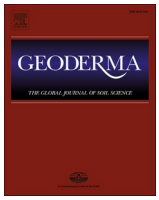
Geoderma 450 (2024) 117017

Contents lists available at ScienceDirect

Geoderma

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



Which and how many soil sensors are ideal to predict key soil properties: A case study with seven sensors

J. Schmidinger [a,b,*], V. Barkov [a,b], H. Tavakoli [b], J. Correa [b], M. Ostermann [c], M. Atzmueller [a,d,e], R. Gebbers [b,f], S. Vogel [b]

[a] Osnabrück University, Joint Lab Artificial Intelligence and Data Science, Osnabrück, Germany
[b] Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Department of Agromechatronics, Potsdam, Germany
[c] Federal Institute for Materials Research and Testing (BAM), Process Analytical Technology, Berlin, Germany
[d] German Research Center for Artificial Intelligence (DFKI), Research Department Plan-Based Robot Control, Osnabrück, Germany
[e] Osnabrück University, Semantic Information Systems Group, Osnabrück, Germany
[f] Martin Luther University Halle-Wittenberg, Institute of Agricultural and Nutritional Sciences, Halle, Germany

ARTICLE INFO

Keywords:
Proximal soil sensing
Remote sensing
Sensor fusion
Machine learning
Precision agriculture
In-situ measurements

ABSTRACT

Soil sensing enables rapid and cost-effective soil analysis. However, a single sensor often does not generate enough information to reliably predict a wide range of soil properties. Within a case-study, our objective was to identify how many and which combinations of soil sensors prove to be suitable for high-resolution soil mapping. On a subplot of an agricultural field showing a high spatial soil variability, six in-situ proximal soil sensors (PSSs) next to remote sensing (RS) data from Sentinel-2 were evaluated based on their capabilities to predict a set of soil properties including: soil organic carbon, pH, moisture as well as plant-available phosphorus, magnesium and potassium. The set of PSSs consisted of ion-selective pH electrodes, a capacitive soil moisture sensor, an apparent soil electrical conductivity measuring system as well as passive gamma-ray-, X-ray fluorescence- and near-infrared spectroscopy. All possible combinations of sensors were exhaustively evaluated and ranked based on their prediction performances using model stacking. Over all soil properties, data fusion demonstrated a considerable increase in prediction accuracy. Five out of six soil properties were predicted with an $R^2 \geq 0.80$ with the best sensor fusion model. Nonetheless, the improvement derived from fusing an increasing number of PSSs was subject to diminishing returns. Sometimes adding more PSSs even decreased prediction performances. Gamma-ray spectroscopy and near-infrared spectroscopy demonstrated to be most effective, both as single sensors or in combination with other sensors. As a single sensor, RS outperformed three out of six PSSs. RS showed especially potential for fusion with single PSSs but was of limited benefit when multiple PSSs were fused. Model stacking proved to be more robust than using single base-models because sensor performances were less model-dependent.

1. Introduction

Precision agriculture (PA) has been identified as a promising strategy by which latest advancements in technology and data analysis can help mitigate soil depletion, while maintaining high agricultural output (Keesstra et al., 2016). The underlying concept of PA is to site-specifically adopt farming practices that match the spatial variability of soil characteristics within a field (Gebbers and Adamchuk, 2010). This requires a large amount of spatial information about the soil in order to produce reliable, high-resolution soil maps for farmers. Obtaining these data by conventional approaches, which rely on sample collection and consecutive analysis in a laboratory, is usually too expensive and time consuming. Therefore, soil sensing based data acquisition emerged as a cost-efficient alternative for high-resolution soil mapping (Viscarra Rossel and Bouma, 2016). Soil sensors usually do not measure the actual target soil properties directly but proxies related to them. Nonetheless, statistical prediction models can be developed, which describe the quantitative relationship between the proxies (i.e., predictors) and the target soil properties. The concept of spatially modelling the soil with a set of predictors is known as Digital Soil Mapping (DSM) (McBratney et al., 2003).

A large number proximal soil sensors (PSSs) have been developed,

* Corresponding author at: Joint Lab KI-DS, Hamburger Straße 24, 49084 Osnabrück, Germany.
E-mail address: Jonas.Schmidinger@uni-osnabrueck.de (J. Schmidinger).

https://doi.org/10.1016/j.geoderma.2024.117017
Received 22 May 2024; Received in revised form 27 August 2024; Accepted 28 August 2024
Available online 20 September 2024

which can be mounted on terrestrial vehicles for high-resolution soil mapping purposes (Viscarra Rossel and Bouma, 2016). PSSs measure specific soil characteristics within close proximity to the soil surface or through direct contact. Their measurements require minimal to no sample preparation, which allows in-situ measurements i.e., direct deployment and use in the field. Together with short measurement durations, mobile PSSs produce data with very high sampling densities, while simultaneously remaining cost-efficient (Adamchuk et al., 2011). As compared to PSSs, remote sensing (RS) from satellites is limited in terms of measurement principles and depth profile exploration due to measuring from large distances. E.g., RS cannot be done by mechanical, chemical or biological sensors, which require direct contact with the soil. However, RS has huge advantages regarding spatial coverage, cost for data, and field access (Grunwald et al., 2015). In particular, multispectral data from the Sentinel-2 mission are available free of charge and have a relatively high spatial resolution (Drusch et al., 2012), making it potentially attractive for PA.

There is no single soil sensor based on which all relevant soil properties can be reliably predicted (Gebbers, 2018). This is because the pedosphere is a dynamic and complex system, shaped by natural and anthropogenic processes. In addition, measurements of a sensor may be affected and confounded by various factors that are impossible to control within field conditions (Adamchuk et al., 2011). Fusing data from multiple sensors can address these issues, as different sensors ideally provide complementary information. In fact, sensor fusion has proven capable of enhancing the prediction accuracy and −consistency (Adamchuk et al., 2011; Mouazen and Shi, 2021; Shi et al., 2022; Tavakoli et al., 2022; Vogel et al., 2022). Consequently, the integration of different PSSs into a versatile and mobile multi-sensor platform has continuously sparked interest within the DSM and PA community (e.g. Adamchuk et al., 2004; Tavakoli et al., 2022).

Nonetheless, the success of sensor fusion depends on the combination of PSSs and their relation to the target soil properties. Specific PSSs may not create synergistic effects when fused or only do so for a subset of target soil properties (Ji et al., 2019; O'Rourke et al., 2016; Tavares et al., 2021; Vasques et al., 2020; Xu et al., 2019; Xue et al., 2023). Sensor fusion stumbles when the data of the combined PSSs contain mainly redundant information or when the additional PSSs provide information that is not meaningfully related to the target soil property. In such cases, sensor fusion may foster model overfitting, causing a decrease in performance. Additionally, creating unnecessary data is economical inefficient. The number of PSSs on a multi-sensor platform must be restricted, because efforts for investment, operation, maintenance, and data processing must be kept at a minimum in practical applications (Adamchuk et al., 2004; Gebbers, 2018). Therefore, the combination of PSSs must be carefully selected to increase the prediction accuracy for as many key soil properties as possible, while limiting the required efforts for a multi-sensor system.

Multiple studies have examined and explored PSS fusion by comparing a study-specific PSS combination to a single-sensor approach with no fusion (e.g., O'Rourke et al., 2016; Shi et al., 2022; Tavakoli et al., 2022; Vasques et al., 2020; Vogel et al., 2022; Xue et al., 2023). However, there is a lack of studies which compared and ranked the performance of different combinations of PSSs with at least two or more PSSs. Exceptions are Chen et al. (2021), Ji et al. (2019), Xu et al. (2019) and Tavares et al. (2021), which exhaustively tested the model performance of all possible PSS combinations, given their study-specific set of PSSs. Yet, in these studies, a maximum of three to four PSSs were used, which limited the total number of combinations to be examined. A comprehensive evaluation using a wider range of state-of-the-art PSSs within a single study is lacking. The outcome of such a comparative study may provide useful insight for the future development of new multi-sensor platforms.

Data from remote sensing (RS), such as multispectral data of Sentinel-2, is often accessible without acquisition costs for the user. Since RS collects data from huge spatial distances and gather information without direct contact to the soil, it is not surprising that models fitted on PSS data have shown better prediction accuracies compared to models based on RS data for predictions of topsoil organic carbon (Bao et al., 2023; Biney et al., 2021). Nonetheless, fusing PSSs with Sentinel-2 multispectral data may improve model performances, as demonstrated for various soil properties (Bao et al., 2023; Wang et al., 2022; Wang et al., 2024). However, it needs to be further explored how Sentinel-2 data interacts with different types of PSSs.

Our objective was to identify promising sensor combinations for the prediction of multiple key soil properties on a field-scale. To achieve this, we exhaustively tested all possible PSS combinations given a dataset with six PSSs. However, for simplicity reasons, in the following, we only focus on combinations of up to five PSSs. Thereby, we evaluated how the number of PSSs affects the prediction accuracy and explored the capabilities of the different PSS combinations, both with and without consideration of RS data. To our knowledge this is the largest set of distinctive soil sensors that have been tested within a single field-scale case-study.

2. Methodology

2.1. Study area

The study was conducted on a 2.5 ha representative subplot within a larger agricultural field located in eastern Brandenburg, Germany (Fig. 1). The study area mostly consists of sandy soil, apart from increased clay contents in the middle of the field (Schmidinger et al., 2024), which are presumably deposits from a former stream. This small-scale spatial heterogeneity is characteristic for the Northeast German Plain, which was strongly influenced by peri- and interglaciation. Therefore different soil formation processes shaped the area (Vogel et al., 2022). Three parallel transects of approximately 800 to 1,000 m length were established along the main gradient of soil variability for reference sampling. The elevation of the study area ranged from 50 to 80 m a.s.l. The climate is characterized by an average of 550 mm annual rainfall and a mean annual temperature of 9 °C.

2.2. Study design

Six top-layer soil properties were examined in this study: soil organic carbon (SOC), pH, moisture content, plant-available phosphorus (P), magnesium (Mg) and potassium (K) (**see Section 2.3**). Six state-of-the-art PSSs were used for the predictions: ion-selective pH electrodes (pH-ISE), a capacitive soil moisture sensor (CSMoist), an apparent soil electrical conductivity measuring system (ECa) as well as passive gamma-ray- (γ), X-ray fluorescence- (XRF) and near-infrared (NIR) spectroscopy (**see Section 2.4.1**). Bare soil multispectral RS data were obtained from Sentinel-2 (**see Section 2.4.2**). From a single PSS up to five PSSs, all possible combinations of PSSs were generated. For each combination, we created two versions for the analysis, one including and one excluding RS data. This led to a total set of $C=124$ distinct combinations. We exhaustively tested and compared the model performances of all C combinations.

Stacking was used for the modelling, in which Cubist, extreme gradient tree boosting (XGBoost), random forest (RF), support vector regression (SVR) and ridge regression (RR) were embedded as base-models and multiple linear regression (MLR) used as meta model (**see Section 2.5**). The model performance was evaluated through K-fold cross validation (CV), where $K=10$. The prediction error was quantified by the coefficient of determination (R^2) and root mean squared error (RMSE) (**see Section 2.6**). Grid search was run within a nested L-fold CV, where $L=5$. In the grid search, model specific hyperparameters and the number of principal components for NIR data were optimized for the model training in the outer-loop. The systematic workflow for this study is summarized in Fig. 2.

