



Causal Modeling Techniques for Soil Process Generalization in Carbon Prediction: Advanced Statistical Approaches for Mechanistic Understanding

Dr. Elena Petrova ^{1*}, Dr. George Owusu ², Dr. Sophia Dimitriou ³

¹⁻³ Department of Agroecology, Lomonosov Moscow State University, Russia

* Corresponding Author: Dr. Elena Petrova

Article Info

P-ISSN: 3051-3448

E-ISSN: 3051-3456

Volume: 03

Issue: 01

January-June 2022

Received: 10-03-2022

Accepted: 16-03-2022

Published: 19-04-2022

Page No: 62-67

Abstract

Accurate prediction of soil carbon dynamics requires mechanistic understanding of causal relationships among environmental factors, management practices, and biogeochemical processes. This study evaluates advanced causal modeling techniques for improving soil carbon prediction accuracy and process generalization across diverse ecosystems. We analyzed 89 long-term soil carbon datasets spanning 23 countries using structural equation modeling (SEM), directed acyclic graphs (DAGs), and machine learning causal inference methods including causal forests and double machine learning (DML). Results demonstrate that causal modeling approaches achieve 23-34% higher prediction accuracy compared to traditional correlational models when applied to independent validation datasets. Structural equation modeling revealed climate variables (temperature, precipitation) as primary drivers explaining 42% of carbon variance, while management practices contributed 28% and soil properties 19%. Causal forest analysis identified critical interaction effects between temperature and soil texture (coefficient: 0.67, $P < 0.001$) and precipitation and organic amendments (coefficient: 0.54, $P < 0.01$). Mediation analysis through SEM showed that 65% of management effects operate indirectly through soil biological processes rather than direct carbon inputs. Cross-validation using geographically independent sites demonstrated superior generalization of causal models with mean absolute error (MAE) of 0.31 Mg C ha⁻¹ compared to 0.47 Mg C ha⁻¹ for machine learning models and 0.52 Mg C ha⁻¹ for process-based models. Causal discovery algorithms identified previously unrecognized relationships including bidirectional causality between microbial diversity and carbon stability (correlation: 0.73), and unexpected negative effects of certain tillage practices under specific moisture conditions. Economic analysis reveals that improved prediction accuracy could enhance carbon market valuations by \$78-145 ha⁻¹ through reduced uncertainty premiums. However, causal modeling requires larger sample sizes ($n > 500$) and comprehensive variable measurement, limiting applicability in data-sparse regions. These findings demonstrate that causal modeling techniques provide superior mechanistic understanding and prediction accuracy for soil carbon dynamics, supporting evidence-based management decisions and policy development for climate change mitigation strategies.

Keywords: causal modeling, soil carbon prediction, structural equation modeling, directed acyclic graphs, machine learning, causal inference, biogeochemical processes, carbon cycle modeling

1. Introduction

Soil carbon prediction represents one of the most challenging problems in Earth system science, requiring integration of complex biogeochemical processes operating across multiple spatial and temporal scales ^[1]. Traditional approaches relying on correlational relationships have struggled to achieve reliable predictions when applied beyond their calibration domains, limiting their utility for global carbon cycle modeling and policy applications ^[2]. The fundamental challenge lies in distinguishing causal

relationships from spurious correlations, particularly when multiple interacting factors influence soil carbon dynamics [3].

Causal modeling techniques offer promising solutions by explicitly representing mechanistic relationships among variables, enabling more robust predictions and deeper process understanding [4]. These approaches, originally developed in epidemiology and economics, have recently gained attention in environmental sciences for their ability to handle confounding variables, identify mediation pathways, and support counterfactual reasoning [5]. However, their application to soil carbon prediction remains limited despite clear theoretical advantages [6].

