We greatly appreciate the constructive comments from both reviewers and have done our best to address all comments (see below). Our replies to reviewers’ comments are shown in RED.

Anonymous Referee #1

This paper by Hertkorn et al. uses analyses of optical properties (absorbance, EEMPARAFAC, etc), NMR, and FTICR/MS to characterize and compare dissolved organic matter from three subtropical wetlands. The results of these analyses are very comprehensive and reported in detail throughout the manuscript, and on the whole these data are an important contribution to the characterization of wetland DOM. The novelty of combining detailed NMR and FTICR/MS analyses to environmental DOM samples was previously demonstrated in Hertkorn et al., 2013, but the application to a different system (wetland DOM) plus the addition of optical property analyses and comparison between sites makes this a valuable contribution to the field. The NMR results in particular are extremely detailed, especially in comparison to the other two analyses, and as such I recommend some reorganization and editing to the manuscript. My specific comments are detailed below. If these moderate revisions are addressed, I would recommend publication in Biogeosciences.

We appreciate the constructive comments from this reviewer.

Major comments:

The abstract is too long and needs to be shortened - in particular the amount/length of the results presented in the abstract needs to be reduced and/or summarized more efficiently.

Done

Primarily because the results of the NMR analyses are much more extensive than the other two analytical approaches, the results and discussion section seems disjointed and would benefit from reorganization. As sections 3.2 to 3.6 are exclusively about NMR, the headings of these sections should reflect that explicitly - I would suggest three parts to the results and discussion section: optical properties, FTICR/MS, and NMR. Then sections 3.2-3.6 should be subsections under the NMR section (though there are also some issues with a few of these specific subsections which are mentioned below).

Done. The removal of original sections 3c and 3f, which exclusively refer to NMR data, also restores better balance between NMR and FTMS section of this contribution. We now have 4 NMR and 4 FTMS figures in the main text as well as similar word count for both NMR and FTMS description.

Section 3.3 seems out of place within this manuscript. A test of the extraction selectivity is not mentioned in the abstract, introduction, or conclusions as part of this study and is not addressed in the methods section. If the authors wish to include this test in this paper, it needs to be evaluated specifically with regards to wetland DOM. Also the extraction selectivity is only tested for the NMR analyses - what are the implications for
FTICR/MS analysis, which also was performed on PPL extracts? Another option would be to move this discussion to the supplemental material along with Figure 3.

Done. We have removed the original section 3c entirely: Extraction selectivity of DOM. This section had in fact been introduced at a late stage of the manuscript to address the issue of the origin of methoxy groups in SPE-DOM. As it does not directly contribute to the main story of wetland SPE-DOM characterization, this omission has no other consequences on the residual text.

The comparison to an open ocean sample (section 3.6) is useful given the biogeochemical links between wetland and marine systems, but this section again only includes the NMR analyses. A few references to this open ocean sample are included in the FTICR/MS section, but the comparison should ideally include all three analytical approaches and be combined into one separate section along with any biogeochemical implications from the data (possible DOM degradation signatures, terrestrial vs. marine components, etc). If this comprehensive comparison is not available, this section should become a sub-section under the NMR results and be clearly identified as such.

Done. We have moved the section 3f (Comparison of wetland SPE-DOM FCL-L with a South Atlantic open ocean SPE-DOM) entirely to the Supporting Online section, where it resides along with Fig. S5. We rate this comparison still important, because no other 2D NMR spectra of sp²-hybridized carbon in any SPE-DOM are available in the current literature. Nevertheless, this comparison can be clearly separated from the main text without loss of coherence. We have inserted a small text section which leads to the respective SI text.

Is it possible to quantify the differences in NMR analyses between sites? The cluster analysis (Figure 9, panel A) from the FTICR/MS data is a nice visualization of the quantitative differences between sampling sites, and something similar that included all the results of all the analyses (if possible) would really strengthen the authors’ conclusion that DOM composition varied among wetlands. The quantitative extent of differences among samples can be difficult for readers less well-versed in NMR to grasp from simple comparisons (i.e., FCE-L > FCE-S > PAN > OKA). The data presented in Table 2 helps with this.

We have added a ¹H NMR derived hierarchical cluster analysis (HCA) and a principal component analysis (PCA) of the six wetlands SPE-DOM in the new Figure S2.

A summary or conclusions section is missing though section 3.8 gives a broader overview-type perspective to the results. This should be clearly titled as conclusions or a separate conclusions section added with the most pertinent results and their implications (or both; i.e., section 4 ‘Conclusions and biogeochemical significance’).

Done.
Minor comments:

- The entire manuscript should be thoroughly checked for consistent use (or non-use) of Oxford commas. Did best we can.

- Please check that all abbreviations are defined (for example, SUVA254 is not defined in the text, and EEM-PARAFAC is only defined in the abstract). Done.

- pg 13716, lines 5-6: this sentence is repetitive (South/southern Florida mentioned twice). Done.

- pg 13718, line 4: should this reference be Dittmar et al., 2008 in LO Methods? Done.

- pg 13720, line 25: a word is missing here, perhaps for? (‘. . . was acquired for an acquisition time of 1 s, . . .’) Done.

- pg 13725, line 19: should this word be ‘where’ rather than ‘were’? Done.

- pg 13732, line 26: Yuan et al., 2011 does not appear in the reference list. Done.

- pg 13744, line 7: move ‘that’ to before ‘many’ (‘. . . suggest, in agreement with previous reports . . ., that many of the bulk . . .’). Done.

- pg 13752, line 12: Yang et al., 2011 appears in the reference list but is not referenced in the manuscript (perhaps this should be Yuan from pg 13732 or vice versa?). Done.

- pg 13759, Figure 4: panels are out of order (A should be upper left) and the first part of the caption is confusing and should be reworded for clarity (panel B is SPE-DOM from which site?)

**Figure 3 (TOCSY NMR spectra of SPE-DOM) has been rearranged accordingly.**

- pg 13764, Figure 9 caption: extra word in (b) description, remove either ‘in’ or ‘to’ from ‘molecular compositions common in to all six wetlands SPE-DOM’. Done.

Anonymous Referee #2

In general, the paper is hard to follow with so much information describing some data and on the hand lacking a clear conclusion and explanation of the observations and linking it to the geochemical and the environmental conditions in each ecosystem. I think the authors lost an opportunity to compare and link between the NMR and the ICR results. Each section appeared to be separate from the section before it or after it.

The paper needs to be re-organized and focused on the main conclusions. I personally don’t think the optical priorities data are needed, …..
The optical properties data are directly connected to critical ecological functions of wetland organic matter. Absorption of electromagnetic radiation refers to $\pi$- and non-bonding electrons which also govern chemical reactivity of critical functional groups in natural organic matter. In addition, due to availability and ease of operation, many more publications refer to UV/VIS and EEM data than to meaningful FTMS and NMR descriptions of DOM. Hence, this contribution which offers a DOM characterization by all three methods will provide useful leads for future use of optical methods in conjunction with a deeper molecular assessment in DOM research. Besides, the comprehensive optical data in the manuscript indicate a well discussed and reasonable variance in-between DOM, even if intrinsic averaging characteristic of any optical spectroscopy of NOM will inevitably result in a restricted variance of these “bulk” parameters. In addition, studies combining optical techniques with molecular characterizations are becoming more commonplace in the literature, clearly reflecting the needs for detailed comparative studies. This is now highlighted in the Introduction with several references added.

.....as well as the comparison between the terrestrial DOM and the marine DOM samples. It might be worth mentioning how different these samples are from each other (which is not surprising), however I would only spend few sentences on that and not 2-3 paragraphs.

We have moved the section 3f (Comparison of wetland SPE-DOM FCL-L with a South Atlantic open ocean SPE-DOM) entirely to the Supporting Online section, where it resides along with Fig. S5. We rate this comparison still important, because no other 2D NMR spectra of sp$^2$-hybridized carbon in any SPE-DOM are available in the current literature. Nevertheless, this comparison can be clearly separated from the main text without loss of coherence. We have inserted a small text section which leads to the respective SI text.

This paper needs major changes. The data sets on the other hand are extensive and would be a great addition to the DOM society and a substantial contribution to scientific progress.

We appreciate this constructive comment and fully agree about the importance of this contribution. While the DOM community fully agrees that most of the DOM remains molecularly uncharacterized, this paper provides a wealth of such information, which should help in moving this field forward. As to changes in the manuscript, we are doing our best in accommodating the requests by the reviewers the best we can, without giving up important data.

Abstract: I don’t think this sentence add anything to the abstract and a result I suggest deleting it. “Analogies were such that an established excitation emission matrix fluorescence parallel 15 factor analysis (EEM-PARAFAC) model for the Everglades was perfectly applicable to the other two wetlands”).

We believe this sentence contains a very important message, particularly to those working on optical properties of DOM. We have re-phrased it to clarify.
The abstract contains a lot of information. I suggest rewriting it to summarize the main points of the study only.

**The Abstract was shortened.**

Introduction: The authors of the paper fail to include references of main groups that are known for using FT ICR MS, NMR in their organic matter characterization, rather, it appears that the authors are mainly referencing their papers.

We have included a large number of important references from other groups working on FT ICR/MS and NMR of DOM, but obviously can’t include them all. We have added some additional references that may have fallen short in the previous version of the manuscript.

Experimental: What time of the year did the sampling take place in the everglades? It was confusing to me which FCE samples were collected in 2011 and which was collected over the years. “Sampling was performed during 2011 for the Florida coastal Everglades (FCE; for SPE-DOM) and during the summer 2010 for the Pantanal (PAN) and the Okavango Delta (OKA) as part of on-going research programs. For EEM-PARAFAC comparisons, multiple samples collected monthly over several years (Chen et al., 2013) for the FCE, 38 samples collected along a trans-Okavango gradient (Cawley et al., 2012), and 22 samples collected in different sub-environments of the Pantanal wetland (rivers, lagoons, marshes; unpublished) were used to assess differences and similarities.”

How representative are these 6 samples to their ecosystems? The number of samples chosen for this study is too low for such a detailed comparison. Are these the samples that were collected for the ICR? Or these are the summary of the 38 samples?

I am still confused on which samples are included in this study. “For the Florida Coastal Everglades (FCE), samples were collected from the freshwater marsh, peat-soil dominated Shark River Slough (FCE-L) and the marl-soil dominated Taylor Slough (FCE-S), from the Okavango Delta (OKA) seasonal floodplain (OKA-L) and occasional floodplain (OKA-S) along the Boro River (Cawley et al., 2012), and the Paraguay River (PAN-L) and a wetland channel in Pantanal National Park (PAN-S; Chacra de Solange) for the Pantanal (PAN).”

The sampling, site selection and number of samples analyzed for optical properties and for SPE-DOM have been clarified.

How long were the extracts stored in the freezer? From my experience, even if the samples are stored in MeOH, you will see changes in OM composition with time; “the isolated SPE-DOM extracts (referred from here on as DOM for the NMR and FT-ICRMS data) were stored in pre-combusted glass vials and kept in a freezer until analyzed.”

Storage periods ranged from a few weeks to a few months; on reacquisition of $^1H$ NMR spectra of SPE-DOM dissolved in CD$_3$OD we could not observe visible alterations upon direct overlay even in 800 MHz $^1H$ NMR spectra acquired with cryogenic detection on storage time intervals exceeding several months.
How was the concentration of the extracts in an organic solvent measured? “Diluted SPEDOM (5 μgmL−1 15 in methanol) were injected into the electrospray source using a microliter pump at a flow rate of 120 μL h−1 with a nebulizer gas pressure of 138 kPa and a drying gas pressure of 103 kPa”.

Upon preparation of NMR samples, defined aliquots of methanolic SPE-DOM solution are dried and weighed on the analytical balance; this provided the dilution ratio employed in methanol solutions used for FTICR mass spectrometry. In addition, this dried sample in standard methanol is exchanged 3-4 times with CD₃OD (for several times); the integration of residual CD₂HOH and its ¹³C satellites and SPE-DOM ¹H NMR resonances provides an alternative lead to the DOM content.

What was the number of unassigned peaks? “Final formulae were generated and categorized into groups containing CHO, CHNO, CHOS or CHNOS molecular compositions which were used to reconstruct the group-selective mass spectra (Schmitt-Kopplin et al., 2010).”

Because we use rather strict and conservative exclusion criteria, about 50% of mass peaks were assigned to CHO, CHNO, CHOS and CHNOS molecular compositions, respectively. We have abstained from providing the count of total mass peaks; nevertheless, we have added a respective note in the experimental section.

Results section: I don’t think the optical properties added anything to this study. I think the problem is with the low number of samples used for comparison. Considering that the authors spend more time and space explaining the NMR and the ICR data, I would suggest removing the optical properties data to the supporting information.

See above the explanation why we consider this information should stay in the manuscript.

Why was the NMR done on the SPE samples? Why not the original samples? Freeze drying the samples then dissolving them in D2O would be less biased than using SPE.

For two reasons: (a) Field logistics – for both the Pantanal and the Okavango, due to the remoteness of the sampling locations, and logistical problems of air travel with large water samples, the samples were filtered and extracted (PPL) in Brazil and Botswana respectively. (b) While a certain extraction selectivity of PPL solid phase extraction is acknowledged, SPE nevertheless shows appreciable carbon yield and produces materials with excellent properties to be used in FTICR mass spectrometry and NMR spectroscopy. Freeze-drying of natural organic matter solutions on the other hand provides rather variable results depending on the extent of organo-mineral interactions and hydrophobic collapse which may impede complete re-dissolving of organic matter. Residual salt interferes with both NMR spectroscopy (metal coordination of organic ligands affects relaxation time and causes line broadening or disappearance of 2D NMR cross peaks) and FTICR mass spectroscopy (metal salt clusters commonly ionize much better than organic compounds, thereby suppressing mass peaks from original organic matter). Therefore, in the interest of comparability of all wetland samples, SPE was selected as the most appropriate method of DOM isolation in this study.
The IHSS extraction protocol utilizes strong acids and bases which could (or is) altering the composition of organic matter. I personally don’t recommend this protocol.

This section has been removed.

Can you elaborate on what you mean by “identical source materials”. A sample description would be useful here.

This section has been removed.

The whole section “Extraction selectivity of dissolved organic matter” appears out of place. Maybe move to supporting information and summarize main conclusion in the manuscript? Maybe the methods section?

We have removed this section entirely: Extraction selectivity of DOM. This section had in fact been introduced at a late stage of the manuscript to address the issue of the origin of methoxy groups in SPE-DOM. As it does not directly contribute to the main story of wetland SPE-DOM characterization, this omission has no other consequences on the residual text.

The NMR section is very detailed. It is very easy to get lost. Why not focus on the main results and conclusions only.

NMR and FTMS record and express SPE-DOM molecular characteristics from complementary perspectives. (Non-destructive) NMR spectroscopy samples atom-specific data of intact molecules which are reconstructed into SPE-DOM substructures; and several distinct NMR spectra have to be acquired to enable these conclusions. In this contribution, 5 different types of complementary 1D and 2D NMR spectra of wetland were discussed (some more had been acquired and were not even mentioned or discussed at all).

In contrast to the very information-rich NMR spectra, in which even subtle molecular differences are manifest from primary data, FTICR mass spectra suffer from extensive projection of structural diversity onto single mass peaks (which represent entire molecular ions composed of many atoms). Initially, ionization selectivity within very complex mixtures causes non-appearance of mass peaks altogether; otherwise, projection of all isomeric molecules, which in case of SPE-DOM likely range in the hundreds to thousands per any single recorded mass peak, represents a strong case of intrinsic averaging which has a tendency to make all FTICR mass spectra of (very complex) SPE-DOM look rather similar. This intrinsic disparity of SPE-DOM molecular complexity and FTMS resolution makes elaborate mathematical data treatment indispensable, to ensure an adequate compositional distinction of wetland SPE-DOM. This has been done in this contribution and has revealed remarkable distinction of OKA, PAN and FCE wetland SPE-DOM not otherwise available.

With the removal of the NMR based section “Extraction selectivity of SPE-DOM” altogether and transfer of the section “Comparison of wetland and marine SPE-DOM…” into the Supporting Online Material we have restored the balance between NMR and
FTMS sections, both with respect to numbers of figures (4 each) and word count (about equal).

What do you attribute the reason of the high number of unassigned formulas? (table 3) DID you compare the samples using Kendrick plots that utilize all the observed masses?

See reply above

A PCA or NMDS or both the NMR and the ICR data could be useful.

We have added a $^1$H NMR derived hierarchical cluster analysis (HCA) and a principal component analysis (PCA) of the six wetlands SPE-DOM in the new Figure S2.

Biogeochemical significance: The biological significant paragraph fails to explain the main factors behind the differences or lack of differences observed between the samples. A summary of the main conclusions is lacking. Each section in this manuscript appears to be separate from the other and rarely linked to one another. A paragraph that links all the observations from different techniques is lacking.

