Report 2005/05

WISE-based soil parameter estimates for Northeastern Africa

(Version 1.0)

Niels H Batjes

(March 2005)

ISRIC World Soil Information

Disclaimer:

Correct citation:

Batjes NH 2005. WISE-based soil parameter estimates for Northeastern Africa. Report 2005/05. ISRIC – World Soil Information, Wageningen.

Inquiries:

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

Contents

List of Tables

List of Figures

ABSTRACT

This report presents a harmonized set of soil parameter estimates for Northeastern Africa based on soil geographic data of the 1:1M Soil and Terrain Database for Northeast Africa (FAO *et al.* 1998b) and soil profile data held in the ISRIC-WISE soil profile database.

The land surface of Northeastern Africa has been characterized using 1469 unique SOTER units, corresponding with some 7700 polygons. Since the original data set for the region did not include any measured soil profile data, all soil parameter estimates presented here were derived from taxotransfer rules (TTRs). The soil data that underlie the TTR-scheme were derived from analyses of some 9600 soil profiles held in the ISRIC-WISE database (Batjes 2003).

Parameter estimates are presented by soil unit for fixed depth intervals of 0.2 m to 1 m depth for: organic carbon, total nitrogen, pH(H_2O), CEC_{soil}, CEC_{clay}, base saturation, effective CEC, aluminium saturation, $CaCO_3$ content, gypsum content, exchangeable sodium percentage (ESP), electrical conductivity of saturated paste (ECe), bulk density, content of sand, silt and clay, content of coarse fragments (> 2 mm), and available water capacity (-33 to -1500 kPa). These attributes have been identified as being useful for agro-ecological zoning, land evaluation, crop growth simulation, modelling of soil carbon stocks and change, and analyses of global environmental change.

The current parameter estimates should be seen as best estimates, pending the availability of measured soil data for the region in SOTER. The uncertainty attached to some of these parameter estimates, however, can be high. All taxotransfer rules have been flagged to provide an indication of the possible confidence in the derived data.

Results are presented as summary files and can be linked to the 1:1M scale SOTER_IGAD map in a GIS, through the unique SOTER-unit code.

The secondary data set is considered appropriate for exploratory studies at the continental scale. Correlation of soil analytical data should be done more rigorously when more detailed scientific work is considered.

Keywords: soil parameter estimates, Northeastern Africa, environmental modelling, WISE database, SOTER database, secondary data set

1. INTRODUCTION

ISRIC and FAO and UNEP, under the aegis of the International Union of Soil, Sciences (IUSS), are updating the information on world soil resources in the World Soils and Terrain Digital Databases (SOTER) project. Once global coverage has been attained, SOTER is to supersede the 1:5M Soil Map of the World (FAO 1995; FAO-Unesco 1971-1981)

The SOTER methodology has been applied at scales ranging from 1:250 000 to 1:5M. Continental scale SOTER databases are available for Latin America and the Caribbean (Dijkshoorn *et al.* 2005; FAO *et al.* 1998b), Central and Eastern Europe (FAO and ISRIC 2000) and Southern Africa (FAO and ISRIC 2003). Some interim, SOTER-like products have also been published for North and Central Asia (FAO and IIASA 1999) and Northeastern Africa (FAO *et al.* 1998a).

The product for Northeastern Africa was compiled before the SOTER methodology was fully developed (van Engelen and Wen 1995); as such it fell short of a full SOTER product. Apparently, there were no resources at the time to undertake a full characterization with analyzed soil profiles of the large region concerned (FAO *et al.* 1998a). Subsequently, FAO started to collate profiles for the region under consideration albeit not in SOTER format. In that exercise, Dondeyne and Deckers (2003) drew heavily on existing soil databases such as WISE (Batjes 2002a) and the SOTER set for Kenya (KSS 1995), which yielded some 70% of the 'new' soil profile data. Dondeyne and Deckers (2003) further enumerated tasks that should be undertaken before the present soil data set for Northeastern Africa can be upgraded to full SOTER format. Consequently, the present study was based on the original map of the Soils of Northeastern Africa (FAO *et al.* 1998a). Therefore all soil parameter estimates presented here were derived through taxotransfer and expert rules (Batjes 2003).

Typically, primary SOTER databases are composed of two main elements: a geographic and an attribute data component (van Engelen *et al.* 2005). 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.

This report[1] discusses the application of the TTR-scheme to the spatial data set for Northeastern Africa. The study has been carried out in support of FAO-ISRIC project PR-29651 aimed at preparing a harmonized Global Soil Resources Database, with a resolution of 5 arcminutes by 5 arcminutes, using recently produced Soil and Terrain (SOTER) databases.

