https://doi.org/10.17221/115/2026-SWR

# Methodological approaches in the development of a quantified Czech soil classification system

*Tereza Zádorová[1]*, Anna Margoldová[2], Aleš Kučera[3], Miloš Boček[3,4], Jiří Jandák[5], Pavel Šamonil[3,6], Dušan Vavříček[3], Vítězslav Vlček[5], Anna Žigová[7], Daniel Žížala[2], Petr Zapletal[3], Jan Skála[2], Tomáš Hrdlička[1], Aleš Vaněk[1], Vít Penížek[1]*

[1]*Department of Soil Science and Soil Protection, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Prague, Czech Republic*
[2]*Research Institute for Soil and Water Conservation, Prague, Czech Republic*
[3]*Department of Geology and Soil Science, Faculty of Forestry and Wood Technology, Mendel University in Brno, Brno, Czech Republic*
[4]*Czech Forestry Institute, Branch Hradec Králové, Czech Republic*
[5]*Department of Soil Science and Plant Nutrition, Faculty of AgriSciences, Mendel University in Brno, Brno, Czech Republic*
[6]*Department of Forest Ecology, Landscape Research Institute, Brno, Czech Republic*
[7]*Institute of Geology of the Czech Academy of Sciences, Prague, Czech Republic*

**Corresponding author: zadorova@af.czu.cz*

**Citation:** Zádorová T., Margoldová A., Kučera A., Boček M., Jandák J., Šamonil P., Vavříček D., Vlček V., Žigová A., Žížala D., Zapletal P., Skála J., Hrdlička T., Vaněk A., Penížek V. (2026): Methodological approaches in the development of a quantified Czech soil classification system. Soil & Water Res., XX: 00–00.

**Abstract:** Approaches to soil classification have evolved in recent years, driven by a growing need for quantitative, accurate data and detailed soil maps, as well as changes in general understanding of soils and their functions. This study presents methods for innovating the existing Czech soil classification system through various analyses of an extensive database of legacy data from different soil surveys, with the aim of clarifying and refining soil diagnostics and increasing the usability of soil classification in diverse applications. The approaches included establishing precise quantitative limits for diagnostic criteria based on statistical analysis of the dataset, compiling and testing the key to soil types, and evaluating the consistency and taxonomic relationships of the resulting soil units using taxonomic distance analyses. The results demonstrated the importance of legacy data as a basis for quantifying diagnostic criteria, as they allow the setting of thresholds that correspond to actual land cover. Exact limits and the organisation of soil groups into a classification key led to redistribution of studied profiles between units, often in favour of less developed soil types (kambizem, regozem). However, the newly proposed and actual soil units are taxonomically very close; thus, despite the change in the approach to classification, continuity between the two systems has been maintained.

**Keywords:** Czechia; diagnostic criteria; legacy data; soil type; taxonomic distance

Supported by the Technology Agency of the Czech Republic (Project No. SQ01010088), by the Ministry of Agriculture of the Czech Republic (MZE-RO0218) and by the Institute of Geology of the Czech Academy of Sciences (RVO 67985831).

© The authors. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0).

https://doi.org/10.17221/115/2026-SWR

Soil classification represents our current knowledge of soil evolution and functions (Arnold 2002). Shifts in its concepts and approaches to soil diagnostics reflect (*i*) new perspectives on the role of soil in relation to human use of the landscape, including a growing understanding of soil threats, (*ii*) gradually evolving diagnostic and analytical methods usable in soil classification, and (*iii*) the needs and demands placed on the outputs of soil classifications, primarily in terms of their usability in applications based on quantitative data, such as environmental modelling and digital soil mapping. Our understanding of soil is shifting from an emphasis on its productive functions to a multidirectional ecological approach (Kabała et al. 2019a), and demands on soil data are also evolving, particularly with regard to their inclusion in various mapping, predictive dynamic modelling and database platforms at the national and international levels (Batjes 2016; Zádorová et at. 2021; Žížala et al. 2022). Moreover, soil classification plays a key role in a number of decision-making mechanisms at both national and international levels, particularly in terms of soil cover regionalisation and the associated spatial evaluation of soil ecosystem functions, soil bonitation scores and the implementation of soil protection measures. This is reflected, for example, in the recently approved Soil Monitoring Law (Directive (EU) 2025/2360 of the European Parliament and of the Council of 12 November 2025 on soil monitoring and resilience), where soil taxa play a key role in defining soil districts for soil health monitoring (Ginzky & Pieper 2025).

Given the use of soil classification not only in soil survey and mapping, but also as a multi-purpose input in diverse quantitative applications, databases and global maps, it is crucial to have a relevant soil classification system that allows for unambiguous diagnosis of soil units and easy harmonisation with international classification systems (Michéli et al. 2019). The current Czech Taxonomic Soil Classification System (CTSCS; Němeček et al. 2011) adopts a morphogenetic approach, whereby higher taxa (reference classes) are distinguished on the basis of dominating pedogenesis, and classification at lower taxon levels (soil types, subtypes, varieties) is based primarily on the identification of diagnostic characteristics. The classification builds on the Morphogenetic Soil Classification System (MSCS; Hraško et al. 1991), while also reflecting the progress in international classification systems, namely World Reference Base for Soil Resources (WRB; IUSS Working Group 2006). Convergence with this classification is reflected primarily in the establishment of the highest taxonomic category in CTSCS – reference soil class. Several factors reduce the potential for full utilisation of the current classification in diverse applications (Zádorová et al. 2020; 2021). These include: (*i*) inconsistencies in the definition of diagnostic characteristics, particularly frequent unquantified criteria or limits of properties that tend to reflect typical or average values of the given characteristics, (*ii*) possible ambiguity in assignment, where individual taxa are characterised without preference rules, (*iii*) insufficiently developed diagnostics of some units, especially anthropogenic soils or some less widespread units. Revisions to the classification should therefore primarily focus on more accurate settings for diagnostic characteristics based as much as possible on quantified identifiers, enabling unambiguous classification of profiles. The relevant setting of quantitative criteria for diagnostic characteristics should reflect the actual state of the soil cover in each country. In addition to expert knowledge and field experience, it is also advisable to process robust datasets, enabling statistical evaluation of individual soil characteristics and setting their thresholds based on actual soil conditions (Michéli et al. 2014). Currently, such datasets are available primarily through legacy data from national and international soil surveys, long-term monitoring programmes and mapping campaigns, which have been frequently used in the for, among other, modernization (Michéli et al. 2014, 2019) or correlation of classification systems (Láng et al. 2013; Hughes et al. 2017; Zádorová et al. 2021) using methods of taxonomic distance (Van Huyssteen et al. 2014; Láng et al. 2016; Michéli et al. 2016), iterative algorithmic approach (Láng et al. 2024) or multivariate statistical analyses (Reyes Rojas et al. 2025).

The main objective of this study is to show how quantitative analysis can contribute to the optimisation of approaches to soil classification by approximating diagnostic criteria to actual soil cover. Particularly, we aim (*i*) to describe the procedures and methodologies leading to the optimisation of the current Czech soil classification by establishing quantified data-based diagnostic criteria and a classification key, (*ii*) to show how this approach is reflected in the distribution of newly defined soil units, (*iii*) to show how the newly proposed soil units taxonomically relate to their counterparts in the current soil classification and WRB 2022. The study does not provide

https://doi.org/10.17221/115/2026-SWR

proposals for the general concepts of the revised classification, the definition of hierarchical classes, or other aspects of the overall design of the revised classification, nor does it specify the final values of diagnostic criteria or the format of the classification key. These are currently being established by the Commission for Soil Classification of the Czech Society of Soil Science.

## MATERIAL AND METHODS

**Data sources.** For the purposes of the study, robust legacy databases providing data from agricultural and forest soil surveys and mapping campaigns were used. Specifically, these were: (1) systematic soil survey (SSS; Němeček et al. 1967), (2) national forest inventory (NFI), and (3) forest typology (FT). The databases provide comprehensive horizon-based information on soil profiles, including morphological characteristics (colour, structure, rock fragments content, neoformations), analytical properties (soil organic carbon (SOC) and $CaCO_3$ content, texture, soil pH, base saturation (BS), cation exchange capacity (CEC)), information on parent material and soil classification. A total of approximately 46 000 profiles were included in the analysis, of which 36 000 were so-called selective soil pits from the SSS, 3 500 profiles from the NFI, and 6 500 from the FT. For detailed information on the methodology of SSS, NFI and FT, their sampling schemes and analytical approaches, see Zádorová et al. (2020, 2021) and Kučera (2016).