While we do not agree with this statement, we have modified this section hoping it will be more comprehensive now.
Molecular characterization of dissolved organic matter from subtropical wetlands:  
A comparative study through the analysis of optical properties, NMR and FTICR/MS.  

N. Hertkorn¹, M. Harir¹, K. M. Cawley², P. Schmitt-Kopplin¹, R. Jaffé²*  

¹ Helmholtz Zentrum Muenchen, German Research Center for Environmental Health,  
  Research Unit Analytical Biogeochemistry (BGC), Ingolstaedter Landstrasse 1, D-85764  
  Neuherberg, Germany  
² Southeast Environmental Research Center, and Department of Chemistry and Biochemistry,  
  Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA  

* Corresponding author: R. Jaffé (jaffer@fiu.edu; 305.348.2456)  

Abstract: Wetlands provide quintessential ecosystem services such as maintenance of water  
quality, water supply and biodiversity, among others; however, wetlands are also among the  
most threatened ecosystems worldwide. Natural dissolved organic matter (DOM) is an abundant  
and critical component in wetland biogeochemistry. This study describes the first detailed,  
comparative, molecular characterization of DOM in sub-tropical, pulsed, wetlands, namely the  
Everglades (USA), the Pantanal (Brazil) and the Okavango Delta (Botswana), using optical  
properties, high field nuclear magnetic resonance (NMR) and ultrahigh resolution mass  
spectrometry (FT-ICRMS), and compares compositional features to variations in organic matter  
sources and flooding characteristics (i.e. differences in hydroperiod). While optical properties  
showed a high degree of variability within and between the three wetlands, analogies in DOM  
fluorescence properties were such that an established excitation emission matrix fluorescence  
parallel factor analysis (EEM-PARAFAC) model for the Everglades was perfectly applicable to  
the other two wetlands. Area-normalized ¹H NMR spectra of selected samples revealed clear  
distinctions of samples while a pronounced congruence within the three pairs of wetland DOM  
readily suggested the presence of an individual wetland-specific molecular signature. Within  
sample pairs (long vs. short hydroperiod sites), internal differences mainly referred to intensity  
variations (denoting variable abundance) rather than to alterations of NMR resonances.
positioning (denoting diversity of molecules). The relative disparity was largest between the
Everglades long and short hydroperiod samples, whereas Pantanal and Okavango samples were
more alike among themselves. Otherwise, molecular divergence was most obvious in the case of
unsaturated protons ($\delta_H > 5$ ppm). 2D NMR spectroscopy for a particular sample revealed a large
richness of aliphatic and unsaturated substructures, likely derived from microbial sources such as
periphyton in the Everglades. In contrast, the chemical diversity of aromatic wetland DOM likely
originates from a combination of higher plant sources, progressive microbial and photochemical
oxidation, and contributions from combustion-derived products (e.g. black carbon). FT-ICRMS
spectra of both Okavango and Pantanal showed near $57 \pm 2\%$ CHO, $8 \pm 2\%$ CHOS, $33 \pm 2$
CHNO, and $< 1\%$ CHNOS molecules, whereas those of Everglades samples were markedly
enriched in CHOS and CHNOS at the expense of CHO and CHNO compounds. In particular,
the Everglades short hydroperiod site showed a large set of aromatic and oxygen-deficient “black
sulphur” compounds whereas the long hydroperiods site contained oxygenated sulfur attached to
fused-ring polyphenols. The elevated abundance of CHOS compounds for the Everglades
samples likely results from higher inputs of agriculture-derived and sea spray derived sulphate.
Although wetland DOM samples were found to share many molecular features, each sample was
unique in its composition, which reflected specific environmental drivers and/or specific
biogeochemical processes.

Abstract: Wetlands provide quintessential ecosystem services such as maintenance of water
quality, water supply and biodiversity, among others; however, wetlands are also among the
most threatened ecosystems worldwide. They are usually characterized by high levels of natural
dissolved organic matter (DOM), representing a critical component in wetland biogeochemistry.
This study describes the first detailed, comparative, molecular characterization of DOM in sub-
tropical, pulsed, wetlands, namely the Everglades (USA), the Pantanal (Brazil) and the
Okavango Delta (Botswana), using optical properties, high-field nuclear magnetic resonance
(NMR) and ultrahigh resolution mass spectrometry (FT-ICRMS), and compares compositional
features to variations in organic matter sources and flooding characteristics (i.e. differences in
hydroperiod). While optical properties showed both similarities and differences between these
ecosystems, these differences were mainly based on the degree of aromaticity of the DOM.
Analogies were such that an established excitation emission matrix fluorescence parallel factor
analysis (EEM-PARAFAC) model for the Everglades was perfectly applicable to the other two
wetlands. High-field (500 and 800 MHz) NMR spectra with cryogenic detection provided exceptional coverage and chemical description of wetland solid phase extracted (SPE)-DOM. Area-normalized $^1$H-NMR spectra of selected samples revealed clear distinctions of samples along with pronounced congruence within the three pairs of wetland DOM. Within sample pairs (long vs. short hydroperiod sites), internal differences mainly referred to intensity variations (denoting variable abundance) rather than to alterations of NMR resonances positioning (denoting diversity of molecules). The relative disparity was largest between the Everglades long and short hydroperiod samples, whereas Pantanal and Okavango samples were more alike among themselves. Otherwise, molecular divergence was most obvious in the case of unsaturated protons ($\delta_H > 5$ ppm). The larger discrimination observed between $^1$H-NMR spectra of DOM from different wetlands in comparison with the intrinsic variance among DOM within each wetland readily suggests the presence of an individual molecular signature, characteristic of each particular wetland. 2D-NMR spectroscopy for a particular sample revealed a large richness of aliphatic and unsaturated substructures, likely derived from microbial sources such as periphyton in the Everglades. In contrast, the chemical diversity of aromatic wetland DOM likely originates from a combination of higher plant sources, progressive microbial and photochemical oxidation, and contributions from combustion-derived products (e.g. black carbon). In addition, FT-ICRMS spectra allowed far-reaching classifications of wetland DOM. While DOM of both Okavango and Pantanal showed near $57 \pm 2\%$ CHO, $8 \pm 2\%$ CHOS, $32 \pm 2$ CHNO, and $< 1\%$ CHNOS molecules, the mass spectra of Everglades samples were fundamentally different compared to those as well as among long and short hydroperiod samples, as they were markedly enriched in CHOS and CHNOS at the expense of CHO and CHNO compounds. Here, four groups of CHOS molecules were differentiated as (a) saturated sulfolipids, (b) unsaturated sulfolipids, (c) molecularly diverse DOM-type CHOS molecules, (d) and particularly enriched in the Everglades short hydroperiod site, a large set of aromatic and oxygen deficient “black sulphur” compounds. The significantly higher proportion of CHOS compounds in general for the Everglades samples is likely the result of higher inputs of agriculture-derived and sea spray-derived sulphate to this wetland compared to the others. Although wetland DOM samples were found to share many molecular features, each sample was unique in its composition, which reflected specific environmental drivers and/or specific biogeochemical processes.
1) Introduction:

Natural dissolved organic matter (DOM) is a critical component of the global carbon cycle (Battin et al., 2009) and serves as an energy resource fuelling the microbial loop (Amon and Benner, 1996a), acts as a carrier facilitating the mobilization of trace metals and combustion derived products (Yamashita and Jaffé, 2008; Jaffé et al., 2013), and functions as a sun screen for aquatic organisms by limiting light penetration (Blough and Green, 1995; Foden et al., 2008) among other biogeochemical processes. In addition to comprising one of the largest organic matter pools in aquatic environments, DOM is one of the most complex mixtures of OM in natural systems containing millions of organic compounds (Koch et al., 2005; Hertkorn et al., 2008). While the molecular characterization of DOM (Hertkorn et al., 2006 and 2013; Jaffé et al., 2014; Woods et al., 2011 and 2012; Panagiotopolous et al., 2007; Lam et al., 2007; Aluwihare and Repeta, 1999) has significantly advanced our understanding of its composition and ecological functions, a significant portion of this material remains uncharacterized at the molecular level. Although molecular similarities between bulk DOM from vastly different environments have been reported (Repeta et al., 2002; Perdue and Ritchie, 2003; Jaffé et al., 2012; Hertkorn et al., 2013), the variability in composition (quality) among samples can also be quite significant (Jaffé et al., 2008; Zhang et al., 2014) implying differences in the photo- and bio-reactivity of these materials (Amon and Benner, 1996b). Such compositional differences (or similarities) may have important implications with regards to carbon cycling and ecological functioning of DOM. While the characterization of DOM using targeted substrates such as amino acids (Yamashita and Tanoue, 2003), neutral sugars (Panagiotopolous et al., 2007), lignin phenols (Spencer et al., 2012) and others have actively been pursued, much of the bulk DOM remains uncharacterized (Hedges et al., 2000) and broader spectrum analyses are required. As such, multi-analytical approaches for the advanced molecular characterization of DOM are needed to advance this field (Hertkorn et al., 2013; Jaffé et al., 2012; Minor et al., 2014).

During synthesis of natural dissolved organic matter (DOM), common biosignatures characteristic of the respective sources are progressively attenuated by the combined action of biotic and abiotic reactions. While degradative analysis of DOM intentionally destroys the sample in the beginning to recover a suite of known decomposition products, such as amino acids, carbohydrates, lignin phenols and lipids (which typically account for about 5-30% of the
organic carbon, depending on age, environment and diagenesis), non-target molecular-level analyses of DOM attempt to characterize the entire carbon present in DOM by means of information-rich detection methods, such as UV-Vis, fluorescence, NMR spectroscopy and FT-ICRMS. While optical properties have been widely applied for DOM bulk characterizations (Jaffé et al., 2014; Fellman et al., 2010) and allow large sample throughput at low cost, more advanced molecular level characterization techniques such as NMR provide unsurpassed insight into close-range molecular order (Hertkorn et al., 2006 and 2013; Lam et al., 2007; McCaul et al., 2011; Woods et al., 2011 and 2012; Zhang et al., 2014; Bell et al., 2015), while FT-ICRMS provides depiction of the compositional space with exceptional resolution (Kujawinski, 2002; Sleighter and Hatcher, 2007; D’Andrilli et al., 2010; Hertkorn et al., 2008; Hertkorn et al., 2013; Kaiser et al., 2003; Minor et al., 2014; Koch et al., 2005 and 2007). The combination of such techniques in the assessment of DOM dynamics has become more frequent (e.g. Tfaily et al., 2015; Jaffé et al., 2012), and furthermore, the association between optical properties and the molecular characteristics of DOM have recently become an active research endeavour in an attempt to better link these parameters (Stubbins et al., 2014; Kellerman et al., 2015; Wagner et al., 2015). As such, this work should provide further advances in this field.

While significant efforts have been devoted to the detailed characterization of DOM in oceanic, lacustrine and riverine environments (Hertkorn et al., 2013; Kujawinski et al., 2009; Einsiedl et al., 2007; Minor et al., 2012; Jaffé et al., 2012), still little is known about its molecular features in large freshwater wetlands, environments that are critically threatened by anthropogenic influences such as pollution and drainage for flood control, agricultural and urban development. Organic matter dynamics in large wetlands are particularly complex (e.g. Chen et al., 2013; Yamashita et al., 2010; Cawley et al., 2012) due to a high variability in spatial and temporal organic matter sources, concentrations, and diagenetic transformations. These variations are to a large extent driven by interplay between complex hydrological and primary productivity patterns. In this study, DOM samples from three of the largest and most important sub-tropical, pulsed wetlands, the Everglades (USA), the Pantanal (Brazil) and the Okavango Delta (Botswana), were collected and analysed on a comparative basis using optical properties, including EEM-PARAFAC, high field $^1$H NMR, and FT-ICRMS in order to assess similarities and differences in DOM composition and molecular structure in such vital ecosystems.
2) Experimental:

2a) Site descriptions, sample collection and analysis: The Everglades, Okavango Delta and Pantanal are three of the largest sub-tropical, pulsed, freshwater wetlands in the world and represent a wealth of biodiversity (Junk et al., 2006a). The Everglades ecosystem, located in South Florida, USA, is a large (610,483 ha) subtropical wetland located in southern Florida, USA (Figure 1). Annually, the southern section of the system, namely Everglades National Park receives ca. 120 cm of precipitation (50-year averages from 1962-2012) with 21 cm falling during the dry season (December to April) and 99 cm falling during the wet season (May to November) (Southeast Regional Climate Center, http://www.sercc.com). The freshwater area of the Everglades consists primarily of grassy marshes dominated by sawgrass (Cladium jamaciensis) with some small stands of trees on higher ground. The freshwater marshes drain through two main slough areas, namely the peat soil dominated Shark River Slough and the less extensive, marl-soil based Taylor Slough, which are characterized by longer and shorted hydroperiods (time and depth of inundation), respectively.

The Okavango Delta is a large wetland located in semi-arid NW Botswana and is subject to an annual flood event generated by water of the Okavango River flowing south from the highlands of Angola. During the flood event, the inundated area in the Delta expands in size from the annual minimum of 3,500-6,000 km² to the annual maximum of 9,000-13,000 km² (Gieske, 1997; McCarthy et al., 2003). About 88 % of inflowing water leaves the wetland through evaporation (Wolski et al., 2006). Flood water moves in the Okavango Delta as a combination of channel and floodplain flows. Several zones featuring differences in hydroperiod due to the seasonality of inundation are categorized as the panhandle, permanent swamp, seasonal floodplains, and occasional floodplains (Gumbricht et al., 2004; Cawley et al., 2012). The permanent swamp is characterized by extensive peat development and dominated by Phragmites australis and C. papyrus (Ellery et al., 2003; Mladenov et al., 2007). The seasonal floodplains are less peat rich and support mostly emergent sedges and aquatic macrophytes, while the occasional floodplains, characterized by the shortest hydroperiod are dominated by aquatic grasses.

The Pantanal is a large inland wetland of ca. 160,000 km², located mostly in SE Brazil, but also extends into Bolivia and Paraguay (Junk and Cunha, 2005; Junk et al., 2006b). The regional geology (depression) features natural levees along stream channels, and thus the
wetland is comprised of a labyrinth of large river channels, small streams, canals, and lagoons. Climate conditions lead to clear wet and dry seasons creating a monomodal flood pulse system. The wetland discharges about 80% of its water to the Paraguay River in the southern section of the system. The climate is tropical to sub-tropical with a large number of habitats including savannas and dry forests, leading to broad species diversity. While still mostly pristine, the expansion of cattle ranching surrounding the protected national park and hydrological modifications in the greater watershed have been suggested as potential threats to this ecosystem (Junk and Cunha, 2005).

Surface water grab samples for the three above-described wetlands were collected in pre-cleaned, brown plastic bottles (60 ml for DOC and optical properties; 2 L for solid phase extracts - SPE-DOM), placed on ice and filtered through GFF (0.7 μm nominal pore size), pre-combusted glass fiber filters within 6 hours after collection. For EEM-PARAFAC comparisons, multiple samples collected monthly over several years for the FCE (n = 858; Chen et al., 2013), samples collected along a trans-Okavango gradient (n = 38; Cawley et al., 2012), and samples collected in different sub-environments of the Pantanal wetland (n = 22; rivers, lagoons, marshes; unpublished) were used to assess differences and similarities in the fluorescence character of the DOM. Sampling for SPE-DOM was performed during summer 2011 for the Florida coastal Everglades (FCE; for SPE-DOM) and during the summer 2010 for the Pantanal (PAN) and the Okavango Delta (OKA) as part of on-going research programs. Only two SPE-DOM samples from each wetland were selected for detailed NMR and FTICRMS analyses, and consisted of one sample characteristic for long hydroperiod (-L) and one for short hydroperiod (-S) environments for each wetland respectively. For EEM-PARAFAC comparisons, multiple samples collected monthly over several years (Chen et al., 2013) for the FCE, 38 samples collected along a trans-Okavango gradient (Cawley et al., 2012), and 22 samples collected in different sub-environments of the Pantanal wetland (rivers, lagoons, marshes; unpublished) were used to assess differences and similarities. Six subsamples from this set were selected for the NMR and FTICR/MS study at long (-L) and short (-S) hydroperiod sites; two from each wetland.

