



Royal Netherlands Institute for Sea Research

This is a pre-copyedited, author-produced version of an article accepted for publication, following peer review.

de Bar, M.W.; Weiss, G.M.; Yildiz, C.; Rampen, S.W.; Lattaud, J.; Bale, N.J.; Mienis, F.; Brummer, G.-J.A.; Schulz, H.; Rush, D.; Kim, J.-H.; Donner, B.; Knies, J.; Lückge, A.; Stuut, J.-B.; Sinninghe Damsté, J.S. & Schouten, S. (2020). Global temperature calibration of the Long chain Diol Index in marine surface sediments. *Marine Chemistry*, 224, 103815

Published version: <https://doi.org/10.1016/j.orggeochem.2020.103983>

NIOZ Repository: <http://imis.nioz.nl/imis.php?module=ref&refid=323732>

Research data: <https://doi.pangaea.de/10.1594/PANGAEA.912140>

[Article begins on next page]

The NIOZ Repository gives free access to the digital collection of the work of the Royal Netherlands Institute for Sea Research. This archive is managed according to the principles of the [Open Access Movement](#), and the [Open Archive Initiative](#). Each publication should be cited to its original source - please use the reference as presented.

When using parts of, or whole publications in your own work, permission from the author(s) or copyright holder(s) is always needed.

1 **Global temperature calibration of the Long chain Diol Index in marine surface sediments**

2

3 Marijke W. de Bar<sup>a,\*</sup>, Gabriella Weiss<sup>a,1</sup>, Caglar Yildiz<sup>a</sup>, Sebastiaan W. Rampen<sup>b</sup>, Julie  
4 Lattaud<sup>a,2</sup>, Nicole J. Bale<sup>a</sup>, Furu Mienis<sup>a</sup>, Geert-Jan A. Brummer<sup>a,c</sup>, Hartmut Schulz<sup>d</sup>, Darci  
5 Rush<sup>a</sup>, Jung-Hyun Kim<sup>e</sup>, Barbara Donner<sup>f</sup>, Jochen Knies<sup>g,h</sup>, Andreas Lückge<sup>i</sup>, Jan-Berend W.  
6 Stuur<sup>a,c</sup>, Jaap S. Sinninghe Damsté<sup>a,j</sup> and Stefan Schouten<sup>a,j,\*</sup>

7

8 <sup>a</sup>*NIOZ Royal Netherlands Institute for Sea Research, and Utrecht University, P.O. Box 59, 1790*  
9 *AB Den Burg, Texel, the Netherlands*

10 <sup>b</sup>*Universität Göttingen, Geoscience Center, Geobiology Group, Goldschmidtstrasse 3, 37073*  
11 *Göttingen, Germany*

12 <sup>c</sup>*Vrije Universiteit Amsterdam, Faculty of Science, Department of Earth Sciences, De Boelelaan*  
13 *1085, 1081HV Amsterdam, the Netherlands*

14 <sup>d</sup>*Universität Tübingen, Faculty of Science, Mikropaläontologie, Hölderlinstrasse 12, 72074*  
15 *Tübingen, Germany*

16 <sup>e</sup>*Korea Polar Research Institute (KOPRI), 26 Songdomirae-ro, Yeonsu-gu, Incheon 21990, South*  
17 *Korea*

18 <sup>f</sup>*Universität Bremen, Center for Marine Environmental Sciences, Leobener Strasse, 28359*  
19 *Bremen, Germany*

20 <sup>g</sup>*Geological Survey of Norway (NGU), P.O. Box 6315, 7040 Trondheim, Norway*

21 <sup>h</sup>*CAGE - Centre for Arctic Gas Hydrate, Environment and Climate; Department of Geosciences,*  
22 *UiT The Arctic University of Norway, 9037 Tromsø, Norway*

23 <sup>i</sup>*Federal Institute for Geosciences and Natural Resources, Stilleweg 2, 30655 Hannover,*

24 *Germany*

25 <sup>1</sup>*Utrecht University, Faculty of Geosciences, Department of Earth Sciences, P.O. Box 80115,*

26 *3508 TC Utrecht, the Netherlands*

27

28 \* Corresponding author: Marijke W. de Bar ([marijkedebbar@gmail.com](mailto:marijkedebbar@gmail.com)) and Stefan Schouten

29 ([s.schouten1@uu.nl](mailto:s.schouten1@uu.nl))

30 <sup>1</sup> Present address: Pennsylvania State University, Jordan Road, 16563, University Park, PA, USA

31 <sup>2</sup> Present address: ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland

32

33 **ABSTRACT**

34 The Long chain Diol Index (LDI) is a relatively new organic geochemical proxy for sea surface  
35 temperature (SST), based on the abundance of the C<sub>30</sub> 1,15-diol relative to the summed  
36 abundance of the C<sub>28</sub> 1,13-, C<sub>30</sub> 1,13- and C<sub>30</sub> 1,15-diols. Here we substantially extend and re-  
37 evaluate the initial core top calibration by combining the original dataset with 172 data points  
38 derived from previously published studies and 262 newly generated data points. In total, we  
39 considered 595 globally distributed surface sediments with an enhanced geographical coverage  
40 compared to the original calibration. The relationship with SST is similar to that of the original  
41 calibration but with considerably increased scatter. The effects of freshwater input (e.g., river  
42 runoff) and long-chain diol contribution from *Proboscia* diatoms on the LDI were evaluated.  
43 Exclusion of core-tops deposited at a salinity < 32 ppt, as well as core-tops with high *Proboscia*-  
44 derived C<sub>28</sub> 1,12-diol abundance, resulted in a substantial improvement of the relationship  
45 between LDI and annual mean SST. This implies that the LDI cannot be directly applied in  
46 regions with a strong freshwater influence or high C<sub>28</sub> 1,12-diol abundance, limiting the  
47 applicability of the LDI. The final LDI calibration ( $LDI=0.0325 \times SST + 0.1082$ ;  $R^2 = 0.88$ ;  $n =$   
48 514) is not statistically different from the original calibration of Rampen et al. (2012)  
49 (<https://doi.org/10.1016/j.gca.2012.01.024>), although with a larger calibration error of 3 °C. This  
50 larger calibration error results from several regions where the LDI does not seem to have a strong

51 temperature dependence with annual mean SST, posing a limitation on the application of the  
52 LDI.

53

54 *Keywords:* LDI core-top calibration, long-chain diols, SST, freshwater, *Proboscia* diatoms

55

## 56 **1. Introduction**

57 The present-day release of anthropogenic greenhouse gases into the atmosphere has  
58 resulted in warming of the Earth's atmosphere and surface oceans, which is expected to continue  
59 in the coming decades (IPCC, 2014). However, the actual extent of this temperature rise and its  
60 implications for global climate is difficult to accurately predict due to the complexity of the  
61 Earth's climate system. For the prediction of future climate conditions, we typically rely on  
62 computer simulations of ocean-atmosphere circulation models, which in turn rely on time-series  
63 of observational data of various climate parameters. However, instrumental records only extend  
64 back to the last century. To accurately predict climate, it is essential to study natural climate  
65 evolution on geological timescales. For this purpose, a variety of climate proxies need to be used.

66 One of the most important climate parameters is past sea surface temperature (SST),  
67 since oceans make up more than two thirds of the world's surface, and therefore profoundly  
68 influence (and respond to) global climate. Proxies are commonly based on measurements of  
69 either inorganic or organic remnants of organisms preserved in sediment. One of the most  
70 commonly applied inorganic paleotemperature proxies uses the stable oxygen isotopic  
71 composition ( $\delta^{18}\text{O}$ ) of the carbonate shells of foraminifera (e.g., Emiliani, 1955; Shackleton et  
72 al., 1967). Also, the Mg/Ca ratio measured in a foraminiferal shell is correlated with temperature  
73 (e.g., Nürnberg et al., 1996). Organic temperature proxies, on the other hand, are generally based

74 on lipid biomarkers, which are specific for a certain organism or a group of organisms. In  
75 paleoclimate studies there are two biomarker proxies for SST which are frequently applied. The  
76 first is the  $U_{37}^{K'}$  index, based on long-chain unsaturated alkenones, detected in marine sediments  
77 world-wide, which uses the ratio of the di-unsaturated C<sub>37</sub> methyl alkenones over the tri-  
78 unsaturated C<sub>37</sub> methyl alkenones (Brassell et al., 1986; Prahl and Wakeham, 1987). This ratio is  
79 positively correlated with temperature, since the modern-day alkenone producers (mainly  
80 *Emiliania huxleyi* and *Gephyrocapsa oceanica*; e.g., Volkman et al., 1980, 1995; Marlowe et al.,  
81 1984; Conte et al., 1995) synthesize C<sub>37</sub> alkenones, of which the degree of saturation varies with  
82 growth temperature (Brassell et al., 1986; Prahl and Wakeham, 1987). The temperature range of  
83 the proxy is between -2 °C and ca. 29°C (Müller et al., 1998; Conte et al., 2006; Tierney and  
84 Tingley, 2018).

85         The second proxy, TEX<sub>86</sub>, is based on the distribution of isoprenoid glycerol dialkyl  
86 glycerol tetraethers (GDGTs), produced by the archaeal phylum Thaumarchaeota (see Schouten  
87 et al., 2013 for a review). These archaea synthesize GDGTs containing 0–3 cyclopentane  
88 moieties (GDGT-0 to GDGT-3) and crenarchaeol, which contains 4 cyclopentane rings and a  
89 cyclohexane moiety (Schouten et al., 2002; Sinninghe Damsté et al., 2002). The proxy is based  
90 on the relative abundance of GDGT-1, GDGT-2 and GDGT-3 and an isomer of crenarchaeol  
91 (Schouten et al., 2002). The index is positively correlated with annual mean SST, showing an  
92 increase in the number of cyclopentane moieties with increasing temperature, and can be applied,  
93 with caution, at temperatures > 30 °C. All these proxies have advantages but also recognized  
94 uncertainties, and since these uncertainties are proxy-specific, SST reconstructions are ideally  
95 based on multiple proxies. Accordingly, the development of additional proxies is desired.

