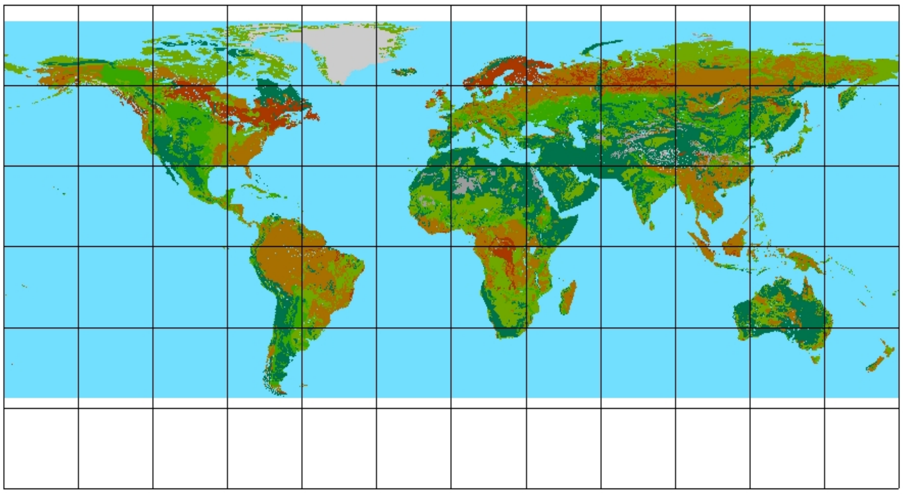
Report 2006/02

ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (version 1.1)

Niels H Batjes

(November 2006)



ISRIC World Soil Information

Disclaimer:

Correct citation:

Batjes NH 2006. ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (version 1.1). Report 2006/02 (available through: http://www.isric.org), ISRIC – World Soil Information, Wageningen (with data set)

Inquiries:

c/o Director, ISRIC – World Soil Information
PO Box 353
6700 AJ Wageningen
The Netherlands
Telefax: +31-(0)317-471700
E-mail: soil.isric@wur.nl
Web: www.isric.org

Front cover: WISE-derived soil pH classes for the World (0-20 cm)

Contents

List of Tables

List of Figures

SUMMARY

This report describes a harmonized dataset of derived soil properties for the world. The data were created using the soil distribution shown on the 1:5 million scale FAO-Unesco Soil Map of the World (DSMW) and soil parameter estimates derived from ISRIC's global WISE soil profile database.

The dataset considers 19 soil variables that are commonly required for agro-ecological zoning, land evaluation, crop growth simulation, modelling of soil gaseous emissions, and analyses of global environmental change. It includes "best" estimates for: soil drainage class, organic carbon content, total nitrogen, C/N ratio, pH(H_2O), CEC_{soil}, CEC_{clay}, effective CEC, base saturation, aluminium saturation, calcium carbonate content, gypsum content, exchangeable sodium percentage (ESP), electrical conductivity, particle size distribution (i.e. content of sand, silt and clay), content of coarse fragments (> 2 mm), bulk density, and available water capacity (-10 to -1500 kPa). Parameter estimates — median values — for each of these variables are presented by FAO soil unit for fixed depth intervals of 20 cm up to 100 cm depth (or less when appropriate). These were derived from analyses of some 10,100 profiles held in WISE, using a scheme of taxonomy-based taxotransfer rules complemented with expert-rules. The type of rules used to fill gaps in the measured soil data have been flagged in the database.

Most map units on the DSMW are complex, comprising up to eight different soil units. Assessments and model applications that use the derived soil properties therefore should consider the full map unit composition.

Three methods are discussed for presenting results spatially using GIS. Examples of un-binned and classified data — e.g. spatially dominant pH class at 0-20 cm depth in grid cell — are included in the GIS data set for general usage.

The soil parameter values and derived maps presented here are should be seen as best estimates based on the current selection of soil profiles in WISE, the procedure for clustering the measured data, taxotransfer

scheme used for filling gaps in the measured data, and the spatial data of the digital Soil Map of the World. This information is considered appropriate for exploratory assessments at scale < 1:5 million, pending the global update of the information on world soil resources in the World Soil and Terrain Database (SOTER) programme.

Keywords: soils, derived data, environmental modelling, ISRIC-WISE database, FAO-Unesco Soil Map of the World

1 INTRODUCTION

This report describes the sources and procedures used to develop a global data set of derived soil properties, with a spatial resolution of 5 by 5 arc-minutes. The spatial data were derived from the digital Soil Map of the World (DSWM, see FAO 1995b), at scale 1:5 million. Soil analytical data to characterize the main soil units of the DSMW were taken from the ISRIC-WISE database (from now on referred to as WISE). Version 3.0 of WISE holds some 10,100 globally distributed soil profiles, collated from disparate sources.

The DSMW was derived from the printed version of the Soil Map of the World (FAO-Unesco 1971-1981), with minor corrections and updates (FAO 1995a). Compilation of the Soil Map of the World was a huge task involving collection and correlation of soil information from all regions of the world. Most of this information was collated prior to the 1970s; the reliability thereof is known to vary considerably between different areas. Part of the information is outdated. Further, international soil classification systems have since then evolved through the Revised Legend (FAO 1988) to the World Reference Base for Soil Resources (IUSS Working Group WRB 2006). Hence the ongoing update of the information on world soil resources in SOTER, the World Soil and Terrain Database programme (Nachtergaele and Oldeman 2002; Oldeman and van Engelen 1993). Ultimately, a global SOTER product at scale 1:5 million is scheduled to supersede the digital Soil Map of the World.

As a first step in this direction, ISRIC staff compiled a harmonized global soil resources data base, with a resolution of 5 by 5 arc-minutes, that combines secondary data derived from all continental SOTER databases completed so far (Batjes 2004, 2005a, b, c, d; Van Engelen *et al.* 2005). Soil parameter estimates were derived using a system of taxonomy-based pedotransfer rules, complemented with expert rules (Batjes 2003a). The secondary data discussed in this report have the same format and were derived using similar methods as used for developing secondary SOTER data sets. As such, they may be used as best estimates for those sections of the world that are not yet represented in SOTER.

The present data complement earlier WISE-derived datasets that were based on a smaller selection of globally distributed soil profiles, different procedures for aggregating the soil analytical data and filling gaps in the measured data, and aggregated at a resolution of 0.5 by 0.5 degrees (Batjes 2002b, 2003b; Batjes 2005e; Batjes and Bridges 1994; Batjes *et al.* 1995).

The materials and methods are described in Section 2. Results are presented and discussed in Section 3, while conclusions are drawn in Section 4. The structure of the various output tables, GIS-legend files, and installation procedure are presented in the Appendices.

2 MATERIALS AND METHODS

2.1 Primary soil data

2.1.1 Spatial data

The spatial distribution of soil units, per 5 by 5 arc-minute grid cell, was taken from the 1:5 million scale, digital Soil Map of the World. The base map of the DSMW is based on ETOPO5[1] (Earth Topography — 5 arc minute). ETOPO5 was assembled from several uniformly gridded databases into a worldwide data set with a cell size of 5 minutes of latitude by 5 minutes of longitude. Cells are written row-wise from the West to the East, starting at the Northwest corner. At the equator, the cell size corresponds with about 0.08333 decimal degrees (see FAO 1995b).

The ERDAS files (WORLD.GIS and WORLD.TRL), prepared by FAO (1995a), were first converted to ArcGis® GRID format. The spatial data are bound by longitudes -180°W and +180°E and latitudes +84°N and -56.50°S. This corresponds to 4320 columns and 1686 rows or 7,283,520 grid cells in total. The Arctic, Antarctica and some islands are not included on the DSMW (see FAO 1995b).

The DSMW legend is comprised of 4931 different map units. These consist of soil units or associations thereof. Complex map units are comprised of one dominant soil unit and up to seven component soils. According to FAO's composition rules, the latter include associated soils (>20% of the map unit) and inclusions (<20% of the map unit).

Prior to starting with the analyses, the expansion files where checked for possible inconsistencies, also *vis a vis* the spatial, raster data. Several map units, while described in the expansion file, did not occur on the raster GIS map (SNUM 699, 1213, 1874, 1875, 1876, 1877, 1881, 1890, 1898, 1912, 1914, 1915, 1918, 1931, 3587, 6208, 6209, 6232, 6243, 6269, 6272, 6283, and 6324). In 35 cases, the dominant soil unit as given in the DSMW expansion files proved to be incorrect. For example, in map unit number (SNUM) 1823, coded "Yh10-a", the fifth unit ($soil_5$; DS= dune sands) was estimated to cover 50% of the map unit and the so-called dominant soil ($soil_1$, Yh) only 10%. In all instances, this related to miscellaneous soil units with more than 50% coverage and these were

[1] http://www.ngdc.noaa.gov/mgg/global/etopo5.HTML

always listed as the last component soil for the given map unit. These minor inconsistencies have been corrected. The original map unit codes (SNUM), however, were maintained to preserve consistency with the original codes used on the DSMW.

In seven instances (SNUM 3075, 3076, 3663, 4205, 5018, 5101, and 5211), the same soil unit was listed twice in FAO's expansion files. In such cases, the area of identical soil units within the map unit was summed and the expansion files were updated accordingly.

Some map units have been described as comprising say 50% of soil unit 1 and 50% of soil unit 2. In such cases, it will be difficult to select the so-called dominant soil unit when spatially aggregating the derived data (see method-B in Section 3.4).

Statistics for the proportion of the dominant and component soils — based on the updated expansion files — are given in Table 1. The median relative area of the dominant soil unit ($soil_1$) within a map unit is 60%, with lower and upper quartiles of 50 and 70% respectively. The median, relative area for $soil_1$, $soil_2$, and $soil_3$ combined is 100% with lower quartile of 90% and a minimum of 60%.

Table 1. Relative area of dominant and component soils within map units of the Soil Map of the World

Descriptive statistics	Relative area[a] of dominant and component soils (%)							
	Soil1[b]	Soil2	Soil3	Soil4	Soil5	Soil6	Soil7	Soil8
Minimum	24	0	0	0	0	0	0	0
1st Quartile	50	20	0	0	0	0	0	0
Median	60	20	10	0	0	0	0	0
3rd Quartile	70	30	20	10	0	0	0	0
Maximum	100	50	34	25	10	10	5	4

[a] The actual area within a 5 by 5 arc-minutes grid will vary with latitude — the grid cell size is some 9 by 9 km at the equator and will decrease gradually to the poles according to a cosine function of latitude. [b] $Soil_1$ is the dominant soil and $Soil_2$ to $Soil_8$ are the component or associated soils.

Consideration of the dominant soil unit only will often ignore the inherent complexity of many map units (Figure 1). Therefore, the full map unit composition should be considered when using the derived data in model applications!

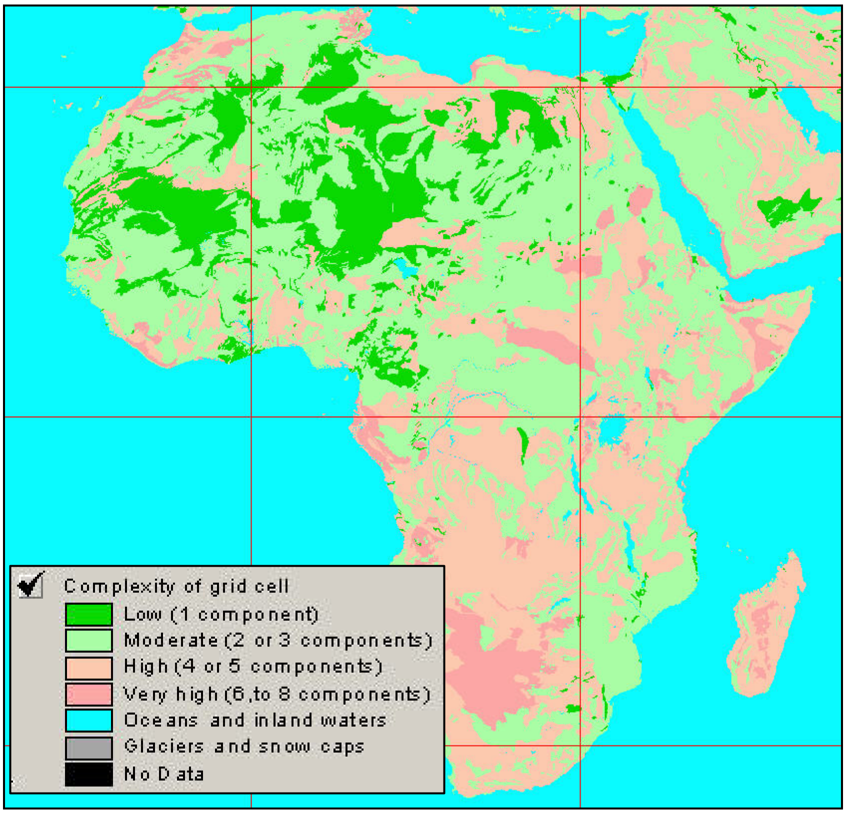


Figure 1. Complexity of grid cells for Africa according to the digitalSoil Map of the World

2.1.2 Profile data

Unlike for SOTER, the component soils of individual DSMW units have not been characterised by typical profiles. Therefore, all parameter estimates had to be derived on a soil unit basis, irrespective of the soil unit's location in the world. This was done using a system of taxonomy-based pedotransfer and expert-rules (Batjes 2003a). The classification code was used to define a *virtual* profile for each of the 106 soil units of the FAO Legend (e.g. WD-Fr for Rhodic Ferralsols).

