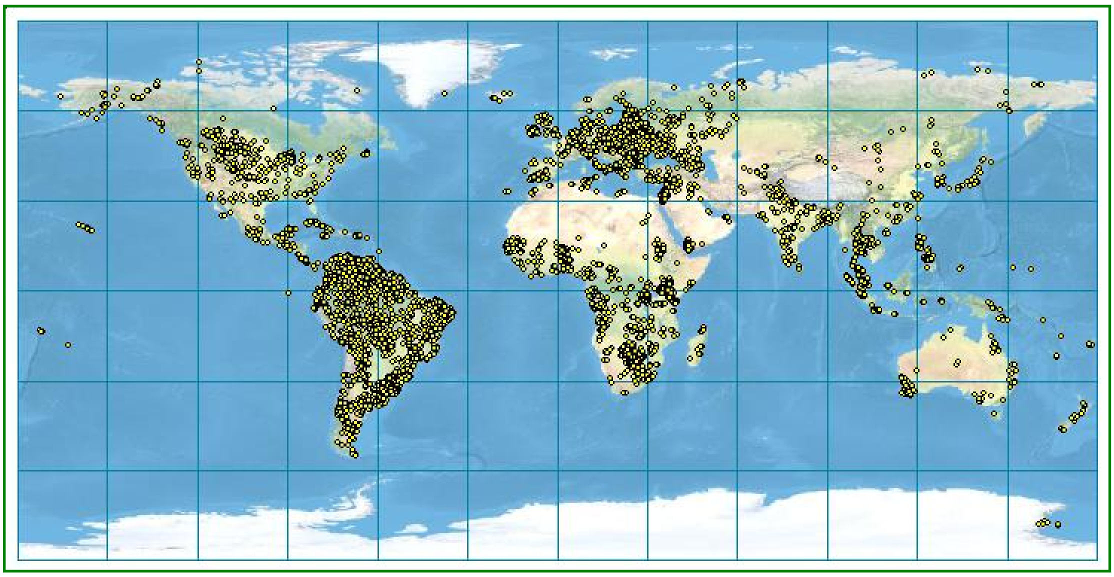
Report 2008/02

ISRIC-WISE Harmonized Global Soil Profile Dataset
(Ver. 3.1)

Niels H Batjes
(December 2008)



World Soil Information

Disclaimer:

Citation:

Technical documentation:
Batjes NH 2008. ISRIC-WISE Harmonized Global Soil Profile Dataset (Ver. 3.1). Report 2008/02, ISRIC – World Soil Information, Wageningen (with dataset).
(http://www.isric.org/isric/Webdocs/Docs/ISRIC_Report_2008_02.pdf)

Article in primary Journal (preferred citation):
Batjes, N.H., 2009. 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)

Inquiries:
c/o Director, ISRIC – World Soil Information
PO B08 353
6700 AJ Wageningen
The Netherlands
Telefax: +31-(0)317-471700
E-mail: soil.isric@wur.nl
Web: www.isric.org

Front cover: Global distribution of geo-referenced soil profiles in WISE3

Contents

Appendices

Tables

Figures

SUMMARY

Version 3.1 of the ISRIC-WISE database (WISE3) was compiled from a wide range of soil profile data collected by many soil professionals worldwide. All profiles have been harmonized with respect to the original Legend (1974) and Revised Legend (1988) of FAO-Unesco. Thereby, the primary soil data — and any secondary data derived from them — can be linked using GIS to the spatial units of the Soil Map of the World as well as more recent Soil and Terrain (SOTER) databases through the soil legend code.

WISE3 holds selected attribute data for 10,253 soil profiles, with some 47,800 horizons, from 149 countries. Individual profiles have been sampled, described, and analyzed according to methods and standards in use in the originating countries. There is no uniform set of properties for which all profiles have analytical data, generally because only selected measurements were planned during the original surveys. Methods used for laboratory determinations of specific soil properties vary between laboratories and over time. Sometimes, results for the same property cannot be compared directly. WISE3 will inevitably include gaps, being a compilation of legacy soil data derived from traditional soil survey. These can be of a taxonomic, geographic, and soil analytical nature. As a result, the amount of data available for modelling is sometimes much less than expected. Adroit use of the data, however, will permit a wide range of agricultural and environmental applications at a global and continental scale (1:500 000 and broader).

Keywords: soil profiles, legacy data, global dataset, modelling, ISRIC-WISE database

1 INTRODUCTION

The compilation and processing of broad-scale datasets of the world's environmental resources, using well-documented procedures and standards, are crucial for many assessments at global and continental scale. Between 1991 and 1995, staff at ISRIC has developed a uniform methodology for a global soil profile database in the framework of a project entitled "World Inventory of Soil Emission Potentials" (WISE, see Batjes 2002c; Batjes and Bridges 1994; Batjes *et al.* 1995).

Profiles in WISE have been classified according to the original and the revised FAO-Unesco Legend (FAO-Unesco 1974; FAO 1988). Thereby, interpretations derived from WISE can be linked to the spatial units of the digitized Soil Map of the World (DSMW, see FAO 1995), which uses the original Legend, as well as to more recent digital soil and terrain (SOTER) databases that use the Revised Legend.

The preceding public domain release of WISE comprised data for some 4300 profiles, while there were some 9600 profiles in an interim, in-house version (Batjes 2002a; Batjes *et al.* 2007). These *primary* data have been used for a range of applications. For example, the development of harmonized sets of derived soil properties for the main soil types of the world, gap-filling in primary SOTER databases using pedotransfer procedures, global modelling of environmental change, analyses of global ecosystems, up-scaling and down-scaling of greenhouse gas emissions, and crop simulation and agro-ecological zoning; detailed references are given on the ISRIC-website (ISRIC-WISE 2008).

This report describes the structure and contents of version 3.1 of the ISRIC-WISE database (hereafter referred to as WISE3), which is available on-line at www.isric.org. It consists of five Sections and seven Appendices. Section 2 describes the structure of the relational database (2.1), criteria for accepting profiles in WISE (2.2), main sources of data (2.3), and the comparability of soil analytical data (2.4). Section 3 discusses the contents of the database in terms of its geographic (3.1), taxonomic coverage (3.2) and temporal coverage (3.3), and provides a summary of the available data (3.4). The applicability and limitations of the dataset are discussed in Section 4, while concluding remarks are made in Section 5. The database structure, protocols for coding data, data control procedures, and detailed summaries of the available data are presented in Appendix 1 to 8.

2 MATERIALS AND METHODS

2.1 Data structure

WISE3 is a relational database, compiled using MSAccess®. It can handle data on: (a) soil classification and site data; (b) soil horizon data; (c) source of data; and (d) methods used for determining analytical data. It also includes a code-definition translation file or look-up table.

The attributes considered in WISE3 are listed in Table 1, while the structure of the various attribute tables are presented in Appendix 1 — data coding protocols are given in Appendix 2.

Data for aluminium saturation, effective CEC, CEC of the clay-size fraction, exchangeable sodium percentage and profile available water capacity, for example, are not included in WISE3 as these may be derived from the measured data.

2.2 Criteria for accepting soil profile data

Strict criteria have been defined for accepting profiles into WISE: (a) completeness and *apparent* reliability of data; (b) traceability of source of data; and (c) classifiable in the original and revised FAO-Unesco Legend (1974; 1988). Where possible, profiles are geo-referenced within defined limits. Upon their entry into WISE, the data have been screened for inconsistencies using visual and automated procedures. These include, for example, checks on consistency of successive horizon depths, whether C/N ratio's and pH values are within the normal range for most soils, and if the sum of sand, silt and clay is equal to 100 percent (Appendix 3; Batjes 1995, p. 52).

Table 1. List of WISE3 attribute data

Site Data	Horizon Data
WISE3_ID [a]	***WISE3_ID + horizon_No*** [b]
General:	***General:***
Soil profile description status	horizon number
Description, year of	depth, top
Description, month of	depth, bottom
Depth of soil (described/sampled)	horizon designation
Number of horizons	matrix colour, moist
	matrix colour, dry
Soil classification:	
FAO-Unesco 1974 Legend	***Chemical attributes:*** [c]
phase (1974)	organic carbon
FAO-Unesco 1988 Legend	total N
phase (1988)	electrical conductivity (EC_x)
WRB 2006 Reference Soil Group	free $CaCO_3$
USDA Soil Taxonomy	$CaSO_4$
Edition (year) of USDA Soil Taxonomy	pH-H_2O
National system	pH-KCl
	pH-$CaCl_2$
	exchangeable Ca^{2+}
Location:	exchangeable Mg^{2+}
Country (ISO code)	exchangeable Na^+
Latitude (N/S, deg/min/sec)	exchangeable K^+
Longitude (W/E, deg/min/sec)	exch. Al^{3+}+ H^+ (exch. acidity)
Location (descriptive)	exch. Al^{3+} (exch. aluminium)
	cation exchange capacity (CEC)
General site data:	base saturation (as % of CEC)[d]
Climate (Köppen)	
Altitude	***Physical attributes:*** [c]
Major landform	weight % sand
Landscape position	weight % silt
Slope	weight % clay
Parent material	volume % > 2 mm [d]
Drainage class	bulk density
Land use	volume per cent water held at specified tensions [d]
Source of data:	
Source of data	
Pit identifier	
Laboratory identifier	
Country name (in full)	

[a] Unique identifier (primary key) for profile in WISE3; [b] unique reference number for horizon within a profile; [c] analytical methods can be found through a separate key-attribute look-up table (see App. 1 under *WISE3_coding_conventions*). [d] Derived or calculated values.

2.3 Data sources and methods

Profile data in WISE3 originate from over 260 different sources, both analogue and digital. Some 40% of the profiles were extracted from auxiliary datasets (Table 2), including various Soil and Terrain (SOTER) databases and the FAO Soil Database (FAO-SDB), which, in turn, hold data collated from a wide range of sources.

Table 2. Source of profile data held in WISE3

Source of data	Number of profiles
SOTER	4252
Analogue sources [a]	2985
SDB-FAO	1570
NRCS-USDA	761
ISIS-ISRIC	685
Total	10253

[a] Data collated from a wide range of national contributions as well as soil survey reports represented in the ISRIC – World Soil Information Database (ISRIC 2007).

Individual profiles were selected as representative of a specific FAO soil unit rather than any particular location. The legacy profiles were described and analyzed using various methods; the latter, however, were not always cited comprehensively in the source materials. Similarly, numerous profiles were not or poorly geo-referenced in the source materials.

Profiles from the ISRIC - Soil Information System (ISIS) have been compiled specifically to be representative of the map units of the Soil Map of the World, with emphasis on the tropics. They have been described using the *Guidelines for Soil Description* (FAO 2006a; FAO/ISRIC 1990) and analysed in a standardized manner in the ISRIC laboratory (van Reeuwijk 1995, 2002).

Profile descriptions from the Natural Resources Conservation Service (NRCS-USDA) follow the methodology of the *Soil Survey Manual* (Soil Survey Staff 1983) and *Field book for describing and sampling soils* (Schoenberger *et al.* 2002); analyses have been made at the Lincoln laboratory (Soil Survey Staff, 1996). Data were derived from a magnetic tape specifically prepared for the WISE Project (through John Kimble) as well as the NSSC Soil Survey Laboratory Soil Characterization Database (USDA-NRCS 2008). In general, analytical methods used by USDA-NRCS compare well with those used for ISIS-ISRIC (John Kimble and Piet Van Reeuwijk, *pers. comm.* 1994)(Vogel 1994).

