

Digital soil mapping of copper in Sweden: Using the prediction and uncertainty as decision support in crop micronutrient management

Karl Adler^{a,*}, Kristin Piikki^a, Mats Söderström^a, Jan Eriksson^b

^a Department of Soil and Environment, Swedish University of Agricultural Sciences, SE-53223 Skara, Sweden

^b Department of Soil and Environment, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden



ARTICLE INFO

Keywords:

Digital soil mapping
Micronutrient
Copper
Machine learning
PXRF
Cambisol
Regosol

ABSTRACT

Digital soil mapping (DSM) of topsoil copper (Cu) concentrations and prediction intervals covering 90% of agricultural land in Sweden was performed, in order to identify areas at risk of Cu deficiency. A total of 12,527 soil samples were used to calibrate the DSM model, using airborne gamma radiation data, climate data, topographical data and soil texture class data. Among the samples included, 11,093 had no laboratory-analysed Cu concentrations, so their Cu concentrations were predicted using portable X-ray fluorescence (PXRF) measurements. Cross-validation of the PXRF model resulted in Nash-Sutcliffe model efficiency coefficient (E) of 0.66 and mean absolute error (MAE) of 3.3 mg kg^{-1} . Cross-validation of the DSM model showed somewhat lower performance ($E = 0.57$, MAE = 4.1 mg kg^{-1}). Based on the lower bound of the prediction interval (5th percentile), 48% of agricultural soils in Sweden are most likely not at risk of Cu deficiency ($>7 \text{ mg kg}^{-1}$). The Cu map was also validated against concentrations in soil samples from five fields (25–47 ha in size; four samples per ha). The field means were predicted with a MAE of 1.0 mg kg^{-1} and within-field variation was reproduced with a field-wise squared Pearson correlation coefficient (r^2) of 0–0.36. The classification metric ‘recall’ showed that the map of soil Cu concentrations might not predict all possible areas at risk of being Cu deficient, as observational data indicates that about 22% of soils in the mapped area should have Cu concentrations below the risk limit. However, the metric ‘precision’ showed that when the soil map predicted a concentration at or below 7 mg kg^{-1} , it was generally correct. Increasing the limit resulted in the recall and precision increasing rapidly. The remaining 52% of agricultural soils at risk of being below the Cu concentration limit can be targeted by laboratory analysis or monitoring.

1. Introduction

Copper (Cu) is an important micronutrient for agricultural crops. Low concentrations of plant-available Cu in agricultural soils can limit plant growth and lead to malformation of crop organs (Oorts, 2013). Conversely, high concentrations can be toxic and limit root length, shoot length, plant height and leaf area (Cook et al., 1997; Michaud et al., 2008). Copper deficiency in plants can be related to low total background levels of Cu in the soil, which is often the case for sandy soils (Oorts, 2013). In Sweden, agricultural soils in which the total Cu concentration is $<7 \text{ mg kg}^{-1}$ are regarded as being at risk of crop Cu deficiency (Swedish Board of Agriculture, 2020). Knowledge of the Cu concentration in agricultural soils is therefore important for decisions on crop micronutrient management.

Previous research has indicated that the lowest soil Cu

concentrations in Europe are found in the Scandinavian countries, particularly Sweden, although some of these findings relates to various types of soil and not only agricultural soil (Albanese et al., 2015; Ballabio et al., 2018; Panagos et al., 2018). Eriksson et al. (2017) showed that 22% of Swedish agricultural soils have Cu concentrations below 7 mg kg^{-1} . Hence, agricultural soil that is at risk of being Cu deficient is not uncommon in Sweden.

Eriksson et al. (2017) created maps of soil Cu concentrations using data on soil samples collected in 2007 and earlier, but the maps produced had relatively coarse resolution of $10 \text{ km} \times 10 \text{ km}$. Maps with higher resolution exist, e.g. maps based on the LUCAS dataset encompassing Europe at a resolution of $500 \text{ m} \times 500 \text{ m}$, but the models were calibrated on topsoil samples from many land use types and not specifically agricultural soils (Ballabio et al., 2018). Such large-scale maps can provide insightful information, but they may not be optimal as

* Corresponding author.

E-mail addresses: Karl.Adler@Slu.se (K. Adler), Kristin.Piikki@Slu.se (K. Piikki), Mats.Soderstrom@Slu.se (M. Söderström), Jan.O.Eriksson@Slu.se (J. Eriksson).

decision support for crop production at field or within-field level. Hence, there is a need for high-resolution maps of topsoil Cu concentrations that can be used at field level. Digital soil mapping (DSM), or predictive soil mapping, is a popular framework for predicting soil variables in space using covariates (McBratney et al., 2003; Scull et al., 2003). Mapping is often done using either classical geostatistical methods such as kriging or, more popular of late, machine learning methods such as random forest (McBratney et al., 2003; Lamicchane et al., 2019; Wadoux et al., 2020), alone or in combination with a geostatistical method. Machine learning in DSM can be argued to be more data-driven than theory-driven modelling. Being more data-driven, it can provide an opportunity for hypothesis generation (Wadoux et al., 2021), e.g. by using covariate importance to explore the relationships between the environmental covariates and the soil property of interest, but also their role in predicting the soil property in question (Arrouays et al., 2020).

Representative calibration data of the study area, i.e. soil observations, are needed if DSM of trace elements is to be successful (McBratney et al., 2003). This commonly means using soil datasets with data from laboratory measurements, e.g. wet chemistry analysis of trace elements. However, conventional wet chemistry can be time-consuming and expensive for large-scale studies. Portable X-ray fluorescence (PXRF) measurements have been shown to be a successful alternative for rapid and accurate quantification of concentrations of trace elements such as Cu (Adler et al., 2020; Hu et al., 2014; Shresta et al., 2022), especially when used in an ex situ setting with soil sample preparation steps such as drying, sieving and homogenisation (Goff et al., 2020).

Specific objectives of the present study were to:

- use PXRF measurements to predict Cu concentrations for a set of soil samples lacking laboratory-measured Cu concentrations, in order to extend the total calibration dataset for the subsequent DSM model
- create a digital map of topsoil Cu concentrations from various covariates using a national DSM model ($50\text{ m} \times 50\text{ m}$ grid) encompassing >90% of agricultural land in Sweden
- determine the most important covariates in the DSM model for Cu using two different importance metrics
- assess whether the DSM model can be used to depict within-field variation using an independent validation set from five agricultural fields.
- assess whether the map produced can be employed to identify areas at risk of Cu deficiency by using classification metrics and prediction interval information.

2. Materials and methods

2.1. Soil samples and study area

2.1.1. National datasets

Topsoil samples from two surveys were used in the present study. The first, denoted NV, derived from the Swedish national survey of agricultural topsoil (one sample per 1300 ha), funded by the Swedish Environmental Protection Agency (Eriksson, 2021). From this collection, 1434 soil samples generally evenly spaced across agricultural land were used in this study. The second set of topsoil samples, denoted JV, was collected within a project led by the Swedish Board of Agriculture that aimed to complement the NV soil sampling in terms of geographical coverage. A total of 11,093 soil samples from the project (one sample per 100 ha) were used in this study (Fig. 1a) (Swedish board of Agriculture, 2015). In both surveys, each soil sample was a composite sample consisting of nine subsamples taken at 0–20 cm depth within a circle of 3 m radius from the coordinate point. Our study area encompassed >90% of agricultural land in Sweden and had the same geographical spread as the JV samples. Given the uneven distribution of agricultural land in Sweden, the study area covered the southern half of Sweden. The study area and soil samples refer to soils with <20% organic matter content.

Laboratory-analysed pseudototal Cu concentrations are available for the NV soil samples, but not the JV samples. The Cu concentrations were determined for the NV samples using inductively coupled plasma mass spectrometry (ICP-MS), after extraction with 7 M HNO₃ in an autoclave at 120 °C for 30 min, according to Swedish standard SS 28311 (Swedish Institute for Standards, 2017).

2.1.2. Field dataset

A set of soil samples from Bjertorp Farm, 130 km north-east of Gothenburg, collected in the year 2000 were used to evaluate how well the produced digital soil map reproduced field-average Cu concentrations and spatial variations in Cu concentrations within fields. The samples were also used to assess the validity, i.e. coverage of observations, of the 90% prediction interval. The set comprised 617 topsoil samples from five different fields (25–47 ha), with roughly four samples per hectare in an approximate square grid (Fig. 1b). These fields are located in an intensively cultivated, relatively flat area with soil texture varying from sand and loamy sand to silty clay and clay. Each soil sample was a composite, consisting of nine subsamples taken at 0–20 cm depth within a circle of 3 m radius from the coordinate point. Pseudototal Cu concentrations in the soil samples were determined using

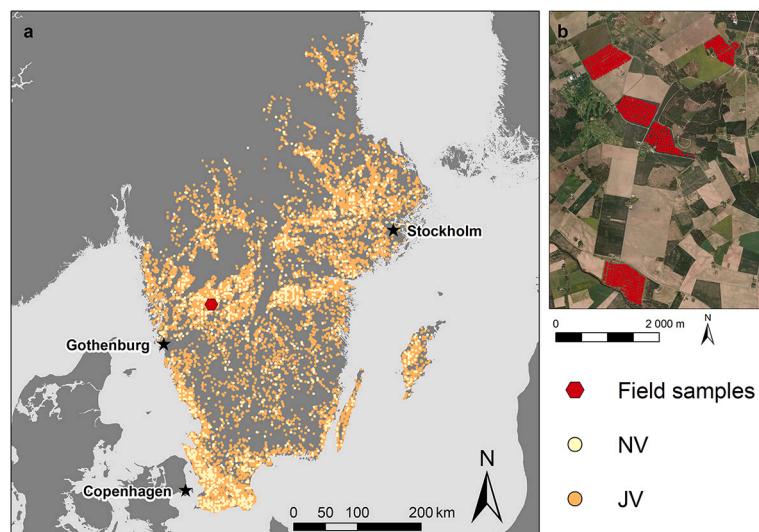


Fig. 1. (a) Map of soil sample locations for the NV, JV and field datasets in southern Sweden and (b) aerial photo showing sampling sites in the five validation fields. Three geographical names are shown for spatial reference. Basemap in (b) courtesy of ESRI, Redlands, CA, USA.

inductively coupled plasma atomic emission spectroscopy (ICP-AES) after extraction with 2 M HCl in containers immersed in boiling water for two hours.