Taxotransfer rules were developed using some 10,100 profiles held in WISE. In absence of more detailed textural data, all virtual profiles were assumed to have the modal textural class of the corresponding soil unit.

Most profiles in WISE are from Africa, Latin America and the Caribbean and Europe (Table 2) — the corresponding soil descriptions and analytical data were largely derived from field surveys carried out between 1960 and 2000.

Table 2. Global distribution of soil profiles held in WISE

Region	Number
Africa	4043
Australia and Pacific Islands	152
China, India, Indonesia & Philippines	931
Europe (West of Ural Mountains)	1348
North America	450
South America and Caribbean	2195
SW and Northern Asia (Siberia)	970
World	10089

Several soil profiles were re-classified subsequent to new checks of the preceding version. In particular, it appeared that several poorly drained profiles with argillic horizons, from Western Africa, were erroneously classified as Gleysols in the source documents. Consistency of soil classification with respect to base saturation (i.e. dystric *versus* eutric) and clay activity were also checked using simple, automated queries. In some cases, it has been difficult to assess the soil classification in detail based on the available data (Table 3).

Most profiles for WISE have been collated from soil survey reports or imported from external databases such as those of USDA-NRCS or FAO (Batjes 1995b). WISE version 3.0 includes many new profiles from Asia, in particular Thailand, as well as updates or additions from other parts of the world. The later include updated soil chemical data for large sections of Latin America and the Caribbean (Dijkshoorn *et al.* 2005) and additions from ISIS (2006), version 6.0.

Detailed information about the quality of the various source data was often lacking. The inferred quality of the profile descriptions ranged from good, for pedons described and analyzed according to international standards (Soil Survey Staff 1996; van Reeuwijk 2002), to poor for so-called "incomplete pedons". The later typically have numerous gaps in the soil analytical data (Table 3). Incomplete characterization data were available for many soil profiles because only selected measurements were undertaken during the original surveys. Further, every country has its own methods and these methods may vary for one laboratory to the next within one country. Consequently, issues of quality and comparability of analytical data, collated from disparate sources, are critical in any analysis of soil profiles. Yet, there are no straightforward solutions (Batjes 1999, 2002a; Dobos *et al.* 2006; Pleijsier 1989; van Reeuwijk 1998).

Consequently, these issues have been addressed pragmatically in this study similar to what has been the case with the Soil Map of the World (e.g. FAO-Unesco 1981, p. 91; FAO 1995a). Correlation of soil analytical data, however, should be done more rigorously when more detailed scientific work is considered.

Table 3. Description status of soil profiles held in WISE

Description status	Number	Description
Reference profile	1747	ISRIC-ISIS, USDA-NRCS or other Reference Pedons
Routine description	4958	Routine profile description in which no essential data are lacking from the description, sampling, or analysis. The data give a good indication of the nature of the soil in the FAO-Unesco (1974) Legend
Partial description	3301	Incomplete description in which certain relevant elements are missing from the description, an insufficient number of samples collected, or the reliability of the analytical data do not permit a complete characterization of the soil. The description/data, however, is useful for specific purposes and provides a satisfactory indication of the nature of the soil in the FAO-Unesco (1974) Legend
Incomplete description[a]	83	Other descriptions in which essential elements are lacking from the description preventing a satisfactory soil characterization and classification

[a] Generally not accepted for inclusion in WISE unless the corresponding soil unit is grossly under-represented (Batjes 1995a); may be deleted when better profiles become available.

The relative number of soil profiles available for each major soil group of the FAO Legend is shown in Figure 2. Luvisols, for example, account for some 15% of the total number of profiles in WISE and some 7% of the total extent of soils on the DSMW. Conversely, shallow Lithosols account for some 11% of the total extent, yet less than 1% of the profiles under consideration. The distribution of profiles in Figure 2 reflects that routine soil surveys largely focussed on describing agriculturally promising soils.

As detailed in Appendix 1, there are still few (<10) profiles for various soil units, in particular Takyric Yermosols, Gleyic Greyzems, Glossic Chernozems, Gelic Histosols, Takyric Solonchaks, Mollic Solonchaks, Gelic Regosols and Dystric Podzoluvisols. Typically, however, in order to present robust parameter estimates (medians) for a given soil variable, at least 30 profiles — with complete and comparable sets of soil analytical data — would be needed for each soil unit. This ideal situation, however, will seldom occur as reflected by ongoing discussions on the comparability of soil analytical data in Europe (Dobos *et al.* 2006).

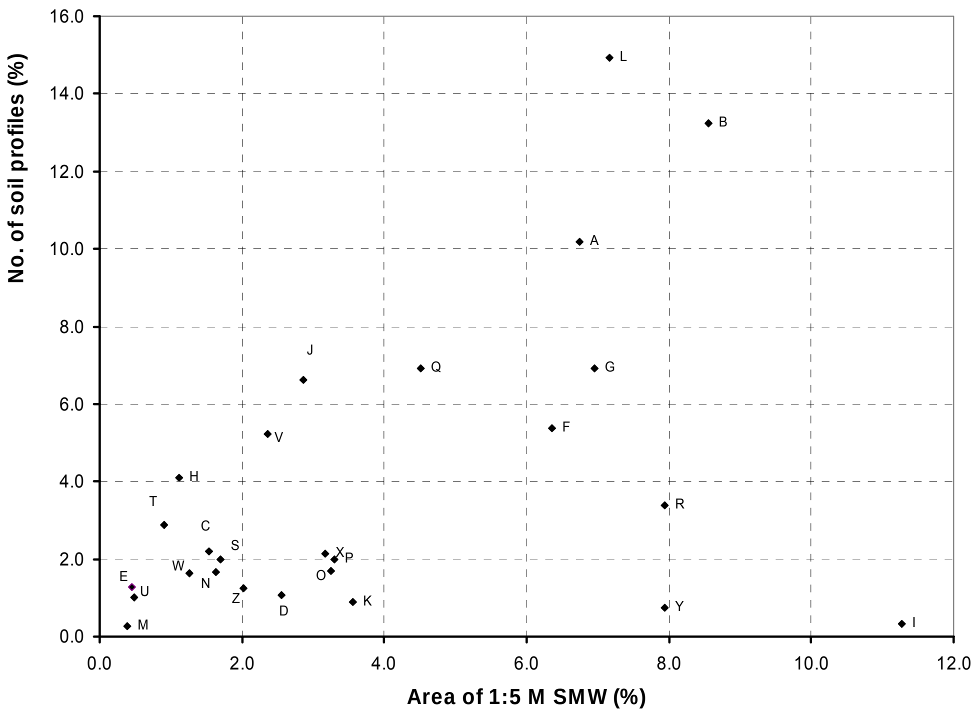


Figure 2. Representation of major soil groups in the profile data set relative to their extents on the 1:5M Soil Map of the World[2]

2.2 Derived soil data

This study considers nineteen soil attributes (Table 4, commonly required in studies of environmental change, for which primary or measured data are collated in WISE. Soil parameter estimates (medians) for each of these — by soil unit, depth range of 20 cm up to 100 cm depth, and five textural classes (CEC 1985; FAO 1988) — were derived from statistical analyses of the measured data. The procedure, which includes an outlier-rejection scheme to screen the available data, is similar to the one used to generate secondary SOTER databases (Batjes 2003a). However, it was

[2] Relative area is expressed as percentage of total area of all major soil groups considered, that is excluding miscellaneous units. Soil codes are accroding to the FAO-Unesco (1974) Legend.

adapted to run with soils classified according to the old Legend (FAO-Unesco 1974).

Table 4. List of soil variables for which parameter estimates are presented

Organic carbon
Total nitrogen
C/N ratio ‡
Soil reaction (pH_{H2O})
Cation exchange capacity (CEC_{soil})
Cation exchange capacity of clay size fraction (CEC_{clay}) • ‡
Base saturation (as % of CEC_{soil}) ‡
Effective cation exchange capacity (ECEC) † ‡
Aluminium saturation (as % of ECEC) ‡
$CaCO_3$ content
Gypsum content
Exchangeable sodium percentage (ESP) ‡
Electrical conductivity (ECe)
Bulk density
Coarse fragments (volume %)
Sand (mass %)
Silt (mass %)
Clay (mass %)
Available water capacity (AWC; from -33 to -1500 kPa; cm m^{-1}) ‡ □

‡ Calculated from other measured soil properties.
† ECEC is defined as exchangeable (Ca^{++}+Mg^{++}+K^{+}+Na^{+}) + exchangeable (H^{+}+Al^{+++}) (van Reeuwijk 2002).
• CEC_{clay} was calculated from CEC_{soil} by assuming a mean contribution of 350 $cmol_c$ kg^{-1} OC, the common range being from 150 to over 750 $cmol_c$ kg^{-1} (Klamt and Sombroek 1988).
□ The soil water potential limits for AWC conform to USDA standards (Soil Survey Staff 1983). Values shown have not been corrected for the presence of fragments < 2 mm.

There are six miscellaneous units on the Soil Map of the World for which there are no measured data. The following assumptions were used:

- Dune sands (DS): soil parameter estimates for Cambic Arenosols (Qb) were used as default, except for: organic carbon content that was set to 2 g C kg^{-1} for the topsoil and 1 g C kg^{-1} for the subsoil; the content of sand was set at 98%; soil drainage at excessively drained.

- Rock outcrops (RK): Being non-soil, all soil parameter estimates were set at -7. Similarly, for shallow soils, including Lithosols, Rendzinas and Rankers, all parameter estimates for the rocky subsoil were set at -7.

- Not Determined (NA): as above, but using soil parameter estimates for Cambisols (B) as the default.

- Salt Flats (ST): as above, but using soil parameter estimates for Orthic Solonchaks (Z) as the default.

- Oceans and Inland Waters (WR): all parameter estimates were set at -1 to permit visualisation using GIS (SNUM= 6997 and 1972 for Africa).

- Glaciers and snow caps (GL): all parameter estimates were set at -2 (SNUM= 6998).

3 RESULTS AND DISCUSSION

3.1 General

The compilation of the DSMW, at scale 1:5 million, encompassed a marked degree of data integration, the aim being to simplify the geographical distribution of soil types to a regionally representative pattern. This aggregation was based on soil survey data collated prior to the 1970's, using the original Legend (FAO-Unesco 1971-1981). Inherently, all mapping units will include a number of impurities, often in excess of 15% (see Landon 1991), which cannot be mapped at the given scale.

3.2 Soil unit composition

Each grid cell may contain up to eight component soils. The relative extent thereof — estimated according to FAO's expert-based composition rules — has been clustered in five classes to arrive at a compact map unit code: 1 – from 80 to 100 per cent; 2 – from 60 to 80 per cent; 3 – from 40 to 60 percent; 4 – from 20 to 40 per cent, and 5 – less than 20 percent (Appendix 2). The original map unit code, FAOSOIL, has also been maintained in the data set for ease of reference.

3.3 Soil parameter estimates

3.3.1 Calculation procedure

All derived soil data have been stored in an MS Access® table (Appendix 3). The type of taxotransfer rules (TTRs) used were flagged by profile and depth layer (Appendix 4). Data listed under *TTRsub* indicate that the data substitution for a given attribute was based on WISE-derived parameter estimates for similar soil units. Otherwise, if the corresponding population in WISE was too small ($n_{WISE} < 5$) for the substitution to be considered meaningful, the rules were flagged under *TTRmain*. Methodological details may be found in Batjes (2003a).

Each flag consists of a sequence of letters followed by a numeral (see under *TTRsub* and *TTRmain*). The letters indicate soil attributes for which a TTR has been applied (Figure 3). The number code reflects the size of the sample population, after outlier-rejection, used for the statistical analyses (Table 5).

