Review article: A systematic review of terrestrial dissolved organic carbon in northern permafrost

Liam Heffernan¹, Dolly N. Kothawala¹, Lars J. Tranvik¹

¹Limnology/Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, Uppsala 75236, Sweden

Correspondence email: liam.heffernan@ebc.uu.se
Abstract

As the permafrost region warms and permafrost soils thaw, vast pools of soil organic carbon (C) become vulnerable to enhanced microbial decomposition and lateral transport into aquatic ecosystems as dissolved organic carbon (DOC). The mobilization of permafrost soil C can drastically alter the net northern permafrost C budget. DOC entering aquatic ecosystems becomes biological available for degradation as well as other types of aquatic processing. However, it currently remains unclear which landscape characteristics are most relevant to consider in terms of predicting DOC concentrations entering aquatic systems from permafrost regions. Here, we conducted a systematic review of 111 studies relating to, or including, concentrations of DOC in terrestrial permafrost ecosystems in the northern circumpolar region published between 2000 – 2022. We present a new permafrost DOC dataset consisting of 2,276 DOC concentrations, collected from the top 3 m in permafrost soils across the northern circumpolar region. Concentrations of DOC ranged from 0.1 – 500 mg L\(^{-1}\) (median = 41 mg L\(^{-1}\)) across all permafrost zones, ecoregions, soil types, and thermal horizons. DOC concentrations were greatest in the sporadic permafrost zone (101 mg L\(^{-1}\)) while lower concentrations were found in the discontinuous (60 mg L\(^{-1}\)) and continuous (59 mg L\(^{-1}\)) permafrost zones. The highest median DOC concentrations of 66 mg L\(^{-1}\) and 63 mg L\(^{-1}\) were found in coastal tundra and permafrost bog ecosystems, respectively. Coastal tundra (130 mg L\(^{-1}\)), permafrost bogs (78 mg L\(^{-1}\)), and permafrost wetlands (57 mg L\(^{-1}\)) had the highest DOC concentrations in the permafrost lens, representing a potentially long-term store of DOC. Other than in Yedoma ecosystems, DOC concentrations were found to increase following permafrost thaw and were highly constrained by total dissolved nitrogen concentrations. This systematic review highlights how DOC concentrations differ between organic- or mineral-rich deposits across the circumpolar permafrost region and identifies coastal tundra regions as areas of potentially important DOC mobilization. The quantity of permafrost-derived DOC exported laterally to aquatic ecosystems is an important step for predicting its vulnerability to decomposition.
1. Introduction

Persistent freezing temperatures since the late Pleistocene and Holocene has led to the accumulation and preservation of 1,460 – 1,600 Pg of organic carbon (C) in northern circumpolar permafrost soils (Hugelius et al., 2014; Schuur et al., 2018). However, in recent decades, there has been an amplified level of warming at high latitudes, occurring at four-times the speed of the global average (Rantanen et al., 2021). This is leading to widespread and rapid permafrost thawing. Under the high C emissions representative concentration pathway (RCP8.5), 90% loss of near-surface permafrost is projected to occur by 2300, with the majority of loss occurring by 2100 (McGuire et al., 2018). Increasing temperatures and widespread thaw exposes permafrost C to heterotrophic decomposition, potentially leading to enhanced emissions of greenhouse gases to the atmosphere in the form of carbon dioxide (CO$_2$; Schuur et al., 2021) and methane (CH$_4$; Turetsky et al., 2020). Alternatively, previously frozen soil organic carbon may be mobilized into the aquatic network as dissolved organic carbon (DOC), the quantity and quality of which will likely depend on local and regional hydrology, and landscape characteristics (Tank et al., 2012; Vonk et al., 2015). At high latitudes (>50°N), lakes and rivers of various sizes cover 5.6% and 0.47% of the total area, respectively (Olefeldt et al., 2021), and the landscape C balance at these high latitudes is highly dependent on aquatic C processing (Vonk & Gustafsson, 2013). The increased leaching of recently thawed DOC from permafrost soils will not only increase the currently estimated 25 – 36 Tg DOC year$^{-1}$ exported into the freshwater system, and subsequently into the Arctic Ocean (Holmes et al., 2012; Raymond et al., 2007), but will also likely lead to enhanced greenhouse gas emissions from freshwater ecosystems (Dean et al., 2020). However, uncertainty remains as to which terrestrial ecosystems are likely to contribute the highest concentrations of laterally transported permafrost DOC and of this, which is expected to contribute the DOC most vulnerable to mineralization.

The contribution of mineralized permafrost C to atmospheric CO$_2$ and CH$_4$ balances, known as the permafrost C feedback (Schaefer et al., 2014), remains poorly constrained due to uncertainty of the magnitude and location of permafrost C emissions (Miner et al., 2022). The lateral transport of DOC represents a source of terrestrial C that can potentially play an important role in both terrestrial and aquatic biogeochemical cycles and is thus an important fraction of the permafrost C feedback. Warming of near surface permafrost causes widespread thawing (Camill,
2005; Jorgenson et al., 2006), which can lead to drastic changes in hydrology, vegetation, and soil carbon dynamics (Liljedahl et al., 2016; Pries et al., 2012; Varner et al., 2022). When permafrost is present, the lateral transport of DOC is restricted to flow paths within the unfrozen, organic rich active layer (Woo, 1986). Deeping of the seasonally thawed active layer due to top-down permafrost thaw can lead to longer flow paths for DOC, allowing for enhanced decomposition or adsorption to mineral particles, resulting in reduced DOC export (Kicklighter et al., 2013; Striegl et al., 2005). Alternatively, thermokarst formation can affect the entire soil profile, leading to surface inundation, and shifting ecological conditions and vegetation communities associated with greater DOC production (Turetsky et al., 2007). This can cause greater hydrological connectivity, resulting in increased runoff in permafrost peatlands (Connon et al., 2014) or increased connectivity to regional hydrology through thermo-erosion gullies or thaw slumps (Kokelj & Jorgenson, 2013) in tundra ecosystems. Permafrost landscape dynamics, including the mode of permafrost thaw and ecological conditions present following thaw, will play a key role in the biogeochemical and ecohydrological processes that constrain DOC mobilization, i.e., export and mineralization upon export. The freshwater DOC pool represents a mix of C derived from a variety of ecosystem types and sources, and the ecological conditions of each source will have a significant impact on the quantity and quality of this mobilized DOC. Determining the relative contribution and impact on mineralization of these DOC sources represents a potentially important step in reducing uncertainty in the permafrost climate feedback.

Here, we conduct a systematic review and compiled 111 studies published between 2000 – 2022 on DOC concentrations in the top 3 m of terrestrial ecosystems found in the northern circumpolar permafrost region. A quantitative assessment of studies pertaining to DOC concentrations in permafrost soils can identify evidence-based recommendations for future topics and areas of research to improve our understanding on terrestrial and aquatic biogeochemical cycling in northern permafrost regions. Our database contains ancillary data describing the geographical and ecological conditions associated with each DOC concentration, allowing us to reveal patterns in DOC concentrations and lability measures for 562 sampling sites across multiple ecosystem types and under varying disturbance regimes. This study represents the first systematic review of DOC concentrations within terrestrial permafrost ecosystems found in the
circumpolar north. As such, it provides unique and valuable insights into identifying ecoregions, or landscape characteristics, associated with the highest DOC concentrations, and thus regions with the greatest potential for DOC mobilization. Mobilization rates represent DOC loss and include specific discharge of DOC (g DOC m$^{-2}$), export rate of DOC per day (g C m$^{-2}$ day$^{-1}$) and per year (g C m$^{-2}$ year$^{-1}$), and biodegradable DOC (BDOC; %). We hypothesized that (i) the highest DOC concentrations would be found in organic rich wetland ecosystems, (ii) disturbance would lead to increased export and biodegradability of DOC, and (iii) the most biodegradable DOC would be found in Yedoma and tundra ecosystems.

2. Methods
This systematic review used a methodological framework proposed by Arksey & O’Malley (2005) and follows five steps: 1) develop research questions and a search query; 2) identify relevant studies; 3) study selection; 4) data extraction; and 5) data analysis, summary, and reporting. The literature search was guided by four research questions: 1) what are the concentrations of DOC found in terrestrial ecosystems across the northern circumpolar permafrost region?; 2) what are the rates of export and/or degradation (mobilization) of DOC within these ecosystems?; 3) What are the major controls on DOC concentrations and rates of mobilization?; and 4) how are concentrations and mobilization rates impacted by thermokarst formation?