For the Florida Coastal Everglades (FCE), samples were collected from the freshwater marsh, peat-soil dominated Shark River Slough (FCE-L) and the marl-soil dominated Taylor Slough (FCE-S), from the Okavango Delta (OKA) seasonal floodplain (OKA-L) and occasional floodplain (OKA-S) along the Boro River (Cawley et al., 2012), and the Paraguay River (PAN-
and a wetland channel in Pantanal National Park (PAN-S; Chacra de Solange) for the Pantanal (PAN). **Representative sample selection for long and short hydroperiod sites was based on previous reports for the FCE and OKA** (Chen et al., 2013; Cawley et al., 2012), **and advise from local wetlands scientists for the PAN** (C. Nunes da Cunha personal communication). The filtered samples were subjected to SPE isolation (Dittmar et al., 2008). Briefly, samples were acidified to a pH 2 using concentrated HCl. DOM in the acidified samples was extracted using PPL (Varian Bond Elut) cartridges and eluted with methanol (Optima, Fisher). The isolated SPE-DOM extracts (referred from here on as DOM for the NMR and FT-ICRMS data) were stored in pre-combusted glass vials and kept in a freezer until analyzed. Milli-Q water was used as a procedural blank and no contamination was observed. DOC measurements were made within three weeks of sample collection at the Southeast Environmental Research Center’s water quality lab at Florida International University with a Shimadzu TOC-V CSH TOC analyzer using a high temperature combustion method.

2b) **Optical properties analyses:** UV-Vis absorbance scans for filtered samples were collected on a Varian Cary 50 Bio spectrophotometer and collected over a range of 200 nm to 800 nm in a 1-cm quartz cuvette. The optical proxy for molecular weight ($S_R$) and the fluorescence index (FI) were determined as described in the literature (Helms et al., 2008; McKnight et al., 2001 respectively), where the $S_R$ value is inversely proportional to the DOM molecular weight, and FI values determined by the ratio of 470/520 nm emission at 370 excitation (Jaffé et al., 2008) can range between 1.4 and 1.9 for soil/terrestrial higher plant and microbial DOM sources, respectively. A blank scan (Milli-Q water) was subtracted from each sample spectrum and spectra were baseline normalized using the average absorbance between 700-800 nm. The absorbance at 254 nm ($A_{254}$) was also determined and normalized to DOC to obtain standard UV absorbance values ($SUVA_{254}$; values (Weishaar et al., 2003). Samples were analyzed for fluorescence within two weeks of collection. Fluorescence EEMs were collected on a Horiba Jobin Yvon SPEX Fluoromax-3 spectrofluorometer using the methods of Maie et al. (2006) and Yamashita et al. (2010). Briefly, EEMs were collected over an excitation wavelength ($\lambda_{ex}$) range of 240 – 455 nm with an increment of 5 nm and an emission range of $\lambda_{em}$ + 10 nm to $\lambda_{em}$ + 250 nm with an increment of 2 nm in a 1 cm quartz cuvette. The excitation and emission slit widths were set to 5.7 nm and 2 nm, respectively. Fluorescence scans were collected in signal/reference...
ratio mode with an integration time of 0.25 s and reported in quinine sulfate units (QSU). EEMs were corrected for instruments optics and inner-filter effects according to Ohno (2002) and Raman normalized and blank subtracted using Matlab v2009a software. EEMs were modeled using Matlab v2009a and fit to an eight component PARAFAC model described in Chen et al. (2010) and Yamashita et al. (2010) that was comprised of FCE samples only.

2c) Nuclear magnetic resonance spectroscopy (NMR): $^1$H NMR detected spectra of methanolic DOM extracts were acquired with a Bruker Avance NMR spectrometer at 500.13 (1D NMR only) / 800.13 MHz ($B_0 = 11.7 / 18.7$ T) at 283 K from a few mg of solid obtained by evaporation of original methanol-$h_4$ solution, dissolved in approx. 130 µL CD$_3$OD (Merck. 99.95%). $^2$H solution with a 5 mm z-gradient $^1$H / $^{13}$C / $^{15}$N / $^{31}$P QCI cryogenic probe (90° excitation pulses: $^{13}$C ~ $^1$H ~ 10 µs) in sealed 2.5 mm Bruker MATCH tubes. $^{13}$C NMR spectra were acquired with a Bruker Avance NMR spectrometer at 500.13 / 800.13 MHz ($B_0 = 11.7 / 18.7$ T) at 283 K from a few mg of solid obtained by evaporation of original methanol-$h_4$ solution (1 s acquisition time, 14 or 19 s relaxation delay; Table S3). 1D $^1$H NMR spectra were recorded with a spin-echo sequence (10 µs delay) to allow for high-Q probe ringdown, and classical presaturation to attenuate residual water present “noesypr1d”, typically 512-2048 scans (5 s acquisition time, 5 s relaxation delay, 1 ms mixing time; 1 Hz exponential line broadening).

A phase sensitive, gradient enhanced TOCSY NMR spectrum with solvent suppression (dipsi2etgpsi19) was acquired for an acquisition time of 1 s, a mixing time of 70 ms, and a relaxation delay of 3 s. The one bond coupling constant $^1$J(CH) used in 2D $^1$H,$^{13}$C DEPT-HSQC spectra (hsqedetgpsi2.2) was set to 145 Hz; other conditions: $^{13}$C 90 deg decoupling pulse, GARP (70 µs); 50 kHz WURST 180 degree $^{13}$C inversion pulse (Wideband, Uniform, Rate, and Smooth Truncation; 1.2 ms); F2 ($^1$H): spectral width of 5981 Hz (11.96 ppm); 1.25 s relaxation delay; F1 ($^{13}$C): SW = 17607 Hz (140 ppm). HSQC-derived NMR spectra were computed to a 4096 × 512 matrix. Gradient (1 ms length, 450 µs recovery) and sensitivity enhanced sequences were used for all 2D NMR spectra. Absolute value JRES/COSY and phase sensitive echo-antiecho TOCSY spectra (with solvent suppression: jresgpprf, cosygpph19, dipsi2etgpsi19) used a spectral width of 9615.4 Hz [JRES (F1) = 50 Hz] and were computed to a 16384 × 2048 matrix [JRES (F1) = 128]. Similarity of $^1$H NMR spectra was computed from 0.01 ppm section integrals in the range $\delta_H = 0.5 – 9.5$ ppm, with exclusion of methanol and residual water (Bruker
AMIX software, version 3.9.4) with Hierarchical Cluster Explorer (HCE); similarity versus
distance metrics used Pearson correlation coefficients; minimum similarity values are provided
in Fig. S2A. Other NMR acquisition conditions are given in Supporting Online Tab. S3.

2d) FTICR mass spectrometry: Ultrahigh-resolution Fourier transform ion cyclotron mass
spectra (FT-ICRMS) were acquired using a 12 T Bruker Solarix mass spectrometer (Bruker
Daltonics, Bremen, Germany) fitted with an electrospray ionization source in negative mode.
Diluted SPE-DOM (5 µg/mL in methanol) were injected into the electrospray source using a
micro-liter pump at a flow rate of 120 µL/h with a nebulizer gas pressure of 138 kPa and a
drying gas pressure of 103 kPa. A source heater temperature of 200°C was maintained to ensure
rapid desolvatation in the ionized droplets. Spectra were first externally calibrated on clusters of
arginine in MeOH (0.57 µmol/L) and internal calibration was systematically done in the
presence of natural organic matter reaching accuracy values lower than 500 ppb. The spectra
were acquired with a time domain of 4 megawords and 1000 scans were accumulated for each
spectrum. Calculation of elemental formulas for each peak was done in a batch mode by an in-
house written software tool. The generated formulae were validated by setting sensible chemical
constraints [N rule – O/C ratio ≤ 1 – H/C ratio ≤ 2n + 2 (CₙH₂ₙ₊₂). Element counts: C ≤ 100, H ≤
200, O ≤ 80, N ≤ 3, S ≤ 2, P ≤ 1 and mass accuracy window (set at ± 0.5 ppm)]. Final formulae
were generated and categorized into groups containing CHO, CHNO, CHOS or CHNOS
molecular compositions which were used to reconstruct the group-selective mass spectra
(Schmitt-Kopplin et al., 2010). The computed average values for H, C, N, O and S (atom %) and
the H/C and O/C ratios were based upon intensity-weighted averages of mass peaks with
assigned molecular formulae, which comprised ~50% of observed mass peaks.

3) Results and Discussion:

3a) Optical properties: Optical properties consisting of A₂₅₄, SUVA₂₅₄, Fl, S₉, and EEM-
PARAFAC for the six samples for SPE-DOM analysis are presented in Table 1. Although large
differences in DOC and A₂₅₄ were observed for the different samples (DOC range of 5.8 to 28.6
ppm; A₂₅₄ from 0.202 to 0.844) some of the qualitative optical parameters such as the S₉ values
(range 0.91 to 0.98) and the Fl values (range 1.30 to 1.44) all fell into a relatively narrow range.
In contrast, the SUVA\textsubscript{254} values covered a larger range from 2.72 to 5.11. A linear correlation was observed between the DOC and the A\textsubscript{254} (r\textsuperscript{2}=0.95).

S\textsubscript{R} and FI values were quite similar for all sample pairs and among samples, suggesting that the molecular weight distribution and the soil/higher plant vs. microbial contributions were quite similar among these samples, or that the mineralization of wetland DOM leads to similar compositional features for systems with different organic matter sources. Detailed molecular characterizations of DOM in headwater streams from different climatic regions (biomes) have been reported to exhibit remarkably similar bulk characteristics, although site-specific features were also identified in each case (Jaffé et al., 2012). However, in the case of the SUVA\textsubscript{254}, some clear compositional variations between different wetland DOM became apparent, where the samples from the more strongly soil-OM (or peat) influenced, long hydroperiod sites, featured higher SUVA\textsubscript{254} values compared to those with larger microbial and emergent/aquatic plant influence, short hydroperiod sites. Indeed, SUVA\textsubscript{254} showed higher values for the peat-based FCE-L compared to the marl-based FCE-S, the Paraguay River sample PAN-L compared to the wetland channel PAN-S, and in the seasonally flooded Boro River floodplain OKA-L compared to the occasional floodplain OKA-S. Although the differences in FI for the PAN samples were not significant, FI values for the other sites were expectedly inversely correlated to SUVA\textsubscript{254} (Jaffé et al., 2008).

The application of the FCE PARAFAC model to the OKA and PAN samples resulted in an excellent fit leaving no significant residues and was properly validated. The application of the FCE PARAFAC model to assess fluorescence characteristics of DOM from other wetlands was previously reported for the Okavango Delta (Cawley et al., 2012). In addition, the distribution of the EEM-PARAFAC components was also surprisingly similar among the six samples with C1 being dominant, followed by C3 > C5 > C4 > C6 and C7 and with C2 and C8 showing the lowest relative abundance. This trend is consistent with previous reports for the greater Everglades ecosystem (Yamashita et al., 2010; Chen et al., 2013). In addition, the combination of the entire dataset for the larger sampling schemes (see above; Chen et al., 2013; Cawley et al., 2012) for the EEM-PARAFAC-database results for the three wetlands are shown in Figure 1. The data for FCE are presented as two separate sub-groups representing Everglades National Park (ENP) sites and the Water Conservation Area 2 (WCA2), an area located north of the ENP boundary where water resources are heavily managed and agricultural runoff is significant (Yamashita et
In general terms, no significant differences were observed in the PARAFAC component distributions between FCE, OKA and PAN, and the only difference of significance was the relative abundance of the C2 PARAFAC component which was higher in the WCA2 compared to all other study regions. In agreement with the above, comparing the EEM-PARAFAC distributions between and among the six stations, and amongst the larger datasets of collected surface water samples (Table 1; Figure 1), no statistically significant differences were observed, although the range in values was large. Component C2 has been suggested to be photo-chemically stable or possibly a photo-degradation product (Chen et al., 2010; Cawley et al., 2012; Chen and Jaffé, 2014) and has also been identified as derived from the oxidation of soil OM, being exported from the Everglades Agricultural Area (EAA; located to the north of the WCA; Yamashita et al., 2010). As such it is not surprising that the levels for C2 are enriched in waters from the WCA2, which receives significant canal inputs from the EAA. In the other wetlands and at freshwater marshes in more distant regions of the Everglades, C2 is only a relatively minor component of the DOM fluorescence signal. However, the most interesting aspect of this comparison is that the FCE-based PARAFAC model provided a perfect fit for both the OKA and PAN samples, suggesting that the overall fluorescent properties of the DOM in the three wetlands are quite similar.

3b) NMR study

NMR spectra of SPE-DOM: High field (800 MHz) NMR spectra with cryogenic detection performed on six samples (paired long and short hydroperiod sites from each wetland) revealed an exceptional coverage and chemical description of wetland organic proton and carbon chemical environments. The \(^1\)H NMR spectra of wetland DOM acquired with solvent suppression showed the prevalence of rather smooth bulk signal envelopes reflecting intrinsic averaging from massive signal overlap with a considerable variance in abundance for all major chemical environments. In addition, rather minor superimposed sharp individual NMR resonances were indicative of biological signatures and occurred in the order PAN > OKA > FCE (Fig. 2; Fig. S1). From higher to lower field (from right to left), abundant (a) aliphatics, (b) “acetate-analouges”, (c) carboxyl-rich alicyclic molecules (CRAM), (d) “carbohydrate-like” and methoxy, (e) olefinic, and (f) aromatic NMR resonances showed well visible and rather broad maxima (letters given according to Fig. S1). 

al., 2010).
Superimposed small NMR resonances indicative of comparatively abundant biological and biogeochemical molecules were most significant in the aromatic section (f), well noticeable in sections (e) and (a) and of continual lesser occurrence in the order c > b > d (Fig. S1). The area-normalized $^1$H NMR spectra of the six DOM samples (Fig. 2) showed more variance than their respective $^1$H NMR section integrals (Table 2), a plausible consequence of intrinsic averaging across sizable chemical shift windows (Hertkorn et al., 2007). One dimensional $^1$H NMR spectra of wetland SPE-DOM revealed clear distinctions according to sample location, with pronounced congruence between the three pairs of samples (Fig. S1). Within sample pairs, internal differences mainly referred to intensity variations (denoting variable abundance) rather than to alterations of NMR resonances positioning (denoting molecular diversity). The relative disparity was largest between both FCE-L and FCE-S whereas PAN and OKA samples were more alike among themselves. Otherwise, molecular divergence was most obvious in the case of unsaturated protons ($\delta_H > 5$ ppm). Subtle relative changes in composition between pairs of samples were readily visualized by superposition NMR spectra in which the relative NMR section integrals of each aromatic and aliphatic substructures had been normalized to 100% (Fig. S2). Subtle relative changes in composition between pairs of samples were readily visualized by superposition and difference NMR spectra in which the relative NMR section integrals of each aromatic and aliphatic substructures had been normalized to 100% (data not shown Fig. S2).

The larger discrimination observed between $^1$H NMR spectra of DOM from different wetlands in comparison with the intrinsic variance among DOM within each wetland already suggested presence of an individual molecular signature, characteristic of each particular wetland. Table 2 shows the respective $^1$H NMR section integrals for the six samples under study. Generally, the OCH, XCH and CCCH aliphatic chemical environments represented nearly equal contributions to make up approx. 90% of the spectrum with the CCCH units consistently exceeding 30%. Carboxyl-rich alicyclic molecules (CRAM) and functionalized and pure aliphatics followed the order FCE (L > S) > PAN $\approx$ OKA. Molecular divergence was most noticeable in the chemical environment of unsaturated protons, where the ratio of aromatic to olefinic protons declined in the order FCE $>$ PAN $>$ OKA. Here, H$_{ar}$ ($\delta_H > 7$ ppm) and C=CH, O=C=CH ($\delta_H : 5.3 – 7$ ppm) contributed less than 5% each to the overall spectra. Difference NMR spectra (L-S) obtained for FCE, OKA and PAN wetland SPE-DOM were computed from area-
normalized NMR spectra (Fig. S2C, S2D) and indicated congruent behaviour for OKA and PAN SPE-DOM in the purely aliphatic section ($\delta_H < 3$ ppm), with moderate increase of $C_nCH$ groups ($n > 1$; $\delta_H < 1.6$ ppm). The alterations in FCE-based aliphatics were governed by a marked increase of CRAM whereas the abundance of $C_nCH$ decreased (Fig. S2D). Interestingly, rather concordant decline of methoxy groups (primarily methyl esters; Fig. 5) was observed for both FCE and PAN (Fig. S2D). Polycarboxylated and PAH-derived aromatics ($\delta_H > 8$ ppm) were markedly increased in FCE-L as compared with FCE-S (cf. below).

For improved assessment of unsaturated protons, the respective chemical shift range was divided into several sections, comprising ($f_1$; letters according to Fig. S1) polycyclic and polycarboxylated aromatics as well as six-membered nitrogen heterocycles ($\delta_H > 8$ ppm); ($f_2$) electron withdrawing substituents (COX; Perdue et al., 2007; $\delta_H \approx 7.3 - 8.0$ ppm); ($f_3$) electroneutral substituents (alkyl, H, R; $\delta_H \approx 7.0 - 7.3$ ppm); ($f_4$) electron-donating substituents (OR, OH, phenolics; $\delta_H \approx 6.5 - 7.0$ ppm); ($e_1$) polarized and conjugated olefins; ($e_2$) isolated olefins ($\delta_H \approx 5.0 - 5.5$ ppm), this section features however contributions from anomeric protons and certain ester groups (cf. discussion of 2D NMR spectra). The relative and absolute abundance of electroneutral substituted and phenolic aromatic compounds were maximal in OKA, and declined through PAN to FCE. The ratio of conjugated olefins and aromatics was similar in FCE and PAN; however, the abundance of these units was lower by ca. 30% in FCE. DOM from FCE-L showed higher proportions of isolated olefins and, possibly, anomeric positions within carbohydrates.

