



1 **Modelling and Interpreting Thermal Stability Indices to Understand Soil Carbon Stabilization**
2 **Using Soil Properties Data**

3
4 **Kingsley John^{*1}, Travis Pennell^{1,2}, Louis-Pierre Comeau², Derek H. Lynch¹, Moazame Mesgar³,**
5 **Adam W Gillespie³, Brandon Heung¹**

6
7 ¹Department of Plant, Food, and Environmental Science, Dalhousie University, Truro, NS,
8 Canada

9
10 ²Fredericton Research and Development Center, Agriculture and Agri-Food Canada, Fredericton, NB,
11 Canada

12
13 ³School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada

14
15 *Corresponding Author: **Kingsley John^{*1}**(john.kingsley@dal.ca)

16



17 **ABSTRACT**

18 Soil organic carbon (SOC) sequestration and nutrient cycling are related to the susceptibility of soil
19 organic matter to biological decomposition. Several studies have demonstrated associations between
20 biological stability and thermal stability, as assessed using programmed pyrolysis. We sought to develop
21 parsimonious machine learning (ML) models to predict SOC stability indices from measured soil
22 properties. The study analyzed indices such as S_1 , S_2 , and S_3 ; the oxygen index (OI); the hydrogen index (HI);
23 and T_{50} , which reflect SOC composition, thermal behaviour, and stability. A total of 203 soil samples
24 collected at 0–15cm depth increments from agricultural, forest, and wetland landscapes in New
25 Brunswick, Canada, were analyzed. Feature selection techniques optimized predictive models, and a
26 random forest (RF) was used to develop one. Correlation results revealed that HI was
27 negatively associated with pH ($r = -0.35$) and bulk density ($r = -0.33$), whereas OI showed a
28 positive correlation with pH ($r = 0.34$). Thermal indices were more strongly related to soil chemistry and
29 texture, with S_1 closely correlated with total carbon ($r = 0.88$) and nitrogen, and S_2 negatively associated
30 with sand ($r = -0.66$) but positively associated with clay ($r = 0.24$) and POXC ($r = 0.84$). T_{50} showed
31 positive correlations with both pH ($r = 0.48$) and bulk density ($r = 0.36$), indicating greater thermal
32 stability in higher-pH, compacted soils, though these patterns varied by land use. Random Forest (RF)
33 model predicted S_1 - S_3 indices with high accuracy (CCC = 0.83–0.86), while HI and OI were more
34 difficult to model (CCC = 0.44–0.48), suggesting missing biological or environmental predictors; NH_4^+
35 and POXC emerged as key predictors. Structural equation modeling (SEM), after addressing
36 multicollinearity, supported a hypothesis-driven model that explained ~54% of T_{50} variation. Clay
37 dissolved organic carbon, pH, and aluminum showed significant direct associations with T_{50} (β for pH =
38 0.44), whereas bulk density showed no meaningful relationship. Our study demonstrates that ML and
39 SEM can reveal patterns and associations between soil properties and thermal stability indices, offering
40 insight into understanding the SOC stability under a changing climate as well as presenting a framework
41 for rapid estimation of SOC stability proxies.

42

43 **Keywords:** Carbon Stabilization; Machine Learning; Geochemical Indicators; Thermal Stability;
44 Rock-Eval Pyrolysis; Structural Equation Model

45

46

47



48 1. INTRODUCTION

49 Understanding soil organic carbon (SOC) stability is essential for maintaining the global nutrient cycle
50 and advancing climate-smart agricultural practices (Barreto et al., 2021). SOC stability is referred to as
51 the capacity of organic matter to resist microbial decomposition (Gregorich et al., 2015) and directly
52 influences nutrient cycling, greenhouse gas emissions, and carbon sequestration. The decomposability
53 of SOC, measured as the mineralized carbon per unit of total SOC, is influenced by environmental
54 factors, soil type, microbial activity, and interactions between organic matter and soil minerals
55 (Gregorich et al., 2015).

56 Thermal analysis is a widely used method for assessing SOC stability, involving controlled
57 heating of soil samples to measure mass loss and energy exchange associated with decomposition of
58 organic bonds and organo-mineral associations (Gillespie et al., 2014; Peltre et al., 2013; Plante et al.,
59 2009). Programmed pyrolysis (PP), as implemented in Rock-Eval or HAWK analyzers (Wildcat
60 Technologies, Humble, Texas, USA), operates by heating whole soil under an inert atmosphere,
61 generating diagnostic indices including S_1 (volatile hydrocarbons), S_2 (hydrocarbons from thermal
62 cracking of non-volatile OM), S_3 (CO_2 from oxygen-containing functional groups), and the derived
63 hydrogen index ($HI = S_2/TOC$) and oxygen index ($OI = S_3/TOC$) are commonly used Rock-Eval-derived
64 indicators of organic matter composition (Gillespie et al., 2014; Mesgar et al., 2024b; Plante et al., 2009).
65 Additional thermal stability metrics include T_{50} , the temperature at which 50% of SOC is volatilized
66 (Cécillon et al., 2018; Delahaie et al., 2023). Other thermal analytical approaches, such as
67 thermogravimetry (TG), differential scanning calorimetry (DSC; Gill et al., 2010; McElhaney, 1982),
68 and evolved gas analysis (EGA; Fernández et al., 2011; Risoluti et al., 2022; Verchovsky et al., 2020;
69 Barros et al., 2011; Lorenz et al., 2024), provide complementary information on mass-loss patterns,
70 energy fluxes, and evolved gas composition. Pyrolysis–gas chromatography/mass spectrometry
71 (Py-GC/MS) further enables molecular-level characterization of organic compounds (Fernández et al.,
72 2012; Rombolà et al., 2016). Although these techniques operate on different principles and platforms,
73 integrating their outputs can provide a more complete assessment of SOC composition and persistence
74 (Doležalová-Weissmannová et al., 2023; Gao et al., 2015; Kučerík et al., 2013; Plante et al., 2009).

75 Previous studies demonstrated that SOC thermal stability was correlated with biological stability,
76 supporting the deployment of thermal analysis as a proxy for decomposability (Gregorich et al., 2015;
77 Peltre et al., 2013). Furthermore, recent studies have linked thermal stability indices to landscape
78 position, management practices, and soil properties; for example, Shi et al. (2024) showed that SOC



79 thermal stability was higher in depositional landscapes, in comparison to eroded ones, due to the
80 increased clay and silt contents, cation exchange capacity (CEC), and specific surface area. An et al.
81 (2023) found higher thermal stability in hedgerow croplands than adjacent forested areas, suggesting
82 management-induced changes in organo-mineral associations. In permafrost regions, Fang et al. (2021)
83 observed that drier conditions promoted stronger SOC stability via enhanced mineral-organic
84 associations. Collectively, these studies indicate that soil physicochemical properties, specifically
85 texture, mineralogy, and nutrients, are closely linked to the thermal behaviour of SOC.

86 Despite these advances, no studies have sought to systematically understand the relationships
87 between routinely measured soil properties and thermal stability indices. Traditional methods for
88 assessing SOC stability (e.g., PP, Py-GC/MS, NMR) remain labour-intensive and costly, limiting their
89 application in large-scale monitoring. Pedotransfer functions (PTFs), predictive models estimating hard-
90 to-measure properties from readily available data, offer a potential solution. Originally developed for
91 soil physical properties (Bouma, 1989), PTFs have been extended using machine learning (ML)
92 algorithms such as Random Forest, Cubist, and support vector regression to capture nonlinear
93 relationships between soil properties (Arbor et al., 2023, 2024; Laurence et al., 2023; Padarian et al.,
94 2020; Van Looy et al., 2017). Although PTFs are typically used for predictive purposes, their structure
95 also provides insight into the soil relationships themselves. For example, interpretable ML frameworks,
96 incorporating feature selection (e.g., recursive feature elimination) and *post-hoc* analyses (e.g., variable
97 importance, partial dependence plots) have been advocated to both predict and understand soil-
98 environment relationships (Kasraei et al., 2024; Laurence et al., 2024; Paul et al., 2023; Wadoux &
99 Molnar, 2022). These approaches are highly applicable to developing PTFs for SOC thermal stability
100 while simultaneously elucidating the underlying drivers.

101 Structural equation modeling (SEM) provides a complementary framework to machine learning
102 by enabling explicit hypothesis testing of theory-derived pathways linking soil properties to thermal
103 stability indices. Unlike black-box prediction algorithms, SEM allows the specification and evaluation
104 of hypothesized relationships among variables, the distinction between direct and indirect associations,
105 and the identification of potential mediation patterns (Grace et al., 2016; Shipley, 2016). This capability
106 is particularly important for understanding how soil properties may interrelate in an integrated system;
107 for example, whether clay content shows a direct association with T_{50} , or whether its relationship with
108 T_{50} is mediated through carbon quality and nutrient availability. Recent studies have demonstrated
109 SEM's utility in revealing complex biogeochemical patterns, including the pathways by which land use



110 change relates to organic matter stability (Li et al., 2022) and how mineralogical controls correlate
111 with carbon persistence across climatic gradients (Rasmussen et al., 2018). However, no study has
112 systematically integrated SEM with interpretable machine learning to both predict thermal stability
113 indices and explore their multivariate relationships with soil physicochemical properties.

114 Here, we employ a two-stage analytical framework: first, random forest modelling with recursive
115 feature elimination to identify key predictors and quantify their importance; second, using SEM to test
116 *a priori* hypotheses about causal pathways linking these predictors to T_{50} . This integrated approach goes
117 beyond single-method studies by leveraging machine learning for predictive accuracy and SEM for
118 mechanistic inference, thereby addressing both practical prediction needs and fundamental
119 understanding of SOC stabilization processes.

120 We hypothesize that soil physical and chemical properties, as key controls on stabilization
121 mechanisms such as organo-mineral association, aggregation, and organic matter chemistry, can serve
122 as proxies for predicting SOC thermal stability indices. Specifically, we expect that soil properties
123 associated with mineral reactivity (clay content, Fe/Al oxides), chemical characteristics (pH, CEC), and
124 labile carbon pools (permanganate oxidizable carbon, dissolved organic carbon) will be significant
125 predictors. Our objectives were as follows: (1) to develop parsimonious machine learning (ML) models
126 to predict a suite of thermal stability indices (S_1 , S_2 , S_3 , HI, OI, and T_{50}) from routinely measured soil
127 properties; (2) to identify the most influential predictors using feature selection and variable importance
128 analysis; and (3) to explore causal pathways linking soil properties to thermal stability using SEM. By
129 leveraging interpretable ML, this study aimed to advance mechanistic understanding of how soil
130 properties inform SOC thermal stability patterns and provide a framework for rapid, low-cost estimation
131 of HAWK pyrolysis thermal stability indices, which explain SOC stability.

132 2. MATERIALS & METHODS

133 2.1. Study Area

134 This research was conducted in the Province of New Brunswick, in the Atlantic Maritime Ecozone of
135 Eastern Canada (**Fig. 1**). The province of 71,388 km² receives an annual average precipitation of 1100
136 mm. Geomorphological features include mountainous regions in the North, lowlands and ocean coasts
137 in the East, and the Saint John River in the South, flowing towards the Bay of Fundy (Pronk and
138 Ruitenbergh, 1991). Geological deposits comprise mainly glacial till, glaciofluvial, glaciolacustrine, and
139 glaciomarine deposits (Furze and Arp, 2020; Rampton, 1984). In addition, the area is characterized by
140 outcrops and relatively shallow soils (< 25 cm) across an extended region of the province (Furze et al.,



156 hypercube sampling (cLHS; Minasny and McBratney, 2006) and opportunistic sampling. For the cLHS
157 approach, locations were restricted to selected areas that fell within a 200 m buffer of the road for ease
158 of sampling. Opportunistic sampling locations were selected at the discretion of field technicians to
159 supplement cLHS points, ensuring that each sampling point maintained a minimum distance of 10 km
160 between points with a shared land cover classification.

161 2.2.2. Soil Physical and Chemical Properties

162 All laboratory analysis methods were determined using standard soil testing protocols as reported in
163 Vallotton et al. (2025). These included sand, silt, clay (particle size fractions, %); BD (bulk density, g
164 cm⁻³) and BD_FINE (bulk density corrected for coarse fragments, g cm⁻³); and Agg_stab (aggregate
165 stability, %) at 0 -15cm. Soil carbon and nitrogen pools comprised: TC (total carbon, %), TN (total
166 nitrogen, %), and C:N (carbon-to-nitrogen ratio); POXC (permanganate oxidizable carbon, mg kg⁻¹), a
167 measure of labile carbon; DOC (dissolved organic carbon, ppm); TDN (total dissolved nitrogen, ppm);
168 TON (total organic nitrogen, ppm); TIN (total inorganic nitrogen, ppm); NH₄⁺ (ammonium-nitrogen,
169 ppm); and NO₃⁻ (nitrate-nitrogen, ppm). Soil chemical properties and nutrients included: pH (soil pH);
170 CEC (cation exchange capacity, meq 100 g⁻¹); Base_sat (base saturation, %); P₂O₅ (phosphate, ppm);
171 K₂O (potash, ppm); and extractable nutrients B (boron), Cu (copper), Zn (zinc), S (sulfur), Mg
172 (magnesium), Fe (iron), Ca (calcium), Mn (manganese), Al (aluminum), and Na (sodium), all in ppm.
173 Exchangeable cation saturations were expressed as percentages: K_sat (potassium saturation), Mg_sat
174 (magnesium saturation), Ca_sat (calcium saturation), H_sat (hydrogen saturation), and Na_sat (sodium
175 saturation).

176 2.2.3. Thermal Stability Indicators

177 SOC thermal stability was assessed using programmed pyrolysis (PP), as described by Mesgar
178 et al. (2024). Analyses were performed using a HAWK pyrolysis instrument (Wildcat Technologies,
179 Humble, Texas, USA) at the Geological Survey of Canada, Calgary. This approach is analogous to Rock-
180 Eval© pyrolysis, involving two controlled heating stages under anaerobic (pyrolysis) and oxidative
181 conditions.

182 In the initial pyrolysis phase, 70 mg of finely ground (< 250 µm) soil was held isothermally under
183 N₂ at 300 °C for 3 min. During this stage, the hydrocarbons (HC) released were detected using a Flame
184 Ionization Detector (FID) and recorded as the S₁ peak (mg HC g⁻¹ of sample). Following the SI step, the
185 temperature was increased at a rate of 20 °C min⁻¹ to 650 °C to thermally decompose the more thermally



186 stable fraction of OM. The HC released in this phase was measured as the S₂ peak (mg HC g⁻¹ of sample),
 187 also recorded via the FID. Together, S₁ and S₂ represent the organic C liberated during thermal
 188 decomposition (Table 1). Concurrently, evolved gases, including carbon dioxide (CO₂) and carbon
 189 monoxide (CO), released during pyrolysis were monitored using an infrared (IR) detector and are
 190 represented as S₃' (mg CO g⁻¹ of sample; Baudin et al., 2015; Table 1).

191 From these measurements, several derived indices were calculated. The Hydrogen Index (HI, mg
 192 HC/g TOC) and Oxygen Index (OI, mg O₂/g TOC) are calculated by normalizing the S₂ and S₃ values,
 193 respectively, to the TOC content of the sample. The HI reflects the hydrogen-rich, energy-dense portion
 194 of SOC and is often used as a proxy for OM quality (Gillespie et al., 2014; Gregorich et al., 2015). The
 195 OI characterizes the relative oxygen content in OM and its potential oxidizability. The T₅₀ is a thermal
 196 stability index denote the temperature at which 50% of the total pyrolyzable organic matter is released.
 197 T₅₀ captures the median decomposition temperature across the full temperature range (Barros et al.,
 198 2011; Gregorich et al., 2015). Higher T₅₀ values indicate greater SOC thermal resistance and are typically
 199 associated with mineral-associated organic matter or chemically recalcitrant compounds.

