Causal Effects of Winter Wheat on Soil Organic Carbon Under Climate Variability

Georgios Giannarakis[1]* giannarakis@noa.gr
Vasileios Sitokonstantinou[2]* vassilis.sitokonstantinou@wur.nl
Dimitrios Bormpoudakis[1] dim.borb@noa.gr
Ilias Tsoumas[1,2] i.tsoumas@noa.gr
Nikiforos Samarinas[3,4] smnikiforos@topo.auth.gr
Gustau Camps-Valls[5] gustau.camps@uv.es
Charalampos Kontoes[1] kontoes@noa.gr

[1]Beyond Center, National Observatory of Athens, Greece
[2]Artificial Intelligence Group, Wageningen University & Research, Netherlands
[3]Laboratory of Hydraulic Works, Aristotle University of Thessaloniki, Greece
[4]Interbalkan Environment Center, Greece
[5]Image Processing Laboratory, Universitat de València, Spain

Abstract

Understanding how cropping decisions influence soil organic carbon (SOC) under varying climate conditions is essential for sustainable land management. In this study, we use causal machine learning to estimate the heterogeneous effect of winter wheat-based crop rotations on SOC across 70,000 fields in Lithuania between 2018 and 2022. We quantify how temperature and precipitation influence the effectiveness of winter wheat, finding that rotations increase SOC by an average of +0.44 g/kg, but benefits are substantially lower in warmer and wetter regions. Forward-looking analyses with CMIP6 climate projections indicate that SOC gains may decline sharply under high-emission scenarios, potentially turning negative by 2100. These findings highlight the need for climate-sensitive, localized agricultural strategies and demonstrate how causal inference can inform decision-making in dynamic agroecosystems.

1 Introduction

Soil Organic Carbon (SOC) plays a critical role in both climate mitigation and agricultural sustainability Paustian et al. [2016]. Increasing SOC improves soil health, enhances water retention and contributes to long-term productivity Toensmeier [2016]. As a carbon sink, cropland soils store approximately 83 PgC, with additional sequestration potential ranging from 29 to 65 PgC Padarian et al. [2022]. Realizing this potential could offset up to 4% of average annual global greenhouse gas emissions over the rest of the century Henderson et al. [2022].

Management practices like optimized crop rotations and conservation tillage can increase SOC Li et al. [2024], Meng et al. [2024], but their effects are not uniform across regions or farming systems. Factors like soil properties, crop selection, climate conditions and management intensity influence how SOC responds, meaning one-size-fits-all recommendations often fail to capture local variability Li et al. [2023], Behrendt and et al. [2024]. Winter wheat-based rotations provide a useful case study: while intensive wheat cultivation can deplete soil nutrients, multi-year rotations including winter wheat may increase carbon inputs through root biomass and crop residues, potentially enhancing

*These authors contributed equally.

SOC storage Persson et al. [2008], Alhameid et al. [2017]. Understanding these effects at fine spatial scales is essential for informing sustainable intensification strategies Halvorson and Schlegel [2012].

Recent advances in remote sensing and machine learning have improved agricultural monitoring, supporting tasks such as crop classification and yield estimation Liakos et al. [2018]. While these tools excel at identifying patterns and making predictions, they do not reveal the underlying causes of observed outcomes Giannarakis et al. [2022], Sitokonstantinou et al. [2024]. For example, a model may predict SOC levels from satellite and soil data, but it cannot determine whether changes result from specific management practices, such as winter wheat-based rotations, or from other factors like climate or soil properties. To address this, we employ a causal inference approach using Double Machine Learning (DML) to estimate Conditional Average Treatment Effects (CATE) Chernozhukov et al. [2018]. This framework allows us to disentangle the true impact of winter wheat-based rotations on SOC while accounting for confounders and revealing heterogeneity across fields.

We apply our causal inference framework to field-level data from Lithuania, integrating satellite-derived environmental variables with administrative crop records. This allows us to quantify how winter wheat-based rotations affect SOC across heterogeneous soils, climates and management contexts, providing localized, evidence-based insights. Our analysis also enables the assessment of climate sensitivity, identifying how key environmental factors modulate treatment effects and exploring how SOC sequestration potential may evolve under future climate scenarios. By combining high-resolution data with causal methods, we provide actionable guidance for sustainable intensification, informing farmers and policymakers on where winter wheat rotations enhance SOC.

