Highlights

Predicting Spatiotemporal Soil Organic Carbon Responses to Management Using EPICIIASA Meta-Models

Tara A. Ippolito,Juraj Balkovič,Rastislav Skalsky,Christian Folberth,Tamás Krisztin,Jason Neff

- Meta-model approach used to simplify process based crop model
- Exploration of the determinants of soil organic carbon responses to management
- Meta-model predictions validated with regional and local data
- Meta-models largely capture broad dynamics, underestimate magnitude of responses
- More accurate input data, calibration, and validation needed to improve accuracy

Predicting Spatiotemporal Soil Organic Carbon Responses to Management Using EPIC-IIASA Meta-Models

Tara A. Ippolito[a,*], Juraj Balkovič[b], Rastislav Skalsky[b], Christian Folberth[b], Tamás Krisztin[b,c] and Jason Neff[a]

[a]*University of Colorado, Department of Environmental Studies, 4001 Discovery Drive, Boulder, 80309, CO, U.S.A.*

[b]*International Institute for Applied Systems Analysis, Biodiversity and Natural Resources Program, Schlossplatz 1, Laxenburg, A-2361, Austria*

[c]*Paris Lodron University of Salzburg, Department of Economics, Kapitelgasse 4-6, Salzburg, A-5020, Austria*

ARTICLE INFO

Keywords:

Crop Management

EPIC-IIASA

Gridded Agronomic Modeling

Soil Organic Carbon

ABSTRACT

The management of Soil Organic Carbon (SOC) is a critical component of both nature-based solutions for climate change mitigation and global food security. Agriculture has contributed substantially to a reduction in global SOC through cultivation, thus there has been renewed focus on management practices which minimize SOC losses and increase SOC gain as pathways towards maintaining healthy soils and reducing net greenhouse gas emissions. Mechanistic models are frequently used to aid in identifying these pathways due to their scalability and cost-effectiveness. Yet, they are often computationally costly and rely on input data that are often only available at coarse spatial resolutions. Herein, we build statistical meta-models of a multifactorial crop model in order to both (a) obtain a simplified model response and (b) explore the biophysical determinants of SOC responses to management and the geospatial heterogeneity of SOC dynamics across Europe. Using 5,600 unique simulations of crop growth from the gridded Environmental Policy Integrated Climate-based Gridded Agricultural Model (EPIC-IIASA GAM) covering 86,000 simulation units across Europe, we build multiple polynomial regression ensemble meta-models for unique combinations of climate and soil across Europe in order to predict SOC responses to varying management intensities. We find that our biophysically-explicit meta models are highly accurate (R^2 = .97) representations of the full mechanistic model and can be used in lieu of the full EPIC-IIASA GAM model for the estimation of SOC responses to cropland management. Model stratification by means of climate and soil clustering improved the performance of the meta-models compared to the full EU-scale model. In regional and local validations of the meta-model predictions, we find that the meta-models largely capture broad SOC dynamics such as the linear nature of SOC responses to residue application, yet they often underestimate the magnitude of SOC responses to management. Furthermore, we find notable differences between the results from the biophysically-specific models throughout Europe, which point to spatially-distinct SOC responses to management choices such as nitrogen fertilizer application rates and residue retention that illustrate the potential for these models to be used for future management applications. While more accurate input data, calibration, and validation will be needed to accurately predict SOC change, we demonstrate the use of our meta-models for biophysical cluster and field study scale analyses of broad SOC dynamics with basically zero fine-tuning of the models needed. This work provides a framework for simplifying large-scale agricultural models and identifies the opportunities for using these meta-models for assessing SOC responses to management at a variety of scales.

1. Introduction

Soil Organic Carbon (SOC) contains a substantial portion of global carbon stocks with roughly 1,500-2,400 Gt C (5500-8800 Gt CO2) globally (Sanderman et al., 2017). Agriculture has historically contributed to a loss in SOC, primarily through conversion of native soil to agricultural uses as soil carbon stocks decline substantially when cropland replaces native forest (-42%) and pasture (-59%) (Guo and Gifford, 2002). Once soils have been converted for cultivation, further soil carbon losses may occur as a result of management practices such as tillage, choice of crop, and nutrient inputs (Lal, 2004). Although agricultural soils are often a source of carbon emissions, they can also serve as sinks for atmospheric CO2, depending on the interaction of factors such as soil properties, climate, and management

*Corresponding author taip7142@colorado.edu (T.A. Ippolito)

ORCID(s): 0000-0003-0618-9249 (T.A. Ippolito); 0000-0003-2955-4931 (J. Balkovič); 0000-0002-0983-6897 (R. Skalsky); 0000-0002-6738-5238 (C. Folberth); 0000-0002-9241-8628 (T. Krisztin); 0000-0002-8290-1472 (J. Neff)

choices (Eglin et al., 2010). There has been a renewed focus on global soil carbon sequestration as global policy makers attempt to mitigate climate change effects. A recent analysis suggests a soil C sequestration annual technical potential of .79-1.54 Gt CO_2/year (Amelung et al., 2020). Smith et al. (2020) highlighted the importance of improving our understanding of how SOC changes are influenced by climate, land use, management and edaphic factors as these processes and properties control the mechanisms of SOC changes (Attard et al., 2016). As a result, there is a great deal of complexity and spatial variability in potential SOC changes especially as it relates to the effects of management practices. One way to better understand the complexity and spatial variability of SOC changes is through the use of process-based models.

Process-based models, which simulate daily crop growth and SOC dynamics under a variety of conditions, are frequently used for a wide range of applications such as assessment of policy goals (Minasnyet al.,2017) and estimation of landscape-scale SOC dynamics (Pennock and Frick, 2001). There are a variety of models and SOC quantification methodologies which vary in the processes they represent and how they are structured mathematically (Manzoni and Porporato, 2009; Whittaker et al., 2013). Among others, gridded agricultural models (GAM), including the gridded model EPIC-IIASA (Balkovič et al., 2014), have been evaluated as tools for agriculture sector assessments at large scales, including globally (Jägermeyr et al., 2021; Müller et al., 2017). These models are increasingly used in the EU-scale assessments to support land use policies, such as carbon emissions and removals from land use and land use change (Frank et al., 2015; Petrescu et al., 2021). One major benefit of using GAMs like EPIC-IIASA for estimation of SOC changes is the ability to simulate both existing and potential agricultural practices across large areas. As the effects of climate change continue to affect farmers, management practices may change to reflect new growing conditions (Iglesias et al., 2012). For example, a northward shift in cultivable zones may allow farmers in some areas of Europe to grow a wider variety of crops while other areas of Europe may see a reduction in crop growing potential of certain varieties (Ewert et al., 2005). Models like GAMs allow for the investigation of such scenarios and the SOC dynamics of these novel management choices.

While there are many benefits to using process-based models, the robustness and accuracy of these models are limited by the availability of reliable calibration and validation data and by the structural representation of the processes within the model (Keel et al., 2017; Toudert et al., 2018; Jones et al., 2017). Some process-based models, such as reduced-form models, are relatively simplistic and require less data than others depending on the structure or the application (Jones et al., 2017, 1999). A model's structure, computational complexity, and data requirements are often dependent on factors such as the purpose of the model and the data available for calibration and validation (Jones et al., 2017). Differences in the structures of process-based models can result in different responses when used to simulate the same experiments (Jones et al., 2017). Additionally, GAM applications are often computationally costly (Khabarov et al., 2020) and are limited by accuracy of often granular spatial inputs including initial SOC stock and crop management (Balkovič et al., 2020). In order to navigate these limitations, users sometimes build meta-models which train statistical models on simulation data output from process-based models in order to emulate process-based models in a simplified format.

Statistical meta-models take complex, spatially-explicit simulations and generate a simplified framework which distills complex interactions. These meta-models can be used to identify relationships of interest and the characteristics which drive these relationships. Meta-models also help reduce the substantial data requirements and computational cost of running, calibrating, and validating a complex process-based model, while preserving their robust scientific capabilities. These meta-models, which so far focus foremost on crop yield responses, have been utilized to evaluate differences between process-based models (Ringeval et al., 2021), explore yield responses to climate change (Blanc, 2017; Franke et al., 2020; Oyebamiji et al., 2015), and downscale process-based model yield estimates (Folberth et al., 2019). A meta-model framework that allows parameterization for different regions, while still accounting for the main biophysical and management determinants of SOC dynamics, would be a significant improvement for regional SOC modeling. This framework could be used to help identify the relationship between management interventions and SOC outcomes and the biophysical conditions under which SOC may be the most responsive to interventions.

In this paper, we present (1) the development of statistical meta-models built on a multifactorial implementation of EPIC-IIASA GAM for predictions of SOC responses to management change, climate, and soil properties, (2) the evaluation of the meta-models' SOC change estimation across regional biophysical clusters (climate x soil zones) in the EU and against field study scale results across Europe, and (3) a demonstration of the utility of this meta-modeling approach in identifying management strategies which may have the best outcomes for increasing SOC and the soil and climate conditions where these strategies may be the most effective.

2. Methodology and Data

2.1. Methodology Overview

The framework we utilize to build the statistical meta-models consists of 4 major components: (i) multifactorial gridded EPIC-IIASA modeling (referred to as EPIC Hypercube), (ii) biophysical clustering of spatially-explicit, gridded simulation units (SimU; homogenous simulation units are unique combinations of soil properties, weather, topography and management at 1km resolution) across Europe, (iii) cluster-specific regression meta-models, and (iv) evaluations at the level of biophysical cluster and field studies. We then demonstrate the use of our meta-models in identifying relationships between management interventions and changes in SOC and highlight the soil and climate conditions where these interventions may be the most effective.

2.2. Multifactorial Gridded EPIC-IIASA Modeling

2.2.1 Model Description

The Environmental Policy Integrated Climate (EPIC) Model is a process-based field scale model which simulates, with a daily time step, crop growth, hydrological, nutrient and carbon cycling, soil temperature and moisture, soil erosion and plant environment control under a wide range of crop management options such as tillage, fertilization, irrigation, pesticides, and liming (Izaurralde et al., 2006; Williams and Singh, 1995). In EPIC, the coupled organic C and nitrogen (N) module (Izaurralde et al., 2006) calculates transformations of five organic matter compartments as regulated by the soil environment, including soil moisture, temperature, oxygen, tillage, lignin content, and N supply. The EPIC model has been used in a variety of studies investigating soil organic matter cycling (Izaurralde et al., 2006), irrigation timing (Bryant et al., 1992), and the impact of climate change on the agricultural sector (White et al., 2011) among others (Gassman et al., 2004).

The EPIC-IIASA GAM (the EU version) was built by coupling EPIC (v. 0810) with EU-wide datasets on land cover (cropland), soils, topography, territorial units (NUTS2), and crop management practices aggregated at a 1x1 km grid covering European countries (Balkovič et al., 2013, 2018). All homogeneous gridded areas, i.e. spatially-unique simulation units (SimU), are assigned with “dominant” farmland fields, cropping systems, and crop management (Skalsky et al.` , 2008). The EPIC-IIASA GAM is one of 14 models included in the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) which provides ensemble projections of climate change impacts on agriculture (Jägermeyr et al., 2021; Warszawski et al., 2014).

2.2.2 Multifactorial Modeling Framework

We designed a multifactorial EPIC-IIASA simulation across 86,000 gridded SimUs covering Europe and a variety of crop management factors, we refer to the collection of multifactorial simulations as the EPIC Hypercube. For each SimU, we model annual crop yield (in tDM/ha), crop residue carbon (in kgC/ha) and annual change in topsoil SOC (tC/ha/year) for the time period 1980-2019. For each SimU x year, we simulate a factorial combination of the following management factors:

1. Crop type of maize, rape, rice, barley, soya, sunflower, rye, wheat
2. Maximum annual nitrogen fertilizer application of 0, 50, 100, and 250 kgN/ha, plus a crop-specific business-asusual (BAU) N application rate (see Balkovič et al. (2013))
3. Retention of 0, 30, 60, and 90% of crop residues at harvest

We discard the first 5 years (1980-1985) of simulation data to allow for model spin-up. Across the 86,000 SimU, 35 years, 8 crop types, 5 nitrogen application rates, and 4 residue management scenarios, the multifactorial simulation results in over 481 million unique crop growth records across Europe. For detailed information on the EPIC Hypercube input data, including climate, soil, and crop management parameters, please refer to the appendix.

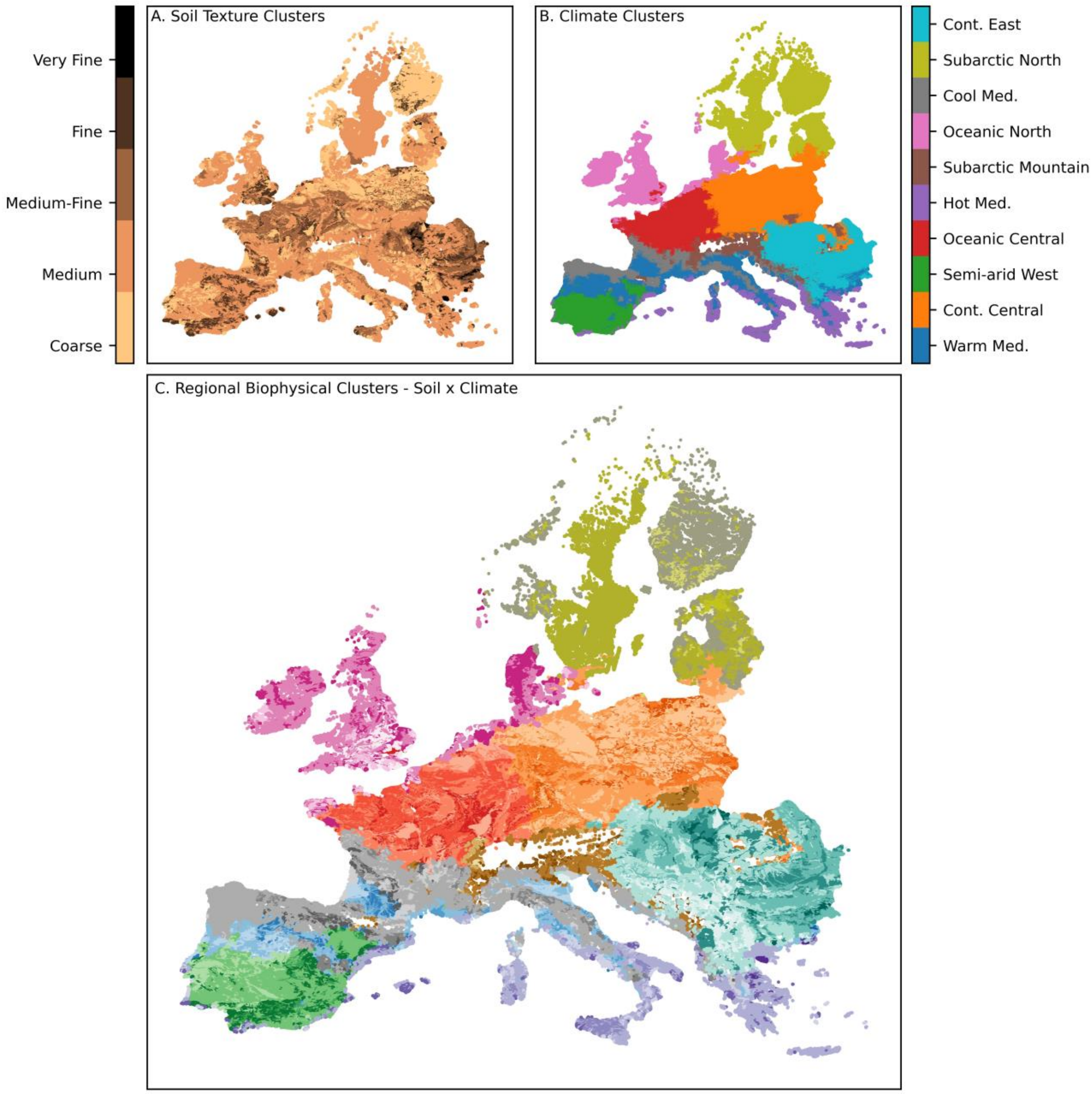

Figure 1: a) Soil clusters of SimU across Europe - clusters consist of 5 discrete classes based on soil texture defined by EPIC-IIASA. b) Climate clusters of SimU across Europe - clusters consist of 10 discrete classes based on 40 years of monthly climate data. "Cont" is an abbreviation for "Continental" and "Med" is an abbreviation for "Mediterranean. c) Regional biophysical clusters generated as a product of soil clusters and climate clusters

2.3. Biophysical Clustering of Simulation Units

We generate regional biophysical clusters of all 86,000 SimU across Europe using a combination of a-priori and unsupervised clustering techniques (Ding and He, 2004). We consider soil properties and climate properties as the two main components of biophysical characterization. Soil clusters and climate clusters are produced separately, and then combined, so that each of the final biophysical clusters is a unique combination of distinct soil and climate conditions. From the 5 SimU soil clusters and 10 SimU climate clusters, there are a total of 50 potential regional biophysical clusters of which 43 are present in the SimU (Fig. 1). Biophysical clusters with less than 1000 simulations were omitted from analysis due to concerns of model overfitting which are raised when there is a lack of sufficient data.

