Geoderma 378 (2020) 114635

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

# Origin of dissolved organic matter in the Harz Mountains (Germany): A thermally assisted hydrolysis and methylation (THM-GC–MS) study

Joeri Kaal[a,b,*], César Plaza[b], Klaas G.J. Nierop[c], Marta Pérez-Rodríguez[a], Harald Biester[a]

[a] Institut fur Geo-Ökologie, Abt. Umweltgeochemie, Technische Universität Braunschweig, Braunschweig, Germany
[b] Instituto de Ciencias Agrarias (ICA), Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
[c] Department of Earth Sciences and Geolab, Faculty of Geoscience, Utrecht University, Utrecht, The Netherlands

ARTICLE INFO

Handling Editor: Ingrid Kögel-Knabner

*Keywords:*
Harz Mountains
Dissolved organic matter
THM-GC–MS
Source assessment

ABSTRACT

Environmental change is increasing the concentration of dissolved organic matter (DOM) in catchments of the Northern Hemisphere. This study aims to assess the causes of high DOM concentrations in streams and reservoirs of the Harz National Park (Germany), by means of molecular characterization using thermally assisted hydrolysis and methylation (THM-GC–MS). In order to formulate proxies of the prevailing origin of the numerous THM products of polyphenols, carbohydrates, proteins, aliphatic macromolecules, resins and other DOM precursors, we created a reference sample set of potential sources (spruce, birch, blueberry, heather, peat moss, soils) from the area. Besides solid-state reference samples (bulk organic matter; BOM) we obtained and analyzed their leachates (water-extractable OM; WEOM). Finally, an existing THM-GC–MS dataset of the DOM from the Oder river, which crosses the boundary between peat and forest biomes in the Harz, was extended and explored chemometrically using Principal Component Analysis (PCA) to test the proxies for stream DOM assessment. The results show large differences between BOM and WEOM, which suggests that the solid-to-leachate transition is highly selective or significantly alters the major biomolecular constituents. THM compounds that tend to be more abundant in WEOM than in BOM are G-type phenolic compounds (1,2-dimethoxybenzenes, from lignin and tannin), nitrogen-containing moieties and benzene carboxylic acids, whereas WEOM is depleted in products of polysaccharides, syringyl lignin and aliphatic macromolecules (cutin and suberin). The lignin fingerprint of the WEOM also differs significantly from that of BOM, being depleted in the vast majority of the typical products of macromolecular lignin (G7, G8, G14, G15) and enriched in the acid moiety (G6, predominantly from vanillic acid), especially for spruce wood. THM chromatograms of DOM from the forest section of the Oder show an extraordinary abundance of G6, most probably from spruce-derived lignin. This may indicate a major role of DOM released from decaying spruce logs and forest soils. The results highlight both the potential and the pitfalls associated with source identification of DOM using THM-GC–MS.

## 1. Introduction

The increasing concentrations of dissolved organic matter (DOM) in the streams and reservoirs of the Harz Mountains in Central Germany (Broder et al., 2017; Broder and Biester, 2015), and in the surface waters of temperate regions of the Northern Hemisphere in general, are largely ascribed to accelerated decay of organic matter from catchment soils, in particular peat deposits, due to global warming or reduced acid rain (Freeman et al., 2001; Vogt, 2003; Eikebrokk et al., 2004; Worrall et al., 2004). In the Harz National Park, increasing temperatures and recent heat waves are having a devastating effect on the main forest biome, as the deterioration of the physiological fitness of Norway spruce (*Picea abies* L. Karst) becomes exacerbated by the bark beetle (*Ips typographus*) (Overbeck and Schmidt, 2012; Knolle and Wegener, 2019). As forest management does not include removal of dead trees, these form a large potential source of DOM. Furthermore, forest clearance is known to accelerate DOM release from litter even if the woody debris would not have been left on-site, because higher soil surface temperatures boost microbial activity (Kalbitz et al., 2004). In short, the DOM increase in the streams and reservoirs of the Harz National Park can be caused by a combination of increased necromass stocks, increased forest soil microbial activity and accelerated decomposition of the peatlands at the upper slopes and flats of the mountains, but the relative importance of these factors is unknown.

* Corresponding author at: Institut fur Geo-Ökologie, Abt. Umweltgeochemie, Technische Universität Braunschweig, Braunschweig, Germany.
*E-mail addresses:* j.kaal@tu-braunschweig.de, joeri@pyrolyscience.com (J. Kaal), cesar.plaza@csic.es (C. Plaza), K.G.J.Nierop@uu.nl (K.G.J. Nierop), m.perez-rodriguez@tu-bs.de (M. Pérez-Rodríguez), h.biester@tu-bs.de (H. Biester).

https://doi.org/10.1016/j.geoderma.2020.114635
Received 24 April 2020; Received in revised form 29 June 2020; Accepted 27 July 2020
Available online 04 August 2020

DOM is one of the most complex and challenging types of natural organic substances (Hedges et al., 2000). Despite of numerous studies on the molecular composition of DOM (reviews by Kalbitz et al., 2000; Nebbioso and Piccolo, 2012), this knowledge is insufficient (Sleighter et al., 2014). Analytical pyrolysis techniques, such as thermally assisted hydrolysis and methylation (THM-GC–MS), have frequently been applied to study the molecular composition of DOM powders (e.g., van Heemst et al., 2000; Frazier et al., 2005; Bardy et al., 2011). For THM, a reagent (often tetramethylammonium hydroxide; TMAH) is added to the sample to produce simultaneous hydrolysis/derivatization, which improves (in comparison with conventional pyrolysis, i.e. Py-GC–MS) the structural information of long-chain aliphatic macromolecules, such as cutin or suberin, and of polyphenolic materials, such as lignin and tannin (Challinor, 1989; Clifford et al., 1995; Hatcher et al., 1995; Del Río and Hatcher, 1998; He et al., 2020). THM-GC–MS has been usually applied to evaluate qualitatively the main DOM components, and with focus on lignin, whereas the variations in relative proportions of THM products have seldom been evaluated numerically (Frazier et al., 2005; Jeanneau et al., 2014, 2015; Denis et al., 2017; Jiang et al., 2017; Gandois et al., 2019). This numerical evaluation is essential to define and validate proxies of DOM sources and transformation. Many common proxies used for particulate organic matter are not or may not be valid for DOM, because of the profound influence of the phase transfer from particulate to dissolved substances (Hernes et al., 2007; Spencer et al., 2012; Matiasek and Hernes, 2019).

Beyond the level of a general fingerprinting from THM chromatograms, the use of proxies of DOM sources is not straightforward. Polyphenols such as lignin and tannin tend to maintain the substitution pattern of the aromatic functional groups. Hence, for lignin, the *p*-hydroxyphenyl, guaiacyl and syringyl units are methylated to 4-methoxybenzenes (from hereon, H-type products), 3,4-dimethoxybenzenes (G-type) and 3,4,5-trimethoxybenzenes (S-type), respectively, and numerous side-chain configurations give rise to an array of THM products (Mulder et al., 1992; Clifford et al., 1995; Del Río et al., 1998; Vane et al., 2001). The balance between the main units provides information of lignin sources (for instance syringyl groups are not metabolized by gymnosperms; Higuchi et al., 1977). Further complexity of lignin is related to acylation (binding through ester groups) by moieties other than the three basic building blocks, such as *p*-hydroxybenzoates, *p*-coumarates and ferulates (Lu et al., 2015). Lignin is often bound to polysaccharides by bridges of this kind, forming the lignin-carbohydrate complex. Lignin is not the only macromolecular source of methoxybenzenes. Tannins in vascular plants occur as condensed and hydrolysable ones. Hydrolysable tannins, which are produced only by angiosperms, are based on gallic acid moieties (trihydroxybenzenes) esterified to a central carbohydrate unit and form mostly S-type products after THM, whereas condensed tannins, metabolized by both gymno- and angiosperms, are built of monomers with a phloroglucinol (1,3,5-trihydroxybenzene) A-ring, which yield various 1,3,5-trimethoxybenzenes upon THM (Phl-type products); and a catechol (1,2-dihydroxybenzene, in procyanidin condensed tannin) or pyrogallol (1,2,3-trihydroxybenzene, in prodelphinidins) B-ring, which produce G- and S-type products upon THM, respectively (Galletti et al., 1995; Garnier et al., 2003; Nierop et al., 2005). Hence, THM of lignin and tannin yield partially overlapping G- and S-type methoxybenzenes. They can be distinguished by using labelled derivatization agents such as $^{13}$C TMAH (Filley et al., 1999; Nierop and Filley, 2008; Klotzbücher et al., 2013), as this implies a mass difference between the THM products of guaiacol and catechol, and of syringol and pyrogallol. However, $^{13}$C labeling complicates the already problematic identification of non-phenolic THM products because comparison with literature and MS databases becomes more cumbersome (giving rise to even larger numbers of unidentified products). Another approach is to use Py-GC–MS (without methylation) to estimate the balance between the two main polyphenols using the ratio of catechol to guaiacol. For instance, a sample that is productive of guaiacol upon pyrolysis and 3,4-dimethoxybenzoic acid methyl ester (G6) upon THM would have G6 that originates predominantly from lignin, whereas a dominance of catechol with Py-GC–MS would imply that THM product G6 originates from tannin or related compounds. Likewise, dominance of G6 upon THM-GC–MS would indicate that unsubstituted guaiacol from Py-GC–MS is a decarboxylation product of vanillic or protocatechiuc acid (Mulder et al., 1992). Hence, by comparing the methods, one can generate a significant body of information that would not be feasible if only one of the two methods is applied. Besides polyphenols, other biopolymers that may generate methoxybenzenes upon THM are polysaccharides, which form 1,4-dimethoxybenzene and 1,2,4-trimethoxybenzene (Fabbri and Helleur, 1999). Albeit that the latter is also formed upon THM of tannin (Nierop et al., 2005), these products are usually only minor products of polyphenols which implies limited interference. Many studies on THM of carbohydrates have been performed (Fabbri and Helleur, 1999; Schwarzinger et al., 2002), and proxies of polysaccharides sources in DOM have been proposed (e.g. Jeanneau et al., 2014), but due to the profound rearrangement of carbohydrate structures during THM and the numerous possible isomeric/stereomeric forms, many products remain unidentified. Long-chain methylene groups based on alkanoic acid moieties are efficiently transmethylated to fatty acid methyl esters (FAMEs) by THM. The substitution patterns of these FAMEs is different for different polymethylene sources such as cutin in plant cuticles (having mid-chain-hydroxy substitution of mostly $C_{16}$- and $C_{18}$-FAMEs), suberin in bark and root materials (forming long-chain FAMEs in the $C_{20}$-$C_{34}$ range, including ω-methoxylated FAMEs and diacid dimethyl esters; DAMEs; Nierop and Verstraten, 2004) and bacterial FAMEs have relatively large proportions of odd-numbered and branched FAMEs such as *iso*- and *anteiso*-$C_{15}$ FAMEs. Suberin has cross-linkages of methylated caffeic and/or ferulic acid (G18; Riley and Kolattukudy, 1975; Filley et al., 2006), cutin has aromatic domains mainly of *p*-coumaric acid (P18; Riley and Kolattukudy, 1975) and all three sources of P18 and G18 have been detected in sporopollenin (Wehling et al., 1989; Nierop et al., 2019), creating some overlap with especially graminoid lignin THM fingerprints. The non-hydrolysable cutan may have Phl-type linkages between the fatty acid moieties (Nip et al., 1986; McKinney et al., 1996; Boom et al., 2005). Unsubstituted FAMEs (especially $C_{16}$- and $C_{18}$-FAMEs) are formed by any source of fatty acids including free/esterified oils, fats and wax esters, and are therefore of little diagnostic value. Finally, terpenoids usually maintain their molecular structure during THM reactions and are useful markers of plant resins in DOM (e.g. van den Berg, 2003).

Clearly, interpretation of THM data demands a rationale behind each product's source identification, and sometimes for each sample differently, and this process can be made easier by (1) taking ecosystem-specific parameters into account, such as vegetation patterns, on the basis of a reference sample set (limiting the range of potential sources), and (2) by improving our scarce knowledge on the different THM fingerprints of solid biological samples and the DOM structures that they may release to the environment by the examination of leachates.

This work presents THM-GC–MS of bulk organic matter (BOM) and water leachates (water-extractable OM; WEOM) of biological materials from the peatland and forest environments of the Harz Mountains, and a limited number of WEOM samples from peat and mineral soils. Semi-quantitative data from BOM and WEOM is then used to improve the source assessment of DOM in surface waters (after expanding the THM dataset of samples from the Oder river; Kaal et al., 2017) and propose a series of proxies for DOM characterization.

## 2. Methods and materials

### 2.1. Selection of potential source materials

The catchments of the Oder (source at 51°46′22″N, 10°33′53″E) and

**Table 1**
List of biological sources studied, estimated dissolved organic carbon (DOC) yield, and the compositions of hypothetical WEOM mixtures A-E (numbers indicate proportions of each source; sum 100%).

| Material | Source | DOC (mg TOC/g sample) | Mix A | Mix B | Mix C | Mix D | Mix E |
|---|---|---|---|---|---|---|---|
| Spruce wood (living) | *Picea abies* | 5.01 | – | – | 11 | 20 | 30 |
| Spruce wood (log 1) | *Picea abies* | 2.22 | – | – | 11 | 20 | 30 |
| Spruce wood (log 2) | *Picea abies* | 3.02 | – | – | 11 | 20 | 30 |
| Spruce bark | *Picea abies* | 4.41 | – | – | – | – | 3 |
| Spruce needles | *Picea abies* | 15.81 | – | – | – | – | 3 |
| Spruce cones | *Picea abies* | 4.84 | – | – | – | – | 3 |
| Birch mixed litter | *Betula* ssp. | 7.58 | – | – | 11 | 20 | – |
| Blueberry mixed litter | *Vaccinium myrtillus* | 18.00 | – | – | 11 | 20 | – |
| Heather mixed litter | *Calluna vulgaris* | 20.82 | – | 25 | 11 | – | – |
| Grasses mixed litter | Poaceae | 4.38 | – | 25 | 11 | – | – |
| Sedges mixed litter | Cyperaceae | 8.71 | – | 25 | 11 | – | – |
| Peat moss mixed litter | *Sphagnum* ssp. | 3.03 | 33 | 25 | 11 | – | – |
| Podzol Ah | forest | 1.73 | – | – | – | – | – |
| Podzol E | forest | 0.31 | – | – | – | – | – |
| Podzol Bhs | forest | 0.23 | – | – | – | – | – |
| Mollisol AE | forest | 1.91 | – | – | – | – | – |
| Mollisol Ah | forest | 1.06 | – | – | – | – | – |
| Mollisol C | forest | 0.13 | – | – | – | – | – |
| Peat 1 (290–295 cm depth) | Odersprungmoor | 0.62 | 33 | – | – | – | – |
| Peat 2 (295–300 cm depth) | Odersprungmoor | 0.86 | 33 | – | – | – | – |

Ecker (51°47′19″N, 10°35′09″E) streams in the Harz National Park are covered by two main landscape units: peatland and forest. Peat moss (*Sphagnum* ssp.) and heather growing on drier hummocks are dominant in the uphill peatlands, whereas the forest is spruce-dominated with abundant blueberry understory and patches of birch and other angiosperm trees.

