

<https://doi.org/10.1038/s44264-026-00125-0>

A digital soil mapping approach to soil carbon monitoring, reporting and verification (MRV)



Alexandre M.J.-C. Wadoux¹, Mitchell Donovan², Peng Fu^{2,3}, Nicholas R. Leach², Julia Maddalena², Rose Rustowicz², David Schurman² & James R. Kellner^{2,4,5} ✉

Soils can sequester carbon and contribute to climate change mitigation, but credible soil carbon offset markets require robust Monitoring, Reporting, and Verification (MRV) programs. Current approaches rely on models with soil sampling or on repeated probability sampling. We propose a digital soil mapping (DSM) framework to estimate soil organic carbon (SOC) stock change and uncertainty. Using physical samples from a hypothetical project across seven fields in two U.S. states, we calibrate, predict, validate, and spatially aggregate SOC estimates. We assess the effect of calibration sample size on field- and project-level means and variances and extend the framework to quantify change. Estimated uncertainty of averaged SOC stocks was $<1 \text{ Mg ha}^{-1}$, yielding a 12% uncertainty deduction for a 5-year, 430-acre project when using a probability of exceedance method. Assuming the sequestration rate is 0.6% of the stock size, simulations suggest deductions $<10\%$ for most projects >5 years and $<5\%$ for large ($>50,000$ acres), long-term projects. This DSM approach offers a cost-effective and scalable alternative to other existing methods.

Soils are one of the largest reservoirs of organic carbon¹. It is estimated that, globally, 2500 Pg of carbon is stored in the first meter of soil, of which 1500 Pg is organic carbon^{2,3}. A large body of work⁴ has shown that soils are a potentially large source of CO₂ emissions in response to management and climate change, while also being a sink able to provide nature-based solutions to atmospheric CO₂ capture. Sanderman et al.⁵ have estimated that 133 Pg of carbon has been lost from 12,000 years of human agricultural land use. However, there is a general agreement that with improved land management practices such as cover cropping, conservation tillage, organic manure application or rotational grazing⁶, soils can potentially store carbon. This would make soils a substantial sink of CO₂ emissions worldwide.

To this end, there is a growing recognition of the potential of soils for carbon sequestration in the broader context of climate change mitigation. The 4p1000 initiative⁷ on Soil for Food Security and Climate, for example, launched at the COP 21, proposed to increase the global soil organic carbon stocks by 0.4% annually as a compensation for global anthropogenic greenhouse gas emissions. More recently, the European Commission has proposed a regulation for carbon removal certification⁸, which, among other options, incentivizes carbon farming and provides rules for certifying carbon removal within Europe. Such initiatives echo the implementation of

several voluntary soil carbon offset markets throughout the world, in which farmers can adopt on-farm practices that promote soil carbon sequestration, earning payments through the creation of voluntary carbon credits. These credits can then be used by emitters to offset their current anthropogenic emissions as part of efforts to minimize the impact of their activities⁹.

A prerequisite for participation in the voluntary carbon market is the availability of tools and techniques capable of monitoring changes in soil organic carbon (SOC) resulting from shifts in land management practices. These efforts are typically framed within the context of Monitoring, Reporting, and Verification (MRV) schemes¹⁰. Broadly, MRV guidelines are designed to quantify and verify changes in soil carbon stocks while accounting for economic considerations, payment systems and market types in relation to the scale of the project. Many MRV frameworks and approved methodologies necessitate the measurement of SOC stocks through baseline sampling and the subsequent crediting of net-removal activities following rigorous verification processes.

Current methodologies to monitor SOC stock change are rooted in the use of soil biogeochemical model simulation along with adaptive baselines and soil sampling (e.g. version 2.1 of the Verra Methodology for Improved Agricultural Land Management (VM0042), and the Climate Action Reserve

¹College of Science and Engineering, James Cook University, Cairns, QLD, Australia. ²Perennial Climate Inc., Boulder, CO, USA. ³School of Plant, Environmental and Soil Sciences, Louisiana State University and Louisiana State University AgCenter, Baton Rouge, LA, USA. ⁴Institute at Brown for Environment and Society, Brown University, Providence, RI, USA. ⁵Department of Ecology, Evolution and Organismal Biology, Brown University, Providence, RI, USA.

✉ e-mail: jim@perennial.earth

Soil Enrichment Protocol). While these methodologies have been adopted in the voluntary carbon market, with 100 out of the 133 projects¹¹ currently using one of the two methodologies¹², there remain significant issues in the use of biogeochemical models in soil carbon crediting. Notably, these models are usually making a forecast, meaning that future changes in SOC are projections under the assumption that the model is correctly representing the processes in question. While this is not a fundamental issue, it is one that DSM approaches address by relying on outcome-based measurements of actual SOC change. Other issues with the use of biogeochemical models pertain to the limitation of the mechanistic models themselves (e.g. boundary conditions, spatial representation, large uncertainties) and the high-cost and labour intensive nature of soil sampling for large projects. For these reasons, very few sustainable agricultural projects have issued credits to date, and soil carbon projects currently represent only a minor fraction of the projects under study in the voluntary carbon market.

Until recently, the only alternative to using a mechanistic model was the use of physical soil samples to directly measure changes in SOC stocks. This can be done if the samples are selected with probability sampling, in which case population parameters (e.g. mean and variance of the project-level SOC stocks) can be estimated with a design-based inference. Addressing change over time with a probability sampling is straightforward, either indirectly with the two means and variances of the two time steps, or directly with space-time sampling. For more information on space-time probability sampling, the reader is redirected to Chapter 14 in de Gruijter et al.¹³. Current methodologies address this possibility by leaving room to monitor SOC stock changes with design-based inference. While this could be the preferred approach because it is statistically sound and allows a robust estimation of the SOC stock change uncertainty, designing a probability sample is often not possible because agricultural soil carbon projects are dynamic (i.e. new fields are added in the lifetime of a project) and the condition of a probability sampling would have to be satisfied for all new added fields. Further, the required sampling effort to reach an acceptable level of uncertainty can be substantial^{14,15}.

When biogeochemical models or design-based approaches are not practical, a model-based approach with an algorithmic solution (e.g. machine learning) or a (geo)statistical model can be used. Several techniques from the field of digital soil mapping have matured and are operational for estimating SOC stock at various spatial supports^{16–20}. In digital soil mapping (DSM), an empirical model is fitted between a sampled measurement of the property of interest and a set of environmental covariates. An attractive aspect of DSM is that it enables a robust estimation of uncertainty of the predicted SOC stocks and a validation of both the prediction and prediction uncertainty. There is also a trend towards developing methods and approaches that use DSM for MRV. The Australian Clean Energy Regulator quantification approach²¹, for example, allows the use of a statistical model for estimating SOC stocks and change. The Verified Carbon Standard recently published VT0014, which brings DSM into compliance with methodologies for quantifying changes in SOC stocks in agricultural land management carbon programs.

This paper demonstrates an end-to-end DSM framework for soil carbon MRV using ATLAS-SOC, a DSM framework developed to quantify organic carbon content in the top 30 cm of soil in non-tree row crops in the continental United States²⁰. This approach is extensible to a variety of covariates and statistical models and can be applied to estimate SOC stocks or changes in SOC stocks over time. We present model calibration, prediction, validation, and spatial aggregation of predictions and their uncertainties after baseline sampling of a hypothetical project. We examine the impact of different project sample sizes on uncertainty and validation metrics, and discuss a possible extension of the framework for quantifying SOC stock change and its uncertainty. The end-to-end DSM approach is developed within the framework of existing VCS methodologies and other initiatives that require spatially explicit estimates of SOC stocks with uncertainty at various spatial supports (e.g. field, project, region, etc.). The requirement for uncertainty deduction from VM0042 is tested on scenarios of SOC stock change with quantified uncertainty.

Results

The end-to-end approach

The QRF model was optimized with three hyperparameters: Mtry set to 7, ntree set to 750, and the minimum node size set to 7. The model, fitted with these hyperparameters, was used for cross-validation (CV) and prediction. The CV predictions generated by the QRF model yielded a ME of 0.01 Mg ha⁻¹, an RMSE of 10.56 Mg ha⁻¹, an R² of 0.61, and a PICP of 91%. These results indicate that the model successfully passed all three evaluation criteria: the R² value was greater than 0, the bias was not significantly different from 0 (as confirmed by a one-sample t-test with *p*-value = 0.79), and the PICP was greater than or equal to 90%.

