

Research Article

Dániel Balla*, Tibor József Novák, Emőke Kiss, Tamás Mester, Gergő Karancsi, Marianna Zichar

Classification and geovisualization process of soil data using a web-based spatial information system

<https://doi.org/10.1515/opag-2020-0054>

received February 28, 2020; accepted August 04, 2020

Abstract: Because of the increased demand for information, the scientific community have published their results on the Internet over the last 20 years, adapting to society requirements. Due to the development of web-based spatial information soil systems, access to data on various themes and of varying quality has become substantially easier. The focus of our paper is to demonstrate a freely accessible and usable web-based soil database and soil information system (Soil Information and Soil Classifier System), which is suitable for geovisualization uploaded soil data and for determining reference soil groups (RSGs) in accordance with the World Reference Base of Soil Resources (WRB; RSG). In order to achieve this, we algorithmized the diagnostic soil classification process of the WRB, then we created decision trees to correspond input soil data to the WRB system. In order to facilitate geovisualization of the spatial data, the Keyhole Markup Language file format supported by Google API was applied.

Keywords: SISCS, soil classification system, visualization, WRB

1 Introduction

Soil classification concerns the grouping of soils with a similar range of properties (chemical, physical and biological) into genetic or diagnostic units (FAO 2006). It follows from this that it is an important element of soil research and data processing, serving mainly as an organizational framework for the description of soil properties (Shi et al. 2010; Gray et al. 2011; Pásztor et al. 2012). The primary reason for creating the World Reference Base of Soil Resources (WRB) is the international contribution to the development of the FAO. The first edition of the WRB was introduced at the 15th World Congress of Soil Science in Acapulco, Mexico in 1994 (Nachtergaele et al. 2000). At the Congress, researchers emphasized that by creating the WRB, the goal is not the development of a new soil classification system, but rather the harmonization of the different national systems (Michéli et al. 2006). The system is based on the diagnostic approach to soil classification with unambiguous definitions and quantitative parameters (IUSS Working Group (WG) WRB 2007–2015). National and international soil classification systems, however, are based on different principles. While national level (German, Russian, Hungarian, French and Chinese) classifications are mainly based on the Dokuchaev soil categorization, international systems (USDA, FAO and WRB) mostly utilize the diagnostic classification. It follows from the different systems that the integration or harmonization of certain national systems with the international systems is often difficult to achieve. Over the past two decades, however, numerous attempts have been made to harmonize the soil databases of different origins with the WRB (Bockheim and Gennadiyev 2000; Albrecht et al. 2004; Reintam and Köster 2006; Roca et al. 2008; Shi et al. 2010; Láng et al. 2013; Huyssteen et al. 2014) and to reclassify the (archived) national soil

* **Corresponding author: Dániel Balla**, Department of Computer Graphics and Image Processing, University of Debrecen, Debrecen, H-4028, Hungary, e-mail: balla.daniel@inf.unideb.hu

Tibor József Novák: Department of Landscape Protection and Environmental Geography, University of Debrecen, Debrecen, H-4032, Hungary, e-mail: novak.tibor@science.unideb.hu

Emőke Kiss: Department of Landscape Protection and Environmental Geography, University of Debrecen, Debrecen, H-4032, Hungary, e-mail: kiss.emoke@science.unideb.hu

Tamás Mester: Department of Landscape Protection and Environmental Geography, University of Debrecen, Debrecen, H-4032, Hungary, e-mail: mester.tamas@science.unideb.hu

Gergő Karancsi: Department of Computer Graphics and Image Processing, University of Debrecen, Debrecen, H-4028, Hungary, e-mail: karancsigeri@gmail.com

Marianna Zichar: Department of Computer Graphics and Image Processing, University of Debrecen, Debrecen, H-4028, Hungary, e-mail: zichar.marianna@inf.unideb.hu

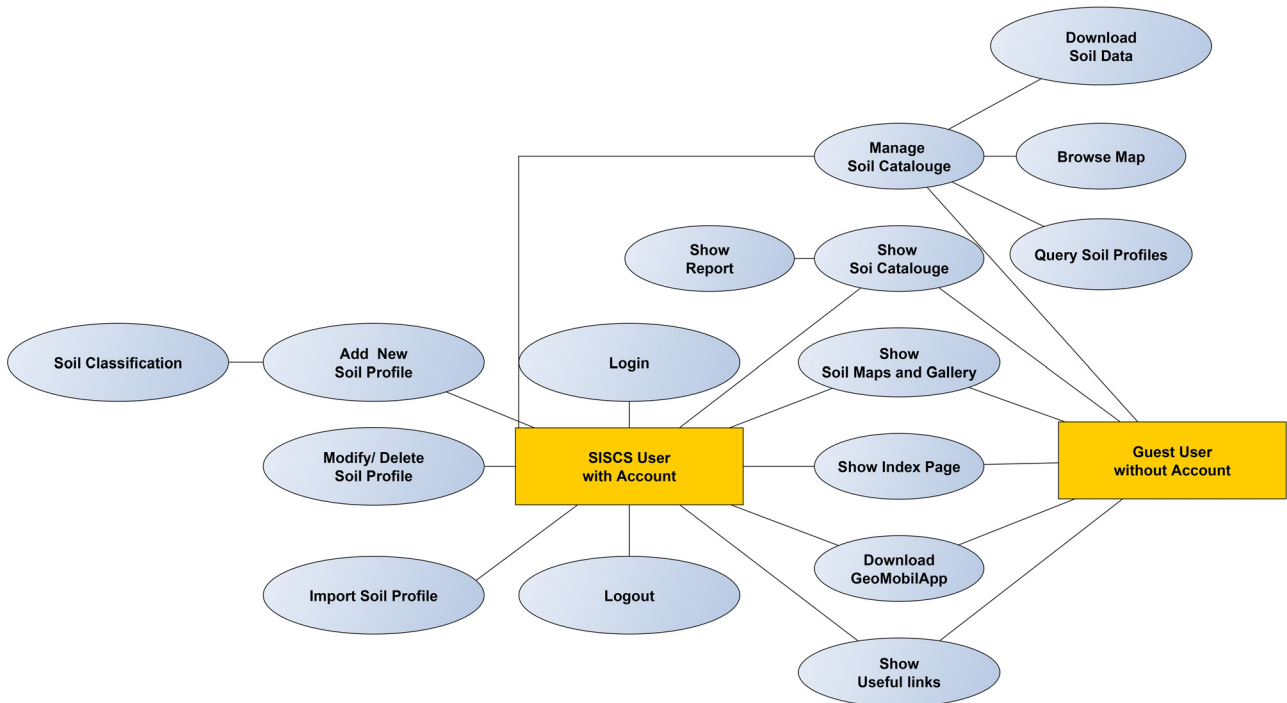


Figure 1: Use Case diagram of SISCS User and Guest.

databases/soil data using simplified algorithms according to the WRB (Krasilnikov et al. 2009; Eberhardt and Waltner 2010; Zádorová and Penížek 2011; Michéli et al. 2014; Michéli et al. 2019).