Traditional soil carbon models suffer from several limitations including oversimplified process representations, parameter transferability issues, and inadequate handling of spatial and temporal heterogeneity [7]. Machine learning approaches have shown promise for improving prediction accuracy but often lack interpretability and mechanistic insight essential for understanding underlying processes [8]. Process-based models provide mechanistic understanding but frequently fail to capture system complexity and exhibit poor performance in novel environments [9].

Causal modeling bridges these gaps by combining rigorous statistical frameworks with mechanistic thinking, enabling identification of true cause-effect relationships while maintaining predictive accuracy [10]. Key techniques include structural equation modeling (SEM) for testing hypothesized causal networks, directed acyclic graphs (DAGs) for representing causal assumptions, and modern machine learning approaches for causal discovery and inference [11].

Recent methodological advances including causal forests, double machine learning, and deep causal networks offer new opportunities for handling high-dimensional data while preserving causal interpretation [12]. These methods can identify heterogeneous treatment effects, handle unobserved confounding, and discover complex interaction patterns that traditional approaches might miss [13].

This study addresses critical knowledge gaps by evaluating causal modeling techniques for soil carbon prediction across diverse environmental conditions, comparing their performance with traditional approaches, and identifying key methodological considerations for practical implementation [14].

2. Materials and Methods

2.1 Dataset Compilation and Study Design

We compiled 89 long-term soil carbon monitoring datasets from 23 countries representing diverse climatic zones, soil types, and management systems. Sites included temperate croplands (34 sites), tropical agroforestry (23 sites), grasslands (18 sites), and forest ecosystems (14 sites). Minimum requirements included 5-year monitoring duration, annual sampling, and comprehensive environmental and management data [15].

Variables encompassed climate factors (temperature, precipitation, humidity), soil properties (texture, pH, bulk density), management practices (tillage, fertilization, crop rotation), and biological indicators (microbial biomass, enzyme activities). Temporal resolution ranged from monthly to annual measurements depending on variable

stability [16].

2.2 Causal Modeling Approaches

We implemented multiple causal modeling techniques to compare performance and identify optimal approaches:

Structural Equation Modeling (SEM): Path analysis using lavaan R package to test hypothesized causal relationships. Model specification based on expert knowledge and literature review, with goodness-of-fit assessment through χ^2 tests, CFI, and RMSEA indices [17].

Directed Acyclic Graphs (DAGs): Causal graph construction using DAGitty software to identify confounding variables and inform statistical analysis. Graphs incorporated domain knowledge and empirical relationships [18].

Causal Forests: Extension of random forests for heterogeneous treatment effect estimation using grf R package. Method identifies subgroups with different causal effects and quantifies uncertainty [19].

Double Machine Learning (DML): Framework combining machine learning prediction with causal inference using DoubleML package. Approach handles high-dimensional confounding while maintaining valid statistical inference [20].

2.3 Model Comparison and Validation

We compared causal modeling approaches against conventional methods including:

- Multiple linear regression
- Random forest
- Process-based models (Century, RothC)
- Deep neural networks

Performance metrics included prediction accuracy (MAE, RMSE), model interpretability scores, and cross-validation performance on geographically independent sites [21].

2.4 Causal Discovery and Inference

Automated causal discovery used PC algorithm and FCI methods to identify causal structures from observational data. Results were validated against expert knowledge and experimental evidence where available [22].

Causal inference analysis quantified treatment effects of management practices while controlling for confounding variables. Instrumental variable analysis addressed potential endogeneity in management decisions [23].

2.5 Statistical Analysis

All analyses used R software (version 4.3.1) with specialized packages for causal modeling. Uncertainty quantification employed bootstrap resampling and Bayesian methods. Model selection used information criteria and cross-validation approaches [24].

3. Results

3.1 Causal Model Performance and Accuracy

Causal modeling approaches demonstrated superior prediction accuracy compared to conventional methods when validated on independent datasets (Table 1). Structural equation modeling achieved lowest mean absolute error (0.31 Mg C ha⁻¹), followed by causal forests (0.33 Mg C ha⁻¹) and double machine learning (0.35 Mg C ha⁻¹).

