On soil districts

Alexandre M.J.-C. Wadoux, Léa Courteille, Dominique Arrouays, Lucas de Carvalho Gomes, Jérôme Cortet, Rachel Creamer, Einar Eberhardt, Mogens Greve, Erik Grüneberg, Roland Harhoff, et al.

Geoderma 452 (2024) 117065

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Discussion paper

On soil districts

Alexandre M.J.-C. Wadoux [a,*], Léa Courteille [a], Dominique Arrouays [b], Lucas De Carvalho Gomes [c], Jérôme Cortet [d], Rachel E. Creamer [e], Einar Eberhardt [f], Mogens H. Greve [c], Erik Grüneberg [g], Roland Harhoff [f], Gerard B.M. Heuvelink [h,i], Ina Krahl [j], Philippe Lagacherie [a], Ladislav Miko [k], Vera L. Mulder [h], László Pásztor [l], Silvia Pieper [j], Anne C. Richer-de-Forges [b], Antonio Rafael Sánchez-Rodríguez [m], David Rossiter [i,n], Bastian Steinhoff-Knopp [o], Stefanie Stöckhardt [j], Gábor Szatmári [l], Katalin Takács [l], Maria Tsiafouli [p], Tom Vanwalleghem [m], Nicole Wellbrock [q], Johanna Wetterlind [r]

[a] LISAH, Univ Montpellier, AgroParisTech, INRAE, IRD, L'Institut Agro, Montpellier, France
[b] INRAE, Info&Sols, 45075 Orléans, France
[c] Department of Agroecology, Aarhus University, Viborg, Denmark
[d] CEFE, Université Paul-Valéry Montpellier 3, Université de Montpellier, CNRS, EPHE, IRD, Montpellier, France
[e] Soil Biology Group, Wageningen University & Research, Wageningen, The Netherlands
[f] Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Germany
[g] Thünen Institute of Forest Ecology, 16225 Eberswalde, Germany
[h] Soil Geography and Landscape Group, Wageningen University & Research, The Netherlands
[i] ISRIC – World Soil Information, Wageningen, The Netherlands
[j] Soil state and soil monitoring, German Environment Agency (UBA), Dessau-Rosslau, Germany
[k] Institute for Environmental Sciences, Faculty of Science, Charles University, Prague, Czech Republic
[l] Institute for Soil Sciences, HUN-REN Centre for Agricultural Research, Hungary
[m] Department of Agronomy, University of Córdoba, Córdoba, Spain
[n] Section of Soil & Crop Sciences, School of Integrative Plant Sciences, Bradfield Hall, Cornell University, Ithaca, NY 14853, USA
[o] Coordination Unit Climate Soil Biodiversity, Thünen-Institute, 38116 Braunschweig, Germany
[p] Department of Ecology, School of Biology, Aristotle University of Thessaloniki, Thessaloniki, Greece
[q] Thünen–Institut für Waldökosysteme (WO), Eberswalde, Germany
[r] Department of Soil and Environment, Swedish University of Agricultural Sciences (SLU), Skara, Sweden

ARTICLE INFO

Handling Editor: Alberto Agnelli

Keywords:
Soil Monitoring Directive
Soil health
Soil quality
Threats to soil
Pedo-ecological regions
Soilscapes
Degradation
Sustainable soil management
World reference base for soil classification

ABSTRACT

In 2023, the European Commission released a legislative proposal for a Directive on Soil Monitoring and Resilience which aims to define a legal framework to achieve healthy soils across the European Union (EU) by 2050. A key component of the initial Directive is the mandate for Member States to establish basic geographic soil governance units, referred to as soil districts, and appoint a district-specific authority to oversee the implementation of soil health assessments. This paper proposes an operational definition of the districts following the conditions outlined in the proposal for the Directive and discusses various attention points for their implementation. Tentative districts were developed for seven EU countries, considering soil type, climate, topography, and land cover factors, starting from the smallest existing administrative unit (i.e. municipalities). Experts were asked to report on the applicability of the proposed districts within well-known pedo-ecological regions and discuss the relevance of the districts for establishing an EU-wide monitoring network and reporting on soil health and degradation. The outcomes highlight the need for detailed soil maps to account for specific soil types when stratifying countries into soil districts. The soilscape approach allows for a consistent method to defining soil districts across Member States. This enables contrasting soils within a district to be managed in a similar manner, with soil degradation/health thresholds applied to each district based on land cover. However, it is unclear whether soil districts as currently formulated in the Directive are in fact the right tool to support local soil management and monitoring of soil health. Districts can help ensure that all soil conditions are covered in a monitoring system, but they may not provide support for soil management or monitoring at a

* Correspondence to: Laboratory of soil-agrosystem-hydrosystem interaction (LISAH), 2 place Pierre Viala, 34090 Montpellier, France.
E-mail address: alexandre.wadoux@inrae.fr (A.M.J.-C. Wadoux).

https://doi.org/10.1016/j.geoderma.2024.117065

Available online 9 November 2024

local scale due to short-scale soil variability and threats affecting soil management within the same soilscape. Beyond the use of districts for designing a European/national scale monitoring system, the districts can help create animations and other educational tools to promote soil literacy and connectivity of users to soils locally.

1. Introduction

The European Commission released a legislative proposal in 2023 for a Directive on Soil Monitoring and Resilience (European Commission, 2023), hereafter referred to as the Soil Monitoring Directive. This is a proposal for a legal framework to achieve that all soils in the European Union (EU) are healthy by 2050, in line with the EU Soil Strategy for 2030 (European Commission, 2021) and the European Green Deal (European Commission, 2020). It is currently estimated that between 60 and 70% of European soils are unhealthy, leading to a loss of ecosystem services estimated to cost more than €50 billion per year (Veerman et al., 2020). The objective of the Soil Monitoring Directive is to reverse this trend and to provide soils with the same level of legal protection in the EU as for the air (Directive 2008/50/EC) and water (Directive 2000/60/EC). To achieve this, ambitious targets are proposed: (i) to provide a common definition of soil health; (ii) to put in place an EU-wide coherent soil monitoring framework; and (iii) to lay down rules for sustainable soil management and the remediation of contaminated sites.

One of the novelties of the initial version of the Soil Monitoring Directive is the establishment of basic soil governance units referred to as "Soil Districts" and defined as "*Part of the territory of a Member State, as delimited by that Member State for the purposes of soil health assessment and management*". Soil Districts are geographical entities seeking for homogeneity in soil types, climatic conditions and land use or land cover. The Directive mandates Member States to establish soil districts throughout their territory, and where an appointed district-specific authority will oversee the implementation of soil health monitoring and reporting (Articles 4 and 5). Soil districts will further complement the existing LUCAS Topsoil survey, which is an EU-wide soil monitoring network (Article 6). The Directive aspires to make districts a unit of reference to monitor efforts in soil health improvement, land take, and for the implementation of soil policy in the EU. While some updates arose in the revision in June 2024, we use hereafter the initial version of the Directive and do not differentiate the districts from the so-called soil units (see Council of the European Union, 2024).

In the scientific literature, however, a definition of what constitutes a soil district does not currently exist. Since the districts are yet to be defined, but could consider soil types, climate and land use or land cover, we may find multiple interpretations or strategies for their implementation across Member States. The soil districts as currently formulated in initial proposal of the Soil Monitoring Directive (see definition above) broadly link to existing conceptual soil entities characterized by homogeneous soil forming factors (e.g. the pedogenon (Román Dobarco et al., 2021) and terron (Carré and McBratney, 2005)) and soil entities characterized by homogeneous local spatial patterns of soil forming factors (e.g. soilscapes, (Hole, 1978) and the small pedological region (Favrot, 1989)), among others. While varying largely in target extent, scope and relevance to management, they represent a geographical area that has homogeneity in soils or soil-landscape relationships defined by a quantitative set of variables representing climate, organisms, topography and parent material grouped together to form relatively homogeneous areas relevant for soil assessment.

A similar definition seems to be proposed in the Soil Monitoring Directive, where the objective is to define districts seeking homogeneity in the following variables: (1) soil type according to the World Reference Base for Soil Resources (WRB), (2) climatic conditions; (3) environmental zones defined in Metzger et al. (2012); and (4) land use or land cover (see European Commission, 2023). Beyond the use of these variables, the initial version of the Soil Monitoring Directive established additional constraints. District borders should not be drawn such that they split administrative units and the number of districts in any Member States must be greater or equal than the number of NUTS 1 territorial units in that Member State. Finally, it is not required but assumed that the districts should be spatially contiguous and ideally be geographically compact.

The first aim of this work is to provide an operational definition that meets the requirements of the initial proposal of the Soil Monitoring Directive. Our second aim is to implement and apply this operational definition to seven example countries to illustrate how it works and learn from that exercise. Our third aim is to interpret and reflect on our initial results and put forward a proposal for a methodology to create districts for the whole EU, which can be applicable to the revised version of the Soil Monitoring Directive where criteria for the definition of soil districts are slightly different.

2. Conceptual framework and pilot studies

We define districts as soilscape units, where a soilscape represents a high-level grouping of soils based on their relationships with the landscape. A soilscape has been described as a *landscape unit including a limited number of soil classes that are geographically distributed according to an identifiable pattern* (Lagacherie et al., 2001), and similarly, as a *cluster of polypedons* (Hole, 1978). A polypedon itself is a collection of pedons: individual soil units that share similar characteristics, forming the basis for defining a soil series or a mapping unit in soil surveys. As a multi-polypedonic entity, a soilscape encompasses sequences of similar soil types and typically associated diverging soils. This similarity in soil types within a soilscape makes it conducive to similar management strategies and policy implementations. More information on this choice is provided in the Discussion.

