

LUCAS Soil, the largest expandable soil dataset for Europe: a review

A. ORGIAZZI¹, C. BALLABIO, P. PANAGOS¹, A. JONES & O. FERNÁNDEZ-UGALDE¹

European Commission, Joint Research Centre (JRC), Sustainable Resources Directorate, Land Resources Unit, Via E. Fermi 2749, I-21027 Ispra (VA) Italy

Summary

Soil is a non-renewable resource that requires constant monitoring to prevent its degradation and promote its sustainable management. The 'Land Use/Cover Area frame statistical Survey Soil' (LUCAS Soil) is an extensive and regular topsoil survey that is carried out across the European Union to derive policy-relevant statistics on the effect of land management on soil characteristics. Approximately 45 000 soil samples have been collected from two time-periods, 2009–2012 and 2015. A new sampling series will be undertaken in 2018, with new measurements included. The organization for the 2018 sampling campaign represents an opportunity to summarize past LUCAS Soil achievements and present its future objectives. In 2009–2012 and 2015, LUCAS Soil surveys targeted physicochemical properties, including pH, organic carbon, nutrient concentrations and cation exchange capacity. Data from 2009–2012 (ca. 22 000 points) and derived output (more than 20 maps) are available freely from the European Soil Data Centre website. Analyses of samples collected during 2015 are ongoing and data will be available at the end of 2017. In the 2018 LUCAS Soil sampling campaign, additional properties, including bulk density, soil biodiversity, specific measurements for organic-rich soil and soil erosion will be measured. Here we present the current dataset (LUCAS Soil 2009–2012 and 2015), its potential for reuse and future development plans (LUCAS Soil 2018 and over). LUCAS Soil represents the largest harmonized open-access dataset of topsoil properties available for the European Union at the global scale. It was developed as an expandable resource, with the possibility to add new properties and sampling locations during successive sampling campaigns. Data are available to the scientific community and decision makers, thus contributing to both research and the development of the land-focused policy agenda.

Highlights

- LUCAS Soil consists of soil physicochemical and biological properties data from Europe.
- Harmonized and open-access dataset allowing inclusion of soil in large-scale inter-disciplinary assessments.
- LUCAS Soil has a broad pool of users from scientists to policy makers, and applications from map validation to modelling.
- LUCAS Soil confirms the need for open-access and a large-scale dataset for soil properties.

Introduction

The Statistical Office of the European Union (EUROSTAT) undertakes a regular survey of land use, land cover and changes over time across the European Union (EU) (based on Decisions 1445/2000/EC by the European Parliament and 22/5/2000 of the Council, and then updated by Decision 2066/2003/EC).

Correspondence: A. Orgiazzi. E-mail: alberto.orgiazzi@ec.europa.eu
This paper was invited for the 70th Anniversary of the British Society of Soil Science and was received on 9 June 2017; revised version accepted 8 October 2017

This programme is known as the Land Use and Coverage Area Frame Survey (LUCAS) (Palmieri *et al.*, 2011). Since 2008, the legal basis for LUCAS has been provided through its inclusion in EUROSTAT's annual and multi-annual work programme and budget. From 2006, EUROSTAT has organized this survey every 3 years. Sampling is based on a regular grid that is defined by the intersection points of a 2 km × 2 km grid that covers the European territory, resulting in around 1 000 000 georeferenced sampling locations. Each point has been classified in accordance with seven land-cover classes using orthophotos or satellite images. From this overall pool, approximately 270 000 points are visited in the field by surveyors to assess the validity of the classification and

to collect additional information that cannot be assessed remotely. Points were selected through a stratification process that provides coverage of all possible types of land cover and land use identified over the whole study area. The latest LUCAS survey was undertaken in 2015 and covered the 28 EU Member States (MS). *In situ* observations were made at 273 401 points, and the data were used to produce aggregated statistical tables of land cover and land use. The LUCAS selection methodology and field survey data are freely available from the EUROSTAT website (EUROSTAT, 2017).

In 2009 the European Commission decided to include soil assessment as a key component of LUCAS. At that time, topsoil (0–20 cm) physical and chemical properties were assessed in 25 EU MS (Bulgaria and Romania were excluded and Croatia was not an MS at the time of the survey). Approximately 20 000 topsoil samples were collected in that first sampling campaign (Figure 1). The LUCAS Soil was subsequently extended to Bulgaria and Romania in 2012 (generating around 2000 additional points). All samples were analysed for the percentage of coarse fragments, particle-size distribution, pH, organic carbon, carbonates, soluble phosphorous, total nitrogen, extractable potassium, cation exchange capacity, multispectral properties and metals (Table 1). The residues of all soil samples collected in 2009 and 2012 were archived and are available for additional analyses. The main aim of the LUCAS Soil programme was to create the first harmonized and comparable dataset of topsoil properties at the EU scale. This was made possible because (i) all samples were collected following the same sampling protocol and (ii) all samples were analysed by a single laboratory using standard analytical methods (ISO).

LUCAS Soil was created from the outset as a monitoring and dynamic database, thus repetition of measurements, new locations and new properties can be added during subsequent surveys. In 2015, sampling was expanded to cover locations at altitudes above 1000 m. Furthermore, the geographical extent was extended to Albania, Bosnia-Herzegovina, Croatia, Former Yugoslav Republic of Macedonia Montenegro, Serbia and Switzerland. Approximately 26 000 locations were selected (Figure 1). Because of technical and environmental issues, however, samples were actually taken in about 22 000 locations (over 23 000 if we consider countries outside the EU; Figure 1). Electrical conductivity was added to the laboratory analysis (Table 1) (Rhoades *et al.*, 1989) because it is a good indicator of salinity, which can cause fertility problems in soil and can affect crop productivity. Naturally saline soil occurs in certain parts of Europe, but in addition salinization (which is the accumulation of soluble salts in topsoil) from the use of saline water for irrigation, inappropriate irrigation practices and poor drainage conditions is a critical issue, especially in the Mediterranean region and southeastern Europe. Analysis of soil mineralogy by X-ray diffraction (XRD) was also carried out on 400 samples selected according to their texture and land-cover class. The subset is representative for the original LUCAS dataset in terms of the rest of the measured soil properties. Chemical properties of minerals affect both nutrient availability (Guo & Gifford, 2002) and organic carbon turnover (Singh & Schulze, 2015) in soil. Both the standard and new physicochemical analyses of the 2015 samples are ongoing,

and publication of the data is expected by the end of 2017. The third LUCAS Soil survey, scheduled for 2018, is now in preparation. The total number of selected locations will be the same as for 2015 (i.e. ~26 000 points in 28 EU countries; Figure 1). The following properties will be added to the LUCAS Soil 2018 survey (Table 1), reflecting the expandable nature of its design:

- Bulk density (i.e. weight of dry soil in a given soil volume). This property takes into account both the solid part and pore spaces of soil; thus, it reflects soil compaction and is commonly used to assess this threat to soil. It is also required to estimate soil organic carbon stock (Barré *et al.*, 2014). In 2018, bulk density will be measured at 9000 points, with the possibility of adding more locations in future surveys.
- Soil biodiversity. This is a key property because of the contribution of soil biological communities to functions such as food and biomass production, a genetic pool for developing novel pharmaceuticals and climate regulation (Bardgett & van der Putten, 2014). To test the sampling protocol for soil biodiversity it will be measured at 1000 points (Figure 3), with the possibility of increasing the number of points in future surveys.
- Visual assessment of soil erosion. This collects information to improve soil erosion modelling and to improve assessments of land degradation (Pimentel *et al.*, 1995).
- Measurement of the thickness of the organic horizon in organic-rich soil. This aims to establish a baseline to assess the degradation of organic soil, and consequently the decline of organic carbon in organic-rich soil because of changes in land use and soil sealing (Hogg *et al.*, 1992). In addition, this measurement will confirm the depth criteria for the definition of Histosols, which is the classification for organic soil used by IPCC in the calculation of emissions.

