European Journal of Soil Science, 2005, Volume 56 Issue 5, Pages 655 – 671

Estimating Organic Carbon in the Soils of Europe for Policy Support

R.J.A. JONES[1,3], R. HIEDERER[2], E. RUSCO[1] & L. MONTANARELLA[1]

[1]*Institute for Environment & Sustainability, Joint Research Centre, Ispra (VA) 21020, Italy,* [2]*Land Management Unit, Institute for Environment & Sustainability, Joint Research Centre, Ispra (VA) 21020 Italy and* [3]*National Soil Resources Institute, Cranfield University, Silsoe, Bedfordshire MK45 4DT, UK*

Correspondence: R.J.A. Jones. Email: r.jones@cranfield.ac.uk

Running title: *Estimating soil organic carbon for Europe*

Summary

The estimation of soil carbon content is of pressing concern for soil protection and in mitigation strategies for global warming. This paper describes the methodology developed and the results obtained in a study aimed at estimating organic carbon contents (%) in topsoils across Europe. The information presented in map form provides policy makers with estimates of current topsoil organic carbon contents for developing strategies for soil protection at regional level. Such baseline data is also of importance in global change modelling and may be used to estimate regional differences in soil organic carbon (SOC) stocks and projected changes therein, as required for example under the Kyoto Protocol to UNFCCC, after having taken into account regional differences in bulk density.

The study uses a novel approach combining a rule-based system with detailed thematic spatial data layers to arrive at a much-improved result over either method, using advanced methods for spatial data processing. The rule-based system is provided by the pedo-transfer rules, which were developed for use with the European Soil Database. The strong effects of vegetation and land use on SOC have been taken into account in the calculations, and the influence of temperature on organic carbon contents has been considered in the form of a heuristic function. Processing of all thematic data was performed on harmonized spatial data layers in raster format with a 1km x 1km grid spacing. This resolution is regarded as appropriate for planning effective soil protection measures at the European level. The approach is thought to be transferable to other regions of the world that are facing similar questions, provided adequate data are available for these regions. However, there will always be an element of uncertainty in estimating or determining the spatial distribution of organic carbon contents of soils.

Introduction

Following the unprecedented expansion and intensification of agriculture during the 20th century, there is clear evidence of a decline in the organic carbon (OC) contents in many soils as a consequence (Sleutel *et al.*, 2003). This decline in OC contents has important implications for agricultural production systems, because OC is a major component of soil organic matter (OM). OM is an important 'building block' for soil structure and for the formation of stable aggregates (Waters & Oades, 1991, Beare *et al.*, 1994). The benefits of OM are linked closely to the fact that it acts as a storehouse for nutrients, is a source of soil fertility and contributes to soil aeration, thereby reducing soil compaction. Other benefits are related to the improvement of infiltration rates and the increase in storage capacity for water. Furthermore, OM serves as a buffer against rapid changes in soil reaction (pH) and it acts as an energy source for soil micro-organisms. Moreover, soil OM might be sequestered by vegetation and soils, as a possible way of mitigating some detrimental effects of Global Change. These circumstances have heightened the interest in quantifying the OC contents of soils at regional as well as global level. The official Communication 'Towards a Thematic Strategy for Soil Protection' (EC, 2002), adopted in April 2002, is an additional stimulus to studying the geographical distribution of soil OC. The Communication identifies eight main threats to soil, of which declining OM is considered one of the most serious, especially in southern Europe.

There have been several attempts to estimate carbon stocks at regional level in Europe (Howard *et al.*, 1995; Batjes, 1996; Smith *et al.*, 2000a, b; Arrouays *et al.*, 2001). Estimates of organic carbon stock at national level were established, for example for the UK by Howard *et al.* (1995) for land under arable agriculture using OC measurements made during the National Soil Inventories in England & Wales and Scotland (1979-83). Smith *et al.* (2000b) revised the estimates of Howard *et al.* (1995) for the UK using data compiled by Batjes (1996) and a relationship that assumes a quadratic decline in soil OC contents with depth. Arrouays *et al.* (2001) calculated OC stocks in the soils of France using the CORINE land cover database, the 1:1,000,000 scale soil geographical database of France and a geographical database containing OC measurements. Lettens *et al.* (2004) used soil OC data collected during 1950-70 from more than 30,000 soil profiles excavated during the soil survey of Belgium. Despite the size of the sampled data, all these studies have the potential problem of assigning point measurements of OC to polygons representing large areas of land with no additional validation of the OC values assigned. Furthermore, they do not provide a basis for estimating OC of soils at the European level, which is accurate enough for policy support.

In contrast to this study, the primary aim of these investigations was to estimate the carbon sequestration potential of soils in global change research: For example, Batjes (1996, 2002) used the WISE database and calculated OC contents for the major soil groups for the purpose of estimating stocks. However, similar to our study, Batjes (1997) estimated OC contents (%) for FAO Reference Soil Groups and, in an attempt to guide policy makers at European level, Rusco *et al.* (2001) estimated OC in topsoils by applying a pedo-transfer rule (PTR21) to the data stored in the European Soil Database.

Methodology

The objective of this study was to produce a continuous pan-European cover of quantitative OC content in the topsoil, taken as 0-30cm depth. An extrapolation procedure based on sample data was deemed unsuitable for the task. The main reason for developing an alternative method to point-data extrapolation was that, although OC contents have been measured systematically in some countries, for example UK, Denmark, The Netherlands and Slovakia, or non-systematically though comprehensively, for example in Belgium, France, Hungary and Italy, the number of samples analysed at the European level is still insufficient to generate an accurate spatial distribution at the required scale. Furthermore, the sample data from national field surveys are regrettably either insufficiently geo-referenced or not accessible outside the country of origin. Another important reason for developing an alternative method to extrapolating from point data stems from the well-known fact that OC contents can vary within pedologically defined soil units, depending on vegetation and land management. This is clear from the data computed by Batjes (1996, 1997), who determined a coefficient of variation (CV) in topsoil OC contents of between 50 and 150% for the same pedological (Reference) soil group. This tendency for large variation in OC contents increases the difficulty of accurately estimating OC stocks in soils.

To overcome the limitations in data availability and intrinsic variability in soil properties, this study developed a distinct procedure, which centres on the processing of a structured series of conditions for defining topsoil OC in a Geographic Information System (GIS). The principal modules of the procedure are depicted in form of a flow chart in Figure 1.

The main data source for the study is the European Soil Database (ESDB), which originates from national soil surveys, following harmonization to provide a seamless spatial and thematic cover of European soil properties (King *et al.*, 1994). The ESDB consists of two main databases, the Soil Geographic Database (SGDB) and the Pedo-Transfer Rules Database (PTRDB) - see Daroussin & King (1997). Both databases were used to produce a European Raster Database, which contains a selected number of thematic soil properties as spatial data layers in raster format (Hiederer *et al.*, In press).

The PTRDB includes a set of conditions for defining topsoil OC, which are arranged in the pedo-transfer rule No. 21 (PTR 21). This rule has been revised and translated into processing commands, which operate directly on spatial data layers in a Geographic Information System (GIS). The spatial layer was combined with spatial data layers from the raster database (for soil properties), a European Land Cover layer (for land use) and a temperature layer (for OC temperature correction). All input data were processed to produce topsoil OC content layers on a 10-year basis, ranging from 1900 to 1990. The data layer for the decade 1980 to 1989 forms the baseline for calculating topsoil carbon stocks in European soils, since it relates most closely to 1990, the baseline chosen for the Kyoto Protocol. Verification of the final OC estimates obtained from the processing chain was performed by comparing the modelled data with measured values from over 12 000 ground samples, which were available to the study from soil surveys conducted in the UK (England and Wales) and Italy.

Data Sources

Soil: European Soil Database

The European Soil Database v.1.0 (Heineke *et al*., 1998) has been constructed from source material prepared and published at a scale of 1:1 000 000 (CEC, 1985). The resulting soil data have been harmonised for the whole area covered, according to a standard international classification (FAO-UNESCO, 1974; FAO-UNESCO-ISRIC, 1990), together with analytical data for standard profiles (Madsen and Jones, 1995). The spatial component of this database comprises polygons, which define Soil Mapping Units (SMUs). These spatial elements can be linked to soil attributes, which are referred to as Soil Typological Units (STUs) and stored in a thematic database. Although each STU is unambiguously defined, an SMU may comprise up to 10 STUs. The spatial location of STUs within an SMU is not known, only the proportion of each STU in the SMU. Hence, a soil property can only be diffusely mapped at the resolution of the SMU. While this structure of the

European Soil Database allows relatively efficient data storage, it is not particularly well-suited for spatial analysis or for combining external information. Therefore, a set of attributes in raster format, which were generated from combining SMUs with all linked STUs, was used in the study (Hiederer *et al.*, In press).

Land Use/Cover: European Land Cover Data

The land use data utilized in the study were taken from the European Land Cover Data layer of the Catchment Information System (CIS) (Hiederer, 2001). The layer covers Europe with information according to the CORINE Land Cover (LC) classification codes. The layer was generated by combining specifically adjusted data from the CORINE LC raster dataset combined with data from the Eurasia land cover data derived from the US Geological Survey (USGS) (United States Geological Survey, 2003). To achieve comparable thematic coverage between the data sets, a series of cross-classifications was carried out, in which various USGS data layers were re-assigned or merged. The final layer corresponds to CORINE level 3 classification codes and is spatially fully compatible with the layers of the CIS. For use in the pedo-transfer rule for OC, the European Land Cover data were then re-classed to the four land use types used in the original PTR21 in the interest of simplicity.

Climate: GHCN

An original spatial layer was generated comprising Average Annual Accumulated Temperature (AAAT), expressed in day degrees Celsius (day degrees C). The layer data are based on meteorological data from the Global Historical Climatology Network - GHCN (Easterling *et al.*, 1996). Spatial layers were derived from the point data through a weighted-distance interpolation. The influence of station altitude on temperature observations was adjusted for by applying an adapted moist adiabatic lapse rate. The

AAAT spatial layers were calculated using average monthly temperatures from 1890 to 1990. The AAAT layer for the decade 1970 to 1979 was used to calculate the OC_TOP validation layer because this period covers the decade prior to the ground sampling. The influence of moisture on OC was not specifically modelled though this soil-forming factor is implicitly taken into account in the soil type. For example, a Gleysol by definition is a soil that shows evidence of water logging within 50cm of the surface.

