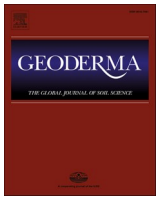
Geoderma 454 (2025) 117199

Contents lists available at ScienceDirect

Geoderma

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



Empirical estimation of saturated soil-paste electrical conductivity in the EU using pedotransfer functions and Quantile Regression Forests: A mapping approach based on LUCAS topsoil data

Calogero Schillaci [a,*], Simone Scarpa [a], Felipe Yunta [a], Aldo Lipani [b], Fernando Visconti [c], Gábor Szatmári [d], Kitti Balog [d], Triven Koganti [e], Mogens Greve [e], Giulia Bondi [f], Georgios Kargas [g], Paraskevi Londra [g], Fuat Kaya [h], Giuseppe Lo Papa [i], Panos Panagos [a], Luca Montanarella [a], Arwyn Jones [a]

[a] JRC European Commission, Ispra, Italy
[b] University College London, United Kingdom
[c] CSIC Centro de Investigaciones sobre Desertificación-CIDE Valencia, Spain
[d] Institute for Soil Sciences, Centre for Agricultural Research, HUN-REN, Budpest, Hungary
[e] Agroecology Dept. Aarhus University, Denmark
[f] TEAGASC, Dublin, Ireland
[g] Department of Natural Resources Management and Agricultural Engineering, Agricultural University of Athens, Athens, Greece
[h] Department of Soil Science and Plant Nutrition, Faculty of Agriculture, Isparta University of Applied Sciences, Isparta, 32260, Türkiye
[i] University of Palermo, Italy

ARTICLE INFO

Handling Editor: Jingyi Huang

Keywords:
Electrical conductivity
$EC_{1:5}$
EC_e
Pedotransfer function
LUCAS soil survey
Predictive modelling
Quantile regression forest
World Reference Base

ABSTRACT

Soil Electrical conductivity (EC) is a measure of the ability of soil to conduct an electric current, which is primarily influenced by the concentration of soluble salts in the soil solution that takes place principally through water-filled pores. Ions (Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, and NH^{4+}, SO_4^{2-}, Cl^{-}, NO_3^{-}, and HCO_3^{-}) from soluble salts dissolved in soil water carry electrical charges and conduct the electrical current. EC is considered a proxy of soil salinity and other soil characteristics, whose monitoring is much needed in the context of climate change, increasing irrigation in agricultural areas and sea level rise. The pan-European LUCAS soil monitoring scheme, established in 2009, provided $EC_{1:5}$ in the topsoil (0–20 cm) in the surveys of the years 2015 and 2018 for almost 20,000 samples. In this work, using the LUCAS 2018 dataset, we provide an empirically-derivedpedotransfer function to convert diluted $EC_{1:5}$ to saturated EC_e using the LUCAS soil texture and soil organic carbon, and a framework for EC_e mapping with a machine-learning algorithm named Quantile Regression Forest. The final model resulted in an R^2 of 0.302 with an RMSE of 0.265 dS m^{-1} for soil samples not used for model calibration. The results are presented as predicted EC_e in the topsoil, and they reveal that in Atlantic and Northern Europe, salts may accumulate in soils through several natural processes, i.e., primary salinization, but in Mediterranean and Southern Europe, they accumulate because of human interventions on the soil water and solute regimes. Among these interventions, seawater intrusion into coastal aquifers, irrigation with waters containing soluble salts, poor drainage of irrigated lands and of naturally saline soils, stand out. Hotspot analysis per country or Nomenclature of Territorial Units for Statistics (NUTS 0) revealed high topsoil EC_e levels occurred in Spain 0.11 %. Increasing EC_e can led to constrained crop productivity in irrigated farming. With this assessment, we try to determine the hotspots for future monitoring and understanding the main drivers for sustainable soil management. Future challenges for EC_e mapping that need to be address are sample numerosity and depth and availability of a consistent set of EC_e data measured to provide a regression based PTF for the use of diluted $EC_{1:5}$.

* Corresponding author.
E-mail address: calogero.schillaci@ec.europa.eu (C. Schillaci).

https://doi.org/10.1016/j.geoderma.2025.117199
Received 28 June 2024; Received in revised form 28 January 2025; Accepted 29 January 2025
Available online 8 February 2025

1. Introduction

The projected availability of land for sustainable food production in the future is marked by considerable uncertainties and variations; as the World's population is expected to increase, there is a need to assess land resources and provide tools for precise and rapid soil monitoring. Expectations are that a considerable portion of this increase will be sustained by crop production (FAO, 2017). Among land degradation and natural pressures, such as drought, extreme weather, and soil contamination by excess nutrients and pollutants, soils are under severe threat and risk of reducing their potential to deliver ecosystem services (Schillaci et al., 2023). Soil salinity is one of the most critical abiotic stressors on crops, affecting 20 % of the World's total cultivated area and causing about 20 % yield losses (Ivushkin et al., 2019). Soil electrical conductivity (EC) can be used as a proxy of dissolved salts in the soil pore water, which are made of Na^+, Mg^{2+}, Ca^{2+}, SO_4^{2-}, Cl^-, HCO_3^- and less often CO_3^{2-} ions, whereas in agricultural lands, there are also NO_3^- and K^+ ions that significantly contribute to the total EC (Visconti & Paz, 2016). EC is recognized as soil parameter that can influence crop growth (Allbed et al., 2014; Tedeschi et al., 2023) and limit the yield (Alkharabsheh et al., 2021; Kaya et al., 2022). The United States Department of Agriculture (USDA) Natural Resources Conservation Service developed technical sheets that summarize the effect of EC on crop (https://www.nrcs.usda.gov), this serves as a resource for understanding the importance of soil EC in various agricultural and environmental applications, including irrigation management, soil fertility assessment, and land classification. Other widely recognized drivers of EC are organic matter (Wong et al., 2010), gypsum (Moret-Fernández & Herrero, 2015), high Ion-exchange capacity (Barbanti et al., 2018; Marzaioli et al., 2010), heavy soil texture (Demir et al., 2019), nutrients (e.g., nitrate and chloride) (Curci et al., 2020; Mechri et al., 2007), water-holding capacity (Marzaioli et al., 2010), and drainage conditions (Alexander, 1980; Vittori Antisari et al., 2020). Salts dissolved in soil water split in ions that carry electrical charges and conduct the electrical current all the more intensely the greater their concentrations. Consequently, the concentration of ions determines the EC of soils and thus the EC has been used in agriculture as an indicator of soil salinity. EC is expressed in deciSiemens per meter (dSm^{-1}) because this unit makes the EC values to mostly range between 0.1 and 10 in agriculture. Apart from the genetic lithological endowment, climate also affects EC. For example, rainfall leaches out soluble salts, especially from topsoil, reducing the amount of mobile ions. Additionally, soil EC increases as the clay content increases because not only are the ions in the soil pore water mobile but the exchangeable ions too (Kim & Park, 2024). Soils with clay dominated by Tetrahedral, Octahedral, Tetrahedral (TOT) type layered phyllosilicate minerals (e.g., smectite and vermiculite) can lead to high cation-exchange capacity (CEC) and therefore an higher EC than those with clay dominated by TO type, low CEC and clay minerals (e.g., kaolinite). Therefore, to avoid the interference of the soil clay, the soil EC is evaluated on an aqueous soil extract. Soils in semiarid and arid regions (with low precipitation levels) are often characterized by high levels of soluble salts and exchangeable sodium percentage, which in turn give rise to extremely high EC. In soils where the water table is high and saline, water rises by capillarity and increases salt concentration, and thereby EC, in the soil surface layers. The soil EC increases with temperature by about 2.2 % per Celsius degree. Therefore, to avoid the interference of temperature, soil EC measurements are corrected to 25 °C. Salinization is recognized among the soil threats identified in the proposal for the Soil Monitoring Law of the European Union (EC COM (2023) 416). However, no continental scale assessment for salt affected soils has been carried out since the compilation of the map "Salt Affected Soils in Europe" by Szabolcs (Szabolcs, 1974). After that, the Joint Research Centre published a map of salinity and sodicity (Fig. 2 in supplementary materials) based on the European Soil Database (ESDB) (https://esdac.jrc.ec.europa.eu). Although the ESDB provides information in greater taxonomic detail about the types of salt-affected soils occurring in the continent, the spatial resolution of the ESDB (1:1M) remains as that of the map of Szabolcs (1974), which was compiled on a scale of 1:1M and published on the scale of 1:5M. In the FAO World saline soil framework the salinity is assessed by combining thee sub-indicators: i) Saturated Paste Electrical Conductivity (EC_e), ii) Exchange Sodium Percentage (ESP) and pH. The present study is aiming at producing the EC_e spatial distribution in EU topsoil.

As the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) underlined: "As halting and reversing land degradation requires location-specific solutions and multi-sectoral collaboration, global and/or regional decision support tools do not provide any prescriptive solutions to combat the degradation problem." The biological or economic loss of value of the land, also called land degradation, (IBPES, 2018), reflects a set of threats, including salinization, especially in the drylands (Tarolli et al., 2024). Reasons for this decline are both anthropogenic, especially the land reclamation and climate change's role (Prăvălie et al., 2021). The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) has proposed a range of adaptation strategies to tackle the issue of land degradation. The IPBES adaptation strategies for combating land degradation should emphasize the importance of soil health (IPBES, 2018). Healthy soils are essential to meet climate and biodiversity UN sustainable development goals (SDGs) under the European Green Deal (Montanarella, 2020; Montanarella & Panagos, 2021). Common Agricultural Policy (CAP) has progressively integrated instruments to support the environment (Pe'er et al., 2019). Plants' needs in terms of water, increase with increasing temperature and aridity, increasing salt contents in soil increase the basal energy expenditure of plants (Munns & Tester, 2008). Furthermore, ions such as Na^+ can contribute to reduce soil permeability (Sumner, 1993), and Cl^- to unbalance plant nutrients and cause toxic effects (Geilfus, 2018). Measuring the EC_e at 25 °C, together with the Exchangeable Sodium percentage (ESP) allows for the assessment of soil salinity.

However, measuring EC_e (saturated soil paste) is cumbersome for routine laboratory work and generally replaced by the measurement of EC at 25 °C in 1:5 soil-to-water extracts ($EC_{1:5}$) (Jones et al., 2020). On a global scale, soil salinization is accelerating due to climate change (Ajilogba & Walker, 2021), intensive farming (Corwin, 2021), use of fertilizers (Mohanavelu et al., 2021), poor drainage of the cultivated land (Mukhopadhyay et al., 2021), and irrigation with low-quality irrigation waters (Obi et al., 2014; Shahid et al., 2018). Apart from specific salt-bearing soil parent materials, geomorphologically low-lying areas and rising sea levels that give birth to primary salinization processes featuring specific soil classes (Solonchak, Solonetz). These factors, in combination with seawater intrusion in coastal aquifers, and poor quality irrigation water are the leading causes of soil salinization. Groundwater quality is highly variable both temporally and spatially, but it is often used for irrigation despite the increased salinization (Delgado et al., 2010; Kurunc et al., 2016; Nikolaou et al., 2020). Specifically, ground waters that cross evaporitic rocks can present high salinity (Gil-Márquez et al., 2017; Li et al., 2020). Indeed, often researches indicate that groundwater quality remarkably vary over short distances (Arslan, 2017; Verma et al., 2017). In the LUCAS soil dataset, the $EC_{1:5}$ soil-to-water extract is the methodology adopted to evaluate soil EC. This has been the preferred method in Australia but not commonly used in the United States, where the $EC_{1:1}$ soil-to-water ratio is the selected methodology, since such ratio is closer to the soil water content at saturation featuring the EC_e measurement (Matthees et al., 2017). Up from a plant's characteristic threshold, high EC_e limits plant growth because of the decreasing potential of the soil water (Nabiollahi et al., 2017). The critical EC_e threshold of the majority of vegetable crops ranges from 1 to 4 dS m^{-1}. A global assessment using machine learning (Hassani et al., 2020) reported the major soil salinity hotspots and produced estimates based on decadal data clusters on the bases of the World Soil Information Service (WOSIS), along with an uncertainty assessment associated with the potential salinity content. However, the

low sample density penalizes this kind of evaluations. (Daliakopoulos et al., 2016) systematically reviewed the models that account for soil salinity assessment and summarized the state of Mediterranean countries as affected by secondary salinity (Stavi et al., 2021) because of seawater intrusion in coastal aquifers, as well as irrigation and drainage issues. In contrast, northern European countries are mostly affected by the primary soil salinization that occurs because of sea level rise. Remote sensing-aided applications were the most studied for global scale mapping (Ivushkin et al., 2019) focusing on irrigated lands (Ivushkin et al., 2018). The LUCAS Soil Module is the first in terms of internal coordination and knowledge sharing soil collection and analysis coherent strategy for developing georeferenced information across the European Union member states (MS). The objective of the present study is to predict the topsoil EC_e spatially and provide an instrument to monitor soils affected by high EC_e at the EU scale using LUCAS soil data with a state-of-the-art machine-learning framework together with environmental data and satellite imagery. Accurate and reliable case-based digital soil mapping (DSM) studies have been performed using various combinations of soil forming factors for a quarter of a century, using a wide variety of machine learning techniques within the framework of DSM (Ivushkin et al., 2019; Szatmári et al., 2020). Quantile regression forest (QRF) was used for the spatial assessment of the EC_e using R quantreg (ranger package); this approach can provide the probability distribution estimated of the prediction (Meinshausen, 2006) (and related uncertainty expressed as per quantiles; Lombardo et al., 2018). The QRF was proven to outperform other ML models as showcased by (Vaysse & Lagacherie, 2017). In their study (27,236 km^2 Mediterranean French region), QRF was used to estimate the uncertainty of pH, organic carbon and clay content at 5–15 cm depth. In this study we implemented an assessment of EC_e at EU level 500 m spatial resolution, using $EC_{1:5}$ data derived from the LUCAS (2018) converted using an empirically-derivedPTF and a set of environmental covariates fitted into a QRF model to produce an EC_e map. The EC_e map and dataset will be provided in the European Soil Data Centre (ESDAC) (Panagos et al., 2022).

