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

EU-JRC Service Contract #936108

Deliverables D.A5

Technical guidelines for populating the 'national' SOTER attribute databases for the EU Danube basin

Niels H. Batjes, Maria Ruiperez Gonzalez and Eloi Ribeiro

March 2019

(rev. February 2021)

Contents

1. Introduction 5
2. SOTER principles 5
2.1 Mapping approach 5
2.2 SOTER requirements 9
2.3 Main steps of data compilation 10
3. Exploring the SOTER (terrain) units for your region 13
4. Prepopulated SOTER terrain-related tables 15
4.1 Terrain 15
4.2 Terrain component 17
4.3 Source map 20
5. Compiling the soil component and soils table 22
5.1 Soil component 22
5.2 Soils 26
6. Compiling the profile and horizon table 28
6.1 General procedure 28
6.2 Profile 29
6.3 Horizon 33
6.4 Source of profile data (profile database) 36
7. Document the laboratory and analytical methods 37
7.1 Laboratory 37
7.2 Laboratory methods 37
Acknowledgments 39
References 40
Appendix 1. SOTER country files and reference manuals 42
Appendix 2. Description of the SOTER Danube basin terrain unit layer 43
Appendix 3. Procedure for populating the terrain table from the GIS shapefile 44
Appendix 4. Relationships enforced in SOTER forms 45
Appendix 5. SOTER look-up table (codes) 46
Appendix 6. Example of procedure for filling the SOTER attribute tables 48
Appendix 7. Pre-harmonizing your data to SOTER conventions 52
Appendix 8. Miscellaneous polygons and ‘non-soil’ units 55

List of figures

Figure 1. Danube region basin 5
Figure 2. Geometric and attribute database structure for the Danube SOTER 6
Figure 3. SOTER Danube attribute database with enforced relationships between tables 7
Figure 4. Excerpt of GIS attribute table provided for each country 13
Figure 5. Selected properties (hypsometry) for terrain units visualized using the html tool 14
Figure 6. Input form for the terrain table 15
Figure 7. Input form for the terrain component table (populated). 18
Figure 8. Input form for the map sources (sourcemap table) 20
Figure 9. Prepopulated terrain, terrain component and soil component tables for SOTER unit GE29142. 23
Figure 10. Example of an error message 24
Figure 11. Input form for the soilcomponent table 26
Figure 12. Input form for the soils table 27
Figure 13. Input form for the profile (site) data. 32
Figure 14. Input form for the source of the profile data (profiledatabase) 36
Figure 15. Form for entering the soil laboratory name 37
Figure 16. Form for entering soil analytical method descriptions 38
Figure 17. Example of a 'non-soil' unit as represented in a hypothetical SOTER unit. 59

List of tables

Table 1. Tables for the SOTER Danube project with purpose and suggested data requirements 11
Table 2. Requirements for the terrain table 16
Table 3. Requirements for the terrain component 18
Table 4. Requirements for the sourcemap table 20
Table 5. Requirement for the soil component table 24
Table 6. Properties to be described in the soils table 27
Table 7. Simplified structure for a correlation table from a hypothetical 'national system' to WRB 2015. 28
Table 8. Properties represented in the profile table 30
Table 9. Properties represented in the horizon table (with units of measurement) 34
Table 10. Requirements for the profiledatabase table 36
Table 11. Documenting soil laboratory names 37
Table 12. Describing the soil laboratory methods 38
Table 13. Codes for 'non-soil' units in the GIS file and attribute database 55

1. Introduction

The European Commission Joint Research Centre (JRC) has proposed soil data collection within the Danube basin, based on SOTER methodologies, to support the Danube strategy. SOTER™ (SOil and TERrain) provides a comprehensive framework for making soil and terrain data available for end users (Van Engelen 2011; van Engelen and Dijkshoorn 2013); application of the SOTER methodology ensures standard procedures for data harmonisation are achieved for the whole Danube basin (Figure 1).

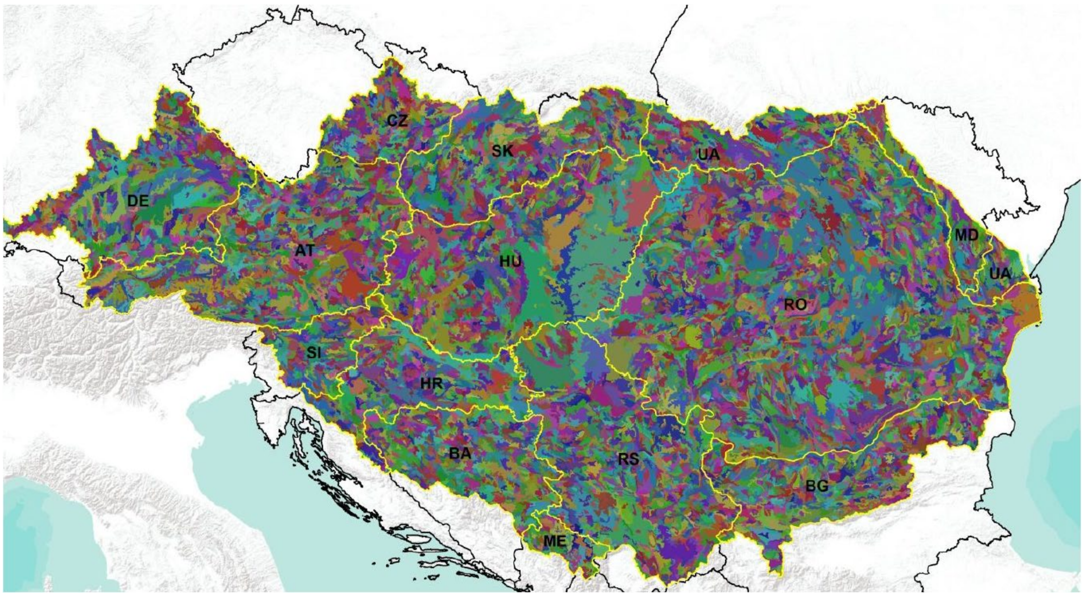


Figure 1. Danube region basin

The SOTER Danube Basin covers parts of 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). Project participants from each country will compile a SOTER database for their part of the basin using national expertise and soil data. Each project participant is provided with a shapefile of terrain units (Ruiperez Gonzalez and Batjes 2019) and an 'empty' SOTER attribute database (Batjes and Ribeiro 2019) specific for the Danube area of the corresponding country. This report provides guidelines for filling the SOTER attribute database.

2. SOTER principles

2.1 Mapping approach

In many respects, the SOTER mapping approach resembles traditional physiographic soil mapping (van Engelen and Dijkshoorn 2013). However, the focus in SOTER is on mapping the terrain-soil relationship, particularly at smaller mapping scales / lower resolutions.

Basic data sources for the construction of SOTER units are topographic, geomorphological, geological and soil maps at regional scale or larger (mostly exploratory and reconnaissance maps), as well as digital data such as Digital Elevation Models (DEM) and satellite imagery. In principle any soil map that is accompanied by sufficient analytical data for soil characterization according to the World Reference Base(IUSS Working Group WRB 2015) can be used for map compilation.

For the Danube basin, we use the terrain unit shapefile created using methodologies developed during the eSOTER project (Appendix 2). This is the basis of the geometric database. The SOTER (terrain) units were mapped at a scale of 1:250,000. The GIS file was partitioned into fourteen 'national' datasets as a basis for the actual data compilation by the participating organisations(Ruiperez Gonzalez and Batjes 2019).

For the Danube SOTER project, all SOTER (terrain) units were mapped as consisting of one terrain component with one or more soil components (also required for the databases to be INSPIRE[1] compliant). This is visualized in Figure 2 for a hypothetical case with three SOTER (terrain) units with one terrain component and one to three soil components. Relationships between the various tables of the SOTER attribute database, as described below, are visualized in Figure 3 (view in high resolution).

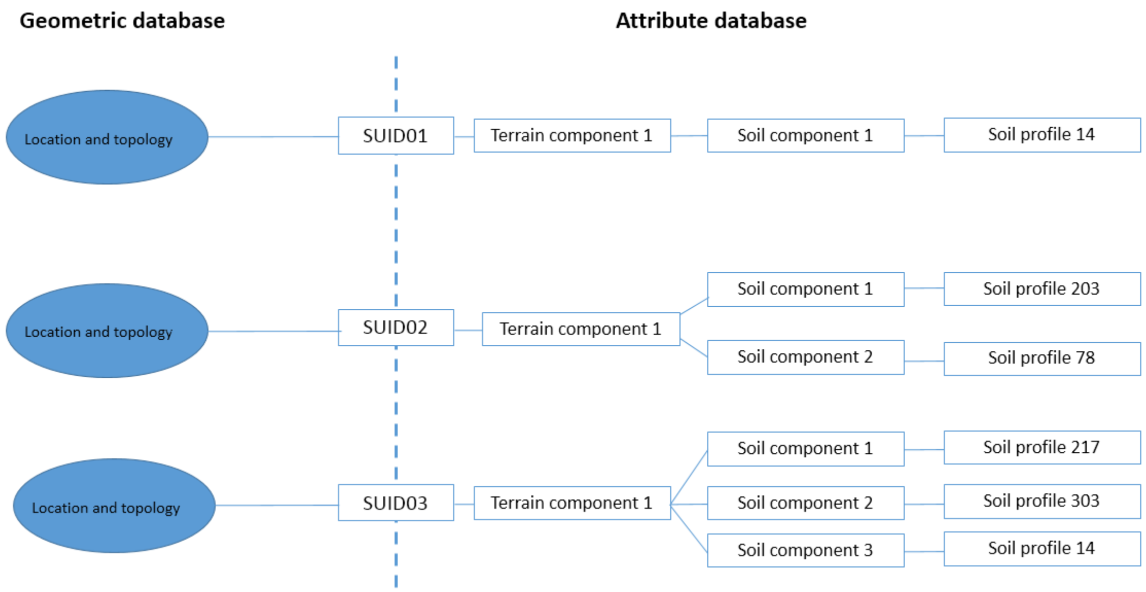


Figure 2. Geometric and attribute database structure for the Danube SOTER

[1] 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

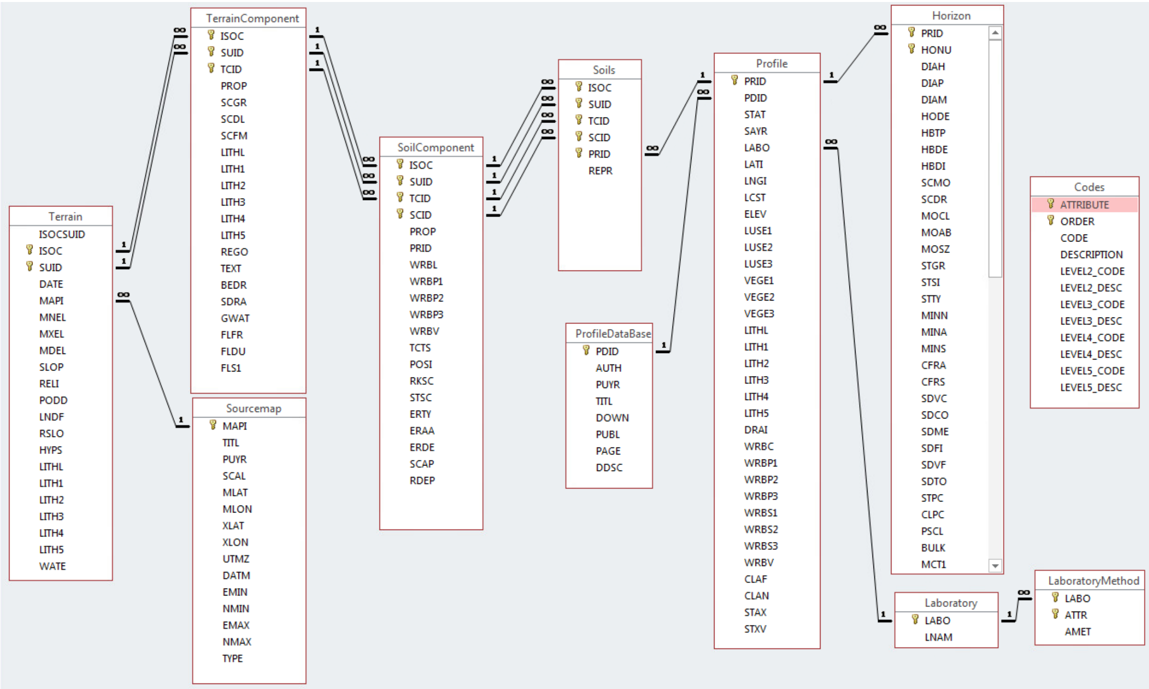


Figure 3. SOTER Danube attribute database with enforced relationships between tables

In short, each SOTER unit is given a unique identification number (*SUID*, SOTer Unique IDentifier) which provides the logical link between the **geometric** database and **terrain** table in the attribute database. Polygons (map units) with identical *SUID*s belong to the same SOTER unit, as characterized in the attribute database. The *ISO*-country code in combination with the *SUID* provides a unique code (e.g., 'HU0001', also called *ISOCSUID*), which is required when combining several national databases into one regional SOTER database.

Each terrain component (**terraincomponent**[2] table) may be comprised of up to 10 soil components; the proportion thereof should be estimated and stored in the **soilcomponent** table. The minimum proportion of a soil component (field *SCID*) is 10% of the corresponding terrain component (*TCID*). The relative position and relationship of soil components vis-à-vis each other within a terrain component is also recorded in the database (see details below).

By definition, the total proportion of soil components within the single terrain component (i.e. Danube SOTER case) will always be 100% (Van Engelen et al., 2013). In some cases, this may be inclusive of small areas of miscellaneous or 'non-soil' units (e.g. water bodies or dunes, see Appendix 8); areas smaller than <10% are considered inclusions[3].

There is an intermediary table that records all the profiles that have been entered into the database (**soils**).

For each soil component, a representative profile (correlated to WRB 2015) has to selected, and its unique indentifier (*PRID* or *Profile_ID*) entered into the **soilcomponent** table. This consists of site information (**profile** table) and measured data of soil properties for each soil horizon for the selected soil profiles (**horizon** table). The source of the profile data must be documented in table **sourcemap**. The analytical

[2] Note: For legibility, all table names have been written in small letters in the text.
[3] Occurrences of soil units smaller than <10% are treated as inclusions; these may be similar or dissimilar. In case of *dissimilar* inclusions, their proportion should be added to that of the largest soil component of the given terrain component. Alternatively, for *similar* inclusions (e.g. same WRB RSG), the PROP should be added to that of a soil component having the same WRB unit if present; otherwise follow the approach used for *dissimilar* inclusions. Again, the sum of PROPs of all soil components within the single terrain component must be 100%.

methods used are stored in table **laboratorymethod**, and the laboratory where the analyses were performed in table **laboratory**.

SOTER is a relational database. This means that, where appropriate, for example, a given same soil profile may be used to characterize different soil components without having to re-enter the data (see Soil Profile 14 in Figure 2).

2.2 SOTER requirements

Source of data

SOTER databases are compiled using national soil data (e.g. maps, survey reports, and digital data sets). In the absence of maps of sufficient detail for a certain study area, or gaps in the available data, it may be necessary to extract the necessary information from broader scale maps. The SOTER approach itself, however, specifically excludes the undertaking of new land resource surveys within its programme (van Engelen and Dijkshoorn 2013).

Correlation to WRB system

Several conventions need to be followed during data compilation. At soil component level, soil types are characterized according to the 'WRB Soil Reference Legend 2014, update 2015' (abbreviated here as WRB 2015, IUSS Working Group WRB 2015). Criteria for separating soil components are based on WRB diagnostic horizons, properties and diagnostic materials. At the present SOTER reference scale of 1:250,000, the soil components should, in general, be characterized up to the third principal qualifier (e.g. Mollic Eutric Stagnic Gleysols, see WRB 2015, p. 14-16 'guidelines for map legends'). Alternatively, the corresponding representative profile (to be described in table **profile**) should be classified in full according to WRB 2015, including the reference soil group (RSG), up to three principal qualifier(s) and up to three supplementary qualifiers (in so far as realistic with the available materials).

Correlation between some national systems (e.g. Feiden *et al.* 2011) and the WRB Legend respectively WRB Classification is not always straightforward (van Engelen and Dijkshoorn 2013). Neither the diagnostic horizons, properties, and classification system(s) in use may fully comply with the WRB conventions. However, conversion to the WRB 2015 Legend is a prerequisite for SOTER Danube and this may require reclassification (correlation) of each selected representative soil profile.

Soil profiles

Every soil component must be characterized by at least one soil profile (see column REPR in table **soils**), but preferably several, fully described and analysed reference profiles. These are to be selected from existing soil information sources (e.g. soil survey reports or digital soil databases). Following careful consideration, <u>one</u> of these reference profiles will be selected as the representative profile for the given soil component. The data for this representative profile must be entered manually, or imported from an external file, into the SOTER database as described elsewhere in this report (see also SOTER 2013, Sections 6.6 and 6.7).

The adopted format is largely based upon the FAO Guidelines for Soil Description (FAO 2006). By implication, profiles described according to earlier FAO Guidelines (FAO-ISRIC 1986; FAO 1977) or the Soil Survey Manual (Soil Survey Staff 1951, 1983), from which FAO has derived many of its criteria, can be entered with little or no reformatting required. Compatibility between the e-SOTER (eSOTER 2012; Kozak and Boruvka 2014) and the relevant parts of the SOTER attribute database should also facilitate transfer of data already stored in databases set up according to FAO, WRB and ISRIC standards.

Horizon notation and depth data

Please adhere to the conventions for entering the upper and lower depth (cm) of a given horizon or layer. Enter the upper and lower depths of the horizon (or layer), measured from the surface, including organic layers (O) and mineral covers, downwards in accord with current conventions (FAO 2006; Schoeneberger *et al.* 2012, p. 2-6). For data sources that are based on 'older' conventions, all depths must be corrected to the new convention.

