https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019)

Niels H. Batjes[1], Eloi Ribeiro[1], Ad van Oostrum[1]

[1]ISRIC – World Soil Information, Wageningen, 6708 PB, The Netherlands

Correspondence to: Niels H. Batjes (niels.batjes@isric.org)

Abstract. The World Soil Information Service (WoSIS) provides quality-assessed and standardised soil profile data to support digital soil mapping and environmental applications at broad scale levels. Since the release of the first 'WoSIS snapshot', in July 2016, many new soil data were shared with us, registered in the ISRIC data repository, and subsequently standardised in accordance with the licences specified by the data providers. Soil profile data managed in WoSIS were contributed by a wide range of data providers, therefore special attention was paid to measures for soil data quality and the standardisation of soil property definitions, soil property values (and units of measurement), and soil analytical method descriptions. We presently consider the following soil chemical properties (organic carbon, total carbon, total carbonate equivalent, total Nitrogen, Phosphorus (extractable-P, total-P, and P-retention), soil pH, cation exchange capacity, and electrical conductivity) and physical properties (soil texture (sand, silt, and clay), bulk density, coarse fragments, and water retention), grouped according to analytical procedures (aggregates) that are operationally comparable. Further, for each profile, we provide the original soil classification (FAO, WRB, USDA, and version) and horizon designations insofar as these have been specified in the source databases. Measures for geographical accuracy (i.e. location) of the point data as well as a first approximation for the uncertainty associated with the operationally defined analytical methods are presented, for possible consideration in digital soil mapping and subsequent earth system modelling. The latest (*dynamic*) set of quality-assessed and standardised data, called 'wosis_latest', is freely accessible via an OGC-compliant WFS (web feature service). For consistent referencing, we also provide time-specific *static* 'snapshots'. The present snapshot (September 2019) comprises 196,498 geo-referenced profiles originating from 173 countries. They

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

represent over 832 thousand soil layers (or horizons), and over 5.8 million records. The actual number of observations for each property varies (greatly) between profiles and with depth, this generally depending on the objectives of the initial soil sampling programmes. In the coming years, we aim to fill gradually gaps in the geographic and feature space, this subject to the sharing of a wider selection of soil profile data for so far under-represented areas and properties by our existing and prospective partners. Part of this work is foreseen in conjunction within the Global Soil Information System (GloSIS) being developed by the Global Soil Partnership (GSP). The 'WoSIS snapshot - September 2019' is archived and freely accessible at https://dx.doi.org/10.17027/isric-wdcsoils.20190901 (Batjes et al., 2019).

1 Introduction

According to a recent review, so far over 800 thousand soil profiles have been rescued and compiled into databases during the past decades (Arrouays et al., 2017). However, only a fraction thereof is readily accessible (i.e. *shared*) in a consistent format for the greater benefit of the international community. This paper describes procedures for preserving, quality-assessing, standardising, and subsequently providing consistent world soil data to the international community as developed in the framework of the Data\WoSIS (World Soil Information Service) project; this collaborative project draws on an increasingly large complement of shared soil profile data. Ultimately, WoSIS aims to provide consistent harmonised soil data, derived from a wide range of legacy holdings as well as from more recently developed soil spectral libraries (Terhoeven-Urselmans et al., 2010; Viscarra Rossel et al., 2016), in an interoperable mode and this preferably within the setting of a federated, global soil information system (GLOSIS, see GSP-SDF, 2018).

We follow the definition of harmonisation as defined by the Global Soil Partnership (GSP, Baritz et al., 2014). It encompasses "providing mechanisms for the collation, analysis and exchange of consistent and comparable global soil data and information". The following domains need to be considered according to GSP's definition: a) soil description, classification and mapping, b) soil analyses, c) exchange of digital soil data, and d) interpretations. In view of the breadth and magnitude of the task, as indicated earlier (Batjes et al., 2017), we have restricted ourselves to the standardisation of soil property definitions, soil analytical method descriptions, and soil property values (i.e. measurement units). We have expanded the number of soil properties considered in the preceding snapshot, i.e. those listed in the GlobalSoilMap (2015) specifications, gradually working towards the range of soil properties commonly considered in other global soil data compilation programmes (Batjes, 2016; FAO et al., 2012; van Engelen and Dijkshoorn, 2013).

Soil characterisation data, such as pH and bulk density, are collated according to a wide range of analytical procedures. Such data can be more appropriately used when the procedures for their collection, analysis, and reporting are well understood. As indicated by USDA Soil Survey Staff (2011), results differ when different analytical methods are used even though these methods may carry the same name (e.g. soil pH) or concept. This complicates, or sometimes precludes, comparison of one set of data with another if it is not known how both sets were collected/analysed. Hence our use of 'operational definitions' for soil properties that are linked to specific methods. As an example, we may consider the 'pH of a soil'. This requires information on sample pre-treatment, soil/solution ratio, and description of solution (e.g. H_2O, 1 M KCl, 0.02 M $CaCl_2$, or 1 M NaF) to be fully understood. pH measured in Sodium Fluoride (pH NaF), for example, provides a measure for the Phosphorus (P) retention of a soil whereas pH measured in water (pH H_2O) is an indicator for soil nutrient status. Consequently, in WoSIS, soil properties are defined by the analytical methods and the terminology used, based on common practice in soil science.

This paper discusses methodological changes in the WoSIS workflow since the release of the preceding snapshot (Batjes et al., 2017), describes the data screening procedure, provides a detailed overview of the database content, explains how the new set of standardised data can be accessed, and outlines future developments. The data model for the underpinning PostgreSQL database itself is described in a recently updated Procedures Manual (Ribeiro et al., 2018); these largely technical aspects are considered beyond the scope of this paper.

Quality-assessed data provided through WoSIS can be, and have been, used for various purposes. For example, as point data for making soil property maps, at various scale levels, using digital soil mapping techniques (Arrouays et al., 2017; Guevara et al., 2018; Hengl et al., 2017a; Hengl et al., 2017b; Moulatlet et al., 2017). Such property maps, for example, can be used to study global effects of soil and climate on leaf photosynthetic traits and rates (Maire et al., 2015), generate maps of root-zone plant-available water capacity (Leenaars et al., 2018) in support of yield gap analyses (van Ittersum et al., 2013), assess impacts of long-term human land use on world soil carbon stocks (Sanderman et al., 2017), or the effects of tillage practices on soil gaseous

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

emissions (Lutz et al., 2019). In turn, this type of information can help to inform the global conventions such as the UNCCD (United Nations Convention to Combat Desertification) and UNFCCC (United Nations Framework Convention on Climate Change), so that policymakers and business leaders can make informed decisions about the environment and human well-being.

2 WoSIS workflow

The overall workflow for acquiring, ingesting, and processing data in WoSIS has been described in an earlier paper (Batjes et al., 2017). To avoid repetition, we will only name the main steps here (Fig. 1). These successively are: a) store submitted data sets with their metadata (including the licence defining access rights) in the ISRIC Data Repository; b) import all datasets 'as is' into PostgreSQL; c) ingest the data into the WoSIS data model, including basic data quality assessment and control; d) standardise the descriptions for the soil analytical methods and the units of measurement, and e) ultimately, upon final consistency checks, distribution of the quality-assessed and standardised data via WFS (web feature service) and other formats (e.g. CSV for snapshots).

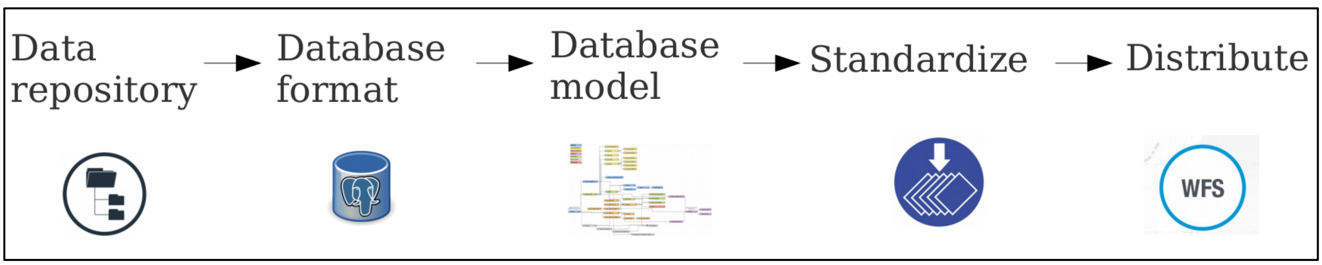

Figure 1. Schematic representation of WoSIS workflow for safeguarding and processing disparate soils data sets.

As indicated, data sets shared with our centre are first stored in the ISRIC Data Repository together with their metadata (currently representing some 452 thousand profiles), in particular the licence and data sharing agreement, this in line with the ISRIC Data Policy (ISRIC, 2016). For the WoSIS standardisation workflow *proper*, we only consider those data sets (or profiles) that have a 'non-restrictive' Creative Commons (CC) licence, as well as the defined complement of attributes (see Appendix. A). 'Non-restrictive' has been defined here as at least a CC-BY (Attribution) or CC-BY-NC (Attribution Non-Commercial) licence. Presently, this corresponds with data for some 196,498 profiles (i.e. profiles that have the right licence and data for at least one of the standard soil properties). Alternatively, some data sets may *only* be used for digital soil mapping *sensu* SoilGrids™, corresponding with an additional 42 thousand profiles. Although the latter profiles are quality-assessed and standardised following the regular WoSIS workflow, they are *not distributed* to the international community in accordance with the underpinning licence agreements; as such, their description is beyond the scope of the present paper. Finally, several data sets have licences indicating that they should only be safeguarded in the repository; inherently, these are not being used for any data processing.

