Collection of soil data in SOTER format from 14 Danube strategy countries, at scale 1:250 000

EU-JRC Service Contract #936108

Deliverables D.A1, D.A2 and DA4

Updated MS-Access SOTER template for the EU Danube basin project with worked examples

Niels H. Batjes and Eloi Ribeiro

February 2019

ISRIC World Soil Information

Cranfield University

Contents

List of figures

1. Introduction

To support the Danube strategy the European Commission Joint Research Centre (JRC) has proposed soil data collection within the Danube basin based on SOTER methodologies. The SOTER™ (SOil and TERrain) methodology is a comprehensive framework making soil and terrain data available for end users. The methodology was developed by ISRIC in collaboration with FAO, UNEP, JRC and a range of national partners (Van Engelen 2011; van Engelen and Dijkshoorn 2013) and implemented in a range of countries[1].

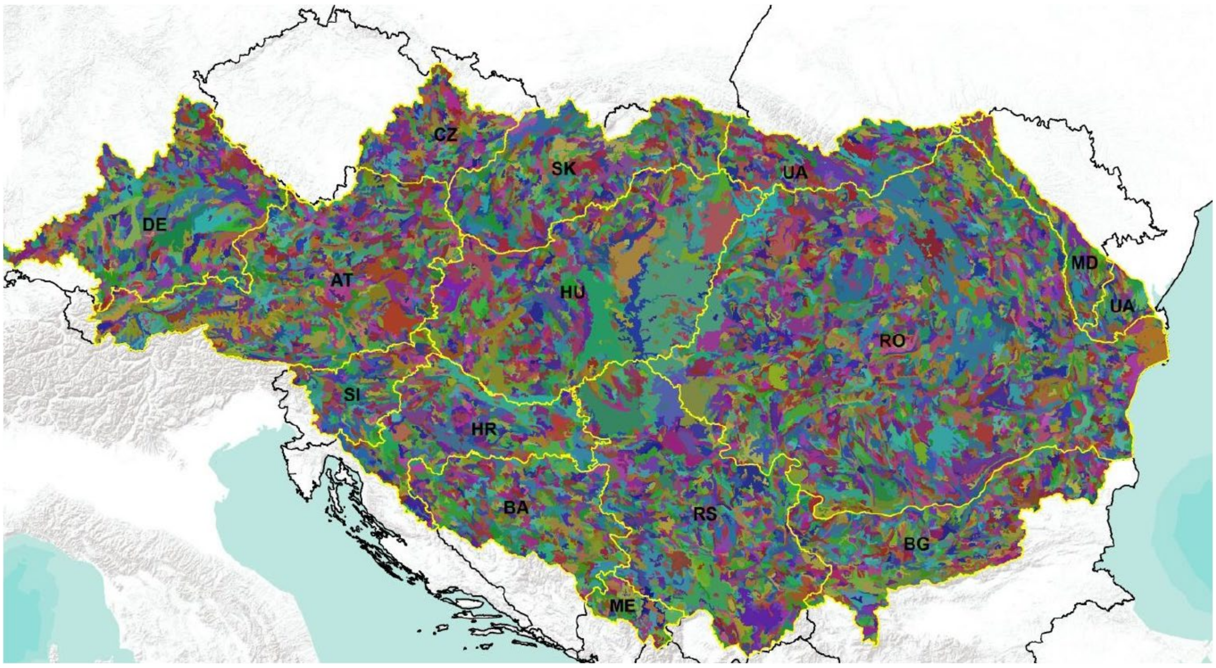


Figure 1. Danube region basin

The 'SOTER Danube Basin' project is led by Cranfield University (Centre for Environmental and Agricultural Informatics) in partnership with ISRIC and 14 regional partners from countries in the Danube basin (Figure 1). These are: Austria (AU), Bosnia and Herzegovina (BA), Bulgaria (BG), Czech Republic (CZ), Germany (DE), Hungary (HU), Croatia (HR), Moldova (MD), Montenegro (ME), Romania (RO), Serbia (RS), Slovakia (SK), Slovenia (SI), and Ukraine (UA).

Use of the SOTER methodology ensures standard procedures for data harmonisation are applied in the contributing countries. This meets the requirements under Pillar 4 of the Global and European Soil Partnership (GSP and FAO 2016) and compliance with INSPIRE (2015) data protocols. The production of a harmonised SOTER database at 1:250,000 scale is aimed to support soil and land use policy development in the Danube basin region.

In short, the project includes preparing 'empty' SOTER database templates and GIS files of terrain units mapped for the participating countries, filling the 'national' SOTER databases with representative soil profiles, standardisation and harmonisation to the World Reference Base for Soil Resources (IUSS Working Group WRB 2015), and ultimately exporting data in SOTER format to GML INSPIRE-compliant format.

[1] https://www.isric.org/projects/soil-and-terrain-soter-database-programme

This report describes the template for the SOTER database and forms for entering data with two worked examples; this corresponds with deliverables D.A1, D.A2 and D.A4 of Work Package 2 (WP2) of the project. The tool is presented in Microsoft-Access[2] format, in accordance with the SOTER procedures Manual (van Engelen and Dijkshoorn 2013), with some structural changes required to improve the database structure itself (Appendix 1), but not the SOTER conventions themselves (Appendix 2).

Worked examples are presented in Appendix 5 for two cases. The first example is for a 'conventional' SOTER database with compound map units; each SOTER (terrain) unit may be comprised of several terrain components having one to several soil components. Each of these soil components is then represented by a regionally representative profile in the database. The second example is for a case where all SOTER units are considered to consist of one terrain component having one or more soil components. Again, each of these soil components is characterised by a representative soil profile. Conceptually, the second approach is similar to the approach used for the eSOTER project[3] and it has been adopted for the Danube basis project.

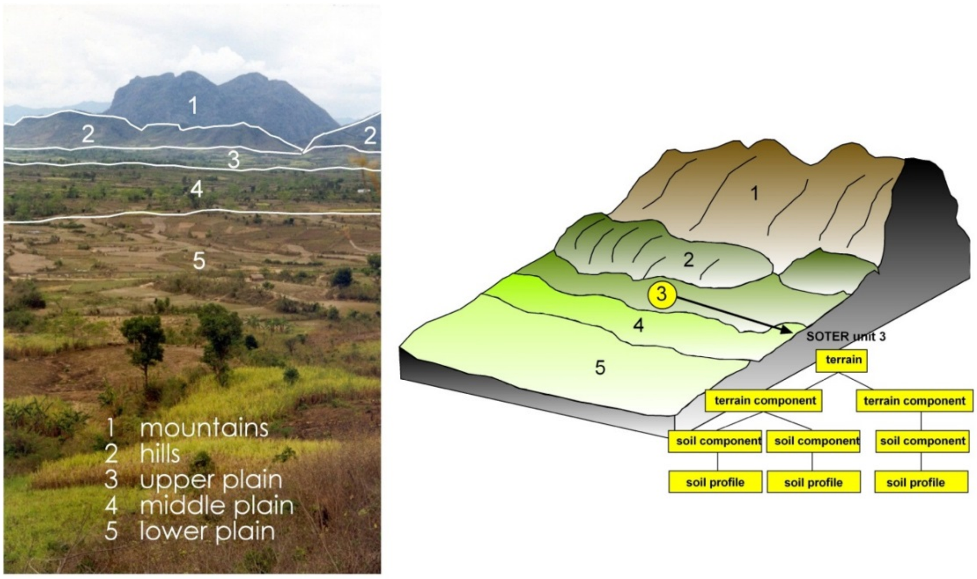
[2] The blank SOTER database structure is presented in MS Access 2002-2003 (mdb) format to ensure overall compatibility with the 14 Danube basin countries and functionality with ArcMap, as agreed with the Project Coordinator. In addition to this, an MsAccess 2010 version (accdb) of the database is also available

[3] https://esdac.jrc.ec.europa.eu/projects/esoter-danube

2. SOTER methodology

SOTER is a land resources information system based on the concept whereby features of the land - in which terrain and soil occur - incorporate processes and systems of interrelationships between physical, biological and social processes over time (van Engelen and Dijkshoorn 2013).

Central to the SOTER methodology is the identification (delineation) of areas of land with a distinctive, often repetitive, pattern of landform, lithology, surface form, slope, parent material, and soil. These are named SOTER units. Each SOTER unit (polygon in the GIS) thus represents one unique combination of terrain and soil characteristics. Figure 2 represents a landscape with five different SOTER units (van Engelen and Dijkshoorn 2013). SOTER unit 3, for example, consists of one terrain type, consisting of an association of two terrain components, the first having two soil components and the second one soil component. Each soil component is characterised using a regionally representative soil profile. The SOTER terrain units for the Danube Basin were mapped earlier, at a scale of 1:250,000, using a methodology similar to the one developed during the eSOTER EU-FP7 project (Pourabdollah *et al.* 2012).[4]



Unit	*SOTER description*
1	one terrain type with one terrain component and one soil component
2	one terrain type consisting of an association of two terrain components each having a particular soil component
3	one terrain type, consisting of an association of two terrain components, the first having two soil components and the second one soil component. Each soil component is characterised using a regionally representative soil profile.
4	one terrain type, consisting of an association of two terrain components, the first having one soil component, the second having an association of three soil components
5	one terrain type with one terrain component, having an association of two soil components

Figure 2. Relation between SOTER units and their composing parts as characterised in the database

[4] https://www.esoter.net/

Each SOTER database is composed of two components:

1) Geometric component: Describes the location and extent of each SOTER unit (polygon) and its topology (shapes, neighbours and hierarchy of delineations). The corresponding information is managed using GIS (Geographic information System), in casu ArcMap (ESRI®) or QGIS (open source).

2) Attribute component: Specifies the characteristics of the geometric object (i.e. given SOTER unit). These attributes, as itemised below and in Appendix 2, are managed in a Relational Database Management System (RDBMS*), in casu* an MS-Access® database.

Information held in the GIS and RDBMS can be linked through the unique label (SUID) for the given polygon or SOTER unit ID (Figure 3). In transnational SOTER databases, unique labels for each polygon will consist of the country code (ISOC) and terrain number (SUID). This field in known as ISOCSUID.