2.3.1. Soil Clustering

We select soil clusters (Fig. 1) a-priori using the EPIC-IIASA defined texture classification. There are 6 discrete soil texture classes within EPIC-IIASA - coarse, medium, medium-fine, fine, very fine, and peat. Peat was excluded from this analysis. Each simulation unit has one soil texture classification.

2.3.2. Climate Clustering

We discover climate clusters (Fig. 1) using an unsupervised machine learning approach - K-means clustering via Principal Component Analysis (PCA) - which has been used for a variety of applications such as clustering of DNA gene expressions and internet news articles (Ding and He, 2004). The climate data used for the climate clusters is the same data which is used as input to the EPIC Hypercube, but aggregated from daily measurements to monthly averages and sums depending on the variable. Climate parameters used in the clustering algorithm include monthly precipitation, temperature minimum and maximum, relative humidity, solar radiation, and potential evapotranspiration over all years (1980-2019). Due to the large number of climate parameters (6 variables x 12 months/year x 40 years), we use PCA as a dimensionality reduction technique to minimize the size of the data while preserving a large amount of information. We calculate principal components and retain those which explain a combined minimum of 90% of the variance in the climate data. We then use the PCA-reduced climate data as input to a K-means clustering algorithm (Pedregosa et al., 2011). The K-means clustering algorithm aims to find cluster centroids by optimizing a squared error function, typically the sum of the squared Euclidean distances (Likas et al., 2003). We tested k = 2 to k = 12 clusters and selected the optimal number of clusters based on a variety of statistical metrics (Inertia, Caliński-Harabasz Score, Davies-Bouldin Score, etc.) in combination with manual inspection and comparison to well known climate classifications such as Köppen-Geiger Climate Classification (Caliński and Harabasz, 1974; Davies and Bouldin, 1979; Peel et al., 2007). We find the optimal number of climate clusters to be 10. We use descriptive labels for the climate clusters in order to refer to them throughout our work - these labels were decided ad-hoc using reference to common climate classifications.

2.4. EPIC-IIASA STAT: Cluster-Specific Regression Meta-Models

2.4.1 Machine-Learning Framework

We build a separate regression meta-model of the EPIC Hypercube for each regional biophysical cluster x crop˙ type in order to investigate SOC responses to management which are unique to given soil and climate conditions. The collection of these statistical meta-models, which together cover all of Europe, is referred to as EPIC-IIASA STAT. By building these meta-models we explore the nature of SOC dynamics under a wide-variety of management choices and highlight the soil and climate conditions where management interventions may be the most effective. We chose to separate out crop types to avoid tangling the signals of different crops as crops may respond to management in unique ways. In EPIC-IIASA STAT, we predict 35-year long term averages of mean annual change in SOC. Since SOC changes occur slowly, long-term averages are more appropriate for our investigation.

For each biophysically-specific meta-model within EPIC-IIASA STAT, we utilize a bagging meta-estimator of multiple polynomial regression models together with inputs and outputs of the EPIC Hypercube simulations (Section 2.2.2). This allows us to build a robust, yet highly interpretable machine learning model for prediction of mean annual change in SOC within a cluster. Ensemble methods such as this combine a large number of base-estimators built with a specific algorithm (here, a multiple polynomial regression model) in order to improve on the performance of a model built with a single estimator. In order to avoid overfitting and the challenges of multicollinearity we utilize n-fold cross validation, L2 regularization, out-of-bag error monitoring, and testing on the 25% of observations held out from the training set (Chong and Jun, 2005; Kohavi et al., 1995). The EPIC Hypercube simulation data for each cluster was randomly split into training (75% of data) and testing (25% of data) sets. The number of simulations used for meta-model building varied by cluster and ranged between 1,222 and 139,003 simulations. Using the same methodology, we also build a meta-model for each crop type without biophysicalspecification (i.e. using all SimU) to test the value of model stratification by soil and climate parameters. In this Europe-wide meta-model, we split the entire EPIC Hypercube into training (75% of data) and testing (25% of data) sets.

2.4.2 EPIC-IIASA STAT Structure

The bagging estimator of multiple polynomial regression models is an ensemble method which estimates a number of base estimators and combines these base estimators to form each meta-model within EPIC-IIASA STAT. The multiple polynomial regression base includes linear and quadratic terms for each parameter, as well as interaction terms between parameters (Eq. 1).

$$\mathrm{f}(x_i) = \alpha + \sum_{j=0}^{J} \beta_j x_j + \sum_{l=0}^{J} \beta_l x_l^2 + \sum_{j \neq l}^{J} \beta_{jl} x_j x_l \tag{1}$$

$$y_i = \mathrm{f}(x_i) + \varepsilon_i \tag{2}$$

Where $\mathrm{f}(x_i)$ is the predicted variable (with i = 1,...,n), x_j (with j = 1,...,J) are the explanatory variables, α is the intercept, β_j are the slope coefficients associated with linear terms, β_l are the slope coefficients associated with quadratic terms, β_{jl} are the slope coefficients with the interaction terms (when $j \neq l$). ε_i is an iid error term from Gaussian distribution with zero mean and $\sigma 2$ variance.This structure was selected for its interpretability and the ability to explore response relationships of interest from the learned model. Our bagging estimator, built with scikit learn, fits 20 base estimators on randomly selected subsets of the training data and then combines the predictions of all estimators into each final meta-model (Breiman, 1996). We tested up to 60 base estimators and found marginal improvements in meta-model performance past 20 base estimators. For discussion regarding the model structure, please see the appendix.

2.4.3. EPIC-IIASA STAT Features

Features which are used to train each EPIC-IIASA STAT meta-model are based on inputs and outputs of EPIC Hypercube simulations. We use a limited set of climate, soil, and management features which are of interest in exploring the biophysically-determined SOC response to management (Table 1). For climate data, we engineer a number of growing season (GS) specific climate variables following the framework of Folberth et al. (2019). For soil data, we engineer full-profile parameters as a weighted sum of topsoil- and subsoil-specific variables. Applied nitrogen fertilizer is calculated as the sum of all applied nitrogen throughout the season as dictated by the automatic application mechanism in EPIC. Applied nitrogen fertilizer does not exceed the specified maximum annual N rate which is specified in the multifactorial simulation. Aboveground residue C is calculated by subtracting the C in dead roots and the C in residues from yield harvest losses from the total amount of residue C - we consider added aboveground residues as the residue treatment. For our target variable we use EPIC-IIASA GAM predicted mean annual change in topsoil organic C (Δ SOC), calculated as the average month to month difference in organic carbon in the 0-15 cm plowing depth over a given year. If mean Δ SOC is positive, soils absorbed C throughout the year, if negative, soils lost C throughout the year. As a final step, we collapse data into 35-year long term averages. For each unique simulation design - SimU x crop type x N application rate x residue management - all features, including soil, climate, management, and target variables, are averaged. Since averaging across 35 years dampens the variability of parameters, we also calculate long-term standard deviations of temporally dynamic features (all features except soil variables) to retain information on variability which may be important to SOC dynamics. Thus each data point used to train the model is a long-term record of a given SimU and cropping scenario.

Table 1
EPIC-IIASA STAT Meta-model input variables

Variable Type	Variable (Unit)
Soil	Profile Sand Content (%)
	Profile Silt Content (%)
	Organic C in topsoil (%)
	Profile Field Water Capacity at 33 kPa (□□^{3}/□□^{3})
Climate	GS Precipitation Sum (mm)
	GS Precipitation Skew (mm)
	GS Temperature Mean (□°)
	GS Temperature Skew (□°)
	GS Radiation Mean (MJ/□^{2})
	GS Radiation Skew (MJ/□^{2})
	GS Potential Evapotranspiration Mean (mm)
Management	Applied nitrogen Fertilizer (FTN, kgN/ha/yr)
	Aboveground residue C content added to the soil (RSDCa, kgC/ha/yr)
Target	Mean Annual Change in SOC (Δ SOC, tC/ha)

2.5. Evaluating EPIC-IIASA STAT

2.5.1. Feature Importance

Weimplement(a)leave-one-outand(b)featuresubsetselectionstrategiesinordertoidentifythefeatureimportance within each meta-model (James et al., 2013). In the leave-one-out feature selection strategy, we test the importance of each feature individually by training the model using all features except the feature of interest. The change in model performance, measured by change in R^2 value, resulting from the exclusion of the feature is the measured importance of that feature. Larger performance losses (negative change in R^2) indicate a feature is important to the accuracy of the model. The main challenge of using this methodology is the effect of multicollinearity on measured feature importance. If there are highly collinear variables in the model, removing one of the collinear variables may not result in a substantial decrease in model performance. Since many features in this model such as climate and soil parameters co-vary, we also implement a feature subset selection where we remove all soil parameters, all climate parameters, all nitrogen management parameters, all residue management parameters, and all nitrogen and residue management parameters from the model to test the importance of these groupings of variables.

2.5.2. Performance of EPIC-IIASA STAT in replicating modeled SOC dynamics

We evaluate each meta-model with the held out testing data (25% of observations) to assess the accuracy of EPICIIASA STAT in predicting EPIC-IIASA mean Δ SOC. We also benchmark each biophysically-specific meta-model against a regression meta-model built on the full set of simulation data to investigate the achieved benefit of soil and climate clustering. The coefficient of determination, □^{2}, was calculated by (Eq. 3) as

$$R^2 = (1 - \frac{u}{v}) \tag{3}$$

Where $u = \sum_{i=1}^{n} \left(\Delta SOC_{true,i} - \Delta SOC_{prediction,\, i}\right)^2 = residual\ sum\ of\ squares$
and $v = \sum_{i=1}^{n} \left(\Delta SOC_{true,i} - \mu_{SOC_{true}}\right)^2 = total\ sum\ of\ squares$

Mean Absolute Error (MAE) was used as a measurement of model bias and is calculated by (Eq. 4) as

$$MAE = \sum_{i=1}^{n} |\Delta SOC_{prediction,\, i} - \Delta SOC_{true,i}|/n \tag{4}$$

2.5.3. Regional and Local Validation of EPIC-IIASA STAT

To evaluate regional scale predictions of SOC change, we utilize a number of published agricultural experiments and meta-analyses from around the world which test and summarize the effects of management on SOC. For a full list of referenced literature, please see (Table3) in the appendix. In this validation, we specifically searched for literature which investigated the effect of inorganic nitrogen fertilizer application and/or residue incorporation using conventional tillage on SOC. Studies of agricultural systems under both continuous cropping and crop rotations were included. Specifically, we look at the linear nature of SOC responses to residue C additions, SOC stock increases with residue incorporation, the effect of N fertilizer on SOC with and without crop residue incorporation, and the geospatial heterogeneity in the combined effect of residue C and nitrogen applications on changes in SOC (see appendix for information on the calculation of specific validation metrics).

To investigate how closely EPIC-IIASA STAT corresponds to local (field) scale SOC dynamics in response to management documented, we selected test sites in the Czech Republic, Italy, Ireland, and Finland (spanning a range of biophysical clusters). In our field study scale reference, we utilize data from four long term experiments across Europe representing a variety of different climates and soils. We investigate how EPIC-IIASA STAT predictions of SOC and the data underlying these predictions diverge from experimental SOC values and site conditions. The experiments utilized are from the Czech Republic (Balkovič et al., 2020), Italy (Triberti et al., 2008), Finland (Singh et al., 2015), and Ireland (van Groenigen et al., 2011). These long-term experiments were chosen to represent 4 contrasting climate and soil conditions (see appendix, Table 2). Data on soil, climate, and SOC dynamics were collected from the publications. We explore how well EPIC-IIASA STAT replicates the reported increase in SOC stock from residue incorporation compared to control as this data is reported in all studies. From the location of each field study, we identify all SimU within a 50km radius of that site and utilize the EPIC-IIASA STAT SOC predictions of these SimU for our comparison.

2.6. Computational Framework

All data engineering, computations, modeling, and plotting were completed with Python 3 software. Data processing and computations were completed using the Numpy and pandas data analysis libraries (Harris et al., 2020; McKinney et al., 2010; Reback et al., 2020). Machine learning models were built and evaluated using the scikit learn package (Pedregosa et al., 2011). Figures were produced using matplotlib (Hunter, 2007). All code available upon request.

3. Results

3.1. Using EPIC-IIASA STAT to Replicate Modeled SOC Dynamics

3.1.1 Accuracy and Model Bias

The accuracy of EPIC-IIASA STAT in predicting EPIC-IIASA GAM simulated mean annual change in SOC is very. high across all regional biophysical clusters and all crop types with low mean bias (mean MAE = 0.005 tC/ha). Meta-models without clustering of biophysical properties achieved an average R^2 of 0.97 across all crops. Meta-models trained on regional biophysical clusters of SimU achieve an average R^2 = 0.99 with all meta-models achieving an R^2 > 0.97 across all crops and clusters. The size of training data has a slight negative impact on the accuracy (R^2) and a positive impact on the bias (MAE) of the meta-models (not shown).

3.1.2. Feature Importance

In our leave-one-out feature importance analysis, we find that on average, mean C content of above-ground residues applied each year (RSDCa) is by a wide-margin the most important variable in predicting Δ SOC when regional biophysical clusters are considered. Across the clusters and crop types, the exclusion of RSDCa results in an average change in R^2 of -0.11, while the average change in R^2 resulting from removal is close to 0 for all other variables. The feature importance resulting from the feature subset selection strategy was more informative in this study, likely due to the covarying nature of features used to train our meta-models. On average across all biophysically-specific meta-models, we find that the combined subset of residue and nitrogen features is the most important in predicting Δ SOC over the 35-year time period followed by residue, climate, soil, and N fertilizer feature subsets (Fig. 2). Exclusion of residue and nitrogen features results in an average change in R^2 of -0.83. In the meta-model built without biophysical stratification, exclusion of residue and nitrogen features resulted in an average change in R^2 of -0.76 across crop types. Exclusion of just residue features results in an average change in R^2 of -0.45 across biophysical clusters and crop types. We find that the subset of residue features also has the highest variation in feature performance across all crops and regional biophysical clusters. Low sensitivity to the exclusion of N fertilization alone is due to its high collinearity with the residue feature. In addition, it should be noted that with biophysical clustering we implicitly lower the importance of soil and climate features aiming to single out the role of crop residue and N fertilization features in the biophysically specific meta-models. As expected, there are distinct differences in feature subset importance among the biophysically specific meta models which are dependent on biophysical cluster properties. For more information on these differences, please refer to the appendix.

3.2. Verification of EPIC-IIASA STAT at Regional and Local Scale

While EPIC-IIASA STAT is highly accurate in replicating EPIC-IIASA GAM's simulated SOC dynamics across all climate and soil clusters, future practical use of a meta model like EPIC-IIASA STAT would be contingent on the ability to reproduce SOC dynamics on the ground. Here we present the results of an EPIC-IIASA STAT comparison to the literature at two scales: (a) an inspection of emulated SOC trends and responses across Europe at the scale of the regional biophysical clusters, and (b) at a field scale by comparing our meta-model against four long term experiments, using predictions within a 50km radius of experimental locations.

3.2.1 Biophysical Cluster Scale Verification

We find that across crop types, management choices, and regional biophysical clusters, EPIC-IIASA STAT captures• a linear relationship between SOC change and the amount of applied C from residues (see appendix Fig. 9, all regressions are statistically significant at P<0.01). From these linear relationships, we identify the C Conversion Efficiency (CE) of residues applied as the ratio between the mean annual change in SOC and the total C applied from residues (including above and below ground residues). Over 35 years across all crop types, management, soils, and climates, EPIC-IIASA STAT shows most 35-year CE ranging from 0 to .05 with an average conversion efficiency of .015 (see appendix Fig. 10). Only 6.9% of the emulated CE, across all crops and N fertilizer rates, are above .05 while the ranges in cited literature are between 0% and 25%.

Many studies report the increase in SOC stock with residue incorporation compared to control treatments. Averaged over crop type and nitrogen fertilizer rates, EPIC-IIASA STAT finds that the majority of increases in SOC stock fall within the range of those from published literature (Fig. 3). The mean increase in SOC stock in EPIC-IIASA STAT is 4.6% across all crops and N-fertilization rates, while 95% of the emulated SOC increases range between 0.7% and 11.7%. The increase in SOC stock with residue incorporation compared to control ranges between 0% and 33.6% in cited literature.