**Adjustment of the databases.** Firstly, all incomplete profiles and erroneous records have been removed from the databases. Secondly, all profiles were classified in CTCSS to enable the statistical evaluation of individual criteria in relation to their association with the soil taxon. The NFI and FT databases already include profiles in the CTCSS classification. The classification into individual taxa was only verified and, if incorrect, reclassified. In the case of SSS, the profiles were classified using the Genetic-Agronomic Classification (GAC; Němeček et al. 1967), which was created specifically for this mapping campaign. While the individual groups in SSS and CTCSS are similar in concept, there are differences in the specific settings of the diagnostic criteria and in how taxa are defined. Therefore, it was necessary to reclassify soil profiles based on their morphological and analytical characteristics. As the CTCSS concept does not allow the use of a classification key or the gradual elimination of irrelevant units, we must select from the entire set of taxa when classifying the profile. The reclassification was therefore performed using a decision tree according to Zádorová et al. (2020), which allows for automated reclassification while preserving the classification structure as much as possible.

**Setting of the diagnostic criteria.** The proposal for criteria for identifying diagnostic characteristics and setting individual taxa was based on the following principles and rationale:

(*i*) In line with the current classification, the basic set of soil properties and characteristics used for identifying diagnostic features (e.g. colour, SOC content, texture and clay content ratio, content of rock fragments) was retained for most soils. This ensures some continuity and predictability of the new classification, and historically classified soils will be able to be reclassified.

(*ii*) The specific limits of the diagnostic criteria were based on quantitative analysis wherever possible. In terms of morphological properties, these included horizon thickness, depth of morphological features, soil colour, content of rock fragments, soil structure, and the occurrence and type of neoformations. The analytical properties evaluated included SOC content and its distribution with depth, $CaCO_3$ content, soil pH, sorption properties (cation exchange capacity, base saturation), texture, bulk density, P retention and Al and Fe forms. The analytical methods differed for certain properties, both between individual databases and within a single database (particularly the NLI, where data were collected over a longer period). In cases where different methods used yielded incompatible results (e.g., the pipette method versus laser granulometry in determining soil texture), statistical analyses were performed only on data analysed by one predominant method. The properties were statistically processed for individual soil types, soil horizons, and, where applicable, individual depth intervals, always depending on the nature of the individual soil taxa. A descriptive statistical analysis was performed (mean, median, variance, variability), and in some cases, developmental diagrams were used to evaluate the extent of changes in properties within the profile as a diagnostic criterion. The rule removes ambiguity from the system and leads to higher accuracy. The effect of changing the criteria on the distribution of soil units was studied by testing various settings for soil property limits.

(*iii*) Given the inclusion of national soil maps and monitoring in international systems and databases,

https://doi.org/10.17221/115/2026-SWR

it is essential to establish links to internationally recognised classification systems, particularly the WRB (IUSS Working Group 2022), at least in a basic framework. In international systems, the inclusion of all the world's soils and geographical environments means that some quantitative criteria are set more broadly. Therefore, national criteria were primarily set using WRB limits in cases of insufficient data from national databases (typically for soil structure, colour, and other morphological characteristics) or ambiguous analysis results. The rule connects Czech taxonomy geographically and fundamentally with the current global concept. This will enable a seamless solution to global scientific topics.

**Designing the key to soil types.** The key to soil types, which represent the backbone of the Czech classification system, is a fundamental condition for the unambiguous assignment of classified soil to a soil taxon. The key was created for soil types, not reference soil classes, because the former form the basic units of the system, are used in soil surveys, and are also the primary mapping units. This taxonomic level also allows for integration with existing data and a more detailed specification of units based on diagnostic characteristics. Unlike the existing classification, which involves selecting from equivalent soil units, the key is based on determining the order of individual taxa. The key was compiled in a way that soil types with specific characteristics and properties (organic soils, anthropogenic soils, soils with specific soil profile development) were placed at the top of the key, while soils with weaker pedogenesis or soils with less specific properties were placed at the end of the key (Table 1). To maintain continuity in soil classification, an effort was made to retain the basic concepts and names of individual soil types without major changes. Several marginal soil types with unconfirmed occurrence (slanec ~ Solonetz RSG in WRB 2022, solončak ~ Solonchak, andozem ~ Andosol) were omitted. Pelozem (~ Cambisol (Clayic) in WRB 2022) were incorporated into kambizem soil type. For several soil types, there were more important shifts in their general understanding (particularly broader definitions of pararendzinas and regozems).

The soil classification key was implemented in R software to enable the automatic classification of soil data according to newly defined diagnostic criteria.

All profiles were classified based on the newly established criteria and classification key. This classification draft was given the title Soil Classification 2026 (SC 2026). This is a working title for the presented classification concept, which will be used solely for the purposes of this publication.

**Assessment of the taxonomical relationship within and between individual soil classifications.** The newly proposed SC 2026 – key-based Czech soil classification system redefined the diagnostic criteria and set new boundaries for soil units. This is potentially shifting the analytical and morphological properties of soil units relative to the previous national system. We therefore evaluated how the units of SC 2026 relate to those of the existing CTSCS and the international system WRB 2022. The comparison was conducted in two complementary steps. First, alluvial diagrams were used to visualise how individual soil profiles were redistributed among corresponding units of all three classification systems. Second, similarities among soil units were assessed quantitatively using taxonomic distance. To assess whether the revised definitions resulted in more clearly differentiated soil units in multivariate soil-property space, we also performed the within-system comparison for SC 2026 and CTSCS.

There are two principal approaches that have been previously used to quantify taxonomic distance between soil classes: conceptual and centroid-based. The conceptual approach represents each class by a set of diagnostic features or soil-forming process. The distance is calculated from their coded presence, absence or intermediate occurrence (Láng et al. 2016). The centroid-based approach, on the other hand, uses mean values of soil analytical properties and calculates distance between class centroids (Minasny & McBratney 2007; Michéli et al. 2016). Previous work (Zádorová et al. 2021) demonstrated that the conceptual taxonomic distance approach can be successfully applied to the harmonisation and conversion of national soil units. In our study, however, the availability of a robust database containing quantitative and morphological information allowed us to use a centroid-based approach.

The original database contained soil properties recorded at the horizon level. To ensure comparability across profiles, the selected soil properties were recalculated for an increment of 20 cm down to a depth of 100 cm. The selected properties were limited to those properties which are widely available across all datasets. This includes SOC, $CaCO_3$, texture, soil pH, base saturation (BS), and cation exchange capacity (CEC). Soil morphological characteristics were represented by binary variables describing the

https://doi.org/10.17221/115/2026-SWR

Table 1. The order of soil types in the working classification key in Soil Classification 2026 (SC 2026), their brief descriptions and approximate counterparts in World Reference Base 2022 (IUSS Working Group 2022)