Field *HODE*, in table **profile**, provides the master horizon (e.g., O or B) with subordinate characteristics (e.g., Ap or Btk1). Organic surface horizons are further identified by the codes 'FO' for folic and 'H' for histic (see **horizon** table under *DIAH*, for diagnostic horizon).

There is one large table, called **codes**, with all the accepted codes and their description (see Appendix 5).

Soil property data

Key soil properties are required for each horizon for the representative soil profiles. These properties must be characterized using measured data only. If there are missing values for some fields, such 'gaps' should not be filled with expert-estimated values. Otherwise, the soil profile data presented in SOTER may not be used meaningfully to underpin digital soil mapping or derive pedotransfer functions.

Software

It is assumed that all consortium partners have licences for MS-Access® and GIS (ESRI ArcMap® or QGIS, that is open source), the two key software components required for the SOTER Danube project.

It is also assumed that all partners will be able to access the necessary base materials and documentation provided by the project co-ordinators and partners from OneDrive and similar on-line platforms (see Appendix 1).

Good practice

It is good practice to make regular copies (backups) of your SOTER database and GIS files. Otherwise, you may accidentally lose any previously compiled information.

Also keep in mind that deletion of some critical content, such as the laboratory identifier (column *LABO* in table **laboratorymethod**), can lead to a 'cascade' of deletions in related tables as database integrity is enforced (see primary keys in Figure 3).

Further, it is good practice to document all data processing steps in a report that accompanies the SOTER database for your region.

2.3 Main steps of data compilation

There are various ways of compiling a SOTER database as described in the 2013 Procedures Manual (van Engelen and Dijkshoorn 2013). Preferred options vary with the country under consideration, regional expertise (e.g. soil science, data management, information technology) and data availability.

For the Danube basin SOTER project, the following steps need to be considered:

- Compile SOTER (terrain) units for your region
- Compile the terrain table
- Document the source map
- Document source of the profile data
- Compile the terrain component table
- Compile the soil component table and soils table
- Compile the profile table
- Compile the horizon table
- Document the soil laboratory and analytical methods

Each step is discussed in a separate section below (Section 3 to 7); selected fields (i.e. *ISOCSUID*, *ISOC*, *SUID*, *TCID* and *SCID*) have already been pre-populated for your study area. Where relevant, reference is made to the corresponding pages in the '2013 SOTER' Procedures Manual (van Engelen and Dijkshoorn 2013), and related SOTER Danube forms (Batjes and Ribeiro 2019).

The general order for entering data in the attribute database is somewhat flexible. However, note that some tables (e.g. **terrain**) will reference another table (e.g. **sourcemap**) to indicate where the source data come from using unique (primary) keys. In such cases, the required information has first to be entered in the referencing tables. Similarly, in order to link a representative profile (table **profile**) to the **soil component** table, the soil components themselves should already be characterized according to the WRB 2015 Legend. The source of the profiles used in **table** profile needs to be specified in table **profiledatabase** and the laboratory where the analytical analyses were made in table **laboratory**, while the laboratory methods themselves are managed in table **laboratorymethod.** Hence, entering references to the source materials used should be done at an early stage in the data entry process.

The different SOTER attribute tables are listed in Table 1. The contents of the shapefile (imported) are stored in table **XX_soter**, where XX is the two-letter ISO country code.

Table 1. Tables for the SOTER Danube project with purpose and suggested data requirements

Tables [a]	**Purpose**	**Suggested data requirements**	**Supporting information**
XX_soter	Geographical delineation of the SOTER (terrain) units for given country (XX)	Based on terrain unit data provided by JRC, subdivided by Danube basin country.	Appendix 2 Ruiperez Gonzalez and Batjes (2019)
Terrain [a]	Table listing all terrain units per country	Populated with country data from the corresponding shape file.	Section 4.1
SourceMap [a]	Specifications for source map	Populated with the information provided by JRC	Section 4.3
TerrainComponent [a]	Registration of terrain components (*TCID*) per SOTER (terrain) unit (*SUID*) in given country (XX).	Pre-populated with the information provided in the **terrain** table (as derived from the GIS shapefile), assumes there is one *TCID* in each terrain unit.	Section 4.2
SoilComponent [a]	Registration of soil components for given terrain component	Derived from integration of national scale soil mapping information to determine the soil components within the mapped terrain unit. Pre-populated from the **terraincomponent** table, assuming there is at least one soil component (and 100% proportion). The number of soil components, and their proportion, within a terrain component may be changed during data compilation, keeping in mind the '100%' area-coverage rule for all soil components combined.	Section 5.1

Tables [a]	Purpose	Suggested data requirements	Supporting information
Soils [a]	List of profiles (WRB 2015) available to represent a given combination of soil component, terrain component and SOTER unit. Provides the basis for selecting a representative profile for each soil component.	Derived from national soil databases or survey reports. Requires correlation to WRB 2015 and harmonization to SOTER conventions.	Section 5.2
Profile [a]	Enter soil profile (site) data	Derived from national soil profile databases or survey reports for the region. Requires harmonization to SOTER conventions.	Section 6.2
ProfileDatabase [a]	Register source(s) from which the profiles were extracted	Derived from sources used.	Section 6.4
Horizon	Enter harmonized properties for given horizon, per profile	Derived from regional soil holdings (analog or digital). Data harmonisation to SOTER conventions (e.g. units of measurement).	Section 6.3
Laboratory [a]	Register laboratory where analyses were made	List of laboratories and their name.	Section 7
LaboratoryMethod [a]	Describe soil analytical methods, used in each laboratory	Concise description of analytical methods derived from laboratory manuals.	Section 7
Codes	Look-up table with SOTER codes and their definitions.	Populated table; should not be altered!	Appendix 5

[a] Tables with forms. Forms in MS-Access® make it easier to view and enter data in the corresponding table.

3. Exploring the SOTER (terrain) units for your region

The shapefile for your country was derived from a terrain unit file for the whole Danube Basin (see Ruiperez Gonzalez and Batjes 2019). It provides the geometry and properties of the different SOTER units. This information has then been imported into the **terrain** table in the MS-Access database for your country.

First, you should familiarize yourselves with the content of the shapefile for your country. For this, open your GIS (e.g. ArcMap® or QGIS) and add file XX_soter. In ArcMap, right click on XX_soter, then open attribute table. The example in Figure 4 is for Austria (AT_soter).

Table

AT_soter

FID	Shape *	ISOCSUID	TUID	HYPSCLAS	SLOPECLAS	RICLASS	SURFCOND	GENETICS	CARBONAT	TEXTURE2	COUNTRY	area
0	Polygon Z	AT100005	1000	4	1	2	Water	N/A		n/a	AT	2.571799
1	Polygon Z	AT100007	1000	4	1	2	Unconsolidated	Aeolian		Sand	AT	41.887122
2	Polygon Z	AT100014	1000	4	1	2	Unconsolidated	Aeolian		Clay	AT	159.388367
3	Polygon Z	AT100059	1000	4	1	2	Unconsolidated	Aeolian		Sand	AT	71.549816
4	Polygon Z	AT100088	1000	4	1	2	Unconsolidated	Aeolian		Loam	AT	32.148753
5	Polygon Z	AT100143	1001	4	1	2	Unconsolidated	Aeolian		Sand	AT	29.536337
6	Polygon Z	AT100153	1001	4	1	2	Unconsolidated	Aeolian		Loam	AT	32.785945
7	Polygon Z	AT100166	1001	4	1	2	Unconsolidated	Aeolian		Sand	AT	69.19295
8	Polygon Z	AT100226	1002	4	1	2	Unconsolidated	Aeolian		Sand	AT	48.655586
9	Polygon Z	AT100299	1002	4	1	2	Unconsolidated	Aeolian		Loam	AT	270.094629
10	Polygon Z	AT100301	1003	4	1	2	Unconsolidated	Aeolian		Loam	AT	82.364835
11	Polygon Z	AT100478	1004	4	1	2	Water	N/A		n/a	AT	11.725667
12	Polygon Z	AT100717	1007	4	1	2	Unconsolidated	Aeolian		Sand	AT	16.083703
13	Polygon Z	AT100861	1008	4	1	2	Unconsolidated	Aeolian		Clay	AT	110.723996
14	Polygon Z	AT100870	1008	4	1	2	Unconsolidated	Aeolian		Clay	AT	174.308812
15	Polygon Z	AT100879	1008	4	1	2	Unconsolidated	Aeolian		Clay	AT	298.691467
16	Polygon Z	AT100889	1008	4	1	2	Unconsolidated	Aeolian		Clay	AT	251.341605
17	Polygon Z	AT100905	1009	4	1	2	Unconsolidated	Aeolian		Clay	AT	99.041918
18	Polygon Z	AT100908	1009	4	1	2	Unconsolidated	Aeolian		Clay	AT	263.069659
19	Polygon Z	AT100945	1009	4	1	2	Unconsolidated	Aeolian		Clay	AT	25.455037
20	Polygon Z	AT100952	1009	4	1	2	Unconsolidated	Aeolian		Clay	AT	250.321294
21	Polygon Z	AT101007	1010	4	1	2	Unconsolidated	Aeolian		Clay	AT	8.011937
22	Polygon Z	AT101052	1010	6	3	4	Unconsolidated	Aeolian		Loam	AT	100.412079
23	Polygon Z	AT101339	1013	7	6	6	Consolidated	Eluvial-colluvial		Loam	AT	80.627856
24	Polygon Z	AT101363	1013	6	3	5	Unconsolidated	Aeolian		Loam	AT	45.833259
25	Polygon Z	AT101401	1014	8	5	6	Consolidated	Eluvial-colluvial		Gravel	AT	84.415269
26	Polygon Z	AT101524	1015	8	6	6	Consolidated	Bare rock		n/a	AT	26.751022
27	Polygon Z	AT101527	1015	8	6	6	Consolidated	Eluvial-colluvial		Gravel	AT	12.380301
28	Polygon Z	AT101580	1015	7	5	5	Consolidated	Eluvial-colluvial		Loam	AT	80.557112
29	Polygon Z	AT101601	1016	6	3	5	Unconsolidated	Aeolian		Loam	AT	34.80165
30	Polygon Z	AT101604	1016	6	3	5	Unconsolidated	Aeolian		Loam	AT	117.876976
31	Polygon Z	AT101610	1016	6	3	5	Consolidated	Eluvial-colluvial		Loam	AT	24.893146
32	Polygon Z	AT101624	1016	7	5	6	Consolidated	Eluvial-colluvial		Clay	AT	47.554244
33	Polygon Z	AT101644	1016	7	5	6	Consolidated	Eluvial-colluvial		Loam	AT	259.073018
34	Polygon Z	AT101646	1016	6	1	3	Unconsolidated	Aeolian		Loam	AT	84.612797
35	Polygon Z	AT101737	1017	7	5	5	Consolidated	Eluvial-colluvial		Loam	AT	64.13195
36	Polygon Z	AT101738	1017	6	1	3	Unconsolidated	Aeolian		Loam	AT	31.02112

1 (0 out of 1020 Selected)

AT_soter

Figure 4. Excerpt of GIS attribute table provided for each country

The most important fields (columns) are *ISOCSUID*; this is a concatenation of the country *ISO* code and number of the given terrain unit (called *TUID* in the shape file). *ISOCSUID* provides the logical link through which the spatial data will be linked to the corresponding SOTER attribute table (XX_soter.mdb).

Other fields include hypsometry (*HYPSCLASS*), slope percentage (*SLOPECLASS*), relief intensity (*RICLASS*), surface conditions (*SURFCOND*), *GENETICS* (e.g. Eluvial-colluvial, Marine, N/A), carbonates (*CARBONAT*), and surface texture class (*TEXTURE2*), as explained in Appendix 2. A selection of these classes can be visualized using html (standard markup language). For the present example (AT), this is shown in Figure 5.

The html files can be accessed through …\html_files\XX_soter.html, where XX is the country ISO code, as described in Ruiperez Gonzalez and Batjes (2019). Clicking on one of the three possible legend classes (upper-right box) will bring you to the given selection (in casu *HYPSCLASS*).

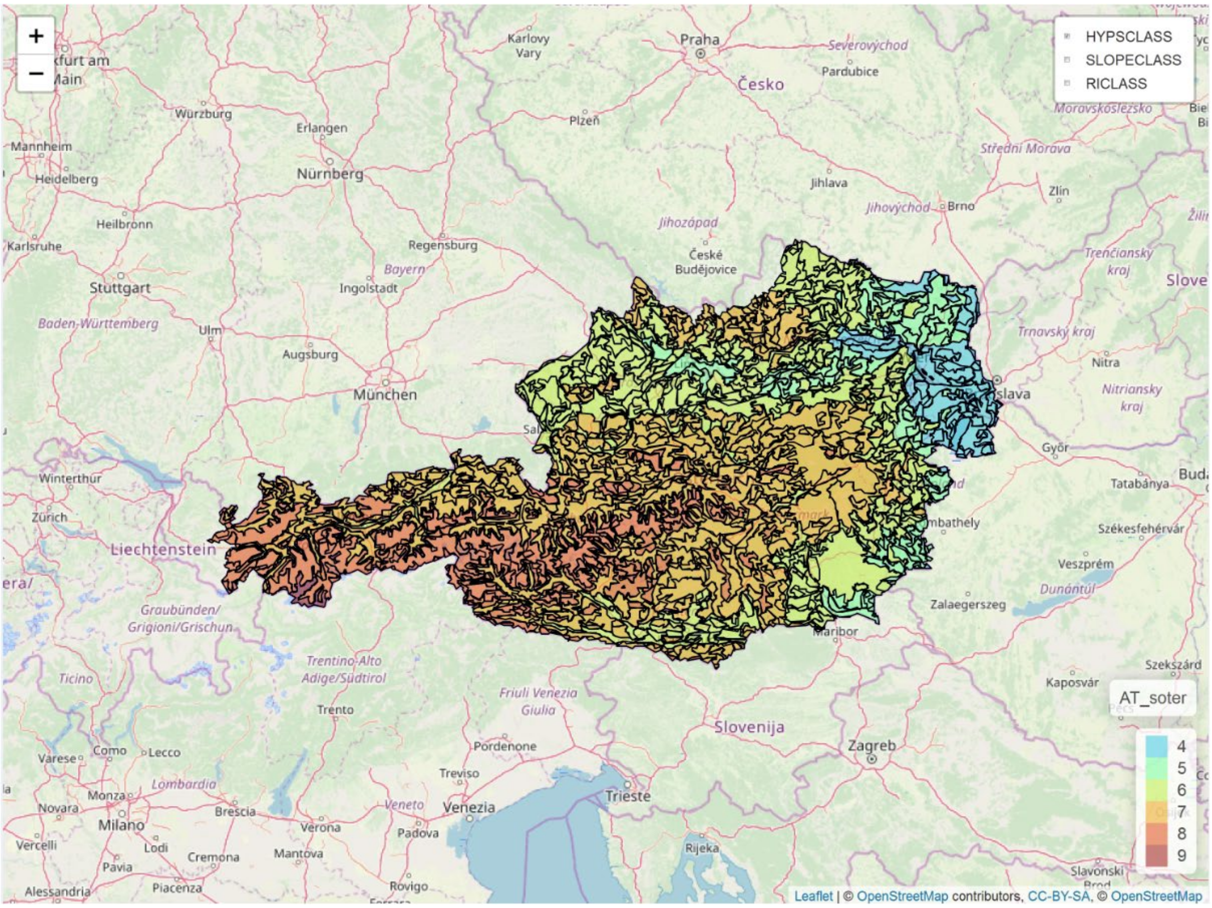


Figure 5. Selected properties (hypsometry) for terrain units visualized using the html tool

4. Prepopulated SOTER terrain-related tables

A number of SOTER attribute tables have already been filled, based on the information provided in the terrain shapefiles (see Section 3). These are the **terrain**, **terraincomponent**, **soilcomponent** and **sourcemap** tables. For the remaining tables, the desired information has to be extracted from available regional sources.

4.1 Terrain

The attribute table for the shapefile (Figure 4) was used to populate the **terrain** table. In Figure 6, we have used the first terrain unit of the populated database for Austria (the first Danube basin country, in alphabetical order) as an example.

It should be noted that many fields (e.g. lithology) are empty as this information was not provided with the original shapefile (see Appendix 2). However, there is no requirement under the Danube project to add further attributes to the **terrain** table, although you may wish to do so if you have the relevant information (e.g. derived from Digital elevation models (DEM)).