Chapter 2 describes the materials and methods with special focus on the procedure for preparing the secondary SOTER sets. Results are discussed in Chapter 3, while concluding remarks are drawn in Chapter 4. The structure of the various output tables is documented in the Appendices, which also include a brief description of the contents of the secondary data file for Northeastern Africa (Appendix 5).

2. MATERIALS AND METHODS

2.1 Source of data

The soil geographical data for Northeastern Africa (FAO *et al.* 1998a) provided the basis for this study. The study area includes Egypt, Sudan, Ethiopia, Djibouti, Somalia, Uganda, Rwanda, Burundi, and Northern Tanzania. Soil parameter estimates for Kenya, for which a full scale SOTER database now exists (KSS 1995), may be found elsewhere (Batjes and Gicheru 2004).

Although the soil map for Northeastern Africa has a generalized scale of 1:1M, the detail and quality of primary information available within the various country can vary greatly (FAO *et al.* 1998a).

[1] Having the same scope, all reports describing secondary SOTER data sets have a similar structure.

In view of differences in approaches used for storing the original soil-geographic data, this information was first converted to the uniform SOTER format (van Engelen and Wen 1995). During this conversion, all SOTER units were given unique codes (NEWSUID) comprised of the country ISO-code and the original reference number (IDNUMBER) used by FAO et al. (1998a), for example "SD1234".

All miscellaneous units, such as water bodies and rock outcrops, were recoded using the naming conventions of the revised SOTER Procedures Manual (van Engelen *et al.* 2005).

2.2 SOTER 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 *et al.* 2005). The approach resembles physiographic or land systems mapping. The collated materials are stored in a SOTER database linked to GIS, permitting a wide range of environmental applications (e.g., Batjes 2004; Batjes and Dijkshoorn 1999; Falloon *et al.* 1998; Mantel *et al.* 2000; Nachtergaele *et al.* 2002) }. The SOTER methodology was mainly applied at scales ranging from 1:250 000 to 1:5M.

Each SOTER database is comprised of two main elements, a geographical component and an attribute data component (Figure 1). The *geographical database* holds information on the location, extent and topology of each SOTER unit. The *attribute database* describes the characteristics of the spatial unit and includes both area data and point data. A geographical information system (GIS) is used to manage the geographic data, while the attribute data are handled in a relational database management system (RDBMS).

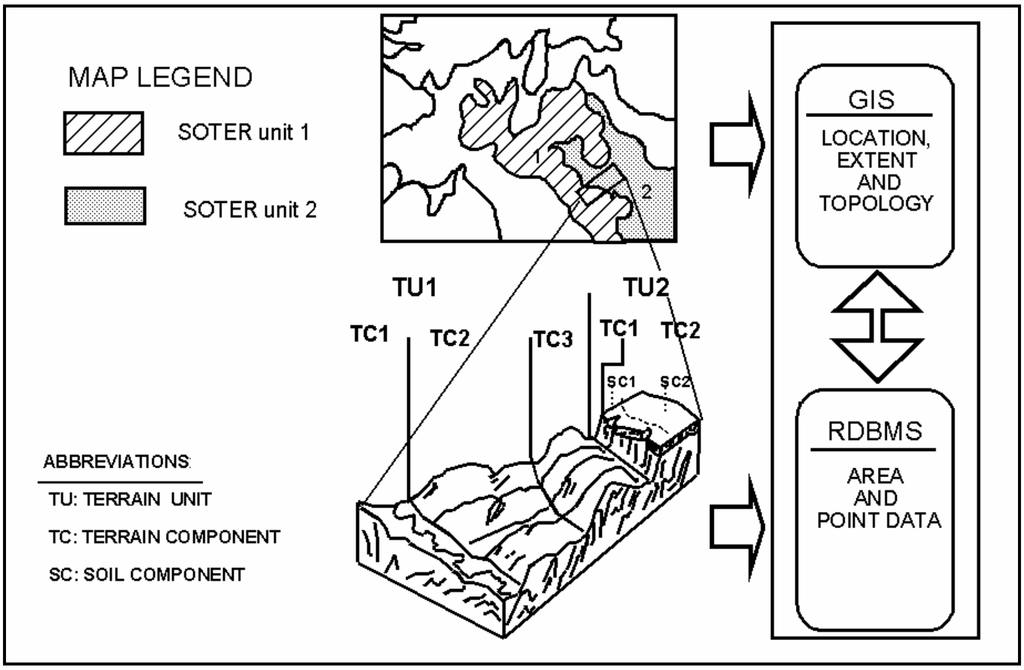

Figure 1. Schematic representation of two SOTER units and their terrain and soil components

Each SOTER unit in the geographic database 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) (see Appendix 4).

