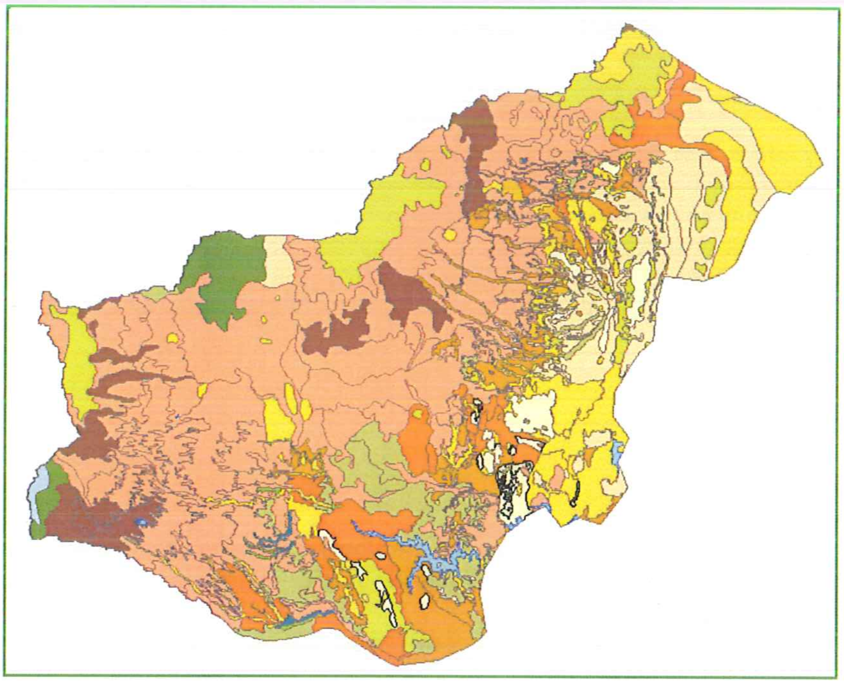
Report 2010/07

Soil property estimates for the Upper Tana, Kenya, derived from SOTER and WISE

(Version 1.0)

Niels H. Batjes

(May 2010)



ISRIC World Soil Information

Correct citation:

Batjes NH 2010. Soil property estimates for the Upper Tana, Kenya, derived from SOTER and WISE (ver. 1.0). Report 2010/07, ISRIC – World Soil Information, Wageningen (41p. with data set) (http://www.isric.org/isric/Webdocs/Docs/ISRIC_Report_2010_07.pdf)

Inquiries:

c/o Director, ISRIC – World Soil Information
PO Box 353
6700 AJ Wageningen
The Netherlands
E-mail: soil.isric@wur.nl
Web: www.isric.org

Front cover: Main FAO soil groupings for the Upper Tana, Kenya

Contents

List of Tables

List of Figures

Foreword

ISRIC – World Soil Information has the mandate to create and increase the awareness and understanding of the role of soils in major global issues. As an international institution, we inform a wide audience about the multiple roles of soils in our daily lives; this requires scientific analysis of sound soil information.

This study presents derived soil properties for the Upper Tana, Kenya, for application in exploratory studies. It draws on two databases developed at ISRIC. First, the Soil and Terrain (SOTER) database for the Upper Tana, Kenya, at scale 1:250 000, compiled in the framework of the Green Water Credits Project. Being dependent on historic data, there are often gaps in the measured analytical data held in SOTER. ISRIC – World Soil Information has therefore developed a uniform, consistent methodology for filling common gaps in *primary* SOTER databases to produce *secondary* (SOTWIS) data sets for general-purpose applications. This taxotransfer rule-based procedure draws heavily on soil analytical data held in the ISRIC-WISE soil profile database.

The consistent taxotransfer procedure has already been applied to SOTER datasets for Latin America and the Caribbean, Central and Eastern Europe, Southern Africa, Central Africa and other areas with SOTER-like databases. These *secondary* databases have been used in support of the 'Harmonized World Soil Database', a collaborative effort of Food and Agriculture Organization of the United Nations (FAO), International Institute for Applied Systems Analysis (IIASA), ISRIC - World Soil Information, Institute of Soil Science - Chinese Academy of Sciences (IISCAS), and Joint Research Centre of the European Commission (JRC).

In order to consolidate its world soil databases, ISRIC – World Soil Information is seeking collaboration with national institutes with a mandate for soil resource inventories.

Dr Ir Prem Bindraban

Director, ISRIC – World Soil Information

SUMMARY

This report describes a harmonized set of soil property estimates for the Upper Tana, Kenya. The data set was derived from the 1:250 000 scale Soil and Terrain Database for the Upper Tana (SOTER_UT, ver. 1.0) and the ISRIC-WISE soil profile database, using standardized taxonomy-based pedotransfer (taxotransfer) procedures.

The land surface of the Upper Tana, Kenya, covering some 15,905 km^2, has been mapped in SOTER using 187 unique SOTER units. Each map unit consists of up to four different soil components. In so far as possible, each soil component has been characterized by a regionally representative profile, selected and classified by national soil experts. Conversely, in the absence of any measured legacy data, soil components were characterized using synthetic profiles for which only the FAO-Unesco (1988) classification is known.

Soil components in SOTER_UT have been characterized using 144 profiles consisting of 108 real and 36 synthetic. The latter were used to represent some 18% per cent of the study area. Comprehensive sets of measured attribute data are seldom available for most profiles (108) collated in SOTER_UT, as these were not considered in the source materials. Consequently, to permit modelling, gaps in the soil analytical data have been filled using consistent taxotransfer procedures. Modal soil property estimates necessary to populate the taxotransfer procedure were derived from statistical analyses of soil profiles held in the ISRIC-WISE database. The current taxotransfer procedure only considers profiles in WISE that: (a) have FAO soil unit names (43) identical to those mapped for the Upper Tana in SOTER, and (b) originate from regions having similar Köppen climate zones (*n*= 5617).

Property estimates are presented for 18 soil variables by soil unit for fixed depth intervals of 0.2 m to 1 m depth: 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 (ECe), bulk density, content of sand, silt and clay, content of coarse fragments (> 2 mm), and volumetric water content (-33 kPa to -1.5 MPa). 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 studies of global environmental change.

The soil property estimates can be linked to the spatial data (map), using GIS, through the unique SOTER-unit code; database applications should consider the full map unit composition and depth range.

The derived data presented here may be used for exploratory assessments at basin scale (< 1:250 000). They should be seen as best estimates based on the current, still limited, selection of soil profiles in SOTER for the Upper Tana and data clustering procedure — the type of taxotransfer rules used to fill gaps in the measured data has been flagged to provide an indication of confidence in the derived data.

Keywords: legacy soil data, taxotransfer procedures, derived soil properties, secondary data set, Upper Tana, Kenya, WISE database, SOTER database

1 INTRODUCTION [1]

ISRIC, 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, a global SOTER is to supersede the 1:5 million scale Soil Map of the World (Nachtergaele and Oldeman 2002; Oldeman and van Engelen 1993).

SOTER databases are composed of two main elements: a geographic and an attribute data component. 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. Methodological details may be found in the SOTER Procedures Manual (van Engelen and Wen 1995).

