland	191	0.19	2.11	1.27	0.20
Bareland to bareland	86	0.06	0.64	0.94	0.35
Cropland to grassland *	498	0.31	0.98	7.49	0.004
Cropland to bareland	316	0.09	0.50	3.05	0.002
Woodland to grassland *	104	0.0009	3.41	0.003	0.99
Woodland to shrubland	86	0.06	3.27	0.18	0.86
Grassland to cropland *	556	0.08	1.23	1.47	0.14
Grassland to woodland *	149	0.83	2.85	3.53	0.0005

(b) Mineral points (<120 g kg^{-1})

LC classes from 2009 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value
Cropland to cropland	6718	0.14	0.64	18.59	<0.0001
Woodland to woodland	3170	0.33	1.29	14.54	<0.0001
Grassland to grassland	2381	0.55	1.21	22.59	<0.0001
Cropland to grassland	492	0.32	0.78	9.09	<0.0001
Woodland to grassland	88	0.24	1.18	1.95	0.05
Grassland to cropland	546	0.08	0.82	2.24	0.02
Grassland to woodland	124	0.40	1.18	3.85	0.0002

Table 15. Results of paired Student t-test at a confidence level of 0.95 for changes in N content in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

(a) All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes from 2012 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	614	0.29	0.35	19.99	<0.0001
Woodland to woodland *	180	0.23	1.16	2.68	0.01
Grassland to grassland *	428	0.19	1.72	2.31	0.02
Shrubland to shurbland	13	0.08	0.76	0.36	0.72
Cropland to grassland	68	0.18	0.41	3.53	0.0007
Cropland to bareland	11	0.17	0.40	1.41	0.19
Grassland to cropland	64	0.20	0.42	3.90	0.0002
Grassland to woodland	11	0.24	0.98	0.80	0.44

(b) Mineral points (<120 g kg^{-1})

LC classes from 2009 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Woodland to woodland	159	0.20	1.08	2.34	0.02
Grassland to grassland	372	0.09	0.84	2.02	0.04

For the C-to-N ratio, we had enough samples to accept significant changes in points that remained in cropland and grassland, and in points that changed from grassland to cropland in the 27 member states between 2009 and 2015 surveys. The C-to-N ratio significantly decreased in these cases (Tables 16a and 16b). This is the result of a decrease of OC content and an increase of N content (even if changes were not significant in most of the cases, Tables 12 to 15). In Bulgaria and Romania, the number of samples was not enough to assume changes in the C-to-N ratio between 2012 and 2015 surveys (Tables 17a and 17b).

Table 16. Results of paired Student t-test at a confidence level of 0.95 for changes in C-to-N ratio in relation to land cover (LC) between 2009 and 2015 surveys in the 27 member states.

Land cover (LC) classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

(a) All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes from 2009 to 2015	***N* revisited points**	**Difference ($g\ kg^{-1}$)**	**SD ($g\ kg^{-1}$)**	**t-value**	***p*-value**
Cropland to cropland *	6657	-1.23	5.82	-26.82	<0.0001
Woodland to woodland *	4072	-2.84	6.76	-28.58	<0.0001
Grassland to grassland *	2515	-1.52	2.93	-26.01	<0.0001
Shrubland to shurbland	184	-1.09	4.13	-3.58	0.0004
Bareland to bareland	75	0.91	5.54	1.42	0.16
Cropland to grassland *	485	-1.07	3.83	-6.14	<0.0001
Cropland to bareland	307	-0.59	4.57	-2.25	0.02
Woodland to grassland *	101	-2.94	4.04	-7.32	<0.0001
Woodland to shrubland	82	-4.52	8.37	-4.89	<0.0001
Grassland to cropland *	547	-1.65	2.94	-13.13	<0.0001
Grassland to woodland *	144	-1.67	5.02	-3.99	0.0001

(b) Mineral points (<120 $g\ kg^{-1}$)

LC classes from 2009 to 2015	***N* revisited points**	**Difference ($g\ kg^{-1}$)**	**SD ($g\ kg^{-1}$)**	**t-value**	***p*-value**
Cropland to cropland	6718	-1.15	5.83	-16.06	<0.0001
Woodland to woodland	3170	-2.65	6.21	-23.39	<0.0001
Grassland to grassland	2381	-1.51	2.83	-26.05	<0.0001
Cropland to grassland	492	-1.03	3.81	-5.91	<0.0001
Woodland to grassland	88	-3.01	3.87	-7.17	<0.0001
Grassland to cropland	546	-1.64	2.96	-12.83	<0.0001
Grassland to woodland	124	-1.03	4.57	-2.46	0.01

Table 17. Results of paired Student t-test at a confidence level of 0.95 for changes in C-to-N ratio in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

All revisited points (LC classes followed by * contain mineral and organic soil points)

LC classes from 2012 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	553	-1.99	2.57	-18.28	<0.0001
Woodland to woodland	160	-1.27	3.68	-4.35	<0.0001
Grassland to grassland	375	-1.36	2.46	-10.70	<0.0001
Shrubland to shurbland	13	-1.45	2.43	-2.15	0.05
Cropland to grassland	58	-1.59	1.74	-6.97	<0.0001
Cropland to bareland	11	-2.46	2.17	-3.76	0.004
Grassland to cropland	55	-2.02	1.97	-7.59	<0.0001
Grassland to woodland	11	-0.61	2.05	-0.98	0.35

(a) Mineral points (<120 g kg^{-1})

LC classes from 2009 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Woodland to woodland	159	-1.30	3.68	-4.45	<0.0001
Grassland to grassland	372	-1.36	2.46	-10.68	<0.0001

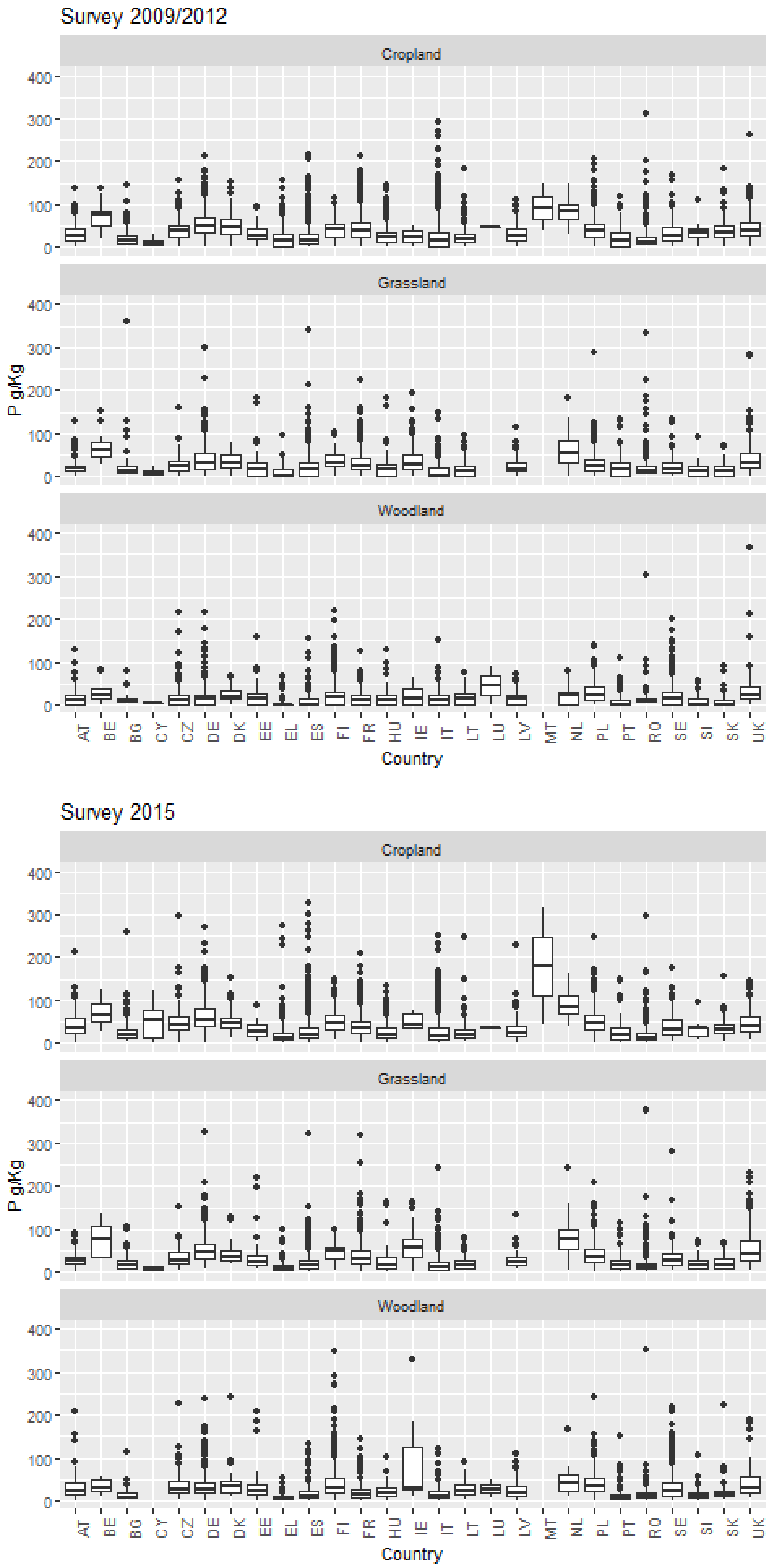
Results of changes in OC and N contents had uncertainties as the large values of standard error showed in cropland, woodland and grassland (Figures 15, 16, and 17). In most of the NUTS 2 regions in UK and in northern Sweden and Finland, the large standard error values in cropland and grassland were mainly due to the reduced number of points per region. In Sweden and Finland, the problems with the removal of litter layer during sampling contributed to the large standard errors in woodland. There are other factors that also contributed to large standard errors. For example, the efficiency of the sampling protocol (differences of sampling location between surveys, difficulties for litter removal), the fact that sampling was carried out by different surveyors or that the analyses were carried out in different laboratories and at different times in each survey. The impact of these factors on the results of OC and N contents and the rest of soil properties should be reduced as we carry out more surveys and we have more data to fine-tune comparison between surveys.

6.2 Changes in P and K between surveys

Phosphorus (P) and potassium (K) levels were higher in cropland and grassland than in woodland in all member states both in the 2009/2012 and the 2015 surveys (Figures 19 and 20). This is due to the larger nutrient supply to grassland and arable land from fertilization. Both P and K contents showed a positive linear relationship between 2009/2012 and 2015 surveys in all member states (Figures 21 and 22). This means that K and P data in 2009/2012 and 2015 moved in the same direction despite the presence of points with large differences in their contents between surveys, especially in MS with great number of woodland points such as Sweden and Finland. This can partly be linked to difficulties to accurately remove the needle-type litter of coniferous forest, a key source of K for the soil. As explained before, a better training of surveyors is needed to ensure an accurate removal of the litter layer in woodland, especially in coniferous forest. This would minimize errors in the laboratory analysis due to the sampling and would improve the comparability of the data between surveys.