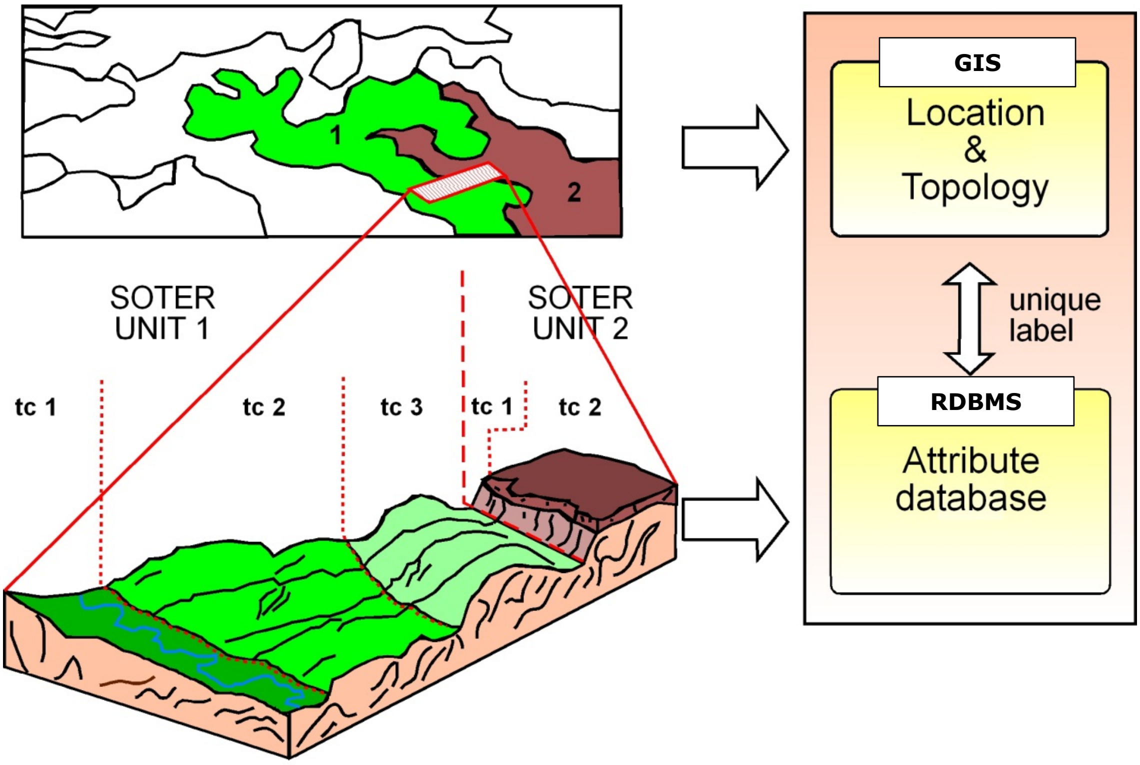


Figure 3. Schematic representation of SOTER database with its geographic and attribute component

Relational databases are one of the most effective and flexible tools for storing and managing spatial and non-spatial attributes. Data are stored in tables, and records are related to each other through specific identification fields (primary keys), such as the SOTER unit identification code (SUID). These codes are essential as they form the logical link between the various tables of the database: **terrain**, **terrain component**, **soil component**, **profile**, **horizon** and related tables (e.g. **laboratory**, **laboratory methods**, **source map**, **codes** see Appendix 3). As such, the primary keys and table formats should not be tampered with; otherwise, the different national SOTER databases cannot be merged into one single product for the Danube basis at a later stage.

Another important characteristic of a relational database is that when two or more components are similar, the associated attribute data need only to be entered once (i.e. they can be referred to through their unique identifiers). This feature is particularly useful, for example, to provide keys to codes used for describing e.g. land use or vegetation classes (see Appendix 4).

Generally, at the considered scale of 1:250,000, a ‘conventional’ SOTER unit will consist of 3 or 4 terrain components, each of these having no more than 3 soil components. It is recommended that the proportion of each terrain component is at least 15%, and each soil component at least 10%, of the corresponding SOTER unit (van Engelen and Dijkshoorn 2013, p. 47). By definition, the total proportion of all soil components within each terrain component adds up to the total proportion of the terrain component within the given SUID. Alternatively, the sum of the relative area (proportion) of the terrain components representing a given SOTER unit will always be 100%.

As a rule of thumb, the minimum size of a single SOTER unit is 0.25 cm^2 on the map (Van Engelen and Dijkshoorn, 2013, p. 29). At a scale of 1:250,000 this corresponds with ~1.6 km^2 in the field. This is the smallest area that can still be cartographically presented at the given scale. However, this limit may be too restrictive for narrow elongated features such as floodplains, ridges and some valleys. Therefore, often the minimum size of 6.25 km^2 is practiced.

Representative soil profiles

As indicated, each soil component is to be characterised by a so-called representative soil profile. Typically, such representative profiles will be selected by local experts from a number of reference profiles having similar characteristics in terms of the WRB Soil Reference Base (IUSS Working Group WRB 2015) classification. For this, SOTER will rely on the selection of reference profiles available for the region under consideration. It is considered good practice that all these reference profiles be stored in the SOTER database, respectively managed in a national soil profile database.

Sometimes, when data are limited, the same reference profile (e.g. for a given WRB 2015 legend unit) may be used to characterise different soil components and SOTER units (see van Engelen and Dijkshoorn 2013, chapter 6.5).

3. SOTER database model for the Danube region

Several changes were made to the '2013 SOTER' database[5] model to improve data storage and handling, as shown in Figure 4. In particular, the 2015 update of the WRB is now used as default system for defining the legend and classifying soil profiles. These changes are summarised in Appendix 1.

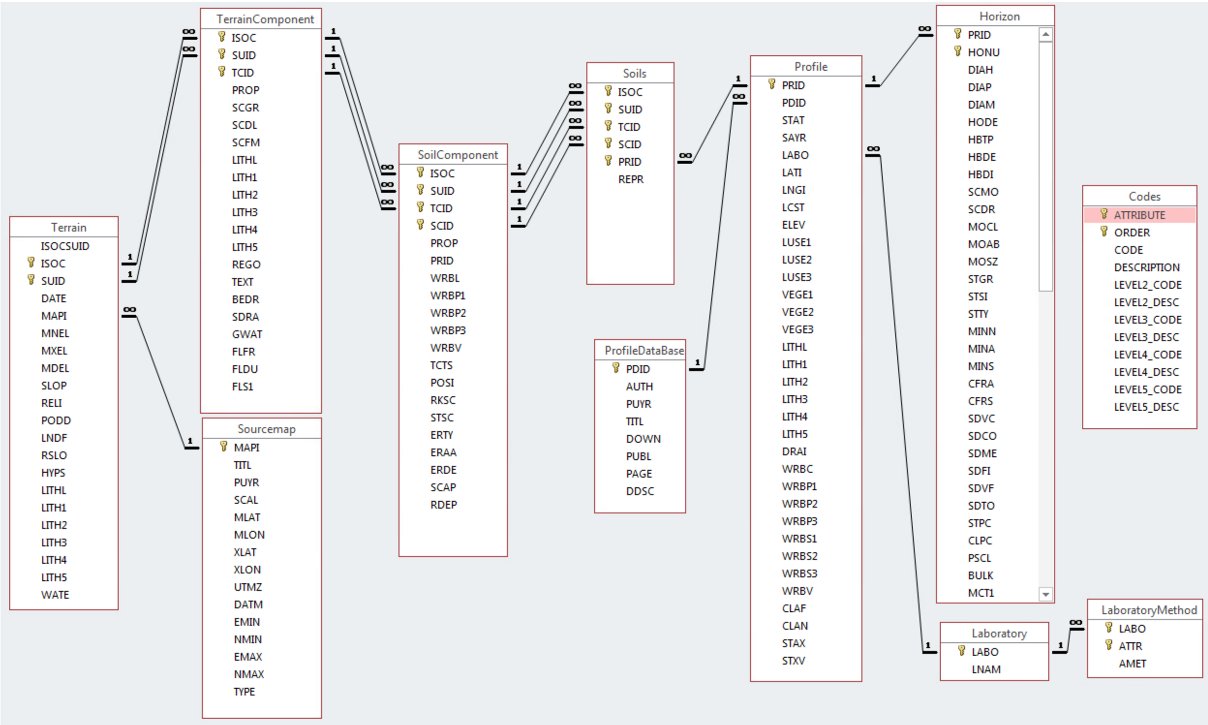


Figure 4. SOTER database structure as modified for EU Danube basin project

The attributes themselves, however, have not been adapted except where necessary to reflect recent changes in soil classification (i.e. WRB2015 versus the earlier WRB2006 version) or some country ISO codes since the 2013 SOTER database release.

Codes and brief descriptions for the complement of attributes for each table shown in Figure 4 are detailed in Appendix 2. The rightmost column in each table shows the corresponding (explanatory) pages and item numbers in the SOTER 2013 Procedures Manual (van Engelen and Dijkshoorn 2013); this information is mainly provided here to facilitate the data compilers.

As indicated, worked examples for two test cases, a 'conventional' and a 'simplified' SOTER database, are presented for an hypothetical country in Appendix 5, this mainly to describe the general workflow. Detailed guidelines for compiling the upcoming 'national' SOTER attribute databases for the Danube project are presented in a separate report (Batjes and Ribeiro 2019).

[5] http://www.isric.org/soter-data-model-v1

4. Next steps

The first component of the Danube basin project fine-tuned the SOTER database model. Worked examples were prepared for two cases, a 'conventional' and a 'simplified' SOTER database to illustrate the overall procedures and workflows. These activities correspond with project deliverables D.A1, D.A2 and D.A4.

In conjunction with the above, terrain unit shapefiles for each of the 14 countries represented in the Danube basin, with documentation, are being generated (D.A3) as a basis for the actual database compilation.

Technical guidelines for data providers to populate their 'national' SOTER database will be developed next for the 'one terrain component with multiple soil components' case (D.A5).

Filling of the 'national' SOTER attribute database(s) themselves will be done by the regional partners, as subcontracted by Cranfield University, in the context of Work Package 3 on data compilation and WRB correlation.

Meanwhile, using the example from D.A5, ISRIC will develop an application to generate GML files from SOTER according to INSPIRE specifications. At a later stage in the project (Work package 3, month 18), ISRIC will generate Geography Mark-up Language (GML) files for the SOTER's compiled by the 14 partners, and the centrally merged SOTER database for the whole Danube basin (D.A6 and D.B2). This in order to assess SOTER compliance with INSPIRE requirements.

Acknowledgments

Funding for the SOTER Danube Basin project was obtained through EU-tender RC/IPR/2017/D.3/0002/OC. The project is implemented under EU-JRC Service Contract #936108.

The SOTER Danube basin consortium is led by Cranfield University (Centre for Environmental and Agricultural Informatics) in partnership with ISRIC – World Soil Information and 14 regional partners from countries in the Danube basin (Austria (AU), Bosnia and Herzegovina (BA), Bulgaria (BG), Czech Republic (CZ), Germany (DE), Hungary (HU), Croatia (HR), Moldova (MD), Montenegro (ME), Romania (RO), Serbia (RS), Slovakia (SK), Slovenia (SI), and Ukraine (UA).

We thank Dr Jacqueline Hannam and Dr Stephen Hallett (Cranfield University) for their editorial comments, and Ir Mark van Liedekerke (EU-JRC) for useful suggestions concerning the worked examples.