2.2. Physical geography of the study area

The Fennoscandian shield, a generally felsic bedrock segment, is the prominent bedrock feature in Sweden (Stephens et al., 1997). Common rocks are gneiss, granite and granodiorite, in most cases 1–2 billion years old. However, younger sedimentary bedrock, e.g. sandstone, limestone and alum shale, overlies this felsic bedrock in relatively small areas in southern Sweden and on the two large Swedish islands in the Baltic Sea (SGU, 2020). The soil sediments in Sweden are young and often relatively thin, and are predominantly a product of the latest glaciation (Fredén, 1994). Hence, a large proportion of the soil parent material is glacial till. However, a large part of agricultural land in Sweden is on glacial or postglacial sorted sediments. In the east, agricultural soils are generally more clayey and heavy clays are frequent. In the west and the south, sandier soils are more frequent. In the very south of the mainland and on the two islands to the east, the textural class is often clay till. Pedologically, most Swedish agricultural soils are Cambisols and Regosols (Soil Atlas of Europe, 2005).

2.3. Software

Modelling was carried out using the Scikit-learn package (version 0.24.2) in the Python programming language (Pedregosa et al., 2011). Results were visualised using the package Datashader (version 0.12.1) and Matplotlib (version 3.1.3) (Hunter, 2007). The overview map was created using ArcGIS (version 10.6, ESRI, Redlands, CA, USA).

2.4. PXRF methodology

2.4.1. PXRF measurements

We used data from the PXRF measurements of the NV samples from Adler et al. (2020). The JV samples were measured at the same time. All soil samples were analysed using a Niton XL3t GOLDD+ instrument with a geometrically optimised large area drift detector, with a Silver (Ag) anode operating at 50 kV and 200 μ A (Thermo Scientific, Billerica, MA, USA). This instrument was used in a laboratory setting, where it was mounted on the static frame provided for the device and connected to a computer. Measurement time was set at 180 s in the “soil mode” of the instrument (Adler et al., 2020). Before the measurements, all soil samples were homogenised, dried and sieved (<2 mm) in accordance with recommendations (Weindorf and Chakraborty, 2020; United States Environmental Protection Agency, 2007). The limit of detection was set at three times the standard deviation of the measurement error. Soil samples were placed on the aperture in a 32 mm XRF sample cup with a transparent polypropylene XRF film with 4 μ m thickness (Adler et al., 2020).

2.4.2. PXRF model

The elements chosen to predict Cu concentration in this study were those frequently showing concentrations above the limit of detection for a majority of the soil samples in Adler et al. (2020). Those elements were lead (Pb), caesium (Cs), manganese (Mn), zinc (Zn), barium (B), vanadium (V), rubidium (Rb), titanium (Ti), potassium (K), iron (Fe), strontium (Sr), zirconium (Zr) and calcium (Ca). As reference and to check the stability of the instrument, Adler et al. (2020) analysed the certified reference soil sample 2709a from the National Institute of Standards and Technology (NIST) four times with the PXRF device during the measurement period. Recovery rates for the selected elements compared with the NIST 2709a reference sample are presented in more detail in Adler et al. (2020). Recovery rates were found to be stable over time and thus usable in modelling.

Two different machine learning models were calibrated on the NV

soil samples and used to predict Cu concentrations in the JV soil samples. Adler et al. (2020) found that non-linear and tree-based models were most accurate for predicting Cu concentrations from PXRF measurements. The two algorithms used in the present study were Extremely Randomized Tree Regression (ERT) and Gradient Boosting Regression (GBR), which are both non-linear and tree-based, but differ in other ways. The ERT algorithm is very similar to that of classical random forest regression except in two specific respects. The first is that the splitting thresholds for each node are drawn at random for each covariate candidate (i.e. element in this case), with the best performing being chosen as the threshold to reduce variance at the expense of bias, while the second is that there is no bootstrapping of samples (Geurts et al., 2006). The GBR algorithm works by fitting shallow (weak/stump) decision trees in sequence to minimise a loss function (Friedman, 2001). Initial testing revealed that both algorithms performed similarly, but with varying accuracy at low or high level of predicted Cu concentration. Furthermore, exhaustive hyperparameter optimisation of a single algorithm was considered too time-consuming and lacking in predictive performance gain. Instead, three GBR and three ERT models were calibrated with different values for a specific hyperparameter to obtain the mean predicted Cu concentration. This could be defined as model averaging but with added hyperparameter variation, such that the PXRF model is a product of a total of six sub-models (Hastie et al., 2009). The chosen hyperparameter was *max_features*, which determines the number of covariates to look for at each split of a node. This hyperparameter is available in both ERT and GBR. Of the most common adjustable hyperparameters (Scikit-learn, 2021), *max_features* had the greatest impact on accuracy in the present study. The values chosen for the fraction of covariates used were 0.3, 0.6 and 1.0. All other settings were set at default apart from *n_estimators*, which was set to 500. The hyperparameter *n_estimator* determines the number of trees created. This ensemble approach enables shorter run time, since no exhaustive hyperparameter search is needed. As a result, after predicting Cu on the JV soil samples, the total number of soil samples available for calibrating the DSM model was 12,527 (JV and NV sets combined).

2.5. DSM methodology

2.5.1. DSM model

The GBR algorithm was used to create the DSM model, which was calibrated on the combined NV and JV datasets. The main reason for choosing the GBR algorithm was that it is possible to choose the kind of loss function that is minimised in model calibration. This makes it possible to estimate the prediction uncertainty, by creating a prediction interval using quantile loss (95th – 5th percentile). A prediction interval is an estimate where future observations will fall, given a probability. In this study, least squared loss was used for the model for predicting Cu concentration, while quantile loss was used to create models to predict the 90% prediction interval. The DSM model prediction was done using model averaging, as in the PXRF model. Hence, a total of three GBR models were calibrated using three different settings of a hyperparameter and these three GBR models made up the resulting DSM model. Since covariate importance was calculated, each covariate needed to be eligible for the construction of the regression trees. Hence, the hyperparameter *max_feature* was omitted and instead the hyperparameter *subsample* was chosen, with three different values for the fraction of samples used for bootstrapping (0.3, 0.6 and 1.0). This reduces the variance at the expense of bias and is often referred to as stochastic gradient boosting (Friedman, 2002).

2.5.2. DSM covariates

A total of four different types of covariates were prepared for this study, derived from (i) a digital elevation model (DEM), (ii) airborne gamma radiation measurements, (iii) climate data and (iv) soil texture class maps. The main dataset of covariates used in this study was the 50 m × 50 m grid dataset created by Piikki and Söderström (2019), which

covers >90% of arable land in Sweden (year 2013). In that study the dataset was used to predict clay and sand content. This dataset contains airborne gamma ray measurements of thorium (^{232}Th), uranium (^{238}U) and potassium (^{40}K). It also includes topographic position index (TPI) with three different neighbourhood areas (5, 50 and 500 ha) and elevation (m above sea level), as well as six binary soil texture classes created by transforming the classes in a Quaternary deposit map (Table 1) (Geological Survey of Sweden (SGU), Uppsala, Sweden). This SGU quaternary map product reflects general soil texture classes at a depth of 0.5 m.

New covariates were added such as climate, soil moisture and new calculations of airborne gamma ray measurements (Table 1). Climate covariates were from the 4 km \times 4 km grid dataset provided by the Swedish Meteorological and Hydrological Institute (SMHI, 2015). That dataset contains data on precipitation and air temperature at 2 m above ground, both annually and seasonally, for two reference periods (1961–1990 and 1991–2013). The seasons are divided into spring (MAM; March, April and May), summer (JJA; June, July and August), autumn (SON; September, October and November) and winter (DJF; December, January and February). All climate covariates were calculated as the mean of the two reference periods.

Raster soil moisture data from the study by Ågren et al. (2021) were used. This raster indicates likelihood of soil moisture, with values ranging from 0 to 100, dry to wet. The raster was re-sampled from resolution of 2 m \times 2 m to 50 m \times 50 m. The original dataset was mainly calibrated and validated on forest soils, but also contains information for agricultural soils in Sweden.

The airborne gamma ray measurements were used to calculate the gamma radiation dose rate (Duval et al., 2005) as:

$$\text{Dose rate} = 13.2 \times K + 5.48 \times U + 2.72 \times Th \quad (1)$$

where K is measured ^{40}K in %, U is measured ^{238}U in mg kg $^{-1}$ and Th is measured ^{232}Th in mg kg $^{-1}$, resulting in a dose rate in nGy hr $^{-1}$. The ratios of K/Th, K/U and Th/U were also calculated.