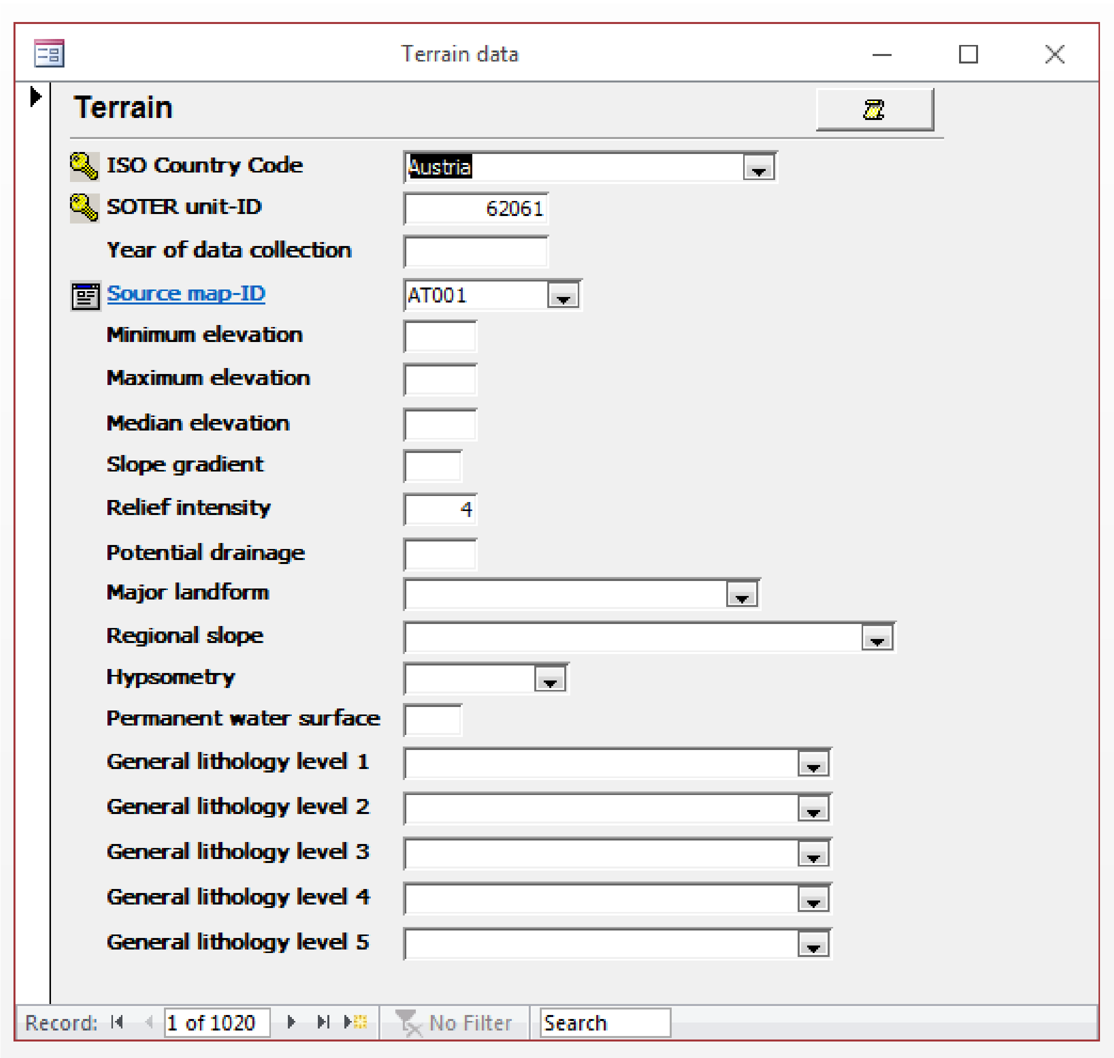


Figure 6. Input form for the terrain table

The structure for the **terrain** table, with both the short and long field name and a short description, is shown in Table 2.

Note: Throughout the report, each SOTER attribute table is shown together with the corresponding form, for ease of reference.

Table 2. Requirements for the terrain table

Field name short	**Field name long**[4]	**Data type**	**Description**	**SOTER 2013**
ISOC[5]	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[6]	--	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[7]	General lithology, level 1	Short Text	A generalized description of the (un)consolidated surficial material, underlying the *larger par*t of the terrain, level 1.	p 37, 14; p37-46, level 1

---

[4] Field names as given on the data entry forms.
[5] Primary keys are indicated in *italics*; these are used to ensure referential integrity within the database.
[6] ISOCSUID can be filled using the following SQL query:
UPDATE Terrain SET Terrain.ISOCSUID = [ISOC]+LTrim(Str([SUID]));
[7] 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**[4]	**Data type**	**Description**	**SOTER 2013**
LITH2	General lithology, level 2	Short Text	A generalized 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 generalized 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 generalized description of the (un)consolidated surficial material, underlying the *larger part* of the terrain, level 4	p37-46, level 4
LITH5	General lithology, level 5	Short text	A generalized description of the (un)consolidated surficial material, underlying the *larger pa*rt of the terrain, level 5	p37-46, level 5

4.2 Terrain component

Similarly, the **terraincomponent** table was populated using the **terrain** table.

The number of terrain components (TCID) has been set at one for all SOTER units (SUID), with proportion (PROP) of 100 %, the default for SOTER Danube (see Batjes and Ribeiro 2019).

If there is readily accessible data for your region, then this may be added to the **terraincomponent** table using the corresponding entry form (Figure 7), but this is not a prerequisite. As indicated earlier, the total proportion of each single terrain component, and its constituent soil components, is always 100 % by definition (Van Engelen et al. 2013).

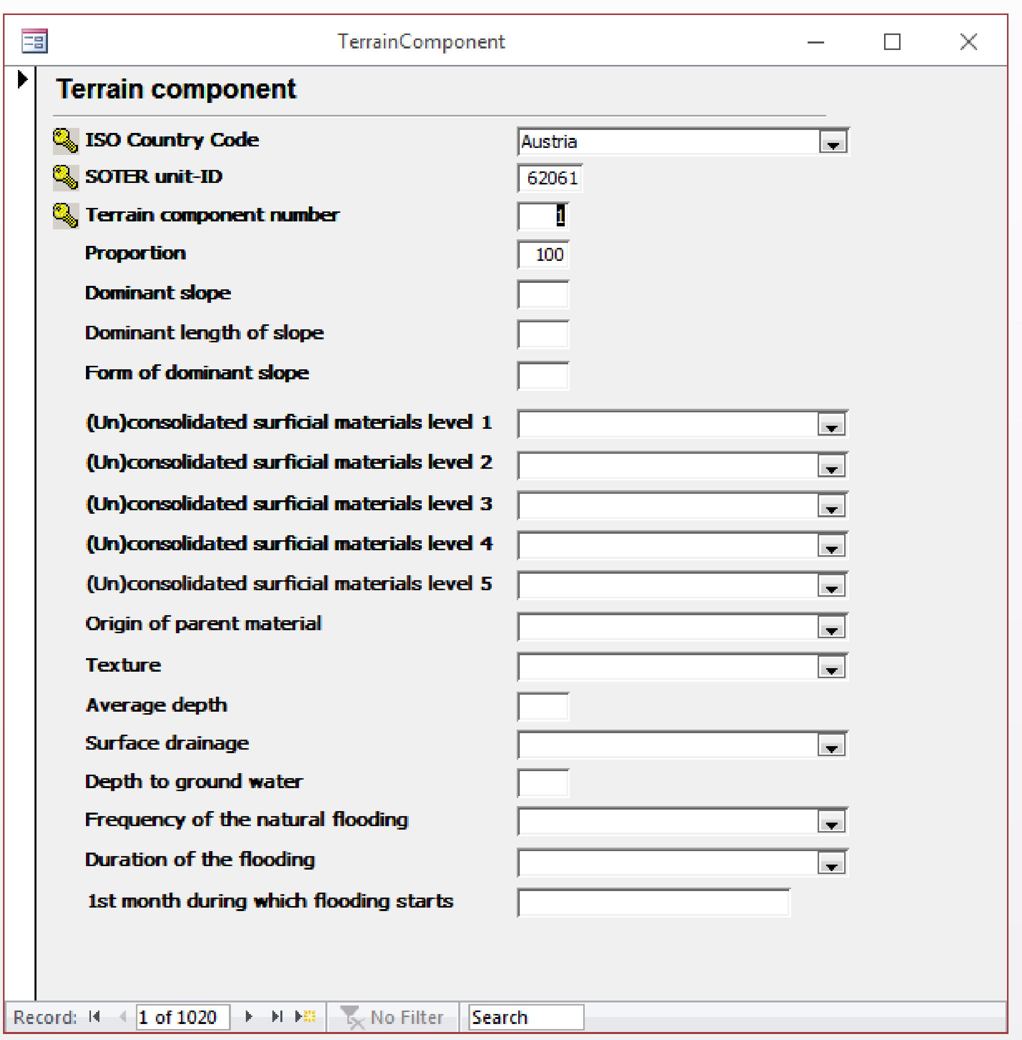


Figure 7. Input form for the terrain component table (populated).

The structure for the **terraincomponent** table with both the short and long field names and short description is shown in Table 3.

Table 3. Requirements for the terrain component

Field name short	**Field name long**[8]	**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	p46, 16
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)	p46, 17
PROP	Proportion	Number	The proportion, as a percentage, that the terrain component occupies within the SOTER unit (For the Danube SOTER DB this is always 100%)	p47, 18

[8] Field names as shown on the data entry forms.

Field name short	Field name long[8]	Data type	Description	SOTER 2013
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 [9]	(Un)consolidated surficial materials, level 1	Short Text	A generalized 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 generalized 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 generalized 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 generalized 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 generalized 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
FLS1	First month during which flooding starts	Number	First month during which flooding of the terrain component normally starts	p51, 32

[9] 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).

4.3 Source map

Similarly, the **sourcemap** table was populated with reference to the terrain unit files provided by JRC. This is entered as 'Danube Basin Terrain Units Layer at scale 1:250,000 (Prep. by JRC)' (Map title), 2015 (Year), and 1:250,000 (denominator of scale), see Figure 8.

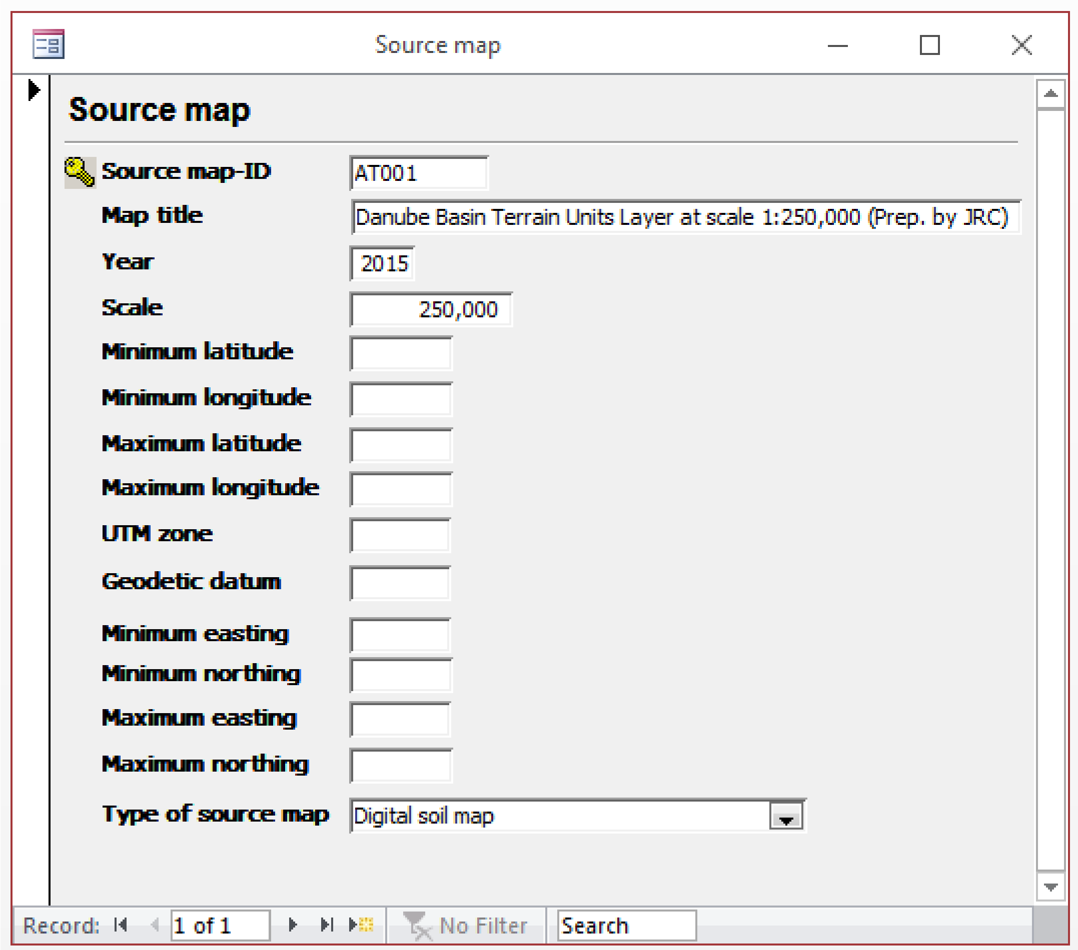


Figure 8. Input form for the map sources (sourcemap table)

The structure for the **sourcemap** table, with both the short and long field names and short description, is shown in Table 4.

Table 4. Requirements for the sourcemap table

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

Field name (short)	Field name (long)	Data type	Description	SOTER 2013
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

5. Compiling the soil component and soils table

The following attribute tables need to be populated for each region: **soilcomponent**, **soils**, **profile**, **horizon**, **profilepatabase**, **laboratory**, and **laboratorymethod** (see Table 1).

Depending on the resources available for each region (e.g. paper maps, survey reports, digital databases or software) and the publication scale of the SOTER product, the data compilation approach may vary (see Dijkshoorn *et al.* 2016; Goyens *et al.* 2007; van Engelen and Dijkshoorn 2013). Yet, the source data must always be processed and harmonized according to SOTER conventions. Divergence from these conventions will hamper future merging of the SOTER databases developed for the respective Danube basin countries.

5.1 Soil component

National scale soil data sources (i.e. soil maps, preferably at scale of ~1:250,000) should be used to determine the soil components within the terrain units. This can be achieved in a GIS environment by intersecting soil mapping data with the terrain units. Each terrain unit must have at least one soil component and can contain up to 10 soil components (each soil component must represent at least 10% of the corresponding terrain unit), depending on the complexity of soils in the terrain unit. As previously described, the SOTER convention for soil classification is WRB 2015 so the criteria for separating the soil components are based on the WRB definitions of diagnostic horizons, properties and materials. Therefore, the number of soil components may be less than those indicated on the national soil map if several national soils are correlated to the same WRB soil type.

The range of attributes to be collated/specified for each soil component in the terrain component is shown in Table 5. These can be added using the **soilcomponent** form (Figure 11). Particularly critical in this table are the records related to the naming of the typical soil type in the area, as characterized according to WRB 2015 Legend conventions. Importantly, the legend unit provides the basis on which the representative soil profile (column *PRID*, table **profile**) can/will be selected[10].

The **soilcomponent** table has been pre-populated using data from the **terraincomponent**, under the assumption that there is one soil component within each terrain component (i.e. *PROP* is 100%). In most cases, however, there will be several soil components within a terrain component; the sum of the proportion of these soil components is always 100%[11] by definition (Van Engelen et al, 2013, p. 51). When necessary, small areas of 'non-soil' (e.g. water bodies or dunes) can be defined at the level of the soil component (see Appendix 8). Conversely, large areas of such 'miscellaneous units' will be mapped in the GIS database only.

The default situation for the pre-populated **terrain**, **terraincomponent** and **soilcomponent** tables for the Danube SOTER is shown in Figure 9.

---

[10] Note: Unlike for the forms, in table view pull-down menus will only work for 1st level entries.

[11] Always start with the dominant soil component; avoid using equal percentages of e.g. 50%-50% in case of two soil components (i.e. rather use 55% and 45%).

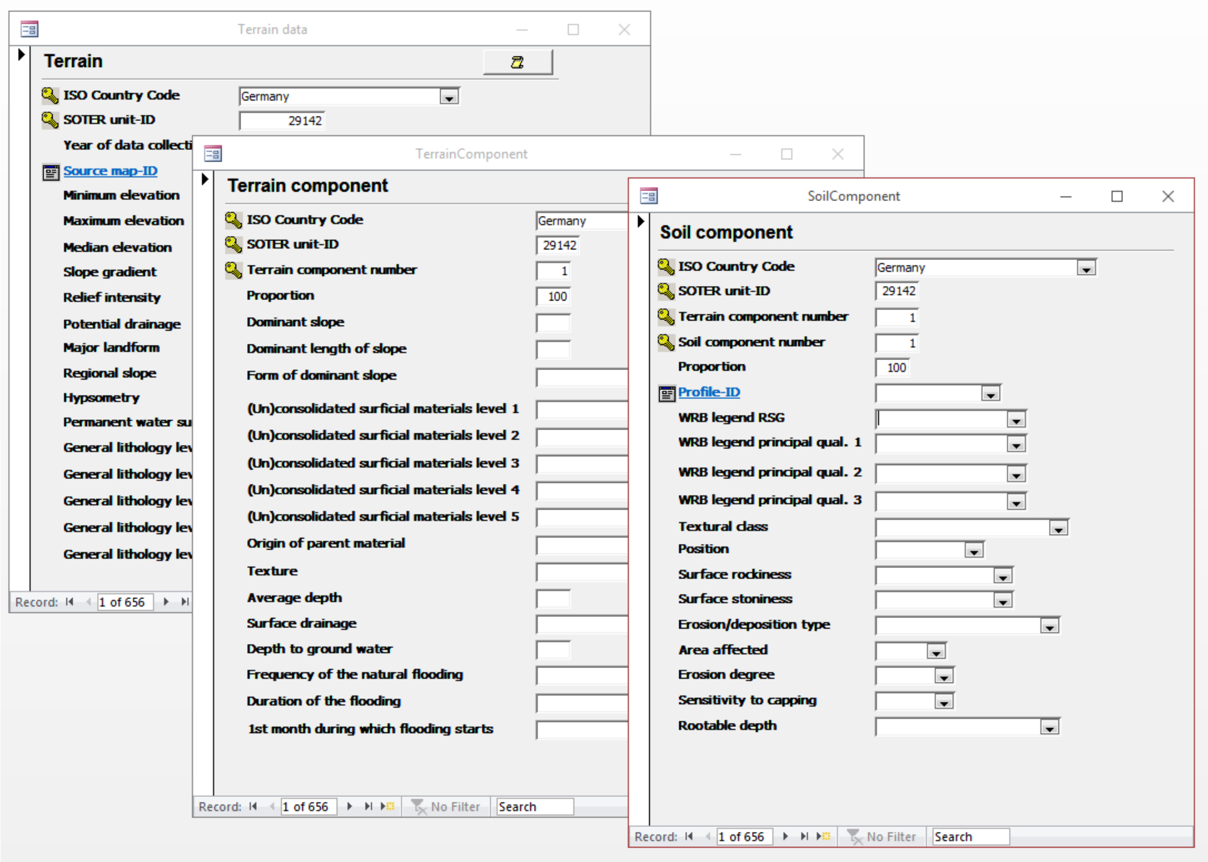


Figure 9. Prepopulated terrain, terrain component and soil component tables for SOTER unit GE29142.