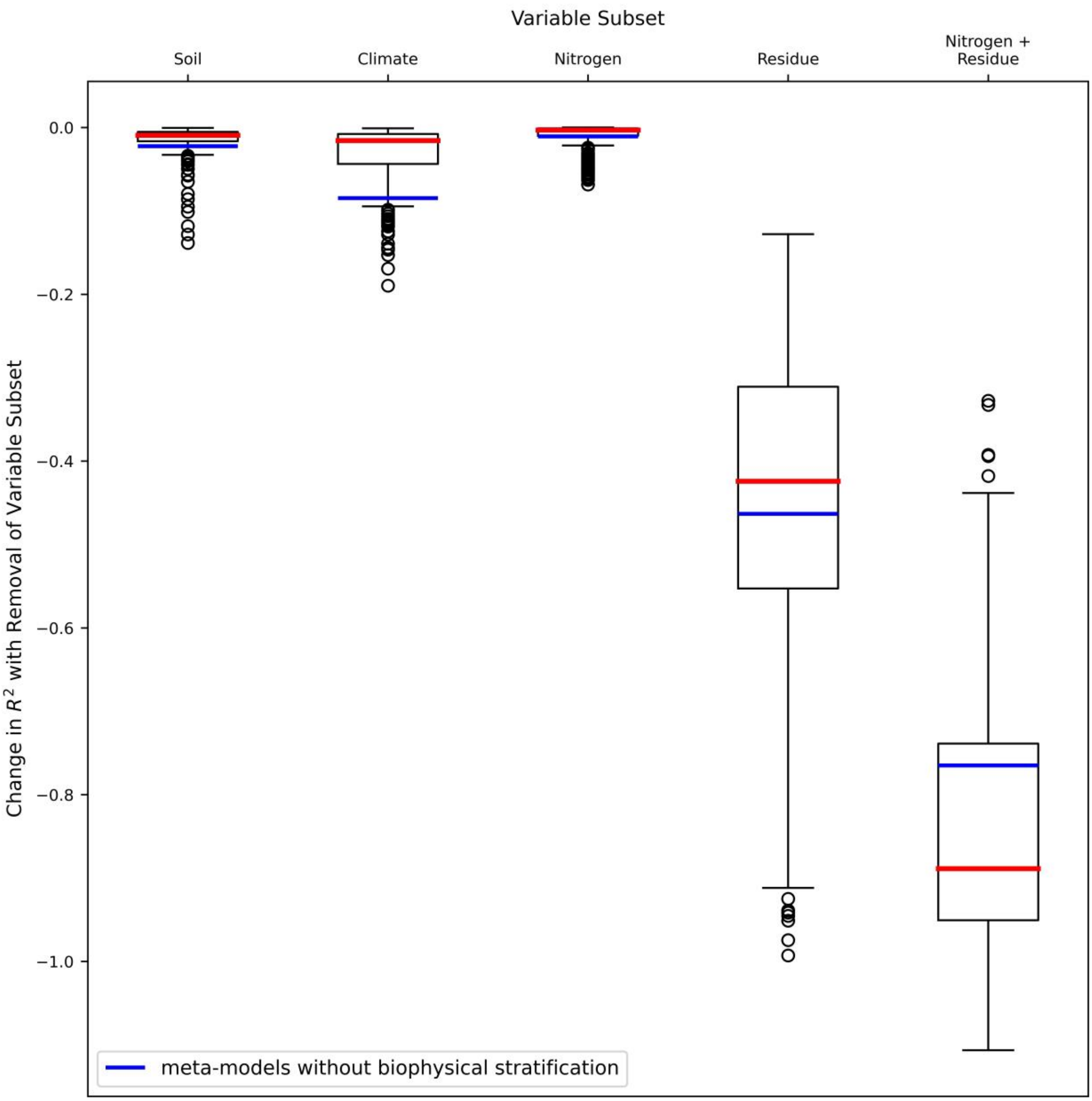

Figure 2: Box plots of feature subset importance for all crops and biophysical-specific meta-models within EPIC-IIASA STAT. Whiskers correspond to 1.5 times the upper and lower quartiles. Box corresponds to the interquartile range, red line corresponds

to median value. Blue lines represent the feature subset importance of the meta-model built without biophysical stratification, averaged across crop types.

EPIC-IIASA STAT also shows a linear effect of nitrogen fertilizer application on SOC change. EPIC-IIASA STAT predictions show that for each crop, N application has a slight positive effect on SOC when residues are not included and the N effect increases as we increase the percentage of residues incorporated (see appendix Fig. 11, all regressions are statistically significant at P<0.01).

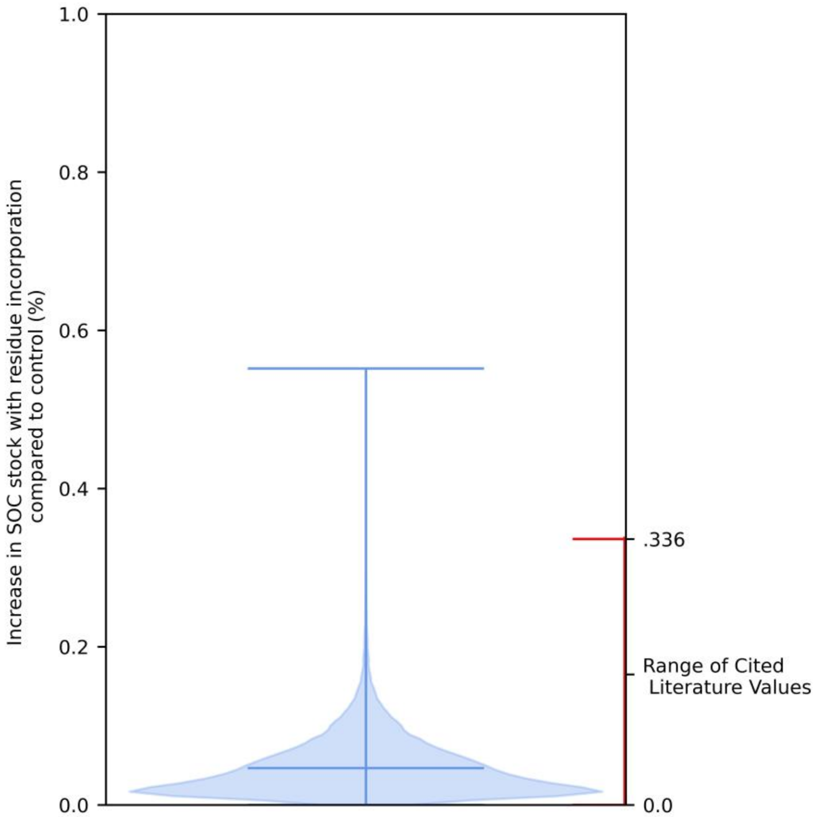

Figure 3: Increases in SOC stock with residue incorporation compared to control averaged over crop type and N inputs for each SimU. Blue lines represent the range of values from EPIC-IIASA STAT, center blue line represents mean of EPIC-IIASA STAT values. Red lines show comparison to values in published literature.

3. 2.2 Field Study Scale Verification

In our field study scale comparison using long-term experiments from across Europe, we find that, on average, EPIC-IIASA STAT predictions underestimate experimental SOC changes for all experiments except for those in Finland (see appendix Fig. 14). In the Czech Republic, EPIC-IIASA STAT captures the range of SOC increases from residue incorporation while the mean of EPIC-IIASA STAT SOC increases (5.5%) is lower than the 7-12% change reported by Balkovič et al. (2020). In Italy, the emulated range of SOC increases also captures the reported value, but the mean EPIC-IIASA STAT increase in SOC of 8.2% is lower than the rate of 19% reported by Triberti et al. (2008). In Ireland, the EPIC-IIASA STAT ranges of 1.3-3.2% are lower than the reported value of 7.3% (van Groenigen et al., 2011). In Finland, our meta-model overestimates the SOC increase from residues as it predicts a mean soil SOC content increase of 1.1% while the experiment reports no measured increase in SOC from residues. Although the model diverges from the measured values of SOC increases from residues, it does capture the spatial variation of experiments with Italy experiencing the largest SOC increases, followed by the Czech Republic, Ireland, and Finland.

For each site, we compare EPIC-IIASA STAT input climate and soil variables to those of published literature which may have impacted the discrepancies in the models ability to replicate SOC responses to management (see appendix, Fig. 15). Again, our ranges capture most of the values in the Czech Republic for most variables with the exception of soil clay content and mean annual precipitation. In Italy, the measured values of soil clay content, soil bulk density, and mean annual temperature are outside the range of EPIC-IIASA STAT values for these variables. In Ireland, the soil sand and clay content are outside the EPIC-IIASA STAT ranges.

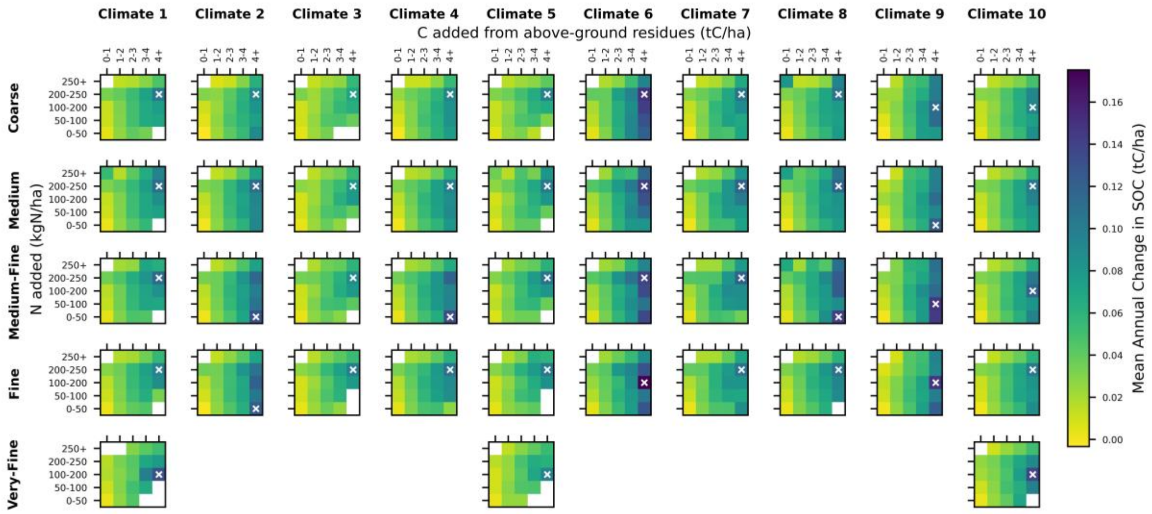

Figure 4: Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster, averaged over all crops except soybean. For crop-specific heatmaps, please see the Appendix. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

3.3. Demonstrating the Use of EPIC-IIASA STAT to Identify Management Interventions

While in general, the addition of N fertilization and residue C lead to larger changes in mean annual SOC (see appendix, 16), the effects of management are highly variable across climatic and soil conditions. The treatment which has the highest variability in SOC accumulation (.031 tC/ha) across clusters is maximum residue addition (>4tC/ha) with minimal N addition (<50kgN/ha) (see appendix, 16). There are distinct responses to treatments depending on the unique combination of soil and climate. For example, with maximal residue addition and minimal N addition, medium fine textured soils in cooler climates (e.g. Subarctic North) are predicted to have the highest SOC accumulation among all soil x climate clusters. Yet with maximal N addition and minimal residue addition, fine textured soils in cooler climates (Subarctic North) are predicted to have the lowest SOC accumulation while coarse soils in temperate climates (Cool Mediterranean) are predicted to have the highest SOC accumulation.

The combination of climate and edaphic parameters results in substantial variability of SOC dynamics across Europe. For each biophysical cluster, we find that EPIC-IIASA STAT predicts varying rates of SOC change as a result of varied application of residue C and nitrogen fertilizer. For most climates x soil combinations, EPIC-IIASA STAT predicts the highest rate of mean annual change in SOC when more than 4tC/ha are applied as residues. Interestingly, for some biophysical clusters EPIC-IIASA STAT shows that high amounts of applied nitrogen are predicted to reduce the mean annual change in SOC (e.g. Subarctic Mountain x fine textured soil in 4). Across all clusters, there is a strong link between a cluster's climatic characteristics and SOC accumulation (see appendix, 17). The average cumulative growing season precipitation of a cluster has a positive linear relationship with average SOC accumulation of a cluster. The relationship between temperature and SOC accumulation is more nuanced - there is a negative quadratic relationship between mean growing season temperature of a cluster and average SOC accumulation of a cluster (see appendix, 17). Across treatments, climate and crop types, soils with medium-fine texture have the highest average SOC accumulation (.054 tC/ha) and the highest variability in SOC accumulation (.031 tC/ha).

4. Discussion

4.1. Utility of EPIC-IIASA STAT in replicating modeled SOC dynamics

The high degree of accuracy achieved by EPIC-IIASA STAT in predicting multifactorial EPIC-IIASA GAM simulations shows the ability to utilize highly interpretable meta-models to simplify and explore the EPIC-IIASA GAM. The multiple polynomial regression model used in this study maintains a high degree of mechanistic information from the original process models compared to the black-box nature of machine learning models (Rudin, 2019) and more importantly, it maintains mechanistic predictions that could be used to support decision making at the local to regional scale. Furthermore, the results of our analysis show that biophysical stratification does modestly improve the accuracy and reduce the bias compared to the meta-models built without biophysical stratification. By stratifying our training data across climatic and soil classes and building separate meta-models on these subsets, EPIC-IIASA STAT may learn more nuanced response relationships which are unique to the biophysical conditions, thus resulting in higher accuracy. This capacity could be further leveraged in future applications of this type of approach. Yet, our regression model has some limitations. Since we set the structure and degree of the multiple polynomial regression base estimator, we must make an assumption about the nature of the SOC response relationships which may lead to inaccuracies. We chose to use quadratic terms for our multiple polynomial regression as quadratic effects of management parameters such as applied nitrogen are well studied and accepted in the literature (Bullock and Bullock, 1994; Puntel et al., 2016). Furthermore, while we attempt to mitigate the effects of multicollinearity in our model, the polynomial regression model is more sensitive to these effects than other machine learning models such as random forests. The feature analysis points to the parameters which are critical to accurately predict mean change in SOC and the regional climate and soil characteristics which vary the importance of these parameters. For further discussion of the feature analysis of EPIC-IIASA STAT, please refer to the appendix.

4.2. Prediction of SOC dynamics

4.2.1 Regional Biophysical Cluster Scale

We find that across crop types and regional biophysical clusters, EPIC-IIASA STAT SOC dynamics agree with the literature on SOC responses to management change (see appendix Table 3 for full list of literature). A number of experiments and meta-analyses have identified a linear relationship between SOC change and the amount of applied C, which is represented well in EPIC-IIASA STAT Campbell et al. (2002); Duiker and Lal (1999); Kong et al. (2005); Rasmussen and Collins (1991); Thomsen and Christensen (2004). Furthermore, there is wide agreement that without residue application, N application has only a little positive effect on SOC, mainly through stimulating higher root biomass (Bertora et al., 2009; Biau et al., 2013; Lugato et al., 2006; Salinas-Garcia et al., 1997; Sandén et al., 2018; Triberti et al., 2008; Searle and Bitnere, 2017). Nitrogen applications have much larger effect when accompanied by crop residue incorporation as evidenced in many studies (Cvetkov and Tajnšek, 2009; Gregorich et al., 1996; Lu et al., 2011; Lugato et al., 2006; Malhi et al., 2011; Mazzoncini et al., 2011; Russell et al., 2009; Šimon et al., 2013; Tajnšek et al., 2013; van Groenigen et al., 2011; Wang and Dalal, 2006; Xu et al., 2007).

At the regional biophysical cluster scale, EPIC-IIASA STAT ranges of C conversion efficiencies and increases in SOC resulting from residue incorporation satisfactorily capture the ranges cited in literature with most of the values falling inside the cited literature ranges. For both of these measurements, EPIC-IIASA STAT predictions trend lower than experimental results. This propensity of EPIC-IIASA to underestimate positive SOC changes has been shown in a handful of published studies (e.g., Balkovič et al. (2020); Izaurralde et al. (2006)). Since our meta-models are trained on the EPIC-modeled gridded dataset, EPIC-IIASA STAT captures the broad patterns of SOC dynamics as well as the ways biophysical parameters affect these dynamics in the EPIC model, including its limitations. The limitations of the biophysical model EPIC and its regional gridded applications in representing SOC dynamics following certain management have widely been discussed in Balkovič et al. (2020). For further discussion of the underlying causes of discrepancies between EPIC-IIASA STAT predictions and published data, please refer to the appendix.

It is widely known that calibration and validation of EPIC-IIASA and other agronomic models is needed to replicate specific rates and SOC measurements at the site scale (Antle et al., 2017; Balkovič et al., 2020; Müller et al., 2017; Silva and Giller, 2020), thus accurate predictions of specific SOC metrics across Europe requires large amounts of ground data and often requires time-consuming work with the model to replicate site conditions. While EPIC-IIASA STAT may have lower than observed mean estimates for the rates of C conversion efficiency and increase in SOC stock over control, it does capture the biophysical variation in SOC responses to management and is consistent with patterns across soils and climate. Although specific rates from EPIC-IIASA STAT may not perfectly match those in cited literature, using EPIC-IIASA STAT to replicate broad patterns of SOC response may still be useful in regional planning efforts as the direction and magnitude of potential changes can be a key issue in decision making (Minasny et al., 2017; Smith et al., 2020; Slessarev et al., 2022). The general SOC response curve and associated variation across biophysical clusters, while less useful in identifying exact rates of SOC change or content, may provide a robust and reliable outlook on the efficacy of management choices across diverse biophysical (climate x soil) settings and could provide a new approach for regional evaluation of the potential efficacy of management interventions to address soil carbon storage.

4. 2.2 Field Study Scale

Field study scale SOC dynamics are of possible utility to farmers who may want to optimize residue harvesting improve soil health, or utilize carbon credits (Antle et al., 2017; Müller et al., 2017; Silva and Giller, 2020). EPICIIASA STAT predicted ranges of SOC stock increases resulting from residue incorporation either fully capture the range of experimental values (Czech Republic and Italy) or come very close to the measured values (Finland and Ireland) (see appendix, Fig 14. The model also captures the broad geographical pattern in SOC changes across the four long term experiments with Italy experiencing the highest SOC increases from residues, followed by the Czech Republic, Ireland, and Finland. EPIC-IIASA STAT shows the closest agreement in SOC changes with the experiment in the Czech Republic.