The maps of the SOC stock predictions using the QRF model are presented in Fig. 1. The predicted values ranged from 19 to 76 Mg ha⁻¹, showing significant differences among fields in terms of both the magnitude of spatial variation and spatial patterns. For example, Pesotum had abrupt variations in SOC stocks over relatively short distances (i.e. less than 50 m). In contrast, Hammond and Villa Grove displayed smoother spatial patterns, with gradients of increasing SOC stocks from east to west and west to east, respectively. Predicted point-level standard deviations are shown in Fig. S2.

Figure 2 shows the correlograms of the seven fields derived from the fitted spherical variograms of the standardized prediction errors. Among the seven fields, two had strong spatial correlation in the standardized prediction errors at short distances. Specifically, Macon and Villa-Grove had ranges of 88, 43, respectively, and nuggets close to 0. The standardized prediction errors for the Brighton, Hammond and Arcola fields showed moderate spatial correlation but at larger distances, with ranges of 108, 77 and 68 m, respectively, and nuggets at 0.12, 0.20 and 0.19 Mg ha⁻¹. In contrast, the remaining fields of Pesotum and Tolono had little spatial correlation in the standardized prediction errors, resulting in large range (i.e. more than 100 m) but a small nugget-to-sill ratio.

Table 1 shows the values of the average SOC stock predictions along with their associated uncertainty (i.e. prediction and standard deviation in Mg ha⁻¹) for the seven fields and at the project level. The uncertainty at the project level was derived using MC sampling across all fields but with the field-specific correlograms. The average SOC stock values ranged from 23.41 Mg ha⁻¹ for Brighton to 64.13 Mg ha⁻¹ for Tolono, while the standard deviation varied from 0.56 Mg ha⁻¹ for Tolono to a maximum of 1.69 Mg ha⁻¹ in Macon.

Effect of the sample size on model calibration

Figure 3 illustrates the results of the sample size experiment, which tested a range of 1 to 30 samples per stratum. Note that the samples were drawn with replacement from the existing sampling locations. As shown in Fig. 3, the estimated average SOC stock remained consistent across different numbers of samples per stratum. This is particularly evident at the project level, where the average SOC stock stayed constant at approximately 60 Mg ha⁻¹. However, the fields of Arcola and Brighton had contrasting patterns in the variation of average SOC stock with varying numbers of samples per stratum. In Arcola, the average SOC stock slightly increased as the number of samples per stratum increased, rising from 60 to 64 Mg ha⁻¹ between 1 and 30 samples. In contrast, Brighton showed a moderate decrease in average SOC stock, declining by 9 Mg ha⁻¹ over the same range of samples per stratum.

Figure 4 displays the standard deviation of the SOC stock for the project and the seven fields. At the project level, the standard deviation exhibited a decreasing trend as the sample size increased, ranging from 2.5 Mg ha⁻¹ for a sample size of 1 to 1 Mg ha⁻¹ for a sample size of 30. The latter value was higher than the standard deviation of 0.51 Mg ha⁻¹ obtained when using all available samples. A similar trend was observed for three fields: Pesotum, Tolono and Villa Grove. In contrast, the standard deviation for the remaining fields appeared to be unaffected by the various sample sizes.

Uncertainty deduction

Figure 5 shows the SOC stock change expressed in tons of CO₂ for the three time intervals of 3, 5 and 10 years, along with the estimated uncertainty deduction. The estimated value of the change at project level (i.e. for 430

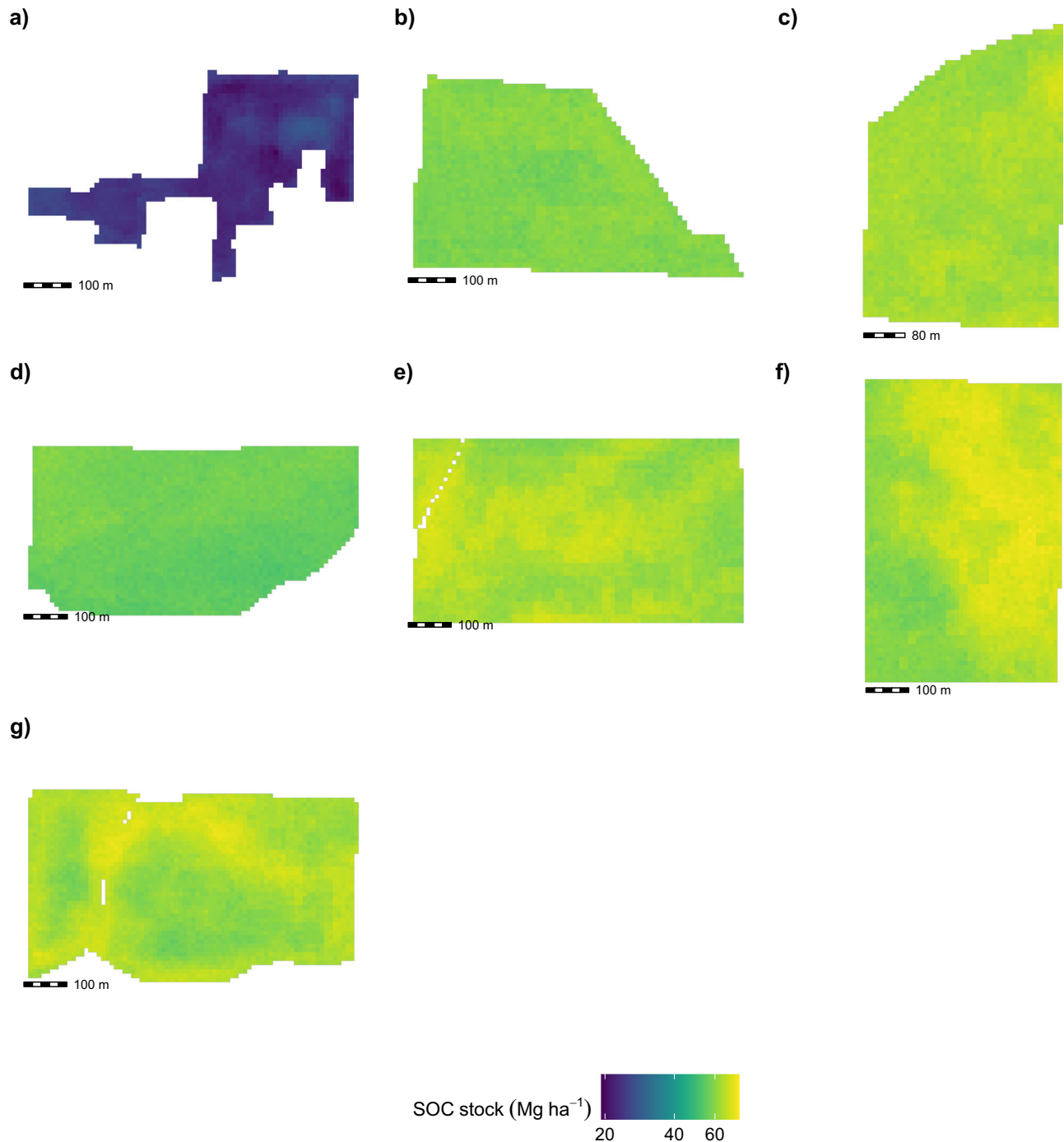


Fig. 1 | Soil organic carbon stock prediction maps. Maps show the SOC stock prediction for the seven fields: **a** Brighton, **b** Hammond, **c** Arcola, **d** Macon, **e** Pesotum, **f** Villagrove, and **g** Tolono.

acres) was 4.01, 6.73 and 13.67 t CO₂ for the interval of 3, 5 and 10 years, respectively. The creditable value obtained using the 33rd percentile of the normal distribution was 3.20, 5.96 and 12.92 t CO₂, which represented an uncertainty deduction of 20.39% for the interval of 3 years, 11.51% for the interval of 5 years and 5.47% for the interval of 10 years.

Figure 6 shows the simulated uncertainty deduction for project sizes up to 100,000 acres. Because the absolute value of the project-level variance of the mean emissions removal remained fixed regardless of duration, increasing the duration of the project, and therefore the total amount of the emissions removal estimate, always led to a decrease of the corresponding uncertainty deduction. However, increasing the size of the project also

reduced uncertainty, even when duration was held constant. For a five year period, for example, the uncertainty deduction for a 10,000 acre project was twice the deduction associated with a 100,000 acre project. Most of the decrease in the uncertainty deduction occurred as projects increased in size from very small to about 50,000 acres. This trend was observed for all time steps. At a given project size, a time step of 3 years always led to larger uncertainty deduction than projects of longer duration. The size of the uncertainty deduction was also dependent on the magnitude of the correlation of the errors. Larger correlations resulted in smaller deductions at a given project size and duration. Additional simulations of uncertainty deduction for different sequestration rates are provided in Figure S3.