Profile data in SOTER, however, have been sampled, described and analysed using a range of different methods (e.g. FAO and ISRIC 2003; FAO and ISRIC 2000; FAO *et al.* 1998); this also applies for data derived from the Soil Map of the World reports (FAO-Unesco 1971-1981) and the SDB database (FAO 1989).

Overall, chemical and physical analyses have taken place in at least 150 laboratories worldwide, using a range of methods. The latter are described in broad terms in WISE3 — details may be found in the various source materials (see Appendix 1). Methods typically vary from one laboratory to the next, even within one country, and may change over time within a single laboratory. These methodological differences complicate the worldwide comparison of soil analytical data (Batjes *et al.* 1997; Breuning-Madsen and Jones 1998; FAO/IIASA/ISRIC/ISSCAS/JRC 2008; Hollis *et al.* 2006; van Reeuwijk 1983; Vogel 1994; Wilson 1979).

2.4 Comparability of soil analytical data

Issues of quality and comparability of analytical data collated from disparate sources are critical in any analysis of soil profile data. Yet, there are no straightforward solutions to data harmonization. These options will vary with the analytical methods and soil attributes under consideration (Hollis *et al.* 2006; Nettleton *et al.* 1996; Pleijsier 1989; Skjemstad *et al.* 2000; van Reeuwijk and Houba 1998). Given the large variation of soil properties, even within soil units that are considered to be taxonomically homogeneous (Batjes 1997; Beckett and Webster 1971; Bernoux *et al.* 1998; Spain *et al.* 1983), many possibly imprecise measurements may be considered more efficient than a few expensive ones carried out in a few reference laboratories (Batjes 2002c; FAO 1995; Minasny and McBratney 2002), particularly for broad-scale applications of the data. An additional legitimate excuse to employ analytical data with inherent uncertainties is the inherent fuzziness of boundary bars, or quantitative class limits, used in international soil classification systems (van Reeuwijk 1983). When classifying soils according to the FAO system, for example, it is generally assumed that all soil analyses have been carried out in accordance with pre-defined laboratory procedures (e.g., FAO 1988, p. 127-131). In practice, however, out of necessity the actual classification is based on the available, albeit locally analysed, data.

Defining too stringent rules on data collection and data comparison would have excluded many legacy data from inclusion in WISE. Conversely, errors associated with measurement of, for example, particle size distribution, organic matter content and bulk density will be reflected in the accuracy of pedotransfer functions derived from the measured data; cross-correlations between uncertain variables determine how uncertainties propagate in a modelling study (Heuvelink and Brown 2006). Accuracy levels considered to be acceptable will vary with the scale and type of questions being asked (Finke 2006; Middelburg *et al.* 1999). Practical solutions that are considered commensurate with the scale of the projected applications of the data will often have to be developed. The ISRIC-WISE profile database was specifically developed for applications at global and continental scale, say 1:500 000 or broader.

3 DATABASE CONTENTS

3.1 Geographic distribution

WISE3 holds data for 10,253 profiles from 149 countries, compared to some 4380 profiles for the preceding, public domain version (Batjes 2002b) and some 9600 profiles for the preceding, interim version (see Batjes 2006; Batjes *et al.* 2007). Most profiles are from Africa (41 %), followed by Asia (18 %), South America (18 %), and Europe (13%) (Table 3); their approximate location is shown in Figure 1.

Table 3. Number of profiles by continent and their description status

Continent	Profile description status[a]				Total
	1	2	3	4	
Africa	421	1337	2392	23	4173
Asia	441	970	426	10	1847
Antarctica	4	6	0	0	10
Europe	225	712	359	20	1316
North America	495	222	127	11	855
Oceania	50	49	106	4	209
South America	149	1380	313	1	1843
Total	1785	4676	3723	69	10253

[a] The number code under "profile description status" refers to the completeness and apparent reliability of the soil profile descriptions and accompanying analytical data for the specified profile in the original source. The status is highest for "1" and lowest for "4" (see under *DESCR* in Appendix 2).

The description status is used here as a *coarse* indicator for the inferred quality and completeness of profile data stored in WISE3; it is indicative of the reliability of soil information entered into a database (FAO 2006a). Essentially, the description status is determined by the confidence the various data compilers had in the various sources of soil profiles that are globally available and accessible (see Appendix 2); analytical data for profiles having a description status of 1 or 2 may be considered as most reliable. Nonetheless, profiles flagged as having a description status of 1 or 2 will commonly show gaps in the measured data, in particular soil physical attributes. Often, even the so-called mandatory analytical attributes required by SOTER (van Engelen and Wen 1995, p. 17) are simply not available — abundance of fragments > 2 mm, particle size distribution (sand, wt% 2.0-0.05 mm; silt, wt% 0.05-0.002 mm; and clay, wt% <

0.002 mm); bulk density; pH-H_20; exchangeable Ca^{2+}, Mg^{2+}, K^+, Na^+, and Al^{3+}; CEC_{soil}; and, organic carbon content. Similarly, descriptive information on essential site data, such as climate, parent material, and land use as well as detailed locations, is not available in many source materials.

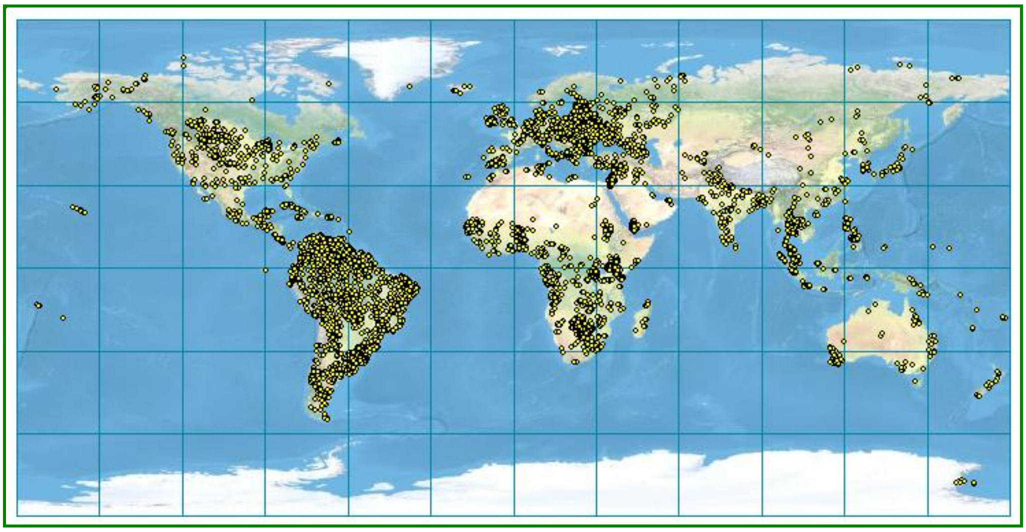


Figure 1. Global distribution of geo-referenced soil profiles in WISE3

About 15 per cent of the profiles in WISE3 were given a description status of 1, corresponding with so called "reference pedons" (Table 3). Some 48 per cent correspond with "routine profile descriptions" in which no essential data are lacking from the description, sampling, or analysis. Category 3, amounting to 36 per cent of the total, corresponds with profiles whose descriptions are useful for specific purposes and provide a satisfactory indication of the nature of the soils in the FAO-Unesco Legend. The fourth and last category (1%) corresponds with incomplete profile descriptions: such profiles are only accepted for inclusion in case of under-represented soil units for which specific measured attributes have been found to be particularly scarce. Typically, status 4 profiles are deleted from WISE when better data become available for the corresponding soil units.

The number of profiles by broad geographical region (UN conventions) and country, with their inferred description status, is listed in Appendix 6 and 7 respectively.

3.2 Taxonomic coverage

The relative number of soil profiles, available for each major soil group of the Legend and Revised Legend respectively, is shown in Figure 2 and 3; details at soil unit level may be found in Appendix 4 and 5. Acrisols (AC), for example, account for about 8% of the total extent of world soils and for some 9% of the total number of profiles in WISE3 (Fig. 2). Conversely, spatially Leptosols (LP) occupy some 13% of DSMW yet less than 4% of profiles in the dataset. Profile representation in WISE3 is not based on an area-weighted basis, but mainly on the availability of sufficiently detailed legacy data.

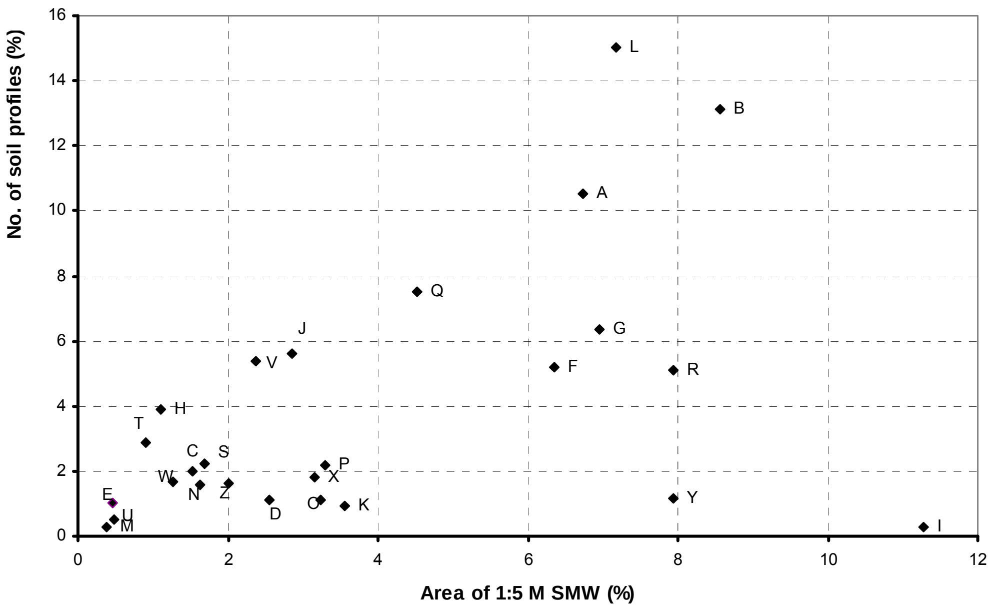


Figure 2. Representation of Major Soil Groups in WISE3 relative to their relative extent on the 1:5M Soil Map of the World[1]

[1] Relative area is expressed as percentage of total area of all major soil groups considered on the 1:5M scale, digital Soil Map of the World (FAO 1995); codes follow the original Legend (FAO-Unesco 1974): A, Acrisols; B, Cambisols; C, Chernozems; D, Podzoluvisols; E, Rendzinas; F, Ferralsols; G, Gleysols; H, Phaeozems; I, Lithosols; J, Fluvisols; K, Kastanozems; L, Luvisols; M, Greyzems; N, Nitosols; O, Histosols; P, Podzols; Q, Arenosols; R, Regosols; S, Solonetz; T, Andosols; U, Rankers; V, Vertisols; W, Planosols; X, Xerosols; Y, Yermosols; Z, Solonchaks.