Similarly to the regional patterns discussed above, differences between EPIC-IIASA GAM and experiments in the plowing depth of residues (20 cm vs 20-40 cm), regional crop rotations and detailed agricultural practices, and C concentration of residues (40% vs 35-45%) may contribute to these differences. Secondly, differences in the soil and climate data which underlie the hypercube training data may also contribute to deviations from experimental values (see appendix, Fig 15). It has been shown that inaccurate localization of soil and climate data may introduce significant bias to regional models (Balkovič et al., 2020; Zhao et al., 2015). Sometimes the input values are missing in the literature data. For example, in Finland, most of the ranges of EPIC-IIASA STAT climate and soil parameters capture the reported values from the long term experiment but the experiment lacked measurements for initial SOC values. Notably, EPIC-IIASA STAT ranges for soil and climate parameters mostly capture the ranges of these parameters in the Czech Republic where EPIC-IIASA STAT SOC dynamics agree most closely with the published study (see appendix, Fig 15). Finally, using predictions from a 50km radius around the site location to compare against experimental data is quite arbitrary. This may lead to a very wide range of impacts, especially in heterogeneous regions, or it also may miss the conditions in long-term experiments entirely. While we show a number of promising successes of EPIC-IIASA STAT, comparison to a larger degree of long-term experiments would help solidify an understanding of the successes and limitations of this approach at the field study scale and the performance of our meta-models under given climate and soil conditions.

4.3. Utilizing EPIC-IIASA STAT to Identify Management Interventions

EPIC-IIASA STAT shows that soil and climate properties may help determine SOC accumulation across Europe and the efficacy of management interventions in improving SOC outcomes. While climates with more precipitation have higher levels of SOC accumulation on average over soil textures and treatments, EPIC-IIASA STAT finds that the importance of management parameters in predicting mean annual change in SOC is the highest in areas where precipitation is between 400 and 600 mm during the growing season (see appendix, 17). These results suggest that these areas may have optimal climate characteristics for effective management of SOC. The relationship between precipitation, temperature, and mean annual change in SOC may be a result of the effects of climate on both SOC processes and yield. Globally, soil carbon is positively correlated with precipitation and negatively correlated with temperature (Jobbágy and Jackson, 2000). Yet yield, which is also correlated with SOC, has more nuanced precipitation and temperature relationships (Agnolucci and De Lipsis, 2020). Additionally, clusters with soils that have high water holding capacity tend to have higher management feature importances and lower climate feature importances (see appendix, 6). Soils with better water storage may more readily accumulate SOC since microbial and enzymatic activity decreases as soils dry (Moyano et al., 2013). EPIC-IIASA STAT results suggest that the interplay of SOC processes and yield may determine the climatic and soil characteristics which are most favorable for improving SOC accumulation.

In some cases and specifically in Continental Central Europe and Subarctic Northern Europe, higher levels of N fertilizer may in fact reduce SOC accumulation. Since nitrogen fertilization can cause substantial adverse effects on the environment, finding areas where SOC accumulation benefits from fertilizer reduction could help direct reduction in usage (Martínez-Dalmau et al., 2021). In Europe, the use of nitrogen fertilizers has been a target of policy intervention in an effort to reduce harmful effects of excessive nitrogen application (Fezzi et al., 2010). While it is often observed that nitrogen fertilization may increase soil C sequestration through an increase in organic carbon inputs via higher productivity, nitrogen fertilization may also affect the soil organic matter decay rates which could outweigh C inputs (Khan et al., 2007; Neff et al., 2002; Russell et al., 2009; Jesmin et al., 2021). The balance between organic inputs and decay rates resulting from N fertilization is unique to the agroecosystem and management choices (Russell et al., 2009), thus for some parts of Europe - e.g. fine-textured soils in Norway - increasing the amount of applied nitrogen may not further increase mean annual change in SOC. EPIC-IIASA STAT shows that cooler climates require the lowest amounts of applied Nitrogen in order to achieve maximum SOC accumulation while warm mediterranean climates require the most applied Nitrogen. Among soil types, EPIC-IIASA STAT finds that coarse textured soils require the most N fertilizer across climates. Coarse textured soils may require the most N fertilizer to achieve maximum SOC accumulation rates due to the inability to form aggregates and mineral protected C, both of which accelerate SOC accumulation rates (Schimel et al., 1994; Xu et al., 2020). The results from EPIC-IIASA STAT could help identify opportunities to balance the reduction of N fertilizers and the improvement of SOC outcomes.

4.4. Conclusions

Understanding the geospatial variation and potentials in SOC responses to management is important to policymakers in evaluating the tradeoffs of SOC management and identifying promising interventions across Europe at a variety of scales. Our approach shows the efficacy of utilizing statistical meta-models, with a particular attention paid to biophysical mechanisms, in understanding and measuring SOC dynamics under different crop and management choices. First, this study identifies a framework to increase the capacity of GAMs for regional SOC modeling by combining multifactorial simulations with statistical response modeling while helping to reduce the challenges of detailed calibration and validation of EPIC-IIASA by providing a simplified version of the model for practical applications. Our novel approach of biophysical-stratification highlights the importance of the soil and climate systems in driving SOC dynamics. Identifying the efficacy of management choices in increasing SOC, which may be influenced by the climate and soil of a given site could help farmers or advisors examine the potential implications of management shifts prior to implementation. The accuracy and divergences in meta-model structure as a result of this stratification signal a benefit of explicitly investigating the ways that biophysical characteristics specify SOC responses. Second, while we highlight the continued necessity of calibration and validation in order to replicate accurate SOC measurements, we demonstrate a robust use of EPIC-IIASA STAT for biophysical cluster and field study scale analyses of broad SOC dynamics. Further utility of our presented approach could also be achieved with improved input climate and soil data. Since the large amount of data and high degree of calibration and validation needed for accurate predictions of farm-level SOC often prohibits the use of bulky process-based models, the successes of a meta-model in capturing experimental SOC values opens the door for targeted usage of these technologies in novel ways. The ability of EPIC-IIASA STAT to capture SOC dynamics across Europe presents an opportunity to provide actionable insights to decision makers and land managers without the need for extensive model calibration and validation, which stays with the GAM development requiring highly specific research tools to deliver reliable results.

5. Acknowledgements

TI acknowledges that this work was performed under the auspices of the Young Scientists Summer Program (YSSP) of the International Institute of Applied Systems Analysis (IIASA), supported by a fellowship from the National Academies of Science, Engineering and Medicine (NASEM) US Committee for IIASA, with funds from the National Science Foundation (NSF Award OISE1663864).TI also acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 2040434. CF, JB, and RS work was supported by the European Union’s Horizon Europe research and innovation action under grant agreement No. 101086179 (AI4SoilHealth). CF was funded by the Austrian Science Fund (grant no. P-36220 N).

CRediT authorship contribution statement

Tara A. Ippolito: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization. **Juraj Balkovič:** Conceptualization, Methodology, Supervision, Data Curation, Software, Writing - Original Draft, Writing - Review & Editing. **Rastislav Skalsky:** Conceptualization, Methodology, Supervision, Writing - Original Draft, Writing - Review & Editing. **Christian Folberth:** Conceptualization, Methodology, Supervision, Writing - Original Draft, Writing - Review & Editing. **Tamás Krisztin:** Supervision, Writing - Review & Editing. **Jason Neff:** Conceptualization, Supervision, Writing - Original Draft, Writing - Review & Editing.

References

Abbas, F., Hammad, H.M., Ishaq, W., Farooque, A.A., Bakhat, H.F., Zia, Z., Fahad, S., Farhad, W., Cerdà, A., 2020. A review of soil carbon dynamics resulting from agricultural practices. Journal of Environmental Management 268, 110319. URL: https://www.sciencedirect.com/science/article/pii/S0301479720302541, doi:https://doi.org/10.1016/j.jenvman.2020.110319.

Agnolucci, P., De Lipsis, V., 2020. Long-run trend in agricultural yield and climatic factors in europe. Climatic Change 159, 385–405.

Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., et al., 2020. Towards a global-scale soil climate mitigation strategy. Nature communications 11, 5427.

Antle, J.M., Jones, J.W., Rosenzweig, C.E., 2017. Next generation agricultural system data, models and knowledge products: Introduction. Agricultural Systems 155, 186–190. URL: https://www.sciencedirect.com/science/article/pii/S0308521X16304887, doi:https://doi.org/10.1016/j.agsy.2016.09.003. Attard, E., Le Roux, X., Charrier, X., Delfosse, O., Guillaumaud, N., Lemaire, G., Recous, S., 2016. Delayed and asymmetric responses of soil c pools and n fluxes to grassland/cropland conversions. Soil Biology and Biochemistry 97, 31–39.

Balkovič,J.,Madaras,M.,Skalsky,R.,Folberth,C.,Smatanová,M.,Schmid,E.,vanderVelde,M.,Kraxner,F.,Obersteiner,M.,2020. Verifiablesoil` organic carbon modelling to facilitate regional reporting of cropland carbon change: A test case in the czech republic. Journal of Environmental Management 274, 111206.

Balkovič, J., Skalsky, R., Folberth, C., Khabarov, N., Schmid, E., Madaras, M., Obersteiner, M., Van der Velde, M., 2018. Impacts and uncertainties` of+ 2 c of climate change and soil degradation on european crop calorie supply. Earth’s future 6, 373–395.

Balkovič, J., van der Velde, M., Schmid, E., Skalsky, R., Khabarov, N., Obersteiner, M., Stürmer, B., Xiong, W., 2013. Pan-european crop modelling` with epic: Implementation, up-scaling and regional crop yield validation. Agricultural Systems 120, 61–75.

Balkovič, J., van der Velde, M., Skalsky, R., Xiong, W., Folberth, C., Khabarov, N., Smirnov, A., Mueller, N.D., Obersteiner, M., 2014. Global` wheat production potentials and management flexibility under the representative concentration pathways. Global and Planetary Change 122, 107–121.

Bertora, C., Zavattaro, L., Sacco, D., Monaco, S., Grignani, C., 2009. Soil organic matter dynamics and losses in manured maize-based forage systems. European Journal of Agronomy 30, 177–186. doi:10.1016/j.eja.2008.09.006.

Bhogal, A., Nicholson, F.A., Chambers, B.J., 2009. Organic carbon additions: effects on soil bio-physical and physico-chemical properties. European Journal of Soil Science 60, 276–286. URL: https://bsssjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2389.2008.01105.x, doi:https://doi.org/10.1111/j.1365-2389.2008.01105.x, arXiv:https://bsssjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2389.2008.01105.x.

Biau, A., Santiveri, F., Lloveras, J., 2013. Stover management and nitrogen fertilization effects on corn production. Agronomy Journal 105, 1264. doi:10.2134/agronj2012.0486.

Blanc, É., 2017. Statistical emulators of maize, rice, soybean and wheat yields from globalgridded crop models. Agricultural and Forest Meteorology 236, 145–161.

Breiman, L., 1996. Bagging predictors. Machine learning 24, 123–140.

Bryant, K., Benson, V., Kiniry, J., Williams, J., Lacewell, R., 1992. Simulating corn yield response to irrigation timings: Validation of the epic model. Journal of Production Agriculture 5, 237–242.

Bullock, D.G., Bullock, D.S., 1994. Quadratic and quadratic-plus-plateau models for predicting optimal nitrogen rate of corn: A comparison. Agronomy Journal 86, 191–195.

Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods 3, 1–27.

Campbell, C., Mcconkey, B., Gameda, S., Izaurralde, R., Liang, B., Zentner, R., Sabourin, D., 2002. Efficiencies of Conversion of Residue C to Soil C. pp. 305–314. doi:10.1201/9781420032291.ch29.

Chong, I.G., Jun, C.H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and intelligent laboratory systems 78, 103–112.

Cvetkov, M., Tajnšek, A., 2009. Soil organic matter changes according to the application of organic and mineral fertilizers within long-term experiments. Acta Agriculturae Slovenica 93, 311–320.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence , 224–227.

Ding, C., He, X., 2004. K-means clustering via principal component analysis, in: Proceedings of the twenty-first international conference on Machine learning, p. 29.

Duiker, S., Lal, R., 1999. Crop residue and tillage effects on carbon sequestration in a luvisol in central ohio. Soil and Tillage Research 52, 73–81. URL: https://www.sciencedirect.com/science/article/pii/S0167198799000598, doi:https://doi.org/10.1016/S0167-1987(99)00059-8.

Eglin, T., Ciais, P., Piao, S.L., Barré, P., Bellassen, V., Cadule, P., Chenu, C., Gasser, T., Koven, C., Reichstein, M., et al., 2010. Historical and future perspectives of global soil carbon response to climate and land-use changes. Tellus B: Chemical and Physical Meteorology 62, 700–718.

Ewert, F., Rounsevell, M., Reginster, I., Metzger, M., Leemans, R., 2005. Future scenarios of european agricultural land use: I. estimating changes in crop productivity. Agriculture, Ecosystems & Environment 107, 101–116.

Feiziene, D., Feiza, V., Slepetiene, A., Liaudanskiene, I., Kadžiene, G., Deveikyte, I., Vaideliene, A., 2011.˙ Long-term influence of tillage and fertilization on net carbon dioxide exchange rate on two soils with different textures. Journal of environmental quality 40, 1787–96. doi:10.2134/jeq2011.0180.

Fezzi, C., Hutchins, M., Rigby, D., Bateman, I.J., Posen, P., Hadley, D., 2010. Integrated assessment of water framework directive nitrate reduction measures. Agricultural Economics 41, 123–134.

Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N., Obersteiner, M., 2019. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. Agricultural and Forest Meteorology 264, 1–15. URL: https://www.sciencedirect.com/science/article/pii/S0168192318303162, doi:https://doi.org/10.1016/j.agrformet.2018.09.021. Frank, S., Schmid, E., Havlík, P., Schneider, U.A., Böttcher, H., Balkovič, J., Obersteiner, M., 2015. The dynamic soil organic carbon mitigation potential of european cropland. Global Environmental Change 35, 269–278.

Franke, J.A., Müller, C., Elliott, J., Ruane, A.C., Jägermeyr, J., Snyder, A., Dury, M., Falloon, P.D., Folberth, C., François, L., et al., 2020. The ggcmi phase 2 emulators: global gridded crop model responses to changes in co 2, temperature, water, and nitrogen (version 1.0). Geoscientific Model Development 13, 3995–4018.

Gassman, P.W., Williams, J.R., Benson, V.W., Izaurralde, R.C., Hauck, L.M., Jones, C.A., Atwood, J.D., Kiniry, J.R., Flowers, J.D., 2004. Historical development and applications of the epic and apex models, in: 2004 ASAE Annual Meeting, American Society of Agricultural and Biological Engineers. p. 1.

Gregorich, E., Liang, B., Ellert, B., Drury, C., 1996. Fertilization effects on soil organic matter turnover and corn residue c storage. Soil Science Society of America Journal - SSSAJ 60. doi:10.2136/sssaj1996.03615995006000020019x.

Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. Global change biology 8, 345–360.

Harris, C.R., Millman, K.J., Van Der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al., 2020. Array programming with numpy. Nature 585, 357–362.

Hunter, J.D., 2007. Matplotlib: A 2d graphics environment. Computing in science & engineering 9, 90–95.

Iglesias, A., Quiroga, S., Moneo, M., Garrote, L., 2012. From climate change impacts to the development of adaptation strategies: challenges for agriculture in europe. Climatic Change 112, 143–168.

Izaurralde, R., Williams, J.R., McGill, W.B., Rosenberg, N.J., Jakas, M.Q., 2006. Simulating soil c dynamics with epic: Model description and testing against long-term data. Ecological Modelling 192, 362–384.

Jägermeyr, J., Müller, C., Ruane, A.C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J.A., et al., 2021. Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. Nature Food 2, 873–885.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. volume 112. Springer.

Jesmin, T., Mitchell, D.T., Mulvaney, R.L., 2021. Short-term effect of nitrogen fertilization on carbon mineralization during corn residue decomposition in soil. Nitrogen 2, 444–460.

Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological applications 10, 423–436.

Jones, J., Kenig, A., Vallejos, C., 1999. Reduced state–variable tomato growth model. Transactions of the ASAE 42, 255.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., et al., 2017. Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural systems 155, 269–288.

Keel, S.G., Leifeld, J., Mayer, J., Taghizadeh-Toosi, A., Olesen, J., 2017. Large uncertainty in soil carbon modelling related to method of calculation of plant carbon input in agricultural systems. European Journal of Soil Science 68, 953–963.

Khabarov, N., Smirnov, A., Balkovič, J., Skalsky, R., Folberth, C., Van Der Velde, M., Obersteiner, M., 2020. Heterogeneous compute clusters and` massive environmental simulations based on the epic model. Modelling 1, 215–224.

Khan, S., Mulvaney, R., Ellsworth, T., Boast, C., 2007. The myth of nitrogen fertilization for soil carbon sequestration. Journal of Environmental Quality 36, 1821–1832.

Kohavi, R., et al., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Ijcai, Montreal, Canada. pp. 1137–1145.

Kong, A., Six, J., Bryant, D., Denison, R., Kessel, C., 2005. The relationship between carbon input, aggregation, and soil organic carbon stabilization in sustainable cropping systems. Reproduced from Soil Science Society of America Journal. Published by Soil Science Society of America Soil Sci. Soc. Am. J. Soil Biology & Biochemistry within 69, 1078–1085. doi:10.2136/sssaj2004.0215.