Within this, the FCE samples showed the lowest proportion of unsaturated protons, and among them, the short hydroperiod site FCE-S was marginally depleted in abundance of carboxylated aromatic protons compared to the longer hydroperiod site FCE-L, possibly due to higher light exposure at the short hydroperiod site. Such differences among samples from PAN and OKA were not significant. Ratios of aliphatic to aromatic signals (CCC/$H_{ar}$; see data in Table 2) were also highest for the FCE samples, suggesting enrichment in microbial-derived DOC (periphyton sources) compared to the PAN and OKA samples, but also featuring differences between long and short hydroperiod sites, where preservation of aliphatics at long hydroperiod sites seemed to be favoured for all wetlands. These differences may at first conflict with previous reports where larger periphyton contributions to DOC at FCE-S compared to FCE-L (Chen et al., 2013) suggested to be related to drying and re-wetting of periphyton mats during
the dry-to-wet transition at FCE-S and higher relative contributions of soil-derived DOM in FCE-L compared to FCE-S. Similarly, in the case of the long and short hydroperiod comparison, the higher CCCH/Har ratios coincided with higher SUVA values for the DOM-L samples, suggesting a difference in the relative contribution of microbial vs. higher plant/soil derived DOM for CDOM compared to bulk DOM. CDOM, often used as a proxy for DOM only represents a small fraction of the bulk DOC and does not include aliphatic molecules as those determined here. As such, while being a convenient and useful proxy for DOC sources, CDOM-based measurements might be less sensitive for the evaluation of and have some limitations when compositional differences between similar samples are not very large.

Methoxy NMR resonances for FCE-S compared to FCE-L were not only more abundant, but were also shifted to lower field, indicating increased fractions of aromatic methylethers and methylesters. FCE-S undergoes periodic drying and thus exposure of soil OM (SOM) to atmospheric conditions and intense sunlight exposure of DOM after high evaporation (drying) conditions. As such, much of the SOM can be aerobically oxidized to CO₂ creating marl soils. It is thus plausible that increased aerobic microbial oxidation and photo-exposure at this short hydroperiod site might enhance DOM oxidation compared to the long hydroperiod site (FCE-L). In addition, while OKA showed an appreciable shoulder at δH > 3.75 ppm indicative of aromatic methyl esters and ethers at however, reduced relative abundance, this distinction was absent in both PAN and FCE (Fig. 2).

In addition to the characteristics described above, the FCE samples showed the largest proportion of aromatic compounds substituted with carbonyl derivatives (most likely carboxylic acids; δH > 7.3 ppm). This pattern is in accordance with the presence of dissolved black carbon (DBC) at these wetland sites, where the highest abundance was reported for the FCE samples (Ding et al., 2014a). The relatively large fraction of protons with very large downfield chemical shift (δH > 8 ppm) suggested the presence of six-membered nitrogen heterocycles as well as that of polycyclic aromatic hydrocarbons (PAH). These units followed the abundance order PAN > OKA > FCE and could be related in part to the presence of dissolved black nitrogen (DBN; Ding et al., 2014b). However, the ratio of olefinic protons (δH ~ 5.2 – 6.8 ppm) to aromatic protons (δH > 6.8 ppm; but see HSQC cross peaks; Fig. 5S3) followed the order FCE > PAN ≈ OKA. The distribution of aromatic protons in OKA indicated elevated abundance of electroneutral (alkyl, H; δH ≈ 7.0 – 7.3 ppm) and electron-donating substituents (OR, OH; δH < 7.0 ppm) in contrast to
both FCE and PAN SPE-DOM which showed similar distribution of aromatic protons with larger proportions of electron-withdrawing substituents (COR; δ_H > 7.3 ppm) at however, different overall abundance (Fig. 2; Table 2). In contrast, the abundance of aromatics with electroneutral (R) or electron-donating substitution (OR) with δ_H ~ 7.3 - 6.6 ppm (Perdue et al., 2007) followed the order OKA > PAN >> FCE (Fig. 2), likely reflecting the enhanced relative contributions of higher plant derived DOM (in different degrees of oxidation) for the OKA and PAN compared to the FCE. In conclusion, one-dimensional 1H NMR spectra show a considerable molecular divergence of aromatic molecules in the DOM of the three wetlands, where the compositional features seem driven by both source strengths and variations in biogeochemical processing.

Although some methoxy groups can be formed by reaction of hydroxyl groups in natural DOM and methanol during storage at ambient temperature (as SPE-DOM; Flerus et al., 2011), the HCO NMR section integral, which was found typically larger by ~ 2 % for the respective short hydroperiod samples (Table 2), might reflect larger abundance of native methyl esters at these sites or larger abundance of DOM methanolysis products (Fig. 3; Fig. S2).

3c) Extraction selectivity of dissolved organic matter: Solid organic matter isolates were prepared from (near) identical source materials FCE-S and FCE-L by means of IHSS extraction using XAD-8 resin (International Humic Substances Society), resulting mainly in a fulvic acid sample (FCE-S FA), by tangential ultrafiltration to produce an ultrafiltered DOM sample (FCE-S UDOM) and by solid phase extraction with PPL cartridges to produce a methanolic eluate (section FCE-S PPL). These three extracts have been dissolved in D2O and CD3OD, and their 1H NMR spectra were compared to evaluate the effects of extraction selectivity and to better assess the origin of methoxy groups in SPE-DOM (Fig. 3). As expected from previous characterization of OM isolates, substantial extraction selectivity was deduced from the 1H NMR section integrals (Table S1) and line shape distribution in the 1H NMR spectra of the individual samples. In addition, variance in NMR section integrals was also observed when 1H NMR spectra have been acquired from identical organic matter dissolved in D2O and CD3OD, respectively. Here, 1H NMR resonances of FCE-S PPL largely coincided in both solvents, whereas the abundance of OCH units in FCE-L FA was decreased in methanolic in comparison with aqueous solution (Fig. 3; Fig. S2; Table S1). Accordingly, functionalized and pure aliphatics of FCE-L FA showed
larger abundance in methanolic solution. This behavior follows the expected solubility characteristics of carbohydrates (preferentially water soluble) and lipids (preferentially methanol soluble) even if no appreciable insoluble residue was recognized at the rather low concentration used for the acquisition of the (very sensitive) 800 MHz $^1$H NMR spectra.

When comparing the different organic matter isolates, UDOM in D$_2$O showed a comparatively large contribution of carbohydrate NMR resonances, probably in conjunction with limited contributions from low amplitude methoxy NMR resonances from aliphatic methyl esters as seen in $^1$H-$^1$C HSQC-NMR spectra (data not shown). In contrast to UDOM, fulvic acid dissolved in D$_2$O and CD$_3$OD was severely depleted in carbohydrates, showing only a marginal hump in the $^1$H-NMR chemical shift section $\delta_H = 3.7–4.1$ ppm. SPE-DOM obtained from PPL resin, which had been exposed to methanol (CH$_3$OH) during initial elution, showed a large NMR resonance in both D$_2$O and CD$_3$OD ranging from $\delta_H = 3.5–4.1$ ppm, which probably were dominated by various methoxy groups as confirmed by HSQC NMR spectra (Fig. S2). This preliminary comparative assessment apparently indicates scavenging of reactive groups present in DOM by methanol during SPE preparation. In a previous experiment, Suwannee river fulvic acid has been reacted with sulphuric acid (H$_2$SO$_4$) to result in expedient formation of CHNOS compounds as detected by FTICR mass spectrometry. Consecutive addition of methanol had shown formation of a considerable abundant and diverse suite of aliphatic methyl esters as confirmed by $^1$H-$^1$C HSQC-NMR spectra (Fig. S16 in Schmitt-Kopplin et al., 2010).

UDOM observed in D$_2$O and CD$_3$OD differed considerably in relative abundance of carbohydrates and aliphatics, most likely a consequence of differential solubility of these respective DOM constituents (Table S1). In terms of relative abundance, UDOM in D$_2$O showed more than double the content of carbohydrate but less than half of that aliphatics; this is very likely a consequence of limited solubility of the respective UDOM fractions in aqueous or organic solvents even at the low concentration applied. In addition, mutual orientation and arrangement of lipid-type UDOM molecules appears rather different in D$_2$O and CD$_3$OD solution. In D$_2$O, the pure aliphatic substructures with $\delta_H < 1.8$ ppm form a rather contiguous and broad NMR resonance, with methyl groups resonating at $\delta_H < 1.1$ ppm producing a (upfield) shoulder of the main aliphatic NMR resonance ($\delta_H = 1.280$ ppm, with a shoulder at $\delta_H = 1.307$ ppm), whereas methyl derived $^1$H-NMR resonances clustered separately at $\delta_H = 0.85$ ppm. Otherwise, apart from the differential solubility of UDOM-related carbohydrates in either D$_2$O
(good solubility) and \( \text{CD}_{3} \text{OD} \) (limited solubility), discrimination in quality of aromatics and CRAM appeared much less pronounced than distinction in lipid abundance and arrangement.

In conclusion, the extraction method of organic matter from natural waters itself imposes a pronounced extraction selectivity which results in a very appreciable distinction of the different organic matter isolates. In addition, differential solubility of the DOM constituents in various solvents might require additional careful evaluation by non-selective analytical characterization.

In NMR spectra, mutual interaction of specific chemical environments in a complex mixture of organic molecules might result in further noticeable \(^1\text{H} \) NMR chemical shift displacements. Overall, methanolysis which is a possibly interfering chemical reaction during solid-phase extraction of organic matter (Flerus et al., 2011), will make otherwise NMR invisible (or inconspicuous) functional groups available for characterization—note, that the amplitude of the \( \text{OCH}_3 \) -singlet NMR resonance will be substantial even at comparatively modest turnover. A clear discrimination of natural and interfering methoxy groups in SPE-DOM would be available from careful labelling studies in conjunction with NMR spectroscopy and mass spectrometry.

\( ^{13} \text{C} \) NMR spectra: \(^{13} \text{C} \) NMR spectra of wetland DOM were not overly conspicuous, with limited variance of spectra appearance and \(^{13} \text{C} \) NMR section integrals (Fig. S2; Table S1). The abundance of non-functionalized aliphatics followed the order FCE-L > FCE-S > PAN > OKA, whereas aromaticity followed a near reverse order FCE-L ≈ FCE-S < OKA ≈ PAN. DOM from FCE-L showed depletion of carbohydrates and increase of lipid-like compounds (Table S2S1). The near invariant abundance of carbonyl derivatives (most likely carboxylic acids) for all DOM could imply that a sizable proportion of low field \(^1\text{H} \) NMR resonances with chemical shift \( \delta_1 > 7.3 \) ppm, which were more abundant in PAN than in the others (Fig. 2; see also aromatic TOCSY cross peaks, Fig. 43), actually represented (substituted) PAH (with \( \delta_c < 140 \) ppm; Hertkorn et al., 2013) rather than (poly)carboxylic aromatics (with \( \delta_c \sim 167 - 187 \) ppm; Fig. 2; Tab. 2; Fig. S3S2; Table S1). Computed average H/C ratios from a basic reverse \(^{13} \text{C} \) NMR based mixing model ranged in the order FCE-L > FCE-S > PAN-S = OKA-L (\(^{13} \text{C} \) NMR spectra of PAN-L and OKA-S were not acquired) and primarily reflected variable content of aliphatic structures (\( \delta_c \sim 0 - 47 \) ppm). The computed O/C ratio was near equal for the OKA, PAN and FCE-S samples, whereas that of FCE-L was lower by \sim 0.07 units. Here, a reduced abundance of oxidized
aliphatic units \( (\text{HC}_n\text{O}) \) was primarily responsible, because phenolic and carboxylic content followed the order \( \text{OKA-L} \approx \text{PAN-S} > \text{FCE} \).

\[ \text{2D NMR spectra:} \] The 2D NMR spectra provided remarkable richness in detail and refined preliminary assignment proposals from the one-dimensional \( ^1\text{H} \) and \( ^1\text{C} \) NMR spectra. TOCSY NMR spectra (Fig. \text{4A}) revealed a wide range of methyl groups \( \text{(H-C-H-X; X: C, O; Fig. 4A, section a); a contiguous, ill resolved cross peak reflected a large number of intra-aliphatic correlations (C-H-C-H-C; n = 0 - 2; Fig. 4A, section b), and fewer cross peaks in-between oxygenated aliphatics (O-C-H-C-O; Fig. 4A, \( \delta_H > 3.4 \text{ ppm} \)). Protons bound to sp\(^2\)-hybridized carbon produced better resolved TOCSY cross peaks and were part of various \( \alpha, \beta \)-unsaturated \( \text{olefin compounds} \) (Fig. \text{4B}, section c, d), oxygenated and carbonyl (COX) derivatives of benzenes with up to three COX substituents (Fig. \text{4C}, section f, g, h) as well as six-membered nitrogen heterocycles and more extended aromatic systems with up to several aromatic rings (Fig. \text{4B, 4C}, section f, g, h). As mentioned earlier, such compounds might be related to the presence of combustion-derived compounds such as DBC and DBN (Ding et al., 2014a and b) and even black sulfur DBS (Hertkorn et al., 2013; see attendant discussion of FTICR mass spectra). In contrast to common five-membered heterocycles, (di)benzothiophene derivatives exhibit NMR resonances ranging from \( \delta_H: 7.4 - 8.1 \text{ ppm; corresponding HSQC cross peaks of DBS would appear in section g, Fig. 4).} \]

HSQC NMR spectra of PAN and OKA did not show peculiar features which were observed in those of both FCE samples and therefore will be not discussed here. The HSQC NMR spectra of both FCE-S and FCE-L were remarkably similar and produced near-identical overlay NMR spectra with some discernible variance in HSQC cross peak amplitude rather than positioning (data not shown). This behavior is expected from comparison of the one-dimensional \( ^1\text{H} \) NMR spectra. These display differences in relative amplitude rather than positioning of NMR resonances which is indicative of variance in abundance of certain molecules rather than variance in molecular diversity (see, however, discussion of CHOS compounds present in FCE DOM as derived from FTICR mass spectrometry). About 90% of overall HSQC cross peak integral resided in a contiguous expansive superimposed assembly of HSQC cross peaks originating from protons bound to sp\(^3\)-hybridized carbon (Fig. S4).
The resolution of these expansive aliphatic HSQC cross-peaks of FCE-S (Fig. S4S3) could be remarkably improved by spectral editing according to carbon multiplicity (Fig. 6S). The combination of methyl- and methylene-selective DEPT-HSQC NMR spectra revealed well discriminated cross peaks for all three types of protonated carbon; i.e. methyl, methylene and methine (Fig. 6S). The chemical diversity of X-CH₃ groups as indicated by DEPT HSQC cross peaks (Fig. 4A; section a, Fig. 6S) was noteworthy, and the near Gaussian distribution of C-CH₃ cross peak amplitudes in ¹H and ¹³C NMR frequencies indicated near maximum diversity of aliphatic chemical environments associated with these methyl groups. However, classical methyl groups terminating extensive, purely aliphatic units (δH < 1.0 ppm; CCCCH₃ units) contributed less than 20% to the total CCH₃ HSQC cross peak integral. The large majority of C-CH₃ units was sufficiently proximate to carbonyl derivatives (i.e., most likely carboxylic acids) to let those experience downfield chemical shift anisotropy from these nearby carbonyl groups, resulting in chemical shifts ranging from δH ~ 1.0 - 1.7 ppm, respectively (cross peak a; Fig. 6S). Alicyclic structures (e.g. CRAM; Hertkorn et al., 2006) facilitate clustering of chemical environments as shorter paths of chemical bonds between different substituents are realized in rings rather than in open chains. Another ~20% of CCH₃ in FCE was bound to olefins [C=C-CH₃], with a possible contribution of S-CH₃ groups (section b; Fig. 6S).