Soil components of individual SOTER units are characterized by a representative soil profile (Figure 1). These legacy data are selected from available soil survey reports, as the SOTER program does not involve new ground surveys. As a result, there are often gaps in the measured (i.e. *primary*) analytical data, in particular the soil physical data. This precludes the direct use of *primary* SOTER data in models. ISRIC has therefore developed a uniform, consistent methodology for filling common gaps in *primary* SOTER databases to produce *secondary* (SOTWIS) data sets for general-purpose applications (Batjes 2003; Batjes *et al.* 2007). This taxotransfer rule-based procedure draws heavily on soil analytical held in the ISRIC-WISE soil profile database (Batjes 2009). So far, the consistent taxotransfer procedure has been applied to SOTER data for Latin America and the Caribbean, Central and Eastern Europe, Southern Africa, Central Africa and other areas with SOTER-like databases (see www.isric.org for details). The approach has also been used in support of the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC 2009).

[1] Note: *Reports that describe secondary SOTER (SOTWIS) databases have similar structure and content, the main difference being the region-specific information presented in each document [NHB].*

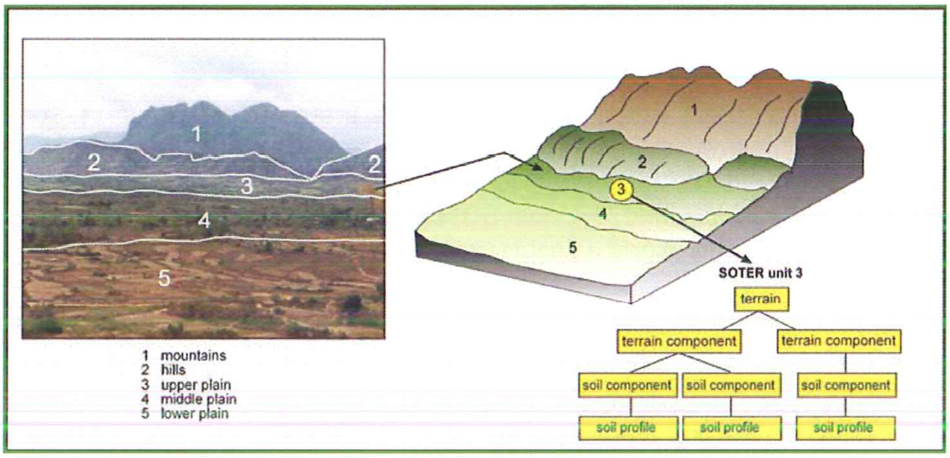


Figure 1. Representation of SOTER units and their conceptual structure

This report discusses the application of the taxotransfer procedure to the *primary* SOTER data for The Upper Tana, Kenya (hereafter referred to as SOTER_UT). Chapter 2 describes the materials and methods with special focus on the procedure for preparing the *secondary* SOTER data. Results are discussed in Chapter 3, while concluding remarks are drawn in Chapter 4. The structure of the various output tables and installation procedure are documented in the Appendices.

2 MATERIALS AND METHODS

2.1 Source of primary SOTER data

The SOTER database covering the Upper Tana, Kenya, compiled in the framework of the Green Water Credits Project (GWC), provided the basis for this study. The available soil geographical and attribute data were collated into SOTER format using the base materials described in Dijkshoorn et al. (2010). Although the map has a generalized scale of 1:250 000 million, the detail, and quality of the primary information varies widely within the study area.

All profiles in SOTER_UT were characterised according to the Revised Legend of FAO (1988) and World Reference Base for Soil Resources (FAO 2006). The Revised Legend, however, was used to display/map the soil units using GIS to ensure global consistency with earlier SOTER databases (e.g. FAO and ISRIC 2003; FAO *et al.* 1998).

2.2 Preparation of secondary SOTER data

2.2.1 Checking of primary data

The *primary* SOTER database was first screened for possible inconsistencies using automated integrity checks developed for WISE (Batjes 1995). All soil classifications were taken at face value; small inconsistencies in the analytical data, however, were corrected. The screened set provided the basis for the current analyses.

The screened dataset includes 144 so-called representative profiles, consisting of 108 real profiles, of which 100 are geo-referenced, and 36 virtual profiles. These profiles are physically linked to the spatial data in accord with SOTER standards.

In accord with SOTER conventions (van Engelen and Wen 1995), so-called virtual profiles have been introduced when the FAO classification for a given soil unit was known from soil maps for the region, but there are no real profiles (i.e. measured data) yet to characterize these units (see Dijkshoorn and others 2010). For each virtual profile, the soil drainage class was inferred using expert judgement.

Several map units consist of inland waters (KEns1) — these are only characterized in the GIS-file.

2.2.2 Filling gaps in the measured soil data

Being based on available soil survey reports, there are always gaps in the soil analytical data — the limited set of so-called "mandatory SOTER attributes" simply is not available for most profiles in SOTER_UT.

Gaps in the attribute data were filled here using consistent taxotransfer procedures (Batjes 2003; Batjes *et al.* 2007). The soil variables considered in the procedure are detailed in Section 2.3.3. The soil property estimates required to run these procedures were derived from statistical analyses of 5617 profiles extracted from version 3.1 of the ISRIC-WISE database (Batjes 2009). This selection only included those profiles in WISE that: (a) have similar FAO (1988) classification as mapped for SOTER_UT (Table 1), and (b) originate from areas that have similar Köppen-Geiger climates as mapped for the Upper Tana by Kottek et al. (2006).

Table 1. FAO soil units mapped in SOTER_UT and number of similar soil profiles in WISE used for taxotransfer rule development

FAO soil units	SOTER_UT [a]	WISE [b]
ALh	1 (1/0)	60
ARc	2 (0/2)	55
ARh	3 (0/3)	240
ATc	1 (1/0)	5
CLh	1 (1/0)	152
CLl	5 (5/0)	44
CLp	7 (6/1)	54
CMc	6 (5/1)	176
CMe	2 (1/1)	205
CMg	1 (0/1)	78
CMu	1 (1/0)	69
CMv	3 (3/0)	64
CMx	2 (1/1)	68
FLc	4 (2/2)	130
FLe	1 (0/1)	133
FLs	1 (0/1)	11
Gle	3 (0/3)	163
GLk	1 (0/1)	13

FAO soil units	SOTER_UT [a]	WISE [b]
GYh	3 (1/2)	11
GYk	2 (0/2)	21
GYp	3 (1/2)	11
KSh	1 (0/1)	25
KSk	3 (3/0)	52
LPe	3 (0/3)	70
LPk	2 (0/2)	33
LVj	1 (1/0)	58
LVk	3 (3/0)	127
LVv	1 (1/0)	34
LXh	2 (2/0)	199
PHl	1 (1/0)	129
PLd	1 (1/0)	31
PZh	1 (1/0)	68
RGc	5 (3/2)	97
RGe	3 (1/2)	164
RGy	3 (0/3)	5
SCg	2 (0/2)	35
SCn	2 (1/1)	33
SCy	3 (1/2)	33
SNh	1 (0/1)	96
SNk	1 (1/0)	48
VRe	2 (2/0)	311
VRk	4 (4/0)	148
VRy	1 (0/1)	7

[a] First number is for total number of soil profiles linked to the SOTER-GIS map units; the first number in brackets is for measured profiles, the second for virtual profiles (i.e. profiles for which there are no measured data; these have codes like KE*syn*15 or KE*syn*LPe)

[b] Number of profiles from WISE considered in the taxotransfer scheme (*n* = 5617); for details see text.

