Report 2003/03

A taxotransfer rule-based approach for filling gaps in measured soil data in primary SOTER databases

(Version 1.1)

(GEF-SOC Project: GFL-2740-02-4381)
(VROM Project: DGM-200207282)

Niels H Batjes

(November 2003)

Global Environment Facility

United Nations Environment Programme

Netherlands Ministry of Housing, Spatial Planning and the Environment

ISRIC World Soil Information

Disclaimer:

Correct citation:

Batjes NH 2003. A taxotransfer rule-based approach for filling gaps in measured soil data in primary SOTER databases (ver. 1.1; GEFSOC Project). Technical Report 2003/03, ISRIC – World Soil Information, Wageningen

Last revision: 8 March 2005

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

Contents

List of Tables

List of Figures

SUMMARY

ISRIC, FAO and UNEP under the aegis of IUSS are updating the information on world soil resources in the World Soils and Terrain Digital Databases (SOTER) project. Primary SOTER databases are composed of two main elements: a geographic and an attribute data component. The first shows the delineations of the SOTER units, while the second holds information on their composition in terms of main soil types described by a suite of representative profiles.

Representative soil profiles for SOTER are selected from existing soil survey reports. Often there are gaps in the associated soil analytical data. This often precludes the direct use of primary SOTER data in environmental models. This study presents a methodology for filling gaps in the primary soil analytical data.

There are three main stages: 1) collating additional measured soil analytical data to consolidate the existing primary data sets; 2) filling gaps using national expertise and common sense; and 3) filling the remaining gaps using taxotransfer rules.

The current report focuses on the taxotransfer scheme. It draws on the soil analytical data held in an auxiliary global soil profile database (WISE – World Inventory of Soil Emission Potentials). Soil parameter estimates by soil unit (FAO Revised Legend) are presented for fixed depths intervals of 0.2 m each (up to 1 m depth) and with reference to the soil textural class. They include organic carbon, total nitrogen, pH(H_2O), CEC_{soil}, CEC_{clay}, base saturation, effective CEC, aluminum saturation, $CaCO_3$ content, gypsum content, exchangeable sodium percentage (ESP), electrical conductivity of saturated paste (ECe), bulk density, sand, silt, clay, content of coarse fragments, and available water capacity. This list has been identified as being useful for agro-ecological zoning (AEZ), land evaluation, crop growth simulation, studies of soil carbon stocks and change, and analyses of global environmental change.

Results are presented digitally. This information can be linked to the soil geographic component of SOTER through the unique SOTER unit and profile identifiers. All taxotransfer rules have been flagged to provide an indication of the confidence in the derived data.

The present approach is considered appropriate for studies at scales smaller than 1:250 000. Correlation of soil analytical data, however, must be done more rigorously when more detailed scientific work is considered.

In the first instance, the methodology will be applied to primary SOTER databases for Brazil, India, Jordan and Kenya in the framework of the GEF co-funded project on *Assessment of Soil Organic Carbon Stocks and Change at National Scale*.

Keywords: soil parameter estimates; environmental modelling; soil carbon; WISE database; SOTER database; taxotransfer

1. INTRODUCTION

Requirements for soil information include the need for an up-to-date geographical coverage, access to secondary soil information obtained via transfer functions or models from the primary (measured) soil data, and monitoring of changes in soil characteristics as associated, for example, with changes in land use systems and processes of global change (Batjes 2002b; Baumgardner 1999; Bullock 1999). With the digitizing of the Soil Map of the World (FAO 1995; FAO-Unesco 1974) and completion of the first version of the global WISE database at ISRIC (Batjes 1997; Batjes and Bridges 1994), it became possible to list soil parameter estimates for each soil unit. Staff at ISRIC, the International Institute for Applied Systems Analysis (IIASA), and FAO subsequently developed a common methodology for deriving soil parameters identified as being important for land evaluation in the context of regional and global agro-ecological zoning (Batjes *et al.* 1997). They also documented the main geographic and taxonomic gaps present in WISE version 1.0. This provided the basis for a follow-up study that generated revised soil parameter estimates for the main soil types of the world in collaboration with the International Food Policy Research Institute (IFPRI) (Batjes 2002b).

The preceding studies made use of the now obsolescent FAO-Unesco (1974) Legend because it is the sole Legend that permits linkage of derived soil parameter estimates with the soil geographic information displayed on the digital Soil Map of the World (SMW). Although the 1:5 million scale SMW (FAO-Unesco 1974-1981) remains the only worldwide, consistent, harmonized soil inventory available in digital format (FAO 1995), the associated spatial data and attribute data are out of date.

In recognition of the above ISRIC, FAO and UNEP under the aegis of the International Union of Soil Sciences have developed a methodology for updating the information on world soil resources in the World Soils and Terrain Database (SOTER) project (Baumgardner and Oldeman 1986; Nachtergaele 1999; Van Engelen 1999). Comprehensive, regional updates completed so far include those for Latin America and the Caribbean (FAO *et al.* 1998b), Central and Eastern Europe (FAO and ISRIC 2000) and Southern Africa (FAO *et al.* 2003). Work on a 1:5 million scale SOTER database for

Europe is ongoing (King *et al.* 2002) and a pilot study has been undertaken in the USA (Dobos *et al.* 2002). A full overview of SOTER activities worldwide, both at the regional and national level, can be found elsewhere (Van Engelen 1999).

Although the information for SOTER is collated according to the same uniform methodology (Van Engelen and Wen 1995), the specific detail of information in each region will result in a variable scale and quality of the end products. Being largely based on soil profile information compiled from soil survey reports, measured data are seldom available for attributes such as bulk density, water holding capacity and soil hydraulic conductivity. Sometimes, even, some of the so-called *mandatory* analytical data (sand, silt, clay, pH_{water}, bulk density and organic carbon) are not available. Inherently, this will preclude the direct use of primary SOTER databases in environmental assessments and modelling. Hence the need for standardized procedures that can help fill gaps in the essential measured soil analytical data in secondary SOTER sets, using parameter estimates derived from auxiliary soil databases such as WISE. This is particularly important in the case of *interim* SOTER databases that, so far, only include no or very limited soil profile data (FAO *et al.* 1998a; FAO and IIASA 1999).

The current study expands ISRIC's earlier taxotransfer work with FAO, IIASA and IFPRI (Batjes 2002b; Batjes *et al.* 1997). A guiding criterion is that results must allow linkage with SOTER databases that use the Revised Legend (FAO 1988). Consequently, the approach no longer considers the original Legend of the Soil Map of the World (FAO-Unesco 1974-1981).

In the first instance, the approach will be applied to the SOTER databases for Brazil (FAO *et al.* 1998b), Jordan (NSMLUP 1996) and Kenya (KSS 1996). The resulting secondary data sets will be used to model soil carbon stocks and change, at the national scale, using the RothC and Century models (Falloon *et al.* 1998; Paustian *et al.* 1997). This work is undertaken in the framework of the GEF co-funded project on *Assessment of Soil Organic Carbon Stocks and Change at National Scale* (GFL-2740-02-4381). This project aims to 'develop and demonstrate generic tools which quantify the impact of land management and climate scenarios on change in soil carbon stocks at national and subnational level' (GEFSOC[1]). The study involves national scientists in Brazil, India, Jordan and Kenya working closely with data management and modeller groups in the United Kingdom, Austria, France, the Netherlands and the USA.

2. PRIMARY SOTER DATABASES

2.1 Methodology

The SOTER methodology allows mapping and characterization of areas of land with a distinctive, often repetitive, pattern of landform, lithology, surface form, slope, parent material, and soils (Van Engelen and Wen 1995). In many aspects it resembles physiographic or land systems mapping.

The SOTER approach to mapping and database compilation is applied mainly at scales ranging from 1:5M to 1:250 000 (FAO and ISRIC 2000; FAO *et al.* 1998b; Graef 1999; KSS 1996; NSMLUP 1996). Application at scales up to 1:50 000 is also possible, but this may require minor changes in some of the definitions (De Oliveira and Van den Berg 1992; Mantel *et al.* 1999). Issues of data acquisition, quality control and sharing encountered while implementing SOTER projects have been discussed by Batjes (1999).

Various sources of uncertainty remain in the soil geographic and attribute data that have been used to compile the existing SOTER databases, and these should be corrected gradually in revised versions of the primary data sets. A promising development in this respect will be the use of Digital Elevation Models (DEM) to systematically delineate landscape units (Dobos *et al.* 2002) in SOTER.

[1] See: www.reading.ac.uk/GEFSOC/

2.2 Database structure

Each SOTER database is comprised of two main elements, a geographic component and an attribute data component (Figure 1). The *geographic database* holds information on the location, extent and topology of each SOTER unit – this information is managed using a geographic information system (GIS). The *attribute database* describes the characteristics of the spatial unit and comprises both area data and point data – this information is handled using a relational database management system (RDBMS).

Each SOTER unit has a unique identifier, called SOTER unit ID (SUID). This primary key provides a link to the attribute data for its constituent terrain, terrain component(s) (TCID) and soil component(s) (SCID).