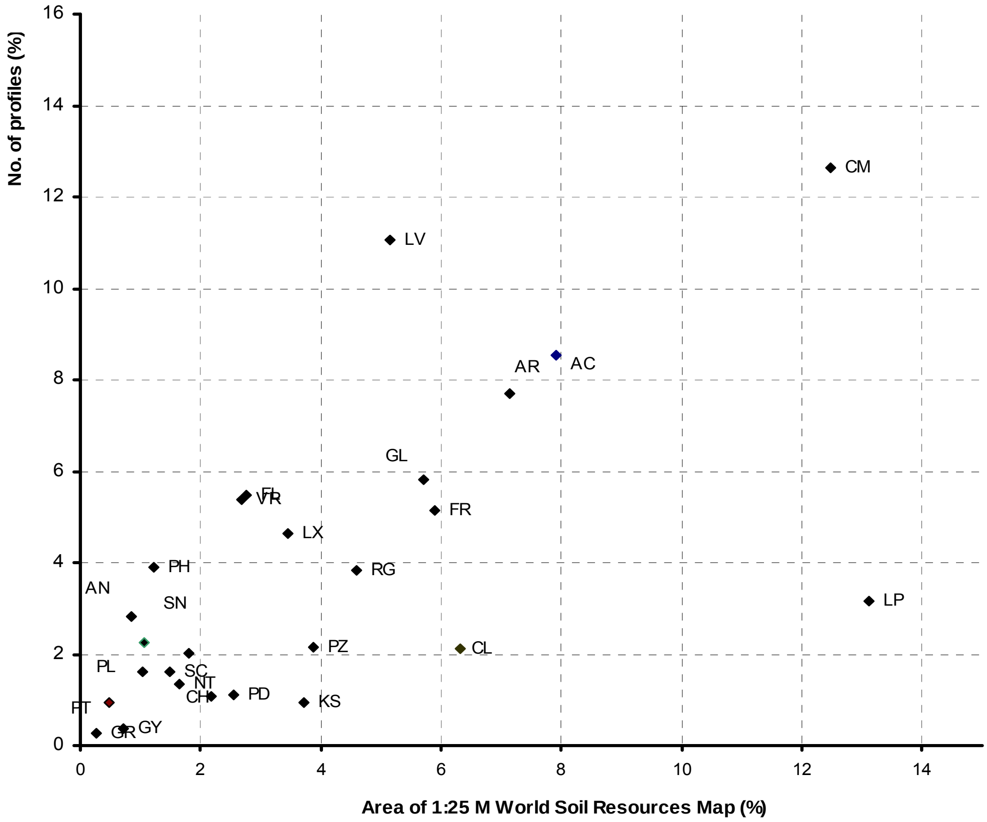


Figure 3. Representation of Major Soil Groups in WISE3 relative to their relative extent on the 1:25M World Soil Resources Map[2].

Several soil units remain under-represented in WISE3, arbitrarily defined here as less than 10 profiles for a given soil unit. With respect to the original Legend, these units are: gleyic Chernozems (Gc); gleyic Greyzems (Mg); ferric Podzols (Pf); humic Planosols (Wh); takyric Yermosols (Yt); and mollic and takyric Solonchaks (Zm and Zt; see Appendix 4). Conversely, with respect to the Revised Legend, under-represented soil units include: stagnic, plinthic and umbric Alisols (ALj, ALp and ALu); gleyic Andosols (ANg), cumulic, fimic and urbic Anthrosols (ATc, ATf and ATu);

[2] Relative area is expressed as percentage of total area of all major soil groups considered on the 1:25M scale World Soil Resources map, which excludes areas of Anthrosols and Alisols (FAO 1993). Codes follow the Revised Legend (FAO 1988): AC, Acrisols; AL, Alisols; AN, Andosols; AR, Areonsols; AT, Anthrosols; CH, Chernozems; CL, Calcisols; CM, Cambisols; FL, Fluvisols; FR, Ferralsols; GL, Gleysols; GR, Greyzems; GY, Gypsisols; HS, Histosols; KS, Kastanozems; LP, Leptosols; LV, Luvisols; LX, Lixisols; NT, Nitisols; PD, Podzoluvisols; PH, Phaeozems; PL, Planosols; PT, Plinthols; PZ, Podzols; RG, Regosols; SC, Solonchaks; SN, Solonetz; VR, Vertisols

gleyic and glossic Chernozems (CHg and CHw); andic and thionic Gleysols (GLa and GLt); gleyic and haplic Greyzems (GRg and GRh); luvic, haplic and petric Gypsisols (GYl, GYh and GYp); gelic, folic and thionic Histosols (HSi, HSl and HSt); gelic Leptosols (LPi); albic and stagnic Lixisols (LXa and LXj); stagnic Podzoluvisol units (PDj); umbric Planosols (PLu); umbric Plinthosols (PTu); ferric Podzols (PZf); gypsic Regosols (RGy); gelic and mollic Solonchaks (SCi and SCm); gypsic Solonetz (SNy); and gypsic Vertisols (VRy; see Appendix 5).

3.3 Temporal coverage

Legacy profile data collated in WISE3 have been sampled, described, and analysed over a long period of years. The year of sampling is known for 6685 profiles, ranging from 1925 to 2005: about 67% thereof originate from the period 1955 to 1995 (Figure 4).

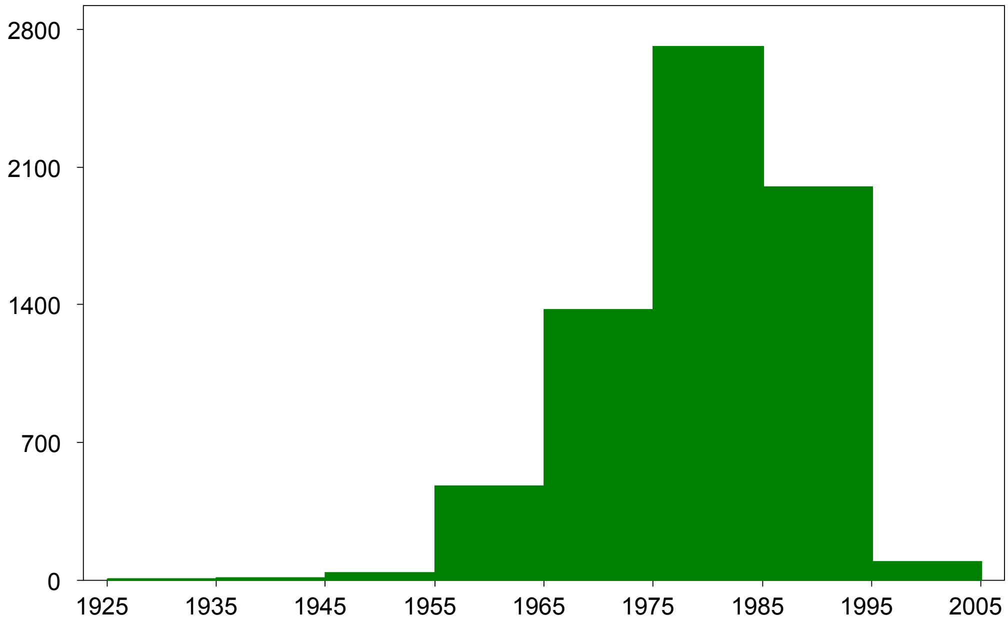


Figure 4. Frequency by sampling year of WISE3 holdings
(6685 cases plotted, 3568 cases no data)

3.4 Attribute data

The percentage of records (see Table 1) that are filled in the site and horizon tables are detailed in Appendix 8. Soil physical attributes, such as soil water retention and bulk density, in particular are under-represented in the database. Such data are seldom collected during routine soil surveys, which provided the bulk of the profile descriptions for WISE.

4 DISCUSSION

Comprehensive site data are often missing in the source materials, for example location (coordinates), landform, parent material, and land use. Similarly, complete sets of soil analytical data are seldom available for each sample/horizon. The number of observations (measured data) for each soil property will vary between soil units and with depth; typically, there are less measurements for the deeper horizons. Further, the available profiles have been analysed according to a range of analytical methods, thus necessitating a screening by analytical procedures. Generally, this process is cumbersome and fraught with uncertainty (see Batjes *et al.* 1997; FAO 1995; Hiederer *et al.* 2006; Jankauskas *et al.* 2006; Johnston *et al.* 2003; Landon 1991; Vogel 1994).

Initial printouts obtained from the NRCS, SDB, and ISIS data sets after transfer into WISE1 contained some distorted soil horizon designations and duplicate horizon depths. This was partly associated with the fact that soil horizon and sample depths were not always defined unambiguously in the source data files. Further, criteria for defining the zero datum for the same horizons have been changed in some systems, such as Soil Taxonomy (Schoenberger *et al.* 2002, p. 2-5) and WRB (FAO 2006a), around 1993. In so far as possible, such data issues have been remedied manually with reference to the original data sheets.

Contents of profile databases such as ISIS, SOTER, FAO-SDB, and NRCS-USDA may be updated as current knowledge increases. As a result, data for some profiles held in digital data files differed from those published elsewhere for the same profiles. This was the case, for example, for some NRCS profiles from Brazil, Korea and Zambia (Spaargaren and Batjes 1995), some SDB profiles from Botswana (FAO 1990), and several ISIS profiles. Overall, the various attribute values in WISE3 reflect the state of the source materials at the time they were incorporated in WISE.

When given in the source materials, the FAO classifications were generally taken at face value — having been assigned by pedologists most familiar with the source materials/regions —, unless (apparent) glaring inconsistencies were observed in the classifications, using various simple cross-checks. In this context, it should be stressed that WISE was never designed to be a classification tool. In addition, it should be noted that even experienced soil classifiers may have different perceptions about how a particular soil profile should be classified[3], in particular when the available soil morphological and soil analytical data are 'limited' (e.g., Goyens *et al.* 2007; Kauffman 1987; Spaargaren and Batjes 1995).

Differences in versions of USDA Soil Taxonomy used in the NRCS source files sometimes formed a difficulty when re-classifying profiles according to the FAO Legend (Spaargaren and Batjes 1995).

Profiles in WISE3 are meant to be taxo-referenced, i.e. considered representative for a given FAO-Unesco soil unit rather than a specific location, so that they may be used for the assessment of derived soil properties and the development of taxonomy-based pedotransfer functions. Many legacy profiles had no coordinates or only descriptive locations, having been described in the pre-GPS era. When possible, in case of missing latitude-longitude references, approximate coordinates have been initially derived from topographic maps or the Times Atlas (2003) and more recently using Google Earth®, using descriptive information on site location (e.g., Machakos, Kenya). As a result, some coordinates in WISE should be seen as being approximate for a profile's location (for details see Appendix 2); this aspect should be kept in mind when DMS coordinates are converted to decimal degrees (DD) for GIS-applications.

When available, information on parent material, climate as well as land use has been presented according to a range of classification systems in the source materials — typically, there is no uniform acceptance of a system for their classification (see FAO 2006a; Kottek *et al.* 2006; Noble *et al.* 2000; Times Atlas 2003). Inherently, this diversity in terminology can lead to some loss of information when the original descriptions or classifications are converted to their nearest equivalents according to the international classification systems adopted for WISE (see Appendix 2). By implication, effects of regional differences in, for example, microclimate, parent material, topography, moisture regime, and land use history on modal properties of similarly classified soils often cannot be studied explicitly using data held in WISE. Yet, the effects of such regional differences can

[3] For example, different estimates for average CEC of organic carbon may have been assumed to calculate CECclay where low clay activity may be a diagnostic property (cf. FAO-Unesco, 1988), the common range being from 150 to over 750 $cmol_c\ kg^{-1}$ (Klamt and Sombroek 1988; Schachtschabel *et al.* 1998). As a result of such differences in assumptions, for example, a given profile may key out as an Acrisol resp. an Alisol depending on the soil classifier. Conversely, some diagnostic criteria, such as the silt/clay ratio of < 0.2 for the ferralic B horizon, may have been applied with different rigour by various experts when the original profiles were classified.

be substantial, for example when estimating soil carbon stocks for specific soil units (e.g., Cerri *et al.* 2000).

Some measured data in the source materials seemed to deviate from what are considered to be "normal" ranges based on expert knowledge, for example bulk density values greater than 1.9 g cm^{-3} (e.g. Landon 1991). Yet, when such values were reported in reputable data sets or published sources they have been stored "as is" in WISE.