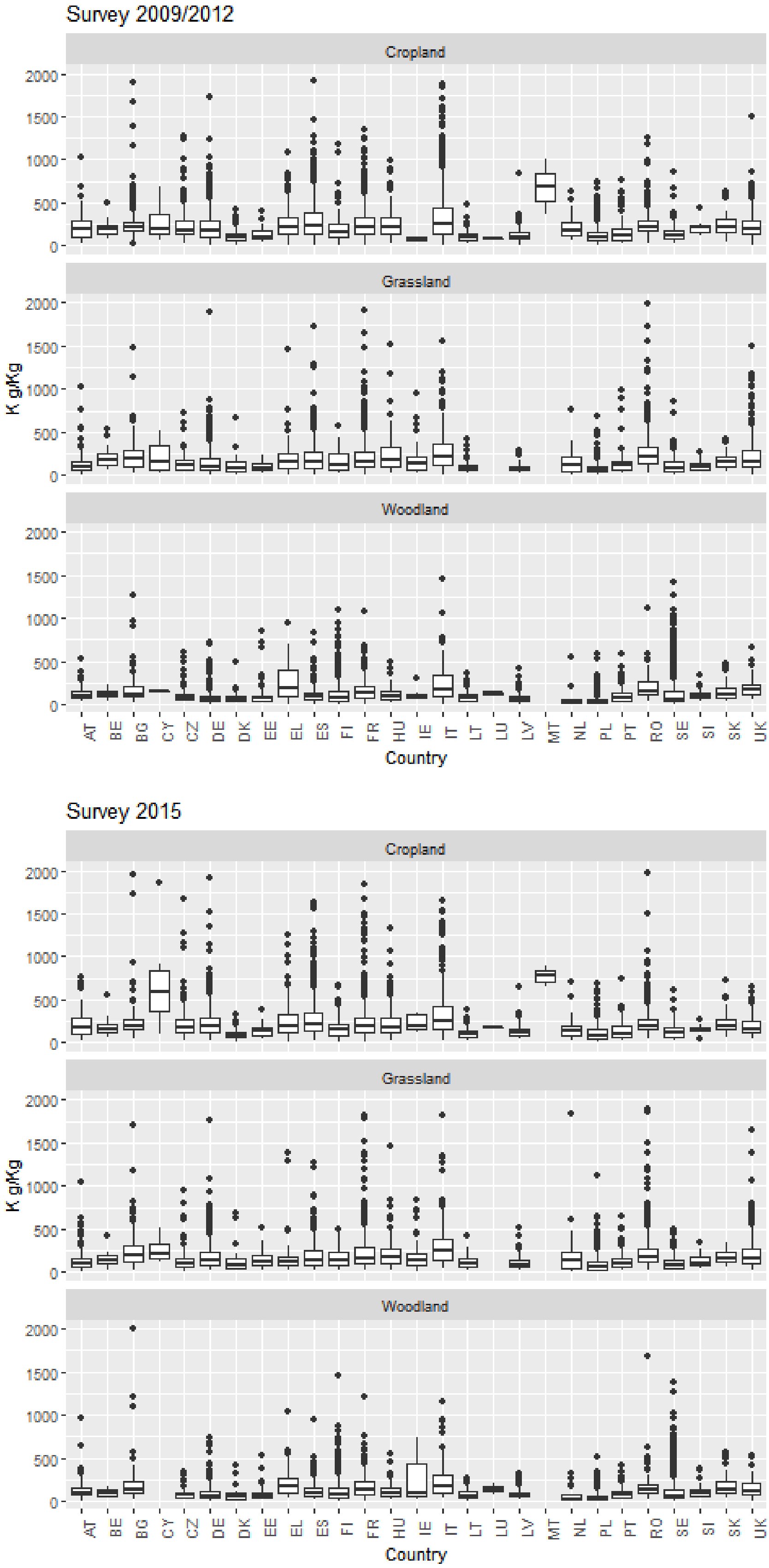
As observed for OC and N contents, changes in P and K contents between the surveys were not affected by the distance between sampling locations of revisited points in the surveys (Figures 23 and 24). Both for P and K contents, the correlation coefficient (*r*) for LLSR between the distance and the soil property was near 0, which indicated that variation in sampling locations very little explained variation in P and K contents between surveys. Only in the United Kingdom, difference on P content slightly decreased while increasing the distance between sampling locations. Thus, the distance can be excluded as a factor to explain changes in these soil properties between surveys at NUTS 2 regional scale. As a result, we did not remove paired samples taken far away (>100 m, see chapter 4 section 4.4) from each other when performing the statistical analysis to assess changes in OC and N contents by LC and NUTS 2 region between surveys.

Figure 19. Box plots of phosphorus (P) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys.



Bold horizontal lines in the boxes are the median values of each member state.

Figure 20. Box plots of potassium (K) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys.



Bold horizontal lines in the boxes are the median values of each member state.

Figure 21. Representation of phosphorus (P) content in the 2009/2012 survey against the content in the 2015 survey

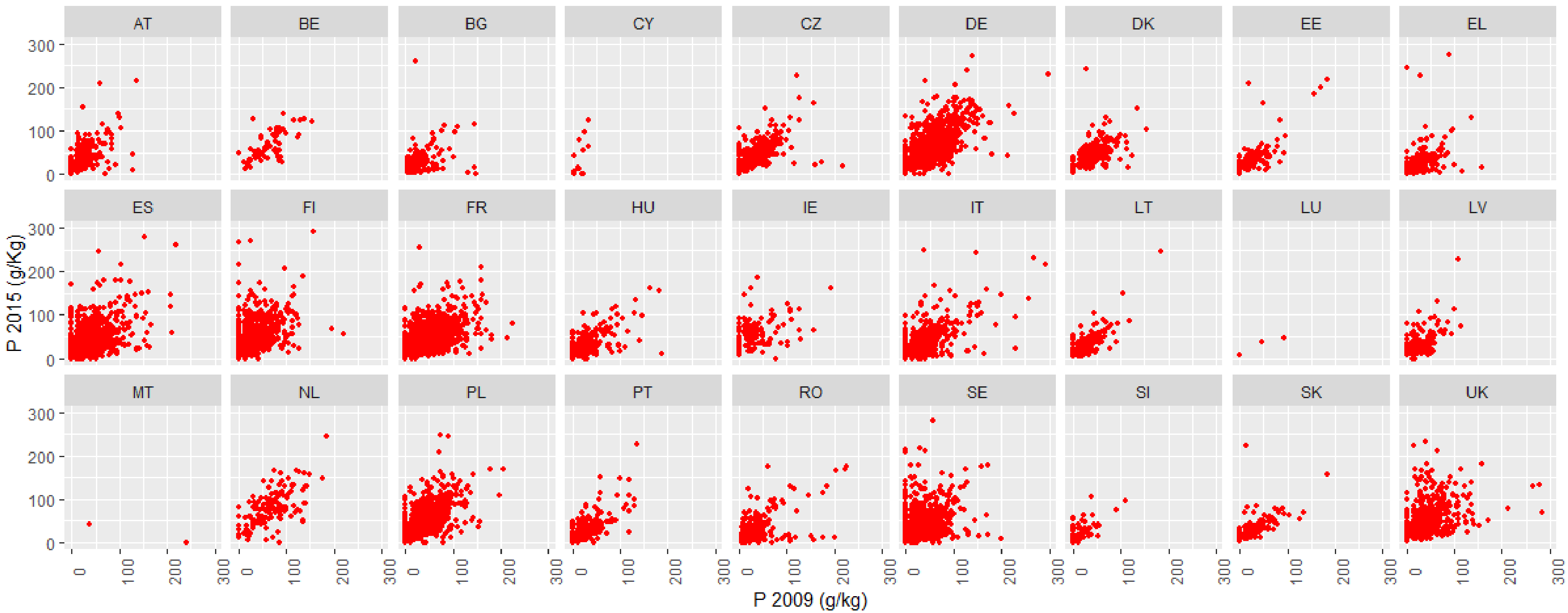


Figure 22. Representation of potassium (K) content in the 2009/2012 survey against the content in the 2015 survey