2.1 Literature Search
Based on a priori tests, we used the following search query string to find papers using information found in their title, abstract, and keywords: (“dissolved organic carbon”) AND (permafrost OR thermokarst OR “thaw slump”) AND (soil OR peat) AND (export OR degrad* OR decomposition OR mineralization). We used Web of Science, Science Direct, Scopus, PubMed, and Google Scholar to generate a database of tier 1, peer-reviewed articles published between 2000 – 2022. The search function on Science Direct does not support the use of wildcards such as “*”, so “degrad*” was changed to “degradation”. We removed duplicate references found across multiple databases using Mendeley© referencing software (v1.17.1, Mendeley Ltd. 2016). Once this initial database was compiled, we used the same search query string as above to search for additional articles on the first 15 pages of Google Scholar.
resulted in the addition of a further 150 articles to be included in our systematic screening process.

### 2.2 Systematic Screening of Peer-Reviewed Publications

The selection of relevant studies was comprised of inclusion criteria and relevance screening in three steps. In the first step we placed limits on initial study searches in the electronic databases mentioned above. Studies were included in the review if they were primary research, published in English, and published between 2000 – 2022 (Table 1). Only quantitative studies conducted in terrestrial ecosystems within the northern circumpolar permafrost region, as defined by Brown et al., (1997), and reporting DOC concentration and mobilization rates were included. Studies not meeting these criteria were eliminated and the remaining studies proceeded to the second screening step.

#### Table 1. Summary of criteria used to identify suitable studies in the preliminary screening stage

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timeline</strong></td>
<td>Study published between 2000 – 2022</td>
</tr>
<tr>
<td><strong>Study type</strong></td>
<td>Study published prior to 2000</td>
</tr>
<tr>
<td><strong>Study type</strong></td>
<td>Primary research article published in peer-reviewed journal using quantitative methods</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>Thesis/dissertations and secondary research studies (reviews, commentaries, editorials)</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td>Conducted within the northern circumpolar permafrost region</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td>Conducted outside of the northern circumpolar permafrost region</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>Studies on DOC concentration, export or degradation in permafrost environments</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>Studies not on DOC concentration, export or degradation in permafrost environments</td>
</tr>
</tbody>
</table>

In the second step, the primary relevance of articles was screened, based on article titles, abstracts, and keywords, and the eligibility criteria provided in Table 2. Studies deemed irrelevant were eliminated and the remaining studies proceeded to the third and final screening...
step, or secondary screening stage, which was based on more specific eligibility criteria (Table 2) applied to the full text.

Table 2. Primary and secondary relevance screening tools. Primary screening tool used in the article title, abstract, and keyword screening stage. Secondary screening tool used in full-text screening stage

<table>
<thead>
<tr>
<th>Screening stage</th>
<th>Screening questions</th>
<th>Response details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Does the study involve quantitative data collected from a permafrost environment?</td>
<td>Yes – reports on quantitative data collected from a permafrost environment</td>
</tr>
<tr>
<td>Primary and Secondary</td>
<td>Is the study region within the northern circumpolar permafrost region?</td>
<td>Yes – reports on quantitative data (including field observations and lab data) collected from the circumpolar permafrost environment.</td>
</tr>
<tr>
<td>Primary and Secondary</td>
<td>Is the article in English and NOT a review, book chapter, commentary, correspondence, letter, editorial, case report, or reflection?</td>
<td>Yes – study is in English and is a primary research article that includes quantitative studies (field and lab based), including model-based research as it relies on observational data.*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No – study is not in English and/or is a review, book, editorial, working paper, commentary, conference proceeding, supplementary text, or qualitative study which does not address outcomes relevant to this review</td>
</tr>
<tr>
<td>Primary and Secondary</td>
<td>Does the study involve the concentration, export or degradation of terrestrially derived DOC?</td>
<td>Yes – reports on terrestrial DOC concentration, export, or degradation, including concentrations and characterization</td>
</tr>
<tr>
<td>Secondary</td>
<td>Is the article in English, longer than 500 words, and published between 2000 - 2022?</td>
<td>Yes – study is published between 2000 – 2022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No – study is published prior to 2000</td>
</tr>
</tbody>
</table>

*For model-based studies, the original field/lab data used to parametrise or develop the model was used. If this data was taken from previously published work, then those studies were used and the model-based study removed.
2.3 Database compilation

A database with reported DOC concentrations and mobilization rates i.e., rates of either DOC export or degradation, was compiled using data from all studies that were deemed relevant following the study selection phase. The database was compiled to compare DOC concentrations and mobilization rates between different sites. We define a site as an area where either soil, water, or ice samples were taken from that has similar vegetation composition, water table position, permafrost regime, and was either disturbed or pristine. Site descriptions were derived from the text of each study. Where possible, individual daily measurements of DOC concentrations and mobilization rates were taken. When replicates of the same daily measurement were provided, we used the mean of those replicates, which was relevant for 10 studies within the database, representing 72 DOC concentrations. All data was extracted from data tables, text, supplementary material, or extracted from data figures using WebPlotDigitizer (https://automeris.io/WebPlotDigitizer).

All studies reported measuring DOC concentrations collected from either open-water, pore water, ice, or soil using a median filter pore size of 0.45 μm with first and third quartiles pore size of 0.45 and 0.7 μm. Measurements from all 12 months of the year were included in the database with the majority occurring during the growing season (May – August), a small portion during the non-growing season, and the remaining sampling times were either not reported or are averages over multiple sampling occasions. We included data from studies that were both field and lab based. However, any data where a treatment was applied was excluded, except for temperature treatments during incubation experiments when assessing the biodegradability of DOC. When lab-based studies included an incubation, only Day 0 DOC concentrations were used when comparing DOC concentrations across studies. We chose to remove any DOC concentrations from samples taken below 3 m depth, which represented 3% of all DOC measurements. These measurements were removed for better comparability with the current best estimation of soil organic carbon stocks within the northern circumpolar permafrost zone (Hugelius et al., 2014). We also removed any DOC concentrations greater than 500 mg L⁻¹, which represented 2% of all DOC concentrations. Samples that were above 500 mg L⁻¹ and were sampled below 3 m represented 1% of all DOC concentrations.
Site averaged daily DOC concentrations (mg L$^{-1}$) and mobilization rates were estimated from the average concentration and mobilization rates measured within a single day or sampling occasion. Repeated measurements at a site, either over the growing season or multiyear measurements, were treated as an individual estimate of DOC concentrations and mobilization rates. Other continuous variables that were similarly estimated include soil moisture, water table position, organic layer depth, active layer depth, bulk density of soil, soil carbon content (%), soil nitrogen content (%), carbon:nitrogen, pH, electrical conductivity ($\mu$S cm$^{-1}$), specific UV absorbance at 254 nm (SUVA; L mg C$^{-1}$ m$^{-1}$), total dissolved nitrogen (mg L$^{-1}$), nitrate (mg L$^{-1}$), ammonium (mg L$^{-1}$), chloride (mg L$^{-1}$), calcium (mg L$^{-1}$), and magnesium (mg L$^{-1}$). Mean annual temperatures and precipitation, sampling depth, filter size, the number of days over which sampling took place, how many years following disturbance measurements were taken were also recorded. Several continuous variables other than those mentioned above were also recorded in the database, but not used for analysis if they represented < 20% of the database. We chose 20% as the cut-off point for use in comparison of the relationship between DOC concentrations and mobilization with other site continuous variables.

Categorical variables included in the database were site location within the permafrost zone (continuous, discontinuous, sporadic; Brown et al., 1997) and ecoregion (arctic tundra, sub-arctic tundra, sub-arctic boreal, and continental boreal; (Olson et al., 2001). We included site surface permafrost conditions (present or absent), the thermal horizon layer sampled (active layer, permafrost, permafrost free, water, and thaw stream), and if present what type of disturbance occurred at the site (fire, active layer thickening, thermokarst terrestrial, or thermokarst aquatic). We also included the soil class found at the site (Histel, Histosol, Orthel, and Turbel; USDA, 1999) and whether the DOC was from the organic or mineral soil. To assess the influence of sampling approach and method of analysis, we included method of DOC extraction (centrifugation of soil sample, leaching of soil, dialysis, grab sample, ice core extraction, potassium sulphate extraction, lysimeter, piezometer, pump, rhizons) and DOC measurement method (combustion, persulphate, photometric, or solid-phase extraction).

Sites were classified according to ecosystem type, and these included coastal tundra, forest, peatland, permafrost bog, permafrost wetland, retrogressive thaw slump, upland tundra, and Yedoma. Ecosystem classification is based on the general site description in the article, the
provided ecosystem classification within the article, and site data including vegetation composition, permafrost conditions, and ecoregion. Yedoma sites include pristine forest, upland tundra, and coastal tundra, as well as retrogressive thaw slumps and other thermokarst features found within the Yedoma permafrost domain (Strauss et al., 2021). The ecosystem classification retrogressive thaw slump only includes these thermokarst features found outside the Yedoma permafrost domain. Each ecosystem type was further classified based on the type of permafrost thaw or thermokarst formation that occurred there. These thaw or thermokarst types included thermokarst bog, thermokarst wetland, active layer thickening, retrogressive thaw slump, exposure, thermo-erosion gully, and active layer detachment.