2. Material and methods

2.1. Soil EC data

LUCAS Survey on Land Use and Coverage Statistics is an initiative prompted by EUROSTAT, the European Statistical Office. EUROSTAT has been conducting systematic and uniform surveys in collaboration with the Directorate-General for Agriculture and with technical assistance from the Joint Research Centre. These surveys span all Member States and aim to compile data on land use and land cover. Known as LUCAS (Land Use/Cover Area frame statistical Survey), the approach of the survey involves calculating estimates for different land use and cover types based on observations at over 250,000 points across the EU, rather than exhaustive mapping. Conducted periodically, the survey facilitates the monitoring of land use changes over time.

In 2009, the European Commission expanded the scope of the recurring Land Use/Land Cover Area Frame Survey (LUCAS) to include the analysis of topsoil's soil properties in 23 European Union (EU) Member States including the 2012 extension of the sampling sites in the Romanian and Bulgarian territories (Tóth et al., 2013). This inaugural topsoil survey was designed to create a uniform spatial database of EU soil cover, based on harmonized sampling and analytical methods. A single laboratory was tasked with analyzing all soil samples. From the primary LUCAS grid (around 300.000 land use observation), roughly 20,000 locations were chosen for soil sample collection. Each sample, about 0.5 kg of topsoil (0–20 cm), was obtained using a standardized procedure (ISO 18400–104:2018 Soil quality, Sampling Part 104: Strategies) and subsequently sent to a central laboratory for detailed physical and chemical examination. Following 2012, 2015 and 2018 have been sampled and analysed in a single laboratory, all the information related (raw data) can be found in the ESDAC website (https://esdac.jrc.ec.europa.eu/projects/lucas).

The 2018 LUCAS Soil Module (Orgiazzi et al., 2018) includes $EC_{1:5}$ measurements in 17,915 locations in EU27, thus being the most recent campaign and abundant sample set for the EU.

Diluted $EC_{1:5}$ was measured to implement a spatial-explicit EC dataset. The method for the measurement of the diluted $EC_{1:5}$ is based on the ISO 11265:1994 standard protocol (ISO 11265:1994) in which a 20 g sample of air-dried soil is equilibrated with 100 mL water at 20 ± 1 °C for 30 min by machine-shaking at a soil-to-water ratio of 1:5, and then directly gravity-filtered through paper. The specific EC of the filtered extract is measured, and the result is corrected to a temperature of 25 °C. The equipment comprises the conductivity meter, a 0.01-g precise weighing scale, a shaking machine, 250 cm^3 bottles and ash-less filter paper for quantitative analysis. The reagents needed are ultrapure water of EC at 25 °C below 0.002 dS m^{-1}, and EC potassium chloride standards of 0.10, 0.02, and 0.01 mol L^{-1}. Measure the EC of the filtrates according to the instructions provided by the manufacturer of the conductivity meter with the temperature corrected to 25 °C (Firm, City, State, Country). The extraction of the soil samples was performed along with blank essays assuming a quality standard in which the blank EC should not exceed 0.01 dS m^{-1}. In the LUCAS soil 2018 Report, the determination of EC is expressed in dS m^{-1} with two decimals. A preliminary hotspot analysis on the raw data has been presented in a previous technical report of LUCAS topsoil 2018 survey (Fernandez-Ugalde et al., 2022). These hotspots are showing the diluted $EC_{1:5}$ and the proposed methodology for conversion will be used to get a new hotspot showing the saturated EC_e distribution presented in Fig. 1 and in the following section.

Given the lack of a universally agreed classification for $EC_{1:5}$, various approaches have been developed that incorporate soil texture to provide a more nuanced assessment of soil salinity, such as those that use different threshold values for sandy, loamy, and clay soils, or that combine $EC_{1:5}$ measurements with other soil properties to estimate salinity risks.

According to the Australian Department of Primary Industries and Regional Development's Agriculture and Food classification, an $EC_{1:5}$ measurement is considered highly saline in sandy soils if it slightly exceeds 1 dS m^{-1}, whereas in clay soils it is considered saline if from 1 to 2 dS m^{-1} Table 2 (supplementary materials).

2.2. Methodologies for the conversion of diluted $EC_{1:5}$ to saturated EC_e

For the conversion of the diluted $EC_{1:5}$ to saturated EC_e, several pedotransfer functions have been developed by empirical calibration (He et al., 2013; Kargas et al., 2020; Slavich & Petterson, 1993; Visconti et al., 2010). Therefore, the pedotransfer function is usually tailored for a specific environment because the conversion coefficients in the pedotransfer generally depend on the textural classes, as well as on the soil organic carbon (Lekka et al., 2023), gypsum contents and maybe other soil properties that feature the soil samples that were used for their calibration (Khorsandi & Yazdi, 2007; Slavich & Petterson, 1993). Sometimes these properties are included as correction factors in the most complex pedotransfer functions for the $EC_{1:5}$ to EC_e conversion (Khorsandi & Yazdi, 2011). In particular, soils with at least 0.5 % gypsum are gypsum-saturated in both the soil to water ratio 1:5 and saturated extracts with $EC_{1:5}$ and EC_e over 2.4 dS m^{-1} in both thus hindering the estimation through simple equations of EC_e from $EC_{1:5}$ (Visconti et al., 2010). Several guidelines exist for the conversion of $EC_{1:5}$ to EC_e. There are several examples in the literature where the effect of texture has been included to convert $EC_{1:5}$ to EC_e (Slavich and Petterson, 1993; Kargas et al., 2020). There are no ISO and/ or CEN standards for EC_e determination, only FAO standard but a large amount of soil (200–400 g) are needed for the EC_e determination using GLOSOLAN-SOP-8 for saturated EC. (https://www.fao.org/global-soil-partnership).

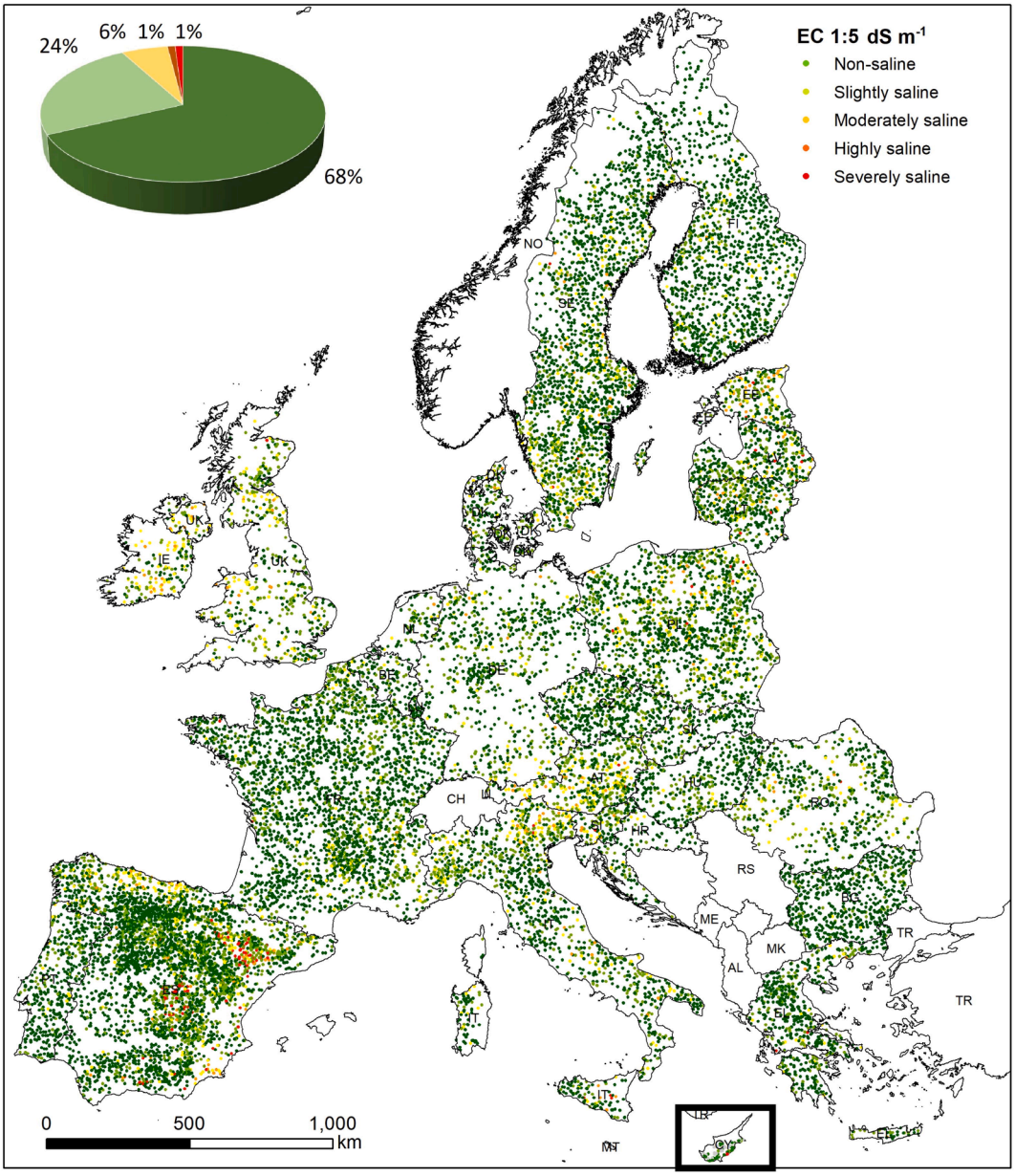


Fig. 1. $EC_{1:5}$ LUCAS 2018 sampling locations, classified according with the Australian Department of Primary Industries and Regional Development's Agriculture and Food, taking in consideration $EC_{1:5}$ range and the different soil textures, https://www.agric.wa.gov.au/soil-salinity/measuring-soil-salinity. The inlet map is done to represent in a smaller extent all EU-27, originally Cyprus is far western than where it is placed.

2.2.1. Pedotransfer function

To fully use the potential of LUCAS soil survey, we developed a pedotransfer function to transform $EC_{1:5}$ to EC_e by taking into account the effect of the texture and the soil organic matter (SOM).

The conversion coefficient from $EC_{1:5}$ to the electrical conductivity in the saturate paste (EC_e) depends on the saturation percentage (SP%) (Slavich and Petterson, 1993; Kargas et al., 2022):

$$ECe = (k + \frac{m}{SP\%}) EC1:5 \quad (1)$$

In Eq. (1), k and m are two empirical coefficients, which were estimated from topsoil samples from (Gharaibeh et al., 2021) with values of k = 1.054 and m = 283.4.

The SP% depends on the soil clay and sand mass fractions and furthermore, on the SOM mass fraction since it affects the SP% of soils where it appears at non-negligible levels. However, the Kargas et al. (2022) soil samples were very low in SOM, and therefore the effect of SOM is not included in the SP% calculation and thus neither in the estimation of the conversion coefficient.

To include the effect of the soil organic matter along with the effect of the clay and sand fractions in the estimation of the SP% and, therefore, the conversion coefficient, the first equation approximated by (Robinson et al., 2022) for the soil pore space may be used:

$$SP\% = 100\varphi = 100\left(1 - \frac{\frac{wSOM}{\rho sOM} + \frac{1-wSOM}{\rho sM}}{\frac{wSOM}{\rho bOM} + \frac{1-wSOM}{\rho bM}}\right) \quad (2)$$

In Eq. (2), w_{SOM} is the soil organic matter mass fraction in g g^{-1}, and the other variables are particle and bulk densities of soil solids, which are defined in Table 1.

Note that in the LUCAS soil survey, the soil organic carbon (SOC) is measured, so to calculate the SOM from the SOC we used the van Bemmelen coefficient (Minasny et al., 2020) in the EC conversion.

Table 1
Definitions and values of the variables in Eq. 3.

Variable	Definition	Value (g cm^{-3})
ρ_{SOM}	Particle density of the soil organic matter	1.40
ρ_{SM}	Particle density of the soil minerals (coarse fraction)	2.65
ρ_{bOM}	Bulk density of the soil organic matter	0.10
ρ_{bM}	Bulk density of the soil minerals	Eq. 3 and 4 (see text)

In Table 1 it is shown that all densities except for the bulk density of the soil minerals, present well-defined values. The bulk density of the soil minerals (ρ_{bM}) can be estimated from the mineral pore space (φ_M):

$$\rho_{bM} = \rho_{sM} (1 - \varphi_M) \quad (3)$$

in which ρ_{SM} is the particle density of the soil minerals for which a value of 2.65 g cm^{-3} can be taken (Table 1). The mineral pore space (φ_M) can be calculated according to Kargas et al (2022) (Eq. 4).

$$\varphi_M = 0.49699 + 0.524\ w_{clay} - 0.339\ w_{sand} \quad (4)$$

where w_{clay} and w_{sand} are the clay and sand mass fractions (g g^{-1}), respectively taken directly from the LUCAS 2018 points.

Therefore, the φ_M given by Eq. 4 can be replaced in Eq. 3 to obtain ρ_{bM}, and this in turn in Eq. (2) to obtain the %SP to be used to calculate the conversion coefficient in Eq. (1) along with the calibrated k and m coefficients. The logic behind the use of this PTF is try to use LUCAS 2018 EC $_{1:5}$ data, in light of the limited amount of soil stored in the archive that do not allow for the analysis of EC_e. Literature analysis reports many examples of conversion between $EC_{1:5}$ to EC_e. Although we agree it is a preliminary assessment, we are trying to develop new laboratory procedures to reduce the sample quantity needed for the EC_e analysis for a better and more direct comparison between LUCAS soils EC $_{1:5}$ and the soil's EC_e.

Table 2
Variables used as predictors in the QRF model, definition, reference, spatial and temporal scale.