Eberhardt and Waltner (2010) suggested an automated methodology which was different from the previous, mainly correlation- and harmonization-based national level projects. The underlying principle behind the suggested

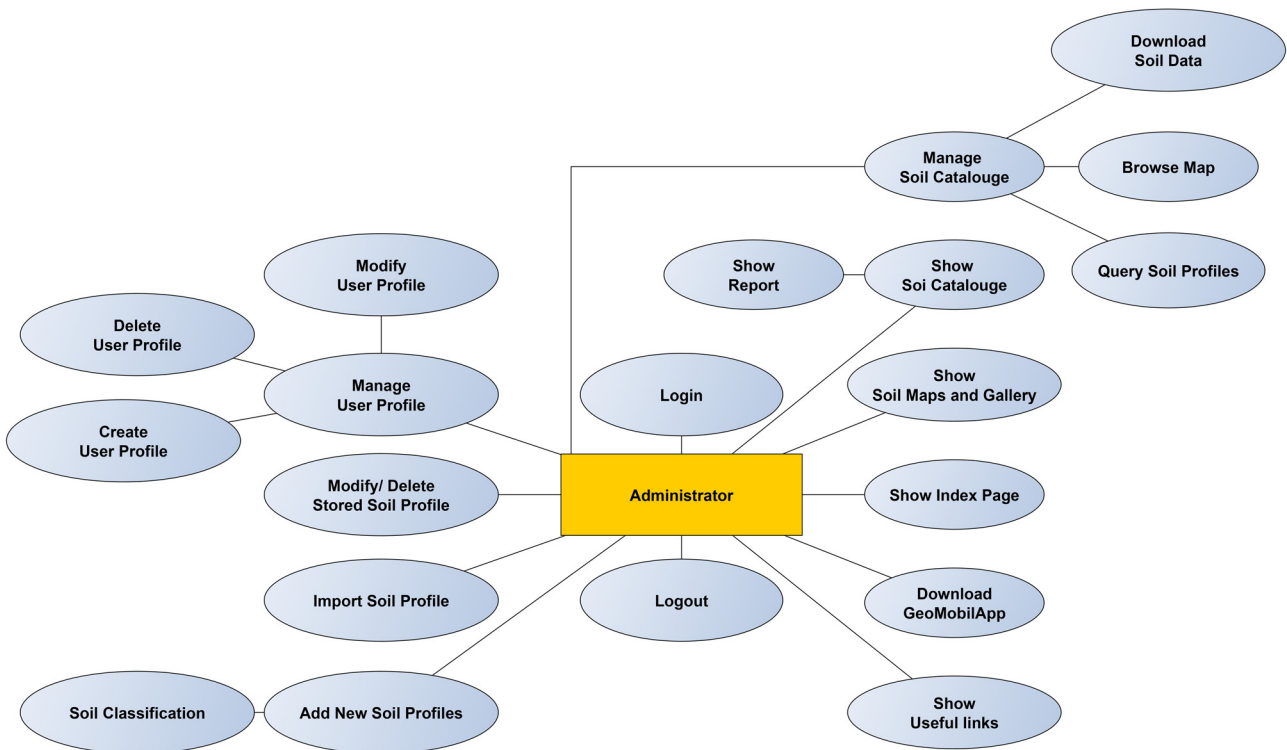


Figure 2: Use Case diagram of Administrator.

Table 2: Description of scripts of the Google API of the database table

Script	Task
mapInitializer.js	Loading of declared map type and KML interface at a given center and zoom scale
mapVisibilityFunctions.js	Modifying the visibility of the kml layer

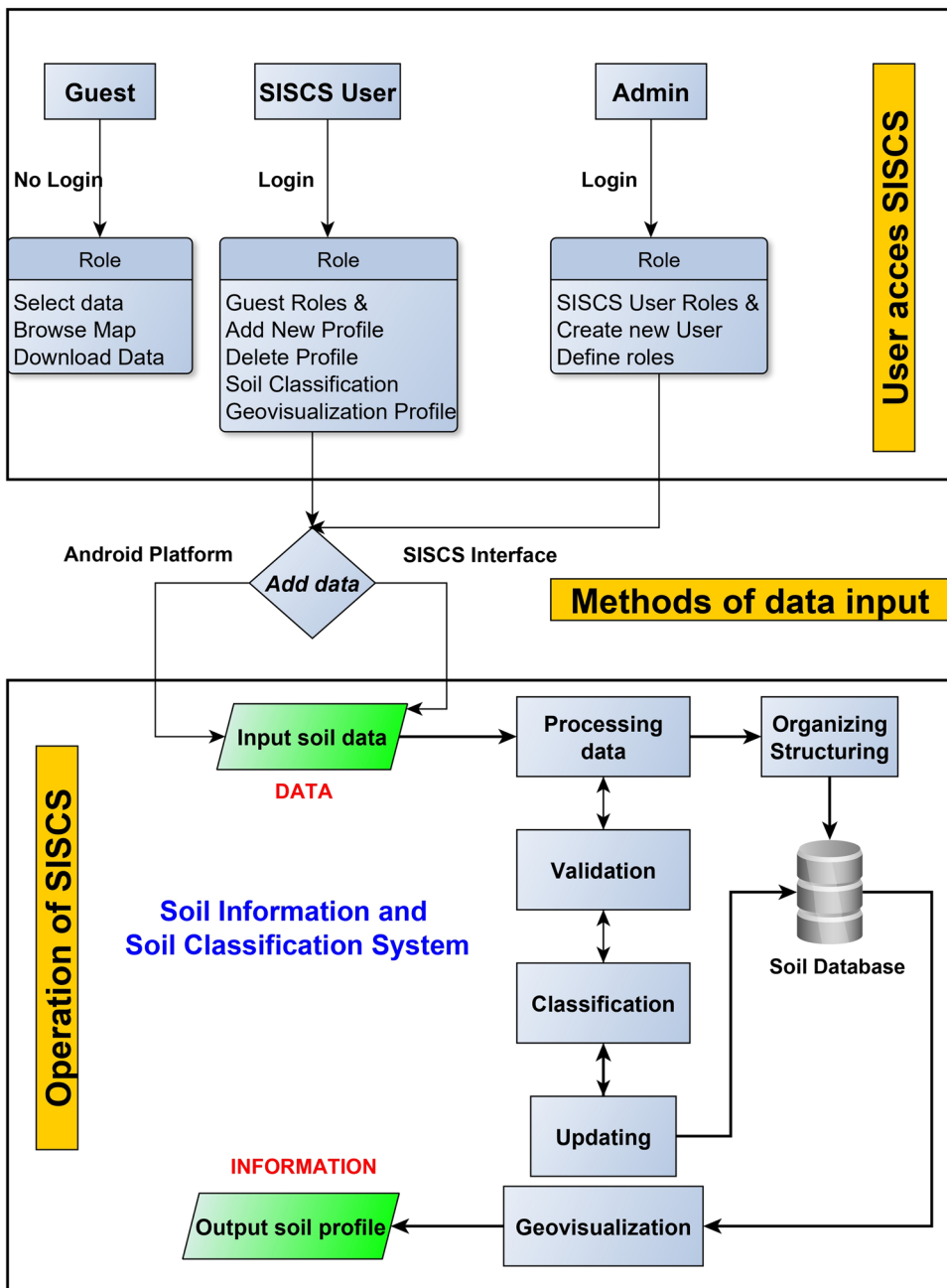


Figure 5: Structure and operation of SISCS (source: Balla et al. 2016 – figure expanded).