2.4 Database analysis

All statistical analyses were carried out in R (Version 3.4.4, R Core Team, 2015). We used Kruskal-Wallis analysis to test for differences in median DOC concentrations among various categorical variables such as permafrost zones, ecoregions, soil class, thermal horizon, and ecosystems. Post-hoc comparisons of median DOC concentrations among these categories were performed using pairwise Wilcox test. We used ANOVAs and Bonferroni post-hoc tests on significant differences in DOC concentrations between methods of DOC extraction and measurement. For regression analysis, data was transformed using a Box Cox transformation and the optimal λ using the MASS package (Ripley et al., 2019). We used analysis of covariance (ANCOVA) to test for differences in DOC concentrations in different thermal horizons (i.e., active layer and intact permafrost lens) between ecosystem types, while controlling for the month in which sampling occurred. Permafrost lens DOC concentrations are determined from soil and pore water within the permafrost layer and extracted via frozen cores, whereas active layer samples are taken from soil cores or porewater that are unfrozen at the time of sampling. We used partial least squares regression (PLS) to assess the performance of continuous and categorical variables in predicting DOC concentrations. Predictor variables were categorized based on their Variable Importance in Projections (VIP) method in the plsVarSel package (Mehmood et al., 2012), whereby variables with a score > 1 are deemed to be significant. PLS was performed using the pls package (Mevik & Wehrens, 2007) and we chose to use PLS as it is tolerant of co-correlation of predictor variable, deviations from normality, and missing values, all...
of which were found within the database. In the PLS ecosystem classes were subdivided into
pristine or disturbed (i.e., impacted by permafrost thaw). Pristine sites were further subdivided
by the thermal horizon in which the DOC concentrations were measured (active layer and
permafrost lens). To evaluate the change in ecosystem DOC concentrations following
thermokarst formation, based on all studies from the systematic review, we calculated the
response ratio using the SingleCaseES package (Pustejovsky et al., 2021). We define thermokarst
as the process by which ice-rich permafrost deposits undergo complete thaw, resulting in surface
subsidence and the formation of a new, thermokarst feature that is ecological different regarding
water table position, redox conditions, and vegetation type, from the preceding pristine
ecosystem. Very few studies in our database report DOC concentrations for both pristine and
thermokarst affected ecosystem (< 20 %). To include as much data as possible we chose an
effect size metric that is unlikely to be influenced by studies with large sample number and
variance. The response ratio is:

\[ \text{Pristine to Thermokarst Effect Response ratio} = \ln\left(\frac{X_P}{X_T}\right) \]  
Eqn. 1

where \( X_P \) = mean DOC concentration of pristine ecosystems and \( X_T \) = mean DOC concentration of
thermokarst effected ecosystems (Lajeunesse, 2011). This represents the log proportional
difference in mean DOC concentrations between thermokarst and pristine ecosystems, where a
positive response ratio indicates a decrease in DOC concentrations following thermokarst. The
distribution of the data was inspected visually and with the Shapiro–Wilk test. We tested
homogeneity of variances using the car package and Levene’s test (Fox and Weisberg, 2011).
We report uncertainty using the interquartile range (lower, median, and upper quartiles), except
for response ratios which we report as ± 95% confidence intervals. We here define the statistical
significance level at 5%.

3. Results

3.1 Database generation

Our initial search using Web of Knowledge, Science Direct, Scopus, PubMed, and
Google Scholar returned a total of 577 unique papers published between 2000 – 2022 that assess
the concentrations and rates of mobilization of DOC in terrestrial ecosystems within the northern
circumpolar permafrost region. Of these initial 577 studies, 111 remained after the systematic screening process (Table 1 & 2). From these 111 studies we generated our database. The final database of 111 studies contained a total of 3,340 DOC concentrations (mg L⁻¹), with 2,845 DOC concentrations between 0 – 500 mg L⁻¹, found within the top 3 m of permafrost soils from field and lab-based studies (using only Day 0 lab-based DOC concentrations). These concentrations were taken from 562 different sampling locations, representing 8 different ecosystem types (Figure 1) across the northern circumpolar permafrost region. All studies except, for one (Olefeldt et al., 2012), reported DOC concentrations.

Figure 1. Map of sampling locations where DOC measurements (n=562) from the top 3 m for each ecosystem type. In many cases, the same sampling location was used in multiple studies leading to some overlap, therefore the number of sampling sites included in the data set (562) are not all clearly identifiable from this map. Retro Thaw Slump = Retrogressive Thaw Slump. Blue shading represents permafrost zonation (Brown et al., 1997).
The final database contained a considerably lower number of DOC mobilization measurements. The database includes 16 measurements of specific discharge of DOC (g DOC m\(^{-2}\)) from 3 studies, 9 export rate of DOC per day (g C m\(^{-2}\) day\(^{-1}\)) and per year (g C m\(^{-2}\) year\(^{-1}\)) measurements were each found in 2 studies. The number of specific discharge, export of DOC per day, and export of DOC per year measurements combined were <1% of the number of DOC concentration measurements. As such they were not considered for analysis of DOC mobilization. A total of 146 BDOC (%) measurements, 4% of the total number of DOC concentration measurements, were found in 14 studies. These measurements of BDOC were from Yedoma (30:5, number of measurements:studies), Upland Tundra (55:5), Forest (18:3), Permafrost Wetland (12:2), and Permafrost Bog (31:5) ecosystems. Given the low number of other forms of DOC mobilization and relatively comparable spread of BDOC measurements across ecosystem types, we chose to include BDOC measurements in our analysis despite a low total number of measurements compared to DOC concentrations, and we consider this lower sample size during our interpretation of results.

Filter size used in studies ranged from 0.15 – 0.7 μm. The majority of studies used a filter size of 0.45 μm (1,375 out of 2,845 DOC measurements), 0.7 μm was the second most common filter size (n = 489), followed by 0.22 μm (n = 332) and 0.6 μm (n = 143). Two studies used a filter size of 0.15 μm totalling 18 DOC measurements and remaining studies (n = 12) did not provide a filter size. DOC concentrations were found to differ between different filter sizes (ANOVA: F\(_{(4, 2339)}\) = 22.9, \(p < 0.001\)). DOC concentrations from samples filtered using 0.7 μm were lower (median = 11 mg L\(^{-1}\)) than 0.45 μm and 0.22 μm filtered samples (median = 53 and 42 mg L\(^{-1}\), respectively). We consider the effects of filter size to be minor. DOC concentrations were found to be significantly different between samples subject to the 11 different extraction methods used (ANOVA: F\(_{(10, 2515)}\) = 21.8, \(p < 0.001\)), and between water based and soil (solid) based extraction methods (ANOVA: F\(_{(1, 2524)}\) = 182.1, \(p < 0.001\)). Median DOC concentrations of the 4 methods of extraction directly from soils (leaching from soil under field moisture conditions, leached from dried soils, centrifuged soils, and extracted using K\(_2\)SO\(_4\)) were 57 mg L\(^{-1}\), with upper and lower quartiles of 20 and 120 mg L\(^{-1}\), respectively. The 7 water-based extraction methods had a median DOC concentration of 24 mg L\(^{-1}\), with upper and lower quartiles of 8 and 59 mg L\(^{-1}\), respectively. DOC concentrations differed (ANOVA: F\(_{(3, 2515)}\) =
36.2, \( p < 0.001 \)) between samples subject to different dissolved organic carbon measurement methods, with median values of 37 and 48 mg L\(^{-1}\) for the combustion, and photometric methods, respectively. Median values measured using the persulphate were higher at 97 mg L\(^{-1}\).

Combustion was the most common method, accounting for 2,170 DOC concentrations, followed by persulphate (\( n = 230 \)) and photometric (\( n = 31 \)). In this study we did not focus on systematically testing the effect of filter sizes, extraction methods, or DOC measurement methods. Our goal was to assess the concentration and mobilization of DOC in terrestrial permafrost ecosystems and the assessment of methods is outside the scope of our study. Rather, we compare DOC concentrations collected from samples using a variety of these methods and suggest that future studies use this information to decide on methods to be consistent with compiled measurements, thus far.