Different horizon designations are used in various national data sets; these have not been harmonized in WISE.

For some older databases, developed using dBaseIV®, it was not yet possible to differentiate between zero (0) values and Null fields. As a result, there may be some ambiguous zero values after dBaseIV® data are imported into MSAccess®. In addition, different conventions have been used to code "trace levels" for soil chemical variables in the various source data.

Data in WISE3 were subjected to a rigorous, computerized data-checking scheme as well as numerous visual checks (Appendix 3). Nonetheless, this large dataset data set is unlikely to be devoid of inconsistencies or even errors; users should always keep in mind such possible limitations when analysing or applying the data. Possible extreme values should be filtered out at the onset of any analyses of the primary data; this may be done, for instance, using a robust statistical outlier-rejection procedure in combination with invaluable expert judgement. Once this has been done, data in WISE can be used for a wide range of applications (see ISRIC-WISE 2008).

5 CONCLUSIONS

Data in WISE3 have been screened carefully. Nonetheless, the dataset is unlikely to be devoid of inconsistencies or even errors; these will be addressed in future updates.

WISE3 is particularly meant for user groups that wish to apply soil data, presented in a consistent format, in applications at a global and continental scale (< 1:500 000). Being a compilation of legacy soil profile data, it will inevitably include gaps; these can be of a taxonomic, geographic, and soil analytical nature. Further, a range of soil analytical methods has been used in the source materials to analyze the same soil properties. Possible methodological differences should be evaluated carefully when using the data; adroit use of the data will permit a wide range of agricultural and environmental applications (see ISRIC-WISE 2008), while keeping in mind the possible limitations of the primary data.

ACKNOWLEDGEMENTS

Special thanks are due to a wide range of land resources institutions and soil scientists who have contributed data for inclusion in WISE (see www.isric.org for details). The author thanks Otto Spaargaren for his editorial comments and useful discussions.

REFERENCES

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

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

Batjes NH 1999. Development of a 0.5° by 0.5° resolution global soil database. In: Sumner (editor), *Handbook of Soil Science*. CRC Press, Boca Ratton, pp Chap. H, p. 29-40

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 PDF

Batjes NH 2002b. *A homogenized soil profile data set for global and regional environmental research (WISE version 1.1)*. ISRIC Report 2002/01, International Soil Reference and Information Centre (ISRIC), Wageningen PDF

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

Batjes NH 2006. *ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1.0)*. Report 2006/02, ISRIC - World Soil Information, Wageningen PDF

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, 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, Ingram, Neue and 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 PDF

Batjes NH, Al-Adamat R, Bhattacharyya T, Bernoux M, Cerri CEP, Gicheru P, Kamoni P, Milne E, Pal DK and Rawajfih Z 2007. Preparation of consistent soil data sets for SOC modelling purposes: secondary SOTER data sets for four case study areas. *Agriculture, Ecosystems and Environment* 122, 26-34

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

Bernoux M, Arrouays D, Cerri C, Volkoff B and Jolivet C 1998. Bulk densities of Brazilian Amazon soils related to other soil properties. *Soil Science Society of America Journal* 62, 743-749

Breuning-Madsen BH and Jones RJA 1998. Towards a European soil profile analytical database. In: Heineke, Eckelmann, Thomasson, Jones, Montanarella and Buckley (editors), *Land Information Systems: Developments for planning the sustainable use of land resources*. Office for Official Publications of the European Community, Luxembourg, pp 43-49

Cerri C, Bernoux M, Arrouays D, Feigl BJ and Piccolo MC 2000. Carbon stocks in soils of the Brazilian Amazon. In: Lal, Kimble and Stewart (editors), *Global climate change and tropical ecosystems*. CRC Press, Boca Raton, pp 33-50

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 1988. *FAO-Unesco Soil Map of the World, Revised Legend, with corrections and updates*. World Soil Resources Report 60, FAO, Rome; reprinted with updates as Technical Paper 20 by ISRIC, Wageningen,1997

FAO 1989. *FAO-ISRIC Soil Database (SDB)*. World Soil Resources Report 60 (Reprinted), Food and Agriculture Organization of the United Nations, Rome

FAO 1990. *Explanatory note on the Soil Map of the Republic of Botswana*. Report AG-BOT/85/011, FAO, UNDP and Republic of Botswana, Gaborone

FAO 1993. *World Soil Resources: An explanatory note on the FAO World Soil Resources map at 1:25,000,000 scale*. World Soil Resources Report 66 (rev. 1), Food and Agriculture Organization of the United Nations, Rome

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

FAO 2006a. *Guidelines for soil description (Fourth ed.)*, FAO, Rome

FAO 2006b. *World Reference Base for soil resources - A framework for international classification, correlation and communication*. World Soil Resources Reports 103, International Union of Soil Sciences, ISRIC - World Soil Information and Food and Agriculture Organization of the United Nations, Rome

FAO and ISRIC 2000. *Soil and terrain database, soil degradation status, and soil vulnerability assessments for Central and Eastern Europe (scale 1:2.5 million; ver. 1.0)*. Land and Water Digital Media Series 10, FAO, Rome

FAO and ISRIC 2003. *Soil and Terrain database for Southern Africa (1:2 million scale)*. FAO Land and Water Digital Media Series 25, ISRIC and FAO, Rome

FAO, ISRIC, UNEP and CIP 1998. *Soil and terrain digital database for Latin America and the Caribbean at 1:5 million scale*. Land and Water Digital Media Series No. 5, Food and Agriculture Organization of the United Nations, Rome

FAO/IIASA/ISRIC/ISSCAS/JRC 2008. *Harmonized World Soil Database (version 1.0)*, Prepared by Nachtergaele FO, van Velthuizen H, Verelst

L, Batjes NH, Djjkshoorn JA, van Engelen VWP, Fischer G, Jones A, Montanarella L., Petri M, Prieler S, Teixeira E, Wiberg D and Xuezheng Shi. Food and Agriculture Organization of the United Nations (FAO), 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 PDF

FAO/ISRIC 1990. *Guidelines for soil description (3rd rev. ed.)*, FAO, Rome

Finke P 2006. Quality assessment of digital soil maps: producers and users perspectives. In: Lagacherie, McBratney and Voltz (editors), *Digital soil mapping: An introductory perspective*. Elsevier, Amsterdam, pp 523-541

Goyens C, Verdoodt A, Van de Wauw J, Baert G, van Engelen VWP, Dijkshoorn JA and Van Ranst E 2007. Base de données numériques sur les sols et le terrain (SOTER) de l'Afrique Centrale (RD Congo, Rwanda et Burundi). *Etude et Gestion des Sols* 14, 207-218

Heuvelink GBM and Brown JD 2006. Towards a soil information system for uncertain soil data In: Lagacherie, McBratney and Voltz (editors), *Digital soil mapping: An introductory perspective*. Elsevier, Amsterdam, pp 97-106

Hiederer R, Jones RJA and Darroussin J 2006. Soil Profile Analytical Database for Europe: Reconstruction and Validation of the data. *Danish Journal of Geography* 106, 71-85

Hollis JM, Jones RJA, Marshall CJ, Holden A, van de Veen JR and Montanarella L 2006. *SPADE-2: The Soil Profile Analytical Database for Europe (version 1.0)*. European Soil Bureau Research Report No.19 (EUR 22127 EN), Office for Official Publications of the European Communities, Luxembourg PDF

ISRIC-WISE 2008. *References and examples of WISE database applications. PDF* ISRIC 2007. *ISRIC - World Soil Information Database* PDF

Jankauskas B, Jankauskiene G, Slepetiene A, Fullen MA and Booth CA 2006. International comparison of analytical methods of determining the soil organic matter content of Lithuanian eutric Albeluvisols. *Communications in Soil Science and Plant Analysis* 37, 707-720

Johnston RM, Barry SJ, Bleys E, Bui EN, Moran CJ, Simon DAP, Carlile P, McKenzie NJ, Henderson BL, Chapman G, Imhoff M, Maschmedt D, Howe DÂ, Grose C, Schoknecht N, Powell B and Grundy M 2003. ASRIS: the database. *Australian Journal of Soil Research* 41, 1021-1036

Kauffman JH 1987. *Comparative classification of some deep, well-drained red clay soils of Mozambique*. Technical Paper 16, International Soil Reference and Information Centre (ISRIC), Wageningen PDF

Klamt E and Sombroek WG 1988. Contribution of organic matter to exchange properties of Oxisols. In: Beinroth, Camargo and 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

Kottek M, Grieser J, Beck C, Rudolf B and Rubel F 2006. World Map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* 15, 259-263 (DOI: 10.1127/0941-2948/2006/0130)

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

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

Minasny B and McBratney AB 2002. Uncertainity analysis for pedotransfer functions. *European Journal of Soil Science* 53, 417-429

Munsell 1975. *Munsell Soil Colour Charts*, Baltimore

Nettleton WD, Hornsby AG, Brown RB and Coleman TL (editors) 1996. Data reliability and risk assessment in soil interpretations, SSSA Special Publication No. 47. Soil Science Society of America, Madison WI

Noble IR, Apps M, Houghton R, Lashof D, Makunidi W, Muridyarso D, Murray B, Sombroek M and Valentini M 2000. Implications of different definitions and generic issues. In: Watson, Noble, Bolin, Ravindranath, Verardo and Dokken (editors), *Land use, Land-Use Change, and Forestry*. Published for the Intergovernmental Panel on Climate Change by Cambridge University Press, Cambridge, pp 55-126

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

Schachtschabel P, Blume H-P, Brummer G, Hartge KH and Schwertmann U 1998. *Lehrbuch der Bodenkunde (Scheffer/Schachtschabel, 14th Ed.)*. Ferdinand Enke Verlag, Stuttgart, 494 p

Schoenberger PJ, Wysicki DA, Benham EC and Broderson WD 2002. *Field book for describing and sampling soils (ver. 2.0)*. National Soil Survey Center, Natural Resources Conservation Service, USDA, Lincoln NE PDF

Skjemstad JO, Spouncer LR and Beech A 2000. *Carbon conversion factors for historical soil carbon data*, Australian Greenhouse Office, Canberra PDF

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

Spaargaren OC and Batjes NH 1995. *Report on the classification into FAO-Unesco soil units of profiles selected from the NRCS pedon database for IGBP-DIS*. Work. Pap. 95/01, ISRIC, Wageningen

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

Times Atlas 2003. *The Comprehensive Times Atlas of the World*. Times Books, Harper Collins Publ., London

UN Statistics Division 2008. Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings PDF

USDA-NRCS 2008. *NSSC Soil Survey Laboratory Soil Characterization Database*, United States Department of Agriculture, Natural Resources Conservation Service, Lincoln PDF

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 KDH and Huting JRM 2005. *Harmonized Global Soil Resources Database (Final Report)*. Report 2005/06, FAO and ISRIC - World Soil Information, Wageningen PDF

van Reeuwijk LP 1983. *On the way to improve international soil classification and correlation: the variability of soil analytical data*. Annual Report 1983, ISRIC, Wageningen PDF

van Reeuwijk LP 1995. *Procedures for soil analysis (5th ed.)*. Technical Paper 9, ISRIC, Wageningen

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

van Reeuwijk LP and Houba VJG 1998. External quality control of data. In: Van Reeuwijk (editor), *Guidelines for quality management in soil and plant laboratories*. ISRIC and FAO, Rome, pp 121-127

Vogel AW 1994. *Comparability of soil analytical data: determinations of cation exchange capacity, organic carbon, soil reaction, bulk density, and volume percentage of water at selected pF values by different methods*. Work. Pap. 94/07, ISRIC, Wageningen PDF