Fig. 1. Study area and sampling locations (n = 159) based on regular grid sampling in eastern Brandenburg, Germany. The area in which sensor measurements were taken is highlighted within the larger test-field on a ©2018 Google satellite base map.

Generate C combinations, where C is the number of all possible combinations of PSSs both with and without RS data;
for $c = 1$ **to** C **do**
  Split dataset fully randomly into $K = 10$ subsets of approximately equal size to prepare outer 10-fold CV;
  for $k = 1$ **to** K **do**
    (1) Hold out k-th subset as test set and compile remaining $K - 1$ subsets as outer training set
    (2) Split outer training set fully randomly into $L = 5$ subsets of approximately equal size to prepare nested 5-fold CV;
    for $l = 1$ **to** L **do**
      (1) Hold out l-th subset as validation set and compile remaining $L - 1$ subsets as inner training set
      (2) Grid search: train base-models (i.e. Cubist, XGBoost, RF, SVR and RR) with inner training set using different distinct hyperparameter combinations. If c contains NIR, different numbers of principal components are included in the grid
      (3) Predict values for l-th validation set with every trained base-model;
    end
    (1) Compute inner loop RMSE for every base-model over all L validation sets
    (2) Identify and select hyperparameter combinations of the base-models with the lowest inner loop RMSE
    (3) Train meta model with MLR, in which the predictions for the L validation sets from the selected base-models are used as features
    (4) Train new base-models with the outer training set using the selected hyperparameter combination
    (5) Predict values for k-th test set with these trained base-models
    (6) Feed these predictions to the meta model to make final predictions for k-th test set;
  end
  Compute R^2 and outer loop RMSE for the meta model predictions over all K test sets;
end
Compile R^2 and outer loop RMSE values from all C combinations as final results

Fig. 2. Systematic workflow used for the modelling in this study.

2.3. Target soil properties

The soil samples were collected along three parallel transects, running from the south-west to the north-east. These transects were 12 m apart, closely following the typical path taken by field machinery. Along each transect, we gathered 53 samples at 15-meter intervals (i.e., regular grid sampling), resulting in a total of 159 sampling points. The soil samples were collected from a depth of 0 to 30 cm (Fig. 1) where each soil sample was composed of three subsamples taken in a radius of 0.5 m around the center. Soil samples analyzed for soil organic carbon (SOC) were collected in May 2017. Samples analyzed for soil pH, moisture content as well as concentrations of plant-available P, K and Mg were taken in August 2021. SOC was analyzed through dry combustion of air-dried soil (DIN ISO 10694). In a suspension with 0.01 M $CaCl_2$ and air-dried soil, pH was determined with a glass electrode after an equilibration time of 60 min (DIN ISO 10390). Soil moisture was measured by the weight difference before and after drying fresh soil at 70 °C. P and K were extracted in a calcium acetate-lactate solution and measured photometrically (VDLUFA I A 6.2.1.1). Mg was extracted with calcium chloride solution and measured through atomic absorption spectrophotometry at 285.2 nm (VDLUFA I A 6.2.4.1). Appendix A includes information about the intercorrelation (Fig. 1A), values at the sampling locations (Fig. 2A) and descriptive statistics (Table 1A) of the six target soil properties.

2.4. Sensor data

2.4.1. Proximal soil sensing

Two mobile multi-sensor platforms were used for in-situ measurements in this study: Veris MSP3 (VERIS Technologies, Salinas, USA) and RapidMapper (Tavakoli et al., 2022). Their measurements were taken around the three transects. There was a disparity between the reference soil samples and sensor measurements in terms of number and location. Therefore, the sensor data were interpolated to the 159 sampling locations of the target soil properties (Fig. 1). For the spatial interpolation, we used ordinary kriging from the package *automap* (Hiemstra et al., 2008) in R (R Core Team, 2023). Soil samples for XRF sensor measurements were directly taken at the sampling locations and analyzed in

the laboratory. Soil samples for XRF were not processed before lab measurement i.e., contained fresh field moisture content to simulate in-situ field conditions. Further information about the data collection can be found in Table 1 and about the measurement locations and inter-correlation in Figs. 3A and 4A, Appendix A.

The RapidMapper is a multi-sensor platform mounted with different PSSs. Three PSSs of the platform included γ, ECa and NIR. The passive γ spectrometer (model MS-2000-CsI-MTS, Medusa Radiometrics BV, Groningen, Netherlands) with its accompanying software was used to detect the naturally occurring radionuclides: Potassium-40 (^{40}K), Uranium-238 (^{238}U) and Thorium-232 (^{232}Th), the human-made radionuclide Caesium-137 (^{137}Cs), as well as the total gamma counts. An NIR spectrometer (C11118GA, Hamamatsu Photonics K.K., Shizuoka Prefecture, Japan) with an average resolution of 15 nm measured subsurface soil diffuse reflectance at a depth of 10 to 15 cm through a furrow opening 'shoe'. The spectral data was harmonized to 1 nm intervals through interpolation and the noisy edges were discarded, which led to a spectrum ranging from 1,000 to 2,400 nm. We used principal component analysis (PCA) for dimensionality reduction instead of using all resulting 1,400 bands. The number of principal components used for the modelling ranged from 5 to 20 and was determined through grid search. ECa of the RapidMapper is based on the galvanic contact resistivity technique developed by the Geophilus company (Caputh, Germany). The effective depth of ECa investigation is 0 to 50 cm.

The Veris MSP3 platform also consists of multiple integrated PSSs but for this study, only pH-ISE was examined. pH-ISE is based on the 'pH manager', which measures the pH value on-the-go using two ion-selective antimony pH electrodes on naturally moist soil samples that are taken while driving across the field. If there were deviations (≥0.4) between the two measurements at a sampling location, the sampling location was discarded from the data set. The Veris MSP3 was additionally equipped with CSMoist originally from the Veris iScan to measure the volumetric soil moisture.

A handheld XRF sensor (model Vanta M Series, Olympus, Waltham, USA) has been used for XRF measurements. During XRF measurements, soil samples are exposed to high-energy X-rays from a 50 kV Rhodium anode. The consequential excitement of electrons and the succeeding relaxation emits secondary (fluorescence) x-rays, which can be measured and associated with specific elements. A sensor-specific calibration mode called GeoChem (2-beam) has been used to obtain an initial estimation of element contents. For the estimation, GeoChem uses theory-based parameters and corrects for interelement effects. Ultimately, element contents of Al, Si, P, Ca, Mn, Fe, Zn, Pb and Mg have been used for this study. Other elements fell below the detection threshold.

This set of PSSs includes specialized sensors such as CSMoist and pH-ISE that are known for their direct relationship to one specific soil property. Yet, it also includes sensors that could be more multifunctional, as they relate to various soil properties, like NIR, XRF, ECa and γ (Gebbers, 2018).

Table 1
Overview of the measurements with different sensors in this study.

Sensor	Number of predictors	Pre-processing method	Data collection	Date of data collection
ECa	1	None	In-situ (RapidMapper)	August 2021
γ	4	Moving average	In-situ (RapidMapper)	June 2020
XRF	9	GeoChem (2-beam)	Laboratory with fresh soil samples	August 2021
NIR	5 – 20*	Omitting noisy edges of the spectrum Interpolation to 1 nm PCA	In-situ (RapidMapper)	April 2020
CSMoist	1	None	In-situ (Veris MSP3/iScan)	August 2021
pH-ISE	1	Double measurement error detection	In-situ (Veris MSP3)	September 2017
Sentinel-2	10	Level-2A processing e.g. atmospheric and terrain corrections (Louis et al., 2016)	Spaceborne	September 2021

* Range of searched principal components in the grid-search.

2.4.2. Remote sensing

Sentinel-2 is a space mission consisting of two twin satellites launched by the European Space Agency (Drusch et al., 2012). Both satellites are equipped with an optical sensor for acquisition of multi-spectral data at 13 spectral bands. Due to the small size of the study area (**see Section 2.1**), we excluded three bands with the lowest resolution of 60 m. The ten remaining spectral bands had a spatial resolution of 10 to 20 m. Their spectral measurements were scattered within the 490 to 2,190 nm range. A cloudless bare-soil image of the study area from the 10th of September 2021 during the vegetation free period was used. The spectral values were extracted from the raster map at the corresponding sampling locations. We retrieved the processed data from Copernicus Data Space Ecosystem as level-2A product (Louis et al., 2016).

2.5. Predictive modelling

By implementing model stacking, we aimed to achieve less model-dependent results and derive more generalizable conclusions about the effectiveness of the different PSS combinations. Additionally, model stacking has shown capable of further increasing the prediction accuracy (Leblanc and Tibshirani, 1996) because a set of learning algorithms can capture different patterns of the data. Five base-models that are commonly used in DSM (Wadoux et al., 2020) were embedded in the stacking, namely Cubist, XGBoost, RF, SVR and RR. After determining optimal hyperparameters for the base-models through grid search in the nested CV, the predictions of the inner validation folds from the base-model were aggregated. Lastly, the meta model based on MLR was fitted on these aggregated predictions. Additional information about the optimized hyperparameters, the selection of the numbers of principal components for NIR data as well as the associated R-packages for the implementation of the base-models can be found in Table 1B, Appendix B and in the published code within a GitHub repository (https://github.com/JonasSchmidinger/soil-sensor-fusion/tree/main).

XGBoost (Chen and Guestrin, 2016) is an adapted implementation of gradient boosted decision trees (GBDTs). Boosting is an ensemble learning method based on a principle of combining a set of weak learners into a single stronger predictive model. A decision tree is a hierarchical model that predicts the target variable by learning decision rules from the predictors. It can be understood as a path of conditional if-then-else statements. Gradient boosting is a generalization of boosting where the process of generating weak learners is sequential and formalized as a gradient descent over a differentiable loss function. Hence, GBDTs is an ensemble of decision trees based on gradient boosting. Lastly, XGBoost introduces multiple regularization parameters which further control the model complexity which distinguishes it from standard GBDTs.

Cubist merges the concepts of decision trees and linear regression. It originates from the M5 model tree introduced by Quinlan (1992) and was formalized by Kuhn and Johnson (2013). A single decision tree is grown in an initial step. In each node of the tree up to the terminal nodes, individual linear regression models are fitted. The different paths along the tree are then summarized to an initial set of so-called rules. Each rule consists of a 'smoothed' regression model which is a function of the regression models along the associated path in the tree. Cubist further allows implementation of a boosting-like ensemble referred to as committee and a k-nearest-neighbors-like adoption for the regression

models in the final rules.

RF (Breiman, 2001) is a decision tree ensemble learning algorithm based on bootstrap aggregating, referred to as bagging. Bagging is the process of generating multiple equally weighted learners in which each learner is trained on distinct and independent training data sets as well as a random subset of predictors. These distinct sets are produced through bootstrapping. This process generates an ensemble of diverse decision trees yielding a more generalizable model. For regression tasks, a final prediction of RF is obtained by averaging the predictions from all individual decision trees within the ensemble.