### 3.2 DOC concentrations and study regions

Upon inspection of DOC concentrations in the database, we determined that the data was non-normally distributed. The DOC concentrations were skewed toward the lower end of our 0 – 500 mg L\(^{-1}\) range; thus, we report median, upper, and lower quartiles below. Across all studies, within the top 3 m, the median DOC concentration was 41 mg L\(^{-1}\), with upper and lower quartiles of 12 and 86 mg L\(^{-1}\), respectively. DOC concentrations were found to differ among the three permafrost zones (chi-square = 32, df = 2, \( p < 0.001 \); Figure 2a). The highest median DOC concentrations were found within the sporadic permafrost zone (\( n = 83 \); 62 mg L\(^{-1}\)), lower quartile (LQ) and upper quartile (UQ) of 23 and 167 mg L\(^{-1}\), respectively. The lowest median of 33 mg L\(^{-1}\) (LQ and UQ of 11 and 88 mg L\(^{-1}\), respectively) was found in the continuous permafrost zone (\( n = 1,648 \)), with the greatest density of samples having lower DOC concentrations than observed in the violin plots of both he discontinuous and sporadic (Figure 2a). This change in DOC concentration’s along the latitudinal gradient of the permafrost zonation was also seen in the latitudinal gradient associated with ecoregion (chi-square = 78, df = 3, \( p < 0.001 \); Figure 2b). The highest DOC concentrations were found in the continental boreal (\( n = 389 \); 56 mg L\(^{-1}\), LQ = 24 mg L\(^{-1}\), UQ = 80 mg L\(^{-1}\)) and Sub-Arctic Boreal (\( n = 442 \); 58 mg L\(^{-1}\), LQ = 20 mg L\(^{-1}\), UQ = 107 mg L\(^{-1}\)) ecoregions, and lowest in the Arctic Tundra (\( n = 1,209 \); 25 mg L\(^{-1}\), LQ = 9 mg L\(^{-1}\), UQ = 84 mg L\(^{-1}\)) and Sub-Arctic Tundra (\( n = 493 \); 43 mg L\(^{-1}\), LQ = 15 mg L\(^{-1}\), UQ = 76 mg L\(^{-1}\)) ecoregions. Inspection of the distribution of DOC cocnentations across...
the ecoregions highlights that the Arctic Tundra ecoregion had the highest density of samples at the lowest DOC concentration (Figure 2b).

These latitudinal differences are also reflected in the observed differences (chi-square = 20, df = 3, $p < 0.001$) in DOC concentrations found within different soil classes. The highest DOC concentrations are found within organic rich Histosol (n = 37; 61 mg L$^{-1}$; LQ = 32 mg L$^{-1}$; UQ = 71 mg L$^{-1}$) and Histel soils (n = 935; 53 mg L$^{-1}$; LQ = 16 mg L$^{-1}$; UQ = 88 mg L$^{-1}$; Figure 2c), with the distribution of the data from these soils types having a higher density at greater DOC concentrations (Figure 2c). Histel and Histosol soils are the main type of permafrost soil found within the sporadic and discontinuous permafrost zone and both boreal ecoregions (Hugelius et al., 2014). Mineral rich Orthels (n = 741; 38 mg L$^{-1}$; LQ = 11 mg L$^{-1}$; UQ = 102 mg L$^{-1}$) and Turbels (n = 820; 31 mg L$^{-1}$; LQ = 12 mg L$^{-1}$; UQ = 74 mg L$^{-1}$), mineral permafrost soils that have experienced cryoturbation, had the lowest DOC concentrations. The median DOC concentrations found within the top 3 m of these soil classes represent <1% of the soil organic carbon stock found in the top 3 m of each soil class (Hugelius et al., 2014). DOC concentrations also differed within the thermal horizon of these different soil classes (chi-square = 91, df = 3, $p < 0.001$; Figure 2d). The highest DOC concentrations were found in permafrost free sites (n = 202; 57 mg L$^{-1}$; LQ = 47 mg L$^{-1}$; UQ = 69 mg L$^{-1}$), which were largely Histosol soils (19%) or Histel soils (74%) that have experienced thermokarst formation. In areas where permafrost was present, DOC concentrations were highest in the active layer (n = 1,400; 45 mg L$^{-1}$; LQ = 14 mg L$^{-1}$; UQ = 88 mg L$^{-1}$) and the permafrost lens (n = 729; 30 mg L$^{-1}$; LQ = 10 mg L$^{-1}$; UQ = 123 mg L$^{-1}$).
Figure 2. Violin plots of DOC concentrations (mg L⁻¹) found in the top 3 m across (a) permafrost zones, (b) ecoregions, (c) soil classes, and (d) thermal horizons. (a) Dark to light blue shading represents the permafrost zones Continuous, Discontinuous, and Sporadic, according to Brown et al., (1997). (b) Dark to light green shading represents the ecoregions Arctic Tundra, Sub-Arctic Tundra, Continental Boreal, and Sub-Arctic Boreal, according to Olson et al., (2001). (c) Dark to light yellow shading represents the soil classes Histosol, Histel, Orthel, and Turbel, according to the USDA Soil Taxonomy (USDA, 1999). Histosols are organic rich, non-permafrost soil. Histels, Orthels, and Turbels are permafrost-affected soils (Gelisol order). Histels are organic rich, Orthels are non-cryoturbated affected mineral soils, and Turbels are cryoturbated permafrost soils. (d) Dark to light red shading represents the thermal horizons Active Layer, Permafrost Lens, Thaw Stream, and Permafrost Free. Active layer represents the seasonally unfrozen soil layer above the permafrost layer. Permafrost Lens represents the permanently frozen (below 0°C) layer. Thaw Stream represents flowing surface waters following permafrost thaw. Permafrost Free represents areas that are not underlain by permafrost. Black dots on each violin plot represents the median. Black vertical lines represent the interquartile range with the upper and lower limits representing the 75th and 25th percentiles, respectively. Either side of the black vertical line represents a kernel density estimation. This shape shows the distribution of the data, with wider areas representing a higher probability that samples within the database will have that DOC concentrations. The number of samples (n) found in each sub-category is found above each corresponding violin plot.
3.3 Trends in DOC concentrations across ecosystems

Similar to other categorical variables (i.e. permafrost zone, ecoregion, soil class, and thermal horizon data), DOC concentrations within each of the eight ecosystem types were found to be non-normally distributed, with median values skewed toward the lower end of the $0 – 500$ mg L$^{-1}$ range of concentrations (Figure A1). Permafrost bogs and permafrost wetlands were the most represented in the database with regards to DOC concentrations, with a total of 685 concentrations from 38 studies and 679 concentrations from 22 studies, respectively. The majority of permafrost bog measurements came from studies with field sites within Canada (Figure 1), as was the case for upland tundra and retrogressive thaw slump DOC concentration data. The majority of permafrost wetland sample locations were found in Russia, whereas the majority of the 399 coastal tundra sampling locations were in the USA. The least represented ecosystem classes included the peatland ecosystem class, which is not strictly a permafrost ecosystem as the other are, and the Yedoma ecosystem class (118 DOC concentrations from 9 studies). DOC concentrations differed significantly across the eight ecosystem types (chi-square $= 700$, df = 7, $p < 0.001$; Figure 3). The highest DOC concentrations were found in coastal tundra ($66$ mg L$^{-1}$; LQ = $24$ mg L$^{-1}$; UQ = $140$ mg L$^{-1}$) and permafrost bogs ($63$ mg L$^{-1}$; LQ = $36$ mg L$^{-1}$; UQ = $111$ mg L$^{-1}$) ecosystems. The lowest DOC concentrations were found in permafrost wetlands ($7$ mg L$^{-1}$; LQ = $6$ mg L$^{-1}$; UQ = $26$ mg L$^{-1}$) and Yedoma ecosystems ($9$ mg L$^{-1}$; LQ = $2$ mg L$^{-1}$; UQ = $20$ mg L$^{-1}$), both of which had only slightly lower median DOC concentrations than retrogressive thaw slumps ($15$ mg L$^{-1}$; LQ = $7$ mg L$^{-1}$; UQ = $26$ mg L$^{-1}$).
Figure 3. Boxplot and jitter plot of (a) DOC concentrations (mg L\(^{-1}\)), (b) the number of DOC measurements, and (c) number of studies including DOC measurements were taken from the top 3 m for each ecosystem type. Retro Thaw Slump = Retrogressive Thaw Slump. Boxes represent the interquartile range (25 \(-\) 75\%), with median shown as black horizontal line. Whiskers extend to 1.5 times the interquartile range (distance between first and third quartile) in each direction. Jitter points represent the concentration of each individual DOC measurement, with random variation applied to each points location vertically in the plot, to avoid overplotting. Yedoma = dark teal. Coastal Tundra = orange. Retro Thaw Slump = red. Upland Tundra = green. Forest = purple. Permafrost Wetland = light pink. Permafrost bog = yellow. Peatland = brown.