Lal, R., 1997. Residue management, conservation tillage and soil restoration for mitigating greenhouse effect by co2-enrichment. Soil and tillage research 43, 81–107.

Lal, R., 2004. Carbon emission from farm operations. Environment international 30, 981–990.

Lal, R., Kimble, J., 1997. Conservation tillage for carbon sequestration. Nutrient cycling in agroecosystems 49, 243–253.

Lehtinen, T., Schlatter, N., Baumgarten, A., Bechini, L., Krüger, J., Grignani, C., Zavattaro, L., Costamagna, C., Spiegel, H., 2014. Effect of crop residue incorporation on soil organic carbon and greenhouse gas emissions in european agricultural soils. Soil Use Manag. 30, 524–538.

Likas, A., Vlassis, N., Verbeek, J.J., 2003. The global k-means clustering algorithm. Pattern recognition 36, 451–461.

Liu, C., Lu, M., Cui, J., Li, B., Fang, C., 2014. Effects of straw carbon input on carbon dynamics in agricultural soils: a meta-analysis. Global change biology 20, 1366–1381.

Lu, M., Zhou, X., Luo, Y., Yang, Y., Fang, C., Chen, J., Li, B., 2011. Minor stimulation of soil carbon storage by nitrogen addition: A metaanalysis. Agriculture, Ecosystems & Environment 140, 234–244. URL: https://www.sciencedirect.com/science/article/pii/S0167880910003257, doi:https://doi.org/10.1016/j.agee.2010.12.010.

Lugato, E., Berti, A., Giardini, L., 2006. Soil organic carbon (soc) dynamics with and without residue incorporation in relation to different nitrogen fertilisation rates. Geoderma 135, 315–321.

Lugato, E., Panagos, P., Bampa, F., Jones, A., Montanarella, L., 2014. A new baseline of organic carbon stock in european agricultural soils using a modelling approach. Global change biology 20, 313–326.

Malhi, S.S., Nyborg, M., Solberg, E.D., McConkey, B., Dyck, M., Puurveen, D., 2011. Long-term straw management and N fertilizer rate effects on quantity and quality of organic C and N and some chemical properties in two contrasting soils in western canada. Biol. Fertil. Soils 47, 785–800.

Manzoni, S., Porporato, A., 2009. Soil carbon and nitrogen mineralization: Theory and models across scales. Soil Biology and Biochemistry 41, 1355–1379.

Martínez-Dalmau, J., Berbel, J., Ordóñez-Fernández, R., 2021. Nitrogen fertilization. a review of the risks associated with the inefficiency of its use and policy responses. Sustainability 13, 5625.

Maynard, J.J., Yeboah, E., Owusu, S., Buenemann, M., Neff, J.C., Herrick, J.E., 2022. Accuracy of regional-to-global soil maps for on-farm decision making: Are soil maps “good enough”? EGUsphere , 1–35.

Mazzoncini, M., Sapkota, T., Bàrberi, P., Antichi, D., Risaliti, R., 2011. Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. Soil & Tillage Research - SOIL TILL RES 114, 165–174. doi:10.1016/j.still.2011.05.001.

McKinney, W., et al., 2010. Data structures for statistical computing in python, in: Proceedings of the 9th Python in Science Conference, Austin, TX. pp. 51–56.

Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.S., Cheng, K., Das, B.S., et al., 2017. Soil carbon 4 per mille. Geoderma 292, 59–86.

Moyano, F.E., Manzoni, S., Chenu, C., 2013. Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. Soil Biology and Biochemistry 59, 72–85.

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., et al., 2017. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. Geoscientific Model Development 10, 1403–1422.

Neff, J.C., Townsend, A.R., Gleixner, G., Lehman, S.J., Turnbull, J., Bowman, W.D., 2002. Variable effects of nitrogen additions on the stability and turnover of soil carbon. Nature 419, 915–917.

Oyebamiji, O.K., Edwards, N.R., Holden, P.B., Garthwaite, P.H., Schaphoff, S., Gerten, D., 2015. Emulating global climate change impacts on crop yields. Statistical Modelling 15, 499–525.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in python. the Journal of machine Learning research 12, 2825–2830.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the köppen-geiger climate classification. Hydrology and earth system sciences 11, 1633–1644.

Pennock, D., Frick, A., 2001. The role of field studies in landscape-scale applications of process models: an example of soil redistribution and soil organic carbon modeling using century. Soil and Tillage Research 58, 183–191.

Petrescu, A.M.R., McGrath, M.J., Andrew, R.M., Peylin, P., Peters, G.P., Ciais, P., Broquet, G., Tubiello, F.N., Gerbig, C., Pongratz, J., et al., 2021. The consolidated european synthesis of co 2 emissions and removals for the european union and united kingdom: 1990–2018. Earth System Science Data 13, 2363–2406.

Puntel, L.A., Sawyer, J.E., Barker, D.W., Dietzel, R., Poffenbarger, H., Castellano, M.J., Moore, K.J., Thorburn, P., Archontoulis, S.V., 2016. Modeling long-term corn yield response to nitrogen rate and crop rotation. Frontiers in plant science 7, 1630.

Rasmussen, P.E., Collins, H.P., 1991. Long-term impacts of tillage, fertilizer, and crop residue on soil organic matter in temperate semiarid regions, Academic Press. volume 45 of *Advances in Agronomy*, pp. 93–134. URL: https://www.sciencedirect.com/science/article/pii/S0065211308600395, doi:https://doi.org/10.1016/S0065-2113(08)60039-5. Reback, J., McKinney, W., Van Den Bossche, J., Augspurger, T., Cloud, P., Klein, A., Hawkins, S., Roeschke, M., Tratner, J., She, C., et al., 2020. pandas-dev/pandas: Pandas 1.0. 5. Zenodo .

Ringeval, B., Müller, C., Pugh, T.A., Mueller, N.D., Ciais, P., Folberth, C., Liu, W., Debaeke, P., Pellerin, S., 2021. Potential yield simulated by global gridded crop models: using a process-based emulator to explain their differences. Geoscientific Model Development 14, 1639–1656.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature machine intelligence 1, 206–215.

Russell, A.E., Cambardella, C.A., Laird, D.A., Jaynes, D.B., Meek, D.W., 2009. Nitrogen fertilizer effects on soil carbon balances in midwestern us agricultural systems. Ecological applications 19, 1102–1113.

Salinas-Garcia, J.R., Hons, F.M., Matocha, J.E., 1997. Long-term effects of tillage and fertilization on soil organic matter dynamics. Soil Science Society of America Journal 61, 152–159. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj1997.03615995006100010023x, doi:https://doi.org/10.2136/sssaj1997.03615995006100010023x.

Sandén, T., Spiegel, H., Stüger, H.P., Schlatter, N., Haslmayr, H.P., Zavattaro, L., Grignani, C., Bechini, L., D hose, T., Molendijk, L., et al., 2018. European long-term field experiments: knowledge gained about alternative management practices. Soil Use and Management 34, 167–176.

Sanderman, J., Hengl, T., Fiske, G.J., 2017. Soil carbon debt of 12,000 years of human land use. Proceedings of the National Academy of Sciences 114, 9575–9580.

Schimel, D.S., Braswell, B., Holland, E.A., McKeown, R., Ojima, D.S., Painter, T.H., Parton, W.J., Townsend, A.R., 1994. Climatic, edaphic, and biotic controls over storage and turnover of carbon in soils. Global biogeochemical cycles 8, 279–293.

Searle, S., Bitnere, K., 2017. Review of the impact of crop residue management on soil organic carbon in europe. The International Council on Clean Transportation, Working Paper .

Silva, J.V., Giller, K.E., 2020. Grand challenges for the 21st century: what crop models can and can't (yet) do. The Journal of Agricultural Science 158, 794–805. doi:10.1017/S0021859621000150.

Singh, P., Heikkinen, J., Ketoja, E., Nuutinen, V., Palojärvi, A., Sheehy, J., Esala, M., Mitra, S., Alakukku, L., Regina, K., 2015. Tillage and crop residue management methods had minor effects on the stock and stabilization of topsoil carbon in a 30-year field experiment. Science of the Total Environment 518, 337–344.

Skalsky, R., Tarasovicova, Z., Balkovic, J., Schmid, E., Fuchs, M., Moltchanova, E., Kindermann, G., Scholtz, P., 2008. Geo-bene global database` for bio-physical modeling v. 1.0. concepts, methodologies and data. The GEO-BENE database report , 58.

Slessarev, E., Mayer, A., Kelly, C., Georgiou, K., Pett-Ridge, J., Nuccio, E., 2022. Initial soil organic carbon stocks govern changes in soil carbon: Reality or artifact? Global Change Biology 29. doi:10.1111/gcb.16491.

Smith, P., Smith, J., Powlson, D., McGill, W., Arah, J., Chertov, O., Coleman, K., Franko, U., Frolking, S., Jenkinson, D., et al., 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. Geoderma 81, 153–225.

Smith, P., Soussana, J.F., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H., Van Egmond, F., McNeill, S., Kuhnert, M., et al., 2020. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. Global Change Biology 26, 219–241.

Tajnšek, A., Čergan, Z., Čeh, B., 2013. Results of the long-term field experiment iosdv jable at the beginning of the 21st century. Archives of Agronomy and Soil Science 59, 1099–1108. URL: https://doi.org/10.1080/03650340.2012.697996, doi:10.1080/03650340.

2012.697996, arXiv:https://doi.org/10.1080/03650340.2012.697996. Thomsen, I., Christensen, B., 2004. Yields of wheat and soil carbon and nitrogen contents following long-term incorporation of barley straw and ryegrass catch crops. Soil Use and Management 20, 432 – 438. doi:10.1111/j.1475-2743.2004.tb00393.x.

Toudert, A., Braimoh, A., Bernoux, M., St-Louis, M., Abdelmagied, M., Bockel, L., Ignaciuk, A., Zhao, Y., 2018. Carbon accounting tools for sustainable land management .

Triberti, L., Nastri, A., Giordani, G., Comellini, F., Baldoni, G., Toderi, G., 2008. Can mineral and organic fertilization help sequestrate carbon dioxide in cropland? European Journal of Agronomy 29, 13–20. URL: https://www.sciencedirect.com/science/article/pii/S1161030108000166, doi:https://doi.org/10.1016/j.eja.2008.01.009.

van Groenigen, K.J., Hastings, A., Forristal, D., Roth, B., Jones, M., Smith, P., 2011. Soil c storage as affected by tillage and straw management: An assessment using field measurements and model predictions. Agriculture, Ecosystems & Environment 140, 218–225. URL: https://www.sciencedirect.com/science/article/pii/S0167880910003233, doi:https://doi.org/10.1016/j.agee.2010.12.008.

Wang, W., Dalal, R., 2006. Carbon inventory for a cereal cropping system under contrasting tillage, nitrogen fertilisation and stubble management practices. Soil and Tillage Research 91, 68–74. URL: https://www.sciencedirect.com/science/article/pii/S0167198705002783, doi:https://doi.org/10.1016/j.still.2005.11.005.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., Schewe, J., 2014. The inter-sectoral impact model intercomparison project (isi–mip): project framework. Proceedings of the National Academy of Sciences 111, 3228–3232.

White, J.W., Hoogenboom, G., Kimball, B.A., Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. Field Crops Research 124, 357–368. URL: https://www.sciencedirect.com/science/article/pii/S0378429011002395, doi:https://doi.org/10.1016/j.fcr.2011.07.001. Whittaker, C., McManus, M.C., Smith, P., 2013. A comparison of carbon accounting tools for arable crops in the united kingdom. Environmental Modelling & Software 46, 228–239. URL: https://www.sciencedirect.com/science/article/pii/S1364815213000789, doi:https://doi.org/10.1016/j.envsoft.2013.03.015.

Williams, J., Singh, V., 1995. The epic model. computer models of watershed hydrology. Water Resour. Publ. Highl. Ranch Colo 25, 909–1000.

Wösten, J., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of european soils. Geoderma 90, 169–185. URL: https://www.sciencedirect.com/science/article/pii/S0016706198001323, doi:https://doi.org/10.1016/S0016-7061(98)00132-3. Xu, S., Sheng, C., Tian, C., 2020. Changing soil carbon: influencing factors, sequestration strategy and research direction. Carbon balance and management 15, 1–9.

Xu, Y., Chen, W., Shen, Q., 2007. Soil organic carbon and nitrogen pools impacted by long-term tillage and fertilization practices. Commun. Soil Sci. Plant Anal. 38, 347–357.

Zhao, G., Siebert, S., Enders, A., Rezaei, E., Yan, C., Ewert, F., 2015. Demand for multi-scale weather data for regional crop modeling. Agricultural and Forest Meteorology 200, 156–171. doi:10.1016/j.agrformet.2014.09.026.

Šimon, T., Mikanová, O., Cerhanová, D., 2013. Long-term effect of straw and farmyard manure on soil organic matter in field experiment in the czech republic. Archives of Agronomy and Soil Science 59, 1193–1205. URL: https://doi.org/10.1080/03650340.2012.706871, doi:10.1080/03650340.2012.706871, arXiv:https://doi.org/10.1080/03650340.2012.706871.

A. Appendix

A.1. Methodology and Data

A.1.1. Crop Model Input Data

Climate Data

Gridded climate data from the CRU ERA v.2.0 source at a spatial resolution of 0.1° from the time period 1980-2019 (provided by the VERIFY project, https://verify.lsce.ipsl.fr) were utilized for the multifactorial EPIC-IIASA GAM simulations. Variables used include daily precipitation (in mm), minimum and maximum temperature (in ◦), relative humidity (fraction), solar radiation (in MJ m-2), and annual atmospheric CO2 concentration (ppm).

Soil Data

Topsoil (0-30cm) and subsoil (30-100cm) properties were calculated from underlying datasets for each SimU. The soil datasets utilized include the European Soil Bureau Database v. 2.0 (ESBD, https://esdac.jrc.ec.europa.eu), the Database of Hydraulic Properties of European Soils (Wösten et al., 1999), and the map of organic carbon content in the topsoil (Lugato et al., 2014). A total of 13 soil properties were used, following the framework of Balkovič et al. (2013, 2018).

Crop Management

In N fertilization scenarios, mineral N fertilizer is automatically applied based on plant requirements until the maximum annual N application rates (i.e., 50, 100, 250 and BAU kg N/ha) were consumed. In crop retention scenarios, we simulate crop residue harvest with a baler operation to match the retention intensities in Section 2.2.2. Under no N application rate (0 kg/ha), we only simulate 0% residue retained. All simulations were carried out under rainfed conditions only. Crop calendars and model parameterization was adopted from Balkovič et al. (2013, 2018, 2020). We assume a conventional tillage consisting of two soil cultivation operations and a 30-cm deep mouldboard plowing, and an offset disking shortly after harvesting of cereals. In addition, two row cultivations were simulated for maize. Soil erosion was not accounted for in our simulations.

A.1.2. Regional and Local Validation

Table 2
Field study site information.