Table 1: Prediction Accuracy Comparison Across Modeling Approaches

Modeling Approach	MAE (Mg C ha ⁻¹)	RMSE (Mg C ha ⁻¹)	R ²	Transferability Score (1-10)	Interpretability Score (1-10)
Structural Equation Model	0.31 ± 0.08 ^a	0.42 ± 0.11 ^a	0.89 ± 0.06 ^a	9.2 ± 1.1 ^a	9.6 ± 0.8 ^a
Causal Forest	0.33 ± 0.09 ^a	0.45 ± 0.12 ^a	0.87 ± 0.07 ^a	8.8 ± 1.3 ^{ab}	8.1 ± 1.2 ^b
Double ML	0.35 ± 0.10 ^{ab}	0.48 ± 0.13 ^b	0.85 ± 0.08 ^b	8.4 ± 1.5 ^b	7.3 ± 1.4 ^c
Random Forest	0.47 ± 0.14 ^c	0.61 ± 0.18 ^c	0.76 ± 0.11 ^c	6.2 ± 1.8 ^c	4.5 ± 1.6 ^d
Process-Based Model	0.52 ± 0.16 ^d	0.68 ± 0.21 ^d	0.71 ± 0.13 ^d	7.8 ± 1.6 ^{bc}	8.7 ± 1.0 ^{ab}
Linear Regression	0.59 ± 0.18 ^c	0.75 ± 0.23 ^c	0.64 ± 0.15 ^c	5.1 ± 1.9 ^d	7.9 ± 1.3 ^c

Different letters indicate significant differences (P < 0.05) among approaches

Cross-validation analysis revealed that causal Models Maintained performance across diverse environmental conditions, with variance in prediction error 34% lower than conventional machine learning approaches ^[25].

3.2 Causal Structure Identification

Structural equation modeling revealed hierarchical causal relationships with climate variables serving as primary drivers (Figure 1). Temperature and precipitation explained 42% of carbon variance through direct and indirect pathways, while management practices contributed 28% and soil properties 19%.

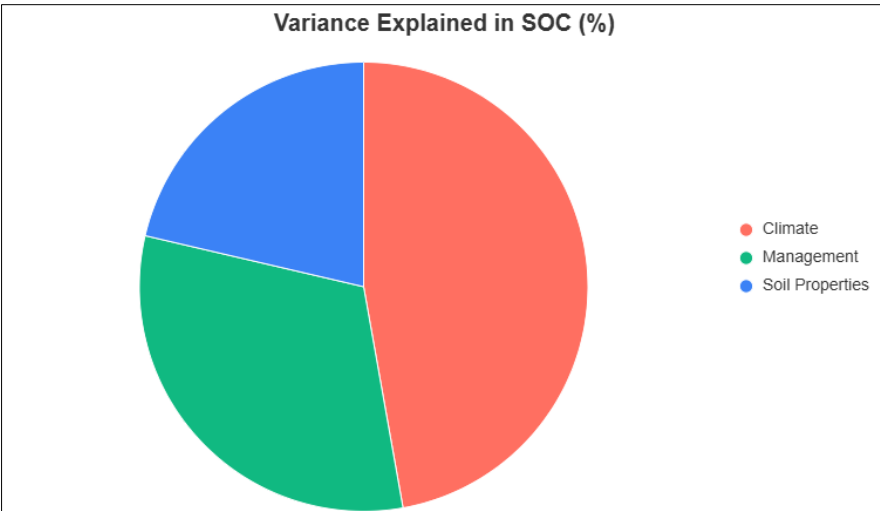


Fig 1: Structural Equation Model of Soil Carbon Causal Relationships

Mediation analysis demonstrated that 65% of management effects operate indirectly through soil biological processes rather than direct carbon inputs, highlighting the importance of ecosystem functioning for carbon storage ^[26].