References

Baritz R, Erdogan H, Fujii K, Takata Y, Nocita M, Bussian B, Batjes NH, Hempel J, Wilson P and Vargas R 2014. *Harmonization of methods, measurements and indicators for the sustainable management and protection of soil resources (Providing mechanisms for the collation, analysis and exchange of consistent and comparable global soil data and information)*, Global Soil Partnership, FAO, 44 p. http://www.fao.org/3/a-az922e.pdf

Batjes NH and Ribeiro E 2019. *Guidelines for compiling a 1:250,000 SOTER database (Deliverabe A5, EU JRC Service Contract #936108, Collection of soil data in SOTER format from 14 Danube strategy countries, at scale 1:250 000)*, ISRIC - World Soil Information, Wageningen

Batjes NH, Ribeiro E, van Oostrum A, Leenaars J, Hengl T and Mendes de Jesus J 2017. WoSIS: providing standardised soil profile data for the world. *Earth Syst. Sci. Data* 9, 1-14. http://dx.doi.org/10.5194/essd-9-1-2017

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

Dijkshoorn JA, Leenaars JGB, Huting J and Kempen B 2016. *Soil and terrain database for Malawi*, ISRIC - World Soil Information, Wageningen, 38 p. http://dx.doi.org/10.17027/isric-wdcsoils.20160001

Dobos E, Daroussin J and Montanarella L 2005. *An SRTM-based procedure to delineate SOTER terrain units on 1:1 and 1:5 million scales*. EUR 21571 EN, European Communities, Institute of Environment & Sustainability, Joint Research Centre, Ispra (IT), 62 p. http://eusoils.jrc.it/ESDB_Archive/eusoils_docs/other/SOTER/SOTER_endre.pdf

FAO 1990. *Guidelines for soil description (3rd rev. ed.)*, FAO, Rome, 45 p

GSP and FAO 2016. *Pillar 4 Implementation Plan -- Towards a Global Soil Information System*, Rome, 36 p. http://www.fao.org/3/a-bl102e.pdf

Hannam JA, Hollis JM, Jones RJA, Bellamy PH, Hayes SE, Holden A, Van Liedekerke MH and Montanarella L 2009 *SPADE-2: Soil Profile Analytical Database for Europe (Version 2.0)*. https://esdac.jrc.ec.europa.eu/Esdb_Archive/eusoils_docs/esb_rr/SPADE-2_Beta_Report.pdf

INSPIRE 2015. INSPIRE Data specifications - Infrastructure for spatial information in the European Community, European Commission Joint Research Centre. http://inspire.ec.europa.eu/index.cfm/pageid/2

IUSS Working Group WRB 2006. *World Reference Base for Soil Resources (2nd ed.)*. World Soil Resources Report 103, FAO, Rome, 145 p. http://www.fao.org/ag/agl/agll/wrb/doc/wrb2006final.pdf

IUSS Working Group WRB 2015. *World Reference Base for soil resources 2014 - International soil classification system for naming soils and creating legends for soil maps (update 2015)*. World Soil Resources Reports 106, Global Soil Partnership, International Union of Soil Sciences, and Food and Agriculture Organization of the United Nations, Rome, 182 p. http://www.fao.org/3/i3794en/I3794en.pdf

Pourabdollah A, Leibovici DG, Simms DM, Tempel P, Hallett SH and Jackson MJ 2012. Towards a standard for soil and terrain data exchange: SoTerML. *Computers and Geosciences* 45, 270-283. http://www.sciencedirect.com/science/article/pii/S0098300411004195

Ruiperez Gonzalez M and Batjes NH 2019. *Country terrain unit shapefiles of the EU Danube basin (Deliverable A3 EU JRC Service Contract No. 936108, Collection of soil data in SOTER format from 14 Danube strategy countries, at scale 1:250 000)* ISRIC - World Soil Information, Wageningen

Soil Survey Staff 2014. *Keys to Soil Taxonomy, 12th ed.* USDA-Natural Resources Conservation Service, Washington, DC.https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=stelprdb1252094&ext=pdf

Van Engelen VWP 2011. Standardizing soil data (*e-SOTER* regional pilot platform as EU contribition to a Global Soil Information System). *International Innovation* 6, 48-49. https://cordis.europa.eu/project/rcn/88557/factsheet/en

van Engelen VWP and Dijkshoorn JA 2013. *Global and National Soils and Terrain Digital Databases (SOTER) - Procedures manual (Ver. 2.0)*. ISRIC Report 2013/04, IUSS, ISRIC and FAO, Wageningen, 198 p. http://www.isric.org/sites/default/files/isric_report_2013_04.pdf

Appendices

Appendix 1. Changes to 2013 SOTER database structure

After consideration, it was concluded that the '2013 SOTER'[6] database model should be simplified to facilitate more efficient handling of the data and in some cases reinforce the referential integrity. The following changes were introduced.

Columns added

Terrain table, added column ISOC (country ISO code) and column ISOCSUID, the combined (unique) code for a terrain unit as represented in the GIS shapefiles.

SoilComponent table, added ISOC column. Further added columns WRBL (for Legend), WRBP1, WRBP2 and WRBP3 to accommodate soil legend descriptions for 1:250,000 scale soil maps (see IUSS Working Group WRB 2015, p. 14). The full legend code will consist of the concatenated strings for WRBL (Reference Soil Group), followed by the first three applicable principal qualifiers (WRBP1 to WRBP3). The most important principal qualifier (WRBP1) is listed closest to the RSG.

Soils table, added REPR (Representative (true/false)) column. When flagged as 'true', this shows that the given profile (with identifier PRID) was selected as the typical profile for the corresponding Soil Component.

Profile table, column SAYR now stands for date (yyyy-mm-dd) instead of only for year and columns WRBC (for RSG classification), WRBP1, WRBP2, WRBP3 for principal qualifiers, and WRBS1, WRBS2 and WRBS3 for supplementary qualifiers. These columns provide the 'building blocks' for the full classification according to WRB 2015. The most important principal qualifier (WRBP1) is listed closest to the RSG; by convention, the supplementary qualifiers are listed alphabetically (see IUSS Working Group WRB 2015, p. 85-116 and Chapter 2).

Profile and **Soil Component** table, added WRBV (WRB version, year) column. Note that the default value is 2015 for this project.

Tables merged

TerrainComponent and **TerrainComponentData** tables were merged.

TerrainComponent, deleted column TCDC (Terrain component data id).

TerrainComponentData , deleted column TCDC (Terrain component data id).

LaboratoryMethod and **AnalyticalMethod** tables were merged.

LaboratoryMethod, added column AMET (Description).

LaboratoryMethod, deleted column AMID (Analysis method id).

AnalyticalMethod, deleted column AMID (Analysis method id).

Updated Codes table

The **codes** table was modified to account for recent changes in the SOTER methodology and to better handle the querying of compound codes.

Major landforms (LNDF): Deleted options level land (L), sloping land (S), steep land (T). Hierarchy now starts with (original) second level coding (see SOTER Procedures Manual 2013 p.35, Table 2).

[6] http://www.isric.org/soter-data-model-v1

Parent material (<u>LITH</u>): The original <u>LITH</u> class was subdivided into five classes (LITH1, LITH2, LITH3, LITH4, and LITH5) to better reflect the hierarchy of the system/coding.

Added columns LITH1, LITH2, LITH3, LITH4, LITH5 to table **Terrain**.

Added columns LITH1, LITH2, LITH3, LITH4, LITH5 to table **TerrainComponent**.

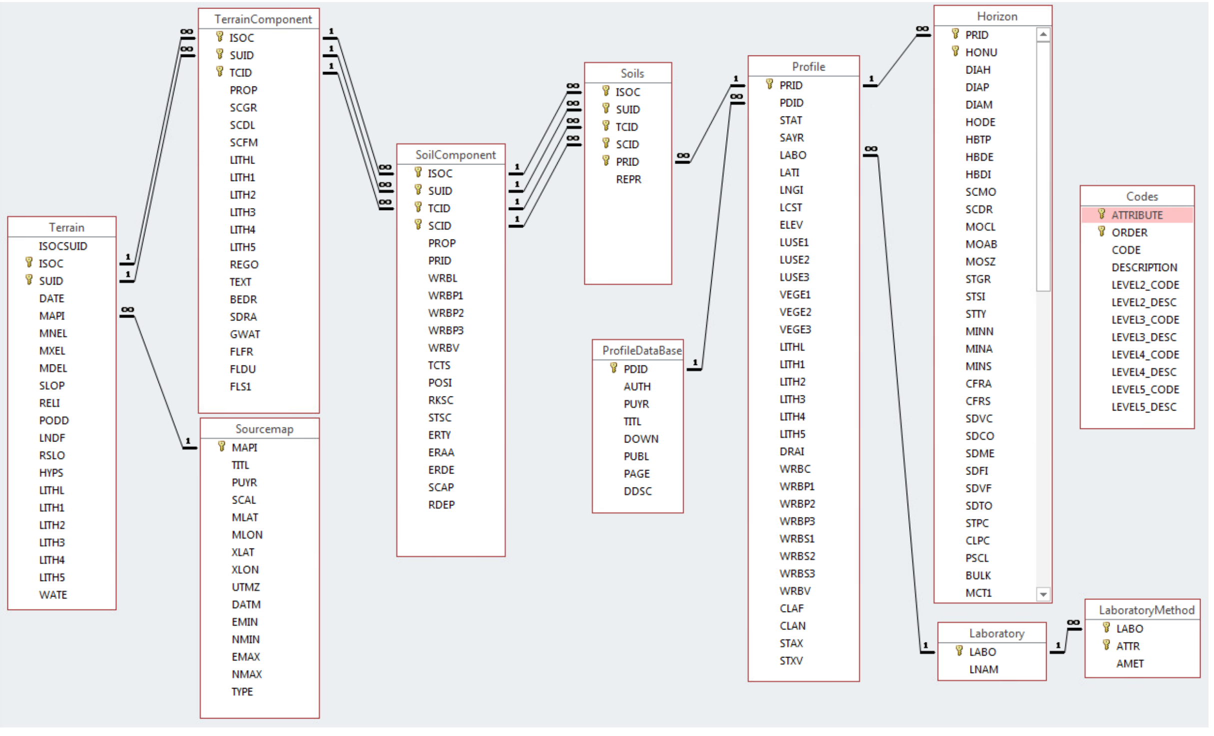
Added columns LITH1, LITH2, LITH3, LITH4, LITH5 to table **Profile**.

<u>Note</u>: Column LITHL is not used in SOTER Danube, though still present as a (redundant) column in table **Terr**ain, **TerrainComponent** and **Profile**. This column should always be blank.

Added columns WRBL, WRBP1, WRB2, WRB3 to table **soil component,** respectively WRBC, WRBP1, WRB2, WRB3, WRBS1, WRBS2, and WRBS3 to table **Profile** to accommodate the WRB 2015 Legend (WRBL), RSG classification (WRBC), and principal (WRBPi) as well as supplementary (WRBSi)qualifiers as appropriate (IUSS Working Group WRB 2015, p. 85-116).

Database model

The structure of the SOTER database model, as modified for the EU Danube Basin project, is repeated below for easy consultation in combination with the tables listed in Appendix 2 (i.e., same as Figure 4 in body of report).