200 **Table 1.** HAWK Rock-Eval pyrolysis thermal indices

Thermal Indices	Abbreviation	Definitions	Unit
S1	S1	Hydrocarbon ¹ released between 0 and 300 °C	HC-mg ¹ / g-rock
S2	S2	Hydrocarbon ¹ produced between 200 and 650 °C	HC-mg ¹ / g-rock
OI	OI	Equal to (S ₃ /TOC ²) × 100	O ₂ -mg / g TOC ²
HI	HI	Equal to ((S ₁ + S ₂) / TOC ²) × 100	HC-mg ¹ / g TOC ²
S3'	S3	Pyrolysis of OM ³	mg-CO ₂ / g-rock
T50	T50	Temperature at which 50% of the loss in SOM ⁴ weight occurs	Celsius (°C)

201 ¹ HC = hydrocarbons; ² TOC = total organic carbon; ³ OM = organic matter; ⁴SOM = organic matter;

202

203



204 2.3. Predictive Modelling Framework

205 The predictive modelling framework follows the approach of Kasraei et al. (2024). It involves four key
206 steps: (1) addressing multicollinearity among predictor variables variance inflation factor (VIF) analysis
207 (**Section 2.3.1**); (2) fitting a Random Forest (RF) model to predict each thermal stability index (**Section**
208 **2.3.2**); (3) applying recursive feature elimination (RFE) to ensure model parsimony by removing non-
209 informative variables (**Section 2.3.3**); and (4) conducting *post-hoc* model assessments, including
210 structural equation modeling (SEM) and model interpretability techniques (Fig. 2; **Section 2.4**).

211 2.3.1. Addressing Multicollinearity

212 Multicollinearity among predictors was assessed and mitigated using the VIF, which quantifies the
213 proportion of variance in each predictor that can be explained by all other predictors (Craney &
214 Surlles, 2002). Here, VIF was calculated in equation (1):

$$215 \quad VIF_i = \frac{1}{1 - R_i^2} \quad (1)$$

216 where r_i^2 is the coefficient of determination from fitting a linear regression between the i -th predictor
217 and all other predictors. VIF can range from 1 to infinity, with a value of 1 indicating no
218 multicollinearity, values between 1 and 5 suggesting a moderate collinearity, and values > 5 indicating
219 a high degree of collinearity (Curto & and Pinto, 2011; Thompson et al., 2017). A stepwise approach
220 was used to address multicollinearity by iteratively removing covariates with the highest VIF until all
221 remaining covariates had a VIF of 5 or lower, reducing the full set of covariates to 21 predictors. The
222 VIF analysis was implemented in R using the *oss.seqVIF* function in the *onsoilsurvey* package (Saurette,
223 2021).

224 2.3.2. Random Forest

225 The RF model is an ensemble machine learning algorithm based on decision trees, which was used to
226 model the relationship between soil physicochemical properties and thermal stability indices. For each
227 tree-based model (i.e., base learner), the training data is split at nodes using the values of the predictor
228 variables to maximize within-node similarity and between-node differences (Breiman et al., 1948). RF
229 enhances this approach by constructing multiple trees, each trained on a randomized bootstrap sample
230 (with replacement) of the dataset, creating a diverse ‘forest’ of decision trees. At each node, a random
231 subset of predictors is selected for splitting. The final RF prediction is determined by aggregating the



232 outputs of all trees, using averaging for regression applications. Two key hyperparameters define the RF
233 model: n_{tree} , which specifies the number of decision trees, and m_{try} , which sets the maximum number of
234 covariates considered at each split (Breiman, 2001).

235 Model performance was assessed using Lin's concordance correlation coefficient (CCC),
236 coefficient of determination (R^2), root mean square error (RMSE), and bias. The best-performing model
237 for each index was selected based on the highest CCC and lowest RMSE values. The models were tuned
238 using a predefined grid of hyperparameters, spanning the range of m_{try} values, and using a 10-fold cross-
239 validation procedure with 20 repeats.

240 2.3.3. Recursive Feature Elimination

241 During the RF modelling process, the 21 uncorrelated variables identified through VIF were
242 further reduced using RFE to identify the minimal set of variables that maintains model performance.
243 RFE was implemented using the *caretFuncs* function within the *caret* R package, which evaluates the
244 contribution of each predictor against the response variable (e.g., thermal stability indicators) based on
245 variable importance scores. Less important features were progressively removed during the iteration.

246 Following this, the RFE output provided various features alongside their corresponding
247 performance metrics. Features resulting in the lowest RMSE were selected for modeling (Kuhn, 2019).
248 Additionally, the RMSE from repeated cross-validation runs was used to compute and plot the 90%
249 confidence intervals (CI) for each RFE model. The model with the reduced set of covariates achieving
250 the lowest RMSE was considered the baseline for comparison with other models. Here, the 'smallest'
251 model, where removing an additional predictor would not significantly increase RMSE, was retained
252 and subsequently used to build the models and SEM for thermal stability indices.

253 2.4. Post-Hoc Model Assessment

254 Several *post-hoc* assessment tools, such as Shapley value, partial dependence plots (PDP), individual
255 conditional expectation (ICE) plots, and accumulated local effects (ALE) plots, have been developed to
256 address the challenge of interpreting ML prediction (Friedman, 2001; Goldstein et al., 2015; Apley and
257 Zhu, 2020; Molnar et al., 2021). This need arises because many ML models are considered 'black box'
258 models. Collectively, the methods help integrate physical knowledge into ML to improve interpretability
259 and transparency.

260



261 *2.4.1. Variable Importance Analysis*

262 We used the *caret* package in R to generate permutation-based, model-agnostic measures of variable
263 importance for predicting thermal indices using RF. This analysis is important as it allows the model to
264 measure the contribution of each predictor (e.g., soil physical and chemical properties) to overall
265 accuracy.

266 *2.4.2. Partial Dependence Plots & Individual Conditional Expectation Plots*

267 The PDPs show the interaction between the features and the prediction (Friedman, 2001). The PDP
268 estimates the mean marginal effect of predictors on the predicted thermal indices, which can be
269 determined value in regression. To plot these effects, the dataset was first split into two groups: the
270 feature of interest and all other features. Following this, each value from this set was systematically
271 substituted into the feature of interest across all observations. This created a synthetic data point where
272 only the feature of interest varied, while all other features remained unchanged. Lastly, the trained RF
273 learner then predicted the target variable (e.g., thermal index) for each synthetic observation. These
274 predictions represented the marginal effect of the feature of interest.

275 Unlike PDPs, which show average effects, ICE plots display the relationship between a feature
276 and the predicted response for each observation (Goldstein et al., 2015). Like PDP, ICE keeps all other
277 features constant while systematically varying the feature of interest across a defined grid. For each
278 observation, synthetic data points are created by replacing the original feature value with values from
279 this grid. The model then predicts the response for each of these synthetic instances. Repeating this
280 process for all observations allows ICE plots to reveal how each feature influences individual
281 predictions, offering a more detailed and transparent view of model behavior.

282 *2.4.3. Structural Equation Modelling (SEM)*

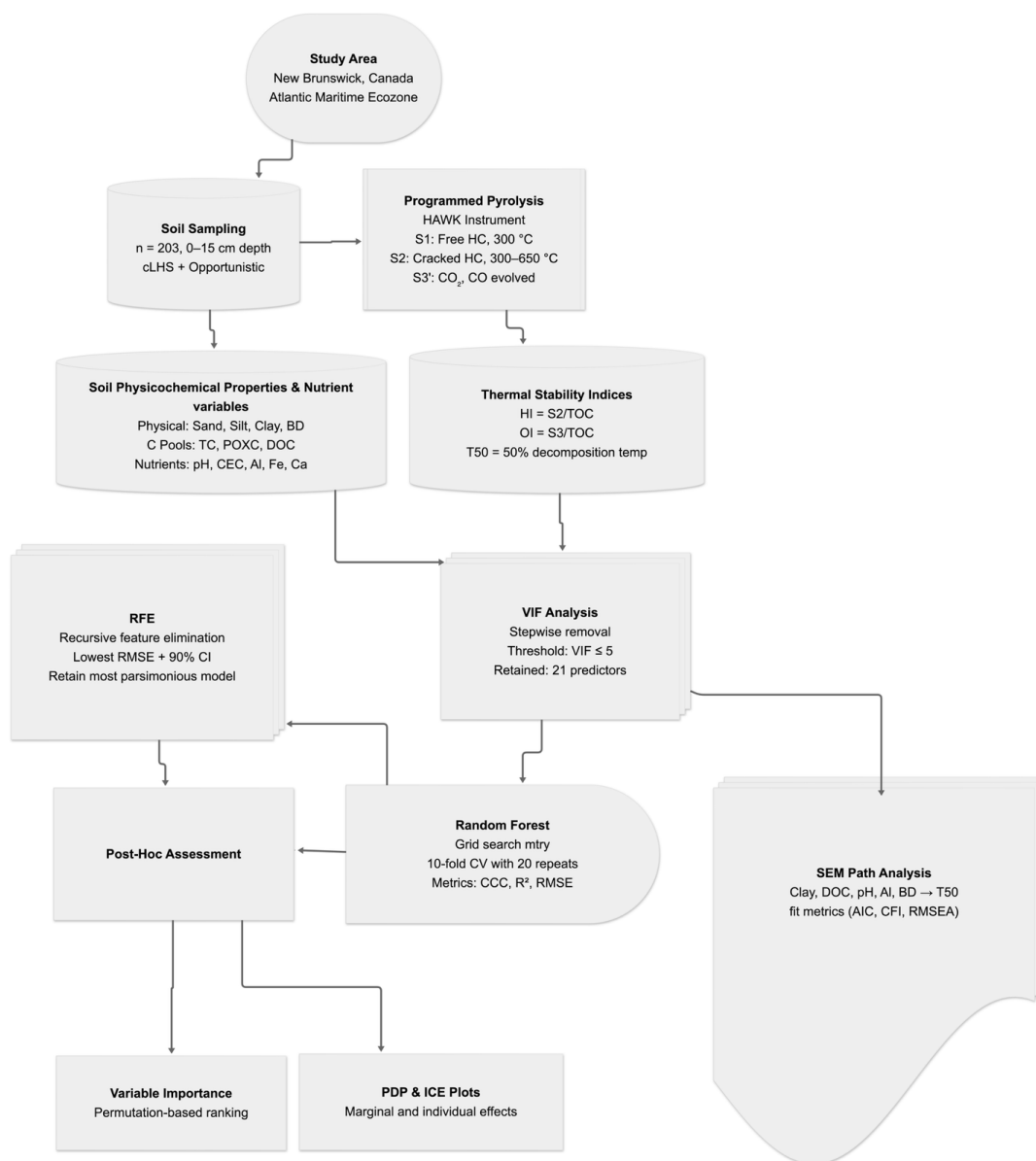
283 We employed structural equation modeling (SEM) to test *a priori* hypotheses about theoretically derived
284 relationships between soil properties and thermal stability of SOC (e.g., T_{50}), complementing predictive
285 machine learning with path analysis that can reveal direct and indirect associations. Before the SEM
286 specification, we addressed severe multicollinearity (Section 2.3.1). Following Wade et al. (2026) on the
287 importance of justifiable *a priori* models, we formulated five hypotheses based on biogeochemical
288 theory, predicting specific directional relationships: clay content would yield a positive association with
289 T_{50} , consistent with physical protection mechanisms (H₁; Six et al., 2004); dissolved organic carbon



290 would yield a negative interaction with T_{50} , reflecting labile carbon pools (H₂; Lehmann & Kleber,
291 2015); soil pH would demonstrate a positive relationship with T_{50} through cation bridging effects (H₃;
292 Rowley et al., 2018); aluminum content would show a positive interaction with T_{50} , consistent with
293 organo-mineral complex formation (H₄; Kleber et al., 2007); and bulk density would show a
294 positive interaction with T_{50} through physical structure effects (H₅; Dungait et al., 2012).

295 Using the *lavaan* package (Rosseel, 2012) with standardized variables, we compared three
296 models: the *a priori* hypothesis model (Model 1; T_{50} as a function of clay, dissolved carbon, pH,
297 aluminum, and bulk density); an exploratory model incorporating additional variables (Model 2;
298 permanganate oxidizable carbon, C:N, P_2O_5) that showed strong bivariate correlations ($|r| > 0.30$) with
299 T_{50} ; and a mediation model testing indirect pathways of pH and clay via dissolved carbon and aluminum
300 (Model 3). Models were fitted using maximum likelihood with robust standard errors and bootstrap
301 resampling ($n = 1000$) for mediation analysis (Preacher & Hayes, 2008). Model fit was assessed using
302 the Akaike information criterion (AIC), Bayesian information criterion (BIC), comparative fit index
303 (CFI ≥ 0.90), and root mean square error of approximation (RMSEA ≤ 0.08 ; Hu & Bentler, 1999; Shi et
304 al., 2019). It is important to note that lower AIC values indicate better fit, but significance may depend
305 on sample size; lower BIC values indicate better fit; RMSEA ≤ 0.08 or lower values (acceptable), ≤ 0.05
306 (good fit); CFI ≥ 0.90 (acceptable), ≥ 0.95 (good fit) (Shi et al., 2019).

307 Model 1 was selected as the primary model based on theoretical grounding, parsimony, and
308 comparable explanatory power to more complex alternatives. Standardized path coefficients (β) were
309 interpreted at $\alpha = 0.05$, with hypotheses considered supported if coefficients were statistically significant
310 and directionally consistent with predictions. We further tested model robustness using 10-fold cross-
311 validation repeated 20 times, with close agreement between calibration and validation (R^2). We present
312 path diagrams generated using *semPlot* (Epskamp et al., 2012) with color-coded edges indicating
313 direction and significance of associations. We explicitly distinguish exploratory Model 2 results from
314 confirmatory hypothesis tests, recognizing that all findings represent correlational evidence consistent
315 with, but not proof of, causal mechanisms (Wade et al., 2026).



316

317 **Figure 2.** Methodological framework for predicting thermal stability indices using soil physicochemical
318 properties and machine learning (ML) and structural equation model (SEM).

319 **3. RESULTS & DISCUSSION**

320 **3.1. Summary Statistics of Thermal Stability Indices**

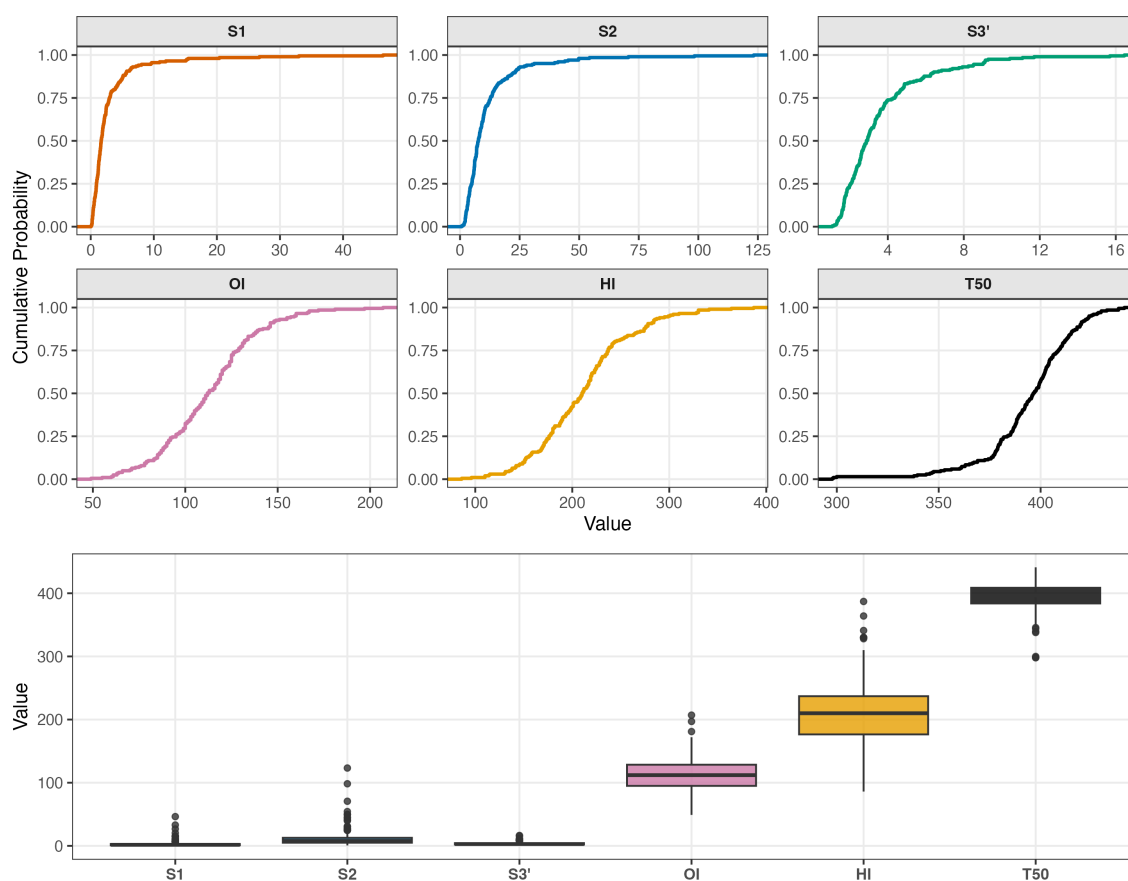


321 **Fig. 3** shows the boxplots and cumulative distribution functions (CDFs) for thermal stability indices
322 derived from HAWK-programmed pyrolysis across New Brunswick soils. These indices include volatile
323 hydrocarbons already present (S_1), pyrolysate (S_2), carbon dioxide (S_3), hydrogen index (HI), oxygen
324 index (OI), and T_{50} (temperature at which 50% of OM is pyrolyzed). Here, the result shows that HI and
325 S_2 have broad interquartile ranges, indicating heterogeneity in thermally stable, hydrogen-rich SOC
326 across sample locations (e.g., land uses; Barros et al., 2011; Gillespie et al., 2014). In contrast, the
327 narrower distributions of OI and S_1 suggest greater uniformity in volatile content and oxidation potential,
328 potentially due to more consistent environmental or soil conditions across sites (Gao et al., 2015; Plante
329 et al., 2009). The high median observed in T_{50} (around 350–420 °C) may suggest a significant presence
330 of thermally stable SOC to readily decomposable SOC, which supports past studies indicating that higher
331 pyrolysis temperatures correlate with stable, mineral-associated SOC in complex soils (Fernández et al.,
332 2011, 2012; Kučerík et al., 2013, 2018; Peltre et al., 2013).