Study	Site	Country	Duration	Soil Texture	Bulk Density (g/cm^3)	Increase in SOC with residues (%)
Balkovič et al. (2020)	Hněvčeves	Czech Rep.	1980-2016	Loam	1.3	7-12%
Balkovič et al. (2020)	Trutnov	Czech Rep.	1966-2009	Sandy Loam	1.4	
Balkovič et al. (2020)	Ruzyně	Czech Rep.	1954-2017	Clay Loam	1.3	
Balkovič et al. (2020)	Uherský Ostroh	Czech Rep.	1972-2017	Loam	1.3	
Triberti et al. (2008)	Cadriano	Italy	1966-2000	Silt Loam	1.2	19%
Singh et al. (2015)	Jokioinen	Finland	1983-2012	Clay	1.3	0%
van Groenigen et al. (2011)	Knockbeg	Ireland	2000-2009	Sandy Loam	1.4	7%

Table 3
Referenced Literature for Regional and Local Validation of EPIC-IIASA STAT

Reference	Study Location or Meta-Analysis
Abbas et al. (2020)	Meta-Analysis
Bertora et al. (2009)	Italy
Bhogal et al. (2009)	United Kingdom
Biau et al. (2013)	Spain
Campbell et al. (2002)	Meta-Analysis
Cvetkov and Tajnšek (2009)	Slovenia
Duiker and Lal (1999)	United States

Feiziene et al. (2011)	Lithuania
Gregorich et al. (1996)	Canada
Kong et al. (2005)	United States
Lehtinen et al. (2014)	Meta-Analysis
Lu et al. (2011)	Meta-Analsis
Lugato et al. (2006)	Italy
Malhi et al. (2011)	Canada
Mazzoncini et al. (2011)	Italy
Rasmussen and Collins (1991)	Meta-Analysis
Russell et al. (2009)	United States
Salinas-Garcia et al. (1997)	United States
Sandén et al. (2018)	Meta-Analysis
Searle and Bitnere (2017)	Meta-Analysis
Šimon et al. (2013)	Czech Republic
Singh et al. (2015)	Finland
Smith et al. (1997)	Meta-Analysis
Tajnšek et al. (2013)	Slovenia
Thomsen and Christensen (2004)	Denmark
Triberti et al. (2008)	Italy
van Groenigen et al. (2011)	Ireland
Wang and Dalal (2006)	Australia
Xu et al. (2007)	China

The C conversion efficiency (CE)is a rate which reflects the slope of the linear relationship between residue C inputs and changes in SOC. There are a limited number of studies which report these rates as it necessitates the application of residues at varying rates in order to identify the effect of increasing amounts of residue C. Furthermore, many studies calculate the CE differently - e.g. some calculate C applied as above and below ground residues while others utilize just above ground residues. To compare our results against these studies, the EPIC-IIASA STAT average CE of residues over 35 years is calculated as

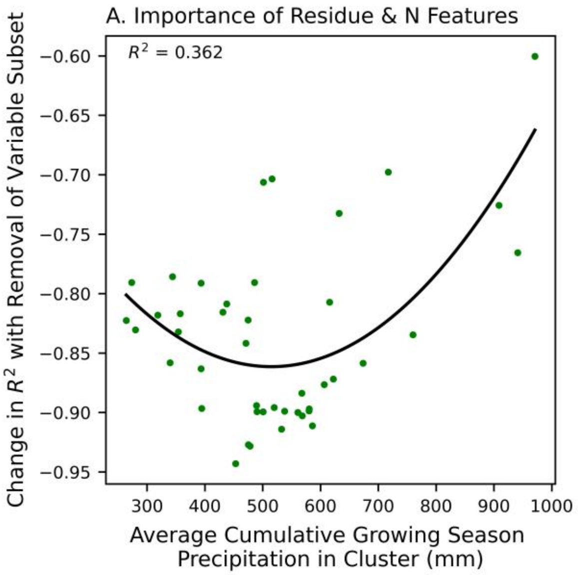

Figure 5: a) Relationship between residue N feature subset importance and average cumulative growing season precipitation of cluster. b) Relationship between climate feature subset importance and average cumulative growing season precipitation of cluster

$$CE = \Delta SOC / RSDC \qquad (5)$$

where ΔSOC is the mean annual change in SOC (tC/ha) and RSDC is the total amount of residue C added from above and below ground residues. Where studies investigating the effect of residues on SOC do not vary the amount of residues added, they utilize a binary residue incorporation (residues included or removed). Studies with binary residue inclusion often report the increase in SOC stock as a result of residue incorporation compared to a control. We calculate the increase in SOC stock after 35 years of residue incorporation compared to the control is calculated for each SimU
x crop x N rate as

$$\frac{final\ SOC_{90\%\ residues\ included} - final\ SOC_{0\%\ residues\ included}}{final\ SOC_{0\%\ residues\ included}} \qquad (6)$$

A.2. Results

A.2.1. Feature Importance

As expected, there are distinct differences in feature subset importance among the biophysically-specific meta models which are dependent on biophysical cluster properties. The importance of residue features is larger among clusters with lower cumulative precipitation during the growing season (Fig 5). Also, the importance of climate features is smaller among clusters with low soil field water-holding capacity (Fig 6). The feature subset importance also varies by crop. The average change in R^2 across biophysically-specific meta-models with removal of the residue feature subset is the largest when modeling soybean growth (change in R^2 = 0.75) and smallest when modeling winter wheat (change in R^2 = 0.29).

A.2.2. Model Structure

The modeled response relationships between features and mean annual ΔSOC are critical to management decisions and optimization of input resources. Due to the interpretable nature of the multiple polynomial regression model, we can unpack the structure of EPIC-IIASA STAT meta-models through model coefficients. While we cannot directly interpret meta-model coefficients as effects on the target variable, we can utilize coefficients to understand the structure

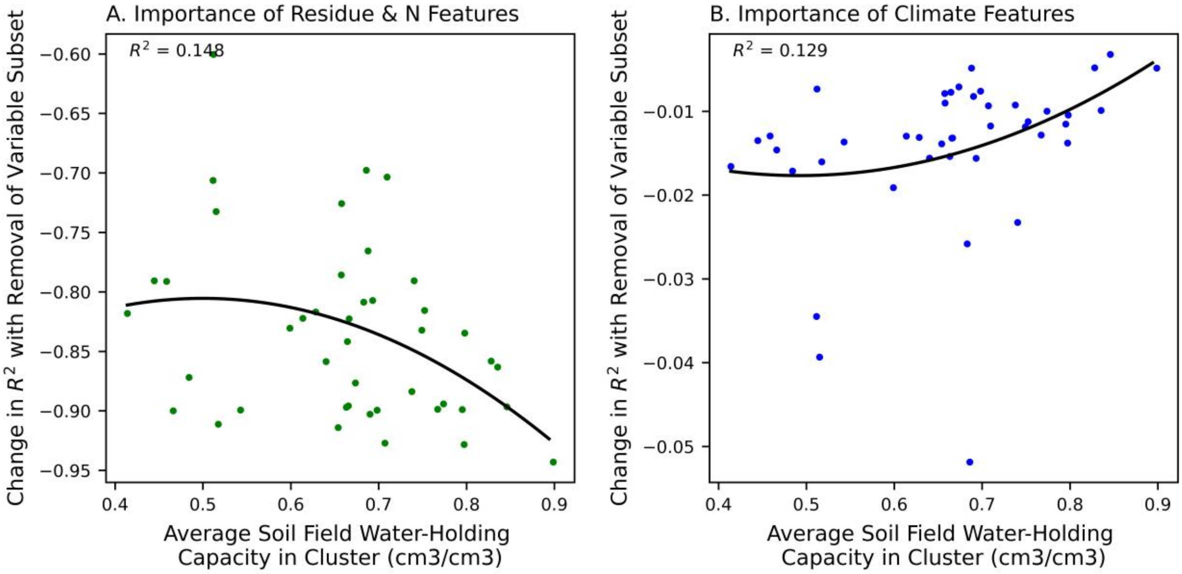

Figure 6: a) Relationship between residue N feature subset importance and average soil field water-holding capacity of cluster. b) Relationship between climate feature subset importance and average soil field water-holding capacity of cluster

of the meta-model and the nature of the feature effects. We find that among all biophysically-specific meta-models and crop types, there is agreement that the nature of the response relationship between mean annual RSDCa and mean annual ΔSOC is linear (Fig 7). We find less overall agreement on the response relationships between mean annual ΔSOC and climatic and soil properties. Response relationships within given climate clusters are often similar, with minor distinctions resulting from the different underlying soil texture clusters (black lines in FIG 8 panels) of the EPIC-IIASA STAT meta-models. For example, the response relationship between mean cumulative growing season precipitation and mean annual ΔSOC is very similar across soils in the continental climate region (denoted in orange, Fig 8) while in other climates such as the coastal Mediterranean climate (denoted in purple, Fig 8) there are wider discrepancies between response relationships depending on the soil texture cluster.

A.2.3. Biophysical Cluster Scale Validation

The trends in C conversion efficiencies highlight the nuances of SOC dynamics represented by EPIC-IIASA STAT (Fig. 12). EPIC-IIASA STAT shows a slight positive correlation between growing season precipitation and CE in agreement with published studies (Rasmussen and Collins, 1991). It is well understood that precipitation helps maintain adequate soil moisture which is related to SOC through a variety of mechanistic controls (Abbas et al., 2020), including the fact that intense rain events can cause SOC losses and oxygen deficit can limit mineralization (Lal and Kimble, 1997). Temperature also plays a role in SOC dynamics as too low or high soil temperature leads to a reduction in mineralization (Lal and Kimble, 1997). Indeed, EPIC-IIASA STAT shows a parabolic relationship between mean growing season temperature and C conversion efficiencies. Another crucial component of the efficacy of SOC management is the initial C content of soils (Slessarev et al., 2022). Residue incorporation may see little effect in soils which already have high levels of SOC as the soils may already be close to saturation (Singh et al., 2015) or the carbon-rich soils may be prone to high mineralization rates on cropland. EPIC-IIASA STAT predictions highlight this dynamic as the CE values are markedly low where initial SOC is high (Fig. 12).

A number of studies have highlighted the increased impact of residues in soils with higher clay content (Liu et al., 2014; Feiziene et al., 2011; Lehtinen et al., 2014). EPIC-IIASA STAT resulted in higher CE values for the medium-fine soils compared to the coarse and medium soils, however, the CE values for clay-rich soils are relatively low (Fig. 13). Lehtinen et al. (2014) report that soils with a clay content >35% show an 8% higher increase in SOC stock with residue incorporation than soils with clay content between 8% and 35%. Feiziene et al. (2011) show that SOC content was 23% higher in loams than in sandy loams after 11 years of varying tillage and fertilization treatments. Liu et al. (2014) find

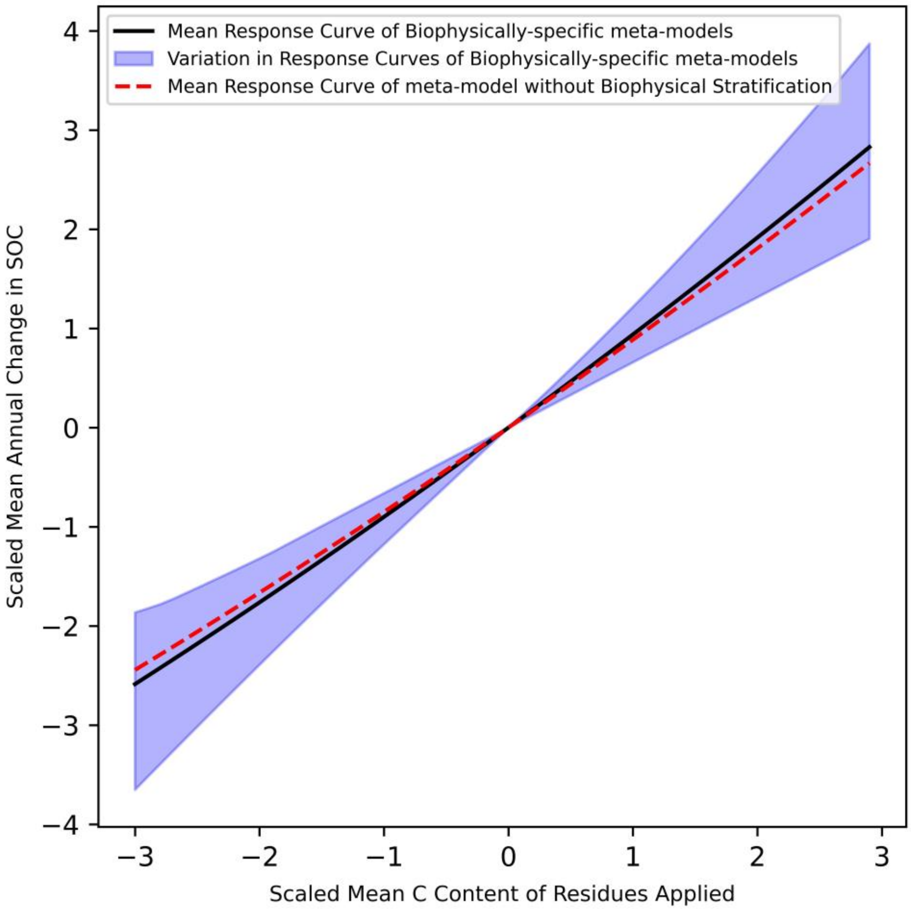

Figure 7: Linear response relationships between mean applied residue C and mean annual change in SOC, variables are scaled to have unit variance and mean centered at zero.

that soil clay content significantly impacts the SOC response to residue inputs. The clay fraction in finer-textured soils generally protects organic matter from mineralization, therefore leading to a higher SOC response to residues in soils with higher clay content (Lal, 1997; Lehtinen et al., 2014). Our results indicate that EPIC-IIASA STAT has certain limitations in capturing the effect of clay soils on SOC response to management.

A.2.4. Field Study Scale Validation

The wider range of predicted SOC response rates (relative to published data) from EPIC-IIASA STAT predictions is likely due in part to the range of climates, soils, crop types, initial C concentrations, and management options modeled in our hypercube. It is also important to note that mapped products, especially for soils, may fail to capture local scale variation in soils that can have large impacts on field-scale carbon change (Maynard et al., 2022). While there are a number of published studies exploring the effects of management on SOC, they do not represent the full range of possible agricultural conditions across Europe - hence the smaller range of reported rates. Discrepancies in SOC response rates may also be the result of a number of differences between modeling and field studies / meta analyses. First, while the EPIC Hypercube simulations underlying EPIC-IIASA STAT assume conventional tillage and continuous cropping only, many studies utilize a range of region-specific crop rotations, management practices, fallow years, or may average SOC responses across multiple tillage practices. Second, studies vary in the depth of soil measurements used for SOC calculations which may skew SOC response rates as different soil horizons experience varying SOC changes. In this study, we calculate SOC from 0-15cm while some experiments sample from 0-10cm or 0-30cm. Third, the GAM-based approaches face accuracy limitations when modelling SOC dynamics (Balkovič et al., 2020). In addition, our fitting algorithm tends to shrink the modelled SOC responses, hence leveling out the response rates further.

A.2.5. Utilizing EPIC-IIASA STAT to Identify Management Interventions

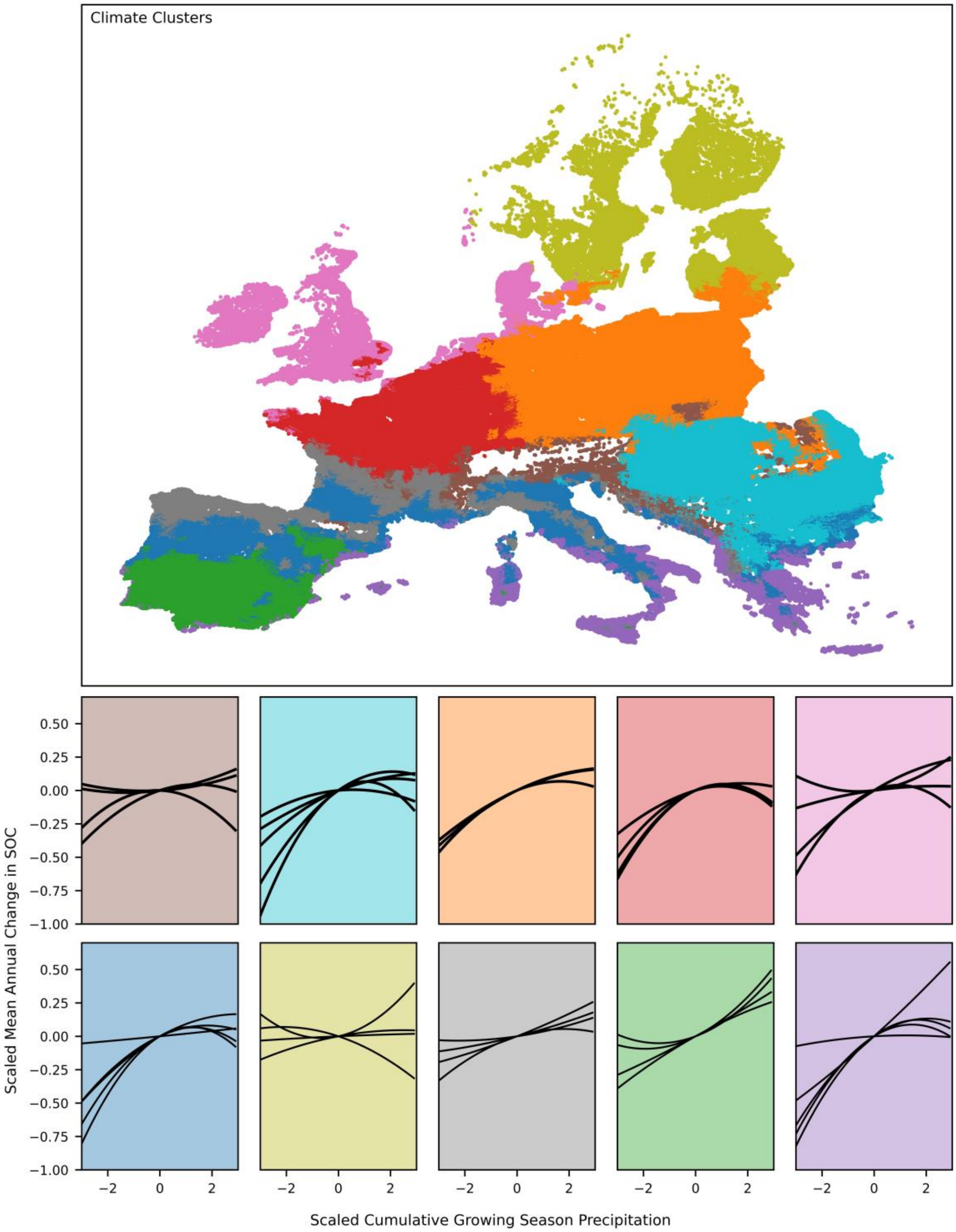

Figure 8: Response relationships between cumulative growing season precipitation and mean annual change in SOC. Each subplot shows five soil specific response curves for the given climate cluster, denoted by color. Variables are scaled to have unit variance and mean centered at zero relative to the cluster rather than global mean. A value of 1 denotes one standard deviation away from the cluster mean.