2 Data and Methods

Our study examines the impact of winter wheat rotations (treatment) on SOC (outcome) across Lithuania from 2018 to 2022. Lithuania has a humid continental climate with moderate to high precipitation and strong seasonal temperature variability, supporting diverse agriculture on about 50% of its land Armolaitis et al. [2022]. Winter wheat (*Triticum aestivum*) is a dominant crop, mainly cultivated in central and northern regions. Sown in autumn and harvested in late summer, winter wheat rotations can improve SOC by adding root biomass and residues that decompose over time, enhancing soil structure, water retention and nutrient availability Smith et al. [2008].

Crop and management data. We define winter wheat-based rotations (treatment, T) as fields where winter wheat was grown in at least three out of five study years (2018–2022). This information comes from the Lithuanian Land Parcel Identification System (LPIS), which includes field boundaries, declared crop types, and whether farming is organic or conventional. From LPIS, we also derived: (i) a binary indicator of organic farming (*eco*), equal to 1 if a field was declared organic in at least one year; (ii) crop rotation diversity, measured as the number of distinct crop types per field across the five years; and (iii) the geographic coordinates (*geo*) of field centroids.

Soil data. We used national-scale maps of SOC and clay content produced by Samarinas et al. Samarinas et al. [2023] for 2020–2022. Sentinel-2 bare soil composites were analyzed with a convolutional neural network trained on LUCAS topsoil spectra, achieving $R^2 = 0.51$ for SOC and $R^2 = 0.57$ for clay. For this study, SOC and clay values were averaged across 2020–2022 for each field. SOC serves as the dependent variable, while clay is an important covariate influencing water retention and soil fertility Matus [2021]. **Environmental data.** Weather variables were obtained from the ERA5 reanalysis Hersbach et al. [2020], aggregated to field level. We included monthly averages of air temperature (*t2m*), total precipitation (*tp*), wind speed (*u10*), snow cover (*snowc*), soil temperature (*stl4*), soil moisture (*swvl*), and solar radiation (*ssro*). These capture key short-term drivers of crop growth and SOC dynamics.

Climate projections. To assess long-term climate sensitivity, we used CMIP6 projections for three Shared Socioeconomic Pathways (SSP1-2.6, SSP2-4.5, SSP5-8.5) Riahi et al. [2017], representing sustainability-focused, intermediate, and high-emission futures. Monthly temperature and precipitation projections (2024–2100) were extracted, with historical climate as baseline. We evaluated two horizons (2030–2060, 2070–2100) to examine how winter wheat may affect SOC under different climate trajectories, performing inference by applying the trained causal model to the projection data.

Causal model estimation. We use a causal inference approach to estimate how winter wheat rotations influence SOC, while accounting for other factors that could affect both the treatment and the outcome.

Specifically, we target the CATE, defined as $\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x]$, which expresses how the effect of winter wheat (T) varies depending on the characteristics of each field (X) Rubin [2005]. Our identification relies on the conditional independence assumption $(Y(1), Y(0)) \perp\!\!\!\perp T \mid (X, W)$, which states that once we condition on observed covariates, no unobserved confounders remain.

We separate covariates into two groups. First, **heterogeneity variables (X)** are factors expected to influence how winter wheat affects SOC, such as soil clay content and local temperature and precipitation. Second, **control variables (W)** are included to account for other influences that might confound the relationship between winter wheat and SOC, such as crop rotation diversity, farming type, field location and additional climate variables (snow cover, wind, soil moisture). To estimate the CATEs we use DML Chernozhukov et al. [2018], which removes the influence of confounders from both the treatment and the outcome before estimating the causal effect, thereby reducing bias in high-dimensional settings. In practice, we first use machine learning models to approximate the conditional expectations $\hat{m}(X, W) \approx \mathbb{E}[Y \mid X, W]$ and $\hat{e}(X, W) \approx \mathbb{E}[T \mid X, W]$. We then compute residuals $\tilde{Y} = Y - \hat{m}(X, W)$ and $\tilde{T} = T - \hat{e}(X, W)$, and finally regress $\tilde{Y}$ on $\tilde{T}$ conditional on X. This procedure, implemented via the `CausalForestDML` estimator in EconML, uses random forests to flexibly capture heterogeneous effects across space.