A methodology was developed to map districts across each Member State, the details of which can be found in the Supplementary Materials. This process starts with the collection of maps categorizing soil types according to the WRB taxonomy, along with maps of land cover, climate (i.e. precipitation and temperature), and topography (i.e. elevation). We started from the smallest existing administrative unit in the EU comprising the municipalities. For these units, we extracted distributions of the aforementioned variables. These distributions were then used to compute a distance metric between units, for categorical variables taking into account both the taxonomic distance between soil types and between land cover classes and the relative area occupied by these classes within each unit. For continuous-numerical variables such as precipitation, temperature and elevation, this distance was calculated between distributions of the variables within the unit. All distances were standardized using the mean and standard deviation of the country and aggregated. They served as the basis for an agglomerative clustering approach which groups units together based on hard spatial contiguity and soft compactness criteria. The weight assigned to the various distances computed during the agglomeration as well as the determination of the number of districts were defined by the experts of the Member State.

In seven Member States, we implemented the methodology, allowing experts to determine the weighting between distances and the number of districts. A range for the final number of districts was suggested for each country (refer to the Supplementary Material), providing experts with options to choose from. Two main scenarios were provided. The first is an agglomeration of municipalities using a weighted combination of WRB soil type, climate, land cover and topography, where the weights were equal to 1, 0.5, 1 and 0.5, respectively.

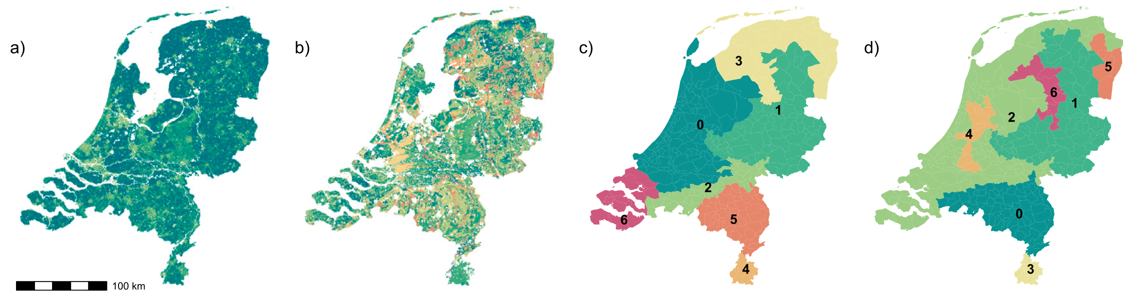

Fig. 1. Map of the Netherlands showing (a) the land use LGN 2022, (b) the Dutch soil map 1:50.000, and 7 districts for (c) the first scenario based the weighted climate, land cover, topography and soil type factors, and (d) the second scenario using soil type in the WRB taxonomy only.

This allows giving a lower weight to the combination of climate and topography which were strongly correlated in our case. In the second scenario, only the WRB soil type (Reference Soil Group, RSG) map is used in the definition of the districts, so the other weights were all set to zero. Experts were also allowed to opt for a custom set of weights. Subsequently, we requested that the experts provide descriptions in 500 words of the resulting districts, focusing on three key elements: (i) pedological/soil geography description that assesses the district's relevance in relation to well-known pedo-ecological regions; (ii) evaluation of the relevance of the variables used for defining the districts, along with suggestions for variables that might be more appropriate in their respective countries; and (iii) analysis of the district's relevance concerning one or more important soil threats and functions: salinization, soil erosion, loss of soil organic carbon (SOC), topsoil compaction, subsoil compaction, excess nutrient content in soil, soil contamination, reduced water holding capacity, acidification, loss of biodiversity, land take, and soil sealing.

2.1. The Netherlands

Clustering with contiguity enforcement was applied to the 342 municipalities, from the minimum four (corresponding to the number of NUTS-1 divisions) to 12 clusters, using the WRB soil type with and without climate, land cover and topography. The latter elements hardly affected the results. Climate has a weak gradient west to east (drier and warmer) also corresponding to a slight topographic rise from the North Sea inland, but these trends are also reflected in the WRB RSG. General land cover (crops, grassland, nature areas) also largely follows the RSG (Fig. 1).

Seven districts from scenario 2 (Fig. 1d) were chosen by expert opinion based on resemblance to known agroecological and soil-geomorphic zones. We describe these with the 12 province names, as to our knowledge these are most likely the preferred administrative units for soil monitoring in the Netherlands.

Districts 1 (eastern Overijssel, the Veluwe and Achterhoek areas of Gelderland, the Drents plateau) and 0 (North Brabant, north Limburg) are dominantly Podzols developed on coversands, with areas of Arenosols on push moraines. District 3 (south Limburg) is almost entirely Luvisols on a hilly terrain. District 4 (parts of Utrecht, North and South Holland) contains mostly Histosols, which are mostly positioned below sea level. District 2, the largest, includes Fluvisols and Cambisols from river sediments and marine clays, as well as dune sands along the coast (Arenosols). The district covers most of Flevoland, all of Zeeland, large parts of North and South Holland, the river districts of Gelderland and the marine border areas of Friesland and Groningen. District 5 is the fen area of northeastern Netherlands (Drenthe and Groningen), dominantly Histosols. These same soils are dominant in District 6, the fen area of northwestern Overijssel, the Northeast Polder of Flevoland, and southwestern Friesland.

In terms of soil health functions and indicators, the three (partially mined) Histosol districts could be merged (Districts 4, 5 and 6), as could the two Podzol districts (Districts 0 and 1), to reduce the number of districts to four, if there were no requirement for contiguity. The Fluvisols/Cambisols district should be split into zones, one dominated by river sediments and one dominated by sea clays. This was not identified in the clustering but has large implications for soil health and associated threats. All Provinces, except Zeeland, contain more than one district. Consequently, monitoring soils at the province level would require applying different indicators to the different districts in their administrative remit.

In terms of threats to soil, soil erosion by water is a concern only in south Limburg. Excess nutrients and groundwater pollution are considered a major problem throughout the Netherlands but especially in the intensively-farmed coversand regions. Acidification is a problem also in forested areas on Arenosols and coarse-sandy Podzols. Subsoil compaction is a problem in river sediment and marine clay regions. The latter also are threatened by salt water intrusions. In recently reclaimed areas, especially Flevoland, land subsidence is a problem. The river sediment areas are also threatened by flooding and heavy metals pollution. Peat shrinkage and land subsidence are the major threats in the fen areas. Many areas are threatened by soil sealing due to urban and infrastructure expansion.

The Netherlands has substantial but widespread areas of Technosols, which are not dominant enough to form a district. These have special soil health and soil threats, and should be accounted for in the Directive. Another peculiarity of the Netherlands is the strict water management, implemented by powerful water boards. Soil health indicators related to threats to water quality and movement across and through the landscape must be defined and monitored, this again in all the districts.

2.2. Denmark

The clustering process was applied at the municipality ($n = 99$) level from the minimum 2 to 12 clusters using the WRB RSG with and without climate, land cover and topography. Climate has a weak gradient west to east (lower precipitation), with croplands primarily in the east and grasslands predominantly in the west. Experts found that 5 soil districts formed with only the WRB soil type (i.e. the second scenario, see Fig. 2) better represented the Danish landscapes. Incorporating climate and land cover did not improve the delineation of districts compared to a delineation solely based on WRB soil type.

The soil districts shown in Fig. 2 represented well the known geological/agroecological zones in Denmark. District 0 (West of Jutland) is a fluvial plain dominated by glacial fluvial sands and the main soil type is Podzols. However, the south part of this district is dominated by moraine landscape with loamy till and the main soil type is Luvisols. This district could be merged with District 1 (East of Jutland), which is also dominated by Luvisols in a moraine landscape with loamy till deposits. District 2 (North of Zealand) is dominated by moraine landscape with sandy till deposits and the main soil type is Podzols. District 3 (Northwest and North of Jutland) is dominated by moraine landscape with loamy till and Luvisols. There is also the eolian sand along the Northwest coast. North of Jutland could hold a unique soil

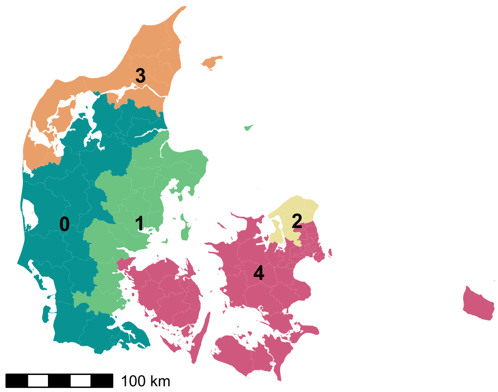

Fig. 2. Map of Denmark showing 5 districts for the second scenario with the agglomeration using WRB soil type.

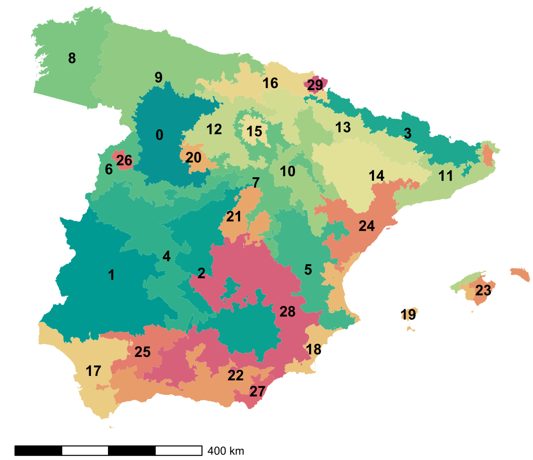

Fig. 3. Map of Spain showing 30 districts for the first scenario based the weighted WRB soil type, climate, land cover and topography factors.

district since it is a mixture of sand till and late glacial marine deposits with Alisols as the main soil type. District 4 is dominated by moraine landscape with loamy till and the soil type is Luvisols.

Soil compaction is the main soil threat across all soil districts in Denmark (Schjønning et al., 2009). District 0 is characterized by higher values of soil organic carbon but also higher nitrogen leaching. In contrast, the Luvisols in Districts 1 and 4 present a lower soil organic carbon content. District 4, dominated by moraine landscape and intensive agriculture, is also more exposed to tillage and water erosion. The intensity of soil threats varies by district and each one may need to focus on the issues causing the most significant economic or environmental impacts. For instance, nitrogen leaching is a major concern in District 0 but not in the southeast of District 4.