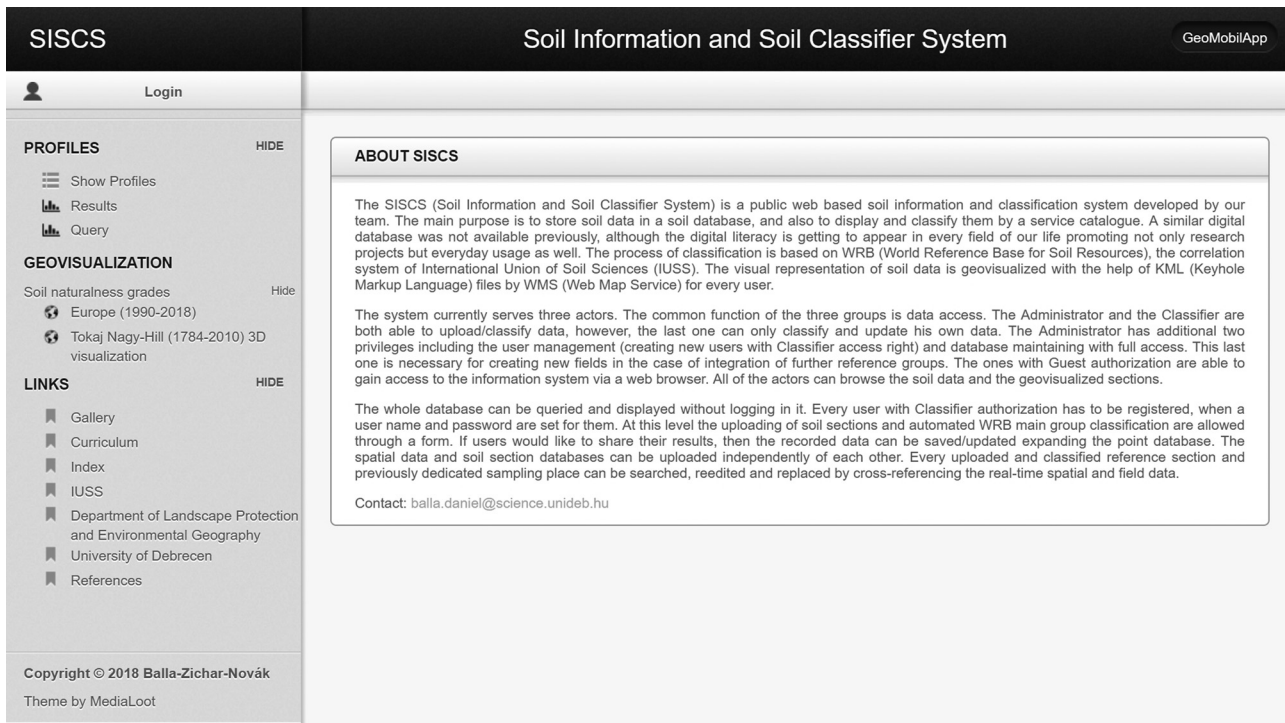


Figure 6: Index page of SISCS.

Our aim was to plan and develop a freely accessible and usable web-based soil database and soil information system (SISCS, Soil Information and Soil Classifier System), which is suitable for the geovisualization of uploaded soil data and for determining reference soil groups (RSGs) in accordance with the WRB (RSG). In order to achieve this, we algorithmize the diagnostic soil classification process of the WRB, then we create decision trees to correspond input soil data to the WRB system.

database is hosted by the MySQL Database system. The visual representation of the soil profiles is achieved via Keyhole Markup Language (KML). The applied technologies support free access to soil profiles, which can be shared among users and geovisualized on Google Maps. To satisfy this modern need for information, our goals also include the development of a classification application that uses chosen soil parameter lists. The classification is based on the correlation system of the International Union of Soil Sciences, the WRB (Balla *et al.* 2016).

2 Methods

2.1 System design and development

The main purpose of this research is to store soil data in a soil database, as well as to be able to display and classify these data by a service catalogue. It can be achieved in a reasonable way by establishing user groups that have different authorizations. Since our data have geospatial components, it is a natural requirement to geovisualize them with, for example, the help of a Google API. The application was created using PHP (PHP Hypertext Preprocessor) and JavaScript languages, as well as Hyper Text Markup Language. The user interface was designed with the free web template of Medialoot. The

2.2 Use case

The use case model demonstrates the user view of the system in which the purpose of the use case is to describe the relationships between the system to be modeled and its environment. This is a functional diagram, that is, the functions to be implemented by the system are placed in the center (Jacobson *et al.* 1995). We used the use case diagram to determine who will use the system and for what purposes. During the planning phase, the creation of three types of actors were considered: Guest, SISCS user and Administrator. Due to the standardized interface, Guests, SISCS users and Administrators have common functions, one of which has three sub-functions and another has one sub-function. SISCS users and Administrators also share

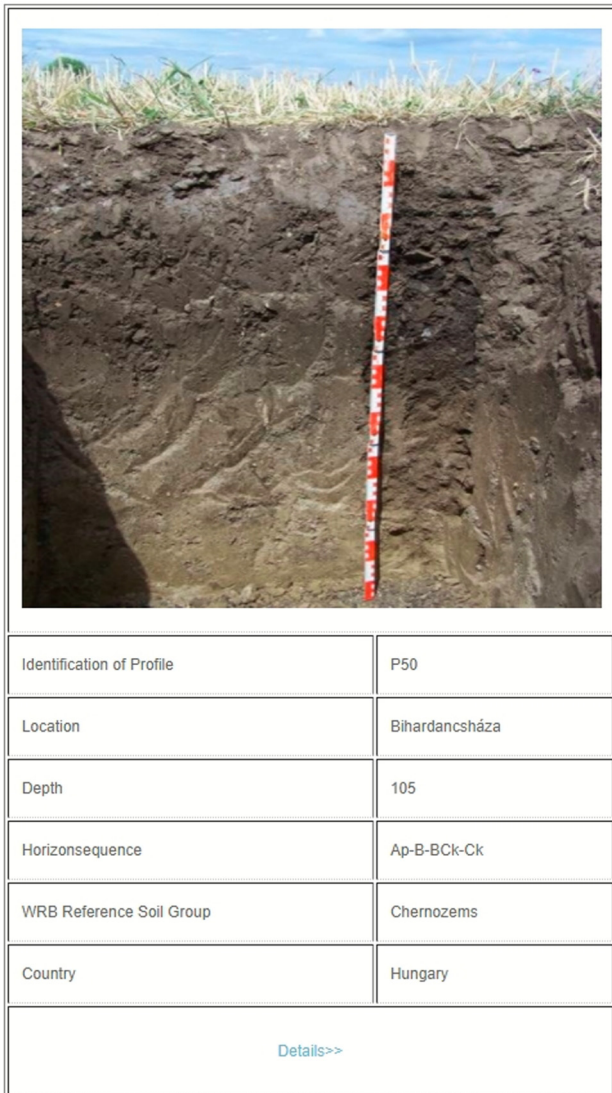


Figure 7: Description of listed soil profile (source: Balla et al. 2016).

functions. Furthermore, three more sub-functions were added for the Administrators because of the administrative tasks of user management (Figures 1 and 2).

2.3 Database scheme

To build the soil database, the variables and identification keys of the FAO have been used (FAO 2006), which contain all necessary information for WRB classification (IUSS WRB 2014). In addition, to describe a soil profile in a complex manner, qualitative and quantitative information about the soil are required, some of which comes from field analysis while the rest comes from laboratory testing (Figure 3 and Table 1). The databases created in

this way store the environmental descriptions of soil profiles, and the genetic and diagnostic information, as well as the physical and chemical attributes of the soil, separately (Balla et al. 2015a–2016).

2.4 Classification according to the WRB

The WRB is the modern soil classification system of diagnostics in Europe. The WRB (2015) is based on the previous classifications of the FAO while including many features of the USDA Soil Taxonomy and can sort soils from all over the world into 32 distinct soil reference groups. The classification is diagnostics based; therefore, it does not sort on the basis of the processes which have resulted in the creation of the soil, but rather on the layers and measurable attributes of the soil itself. These are the diagnostic horizons, diagnostic properties and diagnostic materials, which are clearly defined properties that can be measured or described; therefore, the classification can be transformed entirely into algorithms that are automatic and objective.