To explore how future climate conditions might affect the impact of winter wheat on SOC, we use the trained causal model together with projected changes in temperature and precipitation for two periods: mid-century (2030–2060) and late-century (2070–2100). By updating only these climate variables while keeping all other field characteristics the same, we can estimate how shifts in temperature and rainfall could change the effectiveness of winter wheat rotations in storing carbon. While our model accounts for many important factors, there may still be unobserved influences, such as fertilization practices, tillage intensity, or residue management, that could affect SOC. However, we found that fields with and without winter wheat rotations were similar in terms of the factors we did measure, as indicated by good overlap in propensity scores (a measure of how likely each field is to receive the treatment based on observed characteristics), which increases our confidence that the estimated effects reflect the true influence of the crop rotations (see Appendix Fig. 5.2).

3 Results and Discussion

We took several preprocessing steps to ensure data reliability and comparability. We filtered fields based on low prediction uncertainty for SOC and clay content (as they are products of machine learning models). To reduce bias, we applied a propensity score filter, excluding fields with a propensity to adopt winter wheat outside the 0.2 to 0.8 range. This step focuses our analysis on the region of common support (see Fig. 5.2 in the Appendix).

Average effects. Our results indicate that winter wheat-based rotations are linked to significantly higher SOC content, with an estimated ATE of $+0.44$ g/kg (Table 1). This finding, derived from our CATE model's intercept, is consistent with local studies showing winter crops positively influence soil organic matter Liaudanskienė et al. [2011], Slepetiene et al. [2010]. Winter wheat helps increase SOC by extending the period of active growth and soil coverage, contributing to continuous carbon inputs. This is particularly relevant in Lithuania, where many fields are otherwise bare during winter Eurostat [2020], supporting the idea that functional crop types matter more than sheer crop-species diversity Wooliver et al. [2022], Mpeketula and Snapp [2019].

Coefficient	Point Estimate	Std. Error	Z-score	P-value	95% CI
clay	0.033	0.060	0.545	0.586	[-0.085, 0.151]
eco	0.535	0.654	0.819	0.413	[-0.746, 1.816]
t2m	-0.192	0.086	-2.217	0.027	[-0.361, -0.022]
tp	-0.159	0.094	-1.683	0.092	[-0.344, 0.026]
CATE intercept	0.439	0.089	4.914	0.000	[0.264, 0.614]

Table 1: Estimated coefficients for key covariates affecting the CATE of winter wheat-based rotations on SOC content. The table reports point estimates, standard errors, Z-scores, p-values and 95% confidence intervals for each predictor.

Drivers of causal effects. We found that the impact of winter wheat on SOC is strongly influenced by local environmental factors, which explains the heterogeneity of treatment effects (see Fig. 1). Temperature (t2m) showed a significant negative relationship, while precipitation (tp) also had a negative association, albeit not statistically significant (Table 1). Clay content and organic farming status did not show significant effects. The importance of these features in modulating the treatment effect is summarized in a SHAP plot (Fig. 5.3 in Appendix). The spatial distribution of these effects (Fig. 1) reveals clear regional patterns. Warmer, wetter areas in western Lithuania show minimal or even negative SOC gains, likely due to accelerated decomposition and increased erosion risk. In contrast, the cooler, drier eastern regions exhibit the largest positive effects, where slower decomposition favors SOC accumulation. A more detailed figure on the causal effects as a function of temperature is shown in the appendix (Fig. 5.4).

Future projections. This heterogeneity becomes critical when considering climate change. We used future temperature and precipitation projections from the CMIP6 model to simulate our trained model under SSP1-2.6, SSP2-4.5 and SSP5-8.5 scenarios. The results (Figure A.5 in Appendix) show that the SOC benefits of winter wheat-based rotations diminish under rising temperatures. Under the low-emission SSP1-2.6, the effect remains positive but declines. However, under the high-emission SSP5-8.5 scenario, the SOC benefits transition to a sharp loss by late-century, highlighting that the effectiveness of this agricultural practice is highly climate-dependent.