Remember, that for SOTER Danube, the default number of terrain components (*TCID*) per terrain unit (*SUID*) unit is one by default; this number should not be altered! Alternatively, the number of soil components within a given terrain component may be increased. This may be necessary when the regional source maps show a more complex soil pattern. This may be done using the **soilcomponent** form, but it may be easier to do this in the **soilcomponent** table itself, keeping in mind the logical relationships between the different tables. For this, open the **soilcomponent** table, go to the end of the table, and you will see a '*' followed by an empty record: simply, enter your new data. Move to the next row to actually save your data. Again, referential integrity between related tables, e.g. the **terraincomponent** table, will be enforced!

If an 'unacceptable' combination is entered, a message box will appear on the screen. This is illustrated here using SOTER unit 'AT62016' as a hypothetical example. It has been assumed that this SOTER (terrain) unit, with its single terrain component, consists of three soil components (with sum of portions equal to 100%). Should the data compiler erroneously add a second terrain component (*TCID*) into the **soilcomponent** table, an error box will appear on the screen (Figure 10). The text in the message box succinctly describes the source of error, so that it can easily be remedied by the data compiler.

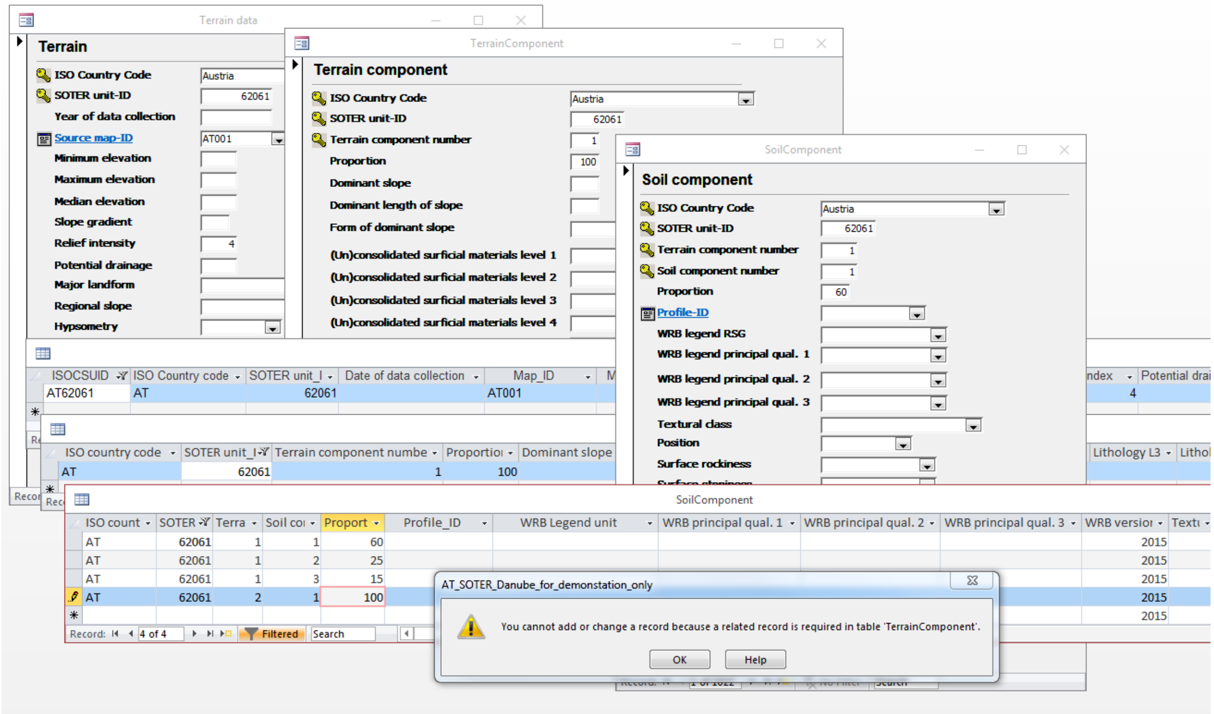


Figure 10. Example of an error message

Table 5. Requirement for the soil component table

Field name	Field name long[12]	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**)	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[13], see WRB 2015, p. 14-16	-

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

Field name	Field name long[12]	Data type	Description	SOTER 2013
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[14]	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
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

Profile_ID is underlined in the form for the **soilcomponent** table (Figure 11). This indicates that a representative profile must be chosen from the **soils** table, which may be done using the pull-down menu. At the beginning of data entry for your region, however, the **profile** table will be empty. It needs to be populated by the regional data compilers first, using the available source materials. Once this has been done, the regionally appropriate *Profile_ID* for the **soilcomponent** table and **soil** table may be selected from the list of so far 'entered' (representative) soil profiles. This can be done using the pull-down menu for *Profile_ID* in the corresponding tables.

[Hint: To see/go to the next entry in a form unit, use the ► button at the bottom of the form; this also ensures that any changes are written to the underpinning table].

The overall procedure for filling the database for your region is illustrated in Appendix 6 using an hypothetical example.

---

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

Initially, the look-up table for *Profile_ID* is empty. Clicking on the *Profile_ID* field will bring you to the (still empty) **profile** table.

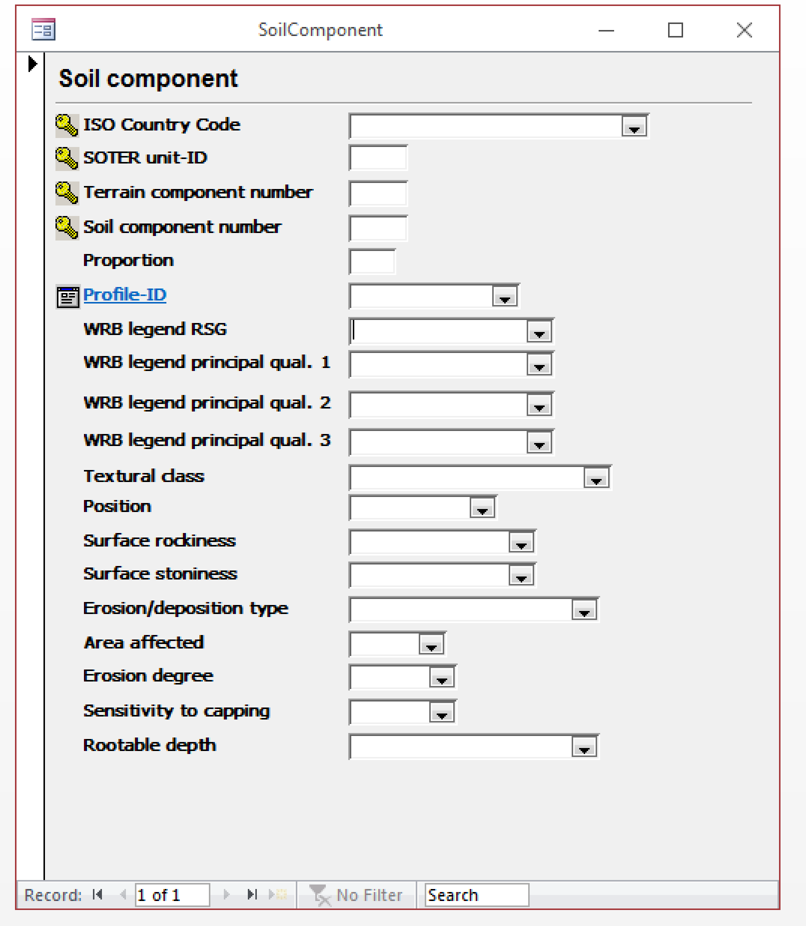
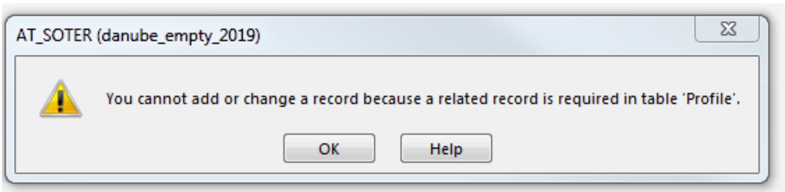


Figure 11. Input form for the soilcomponent table

5.2 Soils

The structure of the **soils** table is shown in Table 6. This table documents all profiles available (i.e. that have been entered) for a given a soil component within a given terrain unit. When the box for Representative is ticked (Figure 12), the corresponding profile is considered to be (most) representative for the given soil component.

In order to be able to fill the **soils** table, i.e. select a representative profile for a given soil component, the **profile** table should already be populated (i.e. populate this table first, see relationships in Figure 3). Otherwise, you may get this type of message:



[Tip: It is good practice to regularly close a form so that the new data entries are saved. Once this has been done, the new entries will become visible/selectable in the related pull-down menus in other tables].

Table 6. Properties to be described in the soils table.

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	Representative	Yes/No	Representative (Yes/No). Yes if selected to represent the given Soil component	New

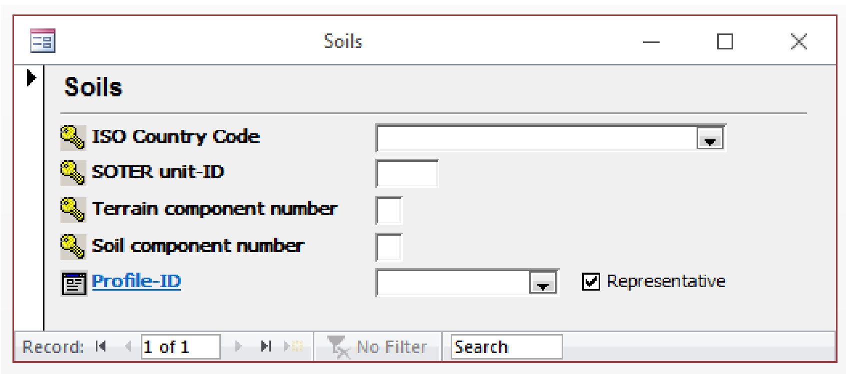


Figure 12. Input form for the soils table

6. Compiling the profile and horizon table

6.1 General procedure

The properties for each (representative) profile have to be entered into the **profile** table. The full range of soil profiles that can be used to characterize the various soil components is documented in the **soils** table (i.e. once they have been entered in the **soils** table, they can be selected using pull down menus to be representative for a given **soilcomponent**, see Figure 3).

Note that the same representative profile may be used to characterize soil components in different SOTER (terrain) units that have the same WRB 2015 soil legend units (see Figure 2, 'soil profile 14').

Each representative profile (*Profile_ID*) should be selected from the available national soil profile database or alternatively be compiled from paper sources; in both cases all profiles will have to be correlated to WRB2015 and the data standardized to the SOTER conventions. Typically, this will require a correlation step from the national classification system to the closest possible equivalent in the WRB 2015 system (i.e., RSG and from 1 to 3 principal respectively supplementary qualifiers, see IUSS Working Group WRB 2015, p. 12-21). Once this has been done, the WRB classification and other data (Table 8) can be entered directly using the corresponding form (Figure 13).

Alternatively, the source data may be imported from existing national databases using e.g. SQL queries. Again, all national data must first be harmonized to the SOTER conventions (e.g. 'land use X' in the source data set correlates best with property *Land use L1,* class A (Agriculture) in SOTER). Once this essential step has been done, the now 'pre-harmonized' data files can be imported into the SOTER database model (see Figure 3). Several options are provided for this by MS-Access (i.e., import text files, excel files, data from other MS-Access tables).

[Tip: See for example https://www.quackit.com/microsoft_access/tutorial/ for guidance when needed]

Correlation to WRB2015 (i.e. legend respectively classification) is necessary at two levels. First, at the level of the **soilcomponen**t to characterize the respective soil components (*SCID*) that occur in a given terrain component (*TCID*) and SOTER unit (*SUID*) in a given region (*ISOC*). This is required, so that a representative profile can be selected from the **profile** table (once this table has been populated!).

Correlation to WRB 2015 can be done on a profile-by-profile basis, with due consideration of all the necessary requirements (recommended as most secure, yet time consuming). However, it may be more pragmatic to develop/compile region-specific look-up tables for this. Such tables make a correlation between e.g. the national soil classification names and the WRB reference soil groups and principal and supplementary qualifiers. A suggested file structure for this is shown in Table 7; examples of existing correlation tables are presented in Krasilnikov et al. (2019).

Table 7. Simplified structure for a correlation table from a hypothetical 'national system' to WRB 2015.

National soil system		**WRB 2015**	**Principal qualifier (WRBP) [c]**			**Supplementary qualifier (WRBS) [c]**			**Confidence in correlation [b]**
Soil name [a]	Year	Reference Soil Group (RSG)	WRBP1	WRBP1	WRBP3	WRBS1	WRBS2	WRBS3	
Calcareous stony soils	1988	Regosols	Leptic			Calcaric	Clayic		High
Poorly drained soils, high base saturation	1988	Gleysols	Mollic	Oxygleyic		Siltic			Medium

National soil system		WRB 2015	Principal qualifier (WRBP) [c]			Supplementary qualifier (WRBS) [c]			Confidence in correlation [b]
Soil name [a]	Year	Reference Soil Group (RSG)	WRBP1	WRBP1	WRBP3	WRBS1	WRBS2	WRBS3	
High fibre organic soil	1988	Histosols	Fibric			-			Medium

[a] Fictional soil names, used as example only.

[b] Rated as L(ow), M(edium) and H(igh) depending on the overall confidence in the translation/correlation.

[c] Note: The input files (i.e. Excel) should have separate columns for the various qualifiers (i.e. WRBP1, WBRP2, WBRP3 and WBRS1, WRBS2, WRBS3), as this is a prerequisite for successfully importing the data into SOTER afterwards.

The required, so-called *site* properties are stored in table **profile** (Table 8). Again, the list is largely self-explanatory. Primary key *PRID* (alias *Profile_ID*) provides the unique identifier for a given representative profile. The recommended way of entering the data is via the provided forms (see sections below), as this will ensure that the referential integrity is preserved, and only the accepted SOTER definitions are used. This approach is to be used by default when the source data are in paper format. Alternatively, for SOTER Danube countries that already have digitized their soil profile data these may be imported into the relevant SOTER tables. However, this should be done with utmost care to ensure referential integrity.

First, the available digital soil data should be converted to their SOTER equivalents, strictly respecting the SOTER coding and naming conventions. This may be done by creating a number of 'intermediary' files. For example, 'national' vegetation classes, drainage classes should be pre-harmonized to the corresponding SOTER conventions using look-up tables of the type earlier (Table 7). This may be done using Excel files and similar. Similarly, all soil profiles must be georeferenced according to WGS 1984 (van Engelen and Dijkshoorn 2013).

As indicated, it is critical that only the SOTER terminology is used in such 'intermediary' files! During the subsequent import to the SOTER database the referential integrity will be enforced/checked for the primary keys only, see Appendix 4. Wrongly coded classes for e.g. land use or WRB2015 may be imported as such corrupting the database!

As indicated, correlation to WRB 2015 is a prerequisite in SOTER Danube. Other classification systems (FAO, USDA or national, and year of publication) are not compulsory, but often useful to have when known/provided in the source profile database.

Column *LABO* holds the unique identifier for the laboratory where the various analyses have been performed. Being a primary key, it is necessary to fill in the required information for laboratories (table **laboratory**) before these can be selected from the **profile** table (using the pull-down menus).

As repeatedly indicated, referential integrity is enforced in SOTER. Therefore, be very careful when deleting certain 'primary key' columns, such as the *PRID* or *LABO* as such a deletion will 'cascade' (i.e. may cause unforeseen and irreversible deletions).

Further, as stressed earlier, you should make regular backups of your SOTER attribute database and any related 'intermediary' files to avoid having to repeat your earlier data compilation work!

6.2 Profile

The site properties that can be accomodated in the **profile** table are listed in Table 8; the corresponding data entry form is shown in Figure 13. The more complete the information that can be provided for each region, the more useful the SOTER data will be later.

Please note that, at this stage, data that do not meet the referential criteria for SOTER, for example a 'not yet registered' *PROFILE-ID* will be 'refused' by the system thereby safeguarding the integrity of the SOTER tables for your region.

Table 8. Properties represented in the profile table

Field name (short) [a]	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	Latitude	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
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.	-

Field name (short) [a]	Field name (long)	Data type	Description	SOTER 2013
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 version	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 2014a)	p59, 73
STXV	Soil Taxonomy version (year)	Number	Soil Taxonomy version (year)	p79, 74

[a] 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. WRBC*).