Wilson AL 1979. Approach for achieving comparable analytical results from a number of laboratories. *The Analyst* 104, 1273-1288

APPENDICES

Appendix 1. Structure of WISE3 attribute tables

Table WISE3_Site

Field Name	Description
WISE3_ID	Unique profile reference number
DESCR	Profile description status, code
DATEYR	Year profile was first described
DATEMON	Month profile was first described
SOLDEP	Reference soil depth (maximum as sampled, in cm)
HORNUM	Number of horizons described for pit
FAO_74	FAO-Unesco (1974), soil classification as code[a]
PHA_74	As above, but code for (main) phase
FAO_90	FAO-Unesco (1988), soil classification as code
PHA_90	As above, but code for (main) phase
WRB2006	WRB Reference Soil Group (ed. 2006)
USCL	USDA Soil Taxonomy classification, descriptive
USYR	Year (version of Soil Taxonomy, e.g., 75, 94)
LOCAL	Local soil classification, descriptive
ISO	ISO code for country of origin
LATIT	Latitude of profile (N/S)
LATDEG	Latitude, degrees
LATMIN	Latitude, minutes
LATSEC	Latitude , seconds
LONGI	Longitude of profile (E/W)
LONDEG	Longitude, degrees
LONMIN	Longitude, minutes
LONSEC	Longitude, seconds
LOCAT	Profile location, descriptive
KOPPEN	Climate, Köppen system
ALTIT	Elevation (m)
LFORM	Landform, code
POSIT	Position, code
SLOPE	Slope at profile site (%)
PARMAT	Parent material, code
DRAIN	Drainage class, code
LANDUS	Land use, code
SOURCE_ID	Unique reference number for source of profile data
PITREF	Profile/page reference in source
LAB_ID	Unique reference number for laboratory

[a] Coding conventions are given in Appendix 2.

Table WISE3_Horizon

Field Name	Description
WISE3_ID	Unique soil profile number
HORIZ	Unique horizon number (in combination with WISE3_ID)
DESIG	Horizon designation, coded according to national system
TOPDEP	Upper depth of horizon (cm)
BOTDEP	Lower depth of horizon (cm)
MCOLOR	Moist matrix colour, Munsell code
DCOLOR	Dry matrix colour, Munsell code
ORGC	Organic carbon (g kg^{-1}) [a]
TOTN	Total Nitrogen (g kg^{-1})
CACO3	Calcium carbonate content (g kg^{-1})
GYPSUM	Gypsum content (g kg^{-1})
PHH2O	pH, measured in water
PHKCL	pH, measured in KCl solution
PHCACL2	pH, measured in $CaCl_2$ solution
ECE	Electrical conductivity (dS m^{-1} or mmho cm^{-1})
EXCA	Exchangeable calcium ($cmol_c$ kg^{-1})
EXMG	Exchangeable magnesium ($cmol_c$ kg^{-1})
EXNA	Exchangeable sodium ($cmol_c$ kg^{-1})
EXK	Exchangeable potassium ($cmol_c$ kg^{-1})
EXACID	Exchangeable acidity ($cmol_c$ kg^{-1})
EXALUM	Exchangeable aluminium ($cmol_c$ kg^{-1})
CECSOIL	Cation exchange capacity ($cmol_c$ kg^{-1})
BSAT	Base saturation, expressed as % of CEC
SAND	Sand content (w/w%)
SILT	Silt content (w/w%)
CLAY	Clay content (w/w%)
GRAVEL	Gravel content (v/v % of fraction > 2 mm)
BULKDENS	Bulk density (g cm^{-3})
VMC1	Volumetric water content at -10 kPa or pF2.0 (v/v% or cm^3 cm^{-3} 10^2)
VMC2	As above, but at -33 kPa or pF2.5 (v/v%)
VMC3	As above, but at -1500 kPa or pF4.2 (v/v%)

[a] Analytical procedures are summarized in table *WISE3_labname*, *WISE3_LABname_codes*, and *WISE3_LABcodes_Description.*

Table WISE3_Source

Field Name	Description
SOURCE_ID	Unique reference number for source of profile data
AUTHOR	Author name and initials
AUTYR	Year of publication
REFTIT	Title of monograph/database, descriptive
REFPUB	Series/publisher/year, descriptive
HasURL	Flag for availability of URL (Y or N)
URL	Link to on-line source (i.e., full text report or dataset)

Table WISE3_Labname

Field Name	Description
LAB_ID	Unique laboratory code
LABNAM	Reference to laboratory, descriptive

Table WISE3_LABname_Codes

Field Name	Description
LAB_ID	Unique laboratory code
ORGC	Organic Carbon (OC__)[a]
TOTN	Total Nitrogen (TN__)
PHH2O	pH in water (PH__)
PHKCL	pH in KCl (PK__)
PHCACL2	pH in $CaCl_2$ (PC__)
ELECON	Electrical conductivity (EL__)
CACO3	Free CaCO3 (CA__)
Gypsum	Gypsum (GY__)
EXBAS	Exchangeable Ca, Mg, K, and Na (EX__)
EXACID	Exchangeable acidity and aluminum (EA__)
CECSOIL	CEC soil (CS__)
BSAT	Base saturation (BS__)
TEXTURE	Particle size distribution (TE__)
BULKDENS	Bulk density (BD__)
MOISTCON	Water content (MC__)

[a] Codes for analytical methods have the general form “XX01”, “XX02”, etc. where X is the variable under consideration; methods are summarized in table *WISE3_LABcodes_Description*.

Table WISE3_LABcodes_Description:

Field Name	Description
KEY	Unique code for laboratory method[a]
LABMETHOD	Brief description of analytical method

[a] Each code for a method consists of two characters and two numbers (e.g. OC01, TN07). The characters refer to the overall type of analyses, such as the determination of organic carbon (OC), while the sequential numbers refer to the specific type of analysis used; see table *WISE3_LABname_Codes* for details.

Table WISE3_coding_conventions:

Field Name	Description
Attribute	Field name (e.g., ISO for country codes)
KEY	Value of given attribute, classified
DESCRIPTION	Description of KEY (e.g., for attribute 'ISO', KEY 'CD' stands for the 'Democratic Republic of Congo')
REGION[a]	Geographical region (e.g., Asia – Southern; used for sorting)
REGIONum[a]	Geographical region according to UN Statistics Division(2008), for example 'Southern Asia'
CONTINENT[a]	Name of continent (only used with ISO codes)

[a] Only used for ISO codes: KEY='AT' —> DESCRIPTION= 'Austria' —> REGION= 'Europe-Western' —> REGIONum= 'Western Europe' —> CONTINENT= 'EU' (Europe)

Appendix 2. WISE3 data coding protocols

Coding protocols for site and soil horizon data in WISE3 are adapted from the "WISE User Manual, ver. 1.0" (Batjes 1995); the complement of attribute data represented in WISE3 is smaller than the one used for version 1.0.

A) *Site data* (Table WISE3_Site)

WISE3_ID

Code for unique reference number for the soil profile in question. It consists of the country's ISO-3166 two-letter code and a number (Example: BR0022).

DESCR

Code, for "soil profile description status," that refers to the apparent completeness and quality of the soil profile descriptions and accompanying analytical data for the specified profile in the original source. The description status is determined after screening of the original information. It may be seen as an indicator of the (likely) reliability of the data as well as the extent to which these data may be used to ascertain the FAO soil unit classification. Four classes are considered in accordance with FAO/ISRIC (1990) conventions:

Status	Description
1	*Reference Pedon Description.* No essential elements or details are missing from the description, sampling or analysis. The accuracy and reliability of the description and analytical results permits the full characterization in the FAO-Unesco (1974) Legend respectively Revised Legend (1990).
2	*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 respectively Revised Legend.
3	*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 is however useful for specific purposes and provides a satisfactory indication of the nature of the soil in the FAO-Unesco Legend respectively Revised Legend.
4	*Other description* in which essential elements are lacking from the description, preventing a satisfactory soil characterization and classification[a].

[a] Code 4 type data are generally not accepted for inclusion in WISE, unless the corresponding soil unit is grossly under-represented in the global profile data set;

such profiles may be purged from subsequent releases of WISE once new status 1 or 2 profiles have been found for the relevant soil units.

DATE

Date on which the profile was described/sampled given as year (DATEyr) and month (DATEmon).

SOLDEP

The reference soil depth of a profile, defined here as the maximum depth to which the profile was described/sampled in the field or documented in the source database, in cm.

HORNUM

Total number of horizons for which analytical data are available for the given profile (i.e. covers the depth range given under SOLDEP).

FAO74

Code for the soil classification according the 1 or 2 letter codes used in the Key to Soil Units (FAO-Unesco 1974, p. 43-53). For example, E stands for a Rendzina and Ge for an Eutric Gleysol. Ultimately, this code forms the key for linking the soil profile data, and any secondary data derived from them, to the mapping units of the Soil Map of the World (FAO 1995).

PHA_74

Code for the main phase, according to FAO-Unesco (1974, p. 5-7):

Code	Description	Code	Description
ST	stony	X	fragipan
PE	petric	MQ	duripan
MK	petrocalcic	Z	saline
LI	lithic	SO	sodic
MY	petrogypsic	CE	cerrado
PH	phreatic	MS	petroferric

FAO_90

FAO-Unesco classification in the Revised Legend, encoded using the 3-letter codes of the Key to Major Soil Groupings and Soil Units (FAO, 1988 p. 74-88). HSf, for example, stands for a Fibric Histosol and ACp for a Plinthic Acrisol. FAO_90 forms the main key for linking soil parameter estimates derived from WISE when filling gaps soil profile data in primary SOTER database (e.g., Batjes *et al.* 2007; van Engelen *et al.* 2005).

PHA_90
Code for the main phase, described according to the Revised Legend (FAO 1988):

Code	Description	Code	Description
AN	anthraquic	PF	petroferric
DU	duripan	PH	phreatic
FR	fragipan	PL	placic
GE	gelundic	SO	sodic
GI	gilgai	RU	rudic
IN	inundic	SA	salic
SK	skeletic	TK	takyric
YR	yermic	LI	Lithic

WRB2006
Code for the Reference Soil Group according to the World Reference Base for Soil Resources (FAO 2006b); see table *WISE3_coding_conventions* for details.

USCL
Soil classification according to USDA Soil Taxonomy, as a text string with a maximum length of 50 characters (abbreviated where necessary); mostly taken as is from source materials.

USYR
Version/year of USDA Soil Taxonomy used, expressed as four characters (e.g., 1975, 1987, 1990, and 1998).

LOCAL
The soil classification according to the National System, with a maximum of 50 characters (abbreviated where necessary).

ISO
Code for the country of origin of the profile, specified according to ISO-3166 (see table *WISE3_coding_conventions* for details).

LOCAT
Description of general location of profile (e.g., town, province), as a text string of maximum 50 characters.

LATIT, LATDEG, LATMIN, LATSEC and
LONGI, LONDEG, LONMIN, LONSEC
The combination of these fields gives the full coordinates of a soil profile. Coordinates are given as degrees, minutes and seconds latitude (LATIT; N or S) and longitude (LONGI; E or W). Unless given in the source description, coordinates were derived from appropriately detailed topographical maps (using descriptive information on location, such as proximity to a large town) and accurate to at least 25 km in view of their original application with a ½ by ½ degree spatial database (Batjes 1999). [*Note:* (a) If only the fields LATIT, LATDEG, LATMIN, LONGI, LONDEG, and LONMIN are filled, this shows profile coordinates are approximate; conversion of such coordinates (WGS_1984) to DMS may suggest a false accuracy of the location; (b) Coordinates (DD) for profiles derived from SOTER were converted to LAT-LON, degrees-minutes and seconds (DMS)]

KOPPEN
Code for the climate at the site classified according to the Köppen system which considers precipitation effectiveness for plant growth as the major classification factor, and uses the appropriate seasonal values of temperature and precipitation to determine the limits of climatic groupings.