Verification: Soil Data from Ground Surveys

Data from national soil surveys were available for the UK (England and Wales) and Italy, thus covering a wide range of European soils and climatic conditions.

England & Wales. Measured OC data from England & Wales were available from ground samples taken during the National Soil Inventory (NSI) in the period 1979-1983 (McGrath & Loveland, 1992). OC was determined by a widely used wet dichromate acid digestion method (Avery & Bascomb, 1982). The sampling procedure was a systematic scheme, using a 5km x 5km grid (McGrath & Loveland, 1992). Sample sites include all land cover types, with the exception of some built-up areas, and the data exist for >5500 points. The systematic nature of the ground samples allows comparison of modelled estimates with measured data over a wide range of soil types, environmental conditions and OC values.

Italy. The measured OC data for Italy were derived from a monitoring network on agricultural land. The 6779 sample locations are strongly clustered in some areas and it is possible that a plot sampled contained grassland as well as arable crops. The data used in this study were compiled by Rusco (In prep.) and analysed by a method similar to that used in England & Wales. The sampling scheme, and the limitations imposed by the location of sample sites, render the Italian ground data unsuitable for the compilation of general statistics for administrative units. However, the data can be used to verify OC estimates for southern European conditions on agricultural land.

Pedo-Transfer Rules

A Pedo-transfer Rule (PTR) forms the basis for calculating OC in the methodology developed during this study (PTR21). The system of PTRs present in the European Soil Database was developed by Van Ranst *et al.* (1995) to extend the range of soil parameters not normally observed or measured during soil surveys, but can be inferred from a combination of soil properties commonly measured or observed. The principal parameters defining a property and the representative value for that property are identified through expert knowledge (Jones & Hollis, 1996). The PTRDB consists of 34 PTRs (Daroussin & King, 1997), each producing values of a single soil parameter as its output. The output values of the parameter are defined through a sequence of conditions, representing, in a structured form, the typical situations found in the field survey data. The conditions use a variety of related environmental parameters. They are applied sequentially, starting from general situations and proceeding to more specific situations. As a consequence, the order in which the conditions are applied is part of the rule.

The common form of using such rules is to apply them to each STU in the European Soil Database to generate a new attribute by STU. This study implements the PTR concept using a different methodology. Firstly, the PTR is not applied to tabulated data, but calculations are performed on spatial data layers directly. Secondly, external data are used for land use and temperature in place of data for these parameters originally stored in the European Soil Database. Furthermore, the influence of temperature on OC content has been removed as a parameter from the revised rule and is now calculated using a mathematical function.

Topsoil OC content defined by PTR 21 uses six input parameters and comprises 150 conditions (Van Ranst *et al.* 1995). The input parameters (see Table 1) are (1) the first character in the FAO code (item SOIL in the database), (2) the second character in item SOIL, (3) the third character in item SOIL, (4) the dominant surface textural class (TEXT), (5) the land use class (USE) and (6) the accumulated temperature class (ATC) of the European Soil Database .

The first step in using the PTR 21 as a basis for estimating topsoil OC was to analyse the existing conditions and to remove any ambiguity in the sequence of application. Following the absence of any conditions differentiating soils with OC content in excess of 6%, the next modification was to define two new OC_TOP classes, one for soils with 18 to 30% OC (very high) and a second for soils >30% OC (extremely high).

Next, values for the USE parameter of the soil database were substituted by those from the European Land Cover data layer. The substitution of the information does not affect the conditions of the PTR, but greatly transforms the method of data processing from computing records in a table to analysing individual pixels of the spatial layer.

The ATC parameter was removed completely from the conditions. This was considered necessary, because the class definitions are rather coarse and version 1.0 of the soil database contains only the class ‘medium’. Thus, any condition using ATC as a defining parameter was effectively ignored in previous applications of PTR 21. In total, 112 modifications were made to the previous rule and 24 new conditions were added. The removal of the ATC parameter from the revised rule requires subsequent processing to account for the influence of temperature (see below).

The revised PTR for OC_TOP has 5 input parameters and comprises 140 conditions, an extract being given in Table 1, which can be translated into programming code as follows:

```
:
17 IF       (SN1=L) AND (SN2=c) AND (TEXT=2) AND (USE=C)  THEN LET OC_TOP=L
18 ↳    IF  (SN1=L) AND (SN2=c) AND (TEXT=2) AND (USE=MG) THEN LET OC_TOP=M
:
68 IF       (SN1=G) AND (SN2=f) AND (SN3=m) AND (TEXT=2) AND (USE=SN) THEN LET OC_TOP=H
69 ↳    IF  (SN1=G) AND (SN2=f) AND (SN3=m) AND (TEXT=3) AND (USE=SN) THEN LET OC_TOP=H
:
77 IF       (SN1=J) AND (SN3=g) THEN LET OC_TOP=M
78 ↳    IF  (SN1=J) AND (SN3=g) AND (TEXT=4) AND (USE=SN) THEN LET OC_TOP=H
:
```

Conditions 17 and 18 of the revised PTR define class values for OC_TOP for *Chromic Luvisols* (*Lc*) with texture class 2 (18% < clay < 35% and sand > 15%, or clay < 18% and 15% < sand < 65%). For such soil under cultivation (USE = C), an OC_TOP class = ‘L’ (1 - 2% OC content) is assigned (Condition 17). Where the soil is under managed grassland, an OC_TOP class =‘M’ (2-6% OC content) is assigned instead (Condition 18). Conditions 68 and 69 are examples of conditions added to the original PTR. They apply to *Molli-fluvic Gleysols* with medium (TEXT = 2) or medium fine (TEXT = 3) texture under semi-natural vegetation (USE = SN). In both cases the OC_TOP class ‘H’ (6-18% OC content) is assigned. The conditions were added to the rule, because the situation was typical and not sufficiently defined in the original PTR. In contrast to the previous conditions, Conditions 77 and 78 are examples of defining OC_TOP going from general to more specific situations and the order of the rule is crucial to the correct functioning of the PTR. In Condition 77, any *Gleyic Fluvisols* are set to medium OC_TOP content. However, when such soils have a fine texture (TEXT = 4) and when land cover is semi-natural (USE = SN), the areas concerned are classified as = class ‘H’ (6 - 18% OC content).

Temperature Effect

The exclusion of the influence of temperature on OC_TOP in the revised PTR necessitated generating adequate information on temperature across the area of interest followed by developing a method to include the data in the evaluation outside the PTR. The first task was accomplished by creating the AAAT data layers. The second task was achieved by substituting the rule-based method with a mathematical function to account for the influence of temperature on OC_TOP. The function was developed in accordance with the established principle that, within belts of uniform moisture conditions and comparable vegetation, the average total OM and nitrogen in soils increase by two to three times for each 10 degrees C fall in mean temperature (Buckman & Brady, 1960, p.152). This is only a very general relationship, but it was thought to be suitable for this pan-European study. Based on this relationship and considerations for mathematically permissible minimum and maximum values, a sigmoidal function of type $y=a\cdot cos(x)^n$ was defined to relate changes in temperature with changes in OC content. The definition of the function parameters was later improved by using data from the ground surveys. The function is graphically presented in Figure 2.

Figure 2 shows the average ratio of the OC values of the ground data to the output of the revised PTR (OC_TOP_{PTR}) for 175 aggregated units. The aggregation was performed, because a display of all 12 275 ratio values produces little discernible information on the form of the relationship and the aggregated values allow a better visual interpretation of the relationship between AAAT and the temperature correction coefficient. Values were aggregated according to land use, FAO soil subgroup and temperature.

To reduce the influence of isolated values on the graphical representation, only those data points, which were defined by more than nine values, are displayed on the graph,. Applying this threshold procedure resulted in 175 aggregated ratio values (managed grassland: 33, semi-natural: 28, cultivated: 103, no information: 11) depicted in Figure 2.

The most obvious outlier in the graph is a value for $TEMP_{cor}$ of 2.25 and a value for AAAT of 5800 day degrees C. The point represents a site in Italy, where the ground samples are classified as *Chromic Vertisols*. Yet the sampled value for OC_TOP averages

31%. Such a large amount of OC precludes defining this soil as a *Vertisol* and thus, for verification purposes, this data point was excluded.

The parameters of the function for $TEMP_{cor}$ are defined in equation 1:

$$TEMP_{cor} = 1.1 \times \cos\left(4.24 \times 10^{-4} AAAT - 1.10\right)^{4} + 0.7 \qquad (1)$$

The equation is applicable within the range of 2200 to 6000 day degrees C. Below and above this range, constant values were used for $TEMP_{cor}$. Estimates of OC_TOP derived from the model (OC_TOP_{MOD}) were calculated by multiplying the OC_TOP_{PTR} value layer with the temperature coefficient layer in the GIS.

The parameters set $TEMP_{cor} = 1.0$ at 4300 day degrees C, i.e. the OC values output by the revised PTR for soil and land use remain unchanged at that temperature. Such AAAT values occur, for example, in southern England, northern France and southern Germany. From approximately 6000 day degrees C upwards a minimum value for $TEMP_{cor}$ of 0.7 is used. Areas with these high temperatures are mainly found in southern Europe. The $TEMP_{cor}$ value of 0.7 was determined by the ground data from Italy alone, where samples were restricted to cultivated land. The maximum value for $TEMP_{cor}$ was set to 1.8 and kept constant for AAAT values of 2200 or less. Areas with AAAT in this range are in northern Europe and in Alpine regions. On the basis of the aggregated mean AAAT in areas below 1800 day degrees C, one could assume a decrease in $TEMP_{cor}$ with decreasing temperature. However, the number of ground data located in such areas was limited to 32 data points of which 9 were located in areas of <1200 day degrees C. Except for one sample all data stem from the Italian survey. Since it would be unusual to have land cultivated under those temperature ranges and the number of samples is relatively low it was decided to exclude these data and to keep the value of $TEMP_{cor}$ constant for areas with AAAT values below 1800 day degrees C.