Appendix 2. Short and long names for non-spatial attributes

Terrain

Field name short	**Field name long**[7]	**Data type**	**Description**	**SOTER 2013**
ISOC[8]	ISO country code	Short text	ISO-3166 country code	p33, 1
SUID	SOTER unit_ID	Number	The identification code of a SOTER unit on the map and in the database	p33, 2
ISOCSUID[9]	--	Short text	Concatenation of ISOC and SUID (as text). Provides the unique identifier for linkage to the GIS shapefiles	-
DATE	Year of data collection	Number	The year in which the original terrain data were collected	p33, 3
MAPI	Source map-ID	Short text	The source map identification code from which the data were derived	p34, 4
MNEL	Minimum elevation	Number	Absolute minimum elevation of the SOTER unit, in metre above sea level	p34, 5
MXEL	Maximum elevation	Number	Absolute maximum elevation of the SOTER unit, in metre above sea level	p34, 6
MDEL	Median elevation	Number	Median elevation, in metres above sea level	p34, 7
SLOP	Slope gradient	Number	The dominant slope angle, as a percentage, prevailing in the terrain	p34, 8
RELI	Relief intensity	Number	The median difference between the highest and lowest point within the terrain per specified distance (m/km)	p34, 9
PODD	Potential drainage density	Number	Potential drainage density - an index for the degree of dissection of the SOTER unit (Dobos *et al.* 2005)	p34,10
LNDF	Major landform	Short Text	Landforms are described foremost by their morphology, not by their genetic origin	p34, 11
RSLO	Regional slope	Short Text	A refining of slope classes compared to those used for major landforms	p36, 12
HYPS	Hypsometry	Short Text	The hypsometric level is an indication of the height above sea level of the local base level	p36, 13
WATE	Permanent water surface	Number	Percentage of the SOTER unit that is largely (> 90%) permanently (> 10 months / year) covered by water	p46, 15
LITH1[10]	General lithology, level 1	Short Text	A generalised description of the (un)consolidated surficial material, underlying the *larger par*t of the terrain, level 1.	p 37, 14; p37-46, level 1
LITH2	General lithology, level 2	Short Text	A generalised description of the (un)consolidated surficial material, underlying the *larger part* of the terrain, level 2	p37-46, level 2
LITH3	General lithology, level 3	Short Text	A generalised description of the (un)consolidated surficial material, underlying the *larger part* of the terrain, level 3	p37-46, level 3
LITH4	General lithology, level 4	Short text	A generalised description of the (un)consolidated surficial material, underlying the *larger part* of the terrain, level 4	p37-46, level 4

[7] Field names as given on the data entry forms.
[8] Primary keys are indicated in *italics*; these are used to ensure referential integrity within the database.
[9] ISOCSUID can be filled using the following SQL query:
UPDATE Terrain SET Terrain.ISOCSUID = [ISOC]+LTrim(Str([SUID]));
[10] In Form mode, the appropriate options at level 2 for say LITH2 will be shown in the 'pull-down' menus based on queries of the **codes** table, the central 'look-up' table. In the view mode, this will only work for level 1 codes.

Field name short	Field name long[7]	Data type	Description	SOTER 2013
LITH5	General lithology, level 5	Short text	A generaliseised description of the (un)consolidated surficial material, underlying the *larger p*art of the terrain, level 5	p37-46, level 5

Terrain component

Field name short	Field name long[11]	Data type	Description	SOTER 2013
ISOC	ISO country code	Short text	ISO-3166 country code	p33, 1
SUID	SOTER unit-ID	Number	The identification code of a SOTER unit on the map and in the database	p33, 2
TCID	Terrain component number	Number	The sequence number of the terrain component in the terrain (largest comes first) (For the Danube SOTER DB this is always one)	p. 46, 17
PROP	Proportion	Number	The proportion, as a percentage, that the terrain component occupies within the SOTER unit (Note: Normally at least 15% of the SOTER unit) (For the Danube SOTER DB this is always 100)	p.47, 18
SCGR	Dominant slope	Number	Dominant slope gradient of the terrain component, as a percentage	p48, 21
SCDL	Dominant slope length	Number	Estimated dominant length of slope (metres)	p48, 22
SCFM	Form of dominant slope	Short text	The form of the dominant slope	p48, 23
LITH1 [12]	(Un)consolidated surficial materials, level 1	Short Text	A generalised description of the (un)consolidated surficial material, underlying the larger part of the terrain component, level 1	p48, 24; p39-43
LITH2	(Un)consolidated surficial materials, level 2	Short Text	A generalised description of the (un)consolidated surficial material, underlying the larger part of the terrain component, level 2	p39-43
LITH3	(Un)consolidated surficial materials, level 3	Short Text	A generalised description of the (un)consolidated surficial material, underlying the larger part of the terrain component, level 3	p39-43
LITH4	(Un)consolidated surficial materials, level 4	Short text	A generalised description of the (un)consolidated surficial material, underlying the larger part of the terrain component, level 4	p39-43
LITH5	(Un)consolidated surficial materials, level 5	Short text	A generalised description of the (un)consolidated surficial material, underlying the larger part of the terrain component, level 4	p39-43
REGO	Origin of parent material	Short text	Origin of non-consolidated parent material (regolith)	p48, 24
TEXT	Texture	Short text	Texture of non-consolidated parent material	p48, 26
BEDR	Average depth	Number	The average depth to consolidated bedrock in metres	p49, 27
SDRA	Surface drainage	Short text	Surface drainage of the terrain component	p49, 28
GWAT	Depth to groundwater	Short text	Depth (metres) of the mean ground water level	p50, 29
FLFR	Frequency of the natural flooding	Short text	Frequency of the natural flooding of the terrain component in classes after (FAO 1990)	p50, 30
FLDU	Duration of flooding	Short text	Duration of the flooding of the terrain component in classes after (FAO 1990)	p50, 31

[11] Field names as given on the data entry forms.
[12] For the first (largest) terrain component in a given SUID, the codes for LITH should be the same as those given in table **terrain** for this SUID (i.e. the predominant lithology).

Field name short	Field name long[11]	Data type	Description	SOTER 2013
FLS1	1st month during which flooding starts	Number	First month during which flooding of the terrain component normally starts	p51, 32

Soil component

Field name	Field name long[13]	Data type	Description	SOTER 2013
ISOC	ISO country code	Short text	ISO-3166 country code	p33, 1
SUID	SOTER ubit_ID	Number	The identification code of a SOTER unit on the map and in the database	p33, 2
TCID	Terrain component number	Number	The sequence number of the terrain component in the terrain (**largest comes first**)	p51, 34
SCID	Soil component number	Number	The sequence number of the soil component in the terrain component (**largest comes first**)	p51, 36
PROP	Proportion	Number	The proportion, as a percentage, that the soil component occupies within the SOTER unit (Normally >10% of a SOTER unit)	p51, 36
PRID	Profile-ID	Number	Code (ID) for the **representative** profile	p53, 42
WRBL	WRB Legend RSG	Short text	World Reference Base – RSG (Reference Soil Group) for Legend unit[14], see WRB 2015, p. 14-16	-
WRBP1	WRB legend principal qualifier 1	Short text	World Reference Base – principal qualifier 1, see WRB 2015, p. 14-16 and 85-116	-
WRBP2	WRB principal qualifier 2	Short text	World Reference Base – principal qualifier 2, see WRB 2015, p. 14-16 and 85-116	-
WRBP3	WRB principal qualifier 3	Short text	World Reference Base – principal qualifier 3, see WRB 2015, p. 14-16 and 85-116	-
WRBV	WRB Legend year	Number	World Reference Base – version (year, 2015 by default)	-
TCTS	Textural class	Short text	Textural class of the topsoil (CEC 1985), SOTER 2013 Fig. 9	p52, 41
POSI	Position	Short text	The relative position of the soil component within the terrain component	p53, 43
RKSC	Surface rockiness	Short text	The percentage coverage of rock outcrops - classes (FAO 1990)	p54, 44
STSC[15]	Surface stoniness	Short text	The percentage cover of coarse fragments (> 2 mm), completely or partly at the surface - classes (FAO 1990)	p54, 45
ERTY	Erosion type	Short text	Characterization of the erosion or deposition type according to (FAO 1990)	p54, 46
ERAA	Area affected	Short text	The area affected by erosion or deposition. Classes according to UNEP-ISRIC (1988)	p55, 47

[13] Field names as given on the data entry forms.

[14] In the WRB approach, at scale 1:250,000, the full legend name consists of the RSG (stored in column WRBL) plus the first three applicable principal qualifiers (stored in column WRBP1, WRBP2 and WRBP3) as applicable. For example, Calcaric Leptic Regosols (i.e. WRBP2 + WRBP1 + WRBL), or Mollic Stagnic Gleyic Vertisols (i.e. WRBP3 + WRBP2 + WRBP1 + WRBL); see WRB 2015, p. 14-16, for details. Rules for the use of codes for naming soils are given on p. 190-191 in WRB 2015.

[15] When different from '< 2mm' this should be indicated in table **laboratory methods**, column AMET

Field name	Field name long[13]	Data type	Description	SOTER 2013
ERDE	Erosion degree	Short text	Degree of erosion (FAO 1990)	p55, 48
SCAP	Sensitivity to capping	Short text	The degree in which the soil surface has a tendency to capping and sealing (FAO 1990)	p55, 49
RDEP	Rootable depth	Short text	Estimated depth to which root growth is unrestricted by physical or chemical impediments - classes after (FAO 1990)	p55, 50

Soils

Field name	Field name long	Data type	Description	SOTER 2013
ISOC	ISO country code	Short text	ISO-3166 country code	p33, 1
SUID	SOTER unit-ID	Number	The identification code of a SOTER unit on the map and in the database	p33, 2
TCID	Terrain component number	Number	The sequence number of the terrain component in the terrain (largest comes first)	p51, 34
SCID	Soil component number	Number	The sequence number of the soil component in the terrain component (largest comes first)	As above
PRID	Profile ID	Short text	Code for a representative profile	p53, 42
REPR		Yes/No	Representative (Yes/No), Yes if selected to represent the given Soil component	New

Profile

Field name	Field name long	Data type	Description	SOTER 2013
PRID	Profile ID	Short text	Code for a representative profile	p53, 42
PDID	Profile database ID	Short text	ID for the owner, institute or organization that holds (part of) the national soil profile database	p56, 57
STAT	Profile description status	Short text	The soil profile description status refers to the inferred quality of the soil description and the completeness of analytical data	p56, 58
SAYR	Sampling date	Number	The year in which the profile was described and sampled (yy-mm-dd)	p57, 59
LABO	Laboratory ID	Short text	ID for the soil laboratory that analysed the samples	p57, 60
LATI	Latitutude	Number	Latitude in decimal degrees. Latitudes in the southern hemisphere are negative (WGS 1984)	p57, 61
LNGI	Longitude	Number	Longitude in decimal degrees. Longitudes in the western hemisphere are negative (WGS 1984)	p57, 62
LCST	Profile location	Short text	The conditions from which the profile locations were derived	p57, 63
ELEV	Elevation	Number	The elevation of the representative profile in metre above sea level	p58, 64
LUSE1	Land use, level 1	Short text	Land use at the (exact) location of the soil profile, level 1	p58, 65

