Food and Agriculture Organization
of the United Nations

cmcc

Global Soil Organic Carbon Database

(at 30 arcsec)

GEOCARBON *Toward an Operational Global Carbon Observing System*

Global Soil Organic Carbon Database

(at 30 arcsec)

Published by
the Food and Agriculture Organization of the United Nations
and
Euro-Mediterranean Center on Climate Change Foundation

The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO), or of Euro-Mediterranean Center on Climate Change concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO, or CMCC in preference to others of a similar nature that are not mentioned.

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO, or CMCC.

ISBN 978-92-5-109762-5
© FAO, 2017

FAO encourages the use, reproduction and dissemination of material in this information product. Except where otherwise indicated, material may be copied, downloaded and printed for private study, research and teaching purposes, or for use in non-commercial products or services, provided that appropriate acknowledgement of FAO as the source and copyright holder is given and that FAO's endorsement of users' views, products or services is not implied in any way.

All requests for translation and adaptation rights, and for resale and other commercial use rights should be made via www.fao.org/contact-us/licence-request or addressed to copyright@fao.org.

FAO information products are available on the FAO website (www.fao.org/publications) and can be purchased through publications-sales@fao.org.

CONTENTS

Figures iv
Tables v
Summary vi
Acknowledgements vii
Acronyms vii

1. Introduction 1

2. Materials and methods 5
2.1 Soil organic carbon stock database 5
2.2 Database processing and SOC stock calculation 6
2.3 Bulk density 9
2.4 Soil type 10
2.5 Land cover 12
2.6 Homogeneous Land Units and analysis area 12
2.7 Quality assessment and uncertainty 14
2.8 Other global datasets 16

3. Results 19
3.1 SOC Geodatabase 19
3.1.1 Identification of geodatabase layers 20
3.1.2 Final outputs 21

4. Discussion 25

5. Next steps and recommendations 31

References 33

FIGURES

1 Global distribution of all soil profiles in database 7

2 Distribution of al profiles used to calculate Top SOC stock 8

3 Distribution of al profiles used to calculate Sub SOC stock 8

4 Total number of soil layers with and without bulk density measurements 9

5 HWSD dominant Soil Types 10

6 Distribution of "mixed soils" 11

7 Distribution of dominant GLC -SHARE Land Cover Database 13

8 SOC spatialization methodology 13

9 Spatial intersection 14

10 Uncertainty factors 14

11 Distribution of uncertainty for Top SOC 15

12 Distribution of uncertainty for Sub SOC 15

13 JRC's "Global Soil Organic Carbon Estimates" (Top SOC) 16

14 Comparison 1 Left JRC Database Right Soil Organic Carbon 30 arc sec Global Database 17

15 Comparison 2 Left JRC Database Right Soil Organic Carbon 30 arc sec Global Database 17

16 Metadata 20

17 SOC 30 arcsec Global database structure 21

18 Distribution of Top soil organic Carbon (Continuous data is aggregated in classes for illustrative purpose) 22

19 Distribution of Sub soil organic Carbon (Continuous data is aggregated in classes for illustrative purpose) 22

20 Classes of Top SOC in percentage. In pie chart global SOC data are aggregated for explicative purpose in Quantile Classes (each class has the same number of observations) 26

21 Classes of Sub SOC in percentage. In pie chart global SOC data are aggregated for explicative purpose in Quantile Classes (each class has the same number of observations) 27

22 Red rimmed: decreasing spatial accuracy using climate data 32

TABLES

Table 1 ______________ 5

Table 2 ______________ 6

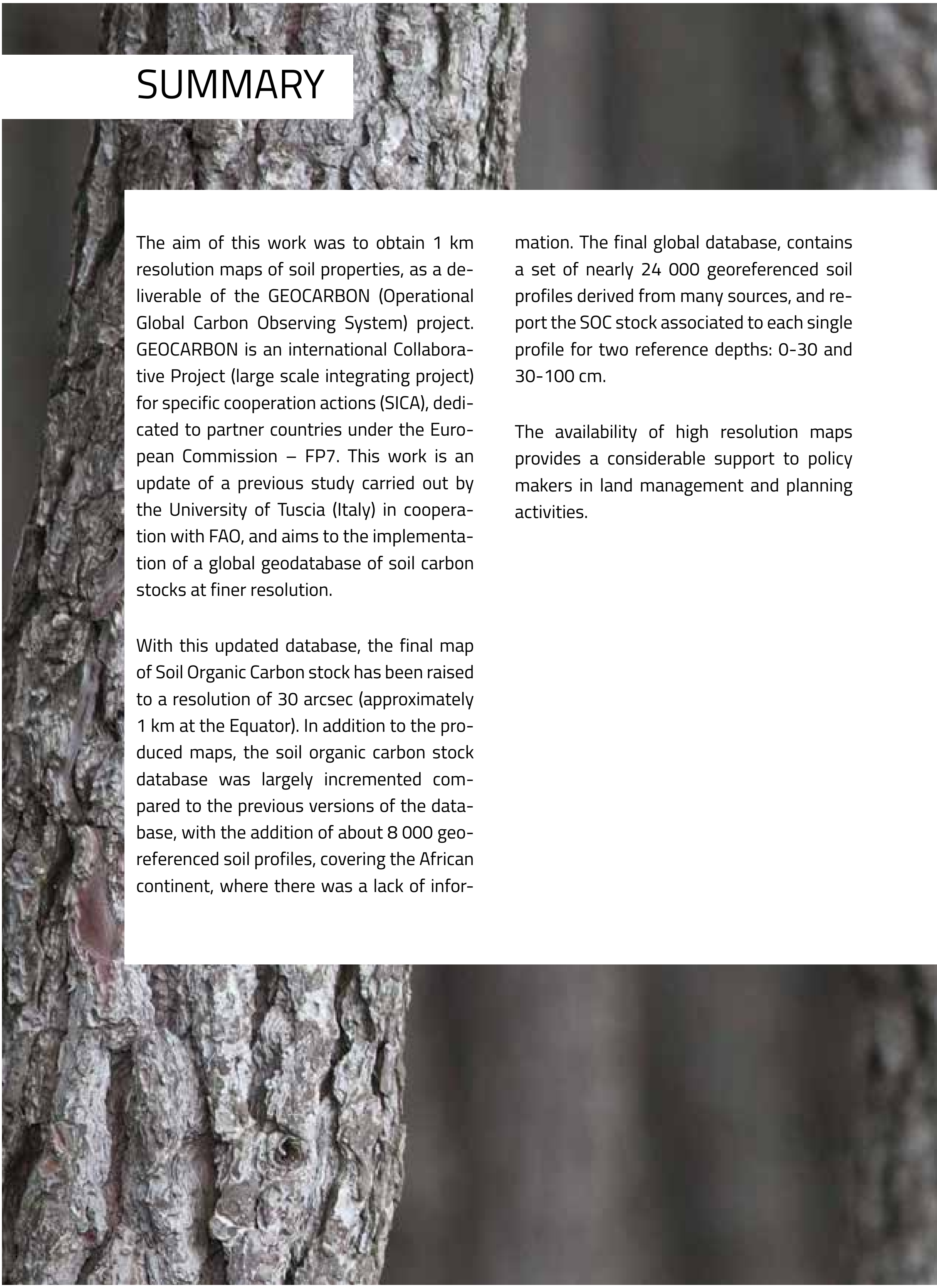
Table 3 _____________ 11

Table 4 _____________ 12

Table 5 _____________ 23

Table 6 _____________ 28

Table 7 _____________ 29



SUMMARY

The aim of this work was to obtain 1 km resolution maps of soil properties, as a deliverable of the GEOCARBON (Operational Global Carbon Observing System) project. GEOCARBON is an international Collaborative Project (large scale integrating project) for specific cooperation actions (SICA), dedicated to partner countries under the European Commission – FP7. This work is an update of a previous study carried out by the University of Tuscia (Italy) in cooperation with FAO, and aims to the implementation of a global geodatabase of soil carbon stocks at finer resolution.