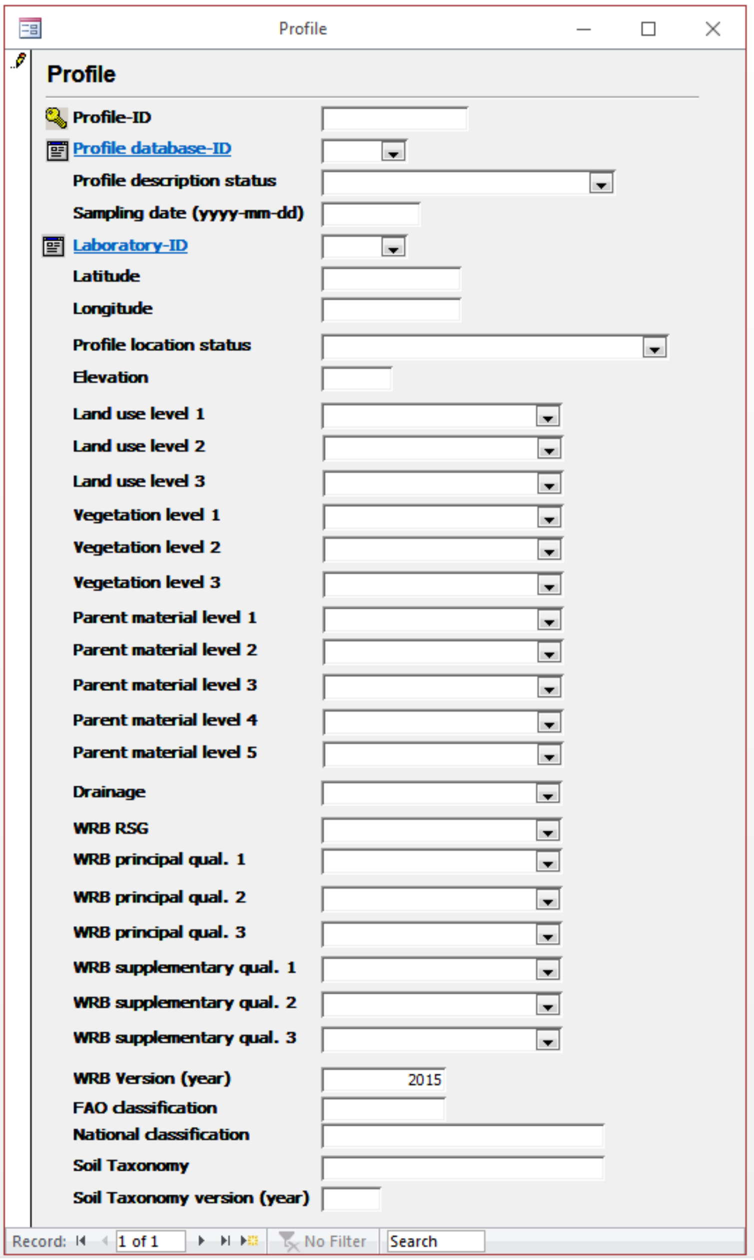


Figure 13. Input form for the profile (site) data.

The selection of profiles for the country, from which the potential representative profile(s) are to be selected, can be managed in table **soils** or extracted from the national soil profile database (after their harmonisation/conversion to SOTER conventions). All representative profiles (i.e., used in the **soil component**) are (to be) flagged in table **soils** (column *REPR* ticked as 'Yes'; Figure 12).

The required fields need to be entered per profile (i.e., define the unique *Profile_ID*). Next, add the source of the data (i.e. click on Profile-database-id) to populate **profile** table. Similarly, details for the laboratory where the analyses were performed must be entered (i.e. populate table **laboratory**). Once the **profile** and **laboratory** and similar 'reference' tables get populated, the corresponding pull-down menus may be used (i.e., for profiles extracted from the same report).

It is good practice to close the data entry forms at regular intervals so that the data are actually stored; subsequently, they will become visible in the related pull-down menus.

6.3 Horizon

Range of soil properties and units of measurement

The range of data available for each region can vary considerably per country and per data source, largely depending on the purposes of the original soil surveys (Landon 1991; Soil Science Division Staff 2017) and the period during which they were carried out. So, it is unlikely that all fields considered in the **horizon** table (Table 9) can be filled for each profile/horizon.

However, there is a minimum set of soil attributes that is generally needed if any realistic interpretation of the horizon data is to be expected. Therefore, selected properties are considered as 'mandatory' for SOTER; soil profiles that do not have this essential information should not be selected as representative profiles for SOTER (however, they are useful to document/compile in an underpinning national soil database).

The following attributes are termed mandatory in SOTER(van Engelen and Dijkshoorn 2013): horizon depth (as defined by the upper and lower limit), matrix colour, structure, texture, pH, CEC, cation composition, $CaCO_3$, organic Carbon and total Nitrogen, with succinct information on the analytical procedures used, as described below (Table 12 and Figure 16, table **laboratorymethod**) and the laboratory where the analyses were performed (Table 11 and Figure 15, table **laboratory**).

Of course, when available, all possible/available soil attributes should be entered into the **horizon** table. Overall, it is good practice to compile an *as complete as possible* set of attributes irrespective of the SOTER table.

As indicated, special attention must be paid to the mandatory units of measurement, for example g kg^{-1} for organic carbon (see Table 9, **horizon** table).

When defining the upper and lower depth for each horizon (or layer), adhere (or convert) to the conventions of FAO Guidelines for Soil Description (FAO 2006) as endorsed in SOTER (see end of Section 2.2).

Data entry

There is no data entry form for the **horizon** table, as these would be extremely long and cumbersome to use. Entry of horizon data is best done in the 'data view' mode in MS-Access. Often, as indicated, there will be no data for many fields in the **horizon** table. For ease of data entry, these fields can be hidden to make the 'views' more compact hence user-friendly.

[*Tip: Go to header of given field in the **horizon** table, right click on it, and then select 'hide'. To reverse the procedure, do the same and select the 'unhide field' option*].

As already discussed, the required data can also be imported from an auxiliary file (e.g. Excel or MS-Access database). However, as indicated in Section 6.2, the corresponding source data must always be 'pre-harmonized' to the SOTER conventions, i.e. follow the terms as defined in the SOTR conventions. Otherwise, 'referential integrity' errors will occur during the import stage! The overall procedure for pre-harmonising descriptive (class) properties is illustrated in Appendix 7 using soil drainage (*DRAI*) as an example.

Table 9. Properties represented in the horizon table (with units of measurement)

Field name (short) [15]	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 [16]	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

[15] 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-harmonized) data from a national soil (profile) database. For paper resources, you should enter the data directly into the profile table (respecting the SOTER conventions).

[16] 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 **laboratorymethod**, column *AMET*.

Field name (short) [15]	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%).

6.4 Source of profile data (profile database)

The source of the profiles considered in the **profile** table must documented table **profiledatabase** (Table 10). This can be done using the form described in Figure 14 or in the table view mode.

Table 10. Requirements for the profiledatabase table

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

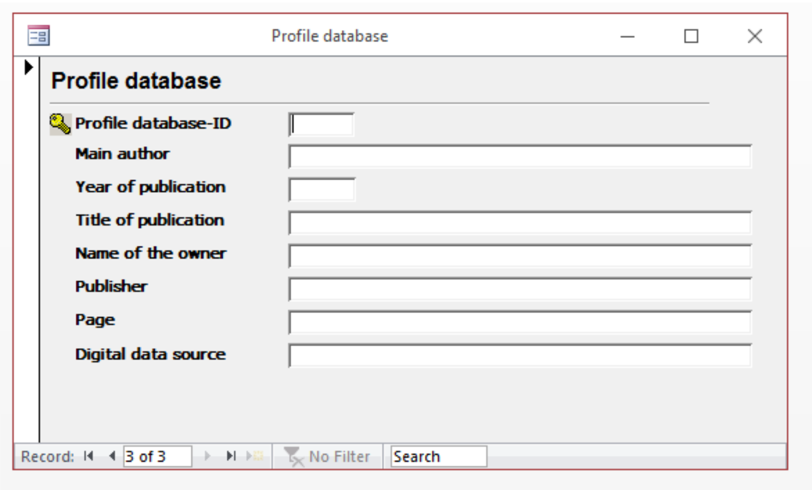


Figure 14. Input form for the source of the profile data (profiledatabase)

7. Document the laboratory and analytical methods

7.1 Laboratory

Soil properties are measured according to a wide range of procedures worldwide; for any given property, these may vary within and between countries, and change over time as new analytical approaches are developed (see Batjes *et al.* 2017). Therefore, it is crucial to document the laboratory of origin (Table 11 and Figure 15) as well as the analytical methods used for analysing the various properties (Table 12 and Figure 16).

Each laboratory is assigned a unique code (*LABO* or *Laboratory_ID*) which consists of the *ISO country code* (ISOC) followed by a number. For example, 'RO002' for the second laboratory represented in the dataset for say Romania (ISOC= 'RO'). The name of the laboratory needs to be provided in full (English spelling) with its acronym.

If there is a major change in some analytical methods in a given laboratory, document/enter this as being a 'new' laboratory. This could be the case, for instance, when the lower size limit for the 'coarse earth fraction' has been changed from '> 1 mm' to '> 2 mm' in the given laboratory. As indicated, it is important to record whether the chemical analyses are based on the '< 2 mm' size fraction (Soil Survey Staff 2014b; van Reeuwijk 2002) or the '< 1mm size fraction' (Katschinski 1956), as commonly applied in many Eastern European countries before the nineties. This in view of possible, future, harmonisation of analytical methods as foreseen in the framework of Pillar 5 of the Global Soil Partnership (Baritz *et al.* 2017).

Table 11. Documenting soil laboratory names

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

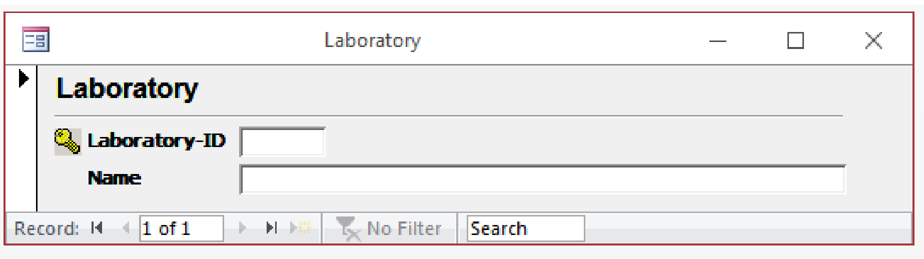


Figure 15. Form for entering the soil laboratory name

7.2 Laboratory methods

For each laboratory entered in the **laboratory** table, the different analytical methods in use must be documented. This is done in table **laboratorymethod**. For each attribute (field *ATTR* or *Attribute)*, a concise description must be provided (*AMET* or *Description)*. For example, soil pH measured in a KCl solution (*ATTR*= 'PHKC') according to the following method (field *AMET*): 'Soil pH measured in a 1:2.5 soil / 0.1M KCl solution (van Reeuwijk 2002, p. 56).

Table 12. Describing the soil laboratory methods

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.	

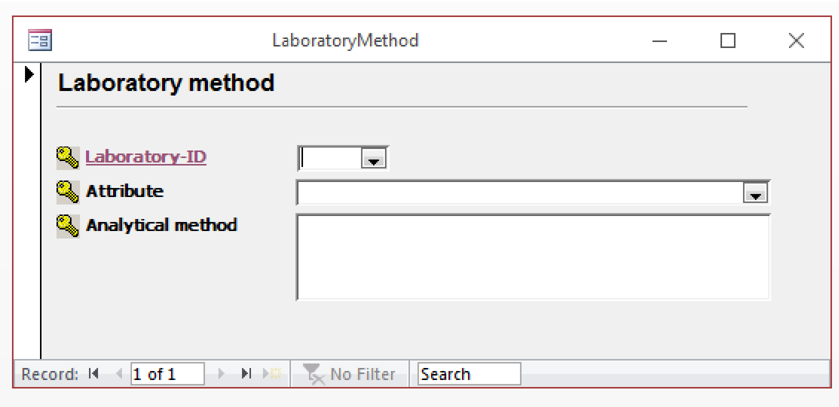


Figure 16. Form for entering soil analytical method descriptions

Acknowledgments

We gratefully thank Dr Jacqueline Hannam and Dr Stephen Hallett (Cranfield University) for their insightful editorial comments.

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 subcontracts to 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).

Last updated: 09 February 2021

References

Baritz R, Erdogan H, Ahmadov H, Ghanma I, Lalljee VB, Wongmaneeroj A, Collins A, Monger C, Ribeiro JL, Bertsch F, Lalljee VB, with, Montanarella L, Comerma J, Khan A, VandenBygaart B, Gaistardo CC, Constantini E, Galbraith JM, Schad P, Lame F, Suvannang N, Hartmann C, Medyckyj-Scott D, Batjes NH, van Liedekerke M and Ziadat F 2017. *Implementation Plan for Pillar Five of the Global Soil Partnership: Providing mechanisms for the collation, analysis and exchange of consistent and comparable global soil data and information [Submitted to ITPS 05/2017]*, Rome, 48 p. http://www.fao.org/3/a-bs756e.pdf

Batjes NH and Ribeiro E 2019. *Updated MS-Access SOTER template for the EU Danube basin project with worked examples (Deliverabes A1, A2 and A4 EU JRC Service Contract Nr. 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

eSOTER 2012. eSOTER database at scale 1:1 million. https://www.esoter.net/products

FAO-ISRIC 1986. *Guidelines for soil description (3rd Edition, Rev.)*. FAO, Rome, 70 p.

FAO 1977. *Guidelines for the description of soils*. FAO, Rome

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

FAO 2006. *Guidelines for soil description (Fourth ed.)*, FAO, Rome, 97 p. http://www.fao.org/docrep/019/a0541e/a0541e.pdf

Feiden K, Kruse F, Řezník T, Kubíček P, Schentz H, Eberhardt E and Baritz R 2011. Best Practice Network GS SOIL Promoting Access to European, Interoperable and INSPIRE Compliant Soil Information. In: Hřebíček J, G Schimak and R Denzer (editors), *Environmental Software Systems. Frameworks of eEnvironment: 9th IFIP WG 5.11 International Symposium, ISESS 2011, Brno, Czech Republic, June 27-29, 2011. Proceedings*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 226-234. http://dx.doi.org/10.1007/978-3-642-22285-6_25

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.

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 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

Katschinski NA 1956. Die mechanische Bodenanalyse und die Klassifikation der Böden nach ihrer mechanischen Zusammensetzung, *Rapports au Sixiéme Congrés International de la Science du Sol*. International Society of Soil Science, Paris (FR), pp 321-327. http://iuss.boku.ac.at/files/vi_e_congres_international_de_la_science_du_sol_volume_b__paris_1956_compressed.pdf#page=323

Kozak J and Boruvka L 2014. The experiences of building CZESOTER-Czech form of SOTER in the scale 1:250,000. *GlobalSoilMap: Basis of the Global Spatial Soil Information System - Proceedings of the 1st GlobalSoilMap Conference*, pp 59-62. http://www.scopus.com/inward/record.url?eid=2-s2.0-84893668225&partnerID=40&md5=cdc38ffe225aa9c88b9ddcab9dbf4c54

Krasilnikov P, Ibánez Marti JJ, Arnold R and Shoba S (editors) 2019. Soil Terminology, Correalation and Classification. Earthscan, London (UK), xviii + 440 p.

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

Ruiperez Gonzalez M and Batjes NH 2019. *Country terrain unit shapefiles of the EU Danube basin (Deliverabe 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

Schoeneberger PJ, Wysocki DA, E.C. Benham and Soil Survey Staff 2012. *Field book for describing and sampling soils (ver. 3.0)*. National Soil Survey Center Natural Resources Conservation Service, U.S. Department of Agriculture, Lincoln (NE) http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_052523.pdf

Soil Science Division Staff (editor) 2017. Soil survey manual. Government Printing Office, Washington, D.C., 639 p. https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=nrcseprd1333016&ext=pdf

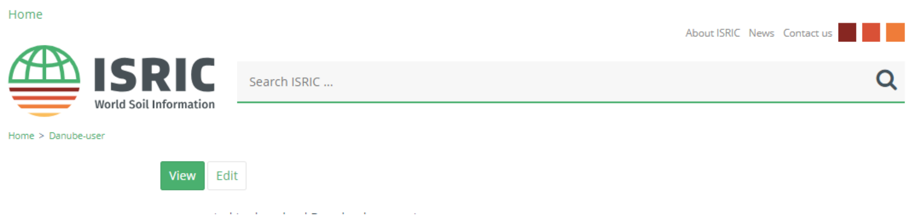
Soil Survey Staff 1951. *Soil Survey Manual*. United States Agriculture Handbook 18, USDA, Washington, 203 p. https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/ref/?cid=nrcs142p2_054262

Soil Survey Staff 1983. *Soil Survey Manual (rev. ed.)*. United States Agriculture Handbook 18, USDA, Washingtonhttp://soils.usda.gov/technical/manual/
Soil Survey Staff 2014a. *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
Soil Survey Staff 2014b. *Kellogg Soil Survey Laboratory Methods Manual (Version 5.0. R. Burt and Soil Survey Staff (ed.))* Soil Survey Investigations Report No. 42, . U.S. Department of Agriculture, Natural Resources Conservation Service, Lincoln (Nebraska), 1001 p. https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1253872.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
van Reeuwijk LP 2002. *Procedures for soil analysis (6th ed.)*. Technical Paper 9, ISRIC, Wageningen, 81 p. https://www.isric.org/sites/default/files/ISRIC_TechPap09.pdf

Appendix 1. SOTER country files and reference manuals

A) The following reports are needed during data compilation:
- WRB 2015 classification system: English PDF; Czech PDF; Russian PDF; Slovene PDF; French PDF
- FAO Guidelines for soil description (2006, 4th edition) PDF
- SOTER Procedures Manual (2013) PDF

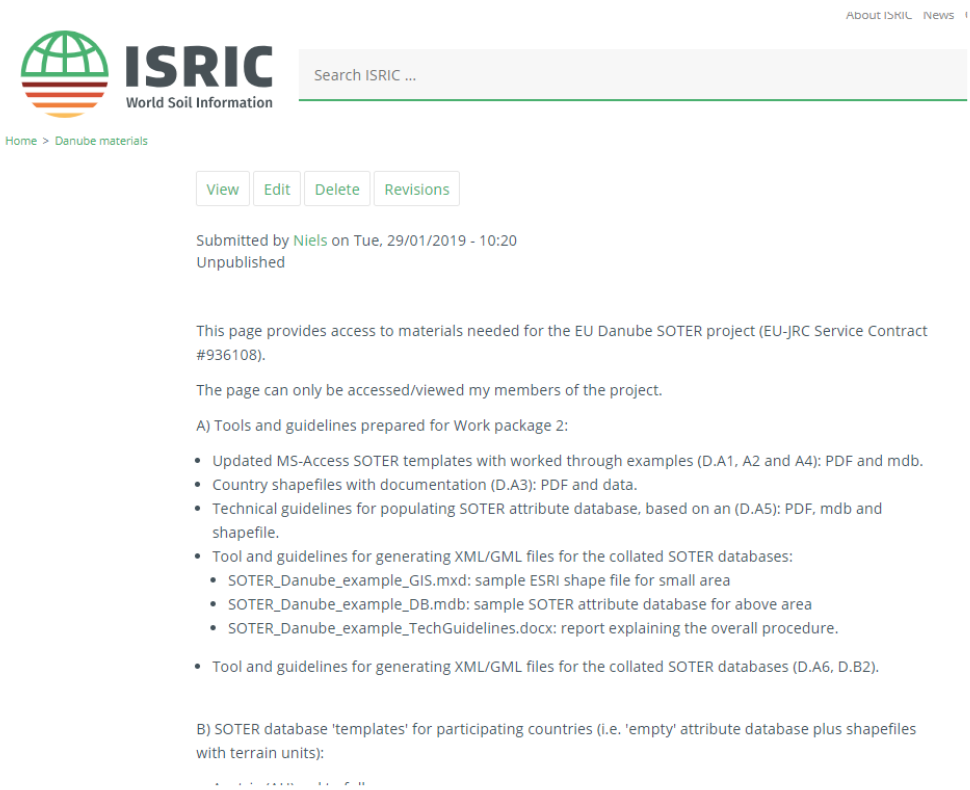
B) The 'country' shape files and 'empty' SOTER databases for the 14 Danube basin countries, as well as the present sample set, can be accessed from the ISRIC website. For this, the user first has to login via https://www.isric.org/user/login. Next use:
User name: Danube-user; Password: xxx (will be provided separately).