The maximum value of 1.8 for the temperature coefficient derived from the ground data ties in with the procedure for estimating OM content from OC_TOP. The maximum estimated OC_TOP content after applying the function was approximately 60%. Thus assuming a relatively stable OC:OM ratio of 1:1.72, the maximum value for estimated OM content is thus 100%. For the purpose of taking temperature into account for estimating OC_TOP from the revised PTR, no specific distinction by land use was made. The distribution of the ratio values depicted in Figure 2 would suggest a coefficient, which could vary by land use and possibly with region. Unfortunately, there is little overlap in the temperature ranges of the areas for which ground data were available to the study. Data from soil samples, including land use other than agriculture, would be required to determine different relationships. This could not be done in the scope of the study, but should be envisaged as a future investigation.

Processing Environment

All processing was performed using spatial data layers, including the SOIL parameter. The rules were converted into processing code of the GIS package used and applied to the spatial data layers. All data – soil, texture, land cover and climate – were compiled as standard 1km x 1km raster data sets for processing as spatial layers conforming to a Lambert Azimuthal Equal Area projection of the CIS. The projection parameters and the spatial frame are in accordance with the Eurostat GISCO database. All data processing was performed in the spatial domain using IDRISI 32 Release 2.

Results

The estimated OC contents in the surface horizon of soils in Europe, produced by applying the revised PTRs and temperature function to 1km spatial data layers of soil, climate and land cover, are shown in Figure 3. For display purposes the data layer of continuous values was grouped into seven classes (Jones *et al*., 2004a,b). The estimates cover an area of 4 947 079 km^2 and includes the following countries: Andorra, Albania, Austria, Bosnia and Herzegovina, Belgium, Bulgaria, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Greece, Croatia, Hungary, Ireland, Italy, Lichtenstein, Lithuania, Luxembourg, Latvia, Monaco, Former Yugoslav Republic of Macedonia, Malta, The Netherlands, Norway, Poland, Portugal, Romania, Serbia and Montenegro, Slovak Republic, Slovenia, Sweden, Switzerland, and United Kingdom.

Verification

To verify the calculated OC values in the surface horizon of European soils, the data were compared with measured OC data from sampling surveys on the ground in the UK (England and Wales) and Italy. The verification was performed for two different types of reference items: (1) *soil-related reference items*, i.e. ground and model data are compared following aggregation at the level of FAO soil subgroup codes and SMU units; (2) *soil-independent spatial items*, i.e. ground and model data are compared following aggregation based on catchments and NUTS (Nomenclature of Territorial Units for Statistics) as used by Eurostat. The use of the reference items required the aggregation of the data into comparable units.

Aggregation Units

1) FAO soil subgroup codes: The use of the FAO soil subgroup code as the aggregating unit allows an evaluation of differences between modelled and measured values using parameters that are also included in the PTR. This permits, to some degree, an assessment of the correctness of a condition within the PTR and can thus serve as a feedback to address any shortcomings in the existing rule-based system.

Because of the construction of the soil database (1:n SMU-STU relationship), it is not possible to generate a definite unambiguous assignment of OC_TOP values to specific soil types. Therefore, the soil type of the dominant STU in an SMU was used as representing the area. For England and Wales, there are 32 different subgroup codes for the dominant soils stored in the database, whereas the Italian ground data covers 22 different subgroup codes.

2) Soil Mapping Units: SMUs are the actual spatial units in the geographical component of the European Soil Database. England and Wales are covered by 75 SMUs, of which four do not contain any ground sample points because of their small extent. For Italy, it was not possible to calculate a meaningful OC value by SMU, because the data collection was concentrated on agricultural land.

3) Catchment Layer: The catchments used in the study were the primary data layer of the Catchment-based Information System (CIS) of the Joint Research Centre (Hiederer & de Roo, 2003). For England and Wales, 159 catchments are defined in the primary layer of the CIS and range in size from 1km^2 to 10 969km^2. The size of the spatial units is of importance, because small units have few or even no ground survey points. Therefore, the study concentrated on primary catchments larger than 1000km^2.

4) Administrative Layer: The aggregation to NUTS spatial units is directed at the implementation of environmental policies, such as protection measures, which are generally implemented across administrative regions. The administrative units used to aggregate the OC data are those of NUTS Level 2. For England and Wales, a total of 32 units is defined at this level, ranging from 322km^2 (Inner London) to 13 122km^2 (West Wales and The Valleys) in the GIS layer.

Ground vs. Modelled Data

The average OC_TOP content in the ground data was calculated using the arithmetic mean of the observed values of all points within a spatial unit. For the ground sample data, 95% confidence levels (CI_{95}) were calculated, as these allow an approximation of the range of values of topsoil OC content that can be expected for a given soil type in the field. Ground data were compared with modelled data separately by region and by land use category. The analysis used only ground data for which modelled data could be calculated. In Italy, only data from ground sample points in *cultivated* land were included, whereas for England and Wales all observations were used.

1) England & Wales - Ground Data vs. Modelled Data by FAO soil subgroup and SMU:

Figure 4 provides a graphical representation of the ground data CI_{95} for OC content and the mean value obtained from the model in England and Wales by FAO soil subgroup. A total of 5 289 points was used in the aggregation and the number of observations per FAO soil subgroup ranges from 5 for *Calcaric Regosols* (*Rc*) to 654 for *Stagno-gleyic Luvisols* (*Lgs*). There is generally an extremely close relationship between the average OC_TOP content of the ground and modelled data. From analysing all land cover classes, it is clear that the model overestimates OC content in the topsoil for *Histosols* (organic soils): *Dystric Histosols* (*Od*) produced a mean topsoil OC content of 36.4% for ground data vs. 45.5%

for model data. For the subgroup *Eutric Histosols* (*Oe*), the mean OC content was calculated as 14.8% for ground data vs. 20.4% for modelled data.

When analysing the results by land use, one should bear in mind that the stratification layer contains inconsistent land classes, either due to classification errors, the attribution of a dominant land use, where the ground sample was taken at a point with sub-dominant land use, or simply a change in land use between observation periods. A total of 1 885 points fell on cultivated land in the land use layer. The most obvious discrepancy between ground and modelled data for cultivated land occurs for *Dystric Histosols* (*Od*), where a mean of 39.9% OC content for ground data contrasts with 17.5% for modelled data. The soil subgroup value was determined by only two ground sample points in an SMU and in which the dominant STU covers 70% of the area (with 30% covered by *Oe*). For *Eutric Histosols* (*Oe*), the model over-estimates the average OC_TOP content by about 4% (15.5% for ground data, from 24 ground sample points) vs. 19.8% for modelled data. By contrast, the model underestimates OC for *Humic Gleysoils* (*Gh*) on cultivated land by 10.6%, though this finding is based on only 4 ground observations.

According to the land use layer, 1 012 ground sample points were located in semi-natural areas. Notable deviations from the generally good agreement between ground and modelled data were found only for *Dystric Histosols* (*Od*) and *Molli-fluvic Gleysols* (*Gmf*). The values for *Od* were determined by data from 95 sample points and the model overestimated the mean OC contents by 10% (38.1% mean ground data OC vs. 48.2% mean modelled OC). The OC values for *Gmf* were determined by just 2 sample points. The mean OC value for the ground data was 18.8%, while the mean modelled OC value was 9.9%. As indicated in the graph, the CI_{95} was also rather large for the soil subgroup and the modelled mean was within the range of the interval by FAO soil unit.

Using SMUs as the aggregation unit, the overall mean OC_TOP is 6.5% for the ground data and 6.4% for the modelled data at the locations of the ground samples. The results of aggregating OC_TOP content by SMUs can be characterized in form of a linear correlation. When relating the mean ground OC_TOP_{GRD} to the mean model OC_TOP_{MOD} at the locations of ground samples, the following regression equation was determined:

$$OC\_TOP_{GRD} = 0.82 * OC\_TOP_{MOD} + 1.45 \tag{2}$$

The coefficient of determination for the relationship (r^2) is 0.95 for the average values from 71 SMUs with data. This indicates a highly significant relationship between the modelled data and the situation found on the ground within the SMUs of England and Wales and suggests that the model predicts OC contents well..

2) England & Wales - Ground Data vs. Modelled Data by Catchment and NUTS:

The results for primary catchments larger than 1000km^2 and NUTS Level 2 units in England & Wales are given in Table 3. For each catchment and NUTS unit, the Table contains the number of ground sample points within the area covered, the mean value of OC_TOP content calculated from the ground survey and two values of mean OC_TOP contents calculated from the modelled OC_TOP content spatial layer. The mean OC_TOP derived from the ground sample data for the whole of England and Wales is 6.7% for catchments and administrative units. The average value calculated from the modelled data at the locations of the ground survey is 6.3% for larger catchments and for administrative units. With an average of 6.1% it is marginally less when using the complete area of either spatial unit. For ground data, the average OC_TOP values for catchments range from 1.5% (2.7% for NUTS) to 19.8% (13.9% for NUTS). The range of values for modelled data for catchments is similar, spanning from 1.5% (2.4% for NUTS) to 19.8% (14.3% for NUTS). The larger range of values in catchments than in the NUTS units can be explained by the number of smaller-sized catchments as compared to NUTS units, i.e. some local particularities are better represented in the smaller spatial units.

A graphical representation of the linear relation between ground observations and modelled data for England and Wales for catchments and NUTS units is given in Figure 5. The graph depicts for each primary catchment the data pair of average OC_TOP content derived from ground data and from modelled data. Filled marker points (●) represent averages from the point aggregation, boxes (☒) relate to values derived from area aggregation. The regression lines show the linear relationship between ground (OC_TOP_{GRD}) and modelled data (OC_TOP_{MOD}) aggregated over catchments >1000km^2 and NUTS Level 2 using point aggregation for all sample points. The mathematical expression of the relation is:

$$\text{Catchments:} \quad OC\_TOP_{GRD} = 0.88*OC\_TOP_{MOD} + 1.11 \tag{3}$$

$$\text{NUTS:} \quad OC\_TOP_{GRD} = 0.89*OC\_TOP_{MOD} + 1.07 \tag{4}$$

The coefficient of determination (r^2) of the relation is calculated as 0.94 for catchments and 0.93 for NUTS units. Determining the regression based on sample points, rather than the spatial units themselves, reduces the influence of varying unit size in the regression analysis. However, the simple calculation of the coefficient assumes that observations are independent. Yet, this is not the case when calculating the coefficient from aggregated sample points, because a fair degree of spatial dependence (auto-correlation) between observations exists, largely overestimating the degrees of freedom.