With this updated database, the final map of Soil Organic Carbon stock has been raised to a resolution of 30 arcsec (approximately 1 km at the Equator). In addition to the produced maps, the soil organic carbon stock database was largely incremented compared to the previous versions of the database, with the addition of about 8 000 georeferenced soil profiles, covering the African continent, where there was a lack of information. The final global database, contains a set of nearly 24 000 georeferenced soil profiles derived from many sources, and report the SOC stock associated to each single profile for two reference depths: 0-30 and 30-100 cm.

The availability of high resolution maps provides a considerable support to policy makers in land management and planning activities.

ACKNOWLEDGEMENTS

This document was the results of the collaborative effort of FAO Natural Land Resources Division (NRL), the University of Tuscia and the CMCC Foundation (Euro-Mediterranean Center on Climate Change) in the frame of their respective organizational and institutional activities for research and/or provision of soil information, including mapping of soil resources, in support to sustainable development.

The team, led from John Latham (FAO-NRL), has involved also: Freddy Nachergaele (FAO) and Matieu Henry (FAO), Rodrigo Vargas (Univ. of Delaware), Sergio Noce (CMCC), Monia Santini (CMCC), Antonio Bombelli (CMCC) and Tommaso Chiti (University of Tuscia).

This work was financially supported jointly by FAO and the European Commission, through the FP7 Project GEOCARBON.

ACRONYMS

BD	Bulk Density
DB	Data Base
GIS	Geographic Information System
HLUs	Homogeneous Land Units
IIASA	International Institute of Applied Systems Analysis
IPCC	Intergovernamental Panel on Climate Change
ISRIC	International Soil Reference and Information Centre
ISIS	ISRIC Soil Information System
NRCS	Natural Resources Conservation Service
NSCN	National Soil Carbon Network
SOC	Soil Organic Carbon
SOTER	Soil and Terrain
USDA	United States Department of Agriculture


© FAO/ K. Boldt

1. INTRODUCTION

Globally, estimates of soil organic carbon (SOC) in the upper meter of soil are ranging from about 1 000 Pg C to almost 2,500 Pg C depending on the database used for the calculation (Scharlemann *et al.* 2014; Hiederer and Köchi 2011; Jobbàgy and Jackson 2000), and most of this SOC is stored in forest soils (Dixon *et al.* 1994; Batjes 1996). Given the ongoing climate change, accurate knowledge about the quantity of carbon stored in soils is important because the large quantity of CO_2 emitted from soil as a consequence of land use change, both through human activities (e.g. agriculture) or natural hazards (e.g. flooding, landslides, erosion). Carbon emissions resulting from land use and land cover changes are the second largest anthropogenic source of C emissions to the atmosphere, after emissions from fossil fuel combustion. The carbon balance of terrestrial ecosystems can be changed markedly by the direct impact of anthropogenic activities, including deforestation, biomass burning, land use change, and environmental pollution, which release trace gases that enhance the 'greenhouse effect' (Intergovernmental Panel on Climate Change, IPCC 1990).

A net carbon loss from soils adds to the increase in the atmospheric CO_2 concentration, probably leading to higher global temperatures (IPCC 2001), which, in turn, could accelerate decomposition of SOC (Cox *et al.* 2000; Jones *et al.* 2005). On the other hand, a net soil CO_2 sequestration could help to mitigate the greenhouse effect and to improve soil quality. Hence, soil carbon sequestration has a great global mitigation potential, both in terms of enhancement of C sinks and reduced C emissions (Smith *et al.* 2007, 2008).

In this purpose, identifying areas suitable for projects aimed to reduce emissions trough SOC sequestration, is of utmost importance for both developed and developing countries.

This report describes the sources and procedures used to develop a global database for SOC content. The spatial data for the analysis were derived from various databases reporting necessary information to calculate the SOC stock, *i.e.* C concentration, bulk density (BD), rock fragment, horizon depth. Besides, the global SOC database was integrated with consistent available Land Use/Cover and Soil Type databases. The main assumption is that carbon content is related to soil type and land use/cover; these two factors create similar condition

for SOC accumulation. The global SOC database is intended to provide an indication at global and regional levels, useful to support large scale C budgets and evaluations. Then, the database can be useful to support decision making and planning.

The approaches and methodologies provided by using Geographic Information Systems (GIS) software, play a crucial role in environment management. In this purpose, an approach combining different kind of data (*i.e.* different spatial resolution or format), and allowing to make queries (spatial or non-spatial) and spatial analysis processes, is of outmost importance to obtain new data.

As in other environmental sectors, the assessment of the soil organic carbon pool at different scales requires remote sensing, GIS and modeling (Lal 2002), such as the activities carried out within the framework of this study.


© FAO/ Giulio Napolitano

2. MATERIALS AND METHODS

2.1 Soil organic carbon stock database

The global database on SOC stock is a relational database compiled using MSAccess (R). It handles data on:

- soil classification and site features;
- source of data;
- SOC stock for the 0-30 and 30-100 cm depth of mineral soil
- information on the vegetation and/or land use to which each single profile refers to.

The attributes considered in the FAO database are listed in Table 1.

The data used to create the database were found after a research to identify the available soil organic carbon databases. The World Soil Information (ISRIC) provided the main sources of information namely version 3.1 of

Site Data	Profile Data
LOCATION: Country (ISO Code) Latitude (N/S, deg/min/sec) Longitude (W/E, deg/min/sec)	GENERAL: SOC stock 0-30 cm SOC stock 30-100 cm
SOURCE OF DATA: Database Identifier	SOIL CLASSIFICATION, e.g: WRB 2006 Reference soil group USDA Soil Taxonomy FAO74/88 classification
GENERAL SITE DATA: Land cover/use Vegetation	

Table 1. List of the FAO database attribute data.

the WISE3 database, the most updated version of the database already used by Batjes (1996) for quantifying the global carbon budget. Besides, several Soil and Terrain (SOTER) databases for different countries were also obtained from FAO-ISRIC: China; Senegal, Gambia, South Africa, Upper Tana (Kenya), Congo, Burundi, Rwanda, Cuba, Argentina and Tunisia. Another relevant source of information for the African continent, was obtained from the last update of the FAO-ISRIC Africa Soil Profile Database ver 1.1 (Leenaars 2013), with about 8 000 soil profiles). Besides, the Natural Resources Conservation Service (USDA-NRCS) and the United States Geological Service databases (Harden 2008), obtained through the National Soil Carbon Network (NSCN 2011), storing mostly data about soils of the United States with Alaska, was also included (Soil Survey Staff 2011). A database specific for European soils, SPADE M version 2 (Hiederer *et al.* 2006), was obtained from the European Soil Bureau, while from IIASA was obtained the "Land Resources of Russia" a database specific for Russian soils (Stolbovoy *et al.* 2002). All of these databases were already used in many publications (Jobbagy and Jackson 2000; Batjes and Dijkshoorn 1999; Batjes 1997) and widely accepted in the scientific community. Besides, this work was enriched with the addition of data never published before such those contained in the database provided by the University of Tuscia, related to soil profiles mostly located in Europe and in a minor part in Central and Western Africa, and in the database provided by FAO holding profiles from North and East Africa (Libia, Somalia) and South America (Bolivia). These two latter databases represent about 7% (1 655 profiles) of the total number of soil profiles used in this study, 23 871. All the available databases used in this work, with the information about the number of profiles are summarized in Table 2. The distribution of the soil profiles in the different continents is showed in Figure 1.

Source	Profiles
ISRIC-FAO	16 806
USDA-NRCS	4 764
Tuscia DB	1 518
Spade	405
Russia	241
FAO DB	137
Total	**23 871**

*Table 2. * Source and number of profiles for each of the databases stratified by continent.*

2.2 Database processing and SOC stock calculation

Strict criteria have been defined for accepting profiles into the database of actual SOC stock:

** include SOTER China (1 403 profiles), Tunisia (39 profiles), Cuba (22 profiles), Argentina (166 profiles), Burundi-Congo Rwanda (153 profiles), Upper Tana (82 profiles), South Africa (566 profiles), Senegal-Gambia (59 profiles), ISRIC Western Europe (486 profiles), Libia (45 profiles), Nepal (125 profiles), WISE3 (6 078 profiles).*

- completeness and apparent reliability of data;
- traceability of source data;
- classificability in the original and revised FAO or Soil Taxonomy nomenclature;
- geo- referencing of the data.