96 Rampen et al. (2012) proposed the Long chain Diol Index (LDI), based on the fractional  
97 abundances of long-chain alkyl diols (LCDs), specifically the C<sub>28</sub> and C<sub>30</sub> 1,13-diols and C<sub>30</sub>  
98 1,15-diols, which contain a hydroxy group at C<sub>1</sub> and a hydroxy group at the C<sub>13</sub> or C<sub>15</sub> position,  
99 respectively:

$$100 \quad \text{LDI} = [\text{C}_{30} \text{ 1,15-diol}] / ([\text{C}_{28} \text{ 1,13-diol} + \text{C}_{30} \text{ 1,13-diol} + \text{C}_{30} \text{ 1,15-diol}]) \quad (1)$$

101 The LDI, based on 161 globally distributed core-top sediments, shows a strong correlation with  
102 SST described by the following transfer function:

$$103 \quad \text{LDI} = 0.033 \times \text{SST} + 0.095 \quad (n = 161; R^2 = 0.97; \text{RE} = 2.0 \text{ } ^\circ\text{C}) \quad (2)$$

104 The main limitation of the LDI is the fact that the producers of the 1,13-diols and 1,15-  
105 diols in the ocean are still unknown. Cultured freshwater and marine eustigmatophyte algae  
106 produce 1,13-diols and 1,15-diols (Volkman et al., 1992; 1999; Gelin et al., 1997; Méjanelle et  
107 al., 2003; Shimokawara et al., 2010; Rampen et al., 2014b), but the LCD distributions observed  
108 in the cultures are dissimilar from the distributions observed in the marine environment.  
109 Moreover, these eustigmatophytes rarely occur in the ocean (e.g., Balzano et al., 2018). In  
110 contrast, Shimokawara et al. (2010) observed that the LCD distributions in the eustigmatophyte  
111 *Nannochloropsis* sp. (containing a dominant C<sub>32</sub> 1,15-diol) were similar to that observed in  
112 sediments of Lake Baikal, suggesting that eustigmatophytes might produce LCDs in lakes.  
113 Additionally, Villanueva et al. (2014) observed similar trends for 18S rRNA gene copy numbers  
114 of (yet unknown) eustigmatophytes with LCD concentrations in an African lake, confirming  
115 eustigmatophytes as potential LCD producers in freshwater. However, Rampen et al. (2014b)  
116 tested the LDI in 62 lakes and found that the correlation with temperature was weak ( $R^2 = 0.33$ ),  
117 which is likely because of the presence of different eustigmatophytes, each possessing different

118 LCD distributions, implying that the applicability of the LDI may be limited to the marine  
119 environment.

120 Besides unknown producers, other issues with the LDI have been recognized. De Bar et  
121 al. (2016) and Lattaud et al. (2017a) observed that the LDI-derived temperatures in surface  
122 sediments near river mouths significantly deviated from satellite-derived SSTs. The diol  
123 distributions are characterized by elevated C<sub>32</sub> 1,15-diol abundance, due to the freshwater input  
124 where the C<sub>32</sub> 1,15 diol occurs in high abundance (Rampen et al., 2014b). Consequently,  
125 applying the LDI in marine regions with riverine input should be done with caution. Rodrigo-  
126 Gámiz et al. (2015) showed that for surface sediments and suspended particulate matter (SPM) in  
127 the subpolar region around Iceland, the LDI underestimated satellite-derived SST. Relatively  
128 high C<sub>28</sub> and C<sub>30</sub> 1,14-diol abundances were observed in this area, which are characteristic for  
129 *Proboscia* diatoms (Sinninghe Damsté et al., 2003; Rampen et al., 2007), although they were  
130 also identified in the estuarine species *Apedinella radians* (Rampen et al., 2011). Accordingly,  
131 the authors hypothesized that *Proboscia* diatoms (at least partially) contributed to the 1,13- and  
132 1,15-diol production, and thereby compromised the LDI. For surface sediments in the Okhotsk  
133 Sea, also a subpolar region, the LDI correlated with SST, but this relationship was statistically  
134 different from the global calibration (Lattaud et al., 2018b). Lastly, down-core applications of the  
135 LDI have shown that the index is promising as a SST proxy but often reveals a slightly larger  
136 glacial-interglacial temperature amplitude than found for  $U_{37}^{K'}$  and TEX<sub>86</sub> records (Rampen et al.,  
137 2012; Lopes dos Santos et al., 2013; Rodrigo-Gámiz et al., 2014; Jonas et al., 2017; de Bar et al.,  
138 2018).

139 Thus, despite promising down-core applications, questions remain about the calibration  
140 of this proxy and in which environments it can be applied. Therefore, in this study we



141 substantially extended the initial global LDI core-top calibration of Rampen et al. (2012) with  
142 literature data and newly generated data, adding 434 data points and considerably increasing  
143 global coverage. Comparison with SST and salinity allowed us to determine the main controlling  
144 factors and identify potential constraints on the applicability of the LDI.

145

## 146 **2. Materials and methods**

### 147 *2.1. Surface sediments*

148 We have combined the global core-top LCD data of Rampen et al. (2012) with other  
149 previously published LCD data and newly acquired core-top data. We re-evaluated the original  
150 LDI core-top dataset of Rampen et al. (2012), consisting of 209 measurements (black dots in Fig.  
151 1), from which 161 LDI data points were used in the original calibration dataset. For this dataset  
152 we quantified additional diols, in particular the C<sub>28</sub> 1,12-diol. Re-integration has led to minor  
153 changes in LDI values of <0.08. For eleven samples, we could not retrieve the original data and  
154 therefore we were not able to reintegrate the LCD peak areas (indicated in the Supplementary  
155 Table S1). Differences in contributions of the selected ions to the total ion counts (*m/z* 50–800)  
156 of saturated vs unsaturated LCDs were considered by applying correction factors as described by  
157 Rampen et al. (2009). For the dataset of Rampen et al. (2012), we applied two different  
158 correction factors to the two mass spectrometer (MS) systems on which the LCDs were analyzed  
159 (Supplementary Table S1). Additionally, we re-integrated some previously published LCD data  
160 (pink dots in Fig. 1), i.e., the core-top sediment data from around Iceland of Rodrigo-Gámiz et al.  
161 (2015), the Iberian margin surface sediment data of de Bar et al. (2016), the Gulf of Lion,  
162 Amazon Basin, Berau delta and Kara Sea data of Lattaud et al. (2017a), the Mozambique  
163 Channel data of Lattaud et al. (2017b), the Okhotsk Sea data of Lattaud et al. (2018a) and part of

164 the Black Sea data of Lattaud et al. (2018b). For the published LCD data of Lattaud et al.  
165 (2017a,b), we have re-evaluated the quality of the raw data (i.e., chromatographic separation,  
166 signal-to-noise levels) and based on this we used 97 of 160 data points. We adopted the LDI data  
167 from the region around Australia (Smith et al., 2013), but did not have the original MS data and  
168 thus were not able to re-evaluate the LCD distributions. Re-evaluation of data of de Bar et al.  
169 (2016) showed that the fractional abundances of the unsaturated LCDs were not corrected for the  
170 differences in contributions of the selected ions to the total mass spectrum, which is corrected  
171 here (Supplementary Table S1). In total, the previously published data comprise 233 sediment  
172 locations. Additionally, we analyzed 105 polar fractions for long-chain diols that had been  
173 analyzed previously by Kim et al. (2008, 2010; white dots in Fig. 1) for the global TEX<sub>86</sub> core-  
174 top calibration. Furthermore, we analyzed 186 new core-tops from several regions for a better  
175 spatial coverage (white dots in Fig. 1). In total, 731 surface sediment samples were considered in  
176 this study.

177

## 178 *2.2. Lipid extraction and instrumental analysis*

179 The 186 new surface sediments (mostly 0–1 or 0–0.5 cm) were freeze-dried and extracted  
180 with an Accelerated Solvent Extractor (ASE 200; Dionex) using a dichloromethane:methanol  
181 (DCM:MeOH) mixture (9:1; v/v) at a temperature of 100°C and a pressure of  $7\text{--}8 \times 10^6$  Pa.  
182 Lipid extracts were dried under nitrogen and separated into three fractions (apolar, ketone, polar)  
183 using activated (2h at 150°C) Al<sub>2</sub>O<sub>3</sub>. Separation was achieved using the eluents hexane/DCM  
184 (9:1; v/v), hexane/DCM (1:1; v/v) and DCM/MeOH (1:1; v/v), respectively (Method 2 in  
185 Supplementary Table S1). The polar fractions were silylated by the addition of pyridine and  
186 N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) and heating at 60°C for 20 min. Prior to

187 injection, ethyl acetate was added. GC–MS analyses were done on an Agilent 7890B gas  
188 chromatograph interfaced with an Agilent 5977A mass spectrometer. Samples were injected on-  
189 column at a starting temperature of 70°C, which was programmed to 130°C at 20°C min<sup>-1</sup>, and a  
190 subsequent gradient of 4°C min<sup>-1</sup> to the end temperature of 320°C, which was kept for 25 min.  
191 The GC was equipped with a fused silica column (25 m × 0.32 mm) with a CP Sil-5 coating  
192 (film thickness 0.12 μm). Helium was used as carrier gas with a constant flow of 2 ml min<sup>-1</sup>, and  
193 the MS operated with an ionization energy of 70 eV. We identified the LCDs in full scan,  
194 scanning from *m/z* 50 to *m/z* 850, based on their characteristic fragmentation patterns (de Leeuw  
195 et al., 1981; Versteegh et al., 1997). Quantification of the LCDs was achieved in selected ion  
196 monitoring (SIM) mode of the characteristic fragmentation ions (i.e., *m/z* 299, 313, 327 and 341;  
197 Rampen et al., 2012). The LDI was calculated according to Rampen et al. (2012) integrating the  
198 relevant peak areas in SIM mode (Eq. 1). For the calculation of fractional abundances, we  
199 applied a correction factor for the relative contribution of the selected fragments during SIM to  
200 the total ion counts for the saturated (16.1%) vs unsaturated (9.1%) LCDs.

