PERGAMON

Quaternary Science Reviews 21 (2002) 455–483

QSR

Apparent long-term cooling of the sea surface in the northeast Atlantic and Mediterranean during the Holocene

Olivier Marchal[a,*], Isabel Cacho[b], Thomas F. Stocker[a], Joan O. Grimalt[c], Eva Calvo[c], Belen Martrat[c], Nicholas Shackleton[b], Maryline Vautravers[b], Elsa Cortijo[d], Shirley van Kreveld[e], Carin Andersson[f], Nalan Koç[g], Mark Chapman[h], Laura Sbaffi[i], Jean-Claude Duplessy[d], Michael Sarnthein[e], Jean-Louis Turon[j], Josette Duprat[j], Eystein Jansen[f]

[a] *Climate and Environmental Physics, Physics Institute, University of Bern, Germany*
[b] *Department of Earth Sciences, Godwin Laboratory, University of Cambridge, UK*
[c] *Department of Environmental Chemistry, Institute of Chemical and Environmental Research, Barcelona, Spain*
[d] *Laboratoire des Sciences du Climat et de l'Environnement, Laboratoire mixte CNRS-CEA, Gif-Sur-Yvette, France*
[e] *Institut für Geowissenschaften, Christian Albrechts Universität, Kiel, Germany*
[f] *Bjerknes Center for Climate Research and Department of Geology, University of Bergen, Norway*
[g] *Norwegian Polar Institute, Tromsø, Norway*
[h] *School of Environmental Sciences, University of East Anglia, UK*
[i] *Istituto di Dinamica Ambientale, University of Urbino, Spain*
[j] *Département de Géologie et Océanographie, Université de Bordeaux, France*

Received 13 March 2001; accepted 5 September 2001

Abstract

Reconstructions of upper ocean temperature (T) during the Holocene (10–0 ka B.P.) were established using the alkenone method from seven, high accumulation sediment cores raised from the northeast Atlantic and the Mediterranean Sea (36°N–75°N). All these paleo-T records document an apparent long-term cooling during the last 10 kyr. In records with indication of a constant trend, the apparent cooling ranges from -0.27 to $-0.15\,^{\circ}\mathrm{C\,kyr^{-1}}$. Records with indication of time-variable trend show peak-to-peak amplitudes in apparent temperatures of 1.2–2.9°C. A principal component analysis shows that there is one factor which accounts for a very large fraction (67%) of the total variance in the biomarker paleo-T records and which dominates these records over other potential secondary influences. Two possible contributions are (1) a widespread surface cooling, which may be associated with the transition from the Hypsithermal interval (~9–5.7 ka B.P.) to the Neoglaciation (~5.7–0 ka B.P.); and (2) a change in the seasonal timing and/or duration of the growth period of alkenone producers (prymnesiophyte algae). The first contribution is consistent with many climate proxy records from the northeast Atlantic area and with climate model simulations including Milankovitch forcing. The second contribution is consistent with the divergence between biomarker and summer faunal paleo-T from early to late Holocene observed in two cores. Further work is necessary, and in particular the apparent discordance between biomarker and faunal T records for the relative stable Holocene period must be understood, to better constrain the climatic and ecological contributions to the apparent cooling observed in the former records.

1. Introduction

The Holocene, i.e. the current warm interglaciation (here from 10–0 kyr before present [AD 1950], ka B.P.), appears as a relatively stable climatic period when viewed in a long-term perspective. The paradigm of climatic stability during the Holocene stems mainly from isotopic records from Greenland ice cores, documenting large temperature changes during, e.g., marine isotope stage 3 and small temperature changes after the last deglaciation (Dansgaard et al., 1993; Grootes and Stuiver, 1997). More recently, a long, continuous sediment record from the subpolar North Atlantic suggested that the Holocene is one of the few

*Corresponding author. Now at Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Department of Geophysics, Woods Hole, MA 02543, USA. Tel.: +1-508-289-3374; fax: +1-508-457-2187.

E-mail address: omarchal@whoi.edu (O. Marchal).