Upon their entry into our database, the data have been screened for inconsistencies using visual and automated procedures. These procedures include the checks on consistency of C concentrations and bulk density data from single soil horizons, to verify that the values were comprised within a realistic range for most soils. To accomplish these purposes, each database was checked to verify the presence or derivability of all the parameters necessary to calculate the SOC stock. Other parameters such as geographical coordinates, soil classification (FAO or USDA soil taxonomy), land use and vegetation necessary to locate and characterize each soil profile were also included in the database. All the profiles with missing C concentrations, rock fragment contents or the geographic coordinates were eliminated.

When the bulk density was the only missing parameter, the profiles were maintained and the bulk density calculated using a pedotransfer function as explained in the next section. Bulk density higher than 1.95 Mg m^3 were excluded from the database, since proctor compaction tests on topsoil and subsoil samples representing a range of textural classes, suggest that BD are unlikely to exceed values of 1.95 Mg m^3 for any textures (Hollis and Woods 1989). Therefore to avoid unrealistic BD predictions an upper maximum limit is set to 1.95 Mg m^3.

The analytical methods used in the different databases to determine the parameters necessary to calculate the SOC stock are fully comparable. The comparability of the USDA-NRCS and of the ISRIC Soil Information System (ISIS) databases has been already documented (Vogel 1994), while for the new data that were never published before the same criteria for data selection as described in the WISE3 report was used (Batjes 2008). The SOC stock was calculated for each soil layer according to equation 1 (Boone *et al.* 1999):

Figure 1. Global distribution of all soil profiles in database.

$$y = a\ b\ c\ d \quad [1]$$

where

y = SOC stock per unit area (Mg C /ha)

a = C concentration in the soil sample (kg C kg^{-1} soil)

b = soil bulk density (Mg soil m^3)

c = depth of the layer (cm)

d = percent in mass of rock fragments [1- (% rock volume/100)].

Then taking into account the thickness of the horizons, the SOC stocks were normalized for the 0-30 cm (TOP SOC) and 30-100 cm (SUB SOC) cm depths. All the databases were merged together so to obtain a single database where the SOC stock from the 0-30 and 30-100 cm depths of each single profile, is connected with the geographic coordinates, the soil classification, the vegetation and the land use. After connection with spatial information two single geographic vector layers were obtained, 23 871 profiles for Top SOC and 17 503 for Sub SOC (see Figures 2 and 3).

Figure 2. Distribution of all profiles used to calculate Top SOC stock.

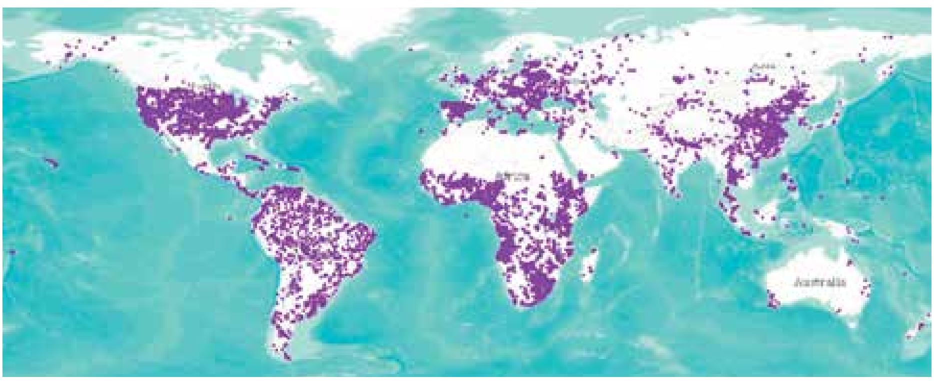


Figure 3. Distribution of all profiles used to calculate Sub SOC stock.

2.3 Bulk density

Bulk densities measurements were not always available for all the layers of the profiles considered in this study (Figure 4). Since the amount of layers with no BD data was considerable, we used some pedotransfer functions (PTF's) to estimate BD. Following a recent work on the global soil organic carbon (Hiederer and Köchi 2011), which was using some of the free available SOC databases (e.g. Wise 3, Spade) we also used for this study, we considered the same approach. In particular, Hiederer and Köchi (2011) calibrated some PTF's for WISE3 and Spade, that represent the largest databases used for this work. For WISE3 database, the selected regression model was different for the layers with a C content higher or lower than 12% (see Hiederer and Köchi 2011 for details). In particular, for layers with a C content lower than 12%, either in the topsoil or subsoil, it was used equation 2 (R^2 = 0.46) which resulted in a value for organic matter (OM = 0.15 g cm^3) close to those reported by Rawls (1983) and Ruehlmann & Körschens (2010) for global data:

$$BD = \exp^{(a*OC+b)} \qquad [2]$$

where

BD = bulk density for topsoil or subsoil

a = -0.034 b = 0.100

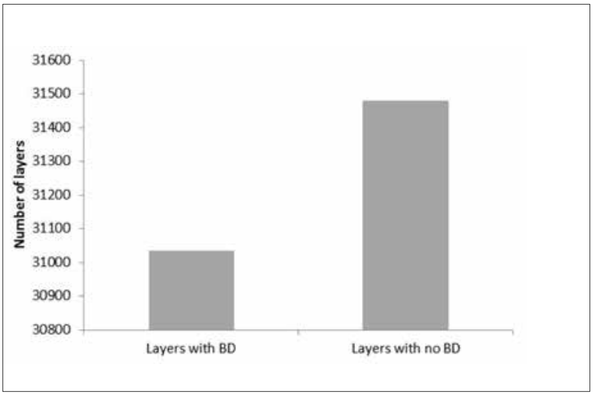


Figure 4. Total number of soil layers with and without bulk density measurements.

For layers with a C concentration higher than 12% it was used the log-transformed mean OC content as defined by equations 3 and 4:

$$BD(0\text{-}30) = -0.285 \times \ln(OC_{>12}) + 1.456\ g\ cm^3 \qquad [3]$$

$$BD(30\text{-}100) = -0.291 \times \ln(OC_{>12}) + 1.389\ g\ cm^3 \qquad [4]$$

where

BD(0-30) = topsoil bulk density (g cm^3)

BD(30-100) = subsoil bulk density (g cm^3)

$OC_{>12}$ = organic carbon content (%) of layer with OC>12%

Similarly, for the SPADE database we used equation 5 for estimating BD since it was the equation leading to the highest and consistent values for the coefficient of determination (*R2=0.85; OM=0.11 g cm³*):

$$BD = exp^{(a*OC+b)} \quad [5]$$

where
a = (-0.042)
b = (0.189) are constant.

Since all the other databases used in this study followed the same rules as for WISE3, when BD was missing we used equations 2, 3 and 4, depending on the OC content and the considered layer. No pedotransfer functions were used for North American profiles and for the Tuscia database, since only profiles with measured bulk density data were selected.

2.4 Soil type

To have an official and updated reference for the soil classification and the distribution of the different soil types, the Harmonized World Soil Database (Figure 5 - 1.21 version *http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/*) was downloaded and utilized.

This database fits perfectly with the resolution expected for the final output maps (30 arcsec). HWSD divides the global surface in a

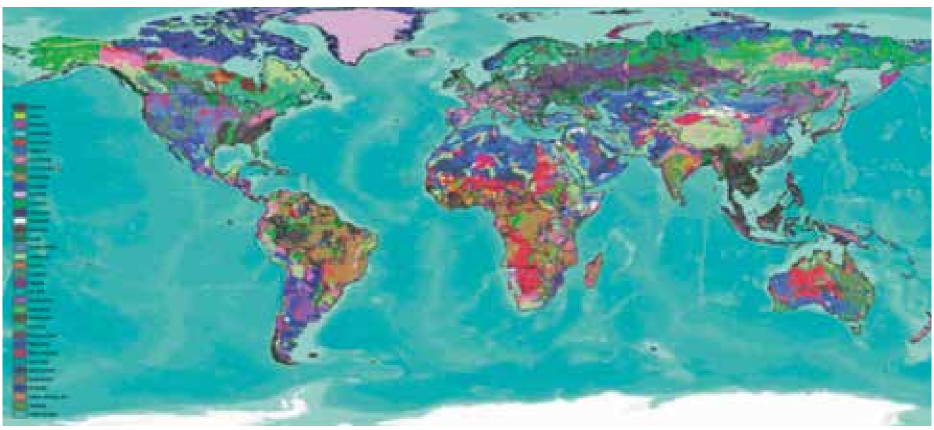


Figure 5. HWSD dominant Soil Types.

