Science of the Total Environment 1047 (2026) 182021

Contents lists available at ScienceDirect

Science of the Total Environment

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

Modeling soil solution electrical conductivity across Europe

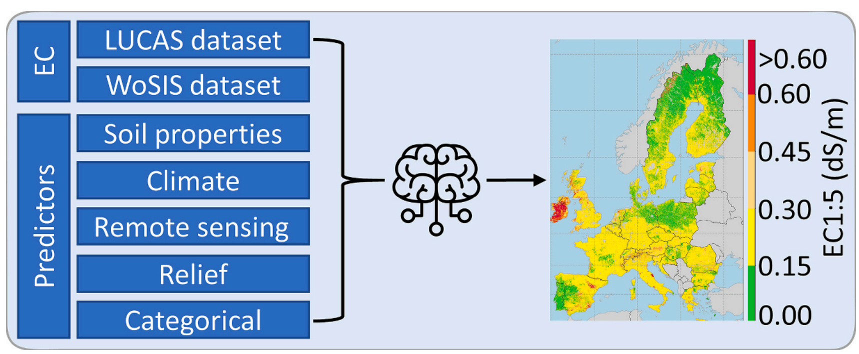
Mohammad Aziz Zarif [a,b], Amirhossein Hassani [c], Mehdi H. Afshar [a,b], Panos Panagos [d], Inma Lebron [e], David A. Robinson [e], Dani Or [f], Nima Shokri [a,b,*]

[a] Institute of Geo-Hydroinformatics, Hamburg University of Technology, 21073, Hamburg, Germany
[b] United Nations University Hub on Engineering to Face Climate Change at the Hamburg University of Technology, United Nations University Institute for Water, Environment and Health (UNU-INWEH), Hamburg, Germany
[c] The Climate and Environmental Research Institute NILU, P.O. Box 100, Kjeller, 2027, Norway
[d] European Commission, Joint Research Centre (JRC), Ispra, Italy
[e] UK Centre for Ecology & Hydrology, Bangor, UK
[f] Department of Civil and Environmental Engineering, University of Nevada, Reno, NV, USA

HIGHLIGHTS

- New models were developed to project soil electrical conductivity (EC) across EU + UK.
- Results include 1 km gridded soil EC data with uncertainty.
- Forward feature selection was applied to find key predictors for EC estimation.
- Models trained on ~40,000 EC points from LUCAS and ~20,000 from WoSIS datasets.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Soil salinization
Electrical conductivity
Machine learning

ABSTRACT

Soil salinization, referring to the excessive accumulation of soluble salts in soils, adversely influences nutrient cycling, biodiversity, soil structure, crop production, soil health, and ecosystem functioning. Accurately assessing soil salinity via electrical conductivity (EC) is key to mitigating its impacts. Thus, developing predictive tools for soil EC at regional and continental scales is essential for sustainable soil management. Here, we apply machine learning models to predict soil EC in the European Union (EU) and United Kingdom (UK) soils using different environmental factors like soil, climate, topography, and satellite data as predictors. The model is trained by ≈40,000 soil EC data points from the 2015 and 2018 Land Use/Cover Area Frame Survey data (LUCAS) surveys, complemented by the EC observations from World Soil Information Services (WoSIS) dataset. To improve the model performance, a forward feature selection technique was used resulting in selection of 17 covariates out of initially 34 predictors. The final selected XGBoost model achieved R^2 values of 0.68, 0.6, and 0.63 for the training, internal testing, and independent validation datasets, respectively. For the year 2018, we estimate ≈21.7 Mha of EU + UK land exceeds an EC of 0.6 dS/m (at a 1:5 soil to water ratio, the so-called $EC_{1:5}$). This estimate should be interpreted as elevated predicted $EC_{1:5,}$ rather than a direct estimate of soils meeting

* Corresponding author at: Institute of Geo-Hydroinformatics, Hamburg University of Technology, 21073, Hamburg, Germany.
E-mail address: nima.shokri@tuhh.de (N. Shokri).

https://doi.org/10.1016/j.scitotenv.2026.182021
Received 18 July 2025; Received in revised form 27 May 2026; Accepted 30 June 2026
Available online 7 July 2026

protosalic diagnostic criteria. The output of the predictive model consists of a gridded dataset that illustrates the spatial distribution of $EC_{1:5}$ throughout the study area for the year 2018, along with an associated uncertainty map with a spatial resolution of 1 km.

1. Introduction

Soil salinization is an increasingly critical issue within the broader context of global soil health and degradation. Saline soils are characterized by excessive accumulation of water-soluble salts (Shokri et al., 2024), while sodic soils are an allied problem specific to the amount of sodium (Na^+) relative to other cations. Soil salinity is often measured through the electrical conductivity (EC) of the soil solution, typically expressed in deciSiemens per meter (dS/m) (Richards, 1954). Soil salinization hinders the growth of natural ecosystems, soil fertility, agricultural productivity, and leads to osmotic stress and nutritional imbalance (Peñuelas et al., 2013). Additionally, it is a critical environmental factor influencing bacterial diversity (Lozupone and Knight, 2007) and can compromise soil structure, affecting its physical characteristics such as porosity and permeability (Rengasamy et al., 2003). Furthermore, soil salinization hinders the soil's capacity to sequester carbon, thereby increasing greenhouse gas emissions that affect the climate (Hassani et al., 2024). The process of soil salinization also affects water evaporation from soil (Shokri-Kuehni et al., 2017) and impacts transpiration rates in salt-stressed plants, thereby affecting the water cycle and a wide range of hydrologic processes (Bresler et al., 2012). Salt storms originating from salinized regions pose significant risks to human health and ecosystems (Hassani et al., 2020a).

While soil salinization is a global environmental issue, it is observed more in arid and semi-arid regions where precipitation is less than evapotranspiration, such as the Middle East, South America, and Australia (Hassani et al., 2020b). Although not a widespread problem in Europe as far as we know, soil salinization is acknowledged as an important factor contributing to soil degradation (Shokri et al., 2025), especially in the Mediterranean region (Daliakopoulos et al., 2016). Moreover, soil salinization is an insidious problem that emerges over time, often unobserved. Saline and sodic soils are difficult to map reliably because many of the processes that create salinity are highly dynamic and strongly seasonal. As a result, soil maps tend to depict only the well-established, persistently saline areas, while more ephemeral expressions of salinity frequently fall outside the diagnostic thresholds at certain times of year. For example, a soil profile may fail to meet the salic-horizon criteria following winter rains, yet the same profile may clearly qualify during the dry summer months when salts re-accumulate at the surface. Consequently, areas with protosalic conditions (ECe $\geq$ 4 dS/m) may be underrepresented in salinity maps when EC values fluctuates seasonally or only narrowly miss the formal diagnostic limits at the time of sampling . This seasonal instability means that current maps may under-represent the emerging extent of soil salinity across Europe and remain biased towards long-standing, strongly developed saline soils.

As indicated in the recent Mission Board's report for Soil Health and Food to the European Commission (Veerman et al., 2020) and the EU Soil Observatory dashboard, 60–70% of the soil in the European Union are considered as "unhealthy" (Panagos et al., 2024) which requires serious actions to "put Europe on a trajectory towards sustainable land and soil management". For soil sustainability in Europe, the Mission Board's implementation plan requests to reduce desertification and salinization through the restoration of 50% of the degraded land by 2030 as the first objective. Achieving this aim requires a thorough understanding of the current and future extent and dynamics of soil salinization across Europe under varying conditions.

Several studies have been undertaken to assess soil salinity throughout Europe. One of the pioneering efforts was by Szabolcs (1974), who created a map of salt-affected soils for Europe at a scale of 1:1,000,000, relying on expert insights and national soil surveys. This map was later revised by Tóth et al. (2008), who used data from the European Soil Database (ESDB, 2004) available in European Soil Data Centre (ESDAC) (Panagos et al., 2022), to enhance the continental soil salinity map. They relied on the distribution of Solonetz and Solonchaks, which have high salinity, and potentially overlooked emerging salinity hotspots. Although these maps give useful information about how soil salinity is spread across the EU, they do not show how it changes over time. Also, they are based on old and limited expert observations with coarse spatial resolution. These shortcomings necessitate an updated mapping approach for soil salinity that provides an analysis framework for both state and ultimately change across the EU.

On the other hand, several recent global studies provide information on the extent of saline soils across the EU. These studies apply ML algorithms and use EC of the saturated soil extract (ECe) from the World Soil Information Service (WoSIS) dataset as the target variable. Ivushkin et al. (2019) developed a categorized soil salinity map with classes ranging from non-saline to extremely saline; however, it did not provide per-cell salinity values. The studies conducted by Hassani et al. (2020b) and Wang et al. (2024) provide gridded datasets of predicted soil salinity at spatial resolutions of 1 km and 10 m, respectively. These studies give global estimates; however, their predictions for Europe have much uncertainty because the data points used to train the ML models are mainly spread out of the EU. These limitations highlight the importance of creating an updated, data-driven soil salinity map utilizing soil salinity measurements across the EU.

Regular monitoring is essential to prevent and address the adverse impacts of soil salinization. Various methodologies have been utilized to detect and monitor soil salinity, including conventional field sampling techniques, geophysical methods, remote sensing imagery (RSI), and the application of machine learning (ML) algorithms. Field measurement represents the traditional approach to soil assessment, which includes collecting samples in the field followed by laboratory analysis (Shokri et al., 2024). This method is effective for tracking changes in soil salinity at a local level or over short periods. However, due to the extensive demands of fieldwork, laboratory analysis, and increased costs, instead of measuring ECe, more rapid methods are sought. $EC_{1:5}$, which is less involved in the laboratory, is seen as a compromise for reconnaissance surveys at national scales.

Numerous studies have employed optical, thermal, radar, and hyperspectral sensors mounted on satellites such as Landsat, Sentinel, MODIS, and PRISMA to develop soil salinity indices and monitor the spatiotemporal variations in soil salinity. For instance, Alqasemi et al. (2021) used Landsat 8 OLI imagery to estimate salinity and vegetation indices, while developing a statistical model to combine the ground truth data and the indices developed to quantify and monitor soil salinity at the coast of Abu Dhabi in the United Arab Emirates. Similarly, Celleri et al. (2022) employed RSI to generate continuous soil salinity maps and to explain the effect of factors such as topography, land use, and climate on salinity distribution.