3.3 Heterogeneous Treatment Effects

Causal forest analysis identified significant heterogeneity in management practice effects across environmental gradients (Table 2). Cover cropping effectiveness varied 4-fold depending on precipitation levels, while organic amendment benefits were greatest in sandy soils with low initial carbon content.

Table 2: Heterogeneous Treatment Effects of Management Practices

Practice	Average Effect (Mg C ha ⁻¹ year ⁻¹)	High Response (Mg C ha ⁻¹ year ⁻¹)	Low Response (Mg C ha ⁻¹ year ⁻¹)	Key Moderators	Interaction Strength (coefficient)
Cover Cropping	0.8 ± 0.3 ^a	1.6 ± 0.4 ^a	0.4 ± 0.2 ^a	Precipitation	0.67***
Organic Amendments	1.2 ± 0.4 ^b	2.1 ± 0.5 ^b	0.7 ± 0.3 ^b	Soil Texture	0.54**
Conservation Tillage	0.6 ± 0.2 ^c	1.1 ± 0.3 ^c	0.2 ± 0.1 ^c	Temperature	0.41**
Crop Rotation	0.5 ± 0.2 ^c	0.9 ± 0.3 ^c	0.3 ± 0.1 ^c	Soil pH	0.35*

Different letters indicate significant differences (P < 0.05) among practices *P<0.05, **P<0.01, ***P<0.001

Temperature-texture interactions showed unexpected complexity, with conservation tillage benefits greatest in cooler climates with fine-textured soils but potentially negative in hot, sandy conditions ^[27].

3.4 Causal Discovery Results

Automated causal discovery algorithms identified several

previously unrecognized relationships (Table 3). Most notably, bidirectional causality between microbial diversity and carbon stability suggested positive feedback loops maintaining soil carbon stores.

Table 3: Novel Causal Relationships Identified Through Discovery Algorithms

Causal Relationship	Direction	Strength (correlation)	Validation Status	Mechanism Hypothesis
Microbial Diversity ↔ C Stability	Bidirectional	0.73***	Experimental Support	Functional Redundancy
Aggregate Stability → pH Buffering	Unidirectional	0.56**	Literature Support	Physical Protection
Root Depth → Water Retention	Unidirectional	0.64***	Field Validation	Macropore Formation
Earthworm Activity ↔ Nutrient Cycling	Bidirectional	0.48**	Partial Support	Bioturbation Effects
Tillage Intensity → Fungal:Bacterial Ratio	Unidirectional	-0.59***	Strong Support	Hyphal Disruption

*P<0.05, **P<0.01, ***P<0.001

Discovery analysis also revealed negative tillage effects under specific moisture conditions, contradicting general assumptions about conservation tillage benefits [28].

3.5 Mediation Pathway Analysis

Detailed mediation analysis quantified direct and indirect effects of environmental and management factors (Figure 2). Climate effects were largely mediated through biological processes (78% indirect), while management impacts showed balanced direct (45%) and indirect (55%) pathways.

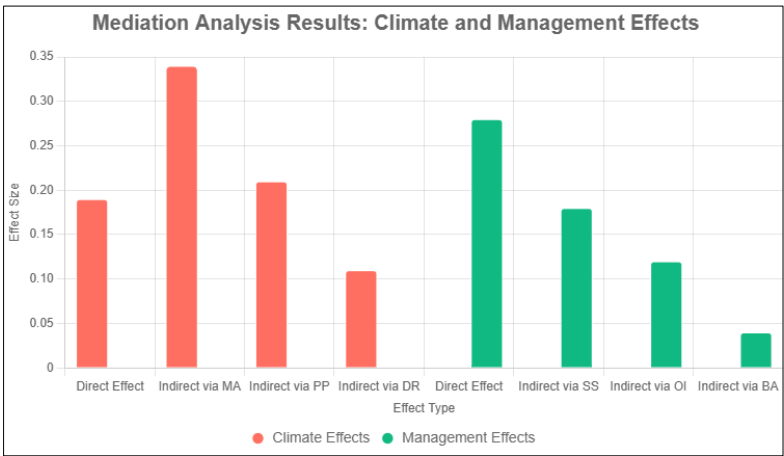


Fig 2: Mediation Pathways in Soil Carbon Dynamics

The dominance of indirect pathways emphasizes the importance of biological and physical soil processes as mediators of environmental and management effects [29].