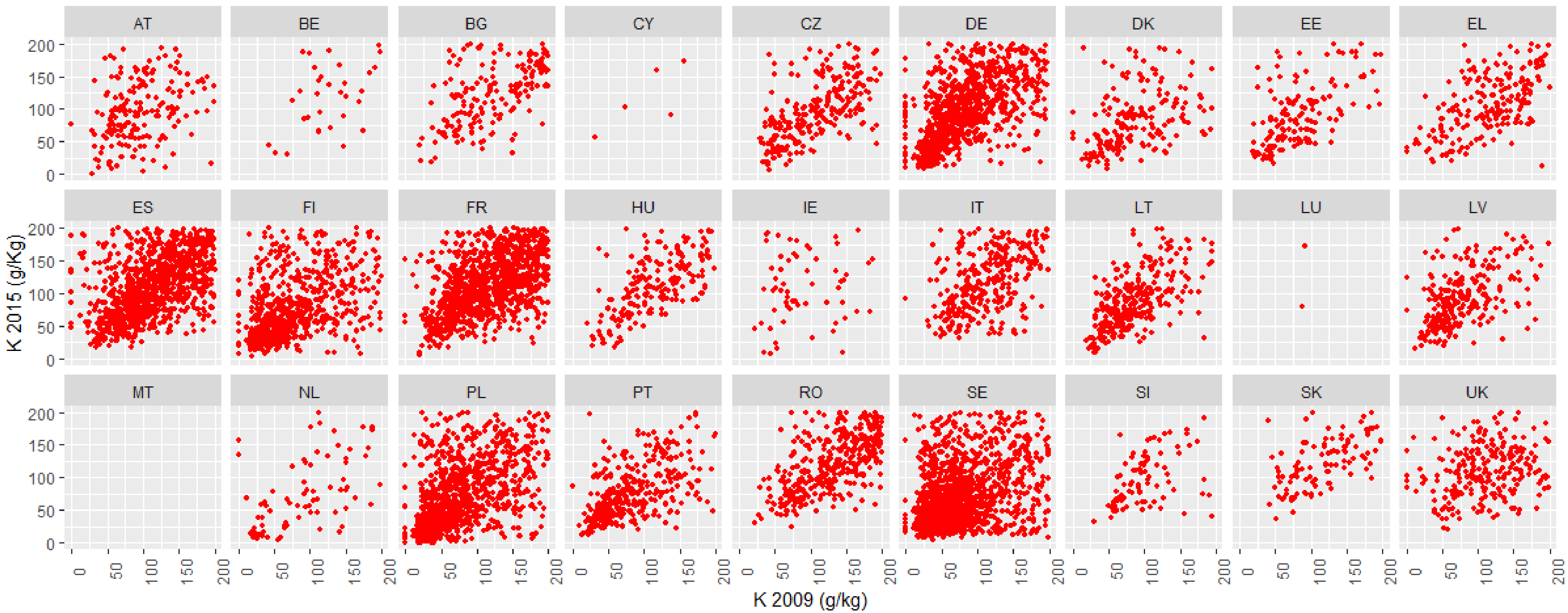
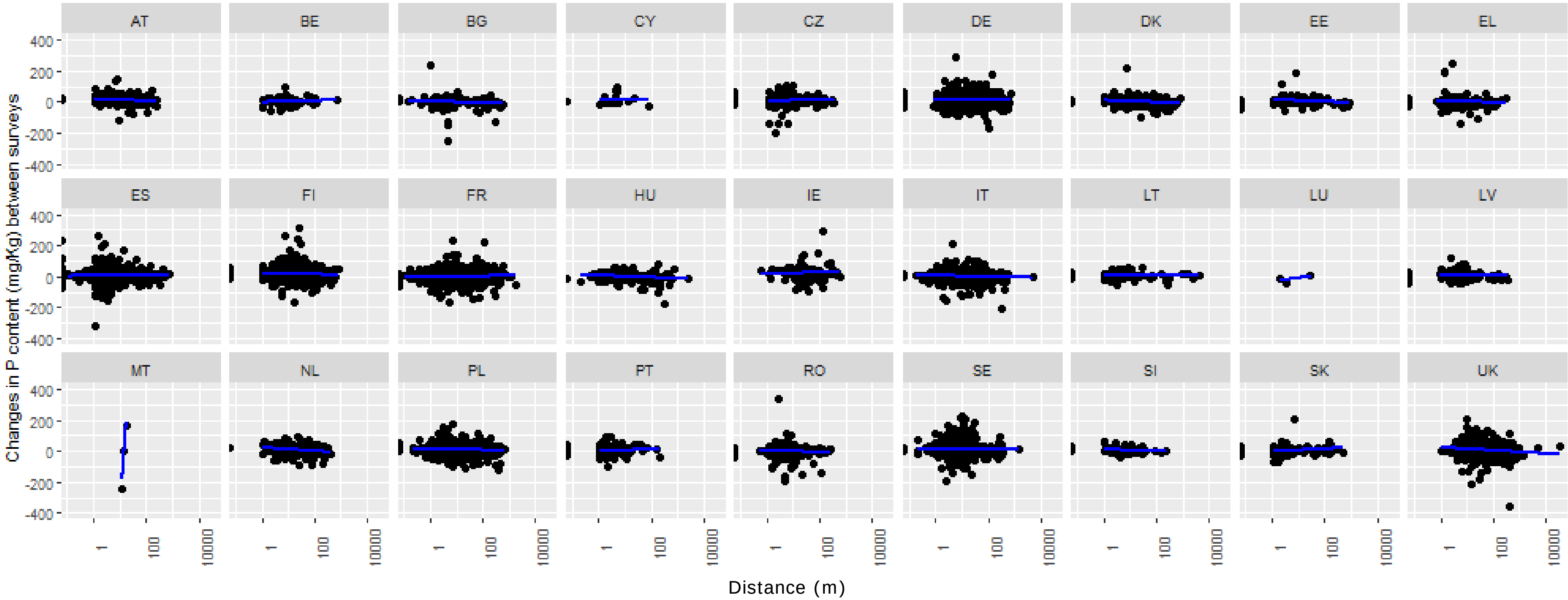
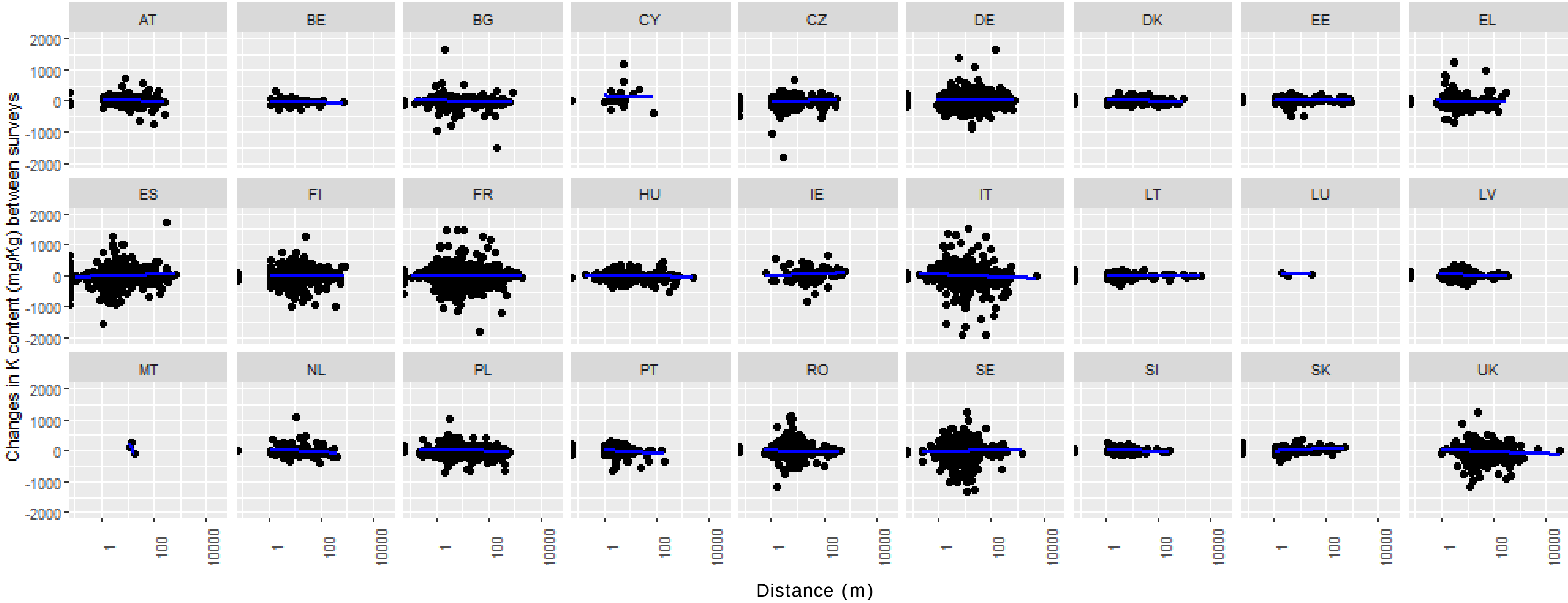


Figure 23. Changes in phosphorus (P) content by member state (NUTS 0) against the distance of sampling locations between surveys



Note that the correlation between changes in P content and distance in Malta and Luxembourg is not real due to the low number of points at each country (*N* points = 3)

Figure 24. Changes in potassium (K) content by member state (NUTS 0) against the distance of sampling locations between surveys



Note that the correlation between changes in P content and distance in Malta and Luxembourg is not real due to the low number of points at each country (*N* points = 3)

Phosphorus content significantly increased in points that remained in cropland, grassland and woodland between 2009 and 2015 surveys in the 27 member states. The number of samples was enough to accept the changes in P with a power of 90 % and a significant level of 0.05 in the three LC classes (Table 18). In points were LC changed between surveys, no tangible changes were observed in the 27 member states. In Bulgaria and Romania, no tangible changes were observed in P content between 2012 and 2015 surveys (Table 19). The changes in P content in woodland can be linked to an inconsistent removal of litter layer. In cropland and grassland, changes in fertilization over the years can explain the changes in P content. Consumption estimates of P-based fertilisers increased between 2009 and 2010, they decreased in 2011 and increased again from 2011 to 2013. Between 2013 and 2016, consumption estimates for P-based fertilisers were quite stable ([11]).

For K content, changes were significant in points that remained in cropland in the 27 member states between 2009 and 2015 surveys. The number of samples was not enough to accept the changes in K with a power of 90 % and a significant level of 0.05 in the rest of the cases (Table 20). In Bulgaria and Romania no tangible changes of K content were observed between 2012 and 2015 surveys (Table 21).

At NUTS 2 level, the Student t-test did not show significant changes on P and K contents in most of the NUTS 2 regions for cropland, grassland and woodland. Despite the lack of statistical significance, P content tended to increase in some northern and eastern regions in the three LC classes. In member states that joined the EU after 2004, such as many eastern member states, a rising trend of consumption of fertilisers has been observed between 2006 and 2016 ([11]). This coincides with increasing trend of P content observed in soil in some eastern NUTS 2 regions. Most of the significant changes (both increase and decrease) on K content were observed in woodland in NUTS 2 region of Sweden. These changes can be linked to the difficulties in removing the litter layer, especially the needle-type litter that is a key source of K (Fernández-Ugalde et al., 2019).

Table 18. Results of paired Student t-test at a confidence level of 0.95 for changes in phosphorus (P) content in relation to land cover (LC) between 2009 and 2015 surveys in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes 2009-2015	***N* revisited points**	**Difference (mg kg^{-1})**	**SD (mg kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	5977	1.65	34.86	3.66	0.0002
Woodland to woodland	2623	6.19	33.55	9.45	<0.0001
Grassland to grassland	2018	7.23	34.89	9.30	<0.0001
Shrubland to shrubland	67	-11.56	75.56	-1.25	0.21
Bareland to bareland	57	3.46	19.11	1.36	0.18
Cropland to grassland	407	1.09	29.81	0.74	0.46
Cropland to bareland	246	1.56	20.81	1.17	0.24
Woodland to grassland	70	0.64	43.65	0.12	0.90
Woodland to shrubland	46	5.63	28.03	1.36	0.18
Grassland to cropland	453	1.88	32.35	1.23	0.22
Grassland to woodland	105	5.54	35.64	1.59	0.11

[11] https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/farming/documents/market-brief-fertilisers_june2019_en.pdf

Table 19. Results of paired Student t-test at a confidence level of 0.95 for changes in phosphorus (P) content in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

LC classes from 2012 to 2015	***N* revisited points**	**Difference (mg kg^{-1})**	**SD (mg kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	452	2.09	23.83	1.87	0.06
Woodland to woodland	87	3.53	15.95	2.07	0.05
Grassland to grassland	255	-2.09	25.40	-1.31	0.19
Shrubland to shurbland	9	1.23	21.55	0.17	0.87
Cropland to grassland	58	-2.37	16.36	-1.10	0.27
Cropland to bareland	10	-5.43	18.39	-0.93	0.37
Grassland to cropland	46	0.44	14.15	0.21	0.83
Grassland to woodland	7	-0.31	3.90	-0.21	0.84

Table 20. Results of paired Student t-test at a confidence level of 0.95 for changes in potassium (K) content in relation to land cover between 2009 and 2015 surveys in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes 2009-2015	***N* revisited points**	**Difference (mg kg^{-1})**	**SD (mg kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	6757	-14.08	163.78	-7.07	<0.0001
Woodland to woodland	4244	-3.91	201.79	-0.26	0.21
Grassland to grassland	2527	4.45	203.59	1.09	0.27
Shrubland to shrubland	191	-10.26	142.79	-0.99	0.32
Bareland to bareland	86	-4.89	80.53	-0.56	0.57
Cropland to grassland	498	-2.73	179.67	-0.34	0.73
Cropland to bareland	316	-3.67	121.29	-0.54	0.59
Woodland to grassland	104	2.85	135.29	0.21	0.83
Woodland to shrubland	86	-25.74	157.29	-1.52	0.13
Grassland to cropland	556	-17.07	188.09	-2.14	0.03
Grassland to woodland	149	10.92	145.29	0.92	0.36

Table 21. Results of paired Student t-test at a confidence level of 0.95 for changes in potassium (K) content in relation to land cover between 2012 and 2015 surveys in Bulgaria and Romania.