The Köppen system figures a shorthand code of letters designating major climate groups, subgroups within these major groups, with further subdivisions to distinguish particular seasonal characteristics of temperature and precipitation (adapted from Strahler, 1969 p. 224; Times Atlas, 1993).

a) Major climate groups
The following six major climate groups are considered:

Code	Classification and description
A	Tropical (rainy) climates: Average temperature of every month is above 18 °C. These climates have no winter season. Annual rainfall is large and exceeds annual evaporation
B	Dry: Potential evaporation exceeds precipitation on the average throughout the year. No water surplus; hence no permanent streams originate in B climate zones
C	Warm temperate (mesothermal) climates: Coldest month has an average temperature under 18 °C, but above -3 °C. The C climates thus have both a summer and a winter season
D	Snow (microthermal) climates: Coldest month average temperature under -3 °C. Average temperature of the warmest month above 10 °C, that isotherm corresponding approximately with pole-ward limit of forest growth
E	Ice climates: A polar climate type with average temperature in no month averaging over 10 °C. These climates have no true summer
H	Mountain/Highland climates

b) Subgroups

Subgroups within the major climate groups are designated by a second letter according to the following code:

Code	Description
S*	Steppe climate, a semiarid climate with about 380 to 760 mm of rainfall annually at low latitudes.
W	Desert climate. Arid climate. Most regions included have less than 250 mm of rainfall annually.
f	Moist. Adequate precipitation in all months. No dry season. This modifier is applied to major climate types A, C and D.
w	Dry season in winter of the respective hemisphere (low-sun season)
s	Dry season in summer of the respective hemisphere (high-sun season)
m	Rainforest climate despite a short dry season in monsoon type of precipitation cycle. Applies only to A climates.

* The letters S and W are applied only to the dry climates (i.e., BS and BW).

From combinations of the two letter groups, 12 distinct climates emerge as follows:

Code	Description
Af	Tropical rainforest (also Am a variant of Af)
Aw	Tropical savanna
BS	Steppe climate
BW	Desert climate
Cw	Temperate rainy (humid mesothermal) climate with dry winter
Cf	Temperate rainy (humid mesothermal) climate moist all seasons
Cs	Temperate rainy (humid mesothermal) climate with dry summer
Df	Cold snowy forest (humid microthermal) climate moist in all seasons
Dw	Cold snowy forest (humid microthermal) climate with dry winter
ET	Tundra climate
EF	Climates of perpetual frost (ice-caps)
H	Mountain/Highland climates (undifferentiated)

c) A third letter may be added to differentiate still more variations:

Code	Description
a	With hot summer; warmest month over 22 °C (C and D climates)
b	With warm summer; warmest month below 22 °C (C and D climates)
c	With cool, short summer; fewer than four months over 10 °C (C and D climates)
d	With very cold winter; coldest months below - 38 °C (D climates only)
h	Dry-hot; mean annual temperature over 18 °C (B climates only)
k	Dry-cold; climates annual temperature under 18 °C (B climates only)

ALTIT

Altitude of the soil profile relative to mean sea level, specified in meters. Unless specified on the original soil description, this information can be derived from a suitably detailed topographical map.

LFORM

Code for the major landform. In accordance with SOTER criteria (van Engelen and Wen 1995), landforms are described principally by the morphology and not by their genetic origin or processes responsible for their shape. The first differentiating criterion is the dominant slope, followed by relief intensity:

Code	Landform Description
L	Level land: Land with characteristic slopes of 0-8 %, and a relief intensity of less than 100 m per km
S	Sloping land: Land with characteristic slopes of 8-30 % and a relief intensity of more than 50 m per slope unit. Areas with limited relief intensity (< 50 m per slope unit) but slopes in excess of 8% are included, as are isolated mountains (relief intensity > 600 m) with slopes of 8-30 %
T	Steep land: Land with characteristic slopes of over 30 % and a relief intensity of mostly more than 600 m per 2 km
C	Land with composite landforms: Land made up of steep elements together with sloping or level land, or sloping land landforms with level land, in which at least 20 % of the area consists of land with the lesser slope

SOTER codes for second level major landforms should be used in WISE:

First level	Second level	Gradient	Relief intensity
L Level land	LP plain	0-8%	< 100 m/km
	LL plateau	0-8%	< 100 m/km
	LD depression	0-8%	< 100 m/km
	LF low-gradient foot slope	0-8%	< 100 m/km
	LV valley floor	0-8%	< 100 m/km
S Sloping land	SM medium-gradient	15-30%	> 600 m/2km
	mountain	8-30%	> 50 m/s.u.
	SH medium-gradient hills	15-30%	< 600 m/2km
	SE medium-gradient	8-30%	> 50 m/s.u.
	escarpment zone	8-30%	> 600 m/2km
	SR ridges	8-30%	> 50 m/s.u.
	SU mountainous highland		
	SP dissected plain		
T Steep land	TM high-gradient mountain	> 30%	> 600 m/2km
	TH high-gradient hill	> 30%	< 600 m/2km
	TE high-gradient escarpment	> 30%	> 600 m/2km
	zone	> 30%	Variable
	TV high gradient valleys		
C Land with	CV valley	> 8%	Variable
composite	CL narrow plateau	> 8%	variable
landforms	CD major depression	> 8%	variable

POSIT

Code for the physiographic position of the site of the profile, described according to FAO/ISRIC (1990, p. 7).

Position	Code	Description
a) Position in undulating to mountainous terrain		
	CR	Crest/top
	UP	Upper slope
	MS	Middle slope
	LS	Lower slope
	BO	Bottom (flat)
b) Position in flat or almost flat terrain		
	HI	Higher part
	IN	Intermediate part
	LO	Lower part
	BO	Bottom (drainage line)

SLOPE

Slope refers to the inclination of the land immediately surrounding the site. The measured or estimated slope angle is specified to the nearest per cent.

PARMAT

Code for the main parent material from which the soil has been formed, coded using SOTER and FAO/ISRIC (1990, p. 14) conventions.

Major class	Group	Type	
I Igneous rocks	**IA** acid igneous	**IA1**	granite
		IA2	grano-diorite
		IA3	quartz-diorite
		IA4	rhyolite
	II intermediate igneous	**II1**	andesite, trachyte, phonolite
		II2	diorite-syenite
	IB basic igneous	**IB1**	gabbro
		IB2	basalt
		IB3	dolerite
	IU ultrabasic igneous	**IU1**	peritonise
		IU2	pyroxenes
		IU3	limonite, magnetite, ironstone, serpentine
M Metamorphic Rocks	**MA** acid metamorphic	**MA1**	quartzite
		MA2	gneiss, magmata
		MA3*	slate, hyalite
		MA4*	schist's

	MB basic metamorphic	**MB1**	slate, hyalite (pelitic rocks)
		MB2	schist
		MB3	gneiss rich in ferro-magnesian min.
		MB4	metamorphic limestone (marble)
S Sedimentary rocks	**SC** clastic sediments	**SC1**	conglomerate, breccia
		SC2	sandstone, greywacke, arkose
		SC3	siltstone, mudstone, clay stone
		SC4	shale
	SO organic	**SO1**	limestone, other carbonaceous rocks
		SO2	marl and other mixtures
		SO3	coals, bitumen and related rocks
	SE evaporates	**SE1**	anhydrite, gypsum
		SE2	halite
U Unconsolidated	**UF** fluvial		
	UL lacustrine		
	UM marine		
	UC colluvial		
	UE eolian		
	UG glacial		
	UP pyroclastic		
	UO organic		
	UX soft laterite and ferruginous materials		
	UY hardened laterite and ferruginous materials		

DRAIN

Code for internal drainage class, coded according to FAO/ISRIC (1990, p. 20).

Code	Description
V	very poorly drained
P	poorly drained
I	somewhat poorly (imperfectly) drained
M	moderately well drained
W	well drained
S	somewhat excessively drained
E	excessively drained

LANDUS

Code for current land use at the site, based on classes of FAO/ISRIC (1990, p. 13), in accordance with SOTER conventions.

S Settlement Industry
- SR - Residential use
- SI - Industrial use
- ST - Transport
- SC - Recreational use
- SX – Excavations

A Crop Agriculture
- AA - Annual field cropping
  - AA1: Shifting cultivation
  - AA2: Fallow system cultures
  - AA3: Ley system cultures
  - AA4: Rainfed arable cultures
  - AA5: Wet rice cultivation
  - AA6: Irrigated cultivation
- AP - Perennial field cropping
  - AP1: Non-irrigated cultures
  - AP2: Irrigated cultures
- AT - Tree and shrub cropping
  - AT1: Non-irrigated tree crop Cultures
  - AT2: Irrigated tree crop cultures
  - AT3: Non-irrig. shrub crop cultivation
  - AT4: Irrigated shrub crop cultivation

H Animal Husbandry
- HE - Extensive grazing
  - HE1: Nomadism
  - HE2: Semi-nomadism
  - HE3: Ranching
- HI: Intensive grazing
  - HI1: Animal Production
  - HI2: Dairying

F Forestry
- FN: Natural forest and woodland
  - FN1: Selective felling
  - FN2: Clear felling
- FP: Plantation forestry

M Mixed Farming
- MF - Agro-forestry
- MP- Agro-pastoralism (cropping and livestock systems)

E Extraction and Collection
- EV - Exploitation of natural vegetation
- EH - Hunting and fishing

P Nature Protection
- PN – Nature & game reserve
  - PN1: Reserves
  - PN2: Parks
  - PN3: Wildlife management
- PD - Degradation control
  - PD1: Without interference
  - PD2: With interference

U Not Used and Not Managed

SOURCE_ID

The unique SOURCE_ID provides an alphanumeric code to the source from which the soil profile data were derived, for example a soil monograph or digital database (e.g., 'JM 1982.04' for a survey monograph held in the ISRIC — World Soil Information Database).

PITREF

The page of the profile description or number of the profile as given in the corresponding source material (SOURCE_ID).

LAB_ID

Unique alphanumeric code to the laboratory where the measurements have been made. Each LAB_ID consists of the country's ISO-code, followed by two numerals (Example: IN02). Summary information on the analytical procedures that have been used to measure a certain attribute is presented in table *WISE3_labcodes_Description*.

B) Horizon data (Table WISE3_Horizon)

WISE3_ID

Unique reference code for the soil profile in question. It consists of the country's ISO-3166 two-letter code followed by a number (Example: BR0022).

HONU

Sequential horizon number, numbered from the top down.

DESIG

Code for the horizon; whenever possible, the horizon designation should be given according to the terminology of the *Guidelines for Soil Description* (FAO 2006a); otherwise, conventions of the national system are maintained.

TOPDEP

Upper depth of horizon (cm); for conventions see FAO (2006a) and Schoenberger et al. (2002). If the original depth of a 'wavy' horizon was given as e.g. 30/40 cm, the harmonized horizon depth must be entered as (30+40)/2= 35 cm.

BOTDEP

Lower depth of horizon in cm. [*Note*: If no lower profile depth is given in the source materials, but analytical data are available for the deepest horizon, it is assumed that this horizon is 15 cm thick. For example, 75^{+} cm would imply an inferred lower depth of 90 cm]

ORGC
Organic carbon content in g kg^{-1}, specified with 2 decimal places.
[*Note*: Codes for measurement methods are documented in table *WISE3_LABcodes_Description*]

TOTN
Total nitrogen (g kg^{-1}), rounded to 2 decimal places.