SVR (Drucker et al., 1996) is a variant of Support Vector Machines developed for regression tasks. SVR aims to find a hyperplane in the multidimensional covariate space, by minimizing a loss function. The loss function in SVR penalizes points that fall outside of a defined tube around the hyperplane. Inside this tube, errors are accepted and not penalized. This seeks to find a balance between model complexity and its tolerance for deviations, aiming for a compromise between accuracy and model simplicity to enhance generalization. Non-linearity in SVR is introduced through the use of kernel functions. Kernel functions work by projecting the input features into a higher-dimensional space, enabling linear separation in cases where the data is not linearly separable.

RR (Hoerl and Kennard, 1970), is an extension of linear regression that introduces an L2 regularization penalty on the coefficients of a linear regression model. This approach helps to address the issue of overfitting in the presence of multicollinearity. The performance of RR is a natural baseline in comparison to the other more sophisticated algorithms. More simple linear models may still prove viable, especially when the training sample size is not too large (Schmidinger et al., 2024).

2.6. Validation metrics

For model validation we used the R^2 and RMSE. RMSE is an absolute validation metric for quantifying the prediction error. It shows how much the predicted values deviate from the observed values:

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(y_i - \hat{y}_i)^2}, \quad (1)$$

where n is the number of testing samples. $\hat{y}_i$ $(i = 1, \cdots, n)$ and y_i are the predicted and the observed values for the target variable y of soil sample i, respectively.

R^2 is a relative error measure, which is equal to one in case of perfect predictions. It becomes negative if predictions are worse than the arithmetic mean of the test data $\bar{y}$:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \quad (2)$$

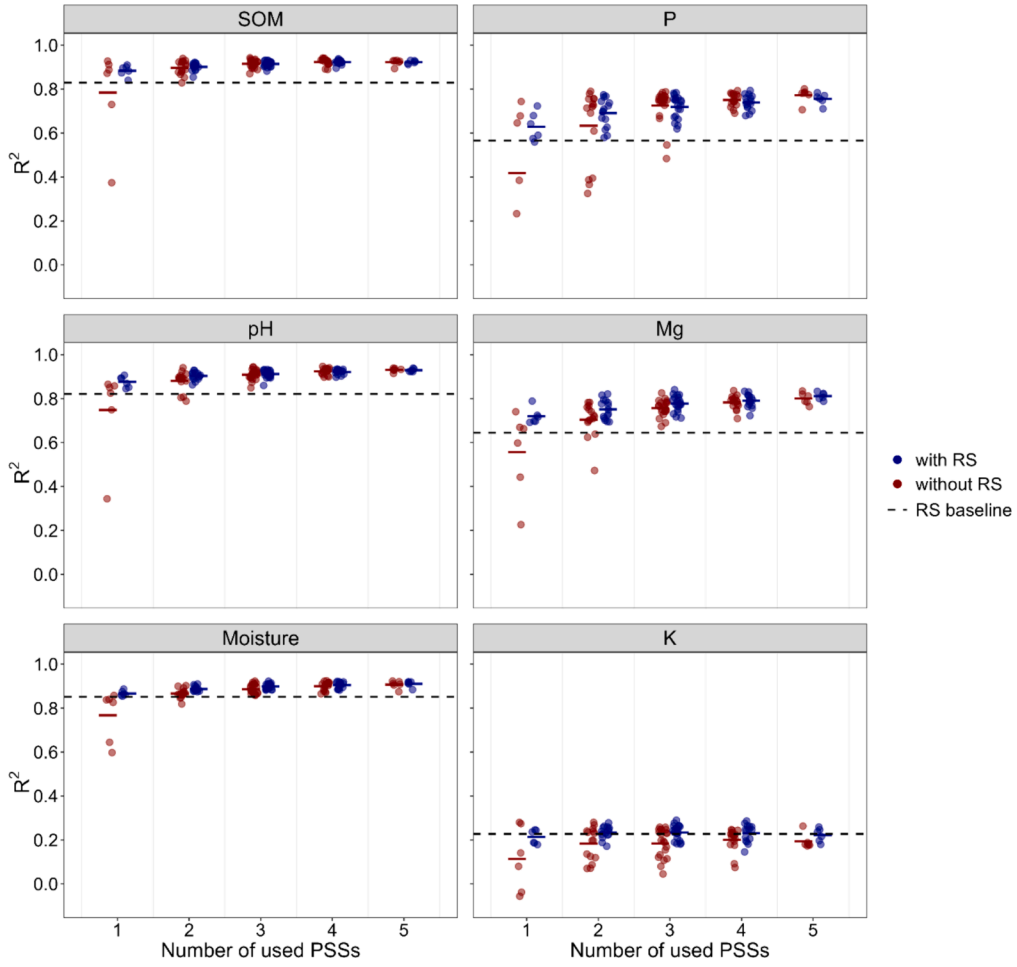


Fig. 3. R^2 values for all combinations of proximal soil sensing (PSS) techniques, evaluated across six target soil properties. Combinations that incorporate remote sensing (RS) data are shown in blue, while those that rely solely on PSS data are shown in red. The RS-baseline represents the performance using only RS data. Interactive plots are provided in the Supplementary data and at https://plotly.com/~Schmidinger_plotly/38/dashboard/. For graphical consistency, the CSMoist prediction for P without RS has been omitted from this figure due to its outlying R^2 value of −0.18.

3. Results

3.1. Sensor fusion

Fig. 3 presents the performance of all PSS combinations for the different target soil properties. The results indicate a varying degree of success with sensor fusion. P, moisture, pH and Mg predictions benefited considerably from synergies through sensor fusion. Their maximum R^2 increased by 0.06 (P), 0.06 (moisture), 0.08 (pH), 0.10 (Mg), respectively, when moving from a single sensor application without RS to the optimal sensor fusion application. In contrast, sensor fusion had a neglectable impact on SOC and K predictions. Their maximum R^2 increased minimally by 0.01. In many cases, adding more PSSs even decreased performances for these two target soil properties. Except for K, all soil properties achieved an $R^2 \geq 0.8$ at some point through sensor fusion.

A similar pattern can be found in the averaged performances of all soil properties (Fig. 4), where the maximum R^2 increased from by 0.07 after optimal fusion. The positive effect of sensor fusion was particularly noticeable when using up to three PSSs but the improvement diminished when going beyond this number. Thus, the overall best performances were achieved with four PSSs instead of five PSSs (Table 2).

We obtained inconclusive results from combining RS with PSSs. The effectiveness varied with the target soil property, the type of PSS and the number of PSSs used (Figs. 4 and 5). Fusion with RS generally had the strongest impact, when using single PSSs. However, this effect was both positive and negative depending on the type of PSS. For example, the mean R^2 for predictions with pH-ISE improved drastically over all soil properties by 0.27, from 0.43 to 0.70 (Table 2 and Fig. 4), whereas it slightly decreased for ECa. Additionally, results varied across the six target soil properties. For instance, including RS data for K predictions with γ decreased the R^2 value by 0.04. Conversely, it increased by 0.06 for Mg predictions.

3.2. PSS performance ranking

While using one to three PSSs, the model performance exhibited significant dependence on the specific PSS combination employed (Figs. 3 and 4). Yet, as more PSSs were incorporated, the variation in model performance decreased. For instance, with single PSSs, the range of mean R^2 values spanned from 0.26 to 0.72. In contrast, when combining three and five PSS techniques, the ranges narrowed down to 0.65 to 0.76 and 0.74 to 0.77, respectively (Fig. 4). Due to this trend, we will focus on identifying the most impactful PSS combinations for using one to three sensors in the following sections.

While a variety of PSS combinations achieved strong performance, specific PSS techniques outperformed others, both individually and when combined (Figs. 3 and 4). NIR and γ emerged as the most effective PSSs. Consequently, at least one of these sensors was consistently included in the top-performing PSS combinations. (Table 2). Although NIR excelled γ for predictions of SOC, Mg, and P, NIR data alone offered no predictive power for K content (Fig. 5). Predicting K proved generally

Table 2
Top three PSS combinations ranked by mean R^2 for each specific number of fused PSSs. The mean R^2 represents the average of all six target soil properties.

Number of PSSs	RS	PSS	Mean R^2	Rank
1	with RS	NIR	0.722	1
	with RS	γ	0.721	2
	with RS	pH-ISE	0.703	3
2	with RS	γ, NIR	0.752	1
	without RS	γ, ECa	0.748	2
	with RS	pH-ISE, NIR	0.747	3
3	with RS	XRF, γ, NIR	0.763	1
	with RS	pH-ISE, γ, NIR	0.762	2
	with RS	XRF, γ, ECa	0.760	3
4	with RS	XRF, γ, ECa, NIR	0.770	1
	with RS	pH-ISE, γ, ECa, NIR	0.768	2
	without RS	XRF, γ, ECa, NIR	0.765	3
5	with RS	XRF, γ, CSMoist, ECa, NIR	0.765	1
	without RS	XRF, pH-ISE, γ, ECa, NIR	0.761	2
	with RS	pH-ISE, γ, CSMoist, ECa, NIR	0.760	3

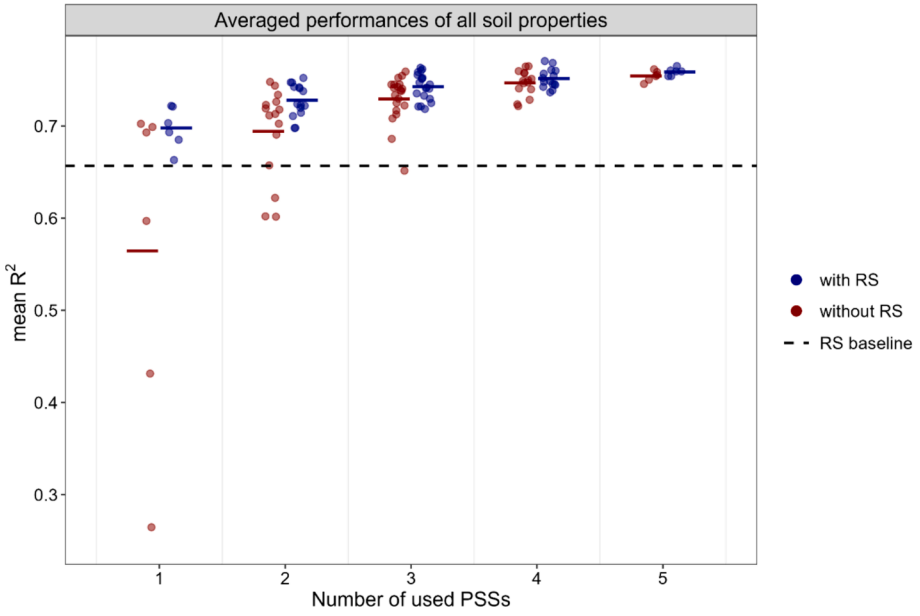


Fig. 4. R^2 values for all combinations of proximal soil sensing (PSS) techniques, averaged over all six target soil properties. Combinations that incorporate remote sensing (RS) data are shown in blue, while those that rely solely on PSS data are shown in red. The RS-baseline represents the performance using only RS data. The interactive plot is provided in the Supplementary data and at https://chart-studio.plotly.com/~Schmidinger_plotly/25.