LC classes from 2012 to 2015	***N* revisited points**	**Difference (mg kg^{-1})**	**SD (mg kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	617	-13.12	144.51	-2.26	0.02
Woodland to woodland	180	-25.94	286.54	-1.21	0.23
Grassland to grassland	428	-14.31	186.46	-1.59	0.11
Shrubland to shurbland	13	-63.96	160.71	-1.43	0.18
Cropland to grassland	69	-19.81	86.22	-1.91	0.06
Cropland to bareland	11	-107.83	211.36	-1.69	0.12
Grassland to cropland	64	1.26	132.12	0.07	0.94
Grassland to woodland	11	20.36	105.59	0.64	0.54

As with C and N, P and K data have large standard deviations at all NUTS 2 regions, which indicates that they are affected by large uncertainties. There are several factors that contribute to the large standard error values: low number of points in some northern regions, the efficiency of the sampling protocol (e.g. differences of sampling location between surveys, difficulties for litter removal), the fact that sampling was carried out by different surveyors and that the analyses were carried out in different laboratories and at different times in each survey. The impact of these factors should be reduced as we carry out more surveys and we have more data to fine-tune comparison between surveys.

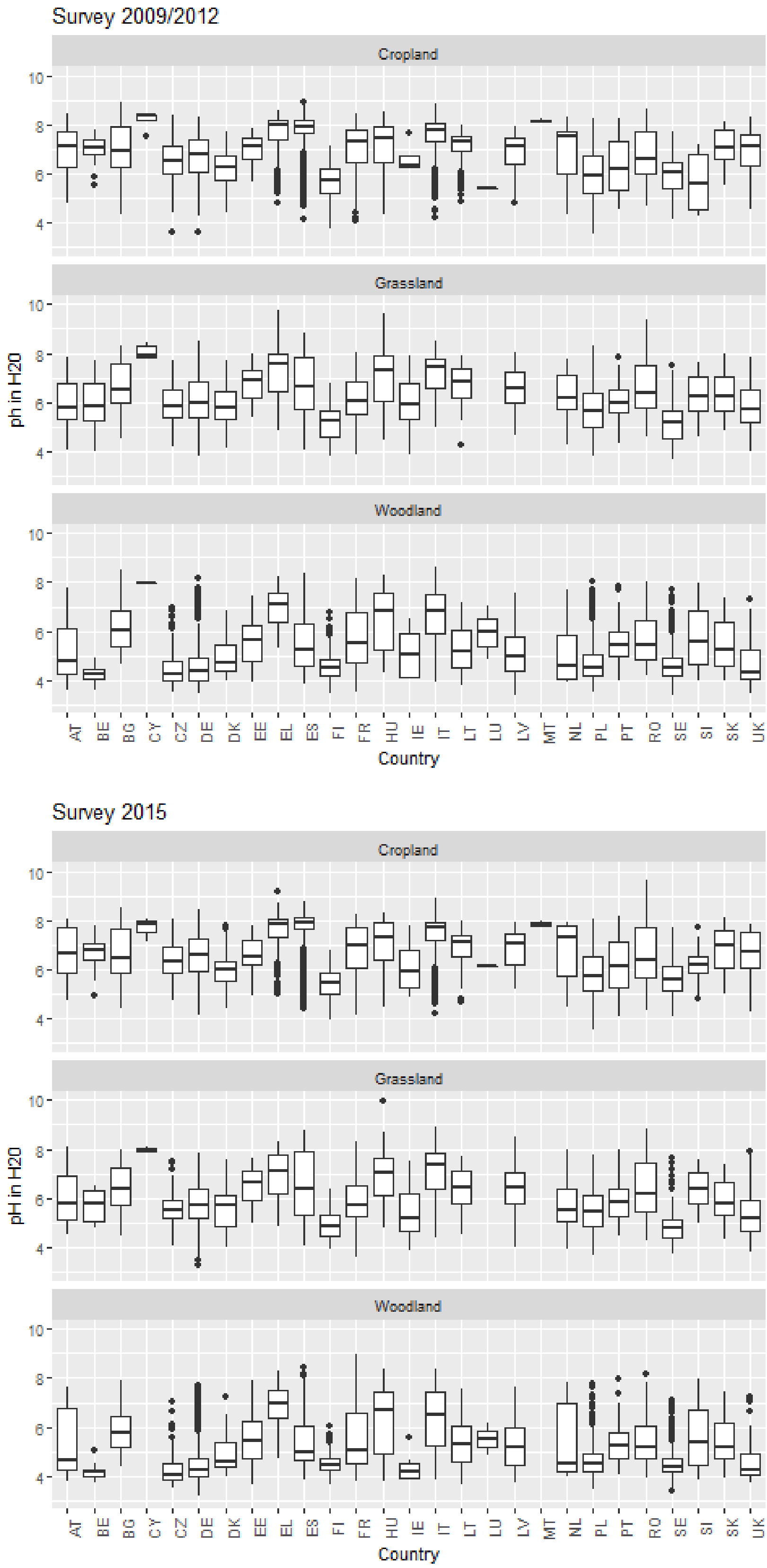
6.3 Changes in $CaCO_3$ and pH between surveys

Overall, pH in H_2O and $CaCl_2$ were lowest in woodland points and highest in cropland points both in the 2009/2012 and 2015 surveys (Figures 25 and 26). This was expected as forest soils are in general acidic and have very low content of carbonates (Figure 27). Cropland and grassland soils in southern/Mediterranean member states (Greece, Spain, Italy and Cyprus, and in a lesser extent France) had large contents of $CaCO_3$ in the 2009 and 2015 surveys. On the contrary, $CaCO_3$ content in northern/Atlantic member states was very low in both surveys (Figure 27). In agreement with $CaCO_3$ content, pH in H_2O and $CaCl_2$ in cropland and grassland in southern/Mediterranean member states was larger than in northern/Atlantic member states in all surveys.

Calcium carbonate, together with pH in H_2O and $CaCl_2$, showed a positive linear relationship between 2009/2012 and 2015 surveys in all member states (Figures 28, 29 and 30), except for $CaCO_3$ in Finland. This is due to the low content of $CaCO_3$, below the LOD, in Finish points in both the 2009 and 2015 surveys.

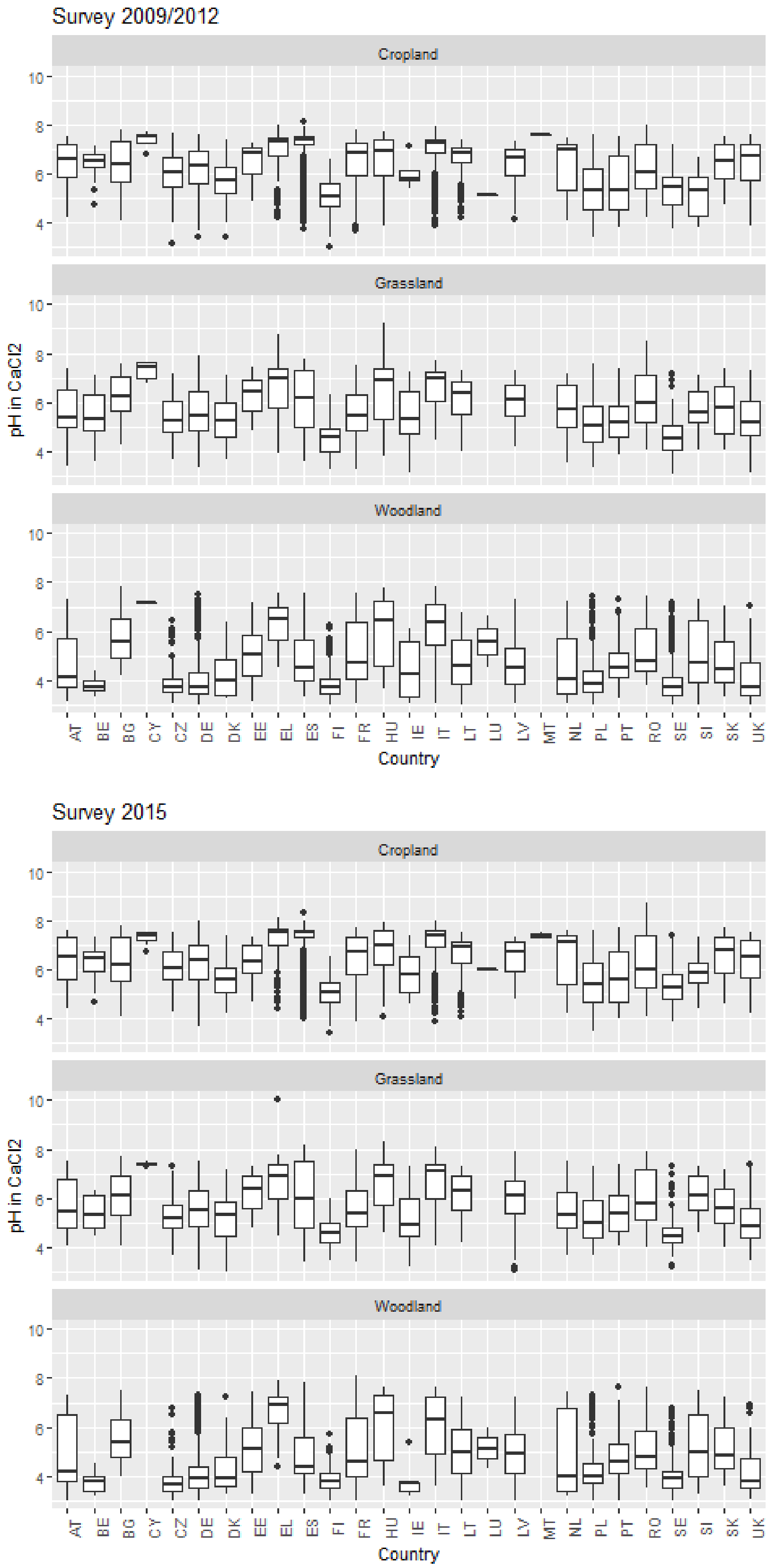
Figures 31 and 32 and 33 show that changes in $CaCO_3$ and pH were not affected by the distance between sampling locations in the 2009/2012 and 2015 surveys. Both for $CaCO_3$ and pH, the correlation coefficient (r) for LLSR between the distance and the soil property was near 0, which indicated that variation in sampling locations very little explained variation in $CaCO_3$ content and pH between surveys. Even in UK and IE, where most of the samples were taken at a distance greater than 10 m, changes in $CaCO_3$ and pH and the distance between sampling locations did not show any correlation with the distance between sampling locations in the surveys. As a result, we did not remove paired samples taken far away (>100 m, see chapter 4 section 4.4) from each other when performing the statistical analysis to assess changes in OC and N contents by LC and NUTS 2 region between surveys.