The carbon bound methylene (C-CH₂-C) cross peak occupied an impressively large area down to δH ~ 3.5 ppm, well into the proton chemical shift range commonly attributed to OCH units. The two major chemical environments discriminated were methylene more distant to COX (C-CH₂-Cn-COX, with n ≥ 1, and δH < 2.1 ppm cross peak d; Fig. 6S), and methylene groups directly proximate to carboxylic groups (in α-position; i.e. C-CH₂-COX, with δH > 2.1 ppm cross peak e; Fig. 6S). The former shows a wider range of remote carbon substitution as indicated by the substantial spread of respective carbon chemical shifts (ΔδC: 24 / 16 ppm, respectively for section d / e HSQC cross peaks; Fig. 6S; see also Fig. 8b in Hertkorn et al., 2013). A wide variety of aliphatic and aromatic methylesters and methylethers were also found, the latter being virtually absent in marine SPE-DOM. Here, aliphatic methyl esters were most abundant (section g₂ in Fig. 6S), aromatic methyl esters (section g₃ in Fig. 6S) and methyl ethers (section g₄ in Fig. 6S); were of similar abundance, and clearly recognizable aliphatic methyl ethers were also present (section g₅ in Fig. 6S). Oxomethylene (OCCH₃) occurred in the form of carbohydrate side chains (section h; Fig. 6S), and a remarkable set of aliphatic oxomethylene...
(OCH₃) HSQC cross peaks ($\delta_{\text{H,C}} \sim 3.4 - 4.0/58 - 72$ ppm; section j in Fig. 65) was present in SPE-DOM FCE-S, which does not correspond to common lignin β-aryl ether units, which resonate in this $^1$H and $^{13}$C NMR chemical shift range, but commonly comprise $\text{C}_\text{ar}-\text{CH}-\text{O}-$, i.e. methine substructures. Analogous oxomethine substructures are also found in phenylcoumaran, resolin and dibenzodioxocin units as well, whereas oxomethylene units with $\delta_H > 4.5$ are rare in common lignins (Ralph et al., 1998; Yelle et al., 2008; Martinez et al., 2008; Wen et al., 2013; Yuan et al., 2011). This peculiar HSQC cross peak was discovered in FCE-S wetland SPE-DOM (section j HSQC cross peak in Fig. 65) but since then has also been observed (in retrospect) with lesser distinction in other SPE-DOM including those from marine sources. The singular positioning of a methylene group in the $^1$H and $^{13}$C NMR chemical shift space strongly restrains the potential diversity of its chemical environments: it has to represent a OCH$_2$ group (methylenic as defined by the phase in $^1$H, $^{13}$C DEPT HSQC NMR spectra; single oxygen because of $\delta_{C}$: any O-CH$_2$-O environment would resonate at $\delta_C > 90$ ppm). Similarly, common O-CH$_2$-N chemical environments would resonate at higher field than observed in both $\delta_{\text{H,C}}$, but cannot be excluded entirely in case of peculiar remote substitution. The most plausible substructure is OCH$_2$C; then, $\delta_H$ from 5.3 – 5.7 ppm warrants presence of an ester group: this implies a -C-(C=O)-O-CH$_2$-C substructure. However, alkylation alone will not produce the necessary low field $\delta_H$ observed. This leaves -C-(C=O)-O-CH$_2$-C=O as a plausible group; possibly confined with a carboxylic group such as -C-(C=O)-O-CH$_2$-COOH or as an ester -C-(C=O)-O-CH$_2$-COOR. Both these substructures have a decent propensity to form enols with variable double bond character -C-(C=O)-O-CH=CH(OH)$_2$. A partial double bond character, which might be possibly controlled by mutual interactions in the complex DOM mixture of molecules, would also explain the observed spread of chemical shift in $^1$H and $^{13}$C NMR frequencies in this HSQC cross peak even if the methylene group itself in -C-(C=O)-O-CH$_2$-COOH is four (carbon) or five (proton) bonds away from the most proximate atom position where substitution may affect its chemical shift.

Several thousands of acid and ester derivatives of acetoacetic acid [H$_3$C-(C=O)-O-CH$_2$-COOH] are known in literature. Here, many of the common esters comprise lipid substructures such as n-alkanes, sterane and other polyalicyclic hydrocarbons, trimethylammonium salts, among others, suggesting a natural origin of these compounds also in wetland SPE-DOM. While substructures with -O-CH$_2$-COOZ (Z: H, R) will produce distinct “oxomethylene (OCH$_2$C)”
cross peaks in \(^1\)H, \(^13\)C DEPT HSQC NMR spectra (section j cross peak; Fig. 65), the derivatives with \(-\text{O-CHCH}_n\text{-COOZ}\) substructures (Z: H, R) will contribute to the \(^1\)H NMR downfield section of the expansive “oxomethine (O\text{CHC})” \(^1\)H, \(^13\)C DEPT HSQC cross peak (section i cross peak; Fig. 65) and will not be readily discerned owing to a larger variance in remote substitution. In addition, oxomethylene units without geminal and vicinal adjacent protons will very likely produce intense singlet NMR resonances, contributing to the enhanced visibility of HSQC cross peaks even at rather limited relative abundance. Further evaluation of aliphatic spin systems in FCE-L provided evidence for massive aliphatic branching in C\text{CCH} units and of large chemical diversity of remote carboxylic substitution (Fig. S65).

TOCSY and HSQC NMR spectra demonstrated presence of olefinic and aromatic unsaturation in all wetland SPE-DOM (Fig. 4-3 and Fig. 54). The FCE-S showed the most informative detail of HSQC cross peaks arising from unsaturated C\text{sp2H} groups (Fig. 54). In comparison with marine SPE-DOM (cf. Fig. S4 and attendant discussion), wetland SPE-DOM displayed a more restricted chemical diversity of conjugated olefins (Fig. 54) whereas all kinds of oxygenated aromatics, i.e. those substituted with electron-withdrawing (e.g. COOH) and electron-donating (e.g. OR) substituents were much more abundant and chemically diverse in (all) wetland DOM (not all data shown). The latter finding is indicative of polyphenol input from vascular plants (e.g. lignin-derivatives) into wetland DOM whereas aromatics in marine SPE-DOM mainly reflect marine natural products (Fig. S54). In general, aromatic unsaturation (as deduced from proton NMR integrals; Table 2) followed the order PAN > OKA > FCE (Fig. S1), whereas olefinic unsaturation followed the order OKA ~ PAN > FCE (Fig. S1). Aliphatic to aromatic ratios changed across the different samples with an order of FCE > PAN > OKA, suggesting higher relative contributions from periphyton in the FCE, whilst the PAN and OKA were more influenced by higher plant-derived organic matter including lignins. The olefinic to aromatic ratios (FCE-L: 0.44; FCE-S: 0.39; 1.5; PAN: 0.384-4; OKA: 0.414-4) were computed from adapted \(^1\)H NMR section integrals [\(\delta_H: 10 – 6.5 \text{ ppm (aromatics) / 6.5 – 5.0 ppm (olefins)}\); Table 2] owing to HSQC cross peak positioning which indicated major contribution of oxygenated aromatics C\text{arO} at \(\delta_H: 7.0 – 6.5 \text{ ppm; Fig. 4}\) and showed in the lower values than range compared to similar determinations for oceanic DOC (Hertkorn et al., 2013), who reported olefinic to aromatic ratios in the range of 1.2 to 3.0. It is likely that this significant difference is due to the contributions of higher plant, lignin-rich carbon in the wetlands compared to marine.
This observation seems to confirm that indeed the PAN and OKA samples are enriched in higher plant derived DOC compared to the FCE samples. The slightly elevated olefin content found in FCE-L is may be the result from the contribution of periphyton-derived DOM in the Everglades (Maie et al., 2005; Chen et al., 2013), which is likely to enhance the olefin content at the expense of aromatic materials (see also above). In addition, all three sites are known for frequent and seasonal fires and have been reported to contain dissolved black carbon (DBC) (Ding et al., 2014a) in abundances close to 10% of their DOC on a global average (Jaffé et al., 2013). However, the DBC content (as %DOC) in the FCE was higher (as high as 20% of DOC) than for the PAN samples (13% and 14% for PAN-L and PAN-S respectively), and these higher than the OKA samples (9.4% and 6.3% for OKA-L and OKA-S respectively) studied here (Jaffé et al., 2013). In addition, the presence of six-membered N-containing heterocycles in these samples might be indicative of the presence of dissolved black nitrogen (DBN), which has previously been reported in the FCE (Maie et al., 2006) and proposed to consist of polyaromatic molecules containing pyrrolic-N, and multiple carboxylic substituents (Wagner et al., 2015); (see also attendant FTMS-based discussion of dissolved black sulfur compounds (DBS) in FCE-S; section 3e). While it is thus not surprising that the wetland samples feature high aromaticity in their structures compared to the oceanic samples, the high olefinic/aromatic ratio of the FCE is surprising unless periphyton is a significant source of olefin. With regards to the degree of oxidation of the aromatic signal, the OKA showed the highest proportion of electron donating groups (Fig. 2; Fig. S1 and S2), such as phenols and ethers, possibly related to lignin oxidation products, while the FCE featured the highest share of electron withdrawing substituents (e.g. carboxyl groups) possibly associated with DBC. Although all three ecosystems have climates leading to high light exposure, the high levels of DOC in the FCE suggest some degree of self-shading, while DOC in the OKA is generally lower and the system is known for its capacity to photo-degrade DOM (Cawley et al., 2012). Thus, the degree of photo-exposure of the DOM combined with combustion by-products such as DBC, may be the driver controlling the oxidation state of the aromatic fraction. The photo-reactivity of DBC in marine environments has recently been shown (Stubbins et al., 2012) and may play a role in the lower DBC levels observed in the OKA samples.
Comparison of wetland SPE-DOM FCE-L with a South Atlantic open ocean SPE-DOM MAX.

Wetland SPE-DOM FCE-L, which showed the most conspicuous signature of \( \text{C}_{\text{sp}2} \)-H HSQC cross peaks (Fig. 5) of all wetland DOM, was compared with Atlantic open ocean SPE-DOM MAX obtained at 48 m depth where the fluorescence maximum was located, supposedly corresponding to maximum biological activity at this open ocean location (Hertkorn et al., 2013). The overall envelope of \( \text{C}_{\text{sp}2} \)-H aliphatic HSQC cross peaks (\( \delta_{\text{H}} \sim 1.5 \) ppm) of South Atlantic FMAX SPE-DOM (at 48 m depth) reached out to larger \( ^{13} \text{C} \) NMR chemical shift values by \( \sim 7 \) ppm on average (Fig. 13b in Hertkorn et al., 2013), suggesting increased extent of (remote) aliphatic branching in marine SPE-DOM compared with wetland SPE-DOM. In contrast, HSQC cross peaks of methyl groups terminating extended aliphatic systems (\( \text{H}_3 \text{C}_{\text{sp}2} \text{Z}; n \geq 2; \text{Z: any heteroatom} \); section a; Fig. 6) occupied a larger area in FCE-L than in marine SPE-DOM FMAX, and extended down to \( \delta_{\text{C}} \sim 32 \) ppm at \( \delta_{\text{H}} \sim 0.9 - 1.2 \) ppm. The abundance and chemical diversity of methyl bound to alkenes (\( \text{C} = \text{CH}_2 \) groups and, possibly, \( \text{CH}_2 \) groups; section b; Fig. 6) was also larger in wetland SPE-DOM FCE-L than in marine SPE-DOM FMAX. Similarly, the methylene HSQC cross peak associated with \( \text{C} = \text{CH}_3 \text{COOH} \) groups in SPE-DOM FCE-L (section e; Fig. 6) extended downfield to \( \delta_{\text{H}} > 3.5 \) ppm, approx. 0.4 ppm further downfield than that of SPE-DOM FMAX (Fig. 8b; Hertkorn et al., 2013). The chemical diversity of methoxy groups (\( \text{OCH}_3 \)) in wetland SPE-DOM FCE-L substantially exceeded that of those identified in marine DOM. In particular, aromatic methyl esters occupied \( ^{1} \text{H} \) NMR chemical shifts down to \( \delta_{\text{H}} \sim 4.1 \) ppm (section g; Fig. 6) as opposed to \( \delta_{\text{H}} \sim 3.85 \) ppm found in marine SPE-DOM FMAX (Fig. 8c; Hertkorn et al., 2013). In addition, aromatic methyl ethers were clearly present in wetland SPE-DOM FCE-L (section g; Fig. 6) but virtually absent in marine SPE-DOM FMAX (Fig. 8c; Hertkorn et al., 2013).

Considerable distinction was observed for HSQC cross peaks which derived from unsaturated protons bound to \( \text{sp}^2 \)-hybridized (methine) carbon for FCE-L and FMAX as presented with an overlay of NMR spectra (Fig. S5). Most remarkably, FCE-L showed a much larger abundance and chemical diversity of substituted single aromatic rings across all substitution patterns compared with FMAX, with corresponding methine HSQC cross peaks covering a larger area in general and specific sections in particular. This suggested presence of aromatics with multiple substituted electron-donating functional groups (HSQC cross peaks with \( \delta_{\text{C}} < 120 \) ppm; cf. below) which were rare or absent in marine SPE-DOM FMAX. Likely
candidates were various plant-derived polyphenols and lignin derivatives such as polyoxygenated guaiacyl and syringyl units (HSQC cross peaks d, e, f in Fig. 5; Kim and Ralph, 2010, Martinez et al., 2008, Wen et al., 2013).

Certain substitution patterns could be proposed from reverse increment analysis (Perdue et al., 2007), suggesting presence of up to three oxygen atoms bound to aromatic rings ($\delta_C < 110$ ppm) as well as similarly high extents of polycarboxylation ($\delta_H > 8$ ppm). At lower field, HSQC cross peaks of FCE extended further than those of FMAX for both $^1$H and $^{13}$C NMR chemical shifts, suggesting more extensive polycarboxylation in SPE-DOM FCE as well as presence of (more extended) PAH derivatives. Otherwise, abundance and chemical diversity of single and conjugated olefins was larger in SPE-DOM FMAX than FCE-L, suggesting elevated contributions of marine natural products to this region. Overall, HSQC cross peak maxima and specific areas were better resolved in wetland than in marine SPE-DOM. This apparent enhanced molecular diversity of wetland SPE-DOM in comparison with marine SPE-DOM is expected to be genuine for lignin-derived chemical environments and might in part reflect maximum molecular diversity in marine SPE-DOM, resulting in more smooth cross peak distributions in both $^1$H and $^{13}$C NMR frequencies as observed in marine SPE-DOM (see discussion in Hertkorn et al., 2013).

FTICR mass spectrometry/MS:

Ultra-high resolution Fourier transform ion cyclotron mass spectra (FTICR/MS) of SPE-DOM may provide several thousands of mass peaks for individual samples (Koch et al., 2005; Kujawinski et al., 2009), of which many hundreds were assigned here to extended CHO, CHNO, CHOS and CHNOS molecular series (Schmitt-Kopplin et al., 2010) based on the technique’s excellent mass accuracy and mass resolution (Fig. 7–6 and Table 3). Although detailed FTICR/MS data are derived only from a few paired SPE-DOM samples (Long and Short hydroperiod) for each wetland, a slightly higher number of mass peaks (relative difference < 6%) and of assigned molecular formulas (relative difference < 1%) was observed for the FCE-L compared to the FCE-S, whereas elevated counts of mass peaks and assigned molecular compositions were found in case of the PAN-S and OKA-S samples (relative difference < 2%; Table 3). Molecular weights ranged in the order FCE-S > FCE-L ~ PAN > OKA (Table 3). This admittedly minor molecular weight difference was not reflected in the $S_R$ values of these samples.
(Table 1) which were quite similar. However, $S_R$ only represents a molecular weight proxy for CDOM and might not be sensitive enough to reflect minor differences accurately. In general, while SPE-DOM of both OKA and PAN showed near $57 \pm 2\%$ CHO, $8 \pm 2\%$ CHOS, $33 \pm 2\%$ CHNO, and $<1\%$ CHNOS molecules, the mass spectra of FCE samples were fundamentally different compared with respect to both OKA and PAN as well as among themselves (Fig. 7; Table 3; see also Fig. 8). Sample FCE-S appeared most distinct from all other samples both with respect to total count of ions, overall mass peak distribution and with respect to molecular diversity within nominal mass ranges (Fig. 7). Here, FTICR mass spectra of both FCE samples showed the conspicuous doublets of CHO/CHOS pairs visible at high resolution ($\Delta m (C_3H_4S) = 2.4$ mDa) indicating a nominal exchange of $H_4S$ against $C_3$ (Schmitt-Kopplin et al., 2010), whereas all other samples showed both lower abundances and diversity of CHOS compounds (Fig. 7–6 and Fig. 8). In case of the FCE samples, CHOS and CHNOS compounds were markedly enriched at the expense of CHO and CHNO compounds. While the proportion of CHNO ($21 \pm 1\%$) and CHNOS ($9 \pm 1\%$) molecules were similar for both FCE samples, the abundance of CHOS molecules in FCE-S was elevated by more than 10%, predominantly at the expense of CHO molecules. The overall abundance of sulphur in the FCE was nearly four-fold when compared with that of the OKA and PAN samples (Table 3).