PHH2O
Soil reaction measured in water at the specified soil:water ratio; rounded to 1 decimal place.

PHKCL
Soil reaction measured in a KCl solution at the specified soil:solution ratio and molarity, rounded to 1 decimal place.

PHCACL2
Soil reaction measured in $CaCl_2$ solution at the specified soil:solution ratio and molarity; rounded to 1 decimal place.

ECE
Electrical conductivity (EC_x) for the horizon, for the specified soil:water ratio (x), in mS cm^{-1} or dS m^{-1}, originally mmho cm^{-1}, at 25 ^{o}C. Rounded to 2 decimal places.

CACO3
Total $CaCO_3$ content (% by weight), rounded off to the nearest integer.

GYPSUM
The content of total gypsum ($CaSO_4.2H_2O$), by weight %, rounded off to the nearest integer.

EXCA, EXMG, EXK and EXNA
Exchangeable bases (i.e., Ca^{2+}, Mg^{2+}, K^+ and Na^+) in $cmol_c$ kg^{-1}.

EXACID
Exchangeable acidity, the sum of exchangeable Al^{3+} and H^{+}, obtained with a percolation of a soil sample with a 1 *M* KCl solution. Exchangeable acidity is measured by titration of the percolate, and exchangeable aluminium is determined separately in the percolate (van Reeuwijk 2002). Exchangeable acidity is specified in $cmol_c\ kg^{-1}$.
[Note: Values for exchangeable acidity, determined in 1 *M* KCl percolate, and extractable acidity, equilibrated with a $BaCl_2$-TEA buffer at pH 8.2, refer to essentially different measurement methods].

EXALUM
Exchangeable aluminium (Al^{3+}), in $cmol_c\ kg^{-1}$, as determined separately in the percolate described above.

CECSOIL
The cation exchange capacity (CEC) of the fine earth fraction, in $cmol_c\ kg^{-1}$, determined according to analytical methods specified in table *WISE3_LABcodes_Description*.

BSAT
Base saturation (BS) calculated as sum of exchangeable cation bases (Ca^{2+}, Mg^{2+}, K^{+} and Na^{+}) divided by CEC_{soil}, measured with the specified analytical method, times 100%.
[Note: BSAT is a derived value – it is rounded to 100% when calculated values exceeded 100%, for example for calcareous and gypsiferous soils].

DCOLOR
The dry matrix colour specified using the Munsell Colour Charts (Munsell 1975). Colour codes have the general form: hue, value, chroma (e.g., 5YR6/4).

MCOLOR
The moist colour according to the Munsell Colour Charts (e.g., 5YR3/2).

SAND, SILT and CLAY
The particle size distribution refers to the fine earth fraction only (< 2 mm). Weight percentages of sand-, silt- and clay-size materials are given as integers. The analytical procedures and equivalent spherical diameters

(esd) for the clay-, silt-, and sand-size fractions are documented in table *WISE3_LABcodes_Description*. For example, “TE01” corresponds with “pipette method, full dispersion; esd: <0.002 mm, 0.002-0.050 mm, and 0.050- 2 mm”.

GRAVEL
This figure generally represents a visual estimate of the percentage of large rock and mineral fragments — coarse fragments — with a diameter larger than 2 mm, rounded off to the nearest 5 per cent (%v/v). Conversely, when wt% are given for the amount of fragments > 2 mm in the source data, these values have been converted to volume % (see Schoenberger *et al.* 2002).

BULKDENS
Bulk density (oven dry) is given as g cm^{-3}, using two decimal places.

VMCx
Soil water retention, or the percentage of water held in the soil horizon, is given for three, commonly used, pre-defined tensions: -10 kPa (VMC1, pF2.0), -33 kPa (VMC2, pF2.5), and -1500 kPa (VMC3, pF4.2).

Water content data in WISE3 are expressed on a percent by volume basis (VMC), similar to ISIS conventions — conversion procedures are described elsewhere (Landon 1991; Schoenberger *et al.* 2002; van Reeuwijk 2002):

$$VMC_x\ (\%v/v) = MC_x\ (\%\ \text{by weight})\ \times\ \text{bulk density}\ (kg\ m^{-3})$$

[Note: Data on VMC_x, for example (VMC_{-33kPa} – $VMC_{-1500kPa}$) and volume percentage of fragments > 2 mm by horizon, depth of horizon, and effective depth of soil may be used to calculate profile available water capacity.]

Appendix 3. WISE3 quality controls

Data in WISE are routinely controlled for possible inconsistencies; the current set of controls includes:

1) Flag, and subsequently check, profiles for which the 2 leftmost characters of the WISE3_ID differ from the country ISO-code.

2) Flag horizons in a profile whose upper depth is larger than its lower depth; update sequential numbering of horizons.

3) Flag horizons for which the sum of clay, sand and silt fractions is < 98% or >102%; otherwise, normalize to 100% if the total is within this range.

4) Flag horizons for which the bulk density is smaller than 0.5 or greater than 2.0 g cm^{-3}. (**Note** for items 4 to 7: all flagged values are potential errors only, and need not be wrong; for example, a low bulk density for an organic layer or a high bulk density for a compact subsoil; flagged values are checked visually against the source data and classification).

5) Flag horizons for which the C/N ratio is either smaller than 5 or greater than 45 (**Note**: high (computed) values are prone to occur in soils with very low Total N contents, where analytical errors in determining total N contents can be large); flagged values may point at inconsistencies in data entry for organic C and Total N content.

6) Flag all horizons for which either the pH-H_2O, pH-KCl or pH-$CaCl_2$ is either less than 2.0 or greater than 11.0, corresponding with the normal pH-H_2O range for most soils (Soil Survey Staff, 1993 p. 192). Flag possible inconsistencies between pH-values measured in different solutions, for instance pH-KCl > pH-H_2O keeping in mind soil classification aspects (e.g. Geric Ferralsols).

7) Flag horizons for which exchangeable aluminum exceeds 0 cmol (+)/kg, when pH-H_2O is greater than 6.0; similar procedures are used to check for the absence of $CaCO_3$ and gypsum in acid soils.

8) Check whether all records associated to a given WISE3_ID exist, and flag any missing records.

9) Check whether all alphanumeric codes in the various database files match allowed definitions in table *WISE3_coding_conventions*, and flag possible errors. (For example, a field for which AF was entered as the FAO-Unesco (1974) legend code for a Ferric Acrisol will be flagged as being incorrect because the correct entry is: Af).

10) Flag laboratories for which analytical procedures have not (yet) been specified.

Appendix 4. Number of profiles by soil unit (1974 FAO Legend)

A: Acrisols (1081)[a]
Af= 390 Ag= 86 Ah= 84 Ao= 392 Ap= 129
B: Cambisols (1343)
Bc= 72 Bd= 200 Be= 336 Bf= 156 Bg= 137 Bh= 131 Bk= 192 Bv= 81 Bx= 38
C: Chernozems (206)
Cg= 6 Ch= 75 Ck= 61 Cl= 64
D: Podzoluvisols (115)
Dd= 22 De= 70 Dg= 23
E: Rendzinas (103)
F: Ferralsols (533)
Fa= 34 Fh= 68 Fo= 210 Fp= 15 Fr= 98 Fx= 108
G: Gleysols (649)
Gc= 16 Gd= 108 Ge= 280 Gh= 73 Gm= 107 Gp= 48 Gx= 17
H: Phaeozems (400)
Hc= 39 Hg= 35 Hh= 134 Hl= 192
I: Lithosols (27)
J: Fluvisols (575)
Jc= 146 Jd= 92 Je= 283 Jt= 54
K: Kastanozems (97)
Kh= 27 Kk= 44 Kl= 26
L: Luvisols (1542)
La= 37 Lc= 227 Lf= 391 Lg= 289 Lk= 119 Lo= 404 Lp= 39 Lv= 36
M: Greyzems (27)
Mg= 3 Mo= 24
N: Nitosols (163)
Nd= 55 Ne= 75 Nh= 33
O: Histosols (112)
Od= 63 Oe= 38 Ox= 11
P: Podzols (222)
Pf= 7 Pg= 43 Ph= 49 Pl= 20 Po= 89 Pp= 14
Q: Arenosols (771)
Qa= 38 Qc= 423 Qf= 163 Ql= 147
R: Regosols (521)
Rc= 139 Rd= 144 Re= 224 Rx= 14
S: Solonetz (226)
Sg= 62 Sm= 30 So= 134
T: Andosols (293)
Th= 148 Tm= 50 To= 35 Tv= 60
U: Rankers (53)
V: Vertisols (549)
Vc= 226 Vp= 323

W: Planosols (171)
Wd= 36 We= 85 Wh= 7 Wm= 15 Ws= 28 Wx= 0
X: Xerosols (185)
Xh= 33 Xk= 76 Xl= 62 Xy= 14
Y: Yermosols (118)
Yh= 20 Yk= 34 Yl= 44 Yt= 1 Yy= 19
Z: Solonchaks (167)
Zg= 44 Zm= 9 Zo= 111 Zt= 3

[a] For details see Legend (FAO-Unesco 1974).

Appendix 5. Number of profiles by soil unit (1988 FAO Legend)

AC: Acrisols[a] (877)
ACf= 268 ACg= 47 ACh= 404 ACp= 90 ACu= 68
AL: Alisols (185)
ALf= 55 ALg= 37 ALh= 68 ALj= 7 ALp= 8 ALu= 10
AN: Andosols (288)
ANg= 7 ANh= 36 ANi= 0 ANm= 45 ANu= 113 ANz= 87
AR: Arenosols (789)
ARa= 34 ARb= 69 ARc= 59 ARg= 61 ARh= 258 ARl= 163 ARo= 145
AT: Anthrosols (19)
ATa= 0 ATc= 8 ATf= 10 ATu= 1
CH: Chernozems (208)
CHg= 9 CHh= 69 CHk= 63 CHl= 64 CHw= 3
CL: Calcisols (218)
CLh= 141 CLl= 36 CLp= 41
CM: Cambisols (1295)
CMc= 191 CMd= 189 CMe= 315 CMg= 135 CMi= 37 CMo= 161 CMu= 121 CMv= 80 CMx=66
FL: Fluvisols (563)
FLc= 140 FLd= 71 FLe= 229 FLm= 42 FLs= 13 FLt= 50 FLu= 18
FR: Ferralsols (528)
FRg= 37 FRh= 213 FRp= 12 FRr= 101 FRu= 56 FRx= 109
GL: Gleysols (596)
GLa= 1 GLd= 105 GLe= 277 GLi= 16 GLk= 14 GLm= 107 GLt= 5 GLu= 71
GR: Greyzems (27)
GRg= 3 GRh= 24
GY: Gypsisols (37)
GYh= 8 GYk= 20 GYl= 2 GYp= 7
HS: Histosols (112)
HSf= 41 HSi= 10 HSl= 3 HSs= 48 HSt= 10
KS: Kastanozems (97)
KSh= 25 KSk= 45 KSl= 27 KSy= 0
LP: Leptosols (324)
LPd= 79 LPe= 86 LPi= 3 LPk= 46 LPm= 40 LPq= 22 LPu= 48
LV: Luvisols (1135)
LVa= 34 LVf= 115 LVg= 180 LVh= 321 LVj= 77 LVk= 131 LVv= 37 LVx=240
LX: Lixisols (474)
LXa= 3 LXf= 189 LXg= 36 LXh= 226 LXj= 1 LXp= 19
NT: Nitisols (139)
NTh= 67 NTr= 45 NTu= 27
PD: Podzoluvisols (115)
PDd= 22 PDe= 70 PDg= 17 PDi= 0 PDj= 6
PH: Phaeozems (401)
PHc= 39 PHg= 30 PHh= 127 PHj= 11 PHl= 194
PL: Planosols (166)
PLd= 39 PLe= 102 PLi= 0 PLm= 19 PLu= 6
PT: Plinthosols (98)
PTa= 20 PTd= 47 PTe= 28 PTu= 3

PZ: Podzols (222)
PZb= 20 PZc= 37 PZf= 8 PZg= 52 PZh= 87 PZi= 18
RG: Regosols (393)
RGc= 104 RGd= 72 RGe= 187 RGi= 14 RGu= 13 RGy= 3
SC: Solonchaks (165)
SCg= 42 SCh= 31 SCi= 4 SCk= 21 SCm= 8 SCn= 33 SCy= 26
SN: Solonetz (231)
SNg= 52 SNh= 90 SNj= 19 SNk= 37 SNm= 29 SNy= 4
VR: Vertisols (551)
VRd= 18 VRe= 367 VRk= 161 VRy= 5

[a] For details see Revised Legend (FAO 1988).