The plant materials (potential primary organic carbon sources), i.e. blueberry (*Vaccinium myrtilus*; branches with leaves and bark), Norway spruce (*Picea abies*; branch wood, needles, cones, bark), sedges (unidentified, leaves), Poaceae (unidentified Ecker floodplain herbs), *Sphagnum* (whole plant from Torfhausmoor), common heather (*Calluna vulgaris)* and birch (*Betula* cf. *B. pendula*; mixed litter) (Table 1), and spruce trunk xylem materials in an intermediate and advanced stage of decomposition, were taken from the Ecker catchment. In addition to BOM and WEOM from these biological samples, WEOM obtained from different horizons of mineral soils from the Ecker catchment was included, i.e. 1) an incipient Podzol (Podsol-Braunerde according to the German Soil Classification; Krasilnikov et al., 2009) under spruce vegetation (51°47′56″N, 10°33′34″E), of which only the topsoil (Ah horizon) provided informative THM-GC–MS data, and 2) a Mollisol taken near the junction of the Abbe and Ecker streams (51°48′31″N, 10°34′10″E), of which topsoil (1A(E)-2Ah) and parent material (2C horizon) were studied. The Mollisol probably corresponds to what was described as “peaty riparian soil” by Broder and Biester (2017). Finally, two peat samples from the Odersprungmoor (51°46′25″N, 10°33′46″E) were studied. This is a *Sphagnum*-dominated peatland (Histosol) and the samples used were taken from the deepest section of the deposit (290–295 cm and 295–300 cm; Blome, 2019) to obtain information of strongly evolved peat material. The plant and peat samples were dried for several days at 50 °C and shredded mechanically to create the BOM. Soil samples were dried only (50 °C).

The Oder river connects the peatland (Odersprungmoor) to the Oderteich reservoir along a transect of about 2.5 km (see map in Broder and Biester, 2017). Due to headward erosion, the source of the Oder is within the peatland. Furthermore, the Oder presents a good opportunity to identify signals of peatland and forest systems due to the proximity of the headwater to the forest boundary (400 m). The DOM in the Oder was studied previously (Broder and Biester, 2015; Kaal et al., 2017). The set consists of nine samples collected during summer baseline discharge, from the headwater (two samples) and forest environments (seven samples, two of which from small tributaries of the Oder). DOM was isolated by filtration (0.45 μm) and freeze-drying of 2 L stream water without use of chemicals (Kaal et al., 2017). The dissolved organic carbon content of the water ranged between 5 and 25 mg/L (Kaal et al., 2017).

### 2.2. Isolation of WEOM

The WEOM was obtained by single batch water extraction of BOM and soil samples (Don and Kalbitz, 2005), using 0.4–0.8 g dry BOM in 25 mL distilled water in 50 mL polyethylene tubes. The extraction was performed on a horizontal shaking device for 24 hrs. The extracts were sonicated, centrifuged and filtrated through 0.45 μm filters, as described in Kaal et al. (2020). Two mL of the extracts were separated for measuring total organic carbon (TOC) content, using a Shimadzu TOC-$V_{CSH}$ analyzer. The WEOM yield was estimated from the amount of TOC per g of sample (mg TOC/g sample; Table 1). The remaining of the WEOM solutions were evaporated in the dark at 40 °C to obtain the solid WEOM material. The Py-GC–MS fingerprints of the BOM and WEOM samples are discussed in Kaal et al. (2020). The fact that WEOM was obtained not only by filtration of the water extracts (leachate) of undisturbed litter or soil material, but included shredding and sonication, implies that the WEOM might include compounds that would not be transferred to the aqueous phase in the natural environment (Zsolnay, 1996; Chantigny, 2003).

The residues after WEOM extraction were also rewetted and incubated for 20 days at 20 °C in the dark (*cf.* Moore and Dalva, 2001), to obtain information on THM product distribution of the potential sources after microbial alteration. Next, the residues were flushed with distilled water for 24 h, centrifuged, filtered through 0.45 μm and dried. For those samples of which sufficient material was obtained, the chromatograms proved similar to those of the WEOM, with the exception of the sedge sample (Supplementary Material S1). The THM-GC–MS relative proportions datasets of both the WEOM and the extracts obtained after incubation of the residues are provided in Supplementary Material S2.

### 2.3. Thermally assisted hydrolysis and methylation (THM-GC–MS)

The THM-GC–MS analyses were performed using a CDS Pyroprobe coupled to an Agilent 6890/5975 GC–MS system. The TMAH (25% in water, from Sigma-Aldrich) was added to sample-containing quartz tubes, assuring that the solution completely soaked the sample and quartz wool, and then inserted into the pyrolysis interface after 30 min.

The setpoint temperature of the Pyroprobe was 650 °C, maintained for 10 s. The GC temperature increased from 70 °C (4 min) to 325 °C (3 min) at a rate of 20 °C/min (runtime 20 min). The GC was equipped with a HP-5MS column, the GC inlet was in 1:10 split mode and the carrier gas was He. The mass spectrometer operated in EI mode (70 eV) scanning in the *m/z* range 50–500 (full scan mode) (Kaal et al., 2017).

### 2.4. Data analysis

The major products in the BOM and WEOM chromatograms were denoted creating a list of 196 products, which were semi-quantified on the basis of peak areas of characteristic *m/z* fragments. Relative proportions of each compound were calculated as % of total quantified peak area (TQPA) in a given sample (sum of all products 100%). The relative proportions of individual compounds, compound groups and the ratios between different (sets of) products, are based on the % TQPA data.

The THM chromatograms of the DOM from the Oder river were originally evaluated on the basis of 47 compounds (Kaal et al., 2017). This dataset was extended, quantifying the remaining 149 compounds, to facilitate comparison between Oder DOM with the BOM and WEOM fingerprints. For the Oder DOM, the relative proportions data (% TQPA) was recalculated to 100%.

The dataset of relative proportions of the Oder DOM was examined by a chemometric approach, using principal components analysis (PCA; using Tanagra software, without rotations), aiming to create proxies of downstream changes in DOM composition. Denis et al. (2017) also used PCA to study changes in DOM composition related to hydrological conditions using THM-GC–MS.

A series of hypothetical mixtures of THM fingerprints of the WEOM samples were calculated (Table 1) to examine whether results obtained from WEOM can account for the THM compound distributions of the DOM in the Oder. This reconstruction gives an idea of the imprint of the potential sources on environmental DOM composition. The calculations are based on the individual peak proportions from the WEOM samples multiplied by the % of the hypothetical admixture (for each sample and for each compound). Five hypothetical WEOM mixtures were calculated, reflecting a shift from peatland to forest sources: Mix A comprised WEOM of *Sphagnum* and the two peat samples (in a 1:1:1 ratio) and represents the peat moss environment. Mix B was calculated from contributions of all selected peatland vegetation members (heather, grass, sedge and peat moss; each 25%). Mix C has all sources combined (except for spruce bark, needles and cones). Mixtures D and E were compiled on the basis of signals from the forest environment (Mix E mainly spruce wood-derivatives; Mix D spruce, birch, blueberry).

## 3. Results and discussion

### 3.1. Potential source materials of DOM

#### 3.1.1. Spruce wood (living and dead xylem)

The fresh BOM sample of spruce wood produces mainly methoxybenzenes from polyphenols (30%; Fig. 1) and carbohydrates from polysaccharides (56% of TQPA; Fig. 2b) upon THM. The carbohydrates include a large peak of trimethyllevoglucosan from intact cellulose (28%; Supplementary Material S2). The G-type products account for 29%, including G4 (G-aldehyde, 7.5%; Fig. 1b) and G6 (G-acid, 7.6%), a lignin dimer (2.9%) and numerous other products (Table 2). The H-type (1.3%; Fig. 1a), S-type (0.2%; Fig. 1c) and Phl-type (Fig. 1d) methoxybenzenes (0.1%) are scarce. Characteristic syringyl lignin products (e.g. S7, S8, S14, S15) were not detected, confirming absence of syringyl lignin in gymnosperms. Hence, the detected S-type products (S1 and S2) should be largely ascribed to B-rings, and the Phl-type products to A-rings (a presence of cutan in wood is unlikely) in condensed tannins. The methylated benzene carboxylic acids (BCA; 0.4%; Fig. 2e), FAMEs (1.3%; Fig. 2a), N-compounds (0.4%; Fig. 2c) and terpenoid products (0.3%; Fig. 2d) are scarce. This THM fingerprint reflects the prevalence of lignocellulose in spruce wood.

The picture is very different for the WEOM obtained from the living spruce wood, with dominance of G6 (72% of TQPA). The total sum of G-type products is 84% (Fig. 1b). The very low ratio between catechol and guaiacol from Py-GC–MS ($Py_{cat/gua}$ 0.01; Fig. 3a) – guaiacol is by far the largest peak from Py-GC–MS (Kaal et al., 2020) – indicates that THM product G6 formed by methylation of vanillic acid moieties in lignin. Vanillic acid is usually associated with oxidation of the α-carbon of the propanoid side-chain of lignin (G6/G4 is often used as a proxy of lignin oxidation; Hedges et al., 1988). It is possible that some abiotic or biological oxidation occurred during the WEOM production procedure (e.g. during shredding, sonication or evaporation), but oxidation at α-carbon also occurs during tree growth (Rencoret et al., 2011). Even though such groups represent only a small part of the lignin in BOM, their preferential release to the leachate could explain predominance of G6 in the WEOM. An alternative explanation of the extraordinary dominance of G6 is deacylation, but acylation of lignin with phenolic acids usually involves *p*-hydroxybenzoates with less, if any, vanillic and sinapic moieties (Lu et al., 2015; Del Río et al., 2020). Furthermore, 2D NMR analyses of *Picea abies* lignin did not reveal intense signals of acylated vanillic acid (Rencoret et al., 2009). It is concluded that the G6 from spruce wood probably corresponds to *in situ* modifications of guaiacyl lignin. Compounds such as *threo*- and *erythro*-1-(3,4-dimethoxyphenyl)-3-methoxyprop-1-ene (G14 and G15; both 1.0%; Supplementary Material S2) indicate that some fragments of the lignin backbone are released to the WEOM in a relatively unaffected state (Clifford et al., 1995), probably from β-*O*-4 interunit linkages (Kuroda et al., 2002). Other derivatives of phenolic groups are H-type (2.1%; Fig. 1a), S-type (0.4%; Fig. 1c) and Phl-type (0.03%; Fig. 1d). Hence, the WEOM does not have a significant contribution of condensed tannin or other (detectable) polyhydroxy-aromatic compounds, and the extremely high G6/G4 ratio of > 50 (compared to 1.0 for the BOM; Fig. 4d) cannot be associated with a contribution of dihydroxybenzoic acid groups as was observed by Klotzbücher et al. (2013) for leachates of spruce needles (see below). Instead, it is a feature of the leachate of fresh spruce wood and G6/G4 should thus not be used to study biological alteration state of lignin in DOM. This is in agreement with Hernes et al. (2007), Spencer et al. (2012) and Godin et al. (2017) who also called for caution in the interpretation of G6/G4 ratio of DOM samples. Carbohydrate products account for only 3.3% (Fig. 2b), indicating that relatively few detectable cellulose-derived compounds are released to the WEOM.

The BOM and WEOM fingerprints of the decomposed spruce log 1 (Table 1) are similar to those of fresh wood (e.g. BOM 67% carbohydrates, WEOM 60% G6). For WEOM, the proportions of carbohydrates (5.3%), FAMEs (5.9%) and terpenoids (2.8%) are slightly higher than for fresh wood (Fig. 2). For the more profoundly decomposed log 2, the BOM indicates significant decay, in particular the much lower proportion of carbohydrate products (11%, against 56–67% for the other spruce wood BOM) and a much higher proportion of G-type products (71%, against 21–28% for the other spruce wood BOM). However, the WEOM of decomposed log 2 was similar to that of the other samples (66% of G6). Hence, the difference in WEOM composition between fresh and decomposed wood is small; both are characterized by vanillic acid dominance. Don and Kalbitz (2005) found that the leaching of aromatic substances from Norway spruce wood and needles increased along the course of a litterbag experiment, arguing that early decomposition vastly increases the release of lignin-derived moieties. Even though it remains to be addressed whether truly unaffected fresh wood can also release these lignin moieties (whether or not the WEOM isolation procedure caused significant alterations of lignin), the fact that lignin-derived moieties are released readily from fresh wood suggests they will also be released in the natural environment.