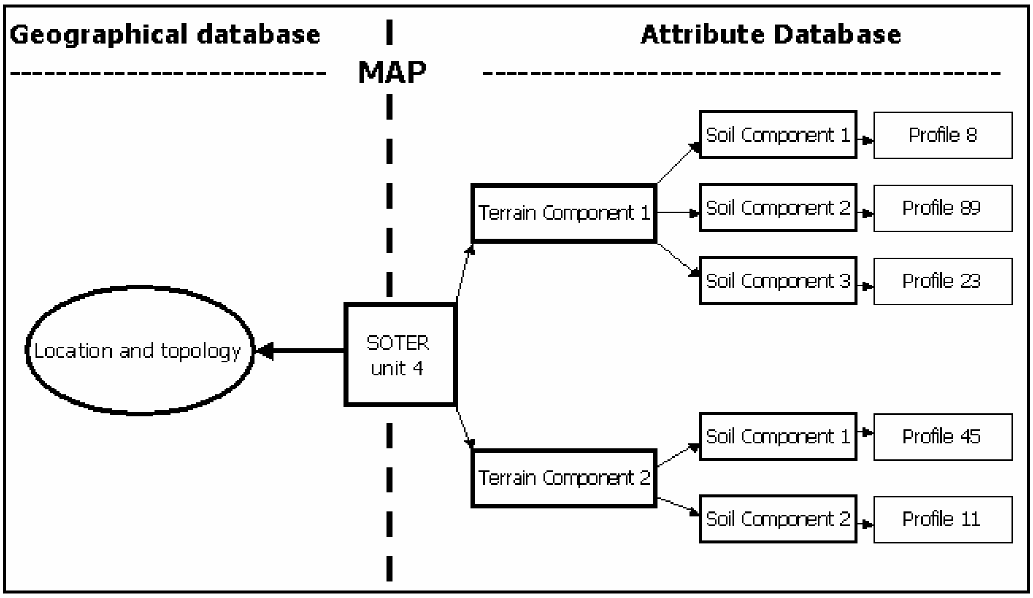

Figure 1. Schematic representation of a SOTER map unit with its geographical and attribute data

Each soil component within a SOTER unit is characterized by a profile, identified as being regionally representative by the national soil experts. These profiles are derived from available soil survey reports, as the SOTER program does not involve any new ground surveys. Characterization according to the

Revised Legend of FAO (1988) is mandatory and classification according to the World Reference Base for Soil Resources (WRB 1998) recommended.

A comprehensive description of the SOTER methodology and coding conventions is given by Van Engelen (1995). The SOTER attribute data are managed with an automated data entry facility (Tempel 2002). In addition, SOTER uses commercially available Access® and ArcView® software.

Table 1. Example of assessments that use SOTER data

Scale	Region	Application and Reference
1:5 M	Latin America and Amazon basin	Soil carbon stocks (Batjes 2000a; Batjes and Dijkshoorn 1998)
1:2.5 M	Central and Eastern Europe	Soil vulnerability to pollution (Batjes 2000b; Nachtergaele *et al.* 2002); soil carbon stocks and change (Batjes 2002a)
1:1 M	Kursk region, Russia	Crop simulation for the analysis of land resources (Savin *et al.* 1997)
	Argentina, Kenya and Uruguay	Impact of water erosion on crop productivity using integrated modelling (Mantel and Van Engelen 1997, 1999)
1:0.5 M	Hungary	Soil vulnerability to pollution (Varallyay *et al.* 1994); soil carbon modelling (Falloon *et al.* 1998)
1:0.25 M	Hainan, P.R. China	Land evaluation and water erosion risk assessments for regional land use planning (Mantel *et al.* 2003)
1:0.20 M	S.W. Niger	Parametric land evaluation; evaluation of cropping systems (Graef 1999)
1:0.05 M	Berau, Indonesia	Inventory of site qualities for forest management planning (Mantel *et al.* 1999)

2.3 Database applications

Data collated in the primary SOTER attribute tables can be linked to GIS, permitting a wide range of applications. These range from land evaluation to soil carbon sequestration (Table 1). However, in these studies, it has been necessary to fill gaps in the measured primary data using tailor-made solutions (e.g. Batjes and Dijkshoorn 1998; Mantel and Van Engelen 1999). A standardized procedure for doing this is elaborated in the next Chapter.

3. SECONDARY SOTER DATABASES

3.1 General approach

The soil profile attribute data for SOTER are derived mainly from soil survey reports and complete soil analytical data sets are seldom available for all profiles.

The following procedure has been developed for filling gaps in the measured soil analytical data. It involves three stages, the desirability of which decreases from highest (a) to lowest (c):

a) Collate additional measured soil data where these exist, in the uniform SOTER format;
b) Use national expert estimates and common sense to fill selected gaps in a secondary data set;
c) Use taxotransfer rules (TTR) to derive soil parameter estimates for similar FAO soil units, as obtained from auxiliary global soil profile databases.

The most appropriate option(s) will vary from country to country, depending on the overall accessibility and quality of the available soil data. Activities (a) and (b) are the key responsibility of the local partners in most national SOTER activities – these experts have the most direct access to the necessary data, should they exist, as well as local knowledge. Discussion of the first two stages is beyond the scope of the present report and they will be discussed in country specific reports (e.g. Batjes *et al.* 2003).

3.2 Comparability of soil analytical data

Generalization of measured soil data involves the transformation of variables that often show a marked spatial and temporal variability. International soil classification, correlation and interpretation require international exchange for comparability of analytical data. Table 2 illustrates the difficulties in comparing analytical data in soil data compilation activities that, *of necessity*, are based on available, historic data. Serious problems are prone to arise with the comparison of soil analytical data originating from disparate laboratories and surveys (Pleijsier 1989). Indeed, seldom have all profile data in a single national or global database been determined according to a uniform set of sampling and analytical procedures as well as in one single laboratory. The few exceptions include ISRIC's Soil Reference Collection and the profiles analysed by NRCS-USDA (Lincoln).

Table 2. Possible sources of variation in soil analytical data

Source of variation	Within laboratory	Between laboratory
Definitions		X
Procedures		X
Execution of procedure	X	
Instruments	X	
Operator	X	
Random error	X	
Calculations	X	X

Source: Pleijsier (1986)

Soil profiles held in large databases often have to be classified with reference to the original methods in use at the national level, and not those explicitly demanded by international classification systems (see CEC 1985; FAO and ISRIC 2000; FAO *et al.* 1998b; FAO-Unesco 1974-1981; Stolbovoi 2001). Therefore, a pragmatic approach to the comparability of soil analytical data has had to be adopted in SOTER – similar to what has been the case for the Soil Map of the World (FAO 1995), the WISE database (Batjes 2002b) and the EU Soil Database (Finke *et al.* 1998; Madsen and Jones 1998). In each case, however, the original soil analytical methods have been documented for possible future reference, pending the availability of appropriate comparability studies at the Regional and Global level.

3.3 Choice of soil parameters

Special attention has been paid to the key parameters required for the spatial runs of the two organic carbon models considered in the GEF-SOC project – RothC and Century. These are: the extent and type of soil, soil drainage status, content of clay, content of organic carbon, and bulk density per depth layer (Falloon *et al.* 1998; Paustian *et al.* 1997). Information on the content of coarse fragments is also needed to calculate carbon stocks (Batjes 1996). The charge and surface density of clay minerals may explain a large part of the variation in the turnover of soil organic matter (Wattel-Koekkoek 2002), for which CEC_{clay} can be a proxy. Besides influencing long-term carbon storage and turnover, C stabilization by non-crystalline minerals – for example allophane in Andosols – also influences nutrient availability and decomposition of labile substrates in surface soils (Torn *et al.* 1997).

Water holding capacity, pH and nutrient retention capacity are, also, important determinants of possible production levels and, thus, inputs of fresh organic materials into the soil. Ideally, these aspects should be considered also in modelling studies of soil organic carbon stocks and change, in addition to changes in climate and land use/management practices.

The basic set of soil parameters required for the RothC and Century models has been expanded to include 18 soil parameters in total (Table 3). These are commonly required in studies of agro-ecological zoning, food productivity, soil gaseous emissions/sinks and environmental change (see Batjes 2004; Bouwman *et al.* 2002a, b; Cramer and Fischer 1997; Fischer *et al.* 2002; Mantel and Van Engelen 1999; Scholes *et al.* 1995).

Table 3. List of soil parameters

Organic carbon
Total nitrogen
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) † ‡
Aluminum saturation (as % of ECEC) ‡
$CaCO_3$ content
Gypsum content
Exchangeable sodium percentage (ESP) ‡
Electrical conductivity of saturated paste (ECe)
Bulk density
Coarse fragments (volume %)
Sand (mass %)
Silt (mass %)
Clay (mass %)
Available water capacity (AWC; from -33 to -1500 kPa; % w/v) ‡ □

‡ Calculated from other measured soil properties.

† ECEC is defined as exchangeable (Ca^{++}+Mg^{++}+K^{+}+Na^{+}) + exchangeable (H^{+}+Al^{+++}) (Van Reeuwijk 1995).

• 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 are not yet corrected for coarse fragments.

Table 3 does not consider soil hydraulic properties. Although these are essential for many simulation studies they are seldom measured during soil surveys. As a result, the corresponding records have seldom been filled in databases such as SOTER and WISE. Information on soil hydraulic properties and pedotransfer functions for Western Europe and the USA may be found in auxiliary databases (see Nemes *et al.* 2003a; Nemes *et al.* 2003b; Wösten *et al.* 1998), but similar work for tropical soils has just begun (Tomasella and Hodnett 1997, 1998; Van den Berg *et al.* 1997).

3.4 Development of taxotransfer scheme

3.4.1 Background

A *taxotransfer* function is a means of estimating soil parameters based on modal soil characteristics of soil units from a combination of their classification name or taxon (which by definition implies a certain range for various soil attributes), expert knowledge and empirical rules, and statistical analysis of a large number of soil profiles belonging to the same taxon (Batjes *et al.* 1997). The elaboration of taxotransfer rules thus requires the availability of large soil profile databases such as WISE (Batjes and Bridges 1994).

The taxotransfer (TTR) approach has been developed in a collaborative study by ISRIC, IIASA and FAO staff (Batjes *et al.* 1997). Subsequently, it has been refined in a follow up study with IFPRI (Batjes 2002b, c).