To define the RSG, the system uses a three-step hierarchy. At the first level of the hierarchy, the RSG is chosen by a (usually complex) system of conditions. The system analyzes the soil profile for the conditions of the RSGs starting from the first one (Histosol) and keeps going through them until the conditions of an RSG are met. If the soil profile does not meet any of the criteria for the RSGs, it gets sorted into the last one (Regosol). After the proper reference group is found, the classification continues to the next hierarchy level, where it looks for the qualifiers that can be attributed to the soil. The qualifiers that can be applied to certain soil groups are limited, as some of these qualities can be ruled out simply by the nature of some groups, while other qualifiers are unnecessary because they are a common trait in every soil of the RSG. The qualifiers that can be used in certain RSGs are located and listed in the charts that contain the definition of the reference groups. Every group has both principal and supplementary qualifiers. The former can be displayed in front of the name of the RSG with a hyphen, while the latter is displayed after the name of the RSG in brackets and separated by a comma.

2.5 Geovisualization of soil profiles

The soil profiles to be displayed on the map are represented by creating a KML file (Figure 4). The KML layer

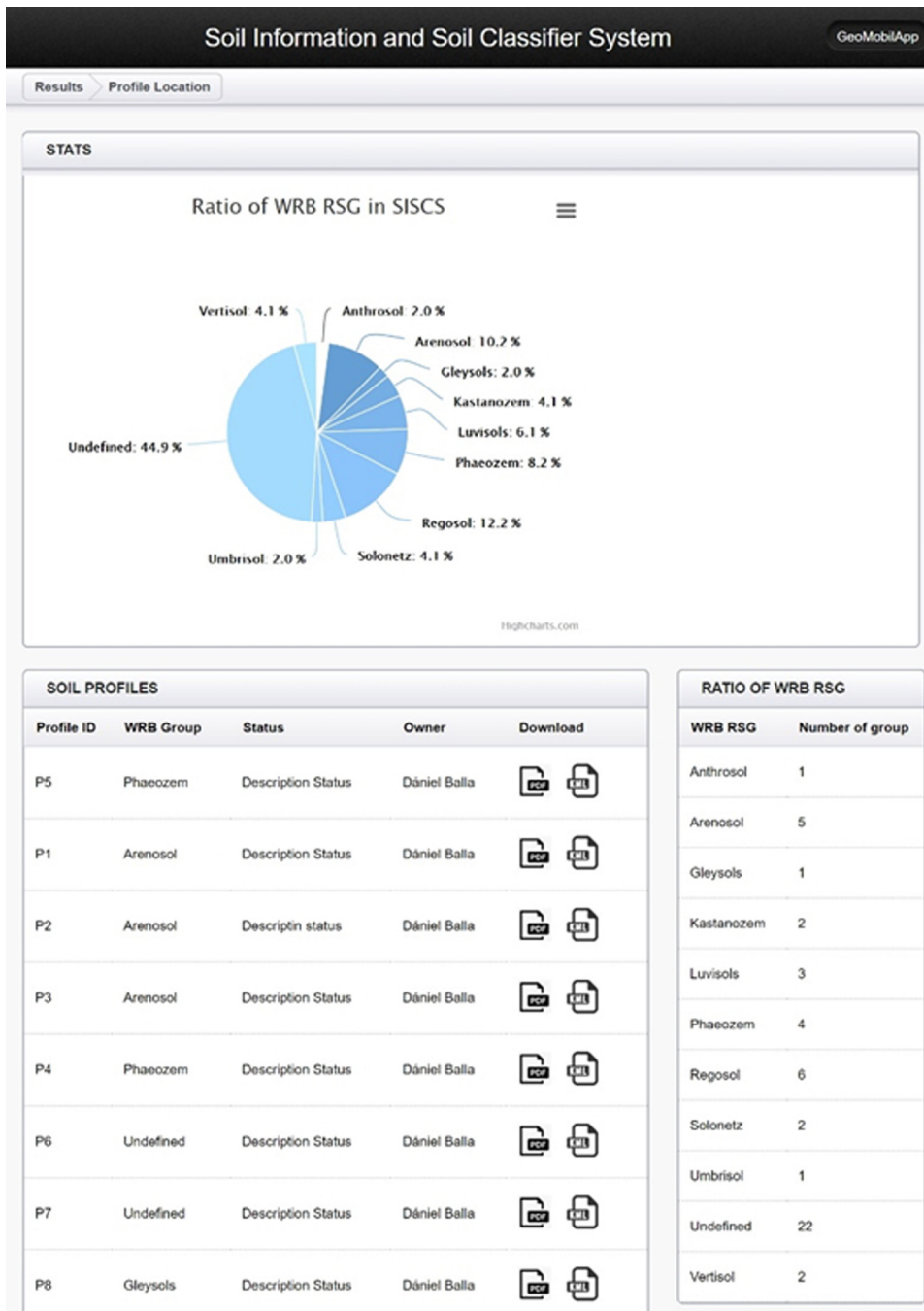


Figure 8: Interface of Result menu.

contains the data of the soil profiles stored in the database, which is expanded when a new soil profile is added. For each position indicator, an information bubble provides the geovisualization of further information available during the soil profiling.

As the last phase of the geovisualization of the soil profiles, a KML template created on the basis of the Google API is displayed in the Google Maps interface integrated into the system (Table 2).

3 Results

The web-based spatial soil information system (SISCS) has separate interfaces associated with the three different user accounts (Figure 5). The first level is the SISCS Guest account, which is assigned to users accessing the system through a web. SISCS Guest users can also list the soil profile data stored in the database and they can even download soil data. Google Maps APIs and their integrated

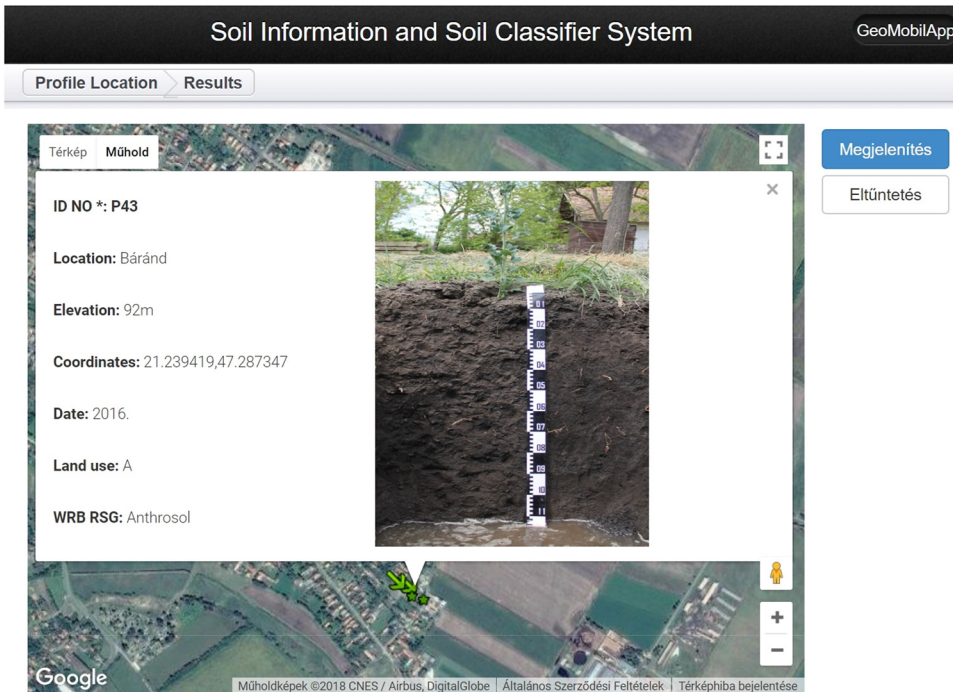


Figure 9: Geovisualized soil profile with the selected data from the database (source of photo of soil profile photo: Mester et al. 2017).

KML operating functions support that stored soil data and spatial soil databases are visualized in the optimum KML format. With the available KML template, the information uploaded from the soil database automatically creates a KML file associated with the soil profile and is automatically geovisualized with other soil profiles when loaded from the Google interactive web map.