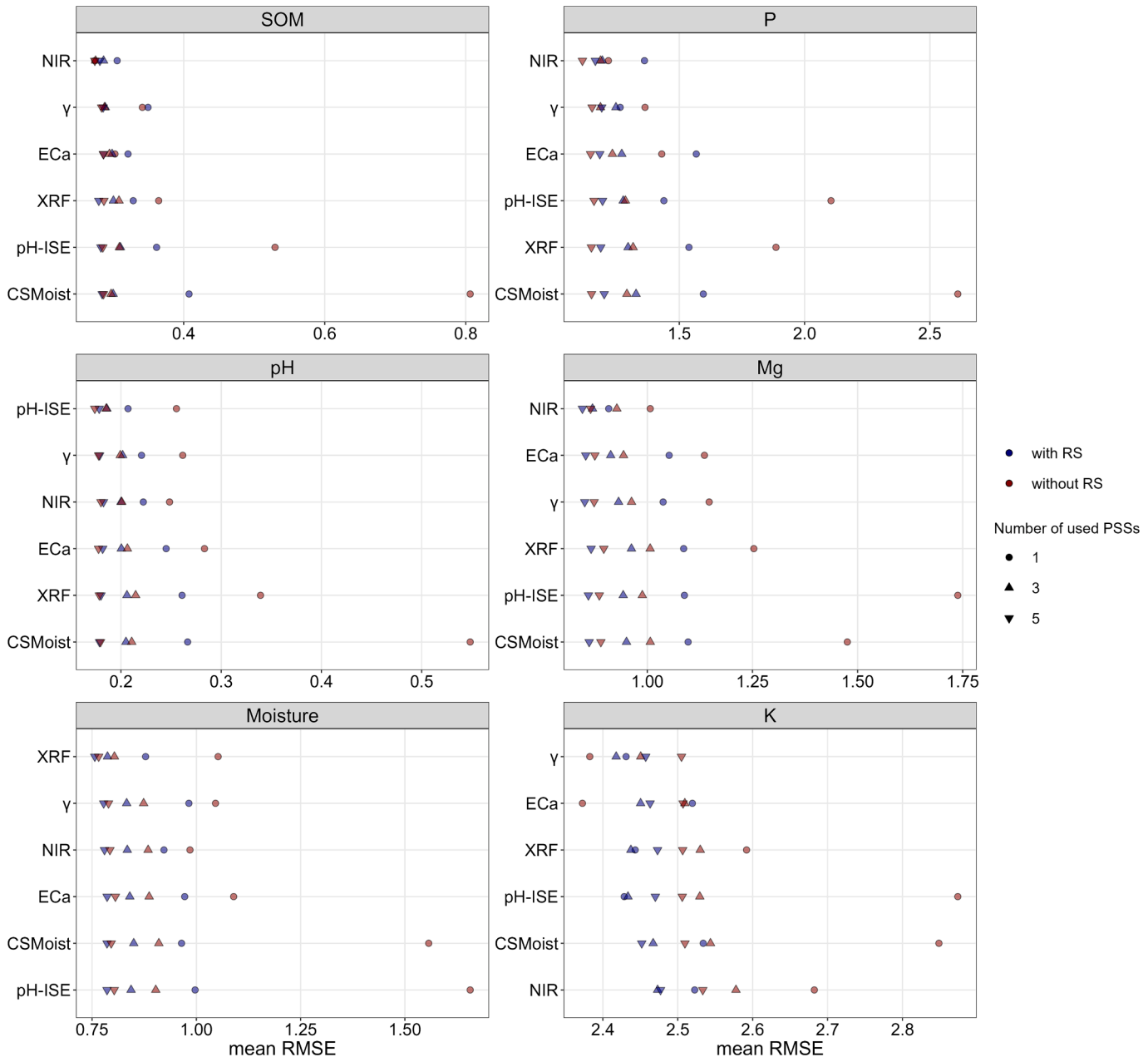


Fig. 5. Mean RMSE values for the six target soil properties across all available PSS combinations that include the specified PSS technique. PSSs are ranked from top to bottom based on their lowest mean RMSE, averaged across all combinations in which the PSS was included. Combinations incorporating remote sensing (RS) data are depicted in blue, while those using only PSS data are shown in red. Symbols indicate the number of PSS techniques used in each combination.

challenging as R^2 values fell below 0.3. However, γ and ECa were the only sensors exhibiting some slight relation to K, obtaining R^2 values of 0.27 and 0.28, respectively. ECa exhibited strong individual performances similar to those of NIR and γ when used without RS. Yet, ECa was not as successful when incorporated into sensor fusion models.

pH-ISE and CSMoist were the least successful standalone PSSs. pH-ISE only showed a strong relation to pH (Fig. 5). However, when pH-ISE was combined with RS data, its performance improved significantly, placing it among the most effective combinations (Table 2). It also demonstrated reasonable synergies with other PSSs.

XRF achieved strong performances for soil moisture but its ability to predict other soil properties was less exceptional, yielding only moderate accuracy (Fig. 5). It generally underperformed compared to NIR, γ and ECa for single-sensor predictions. Nonetheless, it was a frequent component in top-performing combinations when three or more PSS techniques were employed (Table 2), indicating a certain degree of success for XRF in the context of sensor fusion.

CSMoist exhibited the weakest performance among the PSS techniques. Despite being designed specifically for soil moisture measurement, CSMoist was even outperformed by most other PSSs and RS for moisture predictions (Fig. 3). Notably, its R^2 value for moisture predictions (0.64) fell significantly short to the RS baseline (0.85).

Models utilizing RS data alone achieved decent performance across all target soil properties, on average even surpassing XRF, pH-ISE, and CSMoist (Fig. 4). Nonetheless, the most accurate single sensor predictions for each target soil property were always achieved by proximal sensing.

3.3. Model dependency

The results presented earlier were derived from model stacking, as described in **Section 2.5**. This section provides a brief overview on how the results depend on the selected prediction models. We compare the performance of the base-models used within the stacking framework, both against each other and to the final stacked model.

Model stacking yielded superior performance compared to the individual base-models (Fig. 6). Notably, base-models that employed regression trees within their algorithms, such as Cubist, RF and XGBoost, generally achieved stronger results than SVR and RR. Nonetheless, in terms of absolute values, the mean R^2 values were relatively similar among the different prediction techniques. Only RR showed a significantly lower performance. Fig. 1B, Appendix B explores these results in more detail, highlighting potential variations depending on the specific target soil property.

Fig. 7 illustrates the dependence of optimal PSS combination on the chosen modelling approach, as different PSS combinations achieved superior performances with different models. For example, XGBoost obtained the best predictions when used with ECa and without the inclusion of RS data. However, ECa did not rank among the top three performing sensors for the stacked model (Table 2). RS proved to be only helpful for certain modelling techniques. For instance, the combination of γ and ECa produced the most accurate results for Cubist, Random Forest (RF), and XGBoost models. However, only XGBoost successfully incorporated RS data in this PSS combination. Cubist and RF relied solely on the PSSs (i.e., γ and ECa) for their best predictions.

4. Discussion

4.1. Sensor fusion

As few as three properly chosen PSSs yielded accurate predictions for multiple soil properties. Notably, combinations including pH-ISE, γ, NIR and ECa exhibited strong performances. These PSSs have a well-established history in proximal soil sensing and are already incorporated into various multi-sensor platforms (Adamchuk et al., 2011; Tavakoli et al., 2022). This finding highlights the potential of already existing multi-sensor platforms, like RapidMapper, for various agriculturally and environmentally related tasks. Adding more than three PSSs did not substantially enhance the prediction accuracy. This diminishing return suggests that there is a limit to the effectiveness of sensor fusion with common PSSs. It highlights the importance of selecting a few informative and complementary PSSs rather than simply increasing the sensor quantity.

Our models achieved significantly higher absolute performances compared to prior field-based sensor fusion studies (Ji et al., 2019; Vasques et al., 2020), being comparable to lab-based sensor fusion studies (Tavares et al., 2021; Xu et al., 2019). However, our results are consistent with the patterns observed in all of these studies, i.e., sensor fusion may not ultimately lead to significant improvements for the prediction of every individual soil property. This lack of improvement was particularly demonstrated in our study for SOC and K. Simply adding more variables is not always a winning strategy. In fact, the increased dimensionality from data fusion can even be harmful in some cases. For instance, we showed that four PSSs returned slightly better results compared to five PSSs (Table 2). This decrease in performance with an increase in dimensionality is commonly referred to as 'curse of dimensionality' or 'Hughes phenomenon' (Hughes, 1968). Nevertheless, across multiple soil properties, sensor fusion has demonstrated its ability to enhance and stabilize prediction performances. Specifically in our study, predictions of P, moisture, pH and Mg benefited from synergies.

The combination of RS with PSSs gave inconclusive results. In several cases, it improved performances considerably when using single PSSs (e.g., in combination with pH-ISE) but in some cases, it led to a slightly decreased accuracy (e.g., in combination with ECa). Nevertheless, numerous other studies have shown that fusing PSSs with RS data i.e., Sentinel-2, can be more impactful (Bao et al., 2023; Wang et al., 2022; Wang et al., 2024). We hypothesized that sensors like ECa, pH-ISE, and γ would benefit from the incorporation of multispectral RS data, whereas we expected less improvement for NIR due to potential overlap in spectral information. However, this cannot be fully supported by our results. While incorporating RS data enhanced predictions with γ and pH-ISE, it also improved predictions based on NIR. Conversely, ECa did not see any significant improvement with the addition of RS data.

4.2. PSS performance ranking

NIR and γ demonstrated remarkably strong model performances for the majority of soil properties. Strong performances with in-situ γ in comparison to other PSSs were also reported in Vasques et al. (2020) and Ji et al. (2019). Since γ is not causally related to most of the target soil properties, we suspect that soil texture, which is usually well predicted by γ (Meyer et al., 2019; Vasques et al., 2020), was controlling a certain degree of the overall spatial soil variability. We discuss the limitations of exploiting such non-causal secondary-relationships in **Section 4.4**. NIR is a well-established and versatile PSS for multivariate soil mapping (Grunwald et al., 2015). Therefore, it was unsurprising that NIR data proved valuable for various soil property predictions. Especially the strong relation of NIR to SOC is in agreement with the literature (Bai

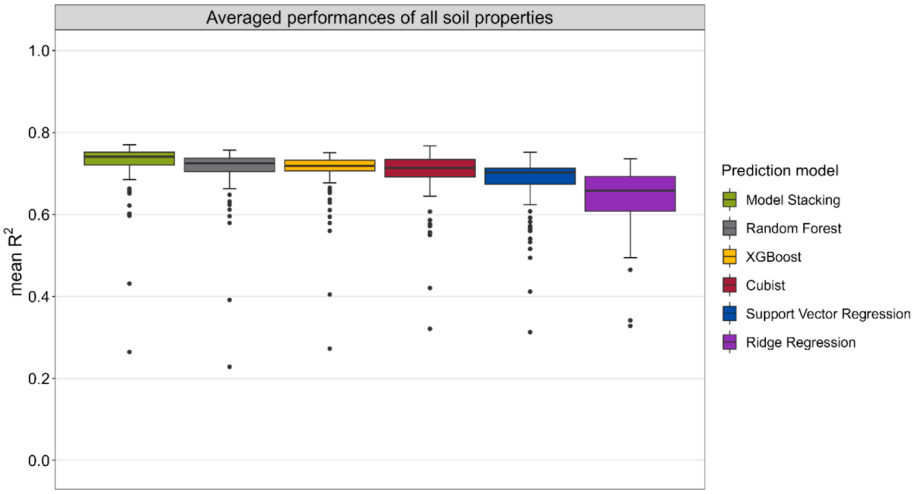


Fig. 6. The mean R^2 values, averaged over all six target soil properties, for all PSS combinations in dependence to the different base-models and model stacking.