The preceding studies focussed on the FAO-Unesco (1974) Legend, as it allows linkage of the soil parameter estimates with the soil units shown on the digital Soil Map of the World (FAO 1995). Aggregation of soil profile data in these studies was in accordance with requirements inherent to the Soil Map of the World – that is consideration of the topsoil (0 - 30 cm) and subsoil (30 - 100 cm), and clustering into 3 *topsoil* textural classes. A similar approach has been applied to the SOTER database for Central and Eastern Europe (Batjes 2000c), instead using the Revised Legend (FAO 1988). The usefulness of soil classes as carriers of information is well documented (Batjes *et al.* 1997; Bouma *et al.* 1998; FAO 1995; Madsen and Jones 1998).

3.4.2 General approach

This study describes an updated taxotransfer approach for use with primary SOTER databases. It considers the Revised Legend (FAO 1988) – in accordance with current SOTER requirements (p. 9 in Van Engelen and Wen 1995) – and uses a more detailed aggregation procedure for the soil profile data. Clustering is now into 5 depth layers and 5 textural classes (see below).

Some 9600 profiles currently held in WISE (Batjes 2002b) provided the basic soil analytical input for TTR development (see Appendix 6).

The approach involves nine stages:

a) Screening by profile of apparent reliability/completeness of the measured soil analytical data – both in SOTER and in WISE – using the automated procedures developed for the WISE database (Batjes 1995).

b) Allocation of individual samples of a profile to five depth intervals: 0 to 20 cm (D1), 20 to 40 cm (D2), 40 to 60 cm (D3), 60 to 80 cm (D4) and 80 to 100 cm (D5). Fixed depth intervals of 20 cm have been used instead of the traditional 0-30 cm and 30-100 cm, as this permits a better representation of changes in soil properties with depth. Horizons below a depth of 100 cm (D6) are not considered yet because there are few measurements for these deeper horizons.

c) Allocation of each depth-layer to one of five textural classes, using commonly used class limits for percent sand, silt and clay (CEC 1985; FAO 1988; Finke *et al.* 1998). The limits of < 15 percent clay and > 70 percent sand for Coarse in Figure 2 – instead of the original < 18 percent clay and > 65 percent sand (CEC 1985; FAO-Unesco 1974) – have been introduced with the Revised Legend (FAO 1988, p. 107).

d) Screening on the comparability of soil analytical data based on criteria developed for the International Laboratory Methods and Data Exchange (LABEX) program. The procedure assumes that the data are normally distributed and rejects the outlying 5 per cent of the data (Pleijsier 1986). The sample populations per soil type, soil attribute, depth layer and textural class which remained after exclusion of the outliers, provided the basis for the following statistical analysis.

e) Statistical analysis of the accepted data: sample size (n_{WISE}), mean, coefficient of variation (CV), median, median of absolute deviations (MAD), standard deviation, variance and extremes. In view of their length, the corresponding files are provided only in digital format (see Appendix 1 for table structure).

f) Appcation of taxotransfer rules to fill gaps in the primary SOTER profile data. The median is used for the substitution since it is a better estimate of the centre of the data than the mean (Snedecor and Cochran 1980; Spain *et al.* 1983). Further details are given in Section 3.4.3.

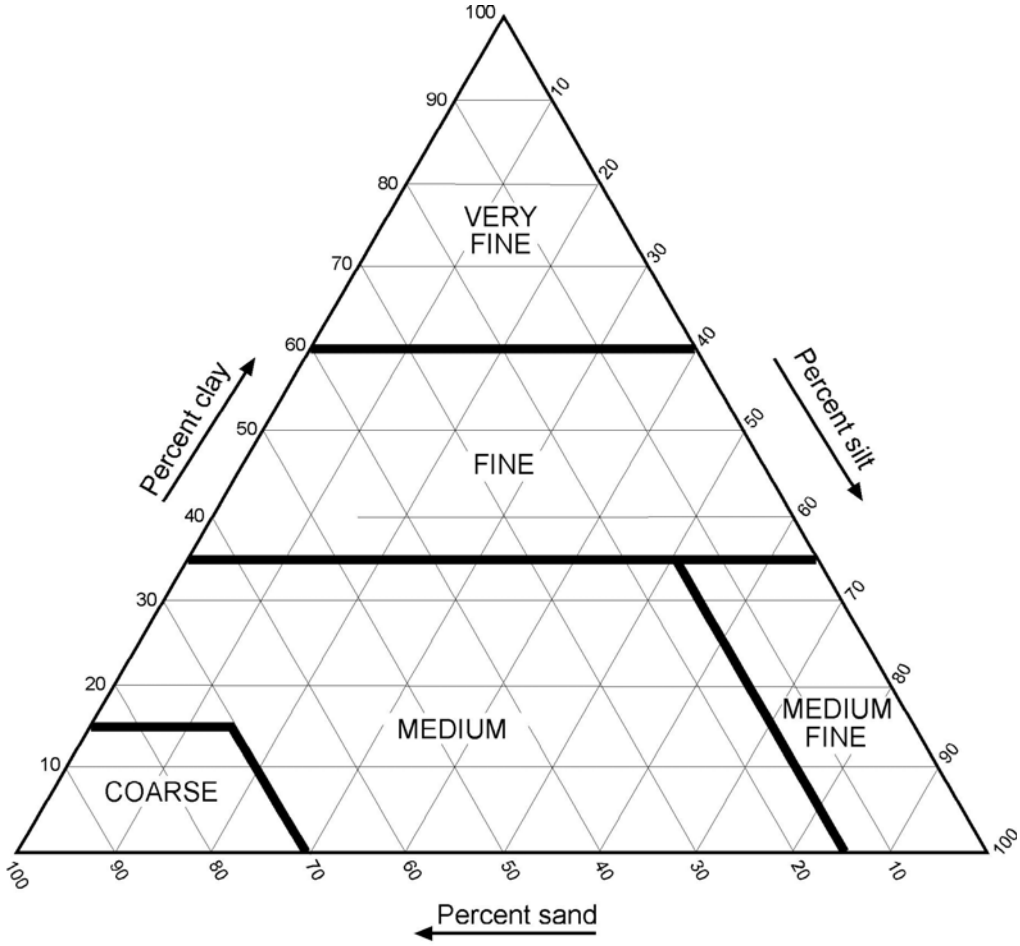

Figure 2. Soil texture classes

g) Depth weighting of the measured and TTR-derived data, per SOTER profile and soil attribute, for each of the five depth intervals under consideration.

h) Storage of the results in a set of secondary SOTER tables, the structures of which are described in the Appendix 3.

i) Linkage of results to the soil geographical component of the corresponding SOTER database using GIS (see Appendix 2; Chapter 4).

3.4.3 Taxotransfer scheme

The TTR scheme is applied to the SOTER profile data by depth layer, with reference to the FAO soil unit, soil textural class and attribute under consideration.

The overall procedure may be summarized as:

a) If measured data exist in the screened primary SOTER data set, this value is maintained in the temporary, secondary file.

b) Otherwise the missing value is replaced with the median value derived from WISE for the considered combination of soil unit, attribute, depth zone and textural class, unless $n_{WISE} < 5$.

   For example, if there are no measured data for, say, the first (0-15 cm; D1) and coarse-textured (M) horizon of a given ferric Acrisols (ACf) – as characterised by its unique profile identifier (PRID) – for CEC_{SOIL} (CECS), then the corresponding missing value is substituted in the secondary SOTER set using the WISE-derived median for the sample set coded ACfCECSD1M.

c) Or else the median for the corresponding combination of soil unit, attribute and depth zone is substituted in the secondary SOTER set, irrespective of the soil textural class, again provided $n_{WISE} < 5$. In the current example, the corresponding combination of soil parameters estimates derived from WISE is flagged ACfCECSD1#.

d) Or else if there are less than 5 samples for the selected combination (e.g. ACfCECSD1M) in WISE, then the median for the corresponding major soil grouping (AC), attribute (CECS), depth zone (D1), and texture class (M) will be used. The corresponding subset is flagged ACCECSD1M in the WISE derived data set. This rule is only applied if $n_{WISE} > 5$.

e) Or else the WISE-derived median for the corresponding major soil grouping, attribute and depth range is used, irrespective of the soil textural class (i.e., ACBCECSD1#), provided again that $n_{WISE} > 5$ for the corresponding subset.

f) Otherwise, no meaningful substitution can be made – the current version of the WISE profile data set does not yet warrant a plausible substitution, as n_{WISE} is still less than 5. This is flagged with a -1 in the secondary SOTER set. This flag also points to the need for additional data collection in field surveys (see 4.4).

g) Next, for each depth layer represented in the profile, the available primary and TTR-derived soil data are depth weighted and stored in the final secondary SOTER data set (see Appendix 3 for attribute coding and table structure).

h) During the above stages, the type of TTR's used is clearly flagged and stored in a separate table. The flags give an indication of the possible confidence in the derived data (see 3.5).

i) Finally, the secondary SOTER table can be joined with a summary file showing the soil geographic information (see Appendix 2), permitting its linkage to and handling in a GIS.

3.5 Tracking TTR-related changes

The type of TTR used, if any, has been flagged by profile and depth layer in a separate table (see Appendix 4). This is illustrated in Figure 3, using results of ongoing work with the SOTER database for Jordan (Batjes *et al.* 2003; NSMLUP 1996).

The flag PTRsub indicates that the data substitution for a given attribute, in the secondary SOTER set, is based on WISE-derived parameter estimates for similar soil units. Otherwise, should the corresponding population in WISE be too small ($n_{WISE} < 5$) for a meaningful substitution, the rules used are flagged under PTRmain.