Table 5. Criteria for defining confidence in the derived data

Code	Confidence level	n_{WISE}[a]
1	Very high	> 30
2	High	15-29
3	Moderate[b]	5-14
4	Low	1-4
-	No data	0

[a] n_{WISE} is the sample size after outlier rejection (see Batjes 2002a).
[b] The cut-off point for applying any TTR is n_{WISE} < 5

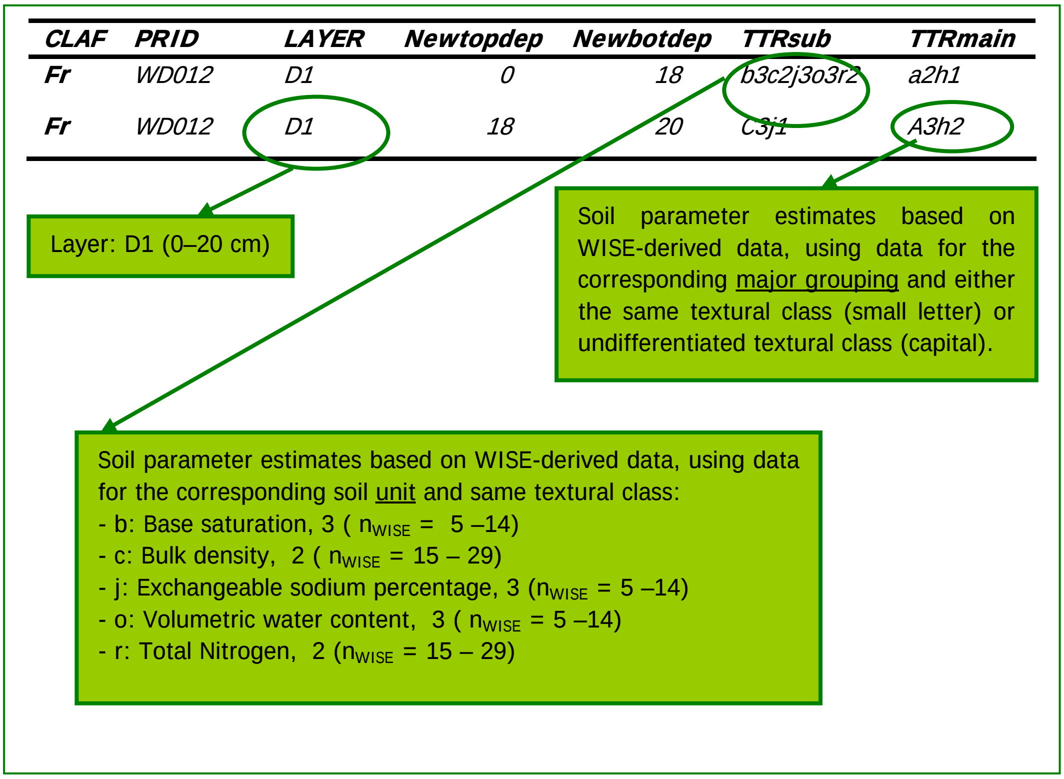


Figure 3. Flagging of taxotransfer rules by profile, depth zone and attribute

When a small letter is used, the substitution considered median data for the corresponding textural class (for example, Fine). Otherwise, when a capital is used, this indicates that the substitution was based on the whole set for the corresponding soil unit and depth layer, irrespective of soil texture (i.e. undifferentiated or “u”). The same coding conventions apply for *TTRmain*. The overall scheme is shown in Figure 3 for a virtual Rhodic Ferralsol (Fr), coded WD012.

3.3.2 Sources of uncertainty

Inherently, the calculation of representative soil parameter estimates for the 106 FAO soil units of the Soil Map of the World remains fraught with uncertainty. Possible effects of regional variation in climate, relief, parent material, land use and management practices on specific soil parameter estimates cannot be considered explicitly with the DSMW. By implication, the same parameter estimates for soil organic carbon content for Lithosols thus had to be used irrespective of whether these soil units are located, for example, in cool and humid parts of Western Europe or in hot and arid regions in the Middle East. The variation observed within each soil unit, however, is known (see Appendix 3). Conversely, this type of regional differences can be accounted for in SOTER.

Each component soil of a SOTER unit is characterized by a soil profile, identified as being regionally representative by regional soil experts (Batjes 2005a; Dijkshoorn, Huting and Tempel 2005; van Engelen and Wen 1995). Moreover, in the case of SOTER databases, the taxotransfer rules can be applied with respect to the textural class of the selected profiles (and 20 cm soil layer), as opposed to the present forced use of the undifferentiated or modal textural class with the DSMW.

It should be stressed that the natural range in measured chemical and physical properties can be considerable at soil unit level, with coefficients of variation often exceeding 50% (Batjes 2002b; Beckett and Webster 1971; Landon 1991; Spain *et al.* 1983). Similarly, the median of absolute deviations (MADs) — the median of the differences between each observation and the median —, are often large pointing at a large spread of the data (see Appendix 3).

The overall assumption in this assessment has been that the confidence in a *TTR*-based parameter estimate should increase with the size of the sample populations present in WISE, after outlier-rejection. In addition, in principle, the confidence in soil parameter estimates listed under *TTRsub* should be higher than for those listed under *TTRmain*. Nonetheless, a high confidence rating does not necessarily imply that the soil parameter estimates shown will be representative for the soil unit under consideration. Profile selection for WISE, as for any other large database, is not probabilistic but based on available data and expert knowledge.

Results may be biased for those soil properties that were recorded as “not observed” or “nil” in the original surveys, for example volumetric gravel content. In such cases, the medians computed here may well give a biased impression of “modal” conditions for some soil units as, probably, only the limiting gravel contents may have been documented in field surveys.

Several soil chemical properties — such as a high aluminium saturation in parts of the subsoil or a high exchangeable sodium percentage in parts of the topsoil — may be “levelled out”” during depth-weighting. When occurring, this will also be reflected in derived values obtained through pedotransfer.

Several of the soil attributes under consideration, such as the presence of fragments > 2 mm and water holding capacity, are not diagnostic in the FAO Legend. Many soil processes and properties are readily modified by changes in land use and management. For example, soil pH and aluminium saturation upon liming; salinity and electrical conductivity upon irrigation or soil drainage; and soil organic matter quantity and quality upon changes in tillage practices, deforestation, or climate change.

The type of rules used for each virtual profile have been documented (see Appendix 3). Since the units of the DSMW are not linked directly to any geo-referenced profiles, all parameter estimates shown here had to be derived using a combination of taxotransfer and expert-rules. Conversely, in the case of SOTER, the taxotransfer scheme is mainly used to fill gaps in measured (primary) data for profiles that have been identified as being regionally representative for the corresponding map unit (e.g., Batjes and Gicheru 2004).

3.4 Linkage to GIS

Comparison of mapping approaches

Soil parameter estimates derived from the *TTR*-scheme can be joined to the rasterized Soil Map of the World using GIS. This linkage, or GIS-based 'join', will be through the unique map unit identifier — coded as VALUE or SUID in the attribute table of raster file *smw5by5min* — which corresponds to FAO's soil map unit number (SNUM; Appendix 4). Similarly, the derived soil data can be joined to individual, continental *vector* sets provided on the CD of the digital Soil Map of the World.

DSMW units are generally comprised of at least one component soil unit, each having its own set of soil parameter estimates for the 5 depth layers under consideration (Figure 4). Clearly, this wealth of information complicates linkage to GIS (Appendix 4). A generalized code for the map unit composition, the FAO Revised Legend code for the spatially dominant soil unit (*soil1*), and the relative proportion of the dominant soil unit (*prop1*) have been 'embedded' in the raster file's attribute table to give an indication of map unit complexity (see Appendix 2 for details).

Elaborate studies will often involve overlays with auxiliary data layers, for example, climate and land cover. Model runs for the resulting spatial units will often have to be performed outside of the GIS. The type of research purpose will determine which soil parameter estimates, and depth layers, will be required for a specific application. Generally, the necessary data selections can best be made with tailor-made programs designed to meet the scope of these applications. The GEFSOC Modelling System©, for example, requires soil data (0-20 cm) and data on climate and land-use to estimate soil organic carbon stocks and changes (Easter *et al.* 2005; Eleanor *et al.* 2006).

SOTERsummaryFile : Table

ISOC	SUID	Layer	TCID	SCID	PROP	CLAF	PRID	Drair	TopDi	BotDe	CFRAG	SDTO	STPC	CLPC	PSCL	BULK	TAWC	CECS	BSAT	CECc	PHAQ	TCEQ	G'
WD	4522	D1	1	1	50	Je	WD-Je	M	0	20	4	43	33	24	M	1.32	12	15	90	50	6.8	1	
WD	4522	D1	1	2	20	Jt	WD-Jt	P	0	20	4	21	35	44	F	1.16	12	20	45	24	4.6	0	
WD	4522	D1	1	3	20	Zg	WD-Zg	P	0	20	4	36	34	30	M	1.37	12	19	100	42	8.4	79	
WD	4522	D1	1	4	10	Re	WD-Re	S	0	20	19	65	21	14	M	1.25	10	7	90	50	6.5	1	
WD	4522	D2	1	1	50	Je	WD-Je	M	20	40	10	43	32	25	M	1.38	13	14	93	56	7.1	23	
WD	4522	D2	1	2	20	Jt	WD-Jt	P	20	40	6	21	33	46	F	1.18	15	21	30	25	4.0	0	
WD	4522	D2	1	3	20	Zg	WD-Zg	P	20	40	11	37	31	32	M	1.43	15	18	100	45	8.4	112	
WD	4522	D2	1	4	10	Re	WD-Re	S	20	40	27	65	20	15	M	1.24	10	6	83	44	6.6	1	
WD	4522	D3	1	1	50	Je	WD-Je	M	40	60	10	43	32	25	M	1.40	10	14	98	52	7.1	23	
WD	4522	D3	1	2	20	Jt	WD-Jt	P	40	60	8	21	34	45	F	1.08	10	22	30	25	3.7	0	
WD	4522	D3	1	3	20	Zg	WD-Zg	P	40	60	11	38	31	31	M	1.62	14	19	100	45	8.4	114	
WD	4522	D3	1	4	10	Re	WD-Re	S	40	60	28	67	19	14	M	1.28	9	6	88	48	6.9	1	
WD	4522	D4	1	1	50	Je	WD-Je	M	60	80	10	45	31	24	M	1.40	11	12	95	51	7.3	24	
WD	4522	D4	1	2	20	Jt	WD-Jt	P	60	80	10	21	33	46	F	1.13	10	23	30	27	3.7	0	
WD	4522	D4	1	3	20	Zg	WD-Zg	P	60	80	5	36	32	32	M	1.62	15	21	100	45	8.3	127	
WD	4522	D4	1	4	10	Re	WD-Re	S	60	80	28	69	19	12	M	1.46	10	6	84	60	6.9	1	
WD	4522	D5	1	1	50	Je	WD-Je	M	80	100	10	45	31	24	M	1.42	11	11	95	50	7.3	23	
WD	4522	D5	1	2	20	Jt	WD-Jt	P	80	100	10	23	32	45	F	1.10	11	22	30	29	3.5	0	
WD	4522	D5	1	3	20	Zg	WD-Zg	P	80	100	4	35	30	35	F	1.62	16	21	100	45	8.4	86	
WD	4522	D5	1	4	10	Re	WD-Re	S	80	100	28	67	20	13	M	1.50	8	7	83	62	7.0	5	

Record: 20 of 20 (Filtered)

Figure 4. Excerpt from summary file for map unit WD4522

Alternatively, there are various "simple" options to spatially aggregate the derived soil data, each of these having their strengths and limitations (Batjes 1996a, 2002c; Carter and Scholes 2002; FAO 1995a; Kern 1995). Three mapping procedures were compared using the content of soil organic carbon (SOC) at 0-20 cm depth (D1), as an example:

Method A:
Consideration of the dominant soil unit only, and subsequent classification of the un-binned data into seven pre-selected SOC classes: 0-5, 50-10, 10-30, 30-100, 100-200, 200-300 and 300-450 g C kg^{-1}.

Method B:
A *priori* classification of the SOC content for each component soil unit, taking into account the full map unit composition, into of one of the seven pre-selected classes. Subsequently, the spatially dominant class was assigned to the grid cell.

Method C:
Area-weighting of the SOC content by grid cell, considering the full map unit composition. Subsequently, the un-binned, area-weighted values were allocated to one of the seven pre-defined SOC classes.

As shown by Figures 5 and 6, the classes mapped can vary widely for methods A, B and C. A *perceived* advantage of Method-C over Method-B is that it would allow modellers to use un-binned values for each grid cell, facilitating model runs. However, area-weighting of soil parameter estimates for a grid cell comprising say 70% of mineral soil units with 20 g C kg^{-1} and 30% of organic soil units with say 400 g C kg^{-1}, between 0-20 cm depth, would lead to a calculated, area-weighted "un-binned" value of 134 g C kg^{-1}. In the case of a logarithmic scale, like for soil pH, area-weighting will be even more misleading!

In view of the above, the attached data set only includes tables derived using method B (binned, for dominant class in grid) and method A (un-binned, for dominant soil unit in grid). Method B, however, is considered most appropriate for visualizing regional differences in derived soil properties at scale 1:5 million.