LUCAS Soil is one of the world's largest and most comprehensive, harmonized continental-scale soil databases to our knowledge because of the range of properties analysed. Furthermore, it is an open-access tool, with data freely available from the European Soil Data Centre (ESDAC) (Panagos *et al.*, 2012). All LUCAS Soil-derived output (e.g. maps of soil properties for the EU) is also available from the ESDAC platform. This dataset offers a unique source of information for researchers; not only pedological and geochemical studies (Crowther *et al.*, 2016), but also biological and ecological analyses, can benefit from the LUCAS Soil database (Ettema & Wardle, 2002). LUCAS Soil-derived outputs also represent valuable tools that are used by policymakers to facilitate the decision processes related to the environmental and agricultural agenda (Keesstra *et al.*, 2016).

The ongoing analysis of samples from LUCAS Soil 2015 and the simultaneous preparation of the LUCAS Soil 2018 campaign provided an opportunity to review the outcomes following full public accessibility of the LUCAS Soil 2009–2012 data for the past 3 years and to discuss future perspectives. Here we present the overall LUCAS Soil structure, and its potential and future development as a tool for analysing soil-related processes.

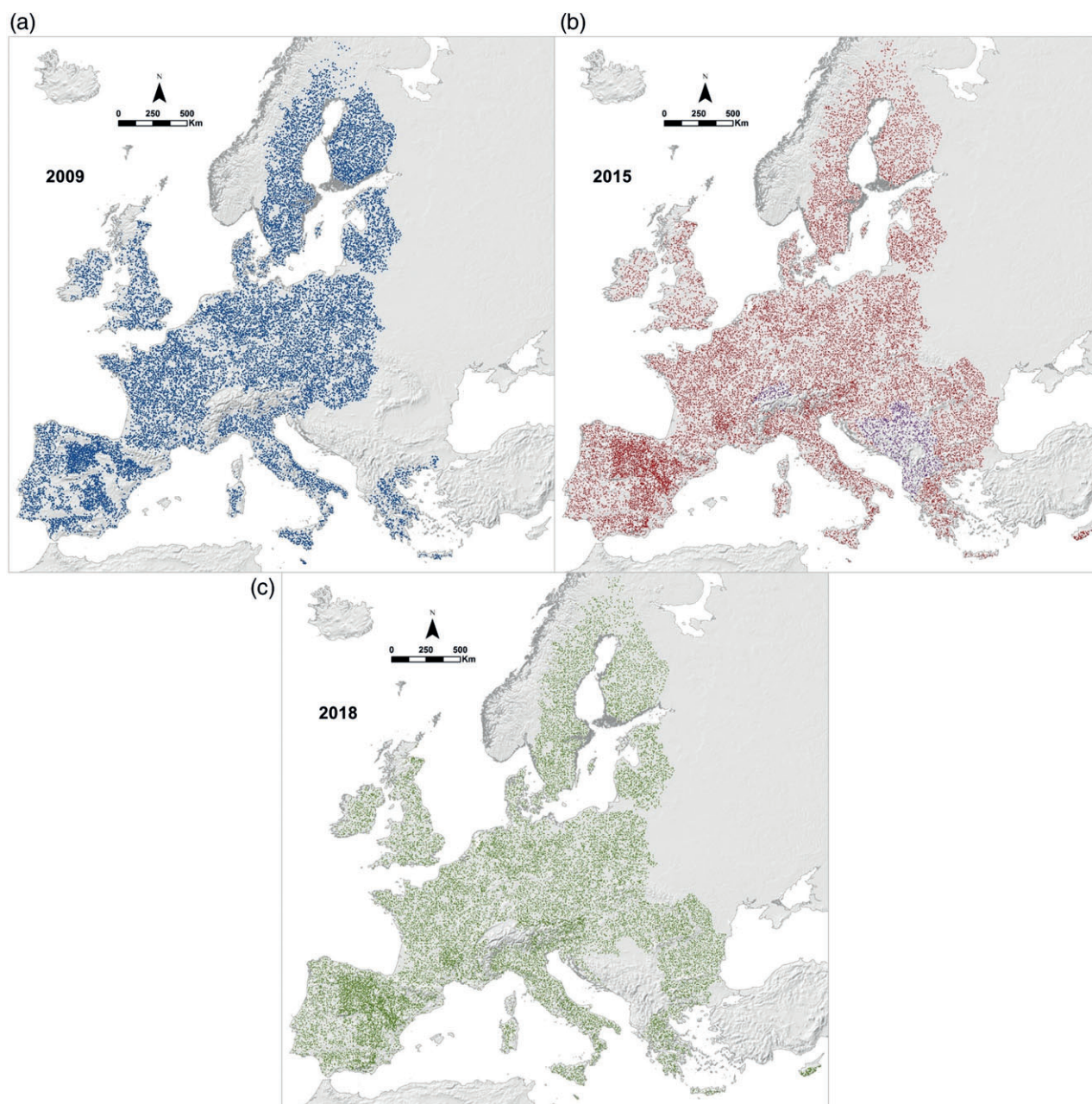


Figure 1 Points where LUCAS soil samples were taken in (a) 2009 (19 879 locations, blue) and (b) 2015 (22 325 locations, red) and will be taken in (c) 2018 (26 014 potential locations, green). In 2015, soil samples were also collected in non-EU countries (not reported on the map): Albania, Bosnia-Herzegovina, Former Yugoslav Republic of Macedonia Montenegro, Serbia and Switzerland (1122 locations, purple in b).

Material and methods

The analyses are split into modules (Table 1) because of the many different properties. First, we describe the methodology used to collect the samples for standard physicochemical analyses in both the 2009 and 2015 surveys (Module 1) and the methods established for the three new variables included in the 2018 survey, namely soil biodiversity (Module 2), bulk density (Module 3) and visual assessment of soil erosion and measurement of the

thickness of the organic horizon in organic-rich soil (Module 4, field measurements).

Point selection and soil sampling

The LUCAS Soil points are a subset of the overall LUCAS survey (Tóth *et al.*, 2013). The metadata related to land cover and use are also available. Moreover, other environmental properties recorded on the ground (for example, irrigation, signs of grazing,

Table 1 Structure of LUCAS Soil survey over the sampling years

MODULE	Type of analysis	Year of survey				
		2009–2012	2015	2018	2021	→
MODULE 1 Physico-chemical properties	Coarse fragments (>2 mm)/%					
	PSD ¹ : clay, silt, sand/%					
	pH (CaCl ₂ , H ₂ O)					
	Organic carbon/g kg ⁻¹					
	Carbonate content/g kg ⁻¹					
	Total nitrogen content/g kg ⁻¹					
	Extractable potassium content/mg kg ⁻¹					
	Phosphorous content/mg kg ⁻¹					
	Cation exchange capacity/cmole(+) kg ⁻¹					
	Electrical conductivity/mS m ⁻¹					
	Metals					
	Multispectral properties					
	Mineralogy					
MODULE 2 Soil biodiversity	Bacteria and Archaea (16S rDNA)					
	Fungi (ITS)					
	Eukaryotes (18S rDNA)					
	Microfauna (nematodes)					
	Mesofauna (arthropods)					
	Macrofauna (earthworms)					
	Metagenomics					
MODULE 3 Bulk density	Bulk density					
	Soil moisture					
MODULE 4 Field measurements	Soil erosion by water and wind					
	Thickness of organic layer in Histosols					
	Soil structure					
MODULE 5 Pollution	Organic pollutants					
	Pesticides residues					
Possibility to include new modules						

PSD¹, particle-size distribution.

Different modules form the overall structure of the survey; each module corresponds to different types of analyses. The analyses are repeated at a standard time interval, namely every 3 years. Types of analyses already established (full colour cells) for the campaign scheduled for 2018. Possible analyses for 2021 (dotted cells) are still under discussion and there will be the opportunity to implement the survey further by including new modules.

stoniness, presence of burned areas and firebreaks) are registered and a photograph of each site and soil sample is taken through a standard protocol that allows the samples to be checked in the laboratory. Land use is the main aspect of the LUCAS survey; therefore, selection of locations for soil sampling was based on this element. In addition, one of the main aims of LUCAS Soil is to provide information on the physicochemical status of the soil in EU countries. The number of points selected was proportional to both the surface area of each country and the percentage of each type of land use and cover in each country, according to the classes proposed by the CORINE Land Cover 2006 dataset (CLC, 2006). Detailed criteria for point selection are available from the ESDAC portal (ESDAC, 2017).