3 Data screening, quality control and standardisation

3.1 Consistency checks

Soil profile data submitted for consideration in WoSIS were collated according to various national or international standards, and presented in various formats (from paper to digital). Further, they are of varying degree of completeness as discussed below. Proper documentation of the provenance and identification of each dataset, and ideally each observation or measurement, is necessary to allow for efficient processing of the source data. In particular, the following need to be specified: feature (x-y-z and time (t) referenced profiles and layers), attribute (class, site, layer-field, and layer-lab), method and value, including units of expression.

To be considered in the actual WoSIS standardisation workflow, each profile has to meet several criteria (Table 1). First, we assess if each profile is geo-referenced, has (consistently) defined upper and lower depths for each layer (or horizon), and data for at least some soil properties (e.g. sand, silt, clay and pH). Having a soil (taxonomic) classification is considered desirable (case 1), though not mandatory (case 2). Georeferenced profiles for which only the classification is specified can still be useful for mapping of soil taxonomic classes (case 3). Alternatively, classified profiles without any geo-reference may still prove useful to develop pedotransfer functions (case 4); however, they cannot be served through WFS (because there is no geometry (x,y)). The remaining three cases (5, 6 and 7) are automatically excluded from the WoSIS workflow. This first, broad consistency check led to the exclusion of over 50,000 profiles from the initial complement of soil profiles.

Table 1. Basic requirements for considering soil profiles in the WoSIS standardisation workflow

Case	(X,Y)	Layer depth	Soil properties[a]	Classification	Keep
1	+	+	+	+	Yes
2	+	+	+	-	Yes
3	+	-	-	+	Yes[a]
4	-	+	+	+	Yes/No[b]
5	+	+	-	-	No
6	-	+	+	-	No
7	+	-	+	-	No

[a] Such profiles may be used to generate maps of soil taxonomic classes using SoilGrids™ (Hengl et al., 2017b).

[b] Such profiles (geo-referenced solely according to their country of origin) may be useful for developing pedotransfer rules. Hence, they are standardised though not distributed with the snapshot.

Consistency in layer depth (i.e. sequential increase of the upper and lower depth reported for each layer down the profile) is checked using automated procedures (see Section 3.2). In accord with current conventions, such depth increments are given as 'measured from the surface, including organic layers and mineral covers' (FAO, 2006; Schoeneberger et al., 2012). Prior to 1993, however, the begin (zero datum) of the profile was set at the top of the mineral surface (the *solum* proper), except for 'thick' organic layers as defined for peat soils (FAO-ISRIC, 1986; FAO, 1977). Organic horizons were recorded as above and mineral horizons recorded as below, relative to the mineral surface (Schoeneberger et al., 2012, p. 2-6). Insofar as possible, such 'surficial litter' layers are flagged in WoSIS so that they may be filtered-out during auxiliary computations of soil organic carbon stocks, for example.

3.2 Flagging duplicate profiles

Several source materials, such as the harmonised WISE soil profile database (Batjes, 2009), the Africa Soil Profile Database (AfSP, Leenaars et al., 2014), and the dataset collated by the International Soil Carbon Network (ISCN, Nave et al., 2017) are compilations of shared, soil profile data. These three datasets, for example, contain varying amounts of profiles derived from the National Cooperative Soil Survey database (USDA-NCSS, 2018), an important source of freely shared, primary soil data. The original NCSS profile identifiers, however, may not always have been preserved 'as is' in the various data compilations.

To avoid duplication in the WoSIS database, soil profiles located within 100 m of each other are flagged as possible duplicates. Upon additional checks concerning the first three layers (upper and lower depth) and their sequential numbering (from top to bottom), as well as range of attribute data (with special attention for sand, silt and clay content), when necessary with some additional visual checks, the duplicates with the least comprehensive component of attribute data are flagged and excluded from further processing. This laborious, yet critical, second screening process (see Ribeiro et al., 2018) led to the exclusion of some 50,000 additional profiles from the initial complement of soil profile data.

3.3 Ensuring naming consistency

A next, key stage has been the standardisation of soil property names to the WoSIS conventions, as well as the standardisation of the soil analytical methods descriptions themselves (see Appendix A). Quality checks consider the units of measurement, plausible ranges for defined soil properties (e.g. soil pH cannot exceed 14) using checks on minimum, average and maximum values for each source data set. The whole procedure, with flowcharts and option tables, is documented in the WoSIS Procedures Manual (see App. D, E and F in Ribeiro et al., 2018).

Presently, we standardise the following set of soil properties in WoSIS:

- Chemical: organic carbon, total carbon (i.e. organic plus inorganic carbon), total nitrogen, total carbonate equivalent (inorganic carbon), soil pH, cation exchange capacity, electrical conductivity, and Phosphorus (extractable-P, total-P, and P-retention),
- Physical: Soil texture (sand, silt, and clay), coarse fragments, bulk density, and water retention.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

It should be noted that all measurement values are reported as recorded in the source data, subsequent to the above consistency checks (and standardisation of the units of measurement to the target units, see Appendix A). As such, we *do not* apply any 'gap filling' procedures in WoSIS, for example, when only the sand and silt fractions are reported, nor do we apply pedotransfer functions to derive soil hydrological properties. This next stage of data processing is seen as the responsibility of the data users (modellers) themselves, as the required functions or ways of depth-aggregating the layer data will vary with the projected use(s) of the standardised data (see Finke, 2006; Hendriks et al., 2016; Van Looy et al., 2017).

3.4 Providing measures for geographic and attribute accuracy

It is well known that 'soil observations used for calibration and interpolation are themselves not error-free' (Baroni et al., 2017; Cressie and Kornak, 2003; Folberth et al., 2016; Grimm and Behrens, 2010; Guevara et al., 2018; Hengl et al., 2017b; Heuvelink, 2014; Heuvelink and Brown, 2006). Hence, we provide measures for the geographic accuracy of the point locations as well as the accuracy of the laboratory measurements for possible consideration in digital soil mapping and subsequent earth system modelling (Dai et al., 2019).

All profile coordinates in WoSIS are presented according to the World Geodetic System (i.e. WGS84, EPSG code 4326). These coordinates were converted from a diverse range of national projections. Further, the source referencing may have been in decimal degrees (DD) or expressed in degrees, minutes, seconds (DMS) for both latitude and longitude. The (approximate) accuracy of georeferencing in WoSIS is given in decimal degrees. If the source only provided degree, minutes and seconds (DMS) then the geographic accuracy is set at 0.01, if seconds (DM) are missing at 0.1, and if seconds and minutes (D) are missing at 1. For most profiles (86 %, see Table 2), the approximate accuracy of the point locations, as inferred from the original coordinates given in the source datasets, is less than 10 m (total= 196,498 profiles, see Section 4). Digital soil mappers should duly consider the geometric accuracy in their applications (Grimm and Behrens, 2010), since the soil observations and covariates may not actually correspond (Cressie and Kornak, 2003).

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Earth System Science Data
Open Access Discussions

Table 2. Approximate accuracy of the profile locations

Decimal places	**Decimal degrees**	**Approximate precision**	**No. of profiles**
7	0.0000001	1 cm	1,345
6	0.000001	10 cm	84,945
5	0.00001	1 m	74,024
4	0.0001	10 m	9,158
3	0.001	100 m	8,108
2	0.01	1 km	10,915
1	0.1	10 km	6,458
0	1	100 km	1,545

After: https://en.wikipedia.org/wiki/Decimal_degrees

As indicated, soil data considered in WoSIS have been analysed according to a wide range of analytical procedures, and in different laboratories. An indication of the measurement uncertainty is thus desired; soil laboratory-specific Quality Management Systems (van Reeuwijk, 1998) as well as laboratory proficiency-testing (PT, Magnusson and Örnemark, 2014; Munzert et al., 2007; WEPAL, 2019) can provide this type of information. Yet, calculation of laboratory-specific measurement uncertainty for a single method, respectively multiple analytical methods, requires several measurement rounds (years of observation) and solid statistical analyses. Overall, such detailed information is not available for the data sets submitted to the ISRIC data repository. Therefore, out of necessity, we have distilled the desired information from the PT-literature (Kalra and Maynard, 1991; Rayment and Lyons, 2011; Rossel and McBratney, 1998; van Reeuwijk, 1983; WEPAL, 2019), in so far as technically feasible. For example, accuracy for bulk density measurements, both for the direct core and the clod method, has been termed 'low' (though not quantified) in a recent review (Al-Shammary et al., 2018); using expert-knowledge, we have assumed this corresponds with an uncertainty (or variability, expressed as coefficient of variation) of 35 %. Alternatively, for organic carbon content the mean variability was 17 % (with a range of 12 to 42 %) and for 'CEC buffered at pH 7' of 18 % (range 13 to 25%) when multiple

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

laboratories analyse a standard set of reference materials using similar operational methods (WEPAL, 2019). For soil pH measurements (log scale), we have expressed the uncertainty in terms of '± pH units'.

Importantly, the figures for measurement accuracy presented in Appendix A represent first approximations. They are based on the inter-laboratory comparison of well-homogenised, reference samples for a still relatively small range of soil types. These indicative figures should be refined once specific, laboratory and method-related accuracy (i.e. systematic and random error) information is provided with/for the shared soil data, for example using the procedures described by Eurachem (Magnusson and Örnemark, 2014). Alternatively, this type of information may be refined in the context of international laboratory PT-networks such as GLOSOLAN and WEPAL. Meanwhile, the present 'first' estimates may already be considered to calculate the accuracy of digital soil maps and of any interpretations derived from them (e.g. maps of soil organic carbon stocks in support of the UNCCD LDN (Land Degradation Neutrality) effort).