Figure 25. Box plots of pH in H_2O by land cover class and member state in LUCAS 2009/2012 and 2015 surveys.



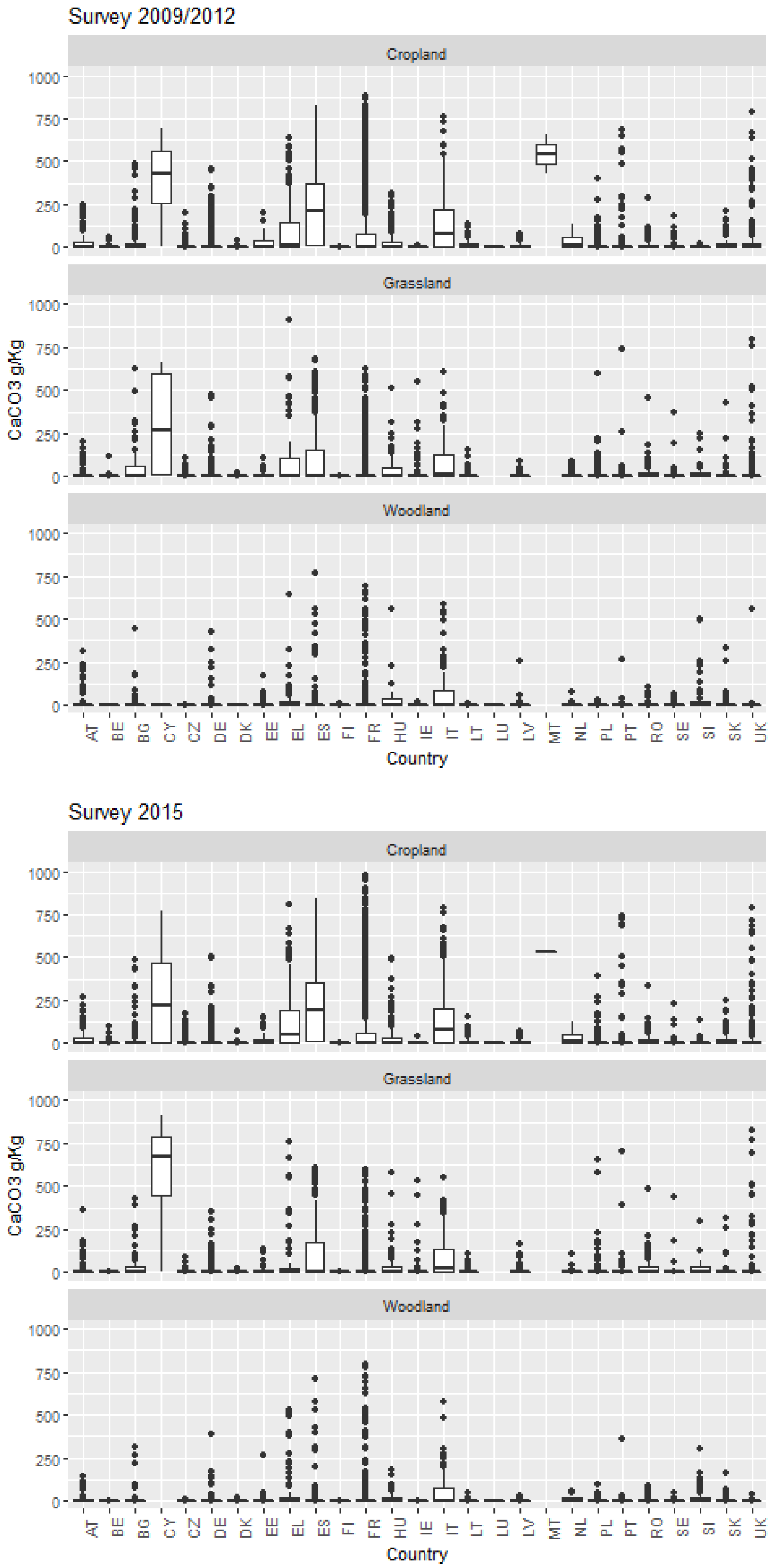
Bold horizontal lines in the boxes are the median values of each member state.

Figure 26. Box plots of pH in $CaCl_2$ by land cover class and member state in LUCAS 2009/2012 and 2015 surveys. Bold horizontal lines in the boxes are the median values of each member state



Bold horizontal lines in the boxes are the median values of each member state

Figure 27. Box plots of carbonates ($CaCO_3$) content by land cover class and member state in LUCAS 2009/2012 and 2015 surveys.



Bold horizontal lines in the boxes are the median values of each member state.

Figure 28. Representation of carbonates ($CaCO_3$) content in the 2009/2012 survey against the content in the 2015 survey

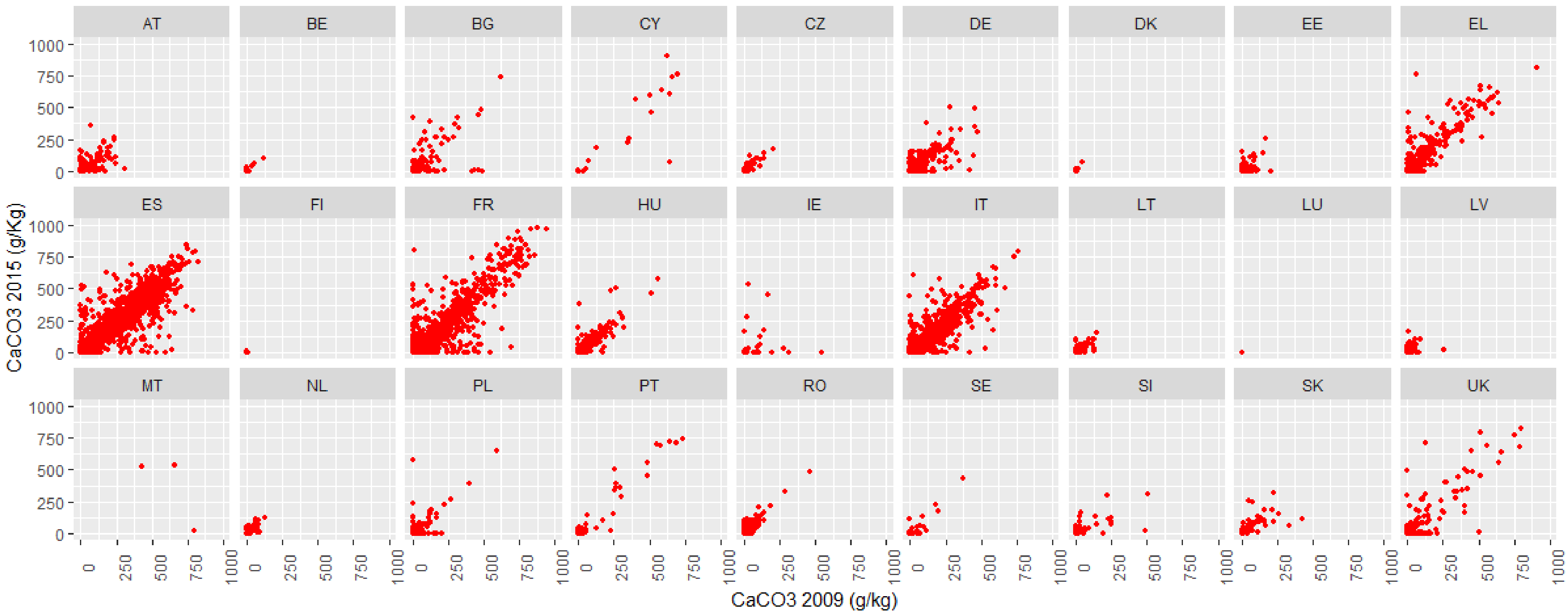


Figure 29. Representation of pH in H_2O in the 2009/2012 survey against the content in the 2015 survey

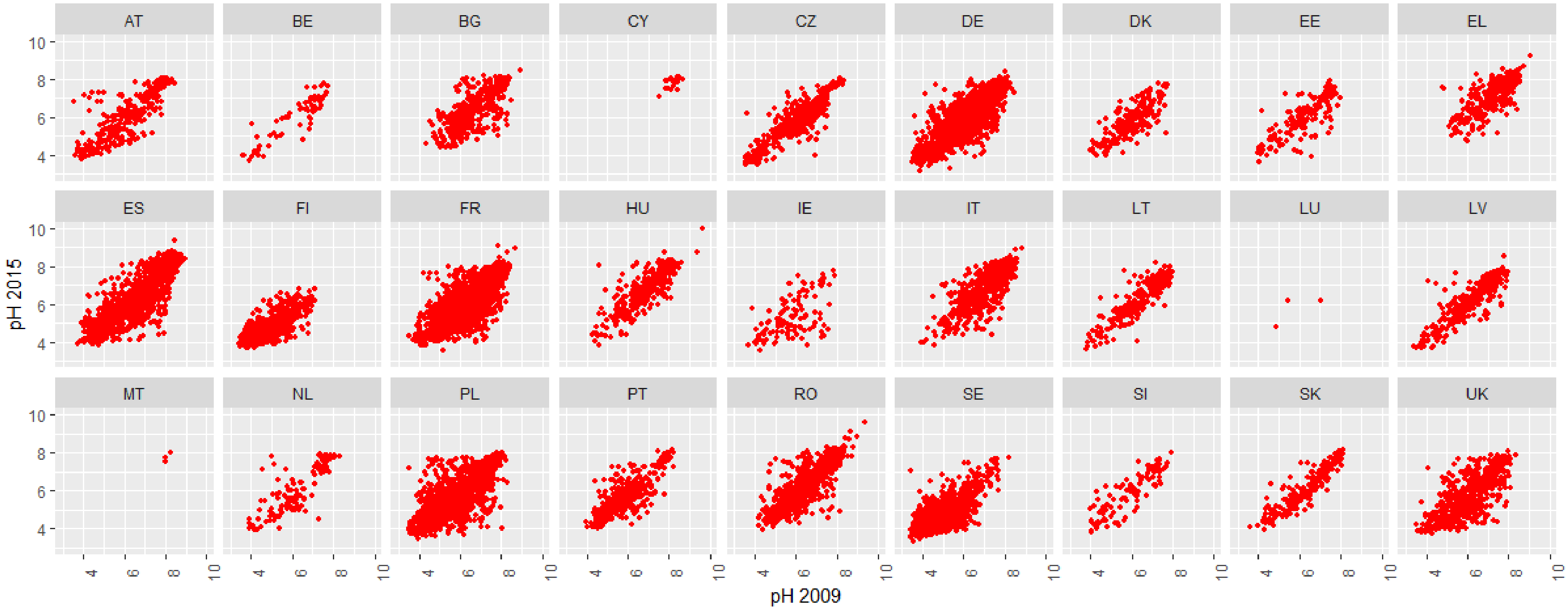


Figure 30. Representation of pH in $CaCl_2$ in the 2009/2012 survey against the content in the 2015 survey