In so-called full SOTER databases, each soil component within a SOTER unit is described by a profile (PRID), identified by the national soil experts as being regionally representative (e.g. FAO and ISRIC 2003). This selection is based on purposive sampling (Webster and Oliver 1990). Profiles are characterised according to the Revised Legend of FAO (1988) and World Reference Base for Soil Resources (WRB 1998). Representative profiles are selected from available soil survey reports, as the SOTER program does not involve new ground surveys. Batjes (1999) reviewed issues of data acquisition, quality control and sharing in the context of SOTER projects. Various sources of uncertainty remain in the available soil geographic and attribute data, and these should be gradually corrected in revised versions of the primary data sets.

2.3 Preparation of secondary SOTER data sets

2.3.1 List of soil parameters

Special attention has been paid to the key attributes (Table 1) commonly required in studies of agro-ecological zoning, food productivity, soil gaseous emissions/sinks and environmental change (see Batjes *et al.* 1997; Bouwman *et al.* 2002; Cramer and Fischer 1997; Fischer *et al.* 2002; Scholes *et al.* 1995).

Table 1. 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) † ‡
Aluminium saturation (as % of ECEC) ‡
$CaCO_3$ content
Gypsum content
Exchangeable sodium percentage (ESP) ‡
Electrical conductivity of saturated paste (ECe)
Bulk density
Coarse fragments (> 2mm, volume %)
Sand (mass %)
Silt (mass %)
Clay (mass %)
Available water capacity (AWC; cm m^{-1}, from -33 to -1500 kPa) ‡ □

‡ Calculated from other measured soil properties.
† ECEC is defined as exchangeable (Ca^{++}+ Mg^{++}+ K^{+}+ Na^{+}) + exchangeable (H^{+}+ Al^{+++}) (van Reeuwijk 2002).
• CEC_{clay} was calculated from CEC_{soil} by assuming a mean contribution of 350 $cmol_{cmm}$ 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); typically, these estimates are smaller than those reported for the -10 to -1500 kPa range commonly used by FAO (Doorenbos and Kassam 1978).

Table 1 does not consider soil hydraulic properties. Although essential for many simulation studies, these properties are seldom measured during soil surveys. As a result, the corresponding records are lacking 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.* 2003; Wösten *et al.* 1998) but similar work for tropical soils has just begun (Tomasella and Hodnett 1997, 1998; van den Berg *et al.* 1997).

2.3.2 Procedure for filling gaps in the measured data

The standardized procedure for filling gaps in key measured data in primary SOTER data sets includes three main stages (Batjes 2003):

a) Collating additional measured soil data where these exist, in uniform SOTER format;

b) Using expert estimates and common sense to fill selected gaps in a secondary data set;

c) Using taxotransfer rule (TTR) derived soil parameter estimates for similar FAO soil units, based on some 9600 profiles held in the global WISE profile database (Batjes 2002b), complemented with the application of expert rules.

The desirability of the above stages decreases from highest (a) to lowest (c). However, within the limited resources available for the current study there was only scope for applying the TTR-based approach (c).

Virtual profiles were defined for all SOTER units. For example, profile BI-ACg represents a gleyic Acrisol (Acg) in Burundi (BI). General information about the likely drainage class and soil depth was derived for the information embedded in the taxonomic names of the Revised Legend (FAO 1988). In total, 317 virtual profiles have been defined for the region.

3. RESULTS AND DISCUSSION

3.1 General

Northeastern Africa has been described using some 1469 unique SOTER units. At the small scale under consideration, most SOTER units are compound. SOTER units in the study region are comprised of up to 5 soil components, with an average of 1.9. End-users should therefore consider all component soil units of a SOTER unit in their assessments or model runs.

Ultimately, the type of research purpose will determine which parameter estimates or single value maps are of importance in a special case. The full map unit composition can best be addressed with tailor made programs depending on the scope of the application.

3.2 SOTER unit composition

A table – *sensu* MS Access® databases – has been generated that shows the full composition of each SOTER unit in terms of its dominant soils – each one characterized by a virtual profile – and their relative extent. The relative extent of each soil unit has been expressed in 5 classes to arrive at a compact map unit code: 1 – from 80 to 100 per cent; 2 – from 60 to 80 per cent; 3 – from 40 to 60 percent; 4 – from 20 to 40 per cent, and 5 – less than 20 percent.