Measured values in WISE that underlie the taxotransfer scheme — like those held in SOTER_UT — will reflect both variations inherent to the soil unit and those that can be ascribed to the methods of sampling and measurement. For reasons outlined earlier (Batjes 2002, p. 6-11), a pragmatic approach to the comparability of soil analytical data had to be adopted for use with small scale SOTER databases. A similar approach has been used with the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC 2009). This type of approach is considered appropriate for soil data applications at broad scale; correlation of soil analytical data, however, should be done more rigorously when more precise scientific research is considered.

The analytical data for each combination of soil unit, texture class and depth layer were screened using a robust outlier scheme, by attribute (see Batjes 2003). The output of the taxotransfer procedure has been stored in a *secondary* data set (known as SOTWIS database); for details see Appendix 1.

Table 2. List of soil variables considered in secondary SOTER data sets

Organic carbon
Total nitrogen
Soil reaction (pH_{H2O})
Cation exchange capacity (CEC_{soil})
Cation exchange capacity of clay size fraction (CEC_{clay}) [a b]
Base saturation (as % of CEC_{soil}) [b]
Effective cation exchange capacity (ECEC) [b c]
Aluminium saturation (as % of ECEC) [b]
$CaCO_3$ content
Gypsum content
Exchangeable sodium percentage (ESP) [b]
Electrical conductivity (ECe)
Bulk density
Coarse fragments (> 2 mm, volume %)
Sand (mass %)
Silt (mass %)
Clay (mass %)
Available water capacity (cm^3 cm^{-3} 10^2 or vol%; -33 kPa to -1.5 MPa) [b d]

[a] CEC_{clay} was calculated from CEC_{soil} by assuming a mean contribution of 350 $cmol_c$ per 100 g OC, the common range being from 150 to over 750 $cmol_c$ per 100 g (Klamt and Sombroek 1988). Similarly, as a rule of thumb, CEC_{OC} values of 300 to 400 $cmol_c$ per 100 g OC (NH_4OAc, pH 7.0), are used by USDA-NRCS (1995 p. 26).

[b] Calculated from other measured soil properties.

[c] ECEC is defined here as exchangeable (Ca^{++}+ Mg^{++}+ K^++ Na^+) + exchangeable (Al^{+++}) in accordance with USDA-NRCS (1995); see also FAO (2006, p. 125).

[d] Limits for soil water potential for Available Water Capacity (AWC) conform to USDA standards (Soil Survey Staff 1983); these values are not corrected for volume percentage of coarse fragments.

2.2.3 List of soil variables

Special attention has been paid to those key attributes (Table 2) that are 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; Easter *et al.* 2007; FAO/IIASA/ISRIC/ISSCAS/JRC 2009; Fischer *et al.* 2002; Scholes *et al.* 1995).

Table 2 does not include soil hydraulic properties because measured data for the latter are generally lacking in the systematic soil survey reports that underlie SOTER and WISE.

3 RESULTS AND DISCUSSION

3.1 Map unit composition

The Upper Tana has been characterized using 190 unique map units or SOTER units in the *SoilComponent* table (Figure 1); these comprise 192 terrain components and 262 soil components. It contains data for four SOTER units that are not shown on the GIS map: KE237, KE239, KE240, and KE411. As a result, the GIS map considers data for 186 unique map units, corresponding with 617 polygons on the map. Some 18% of the study area has been characterized by a synthetic profile.

At the small scale under consideration, many mapping units are compound; they may consist of up to four different soil units. This map unit complexity must be considered when using the data; typically, this will have to be done using software, specifically written for a particular application (e.g. Batjes *et al.* 2007; Easter *et al.* 2007). Overall, the Upper Tana has been characterized using 43 different soil units (FAO Revised Legend).

The full composition of each SOTER unit has been summarized in table *SOTERunitComposition* (Appendix 1). This table lists the name and relative area of the main major FAO soil group for each map unit, as well as the type and relative area of all the component soil units.

3.2 Soil property estimates

The taxotransfer procedure generates soil property estimates for five standardized depth ranges of 20 cm each to 1 m, and 2 standardized depth ranges of 50 cm (100-150 cm and 150-200 cm; (see Batjes 2008). Inherently, property estimates for the deeper layers are considered less reliable than those for the upper layers of soil as they are based on less extensive data sets. Therefore, the current data set only holds derived data up to 1 m depth, or less when applicable (e.g. for shallow Leptosols).

In case of missing measured values in SOTER, the cut-off point for applying any taxotransfer rule is $n_{WISE} < 5$; that is there should be at least 5 cases in the WISE subset for the corresponding combination of soil unit, soil variable, soil layer, and soil textural class in order to apply the substitution procedure. Soil textural classes were defined in accordance with current SOTER standards – coarse, medium, fine, very fine and medium fine (Figure 4, Appendix 6). The taxotransfer procedure is summarized in Figure 2; see also Appendix 3.

Each flag listed under *TTRsub* (where *sub* stands for FAO soil unit) and *TTRmain* (where *main* stands for major soil group) consists of a sequence of letters followed by a numeral, for example *A3h2*. The letters indicate soil attributes for which a taxotransfer rule has been applied; coding conventions are explained in Appendix 3. The number code reflects the size of the sample population in WISE, after outlier rejection, on which the statistical analyses that underlie taxotransfer scheme were based (Table 3).

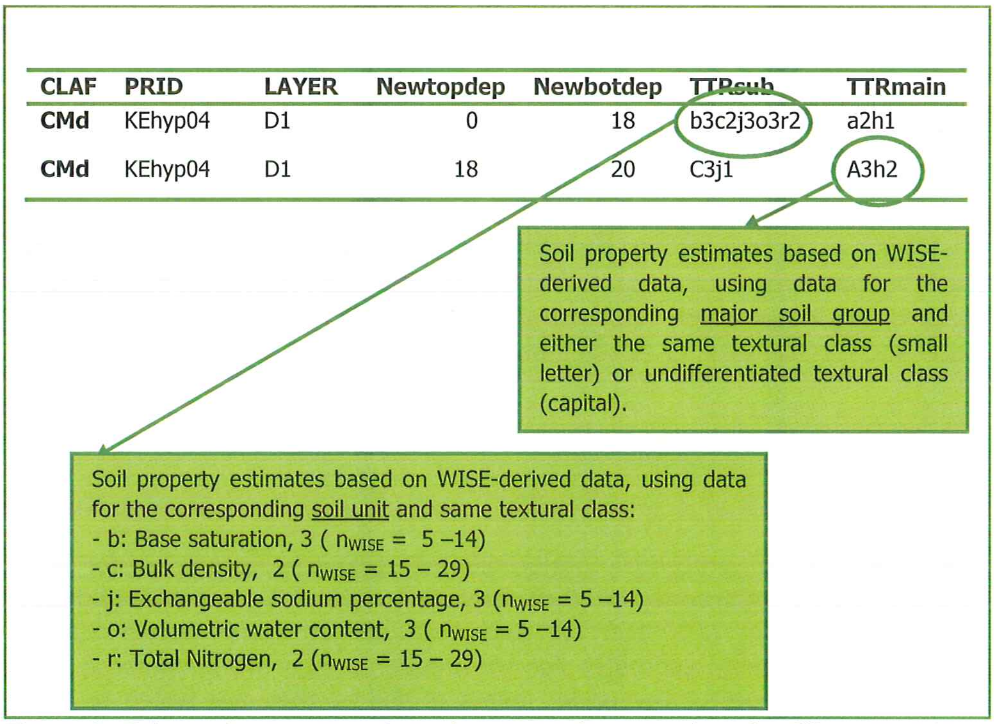


Figure 2. Schematic representation of taxotransfer procedure for filling gaps in SOTER