2.5.3. Validation metrics

Two specific validation metrics were used: (i) Nash-Sutcliffe model efficiency coefficient (E) (Nash and Sutcliffe, 1970) and (ii) mean absolute error (MAE), the latter as it is less sensitive to asymmetric

Table 1

Covariates used in the digital soil mapping (DSM) model, with type and source. Total number of covariates used was 28. Original data source: * Derived from a compilation by Piikki and Söderström (2019). TPI = topographic position index, Lantmäteriet = Swedish Land Survey.

Covariate	Type	Source
U (mg kg $^{-1}$)	Gamma radiation	SGU*
Th (mg kg $^{-1}$)	Gamma radiation	SGU*
K (%)	Gamma radiation	SGU*
Dose rate (nGy hr $^{-1}$)	Gamma radiation	Computed from U, Th and K
K/Th, K/U and Th/U	Gamma radiation	Computed from U, Th and K
TPI 5, 50 and 500	DEM	Computed from elevation*
Soil moisture	DEM	Ågren et al. (2021)
Elevation (m)	DEM	Lantmäteriet*
Precipitation, annual (mm)	Climate	SMHI (2015)
Precipitation, seasonal (MAM, JJA, SON and DJF) (mm)	Climate	SMHI (2015)
Temperature, annual (°C)	Climate	SMHI (2015)
Temperature, seasonal (MAM, JJA, SON and DJF) (°C)	Climate	SMHI (2015)
Soil texture classes (Clay, Clay till, Till, Silt, Sand and Other)	Soil texture	SGU*

distributions of error than root mean square error (RMSE) (Janssen and Heuberger, 1995). The prediction interval coverage probability (PCIP) was used to assess the validity of the prediction interval of the DSM model. This value should ideally be as close to the prediction interval percentage as possible (Shresta and Solomatine, 2006).

2.5.4. Covariate importance

Two different methods were used to determine the most important covariates in the DSM model: mean decrease in impurity (MDI) and permutation importance (PI). Mean decrease in impurity is a measure of how many times a covariate is used for a split in the nodes of the regression trees, and its hierarchy (Breiman, 2001). For instance, the earlier a covariate is used for a split, the more important it is considered by MDI. Permutation importance works by first establishing a baseline performance score, i.e. E or MAE, for the calibrated model against either the calibration data or independent validation data. Then each covariate is permuted (randomly shuffled) while keeping the others fixed, and the deviation in the performance score is noted (Breiman, 2001; Strobl et al., 2008). This is done for a set number of times for each covariate, and the mean deviation is reported. Hence, an important covariate will have a higher deviation in the performance score, and thus a higher PI score. In this study, the number of times a covariate was permuted was set at 10, the performance score was E and PI was calculated on the calibration data. The PI and MDI scores reported for each covariate are the mean of the three GBR models that make up the DSM model. Note that covariate importance gives a metric and ranking for the most valuable covariates in the model, and not for the real-world system the model is trying to predict.

2.6. Validation and risk assessment

Both the PXRF model and the DSM model were validated using five-fold cross-validation. However, the cross-validation for the DSM model was constructed in such a way that validation was always done against the NV dataset, but calibration was done on the NV and JV dataset together. For example, a calibration fold in the DSM model cross-validation consisted of 80% of the NV dataset and the whole JV dataset, and the validation fold was then the remaining 20% of the NV dataset. The PICP was computed during cross-validation of the DSM model.

2.6.1. Evaluation at the field level

The independent field soil sample dataset was used to, in one subarea for which we had data, test how well the map produced performed at field level and how often the prediction interval included the observed Cu concentration. The PICP was used to assess the prediction interval, i.e. map uncertainty, and mean absolute error and E were also used. The squared Pearson correlation coefficient (r^2), i.e. field-wise linear regression between predicted and observed values, was used to assess how well the map explained the variation in Cu concentrations in each field.

2.6.2. Cu deficiency assessment

To quantify how well the DSM model correctly predicted soil samples at or below specific limits of Cu concentration in the soil, the recall and precision metrics were used. Positives in this study refer to soil samples having a concentration of Cu at or below the defined risk limit. Recall, sometimes called sensitivity or true positive rate (TPR), refers to the fraction of actual positives that were predicted as positives, where T_p is the number of true positives and F_n the number of false negatives:

$$\text{Recall} = \frac{T_p}{(T_p + F_n)} \quad (2)$$

Precision, sometimes called positive predictive value (PPV), refers to the fraction of predicted positives that were actual positives, where T_p is the number of true positives and F_p is the number of false positives:

$$\text{Precision} = \frac{\text{Tp}}{(\text{Tp} + \text{Fp})} \quad (3)$$

In the DSM context, a low recall score may indicate that the DSM model cannot accurately predict samples at or below a set limit. However, the precision score can simultaneously be high, indicating that when the DSM model predicts at or below the set limit, it is probably correct.

Precision and recall were calculated for every integer step (1, 2, 3, ..., 60 mg kg⁻¹) of Cu concentrations to create a curve for each respective metric, using the cross-validation predictions of the DSM model. For example, the first integer step set all soil samples with laboratory-measured Cu concentration of 1 mg kg⁻¹ or less as positives, and all others as negatives. The same was done with predictions of the DSM model. Lastly, the recall and precision were calculated. This was done to see how well the DSM model performed at predicting at or below various limits, and not only the deficiency risk limit of 7 mg kg⁻¹. It should be mentioned that the risk limit 7 mg kg⁻¹ refers to Cu concentrations determined after extraction with 2 M HCl, while the Cu concentration from the NV data set used in the modelling were determined after extraction with 7 M HNO₃. Presumably, 7 M HNO₃ is a slightly stronger extractant than 2 M HCl, although few published data to confirm this are available. However, unpublished data from a comparison of different methods on 42 soils from Sweden and Scotland indicate that 7 M HNO₃ extracts nearly the total amount of Cu (HF digestion) and ca 10% >2 M HCl does. For simplicity, we still use 7 mg kg⁻¹ as limit value for high risk for deficiency of Cu in crops.

3. Results

3.1. PXRF modelling and cross-validation

The cross-validation plots of the PXRF model revealed a relatively even spread around the 1:1 line (Fig. 2). The E and MAE value in cross-validation of the PXRF model was 0.66 and 3.3 mg kg⁻¹, respectively. However, there was a slight positive bias at lower concentrations (0–10 mg kg⁻¹) and strong underestimation for five samples with concentrations above 60 mg kg⁻¹ (Fig. 2).

Descriptive statistics revealed a slight positive bias (≈ 1 mg kg⁻¹) in the median, mean, 25th percentile, 75th percentile and minimum of the

predicted Cu concentrations in the JV soil samples compared with the measured Cu concentrations in the NV samples (Table 2). Since both sets of samples were probability samples from the same geographical area, similar frequency distributions of Cu concentration can be expected. The percentage of samples in the NV dataset with values below the risk limit of 7 mg kg⁻¹ was 21%, which is close to the 22% reported in Eriksson et al. (2017) and indicates that a risk of Cu deficiency is relatively common in Swedish agricultural soils. In both cases the percentages of values below the risk limit may be slightly higher than estimated since the estimates are based on Cu-HNO₃ and the risk limit refers to Cu-HCl (see Cu deficiency assessment in Materials and Methods).

3.2. Digital soil mapping and cross-validation

Cross-validation of the DSM model showed more spread around the 1:1 line compared with the PXRF model (Figs. 2 and 3). The E and MAE value for cross-validation of the DSM model was 0.57 and 4.1 mg kg⁻¹, respectively. As with the PXRF model, concentrations in soil samples with measured Cu concentrations above 40 mg kg⁻¹ were generally underestimated (Fig. 3).

Cross-validation of the DSM model only calibrated on the NV dataset, i.e. not NV and JV combined, resulted in E of 0.4 and MAE of 4.6 mg kg⁻¹. Hence, including the JV dataset resulted in a more accurate DSM model. The positive bias seen at lower concentrations (0–15 mg kg⁻¹) in Fig. 3 was a result of inability of the DSM model to predict concentrations lower than 5 mg kg⁻¹, similarly to the PXRF model. The only soil sample with a measured concentration of 130 mg kg⁻¹ was slightly more

Table 2

Descriptive statistics on copper (Cu) concentration in the NV soil samples ($n = 1434$) used as the calibration dataset for the PXRF Cu model and on Cu concentrations predicted on the JV soil samples ($n = 11,093$) by the PXRF Cu model.