Recently, ML algorithms have been used to correlate soil salinity (mainly EC) with various environmental covariates to predict soil salinity levels in regions where direct measurement of EC is not available (Hassani et al., 2020b, 2021; Ivushkin et al., 2019; Wang et al., 2024). The commonly used ML algorithms for this purpose are Random Forest, Support Vector Regression, and XGBoost. ML algorithms are suitable for soil salinity mapping because they capture the relationships between EC and environmental covariates which are often complex and nonlinear. In addition, ML can integrate multisource predictors, handle noisy measurements, and is capable of generating spatially continuous maps in data sparse regions (Wadoux et al., 2020). Moreover, an integrated space-air-ground observation concept is increasingly used in the context of ML to quantify soil salinity, where satellites provide regional coverage and temporal continuity, aerial surveys supports high resolution observations, and ground measurements are used for calibration and validation purposes, resulting in more reliable salinity estimates (Qi et al., 2021).

Only with the recent availability of the Land Use/Cover Area Frame Survey data (LUCAS) https://esdac.jrc.ec.europa.eu/projects/lucas, which provides georeferenced soil observations and soil property information across EU Member States and the UK (Orgiazzi et al., 2018), has it become possible to accurately map soil salinity throughout these regions (Schillaci et al., 2025). The repeat surveys lay the foundation for working towards dynamic temporal assessment, which requires a modeling and mapping framework. Given the significant consequences of soil salinization on soil health and degradation (Afshar et al., 2025), along with the uncertainties regarding both the current situation and future trends of salinization in Europe under various scenarios, this study aims to improve the regional-scale prediction of topsoil $EC_{1:5}$ across EU and UK. The specific aims of the study include identifying the primary factors contributing to soil EC and developing ML models to predict and evaluate the topsoil $EC_{1:5}$ across EU and UK. The results help to find areas with elevated $EC_{1:5}$, which is important for formulating action plans to deal with potential soil salinization.

2. Methodology

2.1. Electrical conductivity data

Soil salinity is typically assessed through the EC of various soil-to-water ratios (Richards, 1954). Salinity offers a broad assessment of electrolyte concentration and can serve to indicate more specific and challenging problems like sodicity. In this research, we have selected EC as the primary variable for our ML model, utilizing data from two sources: LUCAS and WoSIS (Fig. 1a). It is worth mentioning that, initially, the model was trained only on the LUCAS dataset, considering $EC_{1:5}$ as the target variable. However, due to the narrow range and skewed distribution of $EC_{1:5}$ values in the LUCAS database, model performance was weak across training, testing, and validation sets (Fig. S1). To improve the predictive accuracy of the model, ECe observations from the WoSIS dataset were incorporated into the target variable, and the model was retrained. It should be mentioned that the inclusion of the WoSIS data was intended to broaden the EC range available for model training rather than providing a geographically representative European dataset.

We primarily relied on EC measurements from the LUCAS survey conducted in 2015 (Fig. 1c) and 2018 (Fig. 1b), which records $EC_{1:5}$ for the topsoil layer up to 20 cm deep (ISO 11265:1994, 1994). The LUCAS survey, developed by Eurostat, aims to monitor land use and land cover across European Union Member States' territories by providing a harmonized database of around 300,000 observations (Ballabio et al., 2016). The LUCAS Soil Module includes less than 10% of the surveyed points and requests the sampling and analysis of topsoil (0-20 cm).

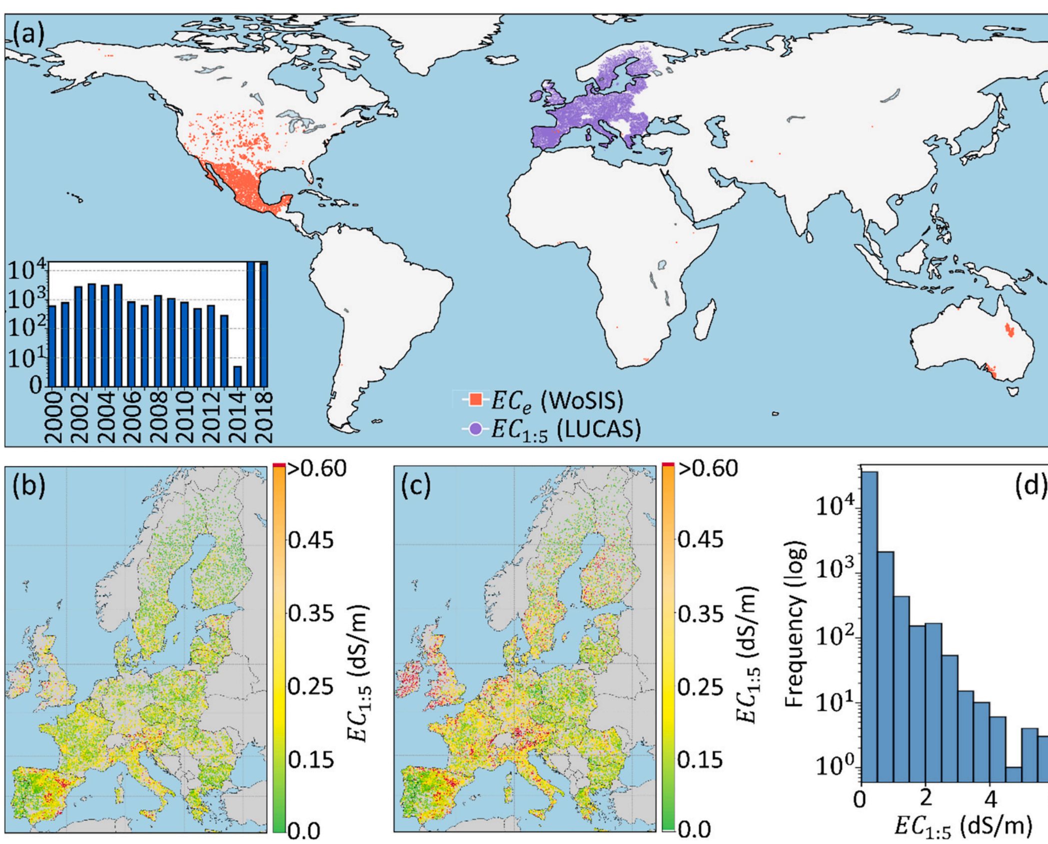


Fig. 1. Spatial and temporal distribution of electrical conductivity (EC) measurements used in this study, derived from the World Soil Information Services (WoSIS) and Land Use/Cover Area Frame Survey (LUCAS) datasets. The map depicts the geographical distribution of EC observation sites, each characterized by varying soil-to-water ratios, indicated by different colours and symbols. The LUCAS observation points, located within the European Union, were measured in a soil-to-water ratio of 1:5, while the WoSIS observation points represents electrical conductivity of the saturated paste extract (ECe) Additionally, the bar graph represents the temporal distribution of these measurements, incorporating data from the LUCAS surveys conducted in 2015 and 2018 (a). Spatial distribution of measured $EC_{1:5}$ from the year 2018 (b) and the year 2015 (c) LUCAS survey, along with the histogram of the measured $EC_{1:5}$ for the years 2015 and 2018 of the LUCAS survey (d).

Conducted every three years since 2009, LUCAS Soil Module has contributed to the unique EU harmonized soil database of physical, chemical and biological properties, modeling and assessments, ecosystem service analysis and development of indicators (Orgiazzi et al., 2018). LUCAS Soil Module provides $EC_{1:5}$ measurements for the years 2015 (21,859 points) and 2018 (18,975 points), however, a considerable number of observation points indicate low $EC_{1:5}$ levels (Fig. 2d), with 94% of the points measuring $EC_{1:5}$ below 0.6 dS/m. Furthermore, the $EC_{1:5}$ data exhibits a skewness of 13. These factors complicate the development of a model capable of accurately predicting salinity levels across the EU under different scenarios, especially in the cases of high soil salinity.

To overcome these challenges, we integrated the global EC of saturated soil extract data (ECe) from the World Soil Information Service (WoSIS) into our model (Batjes, 2025). WoSIS serves as a repository that offers quality-assured soil profile data to facilitate digital soil mapping, inform environmental policy, and support decision-making through data standardization (Ribeiro et al., 2018). This open-access database is regularly updated and includes information on the physical and chemical characteristics of soil. Given the availability of predictors, we opted to include EC measurements collected after the year 2000 from the WoSIS dataset (19,776 points).

2.2. Model predictors

The model predictors were chosen based on their relevance to soil salinity, following the SCORPAN framework, which is commonly used in digital soil mapping to guide the selection of predictors (Schillaci et al., 2025). The SCORPAN model is used to define the relationship between soil properties and environmental factors, such as existing soil properties, climate, organisms, relief, parent material, age, and neighbourhood (McBratney et al., 2003). We have classified the predictors into five main categories: soil physical properties, climate, remotely sensed imagery products, relief, and categorical variables (Table S1).

The SoilGrids 250 dataset was used to obtain soil physical properties, including the percentages of silt, sand, and clay (Hengl et al., 2017). This dataset provides a range of depth averaged soil physical and chemical properties at seven depths. For the prediction phase, soil texture values of 15–30 cm depth were used as a proxy to estimate the values at 20 cm, which is the topsoil depth considered in this study. Elevation, slope degree, aspect, curvature, downslope curvature, upslope curvature, topographic wetness index, topographic positioning index, and valley bottom flatness were downloaded from the Global DEM derivatives dataset (Hengl, 2018) derived from the MERIT DEM (Yamazaki et al., 2018). The Topographic Roughness Index, with a spatial resolution of 1 km, was obtained from Amatulli et al. (2018).

The climatic predictors were extracted from the Terra climate dataset. The dataset provides the monthly climate and climatic water balance variables at a spatial resolution of 4 km (Abatzoglou et al., 2018). For maximum and minimum temperature, wind speed, soil moisture, and the Palmer drought severity index, we calculated the 10-year average of monthly values. On the other hand, for precipitation, potential and actual evapotranspiration, and water deficit, we computed the 10-year average of the yearly accumulated value. MODIS datasets were used to obtain remote sensing imagery attributes. The yearly averaged values of Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) (Didan, 2021), Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) (Myneni et al., 2021), Gross Primary Production (GPP) and Net Primary Production (NPP) (Running and Zhao, 2021), and day time Land Surface Temperature (LST) (Wan et al., 2021) were included as input variables in the ML model.