3) Italy - Ground Data vs. Modelled Data by FAO soil subgroup: The Italian data set contains 6 779 ground measurements of OC content, of which 5 436 points were used to relate ground to modelled data by FAO soil subgroup code. A graphical representation of the OC content for soils is summarized in Figure 6. Because sampling was restricted to agricultural land, the results show generally much smaller values for OC content compared to those found for cultivated land in England and Wales (see Figure 4). Values for the Italian data lie mainly in the range of 1-2% OC. This range is too small to calculate a meaningful coefficient of correlation between ground observations and modelled values. However, the data are ideal for calibrating the AAAT correction function for areas with small OC contents, characteristic of southern Europe. Noticeable is the over-estimation by the model of 6% for *Dystric Histosols* (*Od*) (5.1% ground data vs. 11.1% for model data). The mean value of OC content for *Od* was calculated from 11 points, which is not inappropriately small, but an examination of the location of the points reveals that they are distributed across four spatial elements of a single, spatially non-continuous SMU, two containing one sample, one containing two samples and one including seven sample sites. The values in the ground data included in the SMU vary from 0.8 to 14.0% OC content. The CI_{95} of the soil subgroup ranges from 2.9 to 7.9% and is the largest in the Italian data set. The distribution of soils in the SMU is 45% *Od*, 45% *Eutric Histosol*s (*Oe*) and 10% *Eutric Gleysols* (*Ge*). There are a number of possible explanations for the overestimation. Firstly, the ground samples sites were intentionally selected and clustered at the field scale, which accentuates the situation. Secondly, the Italian part of the European Soil Database was derived from a map of the soils of Italy drawn up in 1966, based on surveys made during the previous decade. The soils identified as *Histosols* during this survey, which subsequently have been sampled for the Italian OC data set, have been cultivated for more than 50 years. In this time the OM content has declined through mineralization to the extent that these soils may no longer be classified as organic. Thirdly, the sites sampled are probably small cultivated areas, which are located within a larger soil mapping unit dominated by pasture and/or semi-natural vegetation and hence not classified as arable in the land use layer.

4) Italy - Ground Data vs. Modelled Data by NUTS:

The results obtained from subjecting the Italian data to an analogous procedure of estimating OC_TOP content for the 20 NUTS Level 2 are summarized in Table 4. The analysis of soil-independent units was restricted to NUTS, because the use of catchments did not give any significantly different answers. The total number of sample points used in the analysis of OC content by NUTS for the Italian data set was 4 500. The number was less than in the analysis of soils because some 1km grid cells contained more than one sample. In those cases the mean of all points within the grid cell was used. The mean values for OC content in the Table are weighted by the portion of arable land by region. The overall mean OC_TOP content for the ground measurements was 1.2%. The mean calculated for the modelled data over the subset of sample points was also 1.2%. This amount is small, but is to be expected for agricultural land in Italy since the dry conditions and high temperatures favour rapid oxidation of OM. The mean OC_TOP content, calculated from the area aggregation of the model data to NUTS units including all land use classes, is estimated at 2.4%. Although the OC values in the Italian data set are restricted by the selection criterion for sample sites, these findings indicate that the modelled data are correct estimates of OC_TOP content for agricultural land in Southern Europe, when aggregated at the NUTS Level 2.

Discussion and Conclusions

Our results demonstrate that the methodology described in this paper represents a realistic alternative to approaches based on direct extrapolation of point observations, either by assigning measured data from a small number of points (deemed to be representative of a particular soil type) to polygons delineated on a soil map that represent much larger areas with no measured values, or by employing a spatial extrapolation procedure of values derived from point data. Even with the apparently large number of ground data points (>12 000 values available to the study) some soils with limited spatial representation are hardly included in the sample data. A stratification of the area by land use further reduces the number of observations per soil type and, as a consequence, lessens the reliability of estimating OC content of a soil type under different land uses from ground data. A sophisticated pedo-transfer rule has been successfully applied to the most detailed (1:1 000 000 scale) harmonized spatial soil data that currently exist for Europe. The conditions defined in the rule are a concentration of expert knowledge in the field soil OC content. The original PTR, defined by Van Ranst *et al.* (1995), was to some extent limited by the data available in the database. Having more detailed data available for land use and temperature has allowed the original rule (PTR 21) to be modified and extended to better distinguish between soils of large OC content. Processing directly in the spatial domain was made possible by technological advances in computer hardware and software. The results are thus encouraging not only because of the detailed quantification of soil OC content at the European scale, but also for demonstrating the viability of using comprehensive spatial databases to generate standardized data layers that can be calibrated by actual measurements (where these are available). There are several other sources of variation that could result in the calculated OC values deviating from the measured data from ground surveys. Firstly, topsoil OC contents are known to vary considerably from place to place because of differing land use history, timing of sampling and small variations in soil drainage conditions. Secondly, the land use at the time of sampling might have been different from that defined by the land cover data set (valid for the period 1988-92). This could be a result of land use change or merely the effect of scale.

However, the results obtained from our study also demonstrate some limits in the detail of OC content estimates presented in the corresponding data layer. One limitation is clearly set by the number of conditions defined in the rule. The more parameters that are taken into consideration the more precisely the conditions have to be defined. Even with one parameter less in the revised PTR, it was found necessary to define 140 conditions to characterize topsoil OC content. Rather than adding more parameters, the rule could be further refined by including more specific conditions. However, extending the detail of the conditions will require a spatial regionalization of their applicable range and, as a consequence, a more complex system. Another limitation is imposed by the accuracy of the data used. The spatial units in the European Soil Database vary in detail depending on the region covered. Soils with very limited extent may not be well represented in areas covered by the database. It would appear that some very organic soils fall into this category.

These limitations in the geographical representation of ground conditions in the database must be considered carefully to avoid misinterpretations when comparing ground with modelled data. This was highlighted during the validation process. The systematic sampling scheme for ground data in England and Wales has by design a tendency to under-estimate the presence of soils with little representation in the area covered. On the other hand, the clustered sampling scheme used in Italy does not provide independent measured values due to auto-correlation of the sample sites. As a result, the areas defined in the database as being soils with large OC content display relatively large ranges of measurements in the ground data located within the spatial units.

Further validations should be performed using measured data from other areas in Europe and for the whole range of land cover types. This will be done when the relevant data sets are made available. There may be scope for further refining the definition of parameters used for the temperature correction. The function parameters were set empirically based on data from very different regions. Additional data could improve the definition of the function, although in its present definition it corresponds to a general relationship of long standing. The research could also be extended to incorporate changes in climatic conditions over longer and different periods, for example 1961-2000 and in decades, for example 1961-70, 1971-80 and 1981-90, thus providing valuable input data for global change modelling. For the purposes of modelling change or future developments, there might also be some merit in adding a correction, based on precipitation and evapo-transpiration data, to account for the effect that moisture may have on crop productivity and OC turnover.

The status of soil OC is known locally in many European countries. However, existing national data must be harmonized and new data collected for regions where OC data are scarce, before a new European map can be produced. The OC map of Europe thus provides the best general picture of the OC/OM status in topsoils throughout the continent at this time.

Acknowledgements

We thank all our colleagues in European Soil Bureau Network for their past and continuing collaboration in providing soil data and expertise for construction of the European Soil Database. We express ours thanks to: the National Soil Resources Institute, Cranfield University and the UK Department of Environment, Fisheries and Rural Affairs, for access to the data from the National Soil Inventory in UK; the Ministero dell’Ambiente, Italy for access to the organic carbon data for Italy; Professor Peter Loveland for fruitful discussions on the results of our research; Dr Peter Smith, University of Aberdeen for reading our manuscript and suggesting improvements; and finally the Institute of Environment and Sustainability, Joint Research Centre, Ispra, Italy, for the support and encouragement in conducting this study.

References

Arrouays, D., Deslais, W. & Badeau, V. 2001. The carbon content of topsoil and its geographical distribution in France. *Soil Use and Management*, **17**, 7-11.

Avery, B.W. & Bascomb, C.L. 1982. *Soil Survey Laboratory Methods*. Soil Survey Technical Monograph No. 6, Harpenden, UK.

Batjes, N.H. 1996. Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, **47**, 151-163.

Batjes, N.H. 1997. A world data set of derived soil properties by FAO-UNESCO soil unit for global modelling. *Soil Use and Management*, **13**, 9-16.

Batjes, N.H. 2002. Carbon and nitrogen stocks in the soils of Central and Eastern Europe. *Soil Use and Management*, **18(4)**, 324-329.

Beare, M.H., Hendrix, P.H. & Coleman, D.C. 1994. Water-stable aggregates and organic matter fractions in conventional and no-tillage soils. *Soil Science Society of America Journal*, **58**, 777-786.

Buckman, H.O. & Brady, N.C. 1960. *The Nature and properties of Soils*. Macmillian, New York.

CEC 1985. *Soil Map of the European Communities, 1:1,000,000*. 124pp. and 7 maps. Office for Official Publications of the European Communities, Luxembourg.

Daroussin, J. & King, D. 1997. A pedotransfer rules database to interpret the Soil Geographical Database of Europe for environmental purposes. In: *The Use of Pedotransfer Functions in Soil Hydrology Research in Europe*. (eds A. Bruand, O. Duval, H. Wosten & A. Lilly). European Soil Bureau Research Report No.3. EUR 17307 EN, pp. 25-40. INRA, Orleans, France.

Easterling, D.R., Thomas, C.P. & Thomas, R.K. 1996. On the development and use of homogenized climate data sets. *Journal of Climate*, **9**, 1429-1434.

EC 2002. Communication of 16 April 2002 from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions: *Towards a Thematic Strategy for Soil Protection* [COM (2002) 179 final]. (At: http://europa.eu.int/scadplus/printversion/en//lvb/l28122.htm; last accessed: 11.11.2004).

FAO-UNESCO 1974. *FAO-UNESCO Soil Map of the World*: *Vol. 1, Legend*. UNESCO, Paris.

FAO-UNESCO-ISRIC. 1990. *FAO-UNESCO Soil Map of the World: Revised Legend*. World Soil Resources Report 60. FAO, Rome.

Heineke, H.J., Eckelmann, W., Thomasson, A.J., Jones, R.J.A., Montanarella, L. & Buckley, B. (eds). 1998. *Land Information Systems: Developments for Planning the Sustainable Use of Land Resources*. European Soil Bureau Research Report No.4, EUR 17729 EN, 545pp. Office for Official Publications of the European Communities, Luxembourg.