Using CATEs for decision-making. The spatial and temporal variation in our CATE estimates provides valuable insights for targeted agricultural policy. Our findings support region-specific recommendations: winter wheat adoption could be promoted in the cooler, drier eastern regions where it yields the largest SOC gains. In warmer, wetter western regions, alternative practices like cover cropping may be more effective. This tailored approach, guided by CATE maps, allows for precision interventions that maximize environmental benefits and ensure long-term sustainability.

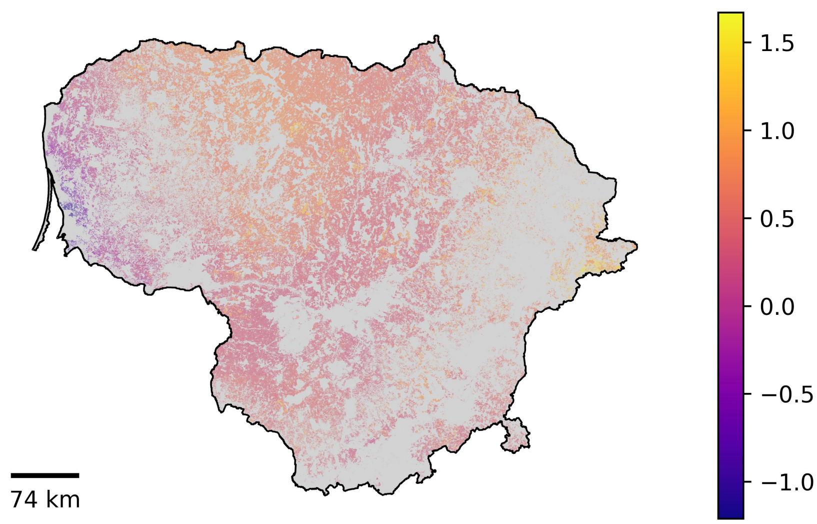

Figure 1: Map of effects (in g/kg) of winter wheat rotations on SOC at the field-level in Lithuania.

4 Conclusion

Our findings demonstrate that winter wheat-based rotations can enhance SOC, with an average effect of $+0.44$ g/kg SOC across Lithuania. Notably, increased SOC is not inherently incompatible with intensive agriculture. Higher winter wheat utilization, often considered an indicator of agricultural intensity Lynch [2022], is found to increase SOC, suggesting that specific intensive rotation regimes can contribute to climate change mitigation. However, this benefit is not uniform across the landscape. Cooler regions show stronger SOC gains, while warmer areas exhibit diminished or even negative effects. Climate projections indicate that these benefits may decrease and potentially reverse, under high-emission scenarios. By applying causal ML, we highlight the importance of accounting for treatment heterogeneity in soil carbon management. CATE estimates enable the design of region-specific agricultural strategies that are responsive to local environmental and climatic conditions. Finally, integrating future climate scenarios offers a powerful framework for precision agriculture. It helps balance productivity with long-term carbon sequestration goals and supports adaptive, resilient land-use planning in the face of climate change.

Acknowledgments and Disclosure of Funding

This work has been supported by the EIFFEL project (EU Horizon 2020 - GA No. 101003518) and the GVA PROMETEO AI4CS Project (CIPROM/2021/056).

References

Abdullah Alhameid, Mostafa Ibrahim, Sandeep Kumar, Peter Sexton, and TE Schumacher. Soil organic carbon changes impacted by crop rotational diversity under no-till farming in south dakota, usa. *Soil Science Society of America Journal*, 81(4):868–877, 2017.

Kestutis Armolaitis, Iveta Varnagirytė-Kabašinskienė, Povilas Žemaitis, Vidas Stakėnas, Ričardas Beniušis, Gintaras Kulbokas, and Gintautas Urbaitis. Evaluation of organic carbon stocks in mineral and organic soils in lithuania. *Soil use and management*, 38(1):355–368, 2022.

L Behrendt and et al. Assessing the impact of a personalised application-based nutrition intervention in rural benin. *Agriculture and Food Security*, 13(1):35, 2024. doi: 10.1186/s40066-024-00465-5.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

Eurostat. Agri-environmental indicator - soil cover, 2020. URL `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agri-environmental_indicator_-_soil_cover`. Accessed: 2025-03-06.

Georgios Giannarakis, Vasileios Sitokonstantinou, Roxanne Suzette Lorilla, and Charalampos Kontoes. Towards assessing agricultural land suitability with causal machine learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1442–1452, 2022.