Statistic	NV (mg kg ⁻¹)	JV (mg kg ⁻¹)
Maximum	129.9	87.0
75th percentile	17.9	18.3
Median	10.9	11.9
Mean	13.9	14.4
25th percentile	7.6	8.9
Minimum	1.7	2.8
Standard deviation	9.8	7.7

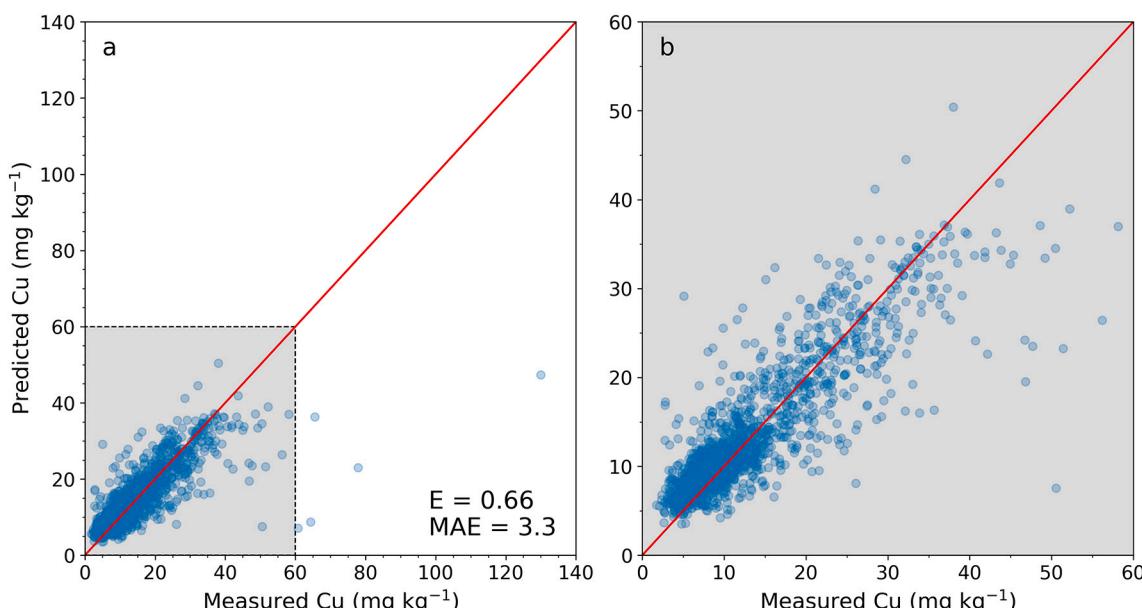


Fig. 2. Results from cross-validation, using five folds, of the PXRF model (a) including the full range of measured copper (Cu) concentrations and (b) zoomed in. The grey area represents the extent of the zoom-in and the red line is the 1:1 line ($n = 1434$; E is Nash-Sutcliffe model efficiency coefficient, MAE is mean absolute error).

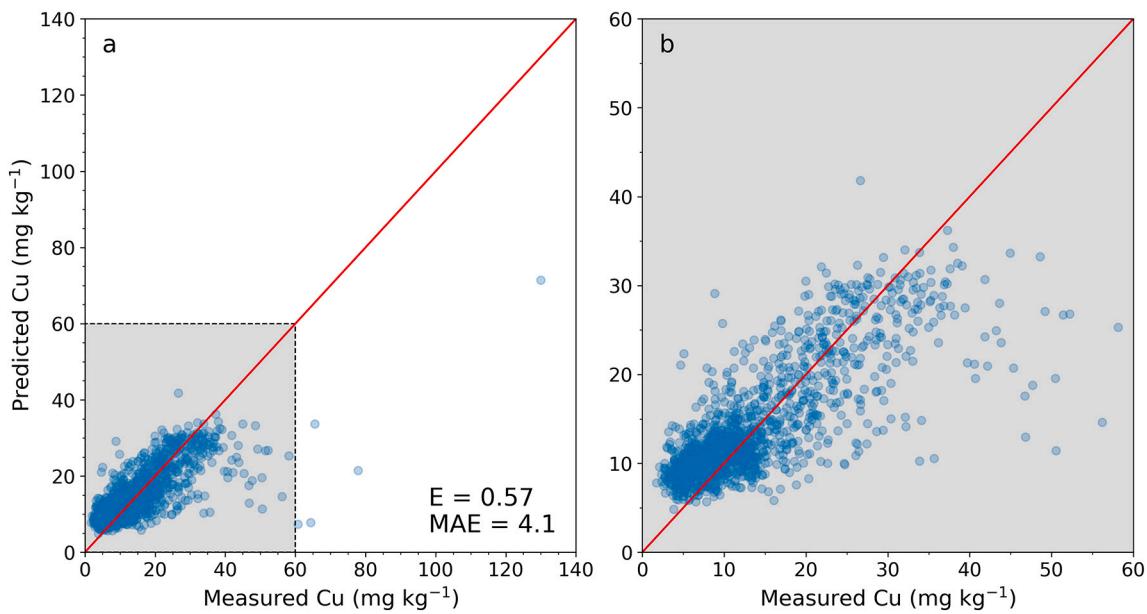


Fig. 3. Results from cross-validation, using five folds, of the DSM model (a) fully zoomed out and (b) zoomed in. The grey area represents extent of the zoom-in and red line is the 1:1 line ($n = 1434$; E is Nash-Sutcliffe model efficiency coefficient, MAE is mean absolute error).

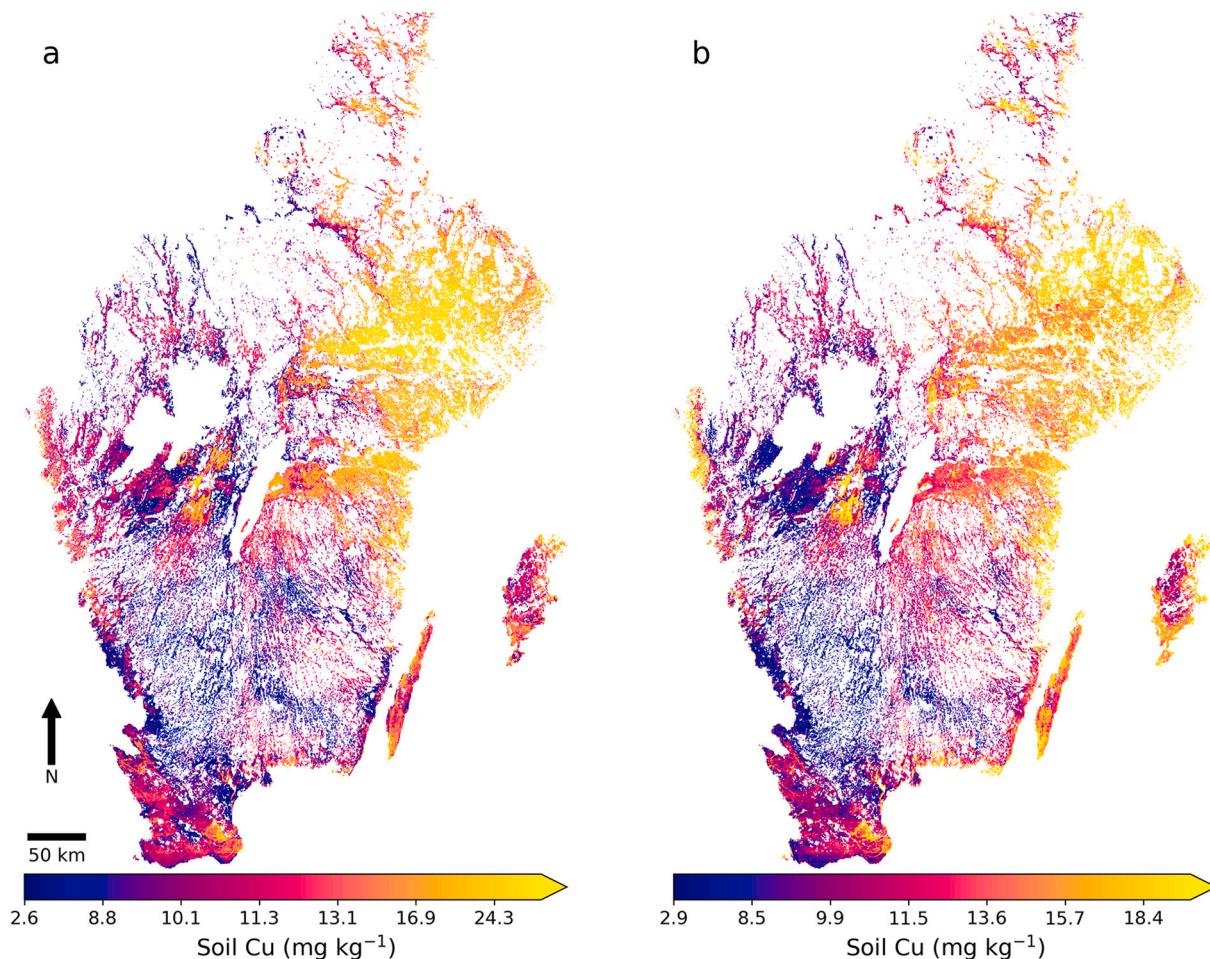


Fig. 4. (a) Map of predicted copper (Cu) concentrations in agricultural soil in southern Sweden and (b) width of the 90% prediction interval. The maximum concentration denoted in the colour bar for (a) and (b) (24.3 and 18.4 mg kg^{-1} , respectively) corresponds to the 90th percentile.

accurately predicted by the DSM model than the PXRF model. Hence, the DSM model might be more accurate at predicting very high Cu concentrations. The PICP from the cross-validation was 76%, indicating that the prediction interval produced was too narrow. The true concentration was below the lower bound 216 times and above the upper bound of the prediction interval 124 times. This indicates that it was more common that the lower bound was not low enough.

Predicted concentrations of Cu were generally high in the north-eastern part of the study area, whereas lower concentrations were seen in the mid-southern part (Fig. 4a). However, in some areas concentrations varied from low to high within a short distance, e.g. in the mid-western and southern part of the study area and on Gotland, the eastern-most of the two islands (Fig. 4a). The mean and median predicted Cu concentration was 14.1 mg kg^{-1} and 11.8 mg kg^{-1} , respectively, indicating positive skewness in predicted concentrations like that of the calibration dataset (cf. Table 2 and Table 3). Furthermore, the maximum and minimum predicted Cu concentrations could be regarded as rare when compared to the 25th and 75th percentile (Table 3). A total of 3% of all agricultural land was predicted to have a Cu concentration at or below 7 mg kg^{-1} , indicating that the DSM model is not efficient for extremely low values. However, Eriksson et al. (2017) showed that 22% of Swedish agricultural land can be regarded as being below the deficiency risk limit. Hence, there was large underestimation for agricultural soils predicted to be below the deficiency risk limit for Cu.