SOTERflagPTRrules : Table

CLAF	PRID	Layer	Newtop	Newbotd	PTRsub	PTRmain
CMc	JOPA130	D1	0	7	b3c2j3o3r2	a2h1
CMc	JOPA130	D1	7	20	b3c2j3o3r2	a2h1
CMc	JOPA130	D2	20	40	b3c2j2o3r2	a3h2
CMc	JOPA130	D3	40	60	b3c3d3j3o3r3	A2H1
CMc	JOPA130	D4	60	80	B1C1d3j3O2r3	A2H1
CMc	JOPA130	D5	80	100	B1C1D1J1O2R1	A2H1
CLh	JOPA179	D1	0	7	b1c1j1o3	-
CLh	JOPA179	D1	7	20	b1c1j1o3r1	-
CLh	JOPA179	D2	20	40	b2c2j2O2r3	-
CLh	JOPA179	D3	40	47	b2c2j2O2r3	-
CLh	JOPA179	D3	47	60	b2c2j2O2r3	-
CLh	JOPA179	D4	60	80	b2c2j2O3r3	-
CLh	JOPA179	D5	80	86	B1C1J1O3R2	-
CLh	JOPA179	D5	86	100	B1C1d3J1O3R2	-
LPe	JOPA210syn	D1	0	5	b2c3d1e1i1j2k3l1O3p3r1	A4H4
LPe	JOPA210syn	D1	5	20	b2c3d1e1i1j2k3l1O3p3r1	A4H4
LPe	JOPA210syn	D2	20	26	b3c3d2e2i3j3l2O3p3r2	A4H4k3

Record: 1 of 292

Figure 3. Flagging of taxotransfer rules by profile, depth zone and attribute (see text for details)

Each flag consists of a sequence of letters followed by a numeral (Figure 4). The letters indicate for which soil attributes a TTR has been applied. The number reflects the size of the sample population in WISE, after outlier rejection, which provided the basis for the statistical analyses (Table 4). The assumption is that the confidence in a TTR-based parameter estimate should increase with the size of the sample populations present in WISE, after outlier rejection. The cut-off point for using data from WISE is $n_{WISE} < 5$.

SWcodes : Table

SOTnam	WISnam	SoilVariable	PTRflag	Comments
---	TAWK	TAWK	*	volumetric water content (-10 to - 1500 kPa) --> not used
---	---	---	y	PSCL estimated from PTR-derived sand, silt and clay
ALSA	ALSA	ALSAT	a	exchangeable Aluminum percentage
BSAT	BSAT	BSAT	b	base saturation %
BULK	BULK	BULKDENS	c	bulk density
CECC	CECC	CECCLAY	d	cation exchange capacity of clay fraction (corr. for Org. C)
CECS	CECS	CECSOIL	e	cation exchange capacity
CFRAG	GRAV	GRAVEL	f	coarse fragments
CLPC	CLAY	CLAY	g	clay %
ECEC	ECEC	ECEC	h	effective CEC
ELCO	ECE	ECE	i	electrical conductivity
ESP	ESP	ESP	j	exchangeable Na percentage
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	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 nitogen content

Record: 4 of 20

Figure 4. Conventions for coding the various attributes used in the taxotransfer scheme (see PTRsub and PTRmain in Figure 3)

Table 4. Criteria for defining confidence in the derived data

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

* n_{WISE} is the sample size after the screening procedure (see Figure 3).
[†] The cut-off point for the TTR-approach is n_{WISE} < 5.

In the case of profile *JOPA130*, for example, a number of taxotransfer rules have been applied. This will be illustrated for the depth layer from 40 to 60 cm (D4) with corresponding flags of *b3c3d3j3o3r3* for PTRsub and *A2H1* for PTRmain.

In the case of base saturation (*b3*), for example, the TTR-based parameter estimate has been derived from BSAT data using from 5-14 profiles in WISE – the actual number is shown in the corresponding data sets (see Appendix 1). Since a small letter is shown, the substitution used median data for the actual textural class (M). Otherwise, should the *b* have been capitalized, this would indicate that the substitution is based on the whole set for the corresponding soil unit and depth layer, irrespective of soil texture (i.e., undifferentiated or #). The same coding conventions apply when data for the corresponding major soil grouping had to be used (n_{WISE} <5).

The flags for PTRsub and PTRmain may help focus profile data collection in future updates of national scale primary SOTER databases. These new profiles may then be fed into revised versions of the WISE data set, allowing for renewed application of the TTR-scheme to the updated set of primary SOTER profile data. In principle, this should lead to increasingly reliable soil parameter estimates as reflected by a higher rating for confidence. However, a high confidence rating does not necessarily imply that the soil parameters shown will be representative for the soil unit under consideration. Profile selection for SOTER, as for any other global database, is not probabilistic but based on available data and available expert knowledge. Also several of the soil attributes under consideration are not diagnostic in the Revised Legend (FAO, 1988). In addition, some properties are readily modified by changes in land use or management. For example, aluminum saturation changes upon liming, and soil organic matter content upon changes in tillage.

Details about the most recent version of the taxotransfer-scheme, including the type and number of rules currently in use, are given in Appendix 6.

4. LINKAGE TO GIS

4.1 Map unit heterogeneity

At the small scales under consideration in SOTER – mostly from 1<250 000 to 1:5M – most SOTER units will be compound units. Some of the spatially minor soil units may, however, be of particular relevance. For example, small areas of peat soils in the context of carbon stocks and change. Therefore, it is recommended that users consider all component soils of a SOTER unit in their assessments or model runs. So far, this has seldom been the case (e.g. Falloon *et al.* 1998; Mantel and Van Engelen 1997; Mantel *et al.* 2000), particularly when overlays with other biophysical and socio-economic data are necessary.

Scale is also important when considering auxiliary data layers because conclusions regarding a spatial process depend on the detail and resolution of the data sets employed (Cramer and Fischer 1997; Middelburg *et al.* 1999; Schimel and Panikov 1999).

Apart from technical problems of data format conversion, it is often difficult to combine or compare data sets from different sources at the modelling stage. For instance, atmospheric deposition data will be presented on a grid basis, as opposed to data on agricultural inputs that will have to be generated by administrative region, and soil data that will be presented by landscape (SOTER) unit. Analysis and overlay of such diverse data formats in GIS will lead to unconformities and gaps. Uncertainties related to data and models may be significant at the macro scale and their quantification remains difficult (Burrough 1986; Goodchild 1994). Good practice demands that the modeller provides an evaluation of the confidence in the model, possibly assessing the uncertainties associated with outcome (response) of the model itself (Crossetto *et al.* 2000; Falloon and Smith 2003).

4.2. SOTER unit composition file

To facilitate complex analyses, a summary file has been generated that shows the full composition of each SOTER unit in terms of its dominant soils – as characterized by a regionally representative profile – and their relative extent (Figure 5; Appendix 2). It combines information held in the *SoilComponent* and *Profile* tables of SOTER.

SOTERunitComposition : Table

NEWSUID	SoilMapUnit	SOIL1	PROP1	Profile-ID1	SOIL2	PROP2	Profile-ID2	SOIL3	PROP3	Profile-ID3
JO5	RGc3LPq4CLh5 [5]	RGc	55	JOM1672sy	LPq	30	JOPH046	CLh	15	JOPM149
JO6	CMc3RGc3 [6]	CMc	51	JOPT004	RGc	49	JOPA389syn			
JO7	LPe1RGc4 [7]	LPe	80	JOPN051sy	RGc	20	JOPN038			
JO8	ARc2LPe3 [8]	ARc	60	JOPN002	LPe	40	JOPN051syn			
JO9	CLh2RGc4 [9]	CLh	75	JOPM149	RGc	25	JOM1672syn			
JO10	CMc3CLh4GYh4 [10]	CMc	50	JOPN099sy	CLh	30	JOPM149	GYh	20	JOPD056
JO11	RGc3GYh4LPe4 [11]	RGc	50	JOPN009	GYh	30	JOPD056	LPe	20	JOPN051syn
JO12	CMv4VRk4CLh4 [12]	CMv	35	JOPA070	VRk	35	JOPA078	CLh	30	JOPA082
JO13	CLh3LPe3 [13]	CLh	50	JOPA082	LPe	50	JOPA210syn			
JO14	CMc2LPe4 [14]	CMc	70	JOPS062	LPe	30	JOPB073			
JO15	CMc2LPe3 [15]	CMc	60	JOPD097	LPe	40	JOPD098			
JO16	CLh3CMc4LPe4 [16]	CLh	40	JOPG007	CMc	30	JOPT017	LPe	30	JOPW041
JO17	GYk3SCk4SCn4 [17]	GYk	40	JOPA100	SCk	30	JOPN050	SCn	30	JOPT009

Record: 1 of 27

Figure 5. Characterization of SOTER units in terms of their main component soils – with their regionally representative profiles – and their relative extent

To arrive at a compact map unit code, the relative extent of each soil unit has been expressed in 5 classes: 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. Figure 6 shows an excerpt from the spatial summary file for

Jordan (Batjes *et al.* 2003), as an example. The SOTER unit with country ISO code JO and number 16, for example, is presently coded as CLh3CMc4LPe4. It is composed of 40 per cent haplic Calcisols (CLh), 30 per cent calcaric Cambisols (CMc) and 30 per cent eutric Leptosols (LPe).

4.3 Joining soil parameter estimates with the geographic data

Aggregated information about the SOTER unit composition (Appendix 2) and results of the TTR work (Appendix 3) can linked to the SOTER map using GIS. At the national scale, this can be done via the unique SOTER unit identifier (SUID). In transnational databases, however, linkage will be through the NEWSUID, which is a combination of the country's ISO code plus the SUID code (Appendix 5).

SOTER units – at a scale of 1:250 000 to 1:5M – generally include at least two to three soil components. In the primary database, the associated information is stored in a range of relational databases to enhance the efficiency of data storage and management (Van Engelen and Wen 1995). To assist end-users, a new table has been created that incorporates data held in the primary SOTER database and the present information on soil parameter estimates (Figure 6; Appendix 5). Clearly, this new wealth of information, although needed for modelling, complicates linkage to GIS.