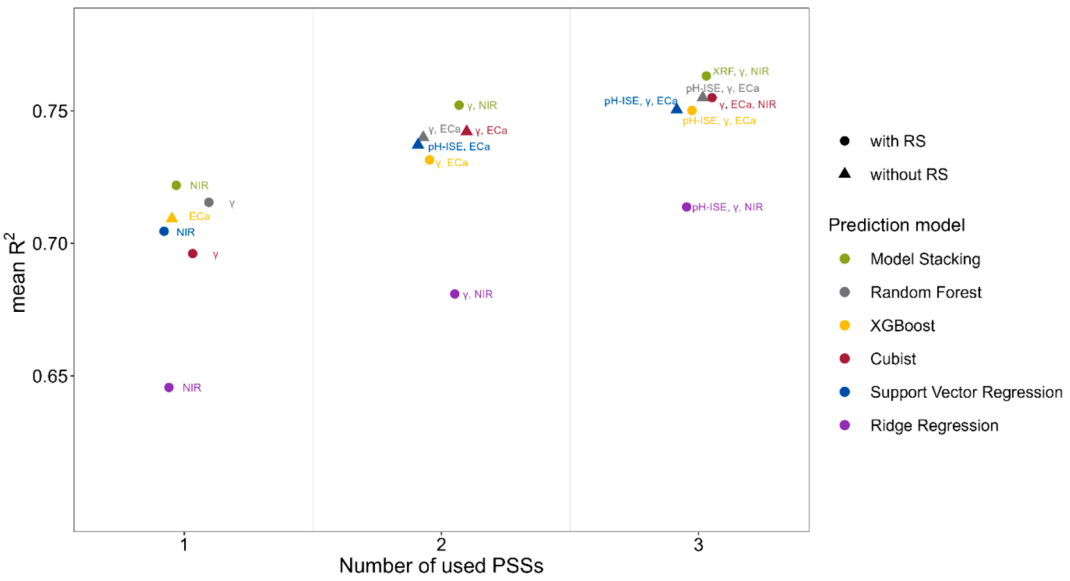


Fig. 7. PSS combinations that achieved the highest mean R^2 values across all six target soil properties. The R^2 values are averaged considering different base-models and model stacking techniques employed. For better visual clarity, only combinations utilizing up to three PSS techniques are included.

et al., 2024; Ji et al., 2019; Tavakoli et al., 2022; Tavares et al., 2021). NIR achieved poor performances for predictions of K, which aligns with previous studies (Ji et al., 2019; Wenjun et al., 2014).

ECa was ranked closely behind γ and NIR, when excluding RS. Being influenced by soil moisture, texture and ion concentration, ECa offers a versatile tool for soil analysis (Grunwald et al., 2015). Hence, ECa performed strongly for multiple soil properties. However, it did not outperform the other PSSs for soil moisture predictions.

As expected, pH-ISE was strongly related to pH but showed weak relations to the other soil properties. However, pH-ISE played a complementary role in predicting other soil properties. Fusing it with various other PSSs enhanced their performance. This highlights that pH-ISE data contributed unique and less redundant information.

Unlike the promising performances of in-situ XRF measurements reported by Vasques et al. (2020), our dataset yielded only moderate performance with XRF. Particularly, the model performances for Mg and P did not meet expectations, as XRF was outperformed by most other PSSs. XRF was the only sensor in this study that can measure elemental P and Mg directly. However, since P and Mg have low atomic numbers, they are more difficult to determine. Furthermore, the target properties were plant available P and Mg, which make only a small portion of the total content. The correlation between the extractable forms and the total content may be affected by several confounding factors (Gebbers, 2018). Surprisingly, strong predictions with XRF were achieved for soil moisture. The natural soil moisture content most likely interfered with XRF measurements, potentially explaining why predictions based on XRF yielded underwhelming results for Mg and P. Alternatively, the relationship between XRF and soil texture (Tavares et al., 2021; Vasques et al., 2020) could also have been exploited for moisture predictions. Note that we obtained much better results with XRF on dried soil samples for Mg and P in the same dataset. However, for comparability between the other in-situ soil sensors and XRF, we only presented results based on moisturized soil samples in this study.

Among the sensors evaluated, CSMoist provided the least value. As discussed by Gebbers (2018), low reliability remains a common issue for capacitive soil moisture sensors. Even models entirely based on RS data outperformed those of CSMoist for soil moisture prediction. Given that RS is freely available, a PSS or PSS combination should be capable to outperform RS for most or at least specific soil properties. We encourage future studies to also include a RS baseline to better evaluate the performance of a PSS. In this study, RS showed decent performances but the trade-off between cost-efficiency and accuracy was not regarded.

4.3. Model dependency

The best PSS or PSS combination can vary depending on the chosen prediction model. While the evaluation of model-dependent outcomes is valuable, we advocate for model stacking in order to achieve more robust and generalizable conclusions. Stacking combines predictions from multiple models, allowing to estimate the overall importance of each sensor from a broader perspective. In contrast, most studies in the literature relied on single-model approaches. Our model-dependent findings suggest that using a single model might lead to less generalizable and potentially misleading interpretations about sensor effectiveness.

4.4. Limitations

Certain sensor measurements showed strong predictive power for soil properties with which they do not have a causal relationship. E.g., in our study, NIR and RS data, while not directly related to pH, could predict pH with high accuracy. In contrast, previous studies (e.g., Ji et al., 2019; Tavares et al., 2021) have reported limited success using NIR for pH predictions. Our dataset showed a strong correlation between pH and SOC (Fig. 1A, Appendix A). This exemplifies that second-order relationships have likely been exploited to a certain degree, as SOC is directly linked to NIR and RS measurements. It is difficult to estimate how reproduceable our results are under different study settings with less intercorrelation. Despite this limitation, this case study offers a novel contribution by employing a diverse range of PSSs. This enabled a comprehensive comparison of multiple PSS technologies commonly used in PA and DSM, revealing both the potential and drawbacks of PSSs and sensor fusion in general.

The study design lacked control over factors influencing sensor performance, hindering a fully standardized comparison. Data collection across different dates (Table 1) and variations in the measuring frequency (Fig. 3A, Appendix A) introduced technically conditioned inconsistencies between the sensors. The time-gap of γ, NIR and pH-ISE compared to the other PSSs, were a result of issues during the 2021 sensing campaign, so that sensor readings from previous sensing campaigns had to be used. This time difference likely introduced a mismatch between the moisture content during sensor readings and the moisture

content of the reference samples, because the absolute soil moisture content is highly dynamic. Nonetheless, the time gap may not always be an issue as the relative spatial moisture patterns in fields are usually stable in time (Alijani et al., 2024) due to the strong relation to texture (Huang et al., 2017). Therefore, it seems that the temporal disparity did not disadvantage γ or NIR. These two sensors still demonstrated strong performances, even for soil moisture. pH-ISE had the largest time-gap compared to the other PSSs. However, the lab-measured pH values did not change considerably since 2017, as shown by a previous sampling campaign (Fig. 2B, Appendix B). For this reason, pH-ISE still provided strong predictive capabilities. Also, the spatial pattern of soil nutrient contents in 2017 was very similar compared to 2021 (Fig. 2B, Appendix B).

This study utilized a high density of training data (>50 samples per hectare) compared to what is typically available to PA providers (Schmidinger et al., 2024). Hence, the reported R^2 values for most soil properties are potentially overoptimistic in comparison to what is achievable with a more realistic number of training samples in practical agriculture. Also, the ratio of training samples to predictor variables has a strong effect on the prediction performance. Therefore, it is important to acknowledge that the validity of our PSS ranking is limited to a scenario with abundant training samples.

We recommend to further studies on the effectiveness of different sensors to consider multiple fields for more robust conclusion and an economic analysis about the cost-effectiveness as in e.g., Chatterjee et al. (2021). We only evaluated Sentinel-2 data from RS but a multi-RS approach (Gasmi et al., 2022) next to PSSs should be further explored.

5. Conclusion

In this study, we investigated the potential of fusing different proximal soil sensors (PSSs) and remote sensing (RS) data. Six different types of PSSs such as ion-selective pH electrodes (pH-ISE), a capacitive soil moisture sensor (CSMoist), an apparent soil electrical conductivity measuring system (ECa) as well as passive gamma-ray- (γ), X-ray fluorescence- (XRF) and near-infrared (NIR) spectroscopy were used to create all possible combinations (fusing one to five PSSs) in order to predict six target soil properties: soil organic carbon (SOC), pH, moisture content as well as plant-available phosphorus (P), −magnesium (Mg) and −potassium (K). Additionally, we explored how incorporating RS data from Sentinel-2 affects prediction accuracies. Thereby, we reached the following conclusions:

Sensor fusion enhanced the overall prediction accuracy for the six target soil properties. However, using more than three PSSs showed only minimal further improvements, indicating that PSS fusion is subject to diminishing returns. Therefore, it is important to select a few informative and complementary PSSs rather than simply increasing the sensor quantity.

After sensor fusion, an $R^2 \geq 0.8$ has been achieved for five out of six target soil properties. Only predictions for K were less successful ($R^2 < 0.3$). NIR and γ proved to be the most successful sensors, both for sensor fusion and as standalone sensors. Nonetheless, there were various combinations of PSSs that demonstrated strong capabilities. Notably, combinations including pH-ISE, γ, NIR and ECa performed strongly. CSMoist was the weakest sensor and showed no real benefit over the other PSSs. While fusing PSS with RS proved mostly successful when having single PSSs, there were also cases in which adding RS decreased performances. Using models entirely fitted on RS already yielded reasonable performances ($R^2 > 0.5$) for all soil properties but K. Depending on the prediction technique, different PSSs or combinations of PSSs performed best. This indicates that model stacking yields more generalizable results. Lastly, model stacking also returned the highest prediction accuracy.

Since there was considerable intercorrelation between the target soil properties, it has to be considered that non-causal second order relations between PSSs and specific target soil properties have been exploited. Nonetheless, this study is unique in its number of distinctive PSSs that were used within one coherent dataset, thereby revealing insightful patterns about sensor fusion in the context of DSM and PA.