2.3. Spain

The municipalities and communes were clustered into between 30 and 146 districts. The classification algorithm using only WRB soil type (i.e. the second scenario) did not yield good results. The precision of the large map scale of soil type (1:1 000 000 map of WRB) seemed not appropriate to differentiate soil districts. Variation in climate, topography and land cover for the agglomeration with the first scenario had an significant impact on the results. We opted for a reduced number of soil districts (i.e. 30 districts) obtained by using all factors: WRB soil type, climate, topography, and land cover and considered it adequate to obtain representative soil districts in Spain (Fig. 3). Increasing the number of soil districts offered no benefit, as it only resulted in the creation of small polygons.

For the selected scenario, the districts fit well the geography of Spain considering the main units of the Iberian geology (Pyrenees, Cantabrian Range, The Spanish Central System, Iberian Cordillera, Betic Cordillera, Sierra Morena), proximity to water bodies (coastal areas and marshes), forests and agroforestry systems and even certain agricultural areas. For example, in the northwest of Spain (Galicia and Asturias) the Umbrisols is the dominant soil type but there are two different soil districts (8 and 9) due to the influence of the topography (Cantabrian Range) and land cover. Furthermore, in the west part of the Spanish Central System, dominated by Calcisols and Leptosols developed on calcareous materials and clay deposits, there are at least six different soil districts (2, 5, 6, 7, 21 and 28) that reflects a combination of topography, climate, and land cover. Calcisols are also prevalent in the Ebro valley and southern Andalusia. However, these areas have entirely different settings, ranging from fertile agricultural areas (olive trees in the south and rainfed farming systems in both areas; Districts 7, 10, 25, 28) to desert areas (Ebro valley, District 14). This is well represented with the current district delineation.

Land cover seems to play a fundamental role in most districts. In the south, for example, the Andalusian countryside (Districts 17, 25 and 28; Calcisols, Vertisols and Fluvisols), mainly dedicated to permanent crops (olive trees) and crop rotations, is framed to the north by Sierra Morena (Districts 1, 2 and 4; Regosols), where livestock farming systems are developed, and to the south by the Cordilleras Béticas, with a higher altitude and where permanent natural vegetation and permanent crops are common (District 22; Cambisols and Leptosols).

In terms of soil threats, soil erosion is probably the most important in Spain. The soil districts represent well the spatial variability in this threat compared to the Spanish national soil erosion map (MITECO, 2022) or the European soil erosion by water map (Panagos et al., 2021). For example, in Southern Spain the central Guadalquivir valley with intensive olive cultivation has some of the highest soil erosion rates > 50 t ha^{1} yr^{1}. This area is covered by three different districts (28, 25 and 17) which represent well an east–west gradient in erosion rates. On the other hand, soil districts just north of this area (i.e. 1, 2 and 4), in the Sierra Morena, have much lower erosion rates below 2 t ha^{1} yr^{1}. In spite of having high slopes, soils in this area are well protected by their stoniness and vegetation cover.

The high salinity areas of Spain, located in the south-west and east of the Iberian Peninsula (mouth of the Guadalquivir rivers, Doñana National Park, the Ebro delta, and the coastal area from Valencia to Almería; Districts 17, 18, 24 and 27), are often associated with rice paddies, intensive agriculture (including greenhouses) and marshlands.

2.4. France

Clustering was applied at the municipality level (i.e. about 35,000 municipalities), for a number between 30 to 150 clusters and the two scenarios (Fig. 4). None of the resulting clusters were fully satisfactory from an expert-based opinion: clustering based on WRB soil type alone for a small number of clusters resulted in large and inconsistent clusters (i.e. Districts 2, 3 and 10 in the WRB 30 districts map, Fig. 4c). When the number of clusters exceeded 120, the delineation did not bring new relevant information and seemed to add noise to the map. Geomorphological variables, extreme pedological conditions, and more detailed climatic variables could have been added to improve clustering. In addition to the WRB soil type, maps of soil properties could also enhance the delineation of the districts.

With a large number of clusters (i.e. larger than 100), the scenario with WRB soil type as clustering variable was efficient to delineate specific soilscapes (e.g. maritime marshlands Solonetzs with District 59, Podzols in South-West France with District 46 in Fig. 4d). The map with 30 clusters including all variables mostly reflected climate and topography and a few relevant specific soilscapes (e.g. Vosges in District 2, Alsace in District 23, Champagne in District 24, Jura in

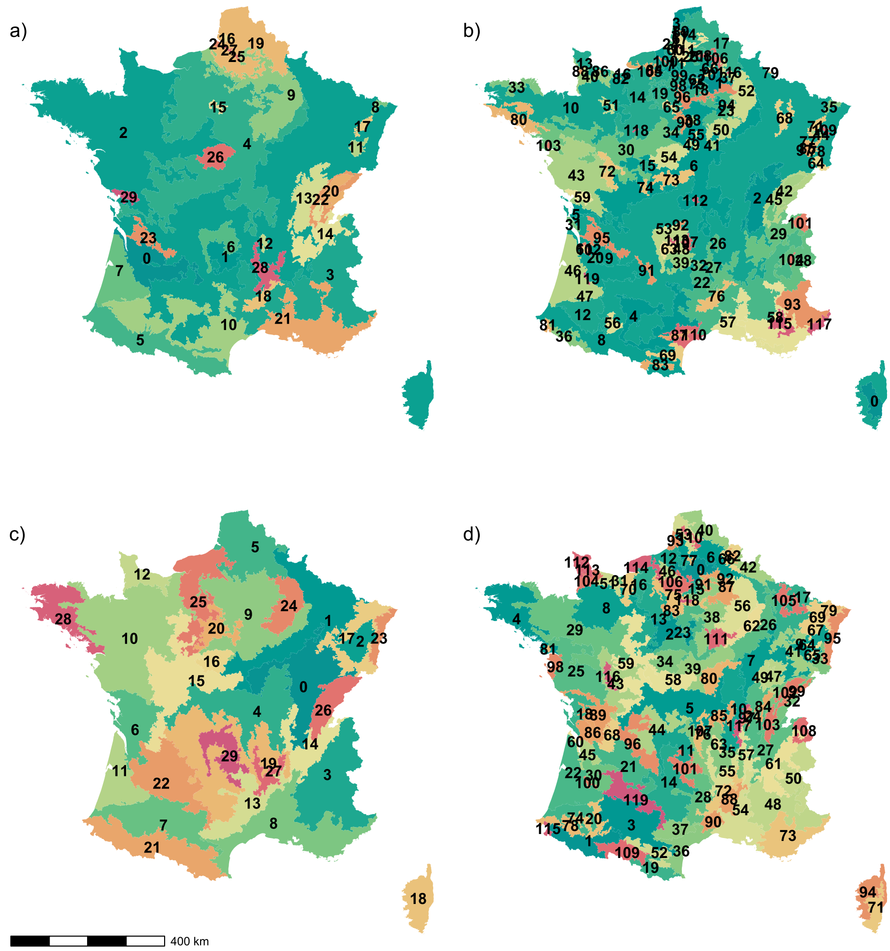

Fig. 4. Maps of France showing 30 (left) and 120 (right) districts for the first scenario using WRB soil type, land cover, elevation and climate (a–b) and the second scenario based on WRB soil type (c–d).

District 26, Landes de Gascogne in District 11). Except for forests, the land cover appeared to have a rather low influence on district delineation.

The delineation best reflecting pedological/soil geography was obtained using 120 districts and including all variables (i.e. the first scenario). It is shown in Fig. 4b This clustering allowed to discriminate some small soilscapes having very specific soil types, land covers and topography combinations. Below 120 districts, the maps showed a loss of pedological meaning for specific districts.

Considering the 120 districts map and the scenario with all variables, similar indicators (e.g. pH threshold values, metabarcoding of bacteria, fungi, protists and animals, organic carbon stocks) could be used to monitor soil health in the most acid, sandy and forest soils. In terms of soil threats, some depressed coastal zones are clearly characterized by flooding and salinization risk induced by sea level rise, extreme rainfall events and submergence by maritime storms (e.g. Districts 18, 98 and 93). Some districts are especially subjected to erosion risks (e.g. Districts 114, 20, 3, 78, 73, 45, 30, 37, 90 and 33), whereas compaction is a threat in many districts, but difficult to characterize because they are heterogeneous in terms of soil texture and land covers. In addition, districts as currently suggested are not relevant to detect differences induced by soil management, cropping systems and related and mechanical pressure on soil. Excess nutrient content in soil is especially a threat in Brittany (District 4 and the western part of 29). It may also be a threat to groundwater in the Alsace region (i.e. District 95) and other intensively cultivated districts (e.g. Districts 2, 13, 75, 38, 87 and the north of District 23). No clear districts grouping appeared to be relevant according to contamination except for copper in some vineyards (e.g. Districts 45, 30, 36 and some districts along the Rhône and Saône valleys). The districts located in mountain areas are threaten by landslides, valley flooding and by loss of SOC under climate change. Very few districts are relevant for soil sealing except for small clusters close to Paris. Note, however, that in the case of the map of 120 districts made with the WRB soil type map (Fig. 4b) the cluster 57 located in southern France is highly relevant for soil sealing due to urbanization, although we do not know whether this is indeed reflected in the variables included in the clustering process.

2.5. Hungary

The representation of Hungary's soil cover on internationally available WRB based soil maps is problematic for several reasons. As a consequence, the European scale WRB soil type map proved to be the least relevant variable in a first round of district definition. Introduction of a national WRB soil map (Dobos et al., 2019) just slightly improved the representation of soil in the delimitation process. This is probably due to the fact that the WRB classification system at the RSG level is less suitable to capture the high diversity of Hungarian soils. Land cover, topography and climate, with a custom weighting, proved to be more informative too. By using such weighting, it was possible to capture and delineate soil districts to a certain extent. The most acceptable scenario for Hungarian soil district delineation by expert opinion based on agroecological, soil and landscape zones included 15 soil districts (Fig. 5). Hereafter, we used the landscape region/micro region names for the description of soil districts.