Several categorical predictors were included in the model, including WRB soil classes (Hengl et al., 2017), parent material from the lithologic classification (Hartmann and Moosdorf, 2012), and soil-to-water ratio (SWR), which accounts for the EC extraction protocol. EC observations used in this study originate from two laboratory extraction protocols, $EC_{1:5}$ and ECe. These measures are not directly comparable because their relationship varies nonlinearly with soil properties and salt composition. Therefore, we do not apply a fixed conversion factor between $EC_{1:5}$ and ECe. Instead, extraction type was included as a categorical predictor in the model, namely soil-to-water ratio (SWR). LUCAS observations were assigned to the $EC_{1:5}$ category, and WoSIS observations were assigned to the ECe category. This allowed the model to account for the method-related differences while learning relationship between EC and the environmental covariates. For spatial prediction, the extraction-type predictor was set to the $EC_{1:5}$ category for all prediction cells; therefore, the final mapped product represents predicted $EC_{1:5}$ across EU and UK. Upper and lower depths of the soil sampling profiles were also included as predictors in the model.

The environmental predictors used in this study had different native spatial resolutions, as reported in Table S1. These predictors were not resampled to a common resolution before model training. Instead, each predictor was kept at its native resolution, and values were extracted from the relevant raster layer at the EC observation coordinates. For spatial prediction, a 1 km grid was created across EU and UK. The centroid of each grid cell was used as the predictor location, and predictor values were extracted from the native-resolution raster layers at these centroid points.

Given the relationship between soil organic carbon SOC (Hassani et al., 2024), soil pH (Han et al., 2024), and soil salinity, these factors could have been included as predictors in our model. However, we decided to exclude them due to specific challenges. Although high-resolution gridded datasets for SOC and soil pH are available, they are not provided on an annual basis, which is important for the current model because the observation points used in the analysis are distributed across the period from 2000 to 2018. In addition, these datasets remain limited in both global and continental coverage. Furthermore, SOC and soil pH frequently exhibit correlations with other soil characteristics and environmental variables that we have already incorporated as predictors in our analysis. Therefore, including SOC and soil pH as predictors could introduce multicollinearity, thereby adversely impacting the model's predictive performance.

2.3. Model setup

ML algorithms were utilized to assess soil salinity throughout Europe. This process involves a systematic modeling workflow (see Fig. 2) that begins with the acquisition and preparation of data. The dataset consists of 34 environmental covariates as predictors and one target variable, i.e., EC. A data matrix was created with 60,610 rows, with each row corresponding to an observation point, and 35 columns representing both the predictors and the target variable. To ensure independent final evaluation, 10% of the observation points were first reserved as an independent validation dataset. This subset was not used during model training, hyperparameter optimization, or feature selection. The remaining 90% of the observations formed the model development dataset. This model development dataset was then divided into an 80% training subset and a 20% internal testing subset, corresponding to 72% and 18% of the original dataset, respectively. To ensure statistical consistency, stratified sampling was employed by categorizing EC values into 10 bins (Fig. S2) and selecting training and testing datasets that reflect similar statistical distributions.

To determine the most significant predictors for estimating soil EC, hyperparameter optimization and feature selection were implemented during the model training phase. Previous research either neglected feature selection techniques or utilized recursive feature elimination to discard the least significant variables from the predictor set (Hassani et al., 2020b, 2021; Ivushkin et al., 2019). We employed a more sophisticated method of feature engineering, known as forward feature selection, in our analysis. This begins with an empty variable list and progressively incorporates predictors into the model based on their

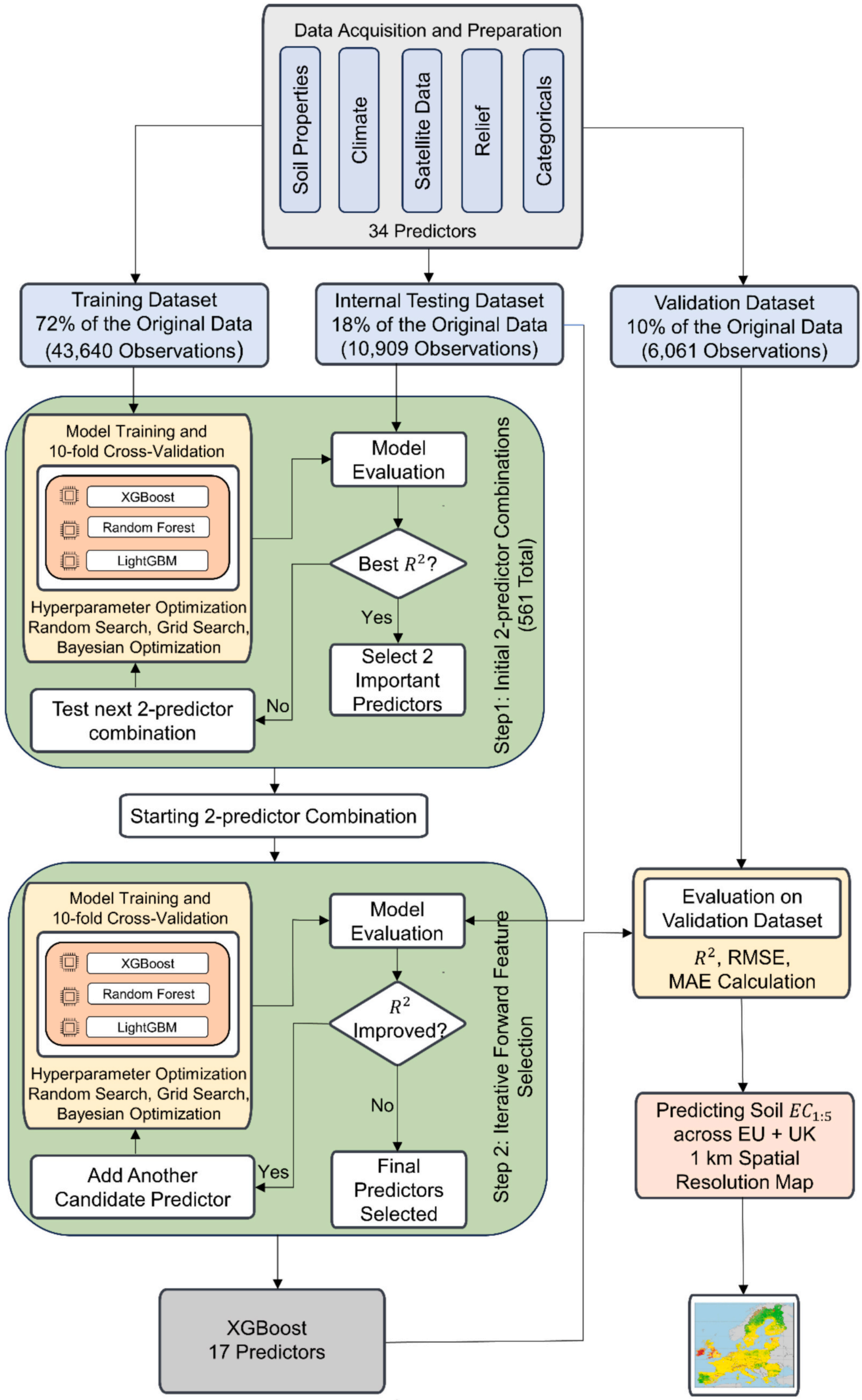


Fig. 2. Machine learning modeling workflow diagram. The diagram illustrates the complete loop of model training, testing, and evaluation that starts with data acquisition and ends with selection of the predictive model.

predictive significance (Meyer et al., 2018). Forward feature selection was performed in two main steps. First, all possible two-predictor combinations were evaluated. With 34 initial predictors, this resulted in 561 two-variable combinations. For each combination, the model was trained using the training subset and evaluated using the internal testing subset. The two-predictor combination that produced the highest internal testing coefficient of determination (R^2) was selected as the starting predictor set. Subsequently, each remaining predictor was individually added to the selected predictor set, one at a time, and the resulting models were evaluated to determine which additional predictor produced the greatest improvement in internal testing R^2. The predictor yielding the highest improvement was retained in the model. This iterative process was repeated until adding another predictor no longer improved model performance. The independent validation dataset was then used for the validation of the final selected model. To mitigate the risk of overfitting, 10-fold cross-validation was used throughout the training process. Ultimately, the trained model was applied to predict soil EC across EU member states and the UK at a spatial resolution of 1 km.

In the workflow explained above, three algorithms were tested to develop the final predictive model: Random Forest (RF), XGBoost, and LightGBM. RF is a supervised ML algorithm based on ensemble learning. It is commonly used for regression and classification tasks, combining bagging and decision tree concepts to reduce overfitting and improve predictive accuracy. During training, RF constructs multiple decision trees, and the final prediction is obtained by averaging the outputs of these trees (Breiman, 2001). XGBoost is built on the gradient-boosted decision tree framework. It employs gradient descent optimization to create a strong predictive model from weak learners, typically decision trees. These trees are built sequentially, with each new tree trained on the residual errors of the previous one, iteratively minimizing error until a predefined stopping criterion is met (Chen and Guestrin, 2016). LightGBM is a boosting algorithm based on decision trees, which is optimized for memory efficiency and speed. Unlike traditional ensemble tree methods, LightGBM employs a leaf-wise tree growth strategy, which reduces computational complexity and training time (Hamou et al., 2024). In this research, hyperparameter optimization was conducted using grid search, random search, and Bayesian optimization techniques. The final trained model's performance was assessed using the calculation of R^2, root mean squared error (RMSE), and mean absolute error (MAE). To identify areas with elevated $EC_{1:5}$, the value of 0.6dS/m was used as an empirical threshold to represent the upper tail of the observed LUCAS $EC_{1:5}$ distribution. As discussed above, this value corresponds to the upper 6% of the observed LUCAS topsoil $EC_{1:5}$ values. The threshold was not used as an indicator of salinity or treated as equivalent to any specific ECe value. Instead, areas above this value were interpreted as having elevated $EC_{1:5}$ values and as regions that may require further investigation.