Standard sampling protocols were developed depending on the types of analyses that were carried out on the topsoil samples

(physicochemical bulk density, soil biodiversity or a combination of the three). The overall workflow of LUCAS Soil sampling and analyses is presented in Figure 2. A composite sampling strategy was chosen. The sample collected at each location comprises five topsoil (0–20 cm) subsamples that are mixed to form a single composite sample. The first subsample is taken at the precise GIS (geographical information system) coordinates of the pre-established LUCAS point, whereas the remaining four are taken 2 m from the central one following the cardinal directions (North, East, South and West). Vegetation residues, grass and litter, if present, are removed from the surface before sampling and from the composite sample. Approximately 500 g of soil are air-dried before being transported to the laboratory for subsequent analyses. In the 2018 survey, an additional 500 g of soil will be stored on ice and transported to a molecular biology laboratory for DNA analysis as part

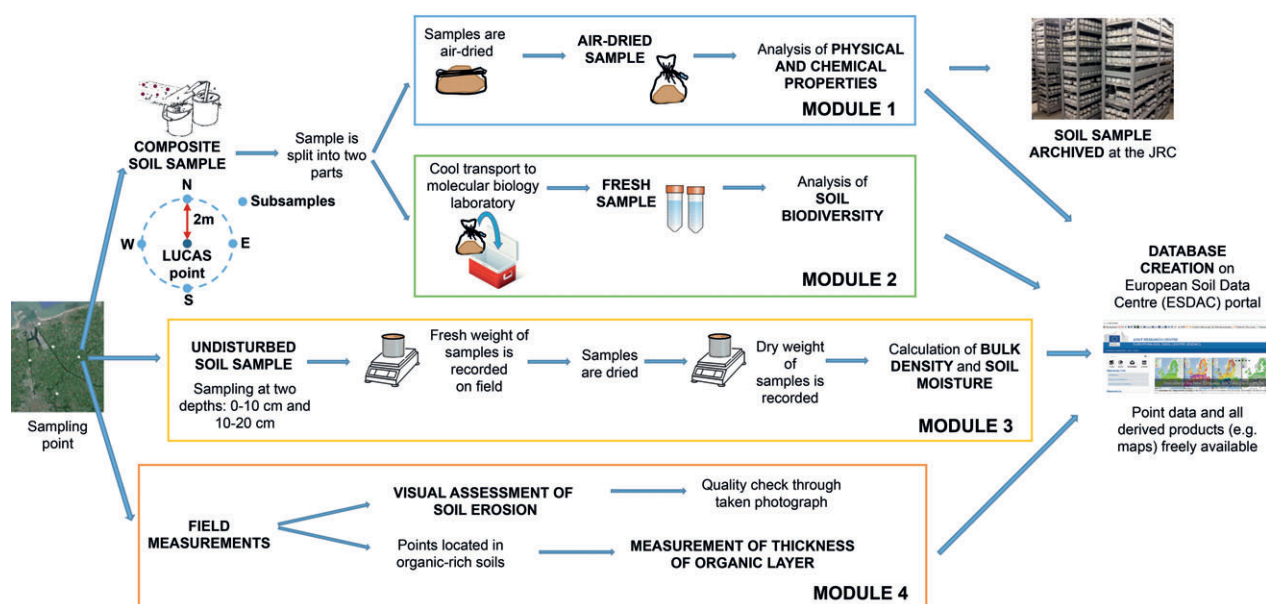


Figure 2 Schematic diagram of LUCAS Soil workflow from sampling to database generation. To organize the workflow better, different analyses were grouped into different modules. Each module can be repeated, or not, in successive sampling campaigns. For instance, in 2009–2012 and 2015 samples were collected for module 1 only (i.e. physical and chemical soil properties), whereas in 2018 modules 2, 3 and 4 will be added to analyse soil biodiversity, bulk density, soil erosion and organic-rich soil.

of the soil biodiversity assessment. For bulk density measurements, undisturbed soil samples will be taken. Land-cover and land-use data are recorded at the same time as soil sampling; they include types of vegetation cover (e.g. coniferous trees, shrubs, lichens and mosses) and land management (e.g. signs of grazing, ploughing, irrigation and presence of crop residues). These data may be used to assess and monitor the effects of land cover and use on soil properties.

Soil physicochemical properties analysis

Analysis of physical and chemical properties represents the core of the LUCAS Soil survey. Each property has its standard protocol and all samples are analysed following the same procedure. This makes the data comparable over time. Here, we briefly describe the analysis of the main soil features included in 2009–2012 and 2015 LUCAS Soil surveys.

Coarse fragments. In all samples, macroscopic roots and all particles, mineral and organic, with a diameter larger than 2 mm, are removed by dry sieving. The mineral particles not passing the 2-mm sieve are weighed separately to determine the content of coarse fragments.

Particle-size distribution. Percentages of clay, silt and sand in mineral soil material are measured by laser diffraction following the ISO (International Organization for Standardization) procedure number 13320:2009 (ISO, 2009). A laser diffraction analyser able to measure particle sizes in the 0.02 µm to 2 mm range is used. The particle-size distribution (PSD) is measured on two or three

subsamples. The third sample is measured only when the two previous samples are significantly different.

Soil pH. Soil pH is measured following the ISO 10390:1994 standard (ISO, 1994a). The method includes the determination of pH in both water and calcium chloride.

Cation exchange capacity. Cation exchange capacity (CEC) is measured following the ISO 11260:1994 protocol (ISO, 1994b). This procedure determines the concentrations of exchangeable sodium, potassium, calcium and magnesium in soil.

Organic carbon content. Organic carbon content is measured following the ISO 10694:1995 protocol (ISO, 1995b). Briefly, the total carbon content of the soil sample is determined after dry combustion with an elemental analyser. The organic carbon content is calculated from this content after correcting for carbonate content in the sample. In the absence of carbonates, total carbon content is equal to organic carbon content of the sample.

Carbonate content. Carbonate content is quantified by the ISO 10693:1994 procedure (ISO, 1995a). This standard reproduces the Scheibler method (ASI, 1999).

Phosphorus content. Phosphorus content is measured by the ISO 11263:1994 protocol (ISO, 1994c). The method calculates the phosphorus soluble in a sodium hydrogen carbonate solution.

Total nitrogen (N) content. Total N is measured in accordance with the ISO 11261:1995 protocol (ISO, 1995c). This method assesses