4 Spatial distribution of soil profiles and number of observations

The present snapshot includes standardised data for 196,498 profiles (Fig. 2), up from some 96,000 profiles for the preceding 'July 2016' snapshot. These are represented by some 832 thousand soil layers (or horizons). In total, this corresponds with over 5.8 million records that include both numeric (e.g. sand content, soil pH, and cation exchange capacity) as well as class (e.g. WRB soil classification and horizon designation) properties. The naming conventions and standard units of measurement are provided in Appendix A, and the file structure in Appendix B.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

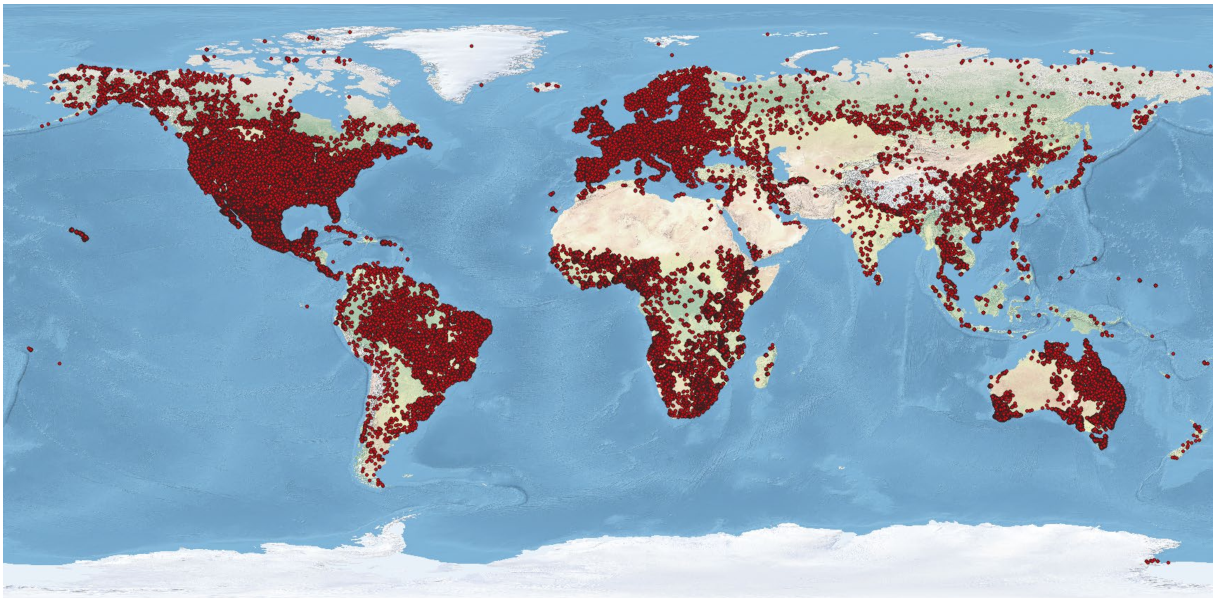

Figure 2. Location of soil profiles provided in the 'September 2019' snapshot of WoSIS.

(See Appendix C for the number and density of profiles by country)

The number of profiles per continent is highest for North America (73,604, was 63,077 for the preceding snapshot), followed by Oceania (42,918, was 235), Europe (35,311, was 1,908), Africa (27,688, was 17,153), South America (10,218, was 8,970), Asia (6,704, was 3,089), and Antarctica (9, was 9). These profiles come from 173 countries; the average density of observations is 1.35 profiles per 1000 km^2. The actual density of observations varies greatly, both between countries (Appendix C) and within each country, with the largest densities of 'shared' profiles reported for Belgium (228 profiles per 1000 km^2) and Switzerland (265 profiles per 1000 km^2). There are still relatively few profiles for Central Asia, South East Asia, Central and Eastern Europe, Russia, and the northern circumpolar region. The number of profiles by biome (Olson et al., 2001b) respectively broad climatic region (Sayre et al., 2014), as derived from GIS overlays, is provided in Appendix D for additional information.

There are more observations for the chemical data than the physical data (see Appendix A) and the number of observations generally decreases with depth, this largely depending on the objectives of the original soil surveys.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Earth System Science Data
Open Access Discussions

Present gaps in the geographic (Appendix C and D) and feature space (Appendix A, last column) will gradually be filled in the coming years, this largely depending though on the willingness or ability of data providers to share (some of) their data for consideration in WoSIS. For the northern Boreal and Arctic region, for example, ISRIC will regularly ingest new profile data collated by the International Soil Carbon Network (ISCN, Malhotra et al., 2019). Alternatively, it should be reiterated that for some regions, such as Europe (EU LUCAS database, see Tóth et al., 2013) and the state of Victoria (Australia), there are holdings in the ISRIC repository that may *only* be used/standardised for SoilGrids™ applications due to licence restrictions. Consequently, the corresponding profiles (~42 thousand) are *not* shown in Figure 2 nor are they considered in the descriptive statistics in Appendix C.

5 Distributing the standardised data

Upon their standardisation, the data are distributed through ISRIC's SDI (Spatial Data Infrastructure). This web platform is based on open source technologies and open web-services (WFS, WMS, WCS, CSW) following Open Geospatial Consortium (OGC) standards, and aimed specifically at handling soil data; our metadata are organised following standards of the International Organization for Standardization (ISO-28258, 2013) and INSPIRE (2015) compliant. The three main components of the SDI are: PostgreSql + PostGIS, GeoServer and GeoNetwork. Visualisation and data download are done in GeoNetwork with resources from GeoServer (https://data.isric.org). The third component is the PostgreSQL database, with the spatial extension PostGIS, in which WoSIS resides; the database is connected to GeoServer to permit data download from GeoNetwork. These processes are aimed at facilitating global data interoperability and citeability in compliance with FAIR principles: the data should be 'findable, accessible, interoperable, and reusable' (Wilkinson et al., 2016). With partners, steps are being undertaken towards the development of a federated, and ultimately interoperable, spatial soil data infrastructure (GLOSIS) through which source data are served and updated by the respective data providers, and made queryable according to a common SoilML standard (OGC, 2019).

The procedure for accessing the most current set of standardised soil profile data ('wosis_latest'), either from R or QGIS using WFS, is explained in a detailed tutorial (Rossiter, 2019). This data set is *dynamic*, hence it will grow when new

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

point data are shared and processed, additional soil attributes are considered in the WoSIS workflow, and/or when possible corrections are required. Potential errors may be reported on-line via a 'google group' so that they may be addressed in the dynamic version (register via: https://groups.google.com/forum/#!forum/isric-world-soil-information.)

For consistent citation purposes, we provide *static* snapshots of the standardised data, in tab-separated values format, with unique DOI's (digital object identifier); as indicated, this paper describes the second WoSIS snapshot.

6 Discussion

The above procedures describe standardisation according to operational definitions for soil properties. Importantly, it should be stressed here that the ultimate, desired full harmonisation to an agreed reference method Y, for example 'pH H_2O, 1:2.5 soil/water solution' for say all 'pH 1:x H_2O' measurements, will first become feasible once the target method (Y) for each property has been defined, and subsequently accepted by the international soil community. A next step would be to collate/develop 'comparative' data sets for each soil property (i.e., sets with samples analysed according to a given reference method (Y_i) and the corresponding national methods (X_j) for pedotransfer function development. In practice, however, such relationships will often be soil type and region specific (see Appendix C in GlobalSoilMap, 2015). Alternatively, according to GLOSOLAN (Suvannang et al., 2018, p. 10) "comparable and useful soil information (at the global level) will only be attainable once laboratories agree to follow common standards and norms". In such a collaborative process, it will be essential to consider the end user's requirements in terms of quality and applicability of the data for their specific purposes (i.e. fitness for intended use). Over the years, many organisations have developed respectively implemented analytical methods, and quality assurance systems, that are well suited for their countries (e.g., Soil Survey Staff, 2014a) or regions (Orgiazzi et al., 2018) and thus, pragmatically, may not be inclined to implement the anticipated GLOSOLAN standard analytical methods.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

7 Data availability

Snapshot 'WoSIS_2019_September' is archived for long-term storage at ISRIC – World Soil Information, the World Data Centre for Soils (WDC-Soils) of the ISC (International Council for Science, formerly ICSU) World Data System (WDS). It is freely accessible at https://dx.doi.org/10.17027/isric-wdcsoils.20190901 (Batjes et al., 2019). The zip file (154 Mb) includes a 'readme first' file that describes key aspects of the data set (see also Appendix B) with reference to the WoSIS Procedures Manual (Ribeiro et al., 2018), and the data itself in CSV format (1.8 Gb, decompressed) resp. GeoPackage format (2.2 Gb decompressed).

8 Conclusions

- The second WoSIS snapshot provides consistent, standardised data for some 196 thousand profiles worldwide. However, as described, there are still important gaps in the geographic and feature space. These will be addressed in future releases in collaboration with our partners.
- We will increasingly consider data derived by soil spectroscopy and emerging innovative methods. Further, long-term time series at defined locations will be sought to support space-time modelling of soil properties, such as changes in soil carbon stocks or soil salinity.
- We provide measures for geographic accuracy of the point data as well as a first approximation for the uncertainty associated with the operationally-defined analytical methods. This information may be used to assess uncertainty in digital soil mapping and earth system modelling efforts that draw on the present set of point data.
- Capacity building and cooperation among (inter)national soil institutes will be necessary to create and share ownership of the soil information newly derived from the shared data, and to strengthen the necessary expertise and capacity to further develop and test the world soil information service worldwide. Such activities may be envisaged within the broader framework of the Global Soil Partnership, and emerging GLOSIS system.