Variable name	Temporal resolution	Type of variable	Original Spatial Resolution in meters	Spatial resolution used in the model	Reference
		Soil properties			
Sand %	2009		500	500	(Ballabio et al., 2016)
Clay %			500	500	
Silt %			500	500	(Ballabio et al., 2019)
	Climate				
Mean Annual Temperature (WorldClim-Bioclim1)	1970–2000		1000	500	(Fick & Hijmans, 2017)
Annual Precipitation (WorldClim-Bioclim12)			1000	500	
Surface Soil Moisture			300	500	https://land.copernicus.eu/global/products/ssm
		Vegetation			
NDVI (Annual Average)			300	500	Copernicus Global Land service web site https://land.copernicus.eu/global/products
		Terrain			
Distance from Sea 10 m asl			90	500	European Digital Elevation Model (EU-DEM) based on SRTM and ASTER
Elevation (DEM)			90	500	
Slope			90	500	
Northness			90	500	
Eastness			90	500	
Plan Curvature			90	500	
Profile Curvature			90	500	
Terrain Wetness Index (TWI)			90	500	
Topographic Position Index (TPI)			90	500	(Trevisani et al., 2023; Wilson et al., 2007)
Terrain Ruggedness Index (TRI)			90	500	
Roughness			90	500	
		Parent material			
Dominant Parent Material (ESDB v2)			1000	500	ESDBv2 Raster Library EUR 19,945 EN
	Soil class				
WRB ESDB dominant Soil Reference Groups*			1000	500	https://esdac.jrc.ec.europa.eu/content/european-soil-database-v2-raster-library-1kmx1km
Land Cover					
CORINE land cover 2018**			100	500	land.copernicus.eu/pan-european/corine-land-cover/clc2018
Irrigated land, derived by CORINE land cover (212)			100	500	
Hydraulic properties EU					
Saturated water content (cm^3 cm^{-3})			250	500	(Tóth et al., 2015)
Water content at wilting point (cm^3 cm^{-3})			250	500	
Water content at field capacity (cm^3 cm^{-3})			250	500	
Saturated hydraulic conductivity (cm day^{-1})			250	500	

*For the WRB dominant Soil Reference Groups Acrisols are not presented in the training moved to no soils.
**The 2018 CORINE land cover classes 335″ Glaciers and perpetual snow, "423″ Intertidal flats, "999″ No Data, "521″ coastal lagoon, "522″ Estuaries were not present in the LUCAS sampling location.

2.3. Covariates

In order to predict the spatial distribution of EC_e, we have considered the inclusion of variables in different proportion (Table 2). It reports all the covariates used in the modelling phase, these categories are derived from the SCORPAN factors (McBratney et al., 2003), to represent the soil-forming factors that can be used to predict soil types and properties:

S – Soil (The textural properties, WRB class (IUSS Working Group WRB 2022), to represent the properties of the soil itself, including the current state and distribution of soil types);

C – Climate (the MAT and MAP to provide data on climatic variables such as temperature, precipitation, and evapotranspiration that influence soil formation and modification);

O – Organisms (Normalized Difference Vegetation Index "NDVI" as a proxy of the plant communities and vegetation to report the influence of flora, fauna, and human activity on soil characteristics).

R – Relief (Based on the Shattle Radar Topography Mission SRTM Digital Elevation Model (Farr et al., 2007), and performing several derivatives primary and secondary (Schillaci et al., 2015)) topography and landscape position, which affect erosion, deposition, and soil moisture regimes.

P – Parent material (Panagos et al., 2022b) the underlying geological material from which the soil has developed, which influences its mineral composition and texture, available online as the ESDAC webpage of the ESDBv2 Raster Library – a set of rasters derived from the European Soil Database distribution v2.0 (published by the European Commission and the European Soil Bureau Network, CD-ROM, EUR 19945 EN); Marc Van Liedekerke, Arwyn Jones, Panos Panagos 2006.

A – Age not considered.

N – Space or position: Considering the distance from the sea and the elevation as the spatial or geographic location of the soil, which can encapsulate the effects of spatial relationships and neighborhood interactions with surrounding soils.

2.4. Descriptive statistics of the observed data of saturated EC_e [dS m^{-1}] dataset, LUCAS 2018

To present the main statistical features of the dataset at NUTS 0 level, count, mean, min, max, median, standard deviation (sd), skewness, kurtosis and variance are reported in Table 3.

2.5. DSM procedure, Quantile regression forest

When looked at the predictors used in previous mapping routines; topography and climate have a significant role. Soil EC_e could affect land productivity, which is the biological productive capacity of the soil, the source of all the food, fuel and fiber that sustains humans (United Nations Statistical Commission 2016). Mapping was done in the EU using a set of environmental covariates, multiple linear regression often considered the benchmark for the prediction of soil parameters and machine learning (ML), at least in recent times, is the approach that yields the best results. We identified the most used ML models from the literature analysis and compared the performances of Random Forest (RF), Gradient Boosting Machine (GBM) and Quantile regression forest (QRF) (Richer-de-Forges et al., 2023; Shi et al., 2023; Wadoux et al., 2020a). The most used ML models in DSM, specifically for EC mapping, are RF, Regression kriging, and QRF. The QRF has achieved the best results regarding Root Mean Square Error (RMSE) and transferability among the analysed papers. Quantile regression forests (QRFs) are a type of machine learning algorithm used for non-parametric regression. It is a type of ensemble learning technique that combines the outputs of many individual regression trees to create more accurate and robust predictions. QRFs are particularly useful when the data contains heterogeneous or nonlinear relationships. In R, the quantregForest package provides an implementation of QRFs. The package allows users to create and tune both parametric and nonparametric models. The package includes functions for fitting, predicting, and visualizing the QRF. It also contains functions for computing various model performance metrics, such as the RMSE and the Coefficient of determination (R^2).

Table 3
Descriptive statistics for all NUTS 0 (n = 17176), NUTS 0 Level 0: Country) refers to the highest hierarchical level in the NUTS classification system, which is a geocode standard developed by the European Union for referencing the subdivisions of countries for statistical purposes. The layer was downloaded from EUROSTAT website NUTS0_2018_EPSG3039.

NUTS0	count	n. of samples $EC_e > 4$ dS m^{-1}	min(EC_e)	max(EC_e)	median(EC_e)	mean(EC_e)	sd(EC_e)	skewness(EC_e)	kurtosis(EC_e)	var(EC_e)
AT	411	5	0.22	7.47	0.94	1.16	0.87	2.48	13.22	0.75
BE	102		0.19	2.09	0.59	0.68	0.35	1.46	5.36	0.12
BG	536		0.13	1.63	0.40	0.48	0.27	1.30	4.77	0.07
CY	60	1	0.19	14.71	0.86	1.11	1.84	6.93	51.74	3.38
CZ	414		0.19	2.26	0.48	0.57	0.29	1.98	9.52	0.09
DE	716	1	0.14	5.74	0.57	0.69	0.48	3.16	23.96	0.23
DK	154		0.17	3.47	0.64	0.73	0.46	2.77	14.22	0.21
EE	192	3	0.11	7.99	0.68	1.01	1.04	2.88	14.64	1.09
EL	570	5	0.11	19.59	0.64	0.76	1.09	11.64	176.90	1.19
ES	3686	125	0.07	25.31	0.71	1.10	1.86	5.65	41.67	3.44
FI	1103	1	0.06	4.21	0.27	0.39	0.35	2.97	19.50	0.12
FR	2563	3	0.12	15.98	0.60	0.69	0.50	12.60	344.38	0.25
HR	101		0.16	2.06	0.65	0.74	0.46	1.01	3.31	0.22
HU	336		0.10	3.32	0.57	0.62	0.36	2.40	14.87	0.13
IE	138	5	0.23	21.55	1.16	1.53	1.97	7.83	79.17	3.87
IT	1162	21	0.18	11.50	0.83	1.05	0.89	4.79	39.84	0.79
LT	377	3	0.12	5.42	0.56	0.70	0.64	3.58	20.34	0.41
LU	33		0.24	1.38	0.53	0.58	0.28	1.14	4.05	0.08
LV	327	5	0.11	5.56	0.49	0.72	0.80	3.46	16.43	0.64
MT	2		0.92	3.85	2.39	2.39	2.08		1.00	4.32
NL	86		0.01	3.15	0.63	0.79	0.52	2.02	8.28	0.27
PL	1293	10	0.11	7.14	0.56	0.69	0.64	4.33	32.06	0.41
PT	408		0.10	2.66	0.29	0.43	0.34	2.15	9.70	0.11
RO	552	1	0.10	7.31	0.66	0.79	0.60	3.48	29.90	0.36
SE	1834	4	0.09	5.32	0.38	0.53	0.52	3.33	19.99	0.27
SI	109	1	0.16	7.15	0.80	1.01	0.95	3.39	19.05	0.90
SK	171		0.16	1.85	0.50	0.60	0.36	1.27	4.54	0.13
UK	480	6	0.16	6.73	0.94	1.19	0.86	2.28	10.72	0.74
EU	17,916	200	0.01	25.31	0.61	0.79	1.06	8.57	111.25	1.11

2.5.1. Grid search of the hyperparameters a permutation approach

The tuning parameters for QRF in R are the following: 1. number of trees (ntree): This parameter specifies the number of trees to use for the creation of the forest. The "mtry" parameter, which specifies the number of variables randomly sampled as candidates at each split. The "node size" parameter, which specifies the minimum size of terminal nodes (leaves) in the forest. The "sampsize" argument of the quantregForest() function controls the size of the subsamples that are used to create the regression trees (Van Eynde et al., 2023). It should be specified as a vector, with the first entry being the number of observations in the first subset, and the second entry being the number of observations in the second subset, and so on. For example, the code available in the supplementary material, will generate a quantile regression forest with 200 trees, using subsamples of size 50, 100, and 150.

2.5.2. Modelling Quantile regression forest

. Quantile Regression Forest gives precious information about the (spatial) uncertainty associated with the QRF prediction, which is relevant for the end-user of digital soil mapping products (Wadoux et al., 2020b). We used the median prediction as the final predicted EC_e concentration. The EC_e were log10 transformed before modelling, because of their skewed distribution. However, the distribution was still slightly skewed (median and mean of log EC_e in the training set were 0.5849 and 0.5803 dS m^{-1}) with smooth tails and samples with relatively high concentrations (supplementary Fig. 3 graphical representation of the distribution of training and testing LUCAS datasets). However, using median or mean prediction did not affect the predictions for the independent test set. Model performance is evaluated by indices R^2, RMSE.

For the QRF, the following hyperparameter were tuned: i) number of trees, ii) number of predictor variables randomly selected at each node, iii) proportion of observations to sample in each regression tree iv) minimum number of observations in a regression tree's terminal node.

The hyperparameters are optimized via cross-validation. In practice, there is no need to tune the number of decision trees; it is usually recommended to set it to a large number, allowing the convergence of the prediction error to a stable minimum (Hengl et al., 2018). Through modelling the relations between soil EC_e observations and a collection of predictors (topographic, climatic, vegetative, soil and landscape properties) coincident with the sampling locations, a grid of soil EC_e values was generated at the EU scale for the reference year 2018.

Predicted values have been provided on a continuous scale and reclassified according to the FAO thresholds of EC_e to determine salinity (Batlle-Sales, 2023; Omuto et al., 2022). The Q50, also known as the median, is often chosen as the best representation of the real condition for several reasons: robustness to outliers, representativeness, symmetry, interpretability, and comparability. Additionally, Q50 is in good agreement with the observed data as confirmed by the frequency distribution plot in the supplementary material Fig. 5. The Q50 derived from the EC_e modelling have been used to generate thematic maps of the median EC_e in the EU at ~500 m spatial resolution

3. Results

The following section presents the results of the statistical descriptive analyses, in which the EC_e obtained by converting the $EC_{1:5}$ (diluted) to EC_e (saturated paste) were assessed by country (NUTS 0). Output of the QRF model are presented in the map of topsoil EC_e (Quantile Q50), and uncertainty (Quantiles Q0.05 and Q95), using the classification approach of the FAO EC_e thresholds (Batlle-Sales, 2023; FAO, 2020), and the variable importance for explaining the EC_e distribution and most influential predictors in Europe.

3.1. Conversion of diluted $EC_{1:5}$ to saturated paste EC_e pedotransfer results

The soil properties involved in this conversion are clay, sand and soil organic matter (SOM) percentage. Sandy soils do not hold as much water as do clayey soils. Therefore, the same EC_e, if derived from the diluted $EC_{1:5}$ in loamy soils, the higher the EC_e, as assessed by the EC_e method, in lighter textured soils (sandy soils) than in heavier textured soils (clayey soils). In Fig. 2, it is shown how the conversion coefficient varies with the clay mass fraction at different SOM levels.

3.2. Descriptive EC_e 2018 per NUTS 0

Quantitative analysis of soil EC_e per administrative region is shown in Table 4. Based on the LUCAS topsoil database, the average EC_e in Europe was 0.76 dS m^{-1}, with a standard deviation of 0.92 and median of 0.61. The LUCAS EC_e (n = 17.915) values ranged between 0.01 and 25.3 dS m^{-1}, with skewness of 8.24 and Kurtosis of 104.4. Ninety-nine percent of samples had values below 4 dS m^{-1}. Samples with $EC_e > 4$ dS m^{-1} were mostly located in Spain, Italy, and Ireland (Fig. 2 and Fig. 4 in supplementary materials). After the application of the pedotransfer function, 146 soil samples had EC_e higher than 4 dS m^{-1} (Table 4).

The 17.769 samples that had EC_e below 4 dS m^{-1} (99 %) mostly came from croplands, followed by woodlands and grasslands (Fig. 5 in supplementary materials). Northern European countries (i.e. Sweden and Finland), as well as in Portugal presents lower EC_e values. Topsoil EC_e tended to be higher in rice fields (CLC 213), salt marshes (CLC 411) and lands principally occupied by agriculture with significant areas of natural vegetation (CLC 241), grasslands (CLC 321). In line with these results, soils under agricultural use generally hold the highest EC_e based on median values and the number of samples with values lower than 0.61 dS m^{-1}. However, it should be noted that agriculture is the dominant land use in the LUCAS database, representing 67 % of all soil samples. A distribution map of the $EC_{1:5} > 4$ dS m^{-1} ("hot spots") is presented in Fig. 5 of the supplementary materials.