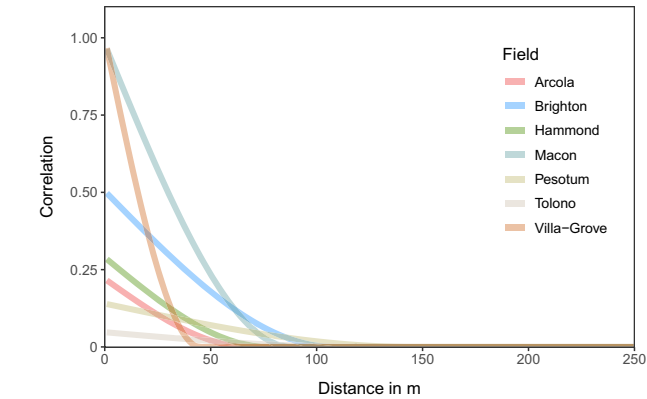


Fig. 2 | Fitted correlograms. Correlograms obtained from the variograms of the standardized prediction errors for the seven fields.

Discussion

The end-to-end DSM framework using ATLAS-SOC for soil carbon MRV yielded realistic field- and project-level estimation of the SOC stock. While the point-level standard deviation ranged between 2.60 and 21.41 Mg ha⁻¹, the standard deviation of the spatially aggregated mean at the project level was 0.51 Mg ha⁻¹. This relatively low standard deviation is expected because averaging reduces uncertainty. For comparison, in a previous study in the Swiss forests, Nussbaum et al.²² calculated SOC stocks using robust external-drift kriging and reported mean SOC stocks with standard errors of 0.15 kg/m² (i.e. 1.5 Mg ha⁻¹) for the 0–30 cm depth interval and 0.24 kg/m² (i.e. 2.4 Mg ha⁻¹) for the 0–100 cm depth interval. However, we found few studies that reported uncertainty of aggregated values of SOC stock in agricultural fields. This is partly because obtaining realistic uncertainty estimates for spatial averages or totals remains challenging. Currently, most maps are produced using machine learning to generate point-level uncertainty estimates, rather than through conventional geostatistical modelling, and

Table 1 | SOC stocks for the 0–30 cm depth interval in Mg ha⁻¹ and uncertainty (i.e. mean and standard deviation) obtained with the QRF model and the correlograms of Fig 2 with the method of Wadoux & Heuvelink³³

	Arcola	Brighton	Hammond	Macon	Pesotum	Tolono	Villa-Grove	Project
Mean	63.44	23.41	59.06	56.11	64.07	64.13	64.11	60.48
Standard deviation	0.95	0.95	1.33	1.69	1.16	0.56	0.92	0.51

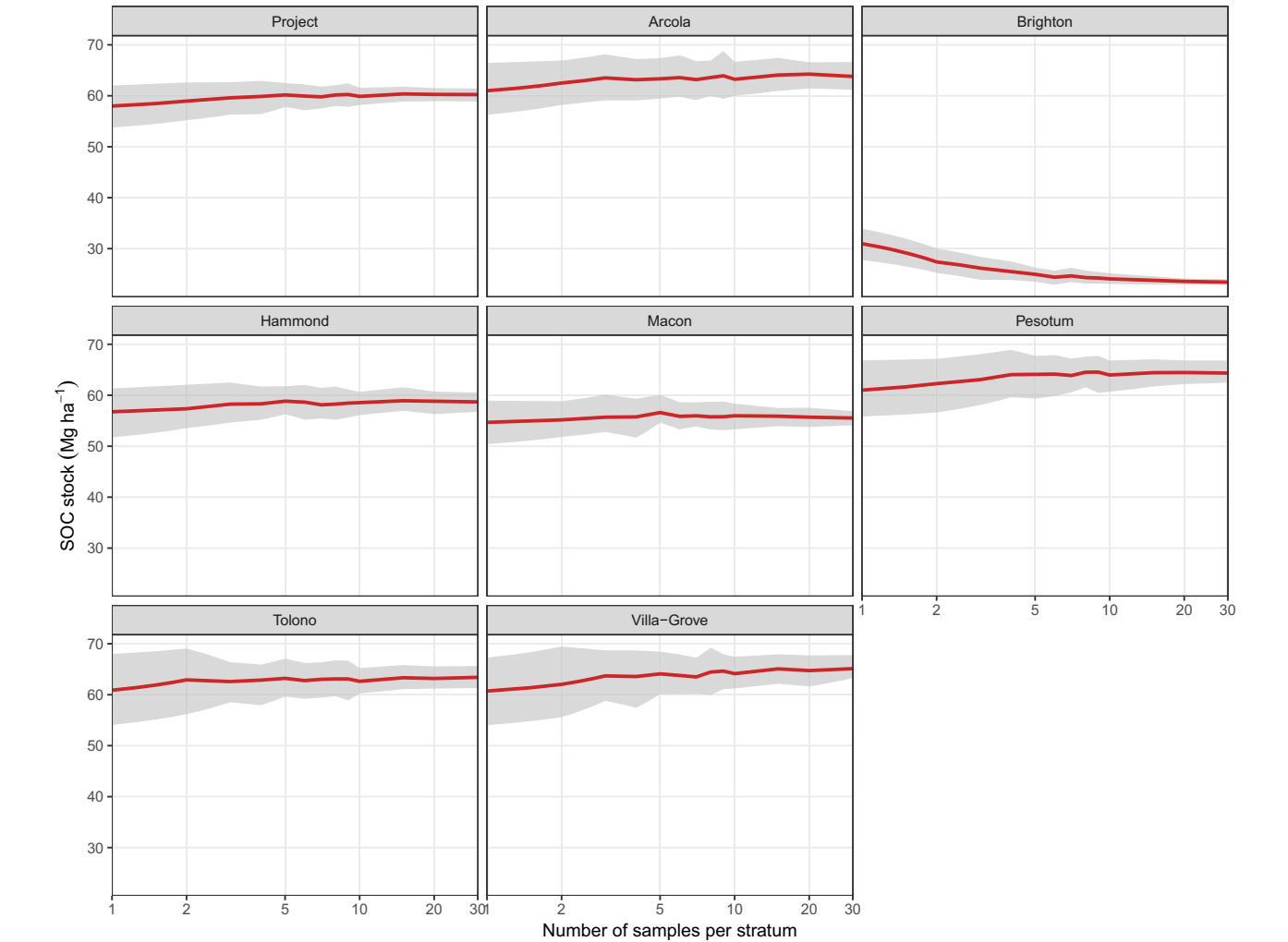


Fig. 3 | Simulation of the model-estimated means. Model-estimated mean SOC stock in Mg ha⁻¹ for the project and the seven fields as a function of the calibration sample size. The grey lines represent the subsampling for the number of samples per stratum and the red line is the average of all subsampling.