Method A — Un-binned data sets

Files generated with method-A present parameter estimates in un-binned format for the dominant soil in each grid cell. The relative area thereof, however, can vary from 24% to 100%, as shown in Table 1. Technical details are given in Appendix 6.

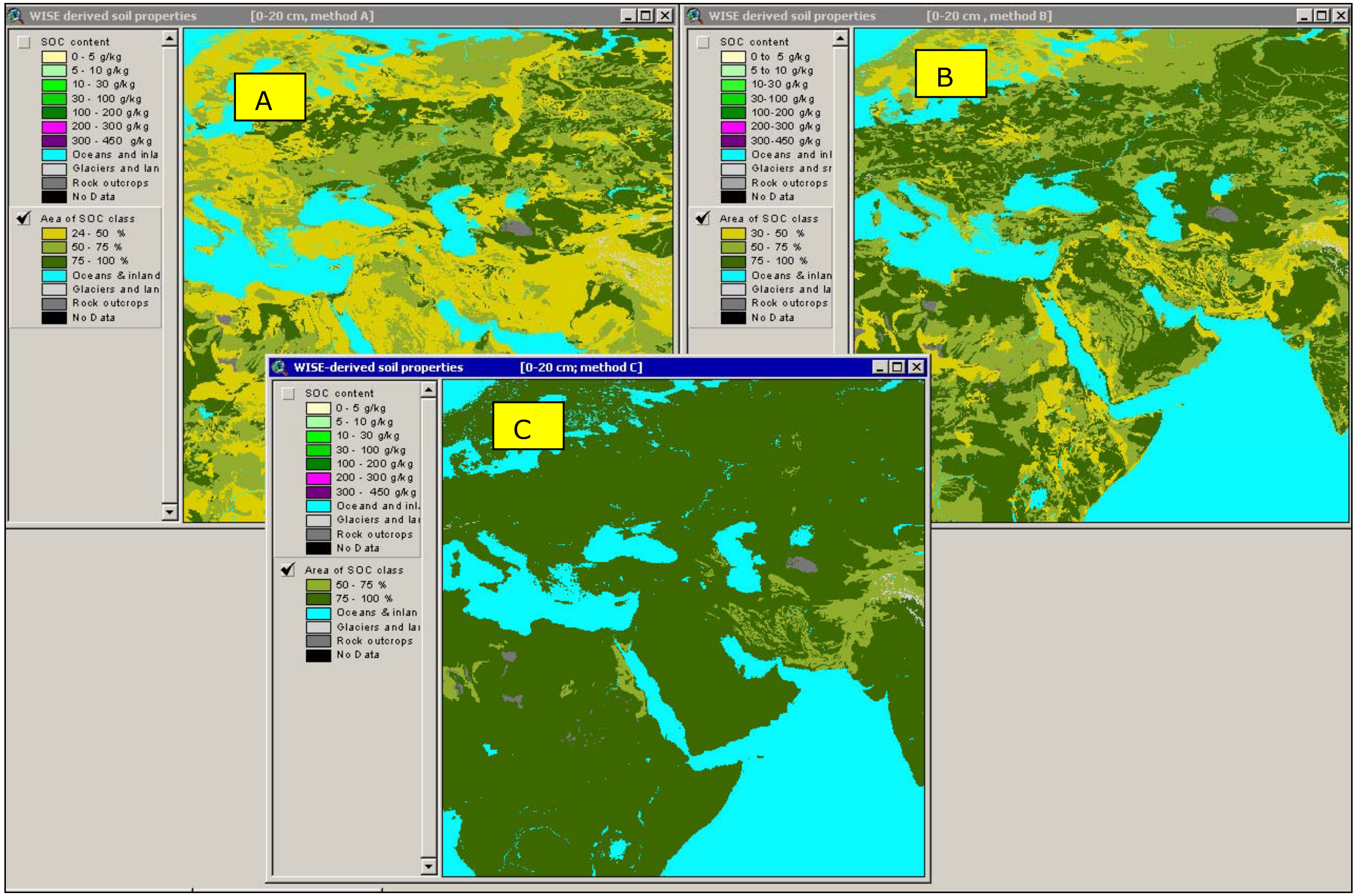


Figure 5. Relative area of grid cell considered when determining SOC classes with methods A, B, and C

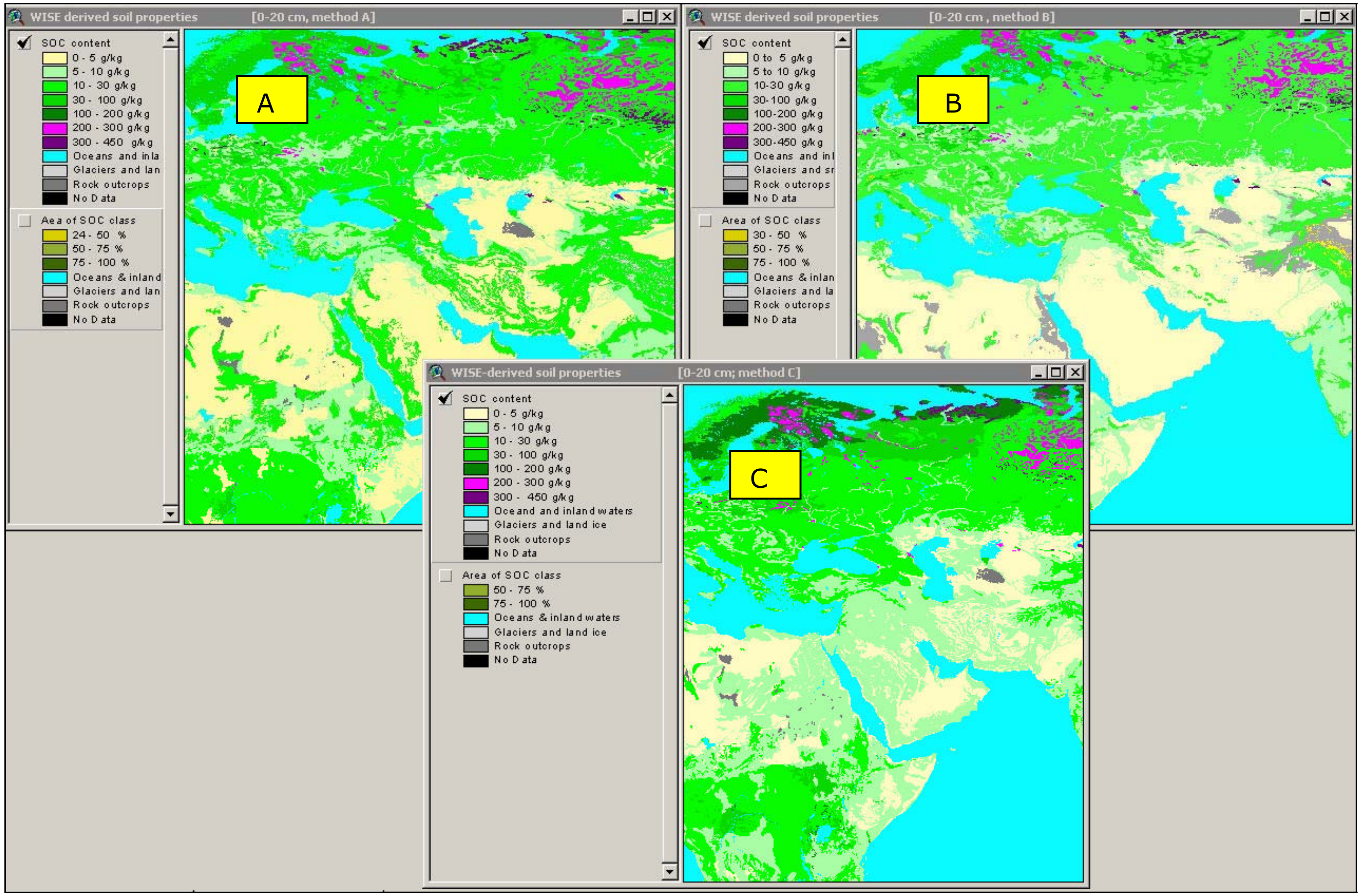


Figure 6. Differences in SOC classes mapped according to method A, B, and C

Method B — Binned data sets

Selection of suitable class limits for the binned or classified data will depend on the uses to which the data are to be put (see Anon. 1984; Chadwick and Kuylenstierna 1990; FAO 1983; Landon 1991; Sys *et al.* 1993). Such class limits are inherently fuzzy (Burrough 1989; McBratney and Odeh 1997) and the spread of the soil parameter estimates is often large. Information about this spread, as reflected by the median absolute deviation from the median (MAD), may be found in the underlying tables (Appendix 3).

The ArcGIS9® project file (*WISE5by5min.mxd*) contains five data frames, one for each 20 cm layer, as examples of possible GIS output. Each frame includes maps of soil pH, organic carbon content, bulk density, base saturation and aluminium saturation for layer D1 to D5 as examples – this selection has been included permanently in the raster file; additional selections can be made using standard GIS-based 'joins'. Besides considering these classes, it is important to look at the relative area that each class covers in a given map unit — for details see table *WISE5binDi* in Appendix 6. In the case of map unit WD600, the bulk density class (BULK) applies to 60% of the map unit (BULKprop; Figure 7). Conversely, the cation exchange capacity (CECs) class represents 90% of this map unit (CECSnum).

The type of GIS files presented here is considered appropriate for exploratory assessments at scale 1:5 million. More detailed studies, however, may need to consider the full map unit composition and soil depth intervals using tailor made programs designed to meet a specific research objective. In some instances, it will be necessary to go back to the primary profile data, make the necessary calculations, and then link results back to the spatial data. For example, when computing soil carbon stocks — which involves combining co-varying data on SOC content, bulk density and content of coarse fragments for each soil horizon to a pre-defined depth (e.g., Batjes 1996b).

Finally, it should be noted that it is better to aggregate model results — for a given spatial and temporal unit — than to aggregate the spatial data before modelling (Bouwman *et al.* 1999; Middelburg *et al.* 1999).

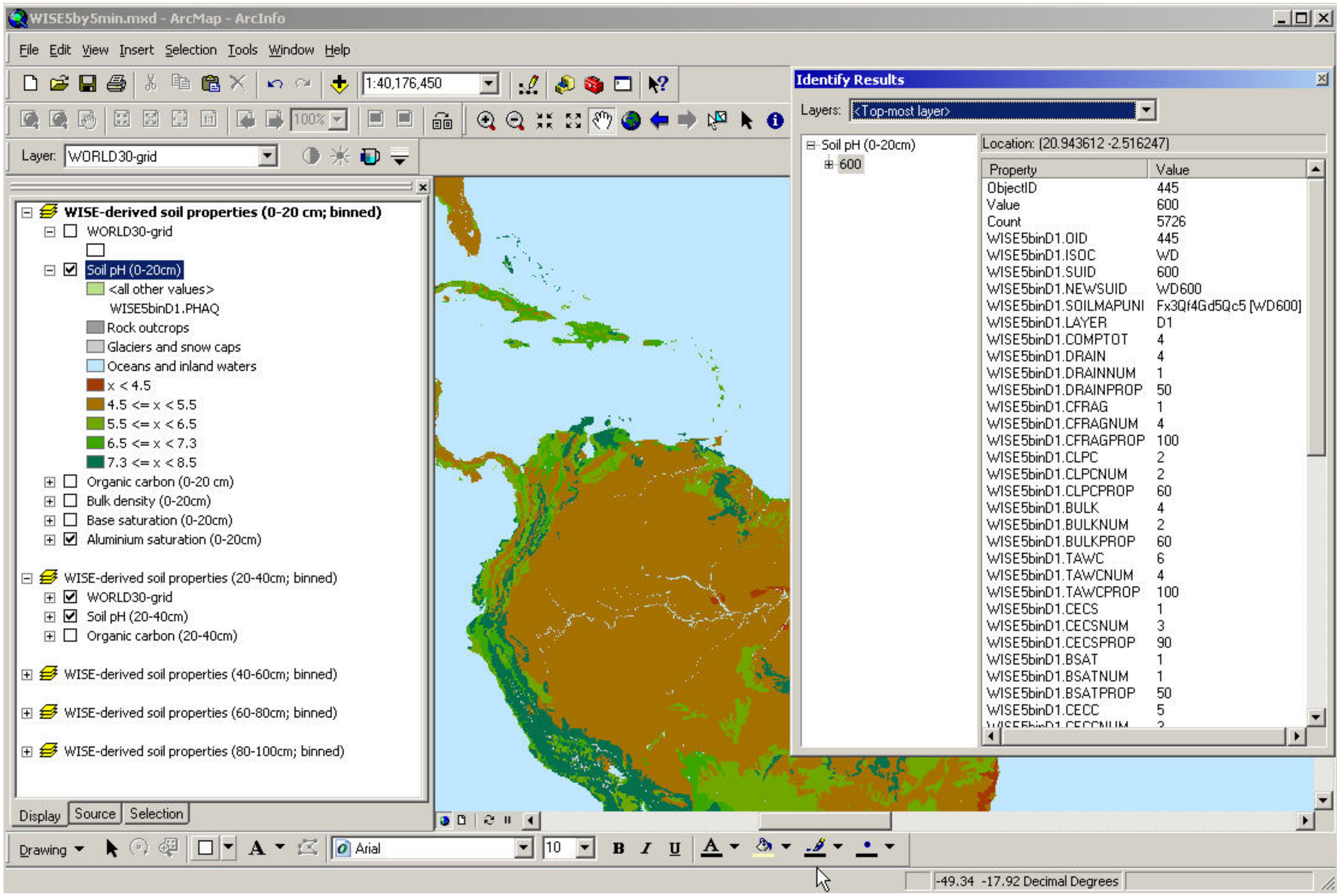


Figure 7. GIS output with details of the attribute data for map unit WD600

4 CONCLUSIONS

Joining WISE-derived soil parameter estimates to the spatial component of the DSMW required generalisation of measured soil (profile) data by soil unit, depth zone and soil textural class. This involved the transformation of soil variables that show a marked spatial and temporal variation. These variables have been determined in a range of laboratories, according to various analytical methods, and over a range of years. Other sources and types of uncertainty that are associated with the spatial data and aggregation procedures, and possible implications for modelling, have been discussed elsewhere (Batjes 1999; Bouwman, Derwent and Dentener 1999; Cramer and Fischer 1997; Dobos *et al.* 2006; Middelburg *et al.* 1999; Oldeman and van Engelen 1993).