| Soil type | | Soil characteristics | Present diagnostic characteristics (according to SC 2026) | Approximate counterparts in WRB 2022 |
|---|---|---|---|---|
| 1 | organozem | soils with deep organic layers | histic horizon | Histosol |
| 2 | kultizem | soils with intensive agricultural use | hortic horizon | Anthrosol |
| 3 | antropozem | man-made soils | anthropogenic material or technic hard layer | Technosol |
| 4 | podzol | soils dominated by eluviation and illuviation of chelates | spodic (humuso-sesquioxidic + sesquioxodic) horizon | Podzol |
| 5 | kryptopodzol | soils with in-situ accumulation of sesquioxides, but no signs of eluviation and iluviation of chelates | spodic (rusty – sesquioxidic) horizon | Podzol, Cambisol |
| 6 | stagnoglej | soils with reducing conditions resulting from long-term surface waterlogging | stagnic horizon and gleyic horizon | Stagnosol |
| 7 | glej | soils developed in predominant reducing conditions resulting from long-term groundwater waterlogging | gleyic horizon | Gleysol |
| 8 | pseudoglej | redoximorphic soils | stagnic horizon | Stagnosol |
| 9 | smonice | clayey soils with shrinking and swelling | tirsic (dark vertic) horizon | Vertisol |
| 10 | černice | soils with deep, dark humus horizons, influenced by shallow groundwater, saturated | chernic horizon, redoximorphic features | Phaeozem, Chernozem, Kastanozem |
| 11 | černozem | soils with deep, dark humus horizon, include vertic and weakly leached soils | chernic horizon, secondary carbonates | Chernozem, Kastanozem |
| 12 | šedozem | soils with clay and humus translocation, dark clay coatings, saturated | organoluvic horizon | Phaeozem, Luvisol |
| 13 | luvizem | leached soils dominated by clay migration, with albic eluvial horizons and/or albeluvic tonguing | luvic horizon, high textural differentiation | Luvisol |
| 14 | hnědozem | soils with clay migration, saturated | luvic horizon | Luvisol |
| 15 | kambizem | soils with brown or reddish horizons of soil weathering | cambic horizon | Cambisol |
| 16 | koluvizem | soils on colluvial sediments, often stratified with uneven humus distribution | colluvial material | Solimovic qualifier |
| 17 | fluvizem | soils on fluvial sediments, often stratified with uneven humus distribution | fluvial material | Fluvisol |
| 18 | litozem | very shallow soils on consolidated rocks | consolidated rock in shallow depth | Leptosol |
| 19 | rendzina | weakly developed soils with rock fragments, on consolidated limestone or dolomite | content of rock fragments and limestone or dolomite parent material | Leptosol, Phaeozem |
| 20 | ranker | weakly developed soils with rock fragments, on non-calcareous parent material | content of rock fragments; on silicate substrate | Leptosol |
| 21 | pararendzina | weakly developed soils on parent material rich in $CaCO_3$ (e.g. marl, calcareous sandstone, loess) | calcareous parent material | Regosol, Leptosol, Phaeozem |
| 22 | regozem | soils with no significant profile developement | other soils | Regosol |

https://doi.org/10.17221/115/2026-SWR

presence or absence of hydromorphic features and consolidated bedrock for each 20 cm interval (Table 2).

Taxonomic distances were quantified at two levels: first among classes belonging to different classification systems, and second, between classes within each of the two national classifications – SC 2026 and CTCSC. The distances were quantified using Mahalanobis distance. All parameters were scaled prior to calculation. For each classification unit, a centroid was calculated as the vector of mean standardised parameter values. The number of observations per class was determined as the number of unique soil-profile identifiers. Only classes represented by at least 10 profiles were kept. For our dataset, the Mahalanobis distance was calculated using the generalised Moore-Penrose pseudo-inverse of a singular matrix:

$$d_{ij} = [(x_i - x_j)^T S^+ (x_i - x_j)]^{1/2}$$

where:
$S^+$ – the Moore–Penrose inverse of the covariance matrix S.

All statistical analyses and visualisations were performed in R software, version 4.6.1 (R core team 2026). The Moore-Penrose generalised inverse of the covariance matrix was calculated using the ginv() function from the MASS package (Venables & Ripley 2002). Alluvial plots illustrating relationships among classification systems were generated using the ggalluvial package (Brunson 2020; Brunson & Read 2026). Maps were prepared in QGIS (Ver. 4.2.0) (QGIS Development Team 2026).

Table 2. List of selected properties used for calculation of centroid of classification SC 2026, CTCSC and WRB 2022 (depth coverage 20 cm interval between 0–100 cm)

| Soil property |
|---|
| Soil organic carbon content (%) |
| Clay content (< 0.002 mm) (%) |
| Silt content (0.002–0.05 mm) (%) |
| Sand content (0.05–2 mm) (%) |
| soil pH (–) |
| $CaCO_3$ content (%) |
| Cation exchange capacity (cmol(+)/kg) |
| Base saturation (%) |
| Presence of hydromorphic features (binary: 0 = absent, 1 = present) |
| Presence of consolidated bedrock (binary: 0 = absent, 1 = present) |

SC 2026 – Soil Classification 2026; CTSCS – Czech Taxonomic Soil Classification System (Němeček et al. 2011); WRB – World Reference Base 2022

## RESULTS AND DISCUSSION

### Quantified diagnostic criteria

The setting of quantitative criteria was tested in various scenarios, e.g., minimum/maximum values of analytically determined soil properties (SOC, soil pH, BS, particle size distribution), horizon thicknesses, depths of the upper limit of feature occurrence (e.g. hydromorphic features), or the representation of soil features in soil taxa (e.g. frequency of neoformations, content of rock fragments). All results of the statistical analysis are presented in the Electronic Supplementary Material (Tables S1–S20). In the following discussion, we highlight a few examples of criterion settings that demonstrate the various types of tasks involved in the quantification of diagnostic criteria.

**Setting diagnostic horizons properties and thicknesses.** Diagnostic horizons are the basis of soil classification, as their characteristics reflect the extent to which various pedogenic processes act on the soil under specific environmental conditions. A genetic horizon is only diagnostic of a given soil taxon if it meets certain criteria, including a minimum thickness; these should be assessed with regard to regional bioclimatic conditions (FAO, ISRIC, ISSS 1998).

We demonstrate the testing of horizon property limits using the example of the chernic horizon, whose presence in the profile is one of the criteria defining černozem. In the current CTSCS classification, as in a number of other classifications (Kabała et al. 2019a; IUSS Working Group 2022; Soil Survey Staff 2022), the minimum thickness of the chernic horizon is set at 30 cm; however, černozem is defined by a chernic horizon thickness of at least 40 cm. Furthermore, the chernic horizon is not defined by a limit of SOC content; for černozem, a minimum humus concentration of 2% (i.e., approximately 1.2% SOC) is specified. This limit thus exceeds the commonly used 1% (Sorokin et al. 2021; IUSS Working group 2022) or even lower threshold values set in the Slovak or Polish classifications (Societas pedologica slovaca 2014; Kabala et al. 2019a), though, according to the findings of Reyes Rojas et al. (2025), Czech černozem profiles are less humic than the global average (Łabaz et al. 2024). Furthermore, it is not specified whether this content applies to part or the

https://doi.org/10.17221/115/2026-SWR

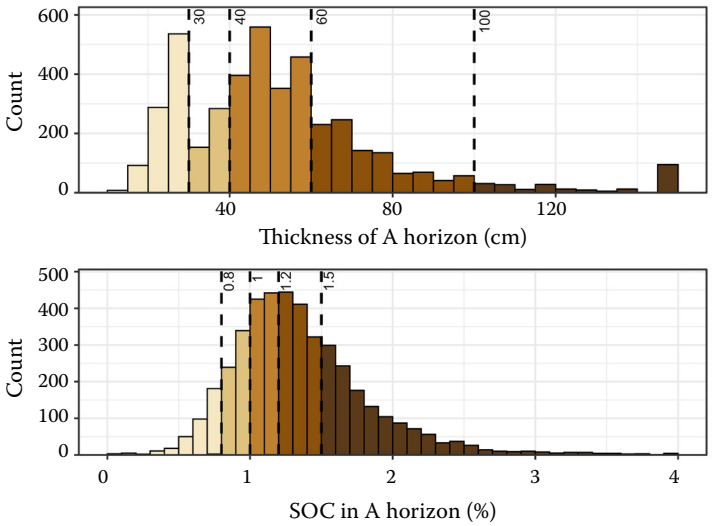



Figure 1. Distribution of the thickness of humus horizons (A) and their soil organic carbon (SOC) content (B) in černozem soil profiles

entire thickness of the horizon. In addition, an eroded form of černozem is distinguished with a reduction in the thickness of the solum and SOC content, which, however, is not defined in terms of quantified criteria.

Figure 1 shows the distribution of thickness and SOC content in the A horizon of černozem, as classified in the SSS database. Although its thickness exceeds 40 cm in most profiles, a relatively large proportion (31%) has a lower thickness. Similarly, regarding SOC content, 30% of profiles in the 0–30 cm depth interval have a value below 1.2%; for the deeper layers, this is the majority of profiles. Therefore, working criteria were set at 30 cm for the thickness of the chernic horizon in černozem and a content of ≥ 1% throughout.