Figure 6. Distribution of "mixed soils".

regular 30 arcsec grid and provides information on the soil unit composition. A soil mapping unit can have up to 9 soil unit/topsoil combination record. Each soil unit has a field called "Share" that identify the % of the soil unit/topsoil texture. In this work was used the dominant soil type. Furthermore to refine the accuracy of SOC calculation where any HWSD soil unit/topsoil texture doesn't reach 50% of "share" a further soil type has been defined; this new class, defined "Mixed soils" is estimated at around 22% of HWSD cells (Figure 6). Table 3 shows some statistics (number of soil profiles, mean and standard deviation) of SOC stock for Top and Sub soil for the soil types in HWSD. Values in Table 4 don't partecipate in any way to SOC calculation fot final maps and are reported only for statistics purpose. The low represented HWSD miscellaneous units (Sand Dunes, Water Bodies, Urban, Salt Flats and Glaciers) although not considered in the Table 3, contributed to the determination of Homogeneous Land Units (see 2.6).

	Top (0-30 cm)			Sub (30-100 cm)		
	Profiles	**Mean**	**St Dev**	**Profiles**	**Mean**	**St Dev**
Acrisols	1 033	58,38	50,69	749	44,06	57,93
Alisols	76	54,97	37,8	49	43,29	37,76
Andosols	244	102,76	79,8	156	111,13	124,63
Arenosols	1 069	26,18	43,31	736	33,54	73,24
Anthrosols	245	41,06	20,82	162	45,13	29,49
Chernozems	403	81,66	39,81	338	84,73	56,53
Calcisols	366	32,59	27,94	214	35,75	36,15
Cambisols	2 017	61,04	59,13	1 503	45,51	53,64
Fluvisols	736	39,34	34,13	571	45,1	47,15
Ferralsols	991	47,6	55,6	871	53,22	83,28
Gleysols	565	67,42	72,23	405	51,53	85,04
Greyzems	81	70,37	23,25	71	61,64	51,49
Gypsisols	36	24,45	28,64	18	61,88	79,95
Histosols	97	127,93	127,41	72	149,77	200,68
Kastanozems	855	54,93	28,65	576	50,86	30,06
Leptosols	1 203	63,99	71,92			
Luvisols	1 993	49,43	45,43	1 356	46,78	60,22
Lixisols	652	30,04	45,02	506	38,13	69,25
Nitisols	634	45,11	45,11	562	53,67	54,73
Podzoluvisols	222	65,6	69,51	163	50,76	113,91
Phaeozems	1 106	71,77	44,17	802	64,73	59,51
Planosols	231	43,13	34,81	149	38,51	36,89
Plinthosols	212	39,01	58,12	151	42,13	92
Podzols	317	89,28	67,44	216	48,15	73,52
Regosols	553	39	41,78	340	36,38	54,78
Solonchaks	77	23,93	18,29	53	31,22	29,43
Solonetz	155	34,49	22,57	110	33,13	29,66
Vertisols	1 000	36,48	40,2	836	50,68	63,61

Table 3. Statistics of SOC (mg/ha) per HWSD soil types.

2.5 Land cover

Information about Land Cover were obtained from the new FAO GLC-SHARE Global Land Cover Database (ver 1.0 Latham *et al.* 2013). This database with a 30 arcsec resolution includes 12 layers in raster format. One layer or each land cover class:

- Artificial surfaces;
- Cropland;
- Grassland;
- Tree covered area;
- Shrubs covered area;
- Herbaceous vegataion;
- Mangroves;
- Sparse Vegetation;
- Baresoil;
- Snow and Glaciers;
- Water Bodies;
- one layer for the dominant land cover.

For SOC stock distribution only the information about dominant land cover were used. Figure 7 shows the GLC-SHARE database with the FAO legend. In Table 4 were reported some statitics (number of soil profiles, mean and standard deviation) for the SOC stocks of the two considered depths of the principal land cover classes. This values don't partecipate in any way to SOC calculation, are reported only for statistics purposes.

	Top (0-30 cm)			Sub (30-100 cm)		
	Profiles	**Mean**	**St Dev**	**Profiles**	**Mean**	**St Dev**
Artificial surfaces	526	50,23	49,35	395	44,89	53,81
Cropland	8 766	48,31	45,46	6 762	50,72	64,22
Grassland	3 791	44,73	53,24	2 581	43,88	67,8
Tree covered area	5 873	66,84	66,75	4 241	57,46	83,83
Shrubs covered area	3 097	46,57	55,34	2 050	51,86	78,59
Herbaceous vegetation	263	55,14	83,49	221	56,22	106,24
Mangroves	16	100,66	96,55	13	106,38	97,89
Sparse vegetation	228	43,66	49,93	134	45,62	68,70
Bare Soil	413	25,15	31,42	245	31,33	37,89

Table 4. Statistics of SOC (Mg/ha) per Land Cover Class.

2.6 Homogeneous Land Units and analysis area

Homogeneous Land Units (HLUs) are necessary to respatialize punctual values of SOC stock to a global surface. In this purpose, as a consequence of the above described activities, HLUs have been extracted at 30 arcsec spatial resolution by spatially combining 35 soil types (34 defined by HWSD and 1 "Mixed soils") and 10 land cover classes. Practically, a unique HLU was assigned

Figure 7. Distribution of dominant GLC-SHARE Land Cover Database.

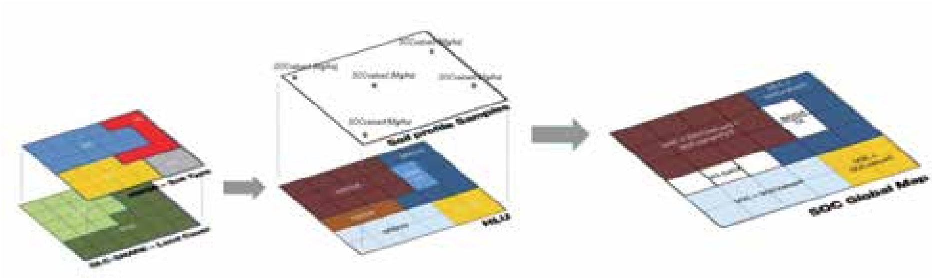


Figure 8. SOC spatialization methodology.

to each unique combination of LC and Soil Type, meaning that land units are homogenous if they correspond to the same soil and the same land cover. The Waterbodies class did not give any contribution to the HLUs identification. As in the previous tasks for single attributes, topsoil and subsoil SOC stock statistics (mean, standard deviation etc.) was calculated inside each HLU. Mean data for each HLU was assigned to all grid cells. All this process was implemented in ESRI ArcGis Spatial Analyst extension and illustrated in Figure 8.

$$SOC = \frac{\sum^{n}_{i=1} x_i}{n} \qquad [6]$$

Figure 9. Spatial intersection.

For each HLU *SOC* (expressed in Mg/ha) is the contents of Soil Organic Carbon and *x* is the SOC contents in i-th soil profile samples and *n* the number of profile samples for HLU. To individuate the area of analysis (therefore create the HLUs) a spatial intersection between Soil Type and GLC-SHARE database was performed (Figure 9). At the end of this process the global surface was divided in 21506'8294 cells (0.0083° x 0.0083°, approx 1 Km x 1 Km at the Equator).