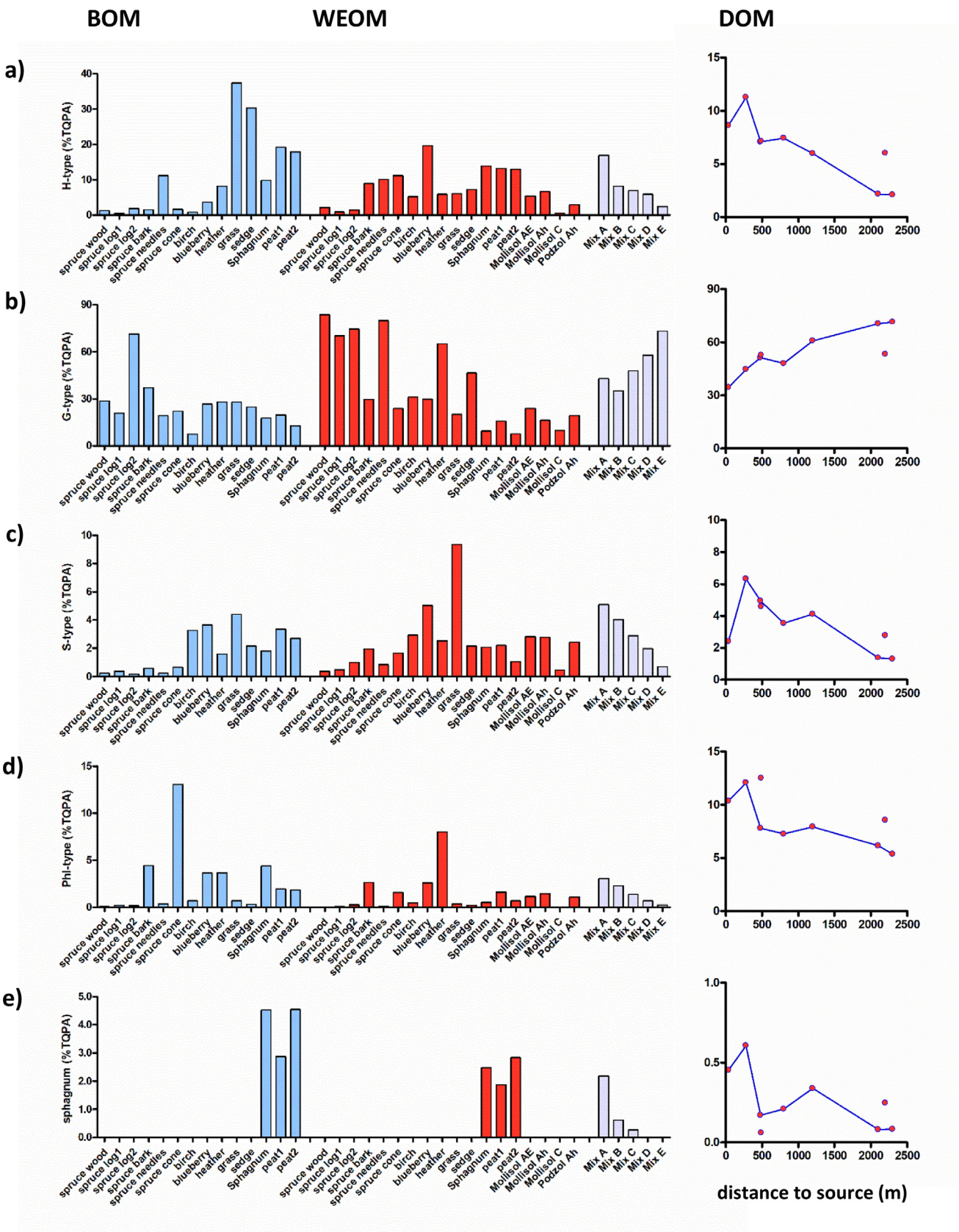


**Fig. 1.** Relative proportions of subgroups of methoxybenzene-based THM products. a) 4-methoxybenzenes (H-type); b) 3,4-dimethoxybenzenes (G-type); c) 3,4,5-trimethoxybenzenes (S-type); d) 1,3,5-trimethoxybenzenes (Phl-type); e) sum of sphagnum acid products (4-isopropenylmethoxybenzene and methyl esters of 4-methoxyphenylbutenoic acids), as % of total quantified peak area (% TQPA). Data corresponds to bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

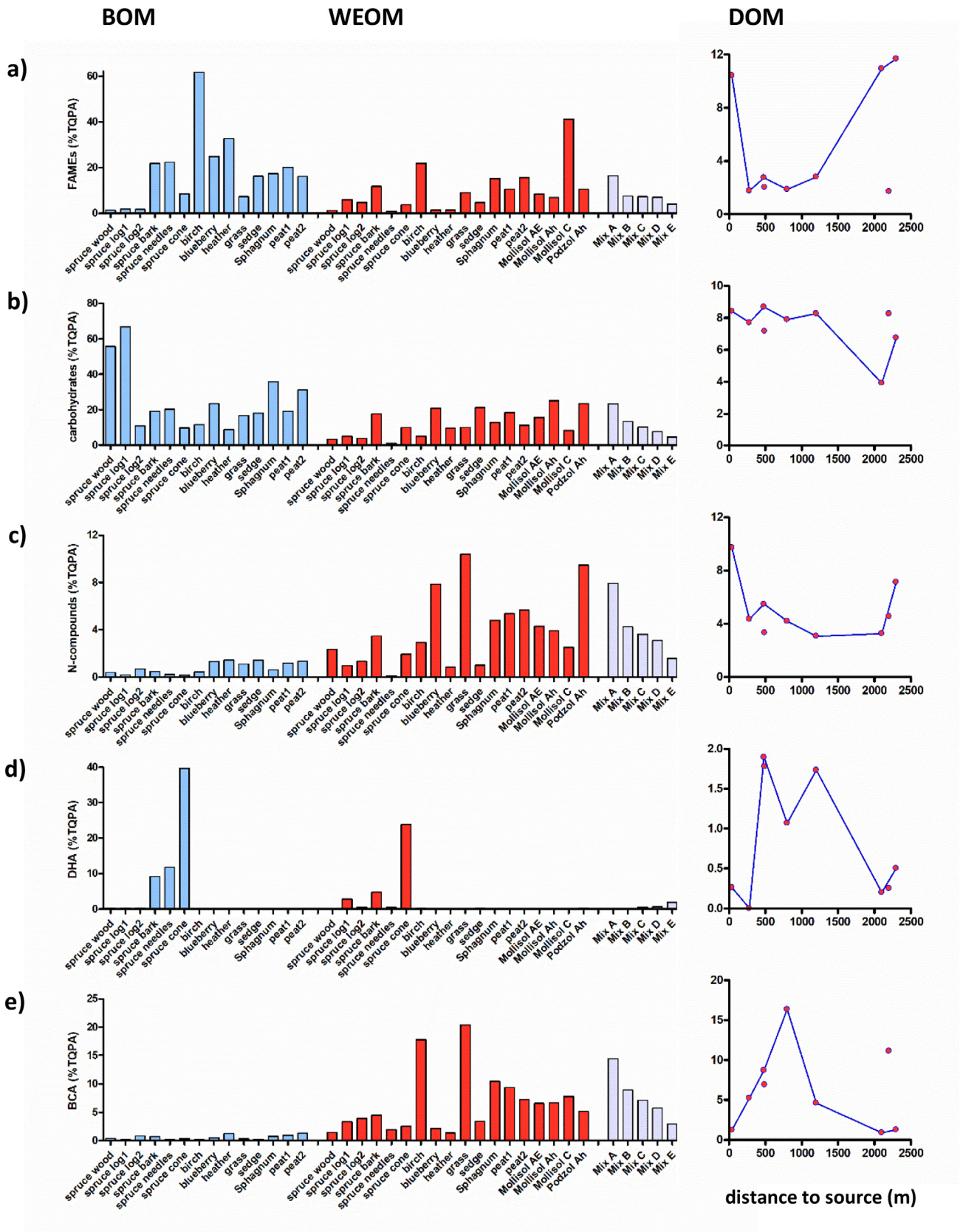


**Fig. 2.** Relative proportions of THM product groups. a) fatty acid methyl esters (FAMEs); b) carbohydrate products; c) nitrogen-containing products (N-compounds); d) dehydroabietic acid derivatives and retene (DHA); e) benzenecarboxylic acid methyl esters (BCA), as % of total quantified peak area (% TQPA). Proportions of diacids (DAMEs) are not shown (provided in Supplementary Material S2). Data corresponds to bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 2**
List of THM products of the methoxybenzene structures (excluding carbohydrate- and plastic-derived methoxybenzenes). ME = methyl ester, H = *p*-hydroxyphenyl, G = guaiacyl, S = syringyl, Phl = phloroglucionol-based (note that dihydroxy- and trihydroxybenzenes contribute to G and S products).

| THM product | Type | Code |
|---|---|---|
| 4-ethylmethoxybenzene | H | C2-P1 |
| 4-methoxystyrene | H | P3 |
| 4-methoxyacetophenone | H | P5 |
| 2/3-methoxybenzoic acid ME | H | P6 isomer |
| 4-methoxybenzoic acid ME | H | P6 |
| 1-(methoxybenzene)-2-methoxyethylene | H | P7 |
| *cis*-3-(4-methoxyphenyl)-3-propenoate | H | P17 |
| *trans*-3-(4-methoxyphenyl)-3-propenoate | H | P18 |
| 4-methoxybenzeneacetic acid ME | H | P24 |
| 1,2-dimethoxybenzene | G | G1 |
| guaiacol | G | G1* |
| 3,4-dimethoxytoluene | G | G2 |
| 4-vinyl-1,2-dimethoxybenzene | G | G3 |
| 3,4-dimethoxybenzaldehyde | G | G4 |
| 3,4-dimethoxyacetophenone | G | G5 |
| vanillic acid ME | G | G6* |
| G6 isomer | G | G6 isomer |
| 3,4-dimethoxybenzoic acid ME | G | G6 |
| $C_1$-dimethoxybenzoic acid ME | G | C1-G6 |
| 2-(3,4-dimethoxyphenyl)-1-methoxyethylene | G | G7 |
| 2-(3,4-dimethoxyphenyl)-1-methoxyethylene | G | G8 |
| 3-methoxyprop-1-ene-3,4-dimethoxyphenyl | G | G10/G11 |
| 3-(3,4-dimethoxyphenyl) propanoic acid ME | G | G12 |
| *trans*-1-(3,4-dimethoxyphenyl)-3-methoxyprop-1-ene | G | G13 |
| *threo*-1-(3,4-dimethoxyphenyl)-1,2,3-trimethoxypropane | G | G14 |
| *erythro*-1-(3,4-dimethoxyphenyl)-1,2,3-trimethoxypropane | G | G15 |
| *cis*-1-(3,4-dimethoxyphenyl)-1,3-dimethoxy-1-propene | G | G16 |
| *trans*-3-(3,4-Dimethoxyphenyl)-3-propenoate | G | G18 |
| 1-(3,4-dimethoxyphenyl)-1-propene | G | G21 |
| 3,4-dimethoxybenzeneacetic acid ME | G | G24 |
| G-dimer (e.g. tetramethoxystilbene compound) | G | G dimer |
| 1,2,3-trimethoxybenzene | S | S1 |
| 3,4,5-trimethoxytoluene | S | S2 |
| 3,4,5-trimethoxybenzaldehyde | S | S4 |
| 3,4,5-trimethoxybenzoic acid ME | S | S6 |
| *cis*-1-(3,4,5-trimethoxyphenyl)-2-methoxyethylene | S | S7 |
| *trans*-1-(3,4,5-trimethoxyphenyl)-2-methoxyethylene | S | S8 |
| *cis*-1-(3,4-trimethoxyphenyl)-1,3-dimethoxy-1-propene | S | S16 |
| 1,3,5-trimethoxybenzene | Phl | – |
| 2-methyl-1,3,5-trimethoxybenzene | Phl | – |
| 2-ethyl-1,3,5-trimethoxybenzene | Phl | – |

### 3.1.2. Other spruce materials (bark, needle, cone)

The BOM of spruce bark material is prolific of G-type structures (37%; Fig. 1b) and FAMEs (22% of TQPA; Fig. 2a). The FAMEs are mainly $C_{22}$ (9.3%) and $C_{24}$ (5.3%), and long-chain ω-methoxy-FAMEs and $C_{20}$- and $C_{24}$-DAMEs (Supplementary Material S2) were also detected, which reflects the high suberin content ($C_{16}$ and $C_{18}$ FAME account for < 1% of TQPA). This suberin is probably also the source of at least part of the caffeic and/or ferulic acid ME (G18; 0.9% *vs*. 0.1–0.2% for spruce wood BOM). Another feature of the bark BOM of spruce is the abundance of methylated dehydroabietic acid and similar products (DHA; 9.3%; Fig. 2d), indicative of diterpenoid resin content (Fig. 2). The Phl-type products account for 4.4% (Fig. 1d), which indicates a significant condensed tannin A-ring content and indicates that an unknown proportion of the G-type compounds originate from the B-ring in procyanidin condensed tannins instead of lignin.

The G-type products (40% of TQPA) are abundant among the THM products of the bark-derived WEOM (Fig. 1b). Contrary to WEOM from wood, there is considerable diversity in G-type products (G6 accounts for only 13%). The S- (1.9%; Fig. 1c) and Phl-type (2.6%; Fig. 1d) products probably reflect some condensed tannin B- and A-rings, respectively, but lignin prevails over tannin as the dominant source of G-type products ($Py_{cat/gua}$ = 0.05; Fig. 3a). Carbohydrates (19%; Fig. 2b), diterpenoids (5.0%; Fig. 2d), N-compounds (3.4%; Fig. 2c) and FAMEs (11%; Fig. 2a) are more abundant than in chromatograms of spruce wood-derived WEOM (Fig. 2). Regarding the FAMEs, the dominant products are $C_{16}$- (5.1%) and $C_{18}$-FAME (3.3%), whereas the long-chain FAMEs account for < 0.5% and long-chain DAMEs were not detected. Hence, the FAME signature differs radically from that of the BOM bark sample, probably due to release of non-macromolecular fatty acids (e.g. oils) to the WEOM and not the phase transfer of the suberin polyester.

The BOM chromatogram of the spruce needles contained FAMEs (22%; Fig. 2a), DAMEs (12%; not shown), terpenoids (12%; Fig. 2d), G-type (19%; Fig. 1b) and H-type products (11%; Fig. 1a) as the main groups. Of the FAMEs, $C_{14}$-$C_{18}$ unsubstituted FAMEs (9.9%), ω-methoxy-$C_{14}$-$C_{18}$ FAMEs (6.1%), 9/16- and 10/16-dimethoxy-$C_{16}$-FAMEs (0.6%) and $C_{14}$-$C_{16}$ DAMEs (12%) clearly reflect the fingerprint of cutin from leaf cuticles. In this case, the methylated *p*-coumaric acid (P18, 3.3%) may originate from cutin. Among the carbohydrate products (18%), several unidentified compounds (*m/z* 155 and 187; Schwarzinger et al., 2002) that are scarce among the THM products of the other samples, indicate a different carbohydrate composition. These results point to the relatively high abundance of cutin (abundance of additionally substituted $C_{16}$ and $C_{18}$ FAMEs) and diterpenoid resin (DHA).