There are often gaps or omissions in the information provided in the soil literature or in available auxiliary databases with respect to several of the input variables required for WISE. The soil parameter estimates and derived soil classes presented here should be seen as "best possible" estimates derived from the present selection of measured soil profile data, scheme of taxotransfer and expert-rules, and spatial data of the DSMW. The overall assumption is that the inferred confidence in the modal parameter estimates shown for a given combination of "soil unit-variable-depth zone-soil textural class" should increase with the size of the corresponding sample population. Nonetheless, the uncertainty attached to individual soil parameters estimates can be large — the types of taxotransfer and expert rules used have been documented in the data set to provide an indication of this uncertainty. Some of the soil variables under consideration are only diagnostic for selected soil units (see FAO-Unesco, 1974).

Soil profile data were clustered according to revised SOTER criteria; by soil unit, variable, textural class (5), and depth range (5). Conversely, earlier assessments of derived soil properties for the Soil Map of the World considered only two broad depth classes (0-30 and 30-100 cm) and 3 *topsoil* textural classes, in accordance with FAO (1995a) protocols (Batjes 2002a; Batjes 2005e; Batjes *et al.* 1997). Inherently, changes in the number, spatial distribution, and type of profiles analyzed as well as differences in data clustering procedures, used for the attribute and spatial data, will lead to different parameter estimates and binned maps for any given soil variable.

The TTR-approach followed in this report was originally developed to create harmonized soil input for use with the GEFSOC Soil Carbon Modelling System[3]. This system couples two dynamic C-models, RothC and Century, and the empirical IPCC method for estimating soil carbon

[3] http://www.nrel.colostate.edu/projects/gefsoc-uk/

stocks and change at national scale to a geographical information system. It requires harmonized soil (SOTER), climate and land-use datasets to act as model input (Easter *et al.* 2005; Milne *et al.* 2006). So far, the system has been tested using SOTER-derived[4] soil data for Amazon Brazil, the Indo-Gangetic Plains of India, Jordan and Kenya. Regional scale SOTER data, having the same format, are also available for Latin America and the Caribbean, Central and Eastern Europe and Southern Africa[5]. SOTER activities are in progress for Western Europe, Central Africa and parts of Asia (Bangladesh, Cambodia).

The dataset described in this report is considered appropriate for exploratory assessments at global scale (< 1:5 million; 5 by 5 minutes), pending the completion of an updated, global SOTER product.

[4] http://www.nrel.colostate.edu/projects/gefsoc-uk/GEFSOC%20SYSTEM.htm

[5] http://www.isric.org/NR/exeres/BC5BB444-DB94-4FE3-B102-2872B0B1A9F1.htm

REFERENCES

Anon. 1984. *Memento de l' agronome (troisieme edition)*. Collection "Techniques rurales en Afrique". Ministere des Relations Exterieures - Cooperation et Developpement -, Republique Francaise, Paris, 1604 p

Batjes NH 1995a. *World Inventory of Soil Emission Potentials: WISE 2.1 - Profile database user manual and coding protocols*. Tech. Pap. 26, ISRIC, Wageningen

Batjes NH 1995b. *A homogenized soil data file for global environmental research: A subset of FAO, ISRIC and NRCS profiles (Version 1.0)*. Work. Pap. 95/10, ISRIC, Wageningen

Batjes NH 1996a. Development of a world data set of soil water retention properties using pedotransfer rules. *Geoderma* 71, 31-52

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

Batjes NH 1999. Soil vulnerability mapping in Central and Eastern Europe: Issues of data acquisition, quality control and sharing. In: Naff T (editor), *Data Sharing for International Water Resource Management: Eastern Europe, Russia and the CIS*. NATO Science Series 2: Environmental Security (Vol. 61). Kluwer Academic Publishers, Dordrecht, pp 187-206

Batjes NH 2002a. *Soil parameter estimates for the soil types of the world for use in global and regional modelling (Version 2.1)*. ISRIC Report 2002/02c, International Food Policy Research Institute (IFPRI) and International Soil Reference and Information Centre (ISRIC), Wageningen

Batjes NH 2002b. Revised soil parameter estimates for the soil types of the world. *Soil Use and Management* 18, 232-235

Batjes NH 2002c. *ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (ver. 2.0)*. Report 2002/03, International Soil Reference and Information Centre (ISRIC), Wageningen

Batjes NH 2003a. *A taxotransfer rule-based approach for filling gaps in measured soil data in primary SOTER databases (GEFSOC Project)*. Report 2003/03, ISRIC - World Soil Information, Wageningen

Batjes NH 2003b. *Parameter estimates for the main soil units on a 0.5 by 0.5 degree global grid (ver. 1.1)*. Report 2003/01b, ISRIC - World Soil Information, Wageningen

Batjes NH 2004. *SOTER-based soil parameter estimates for Southern Africa (ver. 1.0)*. Report 2004/04, ISRIC - World Soil Information, Wageningen

Batjes NH 2005a. *SOTER-based soil parameter estimates for Latin America and the Caribbean (ver. 1.0)*. Report 2005/02, ISRIC - World Soil Information, Wageningen

Batjes NH 2005b. *WISE-based soil parameter estimates for North and Central Asia (East of the Ural) (ver. 1.0)*. Report 2005/03, ISRIC - World Soil Information, Wageningen

Batjes NH 2005c. *WISE-based soil parameter estimates for Northeastern Africa (ver. 1.0)*. Report 2005/05, ISRIC - World Soil Information, Wageningen

Batjes NH 2005d. *SOTER-based soil parameter estimates for Central and Eastern Europe (ver. 2.0)*. Report 2005/04, ISRIC - World Soil Information, Wageningen

Batjes NH 2005e. *ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (ver. 3.0)*. Report 2005/08, ISRIC - World Soil Information, Wageningen

Batjes NH and Bridges EM 1994. Potential emissions of radiatively active gases from soil to atmosphere with special reference to methane: development of a global database (WISE). *Journal of Geophysical Research* 99(D8), 16479-16489

Batjes NH and Gicheru P 2004. *Soil data derived from SOTER for studies of carbon stocks and change in Kenya (ver. 1.0; GEFSOC Project)*. Report 2004/01, ISRIC - World Soil Information, Wageningen

Batjes NH, Bridges EM and Nachtergaele FO 1995. World Inventory of Soil Emission Potentials: development of a global soil data base of process-controlling factors. In: Peng S, KT Ingram, HU Neue and LH Ziska (editors), *Climate Change and Rice*. Springer-Verlag, Heidelberg, pp 102-115

Batjes NH, Fischer G, Nachtergaele FO, Stolbovoy VS and van Velthuizen HT 1997. *Soil data derived from WISE for use in global and regional AEZ studies (ver. 1.0)*. Interim Report IR-97-025, FAO/ IIASA/ ISRIC, Laxenburg

Beckett PHT and Webster D 1971. Soil variability - a review. *Soils and Fertilizers* 34, 1-15

Bouwman AF, Derwent RG and Dentener FJ 1999. Towards reliable global bottom-up estimates of temporal and spatial patterns of emissions of trace gases and aerosols from land-use related and natural resources. In: Bouwman AF (editor), *Approaches to Scaling of Trace Gas Fluxes in Ecosystems*. Elsevier, Amsterdam, pp 3-26

Burrough PA 1989. Fuzzy mathematical methods for soil surveys and land evaluation. *Journal of Soil Science* 40, 477-492

Carter AJ and Scholes RJ 2002. *Generating a global database of soil properties (Available on: Global Soil Data Products, CD-ROM (IGBP-DIS), USDA, CSIR, IGBP-DIS, FAO and ISRIC)*, IGBP Data and Information Services, Potsdam

CEC 1985. *Soil Map of the European Communities (1:1,000,000)*. Report EUR 8982, Office for Official Publications of the European Communities, Luxembourg

Chadwick MJ and Kuylenstierna JCI 1990. *The relative sensitivity of ecosystems in Europe to acidic deposition - A preliminary assessment of the sensitivity of aquatic and terrestrial ecosystems*, Stockholm Environmental Institute, University of York, York

Cramer W and Fischer A 1997. Data requirements for global terrestrial ecosystem modelling. In: Walker B and W Steffen (editors), *Global Change and Terrestrial Ecosystems*. Cambridge University Press, Cambridge, pp 529-565

Dijkshoorn JA, Huting JRM and Tempel P 2005. *Update of the 1:5 million Soil and Terrain Database for Latin America and the Caribbean (SOTERLAC, ver. 2.0)*. Report 2005/01 (available online through http://www.isric.org), ISRIC - World Soil Information, Wageningen

Dobos E, F. C, Hengl T, Reuter HI and Toth G 2006. *Digital soil mapping as a support to production of functional maps*. EUR 22123 EN, Office for Official Publications of the European Comminities, Luxembourg

Easter M, Paustian K, Killian K, Boyack K, Williams S, Feng T, Coleman K, Swan A, Al-Adamat R, Bhattacharya T, Cerri CEP, Kamoni P, Batjes NH and Milne E 2005. *User instructions GEFSOC Soil Carbon Modeling System*, Natural Resources Ecology Laboratory (NREL), Colorado State University, Fort Collins (CO)

Eleanor M, Al-Adamat R, Batjes NH, Bernoux M, Bhattacharyya T, Cerri CC, Cerri CEP, Coleman K, Easter M, Falloon P, Feller C, Gicheru P, Kamoni P, Killian K, Pal DK, Paustian K, Powlson D, Rawajfih Z, Sessay M and Wokabi S 2006. National and regional assessments of soil organic carbon stocks and changes: problems and prospects. *Agriculture, Ecosystems and Environment*, (GEFSOC Special Volume, submitted)

FAO-Unesco 1971-1981. *Soil Map of the World, 1:5,000,000. Vol. 1 to 10*. United Nations Educational, Scientific, and Cultural Organization, Paris

FAO-Unesco 1974. *Soil Map of the World, 1:5,000,000. Vol. 1 - Legend*, United Nations Educational, Scientific, and Cultural Organization, Paris

FAO-Unesco 1981. *Soil Map of the World, 1:5,000,000. Vol. 5 - Europe*. United Nations Educational, Scientific, and Cultural Organization, Paris

FAO 1983. *Guidelines: land evaluation for rainfed agriculture*. FAO Soils Bulletin 52, Food and Agriculture Organization of the United Nations, Rome

FAO 1988. *FAO/Unesco Soil Map of the World, Revised Legend (with corrections and updates)*. FAO World Soil Resources Report 60 (reprinted with updates as ISRIC Technical Paper 20 in 1997), ISRIC, Wageningen

FAO 1995a. *Digital Soil Map of the World and Derived Soil Properties*, Food and Agriculture Organization of the United Nations, Rome

FAO 1995b. *The digital Soil Map of the World - Notes (ver. 3.5)*, FA0, Rome

ISIS 2006. ISRIC Soil Information System, ver. 6.0, ISRIC - World Soil Information (Available through: http://www.isric.org/UK/About+Soils/Soil+data/Geographic+data/Global/ISIS.htm)