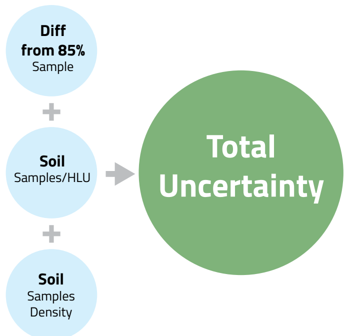


Figure 10. Uncertainty factors.

2.7 Quality assessment and uncertainty

To increase the quality and reliability of the SOC estimation both for Top SOC and Sub SOC the values over the 99.8 percentile were removed and the HLUs with less than 3 values were excluded and set up as no data in the final maps. To obtain a map of uncertainty three factors were considered (Figure 10):

- Variance, expressed in normalized absolute value (0-1) between SOC stock values from all soil samples and also considering a random subset (85% of total samples). No data area were set as 1 (max uncertainty);
- Normalized number of soil samples (0-1) per Homogeneous Lands Units (more soil samples. high accuracy);
- Normalized density of soil samples (0-1)

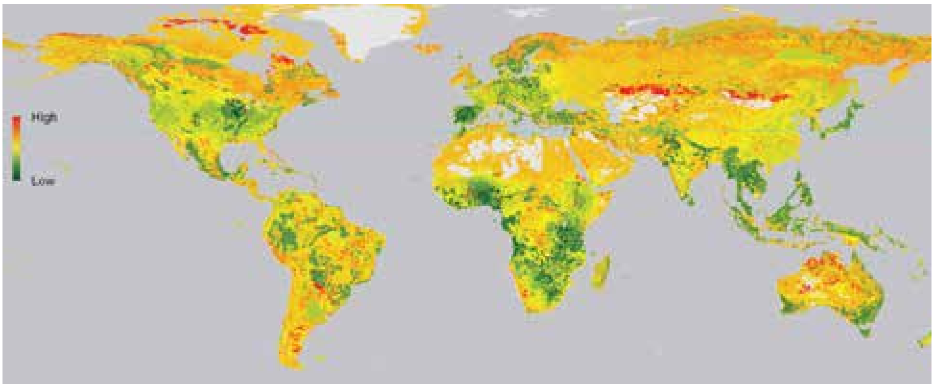


Figure 11. Distribution of uncertainty for Top SOC.

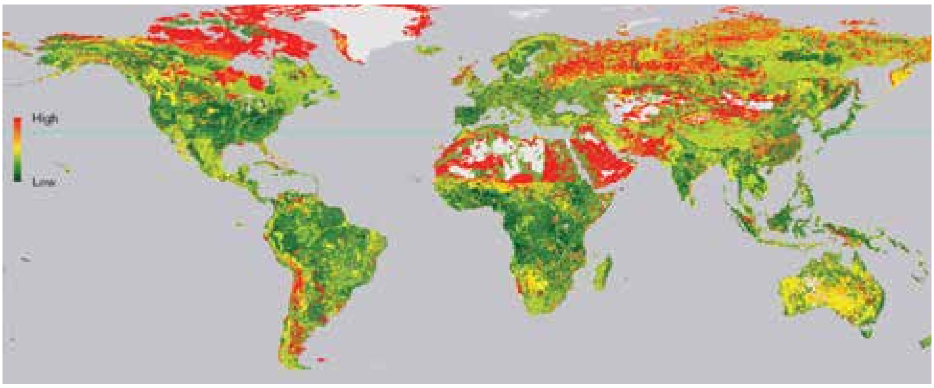


Figure 12. Distribution of uncertainty for Sub SOC.

Total uncertainty was calculated according to Equation 7 assigning a weight to these three factors (60% to 1,20% to 2 and 20% to 3). Figures 11 and 12 show the distribution of total uncertainty derived from this calculation (0 for low uncertainty / high accuracy, 1 for high uncertainty / low accuracy).

$$U = 0.6 \left|\left[\left(\frac{SOC-SOC_{85}}{SOC}\right)\right]\right| + 0.2\left[1-\left(\frac{N-Nmin}{Nmax-Nmin}\right)\right]+0.2\left[1-\left(\frac{D-Dmin}{Dmax-Dmin}\right)\right] \quad [7]$$

where

U = total uncertainty

SOC = Soil Organic Carbon content calculated with the whole soil samples database

SOC_{85} = soil Carbon content calculated with the subset of 85% of the samples,

N = number of soil samples per homogeneous land units

D = density of s.s. per surface units.

2.8 Other global datasets

Other global Soil Organic Carbon datasets are currently available: i.e the Global Soil Organic Carbon Map from U.S. Departement of Agriculture NRCS (2006), the ISRIC- Wise-5by5Min (2006) and many other, among these the JRC's "Global Soil Organic Carbon Estimates" (Hiederer and Köchy, 2011) has a 30 acrsec spatial resolution (Figure 13).

As in the present study the JRC dataset is based on Harmonized World Soil Database, therefore it is possible to find some similarities between the two databases. However the process of SOC spatialization is totally different since the Soil Organic Carbon 30arcsec Global Database is based on the punctual measured values of soil profiles (as described in section 2.2 and following). JRC database is very accurate, but uses a different logical approach; thus making impossible to carry out a cross validation among these databases. Figures 14 and 15 show that although the distribution of Top SOC stock has similar patches, values associated are slightly different.

Figure 13. JRC's "Global Soil Organic Carbon Estimates" (Top SOC).

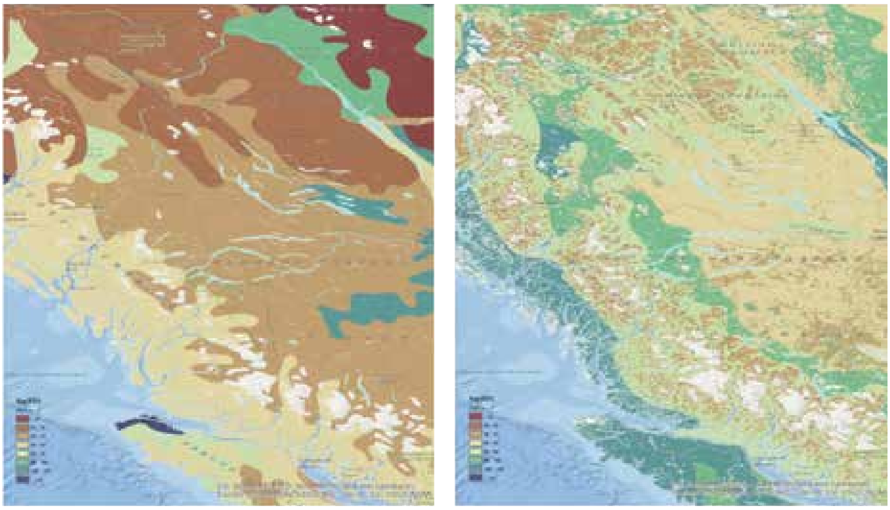


Figure 14. Comparison 1 Left JRC Database Right Soil Organic Carbon 30 arc sec Global Database.

Figure 15. Comparison 2 Left JRC Database Right Soil Organic Carbon 30 arc sec Global Database.

© FAO/ Giulio Napolitano

3. RESULTS

3.1 SOC Geodatabase

An ESRI File Geodatabase was implemented to organize geospatial and non-geospatial data. This geodatabase contains final outputs and intermediate geoprocessing layers in various formats.

All layers are provided of metadata (Figure 16) according to ISO 19115 and ISO 19139 standards. Coordinate System is WGS 84 (geographic – EPSG 4326). In the Soil Organic Carbon 30 arcsec Global Database are stored three types of data:

TABLE

- NON SPATIAL DATA
- STATISTICS
  (Table 5 i.e. SOC per Soil Type)

VECTORIAL

- TOP soil profiles global layers
- SUB soil profiles global layers

RASTER

- SOC STOCK LAYERS
- UNCERTAINTY MAPS
- INTERMEDIATE OUTPUTS
- FINAL OUTPUTS

3.1.1 Identification of geodatabase layers

Each layer in GDB (Figure 15) is identified with a name; this is composed by a suffix and an identifier (X_IDENTIFIER). The suffix "X" indicates:

- **"S"**: source data, i.e. elaborations of HWSD or GLC Share layers, data producted by other Institutes;

Figure 16. Metadata.