The second level of the hierarchy is the SISCS User. The user is a member of the user group who can access the system with a username and a password. This level enables them to upload soil profile data with the use of an information form or to import soil data. For sharing, imported or uploaded soil data can be saved or updated, and the existing database can be expanded. The most valuable function is to determine the automated RSG on the basis of the data entered. The highest account privilege belongs to the Administrator group. SISCS Administrators can manage the entire soil database and create user profiles.

The SISCS is accessed via a web browser through the URL <http://siscs.exitdebrecen.hu/SISCSv2/>. First, a so-called index page is displayed on the interface, which includes a brief description of the system and the contact information. Authorized users can log into their own interface by clicking on Login (Figure 6).

When developing the user interface, user friendliness was a high priority, so we divided the interface into two separate parts. On the left side, we can see the menu

items created on the basis of functional requirements, while on the right side we can see the content, depending on the function selected. The functions of the system and accompanying menu system will be described by authorization types.

3.1 Functional interface of Guest

On this user level personal authentication is not required; therefore, anyone can access the entire database and download the soil data. Here users can browse soil profiles and data uploaded by SISCS users as well as geovisualized databases.

3.1.1 Profiles menu

In functional terms, the Profiles menu item can be divided into three parts as follows: Show Profiles, Results and Query functions.

3.1.2 Show Profiles menu

By accessing Show Profiles, the soil profiles of the database will be listed dynamically (Figure 7). Users can gain

SEARCH FOR PROFILE

Data for profile identification

Location: <input type="text" value="Location"/>	Name of settlement: <input type="text" value="Name of settlement"/>	ID No *: <input type="text" value="Identification of Profi"/>
Closer name of area: <input type="text" value="Closer name of area"/>	WGS84 λ °: <input type="text" value="latitude"/>	WGS84 φ °: <input type="text" value="longtitude"/>
Depth of profile: <input type="text" value="cm"/>	Country: <input type="text" value="Country"/>	

Landform and topography

Geomorphology: <input type="text" value="Choose one..."/>	Slope shape: <input type="text" value="Choose one..."/>	Slope angle: <input style="width: 50px;" type="text" value="%"/>
Exposition: <input type="text" value="Choose one..."/>	Elevation: <input type="text" value="m"/>	

Date

Year: <input type="text" value="Choose one..."/>	Month: <input type="text" value="Choose one..."/>	Day: <input type="text" value="Choose one..."/>
--	---	---

Soil climate

Type of soil climate: <input type="text" value="Choose one..."/>	Type of soil moisture regime: <input type="text" value="Choose one..."/>	Groundwater level: <input type="text" value="cm"/>
--	--	--

Land use and vegetation

Natural vegetation: <input type="text" value="Choose one..."/>	Land use: <input type="text" value="Choose one..."/>	Anthropogenic influences: <input type="text" value="Choose one..."/>
Cultivated plants: <input type="text" value="Cultivated plants"/>		

Parent material

Parent material:

Weather conditions

Weather conditions (current): <input type="text" value="Choose one..."/>	Weather conditions (past): <input type="text" value="Choose one..."/>
--	---

Soil group

WRB Reference Soil Group:

User

Username:

SOIL PROFILES

Profile ID	WRB Group	Status	Owner	Download
P43	Anthrosol	Description Status	Dániel Balla	

Figure 10: Result of query interface after selection.

information on profile identifiers, their geographical locations, the depth of the exposed soil profile, its horizon sequence and reference group. By clicking on the Details link, the digital report regarding the given soil profile can be accessed, which enables users to review the properties of complex profile description (field and laboratory data and recorded diagnostics). Here data cannot be edited.

3.1.3 Results menu

The Results item offers a visual statistical overview on the percentage distribution of the currently stored soil profiles by RSGs, and their occurrence listed in numerical tables.

If the contents of the tables change, for example, due to a deletion of an update, the interface refreshes automatically. The third part of the interface is a dynamically listed table, which includes the Profile ID, reference group, description status and the name of the person who uploaded the soil profile data stored in the database. The soil data regarding soil profiles can be downloaded individually in pdf and kml formats (Figure 8).

Via the Profile Location link users can access the web-based visualization of soil data (Figure 9). Based on the selected data of the uploaded soil records, a KML management function responsible for the geovisualization of the soil profiles automatically displays a kml file on the web map, the visibility of which can be modified by the user.

Table 3: List of queryable parameters

Parameters
Location, Name of Settlement, Id No, Closer name of area, WGS84 λ *, WGS84 φ *, Depth of profile, Country
Geomorphology, Slope shape, Slope angel, Exposition, Elevation
Year, Month, Day
Type of soil climate, Type of soil moisture regime, Groundwater level
Natural vegetation, Land use, Anthropogenic influences, Cultivated plants
Parent material
Weather conditions
WRB RSG

By clicking on the symbol on the map indicating the location of the soil profile, the information interface of the selected soil profiles will appear.

3.1.4 Query menu

The Query item is a query interface on which the soil profiles stored in the database can be listed by setting the search/screening parameters (Figure 10). This is the most interactive part of the Guest interface, since here users themselves can decide which parameters should be applied for screening and displaying data (Table 3).

3.1.5 Links menu

In the Links item any user can open the following links: SISCS index page, IUSS (International Union of Soil Sciences) website, the website of the University of Debrecen, References and the Soil Geography seminar e-learning education material for students enrolled in the Geography BSc program.

Via the Gallery link users can open the profile photos uploaded by registered users (Figure 11).

3.2 Functional interface of SISCS User

The SISCS user is a member of the user group who can access the system with a username and a password. Their most important function is to upload, save and edit soil profile data exposed and recorded during the field sampling on their own interfaces, as well as to determine the automated RSG on the basis of the data entered. Following login, three further subitems appear in the Profiles menu

(Add New Profiles, Import Profiles and Edit Profiles), which also indicates the functions associated with this user level.

3.2.1 Add new Profiles menu

On this interface, the SISCS user can upload the data recorded during the soil profile exposure in a structured format. The interface created can be divided into four larger units. In the Data of Profiles block users should enter data from the soil profile environment description and upload a photo (Figure 12).

The Information Table block helps the filling out of forms; users can access a drop-down list which provides the information required for the interpretation of the selected soil parameters to help screen incorrect parameters during data entry (Figure 13).

The Report and Diagnostic (Diagnostic Horizons, Diagnostic Properties and Diagnostic Materials) blocks are suitable for entering and describing genetic and

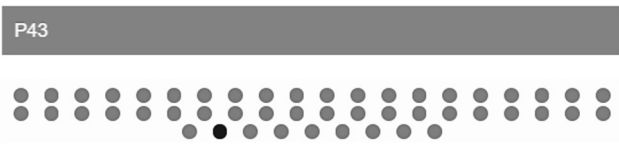


Figure 11: Interface of Gallery menu.

DATA OF PROFILE

Data for profile identification

Location: Name of settlement: ID No *:

Closer name of area: WGS84 λ °: WGS84 φ °:

Depth of profile: Country:

Landform and topography

Geomorphology: Slope shape: Slope angle:

Exposition: Elevation:

Date

Year: Month: Day:

Soil climate

Type of soil climate: Type of soil moisture regime: Groundwater level:

Land use and vegetation

Natural vegetation: Land use: Anthropogenic influences:

Cultivated plants:

Parent material

Parent material:

Weather conditions

Weather conditions (current): Weather conditions (past):

Upload Photo for Profile Nincs fájl kiválasztva

Figure 12: Initializations of soil description.

diagnostic layers determined during field work. With these blocks, the identifier and the depth corresponding to the given profile are critical because the analysis of the recorded samples is also associated with the record with the given ID. The Add/Remove Layer/Horizons/Material/Properties buttons can be used to create or add new

records or delete records from the horizon sequences of the profile (Figure 14).