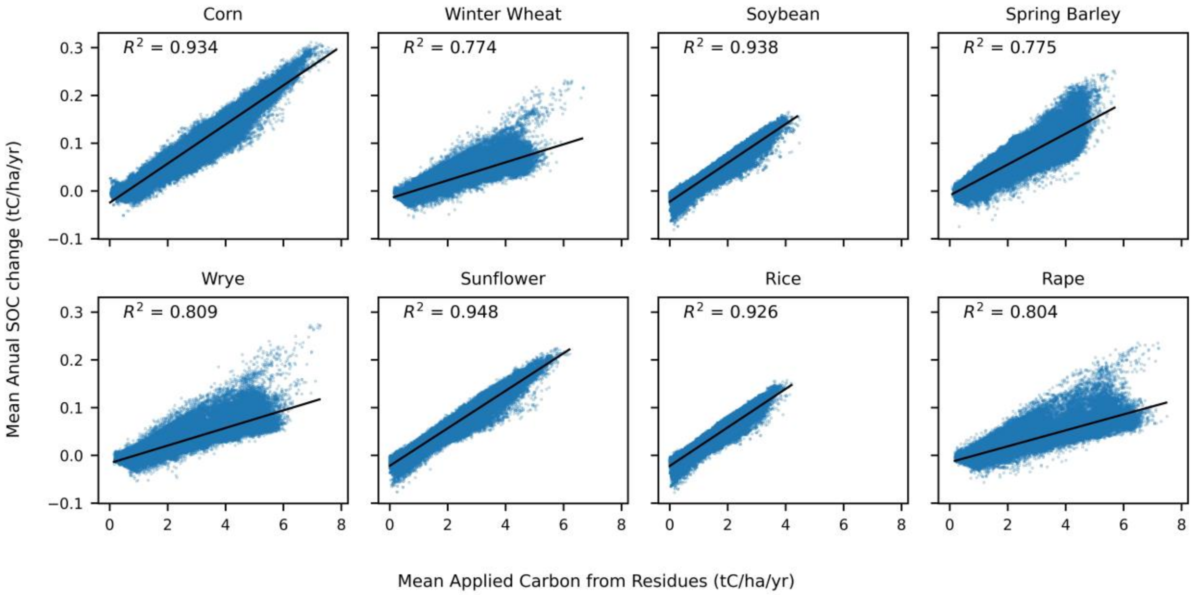

Figure 9: EPIC-IIASA STAT linear relationships between C applied and SOC change for each crop type over all management choices and all regional biophysical clusters

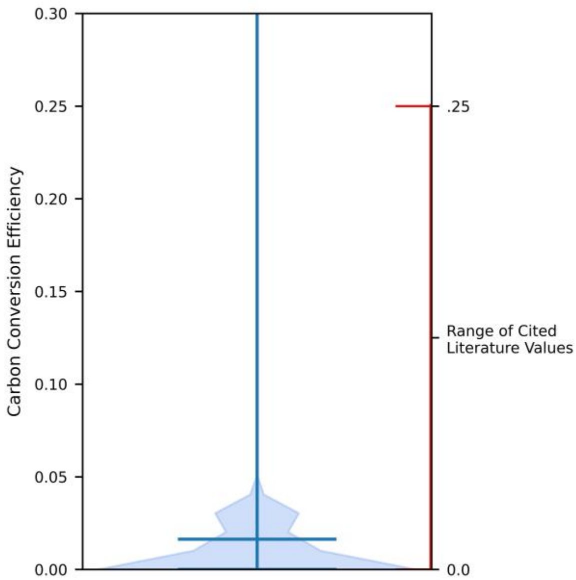

Figure 10: C Conversion Efficiency range from EPIC-IIASA STAT vs cited literature values - 0.03% of EPIC-IIASA STAT predicted C Conversion Efficiencies was above the 30% cutoff.

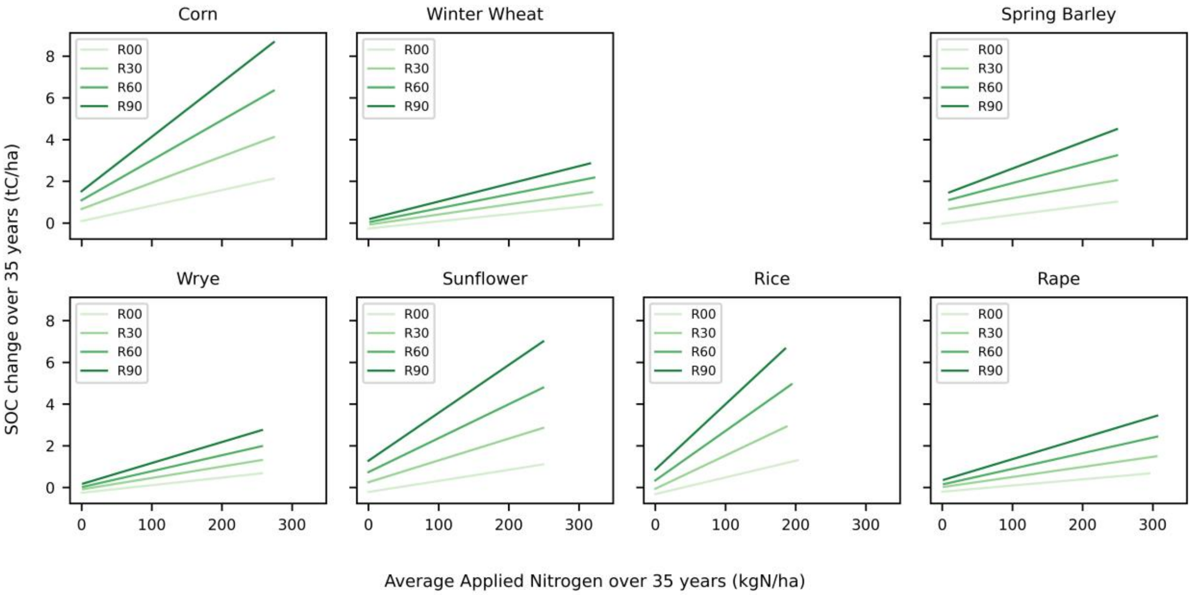

Figure 11: EPIC-IIASA STAT effects of Applied N on SOC change at varying residue levels (R00 = 0% residues retained, R30 = 30% residues retained, R60 = 60% residues retained, R90 = 90% residues retained)

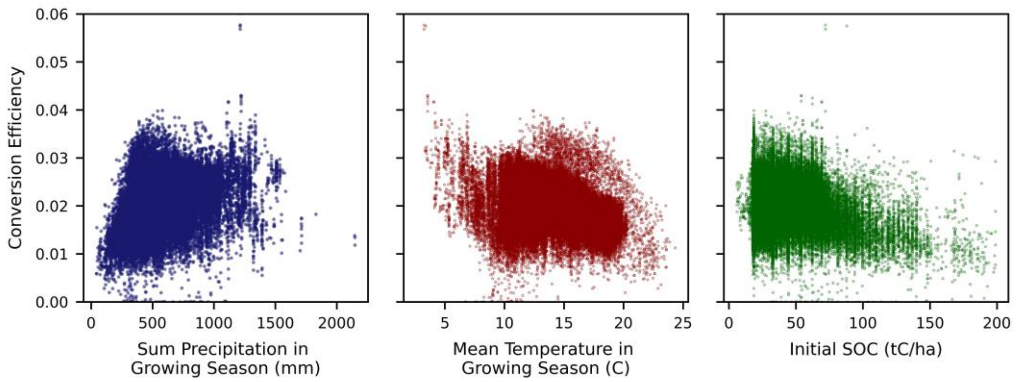

Figure 12: EPIC-IIASA STAT C Conversion Efficiency for each SimU averaged over all crop types and management versus Sum of Growing Season Precipitation, Mean Temperature in Growing Season, and Initial SOC of each SimU

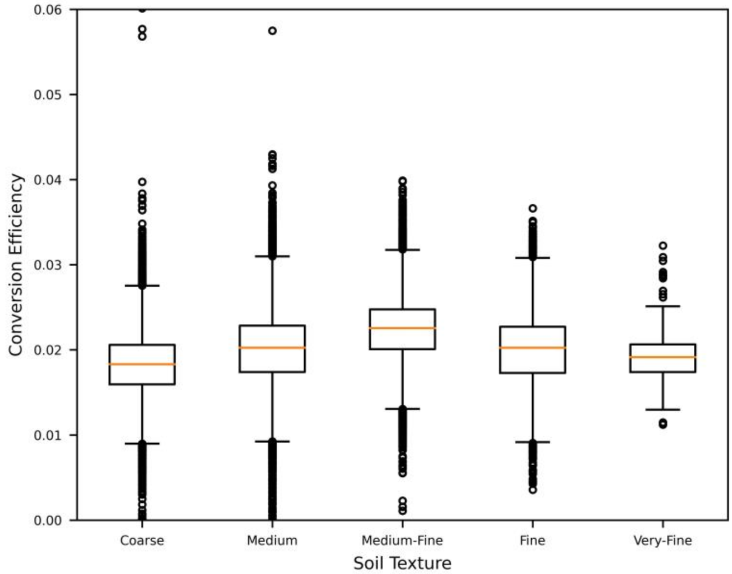

Figure 13: EPIC-IIASA STAT C Conversion Efficiencies of each SimU by soil texture, averaged over crop type and management

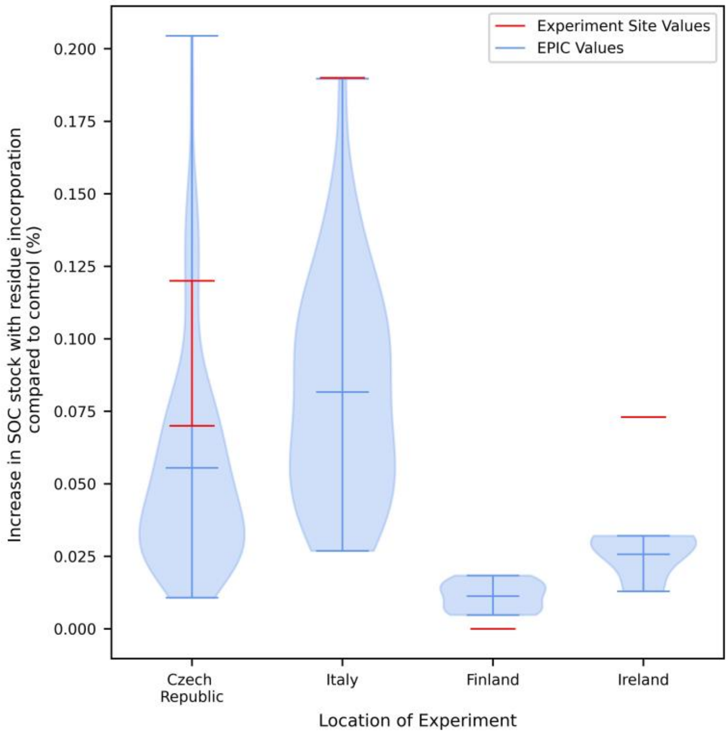

Figure 14: Comparison of increase in SOC stock with residue incorporation compared to control between EPIC-IIASA simulations within 50km of long term experiment and experimental results

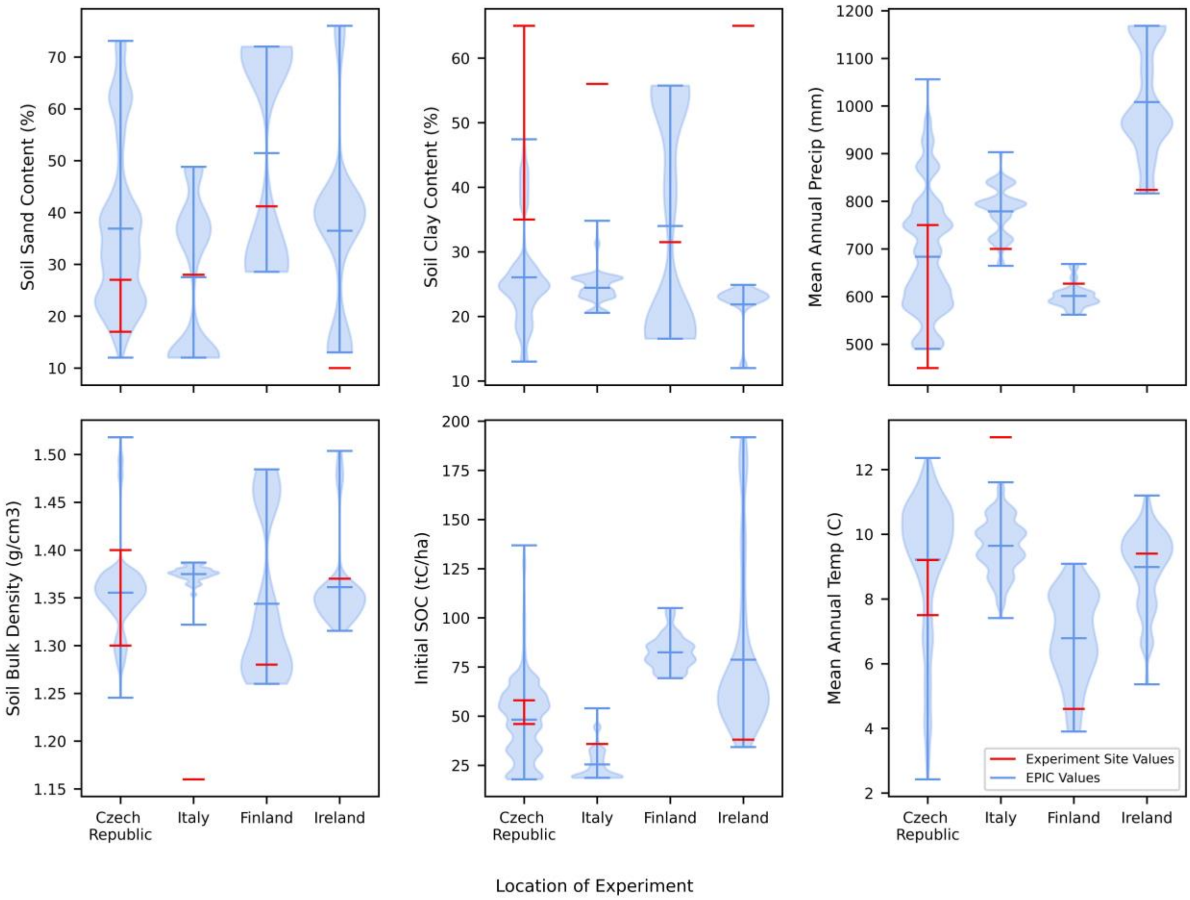

Figure 15: Comparison of sand content, clay content, bulk density, initial SOC, mean annual precipitation, and mean annual temperature between EPIC-IIASA STAT predictions within 50km of long term experiment and experimental values. In Finland, initial SOC values were not measured in the experiment thus reference values are not included in this plot.

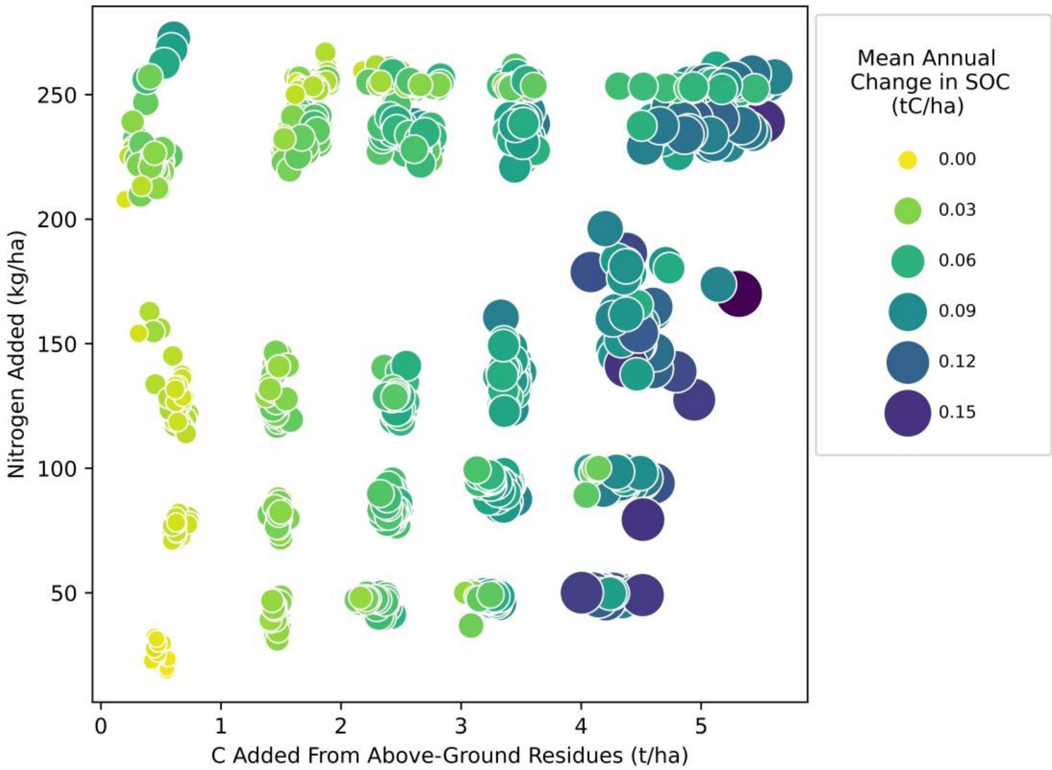

Figure 16: Relationship between applied N (kg/ha) and applied residue C (t/ha) and their combined effect on mean annual change in SOC (tC/ha). Each bubble represents one rate of mean annual change in SOC for each soil x climate cluster, averaged over crop types.