The 90% prediction interval was generally wider in the north-eastern part of the study area and narrower in the mid-southern part (Fig. 4b). A wider prediction interval was more common in areas with predicted high concentrations of Cu, and narrower in areas with predicted lower concentrations (Fig. 4b). However, for the southern-most part, which had medium-high concentrations of predicted Cu, the prediction interval was narrow (Fig. 4b). Hence, the relationship between wide prediction intervals and high predicted Cu concentrations was not consistent. In addition, the median and mean were similar and the distribution of the prediction intervals was less skewed compared with the distribution of predicted Cu concentrations (Table 3).

3.2.1. Covariate importance

Airborne gamma radiation covariates emerged as most important when the rankings for the two covariate importance methods were weighted together (Fig. 5). Specifically, Th, U, K/Th and dose rate ranked high, while K was considered less important by both methods.

Climate covariates, especially seasonal precipitation and temperature, were identified as important, with precipitation in March, April and May at the top of the climate covariates. Generally, the temperature covariates were more important than the precipitation covariates according to the PI scores. Elevation was ranked high by both methods. DEM derivatives such as TPI and soil moisture, as well as texture classes, had lower rankings (Fig. 5).

3.3. Performance at the field level – An example

Field evaluation of the national Cu concentration map revealed a general pattern of low prediction error compared with measured Cu concentrations in the soil samples, apart from in Fields 4 and 5 (Fig. 6).

Table 3

Descriptive statistics on predicted copper (Cu) concentration and 90% prediction interval in agricultural soils in southern Sweden.

Statistic	Predicted Cu (mg kg^{-1})	Prediction interval (mg kg^{-1})
Maximum	129	109
75th percentile	16.9	15.7
Median	11.8	12.2
Mean	14.1	12.9
25th percentile	9.7	9.4
Minimum	2.6	2.9
Standard deviation	6.4	4.8

However, the scatterplot ($E = 0.56$, $\text{MAE} = 2 \text{ mg kg}^{-1}$) and the r^2 values for field-wise linear regressions between predicted and observed values revealed that the map of Cu concentration did not completely capture the variation within each field (Fig. 7, Table 4). Nevertheless, the field mean of measured Cu concentration was generally close to the predicted field mean, with a MAE between predicted and measured field means of 1.0 mg kg^{-1} , and with the largest deviation in Field 3 and 4 (Fig. 6, Table 4). Evaluation of the prediction interval resulted in a PICP of 85% across all fields, indicating that the prediction interval should be slightly wider for this farm. Measured concentrations were found outside the lower and upper bound of the prediction interval.

3.4. Risk assessment and classification accuracy

The DSM model had a low recall score (0.1) when attempting to predict correctly whether a soil sample had a Cu concentration at or below 7 mg kg^{-1} (Fig. 8). This means that if a soil sample had a Cu concentration at or below the deficiency risk limit, the DSM model had a 10% probability of correctly predicting it as such. However, the precision score was much higher (0.76). This means that when the DSM model predicted the Cu concentration of a soil sample to be at or below 7 mg kg^{-1} , it was correct with a 76% probability.

With increasing Cu concentration, the recall and precision increased rapidly (Fig. 8). Thus, if the deficiency risk limit were to be reformulated to 10 mg kg^{-1} , the probability of the DSM model correctly predicting Cu concentrations at or below this concentration in soil samples increased to around 50% (Fig. 8). The graph in Fig. 8 begins where the first recorded score had a value above 0, e.g. at 1 mg Cu kg^{-1} both the recall and precision score were zero.

The low recall and precision scores indicate that only using the Cu concentration map to locate agricultural soils at risk of Cu deficiency is not sufficient. Ideally, the upper bound, i.e. 95th percentile, should be used to locate areas with high probability of Cu concentrations below 7 mg kg^{-1} . However, no points had an upper bound at or below this value. An alternative can be to use the lower bound of the prediction interval, i.e. the 5th percentile, to locate areas with a low risk of soil Cu concentrations below the deficiency risk limit. As shown in Figs. 9, 48% of the agricultural soil had a lower prediction interval bound at or above 7 mg kg^{-1} , which implies that for 48% of the agricultural land in Sweden there is a high probability of soil Cu concentration being above the deficiency risk limit (i.e. whole prediction interval above the risk limit). Conversely, 52% of agricultural soils can have Cu concentrations below the deficiency risk limit, albeit with widely varying probabilities.

4. Discussion

4.1. DSM model performance and cu map applicability

Cross-validation showed that Cu concentrations below 5 mg kg^{-1} and above 40 mg kg^{-1} were difficult to predict (Fig. 3), while evaluation at field level gave similar indications (Fig. 7). The problem of predicting low concentrations was also apparent from the low recall score when predicting concentrations below 10 mg kg^{-1} . Furthermore, the variation in Cu concentration within each field could not be predicted accurately, but the mean Cu concentration in each field was predicted relatively well (Table 4).

Only 3% of agricultural land in the study area was predicted to have soil Cu concentrations at or below 7 mg kg^{-1} , a strong contrast to the 22% reported for 5170 sites in the Swedish national survey of agricultural topsoil (Eriksson et al., 2017). Hence, the Cu concentration map underestimated the amount of agricultural land at or below the deficiency risk limit, as evidenced by the low recall score. However, the precision score showed that predictions at or below 7 mg kg^{-1} were correct with 76% probability. Nevertheless, only using the Cu concentration map as decision support can be insufficient, as shown by the recall and precision scores. However, using the prediction interval map

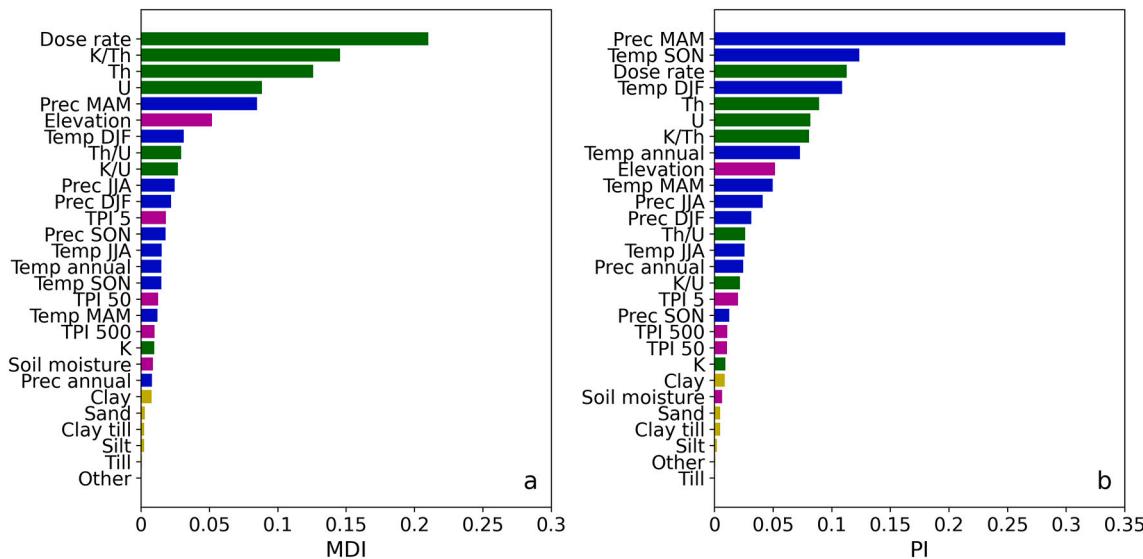


Fig. 5. Covariate importance. Score of (a) mean decrease in impurity (MDI) and (b) permutation importance (PI). Each covariate is coloured according to its covariate type (green: airborne gamma radiation data; blue: climate data; magenta: digital elevation model; yellow: Quaternary deposit texture classes). Prec and Temp refers to precipitation and temperature, respectively. Seasonal suffixes: MAM (March, April, May), JJA (June, July, August), SON (September, October, November), DJF (December, January, February).