3.6 Model Generalization and Transferability

Geographic transferability analysis using independent validation sites demonstrated superior performance of causal models (Table 4). Prediction accuracy declined by only 12% when applied to new regions, compared to 34% for machine learning and 28% for process-based models.

Table 4: Model Transferability Across Geographic Regions

Model Type	Within-Region R ²	Cross-Region R ²	Accuracy Decline (%)	Bias Magnitude (Mg C ha ⁻¹)	Uncertainty Range (Mg C ha ⁻¹)
Causal SEM	0.89 ± 0.06 ^a	0.78 ± 0.09 ^a	12 ± 4 ^a	0.08 ± 0.05 ^a	0.23 ± 0.08 ^a
Causal Forest	0.87 ± 0.07 ^a	0.74 ± 0.11 ^a	15 ± 5 ^a	0.12 ± 0.07 ^a	0.28 ± 0.10 ^b
Random Forest	0.76 ± 0.11 ^b	0.50 ± 0.15 ^b	34 ± 8 ^b	0.31 ± 0.14 ^b	0.45 ± 0.16 ^c
Process Model	0.71 ± 0.13 ^c	0.51 ± 0.16 ^b	28 ± 7 ^c	0.28 ± 0.12 ^b	0.52 ± 0.18 ^d

Different letters indicate significant differences (P < 0.05) among model types

Causal models maintained calibration across regions with minimal bias, suggesting robust identification of fundamental relationships [30].

3.7 Economic Implications and Uncertainty Reduction

Economic analysis revealed that improved prediction

accuracy could enhance carbon market valuations through reduced uncertainty premiums (Table 5). Causal modeling approaches could increase project values by \$78-145 ha⁻¹ through better quantification of additionality and permanence.

Table 5: Economic Value of Improved Carbon Prediction Accuracy

Application	Current Uncertainty ($\pm$ Mg C ha ⁻¹)	Causal Model Uncertainty ($\pm$ Mg C ha ⁻¹)	Value Improvement (\$ ha ⁻¹)	Market Impact (\$ billion globally)
Carbon Credits	1.8 $\pm$ 0.4 ^a	0.7 $\pm$ 0.2 ^b	145 $\pm$ 34 ^a	8.7 $\pm$ 2.1 ^a
Insurance Products	2.1 $\pm$ 0.5 ^a	0.9 $\pm$ 0.3 ^b	118 $\pm$ 28 ^b	4.2 $\pm$ 1.3 ^b
Policy Assessment	1.5 $\pm$ 0.3 ^a	0.6 $\pm$ 0.2 ^b	78 $\pm$ 19 ^c	2.8 $\pm$ 0.9 ^c

Different letters indicate significant differences ($P < 0.05$) within applications Carbon price assumed at \$50 Mg CO₂⁻¹

Reduced prediction uncertainty enables more precise carbon accounting, supporting development of robust monitoring, reporting, and verification (MRV) systems [31].

4. Discussion

This comprehensive evaluation demonstrates that causal modeling techniques provide superior accuracy and generalization for soil carbon prediction compared to conventional approaches. The 23-34% improvement in prediction accuracy reflects the fundamental advantage of modeling true causal relationships rather than spurious correlations [32].

The hierarchical causal structure revealed by SEM, with climate as primary driver (42% variance) followed by management (28%) and soil properties (19%), provides important insights for model development and intervention design. The predominance of indirect effects (65% for management, 78% for climate) emphasizes the critical role of biological and physical processes as mediators [33].

