Finding a way through the maze – WRB classification with descriptive soil data

Einar Eberhardt[A] and István Waltner[B]

[A]Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Stilleweg 2, D-30655 Hannover, Germany, einar.eberhardt@bgr.de
[B]Szént István University, 2103 Gödöllő, Páter Károly utca 1, Hungary, Waltner.Istvan@mkk.szie.hu

Abstract

Huge amounts of soil data are recorded in national databases using national nomenclatures. To use them in international context, they have to be harmonized with an international nomenclature. The World Reference Base for Soil Resources (WRB) provides such a framework. We present our approach to retrieve information from existing soil databases and nomenclatures for classifying the soils according to WRB in the form of graphical algorithms. Our study identifies some general mismatches between typical soil profile database structures and the data requirements for testing the presence or absence of WRB diagnostics. In addition, the difficulties originating e.g. from the data structure for recording horizon-related data, differing class limits in national and FAO soil description systems, and ill-determined rules for soil description have to be addressed for proper identification of WRB diagnostics with existing soil data.

Key Words

Soil profile database, database structure, data harmonisation, soil classification.

Introduction

New soil classifications will always be challenged by existing profile data. This applies particularly to the World Reference Base for Soil Resources WRB (IUSS Working Group WRB 2007), which is intended to be a framework for international correlation, and is likely to be used with existing national data. The huge amount of existing national soil data - e.g. in the European Union at least 330,000 digitally available soil profile datasets recorded using national nomenclatures (Baritz *et al.* 2008) - requires tools for automated transformation into internationally usable data.

Connections between morphogenetic soil classification systems and WRB can often not be established in a clear and unambiguous way (e.g. a correlation between German soil taxa with WRB Reference Soil Groups (RSG) as implemented in Adler *et al.* 2004 can be established for about 50 percent of the taxa only). Levels of detail of soil descriptions vary and are related to national survey guideline design and the purpose of individual survey campaigns for which the data were obtained. Hence, it is necessary to develop specific algorithms for each national description system in order to optimally evaluate the data. Furthermore, challenges arise from the structures in which soil data are recorded on the one hand and what information has to be derived from it for WRB classification purposes on the other.

In the present paper we present our approaches to retrieve information from existing soil databases for classifying the described soils according to WRB. We identify typical mismatches between WRB definitions of diagnostics (diagnostic horizons, materials and properties), the key to the Reference Soil Groups (RSG) and the definitions of qualifiers on the one hand and existing (German and Hungarian) soil data structures and database content on the other.

Source structure

The source of the data used here is taken from typical national soil profile and analytical databases. For the German study, the parameters and code lists of the German soil mapping guideline (Ad-hoc-AG Boden 2005) have been used. Data for the validation of these algorithms comprise: (1) A synthetic dataset that covers a huge number of edge cases, (2) German soil profiles of the European forest soil assessment (ICP Forests programme, ICP Forests 2009), and (3) Soil descriptions (German and according to FAO 2006) with WRB classification in the field obtained for this study and representing widespread, but possibly problematic soil types from various landscapes. For the Hungarian study, validation data (soil profile descriptions and analytical parameters) were collected from various institutions. Number and kind of parameters vary widely between both national studies.

Target structure

Analyses of the definitions of diagnostics, RSGs, qualifiers and specifiers are based on IUSS Working Group WRB 2007. The WRB classification is based on the presence or absence of diagnostics, i.e. diagnostic horizons, diagnostic materials, and diagnostic properties. The key to the Reference Soil Groups gives additional specifications, e.g. depth of diagnostics or stricter colour requirements. Qualifier definitions often ask for presence of a diagnostic only, but may also give further specifications; furthermore, an overall rule defines that a qualifier only applies if not a more specific one that also expresses all its requirements is assigned as well. Specifiers always express further detail, but are not based on diagnostics.

WRB distinguishes between typically associated, intergrade and other qualifiers. For our purpose, it is more important to distinguish between basically two types of qualifiers: the one stating that the relative difference in the expression of a parameter over the profile is above a specific level (relative type), and the other expressing that a profile part fulfils absolute requirements with regard to a specific parameter (absolute type). Complex WRB diagnostics include very often both types of definitions, e.g. by combining an absolute threshold value with a difference to underlying or overlying layers (complex type, e.g. the *argic horizon*).