Funding

This research was supported by the Lower Saxony Ministry of Science and Culture (MWK), funded through the zukunft.niedersachsen program of the Volkswagen Foundation as well as the Federal Ministry of Education and Research of Germany (BMBF) through the i4s project: "I4S – Integrated System for Site-Specific Soil Fertility Management" (grant number 031B1069A).

CRediT authorship contribution statement

J. Schmidinger: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **V. Barkov:** Writing – review & editing, Methodology, Formal analysis, Data curation. **H. Tavakoli:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **J. Correa:** Writing – review & editing, Visualization, Data curation, Conceptualization. **M. Ostermann:** Data curation. **M. Atzmueller:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **R. Gebbers:** Writing – review & editing, Funding acquisition, Data curation, Conceptualization. **S. Vogel:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Data curation, 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.

Data availability

Code referring to the systematic workflow in Fig. 2 and a subset of the processed dataset is published in an associated GitHub repository under https://github.com/JonasSchmidinger/soil-sensor-fusion/tree/main. For demonstration purposes we provided a random subset consisting of 15 samples. The raw data cannot be openly published while the i4s project is ongoing but we may grant access to the whole dataset on request.

Appendix A

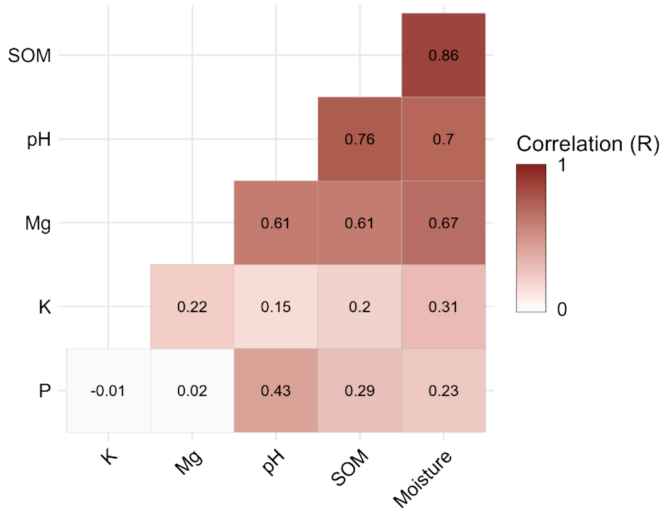


Fig. 1A. Correlation matrix for the six target soil properties, where the correlation is given as Pearson's correlation coefficient (R).

Table 1A
Descriptive statistics of the target soil properties.

Soil property	Mean	Median	Standard deviation	Min	Max	Skewness
SOC (%)	2.04	1.69	1.02	0.81	4.94	1.32
pH	6.31	6.20	0.69	4.90	7.50	0.12
Moisture (%)	9.95	10.21	2.62	4.99	16.95	0.43
P (mg P_2O_5 100^{-1} g soil)	4.44	3.90	2.41	1.00	14.80	1.85
Mg (mg MgO 100^{-1} g soil)	4.50	4.20	1.98	1.40	9.10	0.45
K (mg K_2O 100^{-1} g soil)	10.18	10.20	2.80	3.20	18.50	0.29

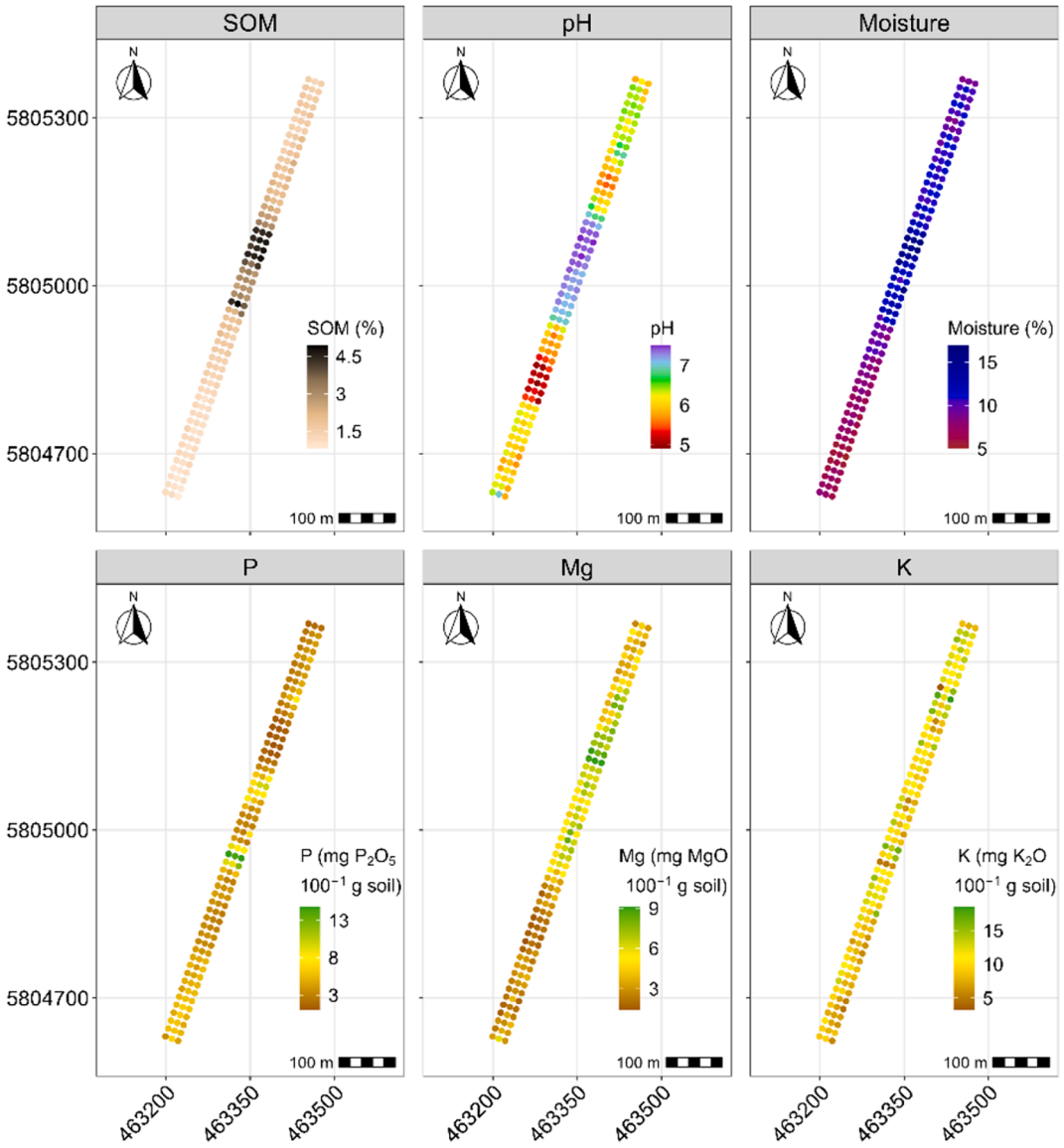


Fig. 2A. Sampling locations of the six laboratory-measured target soil properties and their values.

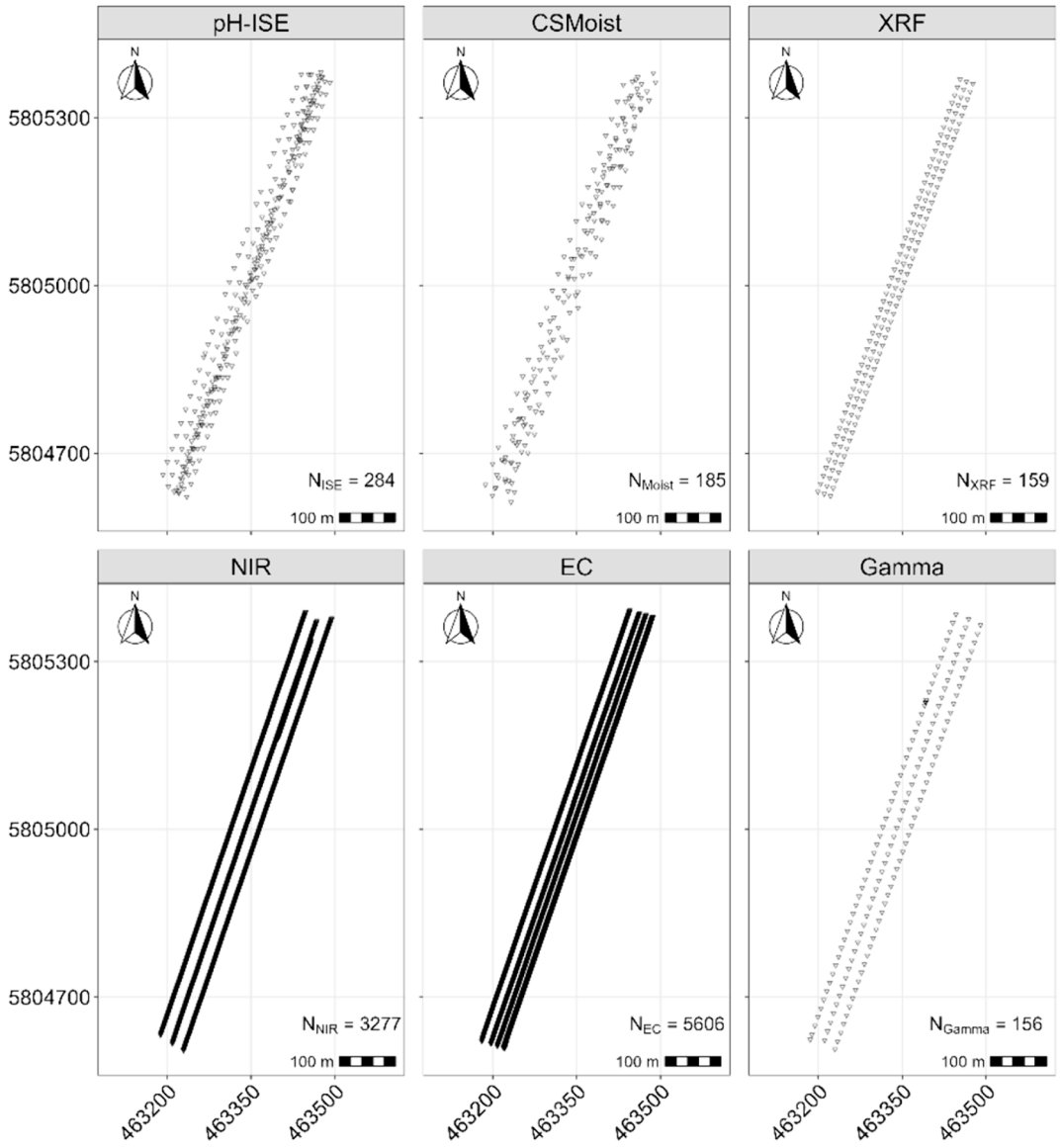


Fig. 3A. Spatial distribution of the PSS measurements after pre-processing.