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

Kern JS 1995. Geographic patterns of soil-water holding capacity in the contiguous United States. *Journal of the American Society for Horticultural Science* 59

Klamt E and Sombroek WG 1988. Contribution of organic matter to exchange properties of Oxisols. In: Beinroth FH, MN Camargo and H Eswaran (editors), *Classification, characterization and utilization of Oxisols. Proc. of the 8th International Soil Classification Workshop (Brazil, 12 to 23 May 1986)*. Empresa Brasileira de Pesquisa Agropecuaria (EMBRAPA), Soil Management Support Services (SMSS) and University of Puerto Rico (UPR), Rio de Janeiro, pp 64-70

Landon JR 1991. *Booker Tropical Soil Manual*. Longman Scientific & Technical, New York, 474 p

McBratney AB and Odeh IOA 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma* 77, 85-113

Middelburg JJ, Liss PS, Dentener FJ, Taminski T, Kroeze C, Malingreau J-P, Noväl M, Panikov NS, Plamnt R, Starink M and Wanninkhof R 1999. Relations between scale, model approach and model parameters. In: Bouwman AF (editor), *Approaches to Scaling of Trace Gas Fluxes in Ecosystems*. Elsevier, Amsterdam, pp 219-232

Milne E, Easter M, Paustian K, Williams SA, Killian K, Swan A, Sessay M, Cerri CEP, Cerri CC, Kamoni P, Gicheru P, Wokabi S, Bhattacharyya T, Pal DK, Gajbhiye KS, Al-Adamar R, Shabbaz M, Rawajfih Z, Khresat S, Powlson D, Coleman K, Batjes NH, Bernoux M, Feller C, Falloon P, Gischer G and Tubiello F 2006. Assessment of Soil Organic Carbon Stocks and Change at National Scale. In: Milne E, M Easter, CE Cerri, K Paustian and S Williams (editors), *Technical Report of the Global Environmental Facility Co-financed Project No. GFL-2740-02-4381*. GEF, UNEP and The University of Reading, Reading (UK), pp 171

Nachtergaele FO and Oldeman LR 2002. World soil and terrain database (SOTER): past, present and future. *Transactions 17th World Congress of Soil Science*. International Union of Soil Sciences (IUSS), Bangkok, pp 653/1-653/10

Oldeman LR and van Engelen VWP 1993. A World Soils and Terrain Digital Database (SOTER) - An improved assessment of land resources. *Geoderma* 60, 309-335

Pleijsier K 1989. Variability in soil data. In: Bouma J and AK Bregt (editors), *Land Qualities in Space and Time*. PUDOC, Wageningen, pp 89-98

Soil Survey Staff 1983. *Soil Survey Manual (rev. ed.)*. United States Agriculture Handbook 18, USDA, Washington

Soil Survey Staff 1996. *Soil survey laboratory methods manual*. Soil Survey Investigations Report 42 (ver. 3), United States Department of Agriculture, Lincoln, NE

Spain AV, Isbell RF and Probert ME 1983. *Soil organic matter*. Soils: An Australian Viewpoint. CSIRO Melbourne Academic Press, London, 552-563 p

Sys IC, van Ranst E, Debaveye IJ and Beenaert F 1993. *Land evaluation (Part I - III)*, Agricultural Publications, General Administration for Development Cooperation, Brussels

van Engelen VWP and Wen TT 1995. *Global and National Soils and Terrain Digital Databases (SOTER): Procedures Manual (rev. ed.)*. (Published also as FAO World Soil Resources Report No. 74), UNEP, IUSS, ISRIC and FAO, Wageningen

Van Engelen VWP, Batjes NH, Dijkshoorn K and Huting J 2005. *Harmonized Global Soil Resources Database (Final Report)*. Report 2005/06, FAO and ISRIC - World Soil Information, Wageningen

van Reeuwijk LP (editor) 1998. Guidelines for Quality management in Soil and Plant Laboratories, Soils Bulletin 74. FAO, Rome, 143 p

van Reeuwijk LP 2002. *Procedures for soil analysis (6th ed.)*. Technical Paper 9, ISRIC, Wageningen

APPENDICES

Appendix 1. Overview of FAO soil units represented in WISE

Soil Unit[a]	Description	Number
Af	Ferric Acrisols	378
Ag	Gleyic Acrisols	76
Ah	Humic Acrisols	109
Ao	Orthic Acrisols	352
Ap	Plinthic Acrisols	113
Bc	Chromic Cambisols	89
Bd	Dystric Cambisols	228
Be	Eutric Cambisols	354
Bf	Ferralic Cambisols	133
Bg	Gleyic Cambisols	137
Bh	Humic Cambisols	112
Bk	Calcic Cambisols	184
Bv	Vertic Cambisols	74
Bx	Gelic Cambisols	25
Cg	Glossic Chernozems	5
Ch	Haplic Chernozems	94
Ck	Calcic Chernozems	71
Cl	Luvic Chernozems	52
Dd	Dystric Podzoluvisols	10
De	Eutric Podzoluvisols	79
Dg	Gleyic Podzoluvisols	20
E	Rendzinas (Ex)	130
Fa	Acric Ferralsols	38
Fh	Humic Ferralsols	76
Fo	Orthic Ferralsols	186
Fp	Plinthic Ferralsols	16
Fr	Rhodic Ferralsols	101
Fx	Xanthic Ferralsols	126
Gc	Calcaric Gleysols	23
Gd	Dystric Gleysols	135
Ge	Eutric Gleysols	284
Gh	Humic Gleysols	66
Gm	Mollic Gleysols	114
Gp	Plinthic Gleysols	63
Gx	Gelic Gleysols	14

[a] According to FAO-Unesco (1974)

(cont.)

Hc	Calcaric Phaeozems	37
Hg	Gleyic Phaeozems	34
Hh	Haplic Phaeozems	159
Hl	Luvic Phaeozems	184
I	Lithosols (Ix)	32
Jc	Calcaric Fluvisols	174
Jd	Dystric Fluvisols	119
Je	Eutric Fluvisols	326
Jt	Thionic Fluvisols	50
Kh	Haplic Kastanozems	29
Kk	Calcic Kastanozems	32
Kl	Luvic Kastanozems	28
La	Albic Luvisols	37
Lc	Chromic Luvisols	228
Lf	Ferric Luvisols	372
Lg	Gleyic Luvisols	225
Lk	Calcic Luvisols	177
Lo	Orthic Luvisols	394
Lp	Plinthic Luvisols	30
Lv	Vertic Luvisols	42
Mg	Gleyic Greyzems	4
Mo	Orthic Greyzems	24
Nd	Dystric Nitosols	67
Ne	Eutric Nitosols	63
Nh	Humic Nitosols	37
Od	Dystric Histosols	85
Oe	Eutric Histosols	81
Ox	Gelic Histosols	6
Pf	Ferric Podzols	6
Pg	Gleyic Podzols	41
Ph	Humic Podzols	45
Pl	Leptic Podzols	20
Po	Orthic Podzols	75
Pp	Placic Podzols	15
Qa	Albic Arenosols	37
Qc	Cambic Arenosols	376
Qf	Ferralic Arenosols	156
Ql	Luvic Arenosols	128
Rc	Calcaric Regosols	56
Rd	Dystric Regosols	88
Re	Eutric Regosols	187
Rx	Gelic Regosols	9

(cont.)

Sg	Gleyic Solonetz	60
Sm	Mollic Solonetz	26
So	Orthic Solonetz	115
Th	Humic Andosols	152
Tm	Mollic Andosols	52
To	Ochric Andosols	34
Tv	Vitric Andosols	53
U	Rankers (Ux)	103
Vc	Chromic Vertisols	210
Vp	Pellic Vertisols	318
Wd	Dystric Planosols	33
We	Eutric Planosols	84
Wh	Humic Planosols	7
Wm	Mollic Planosols	15
Ws	Solodic Planosols	27
Xh	Haplic Xerosols	39
Xk	Calcic Xerosols	66
Xl	Luvic Xerosols	94
Xy	Gypsic Xerosols	18
Yh	Haplic Yermosols	13
Yk	Calcic Yermosols	18
Yl	Luvic Yermosols	20
Yt	Takyric Yermosols	1
Yy	Gypsic Yermosols	23
Zg	Gleyic Solonchaks	36
Zm	Mollic Solonchaks	7
Zo	Orthic Solonchaks	77
Zt	Takyric Solonchaks	6

Note: For computational reasons, the codes for Lithosols (I), Rendzinas (E), and rankers (U) have been changed to Ix, Ex and Ux.

Appendix 2. Soil map unit composition file

This table gives the full composition of each DSMW map unit in terms of its main soil units (FAO-Unesco, 1974), their relative extent, and the identifier for the corresponding virtual soil profile. The contents of this table can be *joined* to the spatial data — using the *VALUE* field in raster file *smw5by5min* — through the *SUID* or *SNUM* field.

Structure of table WISEunitComposition[a]

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994) or WD for World
SUID	Integer	The identification code of a DSMW unit on the map and in the database (corresponds with SNUM on the DSMW)
NEWSUID	Text	Globally unique code, comprising fields ISOC plus SUID (e.g. WD1234)
SoilMapunit[b]	Text	Aggregated code for map unit summarizing the overall composition (see text)
SOIL1[b]	Text	Characterization of the first (main) soil unit according to The FAO-Unesco Legend
PROP1[b]	Integer	Proportion, as a percentage, that the main soil unit occupies within the DSMW unit
PRID1	Text	Unique code for the corresponding virtual soil profile (e.g WD-Fr)
SOIL2	Text	As above but for the next soil unit
PROP2	Integer	As above
PRID2	Text	As above
SOIL3	Text	As above but for the next soil unit
PROP3	Integer	As above
PRID3	Text	As above
SOIL4	Text	As above but for the next soil unit
PROP4	Integer	As above
PRID4	Text	As above
SOIL5	Text	As above but for the next soil unit
PROP5	Integer	As above
PRID5	Text	As above
SOIL6	Text	As above but for the next soil unit
PROP6	Integer	As above
PRID6	Text	As above
SOIL7	Text	As above but for the next soil unit
PROP7	Integer	As above
PRID7	Text	As above
SOIL8	Text	As above but for the next soil unit
PROP8	Integer	As above
PRID8	Text	As above

[a] For the sake of consistency, table structure conventions used for secondary SOTER databases for have been retained here — table WISExxx has the same

structure as table SOTERxxx. SOTER naming conventions are conform those required by the GEFSOC Soil Carbon Modelling System© (Easter *et al.* 2005); parameter estimates presented in this study may be used as default values for those parts of the world for which there are no updates yet in SOTER (keeping in mind the 1:5 million scale of the DSMW).
[b] These fields have been 'embedded' in raster file *smw5by5min* to provide a quick overview of the complexity of each map unit (see text).

Parameter estimates for a given virtual profile can be linked to the above data through the unique profile identifier (PRID, see Appendix 3 and Figure 3).

Appendix 3. Statistical output files

These tables show results of the statistical analyses measured soil data in WISE, by FAO-Unesco (1974) unit, soil attribute, depth layer, and soil textural class. The corresponding information has been stored in 16 tables — in view of their length, these tables are only available in digital format (download via www.isric.org):

- *WISstat9_a_AtoF*: This table holds statistical data by *major soil groupings* (FAO 1974), ranging from Acrisols (A) to Ferralsols (F), considering five soil textural classes: **C**oarse, **M**edium, Medium Fine (**Z**), **F**ine, and **V**ery Fine (Figure 8). Where there is no measured particle size analysis for a given layer – for example for some Histosols - this has been flagged by "-".