The WEOM from the needles produced a chromatogram with very high signal intensity, causing signal overload for some of the main peaks and the abundance of catechol (not quantified), guaiacol (39%), vanillic acid ME (5.5%) and a compound with *m/z* 137, 151 and 196 (tentatively identified as 3,4-dimethoxy-benzenepropanol; 3.8%) shows that methylation was incomplete. G6 was also abundant (23%), and total G-type compounds accounts for almost 90% of TQPA (Fig. 1b). Previous work using $^{13}$C-labelled TMAH on leachates obtained from Norway spruce needles showed predominance of G6 from protocatechuic (catechol moiety), not vanillic (guaiacol moiety) acid (Klotzbücher et al., 2013). That same study showed that G1 was almost completely $M^+$ 140, hence a methylation product of catechol (not guaiacol; which would produce $M^+$ 139 after methylation). The abundance of catechol groups is in agreement with the detection of catechol and G1 (11%) using THM-GC–MS in the present study: these groups should probably be ascribed to condensed tannin moieties (procyanidin B-rings). It is also in agreement with the high $Py_{cat/gua}$ ratio (7.5; Fig. 3a). Hence, the needle material has a high tannin load that is efficiently transferred to the WEOM phase (tannin is known for its extractability and solubility in water; Hernes et al., 2001; Preston et al., 2009), which may explain the relatively high yield of WEOM as well (Table 1). However, the condensed tannin was not prolific of A-ring products in this case, which is perhaps associated with methylation efficiency differences of the different rings: consumption of the available TMAH to methylate the free hydroxylic groups on the more reactive B-ring (Slabbert, 1992). The dominance of dihydroxy- over trihydroxy- B-rings (and hence G1 and G6 over S1 and S6) is in agreement with the prevalence of procyanidin (dihydroxy-B-rings) rather than prodelphinidin (trihydroxy-B-rings) in the condensed tannins of spruce (Nierop et al., 2005), also reported for hot water extracts of Norway spruce (100% procyanidin; Bianchi et al., 2015, 2016). Hence, B-ring products of procyanidin condensed tannin can make a significant contribution to the G-type products (especially G1) even when A-ring products (Phl-type) do not indicate a major tannin content.

The BOM from spruce cones is characterized by the large proportions of terpenoids (40%; Fig. 2d). The Phl-type products are also abundant (Fig. 1d), which indicates a high condensed tannin content. The S-type products are scarce (0.6%; Fig. 1c; indicative of scarcity of prodelphinidin B-rings). The carbohydrates account for 9.9% (Fig. 2b), FAMEs for 8.5% (Fig. 2a) and G-type products for 22% of TQPA (Fig. 1b). Finally, compound P18 accounts for 1.0% of TQPA and may reflect sporopollenin of the pollen grains of the cones.

The WEOM of the cone material produced the highest proportion of terpenoids among the WEOM samples (24% of TQPA; (Fig. 2d), suggesting that the resin of the cones is relatively easily mobilized. Using

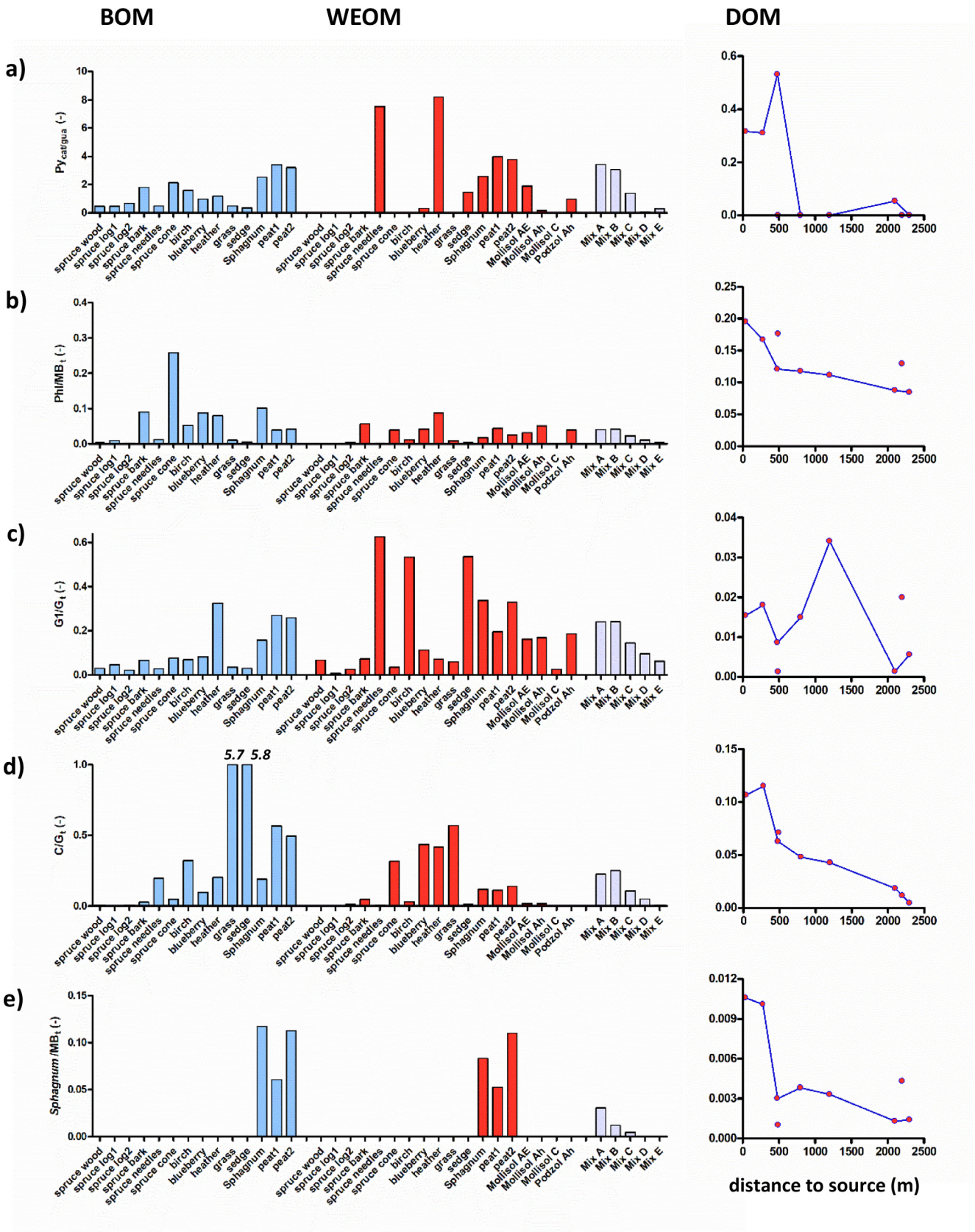


**Fig. 3.** Ratios used for product source assessment of THM-GC–MS. a) catechol/guaiacol ratio ($Py_{cat/gua}$); b) 1,3,5-trimethoxybenzenes/total methoxybenzenes (Phl/$MB_t$); c) 1,2-dimethoxybenzene/total 3,4-dimethoxybenzenes (G1/$G_t$); d) cinnamyl to total 3,4-dimethoxybenzenes (C/$G_t$); e) *Sphagnum* markers to total methoxybenzenes (*Sphagnum*/$MB_t$) (calculations provided in Table 3), for bulk organic matter (BOM), water-extractable OM (WEOM) and dissolved OM from the Oder (DOM). Data corresponds to bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). Values in italics (C/$G_t$ ratio of graminoid sources) indicate values that exceed the maximum of the y-axis. Note that all ratios correspond to THM-GC–MS data except for $Py_{cat/gua}$ (Py-GC–MS; Kaal et al., 2020). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

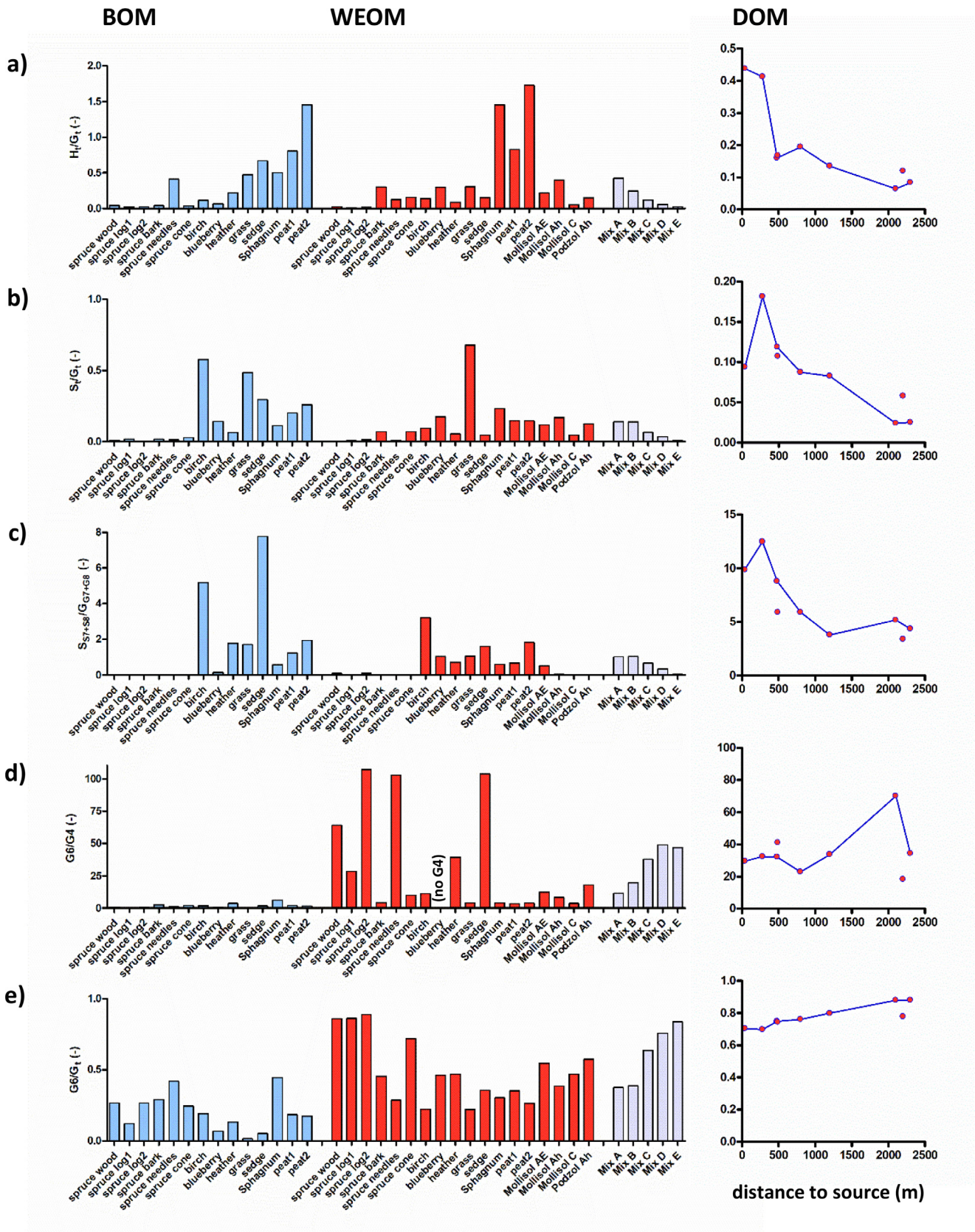


**Fig. 4.** Ratios used for product source assessment of THM-GC–MS. a) 4-methoxybenzenes/3,4-dimethoxybenzenes ($H_t/G_t$); b) 3,4,5-trimethoxybenzenes/3,4-dimethoxybenzenes ($S_t/G_t$); c) methoxyethylene-substituted 3,4,5-trimethoxybenzenes/methoxyethylene-substituted 3,4-dimethoxybenzenes ((S7 + S8)/(G7 + G8)); d) G-acid/G-aldehyde (G6/G4); e) G-acid/total G-type compounds ($G6/G_t$) (calculations provided in Table 3). Data presented for bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Py-GC–MS, catechol was not detected (Kaal et al., 2020), suggesting that the G-type products from THM (24%; Fig. 1b) originate predominantly from lignin. The elevated contribution of P18 (7.4%) may indicate sporopollenin, and even though the pollen grains of Norway spruce are two orders of magnitude larger than the pores of the 0.45 μm filter and degradation or solubilization of the extremely hydrophobic sporopollenin is not expected, P18 moieties may be esterified to sporopollenin and released from the macromolecule during water extraction (Nierop et al., 2019). The FAMEs (3.9%; Fig. 2a) are mostly $C_{16}$-$C_{18}$ FAMEs, and the DAMEs (9.0%; Supplementary Material S2) are almost exclusively $C_7$-DAME, suggesting that cutin and suberin derivatives are not present, or at least not recognizable as such.

In summary, the BOM chromatograms of the different anatomical parts of the spruce material confirm main differences in biopolymer constitution, i.e. the balance between carbohydrates, lignin, tannin, cutin, suberin, resins, protein and possibly sporopollenin. The phase transition to WEOM is highly selective. For WEOM, the cutin and suberin macromolecules are barely identifiable, terpenoids are efficiently released only from cone materials and to a minor extent from bark, the latter of which does have a relatively high fatty acids contribution (not from macromolecules). The bark- and especially needle-derived WEOM generate condensed tannin products. The dominance of lignin-derived G-type products (G6) is very strong for wood-derived material. WEOM of needles, bark and cones have significant contributions of different H-type products (needle P5; bark P24, cone P18). For WEOM, the G6/G4 balance is very high for wood (28–110; Fig. 4d) and moderately high for bark and cones (3.3–4.2), and low for the BOM materials in general (0.3–2.6). The needle-derived WEOM also has high G6/G4 (~100) but here the prevailing source is tannin, not lignin.

#### 3.1.3. Angiosperms, Sphagnum, peat and soil-derived WEOM

The results obtained for the BOM and WEOM samples of the other samples analyzed (Table 1; Figs. 1 and 2), are described in detail in Supplementary Material S3. In summary, the scarcity of syringyl lignin in the WEOM samples of the angiosperms shows that G-type structures are more efficiently released and that low S/G ratios in DOM do not necessarily mean low contributions of angiosperm lignin. Nevertheless, angiosperm polyphenols are likely to be still recognizable from compounds like P18, G18 and syringyl products S7 and S8, among others (Supplementary Material S3). The signal of graminoid lignin (grasses and sedges) may be expected to be recognizable in the THM fingerprints of DOM but woody plants such as heather may create bias if only based on P18 and G18: the high abundance of these compounds in the BOM is only partially reflected in the WEOM. Sphagnum acid-derived phenolics will be detectable if sufficiently abundant. The isopropenylphenol and 3-(4-methoxyphenyl)-butenoic acid ME markers of *Sphagnum* are detected in BOM (4.5% of TQPA) and WEOM (2.8%) (Fig. 1e). *Sphagnum* also releases condensed-tannin like phenolics (van der Heijden, 1994) and H-type products (Williams et al., 1998), in particular H-type products that are not abundant in the graminoids (Supplementary Material S3). None of the WEOM samples showed predominance of G6 at levels near those generated by the spruce wood samples, implying that even though G6 is a universal product of lignin and non-lignin phenolics, very high levels in stream and reservoir waters from the Harz Mountains probably indicates a large contribution of spruce wood-derived DOM.