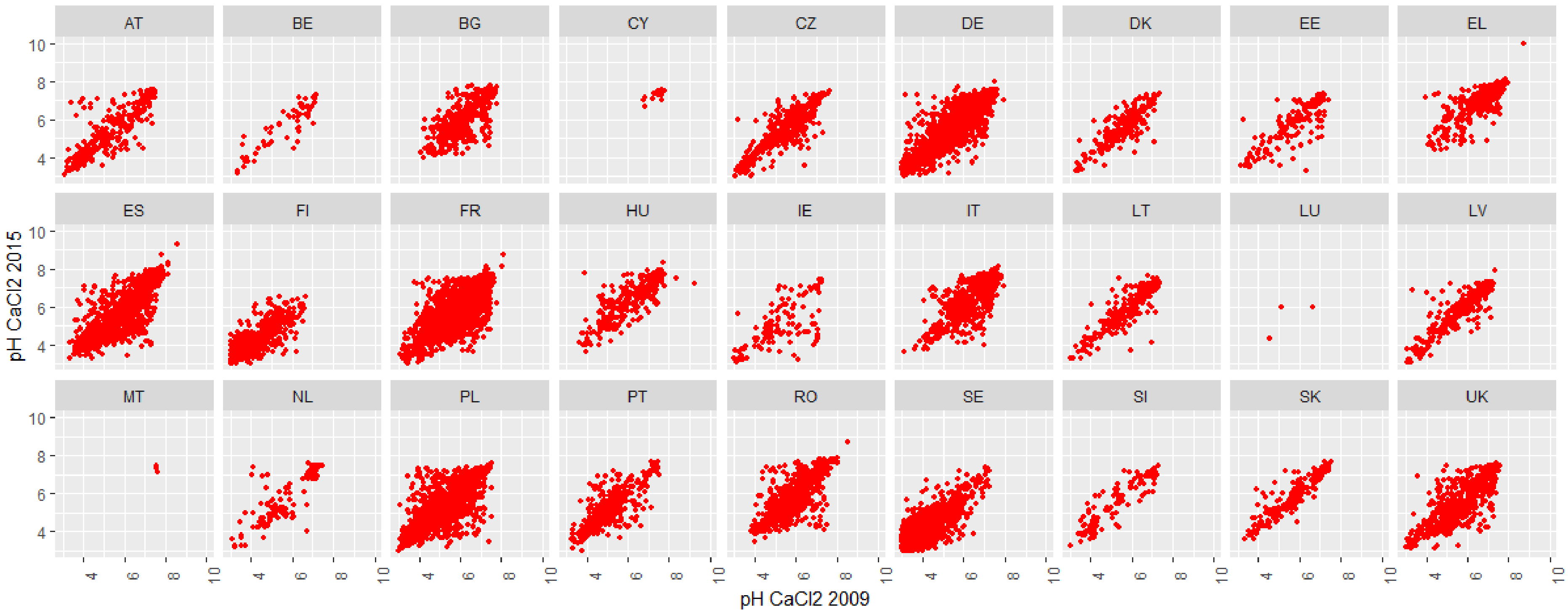
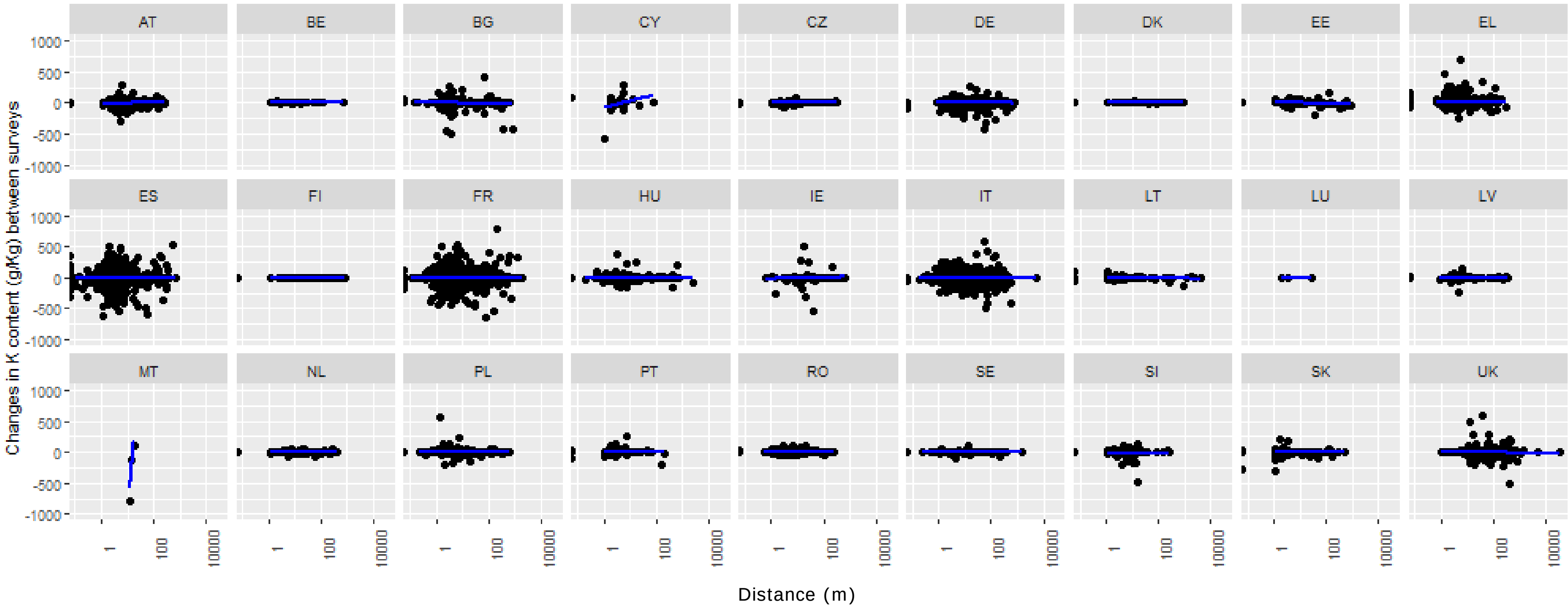
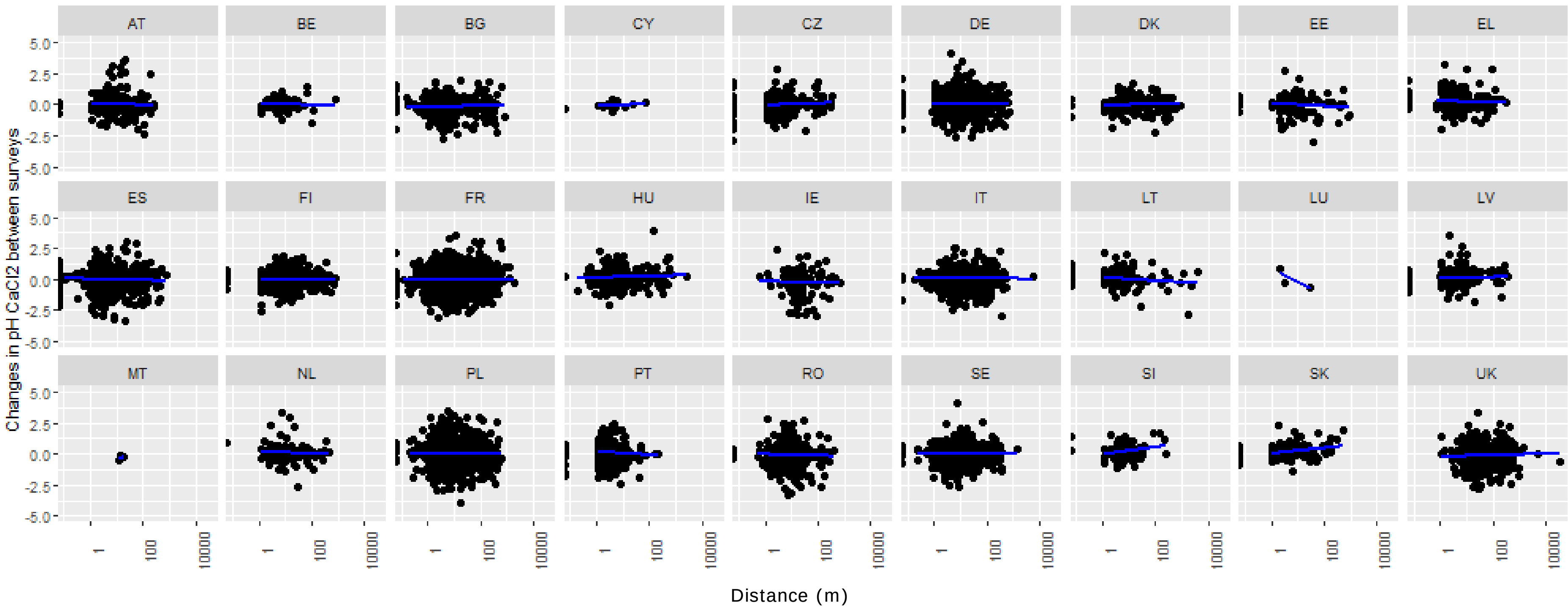


Figure 31. Changes in $CaCO_3$ content by member state (NUTS 0) against the distance of sampling locations between surveys