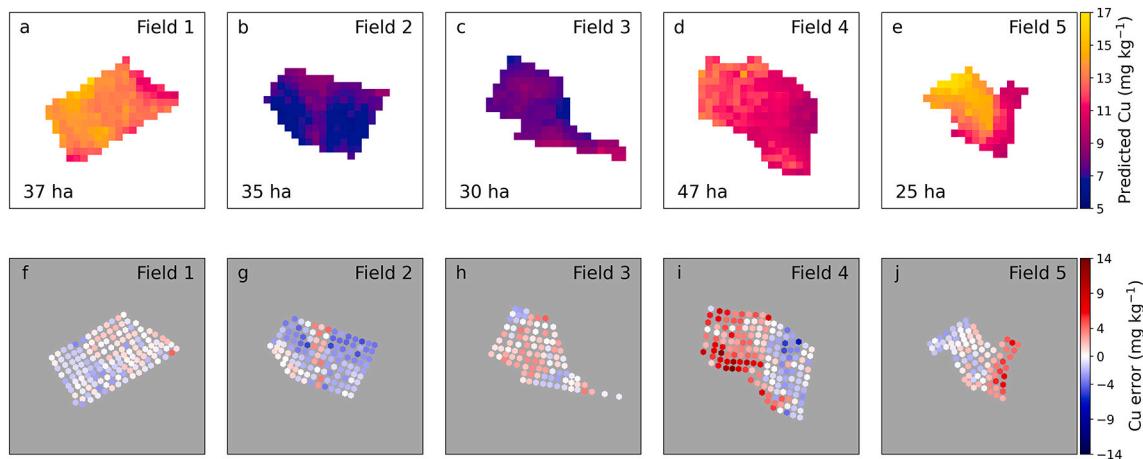


Fig. 6. Maps of (a-e) soil copper (Cu) concentrations and (f-j) prediction error at Bjertorp (zoomed in from map in Fig. 4a), with numbering and size (ha).

with its lower bound can improve the matter. For example, 48% of agricultural land had soils with a predicted lower bound above the 7 mg kg^{-1} (Fig. 9), so the likelihood of Cu concentrations below the deficiency risk limit is very small in these areas. The focus should thus be redirected to the remaining 52% of agricultural land which has a higher risk of Cu concentrations below the deficiency risk limit. End users, such as, advisors and farmers, need information presented in a format suitable for their decisions, e.g. decisions on whether or not to apply fertilizer with Cu, in a specific field. Our evaluation show that Cu maps presented here are not certain enough to directly serve as decision support for this decision (Fig. 8), but by converting the derived information (the lower bound prediction interval) to map showing areas where it is highly likely that the Cu level is above the tabulated value for Cu-deficiency risk, one can use the derived information (Fig. 9) to make decisions on whether or not to take a sample for Cu analyses in the lab (green: most likely above the risk limit – no need to take a sample, grey: uncertain whether the Cu concentration is above or below the risk limit, lab analysis is recommended to determine whether or not to fertilize).

Ideally, the upper bound of the prediction interval should be used to locate areas likely at risk of Cu deficiency. However, this did not work in

the present case due to the width of the prediction interval, as no points had an upper bound below 7 mg kg^{-1} . Hence, the 95th percentile was not usable for locating areas with a high probability of being below the deficiency risk limit. The PICP of 76% from cross-validation of the DSM model was lower compared to that off the independent field evaluation. The prediction interval produced was too narrow and should be wider. The PICP of 85% from the field evaluation showed a more positive view of the validity of the produced map uncertainty. However, the PICP from the field evaluation showed that the prediction interval was also too narrow. Thus, future efforts are needed to optimize the prediction interval to be wider than it is, to encompass more soil samples, as the PICP should ideally be the same as the prediction interval (90%) (Shresta and Solomatine, 2006; Poggio et al., 2021).

Another option could be to use the Cu concentration map and prediction interval map in combination to locate areas with a low Cu concentration and narrow prediction interval. However, using the maps in such a way can be difficult, e.g. it is uncertain how low the predicted concentration of Cu should be and how narrow the prediction interval should be. Robust classification is needed in order to make this option useful as decision support. This needs to be done in cooperation by

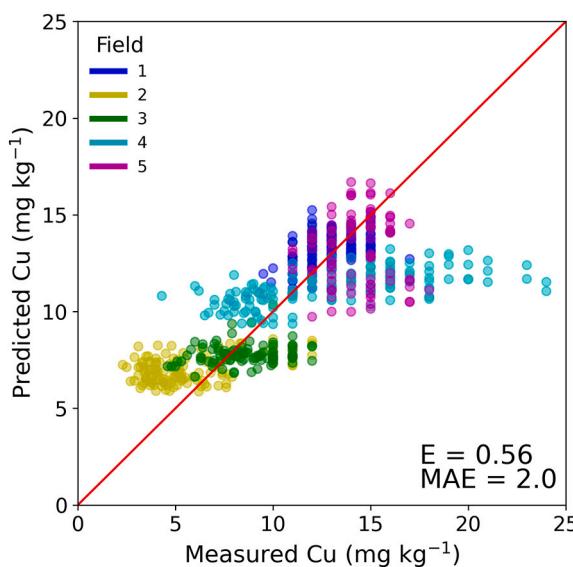


Fig. 7. Scatterplot of predicted concentrations of copper (Cu) according to the Cu map versus measured concentration of Cu in each soil sample ($n = 617$; E is Nash-Sutcliffe model efficiency coefficient, MAE is mean absolute error).

“modellers” and farmers or advisors, due to uncertainty being problematic in terms of communication and understanding (Richer-de-Forges et al., 2019; Arrouays et al., 2020).

As mentioned, using the Cu concentration map alone is not sufficient to locate areas with a risk of Cu deficiency, so information from the prediction interval must be used. Using the 5th and 95th percentile for this purpose may not be enough, since a point is either within or outside these two percentiles. More nuance is needed, so future work should try to predict all feasible percentiles (1, 2, 3 ..., 99th) and not only the 5th and 95th. This would make it possible to query every point and determine the percentile at which the deficiency risk limit is exceeded. For example, a point in the map could have a predicted 20th percentile of $6.9 \text{ mg Cu kg}^{-1}$, roughly indicating a 20% probability of being at or below the deficiency risk limit. However, this option would be computationally demanding, as 297 DSM models ($3 \text{ GBR} \times 99$) would need to be calibrated and applied to create all 99 percentiles, at least for the model averaging method used in this study. The 0th and 100th percentile should not be predicted, in order to avoid stating deterministic claims as they represent the minimum and maximum. Furthermore, it is unclear how all these predicted percentiles should be evaluated. A computationally feasible option could be to calculate PICP in ascending (1st to 2nd, 1st to 3rd ..., 1st to 99th) and descending (1st to 99th, 2nd to 99th ... 98th to 99th) order, to obtain descriptive statistics on the PICP.

4.2. Covariate importance and map interpretation

The geographic distribution of areas with high and low concentrations of Cu is similar to that in the map in Ballabio et al. (2018). More specifically, the Cu concentrations in the map produced here followed

the spatial variation in soil texture, with high predicted Cu concentrations generally in areas with high clay content and lower concentrations in areas with coarser-textured soils. This is in agreement with previous findings by Eriksson et al. (2017) and Piikki and Söderström (2019). The results for covariate importance reinforced this, as airborne gamma radiation covariates were generally regarded as most important (Fig. 5). Measured airborne gamma radiation values are often highly correlated to soil texture (International Atomic Energy Agency, 2003). This is a likely reason why the soil texture classes were somewhat redundant in the DSM model, as the airborne gamma radiation data already contained the necessary information for soil texture.

Climate covariates were also found to be important in the DSM model, especially seasonal climate covariates (Fig. 5), but the reason for this is not clear. Similarly, Hengl et al. (2017) found high importance of climate covariates when mapping Cu (Mehlich-3 extractable) in African soils. The concentration of Cu in a soil is a product of many factors that in turn rely on temperature and moisture at longer timescales (Oorts, 2013). In the present study, it is more likely that the climate data acted as a spatial partitioning covariate (spatial co-occurrence), where the actual spatial pattern was more important than the value itself. Precipitation in Sweden is generally higher in the west, where sandier soils are more common. These sandier soils often have a lower soil Cu concentration. Further investigation is needed to properly assess if it is the spatial partitioning or the actual climate values that are important. Seasonal climate covariates were identified as more important than annual climate covariates, perhaps because there is more spatial variation in the seasonal covariates. Soil moisture ranked low in importance, perhaps due to the precipitation and elevation covariates already containing similar information. It should also be mentioned that the quality

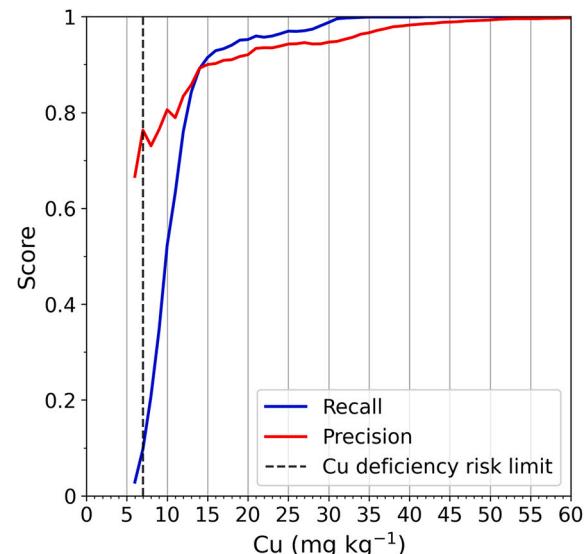


Fig. 8. Recall and precision scores for correct prediction of copper (Cu) concentration in soil samples at or below a certain limit based on cross-validation of the digital soil mapping model. The Swedish Cu deficiency risk limit is at 7 mg kg^{-1} .

Table 4

Measured and predicted field means of soil copper (Cu) concentration and squared Pearson correlation coefficient (r^2). Each field is referred to by the number and colour given in Fig. 7.