CHOS compounds observed in all wetland samples already showed a remarkable chemical diversity (Fig. 9B). However, the chemical dissimilarity of CHOS compounds common to both FCE samples remarkably exceeded that found in OKA and PAN, covering a substantial share of the CHOS chemical space from O/C ratio: 0.3 - 0.8 and H/C ratio: 0.6–1.7, respectively (Fig. 9C). Here, four groups of CHOS molecules were differentiated based on their positioning in H/C against O/C van Krevelen diagrams (Fig. 8F): (a) saturated sulfolipids with H/C ratio > 2 and intermediate O/C ratio, suggesting the presence of sulphur in elevated oxidation states; (b) unsaturated sulfolipids with a rather restricted H/C and O/C ratio; (c) a very large and expansive set of molecularly diverse CHOS molecules with a bandwidth of O/C ratios similar to CHO compounds but reaching out to higher saturation (larger H/C ratio) than the latter (Fig. 8E, Fig. 8F); (d) unique to FCE-S (with traces in OKA-L) was a large set of aromatic and oxygen-deficient “black sulphur” compounds (DBS; section d; Fig. 8F), similarly positioned like CHOS compounds in Atlantic open Ocean abyssopelagic SPE-DOM at 5446 m depth (Fig. S8, in Hertkorn et al., 2013), but covering a larger mass range (Fig. S7F).
Section (b) and (c) CHOS compounds were also observed in PAN and OKA, whereas black sulphur compounds were rare in OKA-L and virtually absent in the other samples except FCE-S. DOM-type CHOS compounds common to all six wetland samples were on average more saturated and oxygenated than their respective CHO and CHNO counterparts, suggesting also here presence of sulphur in elevated oxidation states (Fig. 9B).

The CHOS compounds of both FCE samples not only differed fundamentally from those found in OKA and PAN, but were also remarkably diverse in both FCE-L and FCE-S samples itself. Figure 10 shows the respective mass peaks of both FCE samples which displayed a relative higher intensity, indicating more abundant CHOS compounds present with elevated abundance in either FCE-S (Fig. 10A) or FCE-L (Fig. 10B). The most peculiar feature of FCE-S was a hydrogen-deficient pool of (poly)aromatic CHOS compounds (section a mass peaks; Fig. 10A) in extended molecular series with limited degree of oxidation (O/C ratio < 0.22), ranging from m/z ~ 300 - 600. The positioning in both van Krevelen and mass-edited H/C ratio diagrams (Fig. 10) was in accordance with that of ‘black sulphur’ in abyssopelagic South Atlantic SPE-DOM (Fig. S8; Fig. 16 in Hertkorn et al., 2013), but its signature was more conspicuous and showed larger richness of diverse CHOS compounds in FCE-S. While sulphur can be readily inserted into any C-C and C-H bond, analogous to oxygen, organic sulphur can also occupy oxidation states ranging from -2 to +6, an option not available to oxygen. Nevertheless, the manifest oxygen-deficiency of the proposed highly unsaturated CHOS molecules (section d; Fig. 8F) suggests the presence of reduced sulphur in the form of sulphides. Aromatic CHOS molecules will then most likely occur as benzothiophene derivatives, a chemical environment of sulphur largely favoured in mineral oils (Purcell et al., 2007; Liu et al., 2010; Muller et al., 2012). While both black carbon as well as black nitrogen (Wagner et al., 2015) have been reported in the FCE (Ding et al., 2014 a and b; Maire et al., 2006) the presence of this ‘black sulphur’ was not previously observed at FCE. The environmental factors driving the high abundance of these compounds at FCE-S remains unclear but may be related to the higher fire frequency at short hydroperiod sites and possibly soil charring. A small set of CHOS compounds with more average H/C and O/C ratios (section b mass peaks; Fig. 10A) was accompanied by a rather minor set of highly oxygenated CHNOS compounds, with an O/C ratio > 0.75 (Fig. 10A).
In contrast to the FCE-S, the FCE-L sample displayed an oxygenated set (O/C ratio > 0.6) of a few dozen hydrogen-deficient (H/C ratio < 1.1) CHOS molecules in truncated molecular series and at rather low mass (m/z < 400; Fig. 10B, section d mass peaks). These molecules were most likely composed of oxygenated aromatics connected by (some) ether bridges, which rather likely originate ultimately from plant and/or algal polyphenols. Apart from PAH-derived compounds, which are commonly rather oxygen-deficient (O/C ratio < 0.3), these structures represent one of the most plausible motifs of very hydrogen-deficient molecules found in DOM (H/C ratio < 1). The large extent of average oxygenation makes sulphur functional groups in elevated oxidation states, e.g. sulfones, sulfonates or sulfates, likely candidates for this group of CHOS compounds.

In addition, a rather expansive cloud of abundant CHOS and less common CHNOS compounds at mass range m/z: 200 – 550, with large and variable extent of oxygenation (O/C ratio: 0.4 – 0.95) was prominent in FCE-L and near absent in FCE-S. The sizable expansion of this cloud with a huge range of H/C ratios testified for a rather large overall diversity of these unique CHOS molecules found solely in FCE-L; CHNOS compounds seemed to follow suit but with a lesser overall diversity: highly oxygenated (O/C ratio > 0.8) and hydrogen-rich CHNOS molecules (H/C > 1.6) were missing even if every added nitrogen carried one intrinsic hydrogen into analogous CHO molecular formulas. This higher molecular diversity for the FCE-L site may be driven by higher soil-derived (peat soils) DOM contributions at this site compared to FCE-S (marl soils) (Chen et al., 2013) and a higher degree of DOM preservation at this deeper, less photo-exposed site.

It has to be mentioned that this cloud encircled the common molecular series of several hundreds of CHOS and CHNOS compounds found in both FCE samples (Fig. 8E-7F, 8F-7F, 9C-8C). The significantly higher presence of sulphur-containing molecular formulas for the FCE samples is likely the result of higher inputs of sulphate to the Everglades compared to the Pantanal and Okavango. Firstly, Everglades is a coastal wetland where sea-spray may be an important contributor to sulphate. In addition, it is the most anthropogenically impacted wetland of the three being compared, where runoff from agricultural lands within the Everglades watershed is likely the most important contributing factor to the sulphur load of the system as it is an ingredient of fertilizer applications. The CHO and CHNO components specific for the OKA and PAN samples are shown in Figure 9E-8F and suggest, in agreement with the NMR data, a
higher degree of oxidized, H-deficient materials at these sites compared to the FCE. This is particularly true for the PAN which show unique molecular formulas for oxidized, H-deficient CHO and CHNO components (Fig. 9D), whereas molecular formulas unique for the OKA are relatively few (Fig. 9E).

Comparative analyses of van Krevelen diagrams between the six sites as shown in Figure 9A–8A clearly cluster the FCE samples separately from the OKA and PAN. Cluster analysis showed a clear distinction between the FCE on one hand and the OKA and PAN samples on the other hand, with less pronounced but significant differences between the paired, long and short hydroperiod samples at each site (Fig. 9A). Among pairs of DOM samples, similarity according to FTICR/MS-based cluster analysis was in the order PAN > OKA > FCE (Fig. 9A), whereas in line with the visual overlay of one-dimensional $^1$H NMR spectra clustered according to increasing dissimilarity in the order OKA < PAN < FCE (Fig. S24). This discrepancy is readily explained by the different recognition of aliphatic groups in FTMS (insensible) and NMR spectra (quantitative depiction). CHO and CHNO molecules ionized by negative ESI occupied rather similar expansive regions with near average H/C and O/C elemental ratios (Fig. 9B). This is a common feature of DOM molecular distribution as derived from FTICR mass spectra. Here, the largest number of feasible and chemically reasonable isomeric molecules will project on single mass peaks at average H/C and O/C elemental ratios, contributing to larger overall mass peak amplitude – this applying even more specifically to van Krevelen diagrams, in which different molecular compositions with identical elemental ratios contribute to the same data points (Hertkorn et al., 2007; Lechtenfeld et al., 2014). Analogously, the distribution of CHO, CHOS, CHNO and CHNOS molecular series roughly coincided, with some displacement of CHOS molecules in both FCE samples, towards higher H/C ratio (i.e. higher aliphatic character).

At first glance, the H/C vs. O/C (Fig. 87) as well as the H/C vs. m/z (Fig. S7) plots, showed near uniform fingerprints for OKA and PAN, covering larger areas in the van Krevelen diagrams in case of CHO compared with CHNO compounds (Fig. 9B), suggesting an increased overall chemical diversity of CHO compounds. In addition, the paired wetland samples clustered separately for the high and low hydroperiods respectively, suggesting that molecular compositions differ among sites with different hydrology. The weighted average O/C and H/C values were remarkably similar for PAN and OKA showing rather marginal variance between
different sites or between high and low hydroperiod (Fig. 8A; Table 3). In comparison, FCE-S showed a considerably decreased O/C ratio. While computed O/C ratios of wetland DOM exceeded those found in oceanic DOM by about 0.2 units (Table 3; Table 4 in Hertkorn et al., 2013), the H/C ratio of wetland DOM was approximately 0.15 units higher in comparison. Even if ionization selectivity in negative ESI FTICR mass spectra applied, the ¹H NMR section integrals indicate analogous trends of relative saturation, or alternatively, hydrogen deficiency between wetland and marine DOM. In comparison with average wetland SPE-DOM, average open ocean SPE-DOM showed lesser abundance of aromatics (by 2-3%), lower proportions of OCH chemical environments (by 8%), and, especially, higher abundance of pure aliphatics (i.e. CCCHH units; by 12%). This implies that marine DOM shows lower abundance of hydrogen-deficient (unsaturated) and higher abundance of hydrogen-rich (purely aliphatic)saturated molecules than wetland DOM, in line with the elevated H/C ratio as derived from FTICR mass spectra. Similarly, the higher abundance of oxygen-rich OCH chemical environments in wetland DOM as seen by ¹H NMR section integral (8% relative increase) was in accordance with the increased O/C ratio found in their FTICR mass spectra (increase by 0.2 units).

Comparative analysis of van Krevelen diagrams (Fig. S8) obtained solely from the four PAN and OKA samples confirmed the previously observed higher similarity between PAN-L and PAN-S compared to OKA-L and OKA-S sample pairs (Fig. 87; Fig. 9D8D, 9E–8E and 9F8F). Molecular compositions with unique high abundance when derived from all six wetland samples were sparse and non-significant in case of OKA (Fig. 8E; see also Fig. S8B), whereas molecular compositions with unique high abundance in all four OKA and PAN samples occupied a rather dense, contiguous section of hydrogen-deficient (H/C ratio < 1) and oxygenated (O/C ratio ~ 0.3 – 0.7) CHO and CHNO molecules (Fig. S8F). In agreement with its high degree of photo-oxidation, the OKA contained higher proportions of highly oxygenated CHO (O/C ratio > 0.7) and CHNO (O/C ratio > 0.5) molecules, and a few rather abundant (and easily ionizable) sulfolipids (Fig. S8B), whereas PAN SPE-DOM displayed larger proportions of hydrogen-deficient CHO molecules of considerable chemical diversity and extent of oxygenation (O/C ratio ~ 0.2–0.9), and several dozens of CHNO molecules similarly positioned but with more limited range of oxygenation and hence, overall chemical diversity (Fig. 9D8D). Molecular compositions with unique high abundance when derived from all six wetland samples are sparse.
and non-significant in case of OKA (Fig. 9E; see also Fig. S8B), whereas molecular compositions with unique high abundance in both PAN samples occupied a rather dense, contiguous section of hydrogen-deficient (H/C ratio < 1) and oxygenated (O/C ratio ~ 0.3—0.7) CHO and CHNO molecules (Fig. S8C).

Remarkably, with the exception of a tiny section of CHO molecules (H/C ~ 1.1; O/C ~ 0.4), both CHO and CHNO molecular series for OKA and PAN nearly perfectly superimpose in the H/C against O/C van Krevelen diagram (Fig. 9F). It is very likely that these CHO molecules jointly present in PAN + OKA mainly represent oxygenated aromatic molecules, possibly connected by ether linkages. This is one of the most comprehensive ways to envision such hydrogen-deficient molecules of conceivable natural product origin, and in agreement with the NMR data suggesting a higher degree of oxidized, H-deficient materials at these sites compared to the FCE. This is particularly true for the PAN which shows unique molecular formulas for oxidized, H-deficient CHO and CHNO components, whereas molecular formulas unique for the OKA are relatively few (Fig. 9E, 9F; Fig. S8B and S8C).

Conclusions and biogeochemical significance: Very unprecedentedly detailed molecular analyses of DOM samples from three different sub-tropical freshwater wetlands suggest, that in agreement with previous reports on riverine and marine DOM characterizations, that many of the bulk molecular characteristics in freshwater DOM are shared by ecosystems despite being very different in their environmental settings (Repeta et al., 2002; Jaffé et al., 2012). Nevertheless, organic structural spectroscopy provides evidence for wetland-specific molecular assemblies. NMR and FTMS analysis provided exceptional coverage of wetland SPE-DOM composition and structure, confirming individual wetland organic matter molecular characteristics. Those were directly revealed in 1H NMR spectra, while extensive mathematical analysis was mandatory to discern analogous distinction in FTICR mass spectra, which show extensive projection of structural variance on the primary measured variable. These detailed analyses revealed significant variations in the molecular composition that can, in some cases, be controlled by site-specific environmental conditions. Among those are hydrological drivers such as hydroperiod (lengths and depths of inundation), resulting in variations in light penetration and associated photochemical processes along with seasonal drying of surface soils and associated aerobic oxidation processes. Other drivers include (i) external sources of sulphur, such as...
agricultural activities and sea spray, resulting in the formation of a variety of sulphur compounds in DOM, (ii) fire regime, possibly causing soil OM charring during wildfires, and (iii) natural DOM source variations and source strength in the contribution from vascular plants, grasses, and aquatic vegetation including microbial contributions from periphyton. As such, while not all molecular differences could be explained through one or more of these drivers, this study illustrates for the first time the extensive molecular diversity and compositional complexity of DOM in wetlands, and as such should serve as a database for future characterization efforts. Further detailed molecular-level characterizations of wetland DOM are encouraged as a means to better understand spatial and seasonal variability in sources, transformations and reactivity, which can be ultimately used to aid in constraining carbon cycling models.

Author contributions: N.H. performed NMR study, contributed to the study design, data interpretation and actively participated in the writing of the manuscript; M.H. performed the FT-ICR/MS analyses and data manipulation and participated in data interpretation; K.C. collected samples from Okavango Delta and Everglades and performed optical properties study; P.S.-K. Provided support for the FT-ICR/MS analyses and general data interpretations; R.J. collected Pantanal and Everglades samples, participated in all data interpretations, took the lead in coordinating this study and writing this manuscript, and generated funding in support of this research. All authors provided significant input on the final manuscript.

Acknowledgments: The authors acknowledge the financial support from the George Barley Chair and the Florida Coastal Everglades LTER (DEB – 1237517) for funding this research. We are grateful to C. Nunes, N. Rodriguez, N. Mladenov and P. Wolski who assisted in field logistics and to S. Thaller (HMGU) who expertly assisted in NMR sample preparation and assembly of figures. K.C. thanks Everglades National Park for a Post-doctoral fellowship. This is SERC contribution number ###.
References:


Figure Captions:

Fig. 1. Box plots of the relative abundance (%) of PARAFAC components for the FCE, OKA, and PAN regions. Top of the blue box represents the 75th percentile, middle of the blue box is the median, lower edge of blue box is the 25th percentile, top of the black bar is the maximum value and bottom of the black bar is the minimum value. ENP = Everglades National Park; WCA2 = Water Conservation Area 2.

Fig. 2. $^1$H NMR spectra of six wetland SPE-DOM (500 MHz; CD$_3$OD); overlay: intensities are normalized to total NMR resonance area in the entire chemical shift range shown ($\delta_H = 0$-10 ppm), with exclusion of residual water and methanol NMR resonances (Fig. S1). (A) entire NMR spectrum ($\delta_H = 0$-10 ppm), with section of unsaturated protons ($\delta_H = 5$ - 10 ppm), highlighted in orange; (B) section of unsaturated protons ($\delta_H = 5$ - 10 ppm); (C) section of aliphatic protons ($\delta_H = 0$ - 5 ppm); highlighted in purple colour: vertical expansion of functionalized aliphatic compounds, associated also with CRAM (carboxyl-rich alicyclic compounds).

Fig. 3. $^1$H NMR spectra (800 MHz, CD$_3$OD) of three organic matter preparations from the Florida Coastal Everglades, isolated by means of (panel A) ultrafiltration (FCE-S UDOM), panel B: fulvic acid obtained by the IHSS extraction method (FCE-S FA), and (panel C) solid-phase extracted SPE-DOM by means of PPL cartridges (FCE-S PPL); blue: NMR spectra acquired in D$_2$O; gray: NMR spectra acquired in CD$_3$OD; overlay: $^1$H NMR spectrum depicted in panel A with gray color shows increased amplitude to allow better comparison of non lipid derived NMR resonances (cf. main text).