0277-3791/02/$ - see front matter
PII: S0277-3791(01)00105-6

periods of the last ~0.5 million years characterized by a stable climate (McManus et al., 1999). Other climate proxy records, however, suggest that significant climate shifts occurred within the present interglaciation (for an early study see Denton and Karlen (1973)). The Holocene climate can be subdivided into three main phases (e.g., Nesje and Dahl (1993)). The first is a climatic amelioration, comprising the Preboreal and Boreal chronozones. The conventional radiocarbon ages of the onset of the Preboreal and end of Boreal are ~10 ^{14}C and ~8 ^{14}C ka B.P. (Mangerud et al., 1974), corresponding to calendar ages of ~11.6 and ~9 ka B.P., respectively (Stuiver et al., 1998). The second phase is the Hypsithermal interval, a relatively warm period including the Atlantic chronozone between ~9 and 5.7 ka B.P. (Mangerud et al., 1974; Stuiver et al., 1998). The last phase is the Neoglaciation, comprising the subboreal and subatlantic chronozones between ~5.7 and 0 ka B.P. (Mangerud et al., 1974; Stuiver et al., 1998). The transition from the Hypsithermal interval to the Neoglaciation would be illustrated by a general absence of glacier advances in the early to mid-Holocene and the onset of such advances at ~5 ka B.P. (Grove, 1988). In this paper, we address the possibility that this transition has been recorded in surface water indicators from marine sediment records from the North Atlantic and Mediterranean basins.

Changes in the Earth's precession, obliquity, and eccentricity (Milankovitch forcing) are usually regarded as the ultimate forcing of the long-term (~10 kyr and more) climate changes of the Quaternary. A commonly used indicator of this forcing on boreal hemisphere climate is the extraterrestrial insolation at a high northern latitude during the summer solstice. The insolation at 65°N at this time of the year reached a local maximum at 11–10 ka B.P. (Fig. 1), owing to a combination of summer perihelion and relatively large tilt of the rotational axis on the ecliptic plane. After 11–10 ka B.P. this insolation has declined steadily by ~10%. The amplitude and rate of insolation change are comparable to those characterizing the insolation increase from the last glacial maximum (LGM) at ~21 ka B.P. to the early Holocene. Were Milankovitch forcing exerts a dominant influence on the temperatures of high northern latitudes, a long-term cooling would have occurred at these latitudes during the last ~10 kyr (Broecker, 1998).

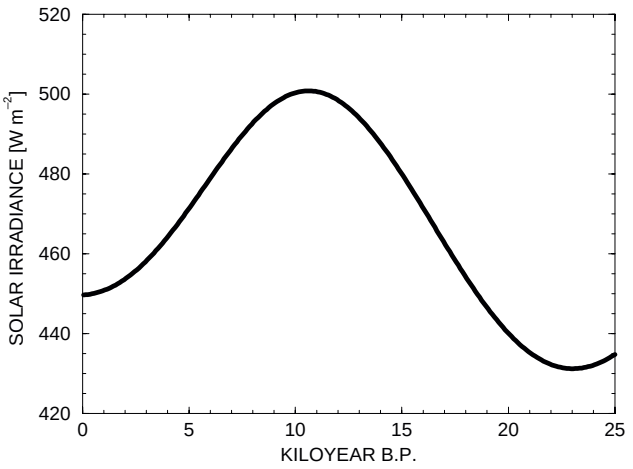

Fig. 1. Evolution of the summer solstice insolation at 65°N over the last 25 kyr (time is in calendar age). The isolation values are calculated according to Berger (1978).

Previous search for a long-term trend in upper ocean temperature (T) over the Holocene focused on the North Atlantic area (Ruddiman and Mix, 1993), as T reconstructions for the LGM show that this area is most thermally reactive on the time scale of 10 kyr (Mix et al., 2001). Temperature estimates for 9 and 6 ka B.P. were produced from sediment records of planktonic foraminiferal assemblage. These estimates did not lead to a coherent distribution of T anomalies at these times, and most of them were regarded as invalid owing to problems of age control and methodology used for quantitative reconstructions (Ruddiman and Mix, 1993). One possible cause of the relative small T anomalies is that ice, which is "widely regarded as one of the major means by which insolation forcing is converted into a climate response on Earth", was obviously less present at 9 and 6 ka B.P. than during the LGM (Ruddiman and Mix, 1993). These results are in apparent conflict with a recent compilation of other T proxies from northern North Atlantic sediments, showing higher summer T at 6 ka B.P. than at present (Kerwin et al. (1999) and references therein).

Here we explore the possibility of a long-term trend in upper ocean temperature during the Holocene, by examining carefully seven T reconstructions based on the more recent alkenone method from high accumulation sediment cores raised from the northeast Atlantic and the Mediterranean Sea. Reconstructions of T based on marine sediment records are contaminated by various pre- and post-depositional effects, in addition to sampling uncertainties and analytical errors. These effects include, e.g., chemical processes during particle settling and early diagenesis, lateral transport, changing sedimentation rates, and bioturbation in the sedimentary column. The contamination effects are particularly acute for records for the Holocene, a period during which T changes were modest compared, e.g., to those of the last deglaciation. In this paper we present an attempt to identify consistent trends in a collection of Holocene T records, using a range of statistical methods. Using different methods allows us to check the validity and stability of the trends, given the weakness of the apparent temperature signals in the Holocene records compared to potential contamination effects.