When grouping all DOC concentrations by ecosystem types and differentiating between the active layer and permafrost lens thermal horizons, we found that DOC concentrations differed between the active layer and permafrost for all ecosystems (ANCOVA: \(F_{(1, 1277)} = 49.8\), \(p < 0.001\)) except for permafrost bogs (chi-square = 0.37, df = 1, \(p = 0.5\)) and Yedoma (chi-square = 3.5, df = 1, \(p = 0.06\)) ecosystems (Figure 4). Within the permafrost lens thermal horizon, the highest DOC concentrations were found in coastal tundra (\(n = 103; 130 \text{ mg L}^{-1}; \text{LQ} = 60 \text{ mg L}^{-1}; \text{UQ} = 179 \text{ mg L}^{-1}\)) and permafrost bogs (\(n = 248; 78 \text{ mg L}^{-1}; \text{LQ} = 19 \text{ mg L}^{-1}; \text{UQ} = 163 \text{ mg L}^{-1}\)) sites, and lowest found in Yedoma sites (\(n = 91; 8 \text{ mg L}^{-1}; \text{LQ} = 6 \text{ mg L}^{-1}; \text{UQ} = 16 \text{ mg L}^{-1}\)).
The highest active layer DOC concentrations were in permafrost bogs (n = 276; 64 mg L\(^{-1}\); LQ = 41 mg L\(^{-1}\); UQ = 102 mg L\(^{-1}\)) and forest (n = 185; 57 mg L\(^{-1}\); LQ = 26 mg L\(^{-1}\); UQ = 110 mg L\(^{-1}\)) sites, and lowest found in permafrost wetland sites (n = 274; 10 mg L\(^{-1}\); LQ = 5 mg L\(^{-1}\); UQ = 47 mg L\(^{-1}\)).

Figure 4. Boxplot of (a) DOC concentrations (mg L\(^{-1}\)) and (b) the number of DOC measurements in the Active Layer and Permafrost Lens thermal horizons of each ecosystem type. Only DOC concentrations from ecosystems with these thermal horizons present is used, thus no peatland or permafrost-free sites are included. Retro Thaw Slump = Retrogressive Thaw Slump. Boxes represents the interquartile range (25 – 75%), with median shown as black horizontal line. Whiskers extend to 1.5 times the interquartile range (distance between first and third quartile) in each direction. Blue boxplots represent DOC concentrations in the active layer. Gren boxplots represent DOC concentrations in the permafrost lens.

3.4 Drivers of DOC concentrations
No continuous variables recorded in the dataset were available for all DOC concentration database entries, with no sites containing data for all continuous variables. This limited our ability to explore relationships between continuous environmental and ecological data and DOC concentrations across the permafrost region. To address drivers of DOC concentrations across the circumpolar permafrost region we used partial least squares regression (PLS) as it is tolerant to missing values. Multiple PLS regressions were run using various combinations of continuous and categorical data with similar model performance throughout. We chose the PLS to predict DOC concentrations using environmental continuous variables and ecosystem type as this contained the lowest background correlation. The most parsimonious PLS regression extracted 5 significant components, captured 79% variation of the predictor variables, and explained 37% of the variance in DOC concentrations in the dataset. The majority of the variance in DOC (35%) is explained along the first two axes of the model. The model was robust and not overfitted as model predictability was moderate ($Q^2 = 0.35$) and background correlation was low (0.006).

The PLS plot (Figure 5a) shows the correlation between DOC concentrations and selected environmental and ecological variables for the first two axes of the model. The two variables with the greatest positive and negative effect on DOC concentrations were total dissolved nitrogen content (mg L$^{-1}$) and C:N ratios, respectively (Figure 5b). The positive relationship between DOC and total dissolved nitrogen soil carbon content (SoilC), and negative relationship with the specific UV absorbance at 254 nm (SUVA), may be a result of ecosystem properties. The aromatic content of organic matter is positively correlated with SUVA (Weishaar et al., 2003), with high SUVA values being used as an indication of high aromatic content (Hansen et al., 2016). Ratios of C:N have been shown to be a good proxy for decomposition (Biester et al., 2014), where high C:N values indicate higher decomposition. The strong negative relationship with C:N ratios indicates that DOC concentrations decrease with increased decomposition. Other than higher soil carbon content (SoilC) in permafrost bogs, there was no clear or obvious trends in SoilC, TDN, C:N ratios, and SUVA across ecosystem types (Figure A3). The PLS demonstrates that ecosystem type strongly affects DOC concentrations, with DOC positively related with the highest ecosystems where the highest DOC concentrations are observed, permafrost bogs and coastal tundra, and negatively related to the lower DOC ecosystems, permafrost wetland and retrogressive thaw slumps (Figure 5). This negative
relationship may be due to the higher latitudes these ecosystems are generally found at, which is supported by the negative relationship with DOC and the climate indicators mean annual temperature (MAAT) and mean annual precipitation (MAP). Additionally, it may be due to the high number of thermokarst affected sites found within these ecosystem classes, particularly retrogressive thaw slumps. There is a clear negative relationship between DOC concentrations and disturbed permafrost wetlands, retrogressive thaw slumps, and permafrost bogs.
Figure 5. Partial least squares regression (PLS) (a) loadings plot explaining 37% of the variability observed in DOC concentrations. PLS component axis 1 explains 28.8% of this variability, whereas PLS component axis 2 explains 5.8%. The remaining axes explain the variability in DOC are not shown for clarity. (b) Bar plot of PLS regression coefficients showing the relative importance of each variable in predicting DOC concentrations. Regression coefficients on y-axis are normalized so their absolute sum is 100, with positive and negative values indicating the direction of the relationship. In the loadings plot squares depict ecosystem classes and the blue triangle represents DOC concentrations. Black circles in the (a) loadings plot and black bars in the (b) bar plot represent continuous environmental data that had at least 20% coverage of DOC data. All continuous data was log transformed, mean centered, and standardized. Continuous data variables are represented by the colour black. CN = carbon:nitrogen ratio. SUVA = the specific UV absorbance at 254 nm (L mg C⁻¹ m⁻¹). MAP = mean annual precipitation (mm). MAAT = mean annual temperature. SoilC = carbon content of soil (g C kg⁻¹). TDN = total dissolved nitrogen (mg L⁻¹). Fe = dissolved iron ((mg L⁻¹)). PermWet.D = disturbed permafrost wetland ecosystem class and is light pink (as in Figure 3) to represent this ecosystem class. RTS = retrogressive thaw slump ecosystem class and is red (as in Figure 3) to represent this ecosystem class. Coast.AL = active layer of coastal tundra ecosystem class and is orange (as in Figure 3) to represent this ecosystem class. PermBog.AL = active layer of permafrost bog ecosystem class and is yellow (as in Figure 3) to represent this ecosystem class. PermBog.P = permafrost lens of permafrost bog ecosystem class and is yellow (as in Figure 3) to represent this ecosystem class. PermBog.D = disturbed permafrost bog ecosystem class and is yellow (as in Figure 3) to represent this ecosystem class.

3.5 Response and mobilization of DOC and BDOC to thermokarst formation

The highest DOC concentrations were found in pristine permafrost bog (75 mg L⁻¹; LQ = 37 mg L⁻¹; UQ = 149 mg L⁻¹, n = 442) and coastal tundra ecosystems (72 mg L⁻¹; LQ = 25 mg L⁻¹; UQ = 151 mg L⁻¹ n = 427; Figure 6a). No thermokarst affected coastal tundra ecosystems were recorded within the dataset. Whereas, in permafrost bogs DOC concentrations were found to differ across different thermokarst disturbances (ANOVA: F(3, 720) = 23.04, p < 0.001), with the lowest found in thermokarst wetlands (10 mg L⁻¹; LQ = 9 mg L⁻¹; UQ = 30 mg L⁻¹, n = 16). DOC concentrations were also found to differ between thermokarst affected and pristine sites in upland tundra ecosystems (ANOVA: F(3, 539) = 5.91, p < 0.001). The highest DOC concentrations in upland tundra ecosystems were found in sites that had experienced active layer thickening (53 mg L⁻¹; LQ = 41 mg L⁻¹; UQ = 80 mg L⁻¹, n = 142), whereas the lowest were found in sites that had experienced active layer detachment (4 mg L⁻¹; LQ = 3 mg L⁻¹; UQ = 5 mg L⁻¹, n = 6). Pristine sites had the highest DOC concentrations in both Yedoma (11 mg L⁻¹; LQ = 6 mg L⁻¹; UQ = 21 mg L⁻¹, n = 114) and forest (49 mg L⁻¹; LQ = 22 mg L⁻¹; UQ = 86 mg L⁻¹, n = 189) ecosystems. However, in permafrost wetland ecosystems pristine sites had the lowest DOC concentrations (7 mg L⁻¹; LQ = 6 mg L⁻¹; UQ = 57 mg L⁻¹, n = 766) with sites that were...
affected by both thermokarst wetland formation (21 mg L$^{-1}$; LQ = 11 mg L$^{-1}$; UQ = 37 mg L$^{-1}$, n = 17) and active layer thickening (41 mg L$^{-1}$; LQ = 34 mg L$^{-1}$; UQ = 47 mg L$^{-1}$, n = 12) having higher DOC concentrations.