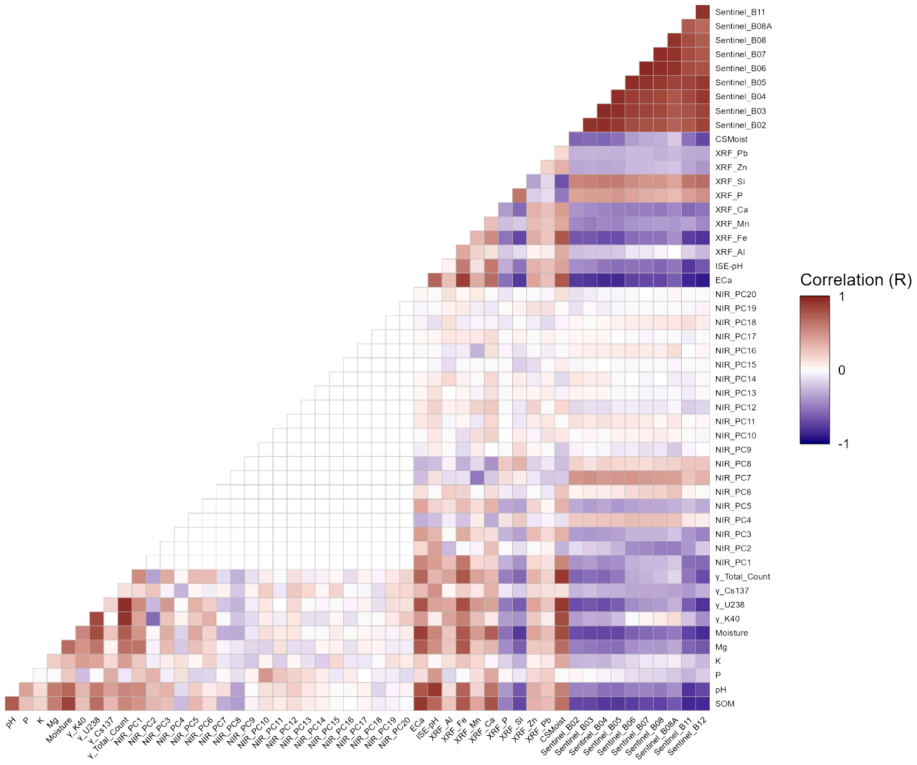


Fig. 4A. Correlation matrix for target soil properties and sensor data used for the modelling.

Fig. 5A. Contribution of different wavelength ranges from the NIR sensor on the principal components.

Appendix B

Table 1B
Overview about the base-models, their associated R-package and the hyperparameters as defined by the R-package. Hyperparameters inputted to the grid search are given in brackets. If no further information is given about the hyperparameters, then the default values associated to the R package were used.

Base-model	R-package	R-package reference	Hyperparameters
RF	randomForest	Liaw and Wiener (2002)	mtry* (0.2, 0.25, 0.33, 0.5, 0.67, 0.8, 1) nodesize (3, 6, 9, 12) samplesize* (0.6, 0.8, 1)

(*continued on next page*)

Table 1B (*continued*)

Base-model	R-package	R-package reference	Hyperparameters
Cubist	Cubist	Kuhn and Quinlan (2023)	committee (80) neigh (0, 1, 2, 4, 6, 8) rules (4, 6, 12, 24) sample* (0.6, 0.8, 1)
XGBoost	xgboost	Chen et al. (2023)	colsample_bytree (0.5, 0.75, 1) eta (0.005, 0.01, 0.05, 0.1) gamma (0.5, 1, 2) max_depth (4, 6, 8, 10) min_child_weight (3) nrounds (1000) subsample (0.6, 0.8, 1)
RR	penalized	Goeman et al. (2022)	lambda (0.001, 0.01, 0.1, 1, 10, 100, 1000)
SVR	e1071	Meyer et al. (2023)	cost (0.1, 1, 10, 100) kernel (linear, radial) gamma** (0.01, 0.1, 1, 10)

* Given as proportion. Actual input in R function took values in another format i.e., percentage or absolute value.** For radial kernel.

If NIR was included in the to be examined PSS combination, different numbers of principal components i.e., 5, 10, 15 and 20, were tested in the grid search for NIR. This was done for all base-models. Hence, the selection of principal components was included as conditional hyperparameter within the grid search (Eggensperger et al., 2014).

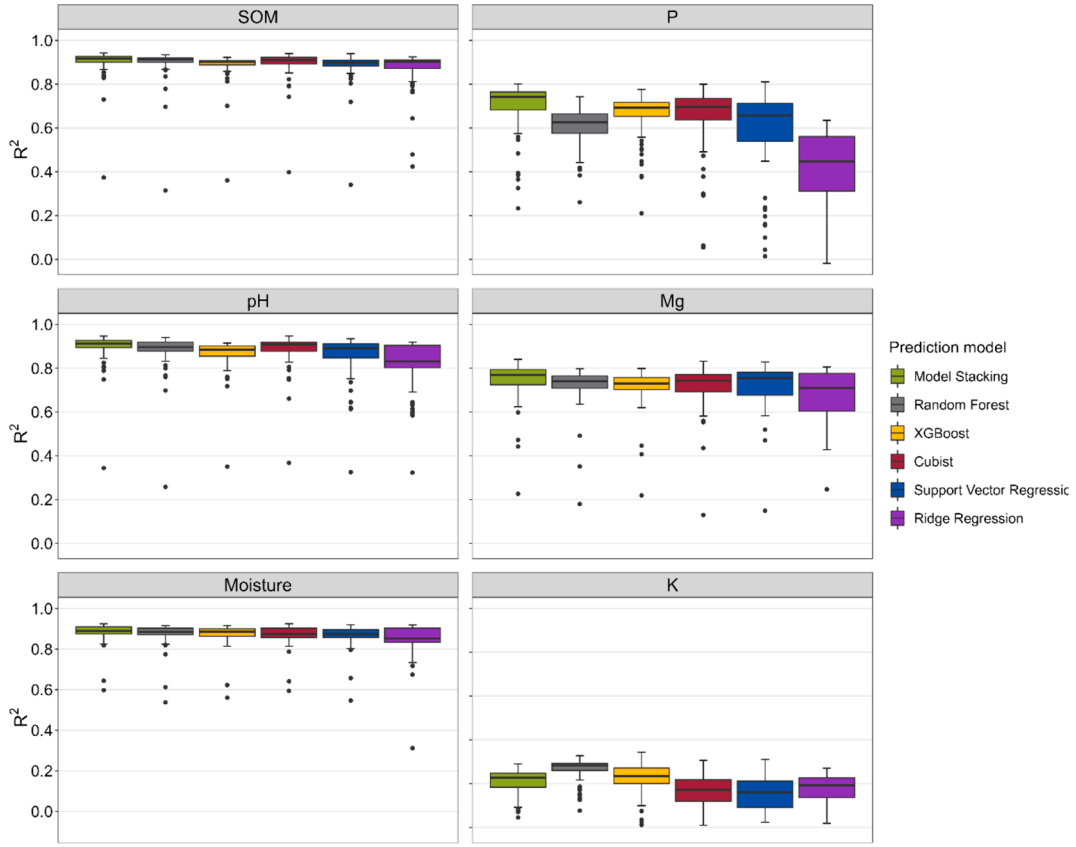


Fig. 1B. The obtained R^2, for each target soil property, for all *C* PSS combinations in dependence to the different base-models and model stacking.

Even though R^2 has some limitations (Alexander et al., 2015), we use it in this study as main metric since it is the most common scale independent model performance measure. In the following we also included a ranking based on the mean ratio of performance to Inter Quartile distance (RPIQ) as compared to the mean R^2 in Table 2. RPIQ is defined as:

$$RPIQ = \frac{Q3 - Q1}{RMSE}. \quad (1B)$$

Q3 and Q1 respond to the 75 % and 25 % quantiles of the testing data, respectively.

Table 2B
Top three PSS combinations ranked by mean RPIQ for each specific number of fused PSSs. The mean RPIQ represents the average of all six target soil properties.

Number of PSSs	RS	PSS	Mean RPIQ	Rank
1	with RS	NIR	3.12	1
	without RS	NIR	3.02	2
	with RS	γ	2.97	3
2	with RS	pH-ISE, NIR	3.41	1
	with RS	XRF, NIR	3.33	2
	with RS	γ, NIR	3.31	3
3	with RS	XRF, γ, NIR	3.52	1
	with RS	pH-ISE, NIR, ECa	3.51	2
	with RS	pH-ISE, γ, NIR	3.49	3
4	without RS	XRF, pH-ISE, γ, NIR	3.64	1
	without RS	XRF, pH-ISE, NIR, ECa	3.60	2
	without RS	pH-ISE, γ, ECa, NIR	3.60	3
5	without RS	XRF, pH-ISE, γ, ECa, NIR	3.67	1
	with RS	XRF, pH-ISE, γ, CSMoist, NIR	3.66	2
	without RS	XRF, pH-ISE, γ, CSMoist, NIR	3.61	3

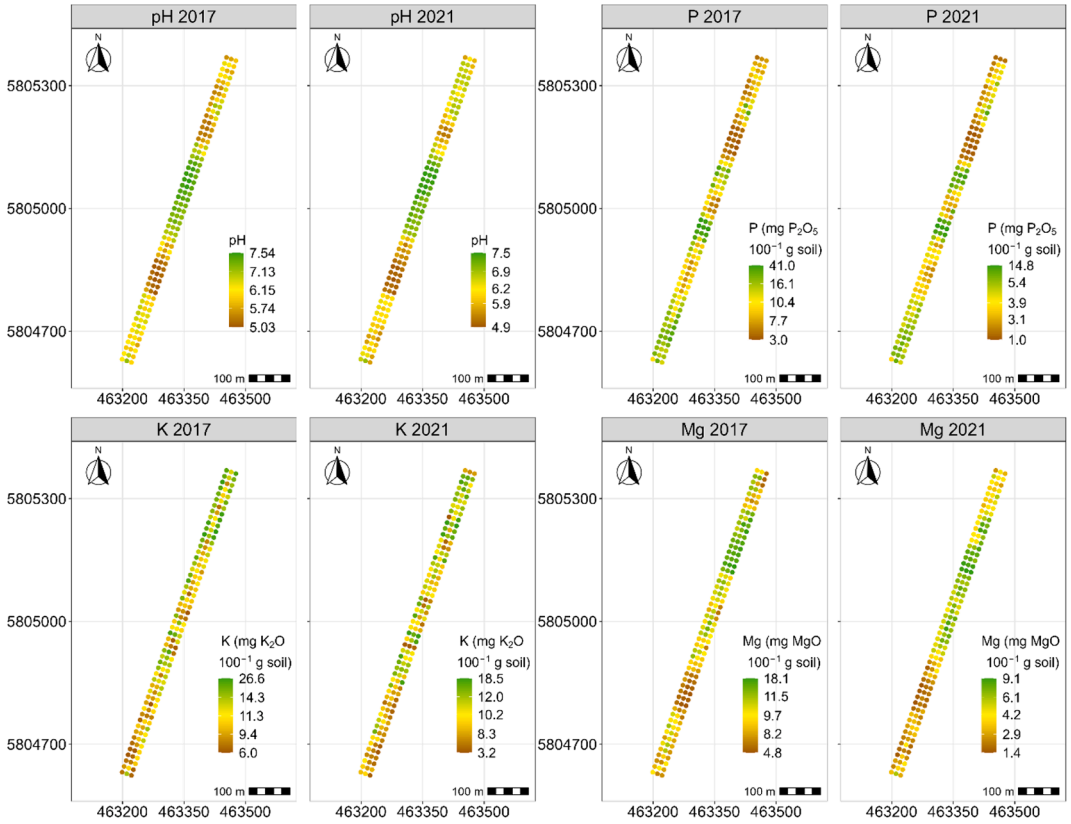


Fig. 2B. Soil property values from earlier sampling campaign (2017) in comparison to soil property values from 2021 campaign used for this study. Coloring-scale is quantile based in contrast to Fig. 2A.