333 In the Supplementary Materials, we present violin boxplots and spatial distribution plots of
334 thermal stability indices obtained from HAWK-programmed pyrolysis across various land use types
335 (Figs. S1-S6). The plots revealed that the forest land use consistently showed higher T_{50} values and
336 moderate-to-high S_2 , suggesting greater thermal stability and a higher proportion of recalcitrant SOC,
337 which may be due to the age of vegetation inputs, acidic conditions, and stronger organo-mineral
338 associations that promote stabilization of thermally resistant carbon fractions (Gregorich et al., 2015).
339 Agricultural land use exhibited lower and more varying T_{50} alongside high S_1 and HI, reflecting
340 disturbance from tillage, enhanced mineralization, and a greater contribution of labile, recently derived
341 carbon pools from crop residues and amendments (Gregorich et al., 2015; Gillespie et al., 2014). The
342 wetland ecosystem displayed a high S_2 with variable T_{50} , confirming the accumulation of fresh, plant-
343 derived organic matter under anoxic conditions that limit oxidative decomposition and support carbon
344 sequestration despite lower thermal maturity in certain fractions (Gregorich et al., 2015). Pastures
345 presented intermediate spatial distribution, with moderate T_{50} and HI influenced by perennial plant root
346 inputs, grazing effects, and partial mineral protection mechanisms that stabilize organic matter to a
347 degree between forest and cropland systems.



348



349

350 **Figure 3. Cumulative distribution function (CDF) plots (top) and boxplots (bottom) for thermal**
 351 **stability indices derived from HAWK-Programmed Pyrolysis.** The indices include the S₁, S₂, S₃,
 352 OI, HI, and T₅₀, which provide insights into the thermal degradation of SOC.

353 3.2. Relationships between Thermal Stability and Physicochemical Properties of Soil

354 3.2.1. S₁

355 A negative correlation with soil pH ($r = -0.33$) suggested that higher free hydrocarbon content is
 356 associated with more acidic soils (Fig. 4). Strong positive correlations with total nitrogen (TN; $r = 0.64$)
 357 and %C ($r = 0.88$) indicate that elevated S₁ values correspond to nutrient-rich environments. Additional
 358 positive correlations with clay ($r = 0.38$), C:N ($r = 0.53$), POXC ($r = 0.70$), DOC ($r = 0.85$), TDN ($r =$
 359 0.31), NH₄⁺ ($r = 0.58$), TON ($r = 0.21$), and CEC ($r = 0.40$) reinforce the association between S₁, labile



360 carbon pools, and overall soil fertility. Collectively, these relationships highlight how free hydrocarbons
361 in SOM reflect interactions among soil acidity, nutrient availability, and physical protection mechanisms
362 (Berthrong et al., 2009, 2012).

363 3.2.2. S₂

364 S₂ was significantly correlated with various soil physicochemical properties reflecting its link to
365 kerogen-like, thermally labile hydrocarbons in SOM (Fig. 4). A moderate negative correlation with soil
366 pH ($r = -0.29$) suggests that acidic soils, where decomposition proceeds more slowly, tend to accumulate
367 more S₂-related organic matter (Disnar et al., 2003). The strong negative correlation with sand ($r =$
368 -0.25) indicates that finer-textured soils protect organic matter and thus have higher S₂ values, consistent
369 with lower retention and stabilization capacities in sandy soils (Entio et al., 2024; Schmidt et al., 2011;
370 Six & Paustian, 2014). Strong positive correlations were observed between S₂ and TN ($r = 0.74$), TC ($r =$
371 0.92), POXC ($r = 0.77$), DOC ($r = 0.81$), NH₄⁺ ($r = 0.49$), clay ($r = 0.41$), Zn ($r = 0.42$) and TIN ($r =$
372 0.31), indicating that S₂ tracks nutrient-rich carbon fractions associated with soil fertility (Culman et al.,
373 2012).

374 In contrast, negative correlations with BD ($r = -0.57$), Al ($r = -0.34$), K_{sat} ($r = -0.22$), and
375 Mg_{sat} ($r = -0.21$) suggest that BD soils and metal-rich environments, often associated with higher
376 temperatures during the S₂ pyrolysis phase (650 °C) and may not favor the accumulation of kerogen-like
377 material. This aligns with evidence that high BD and elevated soil temperatures promote SOM
378 mineralization (Davidson & Janssens, 2006; Disnar et al., 2003).

379 3.2.3 S₃

380 The indicator S₃ exhibited a positive correlation with fine-textured and chemically reactive soil
381 components, such as clay ($r = 0.30$) and silt ($r = 0.32$). This suggests that particle size fractions play a
382 crucial role in protecting the release of CO₂ from organic matter. Conversely, S₃ had a negative
383 relationship with sand fractions ($r = -0.38$) and BD ($r = -0.52$), indicating a higher CO₂ release in coarse-
384 textured soils. Among chemical properties, S₃ correlated positively with CEC ($r \approx 0.56$), TC ($r \approx 0.86$),
385 TN ($r = 0.90$), POXC ($r = 0.84$), and DOC ($r = 0.52$). This highlights that S₃ increases with organic
386 matter content, and soils dominated by readily oxidizable carbon tend to accumulate less thermally stable
387 material. Overall, these patterns suggest that S₃ was likely influenced by the type of organic matter
388 formation in the soil, driven by fine texture, CEC, and labile carbon dynamics (Mesgar et al., 2024b).



389 3.2.4. *Oxygen Index (OI)*

390 OI showed a positive correlation with soil pH ($r = 0.34$), indicating that higher OI values correspond to
391 higher pH. This suggests that neutral to alkaline conditions may favor oxidative transformations of SOC
392 (Wenyika et al., 2025). The negative correlation between OI and C:N ($r = -0.35$) implies that increases
393 were possibly associated with decreases in C:N, which could reflect greater nitrogen availability and
394 enhanced microbial decomposition that promotes nutrient cycling and the accumulation of more
395 oxidized residues (Cotrufo et al., 2013, 2022). However, these patterns contrasted with our study, in
396 which lower C:N ratios (indicating higher nitrogen content and greater organic matter quality) were more
397 often associated with higher HI and lower OI values. Higher nitrogen availability is typically linked to
398 enrichment in labile, hydrogen-rich components such as O-alkyl-C and peptides, which contribute to
399 elevated HI rather than OI (Mesgar et al., 2024b). The observed positive relationship between OI and
400 lower C:N in this study may therefore be site-specific, potentially driven by land-use differences; for
401 example, croplands often exhibit lower C:N and higher OI due to accelerated oxidative processing
402 (Gillespie et al., 2014) or site-specific interactions in New Brunswick's soils. A positive correlation
403 between OI and Al ($r = 0.30$) further suggests that aluminum-rich soils may exhibit greater oxidative
404 signatures, likely due to organo-mineral interactions that influence SOM stabilization and transformation
405 pathways (Mesgar et al., 2024a; Mayer et al., 2023; Sparks, 1995; Wagai & Mayer, 2007).

406 3.2.5. *Hydrogen Index (HI)*

407 In Fig. 4, HI showed a negative correlation with BD ($r = -0.38$), pH ($r = -0.34$), and Al ($r = -0.33$).
408 This suggests that decreases in BD, pH, and Al levels were possibly associated with increases in HI.
409 High HI is often characterized by less stable SOC or rapidly decomposable and unstable SOC, which
410 may control BD variation (Yu et al., 2020). This result suggests that soils with higher compaction tend
411 to contain SOC with lower hydrogen. This pattern may reflect reduced inputs of fresh, O-alkyl-C-rich
412 organic matter and/or advanced decomposition under conditions of limited organic inputs. Increased
413 bulk density is also associated with altered aggregation and mineral protection mechanisms, which can
414 influence SOC composition and apparent stability rather than directly enhancing decomposability (Yu
415 et al., 2020).

416 In contrast, HI showed a positive correlation with TC ($r = 0.40$), TN ($r = 0.30$), POXC ($r = 0.30$),
417 and DOC ($r = 0.36$), reflecting the role of labile carbon fractions in supporting microbial activity, which
418 may in turn enhance HI (Kögel-Knabner, 2002). It is important to note that HI reflects the chemical



419 composition of SOC; higher HI content generally indicates a greater proportion of hydrogen-rich, labile
420 components, such as aliphatic and O-alkyl-C functional groups. In contrast, soils with high SOC can still
421 exhibit low HI if their carbon is dominated by condensed aromatic structures, which are hydrogen-poor
422 and therefore do not contribute to elevated HI values (Moore et al., 2017).

423 A negative correlation between HI and AI ($r = -0.33$) suggests that AI complexes are associated
424 with organic matter of lower hydrogen richness, reflecting differences in molecular composition rather
425 than lower organic matter content. In contrast, the positive correlation with cation exchange capacity
426 (CEC; $r = 0.27$) suggests that organic matter composition, particularly the presence of labile functional
427 groups, may enhance the retention of exchangeable cations.

428 3.2.6. T_{50}

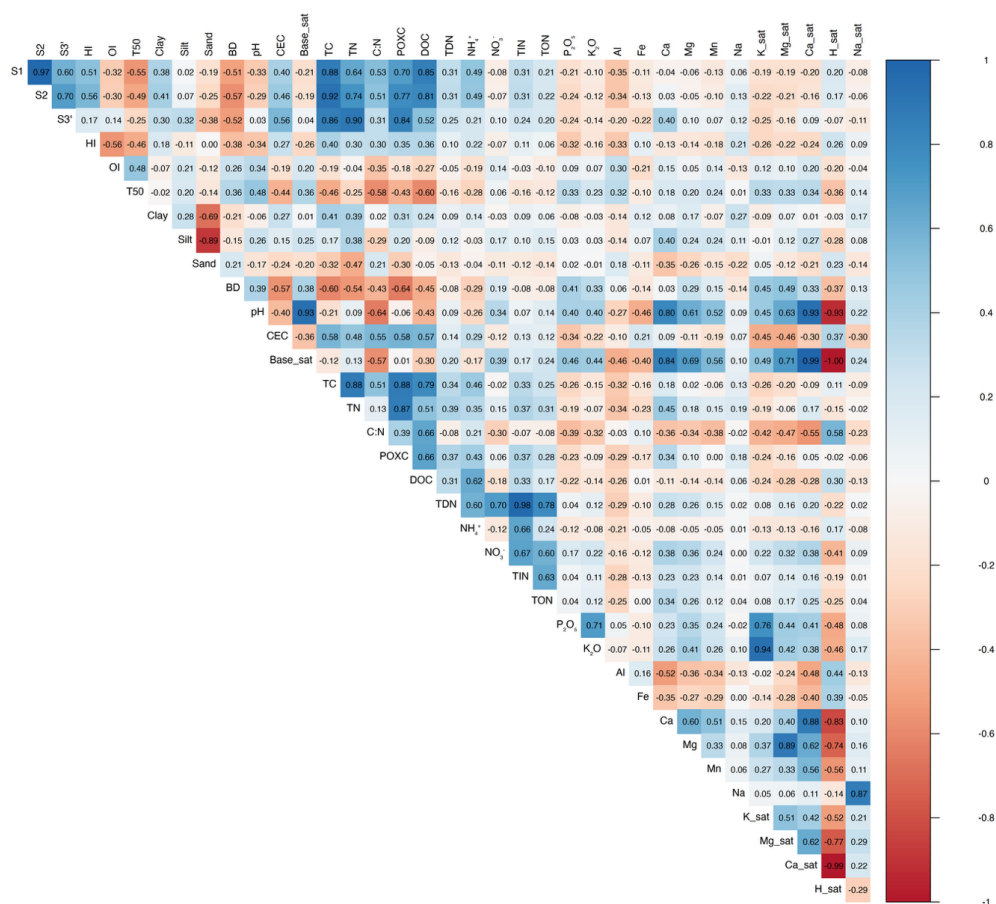
429 We observed a positive correlation between T_{50} and soil pH ($r = 0.48$), which suggests that more
430 thermally stable SOC was prevalent in alkaline soils, likely due to reduced microbial decomposition
431 rates under higher pH conditions, which promote SOC preservation (Almagro et al., 2021; Disnar et al.,
432 2003; Solomon et al., 2012; Tao et al., 2023). However, we also noted that acidic soils could show strong
433 SOC stability due to the presence of Fe and Al, which form strong organo-mineral bonds with organic
434 matter, resulting in higher T_{50} values. Similarly, the positive correlation with BD ($r = 0.36$) indicates
435 that compacted soils tend to contain more thermally stable SOC, potentially because reduced pore space
436 limits microbial access to organic substrates, favoring the accumulation of recalcitrant carbon fractions
437 (Six et al., 2002).

438 Conversely, T_{50} showed negative correlations with TC ($r = -0.46$) and C:N ($r = -0.58$),
439 suggesting that soils richer in labile SOC tend to have lower T_{50} values. Strong negative correlations
440 with DOC ($r = -0.60$) and POXC ($r = -0.43$) further support this trend, implying that high T_{50} values
441 are associated with soils containing fewer readily decomposable SOC pools, likely due to the preferential
442 breakdown of easily accessible organic compounds (Barreto et al., 2021; Briedis et al., 2023; Liu et al.,
443 2020). Additionally, the negative correlations with NH_4 ($r = -0.28$) and CEC ($r = -0.44$) suggest that
444 soils with higher T_{50} values yield a large proportion of recalcitrant SOC that has undergone advanced
445 decomposition and releases its labile constituents (Cremer & Prietzel, 2017; Rakhsh et al., 2017).

446 In contrast, T_{50} showed a positive correlation with P_2O_5 ($r = 0.33$), K_{sat} ($r = 0.33$), Mg_{sat} ($r =$
447 0.33), Ca_{sat} ($r = 0.34$), and base_{sat} ($r = 0.36$), respectively. This is likely due to stabilization
448 mechanisms involving metal oxides and clay minerals, which protect SOC from microbial degradation



449 (Wagai & Mayer, 2007). Additionally, the negative correlation with % H ($r = -0.36$) indicates that more
 450 thermally stable organic matter has a lower hydrogen content, consistent with advanced decomposition
 451 and oxidation processes that remove hydrogen-rich aliphatic compounds (Cecillon et al., 2018;
 452 Gregorich et al., 2015; Li et al., 2020; Sebag et al., 2016). These findings highlight the complex
 453 interactions between thermal stability, soil nutrient availability, and metal oxides, aligning with the
 454 observations of Lehmann & Kleber (2015).



455
 456 **Figure 4.** Correlation matrix between thermal stability indices and soil properties.

457 **3.3. Predictions of Thermal Stability Indices using Random Forest Learner**

458 Using the minimum number of covariates identified by RFE, we developed predictive models for each
 459 thermal stability index (S_1 , S_2 , S_3' , OI, HI, and T_{50} ; see Fig. 5). Accuracy metrics are presented in Table
 460 2, and model residual scatterplots in Fig. 5. The models for S_1 , S_2 , S_3' , OI, HI, and T_{50} showed the highest



461 accuracy, as reflected by their high CCC values (>0.80), low RMSE, and bias, suggesting model
462 consistency in capturing variations in SOC stability using thermal stability indicators and soil properties.
463 In contrast, HI and OI had the lowest model performance ($CCC \leq 0.48$). Additionally, T_{50} showed
464 moderate accuracy, with CCC values ranging from 0.58 to 0.66. While these models provided important
465 insights into SOC stability, they also suggested that incorporating additional variables, such as
466 biological, climatic, and land use factors, could enhance predictive capabilities. These results corroborate
467 previous studies on SOC dynamics and stabilization, where active or readily decomposable SOC
468 fractions, mineral interactions, and soil structure are known to significantly influence carbon cycling,
469 persistence, and storage in soils (Lal, 2004; Wiesmeier et al., 2019).

470 3.3.1. S_1 Model

471 The S_1 model yielded higher accuracy metrics with a CCC of 0.85 and a low RMSE of 1.19 HC-mg/g-
472 rock compared to both HI and OI (Table 2 and Fig. 5). The scatterplot for S_1 showed that the regression
473 line was close to 1:1, indicating a better model was obtained using the selected predictors. In addition,
474 Fig. 6 shows the variable importance plot of the S_1 model with POXC as the most important predictor
475 with the highest importance score ($> 90\%$), indicating that labile organic carbon strongly controls S_1 .
476 This was closely followed by DOC, which yielded an importance score of 85%, indicating that SOC
477 stabilizes via mineral interactions and contributes to S_1 variability. Also, it may be possible that at the
478 S_1 temperature level, partially transformed hydrocarbon compounds become volatilized and hence
479 contribute to the variability of S_1 as an index in understanding the stability of C in soils. BD had a
480 variable importance of 32%, likely due to its effect on microbial access to oxygen, which impacted SOC
481 decomposition (Tao et al., 2023). CEC contributed 28% of the relative importance, suggesting that clay-
482 mineral interactions play a moderate role in thermal indices prediction. The CEC's moderate
483 contribution can vary depending on the type of clay present in the soil (Ma et al., 2024). Conversely, the
484 C:N had an importance of 24%, indicating that, compared with other predictors in the model, it played
485 a moderate role in distinguishing patterns potentially linked to nitrogen availability and the accumulation
486 of more labile carbon. The result highlights that readily decomposable C has a relationship with the S_1
487 index, while BD and other soil properties play a secondary role in stabilizing SOC.

488 Fig. 7 presents PDP and ICE plots for S_1 model predictions. In the S_1 model, DOC and POXC
489 showed a strong positive correlation, suggesting their significant role in hydrocarbon (HC) release below
490 300°C , aligning with Leifeld et al. (2002) da Silva et al. (2025). Conversely, BD exhibited a declining



491 trend, indicating that compacted soil restricts microbial access to organic material, thereby stabilizing
492 SOC and limiting labile carbon availability (Six et al., 2004; Six & Paustian, 2014). CEC and C:N
493 displayed stable trends with slight increases, implying contributions from soil mineral interactions and
494 nitrogen content.

495 3.3.2. S_2 Model

496 The S_2 model demonstrated high predictive accuracy with a CCC of 0.86 and RMSE of 4.01 HC-mg/g-
497 rock (Table 2 and Fig. 5). The scatterplot showed close alignment between predicted and observed
498 values, with the regression line approximating the 1:1 line. Variable importance analysis (Fig. 6)
499 revealed that POXC was the dominant predictor, followed by DOC, BD, CEC, and C:N. This indicates
500 that S_2 , which is hydrocarbons from thermal cracking of non-volatile organic matter, is strongly
501 associated with labile carbon pools and soil physical properties.

502 Partial dependence plots (Fig. 7) showed that POXC had a strong positive relationship with
503 predicted S_2 values, confirming that readily available carbon pools drive the formation of kerogen-like
504 hydrocarbons. Conversely, BD exhibited a consistent negative trend, suggesting that soil compaction
505 limits the accumulation or detection of these thermally labile fractions, likely through physical protection
506 mechanisms that reduce microbial accessibility (Paul et al., 2006; De Neve & Hofman, 2000). CEC
507 showed a moderate positive effect, consistent with organo-mineral interactions that stabilize organic
508 matter decomposing at higher temperatures. The presence of trace elements, such as Mn and Zn among
509 the predictors, suggests that redox-active elements may influence S_2 dynamics in certain environments,
510 aligning with previous findings on trace metal involvement in SOC stabilization (Hall et al., 2016; Hall
511 & Silver, 2015).

512 3.3.3. S_3' Model

513 The S_3 model also exhibited strong predictive performance, with a CCC of 0.85 and RMSE of 1.09 mg-
514 $\text{CO}_2/\text{g-rock}$ (Table 2 and Fig. 5). Variable importance analysis (Fig. 6) identified POXC as the most
515 influential predictor (100%), followed by aggregate stability (61%) and CEC (41%). This indicates that
516 S_3' , representing CO_2 evolved from oxygen-containing functional groups, is primarily associated with
517 labile organic matter, soil structural properties, and cation exchange capacity. These findings align with
518 previous studies highlighting the rapid turnover and microbial accessibility of labile SOC as primary
519 factors in decomposition and stabilization processes (Gillespie et al., 2014; Gregorich et al., 2015;
520 Haynes, 2005). The importance of aggregate stability is supported by Bronick and Lal (2005), who



521 demonstrated that soil macro- and micro-aggregates protect SOC from microbial degradation, thereby
522 reducing CO₂ emissions. Similarly, Six et al. (2002) showed that clay minerals and CEC significantly
523 influence SOC stabilization through mineral-organic interactions.

524 Partial dependence plots (Fig. 7) revealed consistent patterns across the S₃ model: POXC showed
525 a strong positive relationship with predicted values, while BD exhibited a negative trend. Feature
526 interactions in the S₃ model aligned with established frameworks differentiating labile and mineral-
527 associated organic matter pools in soil carbon dynamics (Cotrufo et al., 2013, 2019, 2022; Giannetta et
528 al., 2018; King et al., 2024; Mitchell et al., 2020).

529 3.3.4. OI Model

530 The OI model demonstrated moderate performance, with a CCC = 0.44 and an RMSE = 39.37 mg CO₂
531 /g TOC, explaining approximately 43% of the variance in OI. While the selected predictors contributed
532 to OI estimation, a significant portion of variability remained unexplained. Among the predictors, total
533 organic nitrogen was the most influential, contributing 100%, followed by NO₃ (95%) and Al (76%).
534 This result aligns with previous studies (e.g., Smith et al., 2015; Murphy, 2014), highlighting the
535 challenges of modeling complex soil processes due to soil heterogeneity. The variable importance plot
536 (Fig. 6) further confirmed that organic matter and nutrient indicators, such as TIN, C:N, and DOC, play
537 a crucial role in regulating biochemical transformations and CO₂ release during organic matter
538 decomposition.

539 Additionally, soil texture and structural properties, including silt, BD, and aggregate stability,
540 may control soil pores and diffusion pathways within the soil matrix, consistent with Dexter (2004).
541 Chemical parameters, such as NO₃, Fe, P₂O₅, Al, Zn, Na, and S, provided insights into soil nutrient status
542 and indirectly affected microbial activity and oxygen dynamics. These findings support earlier reports
543 by Havlin et al. (2013) emphasizing the intricate relationship between soil chemistry and biological
544 activity.

545 In the PDP plots (Fig. 7) for the OI model, as total organic nitrogen decreased, NO₃ increased
546 until both reached a stable state, underscoring the role of organic matter decomposition and oxidation.
547 The influence of P₂O₅ further suggests a connection to micronutrient availability, consistent with Kirkby
548 et al. (2013). These results indicate that OI is primarily governed by nitrogen- and phosphate-related
549 properties, mineral interactions, and soil texture, all of which influence soil organic matter oxidation and
550 stability. Future improvements to the OI model could incorporate microbial enzymatic and



551 environmental variables (e.g., drainage) given the significant roles of iron, nitrate, and phosphate
552 predictors.

553

554 3.3.5. HI Model

555 The HI model was developed using the minimum number of predictors: C:N, BD, NH₄, Al, DOC,
556 aggregate stability, P₂O₅, and Na, with a CCC of 0.48 and an RMSE of 39.4 mg HC/g TOC. The CCC
557 of 0.48 for HI models reflects moderate agreement between predicted and observed values. The variable
558 importance plot (Fig. 6) showed that NH₄ was the most important variable, with a 100% relative
559 contribution to predicting HI. This was followed by C:N with 79%, closely marked by P₂O₅ with ~76%,
560 Al with ~57%, BD with ~34%, DOC with ~18%, and Na with ~5%. High HI values indicate less
561 decomposed, hydrogen-rich organic matter, where mineralization can generate NH₄ and limited
562 nitrification allows it to accumulate; thus, NH₄ is more likely a consequence of high-HI than a cause.

563 The C:N, the next important variable in the predictive model for HI, conforms with studies on
564 SOC biogeochemical processes. For example, a lower C: N ratio indicates higher N availability, which
565 can enhance microbial activity and decomposition of organic matter, potentially driving HI (Wu et al.,
566 2023). HI informs the presence of labile functional groups with corresponding narrow C:N; hence, HI
567 could be influenced by narrow C:N, suggesting that high N content in an SOC may lead to more
568 hydrogen-rich compounds, thus driving HI.

569 In Fig. 7 for the PDP plots HI model, C:N exhibited a relatively stable trend with minor
570 fluctuations, indicating a minimal effect on the HI index. In contrast, BD showed a stronger relationship
571 with pyrolysis, likely due to its influence on microbial decomposition and organic matter stabilization.
572 NH₄ initially displayed a sharp increase before stabilizing, suggesting its strong influence on thermal
573 stability at the onset, potentially due to its role in raising pyrolysis temperature before stabilizing as
574 hydrogen molecules were fully released. These findings align with Wang et al., (2023), who reported
575 that ammonium perchlorate decomposition emits gases like HCN and CO₂ at high temperatures.

576 3.3.6. T₅₀ Model

577 In the T₅₀ model, the CCC values ranged from 0.58 to 0.66 (Table 2 and Fig. 5), where the most
578 significant variable was NH₄, which was followed by aggregate stability at 89% and DOC at 85 %. Al
579 had a moderate importance level of 34 %, while permanganate oxidizable C showed no relationship with



580 T₅₀ estimation. These results suggest that higher levels of dissolved carbon likely lower the temperature
 581 at which 50% of SOC degrades because labile dissolved carbon decomposes at lower temperatures,
 582 aligning with previous findings that indicate dissolved carbon can be used to predict SOC stability (Nie
 583 et al., 2019; Zhang et al., 2019).

584 Fig. 7 also shows the PDP for the T₅₀ model, highlighting the contribution of each predictor in
 585 the prediction. As shown, predictors connected with labile carbon pools yielded negative relationships
 586 with T₅₀. The predictors associated with more labile carbon pools generally show negative partial
 587 dependence, explaining that increases in variables such as DOC, C:N, and NH₄ are linked to lower
 588 predicted T₅₀ values. In addition, the model predicts that soils with high amounts of labile organic matter
 589 tend to exhibit lower temperatures of decomposition, reflecting reduced thermal stability. This finding
 590 is also consistent with Hou et al., (2021), who noted that active carbon pools like dissolved C reduce
 591 SOM stability in alpine soils due to their susceptibility to microbial degradation. The increasing trend
 592 indicates that the presence of Al, likely in the form of reactive oxides, contributes to SOC stability
 593 (Gregorich et al., 2015; Schmidt et al., 2011). Further emphasized that SOC stability depends more on
 594 mineral protection than on functional groups (Plante et al. (2011).

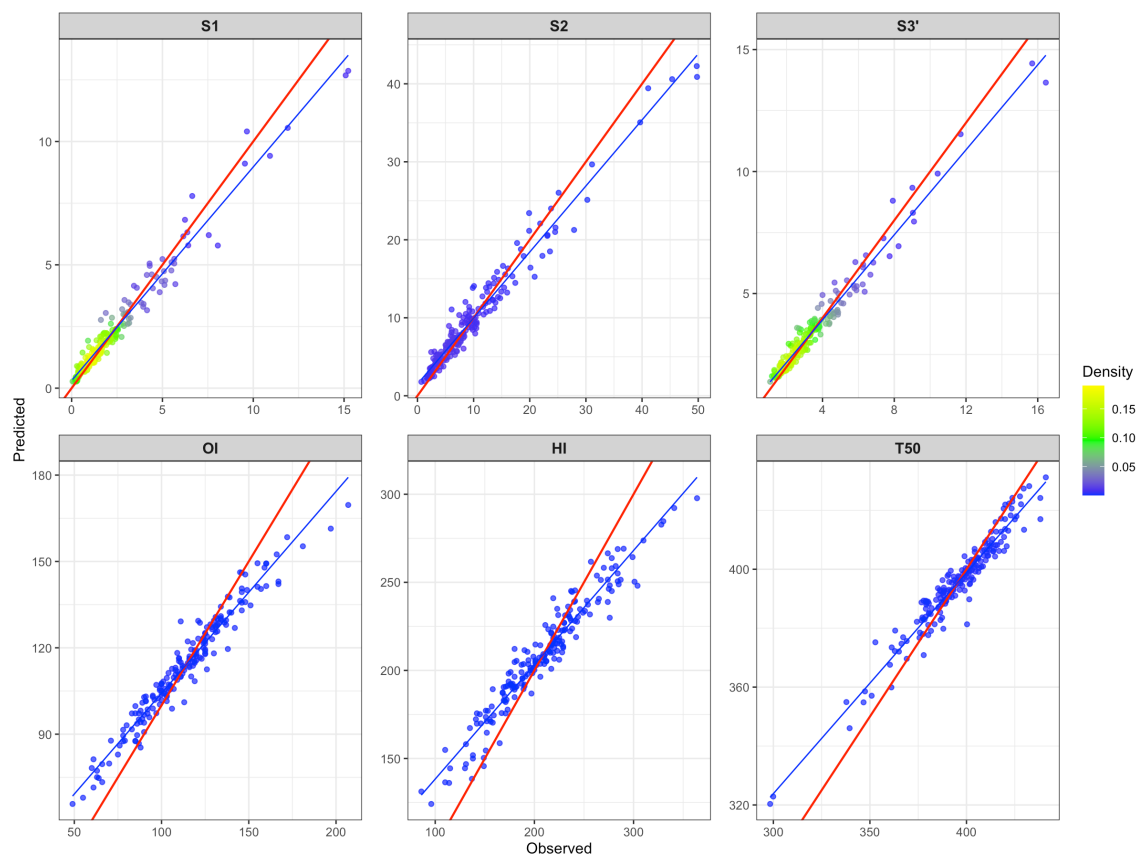
595

596 **Table 2.** Accuracy metrics of the predictive models of thermal stability indices

Thermal Stability Indices	r ²	Adj.r ²	CCC	RMSE	Bias	MAE
S ₁	0.76	0.76	0.85	1.19	-0.01	0.79
S ₂	0.77	0.77	0.86	4.01	0.01	2.78
S ₃	0.77	0.77	0.85	1.09	-0.01	0.78
OI	0.32	0.32	0.44	21.86	0.42	16.52
HI	0.33	0.33	0.48	39.37	0.81	30.57
T ₅₀	0.42	0.42	0.58	15.96	-0.24	11.85

597 **Note:** Coefficient of determination (R²), adjusted coefficient of determination (Adj. R²), Lin's
 598 concordance correlation coefficient (CCC), root mean square error (RMSE), bias, mean absolute error
 599 (MAE).

600

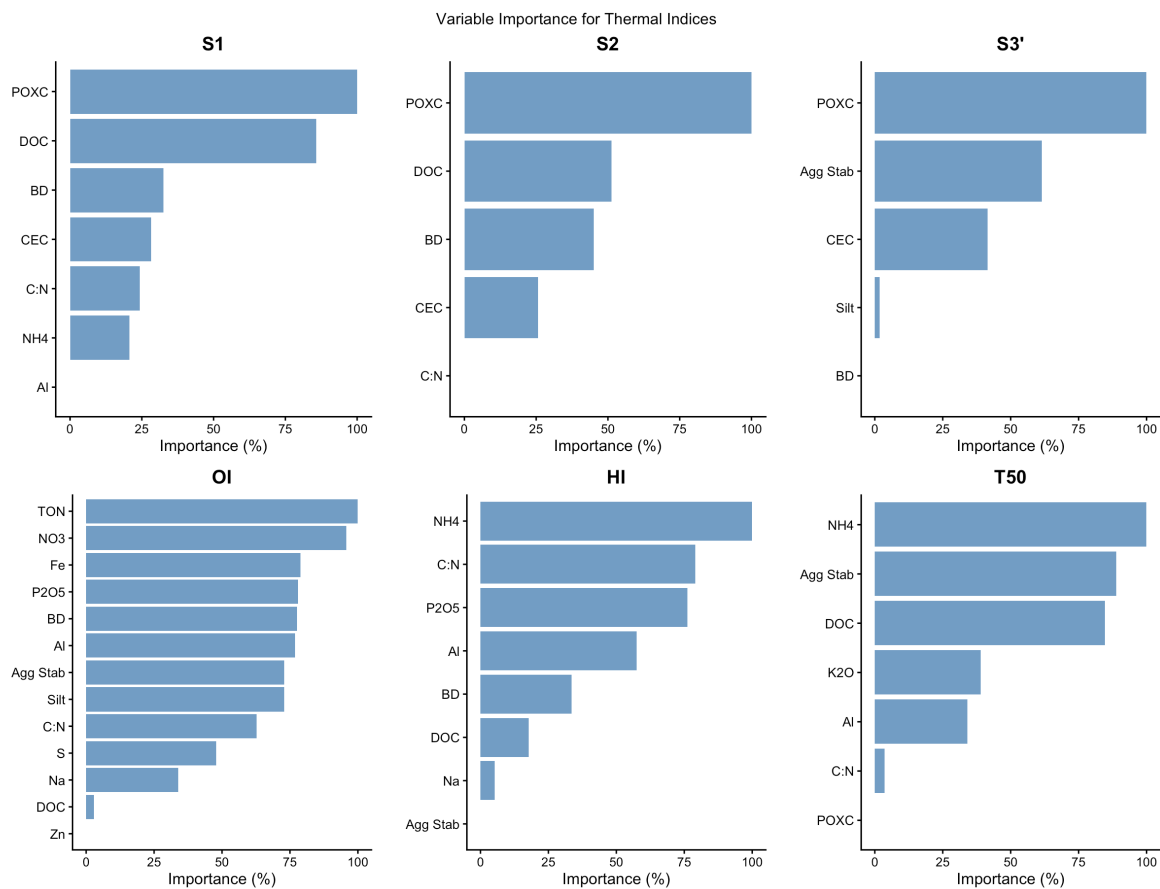


601
602
603
604
605

Figure 5. Scatterplot of the predicted versus observed data of the thermal stability indices: S₁ (HC-mg/g-rock), S₂ (HC-mg/g-rock), S_{3'} (mg-CO₂/g-rock), OI (O₂-mg/g TOC), HI (HC-mg/g TOC), & T₅₀ (°C). The red line represents the 1:1 line, and the blue line represents the line of best fit between observed and predicted values.



606

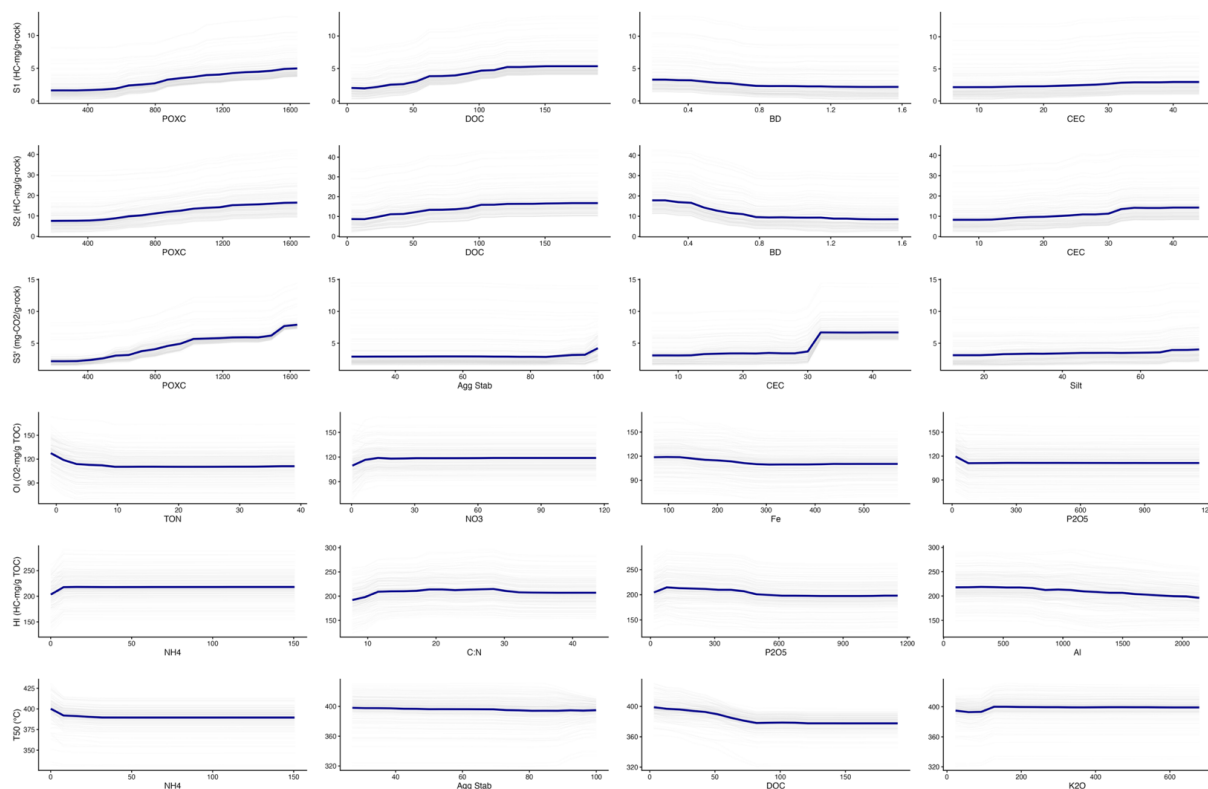


607

608

609

Figure 6. Permutation-based, relative variable importance plots for predicting thermal stability indices using physicochemical properties of soils.



610
 611 **Figure 7. Partial dependence plots generated for the top four most important predictors of**
 612 **thermal stability indices.** The grey lines represent the ICE curves, and the blue line represents the
 613 average of the Individual Conditional Expectation ICE lines.
 614

615 **3.5. Structural Equation Modelling**

616 We present our SEM as a hypothesis-testing framework that reveals patterns of association rather
 617 than demonstrating causation. Our *a priori* model was grounded in soil biogeochemical theory, with
 618 individual paths representing hypothesized relationships informed by existing literature (H₁: physical
 619 protection, Six et al., 2004; H₂: carbon quality, Lehmann & Kleber, 2015; H₃: pH effects, Rowley et al.,
 620 2018; H₄: mineral reactivity, Kleber et al., 2007; H₅: physical structure, Dungait et al., 2012; Tables 3-5
 621 & Fig. 8). Our results demonstrated excellent global fit (CFI = 1.0, RMSEA = 0.0, SRMR < 0.001),
 622 satisfying recommended thresholds (CFI ≥ 0.90, RMSEA ≤ 0.08, SRMR ≤ 0.10; Hu & Bentler, 1999;
 623 Wade et al., 2026). However, we acknowledge that the saturated nature of our model (df = 0) reflects its
 624 simplicity rather than definitive evidence of causal structure.



625 Our *a priori* hypothesis model explained 53.5% of variance in T_{50} , with cross-validation
626 confirming robustness (10-fold CV with 20 repeats $r^2 = 0.538 \pm 0.03$; CCC = 0.73, 95% CI: 0.68–0.78).
627 Hypothesis testing supported four of five predicted directional relationships: H₁ (clay content showed a
628 significant positive interaction with T_{50} , $\beta = 0.152$, $p = 0.002$), H₂ (DOC; dissolved organic carbon
629 showed a significant negative relationship with T_{50} , $\beta = -0.319$, $p < 0.001$), H₃ (soil pH showed a
630 significant positive relationship with T_{50} , $\beta = 0.436$, $p < 0.001$), and H₄ (Al; aluminum showed a
631 significant positive interaction with T_{50} , $\beta = 0.379$, $p < 0.001$). However, H₅ (bulk density) was not
632 supported ($\beta = 0.057$, $p = 0.305$), indicating that physical structure, as measured by density, may be
633 secondary to chemical factors in its relationship with thermal stability.

634 We further tested competing configurations using mediation analysis (H6–H7; Fig. 9) and an
635 exploratory model that included POXC, C: N, and P₂O₅. Clay content showed significant
636 indirect interaction via DOC (indirect $\beta = -0.074$, $p = 0.033$), and pH demonstrated
637 indirect associations through both DOC (indirect $\beta = 0.142$, $p = 0.002$) and aluminum (indirect $\beta = -$
638 0.113 , $p < 0.001$), with a substantial total association ($\beta = 0.501$, $p < 0.001$). These mediation pathways
639 illustrate the complex interplay among soil properties that simple bivariate correlations cannot capture.
640 However, we caution that temporal precedence, a key criterion for causal inference (Hill, 1965; Wade et
641 al., 2026), cannot be established from our single time-point sampling. Thus, while our results are
642 consistent with the hypothesized directional relationships, they represent interaction evidence that
643 contributes to, but does not definitively establish, causal understanding.

644 Model robustness was assessed through multiple approaches: (1) bootstrapped standard errors (n
645 = 1000) for mediation paths, (2) 10-fold cross-validation (CV) with 20 repeats ($R^2 = 0.538$, RMSE =
646 17.92 ± 1.2), and (3) examination of residual correlations (all < 0.10 , indicating acceptable local fit). The
647 exploratory model incorporating POXC, C:N, and P₂O₅ increased explained variance to 57.2% but
648 reduced parsimony, with only POXC showing a significant association ($\beta = -0.259$, $p < 0.001$).
649 Following Wade et al.'s (2026) guidance on distinguishing confirmatory from exploratory analyses, we
650 present these exploratory findings as hypothesis-generating rather than confirmatory, requiring
651 validation in future studies. Collectively, our results demonstrate that pH, aluminum, and dissolved
652 carbon are primary drivers of SOC thermal stability, with clay playing a complex mediating role.
653 Consistent with Wade et al. (2026), who caution that any single analytical approach cannot confirm
654 causal relationships, we position our findings as one line of correlational evidence contributing to an
655 accumulating knowledge base on SOC stabilization.



656
657

Table 3. Standardized path coefficients (β) and bootstrapped standard errors (SE) from SEM showing pathways to SOC thermal stability

Predictors	Std. Coefficient	SE	<i>p</i> -Value
<i>a priori</i> Effects			
pH	0.44	0.06	0.001
Al	0.38	0.06	0.001
DOC	-0.32	0.06	0.001
Clay	0.15	0.05	0.001
BD	0.06	0.06	ns ⁵
Exploratory Effects			
pH	0.46	0.08	0.001
Al	0.36	0.05	0.001
POXC	-0.26	0.08	0.001
Clay	0.16	0.05	0.01
DOC	-0.16	0.07	0.05
C:N	-0.12	0.07	ns ⁵
BD	-0.11	0.07	ns ⁵
P ₂ O ₅	0.04	0.05	ns ⁵

658 Note: ⁵ns = not significant

659

660

661

662

663 **Table 4.** Summary of hypothesis tests evaluating direct and indirect pathways in SEM of SOC thermal
664 stability

Hypothesis	Description	Prediction	Effect Size	<i>p</i> -Value	Confirmed
H ₁	Clay ~ T ₅₀ – Physical protection	positive	0.152	0.00	Yes (direction correct)
H ₂	DOC ~ T ₅₀ – Labile carbon	negative	-0.319	0.00	Yes (direction correct)
H ₃	pH ~ T ₅₀ – Acidity	positive	0.436	0.00	Yes (direction correct)
H ₄	Al ~ T ₅₀ – Mineral reactivity	positive	0.379	0.00	Yes (direction correct)
H ₅	BD ~ T ₅₀ – Structure	positive	0.057	0.31	No (not significant)
H ₆	Clay indirect effects via DOC and Al	mediation	0.048	0.03	Yes
H ₇	pH indirect effects via DOC and Al	mediation	0.241	0.00	Yes



665 Note: Standardized coefficients represent the strength of association. For mediation hypotheses,
 666 coefficients represent the indirect effect magnitude. Supported hypotheses indicate statistical
 667 significance ($p < 0.05$) and direction consistent with predictions
 668

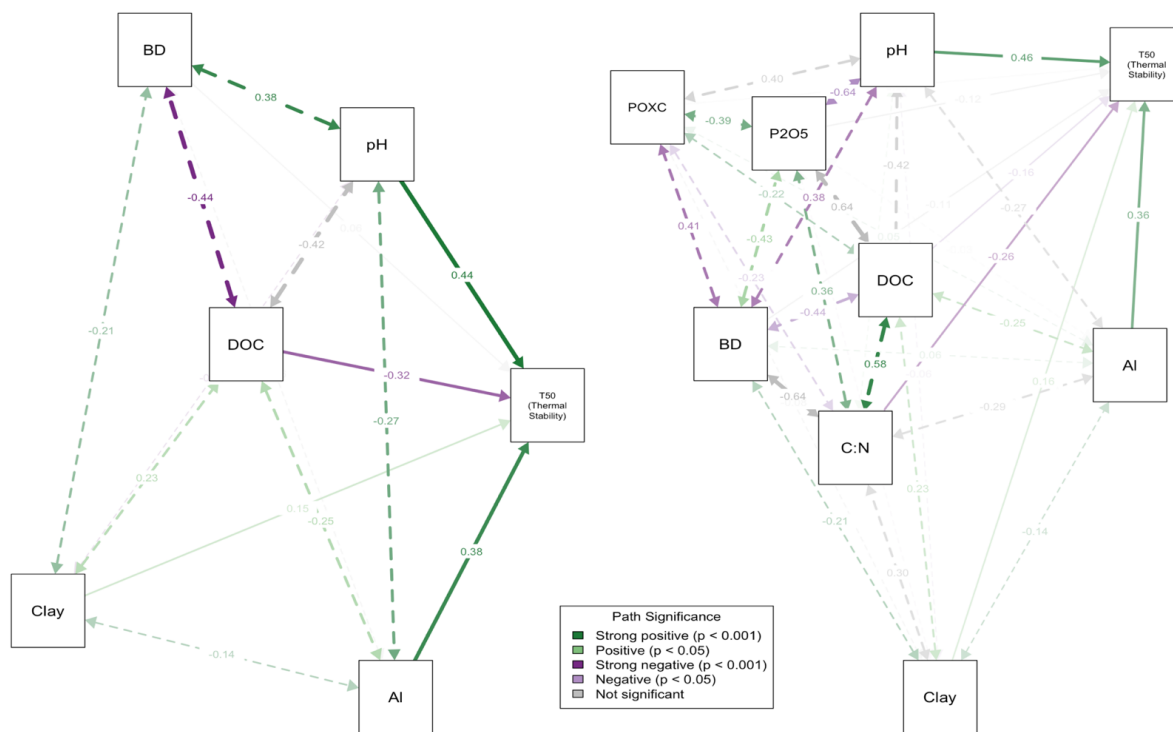
669 **Table 5.** Goodness-of-fit statistics for confirmatory, exploratory, and mediation SEM predicting SOC
 670 thermal stability

Model	Type	⁶ AIC	⁷ BIC	⁸ CFI	⁹ RMSEA	R ²
<i>a priori</i> Hypothesis	Confirmatory (theory-driven)	433.51	456.70	1.00	0.00	0.54
Exploratory	Exploratory	422.84	455.97	1.00	0.00	0.57
Mediation	Confirmatory (mediation)	1520.82	1557.27	0.86	0.41	0.49

671 Note: ⁶Lower AIC values indicate better fit; ⁷Lower BIC values indicate better fit; ⁸CFI ≥ 0.90

672 acceptable, ≥ 0.95 good fit; ⁹RMSEA ≤ 0.08 acceptable, ≤ 0.05 good fit

673



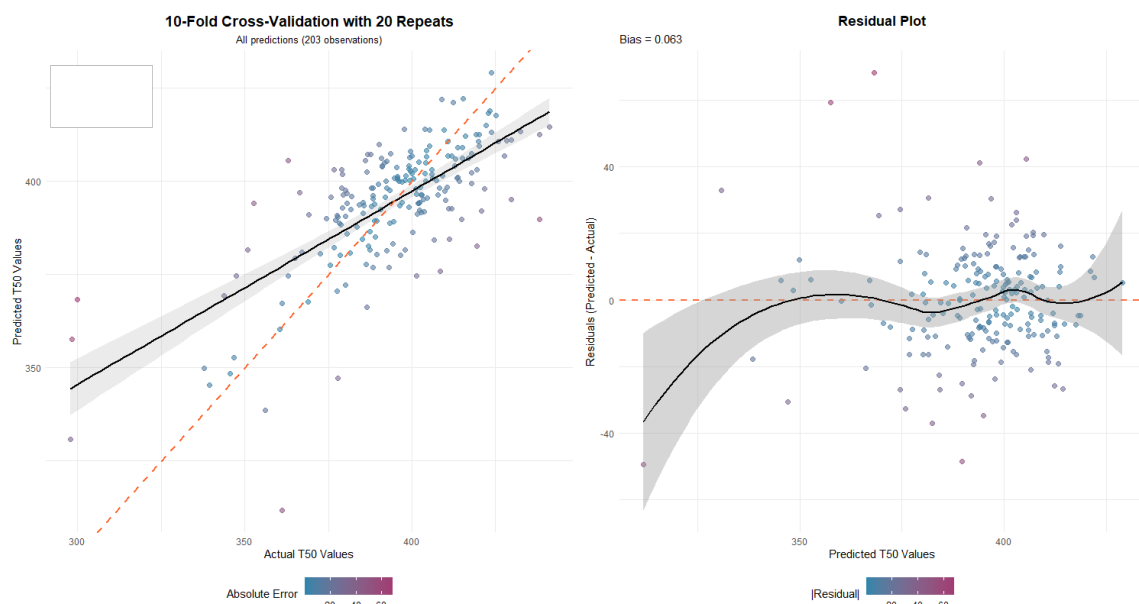
674

675 **Figure 8.** Path diagram from structural equation modeling (SEM) showing the
 676 multivariate relationship between soil properties and thermal stability (T₅₀).

677 Note: Green arrows indicate significant positive relationships ($p < 0.001$, thick; $p < 0.05$, thin); purple
 678 arrows indicate significant negative relationships. Values adjacent to arrows are standardized path
 679 coefficients, representing the strength of association between variables. The model explains 53.5% of
 680 the variance in T₅₀ ($R^2 = 0.54$). Solid arrows represent direct effects, while dashed arrows



681 represent indirect pathways, meaning the predictors influence T_{50} *via* one or more mediators rather
682 than directly.
683



684

685 **Figure 9. Model performance evaluation for the structural equation model (SEM) predicting**
686 **thermal stability(T_{50}) using 10-fold cross-validation repeated 20 times ($n = 203$ sample points).**
687 Left: 1:1 scatterplot of observed vs. predicted T_{50} values with identity line (solid black), regression fit
688 (gray), and 1:1 dashed reference; points colored by absolute error magnitude. Right: Residual plot
689 (predicted – actual) showing low overall bias (0.063) and minimal systematic patterns across predicted
690 values.
691



692 4. CONCLUSIONS

693 In conclusion, assessing SOC stability is essential for understanding carbon sequestration
694 potential and its response to environmental change. Our analysis of thermal stability indices from
695 HAWK-programmed pyrolysis revealed distinct patterns across land uses. Forest soils showed higher
696 T_{50} with moderate to high S_2 , consistent with greater thermal stability and more recalcitrant SOC.
697 Agricultural land use yielded lower and more variable T_{50} with correspondingly high S_1 and HI,
698 suggesting associations with disturbance-driven decomposition and labile carbon pools. Wetlands
699 exhibited high S_2 with variable T_{50} , consistent with fresh organic matter accumulation under anaerobic
700 conditions. Pasture sites showed intermediate values, with moderate T_{50} and HI related to perennial root
701 inputs, grazing, and partial mineral protection.

702 Using random forest modeling, we identified several soil properties—including bulk density,
703 permanganate oxidizable carbon (POXC), cation exchange capacity (CEC), and nutrient content—that
704 showed strong associations with thermal stability indices. Our findings detail how soil properties relate
705 to SOC stability through organo-mineral interactions, microbial activity, and soil structure. The
706 positive correlation between aluminum and thermal stability indices suggests that Al plays a role in SOC
707 stabilization, reflected in higher thermal decomposition temperatures.

708 Partial dependence plots (PDP) and individual conditional expectation (ICE) plots revealed
709 nonlinear relationships between key predictors, enhancing our understanding of thermal
710 decomposition patterns across temperature ranges. Additionally, our SEM analysis supported
711 hypothesized directional relationships: clay and pH showed direct positive associations with T_{50} , while
712 dissolved organic carbon showed a direct negative association with T_{50} . Mediation analysis further
713 revealed indirect pathways through which clay and pH relate to T_{50} via dissolved carbon and aluminum.
714 These results demonstrate that HAWK-programmed pyrolysis thermal stability indices are related
715 to measurable soil properties, capturing complex soil interactions and providing insights into SOC
716 stability.

717 We also identified limitations, particularly the modest predictive performance for OI, suggesting
718 that additional variables (e.g., microbial biomass, enzymatic activity, climate, topography, land use) may
719 better capture SOC oxidation processes. Future research should refine predictive models for lower-
720 performing thermal stability indices using theory-based variables and assess their transferability across
721 regions and soil types. Additionally, we recommend testing hypothesized relationships for other thermal



722 indicators to further elucidate multivariate associations between soil properties and SOC thermal
723 stability. Further investigation into mineral-organic interactions, particularly in soils with high clay
724 content or reactive minerals, will help clarify their role in SOC stabilization. Addressing these gaps will
725 enhance predictive modeling of thermal indices, providing a framework for monitoring SOC
726 sequestration and stability.

727 **CODE AND DATA AVAILABILITY**

728 Data will be made available upon reasonable request but are not available in any online repository to
729 protect the privacy concerns of the participants in this project.

730 **SUPPLEMENT**

731 The supplementary Tables and Figures related to this article are attached along with the manuscript text

732 **AUTHOR CONTRIBUTIONS**

733 KJ, TP, BH, AWG, LPC, and DHL conceptualized the study. LPC contributed to the design of the soil
734 sampling and analytical framework. KJ, TP, and BH conducted the data analysis and modelling, with
735 input from AWG, MM, LPC, and DHL. KJ prepared the original manuscript draft. All authors
736 contributed to the interpretation of the results, reviewed and edited the manuscript, and approved the
737 final version for submission.

738 **DECLARATION OF COMPETING INTEREST**

739 The authors declare that they have no known competing financial interests or personal relationships that
740 could have appeared to influence the work reported in this paper.

741 **ACKNOWLEDGEMENTS**

742 We are very grateful to all landowners who have given us access to their property to collect soil samples.
743 For soil sample collection and general soil analyses, we would like to thank: Technicians Kyle
744 MacKinley, Karen Terry, Mohammad Islam, and Pam MacKinley; Students Vanessa Mossman,
745 Valentin Nugues, Erika Young, Matthew A.C. Milne, Kayli McGarrigle, Allysia Murphy, Simon
746 Bertheleme, Violet Eliza-Sioux James, and Madison White. The authors recognize funding received
747 through the Sustainable Agriculture Research Initiative Grant (ALLRP 588521-23), supported jointly by
748 the Natural Sciences and Engineering Research Council of Canada (NSERC), the Social Sciences and



749 Humanities Research Council (SSHRC), and Agriculture and Agri-Food Canada. This research was
750 undertaken, in part, thanks to funding from the Canada Research Chairs Program awarded to Heung.

751

752 REFERENCES

753

754 Almagro, M., Ruiz-Navarro, A., Díaz-Pereira, E., Albaladejo, J., and Martínez-Mena, M.: Plant residue
755 chemical quality modulates the soil microbial response related to decomposition and soil
756 organic carbon and nitrogen stabilization in a rainfed Mediterranean agroecosystem, *Soil Biol.*
757 *Biochem.*, 156, 108198, <https://doi.org/10.1016/j.soilbio.2021.108198>, 2021.

758 An, Z., Pokharel, P., Plante, A. F., Bork, E. W., Carlyle, C. N., Williams, E. K., & Chang, S. X.: Soil
759 organic matter stability in forest and cropland components of two agroforestry systems in
760 western Canada. *Geoderma*, 433, 116463, 2023

761 Apley, D. W., and Zhu, J.: Visualizing the Effects of Predictor Variables in Black Box Supervised
762 Learning Models, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 82(4), 1059-1086,
763 <https://doi.org/10.1111/rssb.12377>, 2020.

764 Arbor, A., Schmidt, M., Saurette, D., Zhang, J., Bulmer, C., Filatow, D., Kasraei, B., Smukler, S., and
765 Heung, B.: A framework for recalibrating pedotransfer functions using nonlinear least squares
766 and estimating uncertainty using quantile regression, *Geoderma*, 439, 116674,
767 <https://doi.org/10.1016/j.geoderma.2023.116674>, 2023.

768 Arbor, A., Schmidt, M., Zhang, J., Bulmer, C., Filatow, D., Kasraei, B., Smukler, S., and Heung, B.:
769 Filling the gaps in soil data: A multi-model framework for addressing data gaps using
770 pedotransfer functions and machine-learning with uncertainty estimates to estimate bulk
771 density, *Catena*, 245, 108310, <https://doi.org/10.1016/j.catena.2024.108310>, 2024.

772 Barreto, M. S. C., Elzinga, E. J., Ramlogan, M., Rouff, A. A., and Alleoni, L. R. F.: Calcium enhances
773 adsorption and thermal stability of organic compounds on soil minerals, *Chem. Geol.*, 559,
774 119804, <https://doi.org/10.1016/j.chemgeo.2020.119804>, 2021.

775 Barreto, M. S. C., Ramlogan, M., Oliveira, D. M. S., Verburg, E. E. J., Elzinga, E. J., Rouff, A. A.,
776 Jemo, M., and Alleoni, L. R. F.: Thermal stability of soil organic carbon after long-term
777 manure application across land uses and tillage systems in an oxisol, *Catena*, 200, 105164,
778 <https://doi.org/10.1016/j.catena.2021.105164>, 2021.

779 Barros, N., Salgado, J., Villanueva, M., Rodríguez-Añón, J., Proupin, J., Feijóo, S., and Martín-Pastor,
780 M.: Application of DSC–TG and NMR to study the soil organic matter, *J. Therm. Anal.*
781 *Calorim.*, 104(1), 53-60, <https://doi.org/10.1007/s10973-010-1163-4>, 2011.

782 Barré, P., Cécillon, L., and Kanari, E.: Characterization and Evaluation of the Stability of Soil Organic
783 Matter, *The Rock-Eval Method*, 181-207, <https://doi.org/10.1002/9781394256235.ch10>, 2024.

784 Baudin, F., Disnar, J. R., Aboussou, A., and Savignac, F.: Guidelines for Rock–Eval analysis of recent
785 marine sediments, *Org. Geochem.*, 86, 71-80,
786 <https://doi.org/10.1016/j.orggeochem.2015.06.009>, 2015.



- 787 Berthrong, S. T., Jobbágy, E. G., and Jackson, R. B.: A global meta-analysis of soil exchangeable
788 cations, pH, carbon, and nitrogen with afforestation, *Ecol. Appl.*, 19(8), 2228-2241,
789 <https://doi.org/10.1890/08-1730.1>, 2009.
- 790 Berthrong, S. T., Piñeiro, G., Jobbágy, E. G., and Jackson, R. B.: Soil C and N changes with
791 afforestation of grasslands across gradients of precipitation and plantation age, *Ecol. Appl.*,
792 22(1), 76-86, <https://doi.org/10.1890/10-2210.1>, 2012.
- 793 Bouma, J.: Using Soil Survey Data for Quantitative Land Evaluation, *Advances in Soil Science*, 177-
794 213, https://doi.org/10.1007/978-1-4612-3532-3_4, 1989.
- 795 Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: Classification and regression trees,
796 Wadsworth International Group, Belmont, California, 1984.
- 797 Breiman, L.: Random forests, *Machine Learning*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>,
798 2001.
- 799 Briedis, C., de Moraes Sá, J. C., Lal, R., de Oliveira Ferreira, A., Franchini, J. C., and Milori, D. M. B.
800 P.: Preservation of labile organic compounds is the pathway for carbon storage in a 23-year
801 continuous no-till system on a Ferralsol in southern Brazil, *Geoderma Regional*, 33, e00643,
802 <https://doi.org/10.1016/j.geodrs.2023.e00643>, 2023.
- 803 Bronick, C., and Lal, R.: Soil structure and management: a review, *Geoderma*, 124(1-2), 3-22,
804 <https://doi.org/10.1016/j.geoderma.2004.03.005>, 2005.
- 805 Casalicchio, G., Molnar, C., and Schratz, P.: iml: Interpretable Machine Learning, CRAN: Contributed
806 Packages, <https://doi.org/10.32614/cran.package.iml>, 2018.
- 807 Cécillon, L., Baudin, F., Chenu, C., Houot, S., Jolivet, R., Kätterer, T., Lutfalla, S., Macdonald, A., van
808 Oort, F., Plante, A. F., Savignac, F., Soucémarianadin, L. N., and Barré, P.: A model based on
809 Rock-Eval thermal analysis to quantify the size of the centennially persistent organic carbon
810 pool in temperate soils, *Biogeosciences*, 15(9), 2835-2849, [https://doi.org/10.5194/bg-15-2835-](https://doi.org/10.5194/bg-15-2835-2018)
811 2018, 2018.
- 812 Cotrufo, M. F., Haddix, M. L., Kroeger, M. E., and Stewart, C. E.: The role of plant input physical-
813 chemical properties, and microbial and soil chemical diversity on the formation of particulate
814 and mineral-associated organic matter, *Soil Biol. Biochem.*, 168, 108648,
815 <https://doi.org/10.1016/j.soilbio.2022.108648>, 2022.
- 816 Cotrufo, M. F., Ranalli, M. G., Haddix, M. L., Six, J., and Lugato, E.: Soil carbon storage informed by
817 particulate and mineral-associated organic matter, *Nature Geoscience*, 12(12), 989-994,
818 <https://doi.org/10.1038/s41561-019-0484-6>, 2019.
- 819 Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Deneff, K., and Paul, E.: The Microbial Efficiency-
820 Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic
821 matter stabilization: do labile plant inputs form stable soil organic matter?, *Glob. Change Biol.*,
822 19(4), 988-995, <https://doi.org/10.1111/gcb.12113>, 2013.
- 823 Craney, T. A., and Surles, J. G.: Model-Dependent Variance Inflation Factor Cutoff Values, *Quality*
824 *Engineering*, 14(3), 391-403, <https://doi.org/10.1081/qen-120001878>, 2002.



- 825 Cremer, M., and Prietzel, J.: Soil acidity and exchangeable base cation stocks under pure and mixed
826 stands of European beech, Douglas fir and Norway spruce, *Plant Soil*, 415(1-2), 393-405,
827 <https://doi.org/10.1007/s11104-017-3177-1>, 2017.
- 828 Culman, S. W., Snapp, S. S., Freeman, M. A., Schipanski, M. E., Beniston, J., Lal, R., Drinkwater, L.
829 E., Franzluebbers, A. J., Glover, J. D., Grandy, A. S., Lee, J., Six, J., Maul, J. E., Mirksy, S. B.,
830 Spargo, J. T., and Wander, M. M.: Permanganate Oxidizable Carbon Reflects a Processed Soil
831 Fraction that is Sensitive to Management, *Soil Sci. Soc. Am. J.*, 76(2), 494-504,
832 <https://doi.org/10.2136/sssaj2011.0286>, 2012.
- 833 Curto, J. D., and Pinto, J. C.: The corrected VIF (CVIF), *Journal of Applied Statistics*, 38(7), 1499-
834 1507, <https://doi.org/10.1080/02664763.2010.505956>, 2011.
- 835 da Silva, H. F., de Sousa, P. V. F., Escobar, M. E. O., and de Oliveira, T. S. (2025). Changes in the
836 composition of soil organic matter caused by organic and conventional management in the long
837 term. *Journal of Environmental Management*, 374, 124018, 2025.
- 838 Davidson, E. A., and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and
839 feedbacks to climate change, *Nature*, 440(7081), 165-173, <https://doi.org/10.1038/nature04514>,
840 2006.
- 841 Delahaie, A. A., Barré, P., Baudin, F., Arrouays, D., Bispo, A., Boulonne, L., Chenu, C., Jolivet, C.,
842 Martin, M. P., Ratić, C., Saby, N. P. A., Savignac, F., and Cécillon, L.: Elemental
843 stoichiometry and Rock-Eval® thermal stability of organic matter in French topsoils, *SOIL*,
844 9(1), 209-229, <https://doi.org/10.5194/soil-9-209-2023>, 2023.
- 845 De Neve, S. and Hofman, G.: Influence of soil compaction on carbon and nitrogen mineralization of
846 soil organic matter and crop residues, *Biol. Fertil. Soils*, 30, 544–549,
847 <https://doi.org/10.1007/s003740050034>, 2000.
- 848 Dexter, A.R.: Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter,
849 and effects on root growth. *Geoderma*, 120(3-4), pp.201-214, 2004
- 850 Disnar, J., Guillet, B., Keravis, D., Di-Giovanni, C., and Sebag, D.: Soil organic matter (SOM)
851 characterization by Rock-Eval pyrolysis: scope and limitations, *Org. Geochem.*, 34(3), 327-
852 343, [https://doi.org/10.1016/s0146-6380\(02\)00239-5](https://doi.org/10.1016/s0146-6380(02)00239-5), 2003.
- 853 Doležalová-Weissmannová, H., Malý, S., Brtnický, M., Holátko, J., Demyan, M. S., Siewert, C.,
854 Tokarski, D., Kameníková, E., and Kučerík, J.: Practical applications of thermogravimetry in
855 soil science: Part 5. Linking the microbial soil characteristics of grassland and arable soils to
856 thermogravimetry data, *J. Therm. Anal. Calorim.*, 148, 1599–1611,
857 <https://doi.org/10.1007/s10973-022-11709-6>, 2023.
- 858 Dungait, J. A. J., Hopkins, D. W., Gregory, A. S., and Whitmore, A. P.: Soil organic matter turnover is
859 governed by accessibility not recalcitrance, *Glob. Change Biol.*, 18, 1781–1796,
860 <https://doi.org/10.1111/j.1365-2486.2012.02665.x>, 2012.
- 861 Entio, L. J., Taggart, C. B., Muir, J. P., Kan, E., Brady, J. A., and Obayomi, O.: Biochar and Dairy
862 Manure Amendment Effects on *Cynodon dactylon* Performance and Soil Properties, *Plants*,
863 13(2), 242, <https://doi.org/10.3390/plants13020242>, 2024.



- 864 Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D.: qgraph:
865 Network visualizations of relationships in psychometric data, *J. Stat. Softw.*, 48, 1–18,
866 <https://doi.org/10.18637/jss.v048.i04>, 2012.
- 867 Fahmy, S. H., Rees, H. W., Wang, C., and Walker, J. A.: Soil survey reports of New Brunswick,
868 Agriculture and Agri-Food Canada, Fredericton, New Brunswick, 2010.
- 869 Fang, K., Chen, L., Qin, S., Zhang, Q., Liu, X., Chen, P., and Yang, Y.: Mineral and Climatic Controls
870 Over Soil Organic Matter Stability Across the Tibetan Alpine Permafrost Region, *Global*
871 *Biogeochemical Cycles*, 35(12), e2021GB007118, <https://doi.org/10.1029/2021gb007118>,
872 2021.
- 873 Fernández, J. M., Peltre, C., Craine, J. M., and Plante, A. F.: Improved Characterization of Soil
874 Organic Matter by Thermal Analysis Using CO₂/H₂O Evolved Gas
875 Analysis, *Environmental Science & Technology*, 46(16), 8921-8927,
876 <https://doi.org/10.1021/es301375d>, 2012.
- 877 Fernández, J. M., Plante, A. F., Leifeld, J., and Rasmussen, C.: Methodological considerations for
878 using thermal analysis in the characterization of soil organic matter, *J. Therm. Anal. Calorim.*,
879 104(1), 389-398, <https://doi.org/10.1007/s10973-010-1145-6>, 2011.
- 880 Furze, S. and Arp, P.: Amalgamation and harmonization of soil survey reports into a multi-purpose
881 database. *Canadian Journal of Soil Science*, 101(2), pp.222-247, 2020
- 882 Furze, S., O’Sullivan, A.M., Allard, S., Pronk, T. and Curry, R.A., 2021. A high-resolution, random
883 forest approach to mapping depth-to-bedrock across shallow overburden and post-glacial
884 terrain. *Remote Sensing*, 13(21), p.4210, 2021
- 885 Friedman, J. H.: Greedy function approximation: A gradient boosting machine, *The Annals of*
886 *Statistics*, 29(5), <https://doi.org/10.1214/aos/1013203451>, 2001.
- 887 Gao, W., Zhou, T., and Ren, T.: Conversion from Conventional to No Tillage Alters Thermal Stability
888 of Organic Matter in Soil Aggregates, *Soil Sci. Soc. Am. J.*, 79(2), 585-594,
889 <https://doi.org/10.2136/sssaj2014.08.0334>, 2015.
- 890 Giannetta, B., Plaza, C., Vischetti, C., Cotrufo, M. F., and Zaccaro, C.: Distribution and thermal
891 stability of physically and chemically protected organic matter fractions in soils across different
892 ecosystems, *Biology and Fertility of Soils*, 54(5), 671-681, [https://doi.org/10.1007/s00374-018-](https://doi.org/10.1007/s00374-018-1290-9)
893 1290-9, 2018.
- 894 Gillespie, A., Sanei, H., Diochon, A., Ellert, B., Regier, T., Chevrier, D., Dynes, J., Tarnocai, C., and
895 Gregorich, E.: Perennially and annually frozen soil carbon differ in their susceptibility to
896 decomposition: Analysis of Subarctic earth hummocks by bioassay, XANES and pyrolysis, *Soil*
897 *Biol. Biochem.*, 68, 106-116, <https://doi.org/10.1016/j.soilbio.2013.09.021>, 2014.
- 898 Goñi, F. M., and Alonso, A.: Differential Scanning Calorimetry in the Study of Lipid Structures,
899 *Chemical Biology*, 47-66, <https://doi.org/10.1002/9780470319253.ch4>, 2006.
- 900 Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E.: Peeking Inside the Black Box: Visualizing
901 Statistical Learning With Plots of Individual Conditional Expectation, *Journal of*
902 *Computational and Graphical Statistics*, 24(1), 44-65,
903 <https://doi.org/10.1080/10618600.2014.907095>, 2015.



- 904 Grace, J. B., Scheiner, S. M., and Schoolmaster, D. R.: Structural equation modeling, in: Ecological
905 Statistics: Contemporary Theory and Application, edited by: Fox, G. A., Negrete-Yankelevich,
906 S., and Sosa, V. J., Oxford University Press, Oxford, 168–199,
907 <https://doi.org/10.1093/acprof:oso/9780199672547.003.0009>, 2015.
- 908 Gregorich, E., Gillespie, A., Beare, M., Curtin, D., Sanei, H., and Yanni, S.: Evaluating
909 biodegradability of soil organic matter by its thermal stability and chemical composition, *Soil*
910 *Biol. Biochem.*, 91, 182–191, <https://doi.org/10.1016/j.soilbio.2015.08.032>, 2015.
- 911 Hall, S. J., and Silver, W. L.: Reducing conditions, reactive metals, and their interactions can explain
912 spatial patterns of surface soil carbon in a humid tropical forest, *Biogeochemistry*, 125(2), 149–
913 165, <https://doi.org/10.1007/s10533-015-0120-5>, 2015.
- 914 Hall, S. J., Silver, W. L., Timokhin, V. I., and Hammel, K. E.: Iron addition to soil specifically
915 stabilized lignin, *Soil Biol. Biochem.*, 98, 95–98, <https://doi.org/10.1016/j.soilbio.2016.04.010>,
916 2016.
- 917 Haynes, R.: Labile Organic Matter Fractions as Central Components of the Quality of Agricultural
918 Soils: An Overview, *Advances in Agronomy*, 221–268, [https://doi.org/10.1016/s0065-](https://doi.org/10.1016/s0065-2113(04)85005-3)
919 [2113\(04\)85005-3](https://doi.org/10.1016/s0065-2113(04)85005-3), 2005.
- 920 Hill, A. B.: The environment and disease: association or causation?, *Proc. R. Soc. Med.*, 58, 295–300,
921 1965.
- 922 Hu, L. and Bentler, P. M.: Cutoff criteria for fit indexes in covariance structure analysis: conventional
923 criteria versus new alternatives, *Struct. Equ. Modeling*, 6, 1–55,
924 <https://doi.org/10.1080/10705519909540118>, 1999.
- 925 IUSS Working Group WRB: World Reference Base for Soil Resources 2006, first update 2007, *World*
926 *Soil Resources Reports*, 103, FAO, Rome, 2007.
- 927 Kasraei, B., Schmidt, M. G., Zhang, J., Bulmer, C. E., Filatow, D. S., Arbor, A., Pennell, T., and
928 Heung, B.: A framework for optimizing environmental covariates to support model
929 interpretability in digital soil mapping, *Geoderma*, 445, 116873,
930 <https://doi.org/10.1016/j.geoderma.2024.116873>, 2024.
- 931 Kögel-Knabner, I.: The macromolecular organic composition of plant and microbial residues as inputs
932 to soil organic matter, *Soil Biol. Biochem.*, 34(2), 139–162, [https://doi.org/10.1016/s0038-](https://doi.org/10.1016/s0038-0717(01)00158-4)
933 [0717\(01\)00158-4](https://doi.org/10.1016/s0038-0717(01)00158-4), 2002.
- 934 King, A. E., Amsili, J. P., Córdova, S. C., Culman, S., Fonte, S. J., Kotcon, J., Masters, M. D., McVay,
935 K., Olk, D. C., Prairie, A. M., Schipanski, M., Schneider, S. K., Stewart, C. E., and Cotrufo, M.
936 F.: Constraints on mineral-associated and particulate organic carbon response to regenerative
937 management: carbon inputs and saturation deficit, *Soil and Tillage Research*, 238, 106008,
938 <https://doi.org/10.1016/j.still.2024.106008>, 2024.
- 939 Kirkby, C. A., Richardson, A. E., Wade, L. J., Batten, G. D., Blanchard, C., and Kirkegaard, J. A.:
940 Carbon-nutrient stoichiometry to increase soil carbon sequestration, *Soil Biol. Biochem.*, 60,
941 77–86, <https://doi.org/10.1016/j.soilbio.2013.01.011>, 2013.



- 942 Kleber, M., Sollins, P., and Sutton, R.: A conceptual model of organo-mineral interactions in soils:
943 self-assembly of organic molecular fragments into zonal structures on mineral surfaces,
944 *Biogeochemistry*, 85, 9–24, <https://doi.org/10.1007/s10533-007-9103-5>, 2007.
- 945 Kučerík, J., Tokarski, D., Demyan, M. S., Merbach, I., and Siewert, C.: Linking soil organic matter
946 thermal stability with contents of clay, bound water, organic carbon and nitrogen, *Geoderma*,
947 316, 38-46, <https://doi.org/10.1016/j.geoderma.2017.12.001>, 2018.
- 948 Kučerík, J., Čtvrtníčková, A., and Siewert, C.: Practical application of thermogravimetry in soil
949 science, *J. Therm. Anal. Calorim.*, 113(3), 1103-1111, <https://doi.org/10.1007/s10973-012-2849-6>, 2013.
- 950
- 951 Kuhn, M.: caret: Classification and Regression Training, CRAN: Contributed Packages,
952 <https://doi.org/10.32614/cran.package.caret>, 2007.
- 953 Lal, R.: Soil Carbon Sequestration Impacts on Global Climate Change and Food Security, *Science*,
954 304(5677), 1623-1627, <https://doi.org/10.1126/science.1097396>, 2004.
- 955 Laurence, L., Heung, B., Strom, H., Stiles, K., and Burton, D.: Towards a cost-effective framework for
956 estimating soil nitrogen pools using pedotransfer functions and machine learning, *Geoderma*,
957 440, 116692, <https://doi.org/10.1016/j.geoderma.2023.116692>, 2023.
- 958 Laurence, L., Heung, B., Zhang, J., Pennell, T., Nyiraneza, J., Strom, H., Stiles, K., and Burton, D. L.:
959 Integrating multi-year crop inventories as a proxy for soil management within a digital soil
960 mapping framework for predicting nitrogen indices, *Geoderma*, 448, 116944,
961 <https://doi.org/10.1016/j.geoderma.2024.116944>, 2024.
- 962 Lehmann, J., and Kleber, M.: The contentious nature of soil organic matter, *Nature*, 528(7580), 60-68,
963 <https://doi.org/10.1038/nature16069>, 2015.
- 964 Leifeld, J., Siebert, S., and Kögel-Knabner, I.: Changes in the chemical composition of soil organic
965 matter after application of compost, *Eur. J. Soil Sci.*, 53(2), 299-309,
966 <https://doi.org/10.1046/j.1351-0754.2002.00453.x>, 2002.
- 967 Li, F., Gui, X., Ji, W., and Zhou, C.: Effect of calcium dihydrogen phosphate addition on carbon
968 retention and stability of biochars derived from cellulose, hemicellulose, and lignin,
969 *Chemosphere*, 251, 126335, <https://doi.org/10.1016/j.chemosphere.2020.126335>, 2020.
- 970 Liu, M., Qiao, N., Xu, X., Fang, H., Wang, H., and Kuzyakov, Y.: C:N stoichiometry of stable and
971 labile organic compounds determine priming patterns, *Geoderma*, 362, 114122,
972 <https://doi.org/10.1016/j.geoderma.2019.114122>, 2020.
- 973 Li, Z., Xiao, L., Deng, C., Yuan, Z., Liang, C., Xiong, Q., Li, Z., and Nie, X.: Thermal stability of soil
974 organic carbon subjected to water erosion as a function of edaphic factors, *Int. J. Sediment
975 Res.*, 37(1), 26-36, <https://doi.org/10.1016/j.ijsrc.2021.06.002>, 2022.
- 976 Lorenz, M., Maskow, T., and Thiele-Bruhn, S.: Energy stored in soil organic matter is influenced by
977 litter quality and the degree of transformation – A combustion calorimetry study, *Geoderma*,
978 443, 116846, <https://doi.org/10.1016/j.geoderma.2024.116846>, 2024.
- 979 Ma, D., He, Z., Zhao, W., Li, R., Sun, W., Wang, W., Lin, P., Wei, L., and Ju, W.: Long-term effects
980 of conventional cultivation on soil cation exchange capacity and base saturation in an arid



- 981 desert region, *Sci. Total Environ.*, 949, 175075,
982 <https://doi.org/10.1016/j.scitotenv.2024.175075>, 2024.
- 983 Mayer, M., Leifeld, J., Szidat, S., Mäder, P., Krause, H. M., and Steffens, M.: Dynamic stability of
984 mineral-associated organic matter: enhanced stability and turnover through organic fertilization
985 in a temperate agricultural topsoil, *Soil Biol. Biochem.*, 184, 109095,
986 <https://doi.org/10.1016/j.soilbio.2023.109095>, 2023.
- 987 McElhaney, R. N.: The use of differential scanning calorimetry and differential thermal analysis in
988 studies of model and biological membranes, *Chemistry and Physics of Lipids*, 30(2-3), 229-
989 259, [https://doi.org/10.1016/0009-3084\(82\)90053-6](https://doi.org/10.1016/0009-3084(82)90053-6), 1982.
- 990 Mesgar, M., Voroney, R. P., Lo, A., Ardakani, O. H., and Gillespie, A. W.: Chemical composition and
991 thermal stability of topsoil organic carbon: Influence of cropping system and tillage practices,
992 *Eur. J. Soil Sci.*, 75(1), e13459, <https://doi.org/10.1111/ejss.13459>, 2024.
- 993 Minasny, B., and Mcbratney, A. B.: The Neuro-m Method for Fitting Neural Network Parametric
994 Pedotransfer Functions, *Soil Sci. Soc. Am. J.*, 66(4), <https://doi.org/10.2136/sssaj2002.1407a>,
995 2002.
- 996 Mitchell, E., Scheer, C., Rowlings, D., Cotrufo, M. F., Conant, R. T., Friedl, J., and Grace, P.: Trade-
997 off between ‘new’ SOC stabilisation from above-ground inputs and priming of native C as
998 determined by soil type and residue placement, *Biogeochemistry*, 149(2), 221-236,
999 <https://doi.org/10.1007/s10533-020-00675-6>, 2020.
- 1000 Moore, P., Lukenbach, M., Kettridge, N., Petrone, R., Devito, K., and Waddington, J.: Peatland water
1001 repellency: Importance of soil water content, moss species, and burn severity, *Journal of*
1002 *Hydrology*, 554, 656-665, <https://doi.org/10.1016/j.jhydrol.2017.09.036>, 2017.
- 1003 Murphy, B. W.: *Soil Organic Matter and Soil Function – Review of the Literature and Underlying*
1004 *Data*, Department of the Environment, Canberra, Australia, 2014.
- 1005 Nie, X., Yuan, Z., Huang, B., Liao, Y., Zhang, X., Li, Z., and Li, D.: Effects of water erosion on soil
1006 organic carbon stability in the subtropical China, *Journal of Soils and Sediments*, 19(10), 3564-
1007 3575, <https://doi.org/10.1007/s11368-019-02305-7>, 2019.
- 1008 Padarian, J., Minasny, B., and McBratney, A. B.: Machine learning and soil sciences: a review aided
1009 by machine learning tools, *SOIL*, 6(1), 35-52, <https://doi.org/10.5194/soil-6-35-2020>, 2020.
- 1010 Paul, E. A., Morris, S. J., Conant, R. T., and Plante, A. F.: Does the Acid Hydrolysis–Incubation
1011 Method Measure Meaningful Soil Organic Carbon Pools?, *Soil Sci. Soc. Am. J.*, 70(3), 1023-
1012 1035, <https://doi.org/10.2136/sssaj2005.0103>, 2006.
- 1013 Paul, S. S., Heung, B., and Lynch, D. H.: Modeling of total and active organic carbon dynamics in
1014 agricultural soil using digital soil mapping: a case study from Central Nova Scotia, *Canadian*
1015 *Journal of Soil Science*, 103(1), 64-80, <https://doi.org/10.1139/cjss-2022-0012>, 2023.
- 1016 Peltre, C., Fernández, J. M., Craine, J. M., and Plante, A. F.: Relationships between Biological and
1017 Thermal Indices of Soil Organic Matter Stability Differ with Soil Organic Carbon Level, *Soil*
1018 *Sci. Soc. Am. J.*, 77(6), 2020-2028, <https://doi.org/10.2136/sssaj2013.02.0081>, 2013.
- 1019 Plante, A. F., Fernández, J. M., and Leifeld, J.: Application of thermal analysis techniques in soil
1020 science, *Geoderma*, 153(1-2), 1-10, <https://doi.org/10.1016/j.geoderma.2009.08.016>, 2009.



- 1021 Plante, A. F., Fernández, J. M., Haddix, M. L., Steinweg, J. M., and Conant, R. T.: Biological,
1022 chemical and thermal indices of soil organic matter stability in four grassland soils, *Soil Biol.*
1023 *Biochem.*, 43(5), 1051-1058, <https://doi.org/10.1016/j.soilbio.2011.01.024>, 2011.
- 1024 Preacher, K. J. and Hayes, A. F.: Asymptotic and resampling strategies for assessing and comparing
1025 indirect effects in multiple mediator models, *Behav. Res. Methods*, 40, 879–891,
1026 <https://doi.org/10.3758/BRM.40.3.879>, 2008.
- 1027 Pronk, A. G. and Ruitenbergh, A. A.: Bedrock geology of New Brunswick, New Brunswick Department
1028 of Natural Resources and Energy, Minerals and Energy Division, Map NR-1, scale 1:500000,
1029 1991.
- 1030 Rakhsh, F., Golchin, A., Beheshti Al Agha, A., and Alamdari, P.: Effects of exchangeable cations,
1031 mineralogy and clay content on the mineralization of plant residue carbon, *Geoderma*, 307,
1032 150-158, <https://doi.org/10.1016/j.geoderma.2017.07.010>, 2017.
- 1033 Rampton, V. N.: Generalized surficial geology map of New Brunswick, Geological Survey of Canada,
1034 Map 1594A, scale 1:500000, 1984.
- 1035 Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A.,
1036 Blankinship, J. C., Crow, S. E., Druhan, J. L., Hicks Pries, C. E., Marin-Spiotta, E., Plante, A.
1037 F., Schädel, C., Schimel, J. P., Sierra, C. A., Thompson, A., and Wagai, R.: Beyond clay:
1038 towards an improved set of variables for predicting soil organic matter content,
1039 *Biogeochemistry*, 137, 297–306, <https://doi.org/10.1007/s10533-018-0424-3>, 2018.
- 1040 Risoluti, R., Gullifa, G., Barone, L., Papa, E., and Materazzi, S.: On-Line Thermally Induced Evolved
1041 Gas Analysis: An Update—Part 1: EGA-MS, *Molecules*, 27(11), 3518,
1042 <https://doi.org/10.3390/molecules27113518>, 2022.
- 1043 Rombolà, A. G., Fabbri, D., Meredith, W., Snape, C. E., and Dieguez-Alonso, A.: Molecular
1044 characterization of the thermally labile fraction of biochar by hydropyrolysis and pyrolysis-
1045 GC/MS, *J. Anal. Appl. Pyrol.*, 121, 230–239, <https://doi.org/10.1016/j.jaap.2016.08.003>, 2016.
- 1046 Rosseel, Y.: lavaan: An R Package for Structural Equation Modeling, *J. Stat. Softw.*, 48(2),
1047 <https://doi.org/10.18637/jss.v048.i02>, 2012.
- 1048 Rowley, M. C., Grand, S., and Verrecchia, É. P.: Calcium-mediated stabilisation of soil organic
1049 carbon, *Biogeochemistry*, 137, 27–49, <https://doi.org/10.1007/s10533-017-0410-1>, 2018.
- 1050 Saurette, D. D.: onsoilsurvey: Making PDSM in Ontario Better, Ontario Ministry of Agriculture, Food
1051 and Rural Affairs (OMAFRA), Guelph, Ontario, Canada, 2021.
- 1052 Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M.,
1053 Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S.,
1054 and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, *Nature*,
1055 478(7367), 49-56, <https://doi.org/10.1038/nature10386>, 2011.
- 1056 Sebag, D., Verrecchia, E., Cécillon, L., Adatte, T., Albrecht, R., Aubert, M., Bureau, F., Cailleau, G.,
1057 Copard, Y., Decaens, T., Disnar, J. R., Hetényi, M., Nyilas, T., and Trombino, L.: Dynamics of
1058 soil organic matter based on new Rock-Eval indices, *Geoderma*, 284, 185-203,
1059 <https://doi.org/10.1016/j.geoderma.2016.08.025>, 2016.



- 1060 Shi, D., Lee, T., and Maydeu-Olivares, A.: Understanding the Model Size Effect on SEM Fit Indices,
1061 Educational and Psychological Measurement, 79(2), 310-334,
1062 <https://doi.org/10.1177/0013164418783530>, 2019.
- 1063 Shi, J., Lv, J., Peng, Y., Yao, Y., Wei, X., and Wang, X.: Mechanisms controlling the stability and
1064 sequestration of mineral associated organic carbon upon erosion and deposition, *Catena*, 242,
1065 108119, <https://doi.org/10.1016/j.catena.2024.108119>, 2024.
- 1066 Shipley, B.: Cause and correlation in biology: a user's guide to path analysis, structural equations and
1067 causal inference with R, 2nd edn., Cambridge University Press, Cambridge,
1068 <https://doi.org/10.1017/CBO9781139979573>, 2016.
- 1069 Six, J., and Paustian, K.: Aggregate-associated soil organic matter as an ecosystem property and a
1070 measurement tool, *Soil Biol. Biochem.*, 68, A4-A9,
1071 <https://doi.org/10.1016/j.soilbio.2013.06.014>, 2014.
- 1072 Six, J., Bossuyt, H., Degryze, S., and Deneff, K.: A history of research on the link between
1073 (micro)aggregates, soil biota, and soil organic matter dynamics, *Soil and Tillage Research*,
1074 79(1), 7-31, <https://doi.org/10.1016/j.still.2004.03.008>, 2004.
- 1075 Six, J., Conant, R. T., Paul, E. A., and Paustian, K.: Stabilization mechanisms of soil organic matter:
1076 Implications for C-saturation of soils, *Plant Soil*, 241(2), 155-176,
1077 <https://doi.org/10.1023/a:1016125726789>, 2002.
- 1078 Solomon, D., Lehmann, J., Harden, J., Wang, J., Kinyangi, J., Heymann, K., Karunakaran, C., Lu, Y.,
1079 Wirrick, S., and Jacobsen, C.: Micro- and nano-environments of carbon sequestration: Multi-
1080 element STXM–NEXAFS spectromicroscopy assessment of microbial carbon and mineral
1081 associations, *Chem. Geol.*, 329, 53-73, <https://doi.org/10.1016/j.chemgeo.2012.02.002>, 2012.
- 1082 Sparks, D. L.: Chemistry of soil organic matter, *Environmental Soil Chemistry*, 53-80,
1083 <https://doi.org/10.1016/b978-0-12-656445-7.50007-3>, 1995.
- 1084 Smith, P., Cotrufo, M. F., Rumpel, C., Paustian, K., Kuikman, P. J., Elliott, J. A., McDowell, R.,
1085 Griffiths, R. I., Asakawa, S., Bustamante, M., House, J. I., Sobocká, J., Harper, R., Pan, G.,
1086 West, P. C., Gerber, J. S., Clark, J. M., Adhya, T., Scholes, R. J., and Scholes, M. C.:
1087 Biogeochemical cycles and biodiversity as key drivers of ecosystem services provided by soils,
1088 *SOIL*, 1, 665–685, <https://doi.org/10.5194/soil-1-665-2015>, 2015.
- 1089 Tao, F., Huang, Y., Hungate, B. A., Manzoni, S., Frey, S. D., Schmidt, M. W. I., Reichstein, M.,
1090 Carvalhais, N., Ciais, P., Jiang, L., Lehmann, J., Wang, Y. P., Houlton, B. Z., Ahrens, B.,
1091 Mishra, U., Hugelius, G., Hocking, T. D., Lu, X., Shi, Z., Viatkin, K., Vargas, R., Yigini, Y.,
1092 Omuto, C., Malik, A. A., Peralta, G., Cuevas-Corona, R., Di Paolo, L. E., Luotto, I., Liao, C.,
1093 Liang, Y. S., Saynes, V. S., Huang, X., and Luo, Y.: Microbial carbon use efficiency promotes
1094 global soil carbon storage, *Nature*, 618(7967), 981-985, <https://doi.org/10.1038/s41586-023-06042-3>, 2023.
- 1096 Thompson, C. G., Kim, R. S., Aloe, A. M., and Becker, B. J.: Extracting the Variance Inflation Factor
1097 and Other Multicollinearity Diagnostics from Typical Regression Results, *Basic and Applied
1098 Social Psychology*, 39(2), 81-90, <https://doi.org/10.1080/01973533.2016.1277529>, 2017.



- 1099 Tisdale, S. L., and Nelson, W. L.: Soil Fertility and Fertilizers, *Soil Science*, 101(4), 346,
1100 <https://doi.org/10.1097/00010694-196604000-00016>, 1966.
- 1101 Vallotton, J. D., Comeau, L. P., Goyer, C., Wagg, C., and Unc, A.: Land use is reliably predicted by
1102 soil burst respiration profiles, *Appl. Soil Ecol.*, 214, 106315, [doi.org](https://doi.org/10.1016/j.apsoil.2025.106315), 2025.
- 1103 Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasy, B., Mishra, U., Montzka, C., Nemes, A.,
1104 Pachepsky, Y. A., Padarian, J., Schaap, M. G., Tóth, B., Verhoef, A., Vanderborght, J., van der
1105 Ploeg, M. J., Weihermüller, L., Zacharias, S., Zhang, Y., and Vereecken, H.: Pedotransfer
1106 Functions in Earth System Science: Challenges and Perspectives, *Reviews of Geophysics*,
1107 55(4), 1199-1256, <https://doi.org/10.1002/2017rg000581>, 2017.
- 1108 Verchovsky, A., Anand, M., Barber, S., Sheridan, S., and Morgan, G.: A quantitative evolved gas
1109 analysis for extra-terrestrial samples, *Planetary and Space Science*, 181, 104830,
1110 <https://doi.org/10.1016/j.pss.2019.104830>, 2020.
- 1111 Wade, J., Potter, T. S., and Margenot, A. J.: Commentary: Structural equation models and causal
1112 claims in soil science and biogeochemistry – An equation-free “how to”, *Geoderma*, 466,
1113 117702, <https://doi.org/10.1016/j.geoderma.2026.117702>, 2026.
- 1114 Wadoux, A. M. C., and Molnar, C.: Beyond prediction: methods for interpreting complex models of
1115 soil variation, *Geoderma*, 422, 115953, <https://doi.org/10.1016/j.geoderma.2022.115953>, 2022.
- 1116 Wagai, R., and Mayer, L. M.: Sorptive stabilization of organic matter in soils by hydrous iron oxides,
1117 *Geochimica et Cosmochimica Acta*, 71(1), 25-35, <https://doi.org/10.1016/j.gca.2006.08.047>,
1118 2007.
- 1119 Wenyika, P., Enesi, R. O., Gorim, L. Y., and Dyck, M.: Effects of liming on soil biota and related
1120 processes in agroecosystems: a review, *Discover Soil*, 2(1), 37, [https://doi.org/10.1007/s44378-](https://doi.org/10.1007/s44378-025-00063-2)
1121 025-00063-2, 2025.
- 1122 Wiesmeier, M., Urbanski, L., Hobbey, E., Lang, B., von Lützw, M., Marin-Spiotta, E., van Wesemael,
1123 B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H. J., and Kögel-Knabner,
1124 I.: Soil organic carbon storage as a key function of soils - A review of drivers and indicators at
1125 various scales, *Geoderma*, 333, 149-162, <https://doi.org/10.1016/j.geoderma.2018.07.026>,
1126 2019.
- 1127 Wu, G., Gao, J., Li, H., Ren, F., Liang, D., and Li, X.: Shifts in plant and soil C, N, and P
1128 concentrations and C:N:P stoichiometry associated with environmental factors in alpine marshy
1129 wetlands in West China, *Catena*, 221, 106801, <https://doi.org/10.1016/j.catena.2022.106801>,
1130 2023.
- 1131 Yu, G. H., Chen, C. M., He, X. H., Zhang, X. Z., and Li, L. N.: Unexpected bulk density and
1132 microstructures response to long-term pig manure application in a Ferralic Cambisol Soil:
1133 Implications for rebuilding a healthy soil, *Soil and Tillage Research*, 203, 104668,
1134 <https://doi.org/10.1016/j.still.2020.104668>, 2020.
- 1135 Zhang, X., Li, Z., Nie, X., Huang, M., Wang, D., Xiao, H., Liu, C., Peng, H., Jiang, J., and Zeng, G.:
1136 The role of dissolved organic matter in soil organic carbon stability under water erosion,
1137 *Ecological Indicators*, 102, 724-733, <https://doi.org/10.1016/j.ecolind.2019.03.038>, 2019.
- 1138

<https://doi.org/10.5194/egusphere-2026-2849>

Preprint. Discussion started: 12 June 2026

© Author(s) 2026. CC BY 4.0 License.



1139

1140

1141

1142

1143

1144

1145

1146

1147

1148

1149