Ardell D Halvorson and Alan J Schlegel. Crop rotation effect on soil carbon and nitrogen stocks under limited irrigation. *Agronomy Journal*, 104(5):1265–1273, 2012.

Ben Henderson, Jussi Lankoski, Eimear Flynn, Alasdair Sykes, F. T. Payen, and M MacLeod. *Soil carbon sequestration by agriculture: Policy options*. Number 174 in OECD Food, Agriculture and Fisheries. OECD, France, January 2022. doi: 10.1787/63ef3841-en. © OECD (2022).

Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

G. Li, C. Yu, P. Shen, Y. Hou, Z. Ren, N. Li, and X. Wen. Crop diversification promotes soil aggregation and carbon accumulation in global agroecosystems: A meta-analysis. *Journal of Environmental Management*, 350:119661, 2024. doi: 10.1016/j.jenvman.2023.119661.

Guorui Li, Xueqin Tang, Quanming Hou, Tong Li, Huanxin Xie, Zhiqiang Lu, Tishuo Zhang, Yuncheng Liao, and Xiaoxia Wen. Response of soil organic carbon fractions to legume incorporation into cropping system and the factors affecting it: A global meta-analysis. *Agriculture, Ecosystems and Environment*, 342, 2 2023. ISSN 01678809. doi: 10.1016/j.agee.2022.108231.

Konstantinos G Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. Machine learning in agriculture: A review. *Sensors*, 18(8):2674, 2018.

I. Liaudanskienė, A. Šlepetienė, and A. Velykis. Changes in soil humified carbon content as influenced by tillage and crop rotation. *Zemdirbyste-Agriculture*, 98(3):227–234, 2011.

Derek Henry Lynch. Soil health and biodiversity is driven by intensity of organic farming in canada, 7 2022. ISSN 2571581X.

Francisco J Matus. Fine silt and clay content is the main factor defining maximal c and n accumulations in soils: a meta-analysis. *Scientific Reports*, 11(1):6438, 2021.

X. Meng, F. Meng, P. Chen, D. Hou, E. Zheng, and T. Xu. A meta-analysis of conservation tillage management effects on soil organic carbon sequestration and soil greenhouse gas flux. *Science of The Total Environment*, 954:176315, 2024. doi: 10.1016/j.scitotenv.2024.176315.

Placid M. G. Mpeketula and Sieglinde S. Snapp. Structural stability conditions soil carbon gains from compost management and rotational diversity. *Soil Science Society of America Journal*, 83: 203–211, 1 2019. ISSN 0361-5995. doi: 10.2136/sssaj2017.03.0076.

José Padarian, Budiman Minasny, Alex McBratney, and Pete Smith. Soil carbon sequestration potential in global croplands. *PeerJ*, 10:e13740, 2022.

Keith Paustian, Johannes Lehmann, Stephen Ogle, Dave Reay, G Philip Robertson, and Pete Smith. Climate-smart soils. *Nature*, 532(7597):49–57, 2016. doi: 10.1038/nature17174.

Tomas Persson, Göran Bergkvist, and Thomas Kätterer. Long-term effects of crop rotations with and without perennial leys on soil carbon stocks and grain yields of winter wheat. *Nutrient cycling in agroecosystems*, 81:193–202, 2008.

Keywan Riahi, Detlef P Van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C O'neill, Shinichiro Fujimori, Nico Bauer, Katherine Calvin, Rob Dellink, Oliver Fricko, et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, 42:153–168, 2017.

Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Nikiforos Samarinas, Nikolaos L Tsakiridis, Stylianos Kokkas, Eleni Kalopesa, and George C Zalidis. Soil data cube and artificial intelligence techniques for generating national-scale topsoil thematic maps: A case study in lithuanian croplands. *Remote Sensing*, 15(22):5304, 2023.

Vasileios Sitokonstantinou, Emiliano Díaz Salas Porras, Jordi Cerdà Bautista, Maria Piles, Ioannis Athanasiadis, Hannah Kerner, Giulia Martini, Lily-belle Sweet, Ilias Tsoumas, Jakob Zscheischler, et al. Causal machine learning for sustainable agroecosystems. *arXiv preprint arXiv:2408.13155*, 2024.