3.3 Soil Parameter estimates

The depth-weighted primary and TTR-derived data, by layer, for the 18 soil properties under consideration (Table 1) have been stored in a secondary SOTER data set; the cut-off point for applying any TTR was $n_{WISE} < 5$. Appendix 2 shows the structure of the corresponding file.

The type of TTR used, if any, has been flagged by profile and depth layer in table *SOTERflagTTRrules* (Appendix 3). The field TTRsub indicates that the data substitution for a given attribute, in the secondary SOTER set, was 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 were flagged under TTRmain (see Batjes 2003).

Each flag consists of a sequence of letters followed by a numeral (see under TTRsub and TTRmain in Figure 2). The letters indicate soil attributes for which a TTR has been applied, using coding conventions listed in Table 4. The number code reflects the size of the sample population in WISE, after outlier rejection, on which the statistical analyses were based (Table 2).

Table 2. 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 5)
† The cut-off point in the TTR-approach is $n_{WISE} < 5$

When a small letter is used, the substitution considered median data for the corresponding textural class (for example, Fine and $n_{WISE} > 5$). Otherwise, when a capital is used, this indicates 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 'u'). The same coding conventions apply for TTRmain. This is depicted schematically for the upper 0 to 20 cm of a hypothetical profile from country XX (XXhyp04) in Figure 2.

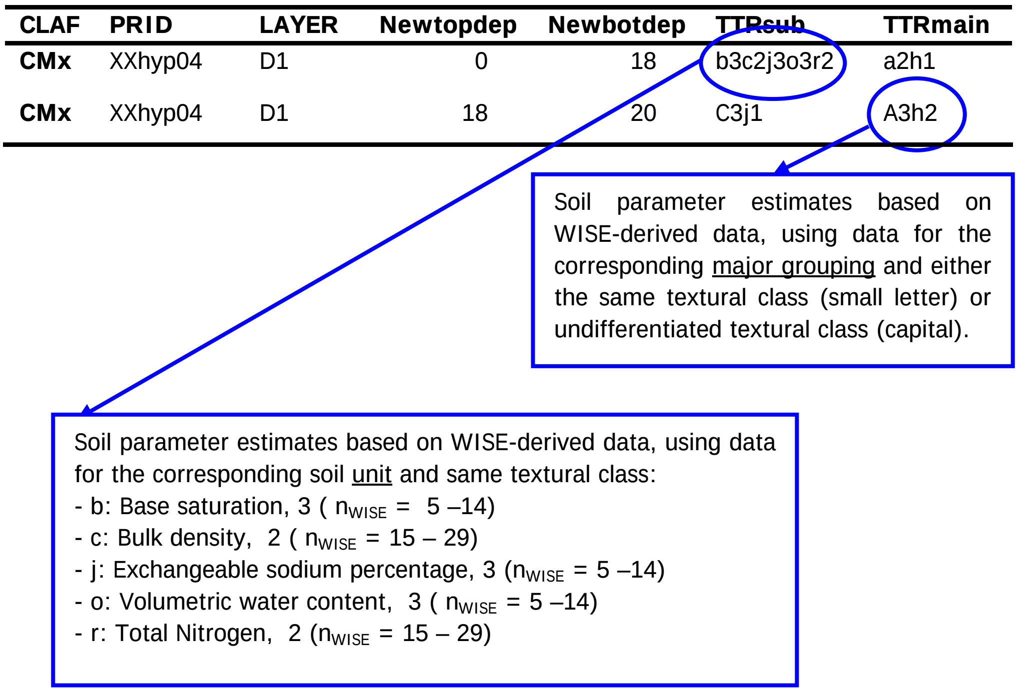

Figure 2. Schematic representation of application of taxotransfer scheme

The present version of the TTR-scheme was based on statistical analyses of some 9600 profiles held in the WISE database (Batjes 2002a, 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 3).

Table 3. Number and type of taxotransfer rules

Taxotransfer rule	Textural class	Number of rules
TTRsub	C, M, F, Z and V	18510
	Undifferentiated	9657
TTRmain	C, M, F, Z and V	7987
	Undifferentiated	2529

Note: For details about codes used for soil textural classes, see Appendix 6. For each parameter (Table 1), TTR-rules are defined per 20 cm depth layer, coded from D1 to D5, up to 1 m where applicable.

The overall assumption in applying the TTR-scheme was 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.

A high confidence rating, however, does not necessarily imply that the soil parameter estimates shown will be representative for the soil unit under consideration. Profile selection for SOTER and WISE, as for any other soil database, is not probabilistic but based on available data and expert knowledge. Several of the soil attributes under consideration in Table 1 are not diagnostic in the Revised Legend (FAO 1988). In addition, several properties are readily modified by changes in land use or management, for example soil pH, aluminium saturation, and organic matter content, while information on land use/management history is seldom available.