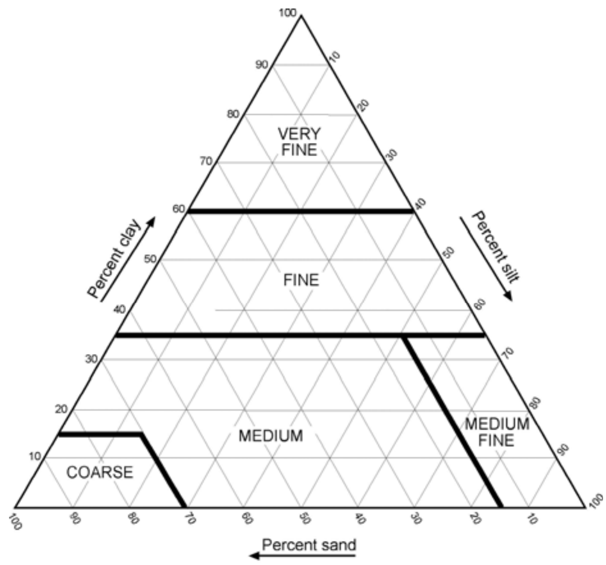


Figure 8. Soil textural classes

- *WISstat9_a_GtoM*: Similar to above but for *major soil groupings* ranging from Gleysols (G) to Greyzems (M)
- *WISstat9_a_NtoS*: Similar to above but for *major soil groupings* ranging from Nitosols (N) to Solonetz (S)
- *WISstat9_a_TtoZ*: Similar to above but for *major soil groupings* ranging from Andosols (T) to Solonchaks (Z)

- *WISstat9_b_AtoF*: This table holds statistical data by *soil unit* (FAO 1988), ranging from Acrisols (A) to Ferralsols (F), considering all five soil textural classes
- *WISstat9_b_GtoM*: Similar to above but for *soil units* ranging from Gleysols (G) to Greyzems (M)
- *WISstat9_b_NtoS*: Similar to above but for *soil units* ranging from Nitosols (N) to Solonetz (S)
- *WISstat9_b_TtoZ*: Similar to above but for *soil units* ranging from Andosols (T) to Solonchaks (Z)

Structure of statistical output tables like *WISstat9_d_YtoZ**:

Name	Type	Description
Short_Id	Text	Code comprising abbreviation for FAO major soil Grouping (resp. soil unit), attribute, depth layer, and soil textural class (e.g., ABSATD1M or AfBSATD2C)
Num0	Integer	Number of observations (before outlier rejection)
Num	Integer	Number of observations (after outlier rejection)
Mean	Single	Mean
STD	Single	Standard deviation
CV	Single	Coefficient of variation
Median	Single	Median
MAD	Single	Median of absolute deviations
Min	Single	Minimum
Max	Single	Maximum
Var	Single	Variance
Fao_74	Text	FAO Legend code (this field is intentionally left blank)

* Notes:

1) Applies to tables: _WISstat9_a_AtoF, _WISstat9_a_GtoM, _WISstat9_c_NtoS, and _WISstat9_c_TtoZ.
2) The structure of tables _WISstat9_b_AtoF, _WISstat9_b_GtoM, _WISstat9_d_NtoS, and _WISstat9_d_TtoZ, is similar to the one listed above except that the first field is called Long_ID. This field differs from Short_ID in that the first two letters refer to the FAO soil unit code (e.g. *AfBSATD1M* for the sample set that relates to base saturation data, BSAT) for Ferric Acrisols (Af) that have medium (M) texture and belong to layer D1 (i.e. 0 to 20 cm).
3) Statistics shown are for depth-weighted data, per layer (from D1 to D5, see text)
4) These tables list results for all analyses, irrespective of the sample size after outlier rejection. The taxotransfer scheme, however, will only consider median (MED) values from the corresponding tables when Num > 5 (see n_{WISE} in text; Appendix 4).

- *WISstat9_c_AtoF*: This table holds statistical data by *soil unit*, ranging from Acrisols (A) to Ferralsols (F), irrespective of soil texture. This has been flagged as class "u", for undifferentiated, which comprises soil textural classes C, M, Z, F and V as well as –
- *WISstat9_c_GtoM*: Similar to above but for *soil units* ranging from Gleysols (G) to Greyzems (M)
- *WISstat9_c_NtoS*: Similar to above but for *soil units* ranging from Nitosols (N) to Solonetz (S)
- *WISstat9_c_TtoZ*: Similar to above but for *soil units* ranging from Andosols (T) to Solonchaks (Z)

- *WISstat9_d_AtoF*: This table holds statistical data by *soil unit*, ranging from Acrisols (A) to Ferralsols (F), irrespective of soil texture. This has been flagged as class "u", for undifferentiated.
- *WISstat9_d_GtoM*: Similar to above but for *major soil groupings* ranging from Gleysols (G) to Greyzems (M)
- *WISstat9_d_NtoS*: Similar to above but for *major soil groupings* ranging from Nitosols (N) to Solonetz (S)
- *WISstat9_d_TtoZ*: Similar to above but for *major soil groupings* ranging from Andosols (T) to Solonchaks (Z)

Appendix 4. Taxotransfer rule-based soil parameter estimates

This table lists soil parameters estimates for all virtual profiles considered in the derived data set . This information can be linked to the geographical component of the DSMW – in a GIS – through the unique profile code (PRID, see Appendix 2).

Structure of table *WISEparameterEstimates*

Name	Type	Description
CLAF	Text	FAO-Unesco (1974) Legend code
PRID	Text	profile ID (as documented in table DSMWComposition)
Drain	Text	FAO soil drainage class
Layer	Text	code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
TopDep	Integer	depth of top of layer (cm)
BotDep	Integer	depth of bottom of (cm)
CFRAG	Integer	coarse fragments (> 2mm)
SDTO	Integer	sand (mass %)
STPC	Integer	silt (mass %)
CLPC	Integer	clay (mass %)
PSCL	Text	FAO texture class
BULK	Single	bulk density (kg dm^{-3})
TAWC	Integer	available water capacity (cm m^{-1}, -33 to -1500 kPa conform to USDA standards)
CECs	Single	cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	base saturation as percentage of CECsoil
CECc	Single	CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	pH measured in water
TCEQ	Single	total carbonate equivalent (g C kg^{-1})
GYPS	Single	gypsum content (g kg^{-1})
ELCO	Single	electrical conductivity (dS m^{-1})
TOTC	Single	organic carbon content (g C kg^{-1})
TOTN	Single	total nitrogen (g kg^{-1})
CNrt	Single	C/N ratio
ECEC	Single	effective CEC ($cmol_c$ kg^{-1})

Notes:
A minus 3 indicates that no meaningful substitution was possible for the specified soil unit and attribute using the present selection of soil profiles, -1 is used for Oceans and inland waters, -2 for Glaciers and snow caps, -7 for rock outcrops (or shallow subsoils) to permit visualization using GIS.

The above table should be consulted in conjunction with table *WISEflagTTRrules* that documents the taxotransfer rules that have been applied (see Appendix 5). Table 6 lists conventions for coding attributes used in the taxotransfer scheme.

Table 6. Conventions for coding attributes in the taxotransfer scheme

SOTnam	WISnam	SoilVariable	TTRflag	Comments
ALSA	ALSA	ALSAT	A	exchangeable Aluminium percentage (% of ECEC)
BSAT	BSAT	BSAT	B	base saturation (% of CECs)
BULK	BULK	BULKDENS	C	Bulk density
CECC	CECC	CECCLAY	D	cation exchange capacity of clay fraction (corrected for organic C)
CECS	CECS	CECSOIL	E	cation exchange capacity
CFRAG	GRAV	GRAVEL	F	coarse fragments % (> 2 mm)
CLPC	CLAY	CLAY	G	clay %
ECEC	ECEC	ECEC	H	effective CEC
ELCO	ECE	ECE	I	electrical conductivity
ESP	ESP	ESP	J	exchangeable Na percentage (as % of CECs)
GYPS	GYPS	GYPSUM	K	gypsum content
PHAQ	PHH2	PHH2O	L	pH in water
SDTO	SAND	SAND	M	sand %
STPC	SILT	SILT	N	silt %
	TAWC	TAWC	O	volumetric water content (-33 to -1500 kPa)
TCEQ	CACO	CACO3	P	Carbonate content
TOTC	ORGC	ORGC	Q	organic carbon content
TOTN	TOTN	TOTN	R	total nitrogen content
CN	CNrt	CN	Z	C/N ratio

Appendix 5. Flagging taxotransfer rules

The type of taxotransfer that have been used when creating table *WISEparameterEstimates* (see Appendix 2) is documented in table *WISEflagTTRrules*. Further details on coding conventions may be found in the text (Section 3.3).

Structure of table *WISEflagTTRrules*

Name	Type	Description
CLAF	Text	FAO Legend code
PRID	Text	Unique identifier for representative profile
Newtopdep	Integer	Depth of top of layer (cm)
Newbotdep	Integer	Depth of bottom of layer (cm)
TTRsub	Text	Codes showing the type of taxotransfer rule used (based on data for soil *units*; see text)
TTRmain	Text	Codes showing the type of taxotransfer rule used (based on data for *major units*; see text)
TTRfinal	Text	Additional flags (based on expert knowledge)

Note: The exchangeable aluminium percentage (ALSA) has been set at zero when pH_{water} is higher than 5.5. Similarly, the electrical conductivity (ELCO), content of gypsum (GYPS) and content of carbonates (TCEQ) have been set at zero when pH_{water} is less than 6.5. Finally, the CEC of the clay fraction (CEC_{clay}) has been re-calculated from the depth-weighted measured and TTR-derived data for CEC_{soil} and content of organic carbon, assuming a mean contribution of 350 $cmol_c$ kg^{-1} OC (Klamt and Sombroek 1988). When applicable, this has been flagged in the field TTRfinal.

Appendix 6. Summary files of derived soil properties

To facilitate access to the derived data, a so-called *summary file* has been created. This file summarizes un-binned parameter estimates —medians — for all component soil units in a given grid cell (e.g. SNUM or SUID) and layer (i.e., D1 to D5), per soil variable.

Structure of table WISE*summaryFile* [6]

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994); WD stands for World
SUID	Integer	The identifier of a DSMW unit on the grid map and in the database (or VALUE in raster file *smw5by5min*, see foonote 4)
NEWSUID	Text	Globally unique map unit code, comprising fields ISOC plus SUID (or SNUM)
TCID	Integer	Number of terrain component in given map unit (Relevant only for SOTER databases; set at 1 by default for the DSMW)
SCID	Integer	Number of soil unit within the given DSMW unit (ranges from 1 to 8 for DSMW)
Layer	Text	Code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm and D5 from 80 to 100 cm)
PROP	Integer	Relative proportion of above in given DSMW unit
CLAF	Text	FAO-Unesco Legend code
PRID	Text	Profile ID (as documented in table *WISE-unitComposition*)
Drain	Text	FAO soil drainage class
TopDep	Integer	Upper depth of layer (cm)
BotDep	Integer	Lower dept of layer (cm)
CFRAG	Integer	Coarse fragments (% > 2mm)
SDTO	Integer	Sand (mass %)
STPC	Integer	Silt (mass %)
CLPC	Integer	Clay (mass %)
PSCL	Text	FAO texture class (see Figure 8)|
BULK	Single	Bulk density (kg dm^{-3})
TAWC	Integer	Available water capacity (cm m^{-1}, -33 to -1500 kPa, USDA standards)
CECS	Single	Cation exchange capacity ($cmol_c$ kg^{-1}) of fine earth fraction
BSAT	Integer	Base saturation as percentage of CECsoil

[6] The structure of this table is indentical to that used for secondary SOTER databases (see: SOTER*summaryFile)*

(cont.)

CECc	Single	CECclay, corrected for contribution of organic Matter (cmol$_c$ kg^{-1})
PHAQ	Single	pH measured in water
TCEQ	Single	Total carbonate equivalent (g C kg^{-1})
GYPS	Single	Gypsum content (g kg^{-1})
ELCO	Single	Electrical conductivity (dS m^{-1})
TOTC	Single	Organic carbon content (g kg^{-1})
TOTN	Single	Total nitrogen (g kg^{-1})
CNrt	Single	C/N ratio
ECEC	Single	Effective CEC (cmol$_c$ kg^{-1})

Notes:
1) These are depth-weighted values, per 20 cm layer.
2) Components within a given DSMW unit are numbered starting with the spatially dominant one. The sum of the relative proportions of all soil units within a DSMW unit is always 100 per cent.
3) Tables having the same structure have also been prepared for the DOMINANT soil unit only, by depth layer (i.e., for layer D1 to D5, see files: *WISEsummaryFile_T1S1D1* etc. to facilitate linkage to GIS).
4) The information in this table can be linked to the grid file (*smw5by5min*) using VALUE resp. *Suid* and the SUID code in the above table; the later corresponds to FAO's soil map unit number (SNUM) (Appendix 4).

In view of the map unit complexity, additional operations will often be needed before results can be visualized using GIS. As indicated in the report, two types of output files were generated to facilitate this linkage:

Method A: Tables showing un-binned soil parameter estimates for the dominant soil unit only, for each 20 cm layer up to 100 cm depth. The structure of this table is similar to that of table WISE*summaryFile* (with as extra selection TCID=1 and SCID=1, and Layer= D1 to resp. D5)

Method B: Tables showing binned or classified data per grid cell, and 20 cm layer, that take into account the full map unit composition. The corresponding tables are called *WISE5binDi*, where *Di* stands for the 20 cm layer under consideration. Criteria for allocating each grid cell to a given legend unit, per soil variable, are presented in Appendix 7.

Structure of tables *WISE5binDi*:

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994), WD for WORLD
SNUM	Integer	The identification code of a DSMW unit, or grid cell, on the map and in the database (or SUID)
MAPUNIT	Text	Globally unique map unit code, comprising fields ISOC plus SNUM, e.g WD123

Layer	Text	Code for depth layer *Di* (from D1 to D5; e.g. D1 is from 0 to 20 cm)
COMPtot	Integer	Total number of soil units or miscellaneous units in grid cell or map unit

(cont.)