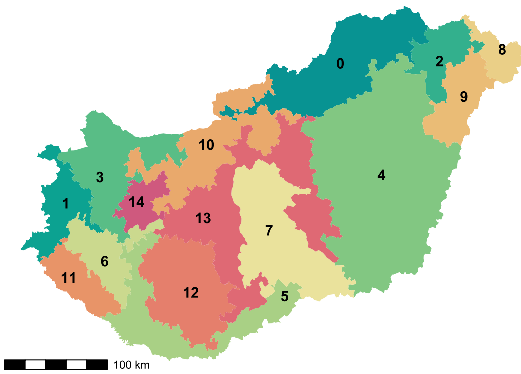

Fig. 5. Map of Hungary showing 15 districts for the second scenario of weighted WRB soil type, climate, land cover, topography factors. Districts are obtained with a custom weighting of all variables.

The Great Hungarian Plain (GHP) consists of 7 districts. The sinking lowland areas and alluvial plains, which material is heavy textured clay, these areas are covered by Vertisols and Gleysols (Districts 8, 2, 5 and 5: Bereg-Szatmár Plain, Zemplén-Szabolcs Plain, Eastern part of the GHP – east from Tisza River –, Lower Drava Plain with East Inner Somogy and Bácska Plain). The tectonically uplifted areas are covered by loess and Chernozem soils (District 13: Mezőföld and eastern part of Danube-Tisza Midland Ridge) or by sand and Arenosols (Districts 7 and 9: Danubian Plain, the western part of Danube-Tisza Midland Ridge and Nyírség). Bácska Plain and Kőrös-Maros Midland are also covered by loess and have Chernozem, but they did not appear separately by the clustering, they were assigned to the alluvial plain districts. The hilly area of East Inner Somogy should be separated from district 5 and merge with district 12.

The Little Hungarian Plain (District 3) is well delineated, where mainly Chernozem soils have formed on loess and loess-like sediments. The low mountainous area of Eastern Alpine Foreland with Luvisols and the alluvial plain of Vas-Sopron Plain with Gleysols (District 1) would be more appropriate to be divided. On the contrary, the Zala Hills, where Luvisols and Stagnosols are characteristic, are separated into 2 districts (Districts 6–11), owing to the differences in precipitation, but they could be combined. The hilly area of Outer Somogy, Tolna Hills, Mecsek and Baranya Hills (District 12) is well delineated. This area is rather heterogenous in terms of topography and climate, resulting in a variety of soils, predominantly Luvisols and Gleysols being characteristic. The Transdanubian range, where Luvisols and Cambisols are the most dominant soils, are assigned to Districts 10 and 14. These districts should be merged and shrunk, where the parent material changes from carbonate to volcanic rocks. The North Hungarian Range (District 0) has very complex geology, the whole range of volcanic stones, carbonate rocks and loess also occur resulting in Leptosols, Cambrisols and Alisols. District 0 is well delineated on the eastern part, but on its western side it should be extended until the parent material changes from volcanic to carbonate.

Areas with salt-affected soils do not appear separately in either scenarios of soil district delineation. This fact is well explicable due to the variables used in delineating districts, because it is difficult to capture salinization and/or sodification as salt-affected soils appear as mosaic like patches in the Hungarian lowlands due to its geological and hydrogeological settings. Hilly and mountainous areas are well recognizable when topography together with land cover is involved and properly represented in the delineation process (Districts 0, 1, 6, 10, 11, 12 and 14). Since soil erosion essentially occurs on arable lands, horticultures and vineyards situated on steep slopes (Pásztor et al., 2016), exposed areas are clearly separated along the derived zones. Some of the districts captured well those areas where either loss of SOC or accumulation of SOC was observed in the past decades (Szatmári et al., 2021). Additionally, they also represent well areas with large SOC deficit meaning potential for additional SOC sequestration (Szatmári et al., 2023).

At the countrywide level, spatially exhaustive information on soil compaction is scarce. However, areas with intensive agricultural activity and at risk of compaction are well delineated by the districts due to the spatial information on land use/land cover. Excess nutrient content in soil and soil contamination being local soil threats, they could be hardly related to any zone structure. Reduced water holding capacity is essentially related to (hydro)physical properties and management practices. The former should and could be represented in the delineation process by the involvement of the available spatial information on physical soil properties (texture, fractions, Laborczi et al., 2019) and water management features (water retention, saturated hydraulic conductivity etc. Szabó et al., 2024). The latter is partly represented by the land cover information, though local differences cannot be taken into account in the present approach.

In the last decades risk of soil acidification decreased due to certain changes in its driving forces, which caused less attention on its spatial assessment and mapping. No spatial assessment or data is available on the occurrence of biodiversity loss, thus its representation in the derived zonalization cannot be evaluated. Land take and soil sealing are dynamic and essentially locally occurring related to urbanization, consequently they might not be considered in the present district approach.

2.6. Germany

The more than 11,000 unequally sized local administrative units of Germany were clustered. Experts decided to opt for 20 contiguous clusters (Fig. 6), numbers around the minimum number (i.e. the 16 NUTS-1 entities) and below the spatially discontinuous 33 soil regions of EUSRM5000 (Baritz et al., 2005), but more than the 174 contiguous soil region entities of the German Soil Landscapes Map (BGR, 2023) that form 12 discontinuous soil regions (integrating geology and topography). Even though there is a notable climatic gradient from west (suboceanic) to east (subcontinental) and a clear altitudinal topographic and climatic differentiation, land cover did not significantly contribute to a better delineation due to its often mosaic-like patterns. The approach using WRB soil type alone largely meets soil regions and is discussed in the following, although there are considerable inconsistencies along most natural boundaries. This is likely due to the use of administrative units in the clustering approach whose boundaries often do not run along the natural boundaries.

Compared to the Soil Regions/Great Soilscapes delimited on the basis of geology and topography, the increase from 20 to 40 soil districts does not seem to improve the homogeneity of districts in a way that justifies the much higher efforts for monitoring. Even the 20 districts could be further combined by soil inventory similarity: (1) the three districts along the Baltic Sea Coast could be combined with the adjacent districts in Schleswig-Holstein (SH) NE of Hamburg (HH) and two in Mecklenburg-Vorpommern (MV) plus the almost individual part of the largest district E of HH, which all encompass mostly Luvisols and Gleysols (or Stagnosols), (2) the large district in the NW (Northrhine-Westfalia - NRW, Lower Saxony - NI) reaching far into the NE (Saxony-Anhalt - ST, Brandenburg - BB, Saxony - SN), could also include the central part of SH and the easternmost part of the following district (its BB and SN parts), is composed of mostly Podzols, Gleysols, current or drained Histosols and – in the east – Retisols, (3) the loess district (NRW, NI, Thuringia – TH, SN and reaching south into Hessia – HE) has foremostly Luvisols and some Phaeozems (more important than the Chernozems). Skipping the contiguity requirement, the Luvisol district in western NRW, the district around Frankfurt (HE), the one in northern Rhineland-Palatinate (RP) and eastern TH could be added here. (4) The large Cambisol/Podzol district in the south-central part (including parts of Saarland – SL, S' NRW, RP, Baden-Wuerttemberg

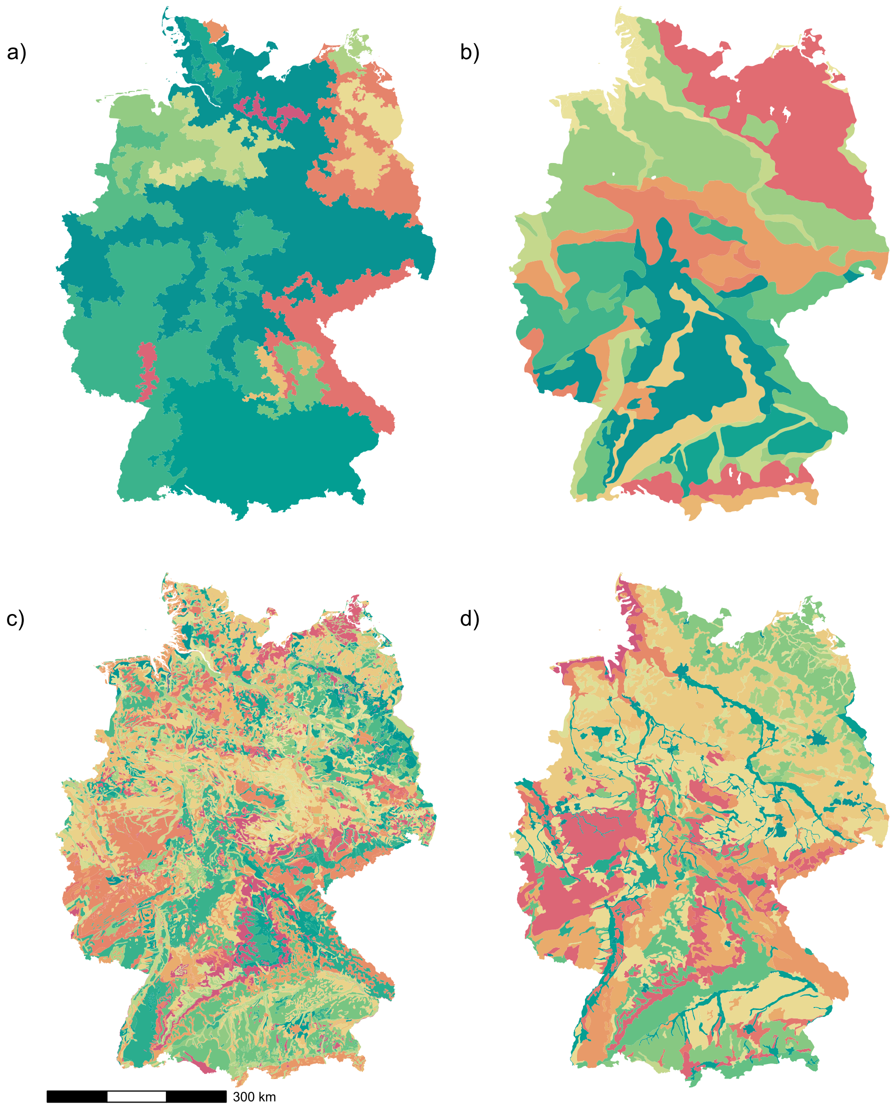

Fig. 6. Maps of Germany showing (a) 20 districts for the second scenario using WRB soil type, (b) the soil region entities of the German Soil Landscapes Map, (c) the soil map of Germany at 1:1,000,000 and (d) the groups of soil parent material.