- **"I"**: intermediate data, result of processing and elaboration, but not final otuputs (i.e. HLU maps, layers composing final maps of uncertainty);
- **"F"**: final outputs, final products of elaborations, i.e. SOC Stock and Uncertainty maps, Soil Profiles layers, statistical table.

Within the database, the associated metadata provide further information, specifics and details of all element.

3.1.2 Final outputs

The Soil Organic Carbon 30 arcsec Global database includes 4 main final outputs and 2 raster layers with uncertainty information. Map of global Top (0-30 cm) SOC stock ("F_TOP_SOC"), ESRI grid format, are shown in Figure 18, while Map of global Sub (30-100 cm) SOC stock ("F_SUB_SOC"), ESRI grid format in Figure 19. Each pixel value of these two maps indicates the amount of SOC stock in Mg/ha. The other two final outputs, stored in "F_SOIL_PROFILES" ESRI Feature Dataset, are "F_TOP_SOC_UPDATE" and "F_SUB_SOC_UPDATE" (Figures 2 and 3).

The latter two datasets, represent vector layers with geolocated soil profiles database and a simple table of attribute (ObjID,SOC in Mg/ha and Unique ID).

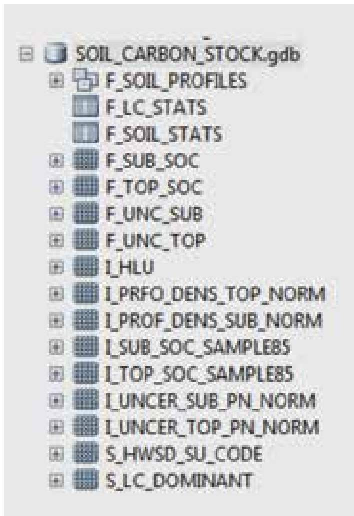


Figure 17. SOC30 arcsec Global database structure.

In the uncertainty maps (calculation process described in 2.7 - Figures 11 and 12) provided in raster format (ESRI grid) and identified with "F_UNC_TOP" and "F_UNC_SUB" codes, each pixel contains a value between 0 and 1, indicating the level of reliabilty of the SOC associated values. More information about data and layer in Table 5.

Figure 18. Distribution of Top soil organic Carbon (Continuous data is aggregated in classes for illustrative purpose).

Figure 19. Distribution of Sub soil organic Carbon (Continuous data is aggregated in classes for illustrative purpose).

Data Name	Format	Product	Description
F_SOIL_PROFILES			Contains the following two datasets
F_TOP_SOC_UPDATE	Vector	Final	Soil profiles database used to calculate Soil Organic Carbon Stock in sub (30-100 cm depth) soil
F_TOP_SOC_UPDATE	Vector	Final	Soil profiles database used to calculate Soil Organic Carbon Stock in top (0-30 cm depth) soil
F_LC_STATS	table	Final	Information about the main stats (number of soil samples, mean and standard deviation) of Soil Organic Carbon (Mg/ha) for each dominant land cover class
F_SOIL_STATS	Table	Final	Information about the main stats (number of soil samples, mean and standard deviation) of Soil Organic Carbon (Mg/ha) for each soil type
F_SUB_SOC	Raster	Final	Global map of Organic Carbon Stock in sub (30-100 cm depth) soil
F_TOP_SOC	Raster	Final	Global map of Organic Carbon Stock in top (0-30 cm depth) soil
F_UNC_SUB	Raster	Final	Uncertainty for global map of SubSoil Organic Carbon Stock (30-100 cm)
F_UNC_TOP	Raster	Final	Uncertainty for global map of Top Soil Organic Carbon Stock (0-30 cm)
I_HLU	Raster	Intermediate	Homogeneous land units (land cover and soil type) on the global surface to re-spatialize soil organic carbon value collected with soil profiles
I_PROF_DENS_TOP_NORM	Raster	Intermediate	Global Density of top soil profiles (Uncertainty factor)
I_PROF_DENS_SUB_NORM	Raster	Intermediate	Global Density of sub soil profiles (Uncertainty factor)
I_SUB_SOC_SAMPLE85	Raster	Intermediate	Differences between the value of Organic Carbon Stock in sub (30-100 cm depth) considering the whole set of soil samples and the 85% (uncertainty factor)
I_TOP_SOC_SAMPLE85	Raster	Intermediate	Differences between the value of Organic Carbon Stock in top (0-30 cm depth) considering the whole set of soil samples and the 85% (uncertainty factor)
I_UNCER_SUB_PN_NORM	Raster	Intermediate	Number of sub soil profiles for each HLU (uncertainty factor)
I_UNCER_TOP_PN_NORM	Raster	Intermediate	Number of top soil profiles for each HLU (uncertainty factor)
S_HWSD_SU_CODE	Raster	Source Data	Harmonized World Soil Database (Rel 1.21) soil dominant group
S_LC_DOMINANT	Raster	Source Data	GLC-SHARE Land Cover Database (Beta Rel 1.0) dominant LC

Table 5. Geodatabase layers description.

© FAO/ Giuseppe Bizzarri

4. DISCUSSION

The Soil Organic Carbon 30 arcsec Global Database map for the 0-30 cm depth (Top SOC) covers approximately 96% of Earth land surface while the 30-100 cm one (Sub SOC) about 94% (excluding Antarctica). Most of this coverage has an acceptable level of uncertainty (Figures 11 and 12).

In general, most of the SOC stored in the 0-30 cm depth (as shown in graphs in Figure 20) ranges between 20 and 120 Mg/ha (almost 80%). Over 20% of global land surface shows high values of C stored in the topsoil (over 120 Mg/ha). These areas refer mainly to Siberia and some areas in the north of Canada. Very low values (< 20 Mg/ha), interest approximately 6% of global surface, and are highly concentrated in the internal zones of Australia, South of Africa and in the Sahara and Gobi deserts according to what observed also by Batjes (1996).

A similar figure is observed for the SOC stored in the sub soil (30-100 cm depth – Figure 21). The only substantial difference with the topsoil is represented by the areas with very low values of SOC (<15 Mg/ha) which are wider (over 8%).

Very interesting is to remark this tendency, whereby high and low values of SOC in top

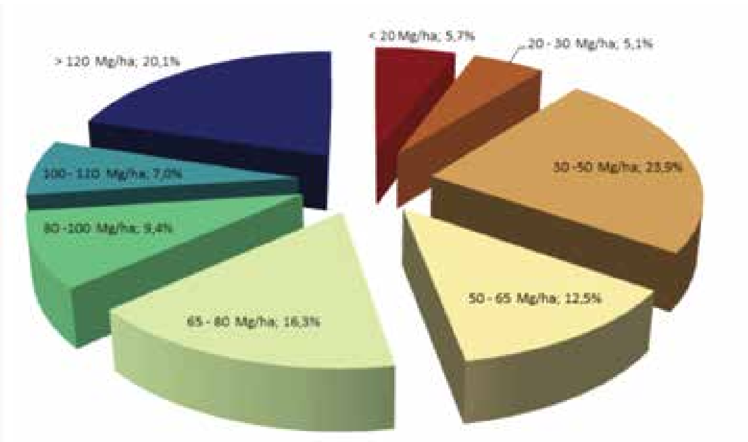


Figure 20. Classes of Top SOC in percentage. In pie chart global SOC data are aggregated for explicative purpose in Quantile Classes (each class has the same number of observations).

and sub soil are situated nearly in the same geographical zones except the polar areas of Canada where high values of top SOC correspond to very low values in sub. This fact confirms the validity of the calculation process adopted for this work.