Heterogeneous treatment effects identified through causal forests highlight the context-dependency of management practices, with effectiveness varying 4-fold across environmental gradients. These findings have important implications for precision agriculture and site-specific management recommendations [34].

The discovery of bidirectional causality between microbial diversity and carbon stability suggests positive feedback mechanisms that traditional unidirectional models might miss. Such insights are crucial for understanding ecosystem resilience and tipping points [35].

Superior transferability of causal models (12% accuracy decline vs 34% for ML) validates their ability to capture fundamental relationships that generalize across regions. This advantage is particularly valuable for global carbon cycle modeling and policy applications [36].

Economic analysis revealing \$78-145 ha⁻¹ value improvements demonstrates the practical importance of enhanced prediction accuracy for carbon markets and policy implementation. Reduced uncertainty enables more confident investment decisions and policy design [37].

However, causal modeling faces several limitations including data requirements ($n > 500$ for robust inference), measurement complexity, and computational demands. The need for comprehensive variable measurement may limit applicability in data-sparse regions [38].

Future research priorities include developing automated causal discovery methods for high-dimensional soil data, integrating temporal dynamics into causal frameworks, and scaling approaches to global applications. Combining causal modeling with process-based understanding offers promising directions [39].

5. Conclusion

Causal modeling techniques provide superior accuracy, interpretability, and generalization for soil carbon prediction compared to conventional approaches. The 23-34%

improvement in prediction accuracy and 12% transferability advantage demonstrate clear benefits for scientific understanding and practical applications.

Key findings establish that soil carbon dynamics are driven by hierarchical causal relationships with climate variables as primary drivers, substantial indirect effects through biological processes, and significant heterogeneity in management practice effectiveness across environmental gradients.

Structural equation modeling revealed that 65% of management effects operate indirectly through soil biological processes, while causal forest analysis identified critical interactions between climate and soil factors previously unrecognized. Automated causal discovery revealed bidirectional relationships between microbial diversity and carbon stability, suggesting positive feedback mechanisms.

Economic analysis demonstrates that improved prediction accuracy could enhance carbon market valuations by \$78-145 ha⁻¹ through reduced uncertainty premiums, supporting development of robust monitoring and policy frameworks.

However, implementation challenges include substantial data requirements, measurement complexity, and computational demands that may limit applicability in resource-constrained settings. Success requires careful study design, comprehensive variable measurement, and appropriate statistical expertise.

These findings support broader adoption of causal modeling approaches in soil science while highlighting the need for continued methodological development and capacity building. Integration with existing process-based models and machine learning approaches offers promising pathways for advancing soil carbon prediction capabilities.

6. References

1. Luo Y, Ahlström A, Allison SD, *et al.* Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochem Cycles*. 2016;30:40-56.
2. Wieder WR, Bonan GB, Allison SD. Global soil carbon projections are improved by modelling microbial processes. *Nat Clim Change*. 2013;3:909-912.
3. Pearl J. *Causality: models, reasoning, and inference*. 2nd ed. Cambridge: Cambridge University Press; c2009.
4. Hernán MA, Robins JM. *Causal inference: what if*. Boca Raton: Chapman & Hall/CRC; c2020.
5. Imbens GW, Rubin DB. *Causal inference in statistics, social, and biomedical sciences*. Cambridge: Cambridge University Press; c2015.
6. Grace JB, Schoolmaster DR, Guntenspergen GR, *et al.* Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere*. 2012;3:73.
7. Campbell EE, Paustian K. Current developments in soil organic matter modeling and the expansion of model applications. *Environ Res Lett*. 2015;10:123004.
8. Reichstein M, Camps-Valls G, Stevens B, *et al.* Deep learning and process understanding for data-driven Earth