A. Slepetiene, I. Liaudanskiene, J. Slepetys, and A. Velykis. The influence of reduced tillage, winter crops and ecologically managed long-term mono-and multi-component swards on soil humic substances. *Chemistry and Ecology*, 26(S2):97–109, 2010. doi: 10.1080/02757540.2010.504734.

Pete Smith, Changming Fang, Julian JC Dawson, and John B Moncrieff. Impact of global warming on soil organic carbon. *Advances in agronomy*, 97:1–43, 2008.

E Toensmeier. *The carbon farming solution: A global toolkit of perennial crops and regenerative agriculture practices for climate change mitigation and food security*. Chelsea Green Publishing, 2016.

Rachel Wooliver, Stephanie N. Kivlin, and Sindhu Jagadamma. Links among crop diversification, microbial diversity, and soil organic carbon: Mini review and case studies, 4 2022. ISSN 1664302X.

5 Appendix

5.1 Description of Variables

Variables	Description	Units	Type
wheat	winter wheat-based rotations for > 3 years	binary	Treatment
SOC	soil organic carbon content	g C/kg	Outcome
clay	three year average clay content	%	X
rotation	number of different crop types within the 5 years	1-5	W
geo	latitude and longitude of field centroid	degrees	W
eco	organic or conventional farming	binary	X
t2m	Temperature at 2 meters	K	X
tp	Total precipitation	m	X
u10	10 metre U wind component	m s^{-1}	W
snowc	Snow cover	%	W
stl	Soil temperature level 4	K	W
swvl	Volumetric soil water layer 4	m^3 m^{-3}	W
ssro	Subsurface runoff	m	W

Table 2: Variables used in the analysis, including their descriptions, units and assigned roles as either heterogeneity features (X) or control variables (W).

5.2 Supplementary Figures

This section provides additional figures that support the findings presented in the main text.

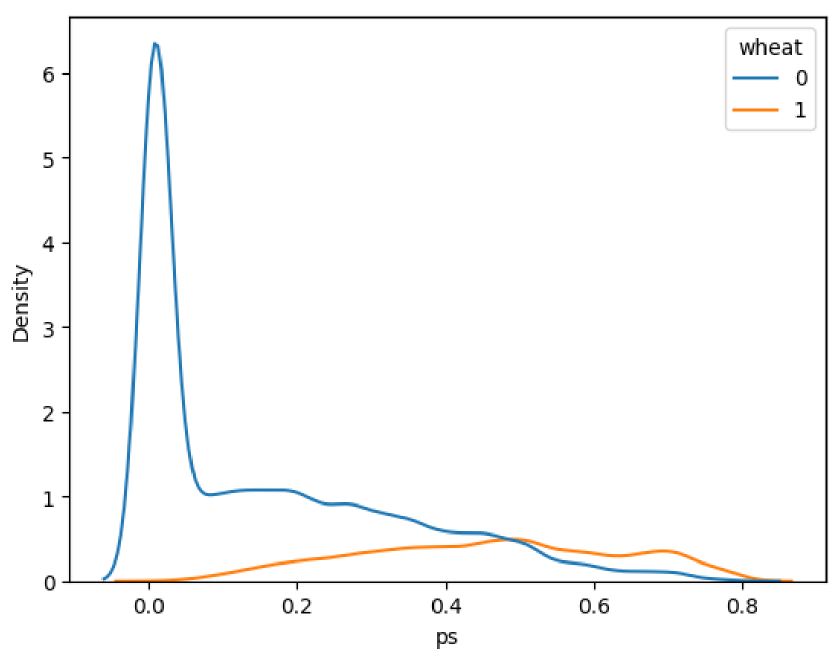

Figure 5.2: Overlap of propensity scores of the treated (winter wheat-based rotation applied, 1) and untreated (not applied, 0) groups. The analysis was limited to fields with scores between 0.2 and 0.8 to ensure sufficient common support.