3. Results and discussion

3.1. Model performance

To characterize variability in the training data, we computed descriptive statistics for EC and the numerical environmental covariates. The summaries of the descriptive statistics are provided in Table S2. To assess the predictive capabilities of the models, the evaluation metrics of the testing phase were compared across all models. The results indicate that the model developed using the XGBoost algorithm performed better than RF and LightGBM, achieving a testing R^2 of 0.6, RMSE of 1.38, and MAE of 0.39, thus being selected for further analysis and mapping. Scatter plots showing the results from each stage for the final XGBoost model are shown in Fig. 3a, b, and c. It should be noted that the model doesn't predict EC values for the observation points with missing values in the list of predictors. As a result, fewer points appear in the scatter plots than were actually used during the training, testing, and validation phases. However, because many WoSIS observations are located outside Europe, the combined-data performance metrics should be interpreted cautiously when assessing the prediction reliability within EU and UK.

The feature importance function integrated within XGBoost was employed to assess the importance of various model predictors, with the results presented in Fig. 3d. The findings indicated that soil texture and remotely sensed variables emerged as key factors on prediction of soil EC. Soil texture influences soil EC by controlling soil water retention and flow pathways, which in turn affect salt transport and accumulation. Sand content ranked highest, consistent with the role of coarse texture material causing infiltration and drainage, and thus increased salt leaching. On the other hand, clay, which is among the five most important variables, typically holds more moisture and supports capillary flow under evaporation resulting in salt accumulation (Corwin and Scudiero, 2019). NPP can be an informative predictor of soil EC because salinity reduces plant productivity. Elevated salt concentrations make it difficult for plants to extract water from the soil resulting in reduced photosynthetic rates. Consequently, plant growth and biomass accumulation decline, which is reflected in lower NPP. Furthermore, the remotely sensed variables obtained from satellite imagery were among the most important predictors within the model.

3.2. ML-driven prediction of $EC_{1:5}$ across the EU

The spatial distribution of soil $EC_{1:5}$ throughout the EU in 2018 is illustrated in Fig. 4. This dataset, which is gridded with a spatial resolution of 1 km, was generated using the XGBoost model developed in this project. As depicted in Fig. 4, certain areas, particularly in Spain, Italy, and Ireland, exhibit higher $EC_{1:5}$ levels. The Bardenas basin in Spain, an agricultural zone, is one such hotspot (Lorenzo-González et al., 2024). In Italy, coastal regions, especially Ancona, Catanzaro, and Naples, are noted for elevated soil salinity levels, a finding supported by Dazzi and Lo Papa (2013). In Ireland, the high $EC_{1:5}$ values were unexpected and are further considered in the discussion. Our model was able to capture the high $EC_{1:5}$ levels in all of these regions.

The soil salinity was analysed considering land cover classes across the EU (Venter and Sydenham, 2021). Our analysis suggests that the wetlands with an area of around 67,000 km^2 are mainly affected by $EC_{1:5}$, followed by bare land and grassland, which cover an area of approximately 40,000 km^2 and 38,000 km^2, respectively. Additionally, based on the USDA soil texture classification for the EU (Ballabio et al., 2016), high values of $EC_{1:5}$ were observed in loamy soil with an area of approximately 132,000 km^2. Sandy loam, clay loam, and silt loam also experience high levels of $EC_{1:5}$ with an area of 30,000 km^2, 24,000 km^2, and 23,000 km^2, respectively. To evaluate the confidence in the model's predictions, we implemented a bootstrap-based uncertainty analysis using the original training dataset. We generate 100 bootstrap resampled training datasets by sampling observation with replacement, while each resample contains the same number of observations as the original training set. For each bootstrap sample, a new model was trained using the previously optimized hyperparameters. These models were then used to predict $EC_{1:5}$ values across the study area at a spatial resolution of 1 km, resulting in 100 predictions per grid cell. Predictive uncertainty was quantified as the standard deviation (SD) of 100 predictions per cell (Fig. 5a). The bootstrap SD map shows higher uncertainty in northern Scandinavia, along the Alp mountain chain, and in parts of Spain in the Ebro basin, while Ireland shows smaller and more localized uncertainty hotspots. Uncertainty patterns were interpreted considering the number of observation points used during the model training per 25 km grid cell and the area of applicability map. In Sweden, high SD occurs where observation density is very low and the AOI indicates that conditions are less covered by the training data. In the Alps, uncertainty increases in high relief areas where more area falls outside the AOI. The high uncertainty in Spain and Ireland are associated with sparse distribution of the observation points. Additionally, the variance, as well as the 5th and

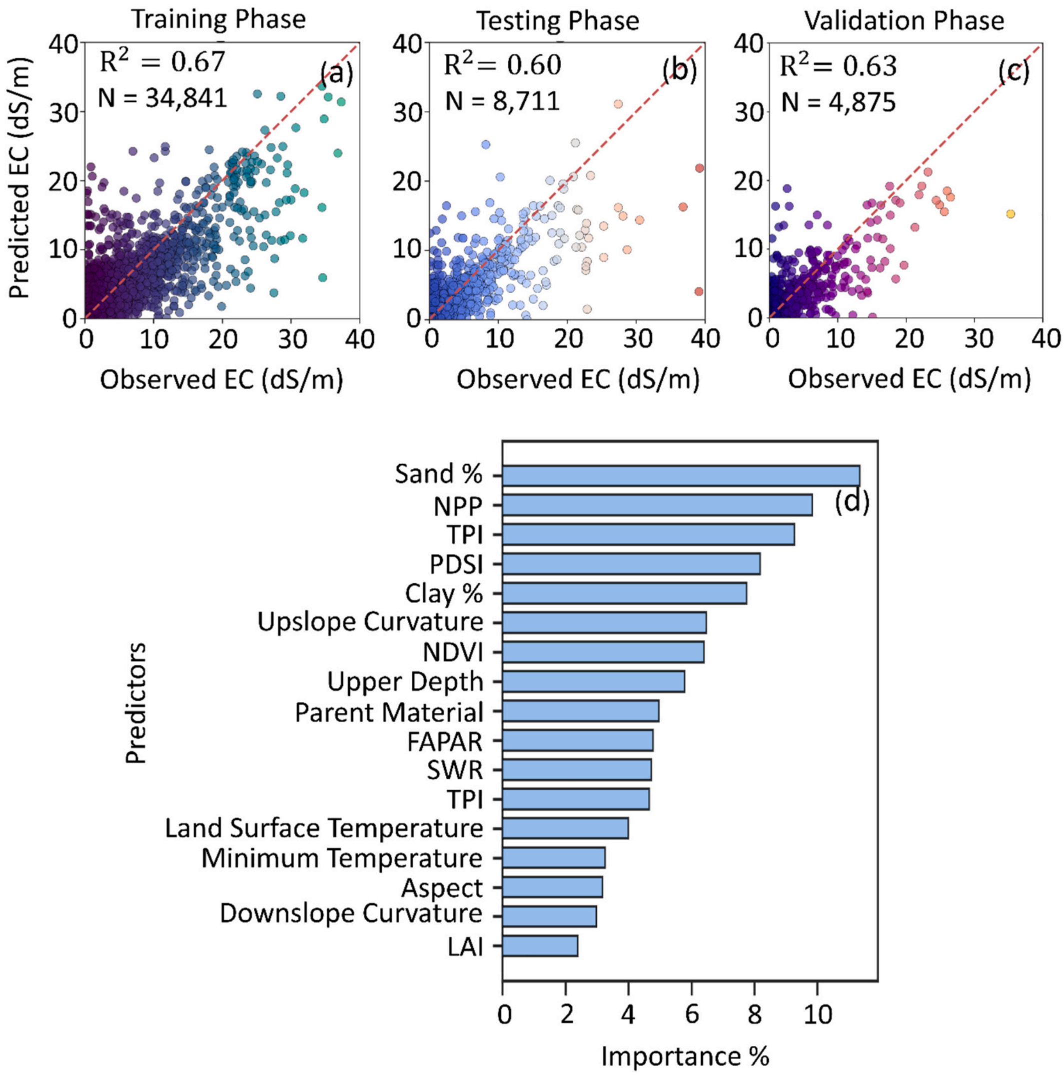


Fig. 3. Scatter plot for the training, testing, and validation phase plotted for the model trained using the XGBoost algorithm (a, b, and c). The electrical conductivity (EC) on these plots includes the $EC_{1:5}$ and ECe measurements from Land Use/Cover Area Frame Survey (LUCAS) and World Soil Information Services (WoSIS) datasets, respectively, used for training, testing, and validation of the final model. Importance of predictors in the final model trained using the XGBoost algorithm (d). NPP: Net primary production, TPI: Topographic positioning index, PDSI: Palmer drought severity index, NDVI: Normalized difference vegetation index, FAPAR: Fraction of Absorbed Photosynthetically Active Radiation, SWR: Soil to water ratio, assigned as the numeric values of 1 and 2 for ECe and $EC_{1:5}$, respectively, TRI: Topographic roughness index, LAI: Leaf Area Index.

95th percentiles of the predictions, are presented in Figs. S3, S4, and S5, respectively.

Because many WoSIS observations are located outside Europe, we also assessed how well the model outputs reflects European $EC_{1:5}$ patterns using the LUCAS 2018 observations. We compared the mean predicted $EC_{1:5}$ with the mean measured $EC_{1:5}$ for each country represented in the LUCAS dataset (Fig. 5b). Most countries lie close to the 1:1 reference line, suggesting that the model captures broad national-scale $EC_{1:5}$ pattern across the EU and UK. This comparison provides a useful check, but it should be viewed as supporting evidence for regional-scale validation, not as a replacement for spatially independent point level validation.

3.3. LUCAS observed EC1:5 - soil property relationships

Fig. 6 presents the the relationship between EC and several soil properties using the data available in the LUCAS database. In Fig. 6(a), $EC_{1:5}$ remains relatively low across most of the pH spectrum, but a cluster of higher values appears around pH ~7.7, consistent with soils containing calcium carbonate, usually in dry areas or with calcareous parent material. At lower pH, such a prominent peak is not observed, but there may potentially be a cluster of points around pH 5 with values up to about 2 dS/m. The data largely reflect the water balance at the global scale, Lebron et al. (2025) showed that European soils had higher pH values in arid regions. This lack of water, to wash out soluble salts, can lead to their build-up in the soil profile, especially if there is a source of salts from parent material, dust containing salts, or irrigation water. It's well established that alkaline soils (pH $>$ 7.5) are more likely to have higher salt concentrations, especially in semi-arid or Mediterranean regions. This aligns with the high pH areas (shown in red) on the pH map. Conversely, acidic soils, on the other hand, are associated with wet conditions and are often leached (Xu et al., 2024), which likely explains the relatively flat $EC_{1:5}$ trend observed at lower pH. The observation of some EC values up to 2dS/m around pH 5 is curious and may be related to weathering of minerals due to acidity or a redox-controlled process. The fact that the data don't rise above 2dS/m indicates perhaps that this process is potentially buffered in the system and hence does not pose a potential risk to plant stress.