201 Lipid extraction methods for re-analyzed polar fractions (from Kim et al., 2010) and the  
202 LCD data which we re-evaluated or adopted, are described in the original literature. We have  
203 classified these methods into five groups, indicated in the Supplementary Table S1 (Methods 1–  
204 5). Generally, these methods differ in extraction protocol, i.e. ASE, ultrasonic extraction, Bligh  
205 and Dyer, or ASE followed by saponification of the extract or in fractionation protocol, i.e. two  
206 (apolar-polar) or three fraction (apolar-ketone-polar) separation using Al<sub>2</sub>O<sub>3</sub>, three fraction  
207 separation using silica gel or separation of core lipids and intact polar lipids over silica gel.

208

### 209 2.3. Oceanographic data

210 The LCD data were compared with temperature, salinity and nutrient data from the World  
211 Ocean Atlas 2013 (WOA13). Annual mean, seasonal sea surface temperatures ( $^{\circ}\text{C}$ ; 0 m depth),  
212 temperatures for different depths, and salinity were obtained from the  $0.25^{\circ}$  grid databases  
213 (decadal averages over the period of 1955 to 2012; Locarnini et al., 2013; Zweng et al., 2013).  
214 Annual mean phosphate and nitrate concentrations ( $\mu\text{mol l}^{-1}$ ) were obtained from the  $1^{\circ}$  grid  
215 WOA13 databases (Garcia et al., 2014b). In case there was no temperature data for the  $0.25^{\circ}$  grid  
216 corresponding to the core-top location, we adopted the SST value for the closest  $0.25^{\circ}$  grid. For  
217 SST, we used data within  $1^{\circ}$  distance; if not available, we did not include the data. Exceptions  
218 are thirteen Antarctic sediments for which we used seasonal SST data of  $1\text{--}2^{\circ}$  degrees away, as  
219 in these regions SST data were generally scarcer. For salinity, we used values within  $1.5^{\circ}$   
220 distance of the core-top location, and for phosphate and nitrate we stayed within a  $3^{\circ}$  radius. For  
221 a principal component analysis, we also obtained oxygen saturation (%), dissolved oxygen ( $\text{ml l}^{-1}$ )  
222 and silicate concentrations ( $\mu\text{mol l}^{-1}$ ) from the World Ocean Atlas 2013 (Garcia et al.,  
223 2014a,b). The oxygen and silicate data were obtained from  $1^{\circ}$  grid databases, and only data of  
224 the grids corresponding to the core-top locations were used; i.e., in case not available, we have  
225 not adopted data from grids nearby.

226

### 227 2.4. Statistical analysis

228 We performed Principal Component Analysis (PCA) on the fractional abundances of the  
229 various LCDs, annual mean SST, salinity, and phosphate and nitrate concentrations, using the  
230 XLSTAT software (Addinsoft, 2018) (Fig. 3). Furthermore, PCA analysis was done on LDI,  
231 dissolved oxygen, oxygen saturation, nitrate, phosphate, silicate, salinity, SST and LDI

232 (Supplementary Fig. S4). When an LCD was not detected, its fractional abundance was  
233 considered zero. In both PCA, data points were not plotted in case data for one or more of the  
234 parameters (i.e., LDI, salinity, SST, phosphate, etc.) was missing, or if the fractional abundance  
235 of one of the long-chain diols was unknown (e.g., due to co-elution). R statistical software was  
236 used for multiple linear regression analyses and for the comparison of different regression slopes  
237 by means of analysis of covariance (ANCOVA).

238

### 239 **3. Results and discussion**

240 We have combined several data sets to extend the core-top calibration of the LDI, as  
241 originally published by Rampen et al. (2012), to re-assess the calibration of the index, and to  
242 identify potential constraints on the proxy. We have re-evaluated the core-top data of Rampen et  
243 al. (2012), and re-evaluated and adopted previously published LCD data (Smith et al., 2013;  
244 Rodrigo-Gámiz et al., 2015; de Bar et al., 2016; Lattaud et al., 2017a,b, 2018a,b) and added new  
245 core-top data, which in total resulted in LCD distributions from 731 core-tops. Of these 731  
246 sediment samples, we have excluded 136 samples because of quantification limit issues, related  
247 to high backgrounds obscuring the signals of the LDI-diols or low abundances of all LCDs,  
248 compromising the reliability of the LDI values. Furthermore, for certain samples fractional  
249 abundances are given but not the LDI (Supplementary Table S1), as the relative diol abundances  
250 were considered too low for index calculation. For instance, some core-tops are dominated by  
251 1,14-diols, for which the Diol Indices were calculated, but not the LDI as the 1,13- and 1,15-  
252 diols were present in too low amounts. In total we obtained 595 LDI data points and cover an  
253 annual mean temperature range of -1.8 °C to 30.3°C (Fig. 2; Supplementary Table S1). Although

254 we have improved global coverage compared to Rampen et al. (2012), it must be noted that  
255 almost 40% of our surface sediments originate from the tropical temperature regime, i.e. > 25°C.

256 We have linearly cross-correlated our LDI core-top data with annual mean SST  
257 (WOA13; Locarnini et al., 2013), resulting in a positive regression and a coefficient of  
258 determination ( $R^2$ ) of 0.82 (Fig. 2), confirming that the LDI contains a strong temperature signal.  
259 The relationship ( $LDI=0.0323\times SST+0.1111$ ) is statistically not different from the slope and  
260 intercept (ANCOVA  $p$ -value > 0.1) of the original LDI-relation as proposed by Rampen et al.  
261 (2012; Eq. 2). Additionally, we performed a Principal Component Analysis to reveal possible  
262 relationships between environmental factors and long-chain diol proxies (Fig. 3a) and between  
263 individual LCDs (Fig. 3b). The first component (PC) in the first PCA (Fig. 3a) explains 51.1% of  
264 the variance, with the strongest positive loadings of annual mean SST and LDI, suggesting that  
265 temperature is the main control of variance on the first component, which is confirmed by the  
266 strong correlation between the Factor 1 scores and SST, with a coefficient of determination ( $R^2$ )  
267 of 0.79 (Fig. 3c). The  $C_{28}$  and  $C_{30}$  1,13-diols load opposite of the  $C_{30}$  1,15-diol on Factor 1 (Fig.  
268 3b). This is consistent with the  $C_{30}$  1,15-diol abundance, and thus the LDI, being higher with  
269 higher SST while the 1,13-diol is more abundant with lower SST (cf. Rampen et al., 2012).  
270 Accordingly, the variance explained by Factor 1 (28%) in Fig. 3b is also likely primarily  
271 temperature, confirmed by the relatively high coefficient of determination ( $R^2 = 0.66$ ) for the  
272 scores of Factor 1 and annual mean SST (Fig. 3e). The  $C_{32}$  1,15-diol has almost no factor loading  
273 on the Factor 1 axis in both plots, suggesting that temperature has minimal influence on the  
274 abundance of the  $C_{32}$  1,15-diol, which is consistent with the results of Rampen et al. (2012).

275 The fractional abundance of the  $C_{32}$  1,15-diol (cf. de Bar et al., 2061) has a strong  
276 negative loading on the axis of Factor 2 (Fig. 3a), where salinity shows the strongest positive

277 loading. The abundance of the C<sub>32</sub> 1,15-diol has previously been linked to freshwater influence  
278 (e.g., Rampen et al., 2014; de Bar et al., 2016; Lattaud et al., 2017a; 2017b). Salinity is used as  
279 an indicator of freshwater input, suggesting the variance explained by Factor 2 is significantly  
280 influenced by freshwater input (and/or salinity), confirmed by the coefficient of determination  
281 (R<sup>2</sup>) of 0.56 (Fig. 3d).

282 A number of data points do not fall close to the calibration line and the decrease in  
283 coefficient of determination (R<sup>2</sup>) from 0.97 of the original calibration of Rampen et al. (2012) to  
284 0.82 in this study. In the following sections we will discuss factors potentially responsible for  
285 this increased scatter in the LDI calibration and evaluate possible constraints on the proxy.

286

### 287 *3.1. Non-marine LCD contributions*

288 Previous studies have shown that river outflow, i.e. freshwater input, can compromise the  
289 LDI (de Bar et al., 2016; Lattaud et al., 2017a), although the effect of riverine input on the LDI is  
290 likely specific for each region. For instance, de Bar et al. (2016) observed lower LDI-derived  
291 SSTs than satellite-derived SSTs in surface sediments close to the river mouths on the  
292 Portuguese margin, whereas Lattaud et al. (2017b) obtained LDI SSTs which were significantly  
293 higher than satellite SSTs for Kara Sea sediments closest to the Yenisei River. Surface sediments  
294 deposited in low salinity environments (<32 ppt), which are impacted by river inflow, are mainly  
295 derived from the Hudson Bay (salinity 26–31 ppt), the Baltic Sea (7–30 ppt), the Black Sea (11–  
296 18 ppt), and the Kara Sea (10–32 ppt). While, despite the riverine input, the annual mean salinity  
297 at the Portuguese margin in the region studied by de Bar et al. (2016) is above 32 ppt. The LDI  
298 values from the Baltic Sea, the Gulf of St. Lawrence and the Black Sea are clearly positioned  
299 above the regression line (Fig. 2a) with temperature differences up to ca. +14.5 °C. This might

300 suggest that the LDI is unlikely to work in low salinity environments, consistent with Rampen et  
301 al. (2014b) who observed that the LDI cannot be applied to lakes. Although there is no  
302 significant relationship between salinity and the LDI or its residual error, plotting the residual  
303 errors of the LDI calibration vs salinity shows that LDI estimates for areas with the lowest  
304 salinities (Supplementary Fig. S1b) overestimate SST. Accordingly, when we exclude surface  
305 sediments deposited at salinities < 32 ppt (57 samples originating from the Baltic Sea, the Black  
306 Sea, the Hudson Bay, the Gulf of St. Lawrence and the Kara Sea), the coefficient of  
307 determination improves ( $R^2 = 0.86$ ; Fig. 2c). However, note that this also results in a change in  
308 calibration slope from 0.0323 to 0.0339 (Fig. 2). The core-tops from the Hudson Bay and the  
309 Gulf of St. Lawrence originate from the dataset of Rampen et al. (2012) but were also excluded  
310 in the original core-top calibration. It remains remarkable but unclear why the samples from the  
311 Portuguese margin, also influenced by freshwater but with salinities > 32 ppt, do not follow this  
312 trend in overestimating SST, but provides lower SST estimates instead.