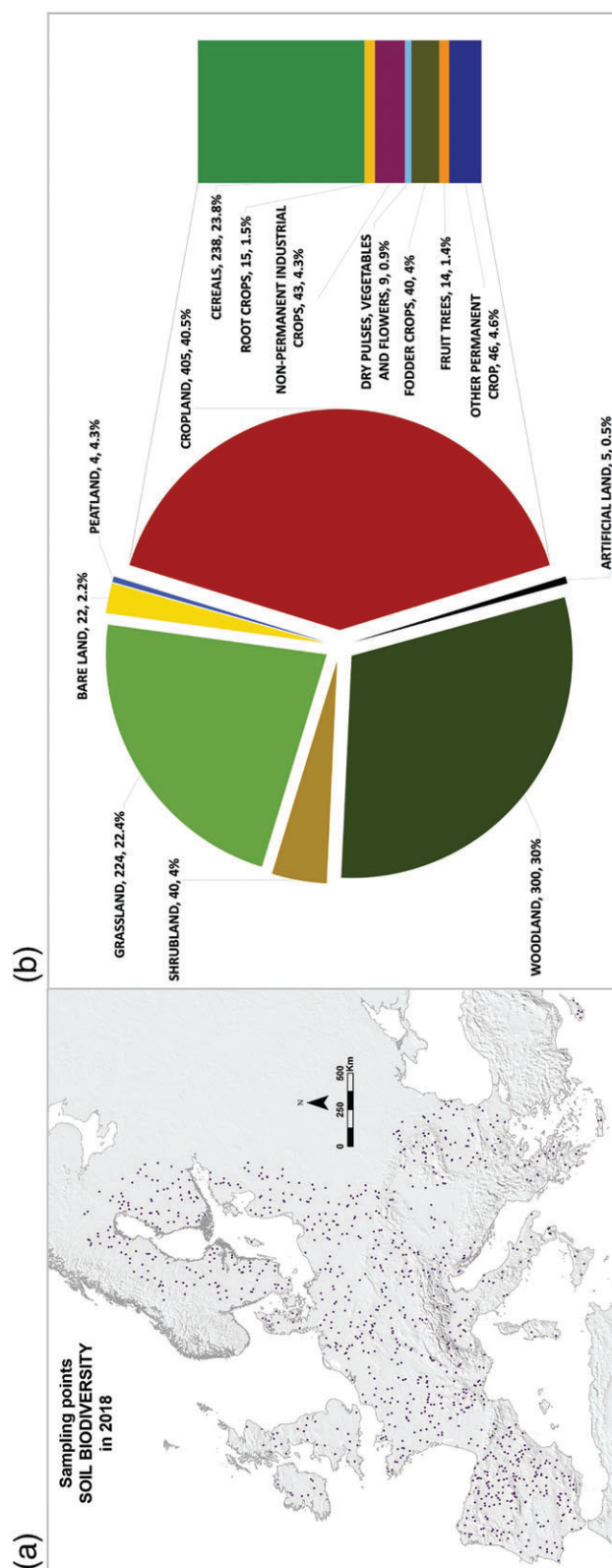


Figure 3 (a) Distribution of locations selected for sampling soil biodiversity in LUCAS Soil 2018. Points were selected by taking into account several properties, including land use and land cover, soil properties (data from LUCAS Soil 2009–2012) and climate conditions. (b) Distribution of points in relation to types of land use and cover (as classified by LUCAS). Values indicate the numbers of samples per land-use and land-cover type and the percentage of the total.

the concentrations of ammonium-N, nitrate-N, nitrite-N and organic N present.

Extractable potassium content. The extractable potassium content is determined following the Soil Survey Laboratory Manual produced by the United States Department of Agriculture Natural Resources Conservation Service (USDA NRCS, 2004).

Electrical conductivity. Electrical conductivity in an aqueous extract of soil is measured as described in the standard ISO 11265:1994 (ISO, 1994d). The measurement indicates the content of water-soluble electrolytes (salts) in a soil.

Metals. A selection of metals is measured by inductively coupled plasma–optical emission spectrometry. These include arsenic (As), cadmium (Cd), cobalt (Co), chromium (Cr), copper (Cu), iron (Fe), magnesium (Mg), manganese (Mn), nickel (Ni), lead (Pb), antimony (Sb), vanadium (V) and zinc (Zn). Detailed explanation of the methodology is reported in Tóth *et al.* (2016).

Multispectral analysis. Diffuse reflectance was determined with a spectroscope that records the full continuous reflectance spectrum from 400 to 2500 nm with a spectral resolution of 2 nm. Details on the methodology are described by Nocita *et al.* (2014).

X-ray diffraction. Mineralogical composition of the clay fraction (<2 µm) is analysed by X-ray diffractometry (XRD). The XRD patterns are obtained under the following conditions: air-drying, ethylene glycolation, heating at 110 °C, 350 °C and 550 °C, saturation with Mg and K ions, and solvation with glycerol. Interpretation and quantification of XRD patterns is carried out with the NEWMOD software.

Other than these physicochemical properties, additional variables can be assessed. For example, soil classification based on visual assessment (Munsell soil colour charts) may be carried out from photographs of a soil sample.

Other soil analyses

In 2018, the LUCAS Soil survey will include the following analyses for the first time.

Soil biodiversity analysis. The most extensive EU assessment of soil biodiversity, based on DNA metabarcoding (Orgiazzi *et al.*, 2015), will be included as part of the LUCAS Soil survey. For this, 1000 points were selected (Figure 3) from the 26 014 total soil data points. The aim of this component of the survey is to obtain data to determine the biological diversity of soil, estimate its relations with environmental features and possibly its spatial distribution. In general, living organisms occupy ecotopes with specific features in terms of chemical and physical properties, for which they are adapted; likewise, different soil biological communities will tend to occupy specific ecotopes ('pedotope') within the soil defined by soil properties, climate and vegetation cover. Thus, the sampling

scheme should aim to capture the range of values of the soil and environmental features. Therefore, conditioned Latin hypercube sampling (CLHS) (Minasny & McBratney, 2006) was chosen as a strategy to select points for soil biodiversity analysis. In CLHS, a multidimensional version of a Latin square (i.e. matrix filled with unique values occurring once in each row and once in each column) is created to sample from a multivariate distribution. This leads to an optimal configuration of samples by applying an iterative optimization process (simulated annealing). To select candidate points for the soil biodiversity survey, the full LUCAS 2009 soil database was taken as the original population from which to sample; the variables considered for the multivariate distribution of soil and environmental features were soil physical and chemical variables, topography (elevation, slope and wetness index), climate (average monthly precipitation and temperatures) and land cover. The sample obtained from the CLHS efficiently replicates (with obvious caveats related to the sample size) the range of environmental and soil variables in the original LUCAS dataset.

The decision to insert assessment of soil biota in future LUCAS surveys resulted from the increasingly recognized importance of soil organisms in the provision of several soil-related ecosystem services, from nutrient cycle regulation to soil erosion control (Barrios, 2007). The biodiversity module of LUCAS Soil aims to cover a broad spectrum of soil-living organisms, from microorganisms to macrofauna. Furthermore, the methodology was established by taking into account other projects with similar research objectives, such as continental assessment of soil biodiversity distribution. Studies in Australia (Bissett *et al.*, 2016) and Africa (Wild, 2016) are currently underway to investigate soil biodiversity with the use of DNA tools. Because of the use of similar protocols, the results obtained may be comparable and could be used to contribute to a global assessment of the distribution of soil biodiversity (Gilbert *et al.*, 2014). In particular, the methodology chosen from DNA extraction to primer set selection and sequencing platform choice will be that proposed by the Earth Microbiome Project (EMP, 2017) and will be applied to the following groups of organisms:

- Bacteria and archaea – target region 16S ribosomal DNA (rDNA).
- Fungi – target region internal transcribed spacer (ITS).
- Eukaryotes other than fungi – target region 18S ribosomal DNA (rDNA).

LUCAS Soil aims to cover as much as possible of the entire soil food web; therefore, DNA metabarcoding will possibly be also applied to other soil organisms:

- Nematodes: primer sets proposed by Sapkota and Nicolaisen (Sapkota & Nicolaisen, 2015) for high-throughput sequencing of nematode communities will be tested.
- Arthropod mesofauna: primer sets targeting the *cox1* gene proposed by Arribas *et al.* (2016) will be used.

- Earthworms: primer pairs developed by Bienert and colleagues (Bienert *et al.*, 2012) and designed on mitochondrial DNA will be used.

The methodology used to analyse the first three target groups of soil organisms is well established (EMP, 2017). Conversely, the most suitable way to explore other groups of soil organisms might be problematic because of technical limits (e.g. the need to extract DNA from a large amount of soil to enhance species detectability). Therefore, here we would like to open a debate on possible strategies for large-scale DNA fingerprinting of soil organisms other than microorganisms and 18S eukaryotes. In this respect, we invite interested researchers to contact us directly with proposals (see corresponding author details).