Data for soil organic carbon (Fig. 6b) and total nitrogen (Fig. 6c) are consistent with this interpretation. Soil organic matter becomes

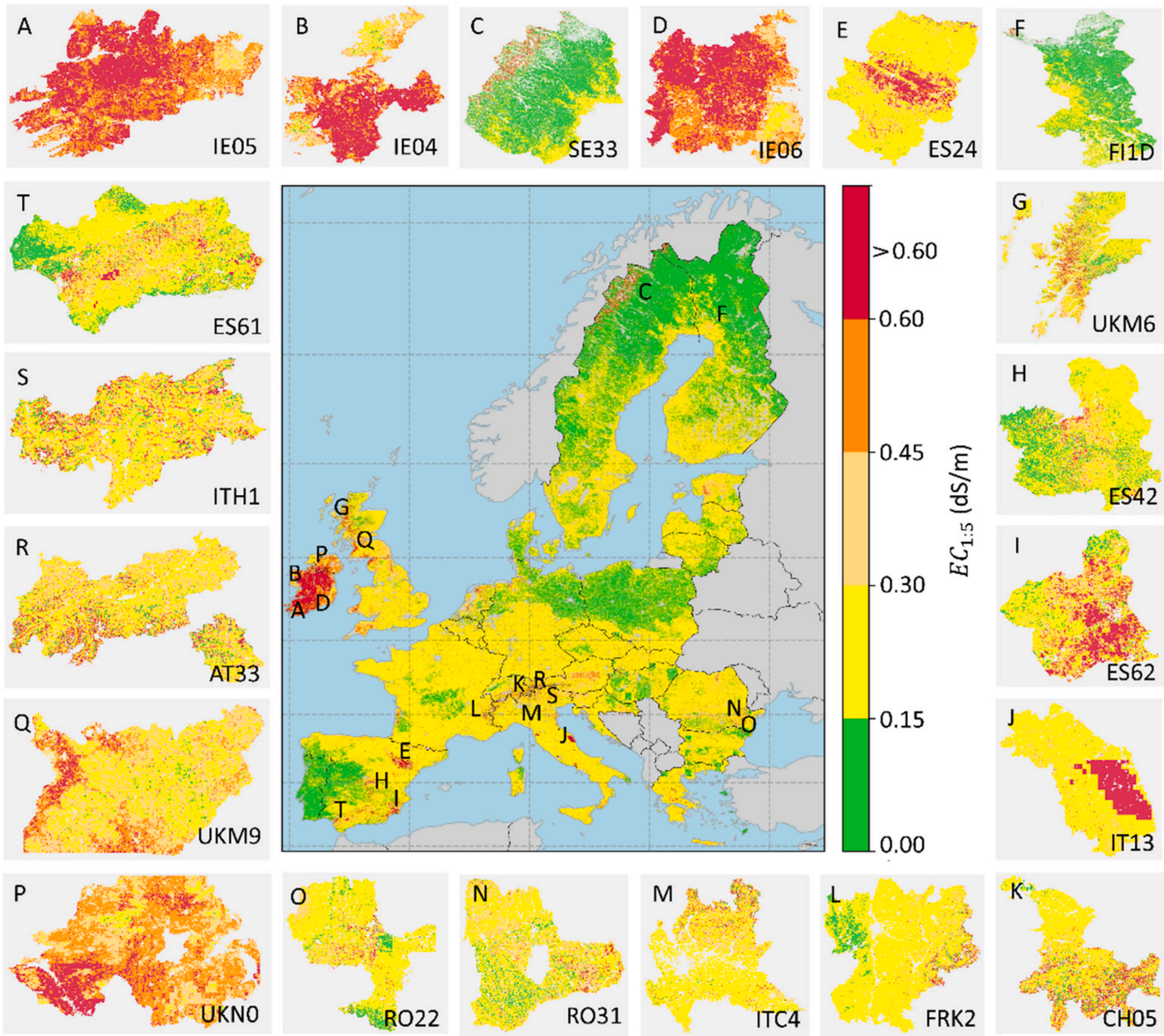


Fig. 4. The predicted gridded dataset of soil electrical conductivity ($EC_{1:5}$) across the EU members and the UK for the year 2018. Also presented are the top 18 NUTS2 regions (Eurostat, 2018) in the EU and the UK with the highest levels of $EC_{1:5}$, based on the area affected by $EC_{1:5}$ values greater than 0.6 dS/m. The NUTS2 IDs are indicated in the subplots.

increasingly unstable under dry conditions, leading to the oxidation of predominantly particulate organic matter, which is the more vulnerable component of the SOC. As the SOC undergoes transformation, the C/N ratio will also change. In dry soils with a pH above 7, relatively little SOM accumulates (Reynolds et al., 2013). What does tend to happen is that in the soils where evaporation is higher than rainfall, mineral carbonates, such as calcium carbonates, build up. Such transitions are widely observed in pedology, for example, the transition from pedocals in the dry western USA (calcium carbonate dominated) to the wetter pedalfers in the western USA, where sesquioxides are more dominant. In the LUCAS data, the presented OC and TN follow the same trend. This is to be expected as they follow the stoichiometry of the organic matter. Higher EC values occur at low OC and low TN, which is consistent with our interpretation that aside from saltmarsh soils, salinity tends to occur in dry soils that tend to be low in organic matter. It is also worth noting that there may be a slight increase in EC, up to 2 dS/m for some soils with OC levels about 400 g/kg. These may be consistent with the pH measurements around 5, showing the same pattern, indicating that in some organo-mineral soils, there may be a process releasing ions.

Although not observed in the LUCAS data, most likely due to the limited extent, soils that are periodically inundated with seawater or occur in salt marshes could have higher organic matter and appear saline. The low OC and TN soils are spatially consistent with drier climates, occurring mostly in Mediterranean soils, especially in Spain and Italy, for example. It must be noted that in all these cases, factors like parent material, clay content, climate, land use, and management practices also significantly influence the relationship between salinity and other soil characteristics. A deeper understanding of these interactions is crucial for accurately describing the processes driving soil salinization and for developing sustainable land management approaches. This could serve as an important direction for future research.

It was clear from both the map (Fig. 4) and the raw data (Fig. 1) that Ireland had some unexpectedly high EC values. These were extracted and plotted in Fig. 7 with EC in relation to SOC. 2015 was an unusually wet year, while 2018 was a normal year. In 2018, the organic soils had low values of EC, but in the wet year, the EC of the organo-mineral soils increased substantially; this was less in the mineral and peat soils. This behaviour is discussed in more detail in the discussion.

4. $EC_{1:5}$ map in the context of soil salinity

In the World Reference Base for Soil Resources (WRB), Solonchaks are defined by the presence of a salic horizon, diagnosed when the electrical conductivity of the saturation extract (ECe) meets WRB

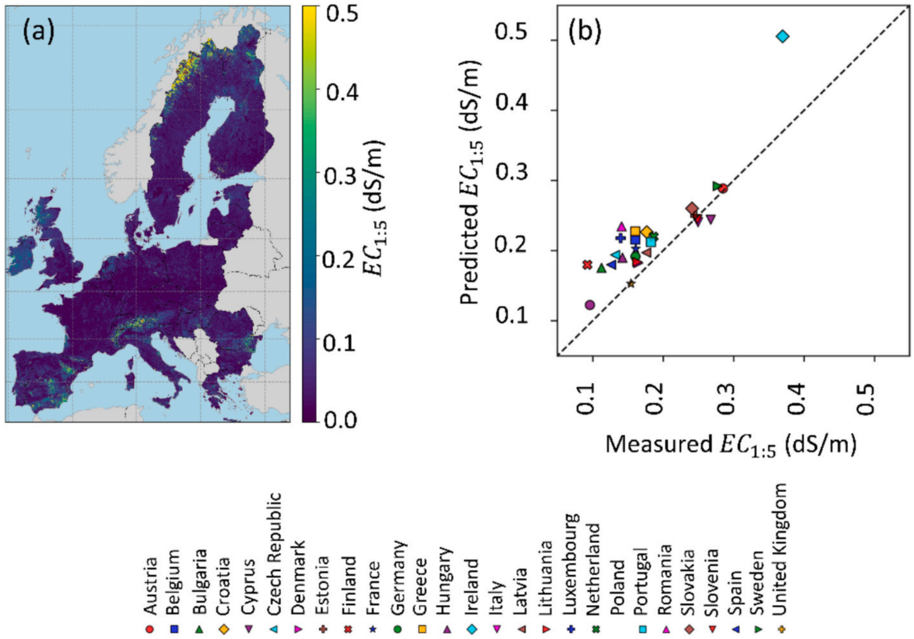


Fig. 5. Uncertainty map showing the standard deviation of the predictions for the year 2018, derived from 100 bootstrap resampled datasets, at a spatial resolution of 1 km (a). Scatter plot illustrating the country-level mean predicted vs mean measured $EC_{1:5}$ for LUCAS observation points for the year 2018 (b).

thresholds—typically ECe ≥ 15 dS/m, or ≥8 dS/m when the saturation extract pH ≥ 8.5, together with the required thickness and ECe–thickness product. When salinity is extremely high, additional qualifiers such as hypersalic (e.g., ECe ≥ 30 dS/m) may be applied. Solonetz soils, in contrast, are defined by a natric horizon, characterized by clay illuviation and ESP ≥ 15 (IUSS, 2022).