Next enter the following string/URL in the search box of your browser:

https://www.isric.org/danube-materials?auHash=n5oCQ8NGoU9wAmOFBUJ4xQo3Is8An32tPJcsSiUX1V0

Then click enter in your browser's search bar. You then are forwarded to the following webpage (which will gradually be filled with data/information):



C) The original mask and terrain shapefile for the Danube basin were made available by JRC:

https://esdac.jrc.ec.europa.eu/projects/esoter-danube (Appendix 2).

Appendix 2. Description of the SOTER Danube basin terrain unit layer

Provided by: Marc van Liederkerke (JRC)

The Danube Basin Terrain Units layer at scale 1:250,000 has been created using a methodology similar to the one that was developed in the eSOTER FP7 project (*http://www.esoter.net*). The resulting layer is a shapefile whose polygons delineate Terrain Units (TU) as defined in the SOTER methodology, at scale 1:250,000.

The shapefile contains a number of attributes of which the TUID, being the unique Terrain Unit identifier is the most important; the other attributes are the results of various input and subsequent computations that have led to this layer. They can be regarded as additional information which may be useful when using this layer of terrain units.

<table>
<tr><td>TUID</td><td>Unique identifier for the Terrain Unit</td></tr>
<tr><td>HYPSCLASS</td><td>
<table>
<tr><td>Hypsometry classes (elevation above sea level in metres)</td><td>hypsclass codes</td></tr>
<tr><td>0-10</td><td>1</td></tr>
<tr><td>10-50</td><td>2</td></tr>
<tr><td>50-100</td><td>3</td></tr>
<tr><td>100-200</td><td>4</td></tr>
<tr><td>200-300</td><td>5</td></tr>
<tr><td>300-600</td><td>6</td></tr>
<tr><td>600-1500</td><td>7</td></tr>
<tr><td>1500-3000</td><td>8</td></tr>
<tr><td>3000-5000</td><td>9</td></tr>
<tr><td>5000+</td><td>10</td></tr>
<tr><td></td><td></td></tr>
</table>
</td></tr>
<tr><td>SLOPECLASS</td><td>
<table>
<tr><td>Slope percentage (%)</td><td>Slopeclass code</td></tr>
<tr><td>0-2</td><td>1</td></tr>
<tr><td>2-5</td><td>2</td></tr>
<tr><td>5-10</td><td>3</td></tr>
<tr><td>10-15</td><td>4</td></tr>
<tr><td>15-30</td><td>5</td></tr>
<tr><td>30-45</td><td>6</td></tr>
<tr><td>45+</td><td>7</td></tr>
</table>
</td></tr>
<tr><td>RICLASS</td><td>
<table>
<tr><td>Relief Intensity (elevation range (m) within 1 km2)</td><td>Class codes</td></tr>
<tr><td>0-5</td><td>1</td></tr>
<tr><td>5-20</td><td>2</td></tr>
<tr><td>20-50</td><td>3</td></tr>
<tr><td>50-100</td><td>4</td></tr>
<tr><td>100-300</td><td>5</td></tr>
<tr><td>300+</td><td>6</td></tr>
</table>
</td></tr>
<tr><td>SURFCOND</td><td>Surface condition, consolidated and unconsolidated parent material

Values: Unconsolidated, Consolidated, Water</td></tr>
<tr><td>GENETICS</td><td>Values: Eluvial-colluvial, Marine, N/A</td></tr>
<tr><td>TEXTURE2</td><td>Surface texture class

Values: Clay, Gravel, Loam, Sand, n/a (refers to polygons with water and peat, where texture is not appropriate)</td></tr>
<tr><td></td><td></td></tr>
</table>

Appendix 3. Procedure for populating the terrain table from the GIS shapefile

The following steps have already been performed for each SOTER Danube country. The procedure below is presented here for information purposes.

Procedure:

Download the GIS shapefile and 'epmpty' SOTER database for your region/country (see Appendix 1) to your working directory.

- Open the MS-Access database (enable content if necessary), click on file in the upperleft of the screen, go to options, 'Current database', and change the *application title* to XX_Soter, where XX is your country's ISO code.

  [Note: Maintain the default file format 'Access 2002-2003' for consistency across all 14 databases. There is no backward compatibility once converted to e.g. MS Access 2006 (acdbb) format]

- Import XX_suid.dbf (as provided with the country shapefile) into your access database; this is done via tab 'external data' and subsequent browsing to file XX_suid.dbf in your own project folder, then click. This imports the data into table **terrain** in your database.

- Next append relevant (available) fields into the terrain table (i.e. populate the terrain table). For the shapefiles provided (see Appendix 2), these fields are *ISOC* and *SUID*. In view of the enforced referential integrity (see Appendix 4), the *TCID* (Terrain component number) must be also filled (i.e. is set to 1). The proportion (*PROP*) is set at 100 %.
  The corresponding query is:
  INSERT INTO TerrainComponent (ISOC, SUID, TCID, PROP) SELECT Terrain.ISOC, Terrain.SUID, 1 AS Expr1, 100 AS Expr2 ;

- Next, table **terraincomponent** must be populated. For SOTER Danube it is assumed that there is one terrain component in each SOTER unit.
  The corresponding query is:
  INSERT INTO TerrainComponent (ISOC, SUID, TCID, PROP)
  SELECT Terrain.ISOC, Terrain.SUID, 1 AS Expr1, 100 AS Expr2
  FROM Terrain;

- Next, table **soilcomponent** can be pre-populated. For this, it is assumed that that there is (at least) one soil component per terrain component in each SOTER unit. Note that the number of soil components (and their proportion, PROP) may be increased/modified when the table is being populated (see example in Figure 10).
  The corresponding query is:
  INSERT INTO SoilComponent (ISOC, SUID, TCID, SCID, PROP) SELECT TerrainComponent.ISOC, TerrainComponent.SUID, TerrainComponent.TCID, 1 AS Expr1, TerrainComponent.PROP FROM TerrainComponent;

  You now have the basis for compiling the SOTER attribute database for your allocated region, as further described in the body of this report.

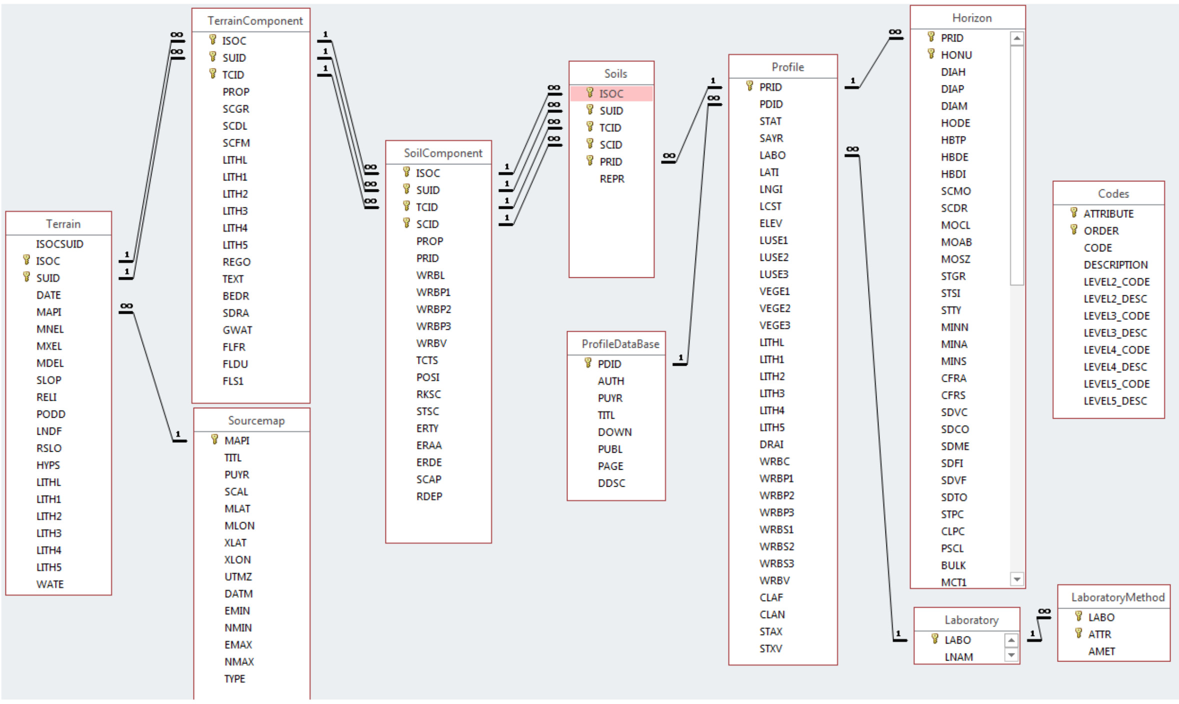
  An overview of main steps and associated requirements is provided in Table 1.

Appendix 4. Relationships enforced in SOTER forms

The attribute data for SOTER are managed in a relational database. This Appendix shows the relationships between the various tables (same as Figure 3, but repeated here convenience). These relations are automatically enforced when using the SOTER entry forms.

The allowed codes for the various properties are listed in table **codes** (Appendix 5).

Great care should be taken to exactly use the SOTER codes (even if they would contain a typing error) when importing data from auxiliary files (e.g. Excel or text files); otherwise the content of the data base may be corrupted or the data to be imported will not be accepted by the SOTER system!



Appendix 5. SOTER look-up table (codes)

There is one large look-up table (**codes**) that 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)

Three examples are given below:

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

ATTRIBL	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

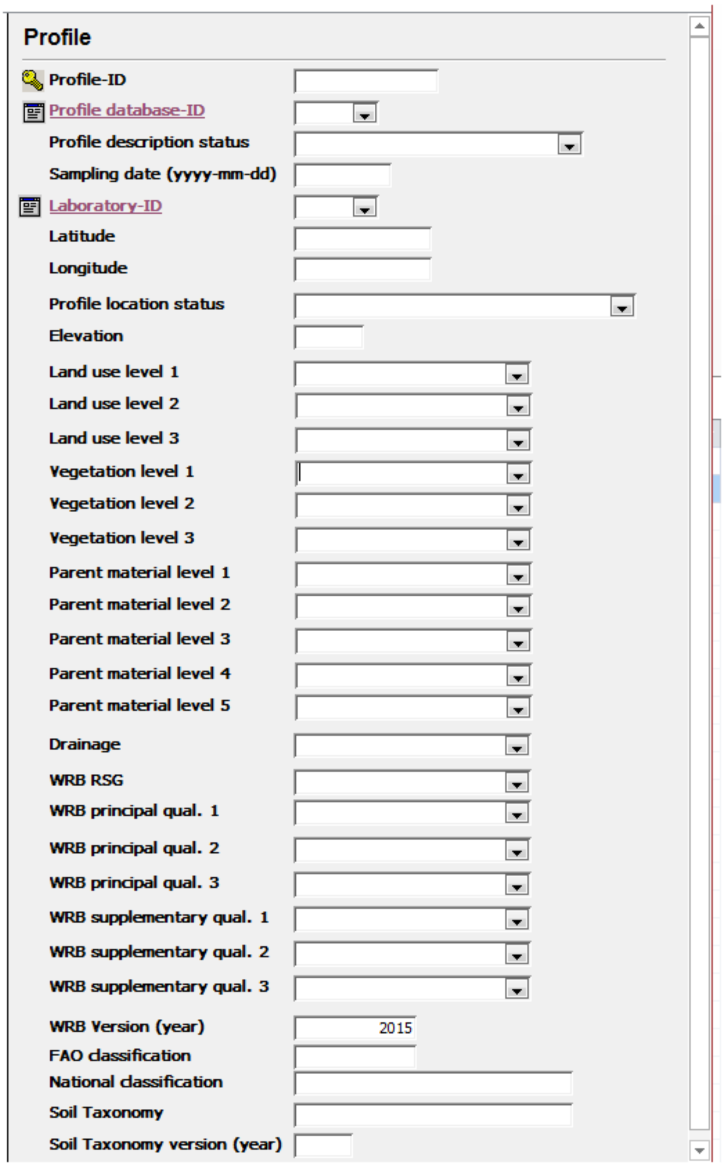
ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_CO	LEVEL2_DESC	LEVEL3_COD	LEVEL3_DESC	LEVEL4_COD	LEVEL4_DESC	LEVEL5_COD	LEVEL5_DESC
LITH	46	C	consolidated	CS	siliceous	CSU	ultrabasic (< 45% SiO2)	CSUM	metamorphic	CSUM1	meta-ultramafic rock
LITH	47	C	consolidated	CS	siliceous	CSU	ultrabasic (< 45% SiO2)	CSUT	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 roc
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 6. Example of procedure for filling the SOTER attribute tables

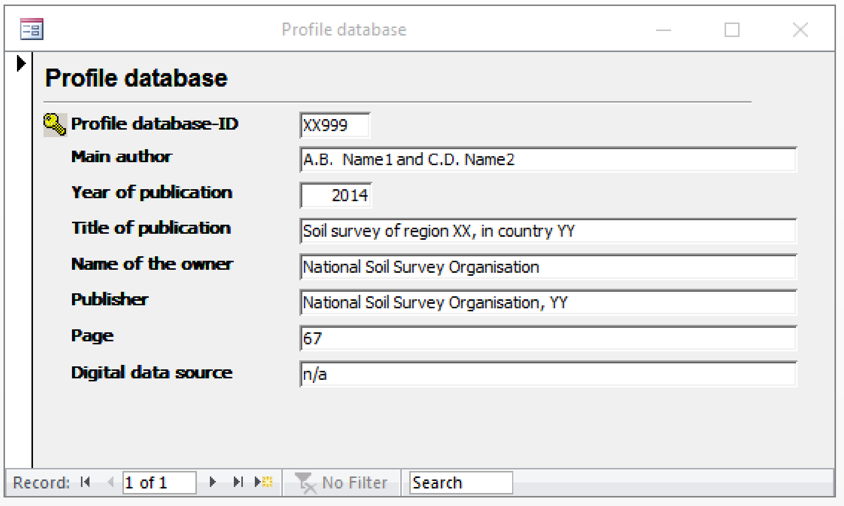
The overall procedure for filling SOTER attribute tables is illustrated below.

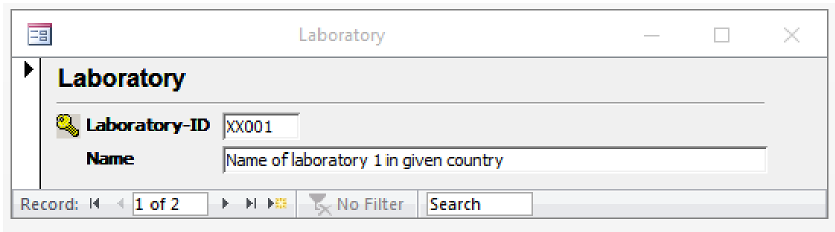
Let us take the **profile** table and open the corresponding form. At the beginning, this form is blank except for the field with WRB version(which is 2015 by default).

The two underlined items (Profile database-ID and Laboratory-ID) refer to the tables that serve to register the source of the data.