313 An alternative way to screen sediments for the impact of freshwater influence could be to  
314 use the abundance of the  $C_{32}$  1,15-diol, which is often elevated near rivers (Versteegh et al.,  
315 1997; 2000; Rampen et al., 2014b; de Bar et al., 2016; Lattaud et al., 2017a; 2017b). However,  
316 there is no clear relation between high  $C_{32}$  1,15-diol abundances and residual errors of the LDI  
317 (Supplementary Fig. S1a) and removal of core-tops with high  $C_{32}$  1,15-diol abundances > 0.3,  
318 only results in a very small improvement in the coefficient of determination ( $R^2 = 0.83$ ;  
319 Supplementary Fig. S1). Accordingly, the fractional abundance of the  $C_{32}$  1,15-diol does not  
320 provide a strong indication for biases in the LDI caused by freshwater input on a global scale.  
321 We also tested if  $C_{32}$  1,15-diol abundances can be used to correct for a possible salinity effect, by  
322 applying multiple linear regression with SST and calculated relative abundances of  $C_{28}$  and  $C_{30}$



323 1,13- and C<sub>30</sub> and C<sub>32</sub> 1,15-diols from the original dataset. The outcome did not result in a higher  
324 coefficient of determination with SST ( $R^2 = 0.83$ ), suggesting C<sub>32</sub> 1,15-diol abundances cannot  
325 be used to correct for salinity effects. In any case, this study, as well as other studies (e.g.,  
326 Rampen et al., 2012; de Bar et al., 2016; Lattaud et al., 2017a), show that the application of the  
327 LDI in low salinity environments or very close near river mouths may be problematic. The weak,  
328 non-significant correlations between the LDI and its residual errors vs salinity or the fractional  
329 abundance of the C<sub>32</sub> 1,15-diol indicates that neither salinity nor C<sub>32</sub> 1,15-diol are conclusive  
330 indications for the compromising influence of freshwater. This influence is most likely specific  
331 for every region, and may depend, amongst others, on the freshwater sources from which long-  
332 chain diols can derive. Rivers, lakes, estuaries, inland seas, wetlands, sea/land ice, etc. may each  
333 contain different diol producers, and therefore contribute different long-chain diol distributions.

334

### 335 *3.2. Influence of Proboscia lipids on the LDI*

336 Our dataset includes the core-top data of Rodrigo-Gámiz et al. (2015) of the subpolar  
337 region around Iceland. However, as mentioned in the introduction, the LDI substantially  
338 underestimates satellite SSTs in this region, likely because *Proboscia* diatoms seem to be at least  
339 a partial source of the 1,13-diols. Since this may also occur in other regions, we screened for the  
340 influence of *Proboscia* diatoms using two Diol Indices based on 1,14-diols vs the 1,15-diol and  
341 vs the 1,13 diols (Rampen et al., 2008; Willmott et al., 2010, respectively). Cross-correlating the  
342 temperature difference between annual mean SST and the LDI-regression based SST (from Fig.  
343 2c) with these indices (Supplementary Fig. S2), does not reveal strong correlations, although  
344 several sediments have a high Diol Index 1 (Rampen et al., 2008) as well as a cold bias in the  
345 LDI-derived SST. However, most core-tops with such a high Diol Index do not reveal this cold

346 bias. Thus, the Diol Indices do not provide an unambiguous indication for a potential bias on the  
347 LDI caused by *Proboscia* LCD contribution.

348 We closely examined the data points which fall well below the LDI regression line, i.e.,  
349 many of the Iceland data-points (pink diamonds), the Okhotsk Sea (brown triangles) and the  
350 North Atlantic Ocean (yellow squares; transect Ireland-Greenland) (Fig. 2a). Examination of the  
351 chromatograms of these respective samples often showed an atypical LCD distribution compared  
352 to other marine sediments (Fig. 4). One sediment from the Gulf of Mexico also shows this  
353 unusual distribution, and interestingly this is the only data-point of this region which also  
354 substantially underestimates SST when compared to the LDI regression (LDI 0.43, SST 24.5°C;  
355 Fig. 2a). These sediments are characterized by high mono-unsaturated and saturated 1,14-diol  
356 abundances, but also relatively high C<sub>26</sub> and C<sub>28</sub> 1,12-diols. The Okhotsk Sea sediments also  
357 contain relatively high abundances of the C<sub>26</sub> 1,13-diol. The C<sub>28</sub> 1,12-diol has been observed in  
358 low amounts in lake sediments (Shimokawara et al., 2010; Rampen et al., 2014a), freshwater  
359 eustigmatophyte algae (Volkman et al., 1999; Rampen et al., 2014a) and in *Proboscia* diatoms  
360 (Rampen et al., 2007) as well as in marine sediments with high 1,14-diol concentrations  
361 (Willmott et al., 2010; Rampen et al., 2007; ten Haven and Rullkötter, 1991; de Bar et al., 2018).  
362 The C<sub>26</sub> 1,12-diol has been observed in cultures of *P. inermis* and *P. indica* (Rampen et al.,  
363 2007), and in Eocene-Oligocene (between ca. 50–30 Ma) sediments from the Falkland Plateau  
364 (southwest Atlantic Ocean; Plancq et al., 2014) and the New Jersey shelf (de Bar et al., 2019). In  
365 sediments where we detected relatively high abundances of the C<sub>26</sub> and C<sub>28</sub> 1,12-diols, we also  
366 detected relatively high abundances of the C<sub>27</sub> and C<sub>29</sub> 12-hydroxy and C<sub>28</sub> and C<sub>30</sub> 13-hydroxy  
367 methyl alkanoates (Fig. 4). Sinnighe Damsté et al. (2003) and Rampen et al. (2007) observed  
368 C<sub>27</sub> and C<sub>29</sub> 12- hydroxy methyl alkanoates in cultures of *P. indica*, *P. alata* and *P. inermis*.

369 Small amounts of the C<sub>28</sub> and C<sub>30</sub> 13-hydroxy methyl alkanoates were detected solely in *P.*  
370 *indica*. Since, to date, *Proboscia* is the only group of organisms known to produce these  
371 hydroxyl methyl alkanoates and 1,14- and 1,12-diols, this is a very strong indication that the  
372 LCDs and mid-chain hydroxy methyl alkanoates in these sediments are produced by *Proboscia*  
373 species. However, the unusual distributions with the high C<sub>26</sub> and C<sub>28</sub> 1,12-diols and C<sub>28</sub> and C<sub>30</sub>  
374 12-hydroxy methyl alkanoates detected in several sediments compared to those of cultures also  
375 strongly suggest that they are sourced by *Proboscia* species that have not yet been cultured and  
376 evaluated for LCDs.

377 The fact that the sediments with unusual distributions of 1,12-diols all reveal very low  
378 LDI values compared to the LDI regression, indicates that particular *Proboscia* species likely  
379 contribute 1,13-diols, thereby compromising the LDI. Consequently, we calculated the fractional  
380 abundance of the C<sub>28</sub> 1,12-diol (with respect to the C<sub>28</sub> 1,12-, 1,13-, 1,14-, C<sub>30</sub> 1,13-, 1,14- and  
381 1,15-diols):

$$382 \quad FC_{28} \text{ 1,12-diol} = [C_{28} \text{ 1,12-diol}] / [C_{28} \text{ 1,12} + C_{28} \text{ 1,13} + C_{28} \text{ 1,14} + C_{30} \text{ 1,13} + C_{30} \text{ 1,14} + C_{30} \text{ 1,15-diols}]$$

383 (3)

384 and subsequently plotted the  $FC_{28}$  1,12-diol against the residual temperature errors of the LDI  
385 calibration (i.e., LDI SST – AM SST; Fig. 5). This showed that the fractional abundance of the  
386 C<sub>28</sub> 1,12-diol is below 0.1 for the large majority of the sediments (~95% of total), but when the  
387 fractional abundance increases, the LDI is biased towards colder temperatures (up to ca. –14 °C  
388 difference). For this reason, the relative abundance of the C<sub>28</sub> 1,12-diol might serve as a better  
389 indication for the influence of certain *Proboscia* species on the LDI than the abundance of 1,14-  
390 diols. Interestingly, the regions in which we observe *Proboscia* influence on the LDI as signified  
391 by the high C<sub>28</sub> 1,12-diol abundance are, except for the one core-top in the Gulf of Mexico,

392 located between 45 and 65 °N, suggesting that these *Proboscia* diatoms thrive mainly in these  
393 high-latitude areas. An alternative manner to correct for the influence of *Proboscia* diatoms is  
394 using multiple linear regression of SST and  $FC_{28}$  1,12-diol. However, this did not result in an  
395 improved correlation, possibly because the relative amount of 1,13-diols produced by *Proboscia*  
396 is not a constant variable but is dependent on several environmental factors such as temperature  
397 and nutrient availability.

398         When we exclude core-top data with a fractional abundance of the  $C_{28}$  1,12-diol  $> 0.1$  (24  
399 core-tops), this improves the coefficient of determination to  $R^2 = 0.88$  (Fig. 6).