Figure 2 shows how the spatial distribution of černozem profiles varies across three classification

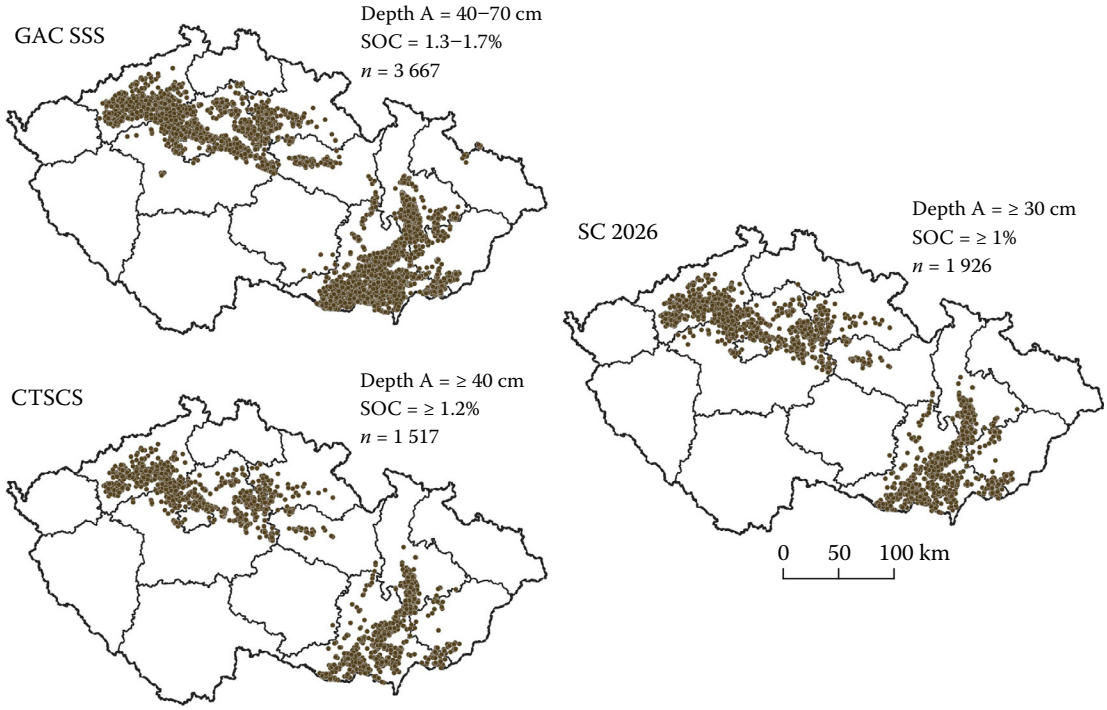



Figure 2. Spatial distribution of černozem soil profiles classified according to the classification criteria of three different taxonomic systems

GAC SSS – Genetic-Agronomic Classification of the Systematic Soil Survey (Němeček et al. 1967); CTSCS – Czech Taxonomic Soil Classification System (Němeček et al. 2011); SC 2026 – Soil Classification 2026; SOC – soil organic carbon

https://doi.org/10.17221/115/2026-SWR

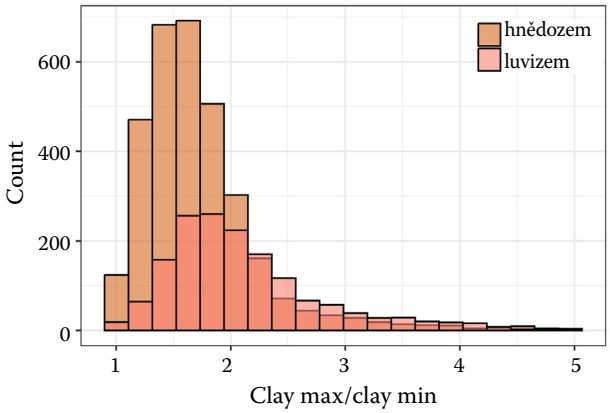



Figure 3. Distribution of textural differentiation (ratio of clay content in illuvial horizon to the overlying coarser layer) in hnědozem and luvizem soil profiles

approaches (the original delineation of černozems in the SSS, classification based on CTSCS criteria, and classification based on SC 2026 working criteria). It is evident that strict adherence to CTSCS criteria results in the exclusion of a significant portion of černozem profiles across the territory. These are not only eroded profiles but, in most cases, soils with lower SOC content (between 1 and 1.2%) in the subsoil. Furthermore, there is a marked reduction in the number of černozem profiles in both modern classifications compared to the historical SSS. This is due, on the one hand, to the original inclusion of šedozem ("greyzems") and smonicas (vertic soils) in the černozem group, and on the other hand, to the broad and non-quantified classification rules in SSS, under which even heavily eroded profiles were classified as eroded černozems.

Another case study involved establishing criteria for soils with the illimerisation process, i.e. physical illuviation of clay and other compounds into the deeper soil layers. In Czechia, three soil types (šedozem, hnědozem, luvizem) undergoing illimerisation are traditionally distinguished, reflecting variations in the intensity of this process. However, the criteria for the luvic (illuvial clay-enriched) horizon and for the classification of individual soil types have gradually changed (e.g. textural differentiation of the luvic/eluvial horizons from 1.2 to 1.5 and back to 1.3; Hraško et al. 1991; Němeček et al. 2001, 2011). To establish appropriate criteria, we tested different settings of textural differentiation and the presence of an eluvial horizon to distinguish the two most related soil types – hnědozem (moderately bleached) and luvizem (bleached eluvial horizon, possible albeluvic tonguing). The thresholds were set primarily based on the distribution of values within individual taxa (Figure 3), as well as on the traditional limits applied in historical Czech and international and national classifications (Němeček et al. 2011; Kabała et al. 2019a; IUSS Working group 2022; Soil Survey Staff 2022).

Figure 4 shows the distribution of the studied soil types under various diagnostic criterion settings. In addition to the actual criteria in CTCSS (textural differentiation of 1.3 for hnědozem and 1.8 for luvizem), thresholds of 1.4 and 1.8, 1.4 and 2.2, and various combinations of the presence and type of eluvial horizons were tested.

The results show that shifts in textural differentiation from values of 1.8 to 2.2 reduce the number of luvizem profiles by about half, with the main changes occurring in agricultural soils that generally have lower leaching rates. The question of whether the bleached eluvial horizon is a required or optional feature appears to be of key importance. This applies primarily to profiles on arable land, where, despite eventual high textural differentiation, the eluvial horizon is often part of the cultivated topsoil.

**Setting the depth limit of diagnostic features.** For soil classification, it is not only the presence of a diagnostic characteristic that is crucial, but also its position within the profile. At higher taxonomic levels, diagnostic characteristics occurring near the soil

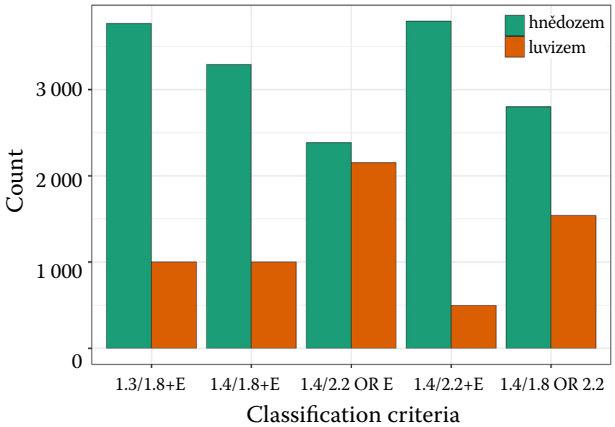



Figure 4. Distribution of hnědozem and luvizem profiles under different working criteria