Beyond these fully developed diagnostic horizons, many soils across Europe exhibit incipient or intermittently elevated salinity that falls below the formal thresholds for a salic horizon. These 'protosalic' soils—a term used in the literature to describe soils with early-stage salt accumulation (often around ECe ≈ 4 dS/m, the FAO threshold for saline soils)—may occur within a wide range of soil groups, not only Solonchaks. However, because their salinity is seasonal, weakly expressed, or spatially patchy, protosalic soils typically fall outside WRB diagnostic criteria and therefore do not appear on continental-scale maps. This results in a systematic underrepresentation of emerging salinity, especially in areas where salts accumulate only during the dry season and are leached below diagnostic depths in wetter months. The mapping framework presented here helps address this gap by using point-based electrical conductivity measurements to identify locations where protosalic conditions are developing, providing early indicators of salinity risks that are not captured in existing WRB-based soil maps. The $EC_{1:5}$ maps (Fig. 4) highlight areas where protosalic soils may occur, but interpreting $EC_{1:5}$ as an indicator of salinity is complicated by the need to convert it to ECe, the standard metric for diagnosing saline horizons (Kargas et al., 2020). Although $EC_{1:5}$ is widely used because it is rapid and inexpensive, no universal conversion factor exists: published relationships vary widely—from roughly 5 to 12 × $EC_{1:5}$—owing to differences in extraction methods, equilibration times (Kargas et al., 2020), and the strong influence of soil texture and mineralogy, particularly clay and smectite content, on paste formation. As a result, $EC_{1:5}$-to-ECe conversions are site-specific approximations, and caution is needed when extrapolating them across diverse European soils.

However, the more significant issue overlooked to date, as far as we can ascertain, is how to convert $EC_{1:5}$ to ECe for non-saline soils. The issue is especially relevant for calcareous soils that may occur in Europe overlaying chalk or limestone parent materials. The dissolution of $CaCO_3$ is small but not inconsequential, it largely depends on the partial pressure of carbon dioxide, log (P_{CO2}). Calculations using PHREEQC (Parkhurst and Appelo, 2013) for an aqueous saturated solution of calcite indicate that the EC of the equilibrium solution changes according to 0.038 dS/m at log P_{CO2} (10^{-5} atm), 0.141 (10^{-3}), and 0.681 (10^{-1}). Equilibrium with the atmosphere is about $10^{-3.4,}$ and so calcareous soils in equilibrium with the atmosphere may be expected to have solution EC values of about 0.2 dS/m, increasing perhaps up to 0.5 dS/m due to CO_2 build-up from root respiration and higher P_{CO2} values in the soil solution. This complicates conversion because, as long as the extract is saturated with calcium carbonate, the concentration in the solution will be the same regardless of dilution. Multiplying the $EC_{1:5}$ by a conversion factor would then result in an erroneous overestimation ECe value. This issue is not generally observed in literature conversions because most analysis is conducted on soils with other ions, particularly sodium or magnesium, in addition to calcium, that are not constrained in the same way. In addition, soils with high amounts of gypsum are also saturated and have an EC of about 2.4 dS/m, but this is not impacted by P_{CO2}. This is why soil salinity is defined using ECe above the influence of carbonate and gypsum.

Relatively high $EC_{1:5}$ values were observed in Ireland (Fig. 4). The one or two highest points are along the coast and likely salt marsh soils, but the rest are mostly inland grasslands. Looking more closely at the data (Fig. 7), they were at their highest in the wetter than average year in 2015 (https://www.climateireland.ie/climate-change/evidence-of-climate-change/), with a median value of 0.79 dS/m, which decreased to 0.29 dS/m in the normal year in 2018. Values hardly rose above 1dS/m in 2018 (3%), but in 2015, 34% were above this value. Some of the increase occurred in mineral soils with a pH above 6, but the majority was in organic soils with a pH around 5.

Ireland is not known for soil salinity in the form of soils with high levels of sodium or magnesium, as is often found in conjunction with calcareous soils in the Mediterranean (Visconti et al., 2009). However, Karst soils with calcium carbonate do occur across large areas of Ireland, which has the highest density of karst wetland habitat globally (Kimberley and Waldren, 2006); moreover, groundwater monitoring also reveals high pH groundwater with relatively high Ca^{2+} concentrations (Hájek et al., 2020). So, the occurrence of EC values of 0.6–0.7 dS/m observed in the pH = 7 soils in 2018 is consistent with high levels of $CaCO_3$ measured in the LUCAS data, and appears to fit with the P_{CO2} carbonate saturation mechanism described previously. In 2015, the

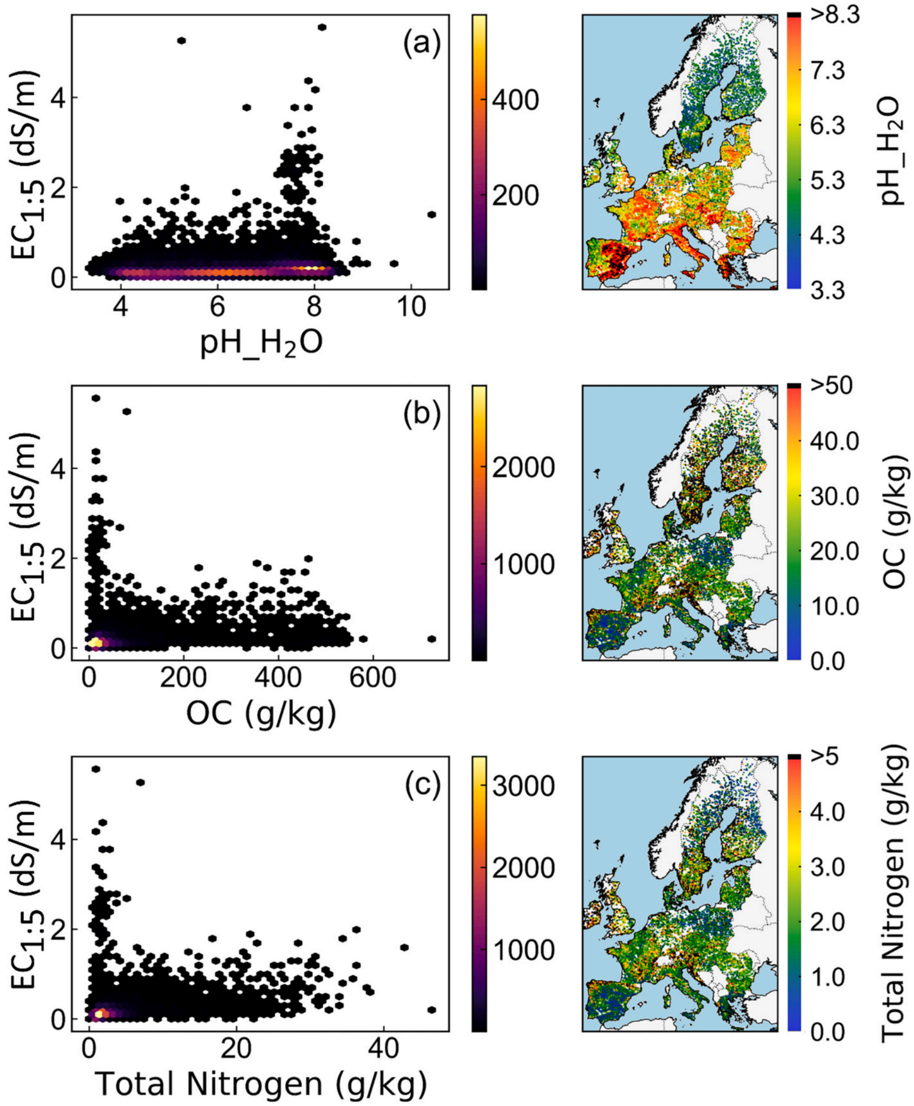


Fig. 6. The density scatter plots illustrate the correlations between $EC_{1:5}$ and pH measured in water (a), organic carbon (b), and total nitrogen (c), based on LUCAS 2018 data points combined. The accompanying maps show the spatial distribution of pH (measured in water), organic carbon, and total nitrogen across Europe, also using LUCAS data.

wetter year, EC values increased a little around pH 7, but the large increases were observed in the more acid soils (pH < 6), which are also high in organic matter. Clustering was also observed in the data, with soils with a SOC content of <400 g/kg having EC values lower than 2 dS/m, but organic soils (> 400 g/kg) showing an almost exponential increase in EC with increasing SOC, reaching EC values of between 2 and 3 dS/m. The precise mechanism for this is unknown, but we can speculate that certain regions may exhibit the establishment of halophytic vegetation while in others the increased water logging and likely organic acids caused ions to be released, possibly in connection with upwelling of ground water, weathering or from the exchange complex, most likely base cations from groundwater, or iron and aluminium for example from redox processes. This may also be exacerbated by ions being forced off the exchange complex, or even perhaps sodium from rainwater that remains to be leached, or ions from land spreading, especially given that cow urine can range from 1 to 33 dS/m (Fazil Marickar, 2010). This must also be an EC saturation mechanism, as multiplying these values by 6 or more for ECe would result in implausible salinity levels, and the grass remains green in Ireland. The implications of these results for ECe remain undetermined without understanding how such a mechanism impacts ECe and whether, to some extent, it's saturated or not. However, we speculate that this mechanism can also be observed in other areas such as C, F, G, Q, L, K, R, and S in Fig. 4. Areas K, R, S, and C are not associated with calcareous materials or ground water (Hájek et al., 2020), hence an alternative explanation to high calcium levels is required.