400 The resulting LDI calibration with annual mean SST is as follows:

$$401 \quad \text{LDI} = 0.00325 \times \text{SST} + 0.1082 \quad (n = 514; R^2 = 0.88; \text{RE} = 3.0 \text{ } ^\circ\text{C}) \quad (4)$$

402 This regression equation is not statistically different from the one originally reported by Rampen  
403 et al. (2012; Eq. 2) (ANCOVA  $p$ -value  $> 0.1$ ; although it should be noted that the data are not  
404 normally distributed).

405

### 406 *3.3 Residual errors in estimation*

407         Although our new correlation between LDI and annual mean SST is strong, there is also  
408 considerable scatter. Indeed, the resulting calibration error, i.e. the standard deviation on the  
409 residual errors, is 3.0 °C, which is higher than the 2.0 °C of the original calibration of Rampen et  
410 al. (2012). The residual errors of the LDI-derived SSTs (LDI SST – annual mean SST) are  
411 between –11.1 and 13.3 °C), without a relationship between the residuals and annual mean SST  
412 (Fig. 6b). This range of residual errors is relative large, potentially limiting the application of the  
413 LDI as an SST proxy. Also, as can be seen in Fig. 6b, the residuals are not randomly distributed,  
414 signifying that the linear regression model does not explain all trends in the dataset. Different

415 statistical models could be more appropriate to define the relationship between the LDI and SST,  
416 however in addition to temperature, other environmental and/or biosynthetic factors may also  
417 control the distribution of the 1,13- and 1,15-diols. LCD sources may have a seasonal  
418 occurrence, thereby registering seasonal instead of annual mean SST, and the LDI-SST  
419 relationship may be affected by regional environmental conditions such as freshwater input,  
420 nutrient conditions, and oxygen concentrations. To assess which water column parameters affect  
421 the LDI, we performed a PCA on the LDI, SST, salinity, phosphate, nitrate and silicate  
422 concentrations, dissolved oxygen, and percent oxygen saturation values from samples from the  
423 final calibration dataset (Supplementary Fig. S4). SST and the LDI show the largest factor  
424 loadings for the first Principal Component which explains 51.8% of the total variance, indicating  
425 that temperature is the most important factor for explaining the variation in the data. Salinity  
426 loads in the same direction as SST and LDI, which may potentially be caused by the global  
427 correlation between salinity and SST. The SST and LDI factor loadings for the second Principal  
428 Component (24.4%) are relatively low and in the same range. Hence, the PCA results provide  
429 support for the idea that SST is the most important parameter affecting the LDI.

430         Nevertheless, this does not exclude the possibilities that in certain regions correlations  
431 between LDI and SST are absent or different. For example, it was shown for the Iberian margin  
432 that these LDI data were likely compromised by river outflow, despite a salinity > 32 ppt (de Bar  
433 et al., 2016), causing a large range of LDI values despite the small range in SST. The cross-  
434 correlation of the LDI with mean seasonal SSTs reveals that the LDI correlates best with summer  
435 temperatures ( $R^2 = 0.90$ ; Supplementary Fig. S3), also providing a more random distribution of  
436 the residuals and a more consistent calibration error over the whole temperature range. The LDI  
437 might be more reflective of seasonal temperatures rather than annual mean temperatures,

438 depending on regional growth seasons of the source organisms which in turn depend on nutrient  
439 and upwelling conditions (e.g., Lattaud et al., 2019). In that case, an increase in scatter of the  
440 calibration is not in the LDI, but in the temperature the LDI is calibrated against. Another issue  
441 might be that the proxy signal is not reflecting surface conditions. However, when correlating the  
442 LDI with annual mean sea temperatures from different water depths, highest coefficients of  
443 determination were observed for temperatures from the upper 30 m of the water column ( $R^2 =$   
444 around 0.88), similar to Rampen et al. (2012). This is also in agreement with Balzano et al.  
445 (2018) who assessed long-chain diol concentrations for different water depths along a  
446 longitudinal transect across the tropical Atlantic, where highest concentrations were observed for  
447 the upper 20–30 m. Moreover, de Bar et al. (2019) calculated LDI temperatures for sediment trap  
448 time series in the tropical Atlantic, the Mozambique Channel and the Cariaco Basin, and for all  
449 three regions the flux-weighted annual mean temperatures agreed well with mean annual surface  
450 temperatures. Thus, whereas seasonality may explain some of the scatter, there are no indications  
451 that different depth habitats play a significant role.

452         To further illustrate that the existence of non-random residuals can partially be explained  
453 by regional differences, the residual errors are plotted in Fig. 7 on a global map. The regions  
454 where the LDI temperatures differ more than  $\pm 2SD$  from the regression line are mainly derived  
455 from the Mediterranean, the Equatorial Pacific and the Iberian margin. The LDI of the surface  
456 sediments in the Mediterranean overestimate annual mean SST (up 6.6°C) while the core-tops of  
457 the equatorial Pacific show severe underestimation (up to  $-11.1^\circ\text{C}$ ). In the Pacific sediments, the  
458 1,14-diols are also relatively high in abundance (30–90%), but the samples do not reveal unusual  
459 LCD distributions, i.e. a high abundance of 1,12-diols. Moreover, the PCA biplot does not show  
460 that the Equatorial Pacific or the Mediterranean data cluster as distinct groups separated from the

461 other data (Fig. 3). The reasons for the cold and warm bias in these regions remain unknown, but  
462 may be due to differences in regional conditions. For instance, in the Equatorial Pacific, the LDI  
463 temperature signal might be related to the seasonal upwelling of cold waters. If the producers of  
464 1,13 and 1,15 LCDs mainly thrive under high-nutrient conditions induced by upwelling, or  
465 during stagnant conditions, then the LDI will reflect temperatures during these times. In general,  
466 if the source organisms are seasonally blooming, then the LDI will likely reflect a seasonal rather  
467 than annual mean SST.

468 Another uncertainty is the age of the core-top material extracted. Although many  
469 sediments represent the upper 0.5 to 1 cm (with a few comprising the upper 2 cm), they can  
470 represent different ages, as sedimentation rates can differ significantly per location, resulting in  
471 different time lengths reflected by the sediment. A more careful assessment of the age  
472 represented by each surface sediment can potentially improve the calibration of the LDI.  
473 The LDI data at the upper end of the calibration ( $>16^{\circ}\text{C}$ ) showed a reduction in slope. However,  
474 application of a third order polynomial equation results only in a moderate increase in coefficient  
475 of determination to 0.90 (data not shown). Therefore, we propose to use equation 4 as the new  
476 calibration of LDI to SST, although we realize that the linear regression model does not  
477 sufficiently explain all data and other models may be needed. More regional and time-series  
478 studies are needed to assess regional influences on the LDI, and future research should also focus  
479 on the identification of the LDI producers. This could lead to a better understanding of the  
480 mechanisms behind the relationship between the LDI and temperature as well as the effect of  
481 differences in the source organisms.

482

483 3.4 Implications for LDI temperature reconstructions

484 Our new extended LDI calibration is statistically similar to the original calibration  
485 proposed by Rampen et al. (2012), suggesting that previous temperature reconstructions based on  
486 the LDI (e.g., Lopes dos Santos et al., 2013; Warnock et al., 2017; Jonas et al., 2017) likely do  
487 not require major adjustments. The maximum temperature difference between the previous SST  
488 calibration and the present calibration is +0.45°C at the lower end of the calibration (LDI=0) and  
489 -0.02°C at the upper end (LDI=1). However, due to the substantial increase in data-points  
490 included in the calibration, the residual error increased from 2 °C to 3 °C, which is larger than  
491 that of the  $U_{37}^{K'}$  (1.5 °C) and the  $TEX_{86}^H$  (2.5 °C). Possibly, differences between proxy values  
492 observed in sediment records may now fall within proxy errors.

493 Importantly, our results provide new constraints on the application of the LDI. Firstly, the  
494 LDI should not be applied in low-salinity environments and environments substantially  
495 influenced by river runoff. These types of conditions can, to some degree, be assessed for past  
496 environments by organic proxies (e.g. the BIT index, the  $\delta D$  of alkenones, dinocyst assemblages)  
497 or inorganic proxies (e.g. mineral composition). Secondly, high abundances of the C<sub>26</sub> and C<sub>28</sub>  
498 1,12-diols and C<sub>27</sub> and C<sub>29</sub> 12-hydroxy and C<sub>28</sub> and C<sub>30</sub> 13-hydroxy methyl alkananoates hint at  
499 LCD contributions from *Proboscia* spp., and we advise against using LDI data when the  
500 fractional abundance of the C<sub>28</sub> 1,12-diol (vs C<sub>28</sub> 1,12-, 1,13-, 1,14-, C<sub>30</sub> 1,13-, 1,14- and 1,15-  
501 diols) is > 0.1. This LCD has probably been ignored in most of the previous studies on long-  
502 chain diols, but our data show it is useful to include it in future studies to assess the potential  
503 influence of *Proboscia* LCD contribution on the LDI. Quantification of the C<sub>28</sub> 1,12-diol does  
504 not require any modification of the SIM analysis method used for the standard 1,13-, 1,14- and  
505 1,15-diols as it will be detected by the  $m/z$  327 ion. Furthermore, caution is advised when



506 applying the LDI in case of high and/or variable 1,14-diol abundances (e.g., Equatorial Pacific).  
507 Since previous LDI records have not reported the abundance of C<sub>28</sub> 1,12-diols we cannot  
508 evaluate whether these were compromised by *Proboscia*-derived LCDs. However, de Bar et al.  
509 (2019) calculated the LDI for the ages of ~11, 18, 33, 41 and 50 Ma for the Bass River core  
510 (New Jersey, USA) and observed that the LDI-derived temperatures did not agree with other  
511 paleotemperature records for this core, with LDI temperatures being between 2 to 14 °C lower as  
512 compared to the *TEX*<sub>86</sub><sup>H</sup>-derived SSTs. Interestingly, the *FC*<sub>28</sub> 1,12-diol varied between 0.2 and 1  
513 for the ages of 18, 33, 41 and 50 Ma, potentially suggesting a *Proboscia* influence on the LDI at  
514 this location.