1.3/1.8 + E: 1.3 – textural differentiation of 1.3 in hnědozem, 1.8 + E textural differentiation of 1.8 and a strongly bleached E horizon is required in luvizem)

https://doi.org/10.17221/115/2026-SWR

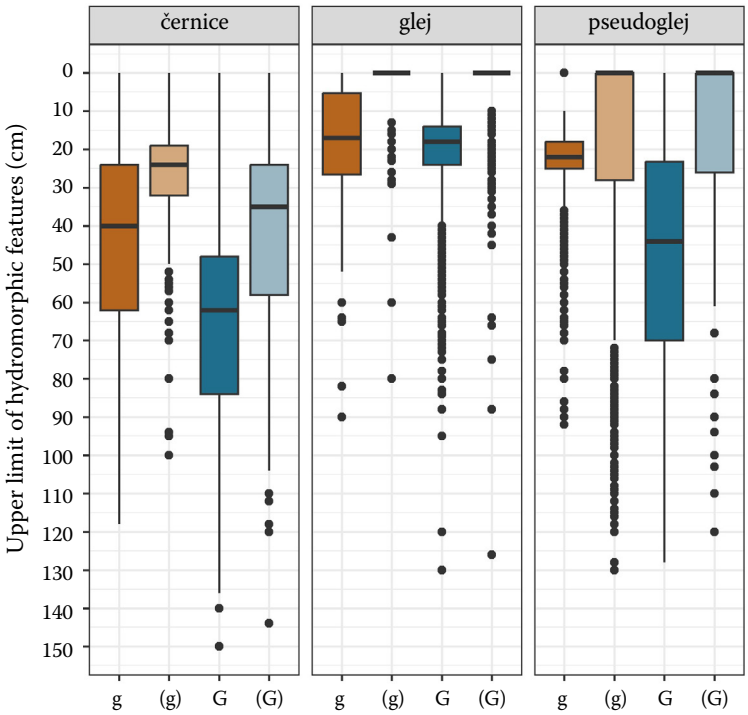



Figure 5. The upper limits of layers with hydromorphic features in different soils
g – pronounced stagnic features; (g) – less developed stagnic features; G – pronounced gleyic features; (G) less developed gleyic features

surface are generally given priority because they have a more significant impact on soil functioning in terms of both ecosystem functions and management (IUSS Working Group 2022). In the CTSCS, the traditional depth limit for identifying diagnostic characteristics and distinguishing taxonomic levels is 0.6 m (Němeček et al. 1967, 2011), and it is especially important for defining hydromorphic soil types and subtypes.

Within the dataset, we tested the upper limits of the occurrence of various levels of hydromorphic features across different taxonomic groups (Figure 5).

It is evident that hydromorphic characteristics occur in stagnic and gleyic soils close to the surface, and the traditionally used depth threshold of 60 cm for the occurrence of these characteristics is adequately set. However, in the case of černice, the results show that, in contrast to the current setting of hydromorphic features occurring right from the surface (Němeček et al. 2011), its occurrence is mostly observed only below the topsoil. This may result primarily from the difficulty of identifying hydromorphic features in the plough layer. For comparison, in the Polish classification (Kabała et al. 2019a), similar units, ciarne ziemie, are characterised by the occurrence of a layer with hydromorphic properties ≤ 80 cm from the soil surface.

**Key to soil types**

The organisation of soils within the key to soil groups significantly simplifies profile classification and prevents ambiguous assignments (Michéli et al. 2019). However, this approach strongly influences both the classification itself and the spatial distribution of individual taxa, as it prioritises certain soil units within the key's order; a simple change in the key's order then results in a significant rearrangement of soil groups, as shown by Láng et al. (2024). Some soil units are more sensitive to this rearrangement than others. For example, podzol, solonchak, and other soils with very specific characteristics do not overlap much with other units, but the classification order is crucial for soils with diagnostic topsoil horizons, hydromorphic features, or for soils with a high content of rock fragments, as they often exhibit characteristics of multiple soil units. Below we show the rearrangement of soil group of shallow and highly skeletal soils (ranker and rendzina) to the top and bottom positions in the key, respectively. Figure 6 shows the resulting rearrangement of affected soil units under these scenarios, as well as changes in the number of profiles within the units. Figure 7 shows changes in their spatial distribution.

The rearrangement most significantly affect kambizem and soils with spodic horizons (podzol, kryptopodzol) in mountainous areas, particularly in forests (8%), where – despite the development of diagnostic horizons – steep slopes very often feature highly skeletal soils formed on the weathered silicate bedrock; these soils will then be clas-

https://doi.org/10.17221/115/2026-SWR

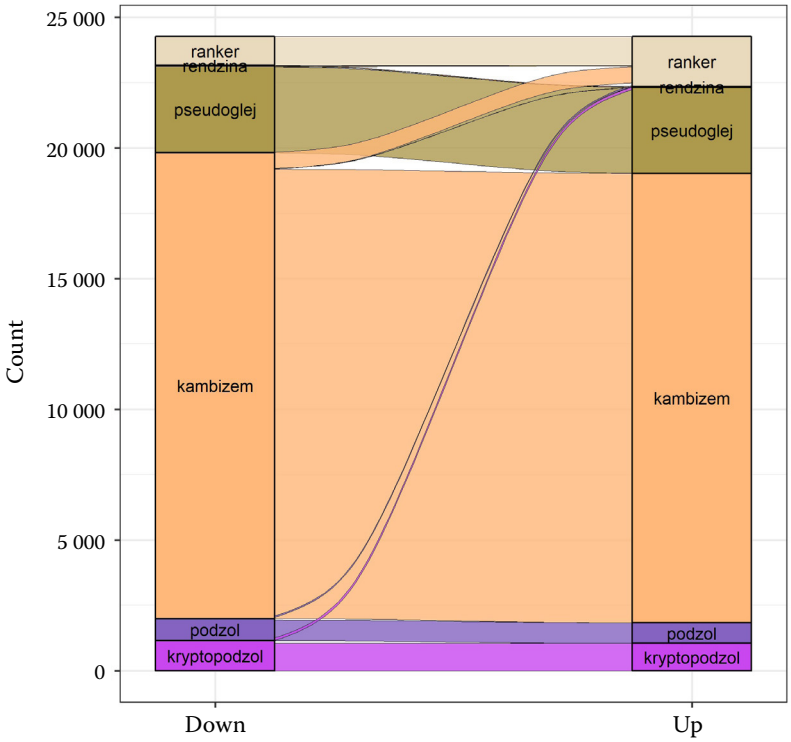



Figure 6. Shifts among affected soil units after rearrangement of ranker and rendzina in the classification key

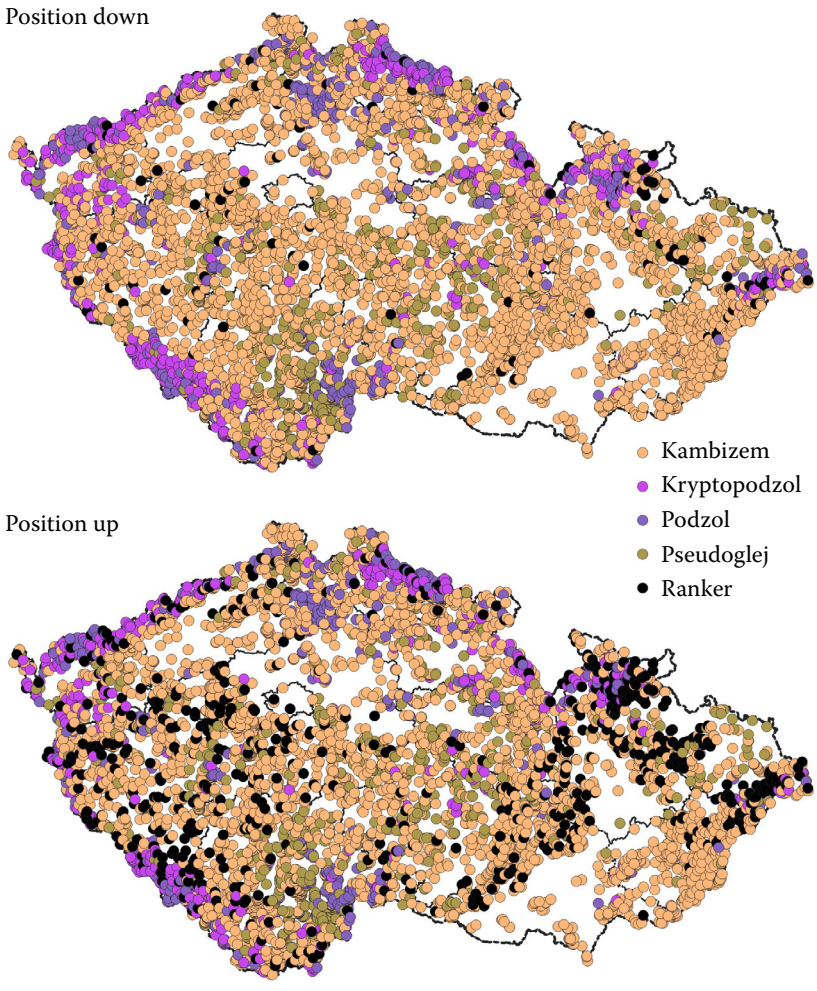



Figure 7. Changes in the spatial distribution of affected soil types after rearrangement of ranker in the classification key