BDOC was found to differ between thermokarst disturbances within ecosystem types in only Yedoma (ANOVA: $F_{(2, 27)} = 23.09$, $p < 0.001$) and permafrost wetland (ANOVA: $F_{(1, 10)} = 15.87$, $p < 0.001$) ecosystems. The highest BDOC was found in both of these ecosystem types also, with 54% (n = 5) in pristine Yedoma sites and 49% (n = 8) in thermokarst wetland affected permafrost wetland sites (Figure 6b), with the latter exhibiting the highest BDOC across all permafrost affected sites followed by thaw slumps (18%, n = 11) in Yedoma ecosystems and active layer thickening (40%, n = 1) in upland tundra sites. The lowest median BDOC of 4% were seen in thermokarst bogs (n = 5) and active layer thickening (n = 3) affected sites, with pristine sites experiencing BDOC of 9% (n = 15). However, not all ecosystem types in the database had BDOC data for both pristine and disturbance sites. For example, only pristine sites data was available for forests, whereas there was no pristine site data available for upland tundra sites. No BDOC data was available for coastal tundra sites.

All ecosystem types that had BDOC data, reported BDOC observed following 40 – 90 incubation days, and this also corresponded to the highest BDOC values for each ecosystem type (Figure A4). When comparing the greatest BDOC observed within this incubation length window, we found that values varied across ecosystem type (ANOVA: $F_{(5, 131)} = 14.6$, $p < 0.001$). The highest loss rates were observed in Yedoma and permafrost wetland ecosystems, whereas the lowest we observed in organic rich forest and permafrost bog ecosystems (Figure A4). Forest (ANOVA: $F_{(1, 16)} = 2.31$, $p = 0.15$) and permafrost bog (ANOVA: $F_{(3, 24)} = 2.49$, $p = 0.09$) BDOC did not differ over incubation length, whereas Yedoma (ANOVA: $F_{(4, 25)} = 24.92$, $p < 0.001$) and permafrost wetland (ANOVA: $F_{(1, 10)} = 15.87$, $p < 0.01$) did differ over time, with their max occurring during this 40 – 90-day incubation length. This suggests that when incubated for the same number of days, we would expect greater BDOC in Yedoma and permafrost wetland ecosystems. Note, for this analysis BDOC values from all thermokarst and non-thermokarst affected sites within an ecosystem type were included.
Figure 6. DOC concentrations (mg L\(^{-1}\)) and biodegradable DOC (BDOC; %) from the top 3 m following disturbance including data from both field based and incubation studies. (a) DOC concentrations from each ecosystem type following disturbance where data was available. (b) Biodegradable DOC (BDOC) from each ecosystem type following disturbance where data was available. BDOC loss was determined following 3 – 304 days of incubation. Data from different incubation lengths was combined due to low sample size. Retro Thaw Slump = Retrogressive Thaw Slump. Boxes represents the interquartile range (25 – 75%), with median shown as black horizontal line. Whiskers extend to 1.5 times the interquartile range (distance between first and third quartile) in each direction, with outlier data plotted individually as black dots. Note colours associated with boxplots in this figure are only relevant for this figure.

Response ratios comparing the change in DOC concentrations between pristine and thermokarst affected sites were calculated from our dataset from 108 studies using Eq. 1 (Figure 7). Only 17 studies provided data for both pristine and thermokarst affected ecosystems, with 87 papers providing DOC concentrations from pristine and 34 from thermokarst affected sites. When considering all ecosystems together we found that response ratios were negative, suggesting that DOC concentrations were higher in thermokarst affected sites compared to pristine sites (Figure 7). These negative response ratios were most evident in permafrost bogs, where they found throughout the entire column and individual thermal horizons. The greatest
increase in DOC concentrations following thermokarst was seen when comparing DOC concentrations in the permafrost lens of permafrost bogs, and to a lesser extent permafrost wetlands (Figure 7). Only in Yedoma ecosystems did we see positive response ratios throughout the entire profile, suggesting a decrease in DOC concentrations following thermokarst formation in Yedoma sites. This was also seen for DOC concentrations within the permafrost lens of upland tundra sites, which include DOC concentrations from retrogressive thaw slumps and thermo-erosion gullies in their thermokarst affected sites. The large confidence intervals for some response ratios suggests high variability in the response of DOC concentrations to thermokarst formation.
Figure 7. Response ratios of DOC concentrations from the top 3 m following thermokarst formation (calculated using Eq. 1). Response ratio means allow for relative comparison of changes in DOC following thermokarst formation between different ecosystem types. Negative values indicate lower DOC concentrations found in pristine ecosystems, whereas positive value indicates a decrease in DOC concentrations following thermokarst. Studies reporting DOC concentrations from Exposures, Retrogressive Thaw Slumps, and Thermo-Erosion Gullies from sites within the continuous permafrost zone were combined into the Upland Tundra ecosystem category. This did not include DOC concentrations from studies within the Yedoma permafrost domain (Strauss et al., 2021). Blue line represent DOC concentrations in the active layer, as per Figure 4. Green lines represent DOC concentrations in the permafrost lens, as per Figure 4.
Brown lines represent DOC concentrations from the entire column (i.e., both active layer and permafrost lens).

4. Discussion

In this systematic review, we evaluated patterns of DOC concentrations in the top 3 m of terrestrial ecosystems across the northern circumpolar permafrost region based on results from 111 studies and 2,845 DOC measurements. We focused on comparing concentrations of DOC in soils across various geographical regions, ecological conditions, and disturbance types. Our synthesis shows that median DOC concentrations across ecosystems range from 9 – 61 mg L\(^{-1}\), which represents similar albeit slightly higher DOC concentrations when compared to the median DOC concentrations found in top soils of other land cover groups below 50°N (25 mg L\(^{-1}\); Langeveld et al., 2020), globally distributed lakes (6 mg L\(^{-1}\); Sobek et al., 2007), and lakes across the permafrost region (11 mg L\(^{-1}\); Stolpmann et al., 2021). In general, we show that organic soils have higher DOC concentrations than mineral soils, and that DOC concentrations are positively related to total dissolved nitrogen concentrations and negatively to C:N ratios, which corroborate previous findings of factors correlating with DOC concentrations (Aitkenhead & McDowell, 2000; Lajtha et al., 2005). Overall, we found that properties associated with ecosystem type are the main constraint on DOC concentrations. Furthermore, disturbance through permafrost thaw has little impact on measured DOC concentrations, however this may be due to the loss of biologically reactive DOC or the loss of an initially larger pulse of DOC having been previously mobilised prior to the timing of sampling.

4.1 Environmental factors influencing DOC

Our database confirmed our first hypothesis that the highest DOC concentrations would be found in organic rich soils. Previous synthesis efforts estimating global distributions of terrestrial DOC concentrations have presented similar findings (Guo et al., 2020; Langeveld et al., 2020). Both of these previous studies also show that some of the highest terrestrial DOC concentrations are found within the northern circumpolar permafrost region, highlighting that these high DOC concentrations found in organic rich permafrost soils are of global significance. Concentrations of DOC in the top 3 m of soils closely mirrored stocks of SOC across the circumpolar permafrost region (Hugelius et al., 2014). Organic rich Histosol and Histel soils contain the greatest SOC.
per km², followed by Turbels and Orthels (Hugelius et al., 2014), as was seen in DOC concentrations across these soil types (Figure 2a). While the highest DOC concentrations are found within organic rich soils, the amount of C found as DOC represent a small amount of the total SOC pool. Using the current best estimates of Histel SOC stocks (Hugelius et al., 2020), the DOC pool represents <1% of the total C stock in permafrost-affected peatlands as has been shown for both permafrost and global soils (Guo et al., 2020; Prokushkin et al., 2008).

### 4.2 Thermal horizons

In many ecosystems, DOC concentrations are greatest in the active layer nearer the surface (Figure 4). This trend has also been observed in the vertical distribution of DOC across global soils, with 50% of the DOC pool found in the top 0 – 30 cm (Guo et al., 2020). The production of DOC is associated with soil microbial activity (Guggenberger & Zech, 1993) and plant inputs (Moore & Dalva, 2001), and the microbial production of DOC via input of labile substrates has been shown to decrease with depth in permafrost (Hultman et al., 2015; Monteux et al., 2018; Wild et al., 2016). Furthermore, the organic matter content decreases and mineral content increases with depth, this depth trend and decrease in DOC with depth is particularly evident between the active layer and permafrost lens in forest ecosystems (Figure 4a). While permafrost and non-permafrost bogs do also see a shift in microbial community with depth (Heffernan & Cavaco et al., 2022; Lamit et al., 2021), the movement of modern, surface derived DOC down into deeper layers has also been observed (Chanton et al., 2008; Estop-Aragonés et al., 2018). These, combined with the large, frozen SOC stores found at depth (Hugelius et al., 2020) and hydrological isolation (Quinton, Hayashi, & Chasmer, 2011), results in a DOC pool that remains relatively similar across thermal horizons in permafrost bogs (Figure 4b). Intriguingly, in both coastal tundra and permafrost wetland ecosystems, DOC concentrations were found to be higher in the permafrost lens than in the active layer. This suggests that DOC within the active layer of these ecosystems experienced some degree of mobilization, either via export to the aquatic network or enhanced decomposition within soils. The higher DOC concentrations found within the permafrost lens of these ecosystems may represent a vulnerable DOC pool to enhanced mineralization following permafrost thaw (Figure 6).