Field name	**Field name long**	**Data type**	**Description**	**SOTER 2013**
LUSE2	Land use, level 2	Short text	Land use at the (exact) location of the soil profile, level 2	p58, 65
LUSE3	Land use, level 3	Short text	Land use at the (exact) location of the soil profile, level 3	p58, 65
VEGE1	Vegetation, level 1	Short text	Vegetation at the (exact) location of the soil profile, level 1	p58, 66
VEGE2	Vegetation, level 1	Short text	Vegetation at the (exact) location of the soil profile, level 2	p58, 66
VEGE3	Vegetation, level 1	Short text	Vegetation at the (exact) location of the soil profile, level 3	p58, 66
LITH1	Parent material, level 1	Short Text	Parent material at the (exact) location of the soil profile, level 1	p58, 67
LITH2	Parent material, level 2	Short Text	Parent material at the (exact) location of the soil profile, level 2	p58, 67
LITH3	Parent material, level 3	Short Text	Parent material at the (exact) location of the soil profile, level 3	p58, 67
LITH4	Parent material, level 4	Short text	Parent material at the (exact) location of the soil profile, level 4	p58, 67
LITH5	Parent material, level 5	Short Text	Parent material at the (exact) location of the soil profile, level 5	p58, 67
DRAI	Drainage	Short text	Present drainage class of a soil component represented by this profile	p58, 68
WRBC	WRB RSG	Short text	World Reference Base – RSG classification, see WRB 2015, p. 12-21 and 85-116.	-
WRBP1	WRB principal qualifier 1	Short text	World Reference Base – principal qualifier 1, see WRB 2015, p. 14-16 and 85-116.	-
WRBP2	WRB principal qualifier 2	Short text	World Reference Base – principal qualifier 2, see WRB 2015, p. 14-16 and 85-116.	-
WRBP3	WRB principal qualifier 3	Short text	World Reference Base – principal qualifier 3, see WRB 2015, p. 14-16 and 85-116.	-
WRBS1	WRB supplementary qualifier 1	Short text	World Reference Base – supplementary qualifier 1, see WRB 2015, p. 12-18 and 85-116.	
WRBS2	WRB supplementary qualifier 2	Short text	World Reference Base – supplementary qualifier 2, see WRB 2015, p. 112-18 and 85-116.	
WRBS3	WRB supplementary qualifier 3	Short text	World Reference Base – supplementary qualifier 3, see WRB 2015, p. 12-18 and 85-116.	
WRBV	WRB vesion	Number	World Reference Base, version (year; 2015 by default)	-
CLAF	FAO classification	Short text	Characterization of the profile - revised legend of the FAO-Unesco Soil Map of the World Legend (FAO, 1988)	p58, 71
CLAN	National classification	Short text	The original national classification of the profile, if different from the FAO 1988 or WRB 2015 classification	p59, 72
STAX	Soil Taxonomy	Short text	Classification according to USDA Soil Taxonomy (e.g., Soil Survey Staff 2014)	p59, 73
STXV	Soil Taxonomy version (year)	Number	Soil Taxonomy version (year)	p79, 74

Horizon[16]

Field name (short) [17]	Field name (long)	Data type	Description	SOTER 2013
PRID	Profile_ID	Short text	Code for a representative profile	p53, 42
HONU	Horizon number	Short text	A consecutive number, starting with the surface horizon	p59, 76
DIAH	Diagnostic horizon	Short text	Diagnostic horizon – According to WRB classification (2015)	p59, 77
DIAP	Diagnostic property	Short text	Diagnostic property - According to WRB classification (2015)	p63, 78
DIAM	Diagnostic material	Short text	Diagnostic materials - reflect partly the properties of the original parent material in which pedogenetic processes have not yet been very active	p64, 79
HODE	Horizon designation	Short text	Master horizon with subordinate characteristics (FAO 2006)	p65, 80
HBTP*	Upper boundary	Short text	The average depth of the upper boundary in centimetre	p67, 81
HBDE*	Lower boundary	Short text	The average depth of the lower boundary in centimetre	p67, 82
SCMO	Moist colour	Short text	The Munsell colour of the moist soil	p67, 84
SCDR	Dry colour	Short text	The Munsell colour of the dry soil	p67, 85
MOCL	Colour of mottles	Short text	Colour of mottles, corresponding to the Munsell colour notation	p68, 86
MOAB	Abundance of mottles	Short text	Abundance of mottles	p68, 87
MOSZ	Size of mottles	Short text	Size of mottles	p68, 88
STGR	Grade of structure	Short text	Grade of structure (FAO,1990)	p68, 89
STSI	Size of structure elements	Short text	Size of structure elements (FAO, 1990)	p69, 90
STTY	Type of structure	Short text	Type of structure (FAO, 1990)	p69, 91
MINN	Nature of concretions and nodules	Short text	The nature of concretions and mineral nodules (FAO 2006, FAO and ISRIC 1990)	p69, 92
MINA	Abundance of concretions and nodules	Short text	Classes of volume% of concretions and/or mineral nodules in the soil matrix (FAO, 1990)	p70, 93
MINS	Size of concretions and nodules	Short text	Size of dominant concretions and/or mineral nodules in classes (FAO, 1990)	p70, 94
CFRA	Abundance of coarse fragments	Short text	Classes of volume% of rock and/or coarse fragments in the soil matrix (FAO, 1990)	p70, 95
CFRS	Size of coarse fragments	Numeric	Size of dominant rock and/or coarse fragments in classes (FAO, 1990)	p70, 96
SDVC	Very coarse sand	Numeric	Weight% of particles 2.0 - 1.0 mm (very coarse sand) in fine earth fraction [18]	p70, 97
SDCO	Coarse sand	Numeric	Weight% of particles 1.0 - 0.5 mm (coarse sand) in fine earth fraction	p70, 98
SDME	Medium sand	Numeric	Weight% of particles 0.5 - 0.25 mm (medium sand) in fine earth fraction	p70, 99
SDFI	Fine sand	Numeric	Weight% of particles 0.25 - 0.1 mm (fine sand) in fine earth fraction	p70, 100
SDVF	Very fine sand	Numeric	Weight% of particles 0.1 - 0.05 mm (very fine sand) in fine earth fraction	p70, 101
SDTO*	Total sand	Numeric	Weight% of particles 2.0 - 0.05 mm (total sand) in fine earth fraction	p70, 102
STPC*	Silt	Numeric	Weight% of particles 0.002-0.05 mm (silt) in fine earth fraction	p70, 103
CLPC*	Clay	Numeric	Weight% of particles < 0.002 mm (clay) in fine earth fraction	p70, 104
PSCL	Particle size class	Numeric	Particle size class as derived from the particle size analysis	p70, 105
BULK	Bulk density	Numeric	The bulk density in kg per cubic dm	p72, 106
MCT1	Soil moisture (-0.1 KPa)	Numeric	Soil moisture (%) at -0.1 KPa tension	p72, 107
MCT2	Soil moisture (-10 KPa)	Numeric	Soil moisture (%) at -10 KPa tension	p72, 107
MCT3	Soil moisture (-20 KPa)	Numeric	Soil moisture (%) at -20 KPa tension	p72, 107
MCT4	Soil moisture (%) at -33 KPa tension	Numeric	Soil moisture (%) at -33 KPa tension	p72, 107
MCT5	Soil moisture (%) at -50 KPa tension	Numeric	Soil moisture (%) at -50 KPa tension	p72, 107

[16] There are no forms for the **horizon** table as data entry into this table is best done in the 'data view' mode respectively using a tailor-made SQL-procedure to import the corresponding data from the national soil (profile) database.

[17] There are no forms for the **horizon** table as data entry into this table is best done using tailor-made SQL-procedures to import the corresponding (pre-harmonised) data from a national soil (profile) database. For paper resources, you should enter the data directly into the profile table (respecting the SOTER conventions).

[18] These are the 'default' fraction size limits as given in the 2013 SOTER Procedures Manual. In practice, different limits may be used for the clay, silt and sand-size fraction in various countries/laboratories. Where this is the case, the exact fraction size limits for the 'clay-size', 'silt-size' and 'sand-size' fraction should be explicitly mentioned in table **laboratory method**, column *AMET*.

Field name (short) [17]	Field name (long)	Data type	Description	SOTER 2013
MCT6	Soil moisture (%) at -100 KPa tension	Numeric	Soil moisture (%) at -100 KPa tension	p72, 107
MCT7	Soil moisture (%) at -330 KPa tension	Numeric	Soil moisture (%) at -330 KPa tension	p72, 107
MCT8	Soil moisture (%) at -1500 KPa tension	Numeric	Soil moisture (%) at -1500 KPa tension	p72, 107
ELEC	Electrical conductivity	Numeric	Electrical conductivity in the supernatant of a 1:2.5 soil-water mixture (dS/m)	p73, 108
PHAQ*	pH (H2O)	Numeric	pH (H_2O) in a supernatant suspension of a 1:2.5 soil - water mixture	p73, 109
PHKC(*)	pH (KCl)	Numeric	pH (KCl) in a supernatant suspension of a 1:2.5 soil - 1M KCl mixture	p73, 110
PHCA	pH (CaCl2)	Numeric	pH ($CaCl_2$) in a supernatant suspension of a 1:2 soil - 0.1M $CaCl_2$ mixture	p73, 111
ELCO	Electrical conductivity (sat. extract)	Numeric	Electrical conductivity of saturation extract (dS /m)	p73, 112
SONA	Soluble Na	Numeric	The soluble Na^+ content of the saturated paste in cmol(c)/ liter	p73, 113
SOCA	Soluble Ca	Numeric	The soluble Ca^{++} content of the saturated paste in cmol(c) / liter	p73, 114
SOMG	Soluble Mg	Numeric	The soluble Mg^{++} content of the saturated paste in cmol(c) / liter	p73, 115
SOLK	Soluble K	Numeric	The soluble K+ content of the saturated paste in cmol(c) / liter	p73, 116
SOCL	Soluble Cl	Numeric	The soluble Cl^- content of the saturated paste in cmol(c) / liter	p73, 117
SSO4	Soluble SO4	Numeric	The soluble $SO4^{--}$ content of the saturated paste in cmol(c) / liter	p74, 118
HCO3	Soluble HCO3	Numeric	The soluble $HCO3^-$ content of the saturated paste in cmol(c) / liter	p74, 119
SCO3	Soluble CO3	Numeric	The soluble CO3- content of the saturated paste in cmol(c) / liter	p74,119
EXCA*	Exchangeable Ca	Numeric	The exchangeable Ca in cmol(c) / kg	p74, 120
EXMG*	Exchangeable Mg	Numeric	The exchangeable Mg in cmol(c) / kg	p74, 121
EXNA*	Exchangeable Na	Numeric	The exchangeable Al in cmol(c) / kg	p74, 122
EXCK*	Exchangeable K	Numeric	The exchangeable K in cmol(c) / kg	p74, 123
EXAL*	Exchangeable Al	Numeric	The exchangeable Al in cmol(c) / kg	p74, 125
EXAC*	Exchangeable acidity	Numeric	The exchangeable acidity, as determined in 1N KCl, in cmol(c) / kg	p74, 126
CECS*	CEC soil	Numeric	The cation exchange capacity of the soil at pH 7.0 in cmol(c) / kg	p74, 127
TCEQ*	Total carbonate equivalent	Numeric	The content of carbonates in g / kg	p74, 128
GYPS	Gypsum	Numeric	The gypsum content in g / kg	p74, 129
TOTC(*)	Total carbon	Numeric	The total content of organic and inorganic carbon of the soil layer in g / kg	p75, 130
ORGC*	Organic carbon	Numeric	The content of organic carbon in g / kg of the soil layer	p75, 131
TOTN*	Total nitrogen	Numeric	The content of total N of the soil in g / kg	p75, 132
P2O5	Available P	Numeric	The available P-content of the soil in mg / kg	p75, 133
TOTP	Total P	Numeric	The total P-content of the soil in mg / kg	p75, 134
PRET	Phosphate retention	Numeric	The phosphate retention, in %	p75, 135
FEDE	Fe, Dithionite extractable	Numeric	The Fe fraction, in weight%, extractable in dithionite citrate	p75, 136
ALOE	Al, Oxalate extractable	Numeric	The Al fraction, in weight%, extractable in oxalate acid	p75, 137
FEOE	Fe, Oxalate extractable	Short text	The Fe fraction, in weight%, extractable in oxalate acid	p75, 138
CLAY	Clay mineralogy		The dominant type of mineral in the clay fraction	p75, 139

* Horizon properties that are mandatory according to the SOTER Procedures Manual 2013, besides the primary keys (see table relationships in Appendix 4), are flagged with a star' (e.g. CLPC* for clay weight%).