The paper is organized as follows. The core chronologies and the methods used to estimate the paleotemperatures are briefly described in Section 2. Section 3 provides an overview of the statistical techniques used to detect the trends; methodological details are reported in Appendixes A–D. The results obtained by applying these techniques to the biomarker paleo-T records are reported in Section 4. In Section 5 we discuss different possible contributions to the T trends in these records. We compare these records with evidence from other temperature proxies from marine and continental records from the northeast Atlantic area. Conclusions follow in Section 6.

2. Temperature reconstructions

2.1. Core chronologies

The T reconstructions examined in this paper come from sixteen high accumulation cores raised from the northeast Atlantic and the Mediterranean Sea (Table 1 and Fig. 2). Core chronologies have been described in previous publications (Table 1) and only a brief account is given here. The age controls are provided mainly by ^{14}C dates on planktonic foraminifera obtained by accelerator mass spectrometry. Auxiliary data, including ^{10}Pb-datings and foraminiferal $\delta^{18}O$ values, are also used to assist the chronology for some cores. Radiocarbon ages are corrected for the surface reservoir effect and calibrated to calendar ages (e.g., Stuiver and Reimer (1993); Stuiver et al. (1998)). The calendar chronologies indicate that the average sedimentation rate over the last 10 kyr ranges from 14 cm kyr^{-1} (core SU8118) to 61 cm kyr^{-1} (MD952011; Table 1).

2.2. Temperature estimates

A total of 26 reconstructions are produced from the 16 cores. The reconstructions are based on downcore variations in the concentration of alkenones, the composition of planktonic foraminifer assemblage, and/or the composition of diatom assemblage (Table 1).

The alkenone method is applied to seven cores (for methodological details see Cacho et al. (1999, 2001)). Different alkenone unsaturation indices (U^K_{37} or $U^{K'}_{37}$) and different calibrations to temperature are used. For core M23258 temperature values are obtained from the U^K_{37} index and a calibration to summer T based on core tops from the northeast Atlantic (Rosell-Melé et al., 1995). For core MD952011 the U^K_{37} index is employed in conjunction with the culture calibration of Prahl and

Table 1
Cores and temperature reconstructions

Core name	Latitude	Longitude	S [cm kyr^{-1}][a]	Alkenone	Summer Faunal	Ref.
M23258	74°59.7′N	13°58.2′E	25	Y	RAM, MAT[b], SIMMAX	[c]
PS21842	69°27.8′N	16°31.5′W	23		DI (I&K)	[d,e]
MD952011	66°58′N	07°38′E	61	Y	WA-PLS[b] and DI (I&K)	[f]
HM796	62°58′N	02°42′E	15		DI (I&K)	[g]
NEAP4K	61°30′N	24°10′W	17		SIMMAX[b]	
MD952015	58°46′N	25°58′W	59	Y	MAT[b]	
NEAP15K	56°22′N	27°49′W	40		SIMMAX[b]	
NA8722	55°29.8′N	14°41.7′W	24		I&K, RAM[b]	[h]
NEAP17K	54°41′N	28°21′W	18		I&K, MAT[b], SIMMAX	
CH7702	52°42′N	36°05′W	52		MAT[b]	
CH6909	41°45.4′N	47°21′W	16		RAM[b]	[i]
BS7938	38°24.7′N	13°34.6′E	22	Y		[j]
BS7933	38°15.7′N	14°01.8′E	16	Y		[j]
SU8118	37°46′N	10°11′W	14		I&K, RAM[b]	[h]
M39008	36°22.9′N	07°4.6′W	33	Y		[j]
MD952043	36°08.6′N	02°37.3′W	36	Y		[k]

[a] Average sedimentation rate over the last ~10 kyr.
[b] Summer faunal reconstruction used in principal component analysis.
[c] Van kreveld S., Sarnthein M., Erlenkeuser H., Grootes P., Pflaumann U., Martrat B., Villanueva J., and Grimalt J.O. Centennial-scale cyclicities of Holocene climate and sediment transport on the western Barents shelf (in prep).
[d] Koç et al. (1993).
[e] Koç and Jansen (1994).
[f] N. Koç, C. Andersen, C. Birks, H.J.H. Birks (unpublished data).
[g] Koç and Jansen (1992).
[h] Duplessy et al. (1992).
[i] Labeyrie et al. (1999).
[j] Cacho et al. (2001).
[k] Cacho et al. (1999).