The BOM and WEOM from the peat samples produced sphagnum acid and H-type products, in addition to the major products detected from living *Sphagnum*, such as carbohydrate products, indicating that ancient *Sphagnum* peat contains and releases marker compounds of sphagnum acid (Fig. 1; Supplementary Material S3). The WEOM of surface horizons of mineral soils (Mollisol AE and Ah, Podzol Ah) produce mainly carbohydrates, a variety of lignin and tannin-derived methoxybenzenes, FAMEs, N-compounds and BCA (Figs. 1 and 2; Supplementary Material S2; Supplementary Material S3). These results indicate a significant polyphenol content, in addition to carbohydrates and N-rich groups of plant and/or microbial origin and aliphatic materials. The Mollisol C horizon did not produce a THM chromatogram in which plant-derived polyphenols were clearly detectable (FAMEs were dominant, possibly from root-derived suberin; Supplementary Material S3). The WEOM from the Podzol Bhs and C horizons did not produce meaningful chromatograms.

### 3.2. Proxies for tracing DOM provenance in the Oder stream

#### 3.2.1. Proxies based on phenolic DOM constituents

The phenolic THM products (represented by H-, G-, S-, Phl- and sphagnum acid products) showed the largest structural diversity for the WEOM and Oder DOM samples. Besides the detection of unequivocal lignin-derived compounds (e.g., G14, G15) from THM of most WEOM samples, several samples showed efficient leaching of tannin from BOM to WEOM (e.g., spruce needles/bark, heather, possibly *Sphagnum*), and many THM compounds can be sourced to both. Unravelling the sources of phenolic compounds requires estimating the balance between lignin and tannin derivatives. As a first approach, we evaluated $Py_{cat/gua}$, which showed elevated levels for the WEOM of spruce needle and heather (Kaal et al., 2020; Fig. 3a). For environmental samples subjected to degradation mechanisms (peat, soil and Oder DOM), catechol is not only a product of tannin but also of demethylated lignin (Haider, 1986; Filley et al., 2002). The DOM samples from the headwater and the first sample of the forest section have $Py_{cat/gua}$ values between 0.3 and 0.6, whereas the downstream samples have negligible levels (Fig. 3a). This suggests that the peatland is the main source of catechol in the system (peat samples have relatively high $Py_{cat/gua}$ levels as well). It is very unlikely that the catechol from the peatland is mainly from demethylated guaiacyl lignin (decay-controlled demethylation of guaiacyl groups would be expected to be more relevant to gymnosperm forest soils under oxic conditions than in the peatland), which implies that $Py_{cat/gua}$ in the Oder DOM is controlled by the contribution of tannin or tannin-like phenolics from the peatland. Among the peatland vegetation sources, heather and *Sphagnum* are more likely sources than grasses and sedges (Fig. 3a). Note that for *Sphagnum* this signal may correspond to an unidentified Phl-type-containing polymer that resembles condensed tannin (Wilson et al., 1989; van der Heijden, 1994). In the forest section, tannin contribution to G-type products is probably negligible.

Phl-type THM products correspond mainly to A-rings in condensed tannins. The contribution of Phl-type products to total methoxybenzenes ($Phl/MB_t$; Fig. 3b) is slightly higher in the headwater section than in the forest section of the Oder, and for WEOM highest levels are found for heather. Species of heather including *Calluna vulgaris* are known for their high condensed tannin contents (Frutos et al., 2002). Hence, even though tannin fluxes from the forest environment, such as DOM from spruce needles, are a significant potential source of condensed tannin, their contribution is not needed to explain the trends in molecular composition observed in the Oder.

Another way to approach tannin contribution to the polyphenols is the G1/total G ($G1/G_t$) ratio (Fig. 3c; Table 3) which may reflect procyanidin A-rings. This applicability of this ratio relies on the fact that for lignin the degradation pathways does not eliminate the α-carbon of the side-chain, whereas for condensed tannin the opening of the C-ring favors the formation of unsubstituted B-rings (Nierop et al., 2005). Hence, even though G1 is partially lignin-derived, prevalence of procyanidin tannin as the main source of G-type products will have relatively high $G1/G_t$. For the BOM, $G1/G_t$ is highest for heather and *Sphagnum* (and *Sphagnum* peat) (Fig. 3c). For the WEOM, $G1/G_t$ is highest for spruce needles (demonstrated high tannin content; Klotzbücher et al., 2013). For the Oder DOM the ratio is below 0.04, which implies that a significant contribution of condensed tannin to G-type products is not evident. Nevertheless, there is a small maximum in $G1/G_t$ in the forest section, perhaps due to the high proportion of procyanidin tannin in spruce materials (Fig. 3c) and in particular the

**Table 3**
List of proxies used to evaluate molecular fingerprints.

| Ratio | Calculation | Meaning | Figure |
|---|---|---|---|
| $Py_{cat/gua}$ * | catechol/guaiacol* | Balance between dihydroxybenzenes (tannin) and methoxyphenols (G-lignin) | 3a |
| $Phl/MB_t$ | 1,3,5-trimethoxybenzenes/total phenolics | Proportion of polyphenols that can be ascribed to condensed tannin A-rings | 3b |
| $G1/G_t$ | 1,2-dimethoxybenzene/total G-type products | Proportion of polyphenols that can be ascribed to condensed tannin B-rings | 3c |
| C/G | cinnamyl (P18 + G18)/vanillyl (all G-type product minus G18) | graminoid/non-graminoid lignin and lignin-like phenolics | 3d |
| *Sphagnum acid*/ total phenolics | sphagnum acid products/total phenolics | index of contribution of *Sphagnum* phenolics | 3[e] |
| $H_t/G_t$ | H-type to G-type methoxybenzenes (except for P18 and G18) | index of polyphenol composition, high H-type contribution for peatland sources | 4a |
| $S_t/G_t$ | total S-type to G-type products (except for G18) | syringyl-to-guaiacol ratio for lignin (biased by non-lignin sources) | 4b |
| $S_{S7S8}/G_{G7G8}$ | methoxyethylene-substituted S to G products | syringyl-to-guaiacol ratio for lignin (less biased by non-lignin sources) | 4c |
| G6/G4 | G-acid (G6) to G-aldehyde (G4) | none, usually used as a proxy of biological alteration lignin | 4d |
| $G6/G_t$ | G-acid (G6) to total G-type products (except for G18) | abundance of spruce wood-derived G-type products (only DOM) | 4e |
| $C_5$MSA/total carbohydrates | $C_5$-metasaccharinic acid (methylated)/total carbohydrate products | index of carbohydrate products from degraded plant-derived polysaccharides | 5a |
| ΣDHA | Sum of dehydroabietic acid derivatives and retene | Abundance of Pinaceae-derived diterpene resin | 2 |
| $\Sigma FAME_{CUTIN}$ | sum of di- and trimethoxy-$C_{16}/C_{18}$-FAMEs | abundance of cutin-derived products (leaf cuticles) | 5b |
| $\Sigma FAME_{SUBERIN}$ | sum of ω-methoxy-FAMEs ($C_{20}$, $C_{24}$, $C_{26}$) and DAMEs ($C_{20}$ and $C_{22}$) | abundance of suberin-derived products (from root and bark materials) | 5c |
| ΣPEAT | sum of THM products with elevated ($< -0.7$) negative loadings on PC2 from PCA | abundance of substances from the peatland environment | 6a |
| ΣFOREST | sum of THM products with elevated ($> 0.7$) positive loadings on PC2 from PCA | abundance of substances from the forest environment (essentially spruce wood) | 6b |
| PEAT/FOREST | Ratio between ΣPEAT and ΣFOREST | balance between DOM from peatland and forest biomes (in Oder stream, ranges from 1.1 to 0.2) | 6c |

* From Py-GC–MS (all other proxies from THM-GC–MS).

release of B-ring moieties from spruce needles. This emphasizes that there is not necessarily a link between release of A-rings ($Phl/MB_t$) and that of B-rings ($Py_{cat/gua}$ and $G1/G_t$). In condensed tannins, monomers are bridged through ethers and C-C bonds of A-ring moieties whereas the B-ring is more reactive and is perhaps more easily liberated from the polymer due to the unstable pyran-like C-ring that connects the A- and B-rings. Combined evidence ($G1/G_t$ in combination with $Py_{cat/gua}$ and $Phl/MB_t$) suggests that *Sphagnum* and heather (both associated with peatland environment) are the main sources of condensed tannin signal, giving rise to the downstream decrease. The role of *Sphagnum* in the signal of the Phl-type abundance is not surprising considering that *Sphagnum* tannin-like phenolics are readily released during early decomposition (Zak et al., 2019). However, future studies of DOM from complex systems should include $^{13}C$ TMAH THM-GC–MS measurements of selected samples as a control assessment, as these ratios are shifty.

The contribution of cinnamyl groups to the phenolic fingerprint can be targeted by the proportion of *p*-coumaric (P18) and ferulic (P18) acids to total G-type (except G18) products ($C/G_t$; Chefetz et al., 2000). For BOM, $C/G_t$ ratio shows the highest levels for grasses and sedges (Fig. 3d), due to the well-known high proportion of these moieties in lignin and lignin-like phenolics in graminoids (Hedges and Mann, 1979). The difference in $C/G_t$ for graminoids and other materials is less clear for WEOM, with values for grass material in the same range as blueberry and heather, which is in agreement with Hernes et al. (2007), who showed a strong increase in $C/G_t$ of woody tissues from BOM to leachate. For sedge-derived WEOM, the ratio is very low, but the WEOM after incubation (Supplementary Material S1) was as high as that of the grass-WEOM (0.6; Supplementary Material S2) suggesting that, for sedge, these moieties required some more time or biological action to become water-extractable. A clear downstream trend in $C/G_t$ is observed in the DOM from the Oder, with a progressive decline from headwater to the reservoir, suggesting that $C/G_t$ is controlled by the abundance of graminoid-derived phenolics, in addition to ericoid sources possibly. Of these ericoids, *Calluna vulgaris* (WEOM with large peak of G18) is an important member of the peatland vegetation and should therefore be taken into account when $C/G_t$ is interpreted. The other ericoid studied, i.e. *Vaccinium myrtilus*, also produced WEOM with a high $C/G_t$, suggesting that forest-derived DOM can also contribute cinnamyl groups, but the tendencies in $C/G_t$ of the Oder river (negligible $C/G_t$ in the samples at largest distance from the headwater) suggests that blueberry-derived cinnamyl groups are irrelevant. The same argument can be used to reject a major contribution of cinnamyl groups delivered by suberin, cutin or sporopollenin from the forest. Finally, the low $C/G_t$ for the Oder samples, even those of the headwater environment, in comparison with WEOM, suggests that fresh plant materials release more cinnamyl moieties than the decaying plant remains present in the environment.

The ratio of sphagnum acid markers (van der Heijden et al., 1997) to total methoxybenzenes (*Sphagnum*/$MB_t$) (Fig. 3e) may be a proxy of the contribution of *Sphagnum*-derived DOM. This ratio was significant for the BOM and WEOM of *Sphagnum*, peat samples and the Oder DOM samples. It is noted that absence of these products does not mean absence of sphagnum acid-derived DOM: it is well known that *Sphagnum* produces of a whole range of phenolic compounds upon THM and the markers are often only minor products (Abbott et al., 2013).

The ratio of H- to G-type products ($H_t/G_t$; excluding P18 and G18) shows that the former are relatively abundant in the *Sphagnum*-derived WEOM and the peat-derived BOM and WEOM samples (Fig. 4a). The Oder DOM shows a clear trend from high $H_t/G_t$ (~0.4) in the peatland environment to increasingly low levels in the forest section ($< 0.2$). This suggests that, in the present system, $H_t/G_t$ is controlled by the abundance of *Sphagnum*-derived compounds. Indeed, most of the phenolic THM signal of *Sphagnum* is of the H-type products (van der Heijden et al., 1997).

The S/G ratio was calculated using all G (except G18) and S compounds (Fig. 4b), and by only using those G- and S-type compounds with a methoxyethylene group (G7, G8, S7 and S8; Fig. 4c). The latter products are more specific of lignin (not found in THM chromatograms of several tannin species; Nierop et al., 2005). Even though these ratios are correlated ($r^2 = 0.64$; $P < 0.001$), contrary to $S_t/G_t$, the $S_{S7+S8}/G_{G7+G8}$ is zero for all spruce materials, which is consistent with its lack of syringyl. Moreover, for $S_{S7+S8}/G_{G7+G8}$ there is a strong correlation for BOM and WEOM ($r^2 = 0.76$; $P < 0.001$), with highest values for birch and sedge, and among the angiosperms, low levels for heather and

blueberry (this may be due to G-enriched lignin of bark materials of many angiosperms; Marques et al., 2006). For the Oder DOM samples, $S_{S7+S8}/G_{G7+G8}$ decreases downslope ($r^2 = 0.75$; $P < 0.005$), which probably reflects the progressive outweighing of angiosperm lignin from the peatland to the gymnosperm lignin in the forest environment. This might indicate that for the spatially simple Oder catchment, the shift from peatland to forest is adequately reflected by both ratios, and that the contribution of tannin phenols to the polyphenolic products does not bias the lignin ratio, but the use of $S_{S7+S8}/G_{G7+G8}$ is probably safer when more complex situations are considered.

As explained above, G6/G4 (Fig. 4d) cannot be used as a proxy of biological alteration of lignin in DOM, but G6 can be useful as an estimation of the proportion of G-type products from spruce wood, which generates WEOM with G6 > 60%. The G6/total G ratio ($G6/G_t$; Fig. 4e) is below 0.5 for all BOM samples, below 0.6 for all angiosperm WEOM samples, and ranges from 0.3 (needles) to 0.9 (wood) for spruce WEOM. The $G6/G_t$ increases steadily from 0.7 to 0.9 in the Oder system ($r^2 = 0.75$; $P < 0.005$). Obviously, $G6/G_t$ is not the kind of ratio that can be applied blindly to any system, but it provides a clue on the abundance of spruce wood-derived lignin in the DOM of the forest environments in the Harz Mountains.