SOTERsummaryFile : Table

NEWSUID	TCID	SCID	Propc	CLAF	PRID	Drain	Layer	TopD	BotDe	CFRAG	SDTO	STPC	CLPC	BULK	TAWC	CECS	BSAT	CECc	PHAQ	T
JO10	1	1	50	CMc	JOPN099syn	W	D1	0	20	8	75	15	10	1.40	12	10	98	56	8.0	
JO10	1	1	50	CMc	JOPN099syn	W	D2	20	40	1	60	15	25	1.44	12	17	100	54	8.1	
JO10	1	1	50	CMc	JOPN099syn	W	D3	40	60	1	60	15	25	1.45	13	18	100	62	8.2	
JO10	1	1	50	CMc	JOPN099syn	W	D4	60	65	1	60	15	25	1.40	16	18	100	61	8.1	
JO10	1	2	30	CLh	JOPM149	M	D1	0	20	8	48	36	16	1.30	9	4	100	25	8.8	
JO10	1	2	30	CLh	JOPM149	M	D4	60	80	28	39	33	28	1.34	9	4	100	60	8.3	
JO10	1	2	30	CLh	JOPM149	M	D5	80	90	28	39	33	28	1.36	8	4	100	42	8.3	
JO10	1	2	30	CLh	JOPM149	M	D2	20	40	4	42	35	23	1.35	9	5	100	60	8.1	
JO10	1	2	30	CLh	JOPM149	M	D3	40	60	15	42	31	27	1.35	9	4	100	56	7.8	
JO10	1	3	20	GYh	JOPD056	W	D1	0	20	0	58	22	20	1.30	9	9	100	76	7.6	
JO10	1	3	20	GYh	JOPD056	W	D2	20	40	0	62	15	23	1.36	12	8	100	44	7.5	
JO10	1	3	20	GYh	JOPD056	W	D3	40	60	1	68	9	23	1.34	11	7	100	42	7.4	
JO10	1	3	20	GYh	JOPD056	W	D4	60	80	3	78	6	16	1.37	9	5	100	36	7.5	
JO10	1	3	20	GYh	JOPD056	W	D5	80	100	10	78	9	13	1.39	14	4	100	30	7.6	
JO11	1	1	50	RGc	JOPN009	S	D1	0	20	1	55	31	14	1.30	27	2	100	82	8.1	
JO11	1	1	50	RGc	JOPN009	S	D2	20	40	1	57	29	14	1.50	14	2	100	58	8.0	
JO11	1	1	50	RGc	JOPN009	S	D3	40	60	16	52	32	16	1.51	8	2	100	74	8.0	
JO11	1	1	50	RGc	JOPN009	S	D4	60	80	60	36	42	22	1.48	3	1	100	59	8.0	
JO11	1	1	50	RGc	JOPN009	S	D5	80	100	52	42	39	19	1.49	16	1	100	34	8.0	
JO11	1	2	30	GYh	JOPD056	W	D5	80	100	10	78	9	13	1.39	14	4	100	30	7.6	
JO11	1	2	30	GYh	JOPD056	W	D2	20	40	0	62	15	23	1.36	12	8	100	44	7.5	
JO11	1	2	30	GYh	JOPD056	W	D1	0	20	0	58	22	20	1.30	9	9	100	76	7.6	
JO11	1	2	30	GYh	JOPD056	W	D3	40	60	1	68	9	23	1.34	11	7	100	42	7.4	
JO11	1	2	30	GYh	JOPD056	W	D4	60	80	3	78	6	16	1.37	9	5	100	36	7.5	
JO11	1	3	20	LPe	JOPN051syn	E	D1	0	16	48	59	15	26	1.35	16	14	90	41	7.2	

Record: 1 of 273

Figure 6. Excerpt of the SOTER summary file for units JO10 and JO11

For visualization and analysis purposes in GIS, it will often be necessary to make an extra selection. For example, in the case of the RothC and Century models, information may be required about the properties of the topsoil – that is layer D1: 0-20 cm – for the dominant soil. In this case, the necessary selection will be for the first Terrain Component (TCID=1), first Soil Component (SCID= 1) and the upper most layer (D1= 1). The corresponding selection is included as a separate table in the secondary database. Further details about the database structure may be found in Appendix 5.

Figure 7 illustrates the overall procedure for linking the various secondary attribute data to the geographical data held in the GIS. For ease of visualization, it only considers the top 20 cm of the spatially dominant (first) soil component of SOTER unit JO19 as an example.

The full SOTER unit composition can best be addressed with tailor-made programmes that vary with the scope of the application (e.g. Batjes 2000a, b).

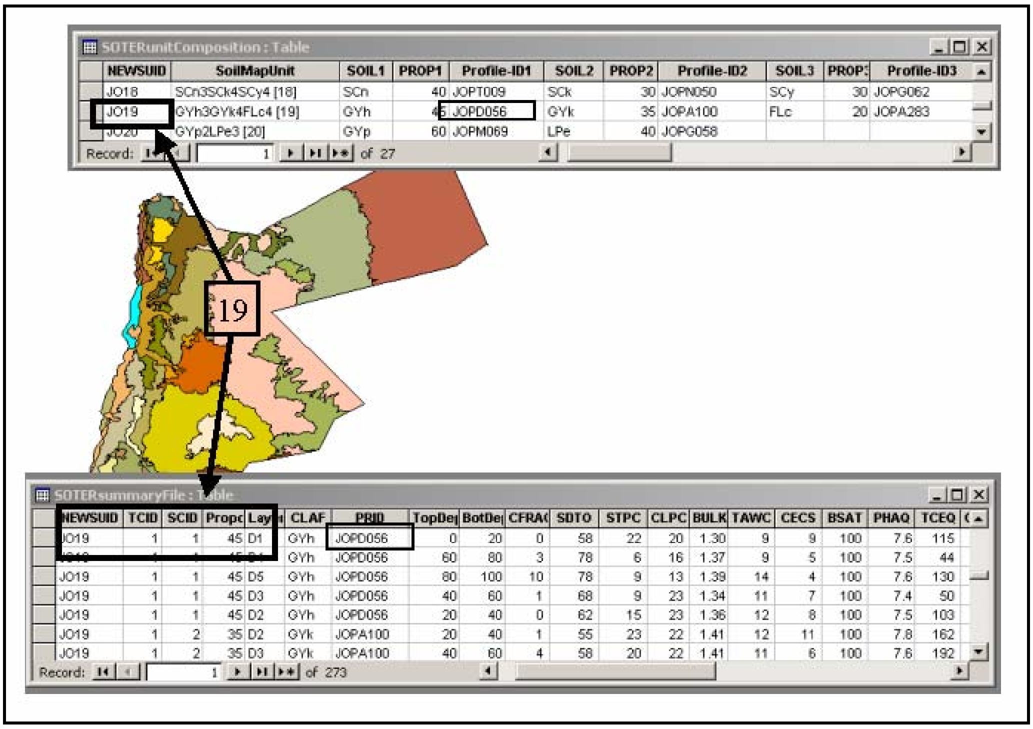

Figure 7. Linking soil parameter estimates for the top 20 cm of the dominant soil (JOPD056) of SOTER unit JO19 with the geographical component of SOTER

4.4 Need for sustained soil data collection

Some interim SOTER products are not yet accompanied by an adequate suite of regionally representative soil profiles. In such cases, the list of soil parameter estimates (Appendix 1) can be linked to the geographical data through the Revised Legend code. This approach is similar to what has been the case for the 1:5 M scale Soil Map of the World (FAO 1995) and its gridded derivatives (Batjes 2002b), albeit using a more elaborate procedure for clustering the available profiles and the Revised Legend.

The need for updating the information on world soil resources in SOTER to supercede the Soil Map of the World, is well recognized (Bouwman *et al.* 1999; ISSS 1987; Nachtergaele 1999). Ideally, questions of necessary accuracy and precision should be determined by the uses to which the data are to be put. However, most national and regional SOTER databases will be based on available data, each with their own characteristics of precision and accuracy. Some of the data received from the collaborating countries will be incomplete. At times, the accuracy of the original information will be questionable. Unfortunately, the systematic collection of new soil data in the field has been aborted in many countries. As a consequence, much of the recent technological progress in soil database development worldwide – at a scale smaller than 1:250 000 – may have to be applied to data that are either outdated or not fully comparable. Exercises such as the present one cannot be regarded other than as stopgaps.

5. CONCLUSIONS

- The specific detail and quality of soil and terrain information available within a country will result in a variable resolution of the primary SOTER data presented.
- A pragmatic approach to the comparability of soil analytical data has been adopted. It is considered appropriate at the exploratory scale of SOTER (< 1:250 000). This correlation must be done more rigorously when more detailed scientific work in considered.
- The soil parameter estimates presented in any secondary SOTER set should be seen as best estimates, based on the currently available profile data held in the corresponding primary SOTER database and the currently available selection of profiles in WISE. Nonetheless, some of these parameter estimates will remain fraught with uncertainty.
- Modellers should familiarize themselves with the assumptions and taxotransfer rules used to develop the set of soil parameter estimates, prior to using these in their models.
- Assessments and model simulation of soil organic carbon stocks and change – like any other environmental study – should consider the full SOTER unit composition, not only the dominant soil component.
- The present procedure may be adopted for use with WRB (1998) when the corresponding nomenclature is standard in new SOTER activities.

ACKNOWLEDGEMENTS

This work contributes to the GEF co-financed project on *Assessment of Soil Organic Carbon Stocks and Change at National Scale* (GFL-2740-02-4381). ISRIC – World Soil Information receives funding from the Netherlands Ministry of Spatial Planning, Housing and the Environment (VROM) Project DGM-200207282.

The constructive comments by Drs David Dent and Eleanor Milne are gratefully acknowledged.

REFERENCES

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

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

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. 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 2000a. Effects of mapped variation in soil conditions on estimates of soil carbon and nitrogen stocks for South America. *Geoderma* **97**, 135-144.