8 Appendices

Appendix A: Coding conventions, property names and their description of soil properties, units of measurement, inferred accuracy, and number of profiles and layers provided in the 'WoSIS September 2019' snapshot. (Soil properties are listed in alphabetical order of the property code)

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
Layer data						
BDFI33	Bulk density fine earth - 33 kPa	kg/dm³	14924	78215	Bulk density of the fine earth fraction[b], equilibrated at 33 kPa	35
BDFIAD	Bulk density fine earth - air dry	kg/dm³	1786	8471	Bulk density of the fine earth fraction, air dried	35
BDFIFM	Bulk density fine earth - field moist	kg/dm³	5279	14219	Bulk density of the fine earth fraction, field moist	35
BDFIOD	Bulk density fine earth - oven dry	kg/dm³	25124	122693	Bulk density of the fine earth fraction, oven dry	35
BDWS33	Bulk density whole soil - 33 kPa	kg/dm³	26268	154901	Bulk density of the whole soil including coarse fragments, equilibrated at 33 kPa	35
BDWSAD	Bulk density whole soil - air dry	kg/dm³	0	0	Bulk density of the whole soil including coarse fragments, air dried	35
BDWSFM	Bulk density whole soil - field moist	kg/dm³	0	0	Bulk density of the whole soil including coarse fragments, field moist	35
BDWSOD	Bulk density whole soil - oven dry	kg/dm³	14588	75422	Bulk density of the whole soil including coarse fragments, oven dry	35

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
CECPH7	Cation exchange capacity - buffered at pH7	cmol(c)/kg	54278	295688	Capacity of the fine earth fraction to hold exchangeable cations, estimated by buffering the soil at 'pH7'	20
CECPH8	Cation exchange capacity - buffered at pH8	cmol(c)/kg	6422	23691	Capacity of the fine earth fraction to hold exchangeable cations, estimated by buffering the soil at 'pH8'	20
CFGR	Coarse fragments gravimetric total	g/100g	39527	203083	Gravimetric content of coarse fragments in the whole soil	20
CFVO	Coarse fragments volumetric total	cm³/100cm³	45918	235002	Volumetric content of coarse fragments in the whole soil	30
CLAY	Clay total	g/100g	141640	607861	Gravimetric content of < X mm soil material in the fine earth fraction (e.g. X = 0.002 mm as specified in the analytical method description) [b c]	15
ECEC	Effective cation exchange capacity	cmol(c)/kg	31708	132922	Capacity of the fine earth fraction to hold exchangeable cations at the pH of the soil (ECEC). Conventionally approximated by summation of exchangeable bases (Ca^{2+}, Mg^{2+}, K^{+}, and Na^{+}) plus 1 N KCl exchangeable acidity (Al^{3+} and H^{+}) in acidic soils	25

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
ELCO20	Electrical conductivity - ratio 1:2	dS/m	8010	44596	Ability of a 1:2 soil water extract to conduct electrical current	10
ELCO25	Electrical conductivity - ratio 1:2.5	dS/m	3313	15134	Ability of a 1:2.5 soil water extract to conduct electrical current	10
ELCO50	Electrical conductivity - ratio 1:5	dS/m	23093	90944	Ability of a 1:5 soil water extract to conduct electrical current	10
ELCOSP	Electrical conductivity - saturated paste	dS/m	19434	73517	Ability of a water saturated soil paste to conduct electrical current (EC_e)	10
NITKJD	Total nitrogen (N)	g/kg	65356	216362	The sum of total Kjeldahl nitrogen (ammonia, organic and reduced nitrogen) and nitrate-nitrite	10
ORGC	Organic carbon	g/kg	110856	471301	Gravimetric content of organic carbon in the fine earth fraction	15
PHAQ	pH H2O	unitless	130986	613322	A measure of the acidity or alkalinity in soils, defined as the negative logarithm (base 10) of the activity of hydronium ions (H^+) in water	0.3
PHCA	pH CaCl2	unitless	66921	314230	A measure of the acidity or alkalinity in soils, defined as the negative logarithm (base 10) of the activity of hydronium ions (H^+) in a	0.3

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
					$CaCl_2$ solution, as specified in the analytical method descriptions	
PHKC	pH KCl	unitless	32920	150447	A measure of the acidity or alkalinity in soils, defined as the negative logarithm (base 10) of the activity of hydronium ions (H^+) in a KCl solution, as specified in the analytical method descriptions	0.3
PHNF	pH NaF	unitless	4978	25448	A measure of the acidity or alkalinity in soils, defined as the negative logarithm (base 10) of the activity of hydronium ions (H^+) in a NaF solution, as specified in the analytical method descriptions	0.3
PHPBYI	Phosphorus (P) - Bray I	mg/kg	10735	40486	Measured according to the Bray-I method, a combination of HCl and NH_4F to remove easily acid soluble P forms, largely Al- and Fe-phosphates (for acid soils)	40
PHPMH3	Phosphorus (P) - Mehlich 3	mg/kg	1446	7242	Measured according to the Mehlich-3 extractant, a combination of acids (acetic [HOAc] and nitric [HNO_3]), salts (ammonium fluoride [NH4F] and ammonium nitrate	25

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
					[NH4 NO3]), and the chelating agent ethylene-diaminetetraacetic acid (EDTA); considered suitable for removing P and other elements in acid and neutral soils	
PHPOLS	Phosphorus (P) - Olsen	mg/kg	2162	8434	Measured according to the P-Olsen method: 0.5 M sodium bicarbonate ($NaHCO_3$) solution at a pH of 8.5 to extract P from calcareous, alkaline, and neutral soils	25
PHPRTN	Phosphorus (P) - retention	mg/kg	4636	23917	Retention measured according to the New Zealand method	20
PHPTOT	Phosphorus (P) - total	mg/kg	4022	12976	Determined with a very strong acid (*aqua regia* and sulfuric acid/nitric acid)	15
PHPWSL	Phosphorus (P) - water soluble	mg/kg	283	1242	Measured in 1:x soil:water solution (mainly determines P in dissolved forms)	15
SAND	Sand total	g/100g	105547	491810	Y to Z mm fraction of the fine earth fraction; Z upper limit as specified in the analytical method description for the sand fraction (e.g. Y = 0.05 mm to Z= 2 mm) [c]	15
SILT	Silt total	g/100g	133938	575913	X to Y mm fraction of the fine earth fraction; X upper limit as	15

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
					specified in the analytical method description for the clay fraction (e.g. X= 0.002 mm to Y = 0.05 mm) [c]	
TCEQ	Calcium carbonate equivalent total	g/kg	51991	222242	The content of carbonate in a liming material or calcareous soil calculated as if all of the carbonate is in the form of $CaCO_3$ (in the fine earth fraction); also known as inorganic carbon	10
TOTC	Total carbon (C)	g/kg	32662	109953	Gravimetric content of organic carbon and inorganic carbon in the fine earth fraction	10
WG0006	Water retention gravimetric - 6 kPa	g/100g	863	4264	Soil moisture content by weight, at tension 6 kPa (pF 1.8)	20
WG0010	Water retention gravimetric - 10 kPa	g/100g	3357	14739	Soil moisture content by weight, at tension 10 kPa (pF 2.0)	20
WG0033	Water retention gravimetric - 33 kPa	g/100g	21116	96354	Soil moisture content by weight, at tension 33 kPa (pF 2.5)	20
WG0100	Water retention gravimetric - 100 kPa	g/100g	696	3762	Soil moisture content by weight, at tension 100 kPa (pF 3.0)	20
WG0200	Water retention gravimetric - 200 kPa	g/100g	4418	28239	Soil moisture content by weight, at tension 200 kPa (pF 3.3)	20

Earth System Science Data
Open Access Discussions

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
WG0500	Water retention gravimetric - 500 kPa	g/100g	344	1716	Soil moisture content by weight, at tension 500 kPa (pF 3.7)	20
WG1500	Water retention gravimetric - 1500 kPa	g/100g	34365	187176	Soil moisture content by weight, at tension 1500 kPa (pF 4.2)	20
WV0006	Water retention volumetric - 6 kPa	cm³/100cm³	9	26	Soil moisture content by volume, at tension 6 kPa (pF 1.8)	20
WV0010	Water retention volumetric - 10 kPa	cm³/100cm³	1469	5434	Soil moisture content by volume, at tension 10 kPa (pF 2.0)	20
WV0033	Water retention volumetric - 33 kPa	cm³/100cm³	5987	17801	Soil moisture content by volume, at tension 33 kPa (pF 2.5)	20
WV0100	Water retention volumetric - 100 kPa	cm³/100cm³	747	2559	Soil moisture content by volume, at tension 100 kPa (pF 3.0)	20
WV0200	Water retention volumetric - 200 kPa	cm³/100cm³	3	9	Soil moisture content by volume, at tension 200 kPa (pF 3.3)	20
WV0500	Water retention volumetric - 500 kPa	cm³/100cm³	703	1763	Soil moisture content by volume, at tension 500 kPa (pF 3.7)	20
WV1500	Water retention volumetric - 1500 kPa	cm³/100cm³	6149	17542	Soil moisture content by volume, at tension 1500 kPa (pF 4.2)	20
Site data						

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Code	Property	Units	Profiles	Layers	Description	Accuracy (± %)[a]
CSTX	Soil classification Soil taxonomy	classes	21314	n/a	Classification of the soil profile according to specified edition (year) of USDA Soil Taxonomy, up to subgroup level when available	-
CWRB	Soil classification WRB	classes	26664	n/a	Classification of the soil profile according to specified edition (year) of the World Reference Base for Soil Resources (WRB), up to qualifier level when available	-
CFAO	Soil classification FAO	classes	23890	n/a	Classification of the soil profile according to specified edition (year) of the FAO-Unesco Legend, up to soil unit level when available	-
DSDS	Depth of soil - sampled	cm	196381	n/a	Maximum depth of soil described and sampled (calculated)	-
HODS	Horizon designation	-	80,849	396,522	Horizon designation as provided in the source database[d]	

[a] Inferred accuracy (or uncertainty), rounded to the nearest 5%, unless otherwise indicated (i.e. units for soil pH) as derived from the following sources (Al-Shammary et al., 2018; Kalra and Maynard, 1991; Rayment and Lyons, 2011; Rossel and McBratney, 1998; van Reeuwijk, 1983; WEPAL, 2019). These figures are first approximations that will be fine-tuned once more specific results of laboratory proficiency tests, resp. national Soil Quality Management systems, become available.