When a *small letter* is used for *TTRsub*, the substitution was based on median data for the corresponding soil unit, depth layer and textural class (for example, Rhodic Ferralsols (FRr), 0-20 cm (D1), Fine and $n_{WISE} > 5$). Otherwise, when a *capital* is used, this indicates that the substitution for the given soil attribute was based on the whole set for the corresponding soil unit and depth layer, irrespective of soil texture (i.e. undifferentiated or *u*). The same coding conventions apply for TTRmain, but substitutions when consider derived soil data for the corresponding major FAO soil group.

Table 3. Criteria for defining confidence in the derived data

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

[a] n_{WISE} is the sample size after the screening or outlier rejection procedure
[b] The cut-off point in the TTR-approach is $n_{WISE} < 5$

Expert rules are applied after the taxotransfer rules to remedy possible pedological inconsistencies (or artefacts) that may have arisen in the TTR-derived data. Such a check is necessary because individual TTR-rules do not consider possible correlations between different soil variables. For example, one expert rule (XR-TCEQ) checks whether there are indeed no carbonates in acid soil layers (pH <5.5). Similarly, another expert rule (XR-BSAT) checks whether base saturation is low in acid soils and so on. In view of the diversity of soils worldwide, however, it remains difficult to account for all possible situations; this should be kept in mind when combining various derived data.

Derived soil data, resulting from the taxotransfer procedure, are presented in table *SOTERparameterEstimates*; see Appendix 2 for details.

3.3 Type and number or taxotransfer rules used

There are numerous gaps in the *primary* soil analytical data in SOTER_UT (see 2.2.1). Table 4 lists how often each *taxotransfer* respective or *expert*-rule has been applied for each attribute as a percentage of the total number of "horizon/layer/depth" combinations in the *secondary* SOTER or

SOTWIS set; details may be found in table *SOTERflagTTRrules* (Appendix 3).

Table 4 shows, for example, that available water capacity (AWC) has been estimated in 89% of the cases, either using data for similar soil units (74% of cases, see under TTRsub) resp. similar major soil groups (15% of cases, see under TTRmain). Further, expert rules for available (XR-AWC) have been applied in 3% of the cases. This shows that in 92% of the cases, AWC for a given profile and fixed-depth layer, had to be estimated in this study due to the limited availability of measured water retention data for the Upper Tana region.

Table 4. Type and frequency of taxotransfer rules (TTR) and expert rules (XR) applied

TTR code	Frequency of occurrence (%)			
(SOTNAM)	TTRsub	TTRmain	TTR total	Expert rules
TTR-BSAT	28	0	28	-
TTR-BULK	64	3	67	-
TTR-CECC	37	0	37	-
TTR-CECS	16	0	16	-
TTR-GRAV	0	0	0	-
TTR-CLAY	16	0	16	-
TTR-ECEC	79	0	79	-
TTR-ELCO	14	6	20	-
TTR-ESP	26	0	26	-
TTR-GYPS	7	11	18	-
TTR-PHAQ	15	0	15	-
TTR-SAND	16	0	16	-
TTR-SILT	16	0	16	-
TTR-TAWC	74	15	89	-
TTR-TCEQ	19	4	23	-
TTR-TOTC	34	0	34	-
TTR-TOTN	92	1	93	-
XR0-Text	-	-	-	16
XR1-Alsa	-	-	-	73
XR2-Bsat	-	-	-	4
XR3-Elco	-	-	-	48
XR4-Gyps	-	-	-	21
XR5-CaCo	-	-	-	27
XR6-CECc	-	-	-	5
XR7-ESP	-	-	-	0
XR8-CFRA	-	-	-	0
XR9-BULK	-	-	-	0
XR10-AWC	-	-	-	3

Note: For definitions of abbreviations see text and Table 4, see also Appendix 3; '-' stands for not applicable.

3.4 Assumptions and limitations

Soil unit classifications (FAO 1988), as presented in the *primary* SOTER_UT database, were taken at face value. Soil experts, however, may classify the same soil profile differently when the available soil morphological and soil analytical data are 'limited' and subjective assumptions have to be made (e.g., Goyens *et al.* 2007; Kauffman 1987; Spaargaren and Batjes 1995). The soil classification code, however, is the primary driver of the taxotransfer procedure (see 2.2.2).

The overall assumption has been that the confidence in a TTR-based property estimate should increase with the size of the corresponding sample populations present in WISE, for the relevant soil units and Köppen climate zones, after outlier-rejection. In addition, the confidence in soil property estimates listed under TTRsub should be higher than for those listed under TTRmain resp. "derived using expert rules.

A high confidence rating for a given property estimate, however, does not necessarily imply that this estimate will be representative for the soil unit under consideration. Profile selection for SOTER and WISE, as for many other small scale soil databases, is not probabilistic, but based on available data and expert knowledge. Several of the soil attributes under consideration in Table 2 are not diagnostic in the Revised Legend (FAO 1988). In addition, some soil properties are readily modified by changes in land use or management, for example soil pH, aluminum saturation, soil salinity, and organic matter content. Information on land use/management history by profile, however, is seldom available in SOTER and, as such, this aspect could not be considered explicitly in the taxotransfer procedure yet.

Finally, it should be noted that adoption of different criteria for clustering data would inherently lead to varying property estimates. For example, selecting a different soil classification system (e.g., FAO 1974, FAO 1988 or WRB 2006), limits for depth layers (e.g., 0-20 cm intervals up to 100 cm *versus* 0-30 cm and 30-100 cm), criteria for defining soil textural classes (e.g., 5 classes in SOTER *versus* 3 classes for the FAO Soil Map of the World), choice of critical limits for applying taxotransfer rules (i.e. reject when $n_{WISE} < 5$ or $n_{WISE} < 15$), as well as the type of outlier-rejection and statistical procedures used, and the number of WISE profiles under consideration. Most importantly, however, the outcome will be determined primarily by the number and quality of the profile data collated in the underpinning, *primary* SOTER database. In particular, their geographic distribution over the region respectively various SOTER units, the degree to which the various data-fields have been filled, and the overall comparability of analytical methods used.