An ECe value of 4 dS/m is commonly considered the threshold distinguishing saline from non-saline soils (FAO, 2021) and is consistent with the development of a protosalic horizon. In this study, however, the mapped target variable is $EC_{1:5}$, and conversion from $EC_{1:5}$ to ECe is not universal and their accuracy depends heavily on local conditions and boundary parameters, as previously discussed. Therefore, rather than inferring ECe from $EC_{1:5}$ using a single conversion factor, we used $EC_{1:5}$ value of 0.6 dS/m as an empirical upper-tail screening threshold derived from the LUCAS $EC_{1:5}$ distribution. Using this threshold, our analysis suggests that approximately 217,000 km^2 of land within the study area showed elevated predicted $EC_{1:5}$ values. This area should therefore be interpreted as a potential priority zone for further investigation, not as the mapped extent of the confirmed saline soils, salic horizons, or protosalic horizons. For the reasons discussed earlier, regarding carbonate soils and saturation, the real value is likely much lower than this. However, it proposes an upper bound, which is ten times higher than the

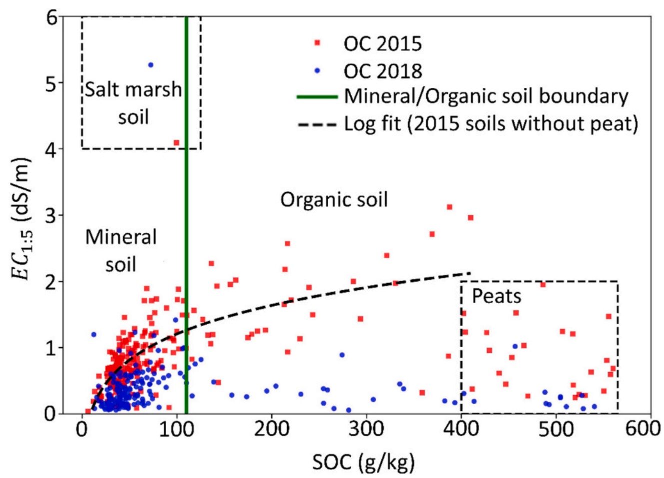


Fig. 7. Raw $EC_{1:5}$ data from the Land Use/Cover Area Frame Survey (LUCAS) 2015 and 2018 data sets. 2018 was a normal year for rainfall, but 2015 was a wet year with rainfall nearly 300 mm above normal. The box with the red dashed line indicates highly organic peaty soils. The soil with an $EC_{1:5}$ of >4dS/m was consistent with saltmarsh. The green line indicates the separation between mineral and organic soils. The red dotted line is a trend line for the mineral and organo-mineral soils in 2015.

mapped area of known Solonetz and Solonchak soils for Europe, which will have salic or hypersalic soil horizons. Even if the extent of soils with protosalic horizons was half the extent predicted, it's still a very significant area. Moreover, the implications for global salinity assessment are profound, indicating that if assessment is based solely on maps of Solonchaks and Solonetz soils, large areas of land with emerging salinity may be going undetected.

5. Limitations of the model and recommendations for future research

A key limitation of this study is that the target variable combines EC measurements derived from two different laboratory extraction protocols. Although we account for method-related differences using the SWR indicator, this is a simplified representation of the complex, soil-specific and potentially nonlinear relationship between EC1:5 and ECe. This simplification can affect model interpretation and spatial inference. Future works could address this limitation through acquiring and modeling paired EC1:5 – ECe measurements.

One limitation is the geographic distribution of the WoSIS observations. Although WoSIS was useful for increasing the range of EC values during model training, many observations are located outside Europe, particularly in North America. This may introduce some transferability uncertainty, because relationships between EC and environmental covariates learned from non-European regions may not fully reflect European soil-forming conditions. Therefore, the combined-data performance metrics should be interpreted mainly as model statistics rather than as an evidence of European spatial prediction accuracy.

Model evaluation was conducted using a train-validation split stratified by the distribution of the target variable. While this approach supports evaluation across the EC range, it does not guarantee spatial separation between training and validation samples. Therefore, extrapolation to unsampled locations should be done cautiously given the lack of spatial cross validation in model validation. Spatially blocked or region holdout validation is recommended in future works.

Our model did not account for factors such as changes in land use, agricultural practices, groundwater levels, and water extraction rates, all of which have impacts on secondary soil salinization. The inclusion of such factors is important to quantify the effects of human activities on soil salinization.

Consequently, future research could consider incorporating diverse sets of environmental variables and employing different modeling approaches (e.g., ML algorithms) to directly predict ECe, using EC measurements at various soil-water ratios as model inputs through integration of LUCAS dataset with other data sources. Also, the small number of salinity observation points in areas with high salinity across the EU makes it difficult to validate the model and affects its reliability. Adding more soil sampling sites, especially in areas with high salinity risk, would improve data coverage and help make predictions more accurate. Another important limitation is the lack of high-resolution geological maps, which makes it difficult to account for variations in parent material as an important factor in soil salinization. This missing information could introduce uncertainty into our predictions. Additionally, the sparse distribution of LUCAS sampling points in some areas limits the model's ability to capture local salinity issues. For example, the model tends to underestimate $EC_{1:5}$ in Hungary and overestimate it in northern Sweden, likely due to insufficient spatial coverage. Moreover, a more detailed investigation of the relationship between soil salinity and other soil properties, such as pH, SOC, and nitrogen, could offer valuable insights into the processes of soil salinization, its consequences, and could help guide the development of effective remediation strategies. A greater mechanistic understanding of what appears to be salinity in organo-mineral soils under wet conditions is required. Warmer, wetter winters due to climate change in areas like Ireland could see more of the patterns we observed, and understanding and attribution are required to determine the potential risk.

6. Conclusions

In this study, we employed three ML algorithms, Random Forest, XGBoost, and LightGBM, to predict soil $EC_{1:5}$ across the EU and UK, using EC values from LUCAS and WoSIS databases as the target variable and a set of environmental covariates as predictors. By applying a forward feature selection algorithm, we identified the most significant predictors, including sand content (in fine earth fraction), net primary production, topographic positioning index, and clay content (among other parameters reported in Fig. 3). Using the best-performing ML algorithm (XGBoost model), spatial patterns of $EC_{1:5}$ values were mapped across the EU for the year 2018. Our results suggest that approximately 21.7 Mha of EU land exceeds an EC of 0.6 dS/m at a 1:5 soil to water ratio ($EC_{1:5}$). We have also concluded that translating $EC_{1:5}$ to ECe cannot be achieved reliably using a single universal conversion factor because $EC_{1:5}$ – ECe relationships are strongly soil dependent and should be treated as site specific, particularly in calcareous soils where standard conversions may lead to overestimation of ECe. This issue is rarely addressed in the literature, as most studies focus on soils containing more soluble ions like sodium or magnesium than calcium.

The proposed framework and findings reported in this study have implications for sustainable soil management and policy development within the EU. Furthermore, the methodologies employed in this study can be adapted and applied to other regions facing similar challenges, contributing to the global efforts for soil salinity management and sustainable agriculture.

CRediT authorship contribution statement

Mohammad Aziz Zarif: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Amirhossein Hassani:** Writing – original draft, Visualization, Validation, Methodology, Investigation. **Mehdi H. Afshar:** Writing – original draft, Visualization, Investigation. **Panos Panagos:** Writing – original draft, Visualization, Investigation. **Inma Lebron:** Writing – original draft, Visualization, Investigation. **David A. Robinson:** Writing – original draft, Visualization, Validation, Investigation, Formal analysis. **Dani Or:** Writing – original draft, Visualization, Investigation. **Nima Shokri:** Writing – original draft, Visualization, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgements

This research is part of the project AI4SoilHealth (Accelerating collection and use of soil health information using AI technology to support the Soil Deal for Europe and EU Soil Observatory), which was funded by the European Union Horizon Europe research and innovation program (Grant No. 101086179). DAR and IL acknowledge funding from the UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10053484) as part of AI4SoilHealth. N.S. acknowledges the Joint Research Centre of the European Commission for the support provided by the contract number CT-EX2024D1075827-101.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2026.182021.