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2024.117017.

References

Adamchuk, V.I., Viscarra Rossel, R.A., A., K., Schulze, P., 2011. Sensor Fusion for Precision Agriculture, in: Thomas, C. (Ed.), Sensor Fusion. Foundation and Applications. IntechOpen.

Adamchuk, V., Hummel, J., Morgan, M., Upadhyaya, S., 2004. On-the-go soil sensors for precision agriculture. Comput. Electron. Agric. 44 (1), 71–91.

Alexander, D.L.J., Tropsha, A., Winkler, D.A., 2015. Beware of R(2): simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. J. Chem. Inf. Model. 55 (7), 1316–1322.

Alijani, Z., Eyre, R., Saurette, D., Laamrani, A., Lindsay, J., Western, A., Berg, A., 2024. An efficient soil moisture sampling scheme for the improvement of remotely sensed soil moisture validation over an agricultural field. 0016-7061 442, 116763.

Bai, Y., Yang, W., Wang, Z., Cao, Y., Li, M., 2024. Improving the estimation accuracy of soil organic matter based on the fusion of near-infrared and Raman spectroscopy using the outer-product analysis. Comput. Electron. Agric. 219, 108760.

Bao, Y., Yao, F., Meng, X., Zhang, J., Liu, H., Mouazen, A.M., 2023. Predicting soil organic carbon in cultivated land across geographical and spatial scales: integrating sentinel-2A and laboratory Vis-NIR spectra. ISPRS J. Photogramm. Remote Sens. 203, 1–18.

Biney, J.K.M., Saberioon, M., Borůvka, L., Houška, J., Vašát, R., Agyeman, P.C., Coblinski, J.A., Klement, A., 2021. Exploring the suitability of UAS-based multispectral images for estimating soil organic carbon: comparison with proximal soil sensing and spaceborne imagery. Remote Sens. (Basel) 13 (2), 308.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Chatterjee, S., Hartemink, A.E., Triantafilis, J., Desai, A.R., Soldat, D., Zhu, J., Townsend, P.A., Zhang, Y., Huang, J., 2021. Characterization of field-scale soil variation using a stepwise multi-sensor fusion approach and a cost-benefit analysis. Catena 201, 105190.

Chen, T., Guestrin, C., 2016. XGBoost, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA. 13 08 2016 17 08 2016. ACM, New York, NY, pp. 785–794.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, T., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Yutian Li, Yuan, J., 2023. xgboost: Extreme Gradient Boosting: R package version 1.7.6.1.

Chen, Y., Gao, S., Jones, E.J., Singh, B., 2021. Prediction of soil clay content and cation exchange capacity using visible near-infrared spectroscopy, portable X-ray fluorescence, and X-ray diffraction techniques. Environ. Sci. Tech. 55 (8), 4629–4637.

Drucker, H., Burges, C.J., Kaufman, L., Smola, A., Vapnik, V., 1996. Support Vector Regression Machines. Advances in Neural Information Processing Systems 9.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36.

Eggensperger, K., Hutter, F., Hoos, H., Leyton-Brown, K., 2014. Surrogate Benchmarks for Hyperparameter Optimization. MetaSel@ECAI.

Gasmi, A., Gomez, C., Chehbouni, A., Dhiba, D., Elfil, H., 2022. Satellite multi-sensor data fusion for soil clay mapping based on the spectral index and spectral bands approaches. Remote Sens. (Basel) 14 (5), 1103.

Gebbers, R., Adamchuk, V.I., 2010. Precision agriculture and food security. Science(New York, N.Y.) 327, 828–831.

Gebbers, R., 2018. Proximal soil surveying and monitoring techniques, in: Stafford, J. (Ed.), Precision agriculture for sustainability. Burleigh Dodds Series in Agricultural Science. Burleigh Dodds Science Publishing, pp. 29–78.

Goeman, J.J., Meijer, R.J., Chaturvedi N, 2022. Penalized: L1 (lasso and fused lasso) and L2 (ridge) penalized estimation in GLMs and in the Cox model: R package version 0.9-52.

Grunwald, S., Vasques, G.M., Rivero, R.G., 2015. Fusion of Soil and Remote Sensing Data to Model Soil Properties, in: , vol. 131. Advances in Agronomy. Elsevier, pp. 1–109.

Hiemstra, P.H., Pebesma, E.J., Twenhoefel, C.J.W., Heuvelink, G.B.M., 2008. Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network. Comput. Geosci.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12 (1), 55.

Huang, J., Scudiero, E., Clary, W., Corwin, D.L., Triantafilis, J., 2017. Time-lapse monitoring of soil water content using electromagnetic conductivity imaging. Soil Use Manag. 33 (2), 191–204.

Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Trans. Inform. Theory 14 (1), 55–63.

Ji, W., Adamchuk, V.I., Chen, S., Mat Su, A.S., Ismail, A., Gan, Q., Shi, Z., Biswas, A., 2019. Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. Geoderma 341, 111–128.

Keesstra, S.D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., Montanarella, L., Quinton, J.N., Pachepsky, Y., van der Putten, W.H., Bardgett, R.D., Moolenaar, S., Mol, G., Jansen, B., Fresco, L.O., 2016. The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. Soil 2 (2), 111–128.

Kuhn, M., Quinlan, R., 2023. Cubist: Rule- And Instance-Based Regression Modeling: R package version 0.4.2.1.

Kuhn, M., Johnson, K., 2013. Regression Trees and Rule-Based Models. In: Applied Predictive Modeling. Springer, New York, New York, NY, pp. 173–220.

Leblanc, M., Tibshirani, R., 1996. Combining estimates in regression and classification. J. Am. Stat. Assoc. 91 (436), 1641–1650.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest: R package version 4.7.1.1.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. SENTINEL-2 SEN2COR: L2A Processor for Users, in: Proceedings Living Planet Symposium 2016. ESA Living Planet Symposium 2016, Prague, Czech Republic. 09 - 13 May 2016. Spacebooks Online, pp. 1–8.

McBratney, A., Mendonça Santos, M., Minasny, B., 2003. On digital soil mapping. 0016-7061 117 (1-2), 3–52.

Meyer, S., Kling, C., Vogel, S., Schroeter, I., Nagel, A., Kramer, E., Gebbers, R., Philipp, G., Lueck, K., Gerlach, F., Scheibe, D., Ruehlmann, J., 2019. Creating soil texture maps for precision liming using electrical resistivity and gamma ray mapping, in: Precision agriculture '19. Papers presented at the 12th European Conference on Precision Agriculture, Montpellier, France, 8-11 July 2019. 12th European Conference on Precision Agriculture, Montpellier, France. 08-11 07, 2019. Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 539–546.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2023. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien: R package version 1.7-14.

Mouazen, A.M., Shi, Z., 2021. Estimation and mapping of soil properties based on multi-source data fusion. Remote Sens. (Basel) 13 (5), 978.

O'Rourke, S.M., Minasny, B., Holden, N.M., McBratney, A.B., 2016. Synergistic use of Vis-NIR, MIR, and XRF spectroscopy for the determination of soil geochemistry. Soil Science Soc. Amer J. 80 (4), 888–899.

Quinlan, R.J., 1992. Learning with Continuous Classes. In: 5th Australian Joint Conference on Artificial Intelligence. World Scientific, Singapore, pp. 343–348.

R Core Team, 2023. R: A Language and Environment for Statistical Computing. Austria, Vienna.

Schmidinger, J., Schroeter, I., Boenecke, E., Gebbers, R., Ruehlmann, J., Kramer, E., Mulder, V.L., Heuvelink, G.B.M., Vogel, S., 2024. Effect of training sample size, sampling design and prediction model on soil mapping with proximal sensing data for precision liming. Precision Agric 1–27.

Shi, P., Six, J., Sila, A., Vanlauwe, B., van Oost, K., 2022. Towards spatially continuous mapping of soil organic carbon in croplands using multitemporal Sentinel-2 remote sensing. ISPRS J. Photogramm. Remote Sens. 193, 187–199.

Tavakoli, H., Correa, J., Vogel, S., Gebbers, R., 2022. RapidMapper–a mobile multi-sensor platform for the assessment of soil fertility in precision agriculture. Proceedings International Conference on Agricultural Engineering. AgEng-LAND. TECHNIK 2022., 351–357.

Tavares, T.R., Molin, J.P., Nunes, L.C., Wei, M.C.F., Krug, F.J., de Carvalho, H.W.P., Mouazen, A.M., 2021. Multi-sensor approach for tropical soil fertility analysis: comparison of individual and combined performance of VNIR, XRF, and LIBS spectroscopies. Agronomy 11 (6), 1028.

Vasques, G.M., Rodrigues, H.M., Coelho, M.R., Baca, J.F.M., Dart, R.O., Oliveira, R.P., Teixeira, W.G., Ceddia, M.B., 2020. Field proximal soil sensor fusion for improving high-resolution soil property maps. Soil Syst. 4 (3), 52.

Viscarra Rossel, R.A., Bouma, J., 2016. Soil sensing: a new paradigm for agriculture. Agr. Syst. 148, 71–74.

Vogel, S., Boenecke, E., Kling, C., Kramer, E., Lueck, K., Philipp, G., Ruehlmann, J., Schroeter, I., Gebbers, R., 2022. Direct prediction of site-specific lime requirement of arable fields using the base neutralizing capacity and a multi-sensor platform for on-the-go soil mapping. Precision Agric 23 (1), 127–149.

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.

Wang, J., Zhao, X., Deuss, K.E., Cohen, D.R., Triantafilis, J., 2022. Proximal and remote sensor data fusion for 3D imaging of infertile and acidic soil. Geoderma 424, 115972.

Wang, J., Zhao, X., Triantafilis, J., 2024. Synergistic use of proximally sensed and time series remotely sensed imagery to map soil sodicity. Comput. Electron. Agric. 216, 108466.

Wenjun, J., Zhou, S., Jingyi, H., Shuo, L., 2014. In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy. PLoS One 9 (8), e105708.

Xu, D., Zhao, R., Li, S., Chen, S., Jiang, Q., Zhou, L., Shi, Z., 2019. Multi-sensor fusion for the determination of several soil properties in the Yangtze River Delta China. European J Soil Science 70 (1), 162–173.

Xue, J., Zhang, X., Chen, S., Lu, R., Wang, Z., Wang, N., Hong, Y., Chen, X., Xiao, Y., Ma, Y., Shi, Z., 2023. The validity domain of sensor fusion in sensing soil quality indicators. Geoderma 438, 116657.