Field	Measured field mean Cu (mg kg^{-1})	Predicted field mean Cu (mg kg^{-1})	r^2
1 (Blue)	12.9	13.2	0.12
2 (Yellow)	5.5	7.0	0.17
3 (Green)	8.6	7.8	0.04
4 (Cyan)	13.0	11.3	0.36
5 (Magenta)	14.3	13.3	< 0.01

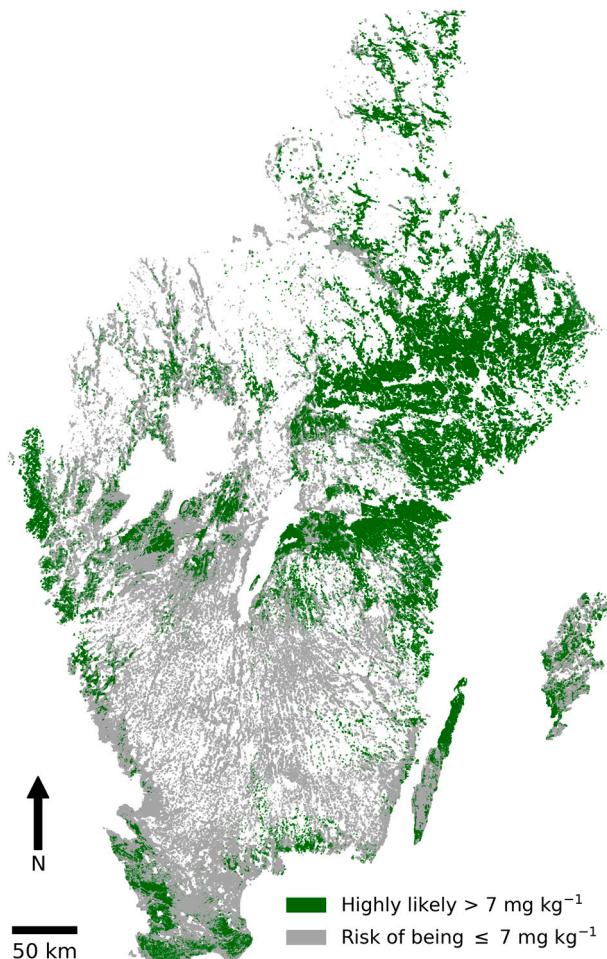


Fig. 9. Map of agricultural areas in Sweden where the soil concentration of copper (Cu) is highly likely to be above the Cu deficiency risk limit of 7 mg kg^{-1} and areas with a higher risk of concentration below that limit. Based on the lower bound of the prediction interval (5th percentile). The map pixels are aggregated for better visibility at this scale.

of the soil moisture map was not evaluated or calibrated for agricultural soils, but for forest soils. Nevertheless, the soil moisture covariate is mainly a product of elevation information and should still provide information better or equal to conventional soil moisture covariates, e.g. TWI (topographic wetness index). Elevation (m above sea level) ranked high in both methods, hinting at potential information redundancy of the DEM derivatives, which ranked lower. Nevertheless, elevation seems to be important in the DSM model. For instance, [Chen et al. \(2012\)](#) found that topographic information determined the spatial distribution of Cu in floodplain soil in China.

Two different methods for assessing covariate importance were used, with the aim of minimising potentially misleading results generated by only using one method. However, mean decrease in impurity importance can lead to misleading results as continuous covariates with many unique values often rank higher than covariates with few values ([Strobl et al., 2007](#)), e.g. the difference between unique values in the dose rate covariate versus the binary classes of soil texture covariates, as the tree-based model can find more splitting points in the dose rate covariate. Furthermore, permutation importance can also be computed on independent validation data to assess how covariates behave on unseen data, rather than the training data. More drawbacks exist with each method, and future focus should be directed towards the multitude of other covariate importance methods, e.g. those listed in [Wei et al. \(2015\)](#). Nevertheless, both methods resulted in similar rankings and both

revealed the importance of airborne gamma radiation covariates, climate covariates and elevation (m above sea level) in the DSM model. Inclusion of more covariates in future versions of the DSM model could probably improve the performance. One such covariate could be information on agricultural management, as it can affect Cu concentrations in agricultural soil ([Vavoulidou et al., 2011](#); [Oorts, 2013](#)). However, it is unclear how that information can be used as a covariate. An option could be to use reflectance data from remote sensing, such as in [Ballabio et al. \(2018\)](#), or crop type information.

The problems with predicting within-field variation could arise because of the underlying spatial resolution of the important covariates from airborne gamma radiation and climate data. For example, the original airborne gamma radiation measurements had a sampling radius on the ground of about four times the flight height (flight height was 30 m and 60 m) ([IAEA, 2003](#)). Hence, capturing within-field variation in measured U, Th and K can be problematic. Lastly, it needs to be emphasised that covariate importance gives indications of what is important to the model and perhaps also the natural system the model is trying to replicate, but by no means indicates any causal connections. However, it can strengthen pre-established causal connections and facts, such as Cu concentrations being linked to soil texture.

4.3. PXRF model performance

The PXRF model was used with the aim of expanding the calibration dataset for the DSM model. If the PXRF model had predicted inaccurate Cu concentrations, the subsequent DSM model would probably have been lacking in performance. Inclusion of predicted PXRF samples made the DSM more accurate, but the PXRF model had problems predicting concentrations below 5 mg kg^{-1} , which was probably accentuated in the DSM model. Hence, the applicability of the PXRF model in expanding the calibration dataset for the DSM model needs to be investigated further. An option could be to calibrate a PXRF model specifically with the intention of predicting lower concentrations of Cu. The high MAE (3.3 mg kg^{-1}) at low concentrations means that the PXRF model might not be suitable for the task of complementary soil analysis when confirmation is needed on whether the Cu concentration in a soil is below the deficiency risk limit or not. Nevertheless, the flexibility of machine learning and the speed of using PXRF measurements make this method an interesting alternative to conventional laboratory analysis.

5. Conclusions

A map of topsoil Cu concentrations encompassing $>90\%$ of Swedish agricultural area was produced, with the aid of 28 spatially extensive covariate datasets and a combination of laboratory-analysed and PXRF-predicted Cu concentrations for point locations. Recall and precision scores of the DSM model showed that the resulting Cu map is only useful to a limited extent for determining whether a specific location is at or below the Cu deficiency risk limit (7 mg kg^{-1}). However, the modelled uncertainty can be used as decision support in crop micronutrient management. The lower bound of the 90% prediction interval can be used to locate areas where Cu concentrations are unlikely to be below the deficiency risk limit, i.e. areas that can be considered safe in terms of Cu deficiency. Areas not in this category need to be assessed further, ideally by complementary soil analysis. In its present state, the Cu concentration map should be used as an explorative tool. At field level, the map could not fully capture the variation in concentration within each field, but the predicted field-average Cu concentration was relatively accurate. Covariate importance assessments showed that airborne gamma radiation covariates were most important for the DSM model, followed by climate covariates and elevation. Using the lower prediction interval bound for ruling out areas where Cu deficiency is unlikely can be a generally applicable approach to derive decision support for micronutrient management.

Data statement

The soil sample datasets used in this study are not freely available. Covariates used are not freely available, but can be accessed on request.

Funding

Research funded by Västra Götaland Region (VGR) and the Swedish University of Agricultural Sciences (SLU), grant number: 2018–00141.

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.

Data availability

The authors do not have permission to share data.

Acknowledgements

We would like to thank the Swedish Board of Agriculture, the Swedish Environmental Protection Agency, the Geological Survey of Sweden and Lantmäteriet for supplying data. We would also like to thank Dr. Omran Alshihabi, Swedish University of Agricultural Sciences, for help in analysing soil samples using the PXRF device, and Lantmännen for providing Cu analysis data for Bjertorp farm.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.geodrs.2022.e00562>.