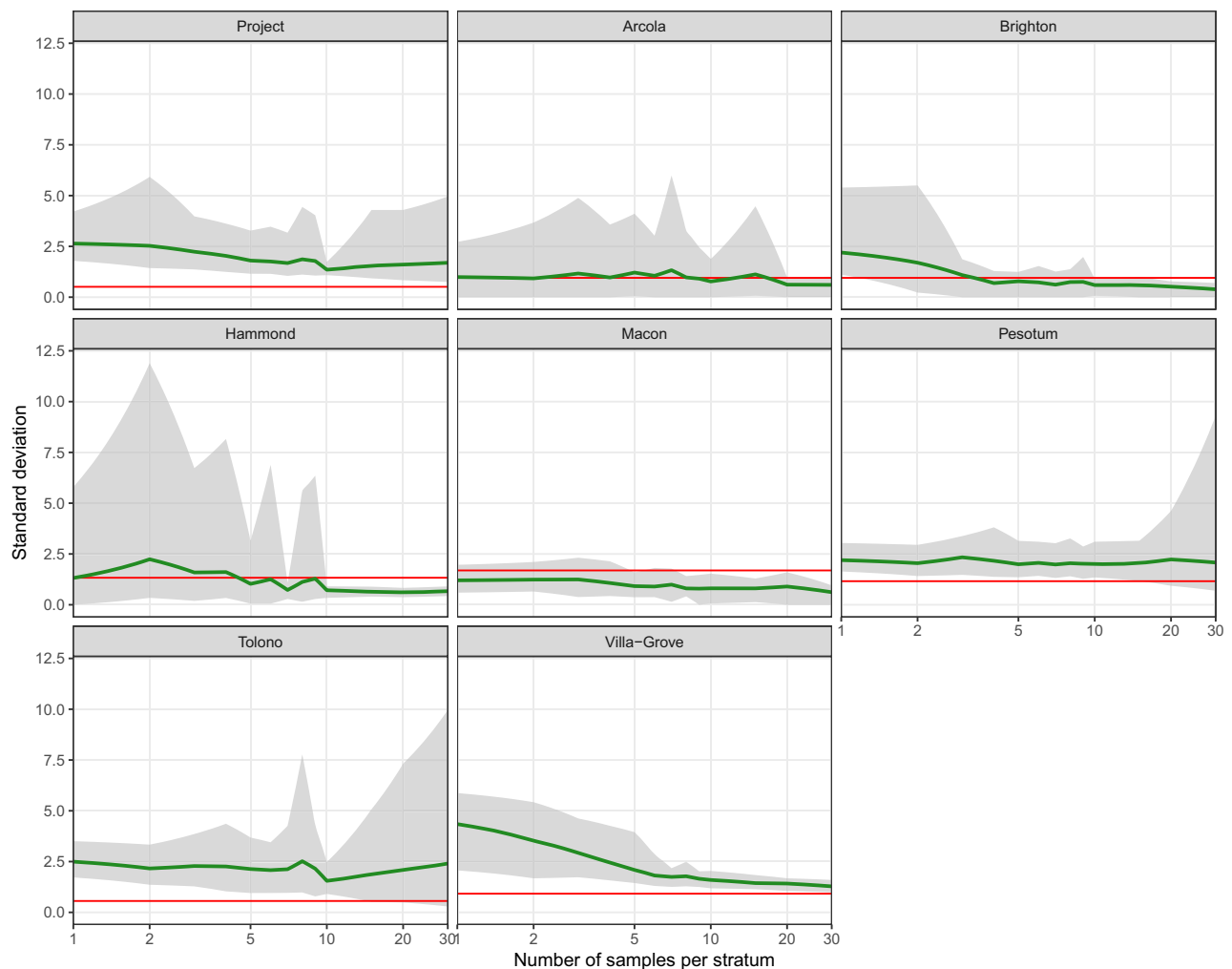


Fig. 4 | Simulation of the model-estimated standard deviations. Standard deviation of the SOC stock in Mg ha^{-1} for the project and the seven fields as a function of the calibration sample size. The grey lines represent the standard deviation value for a number of samples per stratum and the green line is the average of all standard

deviation values for a given number of samples per stratum. The horizontal red line is the value of the standard deviation obtained using all available samples following the results shown in Table 1.

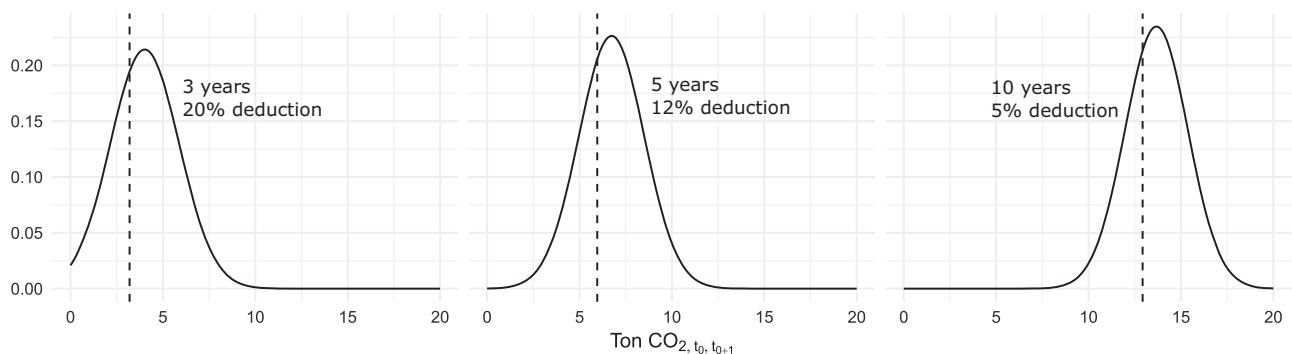


Fig. 5 | Uncertainty deductions for several year intervals. Change in ton of CO_2 emission for a middle-of-the-road sequestration rate and the three time intervals of 3, 5 and 10 years. The mean is obtained with Eq. (6) and the variance with Eq. (7).

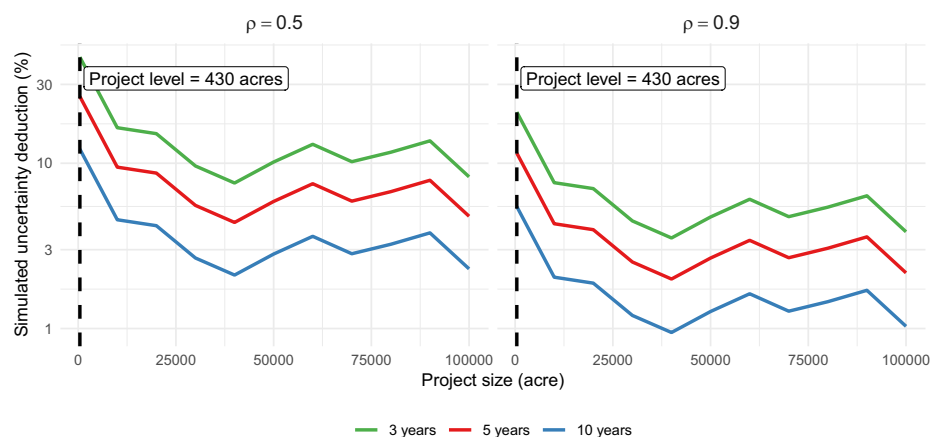
therefore no direct estimation of variance at the project level is straightforward to implement.

One striking result was that the average SOC stock at the project level varied little with different sample sizes used for model training. The standard deviation at the project level did show a slight decrease as the sample size increased. This was expected because uncertainty decreases as the

sample size increases, but it did not decrease further for larger sample sizes. It was also expected because here we did not apply kriging of the residuals to improve the predictions, as is done after regression in universal kriging²³. While the average predicted value at the project level did not change much with increasing sampling density, the model confidence in the prediction did increase. This is an interesting result because, in carbon crediting

Fig. 6 | Simulation of the uncertainty deduction.

The simulation of the uncertainty deduction are obtained for different project sizes for two correlation values of the errors between the time steps. Note the logarithmic scale on the y-axis. The coloured lines represent the time steps for projects of 3, 5 and 10 years duration. The dashed vertical line represents the uncertainty deduction value obtained in this study at project level.



programs, one could find a balance between the required sampling density and the level of uncertainty that one is willing to accept.

The uncertainty deduction was relatively modest for both large projects (i.e. those exceeding 50,000 acres) and longer time intervals between the baseline and evaluation periods (i.e. more than five years). Our simulations typically indicated that for a project covering 100,000 acres over a 3-year period, the uncertainty deduction applied to the estimated change could be close to 5%. A similar deduction was observed for the relatively small project described in our study (i.e. with seven fields and a total of ~430 acres), but with a longer sequestration period of 10 years. This is unsurprising, as averaging uncertainty across large areas reduces overall uncertainty and consequently the uncertainty deduction. This effect is well known but not really addressed in the current literature and current MRV approaches^{10,24}. To reduce the uncertainty deduction when using DSM, project proponents can improve model prediction accuracy, extend the duration of the project (i.e. the sequestration period), or increase the project area. While we tested two different values representing the correlation of errors between points in time, we do not consider this a parameter that project proponents can influence, as it should be empirically estimated, for example, using a cross-correlogram. It was reasonable to assume that this correlation was relatively high, as model errors are often spatially consistent over time (e.g. underestimation of SOC stocks in areas with high SOC stock values and overestimation in areas with low SOC stocks), particularly when the same model is applied at both time steps.

The current experiment relied on subsampling from the existing locations, which was made possible due to the very high sampling density of the fields under study. While this is a common approach in DSM studies^{25–27}, and provides valuable information on the effect of sample size on the uncertainty of spatially aggregated SOC stocks, it treated the existing set of sampling locations as the population of interest. An alternative approach to this experiment would be to use the existing locations to generate a map or a set of possible realities, from which the selection of a sampling design for different sample sizes could be repeated. This would enable testing the sampling density by selecting samples from any points within the geographical domain. Such approaches have been used in various ways in the literature, including treating maps as the population of interest²⁸, using ordinary kriging of the existing data to generate a reference map²⁹, or using geostatistical simulations³⁰ to generate possible realities from which sampling can be repeated.