Bulk density analysis. Bulk density will be measured at 9000 points (Figure 4). Points were selected from the total set based on the heterogeneity of soil texture and organic carbon content, land use and land cover, topography and soil type. A CLHS approach was used to select candidate points, as for the biodiversity. Bulk density data points coincided with soil biodiversity points to explore possible correlation between these properties. At each location, soil cores will be taken at both 0–10-cm and 10–20-cm depths at five locations (LUCAS point plus each of the four cardinal points) following the LUCAS sampling scheme (Figure 2). Bulk density at each depth will be obtained by combining the air-dried weight and volume of the soil cores. The overall topsoil bulk density (0–20-cm depth) will be determined as the average of the bulk density at 0–10-cm and 10–20-cm depths.

The reasons for including bulk density were twofold: first, it is required for calculating soil organic carbon stock, which is a key measure to assess the effect of soil on climate change, and second, it is a good indicator of soil compaction, which is a considerable threat to soil quality (Hamza & Anderson, 2005). Soil compaction increases bulk density and decreases the volume of soil pores. This can lead to restrictions in oxygen supply, water and nutrient availability, root growth and the activity of microorganisms. Compaction of surface soil can also lead to increased water runoff and, as a result, to soil loss.

Visual assessment of erosion. Surveyors will provide a qualitative assessment of soil erosion by indicating the type of erosion (i.e. sheet, rill, gully, mass movement, re-deposition and wind erosion), and the distance and direction from the LUCAS point, together with an estimate of the number of rills or gullies observed. Visual evaluation of the erosion will be carried out at each LUCAS soil sampling location (i.e. ~26 000 points). This assessment will be based on specific guidelines accompanied by pictures of typical examples. Moreover, the surveyor will make a visual assessment of soil conditions and management practices that can favour or hinder soil erosion. These include the direction of ploughing, presence of crop residues and existence of grass margins and stone walls. Both the qualitative assessment of signs of soil erosion and management practices will provide valuable information for improving

models of erosion risk and the effect of mitigation measures, for example those supported by the Common Agricultural Policy to reduce erosion.

Measurement of thickness of organic horizon in organic-rich soil. The thickness of the organic horizon in effectively or potentially organic-rich soil will be measured at 1470 locations. The selection of these points was based on the following criteria:

- 2009–2012 points with organic carbon content greater than 10% (effectively organic-rich soil).
- New 2015 points with LUCAS land-cover and use belonging to one of the following classes: woodland, inland marsh or peat bog (potentially organic-rich soil).

The measurement of thickness of organic horizon will be repeated over time to assess and monitor the risk of degradation of organic and organic-rich soil.

Results and discussion

Data validation, quality control and re-use potential

The LUCAS Soil dataset has a wide range of potential applications:

- Modelling of soil-related properties at the continental scale.
- Assessing the contribution of the soil component to regulation of large-scale environmental phenomena such as climate change and nutrient cycles.
- Validation of small-scale models involving soil properties.
- Comparison and checking of the quality of national and regional soil inventories.
- Assessing the ecosystem services provided by soil.
- Evaluation of the effect of soil on key economic activities such as agriculture.

Apart from raw single-point data, LUCAS Soil data offer a broad spectrum of derived materials that are accessible to users on the ESDAC portal (ESDAC, 2017). This enables LUCAS Soil to be a continuously evolving database with new material included not only when results from a new sampling campaign are obtained (i.e. every 3 years), but also when newly derived products are generated by the European Commission's Joint Research Centre. To ensure quality control, LUCAS Soil-derived maps must be published in peer-reviewed journals before being made available to the user community. To monitor the use of LUCAS Soil data, users are required to fill in a simple registration form. The number of downloads were used to measure the potential re-use of primary and derived data from LUCAS Soil (Table 2). Moreover, total LUCAS land-cover data (not to be confused with LUCAS Soil data) were used to develop additional datasets. For example, assessment of the cover management factor in the EU by Panagos *et al.* (2015a) and the effect of support practices on the reduction of soil erosion in the EU by Panagos *et al.* (2015b) were generated from land-cover and management information gathered by surveyors at all LUCAS

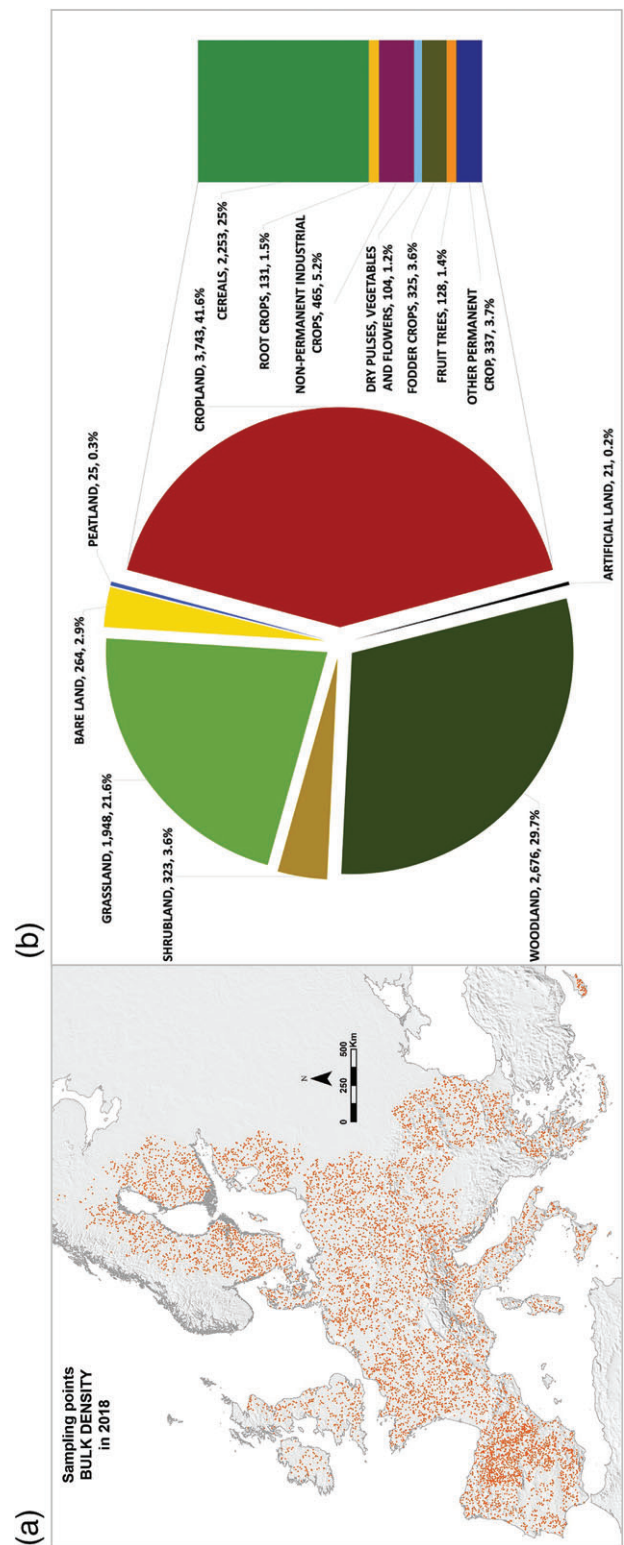


Figure 4 (a) Distribution of locations selected for sampling bulk density in LUCAS Soil 2018. (b) Distribution of points in relation to land-use and land-cover types (as classified by LUCAS). Values indicate the number of samples per land-use and land-cover type and the percentage of the total.