References

- Adler, K., Piikki, K., Söderström, M., Eriksson, J., Alshihabi, O., 2020. Predictions of cu, Zn, and cd concentrations in soil using portable X-ray fluorescence measurements. Sensors 20 (2), 474. <https://doi.org/10.3390/s20020474>.
- Ågren, A.M., Larson, J., Shehkar, S.P., Laudon, H., Lidberg, W., 2021. Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape. Geoderma 404, 115280. <https://doi.org/10.1016/j.geoderma.2021.115280>.
- Albanese, S., Sadeghi, M., Lima, A., Cicchella, D., Dinelli, E., Valera, P., The GEMAS Project Team, 2015. GEMAS: Cobalt, Cr, Cu and Ni distribution in agricultural and grazing land soil of Europe. J. Geochem. Explorat. 154, 81–93. <https://doi.org/10.1016/j.gexplo.2015.01.004>.
- Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A.C., Morgan, C. L.S., Roudier, P., Poggio, L., Mulder, V.L., 2020. Impressions of digital soil maps: the good, the not so good, and making them even better. Geoderma Reg 20, e00255. <https://doi.org/10.1016/j.geodrs.2020.e00255>.
- Ballabio, C., Panagos, P., Lugato, E., Huang, J.-H., Orgiazzi, A., Jones, A., Fernández-Ugalde, O., Borrelli, P., Montanarella, L., 2018. Copper distribution in European topsoils: an assessment based on LUCAS soil survey. Sci. Total Environ. 636, 282–298. <https://doi.org/10.1016/j.scitotenv.2018.04.268>.
- Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Chen, Y., Liu, Y., Liu, Y., Lin, A., Kong, X., Liu, D., Li, X., Zhang, Y., Gao, Y., Wang, D., 2012. Mapping of cu and Pb contaminations in soil using Combined Geochemistry, topography, and remote sensing: a case study in the Le'an river floodplain. China Environ. Res. Public Health 9 (5), 1874–1886. <https://doi.org/10.3390/ijerph9051874>.
- Cook, C.M., Vardaka, E., Lanaras, T., 1997. Concentrations of cu, growth, and chlorophyll content of field-cultivated wheat growing in naturally enriched cu soil. Bull. Environ. Contam. Toxicol. 58, 248–253. <https://doi.org/10.1007/s001289900327>.
- Duval, J.S., Carson, J.M., Holman, P.B., Darnley, A.G., 2005. Terrestrial radioactivity and gamma-ray exposure in the United States and Canada. USGS Open-File Report 2005-1413. <http://pubs.usgs.gov/of/2005/1413/> (accessed 16 Feb. 2022).
- Eriksson, J., 2021. Tillståndet i svensk åkermark och gröda – Data från 2011–2017 (In English: Current status of Swedish arable soils and cereal crops. Data from the period 2011–2017). In: Sveriges Lantbruksuniversitet. https://pub.epsilon.slu.se/23486/1/eriksson_j_210514.pdf (accessed 09 December 2021).
- Eriksson, J., Dahlin, S.A., Sohlenius, G., Söderström, M., Öborn, I., 2017. Spatial patterns of essential trace element concentrations in Swedish soil and crops. Geoderma Reg 10, 163–174. <https://doi.org/10.1016/j.geodrs.2017.07.001>.
- Fredén, C., 1994. Geology. SNA Publishing, Stockholm, Stockholm, National Atlas of Sweden.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>.
- Friedman, J.H., 2002. Stochastic gradient boosting. Computat. Statist. Data Anal. 38 (4), 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2).
- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63, 3–42. <https://doi.org/10.1007/s10994-006-6226-1>.
- Goff, K., Schaetzl, R.J., Chakrabarty, S., Weindorf, D.C., Kasmerchak, C., Bettis, E.A., 2020. Impact of sample preparation methods for characterizing the geochemistry of soils and sediments by portable X-ray fluorescence. Soil Sci. Soc. Am. J. 84 (1), 131–143. <https://doi.org/10.1002/saj2.20004>.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning. Springer-Verlag.
- Hengl, T., Leenars, J.G.B., Shepherd, K.D., Walsh, M.G., Heuvelink, G.B.M., Mamo, T., Tilahun, H., Berkhouit, E., Cooper, M., Fugraus, E., Wheeler, I., Kwabena, N.A., 2017. Soil nutrient maps of sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutr. Cycl. Agroecosyst. 109, 77–102. <https://doi.org/10.1007/s10705-017-9870-x>.
- Hu, W., Huang, B., Weindorf, D.C., Chen, Y., 2014. Metals analysis of agricultural soils via portable X-ray fluorescence spectrometry. Bull. Environ. Contam. Toxicol. 92, 420–426. <https://doi.org/10.1007/s00128-014-1236-3>.
- Hunter, J.D., 2007. Matplotlib: a 2D graphics environment. Comput. Sci. Eng. 9 (3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
- International Atomic Energy Agency, 2003. Guidelines for radioelement mapping using gamma ray spectrometry data. (Vienna, Austria). In: IAEA-TECDOC-1363. https://www-pub.iaea.org/mtdc/publications/pdf/te_1363_web.pdf (accessed 08 March 2022).
- Janssen, P.H.M., Heuberger, P.S.C., 1995. Calibration of process-oriented models. Ecol. Model. 83 (1–2), 55–66. [https://doi.org/10.1016/0304-3800\(95\)00084-9](https://doi.org/10.1016/0304-3800(95)00084-9).
- Lamicchanne, S., Kumar, L., Wilson, B., 2019. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: a review. Geoderma 352, 395–413. <https://doi.org/10.1016/j.geoderma.2019.05.031>.
- McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4).
- Michaud, A.M., Chappellaz, C., Hinsinger, P., 2008. Copper phytotoxicity affects root elongation and iron nutrition in durum wheat (*Triticum turgidum durum* L.). Plant Soil 310, 151–165. <https://doi.org/10.1007/s11104-008-9642-0>.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – a discussion of principles. J. Hydrol. 10 (3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Oorts, K., 2013. Copper. In: Alloway, B. (Ed.), Heavy Metals in Soils: Trace Metals and Metalloids in Soils and their Bioavailability. Springer, Dordrecht, pp. 367–394. https://doi.org/10.1007/978-94-007-4470-7_13.
- Panagos, P., Ballabio, C., Lugato, E., Jones, A., Borrelli, P., Scarpa, S., Orgiazzi, A., Montanarella, L., 2018. Potential sources of anthropogenic copper inputs to European agricultural soils. Sustainability 10 (7), 2380. <https://doi.org/10.3390/su10072380>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12 (85), 2825–2830.
- Piikki, K., Söderström, M., 2019. Digital soil mapping of arable land in Sweden – validation of performance at multiple scales. Geoderma 352, 342–350. <https://doi.org/10.1016/j.geoderma.2017.10.049>.
- Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Riberio, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. <https://doi.org/10.5194/soil-7-217-2021>.
- Richer-de-Forges, A.C., Arrouays, D., Bardy, M., Bispo, A., Lagacherie, P., Laroche, B., Lemercier, B., Sauter, J., Voltz, M., 2019. Mapping of soil and land-related environmental attributes in France: analysis of end-Users' needs. Sustainability 11 (10), 2940. <https://doi.org/10.3390/su11102940>.
- Scikit-learn, 2021. Ensemble Methods. <https://scikit-learn.org/0.24/modules/ensemble.html> (Accessed 10 December 2021).
- Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: a review. Prog. Phys. Geogr.: Earth Environ. 27 (2), 171–197. <https://doi.org/10.1191/030913303pp366ra>.
- SGU, 2020. Bedrock map of Sweden. <https://apps.sgu.se/kartvisare/kartvisare-berggrund-1-miljon.html> (accessed 09 December 2021).
- Shresta, D.L., Solomatine, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. Neural Netw. 19 (2), 225–235. <https://doi.org/10.1016/j.neunet.2006.01.012>.
- Shresta, G., Calvelo-Pereira, R., Roudier, P., Martin, A.P., Turnbull, R.E., Kerezuri, G., Jayakumar, P., Anderson, C.W.N., 2022. Quantification of multiple soil trace elements by combining portable X-ray fluorescence and reflectance spectroscopy. Geoderma 409, 115649. <https://doi.org/10.1016/j.geoderma.2021.115649>.
- SMHI, 2015. Klimatscenarier för Sverige – Bearbetning av RCP-Scenarier För Meteorologiska Och Hydrologiska Effektsstudier (in English: Climate scenarios for Sweden – Processing of RCP-Scenarios For Meteorological and Hydrological Effect Studies). http://www.smhi.se/polopoly_fs/1.165049/Klimatologi_15%20Klimatscenarier%20f%C3%B6r%20Sverige%20-%20Bearbetning%20av%20RCP-scenarier

- ier%20%C3%B6r%20meteorologiska%20och%20hydrologiska%20effektstudier.pdf (accessed 09 December 2021).
- Soil Atlas of Europe**, 2005. European Soil Bureau Network. European Commission. Office for Official Publications of the European Communities, L-2995, Luxembourg, p. 128.
- Stephens, M.B., Wahlgren, C.-H., Weihs, P., 1997. Sweden. In: Moores, E.M., Fairbridge, R.W. (Eds.), *Encyclopedia of European and Asian Regional Geology*. Chapman & Hall, London, pp. 690–704.
- Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measurers: illustrations, sources and a solution. *BMC Bioinformat.* 8, 25. <https://doi.org/10.1186/1471-2105-8-25>.
- Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. *BMC Bioinformat.* 9, 307. <https://doi.org/10.1186/1471-2105-9-307>.
- Swedish board of Agriculture, 2015. Nationell jordartskartering, Matjordens egenskaper i åkermarken (In English: National soil mapping, the properties of agricultural topsoil) (In Swedish). https://www2.jordbruksverket.se/download/18.4288f19214fb7ec78849af18/1441973777932/ra15_19.pdf. (accessed 08 March 2022).
- Swedish board of Agriculture, 2020. Rekommendationer för Gödsling Och Kalkning 2021 (In English: Recommendations for Fertilizing and Liming 2021) (In Swedish). https://www2.jordbruksverket.se/download/18 dc97d8e176cea4b0ec29b80/1609846154443/jo20_12.pdf (accessed 08 March 2022).
- Swedish Institute for Standards, 2017. Soil Analysis – Determination of Trace Elements in Soil by Extraction with Nitric Acid (In Swedish). <https://www.sis.se/en/produkter/evironment-health-protection-safety/soil-quality-pedology/chemical-characteristics-of-soils/ss283112017/> (accessed 08 March 2022).
- United States Environmental Protection Agency, 2007. Method 6200—Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment. Epa.gov/sites/production/files/2015-12/documents/6200.pdf (accessed 08 March 2022).
- Vavoulidou, E., Avramides, E.J., Papadopoulos, P., Dimirkou, A., Charoulis, A., Konstantinidou-Doltsinis, S., 2011. Copper content in agricultural soil related to cropping Systems in Different Regions of Greece. *Commun. Soil Sci. Plant Anal.* 36 (4–6), 759–773. <https://doi.org/10.1081/CSS-200043367>.
- Wadoux, A.M.J.-C., Minasny, B., McBratney, A.B., 2020. 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., Román-Dobarco, M., McBratney, A.B., 2021. Perspectives on data-driven soil research. *Eur. J. Soil Sci.* 72 (4), 1675–1689. <https://doi.org/10.1111/ejss.13071>.
- Wei, P., Lu, Z., Song, J., 2015. Variable importance analysis: a comprehensive review. *Reliabil. Eng. Sys. Saf.* 142, 399–432. <https://doi.org/10.1016/j.ress.2015.05.018>.
- Weindorf, D.C., Chakraborty, S., 2020. Portable X-ray fluorescence spectrometry analysis of soils. *Soil Sci. Soc. Am. J.* 84 (5), 1384–1392. <https://doi.org/10.1002/saj2.20151>.