3.5 Linkage to GIS

SOTER units mapped for the region comprise up to four soil components. The full map unit composition has been summarized in one single table (*SOTERunitComposition*, see Appendix 1). Results of the taxotransfer procedure for each soil component, as typified by the representative profile, are stored in table *SOTERparameterEstimates* (Appendix 2). Results in this table have been linked to the corresponding SOTER units in two tables having the same content, but different data structures: a) *SOTERsummaryFile*, in which data by layer (D_i) are presented vertically by *NEWSUID*, *TCID* and *SCID* (Appendix 4), and b) *SOTERsummaryFile_Prop* in which derived data for layer D_1 to D_5 are data presented horizontally by *NEWSUID*, TCID and SCID (Appendix 5).

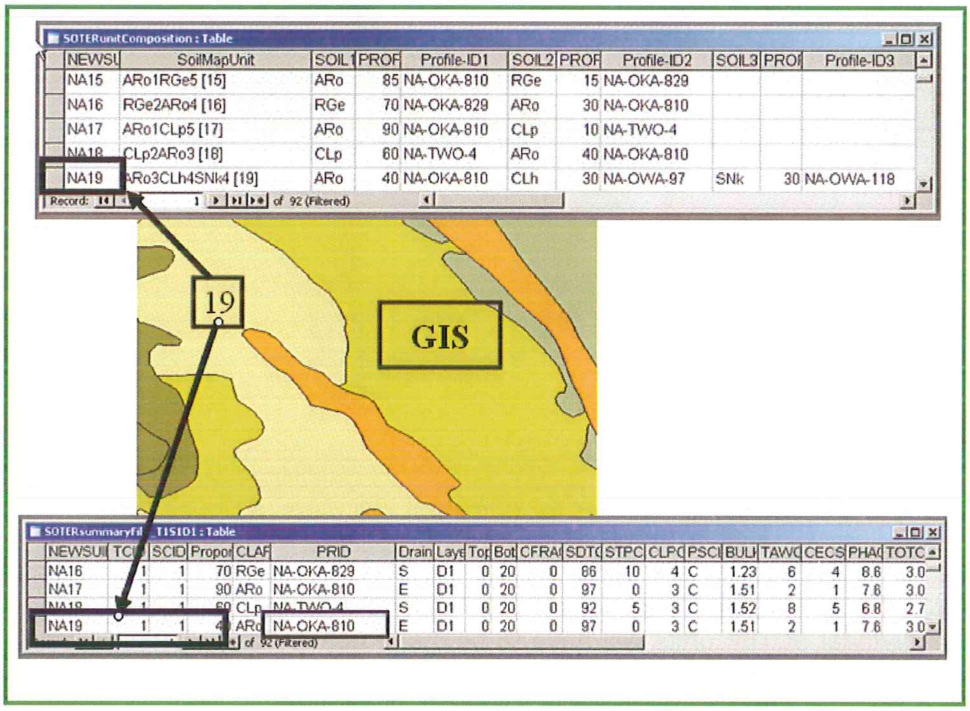


Figure 3. Schematized procedure for linking soil property estimates for the upper layer (D1) of the main soil unit (TCID=1; SCID=1) of a SOTER map unit with the geographical data

Data in the later tables can be linked to GIS through *NEWSUID*, the unique SOTER map unit code. The overall procedure is visualized in Figure 3 for a hypothetical database. It should be noted here that GIS can only be used to display one "set of attributes" at a time per polygon or SOTER map unit. As an example, derived topsoil properties for organic carbon content, bulk density and available water capacity for the dominant soil component in a SOTER unit (i.e. TCID=1, SCID=1 and Layer= D1) are shown in Figure 4; classification is according to natural breaks (Jenks).

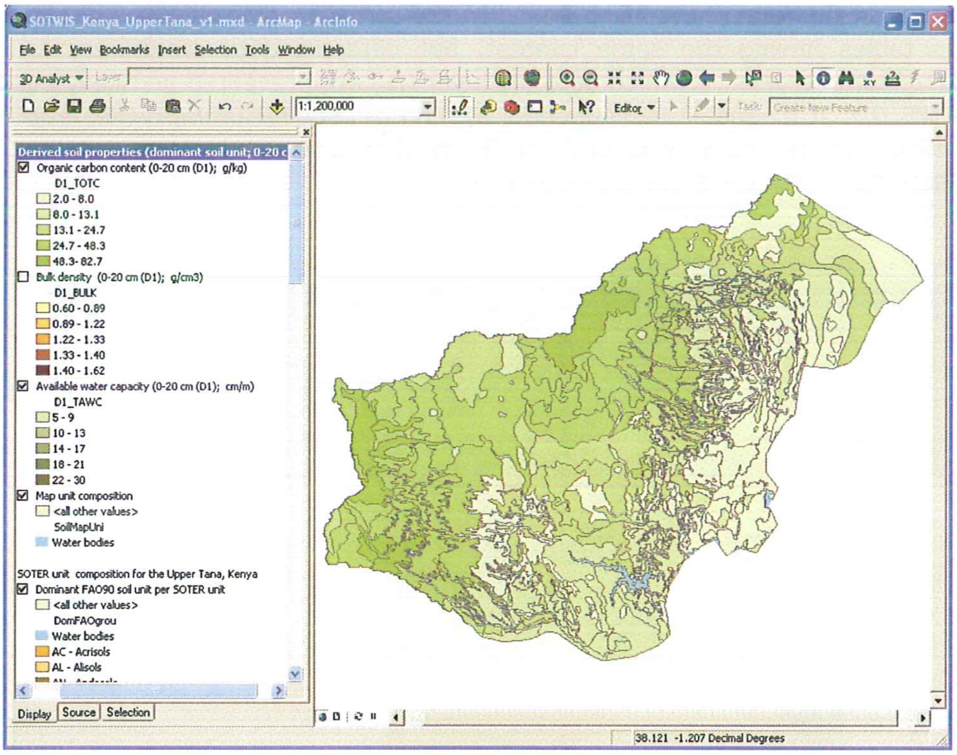


Figure 4. Derived soil properties for dominant soil units in Upper Tana

Typically, specific data selections that consider the full soil unit composition of individual SOTER units will have to be made before "aggregated" model output can be coupled back to the mapping units in the GIS. Details of such an approach may be found in Easter et al. (2007).

4 CONCLUSIONS

- The detail and quality of primary soil and terrain data underpinning SOTER_UT resulted in a variable resolution of the secondary product presented here.
- Linkage between the soil profile data and the spatial component of SOTER_UT required 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 when developing the taxotransfer procedure. Although this is considered appropriate at a present broad scale, such a comparison must be done more rigorously when more detailed scientific work is considered.
- The derived soil data presented here can be used for exploratory assessments at subnational scale — they should be seen as best estimates based on the current selection of soil profiles in SOTER_UT and data clustering procedure. Once additional profiles become available for the region in SOTER format, the present set of derived soil data should be refined.
- End-users should familiarize themselves with the procedures and assumptions that have been used to derive the soil property estimates prior to using them in models — possible uncertainties are documented in the data set.

ACKNOWLEDGEMENTS

Special thanks are due to my colleague Koos Dijkshoorn for useful discussions concerning the primary SOTER database for the Upper Tana and for liaising with staff at the Kenya Agricultural Research Institute (KARI) concerning the primary SOTER data in the context of the Green Water Credits Project.

Green Water Credits is supported by the International Fund for Agricultural Development. The proof of concept was part-financed by the Swiss Agency for Development and Cooperation. The program is implemented by an international consortium that includes: ISRIC - World Soil Information, Stockholm Environmental Institute, International Institute for Environment and Development, and Agricultural Economics Research Institute.