https://doi.org/10.17221/115/2026-SWR

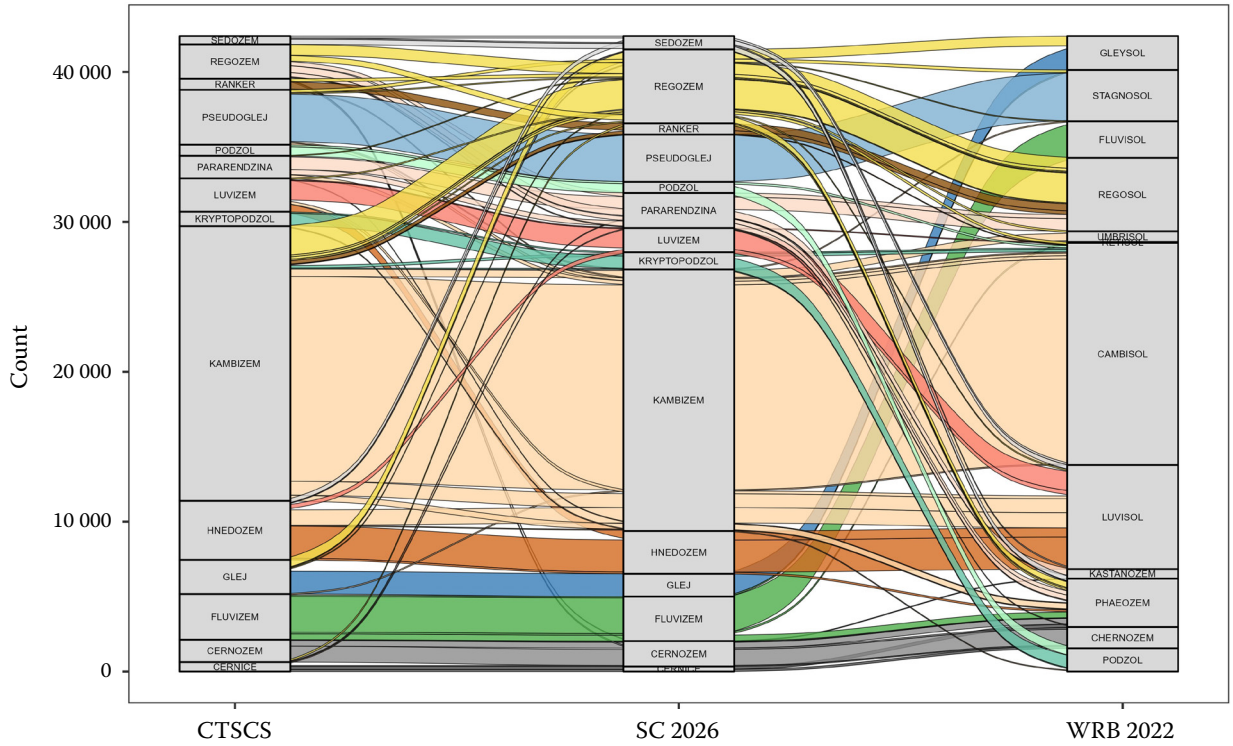



Figure 8. Distribution and shifts among individual soil units in different classifications
CTSCS – Czech Taxonomic Soil Classification System (Němeček et al. 2011); SC 2026 – Soil Classification 2026; WRB – World Reference Base 2022; SOC – soil organic carbon

sified as ranker in the case of their position higher up in the key.

**Taxonomic relationship between soil units in different classifications**

Figure 8 shows changes in the number of soil profiles and their relative proportions within soil classes classified according to various classification systems (CTCSS, SC 2026 and WRB 2022). When comparing the two national classifications, there is considerable agreement among the individual classes (75%). The most significant increase was observed in the case of regozem (50%). This is primarily due to the newly introduced criteria for the thickness of diagnostic horizons. As a result, a number of profiles with thinner diagnostic horizons (e.g., cambic, luvic, and gleyic) now move down the classification key into the regozem group (2900 profiles). Significant shifts are also evident within the groups of illimerised soils, i.e. between hnědozem and luvizem. Some of these soils now belong to the kambizem (20%), which is due to the increased criterion for textural differentiation (1.4 for hnědozem compared to 1.3). When comparing the WRB and SC 2026 classifications, the most notable difference concerns the concept of units with a mollic horizon; in the WRB, these include a wide range of soils with subsoil diagnostic horizons (primarily hnědozem, luvizem, kambizem), but also a significant number of fluvizem and regozem. This is due to the emphasis that the WRB generally places on topsoil characteristics, in contrast to the Czech classification, which tends to prioritise subsoil processes (Reyes Rojas et al. 2025). The newly defined criteria have also significantly reduced the number of Czech černozem classified as Kastanozems in WRB – a common phenomenon in national classifications (Kabała et al. 2019b; Zádorová et al. 2021; Reyes Rojas et al. 2025) – because they have significantly aligned the concept of chernic horizons across both classifications. Thus, the Kastanozem RSG now primarily includes illimerised soils (mainly šedozem) and, in some cases, pararendzina with developed secondary carbonates.

The taxonomic relationships among the individual groups, classified in CTCSS, SC 2026, and WRB 2022 are analysed by the Mahalanobis taxonomic distance.

https://doi.org/10.17221/115/2026-SWR

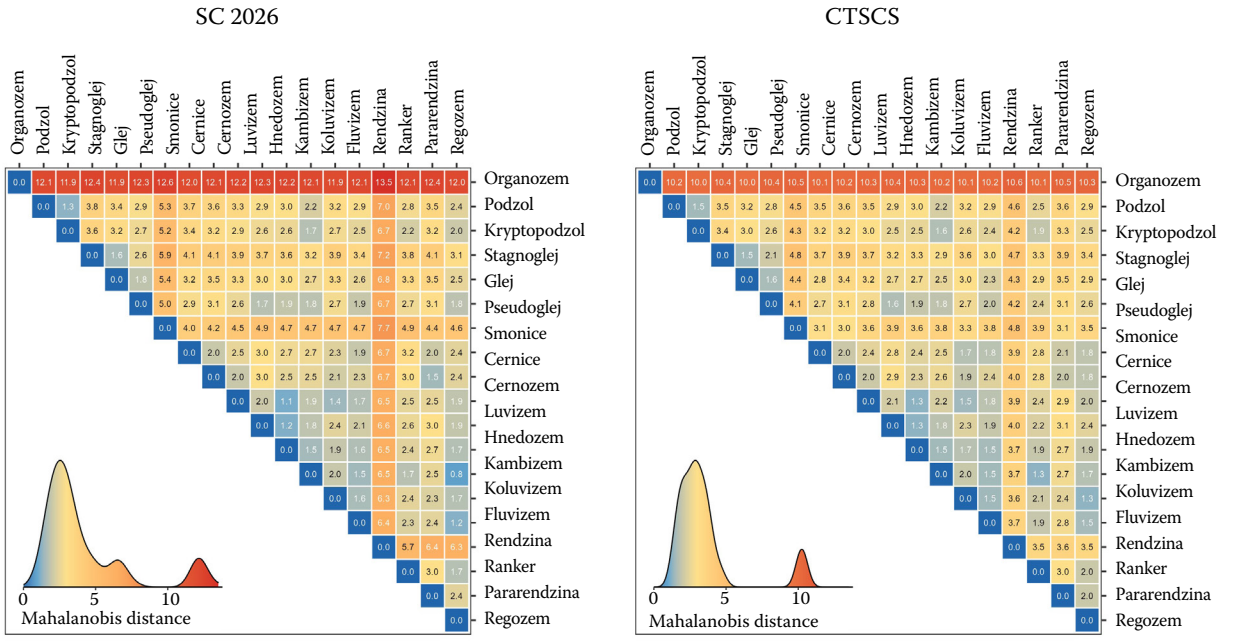



Figure 9. Taxonomic distances between the soil types in Soil Classification 2026 (SC 2026; left) and Czech Taxonomic Soil Classification System (CTSCS; right)

Figure 9 shows the taxonomic distances among the taxa within the SC 2026 and CTCSC to illustrate the impact of the change in classification scheme and criteria on the within-taxon homogeneity of individual soil groups. In principle, the taxonomic distances between individual soil types are similar in both classifications; thus, the new criteria do not significantly affect the consistency of the individual classes. In the case of černozem and černice, their distance from regozem has been emphasised, while they have become more closely aligned with pararendzina. This is due to a change in the concept of pararendzina, which now also include soils on unconsolidated carbonate substrates (e.g., loess) and thus approach the concept of these soils in other Central European classifications (Kabała et al. 2019a; AG Boden 2024). In the case of organozem, smonice and stagnoglej, there was a moderate increase in the distances from other soil types. Conversely, regozem became less distant to the other groups. This is related to its position at the very bottom of the classification system, where it encompasses all soils that, for various reasons, do not meet the criteria of the soils ranked at higher positions in the classification key but may nevertheless exhibit similar characteristics (Michéli et al. 2016). The significant increase in taxonomic distance between rendzina and other soil types in the SC 2026 is likely due to the very small number of profile (49) that are classified as rendzina according to the new criteria (newly introduced threshold of rock fragments content).