Batjes NH 2000b. Soil vulnerability to diffuse pollution in Central and Eastern Europe (SOVEUR project; *ver. 1.0*). Report 2000/03 (also available on CD No.10, Land and Water Digital Media Series, FAO, 2000), ISRIC, Wageningen.

Batjes NH 2000c. Soil data derived from SOTER for environmental studies in Central and Eastern Europe (SOVEUR Project; *ver. 1.0*). Report 2000/02 (also available on CD No.10, Land and Water Digital Media Series, FAO, 2000), ISRIC, Wageningen.

Batjes NH 2002a. Carbon and nitrogen stocks in the soils of Central and Eastern Europe. *Soil Use and Management* **18**, 324-329.

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

Batjes NH 2002c. Soil parameter estimates for the soil types of the world for use in global and regional modelling (Version 2.0). ISRIC Report 2002/02c [Available online at http://www.isric.org], International Food Policy Research Institute (IFPRI) and International Soil Reference and Information Centre (ISRIC), Wageningen.

Batjes NH 2004. Estimation of soil carbon gains upon improved management within croplands and grasslands of Africa. *Environment, Development and Sustainability* **6**, 133-143.

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 Dijkshoorn JA 1999. Carbon and nitrogen stocks in the soils of the Amazon Region. *Geoderma* **89**, 273-286.

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) [Available on-line via http:\\www.iiasa.ac.at]. Interim Report IR-97-025, FAO/ IIASA/ ISRIC, Laxenburg.

Batjes NH Rawajfih Z and Al-Adamat R 2003. Soil data derived form SOTER for studies of carbon stocks and change in Jordan (ver. 1.0). Report 2003/04, ISRIC - World Soil Information, Wageningen.

Baumgardner MF 1999. Soil databases. In: Sumner ME (Editor), *Handbook of Soil Science*. CRC Press, Boca Ratton, pp. H:1-40.

Baumgardner MF and Oldeman LR 1986. Proceedings of an international workshop on the structure of a digital international soil resources map annex data base. International Society of Soil Science (ISSS), Wageningen (20-24 January 1986).

Bouma J Batjes NH and Groot JJR 1998. Exploring land quality effects on world food supply. *Geoderma* **86**, 43-59.

Bouwman AF Boumans LJM and Batjes NH 2002a. Estimation of global NH_3 emissions from synthetic fertilizers and animal manure applied to arable lands and grasslands. *Global Biogeochemical Cycles* **16**, doi: 10.1029/2000GB001389.

Bouwman AF Boumans LJM and Batjes NH 2002b. Modeling global annual N_2O and NO emissions from fertilized fields. *Global Biogeochemical Cycles* **16**, 1080, doi:10.1029/2001GB00812.

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.

Bullock P 1999. Soil information: uses and needs in Europe. In: Bullock P RJA Jones and L Montanarella (Editors), *Soil resources in Europe*. Office for Official Publications of the European Communities, Luxembourg, pp. 171-182.

Burrough PA 1986. *Principles of Geographical Information Systems*. Oxford, Oxford University press.

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

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.

Crossetto M Tarantola S and Saltelli A 2000. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture, Ecosystems and Environment* **81**, 71-79.

De Oliveira JB and Van den Berg M 1992. Application of the SOTER methodology to a semi-detailed soil survey in the Piracicaba Region (Sao Paolo State, Brazil; scale 1:100,000)). SOTER Report No. 6, International Society of Soil Science (ISSS), Vienna and ISRIC, Wageningen.

Dobos E Bliss N Worstell B Montanarella L Johanssen C and Micheli E 2002. The use of DEM and satellite data for regional scale databases, *Transactions 17th World Congress of Soil Science*. International Union of Soil Sciences (IUSS), Bangkok, pp. 649/1-649/12.

Falloon P and Smith P 2003. Accounting for changes in soil carbon inder the Kyoto Protocol: need for improved long-term datasets to reduce uncertainty in model predictions. *Soil Use and Management* **19**, 265-269.

Falloon PD Smith P Smith JU Szabó J Coleman K and Marshall S 1998. Regional estimates of carbon sequestration potential: linking the Rothamsted carbon model to GIS databases. *Biology and Fertility of Soils* **27**, 236-241.

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, 140 pp.

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

FAO IGAD and Cooperazione Italiana 1998a. The Soil and Terrain Database for northeastern Africa: Crop Production System Zones of the IGAD subregion (scale 1:1 M). Land and Water Digital Media Series No. 2, FAO, Rome.

FAO and IIASA 1999. Soil and physiographic database for North and Central Eurasia at 1:5 million scale. Land and Water Digital Media Series No. 7, Food and Agriculture Organization, 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 ISRIC and UNEP 2003. Soil and terrain database for Southern Africa (1:2 million scale), FAO, Rome.

FAO ISRIC UNEP and CIP 1998b. 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-Unesco 1974. *Soil Map of the World, 1:5,000,000. Vol. 1 - Legend*. United Nations Educational, Scientific, and Cultural Organization, Paris.

FAO-Unesco 1974-1981. *Soil Map of the World, 1:5,000,000. Vol. 1 to 10*. United Nations Educational, Scientific, and Cultural Organization, Paris.

Finke P Hartwich R Dudal R Ibanez J Jamagne M King D Montanarella L and Yassoglou N 1998. Georeferenced soil database for Europe. Manual of procedures (ver. 1.0). Report EUR 18092, European Soil Bureau, Joint Research Centre, Ispra.

Fischer G Van Velthuizen HT Shah M and Nachtergaele FO 2002. Global Agro-ecological Assessment for Agriculture in the 21st Fentury: Methodology and Results. RR-02-02, International Institute for Applied Systems Analysis (IIASA) and Food and Agriculture Organization of the United Nations (FAO), Laxenburg.

Goodchild MF 1994. Sharing imperfect data. In: Singh A (Editor), *Proceedings UNEP and IUFPRO international workshop in cooperation with FAO in developing large environmental databases for sustainable development*. UNEP/GRID, Sioux Falls, pp. 102-110.

Graef F 1999. Evaluation of agricultural potentials in semi-arid SW-Niger: a soil and terrain (NiSOTER) study. Heft 54 Thesis, Universitat Hohenheim, Hohenheim, 217 pp.

ISSS 1987. Proceedings of the Second International Workshop on a Global Soils and Terrain Digital Database (18-22 May 1987, UNEP, Nairobi). In: Van de Weg RF (Editor). ISSS, Wageningen.

King D Le Bas C Nachtergaele FO Van Engelen VWP Eimbeck M Jamagne M Lambert JJ Bridges EM R. H and Montanarella L 2002. A method for generalization of a soil geographical database: the example of the transfer of the European database EUSIS at 1:1 M tot the world SOTER program at 1:5 M, *Transactions 17th World Congress of Soil Science*. International Union of Soil Sciences (IUSS), Bangkok, pp. 495/1-495/9.

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 Brasiliera de Pesquisa Agropecuaria (EMBRAPA), Soil Management Support Services (SMSS) and University of Puerto Rico (UPR),, Rio de Janeiro, pp. 64-70.

KSS 1996. The Soil and Terrain Database for Kenya at scale 1:1,000,000 (ver. 1.0), Kenya Soil Survey, National Agricultural Laboratories, Kenya Agricultural Research Institute, Nairobi.

Madsen BH and Jones RJA 1998. Towards a European soil profile analytical database. In: Heineke HJ et al. (Editors), *Land Information Systems: Developments for planning the sustainable use of land resources*. Office for Offical Publications of the European Community, Luxembourg, pp. 43-49.

Mantel S Samsudin S and Tyrie G 1999. Inventory of site qualities for forest management planning. In: Laumonier Y B King C Legg and K Rennolls (Editors), *Data mangement and modelling using remote sensing and GIS for tropical forest land use inventory*. Rodeo International Publishers, Jakarta, pp. 383-395.

Mantel S and Van Engelen VWP 1997. The impact of land degradation of food productivity: case studies of Uruguay, Argentina, and Kenya. Report 97/01, United Nations Environment Programme (UNEP), National Institute of Public Health and Environment (RIVM), and International Soil Reference and Information Centre (ISRIC), Wageningen.

Mantel S and Van Engelen VWP 1999. Assessment of the impact of water erosion on productivity of maize in Kenya: an integrated modelling approach. *Land Degradation and Development* **10**, 577-592.

Mantel S Van Engelen VWP Molfino JH and Resink JW 2000. Exploring biophysical potential and suitability of wheat cultivation in Uruguay at the national level. *Soil Use and Management* **16**, 270-278.

Mantel S ZhangXuelei and Ganlin Z 2003. Identification of potential for bananana in Hainan Island, China. *Pedosphere* **13**, 147-155.

Middelburg JJ Liss PS Dentener FJ Taminski T Kroeze C Malingreau J-P Növäl 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.

Nachtergaele FO 1999. From the Soil Map of the World to the Digital Global Soil and Terrain Database: 1960-2002. In: Sumner ME (Editor), *Handbook of Soil Science*. CRC Press, Boca Raton, pp. H5-17.

Nachtergaele FO Van Lynden GWJ and Batjes NH 2002. Soil and terrain databases and their applications with special reference to physical soil degradation and soil vulnerability to pollution in Central and eastern Europe. In: Pagliaia M and J R. (Editors), *Advances in GeoEcology 35*. CATENA Verlag GMBH, Reiskrirchen.

Nemes A Schaap MG Lei FJ and Wosten JHM 2003a. Description of the Unsaturated Soil Hydraulic Database - UNSODA Version 2.0. *Journal of Hydrology* **251**, 151-162.

Nemes A Schaap MG and Wosten JHM 2003b. Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection. *Soil Sci Soc Am J* **67**, 1093-1102.