REFERENCES

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

Batjes NH 2002. *Soil parameter estimates for the soil types of the world for use in global and regional modelling (Version 2.1)*. ISRIC Report 2002/02c, International Food Policy Research Institute (IFPRI) and International Soil Reference and Information Centre (ISRIC), Wageningen. Available at: http://www.isric.org/isric/webdocs/Docs/ISRIC_Report_2002_02c.pdf; accessed 12/2008.

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 (http://www.isric.org/isric/webdocs/Docs/ISRIC_Report_2003_03.pdf)

Batjes NH 2008. Mapping soil carbon stocks of Central Africa using SOTER. *Geoderma* 146, 58-65

Batjes NH 2009. Harmonized soil profile data for applications at global and continental scales: updates to the WISE database. *Soil Use and Management* 25, 124-127 (http://dx.doi.org/10.1111/j.1475-2743.2009.00202.x)

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 (http://www.iiasa.ac.at/Admin/PUB/Documents/IR-97-025.pdf)

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

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 and Steffen (editors), *Global Change and Terrestrial Ecosystems*. Cambridge University Press, Cambridge, pp 529-565

Dijkshoorn JA and others 2010. *Soil and and terrain conditions for the Upper Tana, Kenya*, Kenya Agricultural Research Institute (KARI) and ISRIC – World Soil Information, Wageningen (in prep.)

Easter M, Paustian K, Killian K, Williams S, Feng T, Al-Adamat R, Batjes NH, Bernoux M, Bhattacharyya T, Cerri CC, Cerri CEP, Coleman K, Falloon P, Feller C, Gicheru P, Kamoni P, Milne E, Pal DK, Powlson D, Rawajfih Z, Sessay M and Wokabi S 2007. The GEFSOC soil carbon modeling system: a tool for conducting regional-scale soil carbon inventories and assessing the impacts of land use change on soil carbon. *Agriculture, Ecosystems & Environment* 122, 13-25

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

FAO 2006. *World Reference Base for soil resources - A framework for international classification, correlation and communication*. World Soil Resources Reports 103, International Union of Soil Sciences, ISRIC - World Soil Information and Food and Agriculture Organization of the United Nations, Rome (http://www.fao.org/ag/agl/agll/wrb/doc/wrb2007_corr.pdf)

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

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

FAO/IIASA/ISRIC/ISSCAS/JRC 2009. *Harmonized World Soil Database (version 1.1)*, Prepared by Nachtergaele FO, van Velthuizen H, Verelst L, Batjes NH, Djjkshoorn JA, van Engelen VWP, Fischer G, Jones A, Montanarella L., Petri M, Prieler S, Teixeira E, Wiberg D and Xuezheng Shi. Food and Agriculture Organization of the United Nations (FAO), International Institute for Applied Systems Analysis (IIASA), ISRIC - World Soil Information, Institute of Soil Science - Chinese Academy of Sciences (ISSCAS), Joint Research Centre of the European Commission (JRC), Laxenburg, Austria (available at http://www.iiasa.ac.at/Research/LUC/luc07/External-World-soil-database/HTML/index.html?sb=1)

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

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

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

http://www.isric.org/isric/webdocs/Docs/ISRIC_TechPap16.pdf; accessed June 2008.

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

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

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

Marsman BA and de Gruijter JJ 1986. *Quality of soil maps: A comparison of soil survey methods in sandy areas*. Soil Survey Papers 15, Netherlands Soil Survey Institute, Wageningen

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

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

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 (Available through: http://soils.usda.gov/technical/manual/; Accessed: June 2008)

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

USDA-NRCS 1995. *Soil Survey Laboratory Information Manual*. Soil Survey Investigations Report No. 45 (version 1.0, May 1995) [Available at ftp://ftp-fc.sc.egov.usda.gov/NSSC/Lab_Info_Manual/ssir45.pdf ; verified 11 January 2007], National Soil Survey Center, Soil Survey Laboratory Lincoln (Nebraska)

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

APPENDICES

Appendix 1. Structure of table SOTERunitComposition

Table *SOTERunitComposition*, in MS-Access® format, gives the full composition of each SOTER unit in terms of its: landform, lithology (parent material), dominant major FAO soil group and its relative extent, then component in soil units with their relative extent, and the identifier for the corresponding representative profile. The relevant information was distilled from three primary SOTER tables, *viz. Terrain*, *SoilComponent,* and *Profile*, to facilitate data processing. The content of this table can be linked to the geographical data in a GIS through the unique SOTER unit code or NEWSUID, a combination of the fields for ISO and SUID.

Structure of table SOTERunitComposition[a]

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994) or WD for World
SUID	Integer	The identification code of a SOTER unit on the map and in the database
NEWSUID	Text	Globally unique code for SOTER unit, comprising fields ISOC plus SUID (e.g. SN15 or GM03)
LNDF	Text	Code for SOTER landforms (see SOTWIS_codes)
LITH	Text	Code for SOTER lithology (See SOTWIS_codes)
NoOfSoilComp	Text	Number of soil components in given SOTER unit
DomFAOgroup	Text	Dominant FAO major soil group in SOTER (Note: This need not always be SOIL1)
PropDomFAOGroup		Proportion of dominant major soil group in SOTER unit (%)
PropSynthProf		Proportion of SOTER unit characterized by a synthethic profile (%)
SoilMapunit	Text	Aggregated code for map unit summarizing the overall composition [b]
SOIL1	Text	Characterization of the first (main) soil unit according to the Revised FAO-Unesco Legend
PROP1	Integer	Proportion, as a percentage, that the main soil unit occupies within the SOTER unit
PRID1	Text	Unique code for the corresponding measured resp. virtual soil profile (e.g. TN_LPe_syn)
SOIL2	Text	As above but for the next soil unit
PROP2	Integer	As above
PRID2	Text	As above
SOIL3	Text	As above but for the next soil unit
PROP3	Integer	As above
PRID3	Text	As above

Name	Type	Description
SOIL4	Text	As above but for the next soil unit
PROP4	Integer	As above
PRID4	Text	As above
SOIL5	Text	As above but for the next soil unit
PROP5	Integer	As above
PRID5	Text	As above
SOIL6	Text	As above but for the next soil unit
PROP6	Integer	As above
PRID6	Text	As above
SOIL7	Text	As above but for the next soil unit
PROP7	Integer	As above
PRID7	Text	As above
SOIL8	Text	As above but for the next soil unit
PROP8	Integer	As above
PRID8	Text	As above
SOIL9	Text	As above but for the next soil component
PROP9	Integer	As above
PRID9	Text	As above
SOIL10	Text	As above but for the next soil component
PROP10	Integer	As above
PRID10	Text	As above

[a] Generally, not all 10 available fields for $SOIL_i$ will be filled in SOTER. In the case of The Upper Tana, Kenya, , up to 4 different soil components have beend defined for each map unit.
[b] These codes have the following format: VRe2GLe4. The relative extent of each soil unit (e.g., VRe) 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.