3.3. Model prediction 2018, EU and NUTS 2-Distribution of EC_e

Optimal model hyperparameters "ntree" 1500, "sampsize" 10000. The final QRF model had relatively deep trees "nodesize" 2, and a high number of variables that were randomly selected at each split "mtry" 15 (Table 2 in supplementary materials). The accuracy of the final model was assessed based on R^2 and RMSE for the independent test set of the LUCAS data (Fig. 3). Whereas the observed-predicted scatter plots are presented in the supplementary Fig. 6. The final model resulted in an R^2 of 0.302 with an RMSE of 0.265 dS m^{-1} for soil samples not used for model calibration. As can be seen in tables 4 (min and max columns), the predicted values cover a slightly smaller range than the actual EC_e measured (Table 3) in the LUCAS topsoil samples. However, since the estimation of EC_e from $EC_{1:5}$ for soils with gypsum over 0.5 % cannot be made with PTF developed for soils with negligible gypsum, i.e. with gypsum under 0.2 % (Visconti et al., 2010), as has been the case in this work, Gypsisols were removed so as to not mislead readers (Fig. 3).The Reference Soil Groups of Europe are derived from the ESDB database and represent a broad group of soils that share similar features, where the dominant soil types do not necessarily have an high share but is the first most represented (Fig. 4). Based on the variable importance, Texture, WRB dominant soils, Mean Annual Precipitation and Average Annual Temperature and CORINE land cover explain more than the 50 % of the variation in EC_e. This result is a first indication of which covariates are most important for explaining soil EC_e in Europe, which is elaborated on in Section 3.4.

Table 5. Reports the prediction ability of the model for EC_e classes according to FAO and the relevant threshold (> 4 dS m^{-1}) of the EU soil monitoring directive to differentiate saline from non-saline soils. Green = the predicted value is within the observed class, yellow = slight different class, red = far from the original.

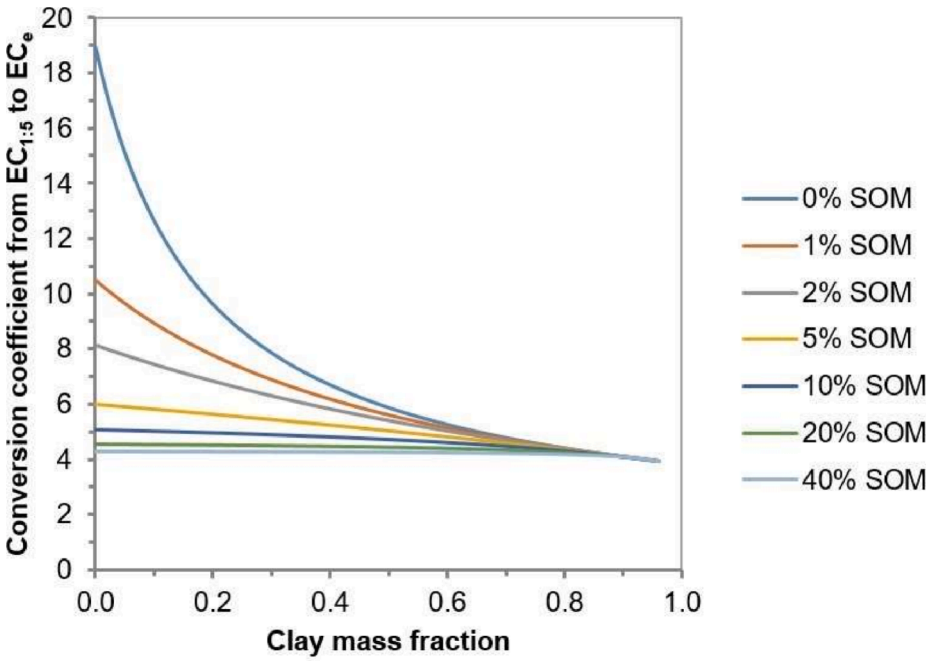


Fig. 2. Conversion coefficient from $EC_{1:5}$ to EC_e with the clay mass fraction at different SOM.

Table 4
NUT0 descriptive predicted EC_e (EC Q50),

NUTS0	count	n. of sample > 4 dS m^{-1}	min	max	median	mean	sd	skewness	kurtosis	var
AT	411	2	0.22	4.96	0.94	1.11	0.72	1.54	6.20	0.52
BE	102		0.19	2.09	0.59	0.68	0.35	1.46	5.34	0.12
BG	536		0.15	1.54	0.40	0.48	0.26	1.20	4.32	0.07
CY	60	1	0.19	8.67	0.86	1.01	1.10	5.92	41.46	1.20
CZ	414		0.20	2.00	0.48	0.56	0.28	1.76	8.08	0.08
DE	716		0.14	2.97	0.57	0.67	0.39	1.33	5.42	0.16
DK	154		0.18	3.47	0.64	0.72	0.43	2.71	14.89	0.19
EE	192		0.12	3.85	0.68	0.94	0.80	1.75	5.61	0.64
EL	570	3	0.13	10.75	0.64	0.72	0.74	9.45	123.32	0.55
ES	3686	116	0.11	25.31	0.71	1.06	1.66	5.54	39.63	2.77
FI	1103		0.06	2.01	0.27	0.38	0.30	1.91	7.23	0.09
FR	2563	1	0.12	15.98	0.60	0.68	0.47	14.33	442.31	0.22
HR	101		0.18	2.03	0.65	0.74	0.45	0.95	3.16	0.20
HU	336		0.10	2.34	0.57	0.61	0.32	1.53	7.75	0.10
IE	138	1	0.23	4.83	1.16	1.37	0.90	1.12	4.05	0.81
IT	1162	10	0.19	9.40	0.83	1.02	0.71	3.53	26.48	0.51
LT	377		0.13	3.88	0.56	0.67	0.51	2.68	12.98	0.26
LU	33		0.24	1.38	0.53	0.58	0.28	1.16	4.08	0.08
LV	327	2	0.13	4.60	0.49	0.69	0.68	3.18	14.82	0.46
MT	2		0.92	3.86	2.39	2.39	2.08	0.00	1.00	4.32
NL	86		0.17	2.51	0.63	0.78	0.47	1.64	5.81	0.22
PL	1293	6	0.12	6.61	0.56	0.66	0.53	3.67	28.20	0.29
PT	408		0.11	1.73	0.29	0.42	0.30	1.52	5.19	0.09
RO	552		0.10	3.58	0.66	0.76	0.48	1.41	6.26	0.23
SE	1834		0.10	3.59	0.38	0.51	0.41	2.23	10.18	0.17
SI	109	1	0.21	7.15	0.80	0.99	0.90	3.60	21.98	0.82
SK	171		0.16	1.85	0.50	0.60	0.36	1.25	4.40	0.13
UK	480	3	0.19	6.73	0.94	1.16	0.78	2.14	10.77	0.61
EU	17,916	146	**0.06**	**25.31**	**0.61**	**0.76**	**0.92**	**8.24**	**104.50**	**0.84**

3.4. Variable importance and uncertainty

The variable importance analysis allows interpreting the machine-learning models and defines which covariate contributed the most to improve the accuracy of the models. For the classification task, the variable importance of the models is shown in Fig. 5.

The variable importance is calculated as the total decrease in node impurities from splitting on the variable, averaged over all trees. The node impurity is measured by residual sum of squares. The type of Variable importance performed is the IncNodePurity, which is the total decrease in node impurities, measured by the Gini Index from splitting on the variable, averaged over all trees.

3.5. EC_e zonal frequency distribution

Most of the sampling points across the EU showed low EC_e, which indicates that the sampled soils display low EC_e (Fig. 2). However, supplementary materials Fig. 5 shows some EC_e hotspots (> 4 dS m^{-1}) in

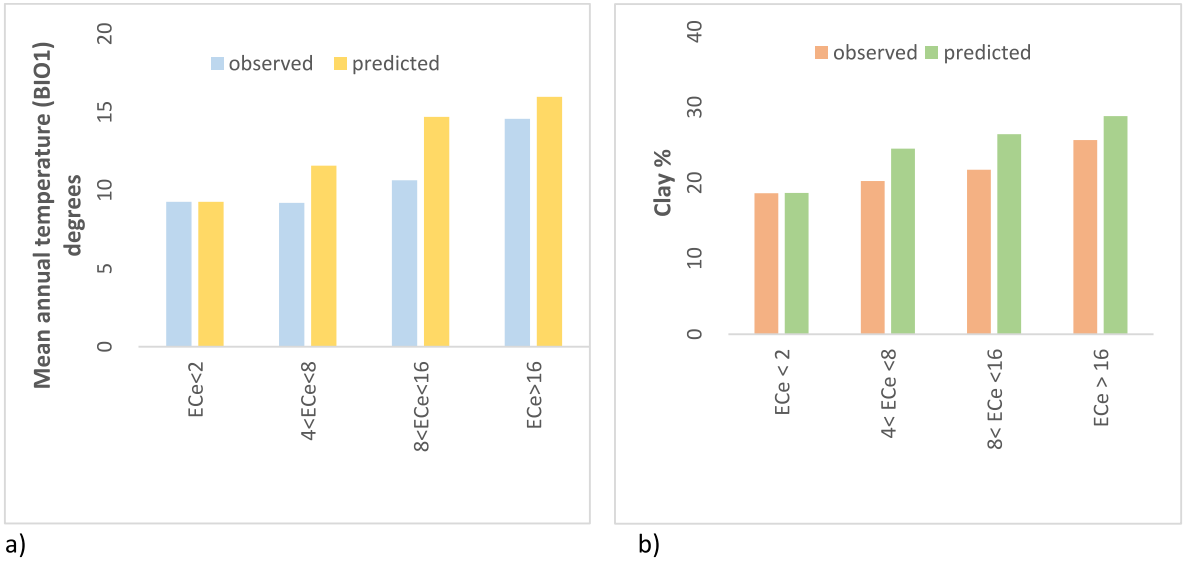


Fig. 3. A) Average annual temperature from WorldClim for LUCAS topsoil EC_e b) Average Clay content for LUCAS topsoil EC_e.

the Ebro Valley and Southeastern Spain, Sicily, Cyprus and Greece. The points on the coasts of Ireland and France reflect coastal salt marshes. Fig. 3 shows the distribution of topsoil EC_e based on QRF modeling. The uncertainty of the topsoil EC_e values map is shown in Fig. 7 and 8 of the supplementary material 0.95 and 0.05 quantiles. Based on QRF modeling, average temperature, clay content and WRB Soil Reference Groups form ESBD were the most important variables explaining topsoil EC_e values in Europe (Figs. 2–5). As a general trend, topsoil EC_e values increase with increasing annual mean temperature and clay content (Figs. 2–5). Zonal statistics are presented in Table 6. The only country that report in their territories the class corresponding to strong EC_e is Spain (0.114 %) (Figs. 6 and 7).

4. Discussions

4.1. Model prediction, variable importance and model uncertainty of EC

Several works have demonstrated the high accuracy of machine learning models and multi-data source approach to achieve high prediction accuracy of soil properties in arid (Wu et al., 2018; Xiao et al., 2023), semi-arid (Kaya et al., 2022), and sub-humid regions (Saygın et al., 2023). Many works have focused on validating the framework, i.e., selecting better set of covariates, data from remote sensing (Ivushkin et al., 2017), model development (Hassani et al., 2020), when instead the scientific community nowadays it is called upon to take a further step, i.e., to make these models available for a wide range of stakeholders (i.e. land managers, local authorities) and to define the best algorithm that generalized enough to predict well at unknown locations. Detailed spatial scale predictors are paramount, especially with entering the Copernicus Sentinel constellation, which delivers high spatial and temporal multi-spectral and radar imagery (Castaldi et al., 2023). The median topsoil EC_e values of 0.61 dS m^{-1} predicted with the proposed approach aligns with findings from literature (Kargas et al., 2020; Kaya et al., 2022; Visconti et al., 2010). The variability in EC_e was large, with clear differences between EU countries (Tables 1 and 2) and soil type (Fig. 3). The lowest topsoil EC_e was found in Finland, Lithuania, Latvia, Sweden, Estonia, Denmark, Poland, Czech Republic, France, Germany, Portugal and Slovakia, with median EC_e below 0.61 dS m^{-1}. The highest topsoil EC_e was found in Ireland, Malta, Cyprus, Spain, Italy, and Romania. The application of the QRF (Meinshausen, 2006) as a nonlinear predictive model in digital soil mapping using LUCAS topsoil database is widely used at continental level (Ballabio et al., 2019; Fendrich et al., 2024; Van Eynde et al., 2023; Vaysse & Lagacherie, 2017), it allowed us to gain insights into the spatial distribution of topsoil EC_e values in Europe and the underlying drivers. Based on this model, we were able to produce a dataset showing EC_e related to LUCAS 2018 distribution across Europe at 500 m resolution (Fig. 3) that will be publicly available at the European Soil Data Centre (ESDAC) of the Joint Research Centre (Panagos et al., 2022) and for which the underlying drivers were quantitatively assessed. The obtained model and the evaluation metrics in line with the literature showed an improvement compared to previous attempts to map EC_e at continental and global level (Hassani et al., 2020; Ivushkin et al., 2019). The evaluation metrics are in line with what has been already found when mapping EC_e and other salts at the country scale, e.g., for Na^+ in the Irish Soil Geochemical Atlas (Fay et al., 2007). The quantitative assessment of the underlying drivers presented in this work, are methodological and operational results that help monitor EC_e. There is indeed a strong influence on sampling locations settings that can be explained while EC_e pattern are captured in some regions in Southern Europe due to the potential proximity effect to the sea, potential secondary salinity due to irrigation, substrate related gypsum (Acosta et al., 2011; Herrero et al., 2022; Visconti et al., 2019), or the vicinity to former and actual mining sites (Pásztor et al., 2018). An interesting finding was the ability of this work to capture high EC_e in arid and semiarid climate, where soils are subject to high evapotranspiration. Particularly, Solonetz WRB Reference Group are soils with high exchangeable sodium percentage, over 15 % with some exceptions, whereas Solonchaks are soils with high soluble salts content (Çullu & Günal, 2018). In the past, Solonetz and Solonchak were jointly considered as the “salt-affected soils” but this is no longer the case because of the very different features one and another present, including low salinity for Solonetz.