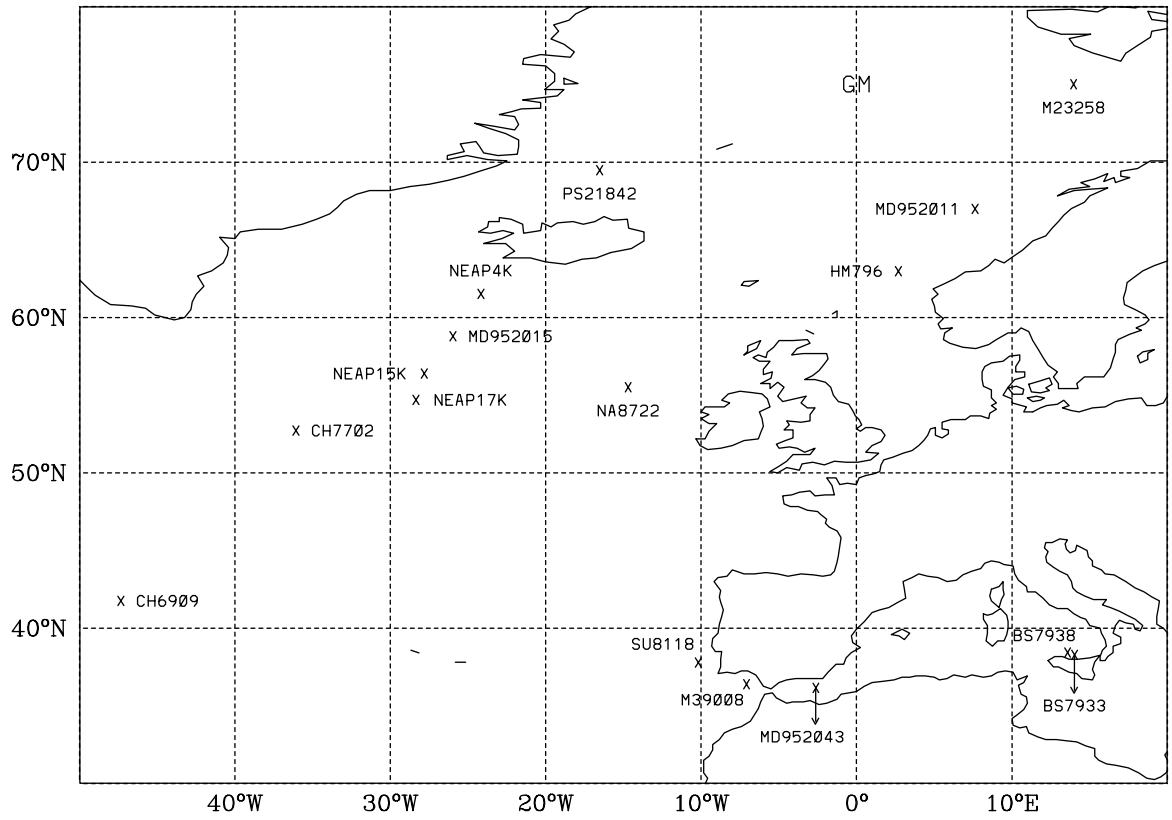

Fig. 2. Marine cores examined in this paper.

Wakeham (1987). For all the other cores (MD952015, BS7938, BS7933, M39008, and MD952043), temperature values are inferred from $U^{K'}_{37}$ using a global core-top calibration to annual mean T (Müller et al., 1998). Rosell-Melé et al. (1995) did not report an estimate of the error in T based on their regional core-top calibration. The error in T recovered from the culture calibration is $\sim$0.5°C (Prahl and Wakeham, 1987). The error reported for the global core-top calibration (scatter in the $U^{K'}_{37}-T$ plot) reaches $\sim$1.5°C (Müller et al., 1998). A main factor responsible for this much larger error is probably the fact that the $U^{K'}_{37}$ values used for the global calibration were measured in about ten laboratories which partly used different methodologies (Müller et al., 1998). By contrast, the alkenones in sediment samples considered here were measured in the same laboratory. Hence, the T error for cores MD952015, BS7938, BS7933, M39008, and MD952043 is presumably much lower than 1.5°C.

Sixteen records of summer T are based on the relative abundance of planktonic foraminifer species in core samples. Five different techniques are used to derive T from the abundance changes: the technique of Imbrie and Kipp (I&K hereafter; Imbrie and Kipp, (1971)), the modern analog technique (MAT; Overpeck et al. (1985)), the SIMMAX (an acronym for a modern analog technique using a similarity index; Pflaumann et al. (1996)), the revised analog method (RAM; Waelbroeck et al. (1998)), and a weighted averaging partial least square regression (WA-PLS; Ter Braak and Juggins (1993)). Waelbroeck et al. (1998) compared the errors in North Atlantic T produced by different faunal methods. The errors for the warm season T amount to $\sim$1.9°C for I&K, 1.1°C for MAT, 1.0°C for SIMMAX, and 0.9°C for RAM (Waelbroeck et al., 1998). No error estimates are available for T obtained from the WA-PLS (Ter Braak and Juggins, 1993). We stress that the transfer-function techniques mentioned above are widely used in paleoceanographic studies and represent the sort of data which is available in the literature.