The end-to-end approach required fitting a variogram to the standardized prediction error in order to derive the correlation function. This required a sufficiently large number of sampling locations. A rule of thumb is that a minimum of 100 to 150 locations is needed for variogram fitting using the Methods of Moments³¹, and 50 locations are needed when (restricted) maximum likelihood estimation is applied³². The sampling locations should also cover a wide range of distances, including short distances. In fact, Wadoux & Heuvelink³³ noted the sensitivity of the spatial

aggregation procedure to variogram fitting and short-distance correlation. Since most correlation occurs at short distances, and correlation is zero for distances larger than the range, it was critical to put effort into estimating the nugget parameter. We found, however, that the range parameter also had a strong influence on the value of the spatially aggregated uncertainty. We tested this by fitting a large number of exponential variograms with slightly different range values (results not shown) and estimating the uncertainty of the mean. The results showed that the values of uncertainty increased linearly with the effective range. This effect was less visible with the nugget parameter when performing the same experiment. We suggest following standard procedures, which are well described in the literature, to fit variograms³⁴, including limiting the range of the variogram function to about half the size of the area and using a bounded variogram.

When sufficient sampling locations are not available to fit the variogram, an average variogram³⁵ or expert knowledge could be used, for which established procedures exist. For example, Kerry et al.³⁶ showed that existing variograms of soil properties from similar parent materials had similar ranges. An average variogram calculated from previous studies can guide soil sampling. Expert elicitation of the variogram³⁷ is another possibility to obtain an a priori variogram in Bayesian geostatistics or soil sampling design optimization when budget constraints make the collection of new observations difficult. Although previous research has demonstrated how an average variogram or a variogram obtained with expert elicitation can be used to guide sampling³⁸, further research is needed to understand the effect of variogram estimation and variogram parameter uncertainty on spatially aggregated uncertainty.

The end-to-end DSM framework developed is scalable and can readily be applied beyond the field and project levels. Spatial aggregation of the mean and variance makes this approach particularly well suited to larger areas such as collections of projects, soil districts, or even nations, provided that adequate sampling or auxiliary information is available. Aggregation of the mean and variance also facilitates the detection of change. This scalability is particularly relevant for emerging national and regional soil carbon monitoring systems, which increasingly rely on the integration of field observations with remote sensing covariates and model-based approaches^{10,20}. In this context, the framework provides a transparent and reproducible way to quantify and report uncertainty, which is a critical requirement for credible MRV frameworks and for alignment with good practice guidance. It also offers practical value for project developers by enabling the assessment of trade-offs between sampling density, project duration, and acceptable levels of uncertainty deduction. Such features make the framework a useful component in the design of scalable, cost-effective, and policy-relevant MRV systems for soil carbon.

This study presented a DSM approach for MRV of SOC stocks, demonstrating its capacity to generate spatially explicit SOC estimates with quantifiable uncertainty. At the project level, the framework achieved

uncertainty of less than 1 Mg ha⁻¹ (e.g. 60.48 ± 0.51 Mg ha⁻¹), and while increasing the training sample size slightly reduced uncertainty, it had little effect on mean SOC stock estimates. We found that spatial aggregation of SOC estimates is particularly sensitive to variogram fitting, underscoring the need for careful parameter estimation. Simulations indicated that uncertainty deductions were 12% over five years in a 430-acre project, decreasing to about 5% with a 10-year monitoring period, and typically falling below 5% in projects larger than 50,000 acres or longer than five years. These findings highlight the scalability of the framework, with spatial averaging reducing uncertainty and improving the detectability of SOC stock changes. The approach can be applied broadly, provided relevant covariates and sampling are available, and the increasing availability of high-resolution remote sensing data further strengthens its potential for use across diverse regions.

Methods

Study area and SOC stocks data

We used the publicly-available dataset from Potash et al.³⁹. The dataset comprises 1,019 georeferenced measurements of SOC and bulk density from seven fields across two midwestern U.S. states. Soil samples were collected with three different sampling designs: gridded, gridded within subplots, and model-based, with an average sampling density of 7 soil samples per hectare. The samples were obtained as intact cores for various intervals between 5 and 30 cm. Bulk density was obtained by dividing the dry mass of soil in grams by its volume in cm³. The samples had no rock fragments. Total carbon concentration was obtained by dry combustion after removing the inorganic carbon for soils with pH >7.2, so the total carbon hereafter is equivalent to the organic carbon concentration. More information on the field protocol and laboratory analyses is provided in Potash et al.³⁹. SOC stocks were calculated by multiplying the bulk density by the SOC concentration and the thickness of the soil layer, and converting to appropriate units (e.g. Mg C ha⁻¹). All samples were depth aligned using simple weighted averaging to conform with a common depth interval of 0–30 cm. The final sampling locations with values of the SOC stocks are shown in Figure S1.

Environmental covariates

We collected a set of 79 environmental variables to be used as covariates in the model. The variables were chosen to reflect a suite of soil, climatic, biologic, or topographic factors known to impact biogeochemical cycling of SOC^{18,20}. In doing so, we not only include the key factors chosen in standard biogeochemical models, but also attributes that have been previously included in previous DSM frameworks. All covariates were resampled to conform with the same spatial resolution of 10 m × 10 m and projected to World Geodetic System 1984 ensemble UTM zone 16N (EPSG:32616) system. The list of covariates along with origin and associated reference, is provided in Table 2 with additional information in Table S1.

End-to-end approach

The end-to-end DSM approach that uses ATLAS-SOC to estimate the SOC stock mean and variance for a field and at project level followed the three steps described hereafter and illustrated in Fig. 7.

A model is fitted between measured SOC stock and a set of covariates. Covariate values for each SOC stock measurement are extracted from mapped covariate values. This follows the ordinary steps of DSM⁴⁰. The approach is extensible to multiple model types and architecture, including the averaging of multiple models, but for the purpose of this study, quantile regression forests (QRF) were used due to their simplicity.

Quantile regression forests^{41,42} was applied to establish the relationship between SOC stock and environmental covariates. QRF is a machine learning model based on an ensemble of decision trees. A single decision tree is built through binary recursive partitioning, where the calibration data, initially grouped in a single root node, are split based on all possible binary partitions. This process repeats for the newly created partitions until a minimum node size value is reached. The final prediction is the average of the values at the tree's terminal nodes. QRF further includes the bagging technique (i.e. bootstrap and aggregating) and random perturbations during tree splits, considering only a subset of covariates. In bagging, an ensemble of regression trees is built from bootstrap samples of the calibration data. The final prediction is obtained through averaging all predictions for the final output. Next to the prediction, QRF also provides an estimate of the uncertainty. This is obtained by the conditional cumulative probability distribution from the data, using the same weights as the QRF model, which allows for quantile estimates. From this conditional distribution, quantiles and variance statistics can be obtained.

The fitting of the QRF model relies on three user-defined parameters: the number of trees (ntree), the number of covariates to select at each split of the tree (Mtry), and the minimum node size. In our implementation, we set ntree to 150, as this provides a balance between model accuracy and computational efficiency. We use Mtry equal to the square root of the total number of covariates. The minimum node size is set to 5. The model is implemented using the ranger package in R⁴³.

The model predictions are assessed using validation statistics obtained with a cross-validation procedure. In principle, the approach is extensible to a variety of cross-validation procedures, and in this study, a leave-one-out (LOOCV) procedure was used. All samples except one were used for model fitting, while the single sample left out was used for validation. The process was repeated for all samples, each time leaving another sample out of the calibration sample without replacement. The validation statistics were calculated from the pairwise set of measured and predicted values. We calculated the mean error (ME), root mean square error (RMSE) and coefficient of determination or modelling efficiency (R²), defined as:

$$ME = \frac{1}{n} \sum_{i=1}^n z(s_i) - \hat{z}(s_i), \quad (1)$$

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

$$R^2 = 1 - \frac{\sum_{i=1}^n (z(s_i) - \hat{z}(s_i))^2}{\sum_{i=1}^n (z(s_i) - \bar{z})^2}, \quad (3)$$

where n is the validation sample size, z and $\hat{z}$ are the measured and predicted values, respectively, at location s_i ($i = 1, \dots, n; s_i \in A$), and A is

Table 2 | Covariates with associated reference

Covariate type	Examples	Source and references
Remote sensing vegetation indices	NDVI and NDTI metrics (e.g., summary statistics, trends, lags, residuals)	48,49
Remote sensing thermal metrics	Surface temperature statistics and trends (e.g., means, offsets, changes)	48
Climate variables	Precipitation and temperature summaries and extremes; drought indicators	50,51
Soil properties and biological indicators	Texture (clay, sand, silt), pH, microbial activity, biomass productivity, carbon use efficiency	52,53
Topographic indices	Elevation, topographic position indices at multiple spatial scales	54

Covariates represent typical inputs for digital soil mapping and were derived from publicly available remote sensing, climatic, edaphic, and topographic databases. More information on covariates is available in Kellner et al.²⁰. All covariates are in projected coordinates grid cells of 10 m × 10 m and coordinate system WGS84 UTM zone 16N (EPSG:32616) and extent: 223,843.7–224,563.9; 4,3275.62–4,328,052. NDTI Normalized Difference Tillage Index, NDVI Normalized Difference Vegetation Index.