515

## 516 **Conclusions**

517 We have extended the global core-top temperature calibration of the Long chain Diol  
518 Index and confirmed that LDI values are strongly correlated with annual mean SST, but with a  
519 considerable increase in scatter and a decrease in the coefficient of determination ( $R^2$  reduced  
520 from 0.97 to 0.82). We observed that most surface sediments with low salinities, in particular  
521 sediments from the Baltic Sea and Black Sea, overestimate LDI-derived temperatures, and  
522 exclusion of these sediments (salinity < 32 ppt) improved the coefficient of determination ( $R^2 =$   
523 0.86). The fractional abundance of the C<sub>32</sub> 1,15-diol is not a consistent indicator for freshwater  
524 influence on the LDI, since its ability to trace riverine input is likely region-specific.  
525 Examination of diol distributions with high input of 1,14-diols from *Proboscia* diatoms shows  
526 that the Diol Indices are not good indicators for identifying biases in the LDI, but that high  
527 abundances of the C<sub>26</sub> and C<sub>28</sub> 1,12-diols, as well as the C<sub>27</sub> and C<sub>29</sub> 12-hydroxy and C<sub>28</sub> and C<sub>30</sub>  
528 13-hydroxy methyl alkanates, are associated with a cold bias in the LDI-based temperatures.

529 Therefore, we have defined a cut-off of 0.1 in the C<sub>28</sub> 1,12-diol fractional abundance, which  
530 further improved the correlation between the LDI and annual mean SST. The new calibration  
531 between the LDI and SST ( $LDI=0.0325\times SST+0.1082$ ;  $n = 514$ ;  $R^2 = 0.88$ ) covers a temperature  
532 range between  $-3.3$  and  $27.4^{\circ}C$  with a calibration error of  $3^{\circ}C$ . However, the linear regression  
533 model does not sufficiently explain all of our data, due to the presence of non-random residuals  
534 in our dataset. More research is needed to constrain local and seasonal influences on the LDI, as  
535 well as to identify the source organism, in order to reduce the calibration uncertainty. The  
536 relationship is statistically similar to the Rampen et al. (2012) calibration, and thus supports  
537 previous down-core LDI applications. Our results confirm that the LDI can be used as a proxy  
538 for the reconstruction of annual mean SST in marine sediment cores, but with caveats, i.e. it  
539 should not be applied in low-salinity/freshwater influenced regions, or when the fractional  
540 abundance of the C<sub>28</sub> 1,12-diol (vs C<sub>28</sub> 1,12-, 1,13-, 1,14-, C<sub>30</sub> 1,13-, 1,14- and 1,15-diols) is  $>$   
541 0.1. Accordingly, re-evaluation of the reliability of the LDI records in terms of freshwater  
542 influence (salinity, C<sub>32</sub> 1,15-diol abundance) and *Proboscia* contribution (high/variable 1,14-diol  
543 abundances, C<sub>28</sub> 1,12-diol abundance) is recommended. Finally, in some regions there seems to  
544 be no, or a weak relation between the LDI and annual mean SST, for reasons which are presently  
545 unclear, thereby limiting the application of the LDI.

546

#### 547 **Acknowledgements**

548 We thank Dr. Andy Revil and John Volkman, two anonymous reviewers and Liz Sikes  
549 for useful comments which improved the manuscript. We are grateful to various people who  
550 have worked-up or provided core-top sediments: Tjerk Veenstra, and Steven D'Hondt for Pacific  
551 core-tops recovered during the R/V Knorr expedition 195-3 (US National Science Foundation

552 grant OCE-0752336), Isla Castañeda (Mozambique Channel), Ivan Tomberg and Roselyne  
553 Buscail (Gulf of Lion), Claudia Zell and David Hollander (Amazon Basin), Kees Booij (Berau  
554 Delta), Cindy de Jonge, Alina Stadnitskaia and Georgy Cherkashov (Kara Sea), Li Lo (Okhotsk  
555 Sea), Marcel van der Meer (Mediterranean), Laura Villanueva (Black Sea) and Zeynep Erdem  
556 (Chilean margin). We thank Allert Bijleveld for statistical advice. This research has been funded  
557 by the European Research Council (ERC) under the European 555 Union's Seventh Framework  
558 Program (FP7/2007-2013) ERC grant agreement [339206] to S.S. S.S. and J.S.S.D. receive  
559 funding from the Netherlands Earth System Science Center (NESSC) through a Gravitation grant  
560 from the Dutch ministry for Education, Culture and Science (grant number 024.002.001).

561

## 562 **Appendices**

563 Supplementary data associated with this article can be found in the online version, at ..., as well  
564 as in the PANGAEA database (DOI: ...).

565

## 566 **References**

567 Anagnostou, E., John, E.H., Edgar, K.M., Foster, G.L., Ridgwell, A., Inglis, G.N., Pancost, R.D.,  
568 Lunt, D.J., Pearson, P.N., 2016. Changing atmospheric CO<sub>2</sub> concentration was the  
569 primary driver of early Cenozoic climate. *Nature* 533, 380.

570 Balzano, S., Lattaud, J., Villanueva, L., Rampen, S.W., Brussaard, C.P.D., van Bleijswijk, J.,  
571 Bale, N.J., Sinninghe Damsté, J.S., Schouten, S., 2018. A quest for the biological sources  
572 of long chain alkyl diols in the western tropical North Atlantic Ocean. *Biogeosciences* 15,  
573 5951–5968.

574 Bijl, P.K., Schouten, S., Sluijs, A., Reichart, G.-J., Zachos, J.C., Brinkhuis, H., 2009. Early  
575 Palaeogene temperature evolution of the southwest Pacific Ocean. *Nature* 461, 776–779.

576 Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U., Sarnthein, M., 1986. Molecular  
577 stratigraphy: A new tool for climatic assessment. *Nature* 320, 129–133.

578 Conte, M.H., Sicre, M.A., Rühlemann, C., Weber, J.C., Schulte, S., Schulz-Bull, D., Blanz, D.,  
579 2006. Global temperature calibration of the alkenone unsaturation index ( $U_{37}^{K'}$  in surface  
580 waters and comparison with surface sediments. *Geochemistry, Geophysics, Geosystems*  
581 7, 1–22.

582 Conte, M.H., Thompson, A., Eglinton, G., 1995. Lipid biomarker diversity in the  
583 coccolithophorid *Emiliana huxleyi* (Prymnesiophyceae) and the related species  
584 *Gephyrocapsa oceanica*. *Journal of Phycology* 31, 272–282.

585 Conte, M.H., Weber, J.C., King, L.L., Wakeham, S.G., 2001. The alkenone temperature signal in  
586 western North Atlantic surface waters. *Geochimica et Cosmochimica Acta* 65, 4275–  
587 4287.

588 de Bar, M.W., Dorhout, D.J.C., Hopmans, E.C., Rampen, S.W., Sinninghe Damsté, J.S.,  
589 Schouten, S., 2016. Constraints on the application of long chain diol proxies in the  
590 Iberian Atlantic margin. *Organic Geochemistry* 101, 184–195.

591 de Bar, M.W., Rampen, S.W., Hopmans, E.C., Sinninghe Damsté, J.S., Schouten, S., 2019.  
592 Constraining the applicability of organic paleotemperature proxies for the last 90 Myrs.  
593 *Organic Geochemistry* 128, 122–136.

594 de Bar, M.W., Stolwijk, D.J., McManus, J.F., Sinninghe Damsté, J.S., Schouten, S., 2018. A  
595 Late Quaternary climate record based on long-chain diol proxies from the Chilean  
596 margin. *Climate of the Past* 14, 1783–1803.

597 de Leeuw, J.W., Rijpstra, W.I.C., Schenck, P.A., 1981. The occurrence and identification of C<sub>30</sub>,  
598 C<sub>31</sub> and C<sub>32</sub> alkan-1,15-diols and alkan-15-one 1-ols in Unit I and Unit II Black Sea  
599 sediments. *Geochimica et Cosmochimica Acta* 45 2281–2285.

600 Emiliani, C., 1955. Pleistocene temperatures. *The Journal of Geology* 63, 538-578.

601 Foster, G.L., Royer, D.L., Lunt, D.J., 2017. Future climate forcing potentially without precedent  
602 in the last 420 million years. *Nature Communications* 8. DOI: 10.1038/ncomms14845  
603 (2017).

604 Gal, J.-K., Kim, J.-H., Shin, K.-H., 2018. Distribution of long chain alkyl diols along a south-  
605 north transect of the northwestern Pacific region: Insights into a paleo sea surface nutrient  
606 proxy. *Organic Geochemistry* 119, 80–90.

607 Garcia, H.E., R.A. Locarnini, T.P. Boyer, J.I. Antonov, O.K. Baranova, M.M. Zweng, J.R.  
608 Reagan, D.R. Johnson, 2014a. World Ocean Atlas 2013, Volume 3: Dissolved Oxygen,  
609 Apparent Oxygen Utilization, and Oxygen Saturation. In: Levitus, S and Mishonov, A.  
610 (Eds.), NOAA Atlas NESDIS 75, 27 pp.

611 Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M.,  
612 Reagan, D.R., Johnson, 2014b. World Ocean Atlas 2013. Vol. 4: Dissolved Inorganic  
613 Nutrients (phosphate, nitrate, silicate). In: Levitus, S and Mishonov, A. (Eds.), NOAA  
614 Atlas NESDIS 76, 25 pp.

615 Gelin, F., Boogers, I., Noordeloos, A.A.M., Sinninghe Damsté, J.S., Riegman, R., de Leeuw,  
616 J.W., 1997. Resistant biomacromolecules in marine microalgae of the classes  
617 Eustigmatophyceae and Chlorophyceae: Geochemical implications. *Organic*  
618 *Geochemistry* 26, 659–675.

619 IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and  
620 III to the In: Core Writing Team, Pachauri, R.K., Meyer, L.A. (Eds.), Fifth Assessment  
621 Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland,  
622 151 pp.