Finally three records of summer T (labelled "DI" hereafter) are obtained from the relative abundance of diatom species. For each record the I&K method is applied to infer T from the changes in floral composition (Koç and Schrader, 1990). The error in warm T produced by the floral method amounts to $\sim$1.5°C (Koç and Schrader, 1990).

It is noteworthy that the T estimated from the biomarker, faunal, and floral methods may not be directly comparable. For instance, alkenones (long-chain, unsaturated ketones) can be synthesized by certain microalgae of the class *Prymnesiophyceae*, in particular by the marine coccolithophorids *Emiliania huxleyi* and *Gephyrocapsa oceanica*, and presumably by

other living and extinct members of the family *Gephyrocapsa*. Accordingly, the sedimentary U_{37}^{K} and $U_{37}^{K'}$ indices should register the temperature in the photic zone at the time of maximum coccolith production (Conte et al., 1995). On the other hand, the T estimated by the faunal and floral methods and considered in this paper, are summer values. The possible differences in the depth and/or seasonal timing of alkenone biosynthesis, foraminifer calcification, and diatom silicification will be considered when interpreting our multiproxy reconstructions of surface paleotemperatures.

3. Statistical methods

Four methods are used to identify trends in the T reconstructions. The first three methods consider each reconstruction separately, i.e., they are univariate. These methods are the linear regression analysis, the Mann test, and the smoothing spline. The fourth method merges several reconstructions into a single procedure to detect common temporal features, i.e., it is multivariate. This method is the principal component analysis. The linear regression analysis and the principal component analysis are so-called parametric methods, whereas the Mann test and the smoothing spline are nonparametric. Using both parametric and nonparametric procedures are a common strategy in applied statistics. Briefly, statistical inference based on the classical (parametric) methods makes fairly specific assumptions regarding the underlying distribution of the data. A frequent cause of departure from the implied distribution is the presence of outliers, which are not uncommon in marine sediment records. Assumptions about the underlying distribution of the data are not required in non-parametric inference. Nonparametric methods are thus useful complements of the parametric methods when these assumptions cannot be verified or when there is doubt about their validity (Gibbons and Chakraborti, 1992).

3.1. Linear regression

The linear regression model is an univariate parametric method (Appendix A). It allows us to examine the possibility of a constant trend in individual reconstructions; the least-square estimate of the paleo-T trend is noted b ($^{\circ}C\,kyr^{-1}$). Strictly, the linear model is known to be valid if the regression errors have zero mean (''randomness''), constant variance (''homoscedasticity''), and zero covariance. In addition, the construction of a confidence interval for the regression parameters relies on the assumption that these errors are normally distributed (Appendix A.1). Here, we assess a posteriori the randomness and normality conditions for each reconstruction using diagnostics based on regression residuals proposed in the statistical literature (Appendix A.2). We consult plots of regression residuals and conduct formal tests. The plots give a visual insight into the randomness and normality of the regression errors. The tests assist these plots by providing a more objective assessment of these two conditions.

For the randomness condition, we consult the plot of regression residuals versus calendar ages and conduct a runs tests (Draper and Smith, 1998). The test is based on the number n_C of sign changes in the sequence of residuals and on the significance level p_C of this number. A low n_C indicates clustering whereas a high n_C indicates mixing of the residuals. The value of p_C, on the other hand, is the probability that the number of sign changes $<n_C$ for perfectly mixed residuals. Thus, a relatively high value of n_C (i.e., compared to the number of data in the reconstruction under consideration) and a high p_C tend to validate the linear model, that is, the occurrence of an underlying constant trend in the reconstruction.

For the normality condition, we produce the normal plot for the standardized deletion residuals (Sen and Srivastasa, 1990; Ryan, 1997) and calculate ''simulation envelopes'' (Atkinson, 1981). A linear pattern in the plot suggests that the regression errors are normally distributed. The envelopes essentially provide a limit on how non-normal the regression residuals can *appear* to be when these errors have a normal distribution. Thus, data points outside the envelopes in the normal plots suggest non-normal residuals. We test the normality assumption by calculating the Shapiro–Wilk statistics W ($0<W<1$) and its significance level p_W (Madansky, 1988; Sen and Srivastasa, 1990). High values of W and p_W indicate normality.

Finally, we attempt to detect data that have a particularly large influence on b for reconstructions for which the linear model is substantiated. We inspect the plot of $(b-b_{(i)})/|b|$ (Sen and Srivastasa, 1990; Ryan, 1997), where $b_{(i)}$ is the trend obtained when the ith observation is removed from the linear regression. A high absolute value of $(b-b_{(i)})/|b|$ indicates that the ith observation has a large influence on the paleo-T trend estimated by the least-square fit.