Appendix 2. Structure of table SOTERparameterEstimates

Table SOTER*parameterEstimates* lists property estimates — depth-weighted by layer — for all soil units (represented by their PRID) that have been mapped for the study region. This information can be linked to the soil geographical data - in a GIS - through the unique profile code (PRID).

Structure of table SOTER*parameterEstimates*

Name	Type	Description
CLAF	Text	FAO-Unesco (1988) Revised Legend code
PRID	Text	profile ID (as listed in *SOTERmapunitComposition*)
Drain	Text	FAO soil drainage class
Layer	Text	code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm etc.)
TopDep	Integer	depth of top of layer (cm)
BotDep	Integer	depth of bottom of layer (cm)
CFRAG	Integer	coarse fragments (vol.% > 2 mm)
SDTO	Integer	sand (mass %)
STPC	Integer	silt (mass %)
CLPC	Integer	clay (mass %)
PSCL	Text	SOTER texture class (see Appendix 6)
BULK	Single	bulk density (kg dm^{-3})
TAWC	Integer	available water capacity (cm^3 cm^{-3} 10^2, -33 kPa to -1.5 MPa conform to USDA standards)
CECs	Single	cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	base saturation as percentage of CEC_{soil}
ESP	Integer	exchangeable Na as percentage of CEC_{soil}
CECc	Single	CECclay, corrected for contribution of organic matter ($cmol_c$ kg^{-1}) [a]
PHAQ	Single	pH measured in water
TCEQ	Single	total carbonate equivalent (g C kg^{-1})
GYPS	Single	gypsum content (g kg^{-1})
ELCO	Single	electrical conductivity (dS m^{-1})
TOTC [b]	Single	organic carbon content (g C kg^{-1})
TOTN	Single	total nitrogen (g N kg^{-1})
ECEC	Single	effective CEC ($cmol_c$ kg^{-1})
ALSA	Integer	exchangeable AL as percentage of ECEC

[a] CEC_{clay} is only calculated for layers where clay content >5%; else CEC_{clay} is set at -9 (see Appendix 3).

[b] Please note that TOTC is a field name used in SOTER representing organic carbon content only, not total carbon!

Contents of table SOTER*parameterEstimates* should be consulted in conjunction with table *SOTERflagTTRrules.* The later lists the taxotransfer rules that have been applied for each profile, by depth layer and soil attribute. Details are given in Appendix 3.

Appendix 3. Structure of table SOTERflagRules

Table SOTER*flagTTRrules* documents the type of taxotransfer rules that have been used to create table *SOTERparameterEstimates* (Appendix 2). Coding conventions are detailed in Table 5.

Structure of table *SOTERflagTTRrules*

Name	Type	Description
CLAF	Text	FAO Legend code
PRID	Text	Unique identifier for representative profile
Layer	Text	code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
Newtopdep	Integer	Depth of top of layer (cm)
Newbotdep	Integer	Depth of bottom of layer (cm)
TTRsub	Text	Code showing the type of taxotransfer rule used (based on derived data for *soil units*; see text)
TTRmain	Text	Code showing the type of taxotransfer rule used (based on derived data for *major units*; see text)
TTRexpert	Text	Additional flags (based on expert-rules)

Note: Expert rules (TTRexpert) are run after the TTR-procedures (see text). For example, exchangeable aluminium percentage (ALSA) has been set at zero when pH_{water} is higher than 5.5. Similarly, the content of gypsum (GYPS) and content of carbonates (TCEQ) have been set at zero when pH_{water} is less than 6.5. Finally, the CEC of the clay fraction (CEC_{clay}) has been re-calculated from the depth-weighted measured and TTR-derived data for CEC_{soil} and content of organic carbon assuming a mean contribution of 350 $cmol_c\ kg^{-1}$ OC, the common range being from 150 to over 750 $cmol_c$ per 100 g (Klamt and Sombroek 1988) — CEC_{clay} values presented here thus are only rough estimates.

Table 5. Conventions used for coding soil attributes in the taxotransfer scheme

TTRflag	SOTnam	WISnam	SoilVariable	Comments
A	ALSA	ALSA	ALSAT	Exch. Aluminum percentage (% of ECEC)
B	BSAT	BSAT	BSAT	base saturation (% of CECs)
C	BULK	BULK	BULKDENS	Bulk density
D	CECC	CECC	CECCLAY	cation exchange capacity of clay fraction
E	CECS	CECS	CECSOIL	cation exchange capacity
F	CFRAG	GRAV	GRAVEL	coarse fragments
G	CLPC	CLAY	CLAY	clay %
H	ECEC	ECEC	ECEC	Effective CEC
I	ELCO	ECE	ECE	electrical conductivity
J	ESP	ESP	ESP	exchangeable Na percentage (% of CECs)

TTRflag	SOTnam	WISnam	SoilVariable	Comments
K	GYPS	GYPS	GYPSUM	gypsum content (g C kg^{-1})
L	PHAQ	PHH2	PHH2O	pH in water
M	SDTO	SAND	SAND	sand %
N	STPC	SILT	SILT	silt %
O	AWC	AWC	AWC	Vol. water content (-33 kPa to -1.5 MPa)
P	TCEQ	CACO	CACO3	carbonate content (g kg^{-1})
Q	TOTC	ORGC	ORGC	organic carbon content (g C kg^{-1})
R	TOTN	TOTN	TOTN	total nitrogen content (g N kg^{-1})
Y	---	---	---	PSCL estimated from TTR-derived sand, silt and clay content (where applicable)

Abbreviations: TTRflag = code for TTR-rule; SOTnam = codes used in SOTER; WISnam= codes used in WISE; SoilVariable= soil variables as described in Table 2 (page 8).

Appendix 4. Structure of table SOTERsummaryFile

Table *SOTERsummaryFile* has been created to facilitate access to the derived data. For each SOTER unit (NEWSUID) on the map, it lists the soil property estimates by component soil unit and depth layer.

Layer data are presented in one single column, i.e. vertically (see also Appendix 5).

Structure of table SOTER*summaryFile*

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994)
SUID	Integer	The identification code of a SOTER on the map and in the database
NEWSUID	Text	Globally unique map unit code, comprising fields ISOC plus SUID
TCID	Integer	Number of terrain component in given map unit
SCID	Integer	Number of soil unit within the given SOTER unit
Layer	Text	Code for depth layer (from D1 to D7; e.g., D1 is from 0 to 20 cm and D7 from 150 to 200 cm)
PROP	Integer	Relative proportion of SCID in given SOTER unit
CLAF	Text	FAO-Unesco Revised Legend code
PRID	Text	Profile ID (see table *SOTERunitComposition*)
Drain	Text	FAO soil drainage class
TopDep	Integer	Upper depth of layer (cm)
BotDep	Integer	Lower dept of layer (cm)
CFRAG	Integer	Coarse fragments (vol. % > 2 mm)
SDTO	Integer	Sand (mass %)
STPC	Integer	Silt (mass %)
CLPC	Integer	Clay (mass %)
PSCL	Text	FAO texture class (see Appendix 6)
BULK	Single	Bulk density (kg dm^{-3})
TAWC	Integer	Available water capacity (cm^3 cm^{-3} 10^2 or vol%, -33 kPa to -1.5 MPa)
CECS	Single	Cation exchange capacity ($cmol_c$ kg^{-1}) of fine earth fraction
BSAT	Integer	Base saturation as percentage of CECsoil
ESP	Integer	Exchangeable Na as percentage of CECsoil
CECc	Single	CEC_{clay}, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	pH measured in water
TCEQ	Single	Total carbonate equivalent (g C kg^{-1})
GYPS	Single	Gypsum content (g kg^{-1})