Therefore, Solonetz have not shown up in this mapping effort simply because there were no sampling point taken in this specific soil class. Additionally, the Hungarian Solonchaks have neither been sampled in this work because soil salts in there are below the LUCAS standard of 20 cm for soil sampling (Szatmári et al., 2020).

When looking at WRB soil type, the potential occurrence of Gypsisols, Calcisols, and Solonchacks was relatively high when selecting only LUCAS samples with EC_e concentrations above 1 dS m^{-1} with Gypsisols on top. Gypsisols were found to be very high in EC_e regarding the other soils because of an artifact derived from the use of a conversion factor from $EC_{1:5}$ to EC_e, which does not take into account the EC buffering effect of gypsum against dilution (Visconti et al., 2010). Therefore, the Gypsisols were removed from the analysis by acknowledging that their EC_e cannot be estimated at this time due to the aforementioned reason.

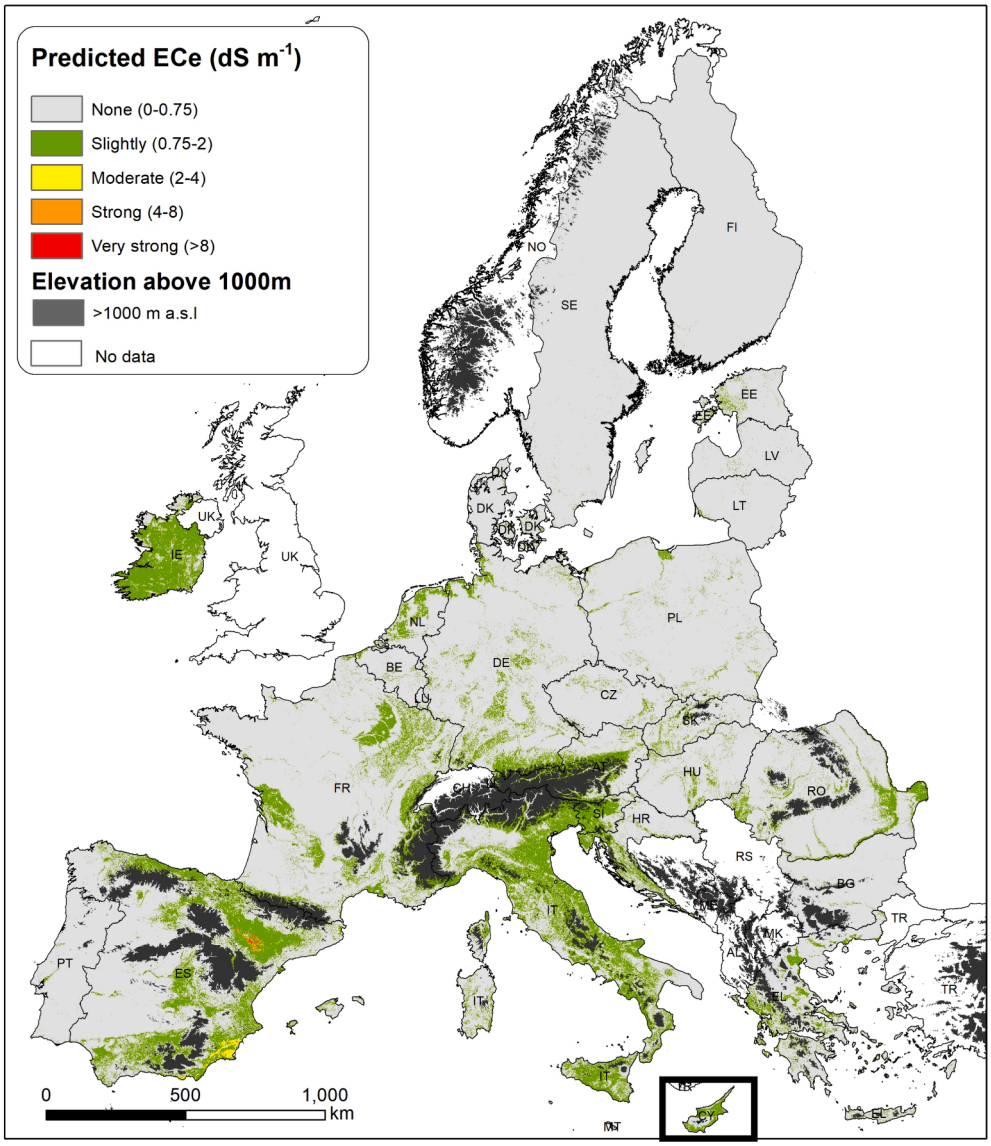


Fig. 4. Predicted Q50 EC_e map classified according to FAO INSAS, EC_e in dSm^{-1} Classes: 0–0.75 None, 0.75–2 Slightly, 2–4 Moderate, 4–8, Strong, >8 very strong.

Table 5
Predicted FAO class error versus observed class in absolute values and percentage.

Observed	Predicted EC_e (dS m^{-1})	# points EC_e (dS m^{-1})	Percentage EC_e (dS m^{-1})
>4 (Strong)	>4	146	0.81 %
	2–4	44	0.25 %
	0.75–2	10	0.06 %
2–4	2–4	418	2.33 %
(Moderate)	0.75–2	125	0.70 %
0.75–2	0.75–2	5550	30.98 %
(Slightly)	<0.75	8	0.04 %
<0.75 (None)	<0.75	11,615	64.83 %

An attempt was made to improve the estimates by including the $CaCO_3$ term into the PTF. However, considering soils with small amount of $CaCO_3$, potential difference arose from the dissolution from the soil, we do not know if the time left to equilibrate the $EC_{1:5}$ soil–water suspensions was long enough to let dissolve all the $CaCO_3$ that could dissolve (Visconti et al., 2010; Visconti and de Paz, 2016). In future works other pedotransfer functions may be shown and tested by analyzing $EC_{1:5}$ and saturation extracts of a widely representative sample of soils. The model can have substantial changes due to increasing resolution of the variable and the setting of a different training testing data partitioning.

4.2. Spatial distribution of the saturated EC_e

In this section, we contextualize the predicted EC_e spatial distribution obtained with the proposed approach and further delve into the reasons why in some countries there are disagreements with the EC_e values reported in national surveys. At EU level both EC_e values > 2 dS m^{-1} and measured $EC_{1:5}$ > 0.8 dS m^{-1} were also found in soils with pH < 7, in soils with low clay content, and no correlations with lime content. Differences between soil EC_e under different land covers could also be related to the differences in soil properties like pH and SOC and climatic factors like precipitation and temperature (Wong et al., 2010). One direct answer is that EC_e is a soil property subjected to spatial variability, both horizontal and vertical, as well as temporal variability. Therefore, i) the sampling density should intensify to capture the horizontal variation at local scale, ii) the sampling should extend to deeper soil layers to capture the soil profile variation, and iii) the sampling should be made at least at two weather-contrasted times during the year to capture the temporal variation. According to the Proposal for a Directive on Soil Monitoring, values of EC_e below 4 dS m^{-1} are not

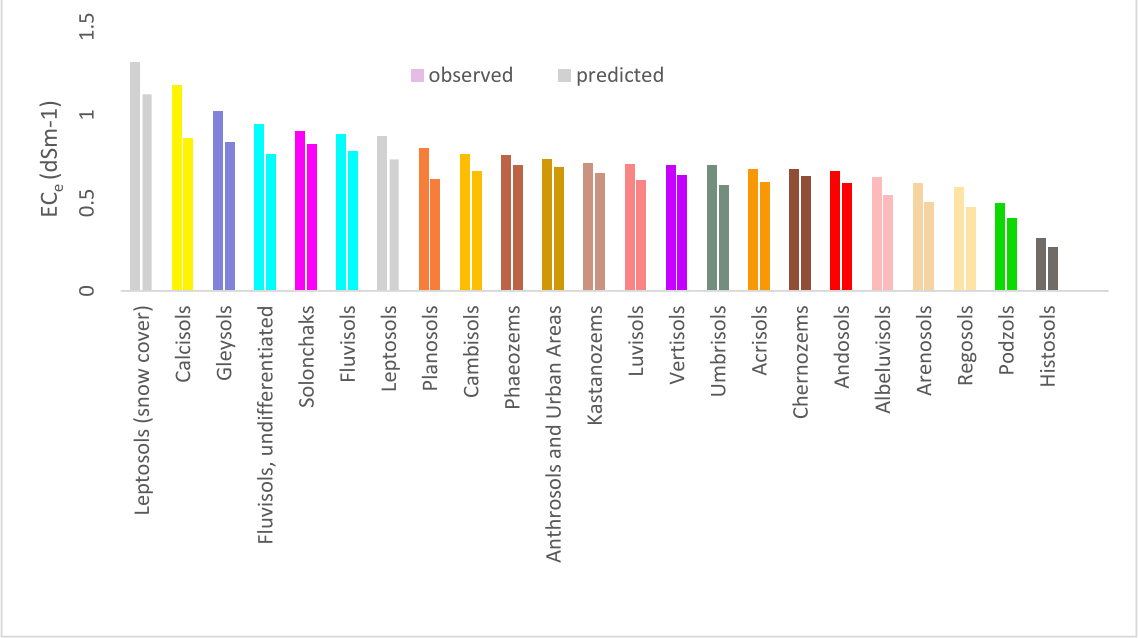


Fig. 5. Average LUCAS topsoil EC$_e$ Observed a) and Predicted b) in each of the Dominant soil type of the WRB (2022) Reference Soil Groups of Europe from the ESDB database (note that no Solonetz LUCAS samples are available in the dataset and Gypsisols have been removed).

Table 6
Share of land under EC$_e$ conditions in percentage (according to FAO Global Salinity Map classification).

NUTS 0	None	Slight	Moderate	Strong
AT	55.734 %	44.262 %	0.004 %	–
BE	95.520 %	4.480 %	–	–
BG	98.362 %	1.638 %	–	–
CY	43.250 %	56.744 %	0.005 %	–
CZ	95.721 %	4.279 %	–	–
DE	87.533 %	12.466 %	–	–
DK	96.391 %	3.609 %	0.001 %	–
EE	92.424 %	7.572 %	0.004 %	–
EL	78.530 %	21.469 %	–	–
ES	72.404 %	26.628 %	0.765 %	0.114 %
FI	99.920 %	0.080 %	–	–
FR	81.399 %	18.601 %	–	–
HR	71.250 %	28.747 %	0.003 %	–
HU	93.428 %	6.572 %	–	–
IE	22.326 %	77.665 %	0.009 %	–
IT	48.781 %	51.205 %	0.014 %	–
LT	99.273 %	0.726 %	0.001 %	–
LU	92.168 %	7.832 %	–	–
LV	99.124 %	0.873 %	0.002 %	–
MT	45.024 %	54.976 %	–	–
NL	76.214 %	23.785 %	0.001 %	–
PL	97.113 %	2.886 %	0.001 %	–
PT	99.068 %	0.932 %	–	–
RO	86.743 %	13.257 %	–	–
SE	99.841 %	0.159 %	–	–
SI	39.056 %	60.882 %	0.062 %	–
SK	87.735 %	12.265 %	–	–

considered affected by salinity. However, a slight saline topsoil, i.e., featuring EC$_e$ between 2 and 4 dS m^{-1}, already poses a constraint to agricultural profitability since many crops are classified as salt-sensitive and hence experience yield losses at such low EC$_e$ values (Kim & Park, 2024). Therefore, tracking the occurrence and distribution soils with slight high EC is crucial for land use planning and sustainable agricultural development (Hassani et al., 2020). Peatland and in general organic rich soils as land cover are not very well represented in the LUCAS sampling and the resulting values predicted in the map cannot be taken as true values for large parts of Europe, which are out of scope of the exercise. Another interesting observation is that many of the areas shown in green, i.e., slightly high EC$_e$ class, in Ireland are found extensively on peatlands, which, rain-fed bogs, are expected to present very low EC$_e$. These areas account for high uncertainty.

In Mediterranean semiarid Europe there are many saline soils cause of primary salinization, as well as in Atlantic and Northern Europe secondary salinization also occurs (Costantini & Lorenzetti, 2013; Daliakopoulos et al., 2016). The case of Ireland could be a clear example of this latter. In Ireland, the soil EC$_e$ distribution pattern can be paired with the Na^+ pattern from the Soil Geochemical Atlas (Fay et al., 2007) with higher Na^+ values for the northern regions of Donegal and the east coast. This seems to be associated to the parent material with fine-grained feldspars in greywackes, schist and granite that contain either Na^+, K^+ or Mg^{2+} (Fay et al., 2007) and other cation and anion ratios are also used. However, the soil salt distribution represented in this map, reflects more the Oceanic deposition through coastal salt marshes that play a significant role in the south west area. Prevailing south-westerly winds can transport oceanic salt inland where it is generally deposited onto gleysols, characterised by heavy textures and high water tables. The interplay between these two factors can prevent salts leaching throughout the soil profile resulting in slightly higher EC$_e$ values.