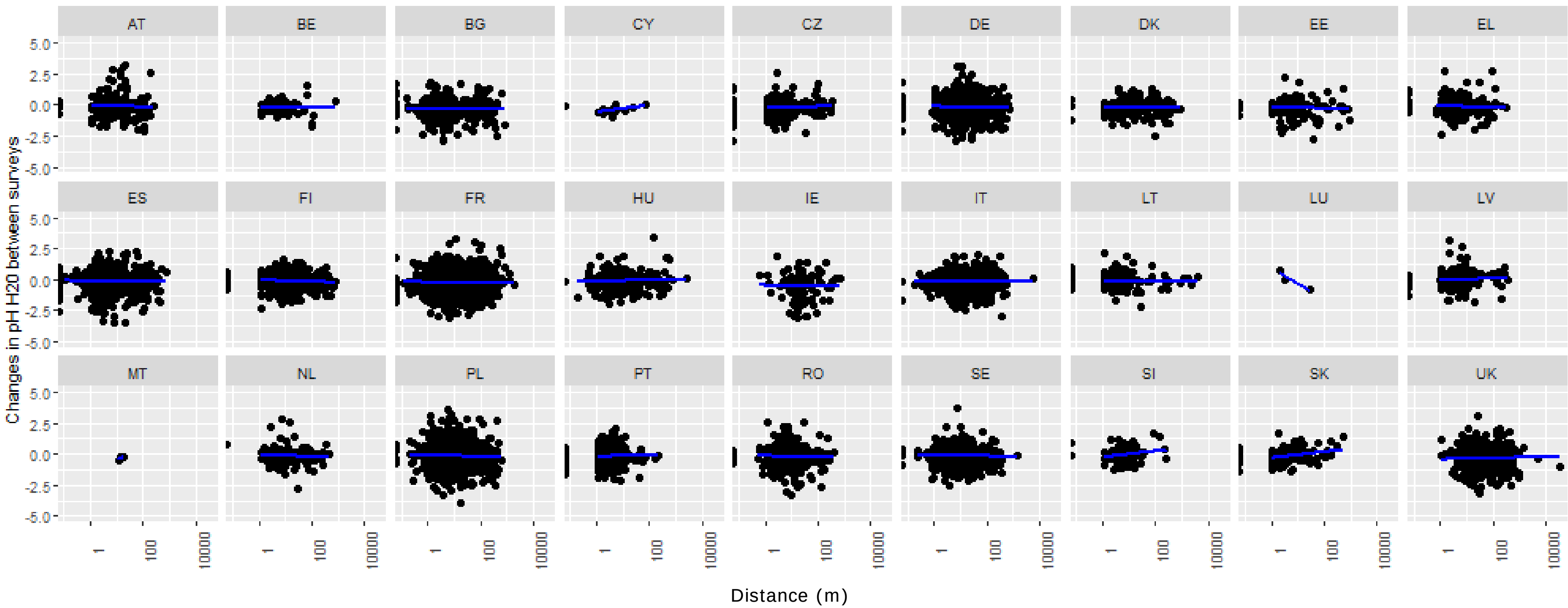
Note that the correlation between changes in P content and distance in Malta and Luxembourg is not real due to the low number of points at each country (*N* points = 3)

Figure 32. Changes in pH in $CaCl_2$ by member state (NUTS 0) against the distance of sampling locations between surveys



Note that the correlation between changes in P content and distance in Malta and Luxembourg is not real due to the low number of points at each country (*N* points = 3)

Figure 33. Changes in pH in H_2O by member state (NUTS 0) against the distance of sampling locations between surveys



Note that the correlation between changes in P content and distance in Malta and Luxembourg is not real due to the low number of points at each country (*N* points = 3)

Statistical analysis showed that pH in $CaCl_2$ significantly increased in points that remained in cropland and woodland, and in points that changed from woodland to shrubland between 2009 and 2015 surveys in the 27 member states. However, pH in $CaCl_2$ decreased in points that remained in grassland in the 27 member states (Table 22). The number of samples was enough to accept these changes with a power of 90 % and a significant level of 0.05 in each case (Table 22). In Bulgaria and Romania tangible changes of pH in $CaCl_2$ were only observed in points that remained in grassland between 2012 and 2015 surveys (Table 23).

Regarding the pH in H_2O, the number of samples was enough to assume a decrease of pH with a power of 90 % and a significant level of 0.05 in the following cases. This was observed in the 27 member states between 2009 and 2015 for points that remained in cropland, grassland, woodland and shrubland, and for points that changed from cropland to grassland and to bareland, from woodland to grassland and from grassland to cropland and to woodland (Table 24). In Bulgaria and Romania this decrease in pH in H_2O was observed in points that remained in cropland, grassland and woodland, and in points that changed from cropland to grassland (Table 25).

This greater detection of changes in pH measured in H_2O than in $CaCl_2$ solution can be linked to the methodology itself. Soil pH measured in $CaCl_2$ is less affected by soil electrolyte concentration and thus provides a more consistent measurement for soils whose salt content may fluctuate over growing seasons and over years (Minasny et al., 2011).

Overall, $CaCO_3$ content did not show significant changes between surveys neither for points that remained in the same LC class nor for points that changed LC class (Tables 26 and 27). $CaCO_3$ content changed significantly only in points that remained in woodland in the 27 member states between 2009 and 2015 (Table 26). The number of samples was enough to assume this change with a power of 90 % and a significant level of 0.05.

Table 22. Results of paired Student t-test at a confidence level of 0.95 for changes in pH-$CaCl_2$ in relation to land cover (LC) between 2009 and 2015 in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes from 2009 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	6757	0.06	0.57	8.26	<0.0001
Woodland to woodland	4244	0.06	0.62	6.20	<0.0001
Grassland to grassland	2527	-0.05	0.67	-3.66	0.0002
Shrubland to shrubland	191	-0.04	0.58	-0.96	0.34
Bareland to bareland	86	0.01	0.55	0.18	0.85
Cropland to grassland	498	-0.0005	0.57	-0.02	0.98
Cropland to bareland	316	0.02	0.55	0.77	0.44
Woodland to grassland	104	0.03	0.63	0.55	0.58
Woodland to shrubland	86	0.24	0.57	3.84	0.0002
Grassland to cropland	556	0.06	0.59	2.24	0.02
Grassland to woodland	149	-0.05	0.58	-1.02	0.31

Table 23. Results of paired Student t-test at a confidence level of 0.95 for changes in pH-$CaCl_2$ in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes from 2012 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value
Cropland to cropland	616	-0.07	0.53	-3.29	0.001
Woodland to woodland	180	-0.15	0.93	-2.12	0.03
Grassland to grassland	428	-0.11	0.59	-3.75	0.0002
Shrubland to shurbland	13	-0.31	0.94	-1.21	0.25
Cropland to grassland	69	-0.08	0.52	-1.31	0.19
Cropland to bareland	11	-0.07	0.63	-0.38	0.71
Grassland to cropland	64	-0.10	0.61	-1.38	0.17
Grassland to woodland	11	0.11	0.67	0.54	0.60

Table 24. Results of paired Student t-test at a confidence level of 0.95 for changes in pH-H_2O in relation to land cover (LC) between 2009 and 2015 surveys in the in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes from 2009 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	p-value
Cropland to cropland	6757	-0.15	0.58	-20.74	<0.0001
Woodland to woodland	4244	-0.12	0.61	-12.82	<0.0001
Grassland to grassland	2527	-0.28	0.66	-20.88	<0.0001
Shrubland to shrubland	191	-0.25	0.55	-6.15	<0.0001
Bareland to bareland	86	-0.16	0.55	-2.74	<0.0001
Cropland to grassland	498	-0.21	0.57	-8.15	<0.0001
Cropland to bareland	316	-0.17	0.55	-5.37	<0.0001
Woodland to grassland	104	-0.23	0.65	-3.63	0.0004
Woodland to shrubland	86	-0.04	0.55	-0.61	0.54
Grassland to cropland	556	-0.18	0.60	-6.99	<0.0001
Grassland to woodland	149	-0.26	0.57	-5.72	<0.0001

Table 25. Results of paired Student t-test at a confidence level of 0.95 for changes in pH-H_2O in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes from 2012 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	606	-0.22	0.56	-10.03	<0.0001
Woodland to woodland	180	-0.22	0.82	-3.59	0.0004
Grassland to grassland	428	-0.22	0.60	-7.59	<0.0001
Shrubland to shurbland	13	-0.48	0.94	-1.84	0.09
Cropland to grassland	69	-0.24	0.53	-3.74	<0.0001
Cropland to bareland	11	-0.26	0.57	-1.52	0.16
Grassland to cropland	64	-0.22	0.60	-2.97	0.004
Grassland to woodland	11	0.07	0.70	0.32	0.76

Table 26. Results of paired Student t-test at a confidence level of 0.95 for changes in carbonates ($CaCO_3$) content in relation to land cover (LC) between 2009 and 2015 surveys in the in the 27 member states.

Land cover classes highlighted in orange are those with enough number of samples to accept the change observed with a power of 90 % and a significant level of 0.05.

LC classes from 2009 to 2015	***N* revisited points**	**Difference (g kg^{-1})**	**SD (g kg^{-1})**	**t-value**	***p*-value**
Cropland to cropland	3869	-0.53	133.75	-0.25	0.80
Woodland to woodland	522	-17.28	110.03	-3.59	0.0004
Grassland to grassland	826	-5.42	95.25	-1.64	0.10
Shrubland to shrubland	108	8.89	195.95	0.47	0.64
Bareland to bareland	66	-14.62	167.06	-0.71	0.48
Cropland to grassland	227	-10.74	116.82	-1.38	0.17
Cropland to bareland	209	10.58	158.56	0.96	0.33
Woodland to grassland	19	37.26	56.36	2.88	0.01
Woodland to shrubland	10	7.2	13.68	1.66	0.13
Grassland to cropland	231	10.09	120.01	1.28	0.20
Grassland to woodland	38	-0.42	79.16	-0.03	0.97

Table 27. Results of paired Student t-test at a confidence level of 0.95 for changes in carbonates ($CaCO_3$) in relation to land cover (LC) between 2012 and 2015 surveys in Bulgaria and Romania.