For both final maps, large areas with no data can be found in the Sahara region, Greenland, Rub al Khali desert (Arabian peninsula), Himalaya and the desert areas of Turkmenistan and Uzbekistan due to lack of soil profiles. It is evident that no data are spread prevalently in deserted areas. Making a comparison between the SOC stored in the topsoil (0-30 cm) in relation to the SOC stock of the whole profile (0-100 cm) for each single soil unit indicate a good agreement with what observed previously by Batjes (1996) working at global scale and

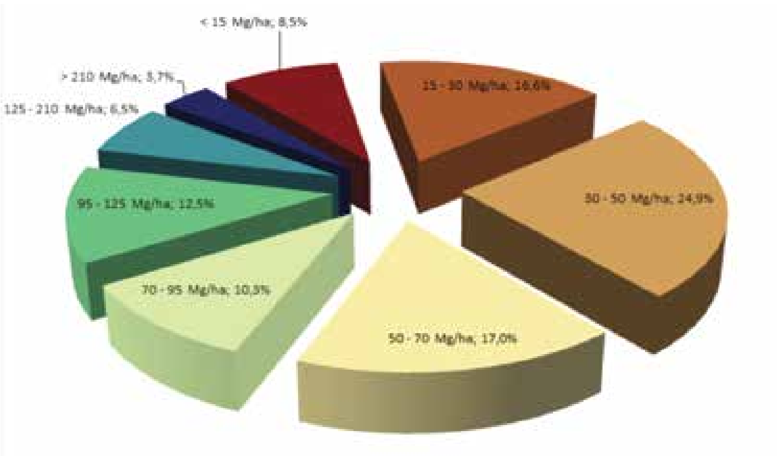


Figure 21. Classes of Top SOC in percentage. In pie chart global SOC data are aggregated for explicative purpose in Quantile Classes (each class has the same number of observations).

using only the profiles contained in WISE3 database. Considering that in the global database we used the number of profiles for each single soil unit is much higher this agreement was not predicted (Table 6).

Besides, comparing the mean SOC stock data for the main FAO soil units, as shown in this study, with the mean values reported by Batjes (1996) for the the same soil units indicate no substantial differences. In the tropical area Ferralsols and Acrisols show a mean SOC stock of 48 Mg C /ha, and 58 Mg C /ha similar to the values reported by Batjes (1996) for the same units: 57 Mg C /ha and 51, respectively. In others climatic regions the same can be observed for Cambisols soil units (61 Mg C /ha vs. 50 Mg C /ha), Luvisols (49 Mg C /ha vs 31 Mg C /ha), Kas-

	This study		Batjes 1996	
	Profiles	**Mean*** *(Mg C /ha)*	**Profiles**	**Mean*** *(Mg C /ha)*
Acrisols	1 033	56	269	54
Cambisols	2 017	57	332	52
Chernozem	403	60	44	50
Podzoluvisols	222	56	7	70
Ferralsols	991	57	228	53
Gleysols	565	56	142	59
Phaeozems	1 106	52	147	53
Fluvisols	736	46	200	44
Kastanozems	855	51	8	40
Luvisols	1 993	51	377	47
Greyzems	81	53	3	57
Nitisols	634	45	67	49
Histosols	97	46	34	37
Podzol	222	64	43	59
Arenosols	1 069	44	166	44
Regosols	553	52	42	49
Solonetz	155	50	39	50
Andosols	244	47	120	48
Vertisols	1 000	41	205	40
Planosols	231	53	28	52
Solonchaks	77	45	42	44

Table 6. Comparison of the relative distribution of SOC as a function of depth in the considered FAO soil units between the results from this study and those resulting from Batjes 1996.

tanozem (55 Mg C /ha vs. 54 Mg C /ha) and Nitisols (45 Mg C /ha vs. 41 Mg C /ha). The main difference concern the soils units of cold regions. For example Histosols seems to be greatly underestimated with a SOC stock of 129 Mg C /ha vs 283 Mg C /ha. The same can be observed for Podzol, 89 Mg C /ha, with a value slightly lower than the estimate of 136 Mg C /ha reported by Batjes (1996). On the other hand, in the same climatic region similar values are reported for Gleysols 67 Mg C /ha vs. 77 Mg C /ha.

In all these cases it has to be considered the large uncertainties that affect the SOC stock in all the soil units (Table 3), resulting in non-significant differences in all the above mentioned cases. Only for Histosols differences are significant. This discrepancy for this soil unit could be related to the PTF's used for calculating the bulk density, despite, as shown by Hiederer and Köchi (2011) the adopted PTF's are those that better fit the organic matter values for organic soils. How-

** The mean value rapresents the ratio of the SOC stock of the 0-30 cm divided by that in the 0-100 cm depth.*

ever, soils of the cold regions are those with the smaller number of representative soil profiles in the database.

This database is greatly enriched in terms of number of soil profiles compared to all the other databases previously used to provide estimates of SOC stock at global level. Despite a possible underestimation of the SOC stock observed in the Histosols soil unit, the agreement of the stocks observed for all the other units suggest a good reliability of the data stored in this database. This is particularly important for the tropical area, where the addition of the recently available African soil profile database allowed for an enhanced coverage of the tropical area and at the same time resulted similar values to those already available in other databases.

In conclusion, this database could be confidently used for identifying areas suitable for SOC sequestration purposes.

	This study	Batjes 1996
	(Mg C /ha)	(Mg C /ha)
Acrisols	58	51
Cambisols	61	50
Chernozem	82	60
Podzoluvisols	66	56
Ferralsols	48	57
Gleysols	67	77
Phaeozems	72	77
Fluvisols	39	38
Kastanozems	55	54
Luvisols	49	31
Greyzems	70	108
Nitisols	45	41
Histosols	129	283
Podzol	89	136
Arenosols	26	13
Regosols	39	31
Solonetz	34	32
Andosols	102	114
Vertisols	36	45
Planosols	43	39
Solonchaks	24	18

Table 7. Comparison of the 0-30 cm SOC stock in the main FAO soil as derived from this study and the values reported by Batjes 1996.


© FAO/ Giuseppe Bizzarri

5. NEXT STEPS AND RECOMMENDATIONS

5

This database should be considered as an implemented version of the Global Soil Organic Carbon database. Nevertheless, future updates, especially concerning the addition of new soil profiles for areas with low or not info on the SOC concentrations (i.e. Canada and Australia) are required. Many institutes (i.e. ISRIC) regularly publish new updates of their soil profiles database both in terms of adding new entities and improving the quality of existing information. These updates are essential to further increase the consistency of this database and decrease the global uncertainty.

An additional step could be the use of climatic information in identifying HLUs, but mandatory is the spatial resolution that must be comparable with 30 arcsec used in this work. Currently, climatic databases with 30 arcsec resolution are not available or unreliable.

During the preliminary phases of this work, tests using 0.5° climate zoning have been performed, but the resolution of this input data affected the spatial quality of the final maps (coarse resolution areas within high resolution. This issue is well described in Figure 22: the red circle around the city of Indianapolis shows a discontinuous distribution of SOC value due to 0.5° resolution of climate information. Using Soil Organic Carbon 30 arcsec Global Database final maps associated with uncertainty maps can be considered a good practice.