### 3.2.2. Proxies from other (non-phenolic) DOM constituents

There are strong differences in carbohydrate fingerprints between BOM, WEOM and DOM. Some compounds such as trimethyllevoglucosan are significant only in BOM, suggesting that they represent intact cellulose or are not transferred to the WEOM for other reasons. Other products are enriched in WEOM and/or Oder DOM, such as the pentose $C_5$-metasaccharinic acid ($C_5$MSA), which is much more abundant in DOM than in BOM or WEOM, which might indicate that it is associated with degraded material. $C_5$MSA increases downstream in the Oder, as do most of the hexose analogues ($C_6$MSA). Jeanneau et al. (2014) suggested a ratio of pentose-to-hexose-based THM products as an indication of the proportion of plant-to-microbial-derived DOM using THM-GC–MS, in the same line of argument as Guggenberger and Zech (1994), but in the present study the $C_5$MSA/deoxy-$C_6$MSA ratio of the BOM samples did not provide meaningful information (not shown). The proportion of $C_5$MSA to total carbohydrate products ($C_5$MSA/total carbohydrates; Fig. 5a) is clearly higher in DOM than in BOM and WEOM (with the expected exception of BOM from decomposing spruce log 2), which could indicate that $C_5$MSA production depends on decomposition (release of degraded plant-derived polysaccharides). This could explain the downstream increase in $C_5$MSA/total carbohydrates in the Oder, as well as the high ratios in mineral soils (Fig. 5a).

Eight diterpene-derivatives (DHA) were identified, either with or without 7-oxo- or 7-methoxy-substitution (functionalized DHA) and with different levels of saturation (2–5 double bonds). They were identified in the BOM and WEOM from spruce tissues, soil-derived WEOM and Oder stream DOM samples. The sum of these compounds (ΣDHA) exceeds 10% in the BOM of the non-woody spruce tissues and is also high (> 2%) in the WEOM of bark and cone materials (Fig. 2d). In the Oder DOM, peatland samples have 0.0 and 0.2% diterpenes, whereas the samples from the forest have 0.2–1.9%. This suggests that ΣDHA is a good proxy of Pinaceae resin (spruce in the study area). We explored multiple possible proxies of abietane diterpene degradation state, on the basis of the number of O-functionalized groups (Colombini et al., 2001; Pastorová et al., 1997; van den Berg, 2003; Lantes Suárez et al., 2018) and the number of double bonds for both 7-oxo-/7-methoxy-DHA and unfunctionalized DHA derivatives. The double bond proxies were inconsistent with alteration state (e.g. lower for soil-derived WEOM than for spruce-derived WEOM). The 7-oxo-DHA was similarly abundant in BOM and WEOM, not present in soil-derived WEOM and scarce in the environmental DOM (not shown). Hence, this product, which forms upon oxidation of DHA structures but is also a native product, should not be ascribed to evolved resin derivatives. For the degree of functionalization, based on the DHA alteration pathway (van den Berg, 2003; Colombini and Modugno, 2009), inconsistent results were obtained as well (decrease from fresh wood to decomposed wood; Oder DOM very low levels). Hence, the calculated parameters to assess differences in resin preservation state, which could have provided an important clue to differentiate between fresh litter- and soil-derived resin, failed.

Brock et al. (2019) used the sum of di- and trimethoxy-$C_{16}/C_{18}$-FAMEs (Fig. 5b) as a proxy of cutin abundance in gymnosperm litter, and the sum of ω-methoxy-FAMEs ($C_{20}$, $C_{24}$, $C_{26}$) and DAMEs ($C_{20}$ and $C_{22}$) as a proxy of suberin (Fig. 5c). The results for the BOM samples support this general balance between major sources of these products (cutin products enriched in spruce needles and birch litter; long-chain DAMEs in spruce bark and birch litter). Cutin products were not identified in any WEOM or DOM sample. The suberin-associated FAME patterns are slightly more intense for the WEOM of spruce bark and *Sphagnum*, but they are much scarcer than in THM chromatograms of BOM. Their pattern in the Oder DOM sequence is unclear (relatively high levels in the headwater and lower forest samples). It is concluded that (1) cutin is virtually absent and only traces of suberin can be recognized in environmental DOM (see also Denis et al., 2017), (2) most of the aliphatic signal cannot be assigned to a specific source, probably because they do not originate from aliphatic macromolecules but free/ester-bound fats (vegetable oils) and waxes.

An unequivocal signal of microbial-derived substances was not found. In the Oder DOM, $C_5$-alkylpyrrole is the most abundant N-containing product whereas BOM and WEOM tend to have larger peaks for proline ME (and no alkylpyrroles) (not shown). This might indicate that the alkylpyrrole is a useful tracer of microbial nitrogen but the analyzed sources cannot be used to test this hypothesis. Py-GC–MS analyses and FTIR did provide information on the abundance of microbial DOM in the Oder stream (Kaal et al., 2017).

### 3.2.3. Peat and forest DOM indices and hypothetical WEOM admixtures

The PCA of the extended THM-GC–MS dataset of the Oder DOM created two main PCs. The PC1 (36% of total variance) reflected the shifts in the balance between FAMEs and DHA (terpenoids) along the stream transect, and the underlying mechanism of this trend (high DHA proportions mid-stream) is unknown. The PC2 (23% of variance) is of more interest here as it shows a clear boundary between headwater and forest biomes at 400 m from the spring, and a progressive further increase downslope (PC2 scores and loadings in Supplementary Material S2). The sums of the products with loadings > 0.7, and those < −0.7, can be used as proxies of DOM constituents that increase or decrease downstream, respectively. Specifically, the proxy of peatland headwater-derived DOM (ΣPEAT; Fig. 6a) is calculated as the sum of the % (of TQPA) of the compounds with a strong negative loading on PC2 (P3, P6, P18, P24, G18, S6, S7, 4-isopropenylmethoxybenzene, 1,3,5-trimethoxybenzene and an unidentified product with *m/z* 240). This set of compounds highlights that the peatland environment releases more tannin, cinnamyl groups, H- and S-type lignin, and sphagnum acid to the DOM. Calculating ΣPEAT for the BOM, WEOM and DOM samples gives high levels for the grass and sedge (BOM) or grass and shrubs (WEOM); and the obligate downstream decrease for the Oder stream DOM. On the other hand, the sum of the products with positive loadings on PC2 (G4, G6, several carbohydrate products; a proxy of mainly spruce forest-derived DOM) are highest for WEOM of spruce wood samples and this sum (ΣFOREST; Fig. 6b) increases progressively in the Oder DOM as it moves from the peatland to and through the forest biome. It appears that the WEOM of the spruce wood samples have higher ΣFOREST than the other spruce organs, and in fact the high levels of Oder DOM are only paired by that of spruce wood, probably because spruce wood is the main source of the G-type lignin in the Oder forest section. The importance of spruce wood-derived DOM in the studied system is a strong indication of a major role of decomposing trees from the deteriorating forests and the underlying lignin-rich forest soils. The ratio between the abundance of compounds from the peat and

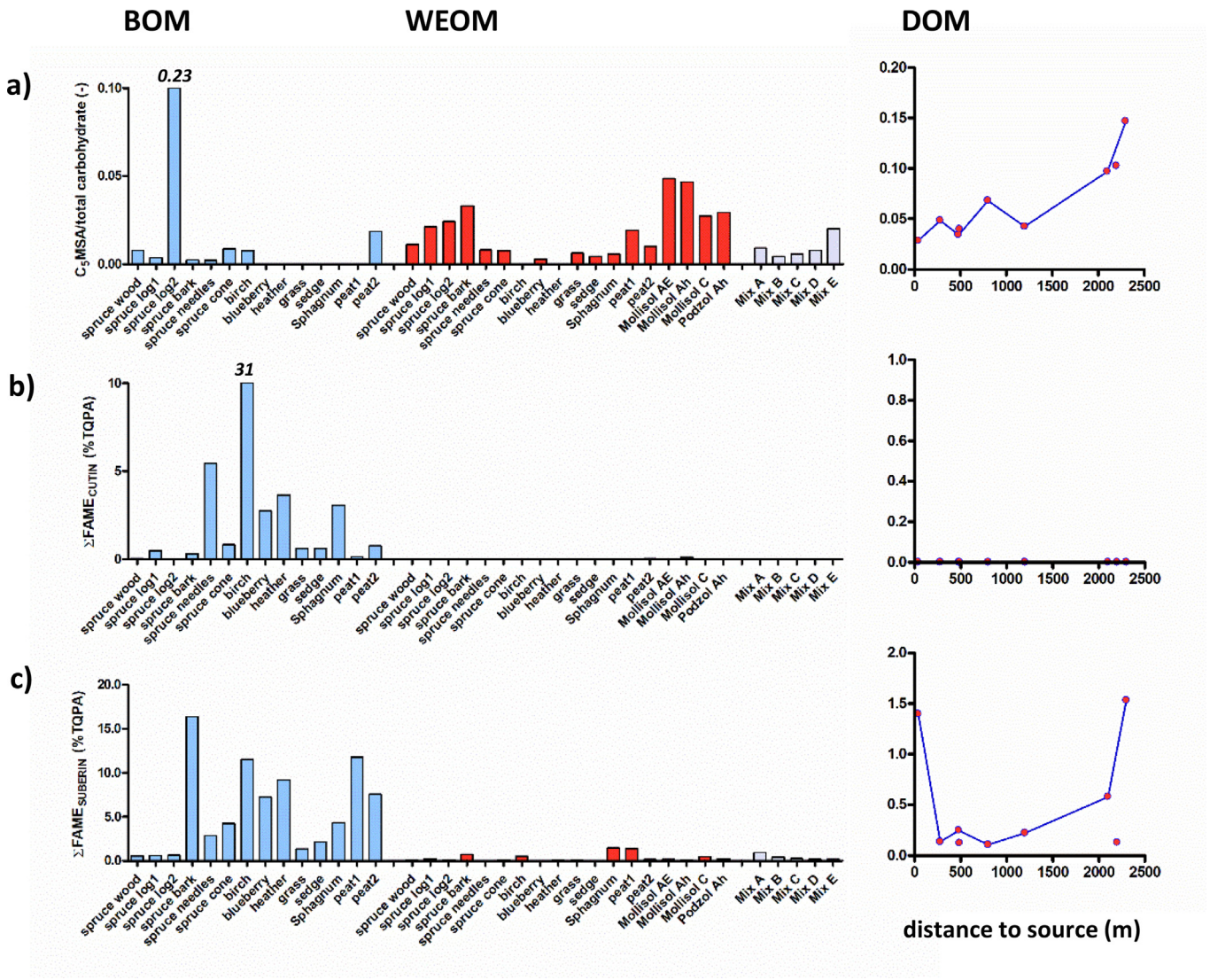


**Fig. 5.** Parameters used for product source assessment of THM-GC–MS, a) methylated $C_5$-metasaccharinic acid/total carbohydrates ($C_5$MSA/total carbohydrate), b) sum cutin products ($\Sigma FAME_{CUTIN}$), c) sum of suberin products ($\Sigma FAME_{SUBERIN}$) (calculations provided in Table 3). Data presented for bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). Values in italics indicate values that exceed the maximum of the y-axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

forest biomes (ΣPEAT/ΣFOREST) ranges between 1.2 and 0.2 and also marks the boundary between the two environments (Fig. 6b). Highest values are obtained for WEOM of grasses, *Sphagnum* and peat material.

The hypothetical DOM admixtures (Table 1) have ΣPEAT values ranging from 23 (Mix A) to 3 (Mix E) (Fig. 6a). ΣFOREST ranges from 15 (Mix B) to 65 (Mix E) (Fig. 6b). For the peat environment, the maximum value for ΣPEAT of the Oder (at the peat outlet) was only matched by the Mix A (*Sphagnum* and peat materials) even though mixtures with a high proportion (70% and higher) of heather to the hypothetical mixture (not shown) could also generate high ΣPEAT values (due to the G18 and condensed tannin signal that marks the peatland DOM). For the forest signal, the lowest ΣFOREST for the Oder DOM (downstream forest section) was only matched by mixtures with high proportions (> 60%) of spruce wood (mixtures D and E), due to the high proportion of G6 in spruce-wood WEOM and Oder DOM fingerprints. These results show that the proxy values are in the same range as those calculated for environmental DOM samples, and are thus useful for provenancing exercises. However, the calculated THM fingerprints of the hypothetical mixtures are compared to those of DOM from summer baseline discharge conditions only. Different hydrological scenarios (e.g. different seasons, or events like snowmelt, storms and drought) may give rise to DOM fingerprints that are not in the range of the mixtures.

Nierop and Filley (2007) showed that "even with the corrections to lignin proxies [on the basis of $^{13}$C-labeling], the soil lignin and polyphenol chemistry remained complex and highlights the limitations of using a few or only one lignin proxy in assessing SOM dynamics". This complexity is also highlighted by the study of potential DOM sources in the Harz Mountains. The habitat is complex and DOM's biogeochemical cycle depends on more than just biological sources. None of the proposed parameters should be carelessly applied to identify the sources of DOM, but careful evaluation of a variety of proxies (Table 3) can allow for the creation of a consistent body of information.