Appendix 3. Reference tables

The so-called reference tables store information on the source materials used for the compilation of SOTER units. These materials include the source maps (table **SourceMap**), laboratories that analysed the data (table **Laboratory**), analytical methods used (table **LaboratoryMethod**), and the organisations responsible for the national soil profile database (table **ProfileDatabase**).

SourceMap

Field name (short)	Field name (long)	Data type	Description	SOTER 2013
MAPI	Source map-ID	Text	The identification code of the source map from which the data were derived	p87, 1
TITL	Map title	Text	The citation of the source map	p87, 2
PUYR	Year	Number	The year of publication of the source map	p87, 3
SCAL	Scale	Number	The scale of the source map as a representative fraction	p80, 4
MLAT	Minimum latitude	Number	Minimum latitude in decimal degrees - latitude South is a negative figure	p80, 5
MLON	Minimum longitude	Number	Minimum longitude in decimal degrees - longitude West is a negative figure	p80, 6
XLAT	Maximum latitude	Number	Maximum latitude in decimal degrees - latitude South is a negative figure	p80, 7
XLON	Maximum longitude	Number	Maximum longitude in decimal degrees - longitude West is a negative figure	p80, 8
UTMZ	UTM zone	Short text	UTM zone (a number in the range 1-60 for a longitudinal belt, followed by a letter in the range C-X for a latitudinal belt - e.g. 43P) Geodetic datum	p80, 9
DATM	Geodetic datum	short text	Geodetic datum	p80, 10
EMIN	Minimum easting	Number	Minimum easting on the map	p80, 11
NMIN	Minimum northing	Number	Minimum northing on the map	p80, 11
EMAX	Maximum easting	Number	Maximum easting on the map	p80, 11
NMAX	Maximum northing	Number	Maximum easting on the map	p80, 11
TYPE	Type of source map	Short text	Type of source map	p80, 11

Laboratory

Field name (short)	Field name (long)	Description	SOTER 2013
LABO	LAB-ID text	Identification code for the laboratory that analysed the reference soil profile	p89, 1
LNAM	Long text	Name of the laboratory, in full	p89, 2

LaboratoryMethod

Field name (short)	Field name (long)	Description	SOTER 2013
LABO	Lab-ID	Identification code for the laboratory that analysed the (reference) soil profile	p89, 3
ATTR	Attribute	ID/code for given property (see table' **codes**, where column ATTRIBUTE is 'ATTR')	p89, 6
AMET	AMET	A concise *description* of the analytical method, including references/URLs.	

ProfileDatabase

Field name (short)	Field name (long)	Data type	Description	SOTER 2013
PDID	Profile database_ID	Short text	ID of the owner, institute or organization that holds (part of) the national soil profile database	p90, 1
AUTH	Author	Short text	Main author	p90, 2
PUYR	Year of publication	Numeric	Year of publication	p90, 3
TITL	Title	Short text	Title of publication	p90, 4
DOWN	Data owner	Short text	Name of the owner, institute or organization of the national soil profile database and address	p90, 5
PUBL	Publisher	Short text	Publisher of the document or the original source	p90,6
PAGE	Chapter / page	Short text	Chapter and/or pages in the document where the description and analytical data can be found	p90, 7
DDSC	Digital data source	Short text	Digital data source	p90, 8

Appendix 4. Look-up table

There is one large look-up table (**Codes**). It gives the code and a short description for the attributes considered in SOTER.

Codes

Field name	Data type	Description
ATTRIBUTE	Short text	Abbreviation for given attribute
ORDER	Numeric	Sequential number
CODE	Short text	Code for given property (Level 1)
DESCRIPTION	Short text	Description of the given attribute (Level 1)
LEVEL2_CODE	Short text	Code for given property (Level 2)
LEVEL2_DESC	Short text	Description of the given attribute (Level 2)
LEVEL3_CODE	Short text	Code for given property (Level 3)
LEVEL3_DESC	Short text	Description of the given attribute (Level 3)
LEVEL4_CODE	Short text	Code for given property (Level 4)
LEVEL4_DESC	Short text	Description of the given attribute (Level 4)
LEVEL5_CODE	Short text	Code for given property (Level 5)
LEVEL5_DESC	Short text	Description of the given attribute (Level 5)

Example of 'simple' codes (e.g. ATTR, soil chemical and physical properties)

ATTRIBU	ORDER	CODE	DESCRIPTION
ATTR	7	ELCO	The electrical conductivity of saturation extract (dS/m)
ATTR	8	EXAL	The exchangeable Al in cmol(+) / kg
ATTR	9	EXCA	The exchangeable Ca in cmol(+) / kg
ATTR	10	EXMG	The exchangeable Mg in cmol(+) / kg
ATTR	11	EXCK	The exchangeable K in cmol(+) / kg
ATTR	12	EXNA	The exchangeable Na in cmol(+) / kg
ATTR	13	EXAC	The exchangeable acidity, as determined in 1N KCl, in cmol(+) / kg
ATTR	14	FEDE	The Fe fraction, in weight%, extractable in dithionite citrate
ATTR	15	FEOE	The Fe fraction, in weight%, extractable in oxalate acid
ATTR	16	GYPS	The gypsum content in g / kg
ATTR	18	PHCA	pH (CaCl2) in a supernatant suspension of a 1:2 soil - 0.1M CaCL2 mixture
ATTR	19	PHAQ	pH (H2O) in a supernatant suspension of a 1:2.5 soil - water mixture
ATTR	20	PHKC	pH (KCl) in a supernatant suspension of a 1:2.5 soil - 1M KCl mixture
ATTR	21	PRET	The phosphate retention, in %
ATTR	22	SOCA	The soluble Ca++ content of the saturated paste in cmol(c) / liter
ATTR	23	SOCL	The soluble Cl- content of the saturated paste in cmol(c) / liter
ATTR	24	SCO3	The soluble CO3- content of the saturated paste in cmol(c) / liter
ATTR	25	HCO3	The soluble HCO3- content of the saturated paste in cmol(c) / liter
ATTR	26	SOLK	The soluble K+ content of the saturated paste in cmol(c) / liter
ATTR	27	SOMG	The soluble Mg++ content of the saturated paste in cmol(c) / liter
ATTR	28	SONA	The soluble Na+ content of the saturated paste in cmol(c) / liter
ATTR	29	SSO4	The soluble SO4-- content of the saturated paste in cmol(c) / liter
ATTR	30	SDVF	Weight% of particles 0.1 - 0.05 mm (very sand) in fine earth fraction
ATTR	31	SDFI	Weight% of particles 0.25 - 0.1 mm (fine sand) in fine earth fraction
ATTR	32	SDME	Weight% of particles 0.5 - 0.25 mm (medium sand) in fine earth fraction
ATTR	33	SDCO	Weight% of particles 1.0 - 0.5 mm (coarse sand) in fine earth fraction
ATTR	34	SDVC	Weight% of particles 2.0 - 1.0 mm (very coarse sand) in fine earth fraction

Example of 'compound' codes (e.g. land use (LUSE), 3 levels)

ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_CO	LEVEL2_DESC	LEVEL3_COD	LEVEL3_DESC
LUSE	1743	A	Agriculture	AA	annual field cropping	AA1	shifting cultivation
LUSE	1744	A	Agriculture	AA	annual field cropping	AA2	fallow system cultivation
LUSE	1745	A	Agriculture	AA	annual field cropping	AA3	ley system cultivation
LUSE	1746	A	Agriculture	AA	annual field cropping	AA4	rainfed arable cultivation
LUSE	1747	A	Agriculture	AA	annual field cropping	AA5	wet rice cultivation
LUSE	1748	A	Agriculture	AA	annual field cropping	AA6	irrigated cultivation
LUSE	1749	A	Agriculture	AP	perennial field cropping	AP1	non-irrigated
LUSE	1750	A	Agriculture	AP	perennial field cropping	AP2	irrigated
LUSE	1751	A	Agriculture	AT	tree & shrub cropping	AT1	non-irrigated tree crop cultivation
LUSE	1752	A	Agriculture	AT	tree & shrub cropping	AT2	irrigated tree crop cultivation
LUSE	1753	A	Agriculture	AT	tree & shrub cropping	AT3	non-irrigated shrub crop cultivation
LUSE	1754	A	Agriculture	AT	tree & shrub cropping	AT4	irrigated shrub crop cultivation
LUSE	1755	E	Extraction/Collecting	EH	hunting and fishing		
LUSE	1756	E	Extraction/Collecting	EV	exploitation of natural vegetation		
LUSE	1757	F	Forestry	FN	exploitation of natural forest and woodland	FN1	selective felling
LUSE	1758	F	Forestry	FN	exploitation of natural forest and woodland	FN2	clear felling
LUSE	1759	F	Forestry	FP	plantation forestry		
LUSE	1760	H	Animal husbandry	HE	extensive grazing	HE1	nomadism
LUSE	1761	H	Animal husbandry	HE	extensive grazing	HE2	semi-nomadism
LUSE	1762	H	Animal husbandry	HE	extensive grazing	HE3	ranching
LUSE	1763	H	Animal husbandry	HI	intensive grazing	HI1	animal production
LUSE	1764	H	Animal husbandry	HI	intensive grazing	HI2	dairying

Example of 'compound' codes (e.g. Soil classification (WRB), 3 levels)

ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_CO	LEVEL2_DESC	LEVEL3_COD	LEVEL3_DESC
WRB	433	CH	Chernozem	ph	Pachic	s	supplementary
WRB	434	CH	Chernozem	rp	Raptic	s	supplementary
WRB	435	CH	Chernozem	sk	Skeletic	p	principal
WRB	436	CH	Chernozem	sl	Siltic	s	supplementary
WRB	437	CH	Chernozem	so	Sodic	s	supplementary
WRB	438	CH	Chernozem	st	Stagnic	s	supplementary
WRB	439	CH	Chernozem	szn	Endosalic	s	supplementary
WRB	440	CH	Chernozem	te	Technic	s	supplementary
WRB	441	CH	Chernozem	tf	Tephric	s	supplementary
WRB	442	CH	Chernozem	tn	Transportic	s	supplementary
WRB	443	CH	Chernozem	to	Tonguic	s	supplementary
WRB	444	CH	Chernozem	tu	Turbic	s	supplementary
WRB	445	CH	Chernozem	vi	Vitric	s	supplementary
WRB	446	CH	Chernozem	vm	Vermic	p	principal
WRB	447	CH	Chernozem	vr	Vertic	p	principal
WRB	448	CL	Calcisol	ab	Albic	s	supplementary
WRB	449	CL	Calcisol	ad	Aridic	s	supplementary
WRB	450	CL	Calcisol	ai	Aric	s	supplementary
WRB	451	CL	Calcisol	ar	Arenic	s	supplementary