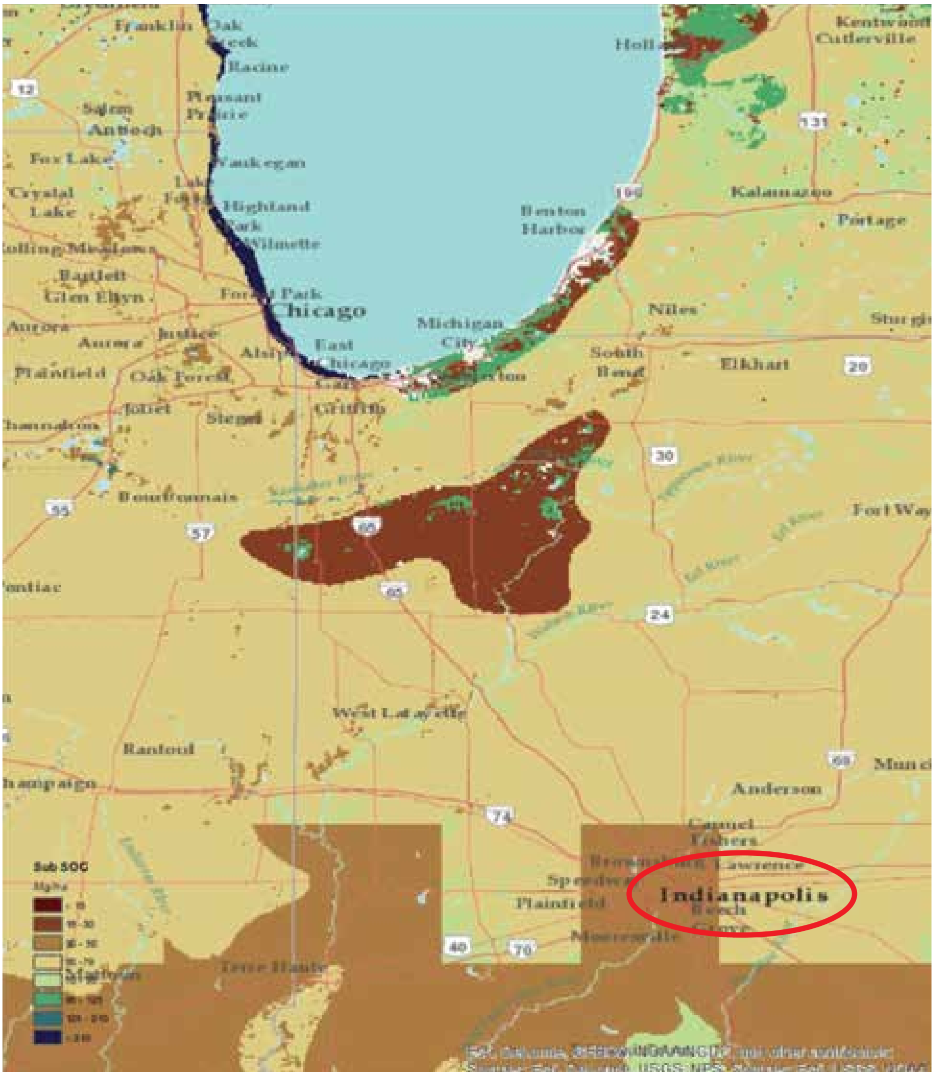


Figure 22. Red rimmed: decreasing spatial accuracy using climate data.

REFERENCES

Adams, W.A., 1973. *The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils*. J. Soil Sci., 24: 10-17.

Batjes, N.H., Dijkshoorn, J.A., 1999. *Carbon and nitrogen stocks in the soils of the Amazon Region*. Geoderma, 89: 273-286.

Batjes, N.H., 1996. *Total carbon and nitrogen in the soils of the world*. Eur. J. Soil Sci., 47: 151-163.

Batjes, N.H., 1997. A *world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling. Soil Use Manage.*, 13: 9-16.

Batjes, N.H., 2008. *ISRIC-WISE harmonized global soil profile dataset (Ver. 3.1)*. Report 2008/2, ISRIC - World Soil Information, Wageningen. (Also available at: http://www.isric.org/isric/webdocs/docs//ISRIC_Report_2008_02.pdf?q=isric/Webdocs/Docs/ISRIC_Report_2008_02.pdf).

Boone, R.D., Grigal, D.F., Sollins, P., Ahrens, R.J., Armstring, D.E., 1999. S*oil sampling, preparation, archiving, and quality control.* In: Robertson GP, Coleman DC, Bledsoe CS, Sollins P (eds) Standard soil methods for long-termecological research. Oxford University Press, New York, pp 3–28.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J., 2000. *Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model.* Nature, 408: 184-187.

Dixon, R.K., Brown, S., Houghton, R.A., Solomon, A.M., Trexler, M.C., Wisniewski, J., 1994. *Carbon pools and fluxes of global forest ecosystems.* Science, 263: 185-190.

Harden, J., 2008. *Delta Junction fire-chronosequence soil database: carbon, nitrogen, bulk density, and other properties.* Bonanza Creek LTER - University of Alaska Fairbanks. BNZ:336, (also available at: http://www.lter.uaf.edu/data_detail.cfm?datafile_pkey=336).

Hiederer, R., & Köchy, M., 2011. *Global soil organic carbon estimates and the harmonized world soil database*. EUR, 25225, 79.

Hiederer, R., Jones, R.J.A. & Daroussin, J., 2006. S*oil Profile Analytical Database for Europe (SPADE): Reconstruction and Validation of the Measured Data (SPADE/M).* Geografisk Tidsskrift, Danish Journal of Geography, 106: 71-85. (also available at http://rdgs.dk/djg/pdfs/106/1/06.pdf).

Hollis, J.M., Woods, S., 1989. T*he estimation of Available Water and Air Capacity at unusually high densities.* SSLRC Research Report for MAFF Project c(iv); Sept. 1989, 73 pp.

IPCC, 1990. *Climate Change* (eds J.T. Houghton, G.J. Jenkins & J.J. Ephraums), Cambridge University Press, Cambridge.

Janssens, I.A., Freibauer, A., Ciais, P., Smith, P., Nabuurs, G.J., Folberth, G. *et al.,* 2003. *Europe's terrestrial biosphere absorbs 7 to 12% of European anthropogenic CO2 emissions.* Science, 300: 1538-1542.

Jobbàgy, E.G., Jackson, R.B., 2000. *The vertical distribution of soil organic carbon and its relation to climate and vegetation.* Ecol. Appl., 10: 423-436.

Jones, C., McConnell, C., Coleman, K., Cox, P., Falloon, P., Jenkinson, D., Powlson, D., 2005. *Global climate change and soil carbon stocks; predictions from two contrasting models for the turnover of organic carbon in soil.* Glob. Change Biol. 11: 154–166.

Lal, R., 2002. *Soil carbon dynamics in cropland and rangeland.* Environ. Pollut., 116: 353-362.

L**atham, J., Cumani R., Rosati I., Bloise M.,** 2013. *Global Land Cover SHARE (GLC-SHARE) database Beta-Release Version 1.0,* FAO.

Leenaars, J. G. B., 2013. *Africa soil profiles database, Version 1.1. A compilation of georeferenced and standardized legacy soil profile data for Sub-Saharan Africa (with dataset).* ISRIC Rep, 3.

Rawls, W.J., 1983. *Estimating soil bulk-density from particle-size analysis and organic matter content.* Soil Science, 135: 123-125.

Ruehlmann, J., Körschens, M., 2009. *Calculating the effect of soil organic matter concentration on soil bulk density.* Soil Sci. Soc. Am. J. (73); pp. 876-885.

Scharlemann, J.P.W., Tanner, E.V.J., Hiederer, R., Kapos, V., 2014. *Global soil carbon: understanding and managing the largest terrestrial carbon pool.* Carbon Manage. 5: 81–91.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, B., *et al.* 2007. *Agriculture. Chapter 8 of Climate change 2007: Mitigation.* In: Contribution of Working group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds B. Metz, O.R. Davidson, P.R. Bosch, R. Dave & L.A. Meyer), pp. 180. Cambridge University Press, Cambridge, UK and New York, USA.

Smith, P., Fang, C., Dawson, J. J., & Moncrieff, J. B., 2008. *Impact of global warming on soil organic carbon.* Adv. Agron., 97: 1-43.

Stolbovoy, V., McCallum, I. *et al.,* 2002. *Land Resources of Russia.* (also available at: http://www.iiasa.ac.at/collections/IIASA_Research/Research/FOR/russia_cd/download.htm).

Vogel, A.W., 1994. *Comparability of soil analytical data: determinations of cation exchange capacity, organic carbon, soil reaction, bulk density, and volume percentage of water selected pF values by different methods.* Work. Pap. 94/07, ISRIC, Wageningen. (also available at: http://www.isric.org/isric/webdocs/docs//ISRIC_Report_1994-07.pdf?q=isric/webdocs/Docs/ISRIC_Report_1994-07.pdf).

This project is supported by the EU under the 7th Framework Programme for Research and Technological Development. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Viale delle Terme di Caracalla 00153 Rome (Italy) - www.fao.org

ISBN 978-92-5-109762-5
9 789251 097625
I7292EN/1/05.17