Appendix 6. Number of profiles by geographical region with their description status

Geographical Region[a]	Description status[b]				Total
	1	2	3	4	
Africa – Eastern	207	172	507	15	901
Africa – Middle	12	34	250	1	297
Africa – Northern	50	68	41	0	159
Africa – Southern	67	788	148	1	1004
Africa – Western	85	275	1446	6	1812
America - Carribean	26	89	69	0	184
America - Central	94	56	4	10	164
America - Northern	375	77	54	1	507
America - South	149	1380	313	1	1843
Antarctica	4	6	0	0	10
Asia - Central	0	0	13	0	13
Asia - Eastern	81	207	48	0	336
Asia - South Eastern	190	276	70	1	537
Asia - Southern	71	196	133	8	408
Asia - Western	99	291	162	1	553
Europe - Eastern	78	508	75	7	668
Europe - Northern	29	127	91	3	250
Europe - Southern	73	43	90	9	215
Europe - Western	45	34	103	1	183
Oceania – Australia & New Zealand	8	39	81	4	132
Oceania – Melanesia	13	4	25	0	42
Oceania – Micronesia	13	3	0	0	16
Oceania – Polynesia	16	3	0	0	19
WISE3	1785	4676	3723	69	10253

[a] Geographical regions are according to UN conventions (UN Statistics Division 2008). Totals by continent are listed in Table 2; the continent of North America comprises Northern America, Caribbean, and Central America. [b] For details see Appendix 2A.

Appendix 7. Number of soil profiles by country with their description status

Country	Description status[a]				Total
	1	2	3	4	
Afghanistan	0	0	5	0	5
Albania	28	1	0	6	35
Algeria	1	4	1	0	6
Angola	0	1	35	0	36
Antarctica	4	6	0	0	10
Argentina	0	212	49	0	261
Armenia	0	0	6	0	6
Australia	7	33	72	4	116
Azerbaijan	0	0	4	1	5
Bahrain	0	0	3	0	3
Bangladesh	0	3	16	0	19
Barbados	0	3	0	0	3
Belarus	0	89	1	0	90
Belgium	4	5	2	0	11
Benin	3	0	856	1	860
Bhutan	0	0	8	3	11
Bolivia	0	70	3	0	73
Botswana	26	747	116	0	889
Brazil	45	594	53	0	692
Bulgaria	23	20	3	0	46
Burkina Faso	1	1	0	0	2
Burundi	12	0	19	1	32
Cameroon	4	31	21	0	56
Canada	14	39	51	1	105
Central African Republic	0	0	32	0	32
Chile	5	21	25	1	52
China, mainland	50	0	10	0	60
Colombia	27	106	30	0	163
Congo, Republic of the	0	0	5	1	6
Congo, The Democratic Republic Of The	2	0	140	0	142
Cook Islands	0	1	0	0	1
Costa Rica	26	3	0	0	29
Côte d'Ivoire	8	16	2	0	26
Cuba	22	28	1	0	51
Czech Republic	0	37	1	0	38
Denmark	0	7	13	0	20
Dominican Republic	0	7	0	0	7
Ecuador	32	30	19	0	81
Egypt	0	14	5	0	19
El Salvador	5	0	0	0	5

Country	Description status[a]				Total
	1	2	3	4	
Estonia	0	48	2	0	50
Ethiopia	0	1	10	0	11
Fiji	0	0	9	0	9
Finland	6	4	9	1	20
France	10	0	47	1	58
French Guiana	0	8	0	0	8
Gabon	6	2	17	0	25
Georgia	0	0	9	0	9
Germany	13	23	18	0	54
Ghana	5	199	23	0	227
Greece	11	1	1	0	13
Greenland	3	0	0	0	3
Grenada	0	4	1	0	5
Guatemala	11	0	0	0	11
Guinea	1	1	6	0	8
Guyana	0	41	2	0	43
Honduras	8	0	0	0	8
Hungary	3	40	0	0	43
Iceland	6	0	0	0	6
India	28	65	69	2	164
Indonesia	79	4	19	0	102
Iran, Islamic Republic of	0	0	0	2	2
Iraq	1	0	14	0	15
Ireland	8	0	28	0	36
Israel	0	0	33	0	33
Italy	15	15	51	3	84
Jamaica	4	14	63	0	81
Japan	8	7	24	0	39
Jordan	14	15	18	0	47
Kenya	87	3	260	0	350
Korea, Democratic People's Republic of	3	10	0	0	13
Korea, Republic of	14	187	14	0	215
Latvia	0	11	0	0	11
Lebanon	0	2	5	0	7
Lesotho	15	0	12	0	27
Liberia	0	16	2	0	18
Libyan Arab Jamahiriya	0	10	7	0	17
Lithuania	0	19	1	0	20
Luxembourg	0	0	1	0	1
Madagascar	20	1	21	12	54
Malawi	0	0	3	0	3
Malaysia	19	5	22	0	46
Mali	18	2	0	0	20
Mauritania	0	11	0	0	11
Mexico	9	53	4	0	66

Country	Description status[a]				Total
	1	2	3	4	
Micronesia, Federated States of	4	1	0	0	5
Moldova, Republic of	0	44	1	0	45
Mongolia	5	3	0	0	8
Morocco	6	2	20	0	28
Mozambique	6	48	6	0	60
Namibia	11	34	3	1	49
Nepal	5	126	23	1	155
Netherlands	18	6	21	0	45
Netherlands Antilles	0	2	0	0	2
New Caledonia	0	1	0	0	1
New Zealand	1	6	9	0	16
Nicaragua	21	0	0	10	31
Niger	11	0	445	5	461
Nigeria	27	10	2	0	39
Norway	3	2	4	0	9
Oman	10	5	0	0	15
Pakistan	35	2	2	0	39
Palau	9	2	0	0	11
Panama	14	0	0	0	14
Papua New Guinea	13	2	16	0	31
Paraguay	0	43	1	0	44
Peru	25	60	29	0	114
Philippines	56	16	1	0	73
Poland	9	27	33	3	72
Portugal	1	24	11	0	36
Puerto Rico	0	17	1	0	18
Romania	11	35	16	0	62
Russian Federation	26	113	14	4	157
Rwanda	14	4	44	0	62
Samoa	16	0	0	0	16
Senegal	0	3	102	0	105
Serbia and Montenegro	0	0	6	0	6
Sierra Leone	0	11	0	0	11
Slovakia	6	28	0	0	34
Solomon Islands	0	1	0	0	1
Somalia	5	2	8	0	15
South Africa	15	7	17	0	39
Spain	18	2	21	0	41
Sri Lanka	3	0	10	0	13
Sudan	27	38	8	0	73
Suriname	0	23	8	0	31
Sweden	5	7	11	2	25
Switzerland	0	0	14	0	14
Syrian Arab Republic	57	2	4	0	63
Taiwan (Republic of China)	1	0	0	0	1

Country	Description status[a]				Total
	1	2	3	4	
Tajikistan	0	0	5	0	5
Tanzania, United Republic Of	0	93	55	1	149
Thailand	36	251	28	1	316
Togo	11	5	8	0	24
Tonga	0	2	0	0	2
Trinidad and Tobago	0	14	3	0	17
Tunisia	16	0	0	0	16
Turkey	13	24	29	0	66
Uganda	12	0	0	0	12
Ukraine	0	75	6	0	81
United Kingdom	1	29	23	0	53
United States	358	38	3	0	399
Uruguay	10	36	72	0	118
Uzbekistan	0	0	8	0	8
Venezuela	5	136	22	0	163
Yemen	4	243	37	0	284
Zambia	31	5	50	0	86
Zimbabwe	20	15	31	1	67
WISE3	*1785*	*4676*	*3723*	*69*	*10253*

[a] For details see Appendix 2A.

Appendix 8. Percentage of records in WISE3 with measured data

a) Site data (*n*= 10253)

Attribute[a]	Type	Number of records filled	
		n	%
WISE3_id	T	10253	100
ISO	T	10253	100
DESCR	T	10253	100
DATEyr	N	6685	65
DATEmon	N	5561	54
SOLDEP	N	10253	100
HORNUM	N	10253	100
FAO_90	T	10253	100
PHA_90	T	773	8
FAO_74	T	10249	100
PHA_74	T	716	7
WRB2006	T	10253	100
USCL	T	3593	37
USYR	T	2295	22
LOCAL	T	3779	37
DRAIN	T	10253	100
KOPPEN	T	4710	46
LANDUS	T	3950	39
LAB_ID	T	10253	100
SOURCE_ID	T	10253	100
PITREF	T	10253	100
LOCAT	T	4775	47
LATIT	T	8189	80
LATDEG	T	8189	80
LATMIN	T	8161	80
LATSEC	T	5555	54
LONGI	T	8189	80
LONDEG	T	8189	80
LONMIN	T	8066	79
LONSEC	T	5500	54
ALTIT	N	6900	67
LFORM	T	3980	39
POSIT	T	2853	28
PARMAT	T	3407	33
ASPECT	T	1298	13
SLOPE	T	3068	30

[a] See Appendix 2 for code definitions; T stands for text field and N for Numeric data.

b) Horizon data (*n*= 47834)

Attribute[a]	Type	Number of records filled	
		n	%
WISE3_ID	T	47834	100
HORIZ	T	47834	100
HONU	N	47834	100
TOPDEP	N	46685	100
BOTDEP	N	47834	100
DESIG	T	41774	87
MCOLOR	T	29090	62
DCOLOR	T	10607	22
ORGC	N	41577	87
TOTN	N	25755	54
CACO3	N	34007	71
GYPSUM	N	26424	55
PHH2O	N	43587	91
PHKCL	N	20714	43
PHCACL2	N	11544	24
ECE	N	32378	68
EXCA	N	35176	74
EXMG	N	34910	73
EXNA	N	29273	61
EXK	N	32946	69
EXALUM	N	23655	49
EXACID	N	26015	54
CECSOIL	N	42714	89
BSAT[a]	N	24898	52
SAND	N	43920	92
SILT	N	43920	92
CLAY	N	43920	92
GRAVEL[a]	N	10304	22
BULKDENS	N	15415	32
VMC1[a]	N	1858	4
VMC2[a]	N	6030	13
VMC3[a]	N	6961	15

[a] See Appendix 2 for code definitions; T stands for text field and N for Numeric data.
[b] Calculated or derived values — see text for details.

World Soil Information

ISRIC - World Soil Information is an independent foundation with a global mandate, funded by the Netherlands Government, and with 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*