Results and Discussion

Starting from existing soil profile and analytical database structures, we developed graphical algorithms that check the presence of WRB diagnostics (diagnostic horizons, properties and materials, Figure 1). The algorithms take into account various levels of data availability and detail, e.g. classified instead of number values. Figure 1 presents the algorithm for the *argic horizon* based on German soil data. Missing lab data can be substituted by the morphological horizon description. This may cause imprecise results, if classified parameters have class boundaries that do not conform to the respective FAO/WRB classes, e.g. national texture class data. Alternative parameter combinations are tested one after another, sometimes with a decreasing reliability. Note that in this case, the definition of the horizon symbol *Bt* is so near to the

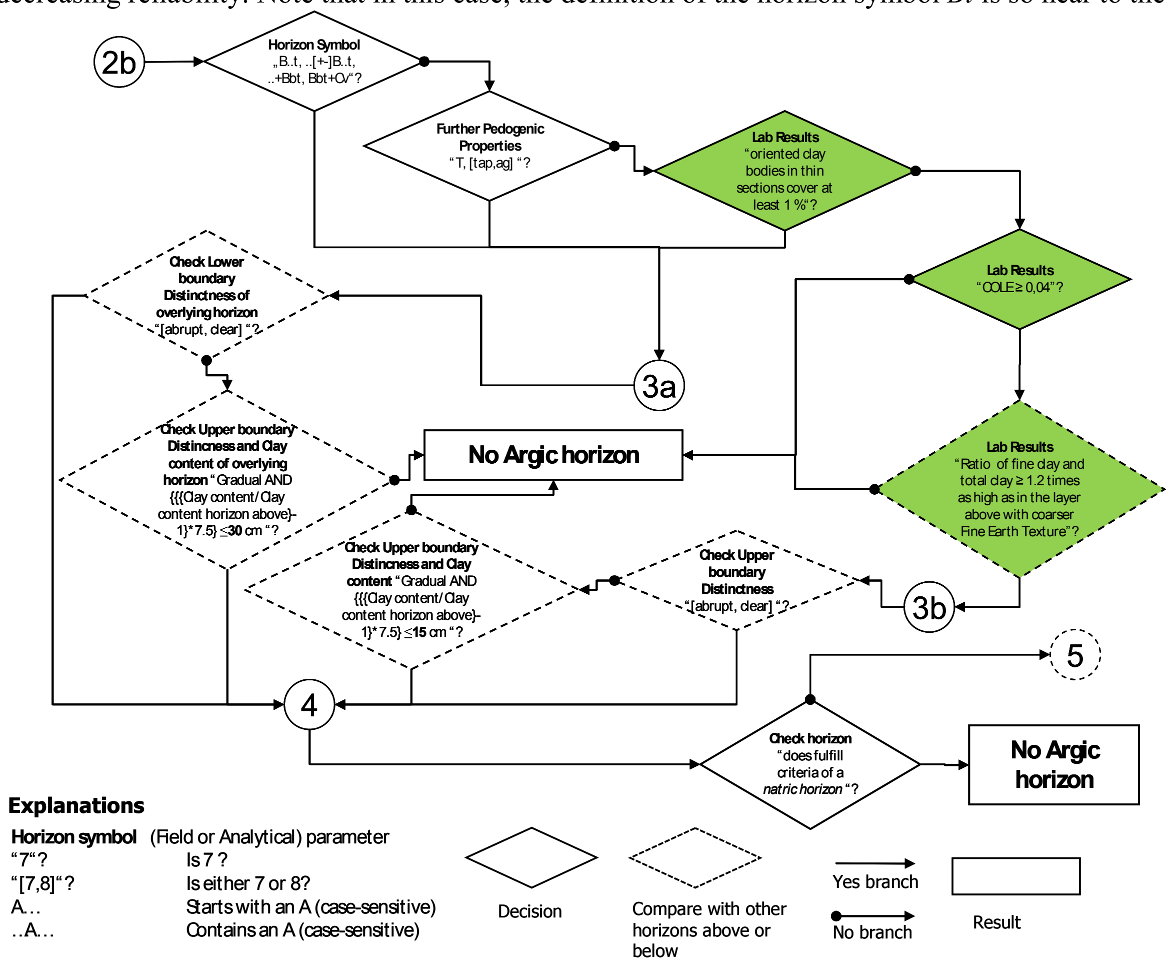

Figure 1. Graphical algorithm. Example: medium part of the algorithm for the *argic horizon* with German soil data, which includes altogether 31 decisions. Lab parameters are highlighted with green colour, numbers relate to criteria of WRB (continuous line marks start of the criterion, dashed line means 'continue with the respective criterion').

definition of the *argic* that it can be assumed almost any *B..t* horizon is an *argic horizon* and therefore, the horizon symbol criterion is put in the beginning of this part of the algorithm. Many other horizon symbol definitions deviate substantially from their FAO analogues or definitions are too wide for a direct correlation, in the German as well as the Hungarian system.