The exceptional similarity between certain taxa raises questions about the reasonableness of distinguishing them at a high taxonomic level. This is particularly true of the podzol-kryptopodzol and hnědozem-luvizem pairs. In such cases, it is important to understand that the selection of identifiers, which is determined by the conceptual framework of the analysis (whether it is horizon-based or interval-based) and the availability of source data, significantly influences the taxonomic distance result (Michéli et al. 2016; Sorokin et al. 2021). For instance, one of the key identifiers that distinguishes between podzol and kryptopodzol is bulk density, but this was unavailable for the analysis. In the case of the luvizem-hnědozem pair, there is no identifier indicating the presence of an albic eluvial horizon, which is one of the criteria for distinguishing between these two classes. The result, therefore, does not only reflect the similarity of these soils, but also the selection and availability of identifiers.

The internal homogeneity of individual groups is also reflected in the taxonomic distances between the studied classifications (Figure 10). More heterogeneous soil types with lacking strong differentiation criteria, such as regozem, kambizem, and fluvizem, are

https://doi.org/10.17221/115/2026-SWR

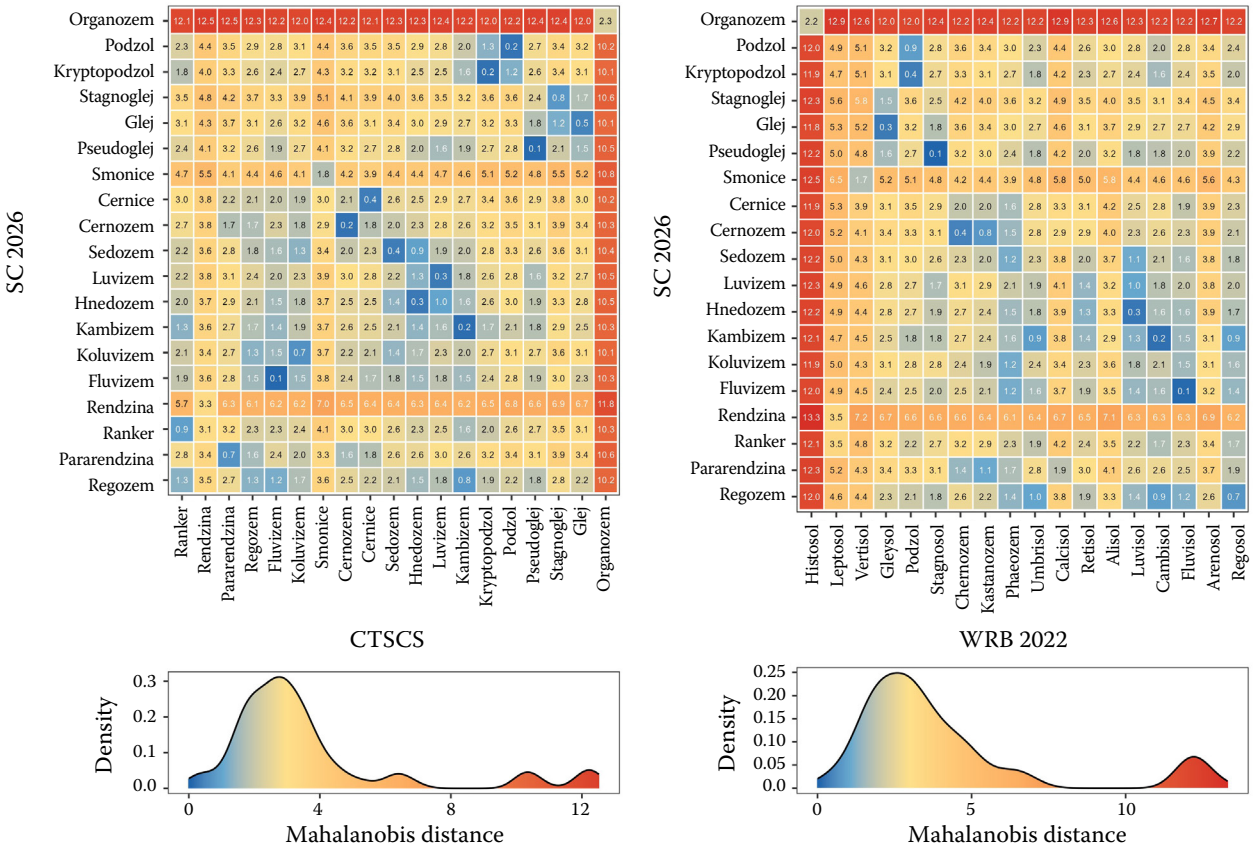



Figure 10 Taxonomic distances between the soil types in Soil Classification 2026 (SC 2026) and the Czech Taxonomic Soil Classification System (CTSCS; left), and soil types in SC 2026 and reference soil groups in World Reference Base 2022 (WRB 2022; right)

closer to a larger number of soil groups. In contrast, more distinctive soil types, such as glej and podzol, have only a small number of closely related counterparts. However, with regard to the taxonomic distance between soil types in both national classifications, the nearest distance between respective soil types (i.e. černozem-černozem, kambizem-kambizem) has been maintained despite changes in the concepts and criteria for individual soil classes. This is important for transferring between the two national classifications and is particularly relevant when linking the classification to existing soil maps, databases, and soil quality evaluation systems.

The taxonomic distances between soil units in SC 2026 and WRB 2022 are influenced by the differing concepts underlying the individual classes. In the WRB, a key factor is the identification of several RSGs with diagnostic topsoil horizons, which are not present at the soil type level in the Czech classification. These soil units then exhibit a number of characteristics specific to other groups as well (e.g., clay illuviation, soil weathering) and are characterised by taxonomic proximity to several different soil types (e.g., Phaeozem is close to černozem, fluvizem, kambizem, hnědozem, and regozem). In contrast, glej, podzol, and kryptopodzol have only one close counterpart. When we compare the results of the presented centroid-based approach with the findings of Zádorová et al. (2021) – in which the taxonomic distance between the Czech and WRB classifications was determined using a concept-based approach (i.e., based solely on the definitions of individual classes) – the results are similar, with only minor deviations in the order of the closest counterparts. This means that the theoretical setting of individual diagnostic criteria and their actual expression in the field are relatively consistent. In contrast, Láng et al. (2013) compared the taxonomic distances between WRB 2006 soil classes and brown forest soils in Hungary, and for some soil groups (brown earths, luvic soils), the two approaches (centroid and concept-based) resulted in completely different closest counterparts.

https://doi.org/10.17221/115/2026-SWR

This was explained by the insufficient representation of certain specific identifiers and the considerable diversity of soils within the units.

## CONCLUSION

The study presents various approaches to optimizing national soil classification using a large database of legacy data and quantitative methods, as illustrated by case studies.

These quantitative approaches enabled more precise and objective threshold value settings for individual diagnostic criteria without significantly altering the consistency and homogeneity of soil taxonomic groups compared to the existing classification.

Using the key to soil types enables precise classification of all soil types, eliminating ambiguous classifications. However, it is associated with the problem of a 'waste' class – that is, the final soil group in the key, which comprises any soil that does not fit into any other category, resulting in a lack of internal homogeneity.