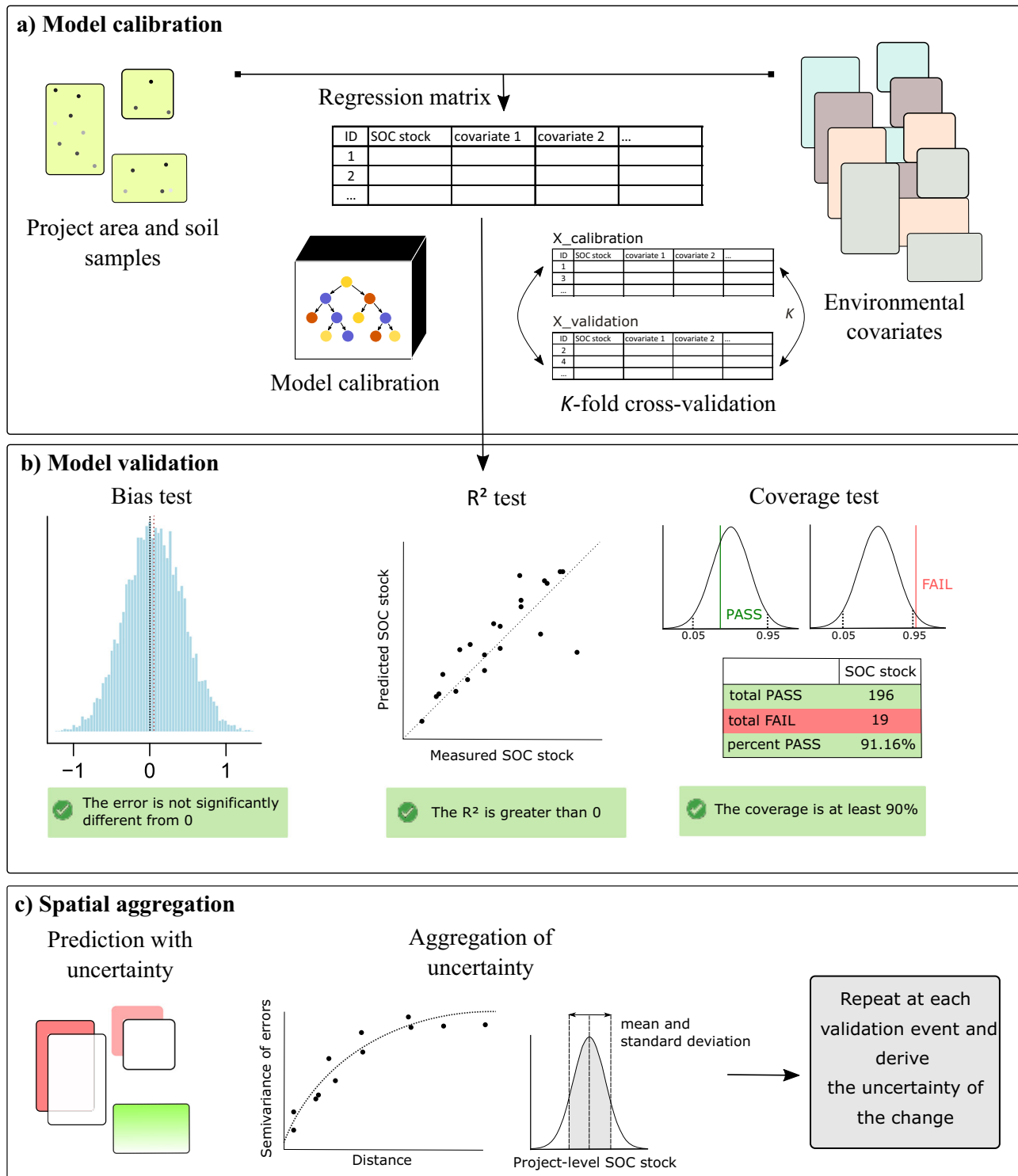


Fig. 7 | Framework for the end-to-end digital soil mapping approach. The end-to-end approach for estimating the project-level spatial average of SOC stocks along with the aggregated uncertainty. The approach is composed of (a) a model calibration where spatially-explicit covariates are overlaid with point estimates of SOC

stocks and a model is fitted and prediction errors obtained with cross-validation, b the predictions are evaluated with three tests that must be passed, and c) spatial aggregation of the mean and variance is made and can be repeated at each time step to estimate the variance of the change and uncertainty deduction.

the study area. The vector of measured values has mean $\bar{z}$. The ME represents the model bias, it has an optimal value of 0. The RMSE represents the magnitude of the error, also has an optimal value of 0 and is expressed in the unit of the measured value. The R² ranges between an optimal value of 1 and can be negative, in which case the mean of the measurements is a better predictor than the model. In addition to validating the predictions, we validate the uncertainty

estimates. For this, we use the prediction interval coverage probability (PICP⁴⁴). The PICP is the percentage of measured values covered by a prediction interval. For example, for the 90% prediction interval, a reliable estimate of uncertainty would entail that 90% of the predictions in the validation dataset fall within this interval. Values higher than the nominal prediction interval represent an overestimation of the uncertainty, whereas a smaller value indicates an underestimation.

A difficulty when calculating validation statistics is to decide whether a model is sufficient for the intended purpose. Following standards described in VCS Tool VT0014, we implemented the three following tests that a model should pass:

- The R^2 should be greater than 0. The justification for this threshold lies in having a model that explains at least some of the variance of the measurements and provides an improvement compared to using the mean of the validation data as predictor for any location in the study area.
- The model ME is not significantly different from 0 under a two-tailed, one-sample t-test with $\alpha = 0.05$.
- The coverage must be at least 90. The model must generate predictions that are within 90% prediction intervals $\geq 90\%$ of the time.

Estimation of the field and project average or total SOC stocks is straightforward with model predictions: this is done by averaging all predicted values within the field. Calculating the field or project-level variance is more complicated because spatial dependence requires the use of geostatistical methods. To obtain the field and project-level uncertainty we follow the procedure described in Wadoux & Heuvelink³³, repeated here briefly.

- Calculate the standardized prediction error at validation locations. The prediction error standard deviation is obtained with QRF for all locations in the field. At every measurement location, the standardized prediction error is obtained by dividing the prediction error obtained by cross-validation by the prediction error standard deviation.
- Estimate a correlogram of the standardized prediction error. This can be done by fitting first a variogram to the standardized prediction error and using the identity $\rho(h) = (sill - \gamma(h))/sill$, where $\rho(h)$ is the correlation function, h is the spatial distance, $\gamma(h)$ is the variogram and $sill$ is the estimated sill of the variogram.
- Obtain the field-average through integration of the variance of the spatial average of the point support prediction errors, see Eq. 7 in Wadoux & Heuvelink³³, defined by:

$$\text{var}\left(\frac{1}{|A|} \int_{s \in A} \varepsilon(s) ds\right) = \frac{1}{|A|^2} \int_{s \in A} \int_{u \in A} \sigma(s)\sigma(u)\rho(|s-u|) ds du, \quad (4)$$

where $\varepsilon(s)$ is the map error at locations s , $\sigma(s)$ and $\sigma(u)$ are the standard deviation at point s and u , and $\rho(|s-u|)$ is the value of the correlation function between the locations s and u . Solving the double integral can be done efficiently using Monte-Carlo integration, where the numerical integration over the two-dimensional area is made by selecting the locations s and u independently from a uniform distribution over A . The number of pairs selected in the Monte Carlo integration must be sufficiently large. This can be estimated with convergence statistics.

Sample size experiment

We further investigate the effect of different sample sizes on model calibration for the field and project-level mean and uncertainty. This is done by subsampling the existing data to various sample sizes using a stratified sampling procedure. Note here the subsampled locations are selected from the existing sample data. Hence, not every location within the area can be selected, as would be the case in a stratified random sampling. The merit and limitation of such an approach are presented in the Discussion. We created compact geographical strata with k -means clustering of the geographic coordinates using the *spcosa* package in R. A total of three compact strata were obtained per field. To test the effect of the sampling density on the uncertainty, we randomly selected various sample sizes per stratum, from 1 to 10 samples per stratum (i.e. from 3 to 30 samples per field). This represents a sampling density of about 0.12 to 1.2 samples per hectare. Because of

the randomness involved in the selection of the samples, this was repeated 100 times to obtain an average estimate of the field and project-level mean and variance.