Some authors have found correlations between high EC$_e$ values and nitrogen salts (NH_4^+; NO_3^-) (Baldi et al., 2020; Tao et al., 2024). Excessive application of fertilizer is indeed considered one of the causes of high soil EC$_e$, particularly in soils under arid and semiarid conditions, and in soils under more humid climates with poor permeability (Gleysols and Stagnosols). We could expect high values of EC$_e$ in Irish topsoils because of the high level of N in soils poor drained. This has been a remarkable issue that brought the agricultural governance to adopt in force the EU Nitrates directive, even if with some waivers (in soils where potential N leaching is low, since Nitrates Directive origins from Directive of water protection. (https://www.teagasc.ie/environment/schemes–regulations/nitrates-derogation/). Of course the grazing system could also affect the nitrogen balance in soil. Unfortunately the EC parameter has not been measured in the survey of Irish Soil Information System (ISIS), so, we cannot make any references of comparisons with that. In Northern Ireland for instance McCormick et al. (McCormick et al., 2003) found correlation of EC$_e$ with sward nitrogen offtake. The 80 % of Irish agricultural land is grassland, in addition, practices of Nitrogen fertilization

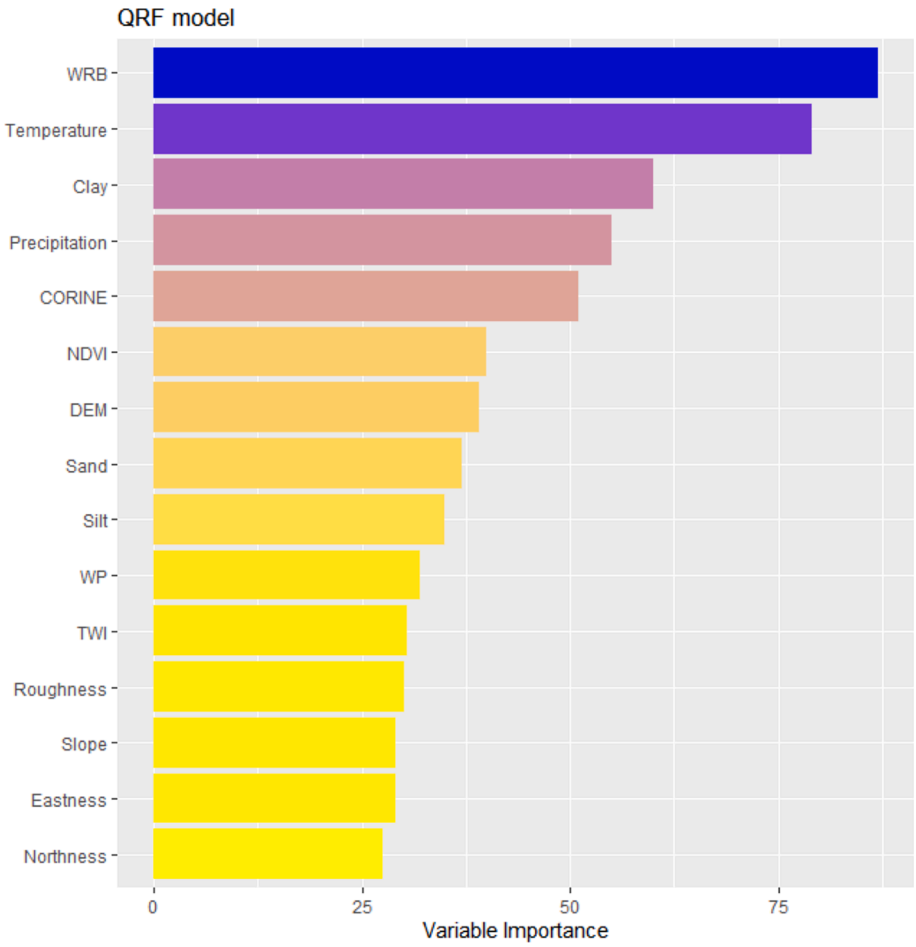


Fig. 6. The 15th most important variables in the machine-learning model.

in Ireland have been extensively adopted to increase the biomass productivity of the grassland, by using fertilizers, slurry, manure and even the sludges from dairy (Micha et al., 2023). According to Binner et al. (Binner et al., 2024), in a recent review of soil information in the Republic of Ireland, EC_e soil parameter has not been never considered. This might be a reflection of limited sampling distribution and density or biased averaging of values, this can be overcome in the future by the integration of national data (Cornu et al., 2023). Overall, the scarce spatial variability of LUCAS sampling points in Ireland coupled with the fact that EC_e can vary significantly over time and across management systems, may represent a source of error. Our results should address for further investigations on eventual correlations between soil EC_e, Nitrogen (and other fertilizers) and soil types.

When salinization occurs, it is often related to parent material, which plays a central role, as well as to soil management and land use (Tedeschi et al., 2023). However, the greatest risk in the context of desertification is secondary salinization, which mainly develops in irrigated areas where, in the most favorable case, the water given to crops is insufficient to leach out the salts it carries and, therefore, they build up into the soil. In the worst cases, however, irrigation water may dissolve salts from saline parent materials and lift groundwater tables with on-site as well as off-site soil salinization effects. In this regard, many areas around the Mediterranean have undergone tectonic uplift during the Tertiary and Quaternary periods, because of the collision between the African and European plates that raised the Alps, Pyrenees, Carpathians and Betic mountain ranges. As a result, there are many land areas with Tertiary or later marine sediments at the surface, and these are geologically young enough to still retain much of their marine salt content, with lowest leaching losses during dry climatic phases. Conversely, older rocks have lost much of their salts, and soils on them are at lower risk of salinization, and generally suffer less from poor groundwater quality. Regarding high groundwater tables, these are associated with low-lying areas, featuring poor lateral drainage towards the sea or the absence of large perennial rivers. Additionally, closeness to sea may play a relevant role in sea-leeward lands since the blowing of salt spray may spread over about 50 km. By using the LUCAS soil survey dataset, we are not able to capture the problematic issues in Hungary, a country in which previous investigation at country and local scale have shown high EC_e values in their studies. Therefore, we need higher sampling density or even targeting more points in which there high EC_e. In the topsoil EC_e map, the lands in Spain that have been revealed as most salt-affected match the ones that were previously identified (de Paz et al., 2004; MAPA, 2002) (supplementary materials Fig. 1d). The Ebro river valley was landlocked during a remarkable span of the Cenozoic era (Riba et al., 1983). Consequently, in the central stretch of the basin a lacustrine system developed, which was fed by the ancient Ebro and tributaries. Because of the arid-to-semiarid conditions in the area, already prevalent during the Tertiary (Cabrera et al., 2002), and the endorheic conditions, evaporites formed. Since the basin opened to the Mediterranean sea, the lake system was drained and the evaporites were mostly buried bellow air- and water-transported sediments throughout the late Cenozoic and Pleistocene. In Denmark, The sandy soils prevalent within the aeolian deposits, Salian moraine and Glacial outwash floodplains are relatively coarse-textured and the younger moraine landscape to the east is occupied by clay-rich (sandy-clay and clayey-sand) soils (Adhikari et al., 2013; Madsen and Jensen, 1992). This boundary between the Moraine and Glacial floodplain is also well captured by the EC_e map, with sandy soils representing lower EC_e values in comparison

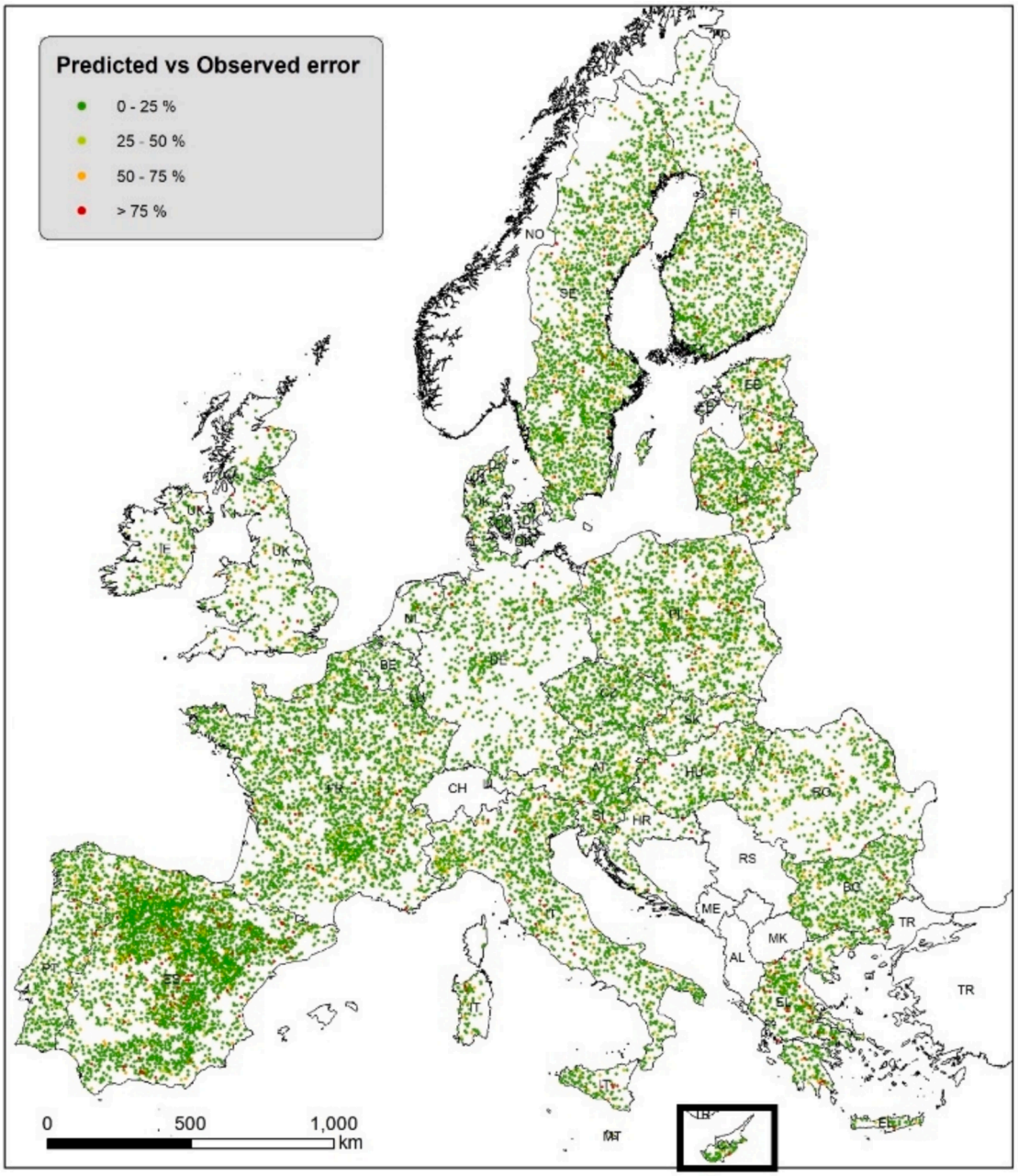


Fig. 7. Spatial distribution of the prediction error.

to the clay-rich soils. The areas of null EC_e values signify the major cities in Denmark. In Italy, potential trends can be explained by parent material effect and confirm that there is a slight increase in EC_e for inherent reasons. In accordance our results Italian soils are affected by high EC_e in almost all the Italian regions (Supplementary material Fig. 1e), (Dazzi & Lo Papa, 2013).

4.3. Limitations and future perspectives

The main limitations of the proposed methodology lies on two aspects, sample numerosity and depth and availability of a consistent set of EC_e data measured to provide a regression based PTF for the use of diluted $EC_{1:5}$. Regarding the sample numerosity and depth, the regional comparison of the EU27 map and the country and local scale map of EC_e and Salinity maps showed the limited ability to find some well recognized high EC_e regions, LUCAS soil module 2018 for the assessment of EC_e with around 18,000 samples for the whole EU27 area. Hence, a tailored sampling design and a further intensification of the sampling density will be required to monitor soil EC_e. However the soil monitoring methods (including stratified sampling) proposed in the Soil Monitoring Law (SML) proposal (an order of magnitude more samples) with the incorporation of the soil monitoring Schemes from Member States will increase the monitoring points and therefore be able to capture the spatial variability. Our findings provides a DSM exercise using state of the art methodology that will inform readers and trigger further research about the potential discrepancies between surface measurements and subsurface conditions. Thanks to the new soil monitoring framework will expect in the near future a substantial increase in data availability that will converge into a more accurate assessment of topsoil EC_e by 2035. To address this challenges, we suggest complementing the LUCAS soil module EC_e data with additional (legacy) measurements from deeper soil layers in areas known to exhibit

such characteristics. Where possible, we will refer to studies or datasets that provide information on sub-surface EC_e in these specific contexts. In a recent review by Paz et al. 2023), various strategies and relevant case studies were examined, focusing on distinct approaches (prevention, mitigation, and adaptation) to address the detrimental effects of salt-affected soils on agricultural productivity.

5. Conclusions

In this study we implemented an assessment of EC_e at EU level 500 m spatial resolution, using $EC_{1:5}$ data derived from the LUCAS (2018) converted using an empirically derived PTF and a set of environmental covariates fitted into a QRF model to produce an EC_e map. The final model (R^2 and RMSE) resulted in line with the literature for the independent test set of a continental wide soil dataset.

The availability of soil data can be sparse and inconsistent across a continental area, leading to an uneven sampling density, this study addressed the need for additional samples in some areas where salinity is considered a soil threat. The complexity of soil-forming factors at continental scale is indeed challenging, and uncertainties may arise due to the complex interactions between these factors and the non-linear nature of soil formation processes. Being aware that the model predictions may not fully capture the complexity of soil systems this research want to trigger further investigation on soil EC_e monitoring.

Policy actions at national, regional, and global scales have addressed the issues related to soil EC_e and limitation to plant growth. On the global scale, the United Nations adopted the Sustainable Development Goals (SDGs). Among these 17 SDGs, soil salinity plays a crucial role in 2, SDG 2 (reaching zero hunger) and 3 (reach good health and well-being) as it can limit crop productions; also SDG 14 to preserve (life below water) and SDG 15 life on land (SDG 15).

The European Commission recently proposed the EU Green Deal to make Europe the first climate-neutral continent by 2050, with a significant role of soils for reaching this ambition. As part of the EU Green Deal, the Farm to Fork Strategy aims to ensure sustainable food security. This assessment can contribute to the identification of potential EC_e "hot spots" across Europe and intensify its monitoring in coastal areas and in specific soil type due to its genesis and parent materials. The limitation of using topsoil data for mapping soil EC_e and also the development of the pedotransfer function, taking into account the first 20 cm of the soil vertical spatial variation can be also assess by using the same methodological approach across all soil horizons as soon as both EC_e and soil properties are available.

Modeling limitations will be addressed in future works including the use of a Regression-based pedotransfer function obtained through regression analysis of a dataset with laboratory measurements of saturated paste EC_e data, that considers also the gypsum so as to improve the soil EC_e characterization. The map of topsoil EC_e produced in this study can be used as input into the EUSO soil degradation dashboard (https://esdac.jrc.ec.europa.eu/esdacviewer/euso-dashboard/).