Despite of the rather large number of TTR-rules already available, it has been necessary to introduce a number of expert-based rules. These rules take into consideration whether certain combinations of soil parameter estimates are considered pedo-chemically feasible or relevant for a given soil unit. 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. So far, 28 expert-rules have been defined (Table 4). Inherently, the scheme of expert-rules was applied after application of the taxotransfer scheme, as a ‘final check’ of the TTR-derived data.

Table 4. 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 (corrected for organic 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 etc. for rules determining the parameter estimates for CECclay). The number of conditions defined so far for each expert-rule is shown in brackets.

Table 5 shows that all soil parameters have been derived through application of taxotransfer roles and expert rules, since the base materials did not contain any measured profiled data (FAO *et al.*

1998b). Details for each virtual profile, layer and soil attribute may be found in table *SOTERflagTTRrules* (Appendix 3).

Table 5. Type and frequency of rules applied

Rules	Code	Frequency of occurrence (%)			
		TTRsub	TTRmain	TTRsummary	TTRfinal
XR1-Alsa	1	-	-	-	81
XR2-Bsat	2	-	-	-	7
XR3-Elco	3	-	-	-	52
XR4-Gyps	4	-	-	-	52
XR5-CaCo	5	-	-	-	74
XR6-CECc	6	-	-	-	3
XR7-Hist	7	-	-	-	3
XR8-LAC	8	-	-	-	2
XR9-ECEC	9	-	-	-	79
TTR-ALSA	A	6	72	78	-
TTR-BSAT	B	89	11	100	-
TTR-BULK	C	89	11	100	-
TTR-CECc	D	87	13	100	-
TTR-CECs	E	92	8	100	-
TTR-GRAV	F	92	8	100	-
TTR-CLAY	G	91	9	100	-
TTR-ECEC	H	78	22	100	-
TTR-ELCO	I	64	36	100	-
TTR-ESP	J	90	10	100	-
TTR-GYPS	K	47	53	100	-
TTR-PHAQ	L	92	8	100	-
TTR-SAND	M	91	9	100	-
TTR-SILT	N	91	9	100	-
TTR-TAWC	O	76	24	100	-
TTR-TCEQ	P	61	39	100	-
TTR-TOTC	Q	92	8	100	-
TTR-TOTN	R	89	11	100	-

Note: For definitions of abbreviations see text and Figure 2; also see the footnote in Appendix 3. The abbreviation '-' stands for not applicable. Cumulative percentages shown under TTRsummary, i.e. TTRsub plus TTRmain, and TTRfinal are not additive.

3.4 Linkage to GIS

Aggregated information about the SOTER unit composition and results of the TTR-work can now be linked to the SOTER map using GIS. In transnational databases, this linkage will be through the

NEWSUID, which is a combination of the country's ISO code plus the SUID code.

SOTER units in the study area are comprised of up to 5 soil components. In the primary database, the associated information is stored in a range of relational databases to enhance data storage and management efficiency. To assist end-users, a new table has been created that incorporates data held in the primary SOTER_IGAD database and the present information on soil parameter estimates (Appendix 4). Clearly, this wealth of information, although needed for the modelling work, complicates linkage to GIS.

For visualization and analysis in GIS, it will often be necessary to make an extra selection. For example, in the case of the RothC and Century carbon models (Falloon *et al.* 2002; Paustian *et al.* 1997), 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). The corresponding selection is included as a separate table, as an example (see Appendix 4).

Figure 3 schematically shows the procedure for linking the various secondary attribute data to the geographical SOTER data held in the GIS. For ease of visualization, it considers only the upper layer (D1) of the spatially dominant (first) soil component of a hypothetical SOTER unit, as an example.

All geographic data in SOTER are presented in vector format. However, should grid-based soil layers be required, these can be generated using the convert-to-grid module of the spatial analyst extension to ArcView (ESRI 1996). The minimum legible delineation implied by the scale of 1:1M is about 25 km^2. Gridding should be based on the NEWSUID field to permit subsequent linkage with the various attribute tables discussed in this report. The procedure will be same as depicted earlier in Figure 3.