Table 2 List of all LUCAS Soil primary and derived data freely available at the European Soil Data Centre

Type of data	Publication	Number of downloads
Point data, Soil chemical and physical properties in EU (European Union)	Tóth <i>et al.</i> , 2013	854
Point data, Multispectral reflectance data in EU soil	Nocita <i>et al.</i> , 2014	854
Map, Soil erodibility (K-factor) in EU	Panagos <i>et al.</i> , 2014	797
Map, Soil organic carbon (SOC) stocks in EU agricultural soils	Lugato <i>et al.</i> , 2014b	249
Map, Soil water erosion in EU	Panagos <i>et al.</i> , 2015c	431
Map, SOC in EU soils / %	de Brogniez <i>et al.</i> , 2015	281
Map, Future SOC stocks based on climate change in EU	Yigini & Panagos, 2016	43
Map, Clay, sand, silt, coarse fragments, bulk density, available water capacity in EU soils	Ballabio <i>et al.</i> , 2016	584
Map, Threats to soil biodiversity in EU	Orgiazzi <i>et al.</i> , 2016	114
Map, Soil erosion in forestland in EU	Borrelli <i>et al.</i> , 2016	101
Map, Soil wind erosion in EU	Borrelli <i>et al.</i> , 2017	192
Map, N ₂ O emissions from EU agricultural soils	Lugato <i>et al.</i> , 2017	31

The number of downloads refers to the period from 10 September 2013 to 3 September 2017.

points (i.e. ~270 000; Palmieri *et al.*, 2011). All these datasets are also available online on the ESDAC (2017) portal.

Statistics on use of LUCAS Soil have been collected since 2013 from user registration information following the publication of the 2009–2012 point data. These data provide a reliable understanding about who is interested in LUCAS Soil datasets and their reasons for downloading the data (Figure 5). In that period, LUCAS SOIL data were downloaded 854 times. Although the majority of requests were from Europe, several also reflect a global interest in the LUCAS Soil dataset (Figure 5a). Moreover, institutions that benefited from LUCAS Soil datasets ranged from universities and research organizations to public administrations and the private sector (Figure 5b). This confirmed the large spectrum of applicability of the LUCAS Soil database and its re-use potential not only for scientific purposes but also for developing soil-related investments and policies.

When downloading the LUCAS Soil point dataset, users can describe their specific need for the data as part of the registration process. Consequently, it was possible to assess common trends by a word cloud analysis (Figure 5c). Results show that research purposes are the main reasons for downloading the data. Interestingly, several users were either Masters or PhD students who needed data to develop projects for their theses. Teaching and educational reasons are also among the justifications for obtaining LUCAS information. Several users stated their intention to use LUCAS Soil as a reference source to compare and validate their own datasets. This was confirmed by frequent repetition on the registration form of words such as ‘compare, comparison’ and ‘validation’. That might result from the large-scale cover (continental) of LUCAS Soil. Furthermore, the word cloud analysis indicated that modelling of soil processes, soil carbon dynamics, soil erosion evaluation, agricultural studies, land-use evaluation, remote sensing analysis and ecosystem service assessment are the main applications leading people to request the LUCAS Soil dataset.

A post-download survey was also developed to evaluate the actual use of LUCAS data. In 2015, all users who had requested LUCAS Soil data during 2013 and 2014 received a simple questionnaire to obtain not only a picture of the real LUCAS Soil use, but also

valuable advice for improving the quality of data. The questionnaire evaluated four aspects: (i) data utility (four possible answers: very useful, somehow useful, not useful or not used), (ii) scale at which LUCAS data were used (global, European or transboundary, national, regional or local), (iii) publication output attributable to LUCAS data and (iv) free comments or suggestions on the LUCAS Soil dataset. In total, 177 of 408 users involved replied to all questions. Among them, 121 confirmed that they actually used the downloaded data. The majority of the users (36%) applied LUCAS Soil to studies at the national scale (Figure 5d). The level of satisfaction was high (55% of the users marked the dataset as ‘very useful’) and many users had produced publications from the data (25% reported publication or submission to scientific journals after 1 year from the download). Free comments were examined and taken into account for both data implementation and format quality. For instance, several users suggested the need to measure bulk density, which will be included in the LUCAS Soil 2018 survey.

Overall, statistics on LUCAS Soil download and use emphasize the value of such a resource. Continuous development of the LUCAS Soil database and derived datasets will ensure further growth of potentially interested users and thus their application in new fields of research. The registration and monitoring approaches adopted by the ESDAC portal (ESDAC, 2017) for requesting data ensure direct contact with the users. They provide constant monitoring of use of the LUCAS Soil data, and user feedback makes it feasible to implement changes in data quality standards according to users’ needs. This kind of self-assessment strategy will be maintained to guarantee a high level of data value and re-use, as well as satisfaction among users.

Future developments and perspectives

The LUCAS Soil dataset was developed as a continuously growing resource in terms of both quantity and quality. Quantitatively, the number of soil samples collected has increased over the time of the surveys because of both the inclusion of new points (points over 1000 m altitude were added in 2015 to the 2009–2012 survey) and

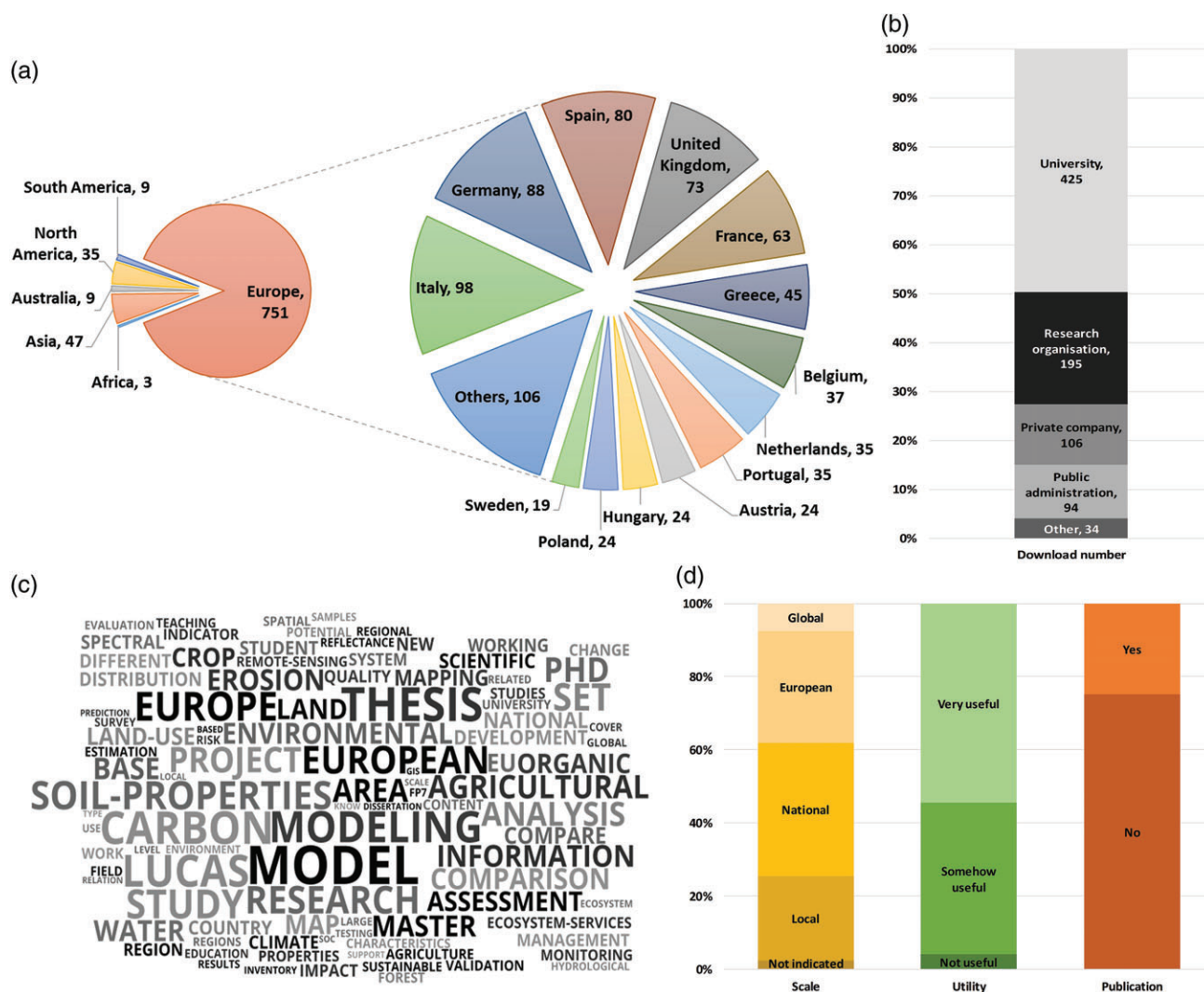


Figure 5 Numbers of downloads of LUCAS 2009–2012 point data (a) per country and (b) per type of organization. Values indicate the number of download(s). (c) A word cloud analysis (Wordle, 2017) of replies regarding data use. The analysis shows the 100 most common words used by the respondents. Word size is proportional to the number of repetitions. (d) Percentages of the LUCAS use potential. In 2015, a post-download questionnaire was sent to all users. Respondents were asked to provide information on three aspects: (i) scale at which data were used, (ii) utility of the data and (iii) publication output associated with LUCAS Soil data.