Hiederer, R., Jones, R.J.A. & Montanarella, L. In preparation. *A European soil raster data set at scale 1:1,000,000.* Special Publication. European Commission Joint Research Centre, Ispra, Italy.

Hiederer, R. & de Roo, A. 2003. *A European Flow Network and Catchment Data Set*. EUR 20703 EN,40pp.. European Commission Joint Research Centre, Ispra, Italy.

Hiederer, R. 2001. European Catchment Information System for Agri-Environmental Issues. In: *Proceedings of EuroConference 'Link GEO and Water Research'*. Genoa, Italy, 7-9 February 2002. (At http://www.gisig.it/eco-geowater/ (last accessed: 21.12.2004).

Howard, P.J.A., Loveland, P.J., Bradley, R.I., Dry, F.T., Howard, D.M. & Howard, D.C. 1995. The carbon content of soil and its geographical distribution in Great Britain. *Soil Use and Management* **11**, 9-15.

Jones, R.J.A, Hiederer, R., Rusco, E., Loveland, P.J. & Montanarella, L. 2004a. *Topsoil Organic Carbon Content in Europe (Ver. 1.2)*. Special Publication Ispra 2004 No. 72, map in ISO B1 format. European Commission Joint Research Centre, Publication Reference No. S.P.I.04.72.

Jones, R.J.A, Hiederer, R., Rusco, E., Loveland, P.J. & Montanarella, L. 2004b. *Topsoil Organic Carbon Content in Europe (Ver. 1.2). Explanation of Special Publication Ispra 2004 No. 72, map in ISO B1 format.* European Soil Bureau Research Report No.17, EUR 21226, 28pp. Office for Official Publications of the European Communities, Luxembourg.

Jones, R.J.A. & Hollis, J.M. 1996. Pedotransfer rules for environmental interpretations of the EU Soil Database. In: *Soil Databases to support Sustainable Development*. (eds C. Le Bas & M., Jamagne). European Soil Bureau Research Report No.2. EUR 16371 EN, pp.125-133. Office for Official Publications of the European Communities, Luxembourg.

King, D., Daroussin, J. & Tavernier, R. 1994. Development of a soil geographical database from the soil map of the European Communities. *Catena*, **21**, 37-26.

Lettens, S., Van Orshoven, J., van Wesemael, B. & Muys, B. 2004. Soil organic and inorganic carbon contents of landscape units in Belgium derived using data from 1950-to 1970. *Soil Use and Management*, **20**, 40-47.

Madsen, H., Breuning & Jones, R.J.A. 1995. Soil profile analytical database for the European Union. *Danish Journal of Geography*, **95**, 49-57.

McGrath, S.P. & Loveland, P.J. 1992. *The Soil Geochemical Atlas of England and Wales*. Blackie Academic and Professional, Glasgow.

Rusco, E., Jones, R.J.A. & Bidoglio, G. 2001. *Organic matter in the soils of Europe: Present status and future trends*. EUR 20556 EN, 17pp. Office for Official Publications of the European Communities, Luxembourg.

Rusco, E. (In preparation). *Carbon sequestration in Italy*. Research Report, European Soil Bureau, European Commission Joint Research Centre, Ispra, Italy.

Sleutel, S., De Neve, S. and Hofman, G. 2003. Estimates of carbon stock changes in Belgian cropland. *Soil Use and Management*, **19**, 166-171.

Smith, P., Powlson, D.S., Smith, J.U., Falloon, P., & Coleman, K. 2000a. Meeting the UK's climate change commitments: options for carbon mitigation on agricultural land. *Soil Use and Management*, **16**, 1-11.

Smith, P., Powlson, D.S., Smith, J.U., Falloon, P. & Coleman, K. 2000b. Revised estimates of the carbon mitigation potential of UK agricultural land. *Soil Use and Management*, **16**, 293-295.

United States Geological Survey (2003) Global Landcover Characteristics Database. (At http://lpdaac.usgs.gov/glcc/globdoc2_0.asp; last accessed: 11.11.2004).

Van Ranst, E., Thomasson, A.J., Daroussin, J., Hollis, J.M., Jones, R.J.A., Jamagne, M., King, D. & Vanmechelen, L. 1995. Elaboration of an extended knowledge database to interpret the 1:1,000,000 EU Soil Map for environmental purposes. In: *European Land Information Systems for Agro-environmental Monitoring*. (eds D. King, R.J.A. Jones & A.J. Thomasson). EUR 16232 EN, p.71-84. Office for Official Publications of the European Communities, Luxembourg.

Waters, A.G. & Oades, J.M. 1991. Organic matter in water stable aggregates. In: *Advances in Soil Organic Matter Research: The Impact on Agriculture and the Environment*. (ed. W.S. Wilson), pp.163-174. Royal Society of Chemistry, Cambridge.

Table 1 Extract of Pedo-Transfer Rule 21(revised for topsoil organic carbon content)

Condition No.	First Character in Item SOIL *SN1*	Second Character in Item SOIL *SN2*	Third Character in Item SOIL *SN3*	Dominant Surface Textural Class *TEXT*	Land Use Class *USE*	Organic Carbon Class *OC_TOP*
:						
17	L	c	*	2	C	L
18	L	c	*	2	MG	M
:						
68	G	f	m	2	SN	H
69	G	f	m	3	SN	H
:						
77	J	*	g	*	*	M
78	J	*	g	4	SN	H
:						

* any value.

Table 2 Mean ratio of ground data OC_TOP over revised PTR OC_TOP for all land use classes aggregated by AAAT

	AAAT Temperature Class									
Group Mean[1]	2063	2551	3039	3516	3994	4552	4965	5492	5927	6340
Ratio Mean[2]	1.80	1.81	1.73	1.59	1.21	0.80	0.81	0.75	0.82	0.72

[1] Mean AAAT value for data within AAAT class of 500 day degree C width.

[2] OC_TOP_{GRD} : OC_TOP_{PTR}..

Table 3 Mean organic carbon content for England and Wales for catchments (>1000km^2) and NUTS Level 2

Catchment Name (>1000km^2)	**Ground Sample Points**	**Mean OC_TOP from Ground Sample**	**Mean Model OC_TOP at Ground Sample**	**Mean Model OC_TOP for NUTS unit**	**Region Name**	**Ground Sample Points**	**Mean OC_TOP from Ground Sample**	**Mean Model OC_TOP at Ground Sample**	**Mean Model OC_TOP for NUTS unit**
	n.	*%*	*%*	*%*		*n.*	*%*	*%*	*%*
Ouse	407	9.4	9.6	8.7	Tees Valley, Durham	109	11.0	12.4	11.0
Thames, above Lea	384	3.8	2.9	2.9	Northumberland, Tyne, Wear	196	13.1	12.6	12.9
Severn	401	5.1	4.7	4.6	Cumbria	254	13.9	14.3	14.2
Trent	377	4.9	3.9	4.0	Cheshire	82	5.1	4.3	4.1
Great Ouse	291	4.1	3.3	3.4	Greater Manchester	41	8.2	7.1	7.5
Wye	163	6.3	7.7	7.9	Lancashire	106	9.3	11.0	10.3
Nene	119	4.6	4.2	4.2	Merseyside	16	6.4	6.8	4.3
Avon	115	5.3	3.2	3.1	East Riding, North Lincolnshire	133	2.7	3.0	3.1
Witham	97	3.9	3.1	3.0	North Yorkshire	324	10.3	10.4	9.6
Tyne	95	19.8	19.8	19.8	South Yorkshire	47	7.5	7.9	7.3
Eden	90	13.4	14.3	14.0	West Yorkshire	68	11.1	11.4	8.7
Mersey	72	10.4	9.6	9.7	Derbyshire, Nottinghamshire	181	5.9	5.2	5.4
Avon	77	5.3	2.7	2.9	Leicestershire, Rutland, Northamptonshire	190	3.7	2.8	2.9
R. Dee	76	12.1	10.2	10.1	Lincolnshire	232	3.5	3.3	3.2
Welland	71	3.7	3.6	3.3	Herefordshire, Worcestershire, Warwickshire	225	3.0	2.8	2.7
Parrett	60	5.4	3.5	3.4	Shropshire, Staffordshire	234	4.9	4.3	4.3
Medway	60	3.7	3.4	3.1	West Wales, The Valleys	498	11.8	11.4	10.9
Exe	53	4.2	4.8	4.5	West Midlands	18	4.4	2.5	2.5
Weaver	51	5.0	5.1	4.3	East Anglia	485	3.7	3.5	3.3
Ribble	51	10.6	13.0	12.2	East Wales	295	9.1	10.0	10.0
Yare	57	1.5	1.5	2.0	Essex	132	3.4	2.4	2.4
River Lea	45	2.3	2.5	2.7	Inner London	2	6.8	3.2	3.2
Usk	49	7.3	9.1	9.7	Outer London	26	4.3	3.0	3.1
River Towy	54	10.5	11.5	11.4	Surrey, East, West Sussex	205	3.7	3.4	3.3
River Tees	52	16.0	17.4	17.3	Bedfordshire, Hertfordshire	112	2.7	2.4	2.5
Test	46	6.2	3.2	3.2	Hampshire, Isle Of Wight	158	5.0	3.3	3.3
Taw	44	5.7	6.7	6.0	Kent	139	3.7	3.0	2.8
Wear	42	11.4	12.7	10.9	Dorset, Somerset	226	5.8	4.1	4.0
Lune	42	19.0	17.4	16.1	Gloucestershire, Wiltshire , North Somerset	285	4.9	3.0	3.0
Arun	39	3.5	3.9	3.8	Cornwall, Isles Of Scilly	141	5.2	5.3	5.2
					Devon	249	6.5	5.8	5.6
					Berkshire, Buckinghamshire, Oxfordshire	220	3.6	2.7	2.7
Total / **Mean***	*3580*	**6.7**	**6.3**	**6.1**	*Total* / **Mean***	*5629*	**6.7**	**6.3**	**6.1**

* Mean: area-weighted average of values aggregated to relative spatial unit.