Example of 'compound' codes (e.g. Parent material (LITH), 5 levels)

Codes

ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_CO	LEVEL2_DESC	LEVEL3_COD	LEVEL3_DESC	LEVEL4_COD	LEVEL4_DESC	LEVEL5_COD	LEVEL5_DESC
LITH	46	C	consolidated	CS	siliceous	CSU	ultrabasic (< 45% SiO2)	CSUM	metamorphic	CSUM1	meta-ultramafic rock
LITH	47	C	consolidated	CS	siliceous	CSU	ultrabasic (< 45% SiO2)	CSUT	metasomatic	CSUT1	serpentinite (43% SiO2), skarn (42% SiO2)
LITH	48	C	consolidated	CS	siliceous	CSX	unspecified	CSXI	igneous	CSXI1	agglomerate, pyroclastic breccia, scoria
LITH	49	C	consolidated	CS	siliceous	CSX	unspecified	CSXI	igneous	CSXI2	tuff-breccia
LITH	50	C	consolidated	CS	siliceous	CSX	unspecified	CSXI	igneous	CSXI3	lapilli-stone, lapilli-tuff
LITH	51	C	consolidated	CS	siliceous	CSX	unspecified	CSXI	igneous	CSXI4	tuff, ignimbrite (welded tuff).
LITH	52	C	consolidated	CS	siliceous	CSX	unspecified	CSXI	igneous	CSXIx	igneous rock (unspecified)
LITH	53	C	consolidated	CS	siliceous	CSX	unspecified	CSXM	metamorphic	CSXM1	suevite, impactite, impact-melt breccias, impact-melt ro
LITH	54	C	consolidated	CS	siliceous	CSX	unspecified	CSXM	metamorphic	CSXM2	cataclasite, mylonite
LITH	55	C	consolidated	CS	siliceous	CSX	unspecified	CSXM	metamorphic	CSXMx	metamorphic rock (unspecified)
LITH	56	C	consolidated	CS	siliceous	CSX	unspecified	CSXS	sedimentary rock	CSXS1	tuffaceous-sedimentary rock, tuffite
LITH	57	C	consolidated	CS	siliceous	CSX	unspecified	CSXS	sedimentary rock	CSXSx	sedimentary rock (unspecified)
LITH	58	C	consolidated	CY	saline	CYX	unspecified	CYXS	sedimentary rock	CYXS1	alkali chloride, earth alkali chloride
LITH	59	S	semi-consolidated	SC	calcareous	SCX	unspecified	SCXS	sedimentary rock	SCXS1	chalk
LITH	60	S	semi-consolidated	SC	calcareous	SCX	unspecified	SCXS	sedimentary rock	SCXS2	tufa
LITH	61	S	semi-consolidated	SF	iron bearing	SFX	unspecified	SFXS	sedimentary rock	SFXS1	laterite, bauxite
LITH	62	S	semi-consolidated	SO	organic	SOX	unspecified	SOXS	sedimentary rock	SOXS1	lignite
LITH	63	S	semi-consolidated	SO	organic	SOX	unspecified	SOXS	sedimentary rock	SOXS2	asphalt
LITH	64	S	semi-consolidated	SS	siliceous	SSA	acid	SSAR	residual deposit	SSAR1	kaolin
LITH	65	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXA	anthropogenic	UCXA1	lime plaster, cement plaster
LITH	66	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXA	anthropogenic	UCXA2	concrete
LITH	67	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXA	anthropogenic	UCXA3	waste combustion ash
LITH	68	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXS	sediment	UCXS1	carbonate sand
LITH	69	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXS	sediment	UCXS2	carbonate mud, carbonate ooze
LITH	70	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXS	sediment	UCXS3	carbonatic diamicton
LITH	71	U	unconsoli-dated	UC	calcareous	UCX	unspecified	UCXS	sediment	UCXS4	carbonatic sediment,marl
LITH	72	U	unconsoli-dated	UF	iron bearing	UFX	unspecified	UFXA	anthropogenic	UFXA1	red mud
LITH	73	U	unconsoli-dated	UF	iron bearing	UFX	unspecified	UFXA	anthropogenic	UFXA2	metal-sludge
LITH	74	U	unconsoli-dated	UF	iron bearing	UFX	unspecified	UFXS	sediment	UFXS1	iron-sediment
LITH	75	U	unconsoli-dated	UG	gypsic	UGX	unspecified	UGXA	anthropogenic	UGXA1	gypsum plaster
LITH	76	U	unconsoli-dated	UG	gypsic	UGX	unspecified	UGXS	sediment	UGXS1	gypsum-mud
LITH	77	U	unconsoli-dated	UO	organic	UOX	unspecified	UOXA	anthropogenic	UOXA1	plaggen

Appendix 5. Worked examples

A5.1 General

As indicated, each SOTER database consists of a geographic component (GIS shapefile) and related SOTER attribute tables (relational database). This Appendix describes worked examples for two *fictional* SOTER databases to illustrate main principles of the SOTER methodology. These examples follow the terminology and conventions adopted for the SOTER Danube basin project.

The first example is for a 'conventional' SOTER database. At the present scale of 1:250,000, each SOTER unit (terrain) is generally comprised of several terrain components, which in turn can consist of one or more soil components. Each of these soil components is characterised by a representative profile. This profile is selected from the national soil profile database or other archives to best represent the given soil component (as characterised by its WRB 2015 Legend unit by regional experts (see column WRBL, WRBP1, WBRP2 and WBRP3 in table **SoilComponent**)).

The second example is for a 'simplified' SOTER database (as will be used for the Danube basin project). In this case, each SOTER (terrain) unit is assumed to comprise only one terrain component. This terrain component can comprise from one to several soil components; each soil component is characterised by one single representative profile.

To visualise the differences in approach, a series of screenshots is provided below for both types of SOTER approaches/databases for an hypothetical study area.

As indicated, guidelines for data providers to compile the 14 'national' SOTER databases for the Danube basin, according to the WRB2015 conventions, are provided in a separate document (Batjes and Ribeiro 2019).

A5.2 GIS file

A (hypothetical) map with SOTER (terrain) units for country XX[19] is shown in Figure 5; the two polygons outlined in blue correspond with SOTER unit 'CZ651'. Each terrain unit on the map, as identified by its unique ISOCSUID, is further characterised in the SOTER attribute tables.

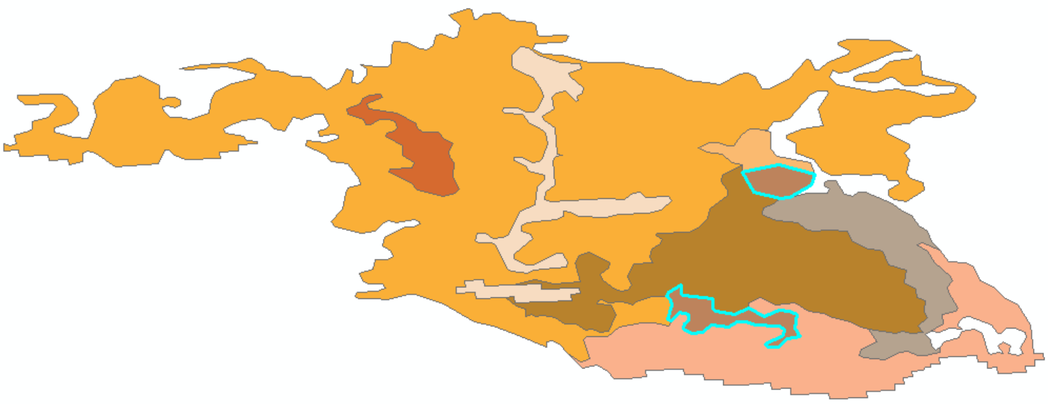


Figure 5. Example of a SOTER unit in the terrain shapefile as characterised by its unique ISOCSUID

[19] Note: In view of referential integrity, an existing ISO code has to be used for the example. Here, 'CZ' is arbitrarily used for this; all data shown/used are hypothetical.

The compound ISOCSUID code provides the logical link to the attribute tables in the corresponding MS Access database (e.g. XX_SOTER, where XX is the country code). Figure 6 shows the attributes that are described in the property table of the corresponding shape file.

FID	Shape *	ISOCSUID	TUID	HYPSCLASS	SLOPECLASS	RICLASS	SURFCOND	GENETICS	CARBONATE	TEXTURE2	COUNTRY	area
0	Polygon	CZ651	651	4	1	2	Unconsolidated	Aeolian		Clay	CZ	27633150
1	Polygon	CZ698	698	7	5	6	Consolidated	Eluvial-colluvial		Clay	CZ	1863773144.1
2	Polygon	CZ715	715	6	2	3	Unconsolidated	Aeolian		Loam	CZ	129417750
3	Polygon	CZ791	791	6	2	3	Unconsolidated	Aeolian		Clay	CZ	66845250
4	Polygon	CZ797	797	6	3	5	Unconsolidated	Aeolian		Loam	CZ	35105400
5	Polygon	CZ804	804	6	3	5	Consolidated	Eluvial-colluvial		Loam	CZ	458183173.83
6	Polygon	CZ815	815	7	5	5	Consolidated	Eluvial-colluvial		Clay	CZ	174960000
7	Polygon	CZ817	817	7	5	5	Consolidated	Eluvial-colluvial		Clay	CZ	410033532.90
8	Polygon	CZ651	651	7	5	5	Consolidated	Eluvial-colluvial		Clay	CZ	44813250
9	Polygon	CZ715	715	6	5	5	Consolidated	Eluvial-colluvial		Loam	CZ	27468785.710

Figure 6. Example of terrain attributes stored in the property table of the shapefile

A5.3 SOTER attribute tables

Case 1: Conventional SOTER with compound terrain components with multiple soil components

This is the 'conventional' approach to developing a SOTER database. As indicated, the procedure is illustrated for a hypothetical example. The **Terrain** table (Figure 7) describes the spatial units shown on the GIS map in greater detail. As indicated, the map units of the **shape file's attribute** table and **Terrain** table can be linked using column ISOCSUID using GIS.

Terrain

ISOCSUID	ISO Countr	SOTER unit	Date of data collection	Map_ID	Minimum ele	Maximum elev	Median eleva	Median slope	Relief index	Potential drai	Major landfo
CZ651	CZ	651	01/01/2011	CEU01	269	518	373	7	55		LD
CZ698	CZ	698	01/01/2011	CEU01	147	640	384	11	84		SE
CZ715	CZ	715	01/01/2011	CEU01	266	475	352	5	37		LL
CZ791	CZ	791	01/01/2011	CEU01	329	859	485	14	114		SE
CZ797	CZ	797	01/01/2011	CEU01	379	1588	806	19	151		SV
CZ804	CZ	804	01/01/2011	CEU01	306	784	497	14	113		SE
CZ815	CZ	815	01/01/2011	CEU01	275	637	455	11	91		SE
CZ817	CZ	817	01/01/2011	CEU01	112	813	356	16	130		SE

Figure 7. Example of attributes stored in the terrain table

Figure 8 shows the above eight SOTER (terrain) units as represented in the **TerrainComponent** table. The first ('CZ651') has two terrain components and the other SOTER units only one. It should be noted here, that 'conventional' SOTER databases may contain terrain units with 'multiple' as well as 'single' terrain components depending on the complexity of the terrain and mapping scale. For illustration purposes, the 1:1 million scale SOTER database for Malawi[20] (Dijkshoorn *et al.* 2016) may be consulted. It should be noted, however, that the Malawi database uses an earlier version of the SOTER database model and considers a superseded version of WRB (IUSS Working Group WRB 2006).