expansion to new countries. In particular, the inclusion of countries outside the EU has further reinforced the value of LUCAS Soil. The number of soil properties being measured has also increased over the period of the sampling campaigns. In addition to the variables already decided upon for the 2018 sampling campaign, issues such as new potential soil pollutants (other than metals), soil metagenomics (Daniel, 2005) and soil structure (e.g. visual evaluation of soil structure, VESS) are being explored for possible inclusion in the future LUCAS Soil. Furthermore, the number of points where new properties (e.g. bulk density and soil biodiversity) are measured will increase in future campaigns, with a future aim to have them for all sampling points.

In addition to the data from the 2015 survey, which we aim to make available by the end of 2017 beginning of 2018, a

major step forward to LUCAS Soil's growth will be the 2018 campaign. Indeed, all the new variables (e.g. soil biodiversity and bulk density) can be combined to provide a better understanding of biogeochemical and ecological processes occurring in the soils of the EU. It is anticipated that the European Soil Data Centre will host a dedicated section for the LUCAS Soil Biodiversity database. We will develop visualization tools that enable maps to be generated based on the presence of soil organisms and to compare soil biodiversity under different types of land use. Furthermore, the entire methodology (from DNA extraction to bioinformatics processing) will be available on ESDAC (ESDAC, 2017) to facilitate its use by external users to ensure compatibility and eventually data sharing. Our proposals could represent a new strength that would allow the LUCAS Soil Biodiversity

database to develop as a self-growing database with inputs not only from LUCAS surveys every 3 years, but also from direct user contributions.

Another key feature of LUCAS Soil is its monitoring attributes, which mean that the analyses are repeated at fixed locations and time intervals (every 3 years). This makes LUCAS Soil a powerful tool to assess environmental changes over time at the continental scale. This also opens the doors to the possible use of LUCAS Soil data for purposes other than scientific research. Soil and its biodiversity are now well recognized as key players in regulating several ecosystem services, such as climate regulation and human health control (Wall *et al.*, 2015). The possibility of having a harmonized continental database of soil chemical, physical and biological properties will be crucial not only for its potential research applications, but also for policy development. For example, outputs from LUCAS Soil are already taken into account to implement the Common Agricultural Policy (CAP) of the EU and to make farmers more aware of the need to manage soil in a more sustainable manner. Recently, Lugato *et al.* (2014a) used the LUCAS data to test management practices for potential carbon sequestration in arable soils in Europe. The results of this research are currently used in both the Common Agricultural Policy (CAP) and climate change policies. Furthermore, LUCAS Soil-derived datasets, such as maps of soil erosion, will be used as one of the indicators for achieving the United Nations' Sustainable Development Goals by the European Union. Finally, there is a certain time lag between the availability of research results and the implementation into policies (Salter & Martin, 2001). The development of environmental policies in the European Union is a complex process, which mixes legal requirements with issues of technical feasibility, scientific knowledge and socio-economic aspects, and which requires intensive multi-stakeholder consultations that are time consuming (Quevauviller *et al.*, 2005). In conclusion, LUCAS Soil will increasingly represent a valid instrument that could be used to facilitate science-driven decisions and to contribute to the implementation of the environmental and economic political agenda of the European Union.

Availability of data

The LUCAS Soil 2009–2012 single point dataset (LUCAS, 2017) and all of the derived material presented in this paper are available in the European Soil Data Centre (ESDAC, 2017). Data from the 2015 survey will be made available on the ESDAC Portal in early 2018. The corresponding LUCAS land-cover and land-use data are also freely accessible at the EUROSTAT website (EUROSTAT, 2017).

Acknowledgements

The LUCAS Survey is coordinated by Unit E4 of the Statistical Office of the European Union (EUROSTAT). The LUCAS Soil sample collection and laboratory analyses are supported by the Directorate-General Environment (DG-ENV), Directorate-General

Agriculture and Rural Development (DG-AGRI) and Directorate-General Climate Action (DG-CLIMA).

References

- Arribas, P., Andújar, C., Hopkins, K., Shepherd, M. & Vogler, A.P. 2016. Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. *Methods in Ecology and Evolution*, **7**, 1071–1081.
- ASI (Austrian Standards Institute) 1999. *ÖNORM L 1084*. Chemical Analyses of Soils – Determination of Carbonate. Austrian Standards Institute, Vienna.
- Ballabio, C., Panagos, P. & Montanarella, L. 2016. Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma*, **261**, 110–123.
- Bardgett, R.D. & van der Putten, W.H. 2014. Belowground biodiversity and ecosystem functioning. *Nature*, **515**, 505–511.
- Barré, P., Fernández-Ugalde, O., Virto, I., Velde, B. & Chenu, C. 2014. Impact of phyllosilicate mineralogy on organic carbon stabilization in soils: incomplete knowledge and exciting prospects. *Geoderma*, **235**, 382–395.
- Barrios, E. 2007. Soil biota, ecosystem services and land productivity. *Ecological Economics*, **64**, 269–285.
- Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.-J. *et al.* 2012. Tracking earthworm communities from soil DNA. *Molecular Ecology*, **21**, 2017–2030.
- Bissett, A., Fitzgerald, A., Meintjes, T., Mele, P.M., Reith, F., Dennis, P.G. *et al.* 2016. Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity database. *GigaScience*, **5**, 1–11.
- Borrelli, P., Panagos, P., Langhammer, J., Apostol, B. & Schütt, B. 2016. Assessment of the cover changes and the soil loss potential in European forestland: first approach to derive indicators to capture the ecological impacts on soil-related forest ecosystems. *Ecological Indicators*, **60**, 1208–1220.
- Borrelli, P., Lugato, E., Montanarella, L. & Panagos, P. 2017. A new assessment of soil loss due to wind erosion in European agricultural soils using a quantitative spatially distributed modelling approach. *Land Degradation and Development*, **28**, 335–344.
- de Brogniez, D., Ballabio, C., Stevens, A., Jones, R.J.A., Montanarella, L. & van Wesemael, B. 2015. A map of the topsoil organic carbon content of Europe generated by a generalized additive model. *European Journal of Soil Science*, **66**, 121–134.
- CLC (CORINE Land Cover) 2006. *CORINE Land Cover 2006*. Copernicus Programme, Copenhagen [WWW document]. URL <http://land.copernicus.eu/pan-european/corine-land-cover> [accessed on 08 September 2017].
- Crowther, T.W., Todd-Brown, K.E.O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B. *et al.* 2016. Quantifying global soil carbon losses in response to warming. *Nature*, **540**, 104–108.
- Daniel, R. 2005. The metagenomics of soil. *Nature Reviews Microbiology*, **3**, 470–478.
- EMP (Earth Microbiome Project) 2017. *Protocols and Standards*. Earth Microbiome Project, La Jolla, CA [WWW document]. URL <http://www.earthmicrobiome.org/protocols-and-standards/> [accessed on 08 September 2017].
- ESDAC (European Soil Data Centre) 2017. *European Soil Data Centre*. Joint Research Centre, Ispra [WWW document]. URL <http://esdac.jrc.ec.europa.eu/> [accessed on 08 September 2017].