Table 4 Mean organic carbon content for Italy by NUTS Level 2

Region Name	**Ground Sample Points** n	**Mean OC_TOP from Ground Sample** %	**Mean Model OC_TOP at Ground Sample Points** %	**Mean Model OC_TOP for NUTS unit** %	**Region Name**	**Ground Sample Points** n.	**Mean OC_TOP from Ground Sample** %	**Mean Model OC_TOP at Ground Sample Points** %	**Mean Model OC_TOP for NUTS unit** %
Piemonte	327	1.2	1.4	3.5	Marche	145	0.8	0.9	1.8
Valle D'Aosta	7	2.3	3.0	5.3	Lazio	295	1.4	1.3	2.0
Liguria	17	1.1	1.8	3.3	Abruzzo	185	0.8	1.1	3.0
Lombardia	198	1.2	1.4	3.1	Molise	117	1.2	1.4	2.3
Trentino-Alto Adige	21	1.9	2.9	5.5	Campania	157	1.7	1.3	1.8
Veneto	294	1.4	1.5	2.5	Puglia	546	1.3	1.0	1.2
Friuli-Venezia Giulia	126	1.6	1.2	2.8	Basilicata	210	1.0	1.1	1.9
Emilia-Romagna	562	1.4	1.6	2.1	Calabria	152	0.9	1.0	1.6
Toscana	214	0.9	1.2	2.2	Sicilia	594	1.1	0.8	1.2
Umbria	169	1.3	1.3	2.1	Sardegna	164	1.1	1.0	1.7
					Total / Mean	**4500**	**1.2**	**1.2**	**2.4**

Table 5 Soil Subgroup Codes and soil names for comparing modelled OC values with ground data in England & Wales, and Italy (see Figures 4 and 6). The FAO soil subgroup code is as used on the The Soil Map of the European Communities (CEC 1985)

Code	Soil Subgroup Name (FAO, 1974)	WRB Reference Group (FAO, 1998)
Bc	Chromic Cambisol	Chromic Cambisol
Bd	Dystric Cambisol	Dystric Cambisol
Bds	Spodo-Dystric Cambisol	Endo-skeletic Umbrisol
Be	Eutric Cambisol	Eutric Cambisol
Bea	Ando-Eutric Cambisol	Eutri-andic Cambisol
Bec	Calcaro-Eutric Cambisol	Calcaric Cambisol
Bef	Fluvi-Eutric Cambisol	Eutri-fluvic Cambisol
Bk	Calcic Cambisol	Haplic Calcisol
Bv	Vertic Cambisol	Vertic Cambisol
Bvc	Calcaro-Vertic Cambisol	Calcari-vertic Cambisol
Bgc	Calcaro-Gleyic Cambisol	Calcari-gleyic Cambisol
Bgg	Stagno-Gleyic Cambisol	Stagnic Cambisol
E	Rendzina	Leptosol
Id	Dystric Lithosol	Dystric Leptosol
Gds	Stagno-Dystric Gleysol	Dystri-stagnic Gleysol
Ges	Stagno-Eutric Gleysol	Eustri-stagnic Gleysol
Gh	Humic Gleysol	Humic Gleysol
Gm	Mollic Gleysol	Mollic Gleysol
Gmf	Molli-Fluvic Gleysol	Fluvi-mollic Gleysol
Jcg	Gleyo-Calcaric Fluvisol	Calcari-gleyic Fluvisol
Jeg	Gleyo-Eutric Fluvisol	Eutri-gleyic Fluvisol
Lc	Chromic Luvisol	Chromic Luvisol
Lg	Gleyic Luvisol	Gleyic Luvisol
Lgp	Plano- Gleyic Luvisol	Gleyic Luvisol
Lk	Calcic Luvisol	Calcic Luvisol
Lgs	Stagno-Gleyic Luvisol	Stagnic Luvisol
Lo	Orthic Luvisol	Haplic Luvisol
Od	Dystric Histosol	Dystric Histosol
Oe	Eutric Histosol	Eutric Histosol
Pg	Gleyic Podzol	Gleyic Podzol
Pgs	Stagno-Gleyic Podzol	Stagnic Podzol
Po	Orthic Podzol	Haplic Podzol
Pp	Placic Podzol	Placic Podzol
Q	Arenosol	Arenosol
Qc	Cambic Arenosol	Haplic Arenosol
Ql	Luvic Arenosol	Lamellic Arenosol
Rc	Calcaric Regosol	Calcaric Regosol
Re	Eutric Regosol	Eutric Regosol
Th	Humic Andosol	Umbric Andosol
Vc	Chromic Vertisol	Chromic Vertisol
U	Ranker	Leptosol

Figure Captions

Figure 1 General procedure for calculating topsoil organic carbon content

Figure 2: Temperature correction coefficient for organic carbon content

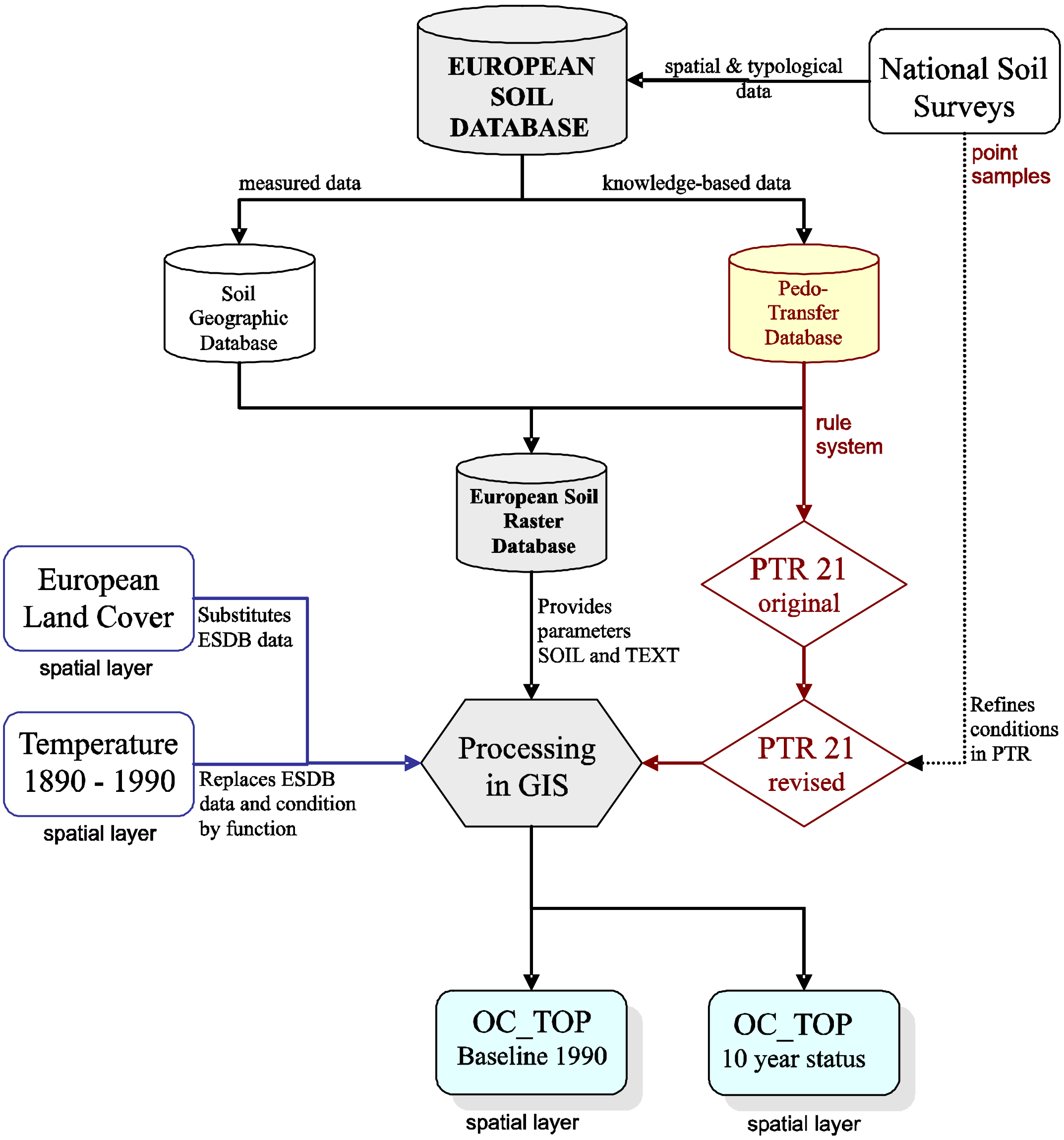
Figure 3 Organic carbon content (%) in the surface horizon of soils in Europe

Figure 4 Ground sample confidence intervals (95%) for topsoil organic carbon content in England and Wales by FAO Soil class (all land cover, semi-natural and cultivated - for explanation of FAO soil subgroup codes, see Table 5)

Figure 5 Relation of topsoil organic carbon between ground and modelled data for England and Wales for CIS primary catchments (>1000km^2) and NUTS Level 2 units

Figure 6 Topsoil organic carbon content in Italy by soil class (cultivated land use class only – for explanation of FAO soil subgroup codes, see Table 5)