– BW, Bavaria – BY, TH and SN) could be combined with the district adjacing to the south, but (5) the Swabian and Franconian Alb should be excluded to form a district on their own (with Eutric Chromic Cambisols and Rendzic Leptosols). (6) The district north of the alpine foreland (BW, BY; Cambisols, Luvisols) could stay as it is, as well as (7) the Alps with their foreland (Leptosols, Cambisols). Because soil protection is in the federal states' responsibility, these districts had possibly to be subdivided along federal state boundaries.

Regarding soil health, the same splitting of coastal soils from floodplain soils as in the Netherlands would make sense. Soil erosion by water threatens soils and soil functions in almost all hilly to mountainous areas of Germany, and by wind in particular in the sandy and drained Histosol areas in the northern lowlands, where excess nutrients are a common threat in those subareas with intensive stock-breeding and dairy farming. Subsoil-compaction can occur on almost all loamy soils all over the country under agricultural land cover. Loss of organic carbon occurs in the mostly intensely drained (i.e. former) fens and bogs in the northern lowlands (NRW, NI, SH, MV, BB) as well as in BY. Acidification is a concern in forested areas on Podzols, Arenosols and Dystric Cambisols in the northern lowlands and mid-range mountains, but is also driven by forest stands. The threat of salt water intrusions and land subsidence is the same as in the Netherlands. Heavy metal contamination is, beside locally threatening soils, to be concerned in Fluvisols and other soils on floodplains of smaller rivers rising in former mining and smelting areas. Soil sealing threatens often high-yield soils (Luvisols, Phaeozems, Retisols, Fluvisols), e.g. in the northern and southern loess areas nearby larger cities. Loss of biodiversity is a threat wherever soils are intensely used for agriculture.

2.7. Sweden

Land cover is the most important variable in Sweden for delineating soil districts for monitoring and especially for implementation of actions, with agricultural soil only making up 7% of the land area and forest corresponding to 68%. The variables used in delineating the soil districts are correlated to different degrees, resulting in similar units independent of weighting. However, larger weight on WRB results in fewer clusters in the northern half of the country that is dominated by Podzols. Sweden covers a large latitudinal range with an old mountain range along the border to Norway, and adding especially climate results in more clusters in the north. Ten districts were chosen and described hereafter for the two scenarios (Fig. 7).

Soil type is not commonly used in Sweden related to agricultural soils with WRB soil type largely following land cover patterns with Cambisols dominating the agricultural land. In forest soils, Podzols are

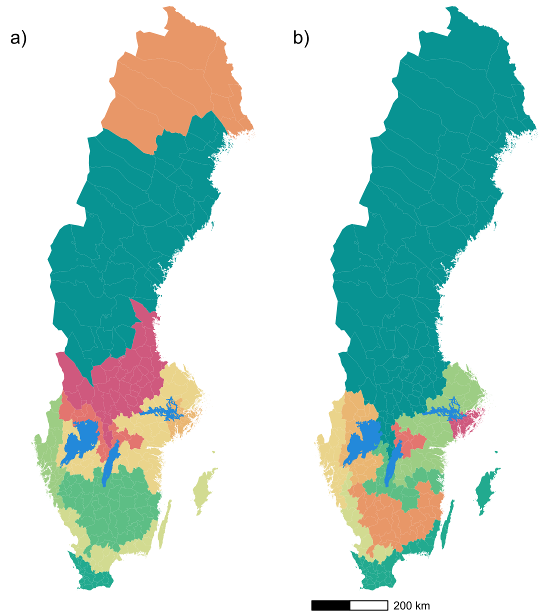

Fig. 7. Map of the Sweden showing 10 districts for (a) the first scenario based the weighted WRB soil type, climate, land cover and topography factors and (b) the second scenario using WRB soil type only.

dominating in the north while Podzols, Arenosols, Leptosols, Regosols and Histosols are common in the south.

The influence of the mountain range, and the climatic east–west gradient in the north, with milder climate along the north east coast, was not easily captured in any of the district delineation, especially not with few districts. However, too many districts resulted in a rather cluttered pattern especially in the southern part of the country, that did not correspond well to any commonly used pedo-ecological regions.

In an agricultural context, Sweden is often divided into 8 production regions. These regions are related to pedo-climatic conditions as well as dominating production forms and production potential. Especially in the southern half of Sweden, the productions regions correlates fairly well with the soil regions described in the "Soil Regions of the European Union and Adjacent Countries" (Baritz et al., 2005). The methodology presented in the Supplementary Material for clustering resulted in districts showing similarities with these production regions. Ten districts using all variables producing the most similar map. However, the delineations are not a perfect match. The clustering delineated the agricultural plains around the lakes and the most southern parts and along the south cost, but with clear irregularities due the borders. Increasing the number of districts did not improve the similarities, indicating that additional information, probably related to agriculture management and production would be needed to increase these similarities.

The relevance of the soil districts to soil health functions and indicators is more difficult to characterize. Many of the threats to soil, for example, are strongly linked to soil management. The threats are also dependent on very local conditions, and varies largely within any proposed soil districts. Studying the variation in some available soil properties in the two Swedish national monitoring systems within the different districts revealed very little reduction within soil districts. Other conditions, such as distance from the cost, lakes or rivers, and production forms such as intense crop or animal production for example, have large impact on soil fertility, degradation and the effects of soil health. With that said, there are regional differences and areas specifically prone to certain threats. These areas are often better captured using the national production regions since they also relate to agriculture production.

However, even these are of course not homogeneous, neither in terms of soil properties nor management. Large enough districts to allow for a substantial amount of soil sampling sites within the soil district without increasing the total amount of national samples would therefore be preferable. Based on this, between 8 and 10 districts, relating also to agriculture and forest production might be reasonable for Sweden. More and smaller districts risk resulting in too few sampling sites within districts or too many sampling sites in total, without really reducing the variation in soil indicators within the districts.

3. Discussion

We delineated soil districts for seven countries of the EU to provide an overview of the feasibility and relevance of developing such geographical units and using them for monitoring and reporting purposes. The main characteristics of the soil districts in the pilot studies are summarized in Table 1. Experts from most countries were positive about the proposed soil district delineation: they found good agreement with well-known pedological regions. The district delineation posed some challenges, which are reflected in the pilot studies in the choice of approaches (i.e., using different sets of variables, either WRB soil type or WRB soil type + covariates), in the choice of the number of districts, and in the size of the soil districts relevant for discriminating threats to soil. We discuss hereafter the challenges that emerged from the methodological framework and regional narratives, and suggest solutions and improvements.

3.1. The WRB soil type map and its limitation

For all of Europe there are soil maps either already classified to the WRB RSG level or national digital soil maps which can be fairly easily correlated with WRB (e.g., the European Soil Geographical Database at 1:1,000,000 scale). The utility of the WRB soil type maps varied significantly depending on the member state. For instance, the WRB soil type map for France is considered outdated and was constructed with limited data, making it less reliable compared to the more detailed 1:250,000 national soil maps, whereas in the Netherlands a more precise national 1:50,000 soil map is available. There is generally a need for sufficiently detailed soil maps to take into account soil specificity when stratifying countries by soil districts. If precise maps are available, there is still some remaining work to conduct to choose the right level of aggregation of these more detailed soils maps, both from a geographical and semantical point of views. In the regional cases, the use of the WRB map did not reveal some expected specific soil type patterns. An example is in France where using WRB only resulted in some very large and sometimes inconsistent clusters, some of them being composed of various Cambisols having very contrasted properties.

The fine-scale national maps of soil types are better suited for incorporating the specific soil types within a region, regardless of whether they are harmonized or adhere strictly to the WRB RSG. Indeed, if WRB is very useful as a common denominator at EU and global level, it is not intended to be a substitute or national soil classification systems (IUSS Working Group, 2022). The primary concern is the accurate representation of soil types and major properties rather than the classification system used. The representation of a soil type in large geographical areas, as in the WRB soil type (RSG) map, poses challenges in addressing the dominance of particular soil types within these classifications and the level of the classification. The Soil Monitoring Directive supports the use of a variable for "soil type", a generic term not specifying a level. In many RSGs some principal qualifiers have large implications for soil management and health, for example in the Luvisols with the Gleyic, Stagnic, Vertic, Calcic, as opposed to the Haplic principal qualifiers. Without a clear understanding of the level to use, the taxonomic distance, which focuses only on the most general class (i.e. RSG), may not adequately reflect the diversity or the specific characteristics of soils in the area. We therefore recommend that it is more advantageous to utilize either the European Soil Geographical Database at 1:1,000,000

Table 1
Summary characteristics for the pilot studies and preferred district zoning options across Europe.

Country	Data input	Number of districts	Average size (km^2)	Accordance with national pedological knowledge	Soil threats discrimination
The Netherlands	WRB soil type	7	5343	good	good
Denmark	WRB soil type	5	8645	very good	fair
Spain	WRB soil type + covariates	30	16 619	good	good
France	WRB soil type + covariates	120	4572	fair	fair
Hungary	WRB soil type + covariates	15	6201	good	weak
Germany	WRB soil type	20	17 861	good	good
Sweden	WRB soil type + covariates	10	44 983	fair	weak

1 WRB soil type or WRB soil type + covariates (i.e. land cover, climate, topography).
2 Weak, fair, good, very good.
3 Weak, fair, good, very good.

scale or national maps that provide a detailed and accurate depiction of soil distributions, supplemented by an elaborate, meaningful and pedologically sound distance matrix calculation. Such distance matrix needs to be adapted to local soil and environmental issues in order to better weight the distances according to the soil classes and their semantic and logical distances regarding critical soil properties. An alternative is to use a soil map that already provides aggregates semantically detailed soil information, as is the case for the "Soil Regions of the European Union and Adjacent Countries" proposed in Baritz et al. (2005) and suggested as one of the variables to define the soil units (here, called districts) in the update of the Soil Monitoring Directive (Council of the European Union, 2024).