Fig. 4. TOCSY NMR spectra (800 MHz, CD$_3$OD) of wetland SPE-DOM, with (A) FCE-S, (B) SPE-DOM, and (C) PAN-L and (D) OKA-S SPE-DOM. Panel (A) depicts TOCSY cross peaks between aliphatic protons ($X$-$C_{sp3}H$-$C_{sp3}H$-$X$; $X$: C, O), whereas panels (B-D) depict TOCSY cross peaks between unsaturated protons ($X$-$C_{sp2}H$-$C_{sp2}H$-$X$; $X$: C, O). Section a: $H_3$C-$C_{ar}$-$X$ cross peaks, with $n = 1$ ($\delta_H > 3$) and $n > 1$ ($\delta_H < 3$); where $X$ is any heteroatom, likely oxygen; section b: -$C\cdot CH\cdot CH\cdot C_{ar}\cdot X$-, intra-aliphatic cross peaks; section c: $\alpha,\beta$-unsaturated and
conjugated double bonds: $\text{HC}_{\text{olefin}}=\text{C}_{\text{olefin}}(\text{C}=\text{O})-\text{X}$; section d: polarized $\alpha,\beta$-unsaturated double bonds: $\text{HC}_{\text{olefin}}=\text{C}_{\text{olefin}}(\text{C}=\text{O})-\text{X}$; section e: congested fjord region in polycyclic aromatics; section f: aromatics $\text{HC}_{\text{aromatic}}-\text{C}_{\text{aromatic}}$ with ortho or/and para oxygenated substituents (classic aromatic substitution of DOM); section g: condensed and strongly electron withdrawing aromatics $\text{HC}_{\text{aromatic}}-\text{C}_{\text{aromatic}}$ (multiply carboxylated, N-heterocycles); section h: (more extended) polycyclic aromatics, polycarboxylated aromatics, N-heterocycles. Panel D: Sections of chemical shift for substituted aromatics as proposed by SPARIA model (substitution patterns in aromatic rings by increment analysis): COR: electron withdrawing substituents; R: electroneutral substituents; OR: electron-donating substituents (Perdue et al., 2007).

Fig. 54. $^1$H, $^{13}$C HSQC NMR cross peaks of FCE-S; section of unsaturated (olefinic and aromatic) protons $\delta_H = 4$–10.5 ppm. Assignment in analogy to South Atlantic SPE-DOM FMAX (Hertkorn et al., 2013) with key substructures denoted as follows: section a: anomeric CH in carbohydrates (sp$^3$-hybridized); section b: isolated olefins; section c: C-conjugated olefins, certain five membered N-, O- and S-heterocycles ($\delta_H < 6.5$ ppm); section d: multiply oxygenated aromatics including oxygen heterocycles, lignin derivatives, syringyl units (S2/6); section e: phenols, classical oxygenated DOM aromatics, lignin derivatives, guaiacyl units (G2), certain admixture of carbonyl derivatives (likely carboxylic units), causing downfield $^1$H NMR chemical shift ($\delta_H > 7.3$ ppm); section f: classical DOM aromatic, lignin derivatives, guaiacyl units (G5/6), para-coumarate (C3/5); section g: classical DOM aromatics with high proportion of carboxylated units; at $\delta_H > 8$ ppm: multiply carboxylated aromatics, classical PAH and certain six-membered nitrogen heterocycles; sterically uncongested PAH; section h: $\alpha,\beta$-unsaturated double bonds for $\delta_C > 140$ ppm, including double bonds adjacent to aromatics: C-$\text{HC}_{\text{olefin}}=\text{C}_{\text{olefin}}(\text{C}=\text{O}), \text{C}_\text{ar}-\text{X}$; section i: nitrogen heterocycles, heteroatom substituted polycyclic aromatics; section j: certain six-membered nitrogen heterocycles, very likely with more than one nitrogen. The green area highlights the HSQC cross peak region accessible to single benzene rings substituted by common electron withdrawing, neutral and electron-donating common substituents of natural organic matter; SPARIA: substitution patterns in aromatic rings by increment analysis (Perdue et al., 2007).
**Fig. 65.** Methylene (CH$_2$) selective $^1$H, $^{13}$C DEPT – HSQC NMR spectrum of SPE-DOM FCE-S with assignment of major substructures; general colours: CH$_3$: red; CH$_2$ green, and CH: gray; section a: C-CH$_3$ cross peaks; section b: C=CH$_3$ and -SCH$_3$ cross peaks; section c: acetate H$_3$C-C(=O)-O-C; section d: C$_2$CH$_2$ cross peaks; section e: -C-CH$_2$-COOH cross peaks; section f: C$_3$CH cross peaks; section g: only methoxy (OCH$_3$) cross peaks are shown here; see insert: section g$_1$: H$_3$COH (HD$_2$COD shows methine carbon); sections g$_2$ and g$_3$: aliphatic (g$_2$) and aromatic (g$_3$) methyl esters H$_3$CO-C(=O)-C; section g$_4$ and g$_5$: aromatic (g$_4$) and aliphatic (g$_5$) methyl ethers H$_3$CO-C-C; section f: C$_3$CH cross peaks; section h: oxomethylene (OCH$_2$) cross peaks, likely from carbohydrates; section i: OC$_2$CH cross peaks; section j: methylene bound to esters –C-H$_3$CO-C(=O)-Z- (cf. main text).

**Fig 76.** Left panel: negative electrospray 12T FTICR mass spectra of Wetlands SPE-DOM (insert in figure 7F show an enlarged mass view of a mass range of 6.0 Da. Right panel: expansion of the mass segment m/z = 465.00-465.20 (Δm = 0.2 Da; asterisk in insert figure 7F), with assignment according to CHO, CHNO, CHOS and CHNOS molecular series. (A) OKA-L; (B) OKA-S; (C) PAN-L; (D) PAN-S; (E) FCE-L; (F) FCE-S.

**Fig. 87.** Van Krevelen diagrams of six wetlands SPE-DOM; (A) OKA-L; (B) OKA-S; (C) PAN-L; (D) PAN-S; (E) FCE-L; (F) FCE-S, obtained from negative electrospray 12T FTICR mass spectra. Only molecular assignments bearing combinations of C,-H,-O,-N, and –S atoms are shown; color coded according to molecular series as follows: CHO-blue, CHOS-green, CHNO-orange, CHNOS-red. Bubble areas reflect the relative intensities of respective mass peaks. Panel F: labels for CHOS compounds correspond to key molecules, section a: saturated sulfolipids; section b: unsaturated sulfolipids; section c: common CHOS compounds in DOM, possibly sulfonated carboxylic-rich alicyclic compounds (CRAM); d: aromatic black sulphur.

**Fig. 98.** Comparative analysis of van Krevelen diagrams derived from negative electrospray 12T FT-ICR mass spectra of all six wetlands SPE-DOM. (A) Clustering diagram based on the similarity values between the spectra of six wetlands SPE-DOM using Pearson correlation coefficient; (B) molecular compositions common to all six wetlands SPE-DOM, (C) unique molecular compositions common in FCE samples (FCE-L and FCE-S); (D) unique molecular compositions with high abundance in both PAN samples; (E) unique molecular compositions
with high abundance in both OKA samples; (F) unique molecular compositions common in all four PAN and OKA. The aromaticity index AI (Koch and Dittmar, 2006) provided denotes single aromatic compounds for AI > 0.5 (bright blue triangle).

Fig. 409. Comparative analysis of (left) H/C vs. m/z and (right) H/C vs. O/C van Krevelen diagrams derived from negative electrospray 12T FTICR mass spectra of the two Florida Coastal Everglades SPE-DOM FCE-S and FCE-L (see also Fig. 8). (A) Molecular compositions with high abundance in Florida Coastal Everglades SPE-DOM FCE-S; section a: oxygen-deficient (poly)aromatic black sulphur; CHNOS: suite of highly oxygenated CHNOS molecules; section b: common CHOS molecules in DOM; section c: saturated sulfolipids. The aromaticity index AI (Koch and Dittmar, 2006) provided in the upper right van Krevelen diagram denotes single aromatic compounds for AI > 0.5 (bright blue triangle) and polyaromatic compounds for AI > 0.67 (bright purple triangle); (B) Molecular compositions with high abundance in Florida Coastal Everglades SPE-DOM FCE-L; section d: a distinct set of oxygen-rich aromatic CHOS compounds, likely associated with ether-linked aromatic units; cf. text.
Supporting Online Information

**Fig. S1.** $^1$H NMR spectra of six wetland SPE-DOM (CD$_3$OD; 500 MHz), acquired with solvent suppression and exclusion regions used in the computation of NMR section integrals and overlay NMR spectra (Fig. 2 and this figure) which denote HD$_2$COD and residual HDO, with section of unsaturated protons ($\delta_H > 5$ ppm) vertically expanded. Intensities are normalized to 100% total integral in the entire chemical shift range shown ($\delta_H = 0...10$ ppm). Fundamental substructures are indicated from higher to lower field (from right to left), (a) aliphatics, $\text{HCCC}$; (b) “acetate-analogue”, $\text{H}_2\text{CC} (=\text{O})\text{-O}$-; (c) carboxyl-rich alicyclic materials (CRAM), $\text{H}(\text{C})\text{-COX}$; (d) “carbohydrate-like” and methoxy, $\text{HCO}$; (e) olefinic, $\text{HCC}=$; and (f) aromatic NMR resonances $\text{HC}_{ar}$ (cf. text). Further division of unsaturated protons provided (f$_1$) polycyclic and polycarboxylated aromatics as well as six-membered nitrogen heterocycles ($\delta_H > 8$ ppm); (f$_2$) electron withdrawing substituents (COX; Perdue et al., 2007; $\delta_H \approx 7.3 – 8.0$ ppm); (f$_3$) electronneutral substituents (alkyl, H, R; $\delta_H \approx 7.0 – 7.3$ ppm); (f$_4$) electron-donating substituents (OR, OH, phenolics; $\delta_H \approx 6.5 – 7.0$ ppm); (e$_1$) polarized and conjugated olefins; (e$_2$) isolated olefins ($\delta_H \approx 5.0 – 5.5$ ppm), with conceivable contributions from anomeric protons and ester groups (cf. discussion of 2D NMR spectra).

**Fig. S2.** $^1$H NMR spectra of wetland SPE-DOM (CD$_3$OD; 500 MHz). Similarity assessment by means of (panel A) cluster analysis (Pearson) and (panel B) PCA as well as (panels C, D) computed difference $^1$H NMR spectra of 3 wetland SPE-DOM pairs (L-S: long minus short hydroperiod) as derived from 0.001 ppm buckets in area-normalized $^1$H NMR spectra; used chemical shift range: $\delta_H = 9.5 – 0.5$ ppm, with exclusion of residual water and methanol NMR resonances. Panels E, F, G: Manual overlay according to identical $^1$H NMR section integral in the respective regions of $^1$H NMR chemical shift shown: (left column) entire NMR spectrum ($\delta_H = 0 - 10$ ppm); (center column) section of unsaturated protons ($\delta_H = 5 - 10$ ppm); (right column) section of aliphatic protons ($\delta_H = 0 - 5$ ppm). Panel E: OKA; panel F: PAN, and panel G: FCE SPE-DOM. Sections f$_n$ of unsaturated protons are denoted as provided in Fig. S1.

**Fig. S2.** $^1$H NMR spectra of wetland SPE-DOM (CD$_3$OD; 800 MHz); visual overlay: intensities are normalized to total area in the entire chemical shift range shown ($\delta_H = 0-10$ ppm), with
exclusion of residual water and methanol NMR resonances; (left column) entire NMR spectrum ($\delta_H = 0$–10 ppm); (center column) section of unsaturated protons ($\delta_H = 5$–10 ppm); (right column) section of aliphatic protons ($\delta_H = 0$–5 ppm). Panel C: FCE SPE-DOM. Manual overlay according to identical $^1$H NMR section integral in the respective regions of $^1$H NMR chemical shift shown. Section of unsaturated protons are denoted as follows (cf. main text): (f1) polycyclic and polycarboxylated aromatics as well as six-membered nitrogen heterocycles ($\delta_H > 8$ ppm); (f2) electron withdrawing substituents (COX; Perdue et al., 2007; $\delta_H \approx 7.3$–8.0 ppm); (f3) electroneutral substituents (alkyl, H, R; $\delta_H = 7.0$–7.3 ppm); (f4) electron donating substituents (OR, OH, phenolics; $\delta_H \approx 6.5$–7.0 ppm); (e1) polarized and conjugated olefins; ($\delta_H \approx 5.5$–6.5 ppm); (e2) isolated olefins ($\delta_H \approx 5.0$–5.5 ppm). Panel D-F: $^1$H NMR spectra (800 MHz) of three organic matter preparations, acquired in $D_2O$ (blue) and $CD_3OD$ (black, gray): overlay with manually adjusted amplitude for respective sections of chemical shift shown, for improved visual assessment of relative similarity and differences. Panel D: ultrafiltered organic matter FCE–S UDOM (due to the very intense lipid NMR resonance at $\delta_H = 1.28$ ppm, vertical expansions are provided with two intensities to also allow for comparison of non-lipid NMR resonances); panel E: fulvic acid FCE–L FA (panel $\Delta$E shows difference NMR spectra D$_2$O minus CD$_3$OD); panel F: SPE–DOM obtained by solid phase extraction with PPL cartridges (FCE–S PPL; panel $\Delta$F shows difference NMR spectra D$_2$O minus CD$_3$OD).

Fig. S3. $^{13}$C NMR spectra of selected wetland SPE-DOM; full spectra computed with 35 Hz exponential line broadening; insert: section of methoxy peaks ($\delta_C = 51$–59 ppm; computed with 2 Hz line broadening); OKA-L and PAN-S: in $^{12}$CD$_3$OD at $B_0 = 11.7$ T; FCE in CD$_3$OD at $B_0 = 18.8$ T.

Fig. S4. $^1$H, $^{13}$C HSQC NMR spectrum of SPE-DOM FCE-S, with regions shown in figures: (A) chemical environments of $sp^3$-hybridized carbon (aliphatic CH$_n$ units; Fig. 6); (B) chemical environments of $sp^2$-hybridized carbon (unsaturated, i.e. olefinic and aromatic CH units; Fig. 5). Sensitivity enhanced apodization is used to emphasize less abundant $sp^2$-hybridized carbon
(overall HSQC cross peak integral <4% of aliphatic units) environments at the cost of resolution in case of aliphatic CH<sub>n</sub> units (n = 1 – 3).

**Fig. S5.** Overlay of <sup>1</sup>H, <sup>13</sup>C HSQC NMR spectra of SPE-DOM FCE-S (dark blue) and South Atlantic SPE-DOM at fluorescence maximum (48 mg, FMAX; orange: Hertkorn et al., 2013), together with region of HSQC NMR cross peaks accessible for single aromatic rings with full range of electron-withdrawing (COX), electroneutral (R, H) and electron donating substitution (OH, OR), shown in green color (SPARIA: Perdue et al., 2007). Wetland SPE-DOM shows more exhaustive coverage of single aromatic rings from contributions of multiply oxygenated aromatics (δ<sub>H</sub> < 7 ppm; δ<sub>C</sub> < 120 ppm), likely originating from plant phenolics but also from polycarboxylated aromatics and PAH derivatives (δ<sub>H</sub> > 8.5 ppm). In contrast, open ocean SPE-DOM FMAX exhibits a larger abundance as well as overall chemical diversity of α,β unsaturated and C-conjugated olefins, likely originating from marine natural products; for assignment of HSQC cross peaks, see Figs. 4 and 5, and Hertkorn et al., 2013.

**Fig. S6.** Further evaluation of aliphatic spin systems of wetland SPE-DOM FCE-L. Panel A: overall <sup>1</sup>H, <sup>1</sup>H JRES NMR spectrum with sections a<sub>1</sub>, a<sub>2</sub>, a<sub>3</sub>, denoting the area of panels B, C, D, which display <sup>1</sup>H NMR projections along JRES and <sup>1</sup>H, <sup>13</sup>C DEPT HSQC NMR spectra (copied from Fig. 6); panel B: section of OCH<sub>3</sub> aliphatic units, demonstrating (section b<sub>1</sub>) presence of intense JRES cross peaks from OCH<sub>3</sub> groups, with absence of J<sub>HH</sub> splittings; panel C: section of aliphatic CCH units, with dominance of HOOC-CH<sub>2</sub>-CH<sub>2</sub>- units (triplet J<sub>HH</sub> splitting; n = 1, 2) over HOOC-CH<sub>2</sub>-CH<sub>2</sub>- units (doublet J<sub>HH</sub> splitting; n = 1, 2) shown in section c<sub>1</sub>; section c<sub>2</sub> indicates panel D; panel D: section of aliphatic CCC<sub>2</sub>H units, showing a remarkable clustering of H<sub>3</sub>C-CH- units at δ<sub>H</sub>: 1.0 – 1.4 ppm, which indicate pronounced aliphatic branching in section d<sub>1</sub> (doublet splitting from J<sub>HH</sub>), whereas ethyl groups H<sub>3</sub>C-CH<sub>2</sub>- dominate the low field section δ<sub>H</sub> < 1 ppm (section d<sub>2</sub>).
blue (CHO), orange (CHNO), green (CHOS) and red (CHNOS). Bubble areas reflect the relative intensities of each mass peak.