LC classes from 2012 to 2015	*N* revisited points	Difference (g kg^{-1})	SD (g kg^{-1})	t-value	*p*-value
Cropland to cropland	245	6.57	65.48	1.57	0.17
Woodland to woodland	39	3.61	45.84	0.49	0.62
Grassland to grassland	142	7.99	63.94	1.49	0.14
Shrubland to shurbland	7	14.00	37.57	0.98	0.36
Cropland to grassland	15	14.67	64.92	0.87	0.39
Cropland to bareland	7	2.14	17.13	0.33	0.75
Grassland to cropland	25	8.00	31.76	1.26	0.22
Grassland to woodland	4	42.00	32.03	2.62	0.07

The Student t-test (at a confidence level of 0.95) showed a significant increase of pH in $CaCl_2$ in some NUTS 2 regions in southern EU member states, especially in arable land in Italy. For the rest of LC classes and regions, no significant changes were observed. pH measurements done in H_2O showed more changes than those in the $CaCl_2$ solution. The Student t-test revealed a significant decrease of pH in several regions all over the EU. As explained before, pH in water detects fluctuations of salt levels in the soils over growing seasons and over years. On the contrary, pH measured in a $CaCl_2$ solution will be more stable over years and growing seasons. When using a $CaCl_2$ solution, Ca^{2+} ions displace the hydronium and aluminum ions from the colloid surfaces in the soil and, as a result, measurements are less dependent on the electrolyte concentration of the soil at the moment of sampling. Regarding $CaCO_3$ content, no significant changes were observed at regional level for any of the LC classes.

It has to be noted that data of $CaCO_3$ content and pH had large standard deviations at all NUTS 2 regions and, consequently, large uncertainties. The main reasons for these uncertainties are the low number of points in some regions, especially in the case of $CaCO_3$ content, the efficiency of the sampling location between surveys, and the fact that the analyses were carried out in different laboratories and at different times in each survey. The impact of these factors should be reduced as we carry out more surveys and we have more data to fine-tune comparison between surveys.

7 Conclusions

From the study to assess the efficiency of the LUCAS soil sampling protocol, we learnt that the spade sampling in LUCAS is an efficient and cost-effective method for topsoil monitoring at large scale. However, some improvements are needed in the control of sampling depth and the accuracy of litter removal in woodland, where many soil properties (especially OC) change rapidly with depth.

We also learnt that the rigor on the labelling of soil samples in the field is another crucial aspect of the LUCAS sampling protocol to ensure the quality of the survey. The need to review the LUCAS sampling protocol is in line with the preliminary results of an OC modelling exercise performed with LUCAS data that showed that the sampling variability is likely to mask small OC changes in the short-term. Accordingly, we have reviewed the sampling protocol for the LUCAS 2018 survey and have taken measures for improving the control of sampling depth, removal of litter in woodland and labelling of samples.

When comparing sampling locations of revisited points between surveys, we observed that changes in soil properties were not significantly affected by the distance between sampling locations in the 2009/2012 and 2015 surveys. Regarding laboratory analysis, the data of the properties analysed showed a coherence from the soil point of view. Organic carbon and N showed a positive correlation, $CaCO_3$ content was minimum in samples were pH was below 7, and the sum of sand, silt and clay percentages was between 99 and 101 in the fine fraction (<2 mm) of all samples.

The assessment of changes in soil properties between the LUCAS 2009/2012 and 2015 surveys provided valuable data for supporting agriculture, environment and climate change policies. Overall, OC and N contents were highest in woodland, followed by grassland and cropland in all member states and surveys. On the contrary, P and K contents were higher in cropland and grassland than in woodland in all member states and surveys. Carbonates content was lowest in woodland from northern member states and highest in cropland from southern member states in the two surveys. In agreement with these results, pH was lowest in woodland than in cropland in the two surveys.

Soil properties showed large standard deviations (both within and between surveys) deriving from sampling. This made necessary a large number of samples for assessing significant changes over time. Unfortunately, some LC classes and LC changes between surveys were under sampled.

Importantly, most soil properties showed limited changes over the six-year period (from 2009 to 2015) in the 27 member states. Changes in Bulgaria and Romania were even less evident over the three-year period (from 2012 to 2015). This confirms that, soil properties change very slowly over time. From a policy perspective, a time lapse longer than six years is necessary to observe small variations in soil conditions, unless an extreme event results in major changes to the soil body.

Despite uncertainties arising from the sampling, it has been possible to draw some conclusions when assessing changes in soil properties between surveys in minerals points (OC <120 g kg^{-1}). Most of the soil properties showed limited changes over the six-year period (from 2009 to 2015) in the 27 member states. Changes in Bulgaria and Romania were even less evident over the three-year period (from 2012 to 2015). This reflects that soil chemical properties vary very slowly over time.

— Taking the revisited points, a statistically significant increase in OC content of 3.74 % was observed in grassland over six years. This is in line with the annual 0.4 % increase in the topsoil (30-40 cm) targeted by the ‘4 per 1000’ initiative. This would contribute to climate change mitigation. However, regional variations in trends are apparent.

— Similarly, for the revisited points in cropland, a statistically significant decrease in OC content of 2.5 % was observed while points that changed from grassland to cropland over six years decreased by 11 %. This suggests that cropland soils are not working as carbon sinks. However, regional variations in trends are apparent.

- In other land cover categories, the number of repeated points was insufficient to assess statistical significance.
- No tangible changes in OC content were observed in Bulgaria and Romania over three years (between 2012 and 2015).
- Nitrogen content increased in cropland, grassland and woodland points, and in points that changed from cropland to grassland and from grassland to woodland over six years in the in the 27 member states. In Bulgaria and Romania, a tangible increase of N content was observed in cropland points and in points that changed from cropland to grassland and vice-versa over three years.
- Phosphorus content increased in cropland, grassland and woodland points over six years. On the contrary, K content decreased in cropland points in the 27 member states. In Bulgaria and Romania, no tangible changes were observed over three years.
- pH measured in H_2O detected more changes between surveys than pH measured in $CaCl_2$ solution. This demonstrate that pH in $CaCl_2$ is a more consistent measurement and is less affected by seasonal fluctuations of electrolyte concentration in soil solution.
- pH in $CaCl_2$ increased in cropland and woodland points, and in points that changed from woodland to shrubland over six years in the 27 member states. On the contrary, pH in $CaCl_2$ decreased in grassland points. In Bulgaria and Romania, pH in $CaCl_2$ decreased in grassland points over three years.

References

Bergaya, F. and Lagaly G., *Handbook of clay science*, Elsevier Ltd, Oxford, 2006.

Bland, J.M. and Altman, D.G., 'Statistical methods for assessing agreement between two pairs of clinical measurements', *The Lancet*, Vol. 327, No 8476, 1986, pp. 307-310.

Bland, J.M. and Altman, D.G., 'Measuring agreement in method comparison studies', *Statistical Methods in Medical Research*, Vol. 8, No 2, 1999, pp. 135-160.

Brogan, J.C., Coulter, B.S. and Kelly, O.D., 'The preparation of soil samples for analysis: a comparison of wetting and drying', *Irish Journal of Agricultural Research*, Vol. 2, No 2, 1963, pp. 125-130.

FAO, *Guidelines for soil description (fourth edition)*, Rome, 2006.

Fernández-Ugalde, O., Jones, A. and Meuli, R.G., 'Comparison of sampling with a spade and gouge auger for topsoil monitoring at the continental scale', *European Journal of Soil Science*, 2019, doi:10.1111/ejss.12862.

Fernández-Ugalde, O., $$$$$$$, LUCAS 2015 Soil Survey – presentation of data and results, EUR $$$$ EN/$, Publications Office of the European Union, Luxembourg, 2020, ISSN $$$$$.

Heiderer, R., Durrant, T., Granke, O., Lambotte, M., Lorenz, M., Mignon, B., Mues, V., Forest Focus Monitoring Database System – Validation methodology, EUR 23020 EN/2, Publications Office of the European Union, Luxembourg, 2008, ISSN 1018-5593.

Hiederer, R., Data evaluation of LUCAS soil component laboratory data for soil organic carbon, Publications Office of the European Union, Luxembourg, 2018, JRC112711.

Jobbágy, E.G. and Jackson, R.B., 'The distribution of soil nutrients with depth: Global patterns and the imprint of plants', *Biogeochemistry*, Vol. 53, No 1, 2001, pp. 51–77.

Lin, L.I., 'A concordance correlation coefficient to evaluate reproducibility', *Biometrics*, Vol. 45, No 1, 1989, pp. 255–268.

Lugato, E., Paniagua, L., Jones, A., De Vries, W., Leip, A., 'Complementing the topsoil information of the Land Use/Land Cover Area Frame Survey (LUCAS) with modelled N_2O emissions', *PLoS ONE*, Vol. 12, No 4, 2017, https://doi.org/10.1371/journal.pone.0176111.

Lugato, E., Leip, A., Jones, A., 'Mitigation potential of soil carbon management overestimated by neglecting N_2O emissions', *Nature Climate Change*, Vol. 8, No 3, 2018, pp. 219-223.

Makó, A., Tóth, G., Weynants, M., Rajkai, K., Hermann, T. and Tóth, B., 'Pedotransfer functions for converting laser diffraction particle-size data to conventional values', *European Journal of Soil Science*, Vol. 68, No 5, 2017, pp. 769-782.

Mills, R.T.E., Gavazov, K.S., Spiegelberger, T., Johnson, D. and Buttler A., 'Diminished soil functions occur under simulated climate change in a sup-alpine pasture, but heterotrophic temperature sensitivity indicates microbial resilience', *Science of The Total Environment*, Vol. 473–474, 2014, pp. 465-472.

Minasny, B., McBratney, A.B., Brough, D.M., Jacquier, D., 'Models relating soil pH measurements in water and calcium chloride that incorporate electrolyte concentration', *European Journal of Soil Science*, Vol. 62, 2011, pp. 728–732.

Tóth, G., Jones, A. and Montanarella, L., *LUCAS Topsoil Survey. Methodology, data and results*, EUR 26102, Publications Office of the European Union, Luxembourg, 2013, doi:10.2788/97922.

Styles, D. and Coxon, C., 'Laboratory drying of organic-matter rich soils: Phosphorus solubility effects, influence of soil characteristics, and consequences for environmental interpretation', *Geoderma*, Vol. 136, Amsterdam, 2006, pp. 120-135.

Vancampenhout, K., Langohr, R., Slaets, J., Buurman, P., Swennen, R. and Deckers, J., ‘Paleo-pedological record of the Rocourt Pedosequence at Veldwezelt–Hezerwater (Belgian Pleistocene loess belt): Part 1 — Evolution of the parent material’, *Catena*, Vol. 107, 2013, pp. 118–129.

Vancampenhout, K., Langohr, R., Slaets, J., Buurman, P., Swennen, R. and Deckers, J., ‘Paleo-pedological record of the Rocourt pedosequence at Veldwezelt–Hezerwater (Belgian Pleistocene loess belt): Part 2 — Soil formation’, *Catena*, Vol. 110, 2013, pp. 8–23.

White, S.L., Sheffield, R.E., Washburn, S.P., King, L.D. and Green, J.T., ‘Spatial and time distribution of dairy cattle excreta in an intensive pasture system’, *Journal of Environmental Quality*, Vol. 30, No 6, 2001, pp. 2180–2187

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: http://europa.eu/contact

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications

You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).

KJ-NA-30147-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office of the European Union

doi:10.2760/5503

ISBN 978-92-76-17430-1