NSMLUP 1996. The Soil and Terrain Database for Jordan at scale 1:500,000 (ver. 1.0; unpublished), National Soil Map and Land Use Project, Department of Afforestation and Forests, Ministry of Agriculture, Amman.

Paustian K Levine E Post WM and Ryzhova IM 1997. The use of models to integrate information and understanding of soil C at the regional scale. *Geoderma* **79**, 227-260.

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.

Pleijsier LK (Editor) 1986. The LABEX programme and its results so far. Proceedings of an International Workshop on the Laboratory Methods and Data Exchange Programme (LABEX), Technical Paper 13. ISRIC, Wageningen, 6-15 pp.

Savin IY Ovechkin S, V. and Aleksandrova EV 1997. The WOFOST simulation of crop growth and its application for the analysis of land resources. *Eurasian Soil Science* **30**, 758-765.

Schimel DS and Panikov NS 1999. Simulation models of terrestrial trace gas fluxes at soil microsites to global scales. In: Bouwman AF (Editor), *Approaches to Scaling of Trace Gas Fluxes in Ecosystems*. Elsevier, Amsterdam, pp. 187-202.

Scholes RJ Skole D and Ingram JS 1995. A global database of soil properties: proposal for implementation. IGBP-DIS Working Paper 10, International Geosphere Biosphere Program, Data & Information System, Paris.

Snedecor GW and Cochran WG 1980. *Statistical Methods (7th. ed)*. The Iowa State University Press, Iowa, 507 pp.

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

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

Stolbovoi V 2001. Soils of Russia: Correlated with the Revised Legend of the FAO Soil Map of the World and World Reference Base for Soil Resources. Report RR-00-13, International Institute for Applied Systems Analysis, Laxenburg.

Tempel P 2002. SOTER - Global and National Soil and Terrain Digital Databases: Database structure (Ver 3.1), ISRIC - World Soil Information, Wageningen.

Tomasella J and Hodnett MG 1997. Estimating unsaturated hydraulic conductivity of Brazilian soils using soil-water retention data. *Soil Science* **162**, 703-712.

Tomasella J and Hodnett MG 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Science* **163**, 190-202.

Torn MS Trumbmore SE Chadwick OA Vitousek PM and Hendricks DM 1997. Mineral control of soil organic carbon storage and turnover. *Nature* **389**, 170-173.

Van den Berg M Klamt E Van Reeuwijk LP and Sombroek WG 1997. Pedotransfer functions for the estimation of moisture retention of Ferralsols and related soils. *Geoderma* **78**, 161-180.

Van Engelen VWP 1999. SOTER: The World Soils and Terrain Database. In: Sumner ME (Editor), *Handbook of Soil Science*. CRC Press, Boca Raton, pp. H19-28.

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

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

Varallyay G Szabo J Pasztor L and Micheli E 1994. SOTER (Soil and Terrain Digital Databases) 1:500,000 and its application in Hungary. *Agrokemia es Talajtan* **43**, 87-108 (in Hungarian).

Wattel-Koekkoek EJW 2002. Clay-associated organic matter in kaolinitic and smectitic soils. PhD No. 3136 Thesis, Wageningen University, Wageningen, 120 pp.

Wösten JHM Lilly A Nemes A and Le Bas C 1998. Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning. Report 156, DLO-Staring Centre, Wageningen.

WRB 1998. World Reference Base for Soil Resources. World Soil Resources Report 84, ISSS, ISRIC, and FAO, Rome.

APPENDICES

Appendix 1: Statistical output files

These tables show results of the statistical analyses of profiles available in WISE, by Revised Legend (FAO 1988) unit, soil attribute, fixed depth layer and soil textural class. The corresponding information is stored in 8 tables – in view of their length the tables will be made available only in digital format (see: www.isric.org)[2].

- *WISstat9_a_ACGY*: This table holds statistical data by *major soil groupings* (FAO 1988), ranging from Acrisols (AC) to Gypsisols (GY), considering all five soil textural classes – these are: Coarse (C): less than 15 per cent clay and more than 70 per cent sand; Medium (M): less than 35 per cent clay and less than 70 per cent sand; more than 15 percent clay if the sand content exceeds 70 percent; Medium Fine (Z): less than 35 per cent clay and less than 15 per cent sand; Fine (F): between 35 and 60 per cent clay; Very Fine (V): more than 60 per cent clay. Where there is no measured particle size analyses for a given layer – for example for some Histosols - this has been flagged by '-'
- *WISSTat9_a_HSVR*: Similar to above but for *major soil groupings* ranging from Histosols (HS) to Vertisols (VR)
- *WISstat9_b_ACGY*: This table holds statistical data by *soil unit* (FAO 1988), ranging from Acrisols (AC) to Gypsisols (GY), considering all five soil textural classes
- *WISstat9_b_HSVR:* Similar to above but for *soil units* ranging from Histosols to Vertisols
- *WISstat9_c_ACGY*: This table holds statistical data by *major soil groupings*, ranging from Acrisols (AC) to Gypsisols (GY), irrespective of soil texture. This has been flagged as class #, for undifferentiated, which comprises soil textural classes C, M, Z, F and V as well as -

[2] After official termination of the GEFSOC project (July 2005).

- *WISstat9_c_HSVR*: Similar to above but for *major soil groupings* ranging from Histosols (HS) to Vertisols (VR)
- *WISstat9_d_ACGY*: This table holds statistical data by *soil unit*, ranging from Acrisols (AC) to Gypsisols (GY), irrespective of soil texture (#)
- *WISstat9_d_HSVR*: Similar to above, but for *soil units* ranging from Histosols (HS) to Vertisols (VR).

Structure of statistical output tables*

Name	Type	Size	Description
Short_Id	Text	25	Code comprising abbreviation for FAO major soil group, attribute, depth layer, and soil textural class (e.g., ACBSATD1M)
Num0	Integer	2	Number of observations (before outlier rejection)
Num	Integer	2	Number of observations (after outlier rejection)
Mean	Single	4	Mean
STD	Single	4	Standard deviation
CV	Single	4	Coefficient of variation
Median	Single	4	Median
MAD	Single	4	Median of absolute deviations
Min	Single	4	Minimum
Max	Single	4	Maximum
Var	Single	4	Variance
Fao_90	Text	3	FAO Revised Legend code (this field is intentionally left blank)

*Notes:

1) Applies to tables named _WISstat9_a_ACGY, _WISstat9_a_HSGY, _WISstat9_c_ACGY and _WISstat9_c_HSVR.
2) The structure of tables _WISstat9_b_ACGY, _WISstat9_b_HSGY, _WISstat9_d_ACGY and _WISstat9_d_HSVR, is similar to the one above except that the first field is called Long_ID. This field differs from Short_ID in that the first three letters refer to the FAO soil unit code (e.g. *ACfBSATD1M* for the sample set that relates to base saturation data (BSAT) for ferric Acrisols (ACf) that have medium (M) texture and belong to layer D1, from 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 use median (MED) values from the corresponding tables when Num > 5 (see n_{WISE} in text).

Appendix 2: SOTER unit composition file

This summary table gives the full composition of each SOTER unit in terms of its main soil units (FAO, 1988), their relative extent, and the identifier for the corresponding representative profile. It contains information aggregated from a number of primary SOTER tables, *viz.* SoilComponent and Profile (see Van Engelen and Wen 1995).

This summary file can be linked to the SOTER geographic data in a GIS through the unique SOTER unit code – NEWSUID, a combination of the fields for ISO and SUID – and linked to the tables holding the soil parameter estimates using the unique profile identified codes (PRID, see Appendix 3).

Structure of table SOTERunitComposition.

Name	Type	Size	Description
ISOC	Text	2	ISO-3166 country code (1994)
SUID	Integer	2	The identification code of a SOTER unit on the map and in the database
NEWSUID	Text	10	Globally unique SOTER code, comprising fields ISOC plus SUID
SOIL1	Text	3	Characterization of the first (main) according to the Revised Legend (FAO, 1988)
PROP1	Integer	2	Proportion, as a percentage, that the main soil occupies within the SOTER unit
PRID1	Text	15	Unique code for the corresponding representative soil profile (as selected by the national soil experts)
SOIL2	Text	3	As above, but for the next soil component
PROP2	Integer	2	As above
PRID2	Text	15	As above
SOIL3	Text	3	As above but for the next soil component
PROP3	Integer	2	As above
PRID3	Text	15	As above
SOIL4	Text	3	As above but for the next soil component
PROP4	Integer	2	As above
PRID4	Text	15	As above
SOIL5	Text	3	As above but for the next soil component
PROP5	Integer	2	As above
PRID5	Text	15	As above
SOIL6	Text	3	As above but for the next soil component
PROP6	Integer	2	As above
PRID6	Text	15	As above

(cont.)

SOIL7	Text	3	As above but for the next soil component
PROP7	Integer	2	As above
PRID7	Text	15	As above
SOIL8	Text	3	As above, but for the next soil component
PROP8	Integer	2	As above
PRID8	Text	15	As above
SOIL9	Text	3	As above, but for the next soil component
PROP9	Integer	2	As above
PRID9	Text	15	As above
SOIL10	Text	3	As above, but for the next soil component
PROP10	Integer	2	As above
PRID10	Text	15	As above

Note: Generally, not all 10 available fields for $SOIL_i$ will be filled in SOTER.

Appendix 3: Taxotransfer rule-based soil parameter estimates

This table lists soil parameters estimates for all representative profiles considered in a given SOTER database. This information can be linked to the geographical component of the SOTER database, in a GIS, through the unique profile code (PRID, see Appendix 2).

The table should be consulted in conjunction with table SOTERflagPTRrules which documents the type of taxotransfer rules that have been applied (see Appendix 4).