The steps can be described as follows:

For each sample size per stratum ($n = 1, \dots, 10$), and iteration ($i = 1, \dots, 100$):

- Train QRF models for SOC stocks using the selected samples and environmental covariates.
- Select randomly with replacement from the existing locations the required sample size per stratum.
- Train QRF models for SOC stocks using the selected samples and environmental covariates.
- Validate the QRF model using leave-one-out cross-validation among the samples that were included in calibration.
- Pass the three validation tests described previously (i.e. bias test, R^2 test and 90% coverage test).
- Predict using the fitted QRF model at all locations in the area.
- Follow the steps for the spatial aggregation procedure described in Section “End-to-end approach”. Estimate a variogram per field on the standardized prediction error. The variograms are recalculated for each sample size. Since variograms cannot be calculated for very small sample sizes, all available sampling locations are used for variogram fitting. For example, for three samples per stratum, a model is fitted on those three samples per stratum, used to predict at all sampling locations within the field, and the resulting standardized error used to calculate the variogram. Note, however, that no cross-validation is made on the variograms. Variograms must be fit in an automated fashion, here using the *automap* R package for all 1000 variograms that need to be fit (100 iterations $\times$ 10 sample sizes).
- For every iteration, report the field and project-level mean and variance of the different sample sizes.

Calculating the uncertainty deduction

Current established standards require that projects account for the uncertainty in their estimates of greenhouse gas (GHG) benefits by deducting a conservative percentage from the calculated credit volume. The approach developed in Section “End-to-end approach” supports estimation of the variance of the spatially-aggregated predicted SOC stock at time t_0 , denoted $\text{var}(\text{SOC}_{t_0})$ hereafter. The approach can be extended to estimate the variance of the mean estimated SOC stock change $\text{var}(\Delta\text{SOC}_{t_0,t_{0+1}})$ between the two time steps t_0 and t_{0+1} :

$$\text{var}(\Delta\text{SOC}_{t_0,t_{0+1}}) = \text{var}(\text{SOC}_{t_0}) + \text{var}(\text{SOC}_{t_{0+1}}) - 2\rho \cdot \sqrt{\text{var}(\text{SOC}_{t_0})} \cdot \sqrt{\text{var}(\text{SOC}_{t_{0+1}})}, \quad (5)$$

where $2\rho \cdot \sqrt{\text{var}(\text{SOC}_{t_0})} \cdot \sqrt{\text{var}(\text{SOC}_{t_{0+1}})}$ is the covariance of the errors over time. In a geospatial context, ρ is the cross-correlation between the standardized prediction errors of the two time steps. This covariance term ensures that we are estimating the uncertainty of the change accurately. Although this term can be ignored, doing so would lead to an overestimation of the uncertainty and a greater uncertainty deduction. It can, however, be estimated using a (pseudo) cross-correlogram⁴⁵ with the same MC integration as done in this study.

The uncertainty deduction is generally applied in units of the carbon dioxide (CO_2) removal. Conversion of the mean and variance in SOC stocks to the unit of the carbon emission removal is done multiplying the SOC stock in Mg ha^{-1} by the conventional conversion factor of $44/12 \approx 3.67$. Accordingly, the mean and variance of the CO_2 removal is obtained as follows.

The mean emission removal:

$$\text{CO}_{2,t_0,t_{0+1}} = (\overline{\Delta\text{SOC}}_{t_0,t_{0+1}} - \overline{\Delta\text{SOC}}_{\text{baseline},t_0,t_{0+1}}) \cdot \frac{44}{12}. \quad (6)$$

The variance of the mean emissions removal:

$$\text{var}(\text{CO}_{2,t_0,t_{0+1}}) = (\text{var}(\overline{\Delta\text{SOC}}_{t_0,t_{0+1}}) + \text{var}(\overline{\Delta\text{SOC}}_{\text{baseline},t_0,t_{0+1}})) \cdot \left(\frac{44}{12}\right)^2, \quad (7)$$

where $\overline{\Delta\text{SOC}}_{\text{baseline},t_0,t_{0+1}}$ and $\text{var}(\overline{\Delta\text{SOC}}_{\text{baseline},t_0,t_{0+1}})$ are the mean and variance of the SOC stock change in the baseline control area between the two time steps. When no control sites are used, the mean is zero and the variance equals the variance of the change estimated with Eq. (5).

Assuming a normal distribution with mean and variance provided in Eqs. (6) and (7), respectively, a probability of exceedance method can be used for the uncertainty deduction, using a defined threshold in the estimated probability density function of the reduction removal source. Given the large sample size ($n = 1019$) and the use of geostatistical aggregation to estimate the variance of the project-area mean, the sampling distribution of the mean was assumed to be approximately normal. We used the probability of exceedance method described in VM0042 to calculate the number of creditable tons of CO_2 . The probability of exceedance method works using Eq. (7) to construct a probability distribution for the mean emissions removal estimate. The credit claim is the 33.3rd percentile of the probability distribution. Thus, greater uncertainty in the estimate of the mean emissions removal reduces the economic value of the credit claim, which can be as small as 0, even when the mean of the emissions removal estimate is positive.

As no data on SOC stock change was available in the current database, the uncertainty deduction is demonstrated with a scenario for a standard sequestration rate taken from Zomer et al.⁴, which reports a middle-of-the-road annual sequestration rate for North American agricultural land that is 0.6% of the stock size. We tested different monitoring periods with an annual sequestration rate that is exponential. The simulated sequestration was therefore calculated as $\text{SOC}_t = \text{SOC}_0 \times 1.006^t$, where t is the number of years, and SOC_0 is the soil organic carbon at the baseline. We tested project durations of 3, 5, and 10 years, which reflect typical MRV cycles⁴⁶ and correspond to timeframes in which SOC changes are generally detectable given turnover rates and spatial variability⁴⁷. The resulting uncertainty deduction is calculated.

The uncertainty deduction obtained at project level for the scenarios for SOC stock changes for 3, 5 and 10 years are compared to the deduction that would be obtained for larger projects. To obtain larger projects, we bootstrapped the seven fields of this study and created projects up to 100,000 acres. Using the correlation functions and properties of the individual fields, the aggregated average standard deviation and the uncertainty deduction was obtained for time steps of SOC stock change of 3, 5 and 10 years, for a correlation value of the error between the time steps of 0.5 and 0.9 (i.e. the variance of the change calculation). We also calculated the uncertainty deduction for smaller sequestration rates, using rate values of 0.1, 0.2, 0.3 and 0.4% of the stock size.

Data availability

The datasets generated and analysed during the current study are available from the study of Potash et al.³⁹ and stored in the repository address: <https://doi.org/10.6084/m9.figshare.23669304.v2>.

Code availability

The underlying code for this study is not publicly available. A reproducible example of spatial aggregation of the uncertainty is provided in <https://github.com/AlexandreWadoux/UncertaintySpatialAggregation>.

Received: 19 July 2025; Accepted: 4 January 2026;