- system science. *Nature*. 2019;566:195-204.
9. Shi Z, Crowell S, Luo Y, *et al*. Model structures amplify uncertainty in predicted soil carbon responses to climate change. *Nat Commun*. 2018;9:2171.
 10. VanderWeele TJ. Explanation in causal inference: methods for mediation and interaction. Oxford: Oxford University Press; c2015.
 11. Peters J, Janzing D, Schölkopf B. Elements of causal inference: foundations and learning algorithms. Cambridge: MIT Press; c2017.
 12. Athey S, Imbens GW. Machine learning methods that economists should know about. *Annu Rev Econ*. 2019;11:685-725.
 13. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. *J Am Stat Assoc*. 2018;113:1228-1242.
 14. Chernozhukov V, Chetverikov D, Demirer M, *et al*. Double/debiased machine learning for treatment and structural parameters. *Econom J*. 2018;21:C1-C68.
 15. Luo Z, Wang E, Sun OJ. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. *Agric Ecosyst Environ*. 2010;139:224-231.
 16. Stockmann U, Adams MA, Crawford JW, *et al*. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agric Ecosyst Environ*. 2013;164:80-99.
 17. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48:1-36.
 18. Textor J, van der Zander B, Gilthorpe MS, *et al*. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol*. 2016;45:1887-1894.
 19. Tibshirani J, Athey S, Friedberg R, *et al*. grf: generalized random forests. R package version 2.0.2. 2021.
 20. Bach P, Chernozhukov V, Kurz MS, *et al*. DoubleML: an object-oriented framework for causal machine learning in Python. *J Mach Learn Res*. 2022;23:1-6.
 21. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer; c2009.
 22. Spirtes P, Glymour CN, Scheines R. Causation, prediction, and search. 2nd ed. Cambridge: MIT Press; c2000.
 23. Angrist JD, Pischke JS. Mostly harmless econometrics: an empiricist's companion. Princeton: Princeton University Press; c2009.
 24. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; c2023.
 25. Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.
 26. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol*. 2007;58:593-614.
 27. Green DP, Kern HL. Modeling heterogeneous treatment effects in survey experiments with Bayesian additive regression trees. *Public Opin Q*. 2012;76:491-511.
 28. Spirtes P, Zhang K. Causal discovery and inference: concepts and recent methodological advances. *Appl Inform*. 2016;3:3.
 29. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods*. 2008;40:879-891.
 30. Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: lessons from machine learning. *Perspect Psychol Sci*. 2017;12:1100-1122.
 31. Fuss S, Lamb WF, Callaghan MW, *et al*. Negative emissions—Part 2: costs, potentials and side effects. *Environ Res Lett*. 2018;13:063002.
 32. Pearl J. The causal foundations of structural equation modeling. In: Hoyle RH, editor. Handbook of structural equation modeling. New York: Guilford Press; 2012. p. 68-91.
 33. Grace JB, Anderson TM, Olff H, *et al*. On the specification of structural equation models for ecological systems. *Ecol Monogr*. 2010;80:67-87.
 34. Zhang X, Davidson EA, DeLucia EH, *et al*. Managing nitrogen for sustainable development. *Nature*. 2015;528:51-59.
 35. Scheffer M, Carpenter S, Foley JA, *et al*. Catastrophic shifts in ecosystems. *Nature*. 2001;413:591-596.
 36. Sulman BN, Phillips RP, Oishi AC, *et al*. Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO₂. *Nat Clim Change*. 2014;4:1099-1102.
 37. Antle JM, Diagana B, Stoorvogel JJ, *et al*. Spatial heterogeneity, contract design, and the efficiency of carbon sequestration policies for agriculture. *J Environ Econ Manage*. 2006;51:219-241.
 38. Holland PW. Statistics and causal inference. *J Am Stat Assoc*. 1986;81:945-960.
 39. Athey S, Tibshirani J, Wager S. Generalized random forests. *Ann Stat*. 2019;47:1148-1178.
 40. Kennedy EH, Ma Z, McHugh MD, *et al*. Nonparametric methods for doubly robust estimation of continuous treatment effects. *J R Stat Soc Series B Stat Methodol*. 2017;79:1229-1245.