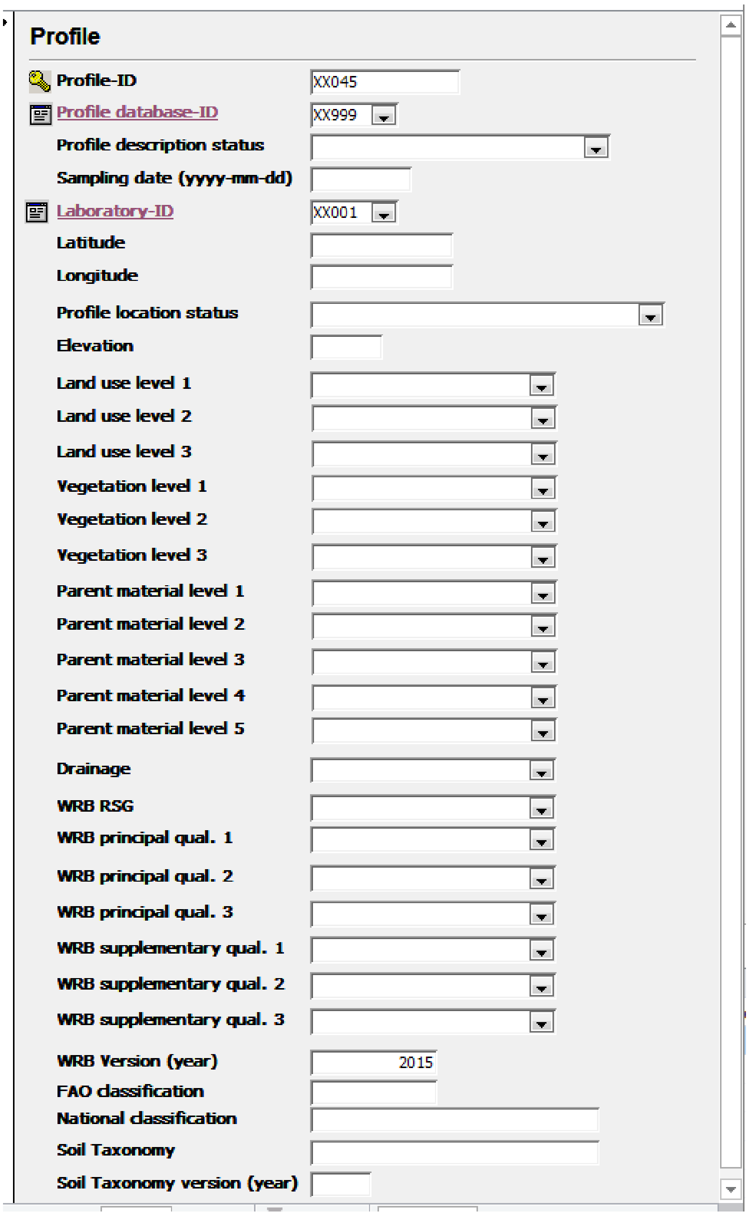
Clicking on these links will bring you to two empty forms (i.e. **profiledatabase** and **laboratory**). You can now enter the required data and save the information.



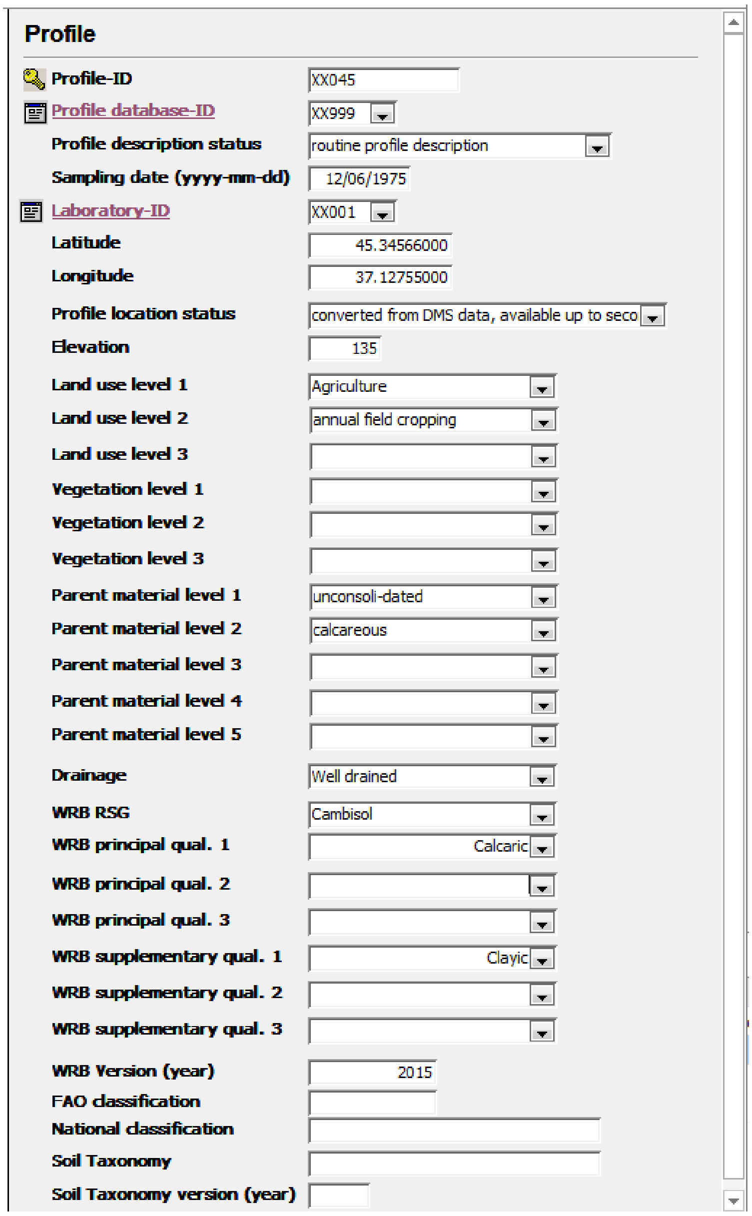


Once this has been done, the objects entered with the above forms can be selected from the pull down menu of the **profile** form.

[Tip: You may have to close the forms first, before the pull-down menus in the related table are populated with your most recent entries/additions]



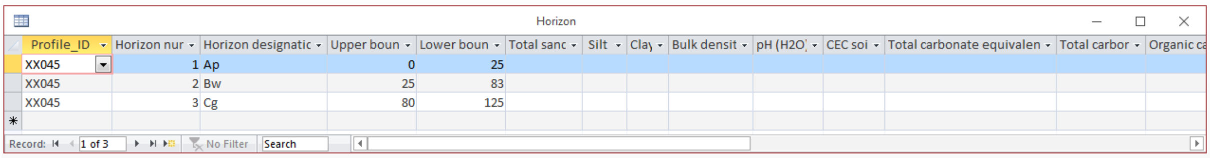
Now, the profile (site) data for say profile XX045 can be entered. We create an hypothetical example for this:



Profile

Profile-ID	XX045
Profile database-ID	XX999
Profile description status	routine profile description
Sampling date (yyyy-mm-dd)	12/06/1975
Laboratory-ID	XX001
Latitude	45.34566000
Longitude	37.12755000
Profile location status	converted from DMS data, available up to seco
Elevation	135
Land use level 1	Agriculture
Land use level 2	annual field cropping
Land use level 3	
Vegetation level 1	
Vegetation level 2	
Vegetation level 3	
Parent material level 1	unconsoli-dated
Parent material level 2	calcareous
Parent material level 3	
Parent material level 4	
Parent material level 5	
Drainage	Well drained
WRB RSG	Cambisol
WRB principal qual. 1	Calcaric
WRB principal qual. 2	
WRB principal qual. 3	
WRB supplementary qual. 1	Clayic
WRB supplementary qual. 2	
WRB supplementary qual. 3	
WRB Version (year)	2015
FAO classification	
National classification	
Soil Taxonomy	
Soil Taxonomy version (year)	

Now enter the associated horizon data. For this, open the **horizon** table. As indicated earlier, one may select only those fields for which data are available for the given region to facilitate data entry (keeping in mind the mandatory properties, see report Table 8). You may use the 'hide' and 'unhide' options for this (URL). Only the horizon number, horizon designation and upper and lower depths have been entered for the example.



Horizon

Profile_ID	Horizon nur	Horizon designatic	Upper boun	Lower boun	Total sanc	Silt	Clay	Bulk densit	pH (H2O	CEC soi	Total carbonate equivalen	Total carbor	Organic ca
XX045	1	Ap	0	25									
XX045	2	Bw	25	83									
XX045	3	Cg	80	125									

Record: 1 of 3 No Filter Search

The overall procedure for entering the other data is similar for the other forms/tables.

Appendix 7. Pre-harmonizing your data to SOTER conventions

This Appendix shows the general procedure for 'pre-harmonizing' your digital data to SOTER conventions, prior to their actual import into the SOTER database model (see Figure 3). It assumes an understanding of MS-Access or similar SQL procedures, and data handling.

The example is for drainage. A similar procedure can be used for the other attributes (*ATTR*) listed in table **codes.**

First, make a table with the SOTER codes for drainage called '**Look_up table_DRAInage**'. This can be done by extracting the corresponding records from the **codes** table. Such tables can be made in Excel or MS-Access as desired. Here, we shall use MS-Access 2003.

[Tip: See for example https://www.quackit.com/microsoft_access/tutorial/ for guidance when needed]

Enter the query (adapt as necessary for the attribute under consideration):

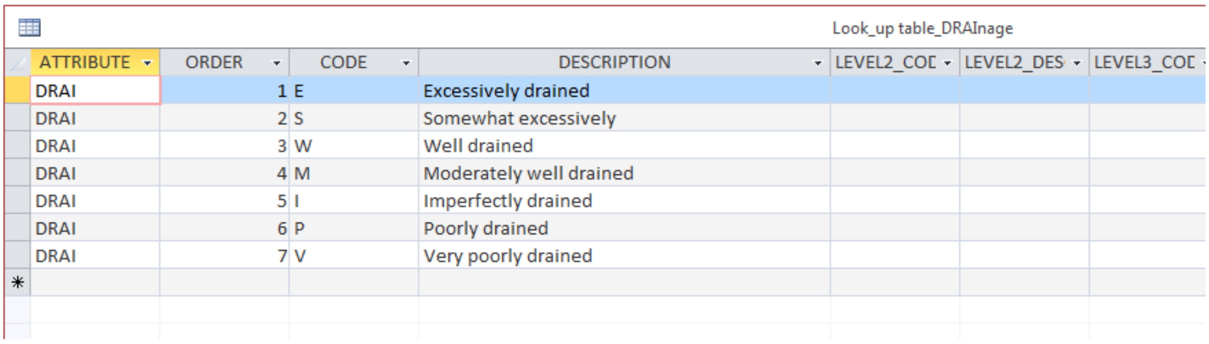
```
SELECT Codes.ATTRIBUTE, Codes.ORDER, Codes.CODE, Codes.DESCRIPTION, Codes.LEVEL2_CODE, Codes.LEVEL2_DESC, Codes.LEVEL3_CODE, Codes.LEVEL3_DESC, Codes.LEVEL4_CODE, Codes.LEVEL4_DESC, Codes.LEVEL5_CODE, Codes.LEVEL5_DESC INTO [Look_up table_DRAInage]

FROM Codes

WHERE (((Codes.ATTRIBUTE)="DRAI"));
```

Run the query.

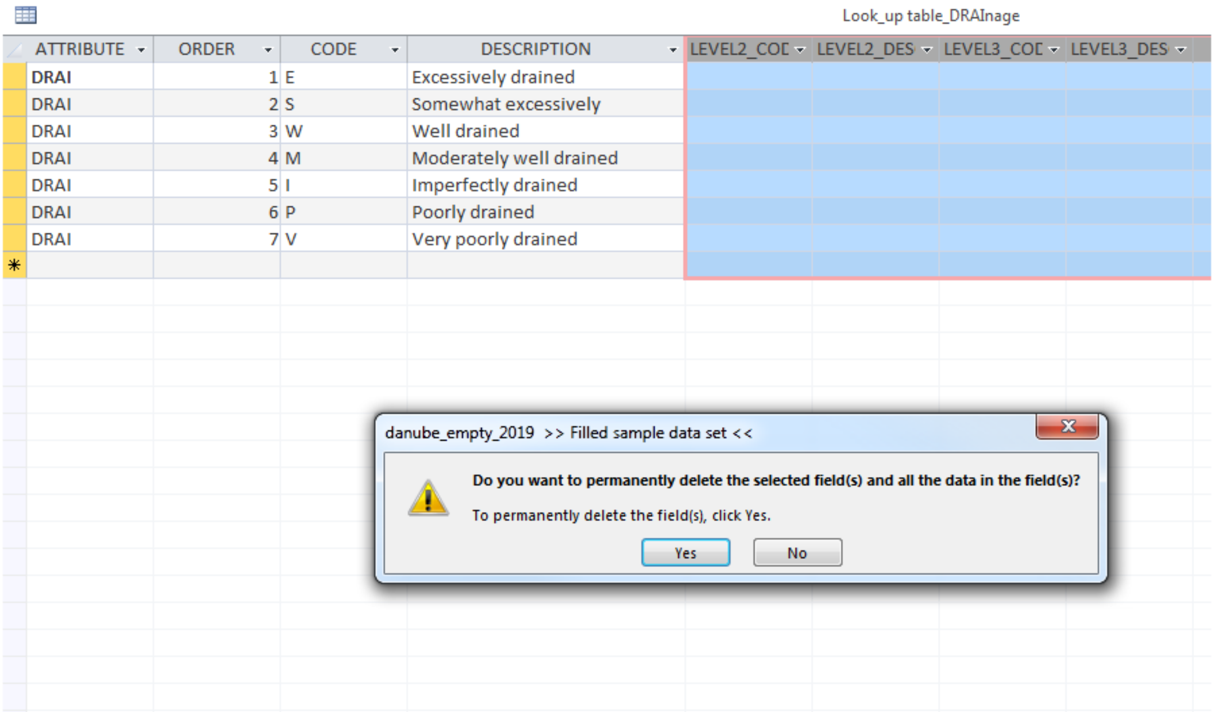
Open the new table: **Look_up table_DRAInage.**


Look_up table_DRAInage

ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_COD	LEVEL2_DES	LEVEL3_COD
DRAI	1	E	Excessively drained			
DRAI	2	S	Somewhat excessively			
DRAI	3	W	Well drained			
DRAI	4	M	Moderately well drained			
DRAI	5	I	Imperfectly drained			
DRAI	6	P	Poorly drained			
DRAI	7	V	Very poorly drained			

For soil drainage, there is only one level in the look-up table (i.e. columns *Level2_code*, *Level2_description* etc. are blank).

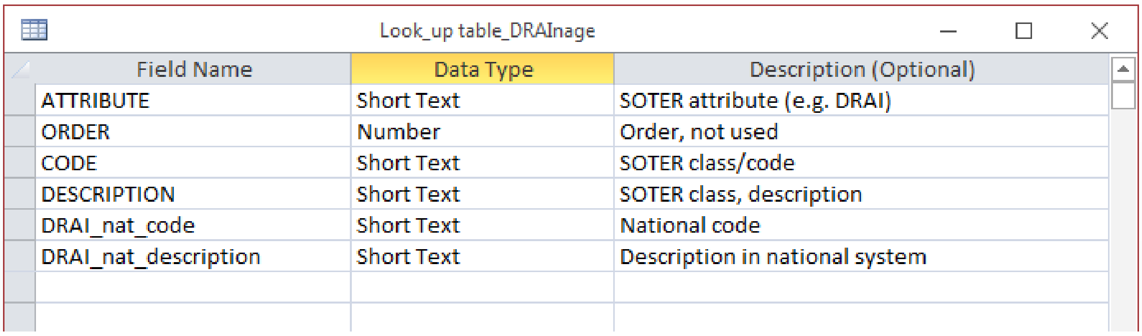
The blank columns are not needed for this particular attribute and thus may be deleted.



Look_up table_DRAInage

ATTRIBUTE	ORDER	CODE	DESCRIPTION	LEVEL2_COD	LEVEL2_DES	LEVEL3_COD	LEVEL3_DES
DRAI	1	E	Excessively drained				
DRAI	2	S	Somewhat excessively				
DRAI	3	W	Well drained				
DRAI	4	M	Moderately well drained				
DRAI	5	I	Imperfectly drained				
DRAI	6	P	Poorly drained				
DRAI	7	V	Very poorly drained				

Enter Yes. Next go to Design view, and enter new fields for the *nat*ional classification system for drainage. Pragmatically, for this example, they will be called *DRAI_nat_code* and *DRAI_nat_description*. It is good practice to fill-in the description field, for ease of reference.



Look_up table_DRAInage

Field Name	Data Type	Description (Optional)
ATTRIBUTE	Short Text	SOTER attribute (e.g. DRAI)
ORDER	Number	Order, not used
CODE	Short Text	SOTER class/code
DESCRIPTION	Short Text	SOTER class, description
DRAI_nat_code	Short Text	National code
DRAI_nat_description	Short Text	Description in national system

Go to Datasheet view and add the corresponding national classes (here, arbitrarily called 'National_description_1', 'National_description_2' etc.) . In some cases (e.g. for vegetation classes or lithology classes), these may be 'best matches'.

Look_up table_DRAInage

ATTRIBUTE	ORDER	CODE	DESCRIPTION	DRAI_nat_c(	DRAI_nat_description
DRAI	1	E	Excessively drained	1	National description-1
DRAI	2	S	Somewhat excessively	2	National description-2
DRAI	3	W	Well drained	3	National description-3
DRAI	4	M	Moderately well drained	4	National description-4
DRAI	5	I	Imperfectly drained	5	National description-5
DRAI	6	P	Poorly drained	6	National description-6
DRAI	7	V	Very poorly drained	7	National description-7

You now have a look-up table for drainage classes that shows the 'correlation' between the national classification system and the codes/terms used in SOTER, *in casu* for soil drainage.

Similar look-up tables table should be prepared for the other categorical attributes, when required. These can then be used to pre-harmonize your national data set to the SOTER conventions, prior to their actual import into the SOTER database for your region.

Also consider the units of measurement for numerical data; these must be converted to the SOTER defaults (see Table 9).

Remember, it is recommended to make regular backups of your work (with versioning, i.e. different names such as XX_SOTER_2019-07-03 etc.).