[b] Generally, the fine earth fraction is defined as being < 2 mm. Alternatively, an upper limit of 1 mm was used in the former Soviet Union and its satellite states (Katchynsky scheme). This has been indicated in file 'wosis_201907_layers_chemical.csv' and 'wosis_201907_layer_physicals.csv' for those soil properties where this differentiation is important (see 'sample pretreatment' in string 'xxxx_method', Appendix B).

[c] Provided only when the sum of clay, silt and sand fraction is ≥ 90 and ≤ 100 percent.

[d] Where available, the 'cleaned' (original) layer/horizon designation is provided for general information; these codes have not been standardised as they vary widely between different classification systems (Bridges, 1993; Gerasimova et al., 2013). When horizon designations are not provided in the source data bases, we have flagged all layers with an upper depth given as being negative (e.g. -10 to 0 cm that is using pre-1993 conventions; see text and WoSIS Procedures Manual 2018, p. 24, footnote 9) in the source databases as likely being 'litter' layers.

Appendix B: Structure of the 'September 2019' WoSIS snapshot

This Appendix describes the structure of the data files presented in the 'September 2019' WoSIS snapshot:

- *wosis_201909_attributes.csv*,
- *wosis_201909_profiles.csv*,
- *wosis_201909_layers_chemical.csv, and*
- *wosis_201909_layer_physicals.csv*.

wosis_201909_attributes.csv: This file lists the four to six letter code for each attribute, whether the attribute is a site or horizon property, the unit of measurement, the number of profiles respectively layers represented in the snapshot, and a brief description of each attribute, as well as the inferred uncertainty for each property (Appendix A).

wosis_201909_profiles.csv: This file contains the unique profile ID (i.e. primary key), the source of the data, country ISO code and name, accuracy of geographical coordinates, latitude and longitude (WGS 1984), point geometry of the location of the profile, maximum depth of soil described and sampled, as well as information on the soil classification system and edition. Depending on the soil classification system used, the number of fields will vary. For example, for the World Soil Reference Base (WRB) system these are: publication_year (i.e. version), reference_soil_group_code,

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

reference_soil_group_name, and the name(s) of the prefix (primary) qualifier(s) respectively suffix (supplementary) qualifier(s). The terms principal qualifier and supplementary qualifier are currently used (IUSS Working Group WRB, 2015); earlier WRB versions used prefix and suffix for this (e.g. IUSS Working Group WRB, 2006). Alternatively, for USDA Soil Taxonomy, the version (year), order, suborder, great group, and subgroup can be accommodated (Soil Survey Staff, 2014b). Inherently, the number of records filled will vary between (and within) the various source databases.

The corresponding field names are listed below:

profile_id	Primary key
dataset_id	Identifier for source data set
country_id	ISO code for country name
country_name	Country name (in English)
geom_accuracy	Accuracy of the geographical coordinates in degrees. Example: If degree, minutes and seconds are provided in the source then geom_accuracy is set at 0.01, if seconds are missing at 0.1, and if seconds and minutes are missing at 1.
latitude	Latitude in degrees (WGS84)
longitude	Longitude in degrees (WGS84)
geom	Point geometry of the location of the profile (WGS84)
dsds	Maximum depth of soil described and sampled (calculated)
cfao_version	Version of FAO Legend (e.g. 1974 or 1988)
cfao_major_group_code	Code for major group (in given version of the Legend),

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

cfao_major_group	Name of major group
cfao_soil_unit_code	Code for soil unit
cfao_soil_unit	Name of soil unit
cwrb_version	Version of World Reference Base for Soil Resources
cwrb_reference_soil_group_code	Code for WRB group (in given version of WRB)
cwrb_reference_soil_group	Full name for reference soil group
cwrb_prefix_qualifier	Name for prefix (e.g. for WRB1988) resp. principal qualifier (e.g. for WRB2015)
cwrb_suffix_qualifier	Name for suffix (e.g. for WRB1988) resp. supplementary qualifier (e.g. for WRB2015)
cstx_version	Version of USDA Soil Taxonomy (UST)
cstx_order_name	Name of UST order
cstx_suborder	Name of UST suborder
cstx_great_group	Name of UST greatgroup
cstx_subgroup	Name of UST subgroup

wosis_201909_layer_chemical.csv and *wosis_201909_layer_physical.csv:* The layer (horizon) data are presented in two separate file in view of their size, one for the chemical and one for the physical soil properties. The file structure, however, is identical:

profile_id	identifier for profile, foreign key to 'wosis_201909_profiles'

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

profile_layer_id	unique identifier for layer for given profile (primary key)
upper_depth	upper depth of layer (or horizon)
lower_depth	lower depth of layer
layer_name	name of the horizon, as provided in the source data
litter	flag (Boolean), indicating whether this is considered a surficial litter layer

Subsequently, the following items are listed sequentially per attribute ('xxxx') as defined under 'code' in file *wosis_201909_attributes.csv*:

xxxx_value	array listing all measurement values for soil property 'xxxx' for the given layer. In some cases, more than one observation is reported for a given horizon (layer) in the source, for example 4 values for TOTC: [1:5.4, 2:8.2, 3:6.3, 4:7.7]
xxxx _value_avg	average, for above (it is recommended to use this value for 'routine' modelling)
xxxx _method	array listing the method descriptions for each value. The nature of this array varies with the soil property under consideration as described in the option tables for each analytical method. For example, in the case of electrical conductivity (ELCO), the method is described using: sample pretreatment (e.g. sieved over 2 mm size, solution (e.g. water), ratio (e.g., 1:5), and ratio base (e.g. weight /volume). Details for each method are provided in the WoSIS Procedures Manual (Appendix D, E and F in Ribeiro et al., 2018).
xxxx _date	array listing the date of observation for each value

xxxx _dataset_id	abbreviation for source data set (e.g. WD-ISCN)
xxxx _profile_code	code for given profile (provides the link to profile_id in *wosis_201909_profiles.csv)*
xxxx _license	licence for given data, as indicated by the data provider (e.g. CC-BY)
(...)	as above, but for the next attribute (for full list see Appendix A)

Format: All fields in the above files are tab-delimited, with double quotation marks as text delimiters. File coding is according to the UTF-8 unicode transformation format.

Using the data: The above csv files can easily be imported into an SQL database or statistical software such as R, after which they may be joined using the unique profile_id. Guidelines for handling and querying the data are provided in the WoSIS Procedures Manual (Ribeiro et al. 2018, p. 45-48); see also the detailed tutorial by Rossiter (2019).

Appendix C: Number of profiles by country and continent.

Continent	**Country_name**	**ISO code**	**No. of profiles**	**Area (km^2)**	**Profile density (per 1000 km^2)**
Africa	Algeria	DZ	10	2308647	0.004
	Angola	AO	1169	1246690	0.938
	Benin	BJ	744	115247	6.456

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	Botswana	BW	994	578247	1.719
	Burkina Faso	BF	2023	273281	7.403
	Burundi	BI	1063	26857	39.58
	Cameroon	CM	1306	465363	2.806
	Central African Republic	CF	88	619591	0.142
	Chad	TD	7	1265392	0.006
	Congo	CG	71	340599	0.208
	Côte d'Ivoire	CI	255	321762	0.793
	Democratic Republic of the Congo	CD	380	2329162	0.163
	Egypt	EG	26	982161	0.026
	Ethiopia	ET	1712	1129314	1.516
	Gabon	GA	47	264022	0.178
	Ghana	GH	432	238842	1.809
	Guinea	GN	128	243023	0.527
	Guinea-Bissau	GW	18	30740	0.586
	Kenya	KE	1601	582342	2.749
	Lesotho	LS	33	30453	1.084
	Liberia	LR	50	96103	0.52
	Libya	LY	14	1620583	0.009
	Madagascar	MG	131	588834	0.222
	Malawi	MW	3049	118715	25.683

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	Mali	ML	884	1251471	0.706
	Mauritania	MR	13	1038527	0.013
	Morocco	MA	113	414030	0.273
	Mozambique	MZ	566	787305	0.719
	Namibia	NA	1462	823989	1.774
	Niger	NE	520	1182602	0.44
	Nigeria	NG	1402	908978	1.542
	Rwanda	RW	2007	25388	79.052
	Senegal	SN	312	196200	1.59
	Sierra Leone	SL	12	72281	0.166
	Somalia	SO	245	632562	0.387
	South Africa	ZA	874	1220127	0.716
	South Sudan	SS	82	629821	0.13
	Sudan	SD	130	1843196	0.071
	Swaziland	SZ	14	17290	0.81
	Togo	TG	9	56767	0.159
	Tunisia	TN	60	155148	0.387
	Uganda	UG	683	241495	2.828
	United Republic of Tanzania	TZ	1915	939588	2.038
	Zambia	ZM	601	751063	0.8
	Zimbabwe	ZW	413	390648	1.057