In the last block of the data entry interface, users can add their own notes regarding profile exposure, which can be useful during later follow-up tasks. Following the filling-out of the record, we can save the soil profile

INFORMATION TABLE

Structure: size

Mark	Name	Granular/platy (mm)	Prismatic/columnar/wedge-shaped (mm)	Blocky/crumbly/lumpy/cloddy (mm)
VF	Very fine/thin	<1	<10	<5
FI	Fine/thin	1-2	10-20	5-10
ME	Medium	2-5	20-50	10-20
CO	Coarse/thick	5-10	50-100	20-50
VC	Very coarse/thick	>10	100-500	>50
EC	Extremely coarse	-	>500	-

Figure 13: Identification keys of soil parameters.

The screenshot displays four stacked form panels for initializing soil profile data:

- REPORT:** Includes tabs for 'Horizons' (selected), 'Color: when dry', 'Color: when moist', 'Primary constituents', 'Structure', 'Cementation/Concretions', 'Biological features', and 'Anthropogenic materials'. Below are sections for 'Basic laboratory tests' and 'Special laboratory tests'. Fields include 'Sample ID' (checkbox), 'ID of sample', 'Start of depth' (cm), 'End of depth' (cm), 'Genetic horizon', 'Distinctness' (Choose one...), and 'Topography' (Choose one...). Buttons: 'Reset', 'Add Layer', 'Remove Layer'.
- DIAGNOSTIC HORIZONS:** Fields include 'Sample ID' (checkbox), 'ID of sample', 'Start of depth' (cm), 'End of depth' (cm), and diagnostic categories: 'Antraquic', 'Argic', 'Calcic', 'Cambic', 'Chernic', 'Cryic', 'Duric' (all set to 'False'). Buttons: 'Add Horizon', 'Remove Horizon'. Page navigation: 'Page 1', 'Page 2', 'Page 3'.
- DIAGNOSTIC PROPERTIES:** Fields include 'Sample ID' (checkbox), 'ID of sample', 'Start of depth' (cm), 'End of depth' (cm), and properties: 'Abrupt textural difference', 'Albeluvic glossae', 'Andic properties', 'Anthric properties', 'Aridic properties', 'Continuous rock', 'Geric properties' (all set to 'False'). Buttons: 'Add Property', 'Remove Property'. Page navigation: 'Page 1', 'Page 2'.
- DIAGNOSTIC MATERIALS:** Fields include 'Sample ID' (checkbox), 'ID of sample', 'Start of depth' (cm), 'End of depth' (cm), and materials: 'Albic material', 'Artefacts', 'Calcaric material', 'Colluvic material', 'Dolomitic material', 'Fluvic material', 'Gypsic material' (all set to 'False'). Buttons: 'Add Material', 'Remove Material'. Page navigation: 'Page 1', 'Page 2'.

Figure 14: Initializations of layers, genetic and diagnostic horizons.

by clicking on the Save button and classify the soil profile based on the data entered according to the WRB by clicking on the Classifier button. The RSG of the soil profile can be determined manually, or it can be classified manually according to the Hungarian classification system if needed (Figure 15). It should be noted that the quality management phase of the implemented classification algorithm ensures that the input data are correct, and in the case of incorrect or inappropriately entered data it notifies the user about the incorrect parameters.

3.2.2 Edit Profiles menu

The Edit Profiles item shows the list of profiles uploaded by a logged-in user (Figure 16). On this interface users can delete or update profiles by clicking on the action buttons on the left side. By clicking on the delete icon the selected soil profiles can be deleted from the database and consequently the dynamically presented interfaces will also be updated (e.g. Results, Profile Location). By

clicking on the Edit icon users can begin to correct, update and save the data of the selected soil profile on the data entry form described above.

3.2.3 Import Profiles menu

The Import Profiles item can be used to import csv files exported from the GeoMobilApp application. Here the environmental and diagnostic data of the profile should be uploaded together with the profile photo (Figure 17).

3.2.4 Operation of GeoMobilApp

After the initialization of the application, users can decide whether they want to review previous soil profiles on their mobile phones or want to define a new sampling profile. In the first case, the database management algorithm will list the previously stored profiles during the startup phase. Here these data can be edited

REMARK

Note

SOIL GROUP

Principal qualifiers

WRB Reference Soil Group

Undefined ▾

Hungarian Soil Classification System

Unknown ▾

Supplementary qualifiers

Results

WRB RSG

Description Status ▾

Figure 15: Interface for remark and soil classification.

SOIL PROFILES						
Profile ID	Latitude	Longitude	Depth of profile (cm)	WRB RSG	Status	Actions
P43	21.239419	47.287347	160	Anthrosol	Description Status	
P1	22.113333	47.698056	180	Arenosols	Description Status	
P2	22.113056	47.698611	140	Arenosols	Descriptin status	
P3	22.114444	47.701389	140	Arenosols	Description Status	
P44	21.677542	47.457261	120	Arenosols	Description Status	

Figure 16: Interface of edit profiles menu.

Figure 17: Interface of import menu.

as well (modification or deletion). In the case of a new sampling, the soil data corresponding to the profile can be determined through a process involving multiple screens (steps). At the end of the process, a picture can be taken of the soil profile using the camera of the device. At the end of every determining process not only is the position of the profile defined automatically but the most important information is also saved in a kml file which can be reviewed later.

3.3 Functional interface of the Administrator

In addition to the previously described functions, on the Administrator interface one more function can be accessed. Users with Administrator authorization can manage the entire soil database and can create user profiles and assign to them either SISCS user or Administrator authorization. Furthermore, Administrators are able to modify, update or delete the data of the soil profile uploaded by any SISCS user.

3.3.1 Add New User menu

The form provided for the creation of user profiles is shown in Figure 18. Following the entry of administrative data, by clicking on the Save button an account is created through which the user can access his/her own interface (depending on their authorization).

Figure 18: Interface of the Add New User menu.

3.3.2 Edit User menu

On this interface, users can delete or update user accounts by clicking on the action buttons on the left side. By clicking on the delete icon, the user will be deleted from the database and therefore will not be able to log in to the system. By clicking on the Edit icon, Administrators can begin to correct, update the authorization level and save the data of the selected account on the data entry form shown in Figure 19.

New User		User List		USERS		
Username	First Name	Last Name	Department	e-mail	Actions	
athos04	Dániel	Balla	DE-TTK	athos04@freemail.hu		
test	X	Y	xyz	aaa@gmail.com		
user99	CCC	BBB	home	home@gmail.com		
debug	Valaki	Valami	Egyetem	ize@freemail.hu		
tomi	Tamás	Mester	DE-TTK	mester.tamas@science.unideb.hu		

Figure 19: Interface of the Edit User menu.