Data availability

The predicted TIFF files of EC1:5 across the EU generated in this study for the year 2015 and 2018 are freely available at the European Soil Data Centre (JRC) (Panagos et al., 2022). All data used to train the model were obtained from publicly available sources, which have been cited appropriately throughout the text. All computer codes required for the regeneration of the main results presented in this paper can be found in the supplementary information.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5 (1), 1–12.
Afshar, M.H., Hassani, A., Aminzadeh, M., Borrelli, P., Panagos, P., Robinson, D.A., Or, D., Shokri, N., 2025. Spatial and temporal assessment of soil degradation risk in Europe. Sci. Rep. 15, 44636. https://doi.org/10.1038/s41598-025-33318-7.
Alqasemi, A.S., Ibrahim, M., Fadhil Al-Quraishi, A.M., Saibi, H., Al-Fugara, A.K., Kaplan, G., 2021. Detection and modeling of soil salinity variations in arid lands using remote sensing data. Open Geosci. 13 (1), 443–453.
Amatulli, G., Domisch, S., Tuanmu, M.N., Parmentier, B., Ranipeta, A., Malczyk, J., Jetz, W., 2018. A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. Sci. Data 5 (1), 1–15.
Ballabio, C., Panagos, P., Monatanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma 261, 110–123.
Batjes, N.H., 2025. World Soil Information Service (WoSIS) - Procedures for Coding and Standardizing Soil Analytical Method Descriptions. ISRIC - World Soil Information, Wageningen. https://doi.org/10.17027/isric-wdcsoils-1nh7-zr51 (51 pp.).
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Bresler, E., McNeal, B.L., Carter, D.L., 2012. Saline and Sodic Soils: Principles-Dynamics-Modeling. Springer Science & Business Media.
Celleri, C., Pratolongo, P., Arena, M., 2022. Spatial and temporal patterns of soil salinization in shallow groundwater environments of the Bahía Blanca estuary: influence of topography and land use. Land Degrad. Dev. 33 (3), 470–483.
Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. https://doi.org/10.1145/2939672.2939785.
Corwin, D.L., Scudiero, E., 2019. Review of soil salinity assessment for agriculture across multiple scales using proximal and/or remote sensors. Adv. Agron. 158, 1–130.
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.
Dazzi, C., Lo Papa, G., 2013. Soil threats. In: Costantini, E., Dazzi, C. (Eds.), The Soils of Italy, World Soils Book Series. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-5642-7_8.
Didan, K., 2021. MODIS/Terra Vegetation Indices 16-Day L3 Global 500m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD13A1.061. Accessed 2025-03-05 from.
ESDB v2.0, 2004. The European Soil Database Distribution Version 2.0. European Commission and the European Soil Bureau Network, CD-ROM, EUR 19945 EN.
Eurostat, 2018. Nomenclature of Territorial Units for Statistics — NUTS 2016/EU-28. Publications Office of the European Union.
FAO, 2021. Global map of salt-affected soils gsasmap v1.0. Retrieved from. https://www.fao.org/documents/card/en/c/cb7247en.
Fazil Marickar, Y.M., 2010. Electrical conductivity and total dissolved solids in urine. Urol. Res. 38 (4), 233–235.
Hájek, M., Jiménez-Alfaro, B., Hájek, O., Brancaleoni, L., Cantonati, M., Carbognani, M., Dedić, A., Dítě, D., Gerdol, R., Hájková, P., Horsáková, V., Jansen, F., Kamberović, J., Kapfer, J., Kolari, T.H.M., Lamentowicz, M., Lazarević, P., Mašić, E., Moeslund, J.E., Pérez-Haase, A., Peterka, T., Petraglia, A., Pladevall-Izard, E., Plesková, Z., Segadelli, S., Semeniuk, Y., Singh, P., Šímová, A., Šmerdová, E., Tahvanainen, T., Tomaselli, M., Vystavna, Y., Biţă-Nicolae, C., Horsák, M., 2020. A European map of groundwater pH and calcium. Earth Syst. Sci. Data Discuss. 2020, 1–41.
Hamou, K.A.B., Jarir, Z., Elfirdoussi, S., 2024. Application of LightGBM algorithm in production scheduling optimization on non-identical parallel machines. Eng. Technol. Appl. Sci. Res. 14 (6), 17973–17978.
Han, J., Wu, N., Wu, Y., Zhou, S., Bi, X., 2024. Spatial effects of shrub encroachment on wetland soil pH and salinity in the Yellow River Delta, China. J. Coast. Conserv. 28 (4), 60.
Hartmann, J., Moosdorf, N., 2012. The new global lithological map database GLiM: a representation of rock properties at the Earth surface. Geochem. Geophys. Geosyst. 13 (12).
Hassani, A., Azapagic, A., D'Odorico, P., Keshmiri, A., Shokri, N., 2020a. Desiccation crisis of saline lakes: a new decision-support framework for building resilience to climate change. Sci. Total Environ. 703, 134718.
Hassani, A., Azapagic, A., Shokri, N., 2020b. Predicting long-term dynamics of soil salinity and sodicity on a global scale. Proc. Natl. Acad. Sci. 117 (52), 33017–33027.
Hassani, A., Azapagic, A., Shokri, N., 2021. Global predictions of primary soil salinization under changing climate in the 21st century. Nat. Commun. 12 (1), 6663.
Hassani, A., Smith, P., Shokri, N., 2024. Negative correlation between soil salinity and soil organic carbon variability. Proc. Natl. Acad. Sci. U. S. A. 121 (18), e2317332121.
Hengl, T., 2018. Global DEM Derivatives at 250 m, 1 km and 2 km Based on the MERIT DEM.
Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLoS One 12 (2), e0169748.
International Organization for Standardization, 1994. ISO 11265:1994 – Soil Quality — Determination of the Specific Electrical Conductivity. ISO, Geneva, Switzerland.
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., Kempen, B., De Sousa, L., 2019. Global mapping of soil salinity change. Remote Sens. Environ. 231, 111260.
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 12 (4), 1010.
Kimberley, S., Waldren, S., 2006. Variations in the soil nutrient properties of temporary limestone lakes (turloughs) located within two contrasting catchment areas in the west of Ireland. ESAI Colloquium Proceedings 2005, p. 74.
Lebron, I., Feeney, C.J., Reinsch, S., Shokri, N., Afshar, M.H., Lofts, S., Griffiths, R., Fidler, D., Jones, B., Panagos, P., Sawicka, K., Keith, A.M., Seaton, F., Robinson, D. A., 2025. Patterns and thresholds for soil pH across Europe in relation to soil health and degradation. Catena 260, 109454.
Lorenzo-González, M.A., Herrero, J., Castañeda, C., 2024. A heritage dataset of soil and water salinity in Bardenas, Spain. Data Brief 54, 110469.
Lozupone, C.A., Knight, R., 2007. Global patterns in bacterial diversity. Proc. Natl. Acad. Sci. U. S. A. 104, 11436–11440.
McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.
Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ. Model. Softw. 101, 1–9.
Myneni, R., Knyazikhin, Y., Park, T., 2021. MODIS/Terra+Aqua Leaf Area Index/FPAR 4-Day L4 Global 500m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD15A3H.061. Accessed 2025-03-05 from.
Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2018. LUCAS soil, the largest expandable soil dataset for Europe: a review. Eur. J. Soil Sci. 69 (1), 140–153. https://doi.org/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. and, Montanarella, L., 2022. European soil data centre 2.0: soil data and knowledge in support of the EU policies. Eur. J. Soil Sci. 73 (6), e13315.
Panagos, P., Broothaerts, N., Ballabio, C., Orgiazzi, A., De Rosa, D., Borrelli, P., Liakos, L., Vieira, D., Van Eynde, E., Arias Navarro, C. and, Breure, T., 2024. How the EU Soil Observatory is providing solid science for healthy soils. Eur. J. Soil Sci. 75 (3), e13507.

Parkhurst, D.L. and, Appelo, C.A.J., 2013. Description of input and examples for PHREEQC version 3—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. In: US Geological Survey Techniques and Methods, vol. 6 (A43), p. 497.

Peñuelas, J., Poulter, B., Sardans, J., et al., 2013. Human-induced nitrogen–phosphorus imbalances alter natural and managed ecosystems across the globe. Nat. Commun. 4, 2934.

Qi, G., Chang, C., Yang, W., Gao, P., Zhao, G., 2021. Soil salinity inversion in coastal corn planting areas by the satellite-UAV-ground integration approach. Remote Sens. 13 (16), 3100.

Rengasamy, P., Chittleborough, D., Helyar, K., 2003. Root-zone constraints and plant-based solutions for dryland salinity. Plant Soil 257, 249–260.

Reynolds, B., Chamberlain, P.M., Poskitt, J., Woods, C., Scott, W.A., Rowe, E.C., Robinson, D.A., Frogbrook, Z.L., Keith, A.M., Henrys, P.A., Black, H.I.J., Emmett, B. A., 2013. Countryside survey: national "soil change" 1978–2007 for topsoils in Great Britain—acidity, carbon, and total nitrogen status. Vadose Zone J. 12 (2), vzj2012-0114.

Ribeiro, E.D., Batjes, N.H., van Oostrum, A.J.M., 2018. World Soil Information Service (WoSIS) - Towards the Standardization and Harmonization of World Soil Data: Procedures Manual 2018. https://doi.org/10.17027/ISRIC-WDCSOILS.20180001.

Richards, L.A. (Ed.), 1954. Diagnosis and Improvement of Saline and Alkali Soils (No. 60). US Government Printing Office.

Running, S., Zhao, M., 2021. MODIS/Aqua Net Primary Production Gap-filled Yearly L4 Global 500m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MYD17A3HGF.061. Accessed 2025-03-05 from.

Schillaci, C., Scarpa, S., Yunta, F., Lipani, A., Visconti, F., Szatmári, G., Balog, K., Koganti, T., Greve, M., Bondi, G., Kargas, G., Londra, P., Kaya, F., Lo Papa, G., Panagos, P., Montanarella, L., Jones, A., 2025. 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. Geoderma 454, 117199.

Shokri, N., Hassani, A., Sahimi, M., 2024. Multiscale soil salinization dynamics from global to pore scale: a review. Rev. Geophys. 62, e2023RG000804. https://doi.org/10.1029/2023RG000804.

Shokri, N., Robinson, D.A., Afshar, M., Alewell, C., Aminzadeh, M., Arthur, M., Broothaerts, N., Campbell, G.A., Eklund, L., Gupta, S., Harper, R., Hassani, A., Hohenegger, C., Keller, T., Kiener, M., Lebron, I., Madani, K., Marwala, T., Matthews, F., Moldrup, P., Nemes, A., Panagos, P., Prăvălie, R., Rillig, M.C., Saggau, P., Shokri-Kuehni, S.M.S., Smith, P., Thomas, A., Wollesen de Jonge, L., Or, O., 2025. Rethinking global soil degradation: drivers, impacts, and solutions. Rev. Geophys. 63, e2025RG000883. https://doi.org/10.1029/2025RG000883.

Shokri-Kuehni, S.M.S., Norouzirad, M., Webb, C., Shokri, N., 2017. Impact of type of salt and ambient conditions on saline water evaporation from porous media. Adv. Water Resour. 105, 154–161.

Szabolcs, I., 1974. Salt Affected Soils in Europe. ISBN 978-94-011-8638-4. Springer Dordrecht.

Tóth, G., Adhikari, K., Várallyay, G., Tóth, T., Bódis, K., Stolbovoy, V., 2008. Updated map of salt affected soils in the European Union. In: Threats to Soil Quality in Europe EUR, 23438, pp. 65–77.

Veerman, C., Correia, T.P., Bastioli, C., Biro, B., Bouma, J., Cienciala, E., Emmett, B., Frison, E.A., Grand, A., Hristov, L., Kriaučiūnienė, Z., Pogrzeba, M., Soussana, J.F., Vela, C.O., Wittkowski, R., 2020. Caring for Soil Is Caring for Life: Ensure 75% of Soils Are Healthy by 2030 for Healthy Food, People, Nature and Climate: Interim Report of the Mission Board for Soil Health and Food: Study.

Venter, Z.S., Sydenham, M.A., 2021. Continental-scale land cover mapping at 10 m resolution over Europe (ELC10). Remote Sens. 13 (12), 2301.

Visconti, F., de Paz, J.M., Rubio, J.L., 2009. Principal component analysis of chemical properties of soil saturation extracts from an irrigated Mediterranean area: implications for calcite equilibrium in soil solutions. Geoderma 151 (3–4), 407–416.

Wadoux, A.M.C., Minasny, B., McBratney, A.B., 2020. Machine learning for digital soil mapping: applications, challenges and suggested solutions. Earth Sci. Rev. 210, 103359.

Wan, Z., Hook, S., Hulley, G., 2021. MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD11A2.061. Accessed 2025-03-05 from. LST.

Wang, N., Chen, S., Huang, J., Frappart, F., Taghizadeh, R., Zhang, X., Wigneron, J.P., Xue, J., Xiao, Y., Peng, J., Shi, Z., 2024. Global soil salinity estimation at 10 m using multi-source remote sensing. J. Remote Sens. 4, 0130.

Xu, D., Ros, G.H., Zhu, Q., Xu, M., Wen, S., Cai, Z., Zhang, F., de Vries, W., 2024. Major drivers of soil acidification over 30 years differ in paddy and upland soils in China. Sci. Total Environ. 916, 170189.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2018. MERIT DEM: Multi-Error-Removed Improved-Terrain DEM. URL. http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM.