Soil organic matter diagenetic state informs boreal forest ecosystem feedbacks to climate change

Allison. N. Myers-Pigg1, a, Karl Kaiser2,3, Ronald Benner4, Susan E. Ziegler1

1 Department of Earth Sciences, Memorial University of Newfoundland, St. John’s, A1B2S7, Canada
2 Department of Marine Science, Texas A&M University, Galveston Campus, Galveston, 77553, USA
3 Department of Oceanography, Texas A&M University, College Station, 77840, USA
4 Department of Biological Sciences and School of the Earth, Ocean and Environment, University of South Carolina, Colombia, 29208, USA

Correspondence to: Susan E. Ziegler (sziegler@mun.ca) and Allison Myers-Pigg (a.myerspigg@gmail.com)

a Current affiliation: Pacific Northwest National Laboratory, Marine and Coastal Research Laboratory, 98382, U.S.A.

Abstract. The fate of soil organic carbon (SOC) in boreal forests is dependent on the integrative ecosystem response to climate change. For example, boreal forest productivity is often nitrogen (N) limited, and climate warming can enhance N cycling and primary productivity. However, the net effect of this feedback on the SOC reservoir and its longevity with climate change remains unclear due to difficulty in detecting small differences between large and variable carbon (C) fluxes needed to determine net changes in soil reservoirs. The diagenetic state of SOC, resulting from the physiochemical and biological transformation of SOC that alters the original biomolecular composition of detrital inputs to soil over time, is useful for tracing the net response of SOC at the timescales relevant to climate change not usually discernible from fluxes and stocks alone. Here, we test two hypotheses using a mesic boreal forest climate transect: (1) SOC diagenetic state is maintained across this climosequence, and (2) the maintenance of SOC diagenetic state is a consequence of coupled soil C and N cycling, signifying the role of enhanced N cycling supporting SOC inputs that maintain SOC stocks within the warmer climate forests. Shifts in nonvascular to vascular plant inputs observed in these and other boreal forests highlighted the need to carefully separate biogeochemical indicators of SOC source from those signifying diagenetic alteration. We thus evaluated and applied lignin biomarkers as a tool to assess the diagenetic alteration of SOC in these boreal forest organic soils, and directly compared the lignin diagenetic state with that of soil organic nitrogen (SON) assessed through amino acid
composition. The lignin diagenetic state remained constant across the climate transect, indicating a balance between the input and removal of lignin in these mesic boreal forests. When combined with previous knowledge of these forest ecosystems, including the diagenetic state of SON and direct measures of C fluxes and stocks, the results indicate a coupled increase in C and N cycling with climate warming that supports forest productivity and maintains SOC stocks. This balance could markedly shift as other factors begin to limit forest productivity (e.g., trace nutrients, water) with further climate change, or affect forest nutrient allocation (e.g., forest age or compositional change). Further application of the approach presented here could be used to detect the limits of this and other ecosystem-climate feedbacks, by providing a tractable and parameterizable index of lignin state across large spatial scales, necessary for ecosystem-scale parameterizations. Our observations are consistent with several lines of evidence in other biomes not limited by water availability, although the mechanisms for the maintenance of SOC stocks during climate warming appear to be ecosystem dependent. And assist in reducing the uncertainty in land-atmospheric C exchange thwarting current Earth system models for climate prediction.
1 Introduction

Terrestrial ecosystems containing globally relevant stores of carbon, such as boreal regions, are rapidly responding to climate-induced change (Soja et al., 2007). Boreal forests can act as a net carbon sink (Vanhala et al., 2016), with 50-75% of the total carbon stock stored within soils (Scharlemann et al., 2014).

The response of soil organic carbon (SOC) stocks to climate stressors in these high latitude regions is a key uncertainty in global Earth system models (Todd-Brown et al., 2013). Productivity is often N limited in temperate and boreal forests, and climate warming can enhance N cycling, primary production, soil C inputs and stores (Melillo et al., 2011; Philben et al., 2016). However, the net effect of this feedback on SOC stocks and its longevity with climate change remains unclear (Melillo et al., 2017; D’Orangeville et al., 2018) yet important to constrain for reducing uncertainty in predictive modeling efforts.

Ecosystem fluxes shape the inputs and losses of SOC stocks on seasonal to annual time scales providing relevant insights about climate feedback controls on SOC. However, detecting small differences between these large and variable fluxes is challenging and makes it difficult to determine net changes in soil reservoirs. Diagenetic signatures of soil organic matter (SOM) are useful for tracing the net response of SOM pools not usually discernable from balancing of soil inputs and losses at the decadal to century timescales relevant to climate change (Billings et al., 2012; Kane et al., 2005). Diagenesis of SOM is the physicochemical and biological transformation (e.g. leaching and decomposition, respectively) of the non-living organic matter in soil that alters the original biomolecular composition of detrital inputs to soil over time. Referred to collectively as the diagenetic state of SOM, the accrual of diagenetic alterations over decades to centuries is observable in the resultant composition of organic matter (Hedges and Prahl, 1993; Hedges and Oades, 1997), and can signify net changes in SOM stocks. For example, increases in the ratio of alkyl to o-alkyl carbon signifies losses of vascular plant-derived carbohydrate and selective retention of less bio-reactive waxes with diagenesis, and thus soil loss exceeding plant inputs (Preston et al., 2009; Baldock et al., 1997). Increases in this ratio have been found to be consistent with a decrease in SOC stocks observed with climate warming in dry continental boreal forests (Norris et al., 2011). However, the ratio of alkyl to o-alkyl C can be complicated by shifting boreal forest vegetation inputs that develop with a warming climate (Guideau et al., 2001; Kohl et al., 2018). Even if primary vegetation sources do not shift, changes in understory composition can alter the soil’s apparent diagenetic state as assessed through broad physicochemical signatures. For example, understory moss inputs can vary as a function of climate change altering the proportion of alkyl to o-alkyl C delivered to soils, preventing the use of this ratio to understand the diagenetic state of SOC across a mesic boreal forest climate transect (Kohl et al., 2018). Thus bulk SOC...
Source specific molecular biomarkers are useful for determination of carbon fate in soils (e.g. Kögel-Knabner, 2002; Rumpel et al., 2002; Otto and Simpson, 2005; Otto and Simpson, 2006a; Otto and Simpson, 2006b) and may help to overcome these challenges encountered when investigating the fate of carbon in boreal forest soils. Molecular transformations of lignin have provided insights about warming impacts on SOC storage and chemistry over timescales of months to years (e.g., Feng et al., 2008), as well as over climate shifts of years to decades represented by climate transects (e.g., Amelung et al., 1999; Guggenberger et al., 2001; Pengert et al., 2017). Lignin is used as an indicator of SOC source and processing in a variety of ecosystems (e.g., see review by Thevenot et al., 2010 for utility in soils, and Table S1 for 12 commonly used indices) because it is an important component of vascular plant material (up to 30 w%), and vascular plants are the sole source of lignin to soils (Kögel-Knabner, 2002). Thus, it is often used as a model compound in climate models (Thevenot et al., 2010; Sainte-Marie et al., 2021). Signatures of lignin phenols include common indicators of microbial degradation and physiochemical processing, such as the ratios of phenolic acids to aldehydes (e.g., vanillyl acids/ aldehydes; V/A), which increase as lignin side-chain oxidation increases (Hedges et al., 1988; See Table S1 for more detail). While certain indicators of lignin composition are confounded by variable vegetation inputs (including mosses) (Williams et al., 1998), multiple lignin phenol ratios can better account for variability due to source heterogeneity (Moingt et al., 2016). The comparison of lignin diagenesis across systems with changing source inputs requires assessment of input and soil lignin chemistry via multiple lignin phenol ratios (Simpson et al., 2008; Benner et al., 1990a, b). Thus, if the source inputs are properly constrained, the diagenetic state of lignin is a useful indicator for estimating ecosystem-level SOC responses to climate change over various timescales (Moingt et al., 2016).

Similarly, the composition of total hydrolysable amino acids has been used to trace the diagenetic alteration of organic N in sediments and soils (Dauwe et al., 1999; Menzel et al., 2015; Philben et al., 2016). Amino acids comprise about half of total soil organic nitrogen (SON), and their composition in detritus is transformed by biological and physiochemical processes leading to characteristic losses of some amino acids (e.g., glutamic acid, lysine) and enrichment or retention of others (e.g., glycine and hydroxyproline). Total hydrolysable amino acids have been used to indicate that SON cycling in mesic boreal forest soils can increase with warming (Philben et al., 2016). Given previous observations of increased temperature sensitivity of soil respiration with climate warming (Podrebarac et al., 2016), the inventories of SON in these forests might be expected to decrease with warming. However, direct measurements suggest that SOC...
stocks are maintained, despite increasing ecosystem-scale C inputs and losses in the warmer climate forests across the transect (Ziegler et al., 2017). This apparent discrepancy highlights the need for a direct assessment of shifts in SOC diagenesis with climate change in these boreal forests which, unlike observations of SON diagenesis (Philben et al., 2018b), is complicated by shifting plant inputs (Philben et al., 2018a, b; Kohl et al., 2018). Soil lignin phenol composition, combined with amino acid composition, is likely to provide a means to assess the extent to which soil C and N cycling is coupled or decoupled in these boreal forests, which is critical for addressing the hypothesis that stimulation of N cycling can offset C losses with climate warming by alleviating N limitation (Medlyn et al., 2000; Strömgen and Linder, 2002), increasing soil inputs and subsequently maintaining SOC stocks.

In this study, we test two hypotheses relevant to understanding the role of enhanced N cycling in supporting the maintenance of SOC stocks in warmer mesic boreal forests across a climate transect (i.e., a climosequence). First, we hypothesize that the SOC diagenetic state is maintained across this transect consistent with enhanced soil inputs keeping pace with increased losses in the warmer climate forests. Second, we postulate that the maintenance of SOC diagenetic state is a consequence of coupled soil C and N cycling, signifying the role of enhanced N cycling supporting productivity and SOC inputs that maintain SOC stocks within warmer climate forests. To test these hypotheses, we employ lignin biomarkers providing novel insights on SOC that are compared with previous measures of SON processing (Philben et al., 2016), SOC inventories and fluxes (Ziegler et al., 2017), and bulk and molecular-level SOM characteristics across this mesic boreal forest climate transect (Ziegler et al., 2017; Kohl et al., 2018).

Maintenance of SOC diagenetic state was tested through the evaluation and use of lignin phenol signatures depicting diagenetic processing free from the influence of shifting sources of soil inputs (e.g., moss) that vary across the transect. The measures of lignin phenol-based SOC diagenetic state were then compared directly with the SON diagenetic state with soil depth and across climate regions to assess whether soil C and N processing was coupled across the forest transect. If increased N cycling does in fact support the maintenance of SOC stocks within these forests, we would expect the ratio of the diagenetic indices of SOC to SON to remain consistent across the forest climosequence.

2 Materials and Methods

2.1 Field Sampling and Sample Preparation

Organic layers from humo-ferric podzols underlying similar stands of mature balsam fir (Abies balsamea) dominated forests located across a climosequence (Newfoundland and Labrador Boreal Ecosystem Latitudinal Transect, NL-BELT) were collected in 2011 (Table 1; Ziegler et al., 2017). Sampling focused
on three regions of the climosequence: the Eagle River, Salmon River and Grand Codroy regions which will be referred to herein as the Cold, Cool and Warm regions (Table 1). The three regions of similarly established forests and soils span approximately 5° in latitude, encompassing a 5.2°C shift in mean annual air temperature (MAT), ~430 mm yr\(^{-1}\) in mean annual precipitation (MAP) and ~180 mm yr\(^{-1}\) in potential evapotranspiration (PET) (Ziegler et al., 2017).

Therefore, the transect of study sites represents a temperature and precipitation gradient congruent with projected climate warming and increased precipitation expected in the region within the next century (Stocker et al., 2013). The use of this climate transect is intended to promote understanding of the combined impact of all ecosystem responses (microbial, plant and hydrologic change) to the warmer and wetter climate predicted for the region and over several decades to a century, rather than immediate responses to warming alone where the soil system is brought far from its equilibrium state. In doing so, we derive insights into the likely responses of these forest soils to climate change over decadal and century time scales within this region (Ziegler et al. 2017). The balsam fir forest sites are primarily underlain by forest floor moss cover in addition to few understory plants, the most common being *Cornus canadensis* and *Trientalis borealis*. Some fern cover occurs very sporadically in the Cold and Cool regions but is more common in the Warm region. The total moss cover represents the main understory difference across the study regions where lowest moss cover occurs in the Warm region (Table 1). *Pleurozium* sp. and *Hylcomium* sp. dominate the moss cover in the Cold and Cool region sites, with some additional coverage by *Ptilium* sp. and *Dicranum* sp. in the Cold region. The moss cover in the Warm region sites are dominated by *Dicranum* sp. and *Rhytidiadelphus* sp.

Each of the three regions along the climosequence contained three forest sites, while at each site soil samples were collected from triplicate plots (total of \(n=27\) plots across the climosequence; Ziegler et al., 2017; Table 1). Samples were collected by cutting out a 20 x 20 cm area of the organic layer with clean serrated knife and clippers, and carefully removing from the surface of the mineral soil surface using a clean trowel. Organic layers were manually separated into three horizons, L, F and H (equivalent to Oi, Oe and Oa, respectively, within the U.S. Soil Classification), and dried at 50°C before being ground and stored for further analysis. Plant end-member samples (forest floor mosses, fresh needle foliage, needle litterfall, roots, wood) subjected to biomarker analysis were collected across one site per region, and were separated visually into taxonomic groups (e.g., moss genus). Balsam fir needles were also collected in litterfall traps in the Spring 2011 (litterfall), as well as roots and wood separated from one half of the organic horizons collected and included as additional endmember plant sources. All end-member samples were dried at 50°C before being ground and stored for further analysis.
2.2 Lignin Analysis

Twelve lignin phenol signatures (see Table S1 for definitions and common usages) were examined across the diagenetic continuum, represented by horizon depth within the total organic layer, and in response to climate warming, represented by one site in each climate region. Samples from each of the three organic horizons (L, F and H) were collected from each of the triplicate organic horizon samples collected from three plots within one site in each of the three regions (total n=27 samples) and were analyzed for lignin phenols released via the cupric oxide oxidation (CuO) method (Hedges and Ertel, 1982), with modifications outlined in Kaiser and Benner (2011) and Louchouarn et al. (2010). Soil samples and the plant endmembers (n=27 moss samples, n=9 needle litter samples) were dried, homogenized, and weighed (~5 mg OC) into monel steel pressure vessels (Prime Focus, Inc.), along with reagents cupric oxide and ferrous ammonium sulfate and a steel ball bearing. Samples were oxidized in 2M NaOH at 155˚C for 3 hrs on a rotating rack to ensure constant sampling mixing. Samples were cooled immediately and spiked with internal standard of trans-cinnamic acid and ethyl vanillin. Sample clean-up was performed by solid phase extraction (SPE) for plant endmember materials or by liquid-liquid extraction with ethyl acetate for oxidized soil samples (Louchouarn et al., 2010; Kaiser and Benner, 2011). SPE eluates or solvent extracts were dried and re-suspended in pyridine for quantification. Samples were analyzed for the twelve common phenols (Fig. S1) via GC-MS-MS (Yan and Kaiser, 2018), utilizing a six-point calibration curve. The percentage of carbon as lignin phenols was determined via methodology outlined in Benner et al. (1990a), correcting for CuO oxidation efficiency.

2.3 NMR

NMR spectra were obtained from Kohl et al. (2018). Briefly, cross-polarization magic-angle spinning solid state $^{13}$C NMR (CP-MAS $^{13}$C-NMR) was performed on the site-level scale from pooled plot-level spatial replicates (Kohl et al., 2018). NMR end-member spectra were weighted based on composite litterfall spectra, moss spectra from Kohl et al. (2018) and Douglas fir wood spectra from the literature (Preston et al., 1998). Previously identified peaks were re-integrated here as determined by the sum of integrated peaks in defined spectral regions. Regions were defined in this study as in Baldock et al. (2004): Alkyl 0-45 ppm, Methoxy/N-Alkyl 45-60 ppm, O-alkyl 60-96 ppm, Di-O-Alkyl 95-100 ppm, Aromatic 110-145 ppm, Phenolic 145-165 ppm, Amide/Carboxyl 165-215 ppm. Carboxyl was assigned as the peak at 173 ppm, and Methoxy was assigned to the peak at 56 ppm.

2.4 Lignin Phenol Diagenetic Index (LPDI) development, application, and validation

To assess the lignin diagenetic state, we created a lignin diagenetic index, modeled after the amino acid index presented and utilized successfully in similar contexts in Dauwe et al. 1999; Menzel et al. 2015 and Philben et al. 2016. We utilized principal component analysis (PCA) as a data reduction tool on measured
lignin phenols datasets, to better track changes in the multiple indices and ratios measured. The development of the lignin diagenetic index for these forests was iterative, with the initial development of a purely informative PCA model used for a posteriori identification of potential confounding variables caused by source variability (i.e., shifts in mass relative to vascular plant inputs) when determining the diagenetic state of lignin (herein referred to as the Source PCA), and a second PCA to then assess variables best representing the diagenetic state of lignin in these soils (i.e., diagenetic PCA). Before using the PCAs, we first explored how the individual phenol ratios varied with depth and/or site, the source, and/or processing of lignin in our datasets. After assessing which variables varied with depth or site, we then performed multiple PCAs targeted at further elucidating source (site) and degradation (depth); details of the steps taken in this approach are below and can also be found in Supplemental Figure S2. Lignin phenol signatures considered in this exercise included the most common phenolic indices used in the literature: ∑6, ∑6:∑8, %C as lignin, C/V, S/V, VAd/Al, SAd/Al, FAD/CAD, diOHBA/V, % side chain alteration, P/(V+S), PON/P (Supplemental Table 1). All data was input to each PCA as a soil carbon normalized molar amount (nmol phenol mgC⁻¹) or ratio (nmol phenol nmol phenol⁻¹), for comparability during multivariate statistical analyses (Panetta and Gélinas, 2009). Values were processed prior to performing each PCA by zero centering and scaling to unit variance for each input variable. The effects of source materials on phenolic signatures were first considered for interpretation of lignin diagenesis in these forests, as they derive from vegetation with highly variable phenol content (e.g., mosses vs. vascular plants).

Therefore, we first completed the Source PCA model, using as inputs lignin phenol signatures that varied significantly with site (n=12). Input variables included the percentage of identifiable source material as woody materials, needles, or mosses contributing to the L horizon soil. ∑6, ∑6:∑8, %C as lignin, C/V, S/V, diOHBA/V, % side chain alteration, P/(V+S), PON/P (Supplemental Table 1). All data was input as a carbon normalized molar amount (nmol phenol mgC⁻¹) or ratio (nmol phenol nmol phenol⁻¹), for comparability during multivariate statistical analyses (Panetta and Gélinas, 2009). Values were processed prior to performing PCA by zero centering and scaling to unit variance for each input variable. In this model, input variables (p=12) included those variables that varied significantly with site. ... 

These included the most common phenolic indices used in the literature: ∑6, ∑6:∑8, %C as lignin, C/V, S/V, VAd/Al, SAd/Al, FAD/CAD, diOHBA/V, % side chain alteration, P/(V+S), PON/P (Supplemental Table 1). All data was input as a carbon normalized molar amount (nmol phenol mgC⁻¹) or ratio (nmol phenol nmol phenol⁻¹), for comparability during multivariate statistical analyses (Panetta and Gélinas, 2009). Values were processed prior to performing PCA by zero centering and scaling to unit variance for each input variable. In this model, input variables (p=12) included those variables that varied significantly with site. 

Based on the information from this PCA model, all source signatures that change significantly with depth (a proxy for diagenesis), but do not appear to be affected by the various source inputs present at these sites (e.g., non-vascular moss source inputs, ...
tested as % moss), and that change significantly with depth (a proxy for diagenesis) were further considered as input variables for the diagenetic index development via PCA (diagenetic PCA model). These variables also best retain original variance while reducing any effect of co-correlation effect on the overall PC loadings, and were chosen via a branch and bound algorithm (p=4; Cadima et al., 2017). The diagenetic PCA model was built using Vad/al, FAD/CAD, %C as Lignin-C, and the L, F, and H samples (total n=27 samples). The lignin phenol diagenetic index (LPDI) was then defined as the first component of the diagenetic PCA model, as it best explained variation in lignin diagenetic state across the samples, while also encompassing the highest explanatory power of a single component. Therefore, higher LPDI scores are interpreted as greater lignin diagenesis, while samples with lower scores are less altered.

To validate the LPDI, including the relationship between the measured phenols and bulk SOC diagenetic state, we created an NMR ratio to represent lignin diagenesis. This ratio is based on previous research indicating the proportion of alkyl and carboxyl carbon increases with increasing lignin diagenetic alteration, whereas the proportion of aromatics (including phenolic components) and methoxy carbon decreases with increasing lignin alteration (Zech et al., 1987). The (alkyl + carboxyl) / (aromatics + methoxy) ratio was then used to relate the alteration of lignin observed with \(^{13}\)C-NMR (n=9) to the state of lignin diagenesis obtained from the site averaged CuO derived phenol-developed LPDI (n=27). The relationship between the LPDI and \(^{13}\)C-NMR spectra was further explored to predict lignin diagenesis in samples not subjected to CuO oxidation. We predicted the LPDI for all 9 study sites along the climate transect (n = 18 predictions) based on the \(^{13}\)C-NMR spectra measured in each site across all three regions (Fig. 2).

### 2.5 Evaluating relative change in diagenetic state of SOC and SON

Lignin and amino acids represent major components of soil C and N, respectively, and the LPDI and AADI (Philben et al., 2016) indicate variability in lignin and amino acid compositions due to increasing alteration with depth. Thus, the ratios of these indices can be used to explore the relative diagenetic state of these compound classes, and to some extent SOC relative to SON as a whole. To compare lignin degradation states with that of amino acids in these soils, we used the amino acid diagenetic index (AADI) data retrieved from Philben et al. (2016). The data were retrieved from Table 1 in Philben et al. (2016), where the amino acid degradation index (AADI) is simply referred to as the degradation index (DI). Because increased diagenesis expressed via the AADI is represented in the literature as a negative loading on its PCI (Dauwe et al., 1999; Philben et al., 2016), in opposition to the LPDI used here, the sign of AADI is reversed in this representation for ease of comparison in this work.

### 2.6 Statistical Analysis

Results from these analyses across the boreal forest sites were used to test two hypotheses: (1) the SOC diagenetic state is maintained across this transect, which is consistent with enhanced soil inputs keeping...
pace with increased losses in the warmer climate forests, and (2) the maintenance of SOC diagenetic state is a consequence of coupled soil C and N cycling, signifying the role of enhanced N cycling supporting productivity and SOC inputs that cumulatively maintain SOC stocks within warmer climate forests. After verifying the LPDI was a robust indicator of diagenetic processing free from the influence of shifting sources of soil inputs (e.g. moss) that vary across the transect, we tested statistically how the LPDI varied with climate region, using analysis of variance (ANOVA) tests. The measures of lignin phenol-based SOC diagenetic state were then compared directly with the SON diagenetic state with soil depth and across climate regions to assess whether soil C and N processing was coupled across the forest transect. If increased N cycling does in fact support the maintenance of SOC stocks within these forests, we would expect the ratio of the diagenetic indices of SOC (i.e., LPDI) to SON (i.e., AADI) to remain consistent across the forest transect.

All statistics were performed in R using R Studio, and R packages ‘tidyverse’ and ‘ggfortify’ were used for data organization and visualization (Tang et al., 2016; Wickham, 2016; RStudio Team, 2016; R Core Team, 2018). A threshold of α=0.05 and β= 0.20 were applied to all ANOVA results. Two-way ANOVA was performed on all soil properties to test for the effects of organic horizon, climate region, and their interactions (n= 27, df= 2 for each). Differences between regions at each organic horizon was determined via Tukey’s Honest Significance Difference tests, when applicable.

As CuO oxidation was only conducted at one site per region, a two-way ANOVA was performed to test for the effects of organic horizon, site, and their interactions for each lignin parameter (n= 27, df= 2 for each). Thus, site level values are reported in Figure 1, and error bars on Figure 1 indicate standard deviation for all spatial replicates at the site level. To further understand if site-level trends were observable on the regional scale, the LPDI was compared to and predicted from NMR spectra obtained at the site level (Fig. 2). Statistically significant changes with depth of all tests are interpreted as changes with diagenesis, while significant changes by site only are considered to be source derived. We further tested the combined measured and predicted LPDI values to differences with region and depth (Fig. 3). Figures 3-5 contain the combined results from the measured and predicted LPDI values for all sites in all regions.

3 Results

3.1 Organic soil lignin composition

Lignin phenol signatures generally display significant variation with horizon depth, a proxy for variability due to diagenesis, and climate site, a proxy for variability due to integrated climate effects (Fig. 1). The depth profile from L to H horizon provides an increasing diagenetic state of organic matter, however, variation often observed in the deepest horizon likely reflects a combination of variation in how degraded the organic matter is, the impact of hydrology (e.g. Hernes et al., 2007) and root inputs or processes (Otto and Simpson, 2006). VAM and FAD/CAD increased only with depth, while S/V did not vary with depth.
but increased with climate warming, and S\textsubscript{Al} displayed no significant trend. Varying significantly with depth and site, the percentage of carbon as lignin (%C as lignin-C), ∑6, and ratio of ∑6 to ∑8 (∑6:∑8) all decreased with depth and were lowest in the Cold region site. The effect of site on the percentage of carbon as lignin carbon is attributed to a lower L horizon value for the Cold region site (Fig. 1e). Percent side chain alteration, diOHBA/V, and C/V all increased with depth and were generally highest in the Cold region site. Indices with p-hydroxyphenols — P/(V+S) and PON/P — were the only variables that exhibited significant variations with depth-site interactions, in addition to variations with depth and/or site and attributed to the decreasing values with depth observed in the Cold region site only.

### 3.2 Lignin Phenol Diagenetic Index (LPDI) model results

The Source PCA results indicated that input variables loaded in unique PC1 vs PC2 space with 70.6% of variance explained by these two components (Fig. S2). Loadings of variables were concurrent with how pure plant end-member samples scored, dependent on sample type (i.e., moss, woody material, and needles; Table S2). For example, the lignin phenol indices impacted by moss inputs were PON/P, diOHBA/V, and C/V, suggesting they would be the most affected by shifting proportion of moss inputs across the study sites (Table 1; Kohl et al., 2018).

The diagenetic PCA model showed that with increasing depth, the relative ratio of Ad/Al, and the ratio of cinnamyl class phenols (ferulic acid/p-coumaric acid; FAD/CAD) increased (Fig. 1) and loaded positively on PC1 (Table S4). The %C as Lignin-C decreased with depth (Fig. 1), and the ratio of carbon-carbon and/or ether-bonded phenols to ester- and ether-bonded phenols in vascular plant materials (V+S:V+S+C or ∑6:∑8; Fig. 1) loaded negatively on PC1 (Table S4). The first principal component explained 64.3% of the variance. The results are consistent with expected declines in percentages of carbon as lignin with increasing diagenesis, and they support increasing lignin diagenesis with depth in the organic soils studied (Fig. 1).

### 3.3 Comparison of LPDI with \textsuperscript{13}C NMR and assessment of LPDI across the climosequence

The LPDI was compared with the determination of soil lignin carbon diagenesis derived from \textsuperscript{13}C NMR, to validate the LPDI across multiple forest sites in each climate region, from which O-horizon NMR data are available. We find good agreement between the CuO derived phenol-developed LPDI and ratio of (alkyl + carboxyl) / (aromatics + methoxy) in our soils for all regions (n = 9, R\textsuperscript{2} = 0.898, p < 0.001; Fig. 2). Because of this strong linear relationship, we determined that in these soils we could further estimate the LPDI for all 9 study sites along the climate transect (n = 18 predictions, for a total of n=27 observations) based on the \textsuperscript{13}C-NMR spectra alone, measured in each site across all three regions (Fig. 2, predicted values are represented with dashed lines). The lignin phenol diagenetic state in these soils, assessed by the measured and predicted LPDI scores, increases with depth, yet is not significantly different by climate region (Fig. 3).
Figure 1. CuO measured lignin phenol content in soils and input materials at one site per climate region. Moss samples in the Warm and Cold regions are composite samples based on dominant species cover, while in the Cool region individual moss species are plotted as the average and standard deviation with equal weights to show variability in moss signatures in phenolic signatures present in these forests. Green, brown, and fresh balsam fir needles from the Warm and Cool regions were also measured. The ratio of vanillic acid to vanillin (VAd/Al) (A) and the ratio of ferulic acid to p-coumaric acid (FAD/CAD) (B) increase with depth, while the syringyl to vanillyl ratio (S/V) (C) is different by site. The sum of syringyl and vanillyl phenols (∑6) (D), %C as lignin-C (E), the ratio of ∑6 to the sum of all cinnamyl, syringyl and vanillyl phenols (∑6:∑8) (F), % Side Chain Alteration (G), the ratio of cinnamyl to vanillyl (C/V) (H), the ratio of dihydroxybenzoic acid to vanillyl (diOHBA/V) (I) all vary with depth and site. The ratio of p-hydroxyacetophenone to p-hydroxyl (PON/P) (J) varied with depth, site, and exhibited a depth x site interaction. The ratio of syringyl acids to aldehydes (SAd/Al) (L) shows no

| Figure 1 | CuO measured lignin phenol content in soils and input materials at one site per climate region. Moss samples in the Warm and Cold regions are composite samples based on dominant species cover, while in the Cool region individual moss species are plotted as the average and standard deviation with equal weights to show variability in moss signatures in phenolic signatures present in these forests. Green, brown, and fresh balsam fir needles from the Warm and Cool regions were also measured. The ratio of vanillic acid to vanillin (VAd/Al) (A) and the ratio of ferulic acid to p-coumaric acid (FAD/CAD) (B) increase with depth, while the syringyl to vanillyl ratio (S/V) (C) is different by site. The sum of syringyl and vanillyl phenols (∑6) (D), %C as lignin-C (E), the ratio of ∑6 to the sum of all cinnamyl, syringyl and vanillyl phenols (∑6:∑8) (F), % Side Chain Alteration (G), the ratio of cinnamyl to vanillyl (C/V) (H), the ratio of dihydroxybenzoic acid to vanillyl (diOHBA/V) (I) all vary with depth and site. The ratio of p-hydroxyacetophenone to p-hydroxyl (PON/P) (J) varied with depth, site, and exhibited a depth x site interaction. The ratio of syringyl acids to aldehydes (SAd/Al) (L) shows no

| Figure 1 | CuO measured lignin phenol content in soils and input materials at one site per climate region. Moss samples in the Warm and Cold regions are composite samples based on dominant species cover, while in the Cool region individual moss species are plotted as the average and standard deviation with equal weights to show variability in moss signatures in phenolic signatures present in these forests. Green, brown, and fresh balsam fir needles from the Warm and Cool regions were also measured. The ratio of vanillic acid to vanillin (VAd/Al) (A) and the ratio of ferulic acid to p-coumaric acid (FAD/CAD) (B) increase with depth, while the syringyl to vanillyl ratio (S/V) (C) is different by site. The sum of syringyl and vanillyl phenols (∑6) (D), %C as lignin-C (E), the ratio of ∑6 to the sum of all cinnamyl, syringyl and vanillyl phenols (∑6:∑8) (F), % Side Chain Alteration (G), the ratio of cinnamyl to vanillyl (C/V) (H), the ratio of dihydroxybenzoic acid to vanillyl (diOHBA/V) (I) all vary with depth and site. The ratio of p-hydroxyacetophenone to p-hydroxyl (PON/P) (J) varied with depth, site, and exhibited a depth x site interaction. The ratio of syringyl acids to aldehydes (SAd/Al) (L) shows no
3.4 Using combined biomarkers to assess relative degree of soil C to N cycling

The LPDI / (-AADI) is similar and near a value of one in all regions (Fig. 4). This ratio also does not change significantly with depth (Fig. 4), indicating that despite increased diagenetic state of lignin with depth (Fig. 3) the relative degree of processing of lignin and protein (i.e., amino acids) remains similar with depth and climate warming in these forest soils. The coupling between the LPDI and the AADI with depth and across climate region contrasts with the decoupling observed between the LPDI and the alkyl to O-alkyl C ratio (A/O-A), a common NMR indicator of SOC diagenetic state (Balock et al. 1997) (Fig. 5).

The greatest discrepancy between the LPDI and A/O-A is observed in Cold and Cool regions soils where moss cover (Table 1) and moss inputs (assessed via moss detritus measured in L horizon soils) are greatest.
Figure 3. The Lignin Phenol Diagenetic Index (LPDI) increases with organic layer depth with individual horizons designated (L, F, and H). The actual p-value for horizon (H), region (R) or horizon x region (HXR) significance is reported with significant values ($\alpha=0.05$) in bold. LPDI is a unitless value.

Figure 4. Ratio of the lignin phenol diagenetic index (LPDI) to the inverse sign of the total hydrolysable amino acid diagenetic index (~AADI) with depth in all regions averaged. Ratios do not change significantly among regions at $\alpha=0.05$, indicating that the relative rates of lignin and amino acid diagenesis are similar across all regions and with depth. Post-hoc tests revealed no significant difference between organic layer depth given as individual horizons (L, F and H) and/or site at the regional level. This ratio is unitless.
Figure 5. Lignin diagenetic state as observed through the lignin phenol diagenetic index (LPDI) is decoupled from SOC diagenetic state as assessed through the ratio of Alkyl-C to O-Alkyl-C (Alkyl-C/O-Alkyl-C) across climate region. Symbols signify the different horizons within the organic layer (L, F and H). This discrepancy appears to be related to the percentage of identifiable moss detritus in the L horizons (depicted with color scaling). Values are given as the site average in each region with error bars representing the standard deviation (n=3). The ratios presented in this figure are unitless.

Discussion

The maintenance of SOC stocks despite increased soil C losses and inputs with climate warming in these mesic boreal forests suggests inputs are keeping pace with losses (Ziegler et al., 2017). This may be a result of increased soil inputs supported by enhanced N cycling and productivity (Melillo et al., 2011; Philben et al., 2016). However, the annual variability and difficulty in capturing all ecosystem fluxes prevent us from being able to assess the net response of SOC stocks using ecosystem fluxes, and thus link their maintenance to N cycling and availability in these forests. Therefore, the diagenetic state of SOM assessed via lignin phenol and amino acid composition was employed here to overcome these challenges, and assess the net effect of these processes on both SOC and SON reservoirs. For example, the diagenetic state of SOC increases with increasing net loss from soil stocks (Kane et al., 2005; Ouideau et al., 2001), while an absence of change in diagenetic state signifies no net loss and thus soil C inputs keeping pace with losses. By coupling previous information about these forests with additional measures of SOM diagenetic state across the climate transect, we support two hypotheses providing evidence for the role of enhanced N cycling in maintaining SOC stocks in the warmer climate boreal forests along this transect. First, we found...
that the diagenetic state of SOC is maintained across climate regions consistent with soil inputs keeping pace with increased losses in the warmer climate forests. Second, we observed a constant ratio of the diagenetic state of lignin to that of amino acids in the organic horizon soils across these forest climate regions indicating that the maintenance of the SOC diagenetic state occurs alongside that of the SON. Therefore, the maintenance of SOC stocks across climate regions appears to be largely supported by the enhanced N cycling and availability that occurs within the warmer climate forests.

4.1 Evaluation of lignin phenol signatures of SOC diagenetic state in mesic boreal forests

Shifts in nonvascular to vascular plant inputs with climate observed in these and other boreal forests (Abolin, 1974; Kohl et al., 2018; Tamm, 1953; Tarkhova and Ipatov, 1975) meant that we had to carefully separate biogeochemical indicators of SOC source from those signifying diagenetic alteration. Though these shifts in nonvascular moss inputs have little impact on total hydrolyzable amino acids (THAA) and their use in tracking SON diagenetic state (Philben et al., 2018b), they do impact some common SOC chemical indicators. For example, the increase in the alkyl to o-alkyl ratio observed with climate warming is consistent with the decreasing moss contribution to soils with warming, and is therefore not indicative of an increase in SOC diagenetic state in the warmer, wetter forests (Kohl et al., 2018). This trend may be due to a slow turnover of structural carbohydrates within moss cell walls (Hobbie et al., 2000; Philben et al., 2018b; Turetsky et al., 2008), thereby affecting the utility of this ratio to trace SOC diagenetic shifts in these systems. However, the lignin phenol diagenetic index (LPDI) developed and applied here was designed to enable the assessment of SOC diagenetic state and the relative differences in soil carbon balance among these forests despite the shifts in nonvascular to vascular plant inputs along this forest climate transect.

The amount of lignin in these soils (as %C as lignin-C) decreases with organic horizon depth at all sites, a relationship which has also been observed in other soil systems (e.g., Guggenberger et al., 1995; Otto and Simpson, 2006b). The same tree species (balsam fir) is dominant across the forests of this transect, but we observed that commonly used lignin parameters, such as source-sensitive ratios (C/V and S/V), are affected by variable vascular to nonvascular understory inputs to these forest soils. These ratios are commonly used in soils to assess diagenetic state (e.g., Otto and Simpson, 2006), yet they are not valid to assess diagenesis of organic C in these soils due to the over-printing of source-related shifts in these forests. The LPDI PCA representation simplifies and integrates the description of observed variability of lignin phenol diagenetic state from multiple indices, allowing for ease of comparability within a given dataset and adaptation for development and implementation in other environments. Testing the impact of mosses on the LPDI index revealed no significant difference in the LPDI by region, moss input, or soil depth (see Table S6). We observed minimal moss impacts on the LPDI, which could be useful for interpretation of phenolic composition in areas with variable moss inputs to SOM, a common scenario with varying climate across...
boreal forest ecosystems (Abolin 1974; Tamminen 1953; Tarkhova and Ipatov, 1975). The combination of the lignin phenol biomarker and NMR approaches (Simpson et al., 2008; Benner et al., 1990a, b) with the measured and predicted LPDI scores indicates this approach provides a robust signature of SOC diagenetic state.

4.2 Consistent lignin diagenetic state indicates a balance between inputs and losses within warmer climate forests.

The lack of change in lignin diagenetic state across these boreal forest sites, despite the +5.2°C MAT range across the transect, contrasts with the increase in the diagenetic state of lignin observed over 14 months of experimental warming in a temperate forest (Feng et al., 2008). This is likely due to the nonequilibrium conditions associated with short-term warming including a lack of additional ecosystem responses to warming such as enhanced soil inputs (Melillo et al., 2011) that are not captured over the shorter time scales associated with most in situ warming experiments. Impacts of a warmer climate on ecosystem properties, such as altered litter inputs (Pisani et al., 2016) and shifts in climate conditions such as MAP (Duboc et al., 2014; Pisani et al., 2014), can serve as drivers of lignin decomposition and its diagenetic state. This likely explains the lack of clear trends in lignin diagenetic state observed with increasing MAT on a continental scale (Amelung et al., 1999; Pisani et al., 2014); including a decrease in the degradative state of lignin across sites spanning ~2°C MAT (Otto and Simpson, 2006b). Losses can be diminished with decreasing water availability associated with increasing MAT, and thus exhibit a decreased diagenetic state of SOC. In the mesic boreal forest climosequence we studied, MAP increases alongside MAT and PET, with decreasing latitude representing a scenario where water limitations on enhanced productivity with warming are not likely (d’Orangeville et al., 2016; Ziegler et al., 2017). Thus, the maintenance of soil lignin diagenetic state across the climate gradient indicates increased inputs are keeping pace with losses over decadal time scales. This is consistent with experimental evidence of a similar balance between input and processing of lignin in temperate forest soils where the maintenance of organic carbon content and lignin composition was observed despite increased litter inputs experimentally added over several decades (Lajtha et al., 2014; Pisani et al., 2016).

4.3 Coupled SOC and SON cycling indicate enhanced N availability supports increased productivity that maintains SOC within warmer climate forests.

Increases in nitrogen availability and litter inputs (Philben et al., 2016; Kohl et al., 2018) in these forest soils appear to sustain the balance between input and processing losses in the warmer climate forests. This is supported by the coupled diagenesis of SOC and SON as revealed here through the comparison of lignin and amino acid diagenetic indices (i.e., the LPDI/ADI). The similarity of diagenetic state observed in the soils across this transect is supported by multiple independent lines of evidence. Increased N availability,
observed as an increase in soil fluxes (Ziegler et al., 2017; Kohl et al., 2018), and maintenance of amino acid diagenetic state (Philben et al., 2016), occurs with increased inputs and losses of C from soils and the lack of change in SOC stocks from the coldest to warmest forest (Ziegler et al., 2017). The lack of change in SOC diagenetic state observed in this study provides previously unseen evidence required for understanding the maintenance of soil C stocks observed, supported by increased soil organic nitrogen cycling (Philben et al., 2016, 2018a), enhanced forest productivity (Table S5), and thereby increased litter inputs to the soil in the warmest relative to coldest climate forests. The SOC and SON diagenetic indicators coupled with flux and stock assessments provided the net result of plant-soil interactions, which gradually respond to climate change and are not easily observed in experiments (Melillo et al., 2011, 2017) or by monitoring of ecosystem fluxes alone.

The sustained SOC stocks observed across this boreal forest climate transect contrasts with SOC dynamics in more water-limited boreal forests in western North America (Norris et al., 2011; Kane et al., 2005), where SOC stocks decline within warmer climate forests. This highlights the heterogeneity of boreal forest responses to warmer climates. The similarity of diagenetic state among major compound classes with depth or climate would not necessarily hold true in soils under non-steady state conditions resulting from responses to environmental change. Soil pools acclimating to increased temperature, for example, display different periods of soil carbon retention and loss over multiple decades of experimental warming (Melillo et al., 2017). On decadal-centennial timescales assessed via climosequences, however, the net result of plant-soil interactions on soil fluxes are more likely adjusted to such differences in climate which likely represent a new steady state. The conservation of soil carbon stocks via increased forest productivity observed in this study region is consistent with the continuity of forest productivity expected in this region, as revealed in decadal tree ring records (D’Orangeville et al., 2016; Charney et al., 2016). These are also consistent with observations of increasing SOC stocks in mineral soils with increasing MAT and MAP and associated net primary production (NPP) across other boreal forests (Callesen et al., 2003). In the forests studied here, surface mineral SOC stocks appear similar, and are likely maintained by the 3-fold increase in dissolved organic carbon (DOC) inputs with climate warming across this transect (Ziegler et al., 2017; Bowering et al., 2022). However, the trajectories of the deeper mineral SOC stocks remain unknown and rely on hydrology and parent material sourcing of reactive metals across these forests (Patrick et al., 2022).

The comparative approach using lignin and amino acid diagenetic states has the potential to be applied over time in these forests and elsewhere to assess the limitations of feedbacks, such as warming enhanced N availability and CO2 productivity, in maintaining ecosystem SOC stocks. For example, if N is not tightly recycled within the terrestrial ecosystem, as is currently observed in our study forests (Philben et al., 2018a), but is instead experiencing net lost via nitrification or increased dissolved N transport, the diagenetic state of the remaining SON would increase relative to that of SOC if those losses are not maintained by additional external inputs (e.g. atmospheric inputs, N-fixation or weathering; Houlton et al.,...
2018). High SOC mineralization rates and low SON mineralization rates, such as those observed in several high-latitude environments (Meyer et al., 2006), would also affect the C balance and be detected as an increase in diagenetic state of SOC decoupled from that of SON, where the diagenetic state would remain relatively unaltered. This in turn would signify a reduction in SOC stocks, similar to what has been observed in response to artificial N fertilization in tundra soils, where losses of SOC were enhanced relative to plant productivity and soil inputs (Mack et al., 2004). Therefore, tracking the relative diagenetic states of SOC and SON may be useful in detecting the impact of declining responses of productivity to increased CO2 (Norby et al., 2010), or water and phosphorus limits on the warming enhanced N availability and productivity on SOC stocks in these ecosystems.

4.4 Conclusion

Past modeling studies have called for improvements to the accuracy with which C and N cycles and their feedbacks are simulated (Thomas et al., 2013). Thus, better observational constraints on C and N cycling and their response to climate change (Meyerholt et al., 2020), such as demonstrated here, are needed. Our observation of maintained SOC stocks across this climate transect of boreal forests supports our first hypothesis and is consistent with stocks in other biomes (Giardina et al., 2014; Sistla et al., 2013) not limited by water availability, although the mechanisms for the maintenance of SOC in response to warmer climate are seemingly ecosystem-dependent. Supporting our second hypothesis, these forests exhibit a coupled increase in biogeochemical cycling of N and C signifying a balance between soil input and loss processes resulting from increased N availability and resulting productivity within the warmer relative to colder climate forests. This balance could markedly shift as other factors begin to limit forest productivity (e.g., trace nutrients, water) with further climate change, or affect forest nutrient allocation (e.g., forest age or compositional change). Further application of the approach presented here could be used to detect the limits of ecosystem-climate feedbacks. Utilization of these coupled diagenetic approaches in modeling studies to represent complex C and N dynamics in ecosystem-scale parameterizations could additionally assist in reducing the large uncertainty in land-atmospheric carbon exchange thwarting current Earth system models for climate prediction.

Data availability. All data not included in the paper in tables and the Supplement are available on figshare at: https://figshare.com/s/bd211f6b1d2d5f4b7ea

Supplement can be found at this link.

Author contributions. AM-P and SEZ designed the study with input from KK and RB. Site identification, design and sampling was led by SEZ and supported by collaborators acknowledged below. AM-P
conducted sample analyses with significant input from KK. AM-P conducted data analyses and construction of figures and tables with input from SEZ and further input from KK and RB on a more finalized version. AM-P prepared the paper with significant input from SEZ and editing from KK and RB.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgments.** We thank Jerome Laganière, Kate Edwards, Andrea Skinner, Darrell Harris, Jamie Warren and Keri Bowering for the essential role they had in sampling campaigns and soil processing within the NL-BELT project. Jamie Warren (MUN), Celine Schneider (MUN), Annie Tamalavage (TAMU), Stephanie Mohan (TAMUG) and Ge Yan (TAMUG) are thanked for their laboratory assistance. Darrell Harris and Andrea Skinner are thanked for procuring the tree productivity dataset. Kate Buckeridge is thanked for information regarding moss coverage and dominant moss species at these sites. Michael Philben and Lukas Kohl are thanked for access to their datasets for amino acids and $^{13}$C-NMR in these organic soils, respectively. We thank Sharon Billings, J. Alan Roebuck, Keri Bowering and the BBERG lab for helpful comments on previous versions of this manuscript. We thank Teri O’Meara and Ben Bond-Lamberty for useful discussions regarding the relevance of this manuscript to Earth System Modeling. Funding for this research was provided by Natural Sciences and Engineering Research Council of Canada award number SPG#79224-15, RGPIN#341863, Canadian Forest Service, Center for Forestry Science and Innovation of the Newfoundland Forestry and Agrifoods Agency, and the Canada Research Chairs program.
References


Barré, P., Durand, H., Chenu, C., Meunier, P., Montagne, D., Castel, G., Billiou, D., Soucemarianadin, L., and Cecilon, L.: Geological control of soil organic carbon and nitrogen stocks at the landscape scale,


R Core Team: R: A language and environment for statistical computing, 2018.


Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and