3.2. Mann test

The Mann test is one of the most popular nonparametric tests for the detection of a trend in univariate time series (for a short review see Bhattacharyya (1984); Appendix B). The test statistics is the rank correlation coefficient Kendall tau between the reconstructed T and the calendar ages; this coefficient is noted as t ($-1\leqslant t\leqslant +1$). The value of t is computed accounting for the possible occurrence of ''ties'' in the reconstructions (Appendix B.1), and its significance level p_t is then determined (Appendix B.2). As t is based on the ranking of the data (not directly on their numerical values), a high $|t|$ indicates a monotonic change of the

paleotemperatures with time, with $t<0$ indicating a negative change (cooling) and $t>0$ indicating a positive change (warming). The value of p_t, on the other hand, is the probability that the Kendall tau $< -t$ or $> +t$ in the case of no underlying rank correlation between the T values and the calendar ages. Thus, a high $|t|$ and a low p_t indicate a downward or upward monotonic trend in the paleo-T records.

Because t is based on ranks, it is much less sensitive to outliers than the linear regression based on least-square fit. For the same reason t is invariant under any rank-preserving transformation of the data. For instance, the value of t would be unchanged if radiocarbon ages or even core depths were used instead of calendar ages. The value of t is therefore insensitive to uncertainties in the time scale of the paleo-T records (different age models for the same core, however, could lead to a different number of T values and calendar ages for the last 10 kyr and thus to a different value of t). We thus use the Mann test as a robust complement of the linear regression model to detect trends in these records. The interpretation of t may be delicate, on the other hand, when the data show indication of a time-variable trend. The test is thus applied only to records showing a random pattern in linear regression residuals.

3.3. Smoothing spline

The smoothing spline is also an univariate non-parametric method (Appendix C). Previous studies used smoothing splines, a form of non-parametric regression, to estimate the time-variable trend in Holocene records (e.g., Stuiver et al. (1991); O'Brien et al. (1995); Indermühle et al. (1999)). Here, we apply a smoothing spline to paleo-T records showing a non-random pattern in linear regression residuals. A smoothing spline is a trade-off between a least-square fit to the data and smoothness, determined by a smoothing parameter ρ. We try different values of ρ. We then inspect the plot of smoothing spline residuals versus calendar ages and perform a runs test (n_C and p_C) on these residuals to assess the quality of the corresponding splines (Appendix C.1). In contrast to the linear model and the Mann test, our approach here is purely descriptive, i.e., we aim at estimating temporal features in individual reconstructions but do not make inference about these features.

3.4. Principal component analysis

The principal component analysis (PCA) is a multivariate parametric method (Appendix D), identical to empirical orthogonal function (EOF) analysis. The PCA has already been used to clarify the covariability between different Holocene records (e.g., O'Brien et al. (1995)). Here we seek to maximize the variance of linear combinations of T reconstructions at different geographic locations; these combinations are the *principal components*. The first component accounts for the largest fraction of the total variance in the reconstructions, the second component accounts for the second largest fraction, and so on. If the reconstructions are strongly correlated, the leading principal components yield a good summary of the variation in the data set. In our attempt to clarify objectively trends shared by several reconstructions, we are thus looking for a situation where the leading principal components capture a substantial portion of the total variance in these reconstructions.

Applying a PCA requires that the T values from different reconstructions are available for identical calendar ages or for identical time intervals. Here, mean T values in the same time intervals are derived from each reconstruction, taking into account the different time resolutions of the reconstructions used in the PCA (Appendix D.1). Principal components are then extracted from either the covariance matrix or the correlation matrix of these values (Appendix E.2). Significance tests are not performed (Appendix D.3), so that our approach again is exploratory and not inferential.

4. Results

All alkenone reconstructions show indication of apparent long-term cooling during the last 10 kyr (solid circles in leftmost panel of Fig. 3a–g). It is noteworthy that apparent cooling is observed in so different oceanographic environments, including the Barents slope (core M23258), off Norway (MD952011), southwest of Iceland (MD952011), the Gulf of Cadiz (M39008), the Alboran Sea (MD952043), and the Tyrrhenian Sea (BS7938 and BS7933). In this section, we apply the four statistical methods described above to constrain the nature and the amplitude of the trends underlying the biomarker paleotemperature records. We consider first the results from the three univariate methods (linear regression, Mann test, and smoothing spline) and then the results from the multivariate method (principal component analysis).