623 Jonas, A.S., Schwark, L., Bauersachs, T., 2017. Late Quaternary water temperature variations of  
624 the Northwest Pacific based on the lipid paleothermometers  $TEX_{86}^H$ ,  $U_{37}^{K'}$  and LDI. Deep  
625 Sea Research Part 1 Oceanographic Research Papers 125, 81–93.

626 Kim, J.-H., Schouten, S., Hopmans, E.C., Donner, B., Sinninghe Damsté, J.S., 2008. Global  
627 sediment core-top calibration of the  $TEX_{86}$  paleothermometer in the ocean. *Geochimica  
628 et Cosmochimica Acta* 72, 1154–1173.

629 Kim, J.-H., van der Meer, J., Schouten, S., Helmke, P., Willmott, V., Sangiorgi, F., Koç, N.,  
630 Hopmans, E.C., Sinninghe Damsté, J.S., 2010. New indices and calibrations derived from  
631 the distribution of crenarchaeal isoprenoid tetraether lipids: Implications for past sea  
632 surface temperature reconstructions. *Geochimica et Cosmochimica Acta* 74, 4639–4654.

633 Lattaud, J., Dorhout, D., Schulz, H., Castañeda, I.S., Schefuß, E., Sinninghe Damsté, J.S.,  
634 Schouten, S., 2017b. The  $C_{32}$  alkane-1,15-diol as a proxy of late Quaternary riverine  
635 input in coastal margins. *Climate of the Past* 13, 1049–1061.

636 Lattaud, J., Kirkels, F., Peterse, F., Freymond, C.V., Eglinton, T.I., Hefter, J., Mollenhauer, G.,  
637 Balzano, S., Villanueva, L., van der Meer, M.T.J., Hopmans, E.C., Sinninghe Damsté,  
638 J.S., Schouten, S., 2018a. Long-chain diols in rivers: distribution and potential biological  
639 sources. *Biogeosciences* 15, 4147–4161.

640 Lattaud, J., Kim, J.-H., de Jonge, C., Zell, C., Sinninghe Damsté, J.S., Schouten, S., 2017a. The  
641 C<sub>32</sub> alkane-1,15-diol as a tracer for riverine input in coastal seas. *Geochimica et*  
642 *Cosmochimica Acta* 202, 146–158.

643 Lattaud, J., Lo, L., Huang, J.-J., Chou, Y.-M., Gorbarenko, S.A., Sinninghe Damsté, J.S.,  
644 Schouten, S., 2018b. A comparison of Late Quaternary organic proxy-based  
645 paleotemperature records of the Central Sea of Okhotsk. *Paleoceanography and*  
646 *Paleoclimatology* 33, 732–744.

647 Locarnini R.A., Mishonov A.V., Antonov J.I., Boyer T.P., Garcia H.E., Baranova O.K., Zweng,  
648 M.M., Paver, C.R., Reagan, J.R., Johnson, D.R., Hamilton, M., Seidov, D., 2013. World  
649 Ocean Atlas 2013, Volume 1: temperature. In: Levitus, S and Mishonov, A. (Eds.),  
650 NOAA Atlas NESDIS 73, 40 pp.

651 Lopes dos Santos, R.A., Spooner, M.I., Barrows, T.T., de Deckker, P., Sinninghe Damsté, J.S.,  
652 Schouten, S., 2013. Comparison of organic ( $U_{37}^{KI}$ ,  $TEX_{86}^H$ , LDI) and faunal proxies  
653 (foraminiferal assemblages) for reconstruction of late Quaternary sea surface temperature  
654 variability from offshore southeastern Australia. *Paleoceanography* 28, 377–387.

655 Marlowe, I.T., Green, J.C., Neal, A.C., Brassell, S.C., Eglinton, G., Course, P.A., 1984. Long-  
656 chain ( $n$ -C<sub>37</sub>–C<sub>39</sub>) alkenones in the Prymnesiophyceae. Distribution of alkenones and  
657 other lipids and their taxonomic significance. *British Phycological Journal* 19, 203–216.

658 Méjanelle, L., Sanchez-Gargallo, A., Bentaleb, I., Grimalt, J.O., 2003. Long chain  $n$ -alkyl diols,  
659 hydroxy ketones and sterols in a marine eustigmatophyte, *Nannochloropsis gaditana*, and  
660 in *Brachionus plicatilis* feeding on the algae. *Organic Geochemistry* 34, 527–538.

661 Müller, P.J., Kirst, G., Ruhland, G., von Storch, I., Rosell-Melé, A., 1998. Calibration of the  
662 alkenone paleotemperature index  $U_{37}^{K'}$  based on core-tops from the eastern South Atlantic  
663 and the global ocean (60°N-60°S). *Geochimica et Cosmochimica Acta* 62, 1757–1772.

664 Naafs, B.D.A., Inglis, G.N., Zheng, Y., Amesbury, M.J., Biester, H., Bindler, R., Blewett, J.,  
665 Burrows, M.A., del Castell Torres, D., Chambers, F.M., Cohen, A.D., Evershed, R.P.,  
666 Feakins, S.J., Gałka, M., Gallego-Sala, A., Gandois, L., Gray, D.M., Hatchr, P.G.,  
667 Honorio Coronado, E.N., Hughes, P.D.M., Hugué, A., Könönen, M., Laggoun-Défarge,  
668 F., Lähteenoja, O., Lamentowicz, M., Marchant, R., McClymont, E., Pontevedra-Pombal,  
669 X., Ponton, C., Poumand, A., Rizzuti, A.M., Rochefort, L., Schellekens, J., de  
670 Vleeschouwer, F., Pancost, R.D., 2017. Introducing global peat-specific temperature and  
671 pH calibrations based on brGDGT bacterial lipids. *Geochimica et Cosmochimica Acta*  
672 208, 285–301.

673 Nürnberg, D., Bijma, J., Hemleben, C., 1996. Assessing the reliability of magnesium in  
674 foraminiferal calcite as a proxy for water mass temperatures. *Geochimica et*  
675 *Cosmochimica Acta* 60, 803–814.

676 O'Brien, C.L., Robinson, S.A., Pancost, R.D., Sinninghe Damsté, J.S., Schouten, S., Lunt,  
677 Alsenz, H., Bornemann, A., Bottini, C., Brassell, S.C., Farnsworth, A., Forster, A.,  
678 Huber, B.T., Inglis, G.N., Jenkyns, H.C., Linnert, C., Littler, K., Markwick, P.,  
679 McAnena, A., Mutterlose, J., Naafs, B.D.A., Püttmann, W., Sluijs, A., van Helmond,  
680 N.A.G.M., Vellekoop, J., Wagner, T., Wrobel, N.E., 2017. Cretaceous sea-surface  
681 temperature evolution: Constraints from TEX<sub>86</sub> and planktonic foraminiferal oxygen  
682 isotopes. *Earth-Science Reviews* 172, 224–247.



683 Prah, F.G., Wakeham, S.G., 1987. Calibration of unsaturation patterns in long-chain ketone  
684 compositions for palaeotemperature assessment. *Nature* 330, 367–369.

685 Rampen, S.W., Schouten, S., Wakeham, S.G., Sinninghe Damsté, J.S., 2007. Seasonal and  
686 spatial variation in the sources and fluxes of long chain diols and mid-chain hydroxy  
687 methyl alkanoates in the Arabian Sea. *Organic Geochemistry* 38, 165–179.

688 Rampen, S.W., Schouten, S., Koning, E., Brummer, G.J.A., Sinninghe Damsté, J.S., 2008. A 90  
689 kyr upwelling record from the northwestern Indian Ocean using a novel long-chain diol  
690 index. *Annual Review of Earth and Planetary Sciences* 276, 207–213.

691 Rampen, S.W., Schouten, S., Schefuss, E., Sinninghe Damsté, J.S., 2009. Impact of temperature  
692 on long chain diol and mid-chain hydroxy methyl alkanoate composition in *Proboscia*  
693 diatoms: Results from culture and field studies. *Organic Geochemistry* 40, 1124–1131.

694 Rampen, S.W., Schouten, S., Sinninghe Damsté, J.S., 2011. Occurrence of long chain 1,14-diols  
695 in *Apedinella radians*. *Organic Geochemistry* 42, 572–574.

696 Rampen, S.W., Willmott, V., Kim, J.-H., Uliana, E., Mollenhauer, G., Schefuss, E., Sinninghe  
697 Damsté, J.S., Schouten, S., 2012. Long chain 1,13-and 1,15-diols as a potential proxy for  
698 palaeotemperature reconstruction. *Geochimica et Cosmochimica Acta* 84, 204–216.

699 Rampen, S.W., Willmott, V., Kim, J.H., Rodrigo-Gámiz, M., Uliana, E., Mollenhauer, G.,  
700 Schefuss, E., Sinninghe Damsté, J.S., Schouten, S., 2014a. Evaluation of long chain 1,14-  
701 alkyl diols in marine sediments as indicators for upwelling and temperature. *Organic*  
702 *Geochemistry* 76, 39–47.

703 Rampen, S.W., Datema, M., Rodrigo-Gámiz, M., Schouten, S., Reichert, G.J., Sinninghe  
704 Damsté, J.S., 2014b. Sources and proxy potential of long chain alkyl diols in lacustrine  
705 environments. *Geochimica et Cosmochimica Acta* 144, 59–71.

706 Rodrigo-Gámiz, M., Martinez-Ruiz, F., Rampen, S., Schouten, S., Sinninghe Damsté, J., 2014.  
707 Sea surface temperature variations in the western Mediterranean Sea over the last 20 kyr:  
708 A dual- organic proxy ( $U_{37}^{Kl}$  and LDI) approach. *Paleoceanography* 29, 87–98.