### 4.3 Variation in DOC amongst permafrost zones and ecoregions
Permafrost soils are estimated to store 1,035 ± 150 Pg C globally within the top 0-3 m (Hugelius et al., 2014), with the highest storage of SOC found in the organic rich Histosols and Histels. While persistent low temperatures are the main common factor which has led to the accumulation of such high SOC amongst all permafrost soils, environmental factors associated with the different ecosystem types are the main driving factors in differences amongst DOC concentrations. The source of the permafrost DOC pool is from recent plant leachate inputs, or from the decomposition and solubilization of SOC. Thus, the molecular composition of the DOC pool is derived from a mixture of current and historical vegetation inputs. There are clear current and historical shifts in dominant vegetation seen in the permafrost region from the south (boreal) to north (arctic tundra), as well as across ecosystem types (upland forest, upland tundra, arctic and boreal wetland). However, the majority of vegetation and its leachates found in the permafrost region are generally found to produce relatively stable DOC (in terms of BDOC) consisting of lignin-derived compounds, highly aromatic polyphenolic compounds, and low molecular weight organic acids (Chen et al., 2018; Drake et al., 2015; Ewing et al., 2015; Selvam et al., 2017). While differences in the stability of different DOC source end-members have been shown (MacDonald et al., 2021), differences in redox conditions are likely a major driver in differences in the accumulation and mineralization of DOC across permafrost ecosystem types (Mohammed et al., 2022).

Similar to their globally significant stores of SOC (Hugelius et al., 2020), the accumulation of high DOC concentrations found in peatlands, permafrost bogs, and permafrost wetlands, is a result of the prevalence of cold and anoxic conditions throughout the Holocene (Blodau, 2002). This leads to a reduction in microbial decomposition, and the accumulation of both the SOC and DOC pool. Our results suggest that the pristine permafrost bog and permafrost wetland DOC pool is relatively stable following permafrost thaw (Figure 6, 7a). Peatland vegetation, in particular Sphagnum mosses, produces litter that has anti-microbial properties and is decay resistant (Hamard et al., 2019; Limpens, Bohlin, & Nilsson, 2017), limiting the amount of SOC that is degraded and assimilated into the DOC pool (Tfaily et al., 2013). This is further enhanced by the build-up of decomposition end products and the thermodynamic constraint on decay observed in anoxic soils (Beer et al., 2008). Permafrost has been continuously present in peatlands across the northern circumpolar permafrost region for the past 6,000 years, with the...
greatest rates of permafrost formation occurring within the past 3,000 years (Treat & Jones, 2018). Thus, a large proportion of the organic matter found peatlands and wetlands in this region were present prior to permafrost aggradation (i.e., permafrost formation), which indicates that permafrost formed epigenetically in these areas. Permafrost aggradation impacts soil biogeochemical properties, leading to potentially less decomposed organic matter with higher C/N ratios than non-permafrost equivalent soils, particularly in permafrost wetlands (Treat et al., 2016). This can lead to the build-up of high DOC concentrations that are vulnerable to potential mobilization following thermokarst. Decomposition in epigenetic permafrost bogs following thermokarst has been shown to be relatively slow (Heffernan et al., 2020; Manies et al., 2021), which further supports our finding (Figure 6) that the large DOC pool found in these systems in relatively stable following permafrost thaw. The permafrost wetland DOC pool that accumulates following thermokarst may represent a potentially labile DOC pool (Figure 7a), but this is likely due to fresh, plant derived inputs rather than the exposure and mineralization of previously frozen organic matter (Figure 7a).

Coastal tundra and forest ecosystems had similarly high DOC concentrations to those found in permafrost bogs (Figure 3a). Coastal tundra and forest ecosystems represented the highest concentrations of DOC in mineral permafrost soils. Concentrations of coastal permafrost DOC were significantly lower in the active layer compared to within the permafrost lens (Figure 4a). This is contrary to findings that deeper coastal permafrost consists of low organic matter Pleistocene marine sediments (Bristol et al., 2021) and the proximity of the active layer to vegetation inputs, although this productivity and inputs are vulnerable to projected climatic warming and regional “browning” and “greening” (Lara et al., 2018). Recent work has shown that DOC in the active layer within the coastal permafrost is more biodegradable that OC in the permafrost lens (Speetjens et al., 2022) and a substantial proportion of organic carbon derived from thawing coastal permafrost is vulnerable to mineralization upon thawing, particularly when exposed to sea water (George Tanski et al., 2021). Export of terrestrial coastal permafrost DOC directly into the Arctic Ocean can significantly influence marine biogeochemical cycles and food webs within the Arctic ocean (Bruhn et al., 2021). Arctic coasts are eroding at rates of up to 25 m yr\(^{-1}\) (Fritz, Vonk, & Lantuit, 2017) and exporting large quantities of terrestrial organic matter export directly to the ocean that is rapidly mineralized (Tanski et al., 2019). Enhanced DOC
export from these coastal tundra ecosystems may disrupt aquatic food webs through altering nutrient and light supply, as has been shown for Swedish coastal systems (Peacock et al., 2022). These coastal tundra sites represent a large DOC pool that is highly vulnerable to enhanced mobilization and deserve further attention.

The remaining ecosystems characterised by mineral soils with an upper organic layer, i.e., Forests, Upland Tundra, and Yedoma, followed a clear latitudinal climate gradient of increasing DOC concentrations from north to south. While not included in the most parsimonious PLS model (Figure 5), Yedoma and Upland Tundra ecosystems were found to negatively correlate with DOC concentrations (Figure A5). The greatest proportions of OC and nutrients used for DOC production are found in shallow organic layers (Semenchuk et al., 2015; Wild et al., 2013) in these ecosystems. Beneath the upper organic horizons in these mineral soils processes such as sorption of DOC to minerals and the formation of Fe-DOC or Al-DOC complexes may remove DOC from the dissolved pool (Kawahigashi et al., 2006) and mechanically protect it from mobilization (Gentsch et al., 2015). In forest ecosystems, large amounts of SOC have accumulated in surface organic layers (Hugelius et al., 2014) through increased vegetative inputs due to warmer and longer growing seasons. This organic layer depth, and the impact of soil temperature, moisture, and pH on SOC found there, strongly influences the production, concentration, and composition of DOC (Neff & Hooper, 2002; Wickland et al., 2007). Furthermore, the sorption of DOC to charcoal (Guggenberger et al., 2008), and high lignin and phenolic input from vegetation (O’Donnell et al., 2016) produce a difficult to degrade DOC pool, leading to the accumulation of the large DOC pool in this ecosystem type.

4.4 Vulnerability of DOC to enhanced mobilization following thermokarst

We define DOC mobilization as DOC lost from an ecosystem either via export or degradation. Our second hypothesis that permafrost thaw would lead to enhanced mobilization of DOC cannot be fully supported by the findings from this database. Using our chosen systematic approach and focusing on data from terrestrial ecosystems, our database was limited to 3 studies which represented <1% of the DOC concentration data. Several previous studies have detailed the export of DOC in Arctic inland waters, see Table 2 in Ma et al., (2019), that have been excluded using this approach. We acknowledge the limitation in our approach regarding the
inclusion of DOC export data. Thus, this database cannot be used to determine how permafrost thaw will influence DOC export from terrestrial ecosystems within the northern circumpolar permafrost region. Currently, Arctic rivers are estimated to export 25 – 36 Tg DOC year\(^{-1}\) (Amon et al., 2012; Holmes et al., 2012), with this being dominated by modern carbon sources (Estop-Aragonés et al., 2020), most likely derived from the top 1 m of terrestrial ecosystems. Using current best estimates of the areal extent and soil organic carbon stores in the top 1 m of Histosols, Histels, Orthels and Turbels (Hugelius et al., 2014), and if we assume that the DOC pool represents ~1% of the SOC pool, we estimate that <1% of the current DOC pool found in the top 1 m of Histosols, Histels, Orthels and Turbels is exported annually to Arctic rivers.

Quantifying the proportion of these DOC pools annually lost, and particularly the proportions lost in headwater streams while being exported to Arctic rivers, is vital to assess the importance of the mobilization of the terrestrial permafrost DOC pool.