SOLnum	Integer	Number of soil units belonging to the given Class
SOLprop	Integer	Relative (total) proportion of above soil units
DRAIN	Integer	FAO soil drainage class
DRAINnum	Integer	Number of soil units considered when rating above class
DRAINrop	Integer	Relative area of above soil units (%)
CFRAG	Integer	Class number for Coarse fragments (see Appendix 7 for Legend)
CFRAGnum	Integer	Number of soil units considered when rating above class
CFRAGprop	Integer	Relative area of above soil units (%)
CLPC	Integer	Class for Clay %
CLPCnum	Integer	Number of soil units considered when rating above class
CLPCprop	Integer	Relative area of above soil units (%)
BULK	Integer	Class for Bulk density (kg dm^{-3})
BULKnum	Integer	Number of soil units considered when rating above class
BULKprop	Integer	Relative area of above soil units (%)
TAWC	Integer	Class for Available water capacity (cm m^{-1}, -33 to -1500 kPa, USDA standards; corrected for presence of coarse fragments)
TAWCnum	Integer	Number of soil units considered when rating above class
TAWCprop	Integer	Relative area of above soil units (%)
CECS	Single	Class for Cation exchange capacity ($cmol_c$ kg^{-1}) of fine earth fraction
CECSNum	Integer	Number of soil units considered when rating above class
CECSprop	Integer	Relative area of above soil units (%)
BSAT	Integer	Class for Base saturation as percentage of CECsoil
BSATnum	Integer	Number of soil units considered when rating above class
BSATprop	Integer	Relative area of above soil units (%)
CECc	Single	Class for CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
CECCnum	Integer	Number of soil units considered when rating above class
CECCprop	Integer	Relative area of above soil units (%)
PHAQ	Single	Class for pH, measured in water
PHAQnum	Integer	Number of soil units considered when rating above class
PHAQprop	Integer	Relative area of above soil units (%)
TCEQ	Single	Class for total carbonate equivalent (g kg^{-1})
TCEQnum	Integer	Number of soil units considered when rating above class
TCEQprop	Integer	Relative area of above soil units (%)

(cont.)

GYPS	Single	Class for gypsum content (g kg^{-1})
GYPSnum	Integer	Number of soil units considered when rating above class
GYPSprop	Integer	Relative area of above soil units (%)
ELCO	Single	Class for electrical conductivity (dS m^{-1})
ELCOnum	Integer	Number of soil units considered when rating above class
ELCOprop	Integer	Relative area of above soil units (%)
TOTC	Single	Class for organic carbon content (g kg^{-1})
TOTCnum	Integer	Number of soil units considered when rating above class
TOTCprop	Integer	Relative area of above soil units (%)
TOTN	Single	Class for Total nitrogen (g kg^{-1})
TOTNnum	Integer	Number of soil units considered when rating above class
TOTNprop	Integer	Relative area of above soil units (%)
CNrt	Single	Class for C/N ratio
CNRTnum	Integer	Number of soil units considered when rating above class
CNRTprop	Integer	Relative area of above soil units (%)

Appendix 7. Legends for binned data sets

In view of the 1:5 million scale, legends for the GIS files comprise from 4 to 6 classes — depending on the soil variable under consideration —, plus three miscellaneous classes:

Class number:	1	2	3	4	5	6	
	<——>	<——>	<——>	<——>	<——>	<——>	
Class limits:	A0	A1	A2	A3	A4	A5	A6
Soil variable:							
TAWC	0.0	4.0	8.0	12.0	14.0	16.0	10000
ALSA	0.0	15.0	30.0	45.0	10000	10000	10000
CECC	0.0	4.0	8.0	16.0	24.0	48.0	10000
TOTC	0.0	5.0	10.0	20.0	100.0	10000	10000
TOTN	0.0	0.5	1.0	2.0	10.0	10000	10000
TCEQ	0.0	20.0	100.0	200.0	10000	10000	10000
GYPS	0.0	20.0	100.0	200.0	10000	10000	10000
PHAQ	3.0	4.5	6.5	7.3	8.5	14.0	10000
ELCO	0.0	5.0	15.0	25.0	10000	10000	10000
CECS	0.0	10.0	20.0	40.0	10000	10000	10000
BSAT	0.0	25.0	50.0	75.0	10000	10000	10000
CLPC	0.0	20.0	40.0	60.0	80.0	10000	10000
CFRAG	0.0	5.0	15.0	25.0	10000	10000	10000
BULK	0.0	0.4	0.9	1.2	1.4	1.6	10000
ESP	0.0	5.0	15.0	30.0	10000	10000	10000
CNrt	5.0	10.0	15.0	20.0	25.0	10000	10000
DRAIN	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0

Notes:

1) Soil drainage was aggregated into five classes using: class 1 for Very poorly and poorly drained soils; class 2 for imperfectly drained soils; class 3 for moderately well drained soils; class 4 for well drained soils; and, class 5 for somewhat excessively and excessively drained soils. Rock outcrops were assigned to the class excessively drained.
2) The value of 10000 has been used as an artificial upper limit to define the highest class (CL_i). The number of classes considered (excluding classes -1, -2 and -7) ranges from 4 to 6. In addition, should there be no estimates for a given soil parameter this has been flagged as class -3, for no data.
3) The same legend has been used for a given soil variable, irrespective of the depth layer under consideration (D1 to D5). Not all these classes, however, need occur for all depth ranges.

Class limits for mapping soil pH are listed below, as an example. In addition to listing the pH class (PHAQ), table *WISE5bin_Di* also lists the number of soil units whose properties fall within the given class (PHAQnum) in the grid cell under consideration, and the relative area

(PHAQprop) thereof, as well as the total number of soil units that occur in the given grid cell (COMPtot; see Appendix 6, under Method B).

Example of classes used for mapping soil pH:

Class 1:	3.0 <= PHAQ < 4.5
Class 2:	4.5 <= PHAQ < 5.5
Class 3:	5.5 <= PHAQ < 6.5
Class 4:	6.5 <= PHAQ < 7.3
Class 5:	7.3 <= PHAQ < 8.5
Class 6:	8.5 <= PHAQ < 14.0
Class -1:	Oceans and inland waters, predominant in grid cell
Class -2:	Glaciers and snowcaps, predominant in grid cell
Class -7:	Rock outcrops, respectively shallow rocky soils, predominant in grid cell for specified depth layer. For example, in case of Lithosols, soil parameter estimates for layer D2 to D5 have been set at -7, while for Rendzinas and Rankers this has been done for layers D3 to D5. (Class -3 stand for No Data).

Appendix 8. Installation procedure

The data set and GIS-files are provided in one single zip file: *WISE5by5min_v1.zip* (about 103Mb compressed and 15Mb decompressed).

By default, the compressed file will be unzipped to folder *X:\WISE5by5min*, where *X* is the actual location. The first time the project file is loaded on a new system, the new folder settings will be automatically updated.

This folder will contain 5 subfolders as described in Table 7. The GIS project file includes selected binned data sets, as examples of possible output. These classified data consider the full map unit composition (see text).

Linkage of the spatial (raster) data to the derived soil data is always through the map unit code or grid cell identifier (*VALUE* or *Suid* in the raster file; *SNUM* code or *SUID* in the soil parameter files). Methodological and technical details may be found in the documentation:

> Batjes NH 2006. ISRIC-WISE global data set of derived soil properties on a 5 by 5 arc-minutes grid (version 1.1). Report 2006/02 (available through: http://www.isric.org), ISRIC – World Soil Information, Wageningen (with data set)

The following software packages are needed to access the data: MS Access® and ArcGIS9®. The GIS file was developed for a 17 inch screen.

Table 7. Overview folders and files in WISE5by5min dataset

Folder	Description
X:\WISE5by5min\DBF	Folder with selected files in dBASE4 format
X:\WISE5by5min\GISfiles	Global raster grids (from ESRI)
X:\WISE5by5min\GRID	Raster file — 5 by 5 arcminutes — derived from the 1:5 million scale Soil Map of the World (Section 2.1); GIS info files
X:\WISE5by5min\LayerFiles	Selected layer files
X:\WISE5by5min\ReadMe	Documentation and Readme1st file
X:\WISE5by5min\	
ISRIC-disclaimer.aux	
ISRIC-disclaimer.tif	ISRIC disclaimer
SMW5by5min.aux	-
smw5by5min.rrd	-
WISE5by5min.mdb	MS Access® database with the derived soil properties
WISE5by5min.mxd	ArcMap® project file
WISE5by5min.mxd.xml	Metadata (ISO 19115) for GIS data set

(cont.)

X:\WISE5by5min\DBF	
WISE5binD1.DBF	Binned data for 0-20cm, in dBASEIV format (App. 6)
WISE5binD1.INF	-
WISE5binD1.MDX	-
WISE5binD1.SUID.atx	-
WISE5binD2.DBF	Binned data for 20-40cm, in dBASEIV format (App. 6)
WISE5binD2.INF	-
WISE5binD2.MDX	-
WISE5binD3.DBF	Binned data for 40-60cm, in dBASEIV format (App. 6)
WISE5binD3.INF	-
WISE5binD3.MDX	-
WISE5binD4.DBF	Binned data for 60-80cm, in dBASEIV format (App. 6)
WISE5binD4.INF	-
WISE5binD4.MDX	-
WISE5binD5.DBF	Binned data for 80-100cm, in dBASEIV format(App.6)
WISE5binD5.INF	-
WISE5binD5.MDX	-
WISEsummaryFile_T1S1D1.DBF	dBASEIV file with the un-binned parameter estimates
WISEsummaryFile_T1S1D1.INF	for the dominant soil unit (0-20cm; App.6)
WISEsummaryFile_T1S1D1.MDX	-
WISEsummaryFile_T1S1D2.DBF	dBASEIV file with the un-binned parameter estimates
WISEsummaryFile_T1S1D2.INF	for the dominant soil unit (20-40cm; App.6)
WISEsummaryFile_T1S1D2.MDX	-
WISEsummaryFile_T1S1D3.DBF	dBASEIV file with the un-binned parameter estimates
WISEsummaryFile_T1S1D3.INF	for the dominant soil unit (40-60cm; App.6)
WISEsummaryFile_T1S1D3.MDX	-
WISEsummaryFile_T1S1D4.DBF	dBASEIV file with the un-binned parameter estimates
WISEsummaryFile_T1S1D4.INF	for the dominant soil unit (60-80cm; App.6)
WISEsummaryFile_T1S1D4.MDX	-
WISEsummaryFile_T1S1D5.DBF	dBASEIV file with the un-binned parameter estimates
WISEsummaryFile_T1S1D5.INF	for the dominant soil unit (80-100cm; App.6)
WISEsummaryFile_T1S1D5.MDX	-
WISEunitComposition.DBF	dBASEIV file describing the map unit (Section 3.2)
WISEunitComposition.INF	-
WISEunitComposition.MDX	-
WISEunitComposition.SUID.atx	-
X:\WISE5by5min\GISfiles	
WORLD30.DBF	15 by 15 arcminutes global grid (from ESRI©)
WORLD30.SBN	-
WORLD30.SBX	-
WORLD30.SHP	-
WORLD30.shp.xml	-
WORLD30.SHX	-
X:\WISE5by5min\GRID\info	
arc.dir	ESRI info files
arc0003.dat	-
arc0003.nit	-
arc0004.dat	-
arc0004.nit	-
arc0005.dat	-
arc0005.nit	-
arc0005r.001	-

(cont.)

X:\WISE5by5min\LayerFiles	
Aluminium saturation (0-20cm).lyr	Binned Al-saturation files; same for 20- 40 cm, 40-60 cm and 80-100 cm; see App. 7 for Legend
Base saturation (0-20cm).lyr	"
Bulk density (0-20cm).lyr	"
Organic carbon (0-20 cm).lyr	"
Soil pH (0-20cm).lyr	"
X:\WISE5by5min\ReadMe	
ISRIC Report 2006_02.pdf	Documentation in PDF format
Readme1st_WISE5by5min.pdf	Readme first file
X:\WISE5by5min\GRID\smw5by5min	
dblbnd.adf	Raster grid, info files
hdr.adf	-
metadata.xml	Metadata for raster data
prj.adf	-
sta.adf	-
vat.adf	-
w001001.adf	-
w001001x.adf	-

Note:
The binned data that have been used as examples have been incroporated to the raster set; the field which is to be used for any other 'joins' is VALUE.

World Soil Information

ISRIC - World Soil Information is an independent foundation with a global mandate, funded by the Netherlands Government. We have a strategic association with Wageningen University and Research Centre.

Our aims:

- *To inform and educate - through the World Soil Museum, public information, discussion and publication*
- *As ICSU World Data Centre for Soils, to serve the scientific community as custodian of global soil information*
- *To undertake applied research on land and water resources*