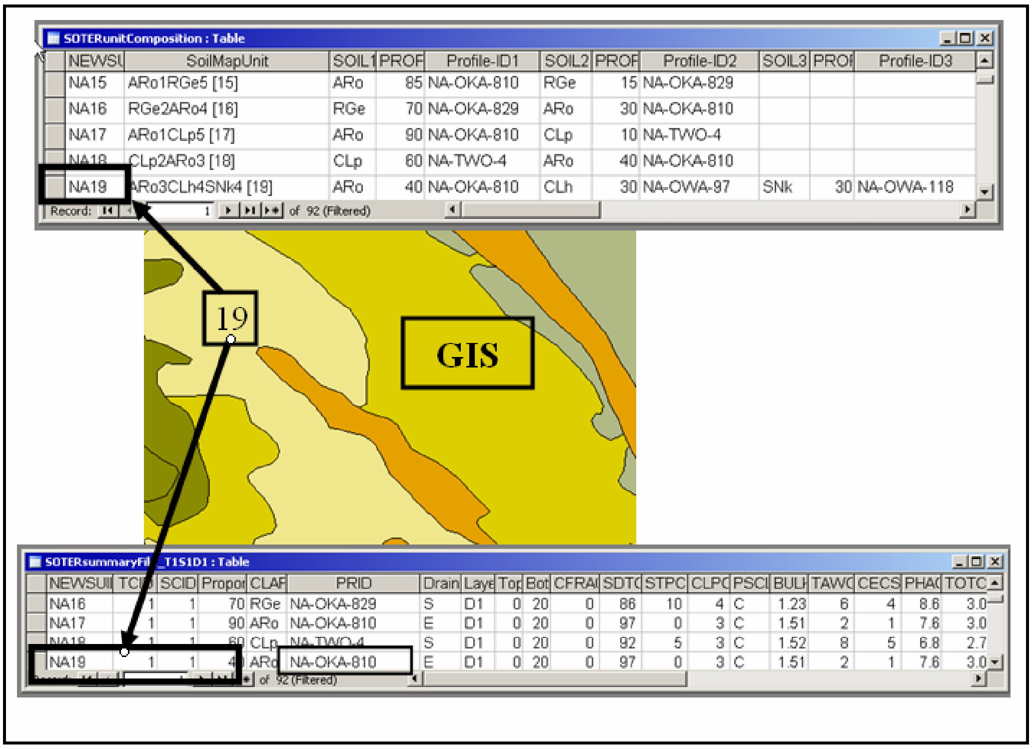

Figure 3. Linking soil parameter estimates for the top 20 cm of the dominant soil of a hypothetical SOTER unit with the geographical component of SOTER

4. CONCLUSIONS

- This study involved generalisation of measured soil (profile) data by soil unit and depth zone. This involved the transformation of variables that show a marked spatial and temporal variation and that have been determined in a range of laboratories, according to various analytical methods.
- A pragmatic approach to the comparability of soil analytical data has been adopted. This was considered appropriate at the present scale of 1:1M, but must be done more rigorously when more detailed scientific work is considered.
- The present set of soil parameter estimates for the Northeastern Africa should be seen as best possible estimates, based on the currently available selection of profiles data held in WISE. The uncertainty attached to some of the present parameter

estimates, however, may be high since there were no measured profile data in the primary data set for the region (see FAO *et al.* 1998a).

- Researchers from the countries under consideration should be encouraged to collect representative soil profiles to help consolidate the primary data set, using SOTER conventions, so as to permit updates of the current secondary data set.
- Modellers should familiarize themselves with the assumptions and taxotransfer rules used to develop the set of soil parameter estimates, before using these in their models.
- The detail and quality of primary information available within the various countries of Northeastern Africa resulted in a variable resolution of the products presented.
- The secondary data set is considered appropriate for studies at national scale, including agro-ecological zoning, land evaluation, and modelling of carbon stocks and changes.

ACKNOWLEDGEMENTS

This study has been carried out in the framework of *ISRIC Project 3.12 — Derived soil parameter estimates for SOTER-based applications*.

I thank Piet Tempel for converting the original data files for Northeastern Africa to a SOTER-compatible format.

REFERENCES

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

Batjes NH 2002a. *Soil parameter estimates for the soil types of the world for use in global and regional modelling (Version 2.0)*. ISRIC Report 2002/02c, International Food Policy Research Institute (IFPRI) and International Soil Reference and Information Centre (ISRIC), Wageningen

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

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

Batjes NH 2004. Soil carbon stocks and projected changes according to land use and management: a case study for Kenya. *Soil Use and Management* 20, 350-356

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

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

Batjes NH, 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

Bouwman AF, Boumans LJM and Batjes NH 2002. Modeling global annual N_2O and NO emissions from fertilized fields. *Global Biogeochemical Cycles* 16, 1080, doi:10.1029/2001GB001812

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

Dijkshoorn JA, Huting JRM and Tempel P 2005. *Update of the 1:5 million Soil and Terrain Database for Latin America and the Caribbean (SOTERLAC, ver. 2.0)*. Report 2005/01 (Release 10 February 2005), ISRIC - World Soil Information, Wageningen