Structure of table SOTERDerivedSoilParameters

Name	Type	Size	Description
CLAF	Text	3	Revised Legend FAO (1988) code
PRID	Text	15	Profile ID (as documented in table: SOTERunitComposition)
Drain	Text	2	FAO soil drainage class
Layer	Text	8	Code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
TopDep	Integer	4	depth of top of layer (cm)
BotDep	Integer	4	depth of bottom of layer (cm)
CFRAG	Integer	2	coarse fragments (> 2mm)

(cont.)

SDTO	Integer	2	sand (mass %)
STPC	Integer	2	silt (mass %)
CLPC	Integer	2	clay (mass %)
PSCL	Text	1	FAO soil texture class (see Figure 2)
BULK	Single	4	bulk density (kg dm^{-3})
TAWC	Integer	2	available water capacity (cm m^{-1}; -33 to -1500 kPa; USDA standards; not yet corrected for coarse fragments)
CECS	Single	4	cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	2	base saturation as percentage of CECsoil
CECc	Single	4	CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	4	pH measured in water
TCEQ	Single	4	total carbonate equivalent (g kg^{-1})
GYPS	Single	4	gypsum content (g kg^{-1})
ELCO	Single	4	electrical conductivity (dS m^{-1})
TOTC	Single	4	organic carbon content (g kg^{-1})
TOTN	Single	4	total nitrogen (g kg^{-1})
ECEC	Single	4	effective CEC ($cmol_c$ kg^{-1})

Notes: These are depth-weighted values, per 20 cm layer.

Appendix 4: Flagging taxotransfer rules

The type of taxotransfer that has been used when creating the table SOTERDerivedSoilParameters (see Appendix 3) is documented in table SOTERflagPTRrules. Further details on coding conventions may be found in the Section 3.5.

Structure of table SOTERflagPTRrules

Name	Type	Size	Description
CLAF	Text	3	Revised Legend (FAO, 1988) code
PRID	Text	15	Unique identifier for representative profile
Newtopdep	Integer	2	Depth of top of layer (cm)
Newbotdep	Integer	2	Depth of bottom of layer (cm)
PTRsub	Text	50	Codes showing the type of taxotransfer rule used (based on data for soil *units*; see text)
PTRmain	Text	50	Codes showing the type of taxotransfer rule used (based on data for *major soil groups*; see text)
PTRfinal	Text	25	Additional flags

Appendix 5: SOTER summary file

Interpretations of a SOTER database, as schematically depicted in Figure 1, and the currently created secondary data sets require a good understanding of relational database handling systems and a sound understanding of the SOTER database structure. This may be an obstacle to end-users with limited programming expertise. Therefore, to facilitate access to the derived data, a SOTER summary file has been generated.

This file incorporates information held in the Terrain, TerrainComponent, Soil Component tables of the primary SOTER database as well as information held in the table with the soil parameter estimates. The resulting table can be directly linked to the soil geographical data using the unique SOTER unit codes (NEWSUID). The structure of this table, called SOTERsummaryTable , is described below.

Information on landform, lithology and slope can be derived from the primary SOTER database and so not been considered in the present study.

Structure of SOTERsummaryFile

Name	Type	Size	Description
ISOC	Text	2	ISO-3166 country code (1994)
SUID	Integer	2	The identification code of a SOTER unit on the map and in the database
NEWSUID	Text	10	Globally unique SOTER code, comprising fields ISOC plus SUID
TCID	Integer	1	Number of terrain component within given SOTER unit
SCID	Integer	1	Number of soil component within given terrain component and SOTER unit
PROP	Integer	3	Proportion of soil component within the given SOTER unit
CLAF	Text	3	Revised Legend FAO (1988) code
PRID	Text	15	Profile ID (as documented in table SOTERunitComposition)
Drain	Text	2	FAO soil drainage class
Layer	Text	8	Code for depth layer (from D1 to D5; e.g., D1 is from 0 to 20 cm)
TopDep	Integer	4	Depth of top of layer (cm)
BotDep	Integer	4	Depth of bottom of layer (cm)

(cont.)

CFRAG	Integer	2	Coarse fragments (> 2mm)
SDTO	Integer	2	Sand (weight %)
STPC	Integer	2	Silt (weight %)
CLPC	Integer	2	Clay (weight %)
PSCL	Text	1	FAO soil texture class (see Figure 2)
BULK	Single	4	Bulk density (kg dm^{-3})
TAWC	Integer	2	Available water capacity cm m^{-1}; -33 to -1500 kPa, USDA standards; not yet corrected for coarse fragments)
CECS	Single	4	Cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	2	Base saturation as percentage of CECsoil
CECc	Single	4	CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	4	pH measured in water
TCEQ	Single	4	Total carbonate equivalent (g kg^{-1})
GYPS	Single	4	Gypsum content (g kg^{-1})
ELCO	Single	4	Electrical conductivity (dS m^{-1})
TOTC	Single	4	Organic carbon content (g kg^{-1})
TOTN	Single	4	Total nitrogen (g kg^{-1})
ECEC	Single	4	Effective CEC ($cmol_c$ kg^{-1})

Notes:
1) These are depth-weighted values, per 20 cm layer.
2) Table SOTERsummaryFile_T1S1D1 only holds information for the main TCID (1), main SCID (1) and first layer (D1) as derived from the above table.

Appendix 6: Latest revisions of the TTR-approach

The TTR-approach was last revised in February 2005. It was based on statistical analyses of some 9600 profiles held in the WISE database (Batjes 2002, 2003), corresponding with over 43,000 horizons. Analyses of these data, so far, permitted to define 38,683 rules in total. These include 28,167 rules for TTRsub and 10,516 rules for TTRmain (Table 5). The cut-off point for defining and applying any TTR was $n_{WISE} < 5$.

The overall assumption in applying the TTR-scheme is 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, the confidence in soil parameter estimates listed under TTRsub should be higher than for those listed under TTRmain (see Section 3.5).

Table 5. Number and type of taxotransfer rules

Taxotransfer rule	Textural class	Number of rules
TTRsub	C, M, F, Z and V	18,510
	Undifferentiated	9,657
TTRmain	C, M, F, Z and V	7,987
	Undifferentiated	2,529

Note: For details about codes used for soil textural classes, see Figure 2. For each parameter (see text), TTR-rules are defined per 20 cm depth layer, coded from D1 to D5, up to 1 m where applicable. The textural class of organic c soils (Histosols) has been coded 'O'.

In spite of the large number of TTR-rules already in use, it has been necessary to introduce a number of expert-based rules. Such expert-rules take into consideration whether certain combinations of soil parameter estimates were considered pedo-chemically feasible or relevant for a given soil unit. Twenty eight expert-rules were used in this revision (Table 6). For example, the aluminium saturation percentage cannot be more than zero in soils with a high pH or, alternatively, a high base saturation is unlikely to occur at low pH values. The scheme of expert-rules was applied after the taxotransfer scheme, as a 'final check' of the TTR-derived data.

Various sources of uncertainty remain in the soil geographic and attribute data that have been used to compile the various SOTER databases; these should be corrected gradually in revised versions of the corresponding primary data sets. After each significant update of the primary data, revised sets of soil parameter estimates may be prepared for the regions concerned, gradually leading to improved estimates for the soil parameters under consideration.

Details about the type of TTR and expert-rules that have been used for each profile, layer, and soil attribute may be found in table *SOTERflagTTRrules* (Appendix 3).

Note: The 'February 2005' version of the TTR-approach has been used to generate revised secondary SOTER sets for the four GEFSOC study areas.

Table 6. Revised conventions for coding taxotransfer and expert rules

Type of rule	Soil Variable	Flag	Description
Taxotransfer:			
TTR-ALSA	ALSAT	A	exchangeable Aluminium percentage (% of ECEC)
TTR-BSAT	BSAT	B	base saturation (% of CECs)
TTR-BULK	BULKDENS	C	bulk density
TTR-CECC	CECCLAY	D	cation exchange capacity of clay fraction (corr. for org. C)
TTR-CECS	CECSOIL	E	cation exchange capacity
TTR-CLAY	CLAY	G	clay %
TTR-ECEC	ECEC	H	Effective CEC
TTR-ELCO	ECE	I	electrical conductivity
TTR-ESP	ESP	J	exchangeable Na percentage (% of CECs)
TTR-GRAV	GRAVEL	F	coarse fragments
TTR-GYPS	GYPSUM	K	Gypsum content
TTR-PHAQ	PHH2O	L	pH in water
TTR-SAND	SAND	M	sand %
TTR-SILT	SILT	N	silt %
TTR-TAWC	TAWC	O	volumetric water content (-33 to - 1500 kPa)
TTR-TCEQ	CACO3	P	carbonate content
TTR-TOTC	ORGC	Q	Organic carbon content
TTR-TOTN	TOTN	R	total nitrogen content
Expert-rule:			
XR1-Alsa	ALSAT	1	Expert rules for ALSAT *vs* soil pH (5 rules)
XR2-Bsat	BSAT	2	Expert rules for BSAT *vs* soil pH (6)
XR3-Elco	ECE	3	Expert rules for ELCO *vs* pH (1)
XR4-Gyps	GYPS	4	Expert rules for GYPSUM *vs* pH (1)
XR5-CaCo	TCEQ	5	Expert rules for CACO3 *vs* pH (5)
XR6-CECc	CECc	6	Expert rules for CECclay (2)
XR7-Hist	HISTO	7	Expert rules for organic soils (for Histosols; 1)
XR8-LAC	LAC	8	Expert rules for CECclay (for Low Activity (LAC) soils; 6)
XR9-ECEC	ECEC	9	Expert rules for effective CEC (for LAC and Andosols; 1)

Note: Codes for taxotransfer-rules start with TTR, while expert-based rules begin with the letters XR. Several subdivisions are possible for each expert-rule; these have been coded with numerals (e.g., 6a, 6b for rules determining the parameter estimates for CECclay). The number of conditions defined so far, for each expert-rule, is shown in brackets.