Our calculated response ratios (Figure 7) for all ecosystems, indicating the difference in DOC concentrations between pristine and permafrost thaw affected sites, partly supports our second hypothesis that disturbance would lead to increased export and biodegradability of DOC. The increase in DOC following thaw observed in permafrost bogs is likely due to increased inputs due to increased runoff and shifts in vegetation following permafrost thaw (Burd, Estop-Aragonés, Tank, & Olefeldt, 2020), a relatively stable soil organic carbon pool at depth due to several millennia of microbial processing (Manies et al., 2021), the prevalence of anoxic conditions, and the potential hydrological isolation of thermokarst bogs (Quinton, Hayashi, & Pietroniro, 2003). While not included in our analysis, DOC found near the surface of the permafrost lens in forest ecosystems has been shown to be more biodegradable than DOC found in the active layer (Wickland et al., 2018), and may represent a decrease in DOC following thermokarst not captured here. Our findings of limited mobilization of permafrost bog DOC upon thawing are supported by the findings that the \(^{14}\)C signature of DOC in Arctic rivers is dominated by modern sources (Estop-Aragonés et al., 2020). However, we do see a reduction in DOC concentrations in thermokarst affected sites at the higher latitude Yedoma upland tundra, and permafrost wetland ecosystems. This reduction in DOC concentrations in these ecosystems may be due to the greater biodegradability and lability of the DOC found there (Figure 6b), supporting our third hypothesis that the most biodegradable DOC would be found in higher
Permafrost DOC in higher latitude ecosystems, particularly Yedoma ecosystems characterised by syngenetic permafrost aggradation which have not undergone centuries to millennia of soil formation and microbial processes, have been shown contain a greater proportion of low oxygen, aliphatic compounds and labile substrates (Ewing et al., 2015b; MacDonald et al., 2021). This leads to a greater biolability and rapid mineralization of DOC (Vonk et al., 2015), potentially causing the reduction in DOC concentrations observed following thaw. If this hypothesis is to be found true across all high latitude ecosystems with further data, it further highlights the vulnerability of the large DOC pool found in coastal tundra ecosystems.

In this study, we focus on the dissolved fraction of the OC pool, however the particulate fraction should also be considered when discussing the mobilization of terrestrial OC in permafrost landscapes. In boreal freshwater networks, particulate organic carbon (POC) represents a small but highly labile fraction of terrestrially derived OC exported to the fluvial network (Attermeyer et al., 2018). The degradation of permafrost derived POC is much slower than that of POC in the boreal freshwater network and POC derived from younger sources along the riverbank (Shakil, Tank, Kokelj, Vonk, & Zolkos, 2020). The DOC pool in Arctic freshwaters in dominated by modern terrestrial sources (Estop-Aragonés et al., 2020), whereas the POC pool has been shown to be dominated by older sources in both permafrost peatland dominated areas (Wild et al., 2019), following the formation of retrogressive thaw slumps (Keskitalo et al., 2021), and in thermokarst affected periglacial streams (Bröder et al., 2022). This older POC has been shown to accumulate following export due to low lability and degradation and mineral association, which suggests that upon thermokarst formation, previously frozen OC exported in the particulate phase is not readily consumed by microbes and that permafrost derived DOC is the more labile fraction of exported terrestrial OC.

4.5 Future considerations for study design

Determining the fate of mobilized terrestrial DOC in both permafrost thaw affected, and pristine sites should be prioritized in future studies to constrain current estimates of the permafrost C climate feedback. There are large spatial gaps in the database, particularly in areas with large stock of permafrost C such as the Hudson Bay Lowlands and Mackenzie River Basin.
both in Canada and two of the three largest deposits of permafrost peatland C in the circumpolar permafrost region (Olefeldt et al., 2021). Similarly, coastal tundra sites, which along with permafrost bog represent the ecosystems with the highest DOC concentrations, were sampled only along the northern shoreline of Alaska and the Yukon (USA and Canada, respectively; Table S1). From our analysis of this database, we determine that DOC mobilization is poorly understood for terrestrial permafrost ecosystems. To address this, the two main needs of future studies are 1) more direct estimates of DOC fluxes and export from terrestrial ecosystems into aquatic ecosystems, and 2) more DOC degradation (BDOC) and mineralization studies. Our results suggest that the high concentrations of DOC in permafrost bogs remains relatively stable upon thermokarst formation, although individual studies do indicate that thawing peat may provide a reactive source of DOC (Panneer Selvam et al., 2017). Whereas the database did not include any studies that reported on the mineralization of DOC from coastal tundra sites. Further sampling and assessing the mineralization of DOC is required to characterize the potential pool of vulnerable DOC in areas with high DOC concentrations. Overall, our database and systematic approach only included 5 studies (Olefeldt & Roulet, 2012, 2014; Olefeldt et al., 2012; Prokushkin et al., 2006; Prokushkin et al., 2005) that explicitly reported rates of DOC discharge, export, or fluxes from terrestrial ecosystems into the fluvial network. Given the importance of terrestrial DOC as a source for CO₂ production within the aquatic network (Weyhenmeyer et al., 2012), and the findings that previously frozen DOC is being exported to the freshwater network (Estop-Aragones et al., 2020), improved estimates of the quantity of terrestrial DOC being exported is essential to determine the potential aquatic greenhouse gas fluxes derived from the mineralization of terrigenous organic matter. To improve current estimates of the permafrost C feedback further studies are needed to determine how much DOC is laterally exported from terrestrial ecosystems, and the mineralization potential of this DOC along the terrestrial-freshwater-aquatic continuum.

Lastly, we suggest that future studies should consider a standardization of methods and approached used to determine DOC concentrations for better comparison across studies. In constructing this database we identified 3 different filter sizes, 11 different extraction procedures, and 4 different measurement methods. The most common filter size used was 0.45 µm and this has previously been described as the cut off to separate DOC from colloid materials (Thurman...
1985; Bolan et al., 1999). In extracting DOC concentrations from soils the mostly commonly used approach (70% of all soil samples) was via soil leaching with no chemical treatment of the soils, although some added filtered water to promote leaching. From the seven approaches identified to extract water samples from terrestrial sites in determining DOC, 48% of samples were collected using a variety of suction devices and 46% done via grab samples. Of the four DOC measurements methods the most common approach was by combustion, with 90% of all DOC concentrations measured using this approach. As such, in order to continue measuring DOC concentrations in terrestrial permafrost ecosystems using the most consistent approach we suggest using 0.45 µm filters, extracting pore water via some type of sucking device or soils via leaching, and using a combustion based method to determine DOC concentrations.

**Data availability**

All data will be made freely and publicly available on an online repository prior to publication.

**Author contributions**

LH, DK, and LT designed and planned the systematic review approach; LH built the database. LH and DK analyzed the data; LH wrote the manuscript draft; DK and LT edited and reviewed the manuscript.

**Competing interests**

The authors declare that they have no conflict of interest.

**Acknowledgements**

We thank Konstantinos Vaziourakis, Mona Abbasi, Elizabeth Jakobsson, Marloes Groeneveld, Sarah Shakil, and Jeffrey Hawkes for helpful discussions throughout the development and writing of this manuscript.

**Financial support**

This work was supported by the Knut and Alice Wallenberg Foundation.
References (in text)


https://doi.org/10.1029/1999GB900083


https://doi.org/https://doi.org/10.1016/j.gca.2012.07.015


https://doi.org/10.1080/1364557032000119616


https://doi.org/10.1029/2018JG004500


Manies, K. L., Jones, M. C., Waldrop, M. P., Leewis, M. C., Fuller, C., Cornman, R. S., &


Landscape Functions (pp. 27–67). https://doi.org/10.1007/978-3-030-71330-0_3


forest: Sources, chemical characteristics, and biodegradability. *Ecosystems*, 10(8), 1323–1340. https://doi.org/10.1007/s10021-007-9101-4


**Studies used to generate database**


*NATURE COMMUNICATIONS*, 11(1). https://doi.org/10.1038/s41467-020-15250-8


unaccounted carbon, nutrient and metal sources. Chemosphere, 266, 128953.
https://doi.org/10.1016/j.chemosphere.2020.128953

https://doi.org/10.5194/essd-8-439-2016

https://doi.org/10.1002/2018JG004461

https://doi.org/10.1016/j.scitotenv.2020.137885


https://doi.org/10.1088/1748-9326/abd971

https://doi.org/10.1016/j.scitotenv.2019.135415

https://doi.org/10.1016/j.scitotenv.2016.04.195


https://doi.org/10.1016/j.soilbio.2019.05.009


Voigt, C., Lamprecht, R. E., Marushchak, M. E., Lind, S. E., Novakovskiy, A., Aurela, M., … Biasi, C. (2017). Warming of subarctic tundra increases emissions of all three important