The output maps represent important soil property baselines for policy recommendations and soil monitoring in the European Union. To allow soil to continue in delivering services to ecosystems, science, societies and economies must work together to tackle the issue more vigorously than we are doing at this moment, we need to highlight the importance of contextualize EC_e trends related to sea water intrusion, irrigation, as well as the importance of soil monitoring on a detailed scale.

CRediT authorship contribution statement

Calogero Schillaci: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Simone Scarpa:** Visualization, Formal analysis, Data curation. **Felipe Yunta:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Aldo Lipani:** Writing – original draft, Software, Methodology, Data curation. **Fernando Visconti:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Gábor Szatmári:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Kitti Balog:** Writing – review & editing, Writing – original draft, Validation, Investigation, Data curation. **Triven Koganti:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Data curation. **Mogens Greve:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Giulia Bondi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology. **Georgios Kargas:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Supervision. **Paraskevi Londra:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology. **Fuat Kaya:** Writing – original draft, Visualization, Validation, Methodology. **Giuseppe Lo Papa:** Writing – review & editing, Validation. **Panos Panagos:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Resources, Supervision. **Luca Montanarella:** Supervision, Writing – review & editing. **Arwyn Jones:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

ESTAT and LUCAS soil funding DGs of the European Commission, and the reviewers for their effort to improve this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2025.117199.

Data availability

The LUCAS topsoil dataset used in this work was made available by the European Commission through the European Soil Data Centre managed by the Joint Research Centre (JRC), http://esdac.jrc.ec.europa.eu/.

References

Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., Greve, M. H., 2013. High-resolution 3-D mapping of soil texture in Denmark. Soil Sci. Soc. Am. J. 77 (3), 860–876.

Acosta, J.A., Faz, A., Jansen, B., Kalbitz, K., Martínez-Martínez, S., 2011. Assessment of salinity status in intensively cultivated soils under semiarid climate, Murcia, SE Spain. *J. Arid Environ.* 75, 1056–1066. https://doi.org/10.1016/j.jaridenv.2011.05.006.

Ajilogba, C.F., Walker, S. 2021. Climate Change Adaptation: Implications for Food Security and Nutrition. *African Handbook of Climate Change Adaptation* 735–754. DOI: 10.1007/978-3-030-45106-6_142.

Alexander, E.B., 1980. Bulk densities of California soils in relation to other soil properties. *Soil Sci. Soc. Am. J.* 44, 689–692. https://doi.org/10.2136/sssaj1980.03615995004400040005x.

Alkharabsheh, H.M., Seleiman MF, Hewedy OA, Battaglia ML, Jalal RS, Alhammad BA, Schillaci C, Ali N, Al-Doss A. 2021. Field crop responses and management strategies to mitigate soil salinity in modern agriculture: a review. *Agronomy 2021,* **11**: 2299. DOI: 10.3390/AGRONOMY11112299.

Allbed, A., Kumar, L., Aldakheel, Y.Y., 2014. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries:

applications in a date palm dominated region. *Geoderma* 230–231, 1–8. https://doi.org/10.1016/J.GEODERMA.2014.03.025.
Arslan, H., 2017. Determination of temporal and spatial variability of groundwater irrigation quality using geostatistical techniques on the coastal aquifer of Çarşamba Plain, Turkey, from 1990 to 2012. *Environ. Earth Sci.* 76, 38. https://doi.org/10.1007/s12665-016-6375-x.
Baldi, E., Quartieri, M., Muzzi, E., Noferini, M., Toselli, M., 2020. Use of in situ soil solution electric conductivity to evaluate mineral N in commercial orchards: preliminary results. *Horticulturae* 6, 39. https://doi.org/10.3390/horticulturae6030039.
Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., Panagos, P., 2019. Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. *Geoderma* 355, 113912. https://doi.org/10.1016/J.GEODERMA.2019.113912.
Ballabio, C., Panagos, P., Monatanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database. *Geoderma* 261, 110–123.
Barbanti, L., Adroher, J., Damian, J.M., Di Virgilio, N., Falsone, G., Zucchelli, M., Martelli, R., 2018. Assessing wheat spatial variation based on proximal and remote spectral vegetation indices and soil properties. *Italian J. Agronomy* 13, 21–30. https://doi.org/10.4081/ija.2017.1086.
Batlle-Sales, J., 2023. Salt-affected soils: a sustainability challenge in a changing world. *Ital. J. Agron.* https://doi.org/10.4081/ija.2023.2188.
Binner, H., Andrade, L., McNamara, M.E., 2024. Assessing synergies between soil research in the Republic of Ireland and European Union policies. *Environ. Challenges* 15, 100881. https://doi.org/10.1016/j.envc.2024.100881.
Castaldi, F., Halil Koparan, M., Wetterlind, J., Žydelis, R., Vinci, I., Özge Savaş, A., Kıvrak, C., Tunçay, T., Volungevičius, J., Obber, S., Ragazzi, F., Malo, D., Vaudour, E., 2023. Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. *ISPRS J. Photogramm. Remote Sens.* 199, 40–60. https://doi.org/10.1016/j.isprsjprs.2023.03.016.
Cornu, S., Keesstra, S., Bispo, A., Fantappie, M., van Egmond, F., Smreczak, B., Wawer, R., Pavlů, L., Sobocká, J., Bakacsi, Z., Farkas-Iványi, K., Molnár, S., Møller, A.B., Madenoglu, S., Feiziene, D., Oorts, K., Schneider, F., da Conceição, G. M., Mano, R., Garland, G., Skalský, R., O'Sullivan, L., Kasparinskis, R., Chenu, C., 2023. National soil data in <scp>EU</scp> countries, where do we stand? *Eur. J. Soil Sci.* 74. https://doi.org/10.1111/ejss.13398.
Corwin, D.L., 2021. Climate change impacts on soil salinity in agricultural areas. *Eur. J. Soil Sci.* 72, 842–862. https://doi.org/10.1111/EJSS.13010.
Costantini, E.A.C., Lorenzetti, R., 2013. Soil degradation processes in the Italian agricultural and forest ecosystems. *Ital. J. Agron.*. Page Press Publications 233–243. https://doi.org/10.4081/ija.2013.e28.
Çullu, M.A., Günal, H., 2018. Solonetz Soils–Solonchaks (Solonetz-Like Soils), 331–346. DOI: 10.1007/978-3-319-64392-2_23.
Curci, M., Lavecchia, A., Cucci, G., Lacolla, G., De Corato, U., Crecchio, C., 2020. Short-term effects of sewage sludge compost amendment on semiarid soil. *Soil Systems* 4, 48. https://doi.org/10.3390/soilsystems4030048.
Fay, D., G. Kramers, C. Zhang DM and EG. 2007. *Soil Geochemical Atlas of Ireland*. Teagasc and the Environmental Protection Agency.
Daliakopoulos, I.N., Tsanis, I.K., Koutroulis, A., Kourgialas, N.N., Varouchakis, A.E., Karatzas, G.P., Ritsema, C.J., 2016. The threat of soil salinity: A European scale review. *Sci. Total Environ.* 573, 727–739. https://doi.org/10.1016/j.scitotenv.2016.08.177.
Dazzi C, Lo Papa G. 2013. Soil Threats (in: "The Soils of Italy"), 205–245. DOI: 10.1007/978-94-007-5642-7_8.
Delgado, C., Pacheco, J., Cabrera, A., Batllori, E., Orellana, R., Bautista, F., 2010. Quality of groundwater for irrigation in tropical karst environment: The case of Yucatán, Mexico. *Agric Water Manag* 97, 1423–1433. https://doi.org/10.1016/j.agwat.2010.04.006.
Demir, Z., Tursun, N., Işik, D., 2019. Effects of different cover crops on soil quality parameters and yield in an apricot orchard. *Int. J. Agric. Biol.* 21, 399–408. https://doi.org/10.1590/1678-4499.2010197.
FAO. 2017. The Future of Food and Agriculture. *Food and Agriculture Organization of the United Nations* 1–52.
FAO. 2020. *Mapping of salt-affected soils*. Rome.
Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The shuttle radar topography mission. *Rev. Geophys.* 45. https://doi.org/10.1029/2005RG000183.
Fendrich, A.N., Van Eynde, E., Stasinopoulos, D.M., Rigby, R.A., Mezquita, F.Y., Panagos, P., 2024. Modeling arsenic in European topsoils with a coupled semiparametric (GAMLSS-RF) model for censored data. *Environ. Int.* 185, 108544. https://doi.org/10.1016/j.envint.2024.108544.
Fernandez-Ugalde O;, Scarpa S;, Orgiazzi A;, Panagos P, Van Liedekerke M, Jones MA&. 2022. LUCAS 2018 Soil Module Presentation of dataset and results.
Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. https://doi.org/10.1002/joc.5086.
Geilfus, C.-M., 2018. Review on the significance of chlorine for crop yield and quality. *Plant Sci.* 270, 114–122. https://doi.org/10.1016/j.plantsci.2018.02.014.
Gharaibeh, M.A., Albalasmeh, A.A., El Hanandeh, A., 2021. Estimation of saturated paste electrical conductivity using three modelling approaches: Traditional dilution extracts; saturation percentage and artificial neural networks. *Catena* 200, 105141. https://doi.org/10.1016/j.catena.2020.105141.
Gil-Márquez, J.M., Barberá, J.A., Andreo, B., Mudarra, M., 2017. Hydrological and geochemical processes constraining groundwater salinity in wetland areas related to evaporitic (karst) systems. A case study from Southern Spain. *J. Hydrol.* 544, 538–554. https://doi.org/10.1016/j.jhydrol.2016.11.062.
Hassani, A., Azapagic, A., Shokri, N., 2020. Predicting long-term dynamics of soil salinity and sodicity on a global scale. *Proc. Natl. Acad. Sci.* 117, 33017–33027. https://doi.org/10.1073/PNAS.2013771117.
He, Y., DeSutter, T., Hopkins, D., Jia, X., Wysocki, D.A., 2013. Predicting EC e of the saturated paste extract from value of EC 1:5. *Can. J. Soil Sci.* 93, 585–594. https://doi.org/10.4141/cjss2012-080.
Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* 6, e5518.
Herrero, J., Castañeda, C., Gómez-Báguena, R., 2022. A heritage agronomic study as a database for monitoring the soil salinity of an irrigated district in NE Spain. *Agronomy* 12, 126. https://doi.org/10.3390/agronomy12010126.
IBPES. 2018. The Assessment Report on Land degradation and restoration. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). DOI: 10.5281/ZENODO.3237393.
IUSS Working Group WRB. 2022. World Reference Base for Soil Resources. International soil classification system for naming soils and creating legends for soil maps. 4th edition. International Union of Soil Sciences (IUSS), Vienna, Austria.
Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., 2017. Satellite Thermography for Soil Salinity Assessment of Cropped Areas in Uzbekistan. *Land Degrad. Dev.* 28, 870–877. https://doi.org/10.1002/LDR.2670.
Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., Bui, E.N., Wilford, J., 2018. Soil salinity assessment through satellite thermography for different irrigated and rainfed crops. *Int. J. Appl. Earth Obs. Geoinf.* 68, 230–237. https://doi.org/10.1016/J.JAG.2018.02.004.
Ivushkin, K., Bartholomeus, H., Bregt, A.K., Pulatov, A., Kempen, B., de Sousa, L., 2019. Global mapping of soil salinity change. *Remote Sens. Environ.* 231, 111260. https://doi.org/10.1016/J.RSE.2019.111260.
Jones, A., Fernandez Ugalde, O., Scarpa S. 2020. LUCAS 2015 Topsoil Survey. JRC Technical Reports.
Kargas, G., Londra, P., Sgoubopoulou, A., 2020. Comparison of soil EC values from methods based on 1:1 and 1:5 soil to water ratios and ECe from saturated paste extract based method. *Water (switzerland)* 12. https://doi.org/10.3390/W12041010.
Kaya, F., Schillaci, C., Keshavarzi, A., Başayiğit, L., 2022. Predictive mapping of electrical conductivity and assessment of soil salinity in a western Türkiye Alluvial Plain. *Land* 11, 2148. https://doi.org/10.3390/land11122148.
Khorsandi, F., Yazdi, F.A., 2007. Gypsum and texture effects on the estimation of saturated paste electrical conductivity by two extraction methods. *Commun. Soil Sci. Plant Anal.* 38, 1105–1117. https://doi.org/10.1080/00103620701278120.
Khorsandi, F., Yazdi, F.A., 2011. Estimation of saturated paste extracts' electrical conductivity from 1:5 soil/water suspension and gypsum. *Commun. Soil Sci. Plant Anal.* 42, 315–321. https://doi.org/10.1080/00103624.2011.538885.
Kim, H.N., Park, J.H., 2024. Monitoring of soil EC for the prediction of soil nutrient regime under different soil water and organic matter contents. *Appl. Biol. Chem.* 67, 1. https://doi.org/10.1186/s13765-023-00849-4.
Kurunc, A., Ersahin, S., Sonmez, N.K., Kaman, H., Uz, I., Uz, B.Y., Aslan, G.E., 2016. Seasonal changes of spatial variation of some groundwater quality variables in a large irrigated coastal Mediterranean region of Turkey. *Sci. Total Environ.* 554–555, 53–63. https://doi.org/10.1016/j.scitotenv.2016.02.158.
Lekka, C., Petropoulos, G.P., Triantakonstantis, D., Detsikas, S.E., Chalkias, C., 2023. Exploring the spatial patterns of soil salinity and organic carbon in agricultural areas of Lesvos Island, Greece, using geoinformation technologies. *Environ. Monit. Assess.* 195, 391. https://doi.org/10.1007/s10661-023-10923-5.
Li, C., Gao, X., Li, S., Bundschuh, J., 2020. A review of the distribution, sources, genesis, and environmental concerns of salinity in groundwater. *Environ. Sci. Pollut. Res.* 27, 41157–41174. https://doi.org/10.1007/s11356-020-10354-6.
Lombardo, L., Saia, S., Schillaci, C., Mai, P.M., Huser, R., 2018. Modeling soil organic carbon with Quantile Regression: Dissecting predictors' effects on carbon stocks. *Geoderma* 318, 148–159. https://doi.org/10.1016/j.geoderma.2017.12.011.
Marzaioli, R., D'Ascoli, R., De Pascale, R.A., Rutigliano, F.A., 2010. Soil quality in a Mediterranean area of Southern Italy as related to different land use types. *Appl. Soil Ecol.* 44, 205–212. https://doi.org/10.1016/j.apsoil.2009.12.007.
Matthees, H.L., He, Y., Owen, R.K., Hopkins, D., Deutsch, B., Lee, J., Clay, D.E., Reese, C., Malo, D.D., DeSutter, T.M., 2017. Predicting soil electrical conductivity of the saturation extract from a 1:1 soil to water ratio. *Commun. Soil Sci. Plant Anal.* 48, 2148–2154. https://doi.org/10.1080/00103624.2017.1407780.
McBratney, A., Mendonça, S.M., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.
McCormick, S., Bailey, J.S., Jordan, C., Higgins, A., 2003. A potential role for soil electrical conductivity mapping in the site-specific management of grassland. *Precis. Agric.*. Brill | Wageningen Academic 393–397. https://doi.org/10.3920/9789086865147_059.
Mechri, B., Attia, F., Braham, M., Elhadj, S.B., Hammami, M., 2007. Agronomic application of olive mill wastewaters with phosphate rock in a semi-arid Mediterranean soil modifies the soil properties and decreases the extractable soil phosphorus. *J. Environ. Manage.* 85, 1088–1093. https://doi.org/10.1016/j.jenvman.2006.11.010.
Meinshausen, N., 2006. Quantile regression forests. *J. Mach. Learn. Res.* 7, 983–999. https://doi.org/10.5555/1248547.
Micha, E., Tsakiridis, A., Ragkos, A., Buckley, C., 2023. Assessing the effect of soil testing on chemical fertilizer use intensity: an empirical analysis of phosphorus fertilizer demand by Irish dairy farmers. *J. Rural. Stud.* 97, 186–191. https://doi.org/10.1016/j.jrurstud.2022.12.018.