783



784
785
786 **Figure 1**
787

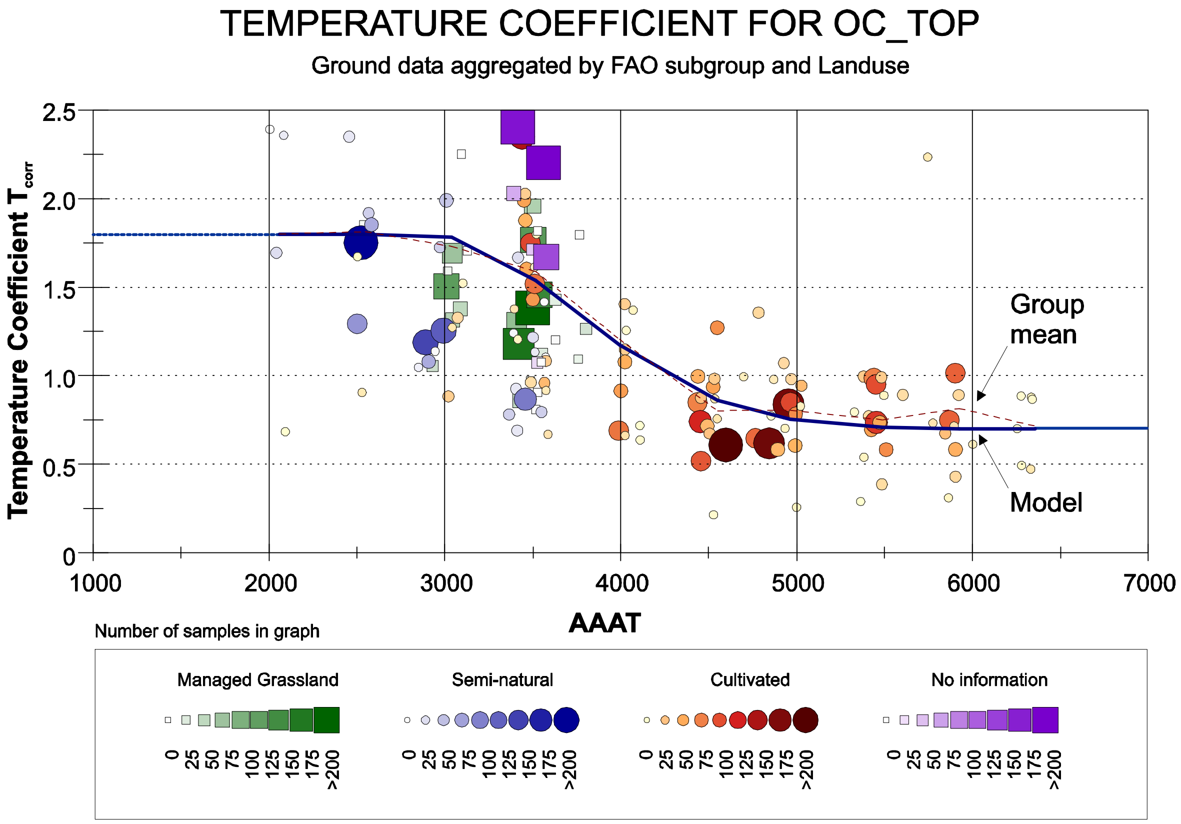

TEMPERATURE COEFFICIENT FOR OC_TOP
Ground data aggregated by FAO subgroup and Landuse
Temperature Coefficient T_{corr}
2.5
2.0
1.5
1.0
0.5
0
1000
2000
3000
4000
5000
6000
7000
AAAT
Group mean
Model
Number of samples in graph
Managed Grassland
Semi-natural
Cultivated
No information
0
25
50
75
100
125
150
175
>200

Figure 2

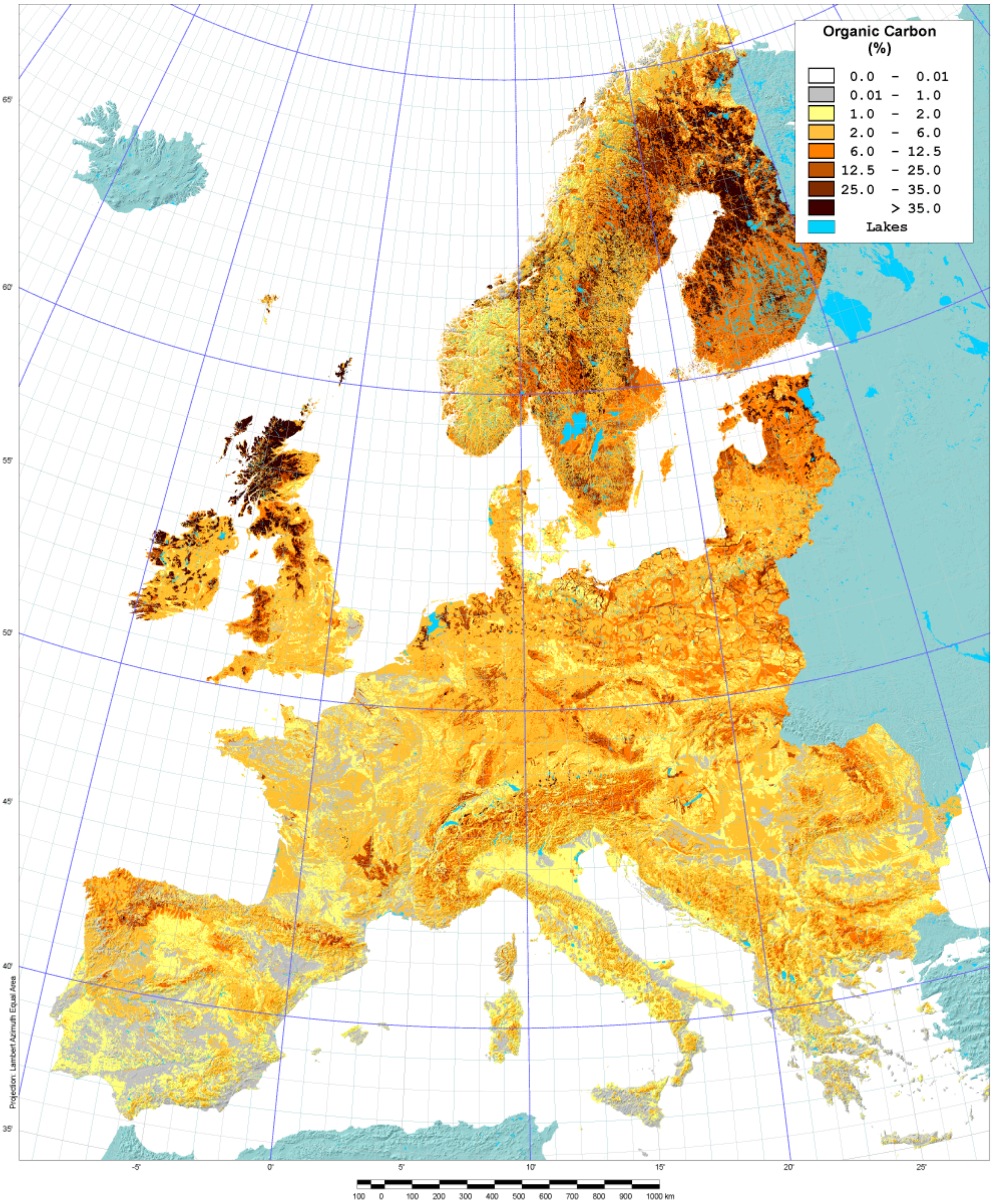

Organic Carbon (%)
0.0 - 0.01
0.01 - 1.0
1.0 - 2.0
2.0 - 6.0
6.0 - 12.5
12.5 - 25.0
25.0 - 35.0
> 35.0
Lakes
Projection: Lambert Azimuth Equal Area
65°
60°
55°
50°
45°
40°
35°
-5°
0°
5°
10°
15°
20°
25°
100 0 100 200 300 400 500 600 700 800 900 1000 km