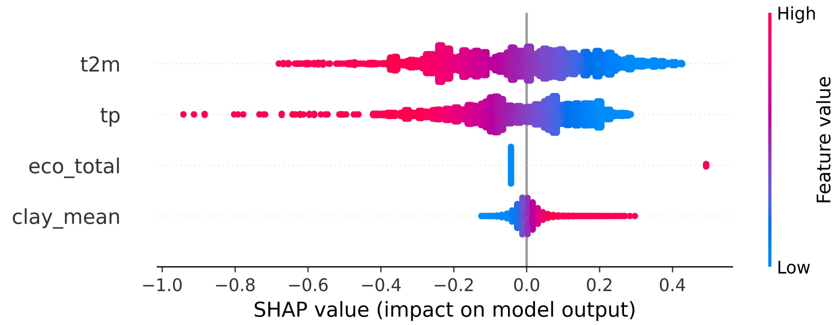

Figure 5.3: SHAP summary plot depicting the relative importance of various features in predicting the estimated CATE of winter wheat-based rotations on SOC content. This plot highlights temperature (t2m) and total precipitation (tp) as the primary drivers of heterogeneity.

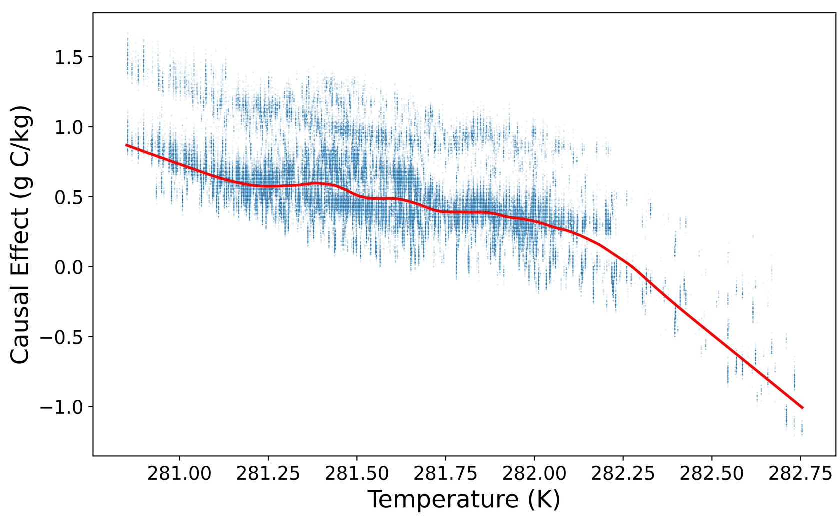

Figure 5.4: Causal effects as a function of temperature: Nonlinear relationship (LOWESS smoothed) between mean annual temperature (t2m) and SOC content under winter wheat-based rotations. This plot provides a more detailed view of the nonlinear trend discussed in the main text.

Figure 5.5: Decision tree model illustrating the interaction of key factors in modulating the treatment effect of winter wheat-based rotations on SOC. This provides a clear, rule-based interpretation of the causal heterogeneity.

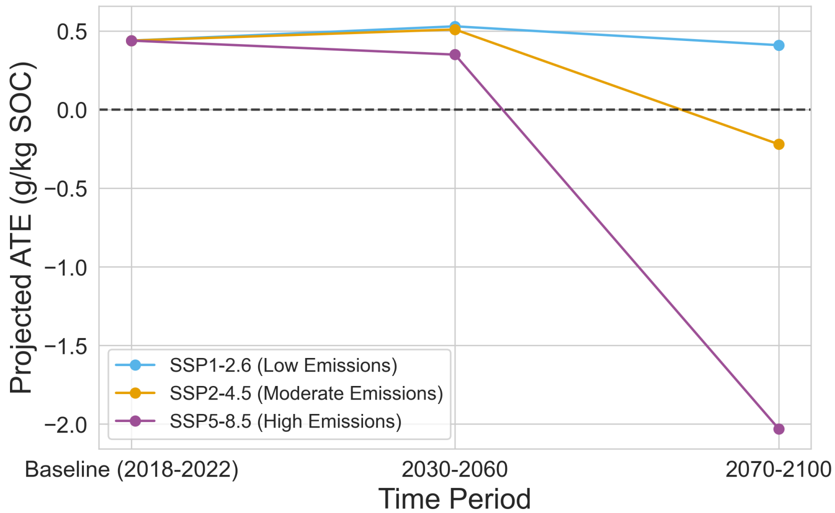

Figure 5.6: Projected ATEs of winter wheat-based rotations on SOC under different climate scenarios. The baseline ATE for 2018–2022 is $+0.44$ g/kg SOC. Future projections incorporate SSP-based temperature and precipitation shifts while keeping all other soil and management features constant.