Dondeyne S and Deckers J 2003. *Updating the soil and terrain database for North Eastern Africa*. Internal Report (unpubl.), Katholieke Universiteit Leuven and FAO, Rome

Doorenbos J and Kassam AH 1978. *Yield response to water*. Irrigation and Drainage Paper 33, FAO, Rome

ESRI 1996. *ArcView GIS*. Environmental Systems Research Institute, Redlands CA, 350 p

Falloon P, Smith P, Szabo J and Pasztor L 2002. Comparison of approaches for estimating carbon sequestration at the regional scale. *Soil Use and Management* 18, 164-174

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

FAO 1995. *Digital Soil Map of the World and Derived Soil Properties*, Food and Agriculture Organization of the United Nations, 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 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, 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, 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 1971-1981. *Soil Map of the World, 1:5,000,000. Vol. 1 to 10*. United Nations Educational, Scientific, and Cultural Organization, Paris

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

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

KSS 1995. *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

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

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 R Jones (editors), *Advances in GeoEcology 35*. CATENA Verlag GMBH, Reiskrirchen

Nemes A, Schaap MG and Wosten JHM 2003. Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection. *Soil Science Society of America Journal* 67, 1093-1102

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

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

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

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

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 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, Dijkshoorn KDH and et al. 2005. *Global and National Soils and Terrain Digital Databases (SOTER): Procedures Manual (2nd rev. ed.)*. (In prep.), UNEP, IUSS, ISRIC and FAO, Wageningen

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

Webster R and Oliver MA 1990. *Statistical methods in soil and land resource survey*. Spatial Information Systems. Oxford University press, Oxford, 316 p

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: SOTER unit composition file

This summary table gives the full composition of each SOTER unit in terms of its main soil units (FAO and ISRIC, 2003), 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*. It can be easily linked to the SOTER geographical data in a GIS through the unique SOTER unit code – NEWSUID, a combination of the fields for ISO and SUID – and linked to the table holding the soil parameter estimates through the unique profile identifier (PRID, see Appendix 2).

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 (sometimes called: ISOCSUID)
SOIL1	Text	3	Characterization of the first (main) soil unit according to the Revised Legend (FAO, 1988)
PROP1	Integer	2	Proportion, as a percentage, that the main soil unit 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

(cont.)

SOIL6	Text	3	As above but for the next soil component
PROP6	Integer	2	As above
PRID6	Text	15	As above
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 2: 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 1).

Structure of table *SOTERparameterEstimates*

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 (cm)
CFRAG	Integer	2	coarse fragments (> 2mm)
SDTO	Integer	2	sand (mass %)
STPC	Integer	2	silt (mass %)
CLPC	Integer	2	clay (mass %)
PSCL	Text	1	FAO texture class (see note at end of this report for codes)
BULK	Single	4	bulk density (kg dm^{-3})
TAWC	Integer	2	available water capacity (cm m^{-1}, -33 to -1500 kPa conform to USDA standards)

(cont.)

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})

Note: These are depth-weighted values.

The above table should be consulted in conjunction with table *SOTERflagTTRrules* which documents the taxotransfer rules that have been applied (see Appendix 3).

Appendix 3. Flagging taxotransfer rules

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

Structure of table *SOTERflagTTRrules*

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)
TTRsub	Text	50	Codes showing the type of taxotransfer rule used (based on data for soil *units*; see text)
TTRmain	Text	50	Codes showing the type of taxotransfer rule used (based on data for *major units*; see text)
TTRfinal	Text	25	Additional flags (based on expert knowledge)

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

(Klamt and Sombroek 1988). When applicable, this has been flagged in the field TTRfinal; coding conventions are given in Table **.

Appendix 4: SOTER summary file

Interpretations of a SOTER database, in combination with the current set of soil parameter estimates requires a good knowledge 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 data and its ultimate linkage to GIS, a SOTER summary file has been created. The structure of the corresponding table is shown below.

Structure of table *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 in given SOTER unit
SCID	Integer	1	Number of soil component within given terrain component and SOTER unit
PROP	Integer	3	Relative proportion of above in 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	Upper depth of layer (cm)
BotDep	Integer	4	Lower dept of layer (cm)
CFRAG	Integer	2	Coarse fragments (> 2mm)
SDTO	Integer	2	Sand (mass %)
STPC	Integer	2	Silt (mass %)
CLPC	Integer	2	Clay (mass %)

(cont.)