Figure 3

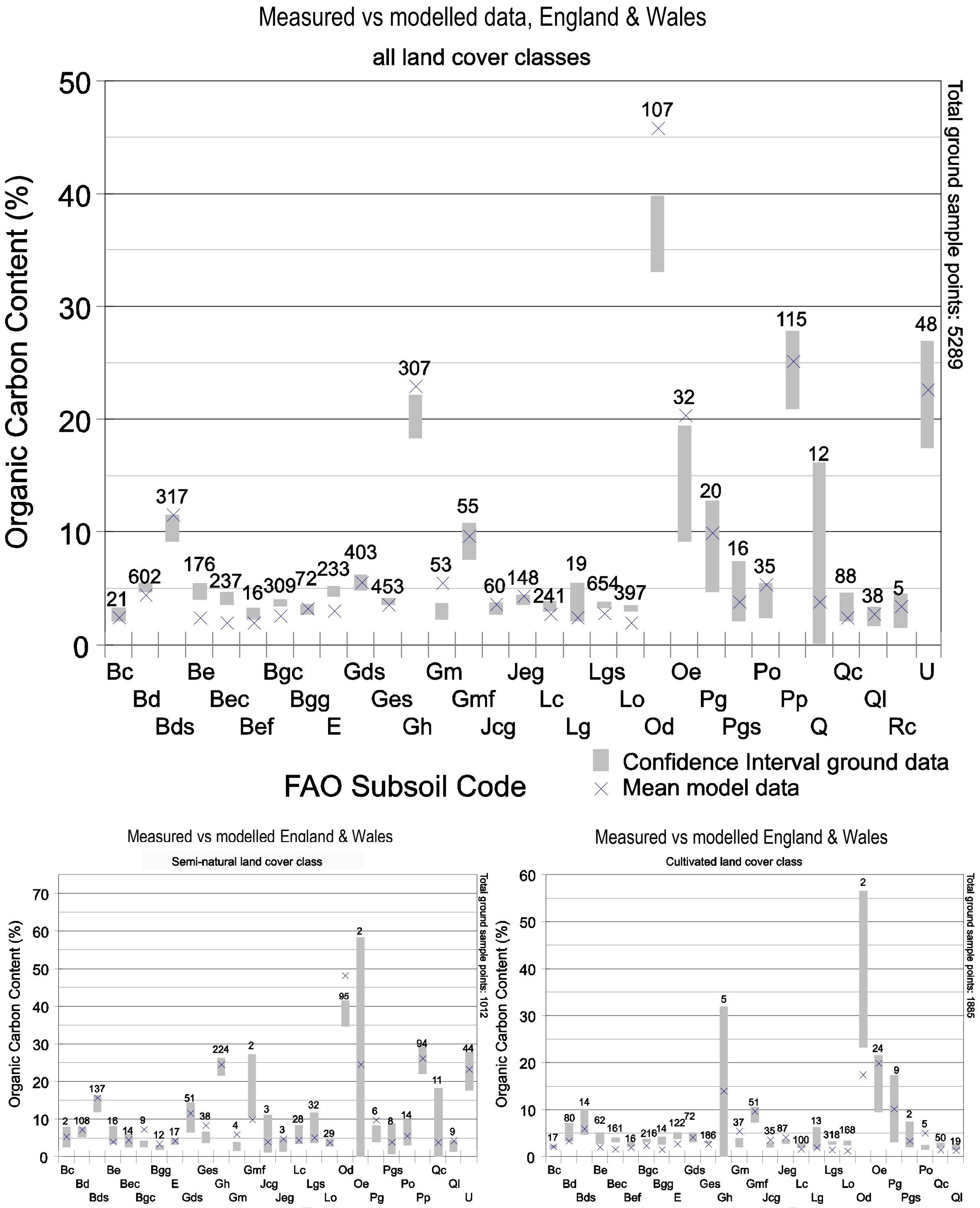

Measured vs modelled data, England & Wales
all land cover classes
Organic Carbon Content (%)
0
10
20
30
40
50
Total ground sample points: 5289
21
602
317
176
237
16
309
72
233
403
453
307
53
55
60
148
241
19
654
397
107
32
20
16
35
115
12
88
38
5
48
Bc
Bd
Bds
Be
Bec
Bef
Bgc
Bgg
E
Gds
Ges
Gh
Gm
Gmf
Jcg
Jeg
Lc
Lg
Lgs
Lo
Od
Oe
Pg
Pgs
Po
Pp
Q
Qc
Ql
Rc
U
FAO Subsoil Code
Confidence Interval ground data
Mean model data
Measured vs modelled England & Wales
Semi-natural land cover class
Organic Carbon Content (%)
0
10
20
30
40
50
60
70
Total ground sample points: 1012
2
108
137
16
14
9
12
17
51
38
224
4
2
3
3
28
32
29
95
2
6
8
14
94
11
9
44
Bc
Bd
Bds
Be
Bec
Bgc
Bgg
E
Gds
Ges
Gh
Gm
Gmf
Jcg
Jeg
Lc
Lgs
Lo
Od
Oe
Pg
Pgs
Po
Pp
Qc
Ql
U
FAO Subsoil Code
Confidence Interval ground data
Mean model data
Measured vs modelled England & Wales
Cultivated land cover class
Organic Carbon Content (%)
0
10
20
30
40
50
60
Total ground sample points: 1885
17
80
14
62
161
16
216
14
122
72
186
5
37
51
35
87
100
13
318
168
2
24
9
2
5
50
19
Bc
Bd
Bds
Be
Bec
Bef
Bgc
Bgg
E
Gds
Ges
Gh
Gm
Gmf
Jcg
Jeg
Lc
Lg
Lgs
Lo
Od
Oe
Pg
Pgs
Po
Qc
Ql
FAO Subsoil Code
Confidence Interval ground data
Mean model data

Figure 4

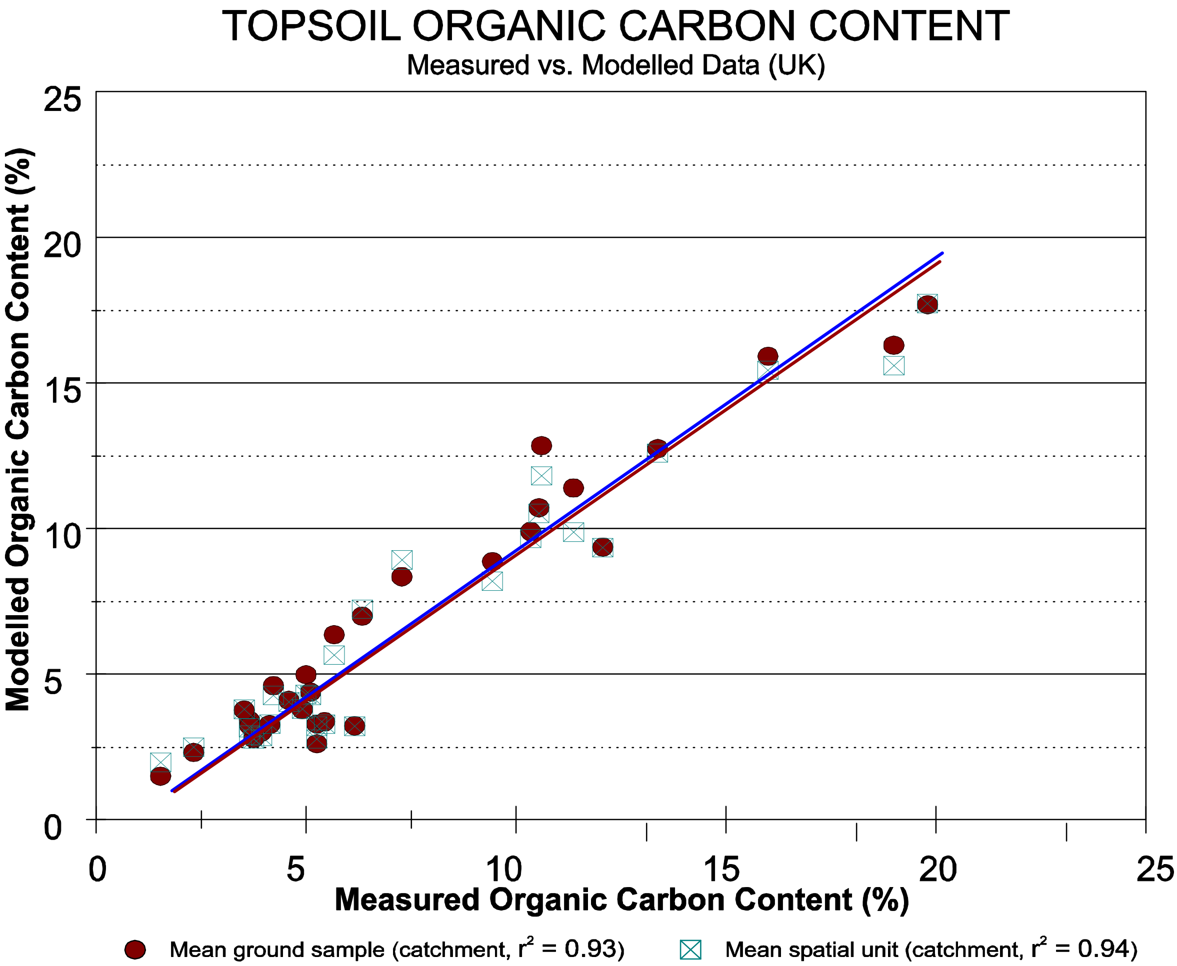

TOPSOIL ORGANIC CARBON CONTENT
Measured vs. Modelled Data (UK)
Modelled Organic Carbon Content (%)
0
5
10
15
20
25
Measured Organic Carbon Content (%)
0
5
10
15
20
25
Mean ground sample (catchment, r^2 = 0.93)
Mean spatial unit (catchment, r^2 = 0.94)

Figure 5

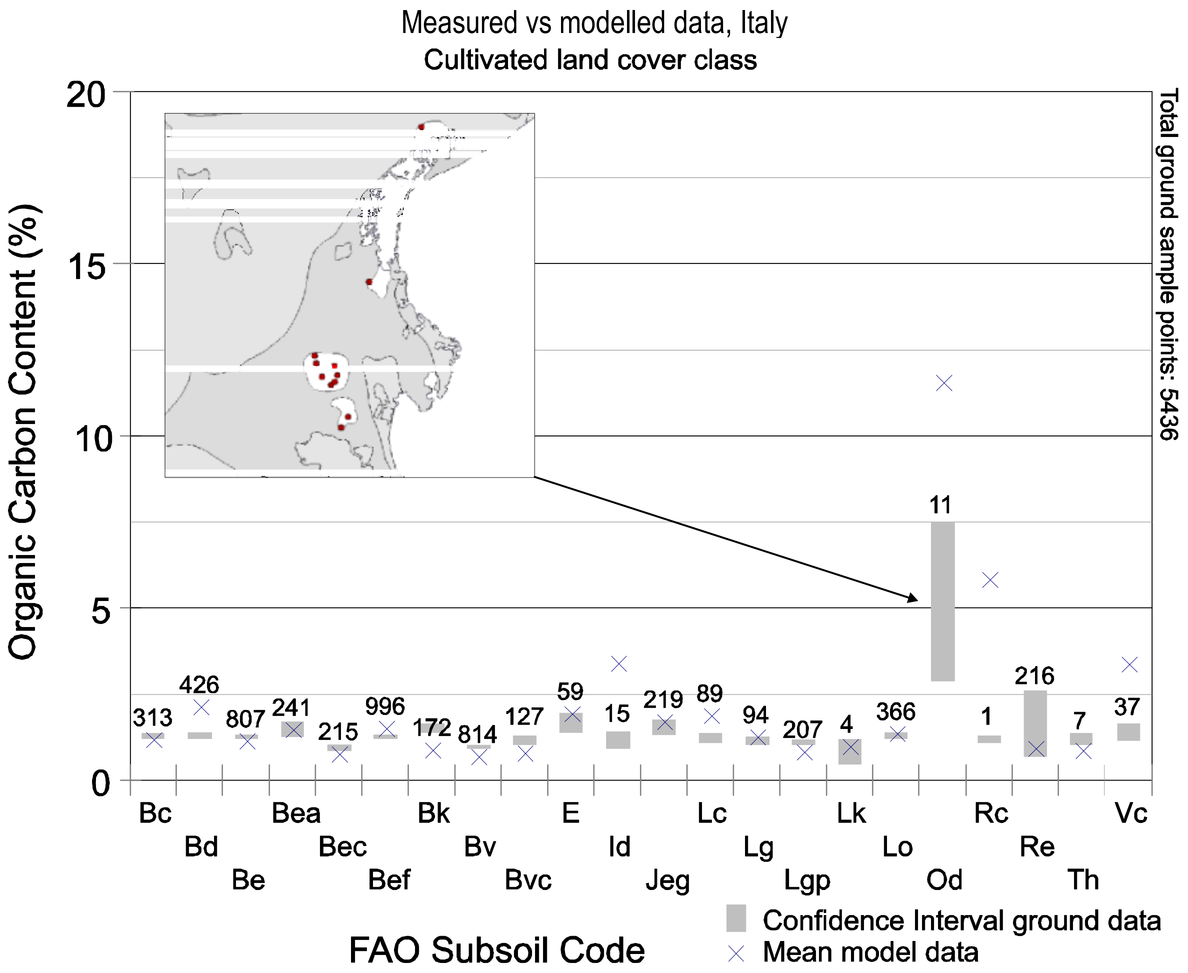

Measured vs modelled data, Italy
Cultivated land cover class
Organic Carbon Content (%)
20
15
10
5
0
Total ground sample points: 5436
313
426
807
241
215
996
172
814
127
59
15
219
89
94
207
4
366
11
1
216
7
37
Bc
Bd
Be
Bea
Bec
Bef
Bk
Bv
Bvc
E
Id
Jeg
Lc
Lg
Lgp
Lk
Lo
Od
Rc
Re
Th
Vc
FAO Subsoil Code
Confidence Interval ground data
Mean model data

Figure 6