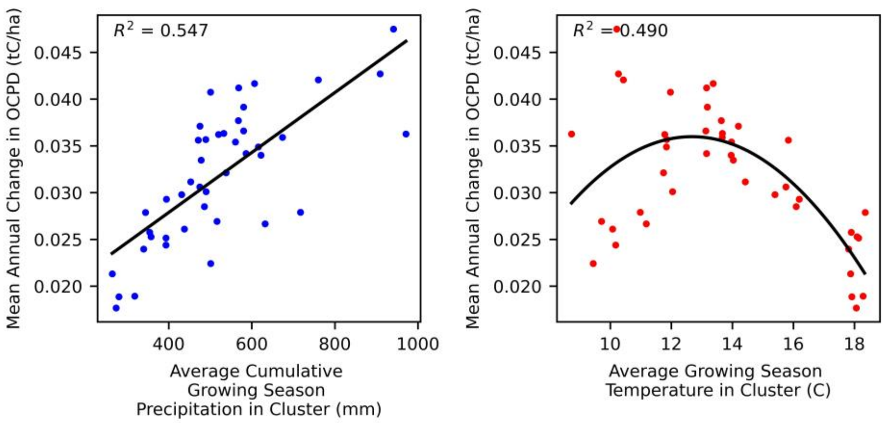

Figure 17: a) Relationship between average cumulative growing season precipitation in cluster and mean annual change in OCPD of cluster b) Relationship between average growing season temperature in cluster and mean annual change in OCPD of cluster

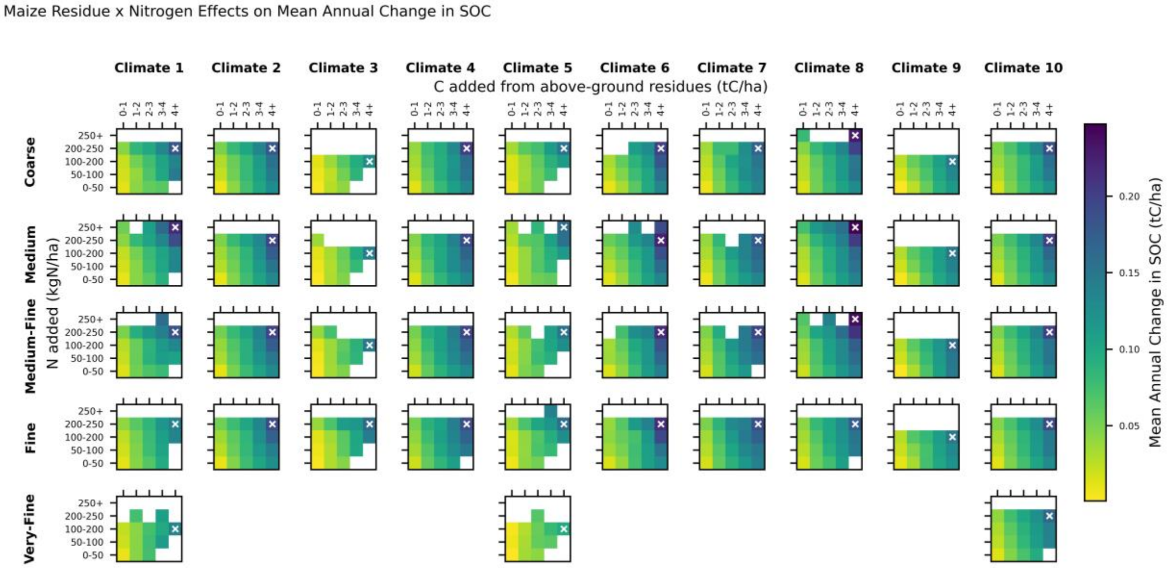

Figure 18: Maize Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

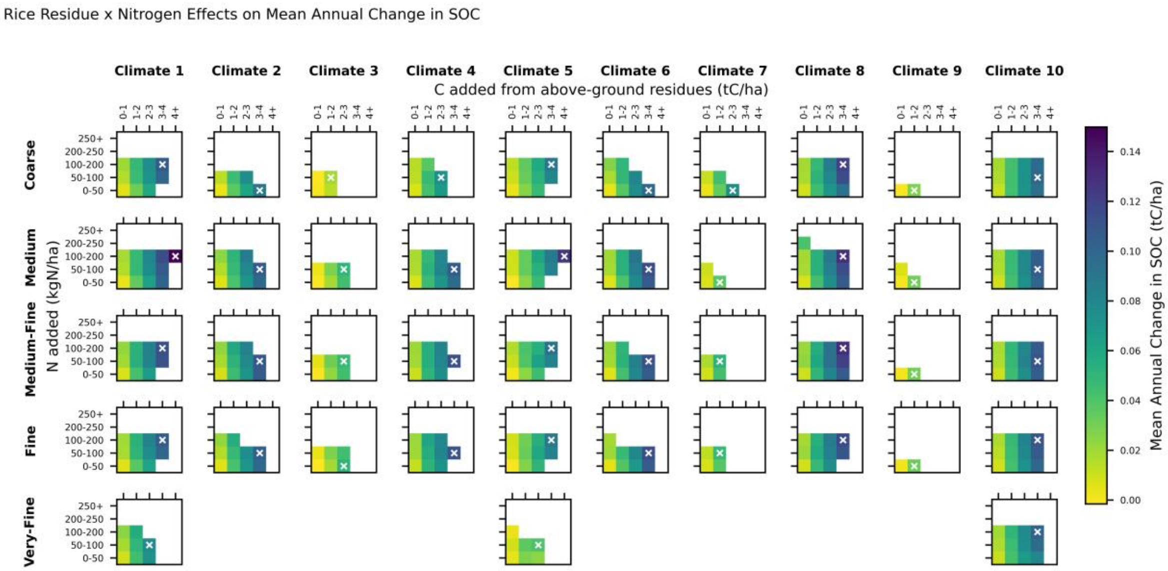

Figure 19: Rice Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

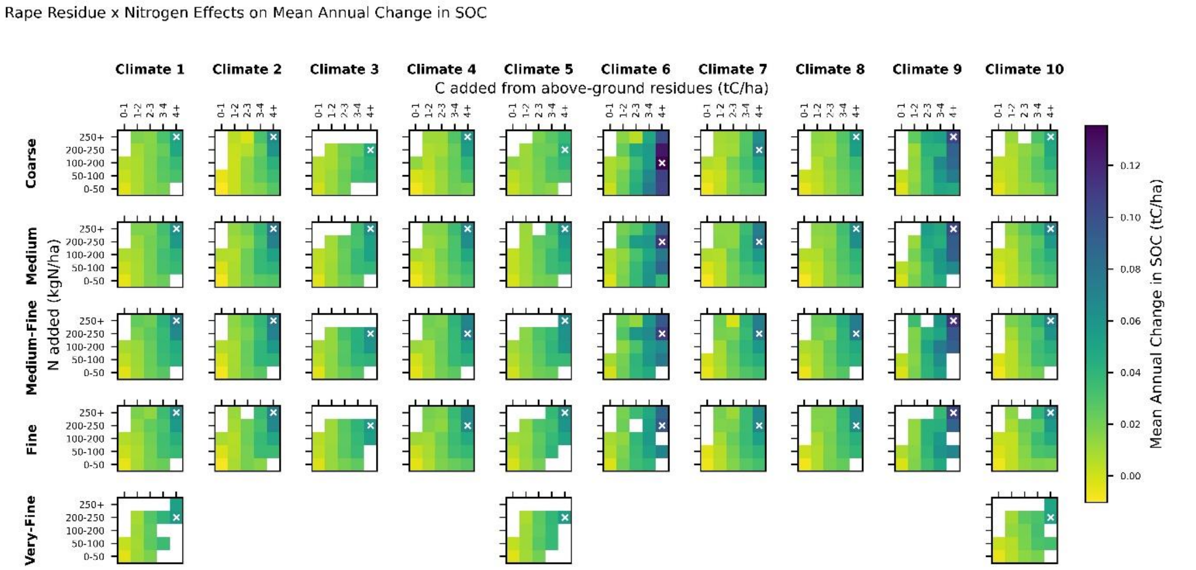

Figure 20: Rape Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

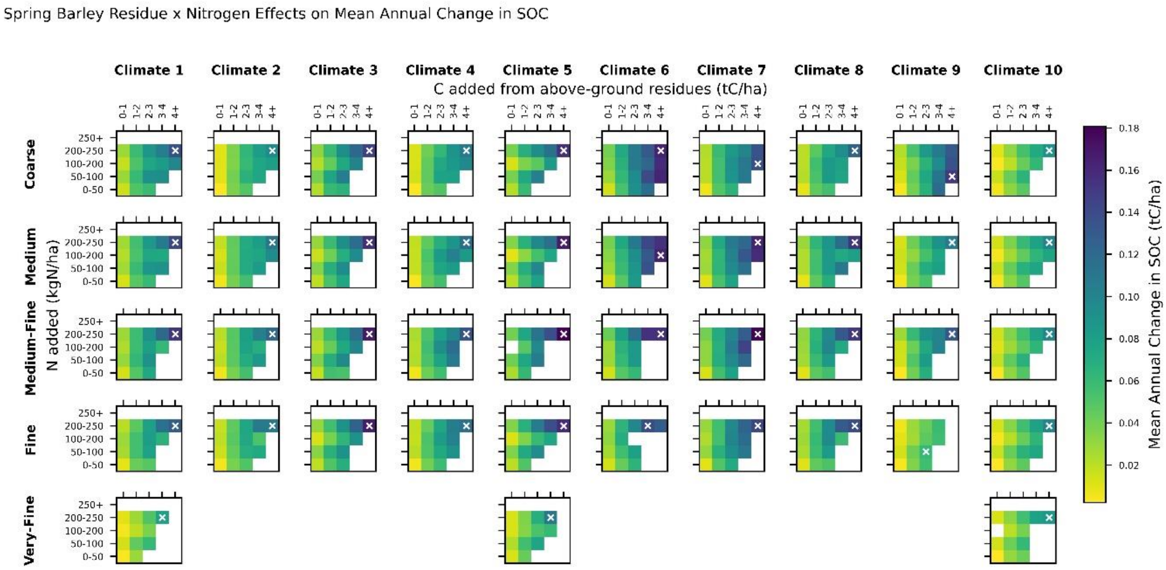

Figure 21: Spring Barley Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

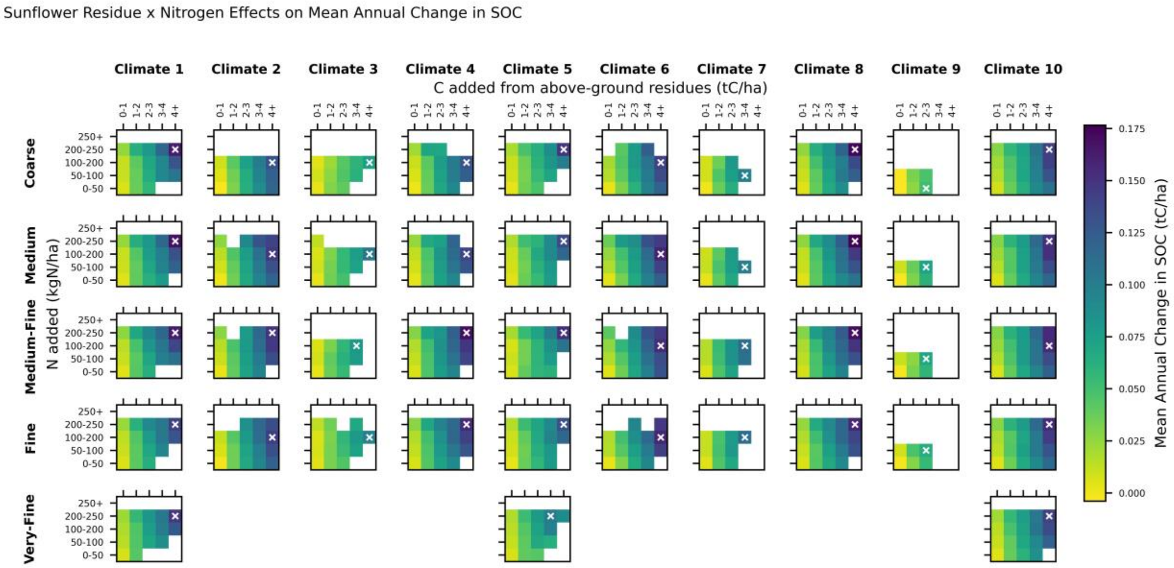

Figure 22: Sunflower Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

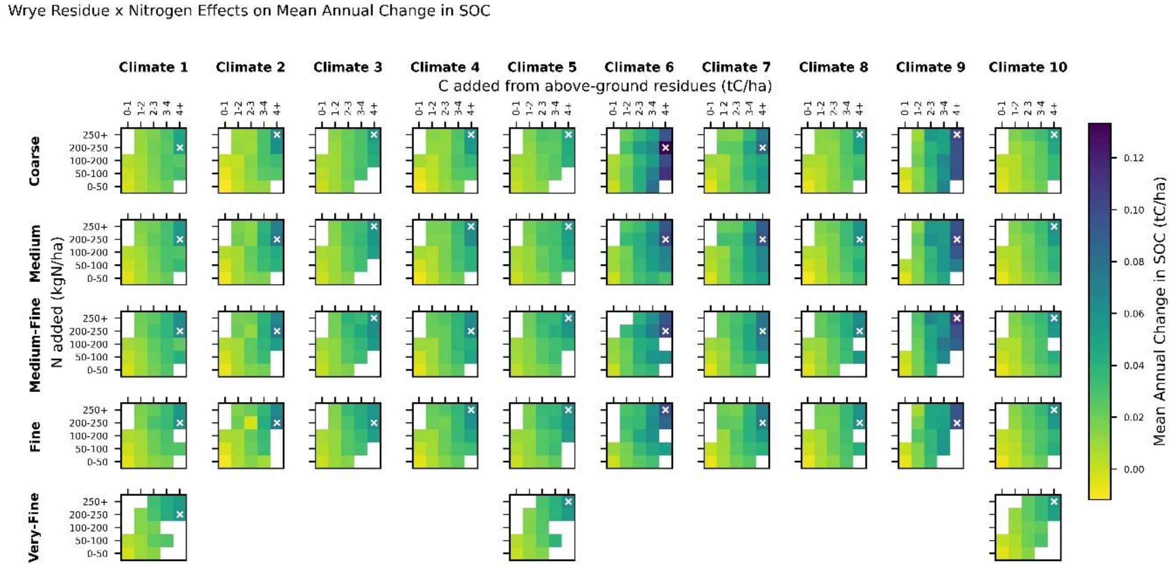

Figure 23: Wrye Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of above-ground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.

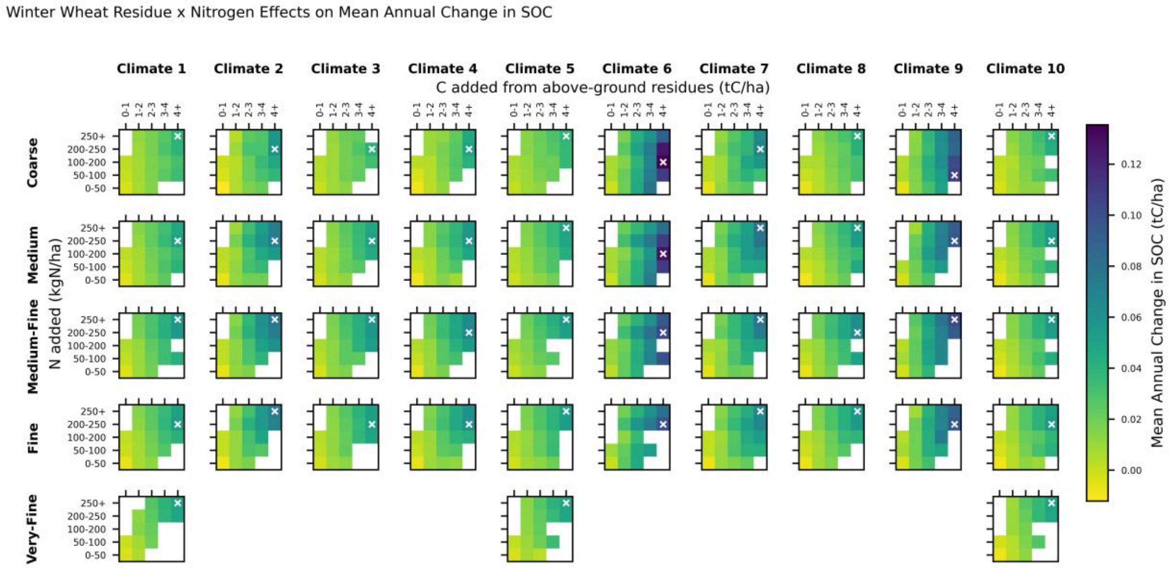

Figure 24: Winter Wheat Heatmap of Mean Annual Change in SOC (tC/ha) achieved with varying applications of aboveground residue C (tC/ha) and nitrogen fertilizer (kgN/ha) for each biophysical cluster. Each biophysical cluster is a unique combination of climate and soil. White X marks the maximum Mean Annual Change in SOC for each biophysical cluster.