Minasny, B., McBratney, A.B., Wadoux, A.-M.-J.-C., Akoeb, E.N., Sabrina, T., 2020. Precocious 19th century soil carbon science. *Geoderma Reg.* 22, e00306. https://doi.org/10.1016/j.geodrs.2020.e00306.

Mohanavelu, A., Naganna, S.R., Al-Ansari, N. 2021. Irrigation Induced Salinity and Sodicity Hazards on Soil and Groundwater: An Overview of Its Causes, Impacts and Mitigation Strategies. *Agriculture 2021, 11, Page 983* **11**: 983. DOI: 10.3390/AGRICULTURE11100983.

Montanarella, L., 2020. Soils and the European green deal. *Ital. J. Agron.* 15, 262–266. https://doi.org/10.4081/ija.2020.1761.

Montanarella, L., Panagos, P., 2021. Soil security for the European Union. *Soil Secur.* 4, 100009. https://doi.org/10.1016/j.soisec.2021.100009.

Moret-Fernández, D., Herrero, J., 2015. Effect of gypsum content on soil water retention. *J. Hydrol.* 528, 122–126. https://doi.org/10.1016/j.jhydrol.2015.06.030.

Mukhopadhyay, R., Sarkar, B., Jat, H.S., Sharma, P.C., Bolan, N.S., 2021. Soil salinity under climate change: challenges for sustainable agriculture and food security. *J. Environ. Manage.* 280, 111736. https://doi.org/10.1016/J.JENVMAN.2020.111736.

Munns, R., Tester, M., 2008. Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681. https://doi.org/10.1146/annurev.arplant.59.032607.092911.

Nabiollahi, K., Taghizadeh-Mehrjardi, R., Kerry, R., Moradian, S., 2017. Assessment of soil quality indices for salt-affected agricultural land in Kurdistan Province, Iran. *Ecol. Indic.* 83, 482–494. https://doi.org/10.1016/J.ECOLIND.2017.08.001.

Nikolaou, G., Neocleous, D., Christophi, C., Heracleous, T., Markou, M., 2020. Irrigation groundwater quality characteristics: a case study of cyprus. *Atmosphere* 11, 302. https://doi.org/10.3390/atmos11030302.

Obi, J.C., Ogban, P.I., Ituen, U.J., Udoh, B.T., 2014. Development of pedotransfer functions for coastal plain soils using terrain attributes. *Catena* 123, 252–262. https://doi.org/10.1016/j.catena.2014.08.015.

Omuto, C.T., Vargas, R.R., Elmobarak, A.A., Mapeshoane, B.E., Koetlisi, K.A., Ahmadzai, H., Abdalla, M.N., 2022. Digital soil assessment in support of a soil information system for monitoring salinization and sodification in agricultural areas. *Land Degrad. Dev.* 33, 1204–1218. https://doi.org/10.1002/ldr.4211.

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde O. 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil Science*. Blackwell Publishing Ltd, 140–153. DOI: 10.1111/ejss.12499.

Panagos, P., Van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., Montanarella, L., 2022. European Soil Data Centre 2.0: Soil data and knowledge in support of the <scp>EU</scp> policies. *Eur. J. Soil Sci.* 73. https://doi.org/10.1111/ejss.13315.

Pásztor, L., Laborczi, A., Bakacsi, Z., Szabó, J., Illés, G., 2018. Compilation of a national soil-type map for Hungary by sequential classification methods. *Geoderma* 311, 93–108. https://doi.org/10.1016/j.geoderma.2017.04.018.

Paz, A.M., Amezketa, E., Canfora, L., Castanheira, N., Falsone, G.G.M., Gould, I., Hristov, B., Mastrorilli, M., Ramos, T., Thompson, R., Costantini, E., 2023. Salt-affected soils: field-scale strategies for prevention, mitigation, and adaptation to salt accumulation. *Ital. J. Agron.* 18, 2166. https://doi.org/10.4081/ija.2023.2166.

Pe'er, G., Zinngrebe, Y., Moreira, F., Sirami, C., Schindler, S., Müller, R., Bontzorlos, V., Clough, D., Bezák, P., Bonn, A., Hansjürgens, B., Lomba, A., Möckel, S., Passoni, G., Schleyer, C., Schmidt, J., Lakner, S., 2019. A greener path for the EU common agricultural policy. *Science* 365, 449–451. https://doi.org/10.1126/science.aax3146.

Prăvălie, R., Patriche, C., Borrelli, P., Panagos, P., Roşca, B., Dumitraşcu, M., Nita, I.A., Săvulescu, I., Birsan, M.V., Bandoc, G., 2021. Arable lands under the pressure of multiple land degradation processes. A global perspective. *Environ. Res.* 194, 110697. https://doi.org/10.1016/j.envres.2020.110697.

Richer-de-Forges, A.C., Chen, Q., Baghdadi, N., Chen, S., Gomez, C., Jacquemoud, S., Martelet, G., Mulder, V.L., Urbina-Salazar, D., Vaudour, E., Weiss, M., Wigneron, J.-P., Arrouays, D., 2023. Remote sensing data for digital soil mapping in french research—a review. *Remote Sens. (Basel)* 15, 3070. https://doi.org/10.3390/rs15123070.

Robinson, D.A., Thomas, A., Reinsch, S., Lebron, I., Feeney, C.J., Maskell, L.C., Wood, C. M., Seaton, F.M., Emmett, B.A., Cosby, B.J., 2022. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. *Sci. Rep.* 12, 7085. https://doi.org/10.1038/s41598-022-11099-7.

Saygın, F., Aksoy, H., Alaboz, P., Dengiz, O., 2023. Different approaches to estimating soil properties for digital soil map integrated with machine learning and remote sensing techniques in a sub-humid ecosystem. *Environ. Monit. Assess.* 195, 1061. https://doi.org/10.1007/s10661-023-11681-0.

Schillaci, C., Jones, A., Vieira, D., Munafò, M., Montanarella, L., 2023. Evaluation of the United Nations sustainable development goal 15.3.1 indicator of land degradation in the European Union. *Land Degrad. Dev.* 34, 250–268. https://doi.org/10.1002/ldr.4457.

Shahid, S.A., Zaman, M., Heng, L. 2018. Soil Salinity: Historical Perspectives and a World Overview of the Problem. *Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques* 43–53. DOI: 10.1007/978-3-319-96190-3_2.

Shi, T., He, L., Wang, R., Li, Z., Hu, Z., Wu, G., 2023. Digital mapping of heavy metals in urban soils: A review and research challenges. *Catena* 228, 107183. https://doi.org/10.1016/j.catena.2023.107183.

Slavich, P., Petterson, G., 1993. Estimating the electrical conductivity of saturated paste extracts from 1:5 soil, water suspensions and texture. *Soil Res.* 31, 73. https://doi.org/10.1071/SR9930073.

Stavi, I., Thevs, N., Priori, S., 2021. Soil Salinity and sodicity in drylands: a review of causes, effects, monitoring, and restoration measures. *Front. Environ. Sci.* 330. https://doi.org/10.3389/FENVS.2021.712831.

Sumner, M., 1993. Sodic soils - new perspectives. *Soil Res.* 31, 683. https://doi.org/10.1071/SR9930683.

Szabolcs I. 1974. *Salt Affected Soils in Europe. The Hague and RISSAC – Budapest*.

Szatmári, G., Bakacsi, Z., Laborczi, A., Petrik, O., Pataki, R., Tóth, T., Pásztor, L., 2020. Elaborating hungarian segment of the global map of salt-affected soils (GSSmap): national contribution to an international initiative. *Remote Sens. (Basel)* 12, 4073. https://doi.org/10.3390/rs12244073.

Tao, Y., Xie, W., Xu, L., Zhang, L., Wang, G., Wang, X., Shi, C., 2024. The characteristics of soil salinization effects on nitrogen mineralization and nitrification in upland fields. *Front. Environ. Sci.* 12. https://doi.org/10.3389/fenvs.2024.1369554.

Tarolli, P., Luo, J., Park, E., Barcaccia, G., Masin, R., 2024. Soil salinization in agriculture: mitigation and adaptation strategies combining nature-based solutions and bioengineering. iScience 27, 108830. https://doi.org/10.1016/j.isci.2024.108830.

Tedeschi, A., Schillaci, M., Balestrini, R., 2023. Mitigating the impact of soil salinity: recent developments and future strategies. *Ital. J. Agron.* https://doi.org/10.4081/ija.2023.2173.

Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., Tóth, G., 2015. New generation of hydraulic pedotransfer functions for Europe. *Eur. J. Soil Sci.* 66, 226–238. https://doi.org/10.1111/ejss.12192.

Tóth G, Jones A, Montanarella L. 2013. *LUCAS Topsoil Survey methodology, data and results*. DOI: 10.2788/97922.

Trevisani, S., Teza, G., Guth, P.L., 2023. Hacking the topographic ruggedness index. *Geomorphology* 439, 108838. https://doi.org/10.1016/j.geomorph.2023.108838.

Van Eynde, E., Fendrich, A.N., Ballabio, C., Panagos, P., 2023. Spatial assessment of topsoil zinc concentrations in Europe. *Sci. Total Environ.* 892, 164512. https://doi.org/10.1016/j.scitotenv.2023.164512.

Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. *Geoderma* 291, 55–64. https://doi.org/10.1016/j.geoderma.2016.12.017.

Verma, D.K., Bhunia, G.S., Shit, P.K., Kumar, S., Mandal, J., Padbhushan, R., 2017. Spatial variability of groundwater quality of Sabour block, Bhagalpur district (Bihar, India). *Appl. Water Sci.* 7, 1997–2008. https://doi.org/10.1007/s13201-016-0380-9.

Visconti, F., de Paz, J.M., Rubio, J.L., 2010. What information does the electrical conductivity of soil water extracts of 1 to 5 ratio (w/v) provide for soil salinity assessment of agricultural irrigated lands? *Geoderma* 154, 387–397. https://doi.org/10.1016/j.geoderma.2009.11.012.

Visconti, F., de Paz, J.M., 2016. Electrical conductivity measurements in agriculture: the assessment of soil salinity. *New Trends Develop. Metrol.* https://doi.org/10.5772/62741.

Visconti, F., Salvador, A., Navarro, P., de Paz, J.M., 2019. Effects of three irrigation systems on 'Piel de sapo' melon yield and quality under salinity conditions. *Agric. Water Manag.* 226. https://doi.org/10.1016/j.agwat.2019.105829.

Vittori Antisari L, Speranza M, Ferronato C, De Feudis M, Vianello G, Falsone G. 2020. Assessment of water quality and soil salinity in the agricultural coastal plain (Ravenna, North Italy). *Minerals 2020, Vol. 10, Page 369* **10**: 369. DOI: 10.3390/MIN10040369.

Wadoux, A.-M.-J.-C., Minasny, B., McBratney, A.B., 2020a. Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth Sci. Rev.* 210, 103359. https://doi.org/10.1016/j.earscirev.2020.103359.

Wadoux, A.M.J.C., Minasny, B., McBratney, A.B., 2020b. Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth Sci. Rev.*. Elsevier b. v. https://doi.org/10.1016/j.earscirev.2020.103359, 103359.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Mar. Geod.* 30, 3–35. https://doi.org/10.1080/01490410701295962.

Wong, V.N.L., Greene, R.S.B., Dalal, R.C., Murphy, B.W., 2010. Soil carbon dynamics in saline and sodic soils: a review. *Soil Use Manag.* 26, 2–11. https://doi.org/10.1111/j.1475-2743.2009.00251.x.

Wu, W., Zucca, C., Muhaimeed, A.S., Al-Shafie, W.M., Fadhil Al-Quraishi, A.M., Nangia, V., Zhu, M., Liu, G., 2018. Soil salinity prediction and mapping by machine learning regression in Central Mesopotamia, Iraq. *Land Degrad. Dev.* 29, 4005–4014. https://doi.org/10.1002/LDR.3148.

Xiao, C., Ji, Q., Chen, J., Zhang, F., Li, Y., Fan, J., Hou, X., Yan, F., Wang, H., 2023. Prediction of soil salinity parameters using machine learning models in an arid region of northwest China. *Comput. Electron. Agric.* 204, 107512. https://doi.org/10.1016/j.compag.2022.107512.