Further, it is good practice to document all 'pre-harmonization' steps in a report that accompanies the SOTER database for your region.

Appendix 8. Miscellaneous polygons and 'non-soil' units

Miscellaneous units in SOTER are areas of land that have a 'non-soil' cover, e.g. an ice mantle or a waterbody. These 'non-soil' areas were mapped as miscellaneous land units (FAO-UNESCO, 1974) and later recoded/harmonised for use with the Harmonised World Soil Database (FAO et al. 2008); these codes were adopted in the '2013 SOTER procedures manual' (Van Engelen et al., 2013).

Non-soil units and non-soil parts

Non-soil units correspond with regions where there is no soil mantle such as bare rock expanses, glaciers, shifting sand dunes, and urban areas. In principle, two situations occur:

a) the entire SOTER unit consist of 'non-soil' units.

b) only a portion of the SOTER unit consists of 'non-soil' units.

a) Differentiation at SOTER unit level

When the entire polygon consists of a non-soil unit (Table 13), a special entry is made in the GIS database for identification in the map legend; however, they are not treated as SOTER units. In the GIS attribute file these polygons are *normally* coded as follows:

ISO (country code) + '-' + (SOTER code), e.g. **CZ-WR** for a lake in Czechoslovakia (CZ).

Table 13. Codes for 'non-soil' units in the GIS file and attribute database

Symbol[a]	**Description**
WR	Lakes, permanent inland water bodies
GG	Glaciers, land ice, permanent snow fields
SP	Salt plains, salt flats
DS	Dunes, shifting sands
RK	Rock outcrops, crumbly rocks
UR	Urban, building areas
QU	Quarry, open air mining and other excavations
SW	Perennial swamps, marshes
SL	Salt lakes
BL	Badlands
FP	Fish ponds
ND	No data

[a] Symbols are adapted from Van Engelen et al. (2013, p. 91). The original code for salt plains (ST) was changed to 'SP' as 'ST' is now used for Stagnosols in WRB 2015; similarly, 'ND' is now used for no data/information (*note*: this applies only to non-soil units), as the former 'NI' code is used for Nitisols in WRB 2015.

In the case of SOTER Danube, however, the pre-populated ISOCSUID's in the GIS file and SOTER attribute table should not be altered as this information came from the underpinning JRC-generated terrain map. When necessary, however, non-soil areas can be indicated in the 'SURFCOND' column of the GIS file (e.g. AT_soter.shp for Austria).

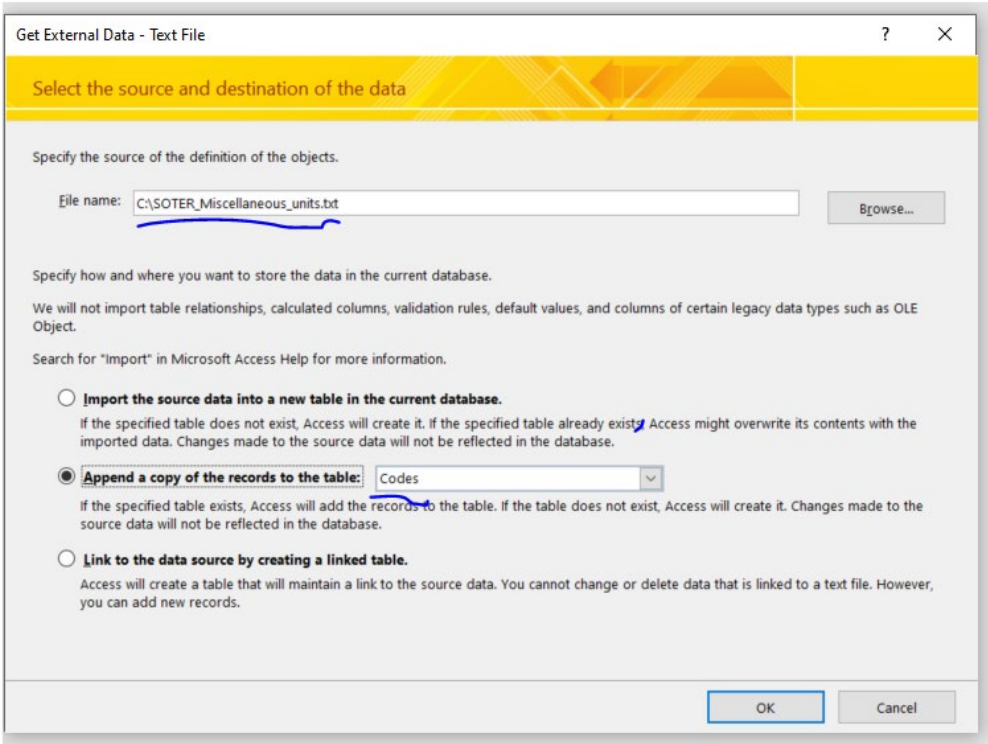
b) Differentiation at soil component level

When 'small' areas of lakes, land ice or rock outcrops cannot be mapped as a terrain unit, i.e. cover only part of the SOTER unit, they need to be characterised at the level of the soil component (Van Engelen et al 2013, p. 31).

The procedure for including non-soil units in a pre-populated SOTER[17] data base, for a given Danube basin country, is straight forward. First, append the standard definitions for non-soil units to the **codes** table. This information is provided as a text file (SOTER_Miscellaneous_units.txt).

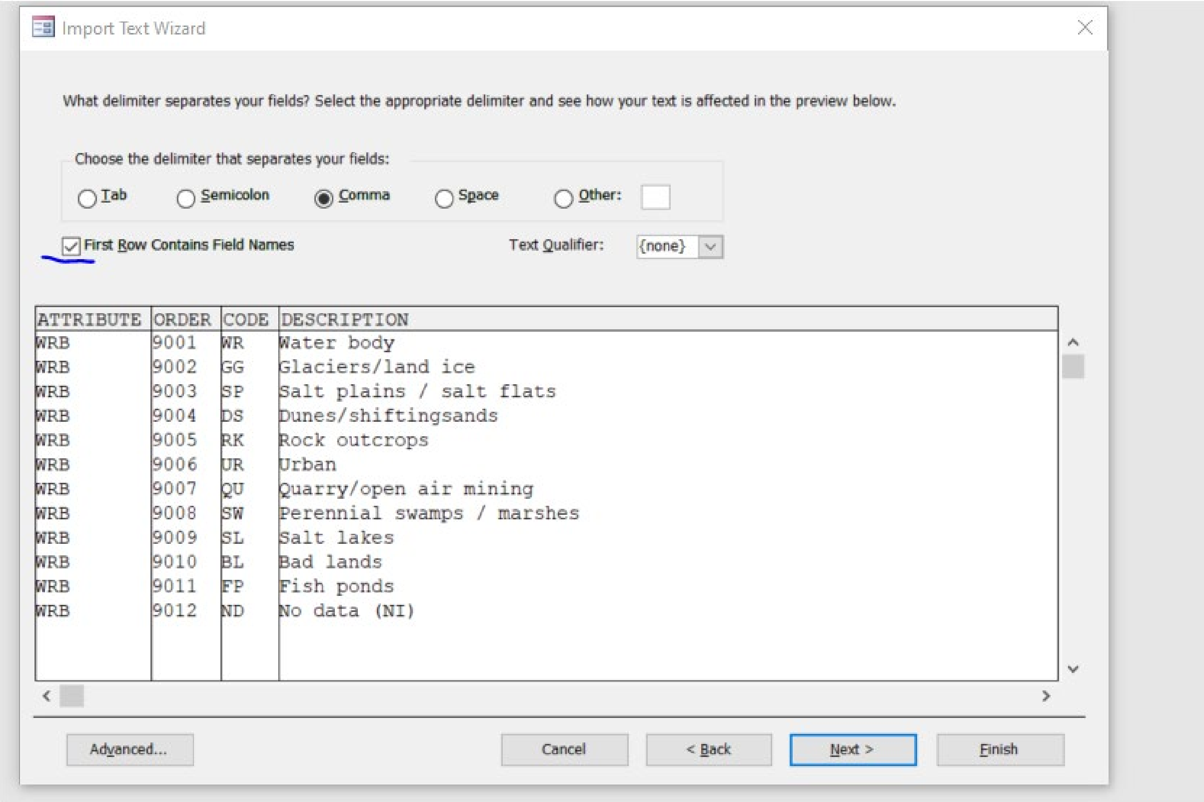
The TXT-file can be imported in your SOTER databases as follows (To be safe, make a backup of the whole database first):

In your MS Access database, go to button 'external data', then select/enter 'new data source', 'from file', select 'Text file', and then select file 'SOTER_Miscellaneous_units.txt', specifying the folder where you have saved the TXT file. Also indicate the table to which the TXT-file should be appended (underscored in blue):



Then enter next (delimited), next (comma; tick 'first row contains field names), next (codes), then enter 'finish'.

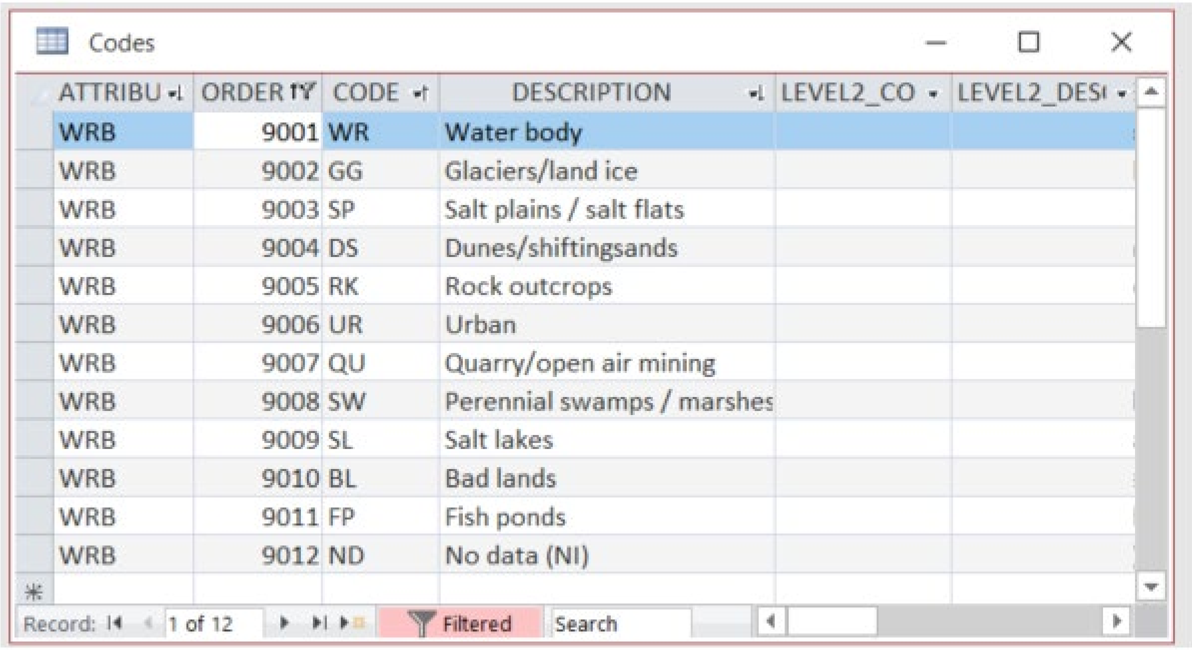
[17] Codes for 'non-soil' units are included in the **codes** table of the 'empty SOTER databases'**.**



ATTRIBUTE	ORDER	CODE	DESCRIPTION
WRB	9001	WR	Water body
WRB	9002	GG	Glaciers/land ice
WRB	9003	SP	Salt plains / salt flats
WRB	9004	DS	Dunes/shiftingsands
WRB	9005	RK	Rock outcrops
WRB	9006	UR	Urban
WRB	9007	QU	Quarry/open air mining
WRB	9008	SW	Perennial swamps / marshes
WRB	9009	SL	Salt lakes
WRB	9010	BL	Bad lands
WRB	9011	FP	Fish ponds
WRB	9012	ND	No data (NI)

*Note: This 'import' procedure should only be done once! Otherwise, you will get 'error messages' as integrity rules will be enforced (i.e. this would create duplicate entries in table **codes**).*

At this stage, the new non-soil records show up in table **code**:



ATTRIBU	ORDER	CODE	DESCRIPTION	LEVEL2_CO	LEVEL2_DES
WRB	9001	WR	Water body		
WRB	9002	GG	Glaciers/land ice		
WRB	9003	SP	Salt plains / salt flats		
WRB	9004	DS	Dunes/shiftingsands		
WRB	9005	RK	Rock outcrops		
WRB	9006	UR	Urban		
WRB	9007	QU	Quarry/open air mining		
WRB	9008	SW	Perennial swamps / marshes		
WRB	9009	SL	Salt lakes		
WRB	9010	BL	Bad lands		
WRB	9011	FP	Fish ponds		
WRB	9012	ND	No data (NI)		

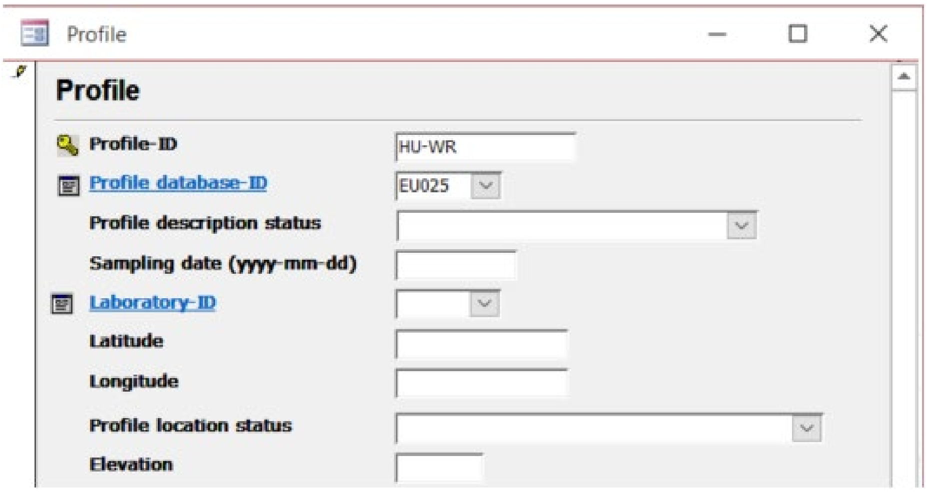
.

Subsequently, 'non-soil' units can be included (as 'artificial profiles') in the **profile** table for each Danube basin SOTER dataset. This can be done either using the **profile** *form* (a) or directly in the **profile** *table* (b) itself. In either case, only the (artificial) Profile-ID (PRID), for example 'HU-SP', and Profile_Database_id (PDID) must be entered. Being a 'non-soil' unit, all other fields can be left blank.

The example below is for 'water bodies' in Hungary. Note that 'EU025' is selected here for the profile-database-ID (PDID).

As indicated in the full report, unique PDID's must be defined in table **SourceMap** at an early stage of the data compilation.

(a) Enter 'non-soil' unit using **profile** *form*:



(b) Enter 'non-soil' unit directly in **profile** *table*:

Profile

Profile_ID	Soil Taxonon	Profile database_I	Profile descripti	Sampling dat
HU-WR		EU025		
CZ-ND		EU025		
CZ-FP		EU025		
CZ-BL		EU025		
CZ-SL		EU025		
CZ-SW		EU025		
CZ-QU		EU025		
CZ-UR		EU025		
CZ-RK		EU025		
CZ-DS		EU025		
CZ-ST		EU025		
CZ-GG		EU025		
CZ-WR		EU025		

Next, the relevant non-soil units can be entered in the **Soils** table and subsequently the **SoilComponent** table.

These steps are illustrated below for hypothetical SOTER unit 'CZ651' in the Czech Republic: 'CZ-WR'; the estimated proportion (PROP) of this non-soil unit (WR, water body) is 25% (Figure 17):

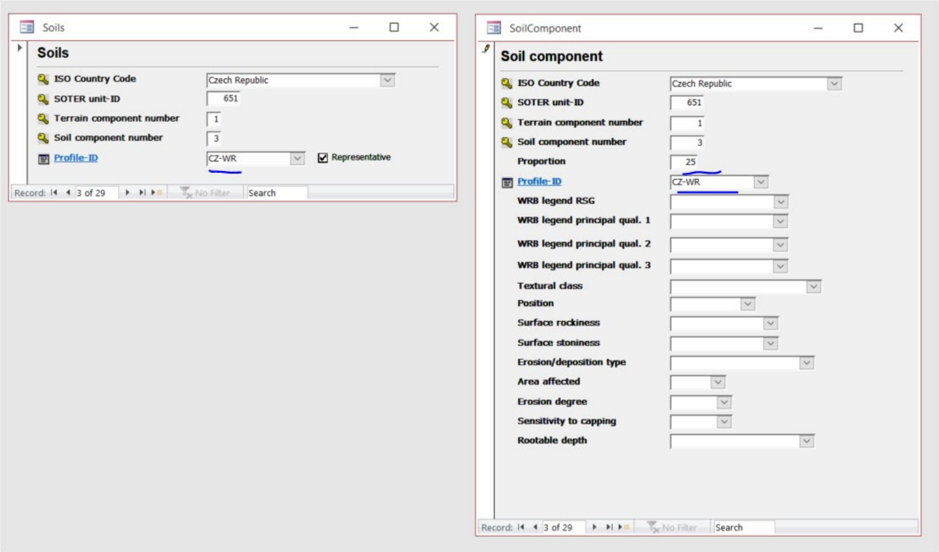


Figure 17. Example of a 'non-soil' unit as represented in a hypothetical SOTER unit.

For good order, note that the sum of proportions of all components (incl. soil and non-soil units) within a given terrain unit must always add up to 100% (Van Engelen et al. 2013).