The criteria for the presence of an *argic horizon* comprise absolute and relative elements. The absolute criteria can be determined easily from a database holding the respective parameters. For each morphogenetic horizon the value of the respective parameter is checked whether it fits the requirements of the diagnostic; for cumulative depth requirements, the thickness of all genetic horizons matching the requirements is calculated. Simple thickness requirements are more difficult to check, because more than one pedon section - each consisting of one or more pedogenetic horizons - might fulfil the requirements of a diagnostic. For each section it thus has to be checked independently whether it meets the thickness requirement. The relative criteria - e.g. difference of clay content for an *argic horizon* - are technically challenging because neither upper nor lower boundary can be fixed from the beginning of the analysis, which means that it is not clear which value (and possibly mean value over several adjoining genetic horizons) is to be compared with the respective value of which genetic horizon above or below. For definitions comprising absolute and relative elements, it is therefore recommended to check absolute requirements first, and only for those genetic horizons that match these requirements, to check the relative requirements afterwards.

Morphogenetic descriptions, particularly with classified parameters, may hinder the identification of WRB diagnostics (Figure 2a). The way how soil horizon related data are recorded in typical database structures may even hide WRB diagnostics (Figure 2b). These problems may originate from the differences between the traditional and the WRB approach to the soil profile: in the former, you identify (morphogenetic) horizons, describe their properties and those of their boundaries, while using the latter, you look for the presence or absence of diagnostics, sometimes applying many and very complex criteria. The former approach provides an inherent, horizon-based structure for the data. In contrast, WRB diagnostics are structurally related directly to the profile as a whole and provide no simple inherent structure for data recording.

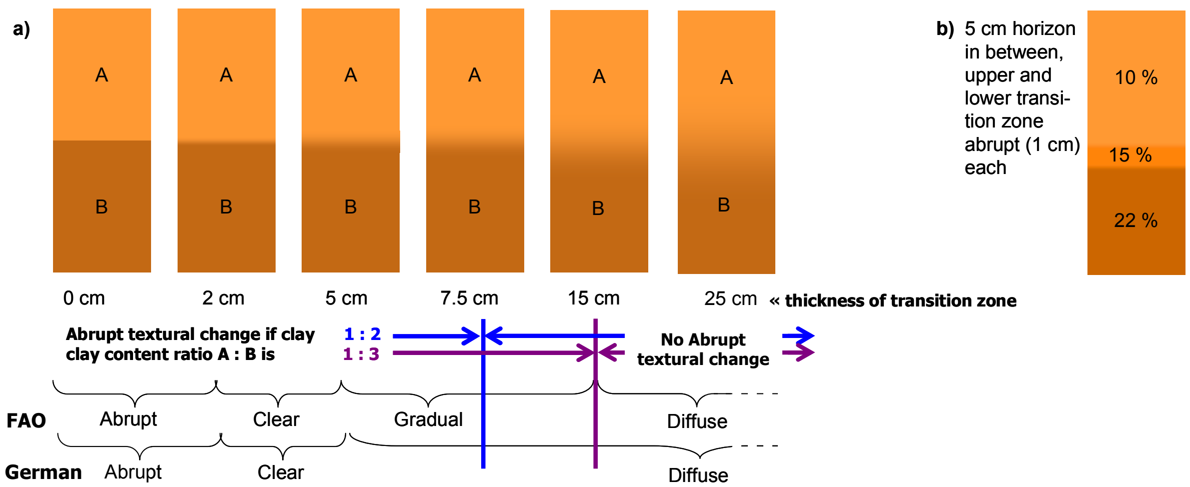

Figure 2. Identification of an *abrupt textural change* (ATC). a) Distinctness of the transition between morphogenetic horizons is recorded with definite classes, while the definition of the ATC is based on relative change. In very few cases, class limits equal the relative threshold values, e.g. when clay content triples (purple threshold line, matching the class limit between *gradual* and *diffuse* horizon transition acc. to FAO 2006). In contrast, the threshold when clay content doubles does not match a class limit (blue line). The German horizon description fails in both cases because the lower class limit for *diffuse* already starts at 5 cm. b) A shallow horizon with intermediate clay content hides the ATC when only adjoining horizons are compared with each other regarding their clay contents.

Different styles in describing soils can produce inconsistent datasets, mostly when detailed survey guidelines are not followed nation-wide or over time (as e.g. in Hungary), or are - although quite elaborated - ambiguous to some extent. E.g. the detection of an *albeluvic tonguing* is relatively easy when the lower boundary of the bleached horizon is described as tongue-like (Figure 3, left); when tongues are long, it is more likely that the surveyor describes a so-called combination horizon with eluviated and illuviated domains (Figure 3, right). In the latter case, the information that the upper horizon penetrates into the *argic horizon* with tongues is lost, because no data on the distribution and shape of the different domains is recorded. No threshold value for the thickness of the transition zone has been defined below which the former and above which the latter way of description shall be used.