Published online: 06 July 2026

References

1. Friedlingstein, P. et al. Global carbon budget 2022. *Earth Syst. Sci. Data* **4**, 4811–4900 (2022).
2. Eswaran, H., Van Den Berg, E. & Reich, P. Organic carbon in soils of the world. *Soil Sci. Soc. Am. J.* **57**, 192–194 (1993).
3. Batjes, N. H. Total carbon and nitrogen in the soils of the world. *Eur. J. Soil Sci.* **47**, 151–163 (1996).
4. Zomer, R. J., Bossio, D. A., Sommer, R. & Verchot, L. V. Global sequestration potential of increased organic carbon in cropland soils. *Sci. Rep.* **7**, 15554 (2017).
5. Sanderman, J., Hengl, T. & Fiske, G. J. Soil carbon debt of 12,000 years of human land use. *Proc. Natl. Acad. Sci. USA* **114**, 9575–9580 (2017).
6. Smith, P. et al. Greenhouse gas mitigation in agriculture. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 789–813 (2008).
7. Minasny, B. et al. Soil carbon 4 per mille. *Geoderma* **292**, 59–86 (2017).
8. European Commission. Proposal for a regulation of the European Parliament and of the Council establishing a union certification framework for carbon removals (2022).
9. Barbato, C. T. & Strong, A. L. Farmer perspectives on carbon markets incentivizing agricultural soil carbon sequestration. *npj Clim. Action* **2**, 26 (2023).
10. Batjes, N. H. et al. Towards a modular, multi-ecosystem monitoring, reporting and verification (MRV) framework for soil organic carbon stock change assessment. *Carbon Manag.* **15**, 2410812 (2024).
11. So, I., Haya, B. & Elias, M. Voluntary Registry Offsets Database, Berkeley Carbon Trading Project (University of California, 2023).
12. Brummitt, C. D. et al. Solutions and insights for agricultural monitoring, reporting, and verification (MRV) from three consecutive issuances of soil carbon credits. *J. Environ. Manag.* **369**, 122284 (2024).
13. Gruijter, J. J., Bierkens, M. F. P., Brus, D. J. & Kotters, M. Sampling for Natural Resource Monitoring (Springer, 2006).
14. Bradford, M. A. et al. Testing the feasibility of quantifying change in agricultural soil carbon stocks through empirical sampling. *Geoderma* **440**, 116719 (2023).
15. Stanley, P., Spertus, J., Chiartas, J., Stark, P. B. & Bowles, T. Valid inferences about soil carbon in heterogeneous landscapes. *Geoderma* **430**, 116323 (2023).
16. Heuvelink, G. B. M. et al. Machine learning in space and time for modelling soil organic carbon change. *Eur. J. Soil Sci.* **72**, 1607–1623 (2021).
17. Yang, R.-M. et al. The effectiveness of digital soil mapping with temporal variables in modeling soil organic carbon changes. *Geoderma* **405**, 115407 (2022).
18. Sothe, C., Gonsamo, A., Arabian, J. & Snider, J. Large scale mapping of soil organic carbon concentration with 3d machine learning and satellite observations. *Geoderma* **405**, 115402 (2022).
19. Helfenstein, A., Mulder, V. L., Heuvelink, G. B. M. & Hack-ten Broeke, M. J. D. Three-dimensional space and time mapping reveals soil organic matter decreases across anthropogenic landscapes in the Netherlands. *Commun. Earth Environ.* **5**, 130 (2024).
20. Kellner, J. R. et al. Digital soil mapping in support of voluntary carbon market programs in agricultural land. *PLoS ONE* **20**, e0327895 (2025).
21. Clean Energy Regulator. Estimating soil organic carbon sequestration using measurement and models method. Australian Carbon Credit Unit Scheme (2025). Retrieved July 18, 2025, from Clean Energy Regulator website: <https://cer.gov.au/>.
22. Nussbaum, M., Papritz, A., Baltensweiler, A. & Walthert, L. Estimating soil organic carbon stocks of Swiss forest soils by robust external-drift kriging. *Geosci. Model Dev.* **7**, 1197–1210 (2014).
23. Hengl, T., Heuvelink, G. B. M. & Rossiter, D. G. About regression-kriging: From equations to case studies. *Computers Geosci.* **33**, 1301–1315 (2007).

24. Wadoux, A. M. J.-C. & Heuvelink, G. B. M. Scientists yet to consider spatial correlation in assessing uncertainty of spatial averages and totals. *Int. J. Appl. Earth Observ. Geoinf.* **139**, 104472 (2025).
25. Somarathna, P. D. S. N., Minasny, B. & Malone, B. P. More data or a better model? Figuring out what matters most for the spatial prediction of soil carbon. *Soil Sci. Soc. Am. J.* **81**, 1413–1426 (2017).
26. Khan, A., Aitkenhead, M., Stark, C. R. & Jorat, M. E. Optimal sampling using Conditioned Latin Hypercube for digital soil mapping: an approach using Bhattacharyya distance. *Geoderma* **439**, 116660 (2023).
27. Sun, X.-L. et al. Variability of soil mapping accuracy with sample sizes, modelling methods and landform types in a regional case study. *Catena* **213**, 106217 (2022).
28. Ma, T., Brus, D. J., Zhu, A.-X., Zhang, L. & Scholten, T. Comparison of conditioned Latin hypercube and feature space coverage sampling for predicting soil classes using simulation from soil maps. *Geoderma* **370**, 114366 (2020).
29. Du, L. et al. Spatial extrapolation of topographic models for mapping soil organic carbon using local samples. *Geoderma* **404**, 115290 (2021).
30. Schmidinger, J. et al. Effect of training sample size, sampling design and prediction model on soil mapping with proximal sensing data for precision liming. *Precis. Agric.* **25**, 1529–1555 (2024).
31. Webster, R. & Oliver, M. A. Sample adequately to estimate variograms of soil properties. *J. Soil Sci.* **43**, 177–192 (1992).
32. Kerry, R. & Oliver, M. A. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma* **140**, 383–396 (2007).
33. Wadoux, A. M. J.-C. & Heuvelink, G. B. M. Uncertainty of spatial averages and totals of natural resource maps. *Methods Ecol. Evol.* **14**, 1320–1332 (2023).
34. Oliver, M. A. & Webster, R. Basic Steps in Geostatistics: the Variogram and Kriging, vol. 106 (Springer, 2015).
35. McBratney, A. A. & Pringle, M. J. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. *Precis. Agric.* **1**, 125–152 (1999).
36. Kerry, R. & Oliver, M. A. Average variograms to guide soil sampling. *Int. J. Appl. Earth Observ. Geoinf.* **5**, 307–325 (2004).
37. Truong, P. N., Heuvelink, G. B. M. & Gosling, J. P. Web-based tool for expert elicitation of the variogram. *Computers Geosci.* **51**, 390–399 (2013).
38. Wadoux, A. M. J.-C., Marchant, B. P. & Lark, R. M. Efficient sampling for geostatistical surveys. *Eur. J. Soil Sci.* **70**, 975–989 (2019).
39. Potash, E. et al. Multi-site evaluation of stratified and balanced sampling of soil organic carbon stocks in agricultural fields. *Geoderma* **438**, 116587 (2023).
40. McBratney, A. B., Santos, M. M. & Minasny, B. On digital soil mapping. *Geoderma* **117**, 3–52 (2003).
41. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
42. Meinshausen, N. Quantile regression forests. *Journal of Machine Learning Research* **7** (2006).
43. Wright, M. N. & Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**, 1–17 (2017).
44. Goovaerts, P. Geostatistical modelling of uncertainty in soil science. *Geoderma* **103**, 3–26 (2001).
45. Wackernagel, H. Multivariate Geostatistics: An Introduction with Applications (Springer Science & Business Media, 2003).
46. Smith, P. et al. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Change Biol.* **26**, 219–241 (2020).
47. Smith, P. How long before a change in soil organic carbon can be detected?. *Glob. Change Biol.* **10**, 1878–1883 (2004).
48. Crawford, C. J. et al. The 50-year Landsat Collection 2 archive. *Sci. Remote Sens.* **8**, 100103 (2023).
49. Jakubauskas, M. & Legates, D. R. Harmonic analysis of time-series AVHRR NDVI data for characterizing US Great Plains land use/land cover. *Int. Arch. Photogramm. Remote Sens.* **33**, 384–389 (2000).
50. Daly, C., Taylor, G. H. & Gibson, W. P. The PRISM approach to mapping precipitation and temperature. In *Proc., 10th AMS Conf. on Applied Climatology*, 20–23 (1997).
51. Leeper, R. D. et al. Characterizing US drought over the past 20 years using the US Drought Monitor. *Int. J. Climatol.* **42**, 6616–6630 (2022).
52. Poggio, L. et al. Soilgrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240 (2021).
53. Fu, P. et al. Accurate quantification of 0–30 cm soil organic carbon in croplands over the Continental United States using machine learning. *Remote Sens.* **16**, 2217 (2024).
54. European Space Agency. Copernicus DEM-Global and European Digital Elevation Model (COP-DEM) (2021).

Author contributions

A.M.J.-C.W. and J.R.K. conceived and designed the study. J.M. and P.F. performed data processing and analysis. D.S., R.R., M.D., and N.R.L. contributed to scientific interpretation, methodology development, and manuscript editing. All authors discussed the results and contributed to the final version of the manuscript.

Competing interests

Perennial Climate Inc. has filed two provisional patent applications currently under consideration, one of which is based on this research. The company provides digital soil mapping as a service and markets products related to this research. There are no additional patents, products in development, or marketed products associated with this work to declare. All authors were employees and/or shareholders of Perennial Climate Inc. during the completion of this work, with the exception of AW, who provided contributions as an independent, compensated consultant. J.R.K. completed this work as Chief Scientist of Perennial Climate Inc. in compliance with the Brown University Policy on Outside Professional Activities for Faculty.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44264-026-00125-0>.

Correspondence and requests for materials should be addressed to James R. Kellner.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026