4.1. Univariate trends

We adopt the following three-step strategy to constrain trends in individual biomarker paleo-T records. In the first step we conduct a linear regression analysis. The statistical properties of the corresponding residuals are examined carefully so as to explore the possibility of a constant trend (b). We regard as valid the b values for records for which residuals show no clustering significant at 5% level ($p_C > 0.05$). We report

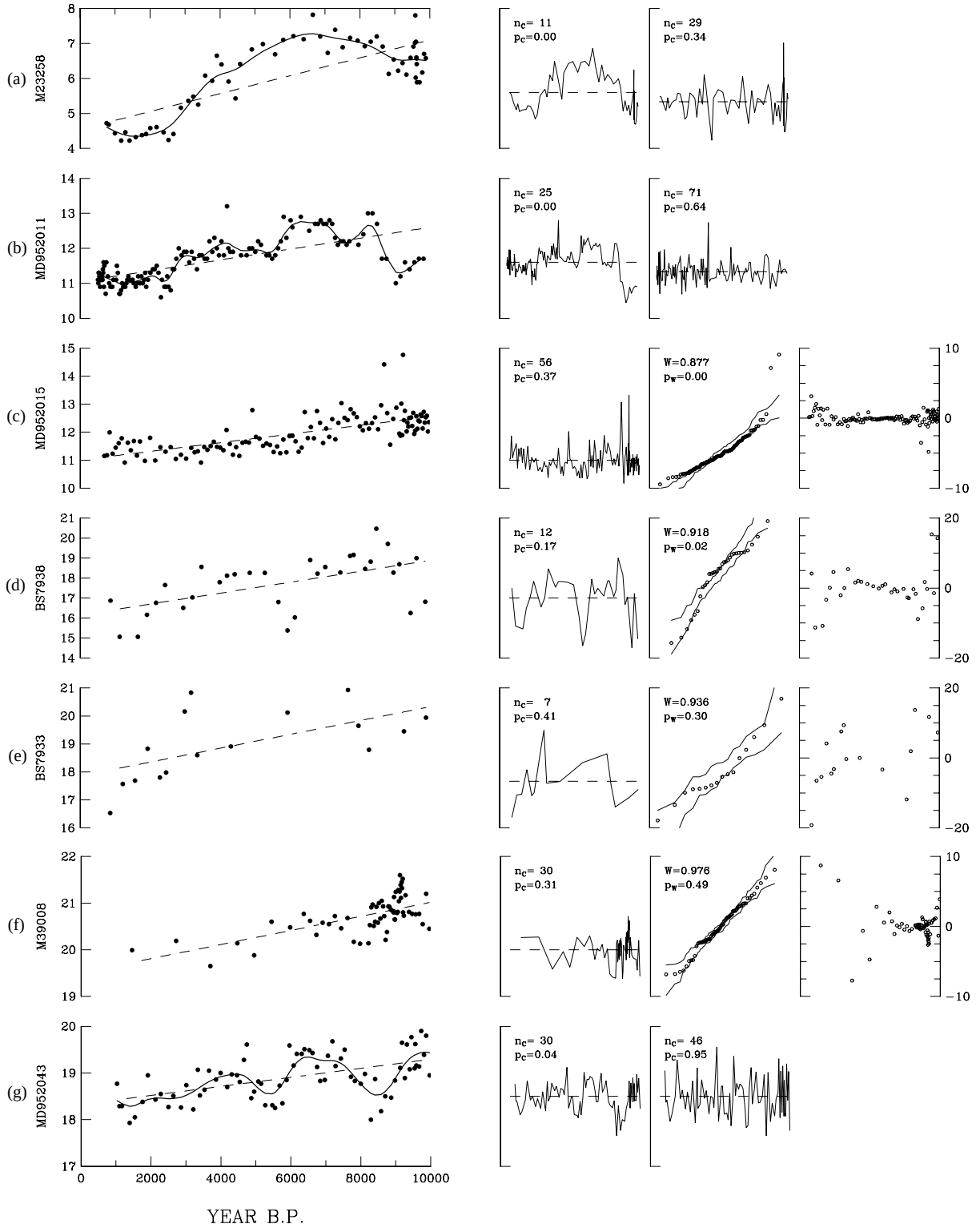

Fig. 3. Records of upper ocean temperature (T) based on alkenones and results from univariate methods (linear regression and smoothing spline). Leftmost panel: Time series of T (solid circles), linear regression fit (dashed line), and cubic smoothing spline (solid line). The vertical scale represents °C. Second panel from left: Linear regression residuals versus calendar ages. n_C is the number of runs in the sequence of residuals and p_C is the significance level of this number. Third panel from left (for cores for which $p_C > 0.05$): Observed versus expected standardized deletion residuals (open circles) and simulation envelopes (two solid lines). W and p_W are the Shapiro–Wilk statistics and its significance level, respectively. Third panel from left (for cores for which $p_C < 0.05$): Smoothing spline residuals versus calendar ages with the corresponding n_C and p_C values. Fourth panel from left (for cores for which $p_C > 0.05$): Ratio $(b - b_{(i)})/|b|$ (in %) versus calendar ages. b is the linear paleotemperature trend and $b_{(i)}$ is the same quantity but when the ith observation is removed from the linear regression. For the second and third panels from left no labels are reported since the information is in the pattern.