## 4. Conclusions

The analysis of biological materials (plant tissues, peat and soils) from the Harz Mountains, by means of THM-GC–MS of BOM and WEOM, proved a useful approach to understand the source and proxy value of the THM products found in DOM samples of the Oder river. The

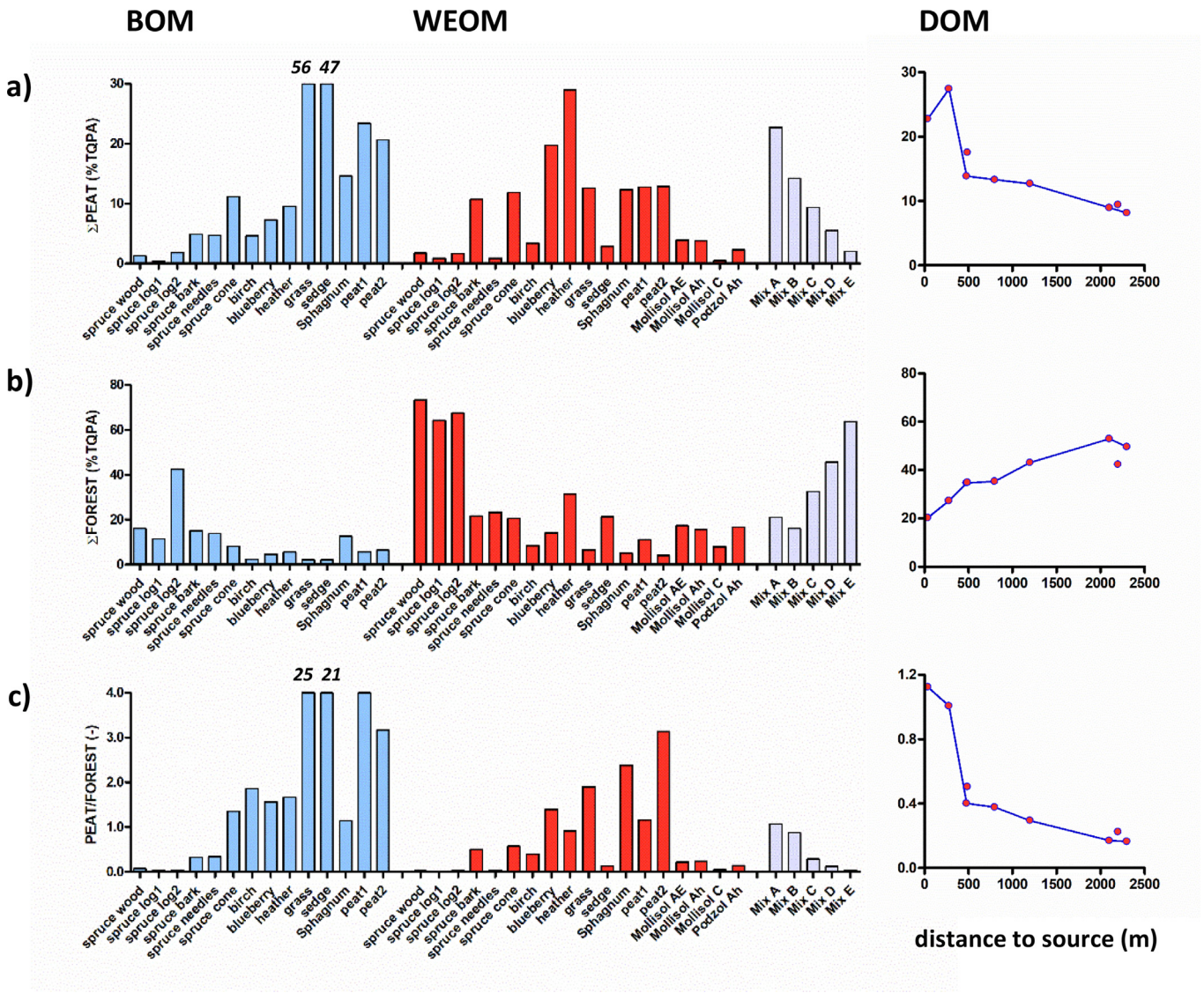


**Fig. 6.** Proxies of predominantly a) peat-derived THM compounds (ΣPEAT) and b) forest-derived compounds (ΣFOREST) (calculations provided in Table 3; numbers reflect summed % of characteristic compounds as defined by PCA); c) ratio between ΣPEAT and ΣFOREST (PEAT/FOREST). Data presented for bulk organic matter (BOM; blue), water-extractable OM (WEOM; red), hypothetical WEOM mixtures A-E (grey) (Table 1) and dissolved OM from the Oder stream (DOM; plotted against distance from headwater peatland). Values in italics indicate values that exceed the maximum of the y-axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

large differences between BOM and WEOM of the same material confirmed that assessment of stream DOM provenance can be improved by using a reference set of leachates. The diversity in THM products of lignin in BOM, including numerous guaiacyl and syringyl products with propanoid side chains, is greatly reduced in WEOM. The most extreme example is found for spruce wood samples, whose WEOM shows exceptional dominance of methylated vanillic acid (G6) derived from lignin. Efficient leaching (BOM to WEOM) of tannin from several samples (heather, birch, spruce needles) demands an effort to distinguish lignin and tannin derivatives in environmental DOM, for which various THM parameters are proposed but further controls are required. Not only graminoid but also peatland ericoid sources of compound G18 (mostly from ferulic acid) such as heather must be considered when interpreting the C/G ratio. The aliphatic biopolymers cutin and suberin show minimal transference to WEOM fractions and their contribution to DOM is negligible. We propose a series of proxies to estimate the balance between DOM from peatland and forest environments (ΣPEAT and ΣFOREST). These proxies will soon be used for the interpretation of THM data for DOM from the Ecker catchment, also in the Harz National Park. The sample set was not suitable to specify sources of microbial DOM. The new insights into the link between BOM and WEOM and between WEOM and DOM are useful to understand THM fingerprints of DOM and identify likely sources. In the Oder stream, the signal of the peatland environment (tannin, H- and S-type lignin, sphagnum acid) which dominates the uphill DOM, is progressively and profoundly outweighed by the G-lignin fingerprint of DOM from spruce wood.

## Declaration of Competing Interest

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

## Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG) BI734/18-1 to HB “The Role of Molecular Dissolved Organic Matter (DOM) Composition to Identify Sources and Release of DOM and

Trace Elements in Catchments of Drinking Water Reservoirs in Mid-Latitude Mountain Areas – (DOMtrace)". We thank Thimo Klotzbücher (Martin-Luther-Universität Halle-Wittenberg) for sharing $^{13}C$ TMAH-THM-GC–MS chromatograms of Norway spruce needle leachates, and two anonymous reviewers for their significant effort and insightful comments.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2020.114635.

## References

Abbott, G.D., Swain, E.Y., Muhammad, A.B., Allton, K., Belyea, L.R., Laing, C.G., Cowie, G.L., 2013. Effect of water-table fluctuations on the degradation of *Sphagnum* phenols in surficial peats. Geochim. Cosmochim. Acta 106, 177–191.

Bardy, M., Derenne, S., Allard, T., Benedetti, M.F., Fritsch, E., 2011. Podzolisation and exportation of organic matter in black waters of the Rio Negro (upper Amazon basin, Brazil). Biogeochem 106, 71–88.

Bianchi, S., Kroslakova, I., Janzon, R., Mayer, I., Saake, B., Pichelin, F., 2015. Characterization of condensed tannins and carbohydrates in hot water bark extracts of European softwood species. Phytochem 120, 53–61.

Bianchi, S., Koch, G., Janzon, R., Mayer, I., Saake, B., Pichelin, F., 2016. Hot water extraction of Norway spruce (Picea abies [Karst.]) bark: analyses of the influence of bark aging and process parameters on the extract composition. Holzforschung 70, 619–631.

Blome, B., 2019. Untersuchung zum Verhalten von Haupt- und Spurenele-menten in der festen und flüssigen Phase am Odersprungmoor. BSc. Thesis, Institut für Geoökologie, Technische Universtät Braunschweig, Abteilung Umweltgeochemie, pp. 68.

Boom, A., Sinninghe Damsté, J.S., de Leeuw, J.W., 2005. Cutan, a common aliphatic biopolymer in cuticles of drought-adapted plants. Org. Geochem. 36, 596–601.

Brock, O., Kooijman, A., Nierop, K.G.J., Muys, B., Vancampenhout, K., Jansen, B., 2019. Disentangling the effects of parent material and litter input chemistry on molecular soil organic matter composition in converted forests in Western Europe. Org. Geochem. 134, 66–76.

Broder, T., Biester, H., 2015. Hydrologic controls on DOC, As and Pb export from a polluted peatland-the importance of heavy rain events, antecedent moisture conditions and hydrological connectivity. Biogeosciences 12, 4651–4664. https://doi.org/10.5194/bg-12-4651-2015.

Broder, T., Biester, H., 2017. Linking major and trace element concentrations in a headwater stream to DOC release and hydrologic conditions in a bog and peaty riparian zone. Appl. Geochem. 87, 188–201.

Broder, T., Knorr, K.-H., Biester, H., 2017. Changes in dissolved organic matter quality in a peatland and forest headwater stream as a function of seasonality and hydrologic conditions. Hydrol. Earth Syst. Sci. 21, 2035–2051. https://doi.org/10.5194/hess-21-2035-2017.

Challinor, J.M., 1989. A pyrolysis-derivatisation-gas chromatography technique for the structural elucidation of some synthetic polymers. J. Anal. Appl. Pyrolysis 16, 323–333.

Chantigny, M.H., 2003. Dissolved and water-extractable organic matter in soils: a review on the influence of land use and management practices. Geoderma 113, 357–380.

Chefetz, B., Chen, Y., Clapp, C., Hatcher, P.G., 2000. Characterization of organic matter in soils by thermochemolysis using tetramethylammonium hydroxide (TMAH). Soil Sci. Soc. Am. J. 64, 583.

Clifford, D.J., Carson, D.M., McKinney, D., Bortiatynski, D.M., Hatcher, P.G., 1995. A new rapid technique for the characterization of lignin in vascular plants: thermochemolysis with tetramethylammonium hydroxide (TMAH). Org. Geochem. 23, 169–175.

Colombini, M.P., Modugno, F., 2009. Organic Mass Spectrometry in Art and Archaeology. John Wiley & Sons Ltd, Chichester.

Colombini, M.P., Lanterna, G., Mairani, A., Matteini, M., Modugno, F., Rizzi, M., 2001. Copper resinate: preparation, characterisation and study of degradation. Ann Chim. 91, 749–757.

Del Río, J.C., Hatcher, P.G., 1998. Analysis of aliphatic biopolymers using thermochemolysis with tetramethylammonium hydroxide (TMAH) and gas chromatography–mass spectrometry. Org. Geochem. 29, 1441–1451.

Del Río, J.C., McKinney, D.E., Knicker, H., Nanny, M.A., Minard, R.D., Hatcher, P.G., 1998. Structural characterization of bio- and geo-macromolecules by offline thermochemolysis with tetramethylammonium hydroxide. J. Chromatogr. A 823, 433–448.

Del Río, J.C., Rencoret, J., Gutierrez, A., Elder, T., Kim, H., Ralph, J., 2020. Lignin monomers from beyond the canonical monolignol biosynthetic pathway – Another brick in the wall. ACS Sustain. Chem. Eng., doi: 10.1021/acssuschemeng.0c01109.

Denis, M., Laurent Jeanneau, L., Petitjean, P., Murzeau, A., Liotaud, M., Yonnet, L., Gruau, G., 2017. New molecular evidence for surface and sub-surface soil erosion controls on the composition of stream DOM during storm events. Biogeosci. Discuss. https://doi.org/10.5194/bg-2017-252.

Don, A., Kalbitz, K., 2005. Amounts and degradability of dissolved organic carbon from foliar litter at different decomposition stages. Soil Biol. Biochem. 37, 2171–2179.

Eikebrokk, B., Vogt, R.D., Liltved, H., 2004. NOM increase in Northern European source waters: discussion of possible causes and impacts on coagulation/contact filtration processes. Water Sci. Technol.: Water Supply 4, 47–54.

Fabbri, D., Helleur, R., 1999. Characterization of the tetramethylammonium hydroxide thermochemolysis products of carbohydrates. J. Anal. Appl. Pyrolysis 49, 277–293.

Filley, T.R., Minard, R.D., Hatcher, P.G., 1999. Tetramethylammonium hydroxide (TMAH) thermochemolysis: proposed mechanisms based upon the application of $^{13}C$ labelled TMAH to a synthetic model lignin dimmer. Org. Geochem. 30, 607–621.

Filley, T.R., Cody, G.D., Goodell, B., Jellison, J., Noser, C., Ostrofsky, A., 2002. Lignin demethylation and polysaccharide decomposition in spruce sapwood degraded by brown rot fungi. Org. Geochem. 33, 111–124.

Filley, T.R., Nierop, K.G.J., Wang, Y., 2006. The contribution of polyhydroxyl aromatic compounds to tetramethylammonium hydroxide lignin-based proxies. Org. Geochem. 37, 711–727.

Frazier, S.W., Kaplan, L.A., Hatcher, P.G., 2005. Molecular characterization of biodegradable dissolved organic matter using bioreactors and [$^{12}C/^{13}C$] tetramethylammonium hydroxide thermochemolysis GC − MS. Environ. Sci. Technol. 39, 1479–1491.

Freeman, C., Evans, C.D., Monteith, D.T., 2001. Export of organic carbon from peat soils. Nature 412, 785.

Frutos, P., Hervás, G., Ramos, G., Giráldez, F.J., Mantecón, A.R., 2002. Condensed tannin content of several shrub species from a mountain area in northern Spain, and its relationship to various indicators of nutritive value. Animal Feed Sci. Technol. 95, 215–226.

Galletti, G.C., Modafferi, V., Poiana, M., Bocchini, P., 1995. Analytical Pyrolysis and Thermally Assisted Hydrolysis-Methylation of Wine Tannin. J. Agric. Food Chem. 1995 (43), 1859–1863.

Gandois, L., Hoyt, A.M., Hatté, C., Jeanneau, L., Teisserenc, R., Liotaud, M., Tananaev, N., 2019. Contribution of peatland permafrost to dissolved organic matter along a thaw gradient in North Siberia. Environ. Sci. Technol. 53, 14165–14174.

Garnier, N., Richardin, P., Cheynier, V., Regert, M., 2003. Characterization of thermally assisted hydrolysis and methylation products of polyphenols from modern and archaeological vine derivatives using gas chromatography–mass spectrometry. Anal. Chim. Acta 493, 137–157.

Godin, P., Macdonald, R.W., Kuzyk, Z.Z.A., Goñi, M.A., Stern, G.A., 2017. Organic matter compositions of rivers draining into Hudson Bay: Present-day trends and potential as recorders of future climate change. J. Geophys. Res. Biogeosci. 122, 1848–1869.

Guggenberger, G., Zech, W., 1994. Dissolved organic carbon in forest floor leachates: simple degradation products or humic substances? Sci. Total Environ. 152, 37–47.

Haider, K., 1986. The synthesis and degradation of humic substances in soil. Transactions XIII Congress ISSS, Hamburg, pp. 644–656.

Hatcher, P.G., Nanny, M.A., Minard, R.D., Dible, S.D., Carson, D.M., 1995. Comparison of two thermochemolytic methods for the analysis of lignin in decomposing gymnosperm wood: the CuO oxidation method and the method of thermochemolysis with tetramethylammonium hydroxide (TMAH). Org. Geochem. 23, 881–888.