709 Rodrigo-Gámiz, M., Rampen, S.W., de Haas, H., Baas, M., Schouten, S., Sinninghe Damsté,  
710 J.S., 2015. Constraints on the applicability of the organic temperature proxies  $U_{37}^{Kl}$ ,  $TEX_{86}$   
711 and LDI in the subpolar region around Iceland. *Biogeosciences* 12, 6573–6590.

712 Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S., 2013. The organic geochemistry of  
713 glycerol dialkyl glycerol tetraether lipids: A review. *Organic Geochemistry* 54 19–61.

714 Schouten, S., Hopmans, E.C., Schefuß, E., Sinninghe Damsté, J.S., 2002. Distributional  
715 variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing  
716 ancient sea water temperatures? *Annual Review of Earth and Planetary Science Letters*  
717 204, 265–274.

718 Shackleton, N., 1967. Oxygen isotope analyses and Pleistocene temperatures re-assessed. *Nature*  
719 215, 15.

720 Shimokawara, M., Nishimura, M., Matsuda, T., Akiyama, N., Kawai, T., 2010. Bound forms,  
721 compositional features, major sources and diagenesis of long chain, alkyl mid-chain diols  
722 in Lake Baikal sediments over the past 28,000 years. *Organic Geochemistry* 41, 753–766.

723 Sinninghe Damsté, J.S., Rampen, S., Rijpstra, W.I.C., Abbas, B., Muyzer, G., Schouten, S.,  
724 2003. A diatomaceous origin for long-chain diols and mid-chain hydroxy methyl  
725 alkanooates widely occurring in Quaternary marine sediments: Indicators for high-nutrient  
726 conditions. *Geochimica et Cosmochimica Acta* 67, 1339–1348.

727 Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Prahl, F.G., Wakeham, S.G., Schouten,  
728 S., 2002. Distribution of membrane lipids of planktonic Crenarchaeota in the Arabian  
729 Sea. *Applied and Environmental Microbiology* 68, 2997–3002.

730 Smith, M., de Deckker, P., Rogers, J., Brocks, J., Hope, J., Schmidt, S., Lopes dos Santos, R.,  
731 Schouten, S., 2013. Comparison of  $U_{37}^{K'}$ ,  $TEX_{86}^H$ , and LDI temperature proxies for  
732 reconstruction of south-east Australian ocean temperatures. *Organic Geochemistry* 64,  
733 94–104.

734 ten Haven, H.L., Rullkötter, J., 1991. Preliminary lipid analyses of sediments recovered during  
735 leg 117 1., *Proceedings of the Ocean Drilling Program, Scientific Results*, pp. 561–569.

736 Tierney, J.E., Tingley, M.P., 2018. BAYSPLINE: A new calibration for the alkenone  
737 paleothermometer. *Paleoceanography and Paleoclimatology* 33 281–301.

738 Versteegh, G.J.M., Bosch, H.J., de Leeuw, J.W., 1997. Potential palaeoenvironmental  
739 information of C<sub>24</sub> to C<sub>36</sub> mid-chain diols, keto-ols and mid-chain hydroxy fatty acids; a  
740 critical review. *Organic Geochemistry* 27, 1–13.

741 Versteegh, G.J.M., Jansen, J.H.F., de Leeuw, J.W., Schneider, R.R., 2000. Mid-chain diols and  
742 keto-ols in SE Atlantic sediments: A new tool for tracing past sea surface water masses?  
743 *Geochimica et Cosmochimica Acta* 64, 1879–1892.

744 Villanueva, L., Besseling, M., Rodrigo-Gámiz, M., Rampen, S.W., Verschuren, D., Sinninghe  
745 Damsté, J.S., 2014. Potential biological sources of long chain alkyl diols in a lacustrine  
746 system. *Organic Geochemistry* 68, 27–30.

747 Volkman, J.K., Barrett, S.M., Blackburn, S.I., 1999. Eustigmatophyte microalgae are potential  
748 sources of C<sub>29</sub> sterols, C<sub>22</sub>–C<sub>28n</sub>-alcohols and C<sub>28</sub>–C<sub>32n</sub>-alkyl diols in freshwater  
749 environments. *Organic Geochemistry* 30, 307–318.

750 Volkman, J.K., Barrett, S.M., Dunstan, G.A., Jeffrey, S.W., 1992. C<sub>30</sub>–C<sub>32</sub> alkyl diols and  
751 unsaturated alcohols in microalgae of the class Eustigmatophyceae. *Organic*  
752 *Geochemistry* 18, 131–138.

753 Volkman, J.K., Eglinton, G., Corner, E.D.S., Forsberg, T.E.V., 1980. Long-chain alkenes and  
754 alkenones in the marine coccolithophorid *Emiliana huxleyi*. *Phytochemistry* 19, 2619–  
755 2622.

756 Warnock, J.P., Bauersachs, T., Kotthoff, U., Brandt, H.T., Andren, E., 2018. Holocene  
757 environmental history of the Angermanalven Estuary, northern Baltic Sea. *Boreas* 47,  
758 593–608.

759 Willmott, V., Rampen, S.W., Domack, E., Canals, M., Sinninghe Damsté, J.S., Schouten, S.,  
760 2010. Holocene changes in *Proboscia* diatom productivity in shelf waters of the north-  
761 western Antarctic Peninsula. *Antarctic Science* 22, 3–10.

762 Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and  
763 aberrations in global climate 65 Ma to present. *Science* 292, 686–693.

764 Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P.,  
765 Garcia, H.E., Baranova, O.K., Johnson, D.R., Seidov, D., Biddle, M.M., 2013. World  
766 Ocean Atlas 2013, Volume 2: Salinity. In: Levitus, S and Mishonov, A. (Eds.), NOAA  
767 Atlas NESDIS 74, 39 pp.

768

769 **Figures**

770

771 **Fig. 1.** Core-top sediment locations investigated in this study. Black circles indicate the surface  
772 sediments of Rampen et al. (2012), pink circles indicate surface sediments analyzed for long  
773 chain alkyl diols in previous studies (de Bar et al., 2016; Lattaud et al., 2017a,b, 2018; Smith et  
774 al., 2013; Rodrigo-Gámiz et al., 2015), and the white circles represent new surface sediment data  
775 obtained in this study. The map of sea surface temperatures (SST) is based on data of the World  
776 Ocean Atlas (2013) (Locarnini et al., 2013) and made in Ocean Data View (Schlitzer, 2015).

777

778 **Fig. 2.** (a) LDI values of all surface sediments vs annual mean SST (0 m depth; WOA13), and  
779 (c) LDI values of marine surface sediments vs annual mean SST when excluding all stations  
780 where surface salinity < 32 ppt. The associated excluded sediments belong to the Hudson Bay,  
781 Black Sea, Gulf of St. Lawrence, Kara Sea (7 of 8 sediments excluded) and the Baltic Sea  
782 (indicated in the legend by strike-through). Panels (b) and (d) show the residual errors associated  
783 with the regression of (a) and (c), respectively. In panel (b), the low-salinity core-tops are  
784 highlighted which were excluded for the calibration plotted in panel (c). In panel (d), core-tops  
785 associated with *Proboscia* long-chain diol contribution are highlighted, which were excluded  
786 from the final LDI-SST calibration (Fig. 6) as discussed in Section 3.2.

787

788 **Fig. 3.** (a) PCA biplot of the long-chain diol proxies, i.e., the LDI and C<sub>32</sub> 1,15-diol fractional  
789 abundance, and environmental parameters, i.e., annual mean sea surface temperature (AM SST;  
790 °C), salinity, and annual mean phosphate and nitrate concentrations ( $\mu\text{mol L}^{-1}$ ); (b) PCA biplot of  
791 the fractional abundances of the different long-chain diols. Panel (c) and (e) show the Factor 1

792 scores of the biplots of panels (a) and (b), respectively, correlated against annual mean SST (°C).  
793 Panels (d) and (f) show the Factor 2 scores correlated against the salinity. In cases where the  
794 fractional abundance of one or more diols was unknown for a core-top, there was no LDI value,  
795 or environmental data (World Ocean Atlas 2013) were not available, this core-top was excluded  
796 from the PCA.

797  
798 **Fig. 4.** Selected ion monitoring chromatograms of two sediments (two upper panels) with ‘usual’  
799 long-chain diol distributions, i.e., with low 1,12-diols. The lower four panels show the  
800 chromatograms of four sediments with relatively high C<sub>26</sub> and C<sub>28</sub> 1,12-diols (blue), as well as  
801 the C<sub>27</sub> and C<sub>29</sub> 12-OH (green) and C<sub>28</sub> and C<sub>30</sub> 13-OH methyl alkananoates (brown). These  
802 sediments reveal LDI values which deviated substantially from the LDI calibration vs annual  
803 mean SST (see Fig. 2). The ‘Greenland’ sediment is classified as ‘North Atlantic Ocean’ in Figs.  
804 2, 5 and 6, and the “Pakistan margin” as “Arabian Sea”.

805  
806 **Fig. 5.** The fractional abundance of the C<sub>28</sub> 1,12-diol vs the residual error in SST estimation  
807 based on the regression plotted in Fig. 2c. The black solid lines indicate a fractional abundance  
808 of 0.1, and a residual error of 0. The dotted lines reflect the negative and positive 2×SD value of  
809 6.8 °C.

810  
811 **Fig. 6.** (a)LDI calibration after exclusion of estuarine sediments (salinity < 32 ppt), as well as  
812 surface sediments in which the fractional abundance of the C<sub>28</sub> 1,12-diol is > 0.1. Regions in  
813 which surface sediments were excluded based on the C<sub>28</sub> 1,12-diol abundance and salinity are  
814 indicated by underline and strike-through in the legend, respectively; (b)Residual SST errors

815 (LDI SST – annual mean SST) against annual mean SST. The black dashed line reflects a  
816 residual error of 0 °C, and the grey dashed lines reflect the standard deviation of the residual  
817 errors, i.e., the calibration error (3.0 °C).

818

819 **Fig. 7.** Residual SST errors (LDI SST – measured annual mean SST) plotted on the global map  
820 (created in Ocean Data View; Schlitzer, 2015).