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
Antarctica	Antarctica	AQ	9	12537967	0.001
Asia	Afghanistan	AF	19	641827	0.03
	Armenia	AM	7	29624	0.236
	Arunachal Pradesh	[a]	2	67965	0.029
	Azerbaijan	AZ	24	164780	0.146
	Bahrain	BH	2	673	2.97
	Bangladesh	BD	207	139825	1.48
	Bhutan	BT	85	37674	2.256
	Cambodia	KH	409	181424	2.254
	China	CN	1648	9345214	0.176
	Cyprus	CY	12	9249	1.297
	Georgia	GE	17	69785	0.244
	Hong Kong	HK	2	1081	1.851
	India	IN	199	2961118	0.067
	Indonesia	ID	180	1888620	0.095
	Iran (Islamic Republic of)	IR	2010	1677319	1.198
	Iraq	IQ	14	435864	0.032
	Israel	IL	17	20720	0.82
	Jammu and Kashmir	[a]	4	186035	0.022
	Japan	JP	198	373651	0.53
	Jordan	JO	47	89063	0.528

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	Kazakhstan	KZ	12	2841103	0.004
	Kuwait	KW	1	17392	0.057
	Kyrgyzstan	KG	1	199188	0.005
	Lao People's Democratic Republic	LA	20	230380	0.087
	Lebanon	LB	10	10136	0.987
	Malaysia	MY	157	329775	0.476
	Mongolia	MN	9	1564529	0.006
	Nepal	NP	142	147437	0.963
	Occupied Palestinian Territory	PS	18	6225	2.892
	Oman	OM	9	308335	0.029
	Pakistan	PK	45	788439	0.057
	Philippines	PH	81	296031	0.274
	Republic of Korea	KR	23	99124	0.232
	Saudi Arabia	SA	7	1925621	0.004
	Singapore	SG	1	594	1.683
	Sri Lanka	LK	72	66173	1.088
	Syrian Arab Republic	SY	68	188128	0.361
	Taiwan	TW	35	36127	0.969
	Tajikistan	TJ	5	142004	0.035
	Thailand	TH	482	515417	0.935
	Turkey	TR	69	781229	0.088

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	United Arab Emirates	AE	12	71079	0.169
	Uzbekistan	UZ	9	449620	0.02
	Viet Nam	VN	29	327575	0.089
	Yemen	YE	284	453596	0.626
Europe	Albania	AL	97	28682	3.382
	Austria	AT	128	83964	1.524
	Belarus	BY	92	207581	0.443
	Belgium	BE	7009	30669	228.536
	Bosnia and Herzegovina	BA	32	51145	0.626
	Bulgaria	BG	136	111300	1.222
	Croatia	HR	78	56589	1.378
	Czech Republic	CZ	664	78845	8.422
	Denmark	DK	74	44458	1.664
	Estonia	EE	242	45441	5.326
	Finland	FI	444	336892	1.318
	France	FR	1037	548785	1.89
	Germany	DE	4345	357227	12.163
	Greece	GR	370	132549	2.791
	Hungary	HU	1420	93119	15.249
	Iceland	IS	11	102566	0.107
	Ireland	IE	125	69809	1.791

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	Italy	IT	575	301651	1.906
	Latvia	LV	102	64563	1.58
	Lithuania	LT	127	64943	1.956
	Luxembourg	LU	141	2621	53.802
	Montenegro	ME	12	13776	0.871
	Netherlands	NL	320	35203	9.09
	Norway	NO	507	324257	1.564
	Poland	PL	618	311961	1.981
	Portugal	PT	460	91876	5.007
	Republic of Moldova	MD	35	33798	1.036
	Romania	RO	104	238118	0.437
	Russian Federation	RU	1410	16998830	0.083
	Serbia	RS	69	88478	0.78
	Slovakia	SK	161	49072	3.281
	Slovenia	SI	67	20320	3.297
	Spain	ES	905	505752	1.789
	Svalbard and Jan Mayen Islands	SJ	4	63464	0.063
	Sweden	SE	583	449212	1.298
	Switzerland	CH	10943	41257	265.238
	The former Yugoslav Republic of Macedonia	MK	20	25424	0.787
	Ukraine	UA	409	600526	0.681

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Continent	**Country_name**	**ISO code**	**No. of profiles**	**Area (km^2)**	**Profile density (per 1000 km^2)**
	United Kingdom	GB	1435	244308	5.874
Northern America	Barbados	BB	3	433	6.928
	Belize	BZ	29	21764	1.332
	Canada	CA	8516	9875646	0.862
	Costa Rica	CR	560	51042	10.971
	Cuba	CU	53	110863	0.478
	Dominican Republic	DO	10	48099	0.208
	El Salvador	SV	38	20732	1.833
	Greenland	GL	6	2165159	0.003
	Guadeloupe	GP	5	1697	2.947
	Guatemala	GT	27	109062	0.248
	Honduras	HN	38	112124	0.339
	Jamaica	JM	76	10965	6.931
	Mexico	MX	7554	1949527	3.875
	Netherlands Antilles	AN	4	790	5.066
	Nicaragua	NI	26	128376	0.203
	Panama	PA	51	74850	0.681
	Puerto Rico	PR	280	8937	31.329
	Trinidad and Tobago	TT	2	5144	0.389
	United States of America	US	56277	9315946	6.041
	United States Virgin Islands	VI	49	352	139.069

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
Oceania	Australia	AU	42758	7687634	5.562
	Cook Islands	CK	1	241	4.142
	Fiji	FJ	9	18293	0.492
	Guam	GU	15	544	27.579
	Micronesia (Federated States of)	FM	78	740	105.397
	New Caledonia	NC	2	18574	0.108
	New Zealand	NZ	53	270415	0.196
	Palau	PW	18	451	39.924
	Papua New Guinea	PG	31	462230	0.067
	Samoa	WS	17	2835	5.996
	Solomon Islands	SB	1	28264	0.035
	Vanuatu	VU	1	12236	0.082
South America	Argentina	AR	244	2780175	0.088
	Bolivia (Plurinational State of)	BO	86	1084491	0.079
	Brazil	BR	8883	8485946	1.047
	Chile	CL	72	753355	0.096
	Colombia	CO	237	1137939	0.208
	Ecuador	EC	94	256249	0.367
	French Guiana	GF	30	83295	0.36
	Guyana	GY	43	211722	0.203
	Paraguay	PY	1	399349	0.003

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Continent	Country_name	ISO code	No. of profiles	Area (km^2)	Profile density (per 1000 km^2)
	Peru	PE	159	1290640	0.123
	Suriname	SR	31	145100	0.214
	Uruguay	UY	132	177811	0.742
	Venezuela (Bolivarian Republic of)	VE	206	912025	0.226

[a] Disputed territories. Country names and areas are based on the Global Administrative Layers (GAUL) database, see: http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691.

5 **Appendix D. Distribution of soil profiles by eco-region and by biome**

A) Number of soil profiles by broad rainfall and temperature zone[a]

Temperature zone / Rainfall zone	Cold	Cool	Warm	Hot
Wet	19,850	3	29,448	3,3151
Moist	2,414	4,308	6,860	10,718
Semi-dry	676	7,098	14,778	22,501
Dry	15	226	1,032	2,673

[a] Bioclimatic zones as defined by Sayre et al. (2014). Arctic zone (not shown in Table), two profiles.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

B) Number of soil profiles by biome[b]

Biome	No. of profiles
Boreal Forests/Taiga	6,129
Deserts & Xeric Shrublands	10,212
Flooded Grasslands & Savannas	779
Mangroves	682
Mediterranean Forests, Woodlands & Scrub	16,759
Montane Grasslands & Shrublands	1,402
Temperate Broadleaf & Mixed Forests	63,912
Temperate Conifer Forests	12,153
Temperate Grasslands, Savannas & Shrublands	25,357
Tropical & Subtropical Coniferous Forests	1,354
Tropical & Subtropical Dry Broadleaf Forests	3,808
Tropical & Subtropical Grasslands, Savannas & Shrublands	34,779
Tropical & Subtropical Moist Broadleaf Forests	16,492
Tundra	1,977
No data	703

[a] Biomes defined according to 'Terrestrial Ecoregions of the World' (TEOW) (Olson et al., 2001a).

9 Competing interests. The authors declare that they have no conflict of interest.

10 Acknowledgements

The development of WoSIS has been made possible thanks to the contributions and shared knowledge of a steadily growing number of data providers, including soil survey organisations, research institutes and individual experts, for which we are grateful; for an overview please see https://www.isric.org/explore/wosis/wosis-cooperating-institutions-and-experts. We thank our colleagues Laura Poggio, Luis de Sousa and Bas Kempen for their constructive comments on a 'pre-release' of the snapshot data.

ISRIC − World Soil Information, legally registered as International Soil Reference and Information Centre, receives core funding from the Dutch Government.

References

Al-Shammary, A. A. G., Kouzani, A. Z., Kaynak, A., Khoo, S. Y., Norton, M., and Gates, W.: Soil Bulk Density Estimation Methods: A Review, Pedosphere, 28, 581-596, https://doi.org/10.1016/S1002-0160(18)60034-7, 2018.