PSCL	Text	1	FAO texture class (see Figure 8)
BULK	Single	4	Bulk density (kg dm^{-3})
TAWC	Integer	2	Available water capacity (cm m^{-1}, -33 to -1500 kPa, USDA standards)
CECS	Single	4	Cation exchange capacity ($cmol_c$ kg^{-1}) of 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) Terrain Components, and their constituent Soil Components, within a given SOTER unit are numbered starting with the spatially dominant one (see Figure 6). The sum of the relative proportions of all Soil Components within a SOTER unit is always 100 per cent.
3) A condensed file showing only soil parameter estimates for the main Terrain Component (TCID= 1) and Soil Component (SCID =1) for the upper layer (D1) is attached as table SoterSummaryFile_T1S1D1 (see Figure7). This type of tables can be created directly in the GIS, in the table mode, using the SQL-connect option.
4) A limited number of TTR-derived records may contain a -1 value; this indicates that it has not yet been possible to plug the corresponding gaps using the taxotransfer scheme. Alternatively, a value of -7 was used for areas of 'non-soil' (van Engelen *et al.* 2005).

General information on landform, lithology, slope and a wide range of other estimated soil properties may be found in the primary data set for Northeastern Africa (FAO *et al.* 1998b).

Appendix 5: Contents of GIS-folder

The SOTER-GIS files for Northeastern Africa, Northeast of the Ural mountains, and soil parameter estimates are provided in one single zip file called: SOTWIS_NCA_ver1.zip.

By default, this compressed file should be unzipped to folder X:\SOTWIS_NCA_ver1.0. The first time the project file is opened on a new system, the folder information (X) will be automatically updated in the apr-file.

The folder contains:

1) The project's apr-file, called SOTWIS_NCA_01.apr. This file can best be accessed from within ArcView.
2) The SOTER shape, legend and documentation files for the study area in three separate subfolders.
3) The access database containing the soil parameter estimates (SOTWIS_NBA_1.mdb; see Appendices 1 to 4).

The current project file only shows a limited number of selections for the upper soil layer (D1= 0 to 20 cm or less for shallow soils) of the dominant soil of a SOTER unit. Should other selections be needed, the underlying MS Access database (SOTWIS_NCA_1.mdb) can be queried via the SQL-connect option of ArcView.

Appendix 6: Limits for soil textural classes

The textural classes (PSCL, see Appendix 2 and 4) used in this study follow the criteria of FAO (1988) and CEC (1985). The following abbreviations are used: C–coarse, M–medium, Z–medium fine, F–fine and V–very fine. The symbol # is used for undifferentiated (i.e. C + M + F + Z + V). The class limits are shown in Appendix 6.

The textural class for all Histosols has been coded "O", for organic materials.

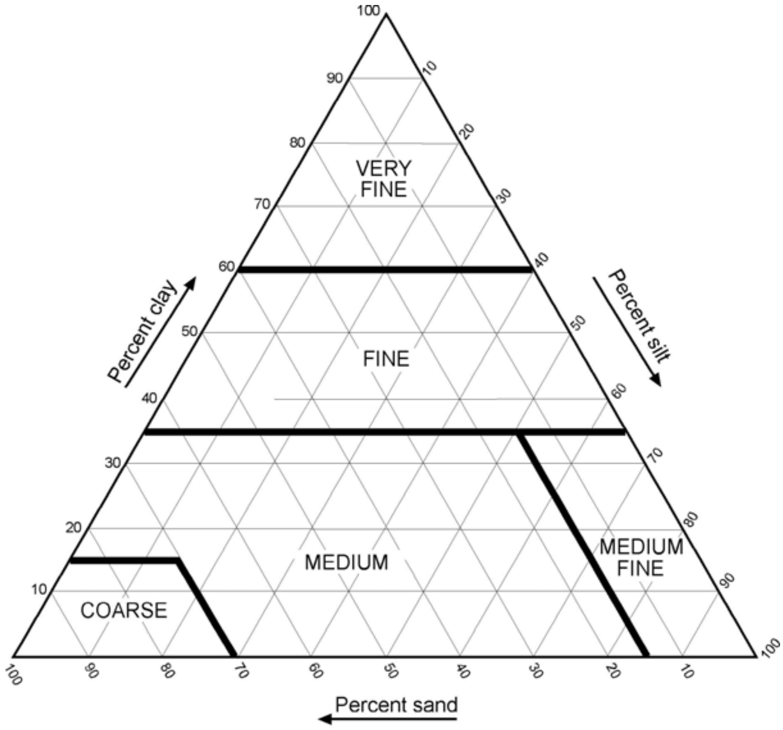

Figure 4. Soil texture classes