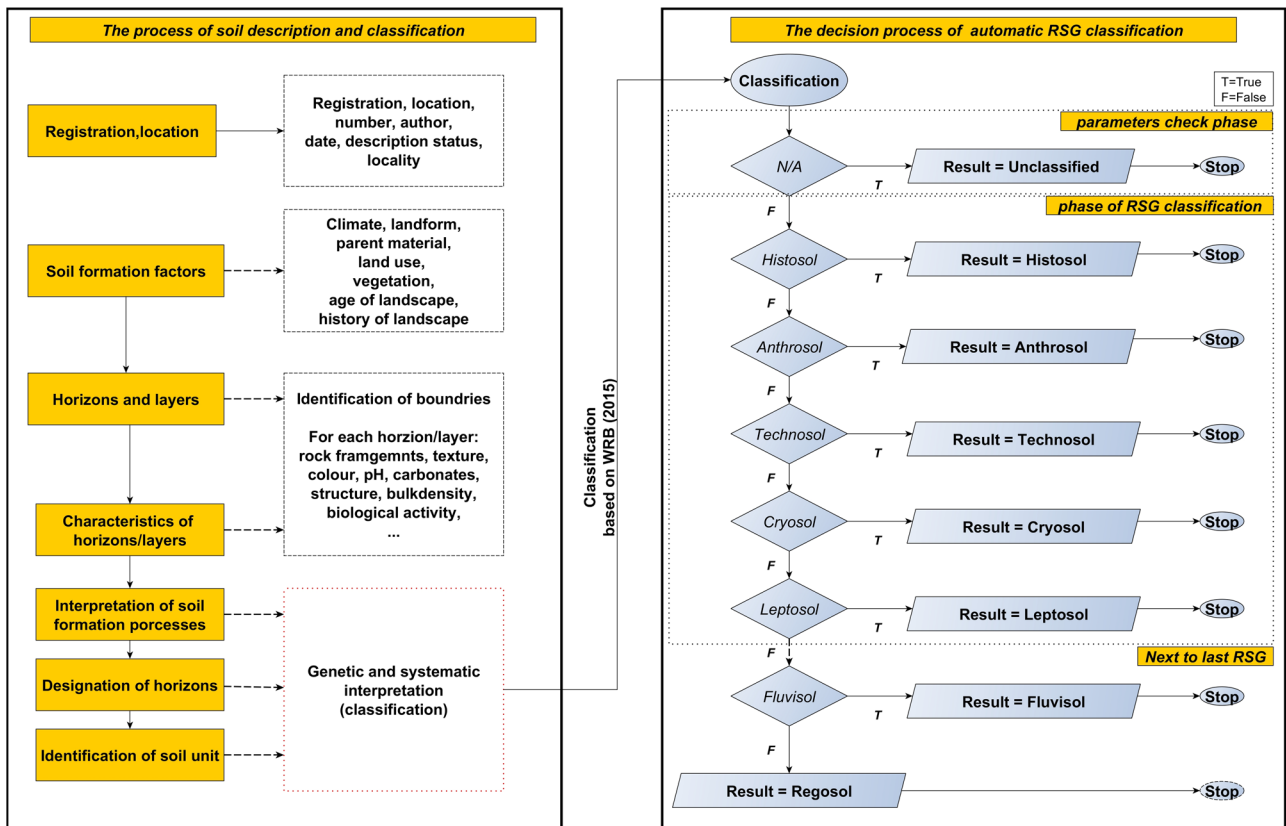


Figure 20: Process of soil description and automatic classification based on WRB 2014 (2015 update).

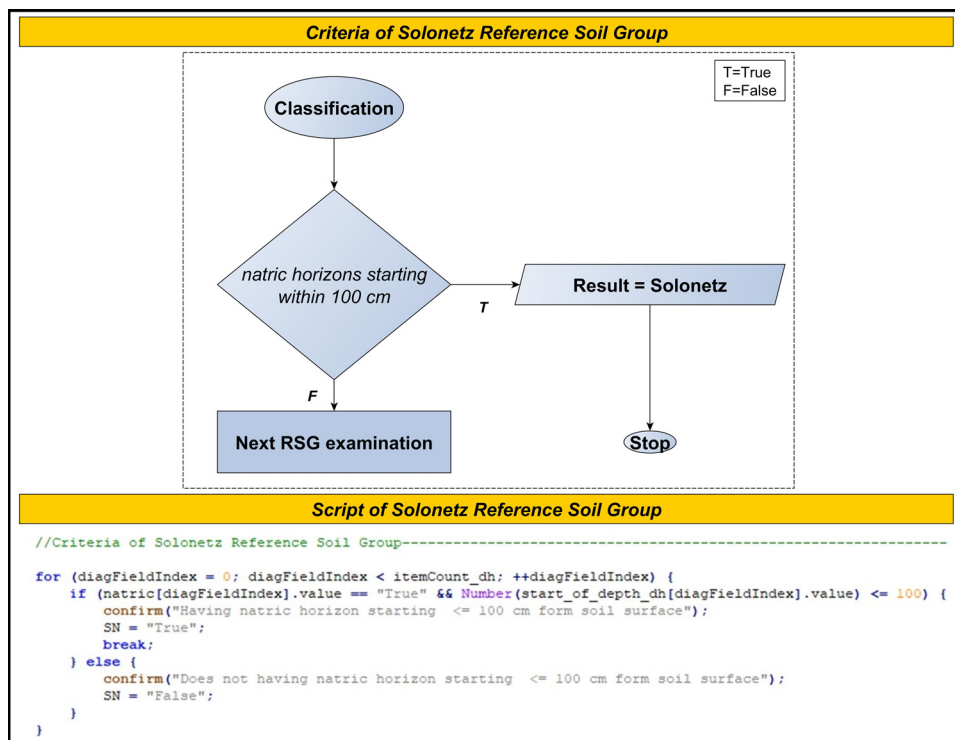


Figure 21: Example script for automatic RSG determination.

3.4 Process of automatic soil classification

The basic idea for the planning of the classification algorithm was to divide it into separate parts. Since the larger tasks (such as handling several RSGs at the same time or separating the diagnostics) cannot be solved all at once, they had to be divided into smaller tasks. Therefore, the classification mechanism now works on a step-by-step basis, starting from the top and fine-tuning so that the smaller tasks (such as the diagnostic properties) are left for last. The specification of the task is therefore to create an algorithm that is capable of classification using the WRB system. The whole of the classification process can be divided into two separate parts. One of them is to develop a proper list of conditions for every reference group, the other is to assign the qualifiers to these groups using a list of conditions for each one, as well. The first step in the classification process is to describe the soil profiles and to store the soil data in the registry. The proper structure of the registry is essential for the accurate recording of the data, as the soil profile can only be classified by using this information. Collecting data is most commonly the result of field work, therefore, getting the most accurate survey of these data is a primary concern. These data are supplemented by the laboratory

measurements that are either too difficult or outright impossible to do in the field. On the second level of the algorithm, the diagnostic horizons, materials and properties must be established by using the lists of criteria associated with each of them (Figures 20 and 21).

In line with Krol et al. (2007), we used Limit check and Internal consistency check to filter extreme or incorrect values and complementary soil data, respectively (Table 4).

4 Conclusions

The main goal of the open-access system is to store data of soil profiles in a soil database, and at same time, also to geovisualize and classify them in an online environment. The purpose of organizing soil data into a database and importing it into the information system was to assess whether, following the establishment of decision rules, the data stored in the soil profiles can be used for international correspondence/harmonization. The implementation of automated classification was performed based on the methodology of Eberhardt and Waltner (2010), since the former, mostly correlation- and harmonization-based activities were usually not performed on the basis of the soil data recorded according to the WRB.

Table 4: List of value ranges of input data

Parameter	Value range
Coordinate (latitude)	−90° to 90°
Coordinate (longitude)	−180° to 180°
Slope angle (%)	0–90
Animal density (%)	0–100
pH (H ₂ O) and pH (KCL)	1–14
BS (base saturation) (%)	0–100
EC _{1:2,5} (dS m ^{−1})	≥0
Humus content (%)	0–100
CaCO ₃ (%)	0–100
Sand (%)	0–100
Silt (%)	0–100
Clay (%)	0–100
CEC (cation exchange capacity) (meq/100 g)	≥0
S-value (meq/100 g)	0–100
K ⁺	0–100
Na ⁺	0–100
Ca ²⁺	0–100
Mg ²⁺	0–100
Al ⁺	0–100
P ₂ O ₅ (mg/kg)	≥0
NO ₃ [−] (mg/kg)	≥0
y1	≥0
P (mg/L)	≥0
SiO ₂ (%)	0–100
C _{org} (total organic carbon) (%)	0–100
Salt (%)	0–100
ESP (exchangeable sodium percentage) (%)	0–100
EC _{SE} (electrical conductivity of the saturation paste) (dS m ^{−1})	≥0
Check texture	Sand (%) + Silt (%) + Clay (%) = 100
Check ESP (%)	$\left\{ \frac{\text{Na}}{\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^{+} + \text{Na}^{+}} \right\} \times 100 \geq 100$

Instead of the recommended harmonization of certain national soil classification units, we used the original soil recording data for the identification of WRB RSGs. The classifying algorithm was based on the principle of deconstruction into parts. During the classification mechanism we developed, we used a top-down algorithm designing method, during which the algorithm first tests whether the requirements of RSGs are met and then evaluates the criterion system of reference groups in accordance with the WRB. With respect to the evaluation of the decision-making rules created, it should be noted that these only represent a “best approximation” for the investigated soil profiles and do not replace the process of field data recording, description and classification conducted in accordance with the detailed WRB methodology. However, until

these data are available in a sufficient amount and with appropriate spatial density, it provides a good opportunity for preliminary estimations.