- Ettema, E.H. & Wardle, D.A. 2002. Spatial soil ecology. *Trends in Ecology & Evolution*, **17**, 177–183.
- EUROSTAT (Statistical office of the European Union) 2017. *Land Cover/Use Statistics (LUCAS) Database*. EUROSTAT, Luxembourg [WWW document]. URL <http://ec.europa.eu/eurostat/web/lucas/data/database> [accessed on 08 September 2017].
- Gilbert, J.A., Jansson, J.K. & Knight, R. 2014. The Earth Microbiome project: successes and aspirations. *BMC Biology*, **12**, 69.
- Guo, L.B. & Gifford, R.M. 2002. Soil carbon stocks and land use change: a meta analysis. *Global Change Biology*, **8**, 345–360.
- Hamza, M.A. & Anderson, W.K. 2005. Soil compaction in cropping systems. A review of nature, causes and possible solutions. *Soil & Tillage Research*, **82**, 121–145.
- Hogg, E.H., Liefers, V.J. & Wein, R.W. 1992. Potential carbon losses from peat profiles: effects of temperature, drought cycles, and fire. *Ecological Applications*, **2**, 298–306.
- ISO (International Organization for Standardization) 1994a. *ISO 10390:1994 Soil Quality Determination of pH*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1994b. *ISO 11260:1994 Soil Quality Determination of Effective Cation Exchange Capacity and Base Saturation Level Using Barium Chloride Solution*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1994c. *ISO 11263:1994 Soil Quality Determination of Phosphorus Spectrometric Determination of Phosphorus Soluble in Sodium Hydrogen Carbonate Solution*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1994d. *ISO 11265:1994 Soil Quality Determination of the Specific Electrical Conductivity*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1995a. *ISO 10693:1995 Soil Quality Determination of Carbonate Content Volumetric Method*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1995b. *ISO 10694:1995 Soil Quality Determination of Organic and Total Carbon after Dry Combustion (Elementary Analysis)*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 1995c. *ISO 11261:1995 Soil Quality Determination of Total Nitrogen Modified Kjeldahl Method*. International Organization for Standardization, Geneva.
- ISO (International Organization for Standardization) 2009. *ISO 13320:2009 Particle Size Analysis Laser Diffraction Methods*. International Organization for Standardization, Geneva.
- Keesstra, S.D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A. et al. 2016. The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *SOIL*, **2**, 111–128.
- LUCAS (Land Use/Cover Area frame statistical Survey) 2017. *LUCAS 2009 TOPSOIL Data*. Joint Research Centre, Ispra [WWW document]. URL <http://esdac.jrc.ec.europa.eu/content/lucas-2009-topsoil-data> [accessed on 08 September 2017].
- Lugato, E., Bampa, F., Panagos, P., Montanarella, L. & Jones, A. 2014a. Potential carbon sequestration of European arable soils estimated by modelling a comprehensive set of management practices. *Global Change Biology*, **20**, 3557–3567.
- Lugato, E., Panagos, P., Bampa, F., Jones, A. & Montanarella, L. 2014b. A new baseline of organic carbon stock in European agricultural soils using a modelling approach. *Global Change Biology*, **20**, 313–326.
- Lugato, E., Paniagua, L., Jones, A., de Vries, W. & Leip, A. 2017. Complementing the topsoil information of the Land Use/Land Cover Area Frame Survey (LUCAS) with modelled N₂O emissions. *PLoS One*, **12**, e0176111.
- Minasny, B. & McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers & Geosciences*, **32**, 1378–1388.
- Nocita, M., Stevens, A., Tóth, G., Panagos, P., van Wesemael, B. & Montanarella, L. 2014. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology & Biochemistry*, **68**, 337–347.
- Orgiazzi, A., Dunbar, M.B., Panagos, P., de Groot, A.G. & Lemanceau, P. 2015. Soil biodiversity and DNA barcodes: opportunities and challenges. *Soil Biology & Biochemistry*, **80**, 244–250.
- Orgiazzi, A., Panagos, P., Yigini, Y., Dunbar, M.B., Gardi, C., Montanarella, L. et al. 2016. A knowledge-based approach to estimating the magnitude and spatial patterns of potential threats to soil biodiversity. *Science of the Total Environment*, **545–546**, 11–20.
- Palmieri, A., Martino, L., Dominici, P. & Kasanko, M. 2011. Land cover and land use diversity indicators in LUCAS 2009 data. In: *Land Quality and Land Use Information* (eds G. Tóth & T. Németh), pp. 59–68. Publications Office of the European Union, Luxembourg.
- Panagos, P., Van Liedekerke, M., Jones, A. & Montanarella, L. 2012. European Soil Data Centre: response to European policy support and public data requirements. *Land Use Policy*, **29**, 329–338.
- Panagos, P., Meusburger, K., Ballabio, C., Borrelli, P. & Alewell, C. 2014. Soil erodibility in Europe: a high-resolution dataset based on LUCAS. *Science of the Total Environment*, **479–480**, 189–200.
- Panagos, P., Borrelli, P., Meusburger, C., Alewell, C., Lugato, E. & Montanarella, L. 2015a. Estimating the soil erosion cover-management factor at European scale. *Land Use Policy*, **48C**, 38–50.
- Panagos, P., Borrelli, P., Meusburger, K., van der Zanden, E.H., Poesen, J. & Alewell, C. 2015b. Modelling the effect of support practices (P-factor) on the reduction of soil erosion by water at European Scale. *Environmental Science & Policy*, **51**, 23–34.
- Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K. et al. 2015c. The new assessment of soil loss by water erosion in Europe. *Environmental Science & Policy*, **54**, 438–447.
- Pimentel, D., Harvey, C., Resosudarmo, P., Sinclair, K., Kurz, D., McNair, M. et al. 1995. Environmental and economic costs of soil erosion and conservation benefits. *Science*, **276**, 1117–1123.
- Quevauviller, P., Balabanis, P., Fragakis, C., Weydert, M., Oliver, M., Kaschl, A. et al. 2005. Science-policy integration needs in support of the implementation of the EU Water Framework Directive. *Environmental Science & Policy*, **8**, 203–211.
- Rhoades, J.D., Manteghi, N.A., Shouse, P.J. & Alves, W.J. 1989. Soil electrical conductivity and soil salinity: new formulations and calibrations. *Soil Science Society of America Journal*, **53**, 433–439.
- Salter, A.J. & Martin, B.R. 2001. The economic benefits of publicly funded basic research: a critical review. *Research Policy*, **30**, 509–532.
- Sapkota, R. & Nicolaisen, M. 2015. High-throughput sequencing of nematode communities from total soil DNA extractions. *BMC Ecology*, **15**, 1–8.
- Singh, B. & Schulze, D.G. 2015. Soil minerals and plant nutrition. *Nature Education Knowledge*, **6**, 1.

- Tóth, G., Jones, A. & Montanarella, L. 2013. *LUCAS Topsoil Survey. Methodology, Data and Results*. JRC Technical Reports Number 26102. Publications Office of the European Union, Luxembourg.
- Tóth, G., Hermann, T., Da Silva, M. & Montanarella, L. 2016. Heavy metals in agricultural soils of the European Union with implications for food safety. *Environment International*, **88**, 299–309.
- USDA NRCS. 2004. *Ion Exchange and Extractable Cations: Potassium Measured with Atomic Absorption Spectrophotometer*. Soil Survey Laboratory Methods Manual Soil Survey Investigations Report 42. United States Department of Agriculture, Natural Resources Conservation Service, Washington, DC.
- Wall, D.H., Nielsen, U.N. & Six, J. 2015. Soil biodiversity and human health. *Nature*, **528**, 69–76.
- Wild, S. 2016. Quest to map Africa's soil microbiome begins. *Nature*, **539**, 152.
- Wordle 2017. *Word Clouds Generator*. Jonathan Feinberg, [WWW document]. URL <http://www.wordle.net/> [accessed on 08 September 2017].
- Yigini, Y. & Panagos, P. 2016. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. *Science of the Total Environment*, **557**, 838–850.