3.2. Potential of digital soil mapping

An alternative to conventional polygon-based soil type maps is the use of soil properties maps, which can serve as a more flexible starting point. These maps can be generated through digital soil mapping (DSM) techniques, where various soil-forming factors are integrated and weighted (McBratney et al., 2003) with a model that links point-measured values of soil properties and environmental factors of which maps are available. DSM is particularly adapted at representing unique soil conditions, which are usually influenced by topography, climate, parent material and vegetation. For example, a pH greater than 8.9 typically indicates sodic and salty soils, while a pH range between > 7 and 8.8 is characteristic of calcareous soils. Using soil properties maps instead of using aggregated WRB soil types would make sense if we consider the list of indicators proposed by the Soil Monitoring Directive. Most of these indicators are based on soil properties, sometimes combined as ratios, and accompanied by a proposed list of threshold values.

DSM could further be used to map rare soil types, such as some intensively cultivated shallow soils in the French Beauce (Chen et al., 2021), which are managed with irrigation and pesticides, or regions with micro-peats and bogs. Mapping of rare soils is a recurrent issue in DSM. Although the health of dominant soils within soil districts must be monitored and protected, changes in local soil management and conditions may have a important effects on other environmental issues and ecosystem services.

Overall, we stress that both approaches, i.e. using aggregated WRB soil type maps or soil property maps as one component for delineating soil districts, have pros and cons. On the one hand, it is valuable to use soil properties stable over time and diagnostic horizons as a basis to define geographical soil strata. On the other hand, some RSG may presently cover a wide range of soil properties, and one may want to base the stratification on properties considered presently as baseline of soil health. Similar strategies are discussed when the aim is to stratify soils according to their capacity or to their condition (McBratney et al., 2014), or according to genoforms or phenoforms (Rossiter and Bouma, 2018).

3.3. Which covariates to add to soil data?

Land cover seemed a useful proxy for soil management in the derivation of soil districts, although in some cases land cover was not necessary because it was strongly correlated with other variables. This is the case in the Netherlands where the land cover relates to soil types. This was also the case in Sweden where in the pilot study land cover was seen more informative than the existing map of soil types. The question that also arose is whether to use static land use information or maps that provide history of land use over the past decades, maps of which are now readily available (e.g. Parente et al., 2021), or combination of several sources of land use information that better account for trajectories of land use change (e.g. Levers et al., 2018).

Topography was not in the list of suggested variables proposed in the Soil Monitoring Directive, but we included elevation in our assessment to further define beyond basic soil groups (e.g. Leptosols or Histosols) and processes linked to altitude and controlling factors of soil formation such as low temperature and high amounts of precipitation, which can significantly affect soil organic carbon content and stock (e.g. in the Alpine regions, Wiesmeier et al., 2014; Mulder et al., 2016). While this was deemed sufficient here, in the future the districts could be defined from a basic set of topographic information which could be composed of slope, curvature and a soil water index such as flow accumulation or an index of valley bottom flatness to help distinguishing specific soil conditions and their relationships to water fluxes.

To avoid redundancy, we disregarded the environmental zones defined in Metzger et al. (2012) and suggested by the Soil Monitoring Directive. These zones were too broad to be useful and redundant of the climate variables. Climate was represented by long-term annual averages of precipitation and temperature. Both variables are strongly correlated with elevation and were down-weighted in the agglomerative clustering procedure. For climate variables we would recommend using variables that accounted for climate seasonality and variables that allow seasonal contrasts (e.g. accumulated precipitation minus evapotranspiration, P-ETP) or link to organisms behaviour (e.g. number of vegetation growing days, number of freezing days, length of drought periods, amount of precipitation in the vegetation period). These might be useful in specific climates such as the Mediterranean and in countries with large climate gradients (south-north, for example in Sweden). It is important to include these, as restrictions on agricultural and forest soil management can influence soil health.

In addition, simple ratios such as P-ETP or aridity indices has been shown useful to distinguish major differences in soil condition (see, for example, Vaysse and Lagacherie, 2015). For instance, at global scale, and without human intervention, soil pH is mostly controlled by the water balance (Slessarev et al., 2016) or exchange frequency of soil water (BGR, 2015). This explains why in the definition of soil districts the controlling factors and the indicators of soil health should be related to climatic factors. Such a control by water balance may explain why Bunce et al. (2002) when conducting a comparison of different biogeographical classifications of Europe, found that "*Geology is considered a major factor that needs to be taken into account when developing land classifications in southern Europe because landform and soil parent material play a more major role in defining soil features and consequently life conditions under a Mediterranean climate than under oceanic and continental temperate climates*". Indeed, one missing variable for delineating soil districts in many countries is the lack of a lithological map. A lithological map could inform not only on soil parent material but also on possible fluxes from soil to the regolith and vice-versa.

It is striking to note that among the main controlling factors of soil formation and evolution specified by the Directive, the effect of vegetation is considered only through land use/land cover. It may be preferable to take into account other characteristics of vegetation that are highly climate- and human-dependent. One example is the net primary productivity or climate/remote sensing proxies which may be related to organic carbon inputs to soil or SOC mineralization rates. Other possible inputs to clustering could be statistics on crop yields, manure spreading, or agricultural practices (e.g. Chen et al., 2019; Martin et al., 2021). It may be argued that these variables will change, especially under the influence of climate change and human practices. This is also the case, however, for land cover and for climate, which are both proposed as bases for soil districts delineation in the Soil Monitoring Directive.

In the updated version of the Soil Monitoring Directive, the list of suggested variables to define the soil units is composed only of soil type and land use. While other variables are also suggested, such as climate, environmental zones, and river basins, we stress that more variables need to be included to do more justice to the underlying factors of soil formation. It is also necessary because the list of variables proposed in the revision does not account for local conditions and rare soil types, or are too redundant to be useful (see, for example, the previous paragraph on the use of the environmental zones defined in Metzger et al. (2012)).

3.4. Determining the number and size of soil districts

The minimum number of districts was set in the Soil Monitoring Directive as equal to or greater than the number of NUTS-1 units. We set an arbitrary maximum number equal to the suggested maximum number of districts at EU level proposed in the Annexe 1 of the Soil Monitoring Directive, which we weighted relative to each country area. Overall, a small number of districts would be preferable to reduce administrative burden and facilitate governance, but the homogeneity within the soil districts would decrease. NUTS-1 units are in some EU member states geographically compact. Here NUTS-1 units will have by definition a limited number of major soil types. In some instances NUTS-1 units may include many RSGs, for example Gelderland Province in the Netherlands with large areas of Fluvisols, Arenosols, Podzols, and Cambisols. The same is true for Lower Saxony in Germany reaching from the coast to the Harz Mountains including a large range of different parent materials and soil types.

There is thus a first choice: should soil districts correspond to a of administrative units or should they be geographically-defined as new units without respect to any administrative unit except NUTS-0 (i.e. the countries)? In the first case, there will be large heterogeneity within most districts but benefits in administration of the monitoring and implementation of measures to increase soil health. In the general approach on the update on the Soil Monitoring Directive reached by the Council of EU on the 17th of June 2024, the administration was partly separated from the geographical soil units, opening up for a more centralized administration in the Member States. In our approach, municipalities were agglomerated to form soil districts. While this is a suitable approach to facilitate a district-level governance, this led in some cases to large heterogeneity within districts because within municipalities there can be subareas with very different soils or soilscapes. In the second case, several administrative entities would have to cooperate to manage a given soil district. This second option would require a new governance specifically for soil districts, or cooperative agreements between existing environmental agencies within the administrative units, whereas in the first case this could be done within an administrative unit with no need for new structures. A third option is to manage all soil districts, however defined, at the national level within the existing administrative structure, by the environmental or agricultural ministry of each country. Thus multiple districts but one administration.

We stress that the observed large variations in district sizes (e.g. see the pilot study for France) are not inherently problematic from a conceptual standpoint, provided there is a relative homogeneity within each district. These discrepancies, however, pose practical challenges for soil monitoring and governance. It is suspected that these issues may be partially attributed to the soil type map used as input, which may not accurately represent the diverse conditions across large areas. As explained earlier in the Discussion, this suggests a potential misalignment between the map granularity and the actual soil variation.

The Soil Monitoring Directive says nothing about geographic continuity or compactness. While we assumed that this was necessary for governance, we could do otherwise by defining a number of districts as a national map (at least as many as NUTS-1 units), and associating each geographic location with a district, without any need for contiguity.

3.5. Defining "homogeneity" within soil districts

The Directive on Soil Monitoring and Resilience suggests to "seek homogeneity within each soil district" with regard to soil type, climatic conditions, environmental zones and land cover. However, these factors are not those used to create NUTS-1 or any other administrative units, except by coincidence. The key words here are "seek" and "homogeneity". For the first, it acknowledges that absolute homogeneity even at a general level of the named factors is not possible, but it should be maximized, given other constraints. For the second, it may be that "homogeneity" could refer to a set of contrasting areas that are themselves homogeneous with respect to the factors. This second also allows for a patterned landscape within a district, where contrasting soils are arranged at close proximity, but can be easily identified and therefore different soil health criteria can be applied to each.

In this study, we proposed to define districts in terms of soilscapes (Wadoux, 2024). This means that we seek a stationarity in the relationships between soil and soil forming factors. We found this was a sensible choice to obtain a similar soil organization within the landscape. The alternative is to define districts as areas that are as homogeneous as possible in soil property or soil types so minimize a weighted sum of variances of soil and other properties within a district, where weights are derived from user perspectives or empirically. We deemed this vision not adapted because a substantial part of the soil variation occurs at short distance (Mulla and McBratney, 2001; Lagacherie et al., 2024) and this would lead to a very large and unmanageable number of districts. Indeed, from a practical point of view, delineating districts on the basis of soil property or soil type homogeneity and considering them as the entities to report on soil health may render a statistically sound soil sampling strategy unfeasible. It may even not be possible to find units that are homogeneous in terms of the different soil properties that we want to monitor, as they usually have a strong field-scale variability. In the revision of the Soil Monitoring Directive, the soil unit was introduced as the geographical unit where the sampling will take place, whereas the district becomes the geographical unit relevant for governance. While the size of the soil units may be smaller than currently formulated in this paper, the downside is the substantial increase that would result in the sample size required to monitor all units.