Acknowledgments: This work was supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund. Research work of Tibor József Novák was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/00448/17/10).

Conflict of interest: The authors declare no conflict of interest.

References

- [1] Albrecht C, Huwe B, Jahn R. Zuordnung des Munsell-Codes zu einem Farbnamen nach bodenkundlichen Kriterien. *J Plant Nutr Soil Sci.* 2004;167:60–5. doi: 10.1002/jpln.200321315.
- [2] Balla D, Zichar M, Kovács Z, Sándor G, Szabó G, Novák T. Open-access soil information and soil classifier system. *Int J Electron Commun Soft Comput Sci Eng (IJECSCSE).* 2015a;4(3):12–5.
- [3] Balla D, Zichar M, Barkóczi N, Varga OG. Cognitive interpretation of different spatial databases in web environment. In: 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom). Győr: 19–21 October, 2015b. p. 159–62. doi: 10.1109/CogInfoCom.2015.7390582.
- [4] Balla D, Varga OG, Barkóczi N, Novák TJ, Zichar M, Karika A. Methods of processing and geovisualization of soil profiles. *J Agric Inform.* 2016;7(2):11–8. doi: 10.17700/jai.2016.7.2.286.
- [5] Bockheim JG, Gennadiyev AN. The role of soil-forming processes in the definition of taxa in soil taxonomy and the World Soil Reference Base. *Geoderma.* 2000;95:53–72. doi: 10.1016/S0016-7061(99)00083-X.
- [6] Eberhardt E, Waltner I. Finding a way through the maze – WRB classification with descriptive data. In: Gilkes JR, Prakongkep N, editors. *Soil Solutions for a Changing World: proceedings of the 19th World Congress of Soil Science.* Brisbane, Australia: International Union of Soil Sciences; 1–6 August 2010. p. 5–8.
- [7] Food & Agriculture Org. Of The United Nations. Guidelines for soil description. Rome: FAO; 2006.
- [8] Gray JM, Humphreys GS, Deckers JA. Distribution patterns of World Reference Base soil groups relative to soil forming factors. *Geoderma.* 2011;160(3–4):373–83. doi: 10.1016/j.geoderma.2010.10.006.
- [9] Huyssteen CW, Michéli E, Fuchs M, Waltner I. Taxonomic distance between South African diagnostic horizons and the World Reference Base diagnostics. *Catena.* 2014;113:276–80. doi: 10.1016/j.catena.2013.08.010.
- [10] IUSS Working Group WRB. World Reference Base for Soil Resources (2006). *World Soil Resources Reports*, vol. 103. Rome: FAO; 2007. p. 93.

- [11] IUSS Working Group WRB. World Reference Base for Soil Resources (2014). World Soil Resources Reports, vol. 106. Rome: FAO; 2014. p. 181.
- [12] IUSS Working Group WRB. World Reference Base for Soil Resources (2015). World Soil Resources Reports, vol. 106. Rome: FAO; 2015. p. 192.
- [13] Krasilnikov P, Arnold R, Michéli E. Soil classification of Hungary. In: Krasilnikov P, Ibáñez, Martí J-J, Arnold R, Shoba S, editors. A Handbook of Soil Terminology, Correlation and Classification. London, Sterling, VA: Earthscan; 2009. p. 170–5.
- [14] Krol B, Rossiter DG, Siderius W. Ontology-based multi-source data integration for digital soil mapping. In: Lagacherie P, McBratney AB, Voltz M, editors. Digital Soil Mapping: An Introductory Perspective. Developments in Soil Science, vol. 31. Amsterdam: Elsevier; 2007. p. 119–33.
- [15] Láng V, Fuchs M, Waltner I, Michéli E. Soil taxonomic distance, a tool for correlation: as exemplified by the Hungarian Brown Forest Soils and related WRB Reference Soil Groups. *Geoderma*. 2013;192:269–76. doi: 10.1016/j.geoderma.2012.07.023.
- [16] Michéli E, Fuchs M, Hegymegi P, Stefanovits P. Classification of the major soils of Hungary and their correlation with the World Reference Base for Soil Resources (WRB). *Agrochem Soil Sci*. 2006;55(1):19–28. doi: 10.1556/Agrokem.55.2006.1.3.
- [17] Michéli E, Fuchs M, Láng V, Szegi T, Kele G. Methods for modernizing the elements and structure of the Hungarian soil classification system. *Agrochem Soil Sci*. 2014;63(1):69–78. doi: 10.1556/Agrokem.63.2014.1.8.
- [18] Mester T, Balla D, Botos Á, Szabó G, Sándor G, Novák T. Az antropogén hatások mértékének és jelentőségének értékelése WRB irányelvek alapján tiszántúli kertek talajában. *Talajvédelem: Klsz*; 2017. p. 179–87.
- [19] Michéli E, Csorba Á, Szegi T, Dobos E, Fuchs M. The soil types of the modernized, diagnostic based Hungarian soil classification system and their correlation with the World reference base for soil resources. *Hungarian Geograph Bull*. 2019;68(2):109–17. doi: 10.15201/hungeobull.68.2.1.
- [20] Nachtergaele FO, Spaargaren O, Deckers JA, Ahrens B. New developments in soil classification World Reference Base for Soil Resources. *Geoderma*. 2000;96:345–57. doi: 10.1016/S0016-7061(00)00023-9.
- [21] Pásztor L, Szabó J, Bakacsi Zs, Laborczi A. Elaboration and applications of Spatial Soil Information Systems and Digital Soil Mapping at Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences. *Geocarto Int*. 2012;28(1):13–27. doi: 10.1080/10106049.2012.685895.
- [22] Reintam E, Köster T. The role of chemical indicators to correlate some Estonian soils with WRB and soil taxonomy criteria. *Geoderma*. 2006;136:199–209. doi: 10.1016/j.geoderma.2006.03.028.
- [23] Roca N, Pazos MS, Bech J. The relationship between WRB soil units and heavy metals content in soils of Catamarca (Argentina). *J Geochem Explorat*. 2008;96:77–85. doi: 10.1016/j.gexplo.2007.04.004.
- [24] Shi XZ, Yu DS, Xu SX, Warner ED, Wang HJ, Sun WX, et al. Cross-reference for relating genetic soil classification of China with WRB at different scales. *Geoderma*. 2010;155:344–50. doi: 10.1016/j.geoderma.2009.12.017.
- [25] Waltner I. Hazai térképi és más talajadatok nemzetközi digitális térképekbe és adatbázisokba illesztésének megalapozása Doktori (PhD) értekezés. Szent István Egyetem Környezettudományi Doktori Iskola. 2013. (in Hungarian); 2013.
- [26] Zádorová T, Penížek V. Problems in correlation of Czech national soil classification and World Reference Base 2006. *Geoderma*. 2011;167–168:54–60. doi: 10.1016/j.geoderma.2011.09.014.