Name	Type	Description
ELCO	Single	Electrical conductivity (dS m^{-1})
TOTC	Single	Organic carbon content (g kg^{-1})
TOTN	Single	Total nitrogen (g kg^{-1})
ECEC	Single	Effective CEC ($cmol_c$ kg^{-1})
ALSA	Integer	Exchangeable Al as percentage of ECEC

Notes:

1) The soil components that occur within a SOTER unit are numbered sequentially, starting with the spatially dominant one. The sum of the relative proportions of all component soil units is always 100 per cent. This total will also include a number of unnamed 'impurities', commonly in excess of 15 to 30 percent of the map unit (Landon 1991 p. 16-17; Marsman and de Gruijter 1986).
2) Each map unit in the geographic database has a unique identifier (NEWSUID) consisting of the country ISO code (ISOC) and the SOTER unit-ID (SUID); this primary key provides a link to the attribute data for the constituent terrain, terrain component(s) (TCID) and soil components (SCID) (see Figure 1).
3) Tables with the same structure have been prepared for the DOMINANT soil unit only, by depth layer (i.e., for layer D1, see for example table *SOTERsummaryFile_T1S1D1*) to facilitate visualization using GIS, as example only. Comprehensive studies, however, should consider the full map unit composition and depth range to 1 m.
4) A limited number of records may contain a negative value (-9); this indicates that it has not yet been possible to plug the corresponding gaps using the current taxotransfer scheme due to a lack of measured data in WISE. Whenever possible, virtual profiles in SOTER should be replaced with real, measured profiles after which new secondary data may be generated. A value of '-8' is used for water bodies or SOTER unit 'TN-9'.
5) Property estimates are depth-weighted values, per 20 cm layer up to 1m depth and per 50 cm from 1 to 2 m (derived soil properties for 100 to 200 cm, however, are not included in the present secondary database, see text).

Appendix 5. Structure of table SOTERsummaryFile_Prop

The field definitions in this table are identical to those used in *SOTERsummaryFile.* The main difference is that derived data for each soil component of a given SOTER unit are now listed in a single row (horizontally); data for a given layer are preceded by a flag for this layer. For example, field *D1_BULK* presents derived values for bulk density for layer D1 (0-20 cm), whereas *D2-BULK* holds data for layer D2 (20-40 cm) and so on. Using this file format, it is easier to query properties of the individual component soil units of a SOTER units using GIS. However, results can only be shown for one soil component, by SOTER unit, at a time (e.g. for TCID=1 and SCID=1).

Structure of table SOTER*summaryFile_Prop*

Name	Type	Description
ISOC	Text	ISO-3166 country code (1994)
SUID	Integer	The identification code of a SOTER on the map and in the database
NEWSUID	Text	Globally unique map unit code, comprising fields ISOC plus SUID
TCID	Integer	Number of terrain component in given map unit
SCID	Integer	Number of soil unit within the given SOTER unit
PROP	Integer	Relative proportion of SCID in given SOTER unit
CLAF	Text	FAO-Unesco Revised Legend code
PRID	Text	Profile ID (see table *SOTERunitComposition*)
Drain	Text	FAO soil drainage class
D1_TopDep	Integer	Upper depth of layer D1 (0-20 cm)
D1_BotDep	Integer	Lower dept of layer D1
D1_varx	Variable	Values (e.g., varx is ORGC, BULK, Clay) for layer D1
D2_TopDep	Integer	Upper depth of layer D2 (20-40 cm)
D2_BotDep	Integer	Lower dept of layer D2
D2_varx	Variable	Values (e.g, varx is ORGC, BULK, Clay) for layer D2
...	...	...
D5_TopDep	Integer	Upper depth of layer D5 (80-100 cm)
D5_BotDep	Integer	Lower dept of layer D5 (0-20 cm)
D5_varx	Variable	Values (e.g, varx is ORGC, BULK, Clay) for layer D5

Note: A table with the same structure has also been prepared for the DOMINANT soil unit only (i.e., TCID= 1 and SCID=1) to facilitate visualization using GIS, as example only (see table *SOTERsummaryFile_PROP_SC1*). Comprehensive studies, however, should always consider the full map unit composition and depth range.

Appendix 6. Soil textural classes

Soil textural classes (PSCL) are in accordance with revised SOTER criteria (Figure 5). The following abbreviations are used: C–coarse, M–medium, Z–medium fine, F–fine and V–very fine. Further, the symbol *u* is used for undifferentiated (i.e., C + M + F + Z + V). In addition, all Histosols data have been flagged as consisting of organic materials (O) even though this may not always be the case for all horizons/layers, in a strict taxonomic sense (see FAO 1988 , p. 39)

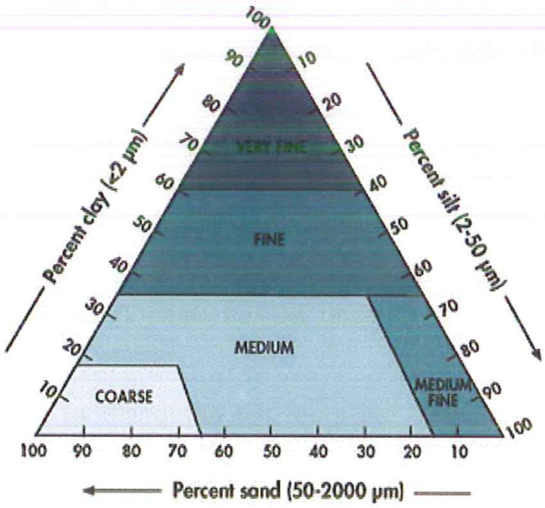


Figure 5. SOTER soil texture classes (Source: CEC 1985)

Appendix 7. Installation

The derived soil data and GIS-files are presented in one single zip file: *SOTWIS_Kenya_UpperTana_v1.zip.*

By default, this compressed file will be unzipped to folder *X:\SOTWIS_Kenya_UpperTana,* where *X* is the actual location (i.e. folder).

This new folder will contain:
- A *Readme1st* file and the documentation (ISRIC Report 2010/0X)
- The project file (*SOTWIS_Kenya_UpperTana_v1.mxd*) with metadata (*SOTWIS_Kenya_UpperTana_v1.mxd.xml)*
- Two subfolders:
  - *GISfiles* with the shape and selected layer files files.
  - *SOTWIS* with the derived soil data in MSAccess® format (*SOTWIS_Kenya_UpperTana_v1.mdb*).

The GIS project file (*.mxd) includes several derived data sets for the top layer (0-20 cm) of the dominant soil unit of each SOTER unit (TCID=1, SCID=1), as examples.

Actual data applications should consider the full map unit composition, in terms of component soil units, and depth range; see text for details.

The dataset has been created using MS-Access® and ArcGIS9/ArcMap9.3®; the shapefiles may also be accessed using ArcView3.3®.

World Soil Information

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

Our aims:

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