The districts are defined at a scale for which pedologists have developed concepts. In Favrot (1981), for example, the concept of reference areas within a regions was defined. The reference area represents the main soil distribution pattern of the region and can be the basis for detailed soil survey and high-density sampling. This enables operational and economical facility in defining major soil types and their association with soil property values and their thresholds for soil health within a larger region. Similar higher association and soil organization levels were also recognized through by number of pedologists acting in different regions of the world. The US pedologists defined the pedological province as "*a part of a region, isolated and defined by its climate and topography and characterized by a particular group of soil*" (Smeck et al., 1983). English pedologists working in Eastern Africa stated that "*in any one landscape there are only a few kind of terrain,* [which] *recur in association with one another in the landscape to give a more or less regular pattern always in the same interrelations*" (Astle et al., 1969). The Russian pedologist Fridland stated that "*Soil combinations consist of elementary soil areas which are genetically linked to various degrees and which produce a definite pattern in the soil mantle [...]. Multiple spatial repetition of a certain soil combination or several soil combinations alternating in a definite order creates various forms of structures of the soil mantle*" (Fridland, 1974).

3.6. Numerical vs. expert-based approaches

The numerical approach is well-suited for defining the soil districts and has several advantages over the qualitative approach; it is reproducible, easy to revise and update. It could also be automated and is faster to produce than an expert-based classification of soils. The numerical approach also facilitates transparency and acceptability of the soil district delineation, as the means by which the results were computed can be displayed and justified to the end-users and decision-makers. Additionally, the numerical approach allows for the incorporation of specific constraints, such as spatial contiguity, if it is required to have compact districts. Should expert-based maps be available, such as those detailing small natural regions or soil fertility classes, these can be effectively integrated as covariates in the numerical framework.

3.7. Establishing soil health thresholds and reference values

In the current Soil Mission, the EU is investing substantial funding for research to come up with solutions for identifying appropriate soil indicators and methods for benchmarking soil health. We stress that the districts should establish the dominant genoforms of the major soils, that is, the most common expression of the soil-forming factors under the dominant land uses, used as the basis for detailed soil mapping (Rossiter and Bouma, 2018). Then the distribution soil health parameters in these genoforms can be quantified. This could be a reference — not an ideal or pre-modern agriculture state neither a reference on which soil health thresholds must be based. Then, phenoforms within the soil district can be identified as "*persistent, non-cyclical variants of a soil genoform with sufficient physical or chemical differences to substantially affect soil functions*" (Rossiter and Bouma, 2018). This is a somewhat broader concept than "soil health" as currently expressed by indicators, but surely the soil health will vary between genoforms. Further, any change in land use will change the genoform without necessarily resulting in unhealthy soils. Therefore, soil genoforms cannot be the reference to which soil health indicators must be compared. We must make the distinction between a "reference" and a "baseline". There will be, obviously, many soil districts in which no more "original" soil genoform exist, i.e., pre-human or at least pre-intensive land use. There is also no guarantee at all that soil genoforms are the healthiest soil if we consider that the aim is to optimize soil condition in such a manner that soil can provide ecosystem services at the level we are expecting them, without increasing the threats to soils. The range of soil health indicators of soil phenoforms within a district can be considered as the present baseline, whether it is healthy or not. This relates to other concepts in soil science, such as that of soil security (Evangelista et al., 2023) which distinguishes between capacity/capability and condition. While the former can be mapped, the latter should be regularly monitored with target values set for improvement. It is up to soil scientists, together with other experts and end-users, to define what are the target values to reach and maintain, and what are the thresholds values under- or over-which action must be taken. Soil management should not, in theory, result in tending towards these threshold values; in other words a well-designed soil monitoring should be able to detect trends before the thresholds are reached and deleterious impacts happen. This means that soil monitoring should be a tool for soil management. As soil management is in practice decided by local actors such as farmers, among others, one aim of the districts should be to provide a source of local advice to end-users.

3.8. Potential role of soil districts in local soil management

Dividing an area into districts enables the delineation of bodies of soil associations that are locally relevant to creating soil typologies that are easily perceivable and understandable by local users. This was done, for instance, in France with drainage reference areas to help in setting specific norms for the design of drainage systems. It is unclear, however, whether districts as currently formulated are useful for local soil management and monitoring of soil health. The intended district uses should guide their implementation. If the intention with the districts is to make sure that all soil conditions are covered in a monitoring system, large districts are sufficient, but these will not be able to provide support for soil management or monitoring at a more local scale. This is because of the short-range soil variability, and because of threats that affect strongly the subsequent soil management within a same soilscape (e.g. water management, compaction, accessibility, structure, and management that compensate climate conditions).

One important aspect for local assessments, recommendations and regulations is to take into account the possibilities of inputs to soils, including atmospheric inputs, or transfers in the soils and between soils themselves, or between soils subject to different land covers and agricultural systems. Soil health assessment should consider other environmental components and issues such as transfer to groundwater, hydrographic networks and the atmosphere. We must recognize that many variables are missing if the aim of delineating districts is to define a local strategy to favour soil health at the district level.

3.9. Importance of considering end-user engagement

Beyond their use in designing a European or national-scale monitoring system, districts can serve as operational units for implementing sustainable soil management practices through a participatory approach involving multiple stakeholders. This is feasible at the district level only if the soils, climate, and vegetation are comparable, and if the districts are small enough. We argue that using administrative units as the basis for aggregation, as we did in our proposed implementation, may facilitate participatory approaches in the future.

Districts will also help in creating animations and other educational tools that promote soil literacy and connectivity at a local scale. This can engage the community and increase awareness about soil and promote participatory soil research and management and citizen engagement. The Living Labs can serve as example for local participatory approaches. This aim implies that districts should have a reasonable size, and that they should also have a reasonable range of soil phenoforms conditions that are easily distinguishable in the landscape (i.e. soilscapes) so that the end-users (e.g. the farmers) can understand not only why but also where they should apply good practices. This has large implications for the people involved at the district scale. Defining the districts as reporting units may have counter-productive effects on end users. Districts should be the place where qualified advisers can disseminate soil monitoring results, alert end-users on negative trends in soil health, and provide relevant recommendations.

4. Conclusion

We proposed an implementation of the soil districts as recently discussed under Article 4 of the initial proposal of the European Union Soil Monitoring and Resilience Directive. The districts implemented here are geographical entities defined from soil type, climate, topography and land cover with municipalities as smallest building block. Experts were asked to report on the consistency of the proposed districts with well-known pedo-ecological regions and discuss the relevance of the districts for establishing an EU-wide monitoring network and reporting on soil health and threats. From the results of the seven country-scale case studies and the Discussion we draw the following conclusions:

- For all of Europe we have sufficiently accurate climate, topography and land use or land cover variables to define the districts. We need, however, sufficiently accurate soil type maps that can account for local soil specificity and rare soil types. National soil map units or more detailed WRB classes could be used for this purpose but they are not currently available for all EU countries.
- The importance of the variables (e.g. soil type, climate, land cover) in defining the districts were strongly dependent on the national case study. The pilot studies revealed specific decisions made by the experts in the choice of and weighting between variables.
- We need topographic variables to define districts. These topographic variables could be, for example, slope and elevation.
- A small number of districts is preferable for governance at EU level. It was also taken as the preferred option by experts in most of our case studies. A small number, however, can lead to a large heterogeneity within districts. While this is not inherently problematic from a conceptual viewpoint, it may cause problems for monitoring and reporting at local level and the implementation of measures increasing soil health. Better soil type maps may help to do more justice to the diverse soil conditions across large areas.
- We propose that soil districts should be defined in terms of soilscapes, allowing for a patterned landscape within a district with contrasting soils that can however be managed in a similar way depending on land use or land cover.
- The districts should establish the dominant genoforms of the major soils, that is, the most common expression of the soil-forming factors under the dominant land uses, used as the basis for detailed soil mapping. Then the distribution of phenoforms and their soil health parameters related to these genoforms can be quantified.
- It is unclear whether districts as currently formulated are useful for local soil management and monitoring of soil health. Defining districts for the purpose of establishing a monitoring network can ensure that all soil conditions are covered in a monitoring system, but districts will not be able to provide support for soil management (or monitoring) at a more local scale. This is because of the short-scale soil variability, and because of threats that affect strongly the subsequent soil management within a same soilscape
- The 2024 revision of the Soil Monitoring Directive proposed to differentiate the soil districts used for governance and soil units for monitoring and reporting. The results presented in this study showed that similar challenges remain to implement soil districts and soil units in terms of input variables, number and size of the districts and units and their relevance for local-scale soil health reporting and monitoring.
- The districts can help in creating animations and other educational tools that promote soil literacy and connectivity at a local scale.

Further multidisciplinary research is needed at the interface between sciences of soil, economics, social and societal issues and science for policy to help Member States on devising how to conciliate the needs for (i) having rather homogeneous environmental conditions, (ii) delineating large enough districts to implement representative soil sampling, (iii) establishing the most favourable conditions to involve all local actors and (iv) choosing the most relevant structure for the governance of the implementation and reporting of soil monitoring.

In addition, as the ultimate goal is not to report, but to maintain or increase soil health, research should be developed to move from soil properties monitoring to assessing soil functions and modelling ecosystem services at scales that are relevant for decision making.

This will imply to include other environmental components (e.g., ground-water, surface-water, air, vegetation), socio-economic contexts (e.g., presence of agrifood companies, farmers' income) and measurements or proxies of exogeneous inputs and pressure to soil (e.g., nutrient inputs, organic amendments, irrigation, pesticides, atmospheric deposits, traffic loading, local sources of contamination).