Arrouays, D., Leenaars, J. G. B., Richer-de-Forges, A. C., Adhikari, K., Ballabio, C., Greve, M., Grundy, M., Guerrero, E., Hempel, J., Hengl, T., Heuvelink, G., Batjes, N., Carvalho, E., Hartemink, A., Hewitt, A., Hong, S.-Y., Krasilnikov, P., Lagacherie, P., Lelyk, G., Libohova, Z., Lilly, A., McBratney, A., McKenzie, N., Vasquez, G. M., Mulder, V. L., Minasny, B., Montanarella, L., Odeh, I., Padarian, J., Poggio, L., Roudier, P., Saby, N., Savin, I., Searle, R., Solbovoy, V., Thompson, J., Smith, S., Sulaeman, Y., Vintila, R., Rossel, R. V., Wilson, P., Zhang, G.-L., Swerts, M., Oorts, K., Karklins, A., Feng, L., Ibelles Navarro, A. R., Levin, A., Laktionova, T., Dell'Acqua, M., Suvannang, N., Ruam, W., Prasad, J., Patil, N., Husnjak, S., Pásztor, L., Okx, J., Hallet, S., Keay, C., Farewell, T., Lilja, H., Juilleret, J., Marx, S., Takata, Y., Kazuyuki, Y., Mansuy, N., Panagos, P., Van Liedekerke, M., Skalsky, R., Sobocka, J., Kobza, J., Eftekhari, K., Alavipanah, S. K., Moussadek, R., Badraoui, M., Da Silva, M., Paterson, G., Gonçalves, M. d. C., Theocharopoulos, S., Yemefack, M., Tedou, S., Vrscaj, B., Grob, U., Kozák, J., Boruvka, L., Dobos, E., Taboada, M., Moretti, L., and Rodriguez, D.: Soil legacy data rescue via GlobalSoilMap and other international and national initiatives, GeoResJ, 14, 1-19, https://doi.org/10.1016/j.grj.2017.06.001, 2017.

Baroni, G., Zink, M., Kumar, R., Samaniego, L., and Attinger, S.: Effects of uncertainty in soil properties on simulated hydrological states and fluxes at different spatio-temporal scales, Hydrol. Earth Syst. Sci., 21, 2301-2320, https://www.hydrol-earth-syst-sci.net/21/2301/2017/, 2017.

Batjes, N. H.: Harmonized soil profile data for applications at global and continental scales: updates to the WISE database, Soil Use and Management, 25, 124-127 http://dx.doi.org/10.1111/j.1475-2743.2009.00202.x (supplemental information: https://www.isric.org/sites/default/files/isric_report_2008_02.pdf), 2009.

Batjes, N. H.: Harmonised soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks, Geoderma, 269, 61-68, http://dx.doi.org/10.1016/j.geoderma.2016.01.034 2016.

Batjes, N. H., Ribeiro, E., van Oostrum, A., Leenaars, J., Hengl, T., and Mendes de Jesus, J.: WoSIS: providing standardised soil profile data for the world, Earth Syst. Sci. Data, 9, 1-14, http://dx.doi.org/10.5194/essd-9-1-2017, 2017.

Batjes, N. H., Ribeiro, E., and van Oostrum, A. J. M.: Standardised soil profile data for the world (WoSIS snapshot - September 2019), ISRIC WDC-Soils, https://dx.doi.org/10.17027/isric-wdcsoils.20190901, 2019.

Bridges, E. M.: Soil horizon designations: past use and future prospects, CATENA, 20, 363-373, https://doi.org/10.1016/S0341-8162(05)80002-5, 1993.

Cressie, N., and Kornak, J.: Spatial Statistics in the Presence of Location Error with an Application to Remote Sensing of the Environment, Statist. Sci., 18, 436-456, https://projecteuclid.org:443/euclid.ss/1081443228, 2003.

Dai, Y., Shangguan, W., Wang, D., Wei, N., Xin, Q., Yuan, H., Zhang, S., Liu, S., and Yan, F.: A review on the global soil datasets for earth system modeling, SOIL, 5, 137-158, https://doi.org/10.5194/soil-5-137-2019, 2019.

FAO-ISRIC: Guidelines for soil description (3rd Edition, Rev.), FAO, Rome, 70 pp., 1986.

FAO: Guidelines for the description of soils, FAO, Rome, 1977.

FAO: Guidelines for soil description (Fourth ed.), FAO, Rome, 97, 2006.

FAO, IIASA, ISRIC, ISSCAS, and JRC: Harmonized World Soil Database (version 1.2), Prepared by Nachtergaele FO, van Velthuizen H, Verelst L, Wiberg D, Batjes NH, Dijkshoorn JA, van Engelen VWP, Fischer G, Jones A, Montanarella L., Petri M, Prieler S, Teixeira E and Xuezheng Shi. Food and Agriculture Organization of the United Nations (FAO),

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

International Institute for Applied Systems Analysis (IIASA), ISRIC - World Soil Information, Institute of Soil Science - Chinese Academy of Sciences (ISSCAS), Joint Research Centre of the European Commission (JRC), Laxenburg, Austria, 2012.

Finke, P.: Quality assessment of digital soil maps: producers and users perspectives, in: Digital soil mapping: An introductory perspective, edited by: Lagacherie, P., McBratney, A., and Voltz, M., Elsevier, Amsterdam, 523-541, 2006.

Folberth, C., Skalsky, R., Moltchanova, E., Balkovic, J., Azevedo, L. B., Obersteiner, M., and van der Velde, M.: Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations, 7, http://dx.doi.org/10.1038/ncomms11872, 2016.

Gerasimova, M. I., Lebedeva, I. I., and Khitrov, N. B.: Soil horizon designation: State of the art, problems, and proposals, Eurasian Soil Science, 46, 599-609, https://dx.doi.org/10.1134/S1064229313050037, 2013.

GlobalSoilMap: Specifications Tiered GlobalSoilMap products (Release 2.4), 52, 2015.

Grimm, R., and Behrens, T.: Uncertainty analysis of sample locations within digital soil mapping approaches, Geoderma, 155, 154-163, https://doi.org/10.1016/j.geoderma.2009.05.006, 2010.

Towards the implementation of GloSIS through a Country Soil Information Systems (CountrySIS) Framework (Concept Note, draft). Prepared by GSP Pillar 4 Working Group: http://www.fao.org/global-soil-partnership/pillars-action/4-information-data/glosis/en/, access: 26 November, 2018.

Guevara, M., Olmedo, G. F., Stell, E., Yigini, Y., Aguilar Duarte, Y., Arellano Hernández, C., Arévalo, G. E., Arroyo-Cruz, C. E., Bolivar, A., Bunning, S., Bustamante Cañas, N., Cruz-Gaistardo, C. O., Davila, F., Dell Acqua, M., Encina, A., Figueredo Tacona, H., Fontes, F., Hernández Herrera, J. A., Ibelles Navarro, A. R., Loayza, V., Manueles, A. M., Mendoza Jara, F., Olivera, C., Osorio Hermosilla, R., Pereira, G., Prieto, P., Alexis Ramos, I., Rey Brina, J. C., Rivera, R., Rodríguez-Rodríguez, J., Roopnarine, R., Rosales Ibarra, A., Rosales Riveiro, K. A., Schulz, G. A., Spence, A., Vasques, G. M., Vargas, R. R., and Vargas, R.: No Silver Bullet for Digital Soil Mapping: Country-specific Soil Organic Carbon Estimates across Latin America, SOIL, 2018, 173-193, https://doi.org/10.5194/soil-4-173-2018, 2018.

Hendriks, C. M. J., Stoorvogel, J. J., and Claessens, L.: Exploring the challenges with soil data in regional land use analysis, Agricultural Systems, 144, 9-21, http://dx.doi.org/10.1016/j.agsy.2016.01.007, 2016.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Hengl, T., Leenaars, J. G. B., Shepherd, K. D., Walsh, M. G., Heuvelink, G. B. M., Mamo, T., Tilahun, H., Berkhout, E., Cooper, M., Fegraus, E., Wheeler, I., and Kwabena, N. A.: Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning, Nutrient Cycling in Agroecosystems, https://doi.org/10.1007/s10705-017-9870-x, 2017a.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS ONE, 12, e0169748, http://dx.doi.org/10.1371/journal.pone.0169748, 2017b.

Heuvelink, G. B. M., and Brown, J. D.: Towards a soil information system for uncertain soil data in: Digital soil mapping: An introductory perspective, edited by: Lagacherie, P., McBratney, A., and Voltz, M., Elsevier, Amsterdam, 97-106, 2006.

Heuvelink, G. B. M.: Uncertainty quantification of GlobalSoilMap products in: GlobalSoilMap. Basis of the Global Spatial Soil Information System, edited by: Arrouays, D., McKenzie, N., Hempel, J., Forges, A. R. d., and McBratney, A., Taylor & Francis Group, London, UK, 335-240, 2014.

INSPIRE Data specifications - Infrastructure for spatial information in the European Community: http://inspire.ec.europa.eu/index.cfm/pageid/2, access: 25.04.2016, 2015.

Soil quality — Digital exchange of soil-related data (ISO 28258:2013(en)): https://www.iso.org/obp/ui#iso:std:iso:28258:ed-1:v1:en, access: 31/01/2018, 2013.

Data and Software Policy: http://www.isric.org/sites/default/files/ISRIC_Data_Policy_2016jun21.pdf access: 15 May 2019, 2016.

IUSS Working Group WRB: World Reference Base for Soil Resources (2nd ed.), FAO, RomeWorld Soil Resources Report 103, 145, 2006.