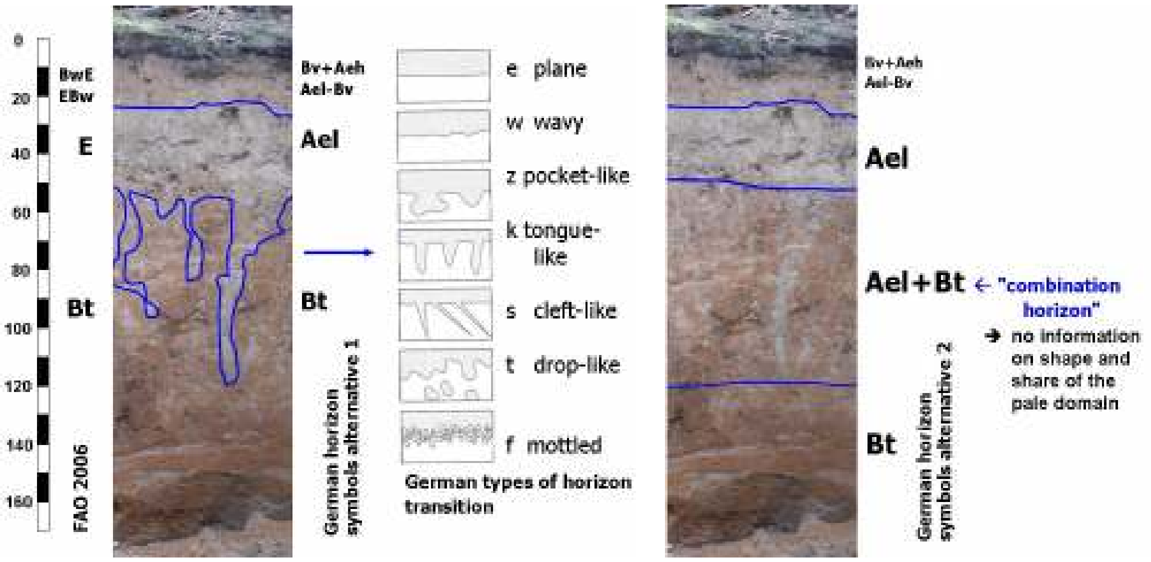

Figure 3. Alternatives for the description of the morphogenetic horizonation according to the German soil mapping guidelines (Ad-hoc-AG Boden 2005) with the possibility to identify *albeluvic tonguing* from the profile description (left) and without (right). Both styles would be compliant with the guideline, the latter case is more common.

Conclusion

Several types of mismatches between the requirements of WRB and the existing data have to be addressed: 1. Missing data. 2. Analytical data obtained with deviating methods. 3. Mismatch of class boundaries for classified values between national and FAO (Food and Agriculture Organization of the United Nations 2006) or WRB nomenclature. 4. Inconsistencies if proxy parameters are used. 5. Description alternatives in guidelines that evoke various rule-conform author's styles. 6. Morphogenetic horizonation may hide the presence of diagnostics. 7. Structure of recording depth information affects identification of diagnostics. While 1-4 are content-related and to be solved individually for each description system, 6 and 7 result from soil data structure and have to be considered in the technical implementation. Many of the inconsistencies can be overcome by proper data analysis.

Acknowledgements

We are indebted to Peter Schad, Erika Michéli and Dana Pietsch for their support. The authorities of several German federal states and Hungarian institutions kindly provided us with soil profile and analytical datasets. Financial support for parts of our studies received from the European Union (project eSOTER, EU's Seventh Framework Programme) is kindly acknowledged.

References

Ad-hoc-AG Boden (2005) 'Bodenkundliche Kartieranleitung. 5th ed.' (Schweizerbart: Stuttgart).

Adler GH, Behrens J, Eckelmann W, Feinhals J, Hartwich R, Krug D (2004) 'Übersetzungsschlüssel zum Transfer von Bodendaten aus den deutschen Klassifikationen KA3 nach KA4 und von KA4 in die internationalen Klassifikationen nach WRB und FAO.' *Arbeitshefte Boden* **1**, 7–49.

Baritz R, Heineke HJ, Eberhardt E, Jones RJA (2008) 'Where is the information from 100 years of soil surveys and research in Europe?' (unpubl. pres., EuroSoil 2008: Vienna).

FAO (2006) 'Guidelines for soil description, 4th ed.' (Food and Agriculture Org. of the UN: Rome).

ICP Forests (2009) International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests. (ICP Forests: Hamburg) www.icp-forests.org/.

IUSS Working Group WRB (2007) 'World reference base for soil resources 2006, update 2007, 2nd ed.' World Soil Resources Reports 103. (FAO: Rome). http://www.fao.org/ag/agl/agll/wrb/doc/wrb2007_corr.pdf, 21.10.2009).