Fig. S8. Comparative analysis of van Krevelen diagrams derived from negative electrospray 12T FTICR mass spectra derived from four Pantanal and Okavango SPE-OM only. (A) Clustering diagram based on the similarity values between the FTICR mass spectra of these four SPE-DOM; (B) Molecular compositions with rather high abundance in both Okavango SPE-DOM; (C) Molecular compositions with rather high abundance in both Pantanal SPE-DOM, with color code according to molecular series (cf. text). The bright blue triangle denotes aromatic compounds, with aromaticity index AI > 0.5 (Koch and Dittmar, 2006); see Fig. 9 and attendant discussion.
### Table 1: DOC and optical properties of the six bulk water samples collected for SPE-DOM

<table>
<thead>
<tr>
<th>Sample</th>
<th>DOC (ppm)</th>
<th>SUVA$_{254}$</th>
<th>Abs$_{254}$</th>
<th>S$_R$</th>
<th>FI (QSU)</th>
<th>TFI</th>
<th>%C1</th>
<th>%C2</th>
<th>%C3</th>
<th>%C4</th>
<th>%C5</th>
<th>%C6</th>
<th>%C7</th>
<th>%C8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FCE-L</td>
<td>28.57</td>
<td>2.95</td>
<td>0.844</td>
<td>0.95</td>
<td>1.34</td>
<td>968.34</td>
<td>28%</td>
<td>9%</td>
<td>23%</td>
<td>12%</td>
<td>14%</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>2</td>
<td>FCE-S</td>
<td>9.67</td>
<td>2.72</td>
<td>0.263</td>
<td>0.98</td>
<td>1.44</td>
<td>360.11</td>
<td>34%</td>
<td>3%</td>
<td>15%</td>
<td>13%</td>
<td>15%</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>3</td>
<td>OKA-L</td>
<td>6.33</td>
<td>3.19</td>
<td>0.202</td>
<td>0.91</td>
<td>1.36</td>
<td>180.99</td>
<td>35%</td>
<td>1%</td>
<td>19%</td>
<td>11%</td>
<td>17%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>4</td>
<td>OKA-S</td>
<td>9.87</td>
<td>2.98</td>
<td>0.294</td>
<td>0.97</td>
<td>1.33</td>
<td>158.16</td>
<td>31%</td>
<td>3%</td>
<td>20%</td>
<td>11%</td>
<td>16%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>5</td>
<td>PAN-L</td>
<td>5.82</td>
<td>5.11</td>
<td>0.297</td>
<td>0.92</td>
<td>1.41</td>
<td>267.59</td>
<td>34%</td>
<td>2%</td>
<td>16%</td>
<td>10%</td>
<td>15%</td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>PAN-S</td>
<td>6.60</td>
<td>4.49</td>
<td>0.296</td>
<td>0.91</td>
<td>1.39</td>
<td>270.05</td>
<td>37%</td>
<td>0%</td>
<td>20%</td>
<td>11%</td>
<td>16%</td>
<td>7%</td>
<td>6%</td>
</tr>
</tbody>
</table>

TFI = total fluorescence (QSU)  
%CX = Relative abundance of PARAFAC component X  
L and S indicate Long or Short hydroperiod
Table 2. $^1$H NMR section integrals for key substructures of natural organic matter (SPE-DOM) as computed from $^1$H NMR bucket NMR integrals (cf. Fig. S2); owing to distribution of HSQC cross peaks, the distinction between aromatic and olefinic molecules was placed at $\delta^H = 6.5$ ppm (cf. Fig. 4).

**TABLE 2.** $^1$H NMR section integrals for key substructures of six wetland SPE-DOM samples.

<table>
<thead>
<tr>
<th>&amp; $\delta^H$ [ppm]</th>
<th>$10.0 \rightarrow 6.5$</th>
<th>$6.5 \rightarrow 5.3$</th>
<th>$4.9 \rightarrow 3.1$</th>
<th>$3.1 \rightarrow 1.9$</th>
<th>$1.9 \rightarrow 0.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>key substructures</td>
<td>H$_{ar}$</td>
<td>C=CH$_2$</td>
<td>O-$CH$</td>
<td>XCC$_H$</td>
<td>CCC$_H$</td>
</tr>
<tr>
<td>OKA-L</td>
<td>7.24</td>
<td>3.04</td>
<td>29.6</td>
<td>29.0</td>
<td>37.4</td>
</tr>
<tr>
<td>OKA-S</td>
<td>7.24</td>
<td>3.04</td>
<td>31.0</td>
<td>26.9</td>
<td>32.9</td>
</tr>
<tr>
<td>PAN-L</td>
<td>7.54</td>
<td>2.94</td>
<td>28.1</td>
<td>28.5</td>
<td>33.1</td>
</tr>
<tr>
<td>PAN-S</td>
<td>7.54</td>
<td>2.94</td>
<td>29.8</td>
<td>27.9</td>
<td>32.4</td>
</tr>
<tr>
<td>FCE-L</td>
<td>3.85</td>
<td>2.16</td>
<td>29.3</td>
<td>28.9</td>
<td>34.7</td>
</tr>
<tr>
<td>FCE-S</td>
<td>3.85</td>
<td>2.16</td>
<td>29.3</td>
<td>28.9</td>
<td>34.7</td>
</tr>
<tr>
<td>Members of Molecular series</td>
<td>OKA-L</td>
<td>OKA-S</td>
<td>PAN-L</td>
<td>PAN-S</td>
<td>FCE-L</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>CHO compounds</td>
<td>1581</td>
<td>1772</td>
<td>1711</td>
<td>1844</td>
<td>1400</td>
</tr>
<tr>
<td></td>
<td>(57.6%)</td>
<td>(60.0%)</td>
<td>(58.8%)</td>
<td>(56.5%)</td>
<td>(37.2%)</td>
</tr>
<tr>
<td>CHOS compounds</td>
<td>266</td>
<td>207</td>
<td>211</td>
<td>260</td>
<td>1127</td>
</tr>
<tr>
<td></td>
<td>(9.7%)</td>
<td>(6.8%)</td>
<td>(7.3%)</td>
<td>(8.0%)</td>
<td>(29.9%)</td>
</tr>
<tr>
<td>CHNO compounds</td>
<td>893</td>
<td>1075</td>
<td>984</td>
<td>1151</td>
<td>864</td>
</tr>
<tr>
<td></td>
<td>(32.5%)</td>
<td>(35.2%)</td>
<td>(33.8%)</td>
<td>(35.3%)</td>
<td>(22.9%)</td>
</tr>
<tr>
<td>CHNOS compounds</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>1581</td>
<td>1772</td>
</tr>
<tr>
<td></td>
<td>(0.2%)</td>
<td>(0.1%)</td>
<td>(0.2%)</td>
<td>(57.6%)</td>
<td>(60.0%)</td>
</tr>
<tr>
<td>total number of mass peaks</td>
<td>2745</td>
<td>3057</td>
<td>2911</td>
<td>3263</td>
<td>3766</td>
</tr>
<tr>
<td>total number of mass peaks</td>
<td>9830</td>
<td>10315</td>
<td>10588</td>
<td>10818</td>
<td>11692</td>
</tr>
<tr>
<td>percent of mass peaks</td>
<td>28%</td>
<td>30%</td>
<td>27.5%</td>
<td>30%</td>
<td>32%</td>
</tr>
<tr>
<td>attributed to CHO, CHOS, CHNO and CHNOS compositions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average H [%]</td>
<td>40.66</td>
<td>41.418</td>
<td>39.3417</td>
<td>39.5448</td>
<td>40.8924</td>
</tr>
<tr>
<td>average C [%]</td>
<td>38.77</td>
<td>37.517</td>
<td>38.7717</td>
<td>38.7717</td>
<td>38.7717</td>
</tr>
<tr>
<td>average O [%]</td>
<td>20.69</td>
<td>20.35</td>
<td>21.45</td>
<td>21.37</td>
<td>21.06</td>
</tr>
<tr>
<td>average N [%]</td>
<td>0.38</td>
<td>0.38</td>
<td>0.40</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>average S [%]</td>
<td>0.16</td>
<td>0.13</td>
<td>0.07</td>
<td>0.07</td>
<td>0.17</td>
</tr>
<tr>
<td>computed average H/C ratio</td>
<td>1.06</td>
<td>1.09</td>
<td>1.01</td>
<td>1.02</td>
<td>1.09</td>
</tr>
<tr>
<td>computed average O/C ratio</td>
<td>0.54</td>
<td>0.53</td>
<td>0.55</td>
<td>0.54</td>
<td>0.56</td>
</tr>
<tr>
<td>computed average C/N ratio</td>
<td>101.2</td>
<td>100.6</td>
<td>97.4</td>
<td>93.2</td>
<td>99.7</td>
</tr>
<tr>
<td>computed average C/S ratio</td>
<td>246</td>
<td>285</td>
<td>877</td>
<td>885</td>
<td>60.5</td>
</tr>
<tr>
<td>mass weighted average</td>
<td>378.2</td>
<td>375.2</td>
<td>388.8</td>
<td>386.0</td>
<td>386.7</td>
</tr>
</tbody>
</table>

**Table 3.** Counts of mass peaks in wetland SPE-DOM as computed from negative electrospray (ESI) 12 T FTICR mass spectra for singly charged ions.
Table S1. ¹H NMR section integral (800 MHz) of three different organic matter preparations in two different solvents, namely D₂O and CD₃OD, acquired at 283 K (cf. text).

<table>
<thead>
<tr>
<th>δ(¹H) ppm</th>
<th>10.0-7.0</th>
<th>7.0-5.3</th>
<th>4.9-3.1</th>
<th>3.1-1.9</th>
<th>1.9-0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>key substructures</td>
<td>H</td>
<td>C=H</td>
<td>O=CH</td>
<td>CH</td>
<td>CH</td>
</tr>
<tr>
<td>FCE-S UDOM</td>
<td>2.6</td>
<td>3.3</td>
<td>3.1</td>
<td>31.8</td>
<td>26.2</td>
</tr>
<tr>
<td>FCE-L EA</td>
<td>4.8</td>
<td>3.0</td>
<td>26.5</td>
<td>34.6</td>
<td>34.0</td>
</tr>
<tr>
<td>FCE-S PPL</td>
<td>2.7</td>
<td>2.0</td>
<td>43.0</td>
<td>26.5</td>
<td>25.8</td>
</tr>
<tr>
<td>key substructures</td>
<td>C=O</td>
<td>COX</td>
<td>C_α-O</td>
<td>C_α-CH</td>
<td>O-CH</td>
</tr>
<tr>
<td>FCE-S</td>
<td>2.5</td>
<td>13.8</td>
<td>2.5</td>
<td>10.3</td>
<td>2.4</td>
</tr>
<tr>
<td>FCE-L</td>
<td>1.6</td>
<td>13.8</td>
<td>2.2</td>
<td>9.5</td>
<td>0.9</td>
</tr>
<tr>
<td>OKA-L</td>
<td>2.2</td>
<td>14.8</td>
<td>5.2</td>
<td>17.2</td>
<td>2.4</td>
</tr>
<tr>
<td>PAN-S</td>
<td>1.8</td>
<td>14.0</td>
<td>5.0</td>
<td>17.2</td>
<td>2.7</td>
</tr>
<tr>
<td>NMR mixing model</td>
<td>C=O</td>
<td>COOH</td>
<td>C_α-O</td>
<td>C_α-CH</td>
<td>O-CH</td>
</tr>
<tr>
<td>H/C ratio</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>O/C ratio</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table S2. (Top): ¹³C NMR section integrals (percent of total carbon) and key substructures of wetland SPE-DOM. Bottom: Substructures used for basic NMR-derived reverse mixing model with nominal H/C and O/C ratios given (Hertkorn et al., 2013).
<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Figure</th>
<th>PK</th>
<th>NS</th>
<th>AQ [ms]</th>
<th>D1 [ms]</th>
<th>NE</th>
<th>WDW1</th>
<th>WDW2</th>
<th>PR1</th>
<th>PR2</th>
<th>SPE- DOM [mg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1$H NMR</td>
<td>2, S1</td>
<td>5TXI</td>
<td>512-1024</td>
<td>5000</td>
<td>10000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>1</td>
<td>3.7 – 9.5 mg</td>
</tr>
<tr>
<td>$^1$H NMR</td>
<td>3, S2</td>
<td>8QCI</td>
<td>512-2048</td>
<td>5000</td>
<td>5000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>1</td>
<td>$0.3 – 2$ mg</td>
</tr>
<tr>
<td>$^{13}$C NMR</td>
<td>S3</td>
<td>5D</td>
<td>74496</td>
<td>1000</td>
<td>14000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>35</td>
<td>OKA-L 4.7 mg</td>
</tr>
<tr>
<td>$^{13}$C NMR</td>
<td>S3</td>
<td>5D</td>
<td>44224</td>
<td>1000</td>
<td>14000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>35</td>
<td>PAN-S 4.2 mg</td>
</tr>
<tr>
<td>$^{13}$C NMR</td>
<td>S3</td>
<td>8QCO</td>
<td>23420</td>
<td>1000</td>
<td>19000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>35</td>
<td>FCE-L 9.5 mg</td>
</tr>
<tr>
<td>$^{13}$C NMR</td>
<td>S3</td>
<td>8QCO</td>
<td>3728</td>
<td>1000</td>
<td>19000</td>
<td>-</td>
<td>-</td>
<td>EM</td>
<td>-</td>
<td>35</td>
<td>FCE-S 9.1 mg</td>
</tr>
<tr>
<td>$^1$H, $^1$H TOCSY</td>
<td>43</td>
<td>5TXI</td>
<td>_24</td>
<td>1000</td>
<td>2500</td>
<td>1024</td>
<td>QS</td>
<td>EM</td>
<td>2.5</td>
<td>2.5</td>
<td>see caption</td>
</tr>
<tr>
<td>$^1$H, $^1$H TOCSY</td>
<td>43</td>
<td>8QCI</td>
<td>_12</td>
<td>1000</td>
<td>2500</td>
<td>1794</td>
<td>QS</td>
<td>EM</td>
<td>2.5</td>
<td>2.5</td>
<td>FCE-S 9.1 mg</td>
</tr>
<tr>
<td>$^1$H, $^{13}$C DEPT HSQC</td>
<td>45</td>
<td>8QCI</td>
<td>_320</td>
<td>250</td>
<td>1250</td>
<td>_256</td>
<td>QS</td>
<td>EM</td>
<td>2.5</td>
<td>2.5</td>
<td>FCE-S 9.1 mg</td>
</tr>
<tr>
<td>$^1$H, $^1$H JRES</td>
<td>S6</td>
<td>8QCI</td>
<td>_3072</td>
<td>1000</td>
<td>_500</td>
<td>_49</td>
<td>QS</td>
<td>QS</td>
<td>0</td>
<td>0</td>
<td>FCE-S 9.1 mg</td>
</tr>
<tr>
<td>$^1$H, $^{13}$C HSQC</td>
<td>45, S4, S5</td>
<td>8QCI</td>
<td>_1600</td>
<td>250</td>
<td>1250</td>
<td>_167</td>
<td>QS</td>
<td>EM</td>
<td>4</td>
<td>7.5</td>
<td>FCE-S</td>
</tr>
</tbody>
</table>

Table S23. Acquisition parameters of NMR spectra, shown according to figures. PK: probeheads used for acquisition of NMR spectra, 8QCI: cryogenic inverse geometry 5 mm z-gradient $^1$H/$^{13}$C/$^{15}$N/$^{31}$P QCI probe ($B_0$ = 18.8 T); 8QCO: cryogenic classical geometry 3 mm z-gradient $^1$H/$^{13}$C/$^{15}$N/$^{31}$P probe ($B_0$ = 18.8 T); 5TXI: cryogenic inverse geometry 5 mm z-gradient $^1$H, $^{13}$C, $^{15}$N probe ($B_0$ = 11.7 T); 5D: cryogenic classical geometry 5 mm z-gradient $^{13}$C, $^1$H probe ($B_0$ = 11.7 T); NS: number of scans (for 2D NMR: F2); AQ: acquisition time [ms]; D1: relaxation delay [ms]; NE: number of F1 increments in 2D NMR spectra; WDW1, WDW2: apodization functions in F1/ F2 (EM/GM: line broadening factor [Hz]; QS: shifted square sine bell; SI: sine bell); PR1, PR2: coefficients used for windowing functions WDW1, WDW2, EM/GM are given in [Hz]; SI/QS derived functions indicate shift by $\pi/n$. 

Formatted Table