IUSS Working Group WRB: World Reference Base for soil resources 2014 - International soil classification system for naming soils and creating legends for soil maps (update 2015), Global Soil Partnership, International Union of Soil Sciences, and Food and Agriculture Organization of the United Nations, RomeWorld Soil Resources Reports 106, 182 2015.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Kalra, Y. P., and Maynard, D. G.: Methods manual for forest soil and plant analysis, Forestry Canada, Edmonton (Alberta), 116, 1991.

Leenaars, J. G. B., van Oostrum, A. J. M., and Ruiperez Gonzalez, M.: Africa Soil Profiles Database: A compilation of georeferenced and standardised legacy soil profile data for Sub Saharan Africa (version 1.2), Africa Soil Information Service (AfSIS) and ISRIC - World Soil Information, WageningenReport 2014/01, 160, 2014.

Leenaars, J. G. B., Claessens, L., Heuvelink, G. B. M., Hengl, T., Ruiperez González, M., van Bussel, L. G. J., Guilpart, N., Yang, H., and Cassman, K. G.: Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa, Geoderma, 324, 18-36, https://doi.org/10.1016/j.geoderma.2018.02.046, 2018.

Lutz, F., Stoorvogel, J. J., and Müller, C.: Options to model the effects of tillage on N2O emissions at the global scale, Ecological Modelling, 392, 212-225, http://www.sciencedirect.com/science/article/pii/S0304380018304034, 2019.

Magnusson, B., and Örnemark, U.: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics (2nd ed.), www.eurachem.org, 2014.

Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., van Bodegom, P. M., Cornwell, W. K., Ellsworth, D., Niinemets, Ü., Ordonez, A., Reich, P. B., and Santiago, L. S.: Global effects of soil and climate on leaf photosynthetic traits and rates, Global Ecology and Biogeography, 24, 706-715, http://dx.doi.org/10.1111/geb.12296, 2015.

Malhotra, A., Todd-Brown, K., Nave, L. E., Batjes, N. H., Holmquist, J. R., Hoyt, A. M., Iversen, C. M., Jackson, R. B., Lajtha, K., Lawrence, C., Vindušková, O., Wieder, W., Williams, M., Hugelius, G., and Harden, J.: The landscape of soil carbon data: emerging questions, synergies and databases, Progress in Physical Geography: Earth and Environment, http://dx.doi.org/10.1177/0309133319873309, 2019.

Moulatlet, G. M., Zuquim, G., Figueiredo, F. O. G., Lehtonen, S., Emilio, T., Ruokolainen, K., and Tuomisto, H.: Using digital soil maps to infer edaphic affinities of plant species in Amazonia: Problems and prospects, Ecol Evol, n/a-n/a, http://dx.doi.org/10.1002/ece3.3242, 2017.

Munzert, M., Kießling, G., Übelhör, W., Nätscher, L., and Neubert, K.-H.: Expanded measurement uncertainty of soil parameters derived from proficiency-testing data, Journal of Plant Nutrition and Soil Science, 170, 722-728, https://onlinelibrary.wiley.com/doi/abs/10.1002/jpln.200620701, 2007.

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

ISCN Database V3-1: https://dx.doi.org//10.17040/ISCN/1305039, 2017.

Soil Data IE (Interoperability Experiment): https://www.opengeospatial.org/projects/initiatives/soildataie, access: 14 June, 2019.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem, K. R.: Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity, BioScience, 51, 933-938, https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, 2001a.

Olson, R. J., Johnson, K. R., Zheng, D. L., and Scurlock, J. M. O.: Global and regional ecosystem modelling: databases of model drivers and validation measurements, Oak Ridge National Laboratory, Oak RidgeORNL/TM-2001/196, 95, 2001b.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., and Fernández-Ugalde, O.: LUCAS Soil, the largest expandable soil dataset for Europe: a review, European Journal of Soil Science, 69, 140-153, http://dx.doi.org/10.1111/ejss.12499, 2018.

Rayment, E. R., and Lyons, D. J.: Soil chemical methods - Australasia, CSIRO Publishing, 495 pp., 2011.

Ribeiro, E., Batjes, N. H., and Van Oostrum, A. J. M.: World Soil Information Service (WoSIS) - Towards the standardization and harmonization of world soil data. Procedures Manual 2018, ISRIC - World Soil Information, WageningenISRIC Report 2018/01, 166, 2018.

Rossel, R. A. V., and McBratney, A. B.: Soil chemical analytical accuracy and costs: implications from precision agriculture, Australian Journal of Experimental Agriculture, 38, 765-775, https://doi.org/10.1071/EA97158, 1998.

Accessing WoSIS latest from R (Expanded tutorial): https://www.isric.org/accessing-wosis-latest-r, access: 08 August 20-19, 2019.

Sanderman, J., Hengl, T., and Fiske, G. J.: Soil carbon debt of 12,000 years of human land use, Proceedings of the National Academy of Sciences, http://dx.doi.org/10.1073/pnas.1706103114 2017.

Sayre, R., Dangermond, J., Frye, C., Vaughan, R., Aniello, P., Breyer, S., Cribbs, D., Hopkins, D., Nauman, R., Derrenbacher, W., Burton, D., Grosse, A., True, D., Metzger, M., Hartmann, J., Moosdorf, N., Dürr, H., Paganini, M., DeFourny, P.,

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Arino, O., and Maynard, S.: A New Map of Global Ecological Land Units — An Ecophysiographic Stratification Approach, Association of American Geographers, Washington DC, 46, 2014.

Schoeneberger, P. J., Wysocki, D. A., E.C. Benham, and Soil Survey Staff: Field book for describing and sampling soils (ver. 3.0), National Soil Survey Center Natural Resources Conservation Service, U.S. Department of Agriculture, Lincoln (NE), 2012.

Soil Survey Staff: Soil Survey Laboratory Information Manual (Ver. 2.0), National Soil Survey Center, Soil Survey Laboratory, USDA-NRCS, Lincoln (NE)Soil Survey Investigation Report No. 45, 506, 2011.

Soil Survey Staff: Kellogg Soil Survey Laboratory Methods Manual (Version 5.0. R. Burt and Soil Survey Staff (ed.)) U.S. Department of Agriculture, Natural Resources Conservation Service, Lincoln (Nebraska), 1001 pp., 2014a.

Soil Survey Staff: Keys to Soil Taxonomy, 12th ed., USDA-Natural Resources Conservation Service, Washington, DC., 2014b.

Suvannang, N., Hartmann, C., Yakimenko, O., Solokha, M., Bertsch, F., and Moody, P.: Evaluation of the First Global Soil Laboratory Network (GLOSOLAN) online survey for assessing soil laboratory capacities, Global Soil Partnership (GSP) / Food and Agriculture Organization of the United Nations (FAO), RomeGLOSOLAN/18/Survey Report, 54, 2018.

Terhoeven-Urselmans, T., Shepherd, K. D., Chabrillat, S., and Ben-Dor, E.: Application of a global soil spectral library as tool for soil quality assessment in Sub-Saharan Africa, A EUFAR Workshop on Quantitative Applications of Soil Spectroscopy (5-16 April 2010), 2010, 15-15,

Tóth, G., Jones, A., and Montanarella, L.: LUCAS Topsoil survey: methodology, data and results Land Resource Management Unit - Soil Action, European Commission Joint Research Centre Institute for Environment and Sustainability, 141, 2013.

USDA-NCSS: National Cooperative Soil Survey (NCSS) Soil Characterization Database, United States Department of Agriculture, Natural Resources Conservation Service, Lincoln, 2018.

van Engelen, V. W. P., and Dijkshoorn, J. A.: Global and National Soils and Terrain Digital Databases (SOTER) - Procedures manual (Ver. 2.0), IUSS, ISRIC and FAO, WageningenISRIC Report 2013/04, 198, 2013.

van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z.: Yield gap analysis with local to global relevance—A review, Field Crops Research, 143, 4-17, http://www.sciencedirect.com/science/article/pii/S037842901200295X, 2013.

Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y., Padarian, J., Schaap, M., Tóth, B., Verhoef, A., Vanderborght, J., van der Ploeg, M., Weihermüller, L., Zacharias, S., Zhang, Y., and Vereecken, H. C. R. G.: Pedotransfer functions in Earth system science: challenges and perspectives, Reviews of Geophysics, n/a-n/a, http://dx.doi.org/10.1002/2017RG000581, 2017.

van Reeuwijk, L. P.: On the way to improve international soil classification and correlation: the variability of soil analytical data, ISRIC, WageningenAnnual Report 1983, 7-13, 1983.

van Reeuwijk, L. P.: Guidelines for quality management in soil and plant laboratories, FAO, Rome, 143, 1998.

Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B. G., Bartholomeus, H. M., Bayer, A. D., Bernoux, M., Böttcher, K., Brodský, L., Du, C. W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C. B., Knadel, M., Morrás, H. J. M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E. M. R., Sanborn, P., Sellitto, V. M., Sudduth, K. A., Rawlins, B. G., Walter, C., Winowiecki, L. A., Hong, S. Y., and Ji, W.: A global spectral library to characterize the world's soil, Earth-Science Reviews, 155, 198-230, http://dx.doi.org/10.1016/j.earscirev.2016.01.012, 2016.

WEPAL: ISE Reference Material - A list with all available ISE reference material samples, WEPAL (Wageningen Evaluating Programmes for Analytical Laboratories), Wageningen, 110+, 2019.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao,

https://doi.org/10.5194/essd-2019-164
Preprint. Discussion started: 16 September 2019

Open Access Earth System Science Data Discussions

J., and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data, 3, 160018, http://dx.doi.org/10.1038/sdata.2016.18, 2016.