Acknowledgements

Alexandre Wadoux has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101059012. We thank Jingyi Huang and Budiman Minasny for providing data to calculate the taxonomic distance between the WRB soil type classes.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geoderma.2024.117065.

References

Astle, W.L., Webster, R., Lawrance, C.J., 1969. Land classification for management planning in the Luangwa Valley of Zambia. J. Appl. Ecol. 6, 143–169.

Baritz, R., Fuchs, M., Hartwich, R., Krug, D., Richter, S., 2005. Soil Regions of the European Union and Adjacent Countries 1: 5,000,000. Technical Report, Version 2.0. Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Online available: https://services.bgr.de/boden/eusr5000.

BGR, 2015. Exchange frequency of water in agricultural soils in Germany. Bundesanstalt für Geowissenschaften und Rohstoffe. WMS, Link: https://services.bgr.de/boden/ahacgl1000.

BGR, 2023. Soil landscapes of the federal Republic of Germany 1:5,000,000. Technical Report, Federal Institute for Geosciences and Natural Resources (BGR), WMS 1.3.0. Online available: https://services.bgr.de/boden/bgl5000.

Bunce, R.G.H., Carey, P.D., Elena-Rossello, R., Orr, J., Watkins, J., Fuller, R., 2002. A comparison of different biogeographical classifications of Europe, Great Britain and Spain. J. Environ. Manag. 65 (2), 121–134.

Carré, F., McBratney, A.B., 2005. Digital terron mapping. Geoderma 128 (3–4), 340–353.

Chen, S., Arrouays, D., Angers, D.A., Chenu, C., Barré, P., Martin, M.P., Saby, N.P.A., Walter, C., 2019. National estimation of soil organic carbon storage potential for arable soils: A data-driven approach coupled with carbon-landscape zones. Sci. Total Environ. 666, 355–367.

Chen, S., Richer-de-Forges, A.C., Mulder, V.L., Martelet, G., Loiseau, T., Lehmann, S., Arrouays, D., 2021. Digital mapping of the soil thickness of loess deposits over a calcareous bedrock in central France. CATENA 198, 105062.

Council of the European Union, 2024. Proposal for a Directive of the European Parliament and of the Council on Soil Monitoring and Resilience (Soil Monitoring Law). Revision. European Commission Brussels, Belgium. Link: https://data.consilium.europa.eu/doc/document/ST-11299-2024-INIT/en/pdf.

Dobos, E., Vadnai, P., Kovács, K., Láng, V., Fuchs, M., Michéli, E., 2019. A novel approach for mapping WRB soil units–A methodology for a global SOTER coverage. Hung. Geogr. Bull. 68 (2), 157–175.

European Commission, 2020. The European Green Deal. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Brussels.
European Commission, 2021. EU Soil Strategy for 2030. Reaping the benefits of healthy soils for people, food, nature and climate. Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions. European Commission Brussels, Belgium.
European Commission, 2023. Proposal for a directive of the European parliament and of the council on Soil Monitoring and Resilience (Soil Monitoring Law). European Commission Brussels, Belgium. Link: https://environment.ec.europa.eu/publications/proposal-directive-soil-monitoring-and-resilience_en.
Evangelista, S.J., Field, D.J., McBratney, A.B., Minasny, B., Ng, W., Padarian, J., Dobarco, M.R., Wadoux, A.M.J.-C., 2023. Soil security-strategising a sustainable future for soil. Adv. Agron. 183.
Favrot, J.C., 1981. Pour une approche raisonnée du drainage agricole en France. La méthode des secteurs de référence. Comptes Rendus de l'Académie d'Agriculture de France, Paris 67 (8), 716–723.
Favrot, J.-C., 1989. Une stratégie d'inventaire cartographique à grande échelle: la méthode des secteurs de référence. Science du Sol 27 (4), 351–368.
Fridland, V.M., 1974. Structure of the soil mantle. Geoderma 12, 35–42.
Hole, F.D., 1978. An approach to landscape analysis with emphasis on soils. Geoderma 21 (1), 1–23.
IUSS Working Group, 2022. World Reference Base for Soil Resources. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps, fourth ed. IUSS, URL http://www.fao.org/3/i3794en/I3794en.pdf.
Laborczi, A., Szatmári, G., Kaposi, A.D., Pásztor, L., 2019. Comparison of soil texture maps synthetized from standard depth layers with directly compiled products. Geoderma 352, 360–372.
Lagacherie, P., Arregui, M., Fages, D., 2024. Evaluating the quality of soil legacy data used as input of digital soil mapping models. Eur. J. Soil Sci. 75 (2), e13463.
Lagacherie, P., Robbez-Masson, J.-M., Nguyen-The, N., Barthès, J.P., 2001. Mapping of reference area representativity using a mathematical soilscape distance. Geoderma 101 (3–4), 105–118.
Levers, C., Müller, D., Erb, K., Haberl, H., Jepsen, M.R., Metzger, M.J., Meyfroidt, P., Plieninger, T., Plutzar, C., Stürck, J., Verburg, P.H., Verkerk, P.J., Kuemmerle, T., 2018. Archetypical patterns and trajectories of land systems in Europe. Reg. Environ. Change J 18, 715–732.
Martin, M.P., Dimassi, B., Román Dobarco, M., Guenet, B., Arrouays, D., Angers, D.A., Blache, F., Huard, F., Soussana, J.-F., Pellerin, S., 2021. Feasibility of the 4 per 1000 aspirational target for soil carbon: A case study for France. Global Change Biol. 27 (11), 2458–2477.
McBratney, A.B., Field, D.J., Koch, A., 2014. The dimensions of soil security. Geoderma 213, 203–213.
McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.
Metzger, M.J., Shkaruba, A.D., Jongman, R.H.G., Bunce, R.G.H., 2012. Descriptions of the European Environmental Zones and Strata. Technical Report, Wageningen, Alterra, Alterra Report 2281, 152 pp.; 12 fig.; 1 tab; 40 ref..
MITECO, 2022. El Inventario Nacional de Erosión de Suelos (INES) 2000–2022. Ministerio para la Transición Ecológica y Reto Demográfico.
Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Martin, M.P., Arrouays, D., 2016. National versus global modelling the 3D distribution of soil organic carbon in mainland France. Geoderma 263, 16–34.
Mulla, D.J., McBratney, A.B., 2001. Soil spatial variability. In: Warrick, A.W. (Ed.), Soil Physics Companion. CRC Press, Boca Raton, FL, pp. 343–373.
Panagos, P., Ballabio, C., Himics, M., Scarpa, S., Matthews, F., Bogonos, M., Poesen, J., Borrelli, P., 2021. Projections of soil loss by water erosion in Europe by 2050. Environ. Sci. Policy 124, 380–392.
Parente, L., Witjes, M., Hengl, T., Landa, M., Brodsky, L., 2021. Continental Europe land cover mapping at 30 m resolution based CORINE and LUCAS on samples. Dataset, v0.1. Zenodo.
Pásztor, L., Waltner, I., Centeri, C., Belényesi, M., Takács, K., 2016. Soil erosion of Hungary assessed by spatially explicit modelling. J. Maps 12, 407–414.
Román Dobarco, M., McBratney, A., Minasny, B., Malone, B., 2021. A modelling framework for pedogenon mapping. Geoderma 393, 115012.
Rossiter, D.G., Bouma, J., 2018. A new look at soil phenoforms–Definition, identification, mapping. Geoderma 314, 113–121.
Schjønning, P., Heckrath, G., Christensen, B.T., 2009. Threats to soil quality in Denmark-A review of existing knowledge in the context of the EU Soil Thematic Strategy. Aarhus University, Faculty of Agricultural Sciences, Denmark.
Slessarev, E.W., Lin, Y., Bingham, N.L., Johnson, J.E., Dai, Y., Schimel, J.P., Chadwick, O.A., 2016. Water balance creates a threshold in soil pH at the global scale. Nature 540 (7634), 567–569.
Smeck, N.E., Runge, E.C.A., Mackintosh, E.E., 1983. Dynamics and genetic modelling of soil systems. In: Wilding, L.P., Smeck, N.E., Hall, G.F. (Eds.), Pedogenesis and Soil Taxonomy, Concepts and Interactions. Elsevier, Amsterdam, pp. 51–81.
Szabó, B., Mészáros, J., Laborczi, A., Takács, K., Szatmári, G., Bakacsi, Z., Makó, A., Pásztor, L., 2024. From EU-SoilHydroGrids to HU-SoilHydroGrids: A leap forward in soil hydraulic mapping. Sci. Total Environ. 921, 171258.
Szatmári, G., Pásztor, L., Heuvelink, G.B.M., 2021. Estimating soil organic carbon stock change at multiple scales using machine learning and multivariate geostatistics. Geoderma 403, 115356.
Szatmári, G., Pásztor, L., Laborczi, A., Illés, G., Bakacsi, Z., Zacháry, D., Filep, T., Szalai, Z., Jakab, G., 2023. Countrywide mapping and assessment of organic carbon saturation in the topsoil using machine learning-based pedotransfer function with uncertainty propagation. CATENA 227, 107086.
Vaysse, K., Lagacherie, P., 2015. Evaluating digital soil mapping approaches for mapping GlobalSoilMap soil properties from legacy data in Languedoc-Roussillon (France). Geoderma Reg. 4, 20–30.
Veerman, C., Correia, T.P., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E.A., Grand, A., Filchev, L.H., et al., 2020. Caring for soil is caring for life: ensure 75% of soils are healthy by 2030 for healthy food, people, nature and climate: interim report of the mission board for soil health and food: study. Technical Report, Luxembourg: Publications Office of the European Union.
Wadoux, A.M.J.-C., 2024. Soil health by 2050: respecting 'soilscapes' is the key. Nature 626 (7997), 33.
Wiesmeier, M., Barthold, F., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Schilling, B., Angst, G., von Lützow, M., Kögel-Knabner, I., 2014. Estimation of total organic carbon storage and its driving factors in soils of Bavaria (southeast Germany). Geoderma Reg. 1, 67–78.