TerrainComponent

ISO country code	SOTER unit_I	Terrain component n	Proportio	Dominant sl	Length	Form of slope	Lithologyl	Litholog	Lithology L2	Litho
CZ	651	1	60	7			UT2E	C	CC	
CZ	651	2	40	13			UT2E	C	CC	
CZ	698	1	100	11			UT2E	C	CS	
CZ	715	1	100	5			PA1	U		
CZ	791	1	100	14			PA1	U		
CZ	797	1	100	19			VP1	C	CS	
CZ	804	1	100	14			UT2E	C	CS	
CZ	815	1	100	11			UQ	C	CC	
CZ	817	1	100	16			SL3	C	CS	

Figure 8. Example of a compound SOTER unit with two terrain components

20 https://data.isric.org/geonetwork/srv/eng/catalog.search#/metadata/60803da0-a15f-4cc5-9cb5-172fa2460af3

Figure 9 is for the same hypothetical example, but now with the corresponding soil components. Each soil component is characterised in terms of its WRB2015 Legend code in table **SoilComponent**. Regional experts have selected profile 'CZ0421/740' as being representative for soil component 1 in terrain component 1 (Figure 10). This is a 'CM' (Cambisol), with 'gl' (gleyic) as principal qualifier and 'dy' (dystric) as second qualifier, hence a dystric gleyic Cambisol (see Figure 11, table **Profile**). Similarly, the other soil components have been characterised using carefully selected representative profiles as characterised in table **Soils** (which serves to flag the representative profile within the full selection of profiles compiled for the country) and table **Profile** that list all the site properties for the given profile, with the related horizon (layer) data stored in table **Horizon**.

SoilComponent

ISO cour	SOTER	Terrain	Soil com	Proport	Profile_ID	WRB Legend unit	WRB principal qual. 1	WRB principal qual. 2	WRB principal qual. 3	WRB versior
CZ	651	1	1	35	CZ0421/740	CM	gl	dy		2015
CZ	651	1	2	15	CZ0513/LB18	RG	cr	dy		2015
CZ	651	1	3	10	CZ0201/513	CM	dy	le		2015
CZ	651	2	1	30	CZ0522/JC28	RG	fc	hg		2015
CZ	651	2	2	10	CZ0205/242	LV	cr			2015

Figure 9. SOTER unit 'CZ651' with its two terrain components, soil components and representative profile

Profile 'CZ0421/740' is flagged in table **Soils** as the representative profile for terrain component 1 and soil component number 1 in SOTER unit 'CZ651' (Figure 10).

Soils

ISO country cod	SOTER unit_ID	Terrain component numbe	Soil component numbe	Profile_ID	Representat
CZ	651	1	1	CZ0421/740	☑
CZ	651	1	2	CZ0513/LB18	☑
CZ	651	1	3	CZ0201/513	☑
CZ	651	2	1	CZ0513/LB18	☑
CZ	651	2	2	CZ0205/242	☑
CZ	698	1	1	CZ0421/739	☑
CZ	698	1	2	CZ0513/LB18	☑
CZ	698	1	3	CZ0511/CL14	☑
CZ	715	1	1	CZ0513/233	☑
CZ	715	1	2	CZ0513/LB06	☑
CZ	715	1	3	CZ0513/LB26	☑
CZ	791	1	1	CZ0513/LB06	☑
CZ	791	1	2	CZ0513/LB26	☑
CZ	797	1	1	CZ0522/JC28	☑
CZ	797	1	2	CZ0514/808	☑

Figure 10. Representative profiles used to characterise the respective soil components

The **Profile** table contains the wider selection of gleyic Cambisols from which this representative profile has been selected (Figure 11; note, some columns or fields are hidden). It also shows the sources from which the profile data were downloaded or entered (see PRID in table **ProfileDatabase**), for example 'EU025'.

Profile

Profile_ID	Profile databas	WRB RSG	WRB princip	WRB principal qual. 2	WRB princip	WRB supplement	WRB supple	WRB supple	WRB versior
CZ0326/668	EU025	CM	gl	eu		lo			2015
CZ0327/270	EU025	CM	gl	cr		lo			2015
CZ0327/TC19	EU025	CM	gl	dy		ce			2015
CZ0421/740	EU025	CM	gl	dy		cl			2015
*									2015

Figure 11. Selecting the representative profile for SOTER unit 'CZ651'

The horizon properties for each profile are characterised in table **Horizon** (Figure 11; note, some columns are hidden).

Horizon

Profile_ID	Horizon numbe	Horizon designatic	Upper boun	Lower boun	Total sanc	Sil	Clay	pH (H2O	pH (KC	pH (CaCl2)	Exchangeable acidity	CEC soi	Organic cark
CZ0421/740	1	Ap	0	20	31.8	57.6	10.65	5.7	5.4		11.48	17.4	2.27
CZ0421/740	2	Bw	20	75	64.9	23.4	11.69	5.6	4.8		7.57	8.7	0.22
CZ0421/740	3	C	75	200	64.9	22.5	12.61	5.3	4.7				0.12

Figure 12. Horizon properties for representative profile 'CZ0421/740'

More specifics about the properties that can be accommodated in the respective tables are provided in Appendix 2.

At an earlier stage of data processing, the laboratory where the various profile (samples) have been analysed needs to be documented (table **Laboratory**, Figure 13).

Laboratory

Lab-ID	Laboratory name
CZ001	Hypothetical laboratory X
CZ002	Hypothetical laboratory Y

Figure 13. Documenting the soil laboratories

Concise, descriptions of the analytical methods (used in a given laboratory) must be provided in table **LaboratoryMethod** (Figure 14). This information is critical to have when analysing soil data derived from various national soil databases in view of their standardisation (Baritz *et al.* 2014; Batjes *et al.* 2017; Hannam *et al.* 2009).

The laboratory methods should be succinctly described in table **LaboratoryMethod** (Figure 14). This should be done at an early stage of the data compilation process.

LaboratoryMethod

Lab-ID	Attribute	Description
CZ001	PHKC	Soil pH measured in a 1:2.5 soil / 0.1M KCl solution (Van Reeuwijk, 2002, p. 56)
CZ001	PRET	More details about the Analytical Method here
CZ001	PSCL	More details about the Analytical Method here
CZ001	SCO3	More details about the Analytical Method here
CZ001	SDCO	More details about the Analytical Method here
CZ001	SDFI	More details about the Analytical Method here
CZ001	SDME	More details about the Analytical Method here
CZ001	SDTO	More details about the Analytical Method here
CZ001	SDVC	More details about the Analytical Method here
CZ001	SDVF	More details about the Analytical Method here
CZ001	SOCA	More details about the Analytical Method here
CZ001	SOCL	More details about the Analytical Method here

Figure 14. Coding and describing the laboratory methods

The geographic basis for the terrain unit map is documented in table **SourceMap** (Figure 15). For the SOTER Danube basin project, there is only one source map (i.e. GIS layer).

SourceMap

Source map-II	Map title	Year	Scale	Geodetic datu	Type of source ma
CEU01	Terrain unit map for Danube Basin project (Provided by JRC)	2019	250,000	WGS 84	D

Figure 15. Documenting the source of the terrain unit maps

Case 2: Simplified SOTER with single terrain components with one or more soil components

As indicated, this approach is adopted for the SOTER Danube basin project. Detailed guidelines for processing source data into this format are provided in a separate document (Batjes and Ribeiro 2019), with accompanying GIS shapefiles (Ruiperez Gonzalez and Batjes 2019) for the respective Danube basin countries.

In essence, the procedure for compiling soil data is similar to that for case 1 above. However, pragmatically (to ensure INSPIRE compliance), for case 2 terrain units are considered to consist of only one terrain component that may comprise from one up to 10 soil components. The example in Figure 16 is for SOTER (terrain) unit 'CZ698' (Figure 13). For SOTER it is mandatory that each soil component represents at least 10% of the total SUID, i.e. has a proportion of at least 10% (van Engelen and Dijkshoorn 2013, p. 30 and 59).

SoilComponent

ISO cour	SOTER	Terrain	Soil com	Proport	Profile_ID	WRB Legend unit	WRB principal qual. 1	WRB principal qual. 2	WRB principal qual. 3	WRB versio
CZ	698	1	1	55	CZ0421/739	CM	do	aq		2015
CZ	698	1	2	30	CZ0513/LB18	RG	cr	dy		2015
CZ	698	1	3	15	CZ0511/CL14	LV	fo	fg	gl	2015

Figure 16. Example of a SOTER unit with one terrain component and three soil components

Schematically, the 'simplified' SOTER approach is visualised in Figure 17. In this example, the representative profile (No. 14) is used to characterise the soil component of terrain component 1 (TCID) of SUID01, has also been used to characterise the third soil component (SCID) of SUID03. This many-to-one relationship is visualised in Figure 4 (see link between **Soil**s table and **Profile** table).

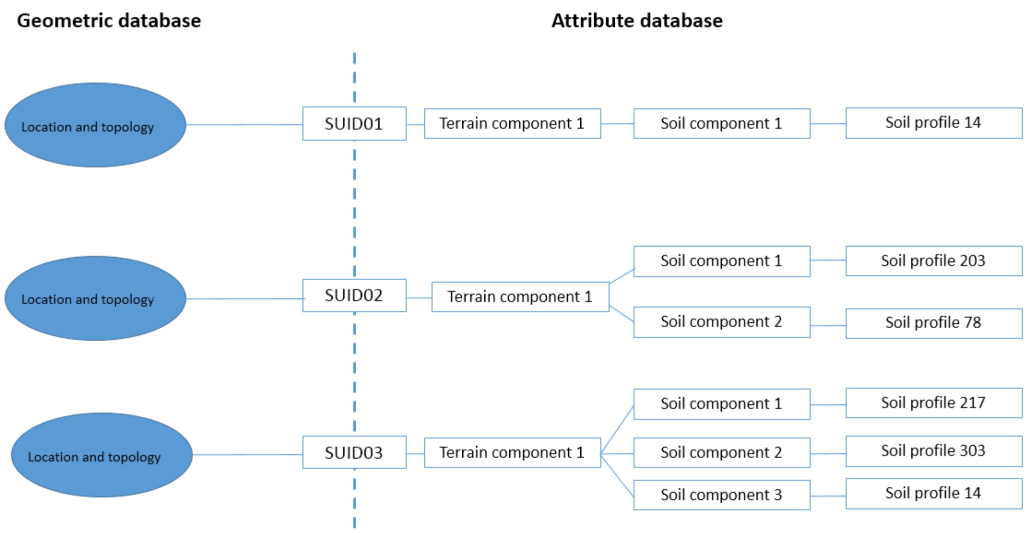


Figure 17. Schematic representation of SOTER DB with its geographic and attribute data component

Further, the overall principles for compiling the database (DB) are similar to those described above for case 1, i.e. a 'conventional' SOTER database. As indicated, the process of data compilation for a 'case 2' type SOTER database is explained in greater detail in separate technical guidelines (Batjes and Ribeiro 2019).