The establishment of quantitative diagnostic thresholds in this study illustrates a broader challenge inherent in revising long-established soil classification systems: balancing empirical rigor with continuity of existing classifications and their applications. This challenge is illustrated by the chernic horizon threshold in černozems, where strict adherence to traditional criteria (40 cm thickness, ~1.2% SOC) would exclude a substantial proportion of profiles historically mapped as černozems, reflecting long-term anthropogenic modification of humus horizons. The working thresholds adopted here (30 cm, ≥ 1% SOC) preserve closer continuity with previous soil maps and applications such as land evaluation. At the same time, calibrating thresholds with reference to the observed distribution of legacy soil units may make the classification less responsive to gradual changes in soil conditions. More broadly, soil classifications are never entirely context-free: diagnostic thresholds inevitably reflect the primary purpose of the system, whether emphasizing continuity with legacy surveys, representation of progressive or regressive pedogenetic sequences, expression of soil degradation, or support for land-use decisions.

Taxonomic distance can serve as a useful guide in final decisions regarding the diagnostic criteria setting and organization and selection of individual taxa representing the soil cover at a high taxonomical level. Nevertheless, it is important to recognize that a given calculation always strongly depends on the choice of identifiers and the number of units used in the calculation.

Once the basic concept, structure, diagnostic criteria and key of the classification have been established, the next essential step is to conduct extensive field testing to evaluate the usability and relevance of the new classification. It is also crucial to have the classification tested and revised by other relevant user groups, i.e. professionals involved in soil survey and mapping, soil evolution research, soil bonity assessment or forest site evaluation, and students, since acceptance of any new soil classification by the broader pedological community is the key to its subsequent active use.

## REFERENCES

AG Boden (2024): Bodenkundliche Kartieranleitung. KA6. Schweizerbart, Hannover.

Arnold R.W. (2002): Soil classification principles. In: Michéli E., Nachtergaele F.O., Jones R.J.A., Montanarella L. (eds.): Soil Classification 2001. ESBR Report No. 7, EUR 20398 EN, Luxembourg: 3–8.

Batjes N.H. (2016): Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. Geoderma, 269: 61–68.

Brunson J.C. (2020): ggalluvial: Layered Grammar for Alluvial Plots. Journal of Open Source Software, 5: 2017.

Brunson J.C., Read Q.D. (2026): ggalluvial: Alluvial Plots in 'ggplot2'. R package version 0.12.6. Available at http://corybrunson.github.io/ggalluvial/

FAO, ISRIC, ISSS (1998): World Reference Base for Soil Resources. Rome, FAO.

Ginzky H., Pieper S. (2025): Soil Legislation by the European Union: An EU "Soil Monitoring Law" on the horizon? In: Ginzky H. (ed.): International Yearbook of Soil Law and Policy 2025. Cham, Springer: 411–433.

Hraško J., Linkeš V., Němeček J., Novák P., Šály R., Šurina B. (1991): Morphogenetic Soil Classification System of Czechoslovakia. Bratislava, VÚPÚ. (in Slovak)

Hughes P., McBratney A.B., Huang J., Minasny B., Michéli E., Hempel J. (2017): Comparisons between USDA Soil Taxonomy and the Australian Soil Classification System I: Data harmonization, calculation of taxonomic distance and inter-taxa variation. Geoderma, 307: 198–209.

IUSS Working Group WRB (2006): World Reference Base for Soil Resources 2006. Rome, FAO.

IUSS Working Group WRB (2022): World Reference Base for Soil Resources 2022. Rome, FAO.

Kabała C., Charzynski P., Chodorowski J., Drewnik M., Glina B., Greinert A., Hulisz P., Jankowski M., Jonczak J.,

https://doi.org/10.17221/115/2026-SWR

Łabaz B., Łachacz A., Marzec M., Mendyk Ł., Musiał P., Musielok Ł., Smreczak B., Sowinski P., Switoniak M., Uzarowicz Ł., Waroszewski J., (2019a): Polish soil classification, 6th edition – principles, classification scheme and correlations. Soil Science Annual, 70: 71–97.

Kabała C., Charzynski P., Czigany S., Novak T.J., Saksa M., Switoniak M. (2019b): Suitability of world Reference Base for soil resources (WRB) to describe and classify chernozemic soils in Central Europe. Soil Science Annual, 70: 244–257.

Kučera M. (2016): Czech Republic. In: Vidal C., Alberdi I., Hernández Mateo L., Redmond J. (eds): National Forest Inventories. Cham, Springer: 307–325.

Łabaz B., Hartemink A.E., Zhang Y., Stevenson A., Kabała C. (2024): Organic carbon in Mollisols of the world – A review. Geoderma, 447: 116937.

Láng V., Fuchs M., Watner I., Michéli E. (2013): Soil taxonomic distance, a tool for correlation: As exemplified by the Hungarian brown forest soils and related WRB reference soil groups. Geoderma, 192: 269–276.

Láng V., Fuchs M., Szegi T., Csorba A., Michéli E. (2016): Deriving World Reference Base Reference Soil Groups from the prospective Global Soil Map product – A case study on major soil types of Africa. Geoderma, 263: 226–233.

Láng V., Fuchs M., Gelsleichter Y.A., Dobos E., Michéli E. (2024): An iterative algorithmic approach for optimizing a diagnostic soil classification system. In: Centennial Celebration and Congress of the International Union of Soil Sciences 2024. Book of Abstracts: 138134.

Michéli E., Fuchs M., Láng V., Szegi T., Szabóné Kele G. (2014): Methods for modernizing the elements and structure of the Hungarian Soil Classification System. Agrokémiai és Talajtan, 63: 69–78.

Michéli E., Láng V., Owens P.R., McBratney A., Hempel J. (2016): Testing the pedometric evaluation of taxonomic units on soil taxonomy – A step in advancing towards a universal soil classification system. Geoderma, 264: 340–349.

Michéli E., Csorba A., Szegi T., Dobos E., Fuchs M. (2019): The soil types of the modernized, diagnostic based Hungarian Soil Classification System and their correlation with the World Reference Base for Soil Resources. Hungarian Geographical Bulletin, 68: 109–117.

Minasny B., McBratney A.B. (2007): Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. Geoderma, 142: 285–293.

Němeček J., Damaška J., Hraško J., Bedrna Z., Zuska V., Tomášek M., Kalenda M. (1967): Survey of Agricultural Soils in Czechoslovakia. Vol. 1. Praha, MZVŽ. (in Czech)

Němeček J., Macků J., Vokoun J., Vavříček D., Novák P. (2001): Czech Taxonomic Soil Classification System. Praha, ČZU. (in Czech)

Němeček J., Mühlhanselová M., Macků J., Vokoun J., Vavříček D., Novák P. (2011): Czech Taxonomic Soil Classification System. 2nd Ed. Praha, ČZU. (in Czech)

Reyes Rojas J., Margoldová A., Skála J., Žížala D., Penížek V., Hrdlička V., Vaněk A., Zádorová T. (2025): Chernozems in Czechia: Soil properties, classification and environmental controls. Geoderma Regional, 40: e00957.

Societas pedologica slovaca (2014): Morphogenetic Soil Classification System of Slovakia. 2nd Ed. Bratislava, VÚPOP. (in Slovak)

Soil Survey Staff (2022): Keys to Soil Taxonomy. 13th Ed. USDA – Natural Resources Conservation Service.

Sorokin A., Owens P., Láng V., Jiang Z.-D., Michéli E., Krasilnikov P. (2021): "Black soils" in the Russian soil classification system, the US soil taxonomy and the WRB: Quantitative correlation and implications for pedodiversity assessment. Catena, 196: 104824.

Van Huyssteen C.W., Michéli E., Fuchs M., Waltner I. (2014): Taxonomic distance between South African diagnostic horizons and the World Reference Base diagnostics. Catena, 113: 276–280.

Venables W.N., Ripley B.D. (2002): Modern Applied Statistics with S. 4th Ed. Springer.

Zádorová T., Žížala D., Penížek V., Vaněk A. (2020): Harmonisation of a large-scale historical database with the actual Czech soil classification system. Soil and Water Research, 15: 101–115.

Zádorová T., Skála J., Žížala D., Vaněk A., Penížek V. (2021): Harmonization of a large-scale national soil database with the World Reference Base for soil resources 2014. Geoderma, 384: 114819.

Žížala D., Minařík R., Skála J., Beitlerová H., Juřicová A., Reyes Rojas J., Penížek V., Zádorová T. (2022): High-resolution agriculture soil property maps from digital soil mapping methods, Czech Republic. Catena, 212: 106024.

Received: August 25, 2026
Accepted: October 8, 2026
Published online: October 29, 2026