He, Y., Buch, A., Szopa, C., Williams, A.J., Milan, M., Guzman, M., Freissinet, C., Malespin, C., Glavin, D.P., Eigendbrode, J.L., Coscia, D., Teinturier, S., lu, P., Cabane, M., Mahaffy, P.R., 2020. The search for organic compounds with TMAH thermochemolysis: from Earth analyses to space exploration experiments. Trends in Anal. Chem. 127, 115896. doi: 10.1016/j.trac.2020.115896.

Hedges, J.I., Mann, D.C., 1979. The characterization of plant tissues by their lignin oxidation products. Geochim. Cosmochim. Acta 43, 1803–1807.

Hedges, J.I., Blanchette, R.A., Weliky, K., Devol, A.H., 1988. Effects of fungal degradation on the CuO oxidation products of lignin – a controlled laboratory study. Geochim. Cosmochim. Acta 52, 2717–2726.

Hedges, J.I., Eglinton, G., Hatcher, P.G., Kirchman, D.L., Arnosti, C., Derenne, S., Evershed, R.P., Kogel-Knabner, I., de Leeuw, J.W., Littke, R., Michaelis, W., Rullkotter, J., 2000. The molecularly-uncharacterized component of nonliving organic matter in natural environments. Org. Geochem. 31, 945–958.

Higuchi, T., Shimada, M., Nakatsubo, F., Tanahashi, M., 1977. Differences in biosyntheses of guaiacyl and syringyl lignins in woods. Wood Sci. Technol. 11, 153–167.

Hernes, P.J., Benner, R., Cowie, G.L., Goi, M.A., Bergamaschi, B.A., Hedges, J.I., 2001. Tannin diagenesis in mangrove leaves from a tropical estuary: a novel molecular approach. Geochim. Cosmochim. Acta 65, 3109–3122.

Hernes, P.J., Robinson, A.C., Aufdenkampe, A.K., 2007. Fractionation of lignin during leaching and sorption and implications for organic matter "freshness". Geophys. Res. Lett. 34, L17401. https://doi.org/10.1029/2007GL031017.

Jeanneau, L., Jaffrezic, A., Pierson-Wickmann, A.-C., Gruau, G., Lambert, T., Petitjean, P., 2014. Constraints on the sources and production mechanisms of dissolved organic matter in soils from molecular biomarkers. Vadose Zone 13 (7).

Jeanneau, L., Denis, M., Pierson-Wickmann, A.-C., Gruau, G., Lambert, T., Petitjean, P., 2015. Sources of dissolved organic matter during storm and inter-storm conditions in a lowland headwater catchment: constraints from high-frequency molecular data. Biogeosciences 12, 4333–4343. https://doi.org/10.5194/bg-12-4333-2015.

Jiang, T., Kaal, J., Liang, J., Zhang, Y., Wei, S., Wang, D., Green, N.W., 2017. Quality of dissolved organic matter (DOM) from periodically submerged soils in the Three Gorges Reservoir areas as determined by elemental and optical analysis, infrared spectroscopy, pyrolysis-GC-MS and thermally assisted hydrolysis and methylation. Sci. Total Environ. 603–604, 461–471.

Kaal, J., Martínez-Cortizas, A., Biester, H., 2017. Downstream changes in molecular composition of DOM along a headwater stream in the Harz mountains (Central Germany) as determined by FTIR, Pyrolysis-GC–MS and THM-GC–MS. J. Anal. Appl. Pyrolysis 126, 50–61.

Kaal, J., Plaza, C., Pérez-Rodríguez, M., Biester, H., 2020. Towards understanding ecological disaster in the Harz Mountains (Central Germany) by carbon tracing: pyrolysis-GC-MS of biological tissues and their water-extractable organic matter (WEOM). Anal. Pyrolysis Lett. 8, 1–17.

Kalbitz, K., Solinger, S., Park, J.-H., Michalzik, B., Matzner, E., 2000. Controls on the dynamics of dissolved organic matter in soils: a review. Soil Sci. 165, 277–304.
Kalbitz, K., Glaser, B., Bol, R., 2004. Clear-cutting of a Norway spruce stand: implications for controls on the dynamics of dissolved organic matter in the forest floor. Eur. J. Soil Sci. 55, 401–413.
Knolle, F., Wegener, U., 2019. Zur Umweltgeschichte der Harzregion am Beispiel von Wald- und Montanwirtschaft seit der Bronzezeit. In: Erb, A., Seyderhelm, B., Volkmar, C. (Eds.), Sachsen und Anhalt: Jahrbuch der Historischen Kommission für Sachsen-Anhalt im Auftrag der Historischen Kommission. Mitteldeutscher Velag, pp. 207–229.
Klotzbücher, T., Filley, T.R., Kalbitz, K., 2013. Processes controlling the production of aromatic water-soluble organic matter during litter decomposition. Soil Biol. Biochem. 67, 133–139.
Krasilnikov, P., Ibáñez Martí, J.J., Arnold, R., Shoba, S., 2009. A Handbook of Soil Terminology, Correlation and Classification. Taylor and Francis, pp. 440.
Kuroda, K., Nishimura, N., Nakagawa-izumi, A., Dimmel, D.R., 2002. Pyrolysis of lignin in the presence of tetramethylammonium hydroxide: a convenient method for S/G ratio determination. J. Agric. Food Chem. 50, 1022–1027.
Lantes Suárez, O., Kaal, J., González Pazos, A., Antón Segurado, R., Fernández Cereijo, I., Mariño Calvo, V., Domínguez Lago, A., 2018. The repentance of San Pedro de Francisco Collantes Restoration and analysis of pigments. Ge-Conservacion 1 (14), 38–51.
Lu, F., Karlen, S.D., Regner, M., et al., 2015. Naturally p-Hydroxybenzoylated Lignins in Palms. Bioenerg. Res. 8, 934–952. https://doi.org/10.1007/s12155-015-9583-4.
Marques, A., Rencoret, J., Gutiérrez, A., del Río, J.C., Moura Pereira, H.C., 2006. Ferulates and lignin structural composition in cork. Holzforschung. https://doi.org/10.1515/hf-2015-0014.
Matiasek, S.J., Hernes, P.J., 2019. The chemical fingerprint of solubilized organic matter from eroded soils and sediments. Geochim. Cosmochim. Acta 267, 92–112.
McKinney, D.E., Bortiatynski, J.M., Carson, D.M., Clifford, D.J., de Leeuw, J.W., Hatcher, P.G., 1996. Tetramethylammonium hydroxide (TMAH) thermochemolysis of the aliphatic biopolymer cutan: Insights to its chemical structure. Org. Geochem. 24, 641–650.
Moore, T.R., Dalva, M., 2001. Some controls on the release of dissolved organic carbon by plant tissues and soils. Soil Sci. 166, 38–47.
Mulder, M., van der Hage, E.R.E., Boon, J.J., 1992. Analytical in-source pyrolytic methylation electron impact mass spectrometry of phenolic acids in biological matrices. Phytochem. Anal. 3, 165–172.
Nebbioso, A., Piccolo, A., 2012. Molecular characterization of dissolved organic matter (DOM): a critical review. Anal. Bioanal. Chem., doi: 10.1007/s00216-012-6363-2.
Nierop, K.G.J., Filley, T.R., 2007. Assessment of lignin and (poly-)phenol transformations in oak (Quercus robur) dominated soils by 13C-TMAH thermochemolysis. Org. Geochem. 38, 551–565.
Nierop, K.G.J., Filley, T.R., 2008. Simultaneous analysis of tannin and lignin signatures in soils by thermally assisted hydrolysis and methylation using 13C-labeled TMAH. J. Anal. Appl. Pyrolysis 83, 227–231.
Nierop, K.G.J., Verstraten, J.M., 2004. Rapid molecular assessment of the bioturbation extent in sandy soil horizons under pine using ester-bound lipids by on-line thermally assisted hydrolysis and methylation-gas chromatography/mass spectrometry. Rapid Commun. Mass Spec. 18, 1081–1088.
Nierop, K.G.J., Preston, C.M., Kaal, J., 2005. Thermally assisted hydrolysis and methylation of purified tannins from plants. Anal. Chem. 77, 5604–5614.
Nierop, K.G.J., Versteegh, G.J.M., Filley, T.R., de Leeuw, J.W., 2019. Quantitative analysis of diverse sporomorph-derived sporopollenins. Phytochem 162, 207–215.
Nip, M., Tegelaar, E.W., de Leeuw, J.W., Schenck, P.A., Holloway, P.J., 1986. A new non-saponifiable highly-aliphatic and resistant bioploymer in plant cuticles. Evidence from pyrolysis and 13C-NMR analysis of present-day and fossil plants. Naturwissenschaften 73, 579–585.
Overbeck, M., Schmidt, M., 2012. Modelling infestation risk of Norway spruce by Ips typographus (L.) in the Lower Saxon Harz Mountains (Germany). Forest Ecol. Manage. 266, 115–125.
Pastorová, I., van der Berg, K.J., Boon, J.J., Verhoeven, J.W., 1997. Analysis of oxidised diterpenoid acids using thermally assisted methylation with TMAH. J. Anal. Appl. Pyrolysis 43, 41–57.
Preston, C.M., Jason R. Nault, J. A. Trofymow, Carolyn Smyth & CIDET Working Group, 2009. Chemical Changes During 6 Years of Decomposition of 11 Litters in Some Canadian Forest Sites. Part 1. Elemental Composition, Tannins, Phenolics, and Proximate Fractions. Ecosystems 12, 1053–1077.
Rencoret, J., Marques, G., Gutiérrez, A., Nieto, L., Santos, J.I., Jiménez-Barbero, J., Martínez, A.T., del Río, J.C., 2009. HSQC-NMR analysis of lignin in woody (Eucalyptus globulus and Picea abies) and non-woody (Agave sisalana) ball-milled plant materials at the gel state. Holzforschung 63 (6), 691–698. https://doi.org/10.1515/HF.2009.070.
Rencoret, J., Gutiérrez, A., Nieto, L., Jiménez-Barbero, J., Faulds, C.B., Kim, H., Ralph, J., Martínez, A.T., del Río, J.C., 2011. Lignin composition and structure in young versus adult *Eucalyptus globulus* plants. Plant Physiol. 155, 667–682.
Riley, R.G., Kolattukudy, P.E., 1975. Evidence for covalently attached p-Coumaric acid and ferulic acid in cutins and suberins. Plant Physiol. 1975 (56), 650–654.
Schwarzinger, C., Tanczos, I., Schmidt, H., 2002. Levoglucosan, cellobiose and their acetates as model compounds for the thermally assisted hydrolysis and methylation of cellulose and cellulose acetate. J. Anal. Appl. Pyrolysis 62, 179–196.
Slabbert, N. (1992). Complexation of Condensed Tannins with metal Ions. In: Hemingway R.W., Laks, P.E. (Eds.), Plant polyphenols: Synthesis, Properties, Significance (1992). Plenum Press, New York, 1053 pp.
Sleighter, R.L., Cory, R.M., Kaplan, L.A., Abdulla, H.A.N., Hatcher, P.G., 2014. A coupled geochemical and biogeochemical approach to characterize the bioreactivity of dissolved organic matter from a headwater stream. J. Geophys. Res.: Biogeosci. 119, 1520–1537. https://doi.org/10.1002/2013JG002600.
Spencer, R.G.M., Hernes, P.J., Aufdenkampe, A., et al., 2012. An initial investigation into the organic matter biogeochemistry of the Congo River. Geochim. Cosmochim. Acta 84, 614–627.
Van den Berg, K.J., 2003. MOLART Report 10, Analysis of Diterpenoid Resins and Polymers in Paint Media and Varnishes (With an Atlas of Mass Spectra). FOM Institute AMOLF Amsterdam, p. 128. Available at: http://wiki.collectiewijzer.nl/images/8/84/Molart_Report_10.pdf.
Van der Heijden, E., 1994. A combined anatomical and pyrolysis mass spectrometric study of peatified plant tissues. PhD Thesis, University of Amsterdam.
Van der Heijden, E., Boon, J.J., Rasmussen, S., Rudolph, H., 1997. Sphagnum acid and its decarboxylation product isopropenylphenol as biomarkers for fossilised Sphagnum in peats. Ancient Biomol. 1, 93–107.
Van Heemst, J.D.H., del Río, J.C., Hatcher, P.G., de Leeuw, J.W., 2000. Characterization of estuarine and fluvial dissolved organic matter by thermochemolysis using tetramethylammonium hydroxide. Acta Hydrochim. Hydrobiol. 28, 69–76.
Vane, C.H., Abbott, G.D., Head, I.M., 2001. The effect of fungal decay (Agaricus bisporus) on wheat strawlignin using pyrolysis–GC–MS in the presence of tetramethylammonium hydroxide (TMAH). J. Anal. Appl. Pyrolysis 60, 69–78.
Vogt, R.D., 2003. Increase in colour and amount of organic matter in surface waters. NORDTEST Position paper 009. Available at http://www.nordtest.org. 11 p.
Wehling, K., Niester, C., Boon, J.J., Willemse, M.T., Wiermann, R., 1989. p-Coumaric acid - a monomer in the sporopollenin skeleton. Planta 179, 376–380.
Williams, C.J., Yavitt, J.B., Wieder, R.K., Cleavitt, N.L., 1998. Cupric oxide oxidation products of northern peat and peat-forming plants. Can. J. Bot. 76, 51–62.
Wilson, M.A., Sawyer, J., Hatcher, P.G., Lerch III, H.E., 1989. 1,3,5-hydroxybenzene structures in mosses. Phytochem 28, 1395–1400.
Worrall, F., Harriman, R., Evans, C.D., Watts, C.D., Adamson, J., Neal, C., Tipping, E., Burt, T., Grieve, I., Monteith, D., Naden, P.S., Nisbet, T., Reynolds, B., Stevens, P., 2004. Trends in dissolved organic carbon in UK rivers and lakes. Biogeochem 70, 369–402.
Zak, D., Roth, C., Unger, V., Goldhammer, T., Fenner, N., Freeman, C., Jurasinski, G., 2019. Unravelling the importance of polyphenols for microbial carbon mineralization in rewetted riparian peatlands. Front. Environ. Sci. https://doi.org/10.3389/fenvs.2019.00147.
Zsolnay, A., 1996. Dissolved humus in soil waters. In: Piccolo, A. (Ed.), Humic Substances in Terrestrial Ecosystems, Elsevier, Amsterdam, pp. 171–223.