confidence intervals of b only for records for which residuals show in addition no signs of non-normality at the same significance level ($p_W > 0.05$). Our decisions about b and its confidence interval are assisted by three diagnostic plots: linear regression residuals versus calendar ages, standardized deletion residuals versus their expected values if these residuals were normally distributed (normal plots including simulation envelopes), and the ratio $(b - b_{(i)})/|b|$ versus calendar ages (influential observations). In a second step, we complement the regression analysis by the Mann test for records showing a random pattern in linear regression residuals. Finally, we apply in a third step a cubic smoothing spline to the other records to estimate the time-variable trend underlying these records.

First, we conduct a linear regression analysis and examine the regression residuals for each biomarker paleo-T record (second panel from left in Fig. 3a–g). The residuals from cores M23258, MD952011, and MD952043, show a tendency to cluster, which is significant at the 5% level ($p_C < 0.05$; Fig. 3a–b and g). Thus, at least four reconstructions can be reasonably approximated by constant trends: MD952015, BS7938, BS7933, and M39008. The apparent cooling rate for these cores ranges from -0.27°C to -0.15°C kyr^{-1} (Table 2). We examine the normal plots for the standardized deletion residuals to (in)validate the confidence interval for these rates (third panel from left in Fig. 3c–f). The p-value of the Shapiro–Wilk statistics $p_W = 0.00$ for MD952015 and 0.02 for BS7938, suggesting non-normally distributed regression errors for these cores. Non-normality in MD952015 is likely due, at least partly, to the two high T values in the early Holocene (leftmost panel in Fig. 3c). These values correspond to the two data points which fall far away from the simulation envelopes in the normal plot (third panel from left in Fig. 3c). They are outliers, but not influential observations as their individual omission from the regression analysis changes the estimate of apparent cooling rate only by ~5% or less (rightmost panel in Fig. 3c). On the other hand, we find no indication of non-normality for BS7933 and M39008 (third panel from left in Fig. 3e–f). It appears therefore reasonable to interpret the confidence interval for the apparent cooling rate in these cores. The rate ± two standard errors amounts to -0.25 ± 0.19°C kyr^{-1} in BS7933 and -0.15 ± 0.05°C kyr^{-1} in M39008 (Table 2). Apparent cooling is thus significant at the 5% level for these cores.

We now complement the regression analysis by the Mann test for records showing a random pattern in linear regression residuals. The association between T and calendar ages in MD952015, BS7938, BS7933, and M39008 is measured by the rank correlation coefficient Kendall tau (t) and the significance of this coefficient (p_t) is determined. The test shows that apparent cooling is significant at the 5% level for all these cores ($p_t < 0.05$; Table 2). The parametric (b) and nonparametric measures (t) of univariate trend produce therefore a consistent geographic picture, with apparent cooling in the open northeast Atlantic, the Gulf of Cadiz, and the Tyrrhenian Sea (Fig. 4).

Finally a cubic smoothing spline is applied to the three reconstructions for which the linear regression diagnostics question the occurrence of a constant trend (M23258, MD952011, and MD952043). For each of these reconstructions, we try different values of the smoothing parameter ρ. We select a value which produces a random pattern in the plot of smoothing spline residuals versus calendar ages and a relatively high value of the runs test statistics n_C and p_C (rightmost panel in Fig. 3a–b and g). The resulting splines clarify the apparent long-term cooling (solid line in leftmost panel of Fig. 3a–b and g). The cooling trend initiates only after ~6.0–6.5 ka B.P. in M23258 and is apparently contaminated by periodic fluctuations in MD952011 and MD952043. According to the splines the peak-to-peak amplitude in T reaches 2.9°C for M23258, 1.8°C for MD952011, and 1.2°C for MD952043 (Fig. 3a–b and g).

Table 2
Univariate trends in temperature reconstructions

Core	Technique	b(°C kyr^{-1})[a]	$t(p_t)$[b]
		T	
MD952015	Alkenones	−0.15	−0.54 (0.00)
BS7938	—	−0.27	−0.46 (0.00)
BS7933	—	−0.25 ± 0.19	−0.52 (0.01)
M39008	—	−0.15 ± 0.05	−0.43 (0.00)
		Summer Faunal T	
NEAP4k	SIMMAX	0.08	0.27 (0.05)
NA8722	I&K	0.09 ± 0.12	0.18 (0.21)
	RAM	0.07 ± 0.07	0.12 (0.14)
CH6909	RAM	0.12	0.27 (0.01)
SU8118	I&K	0.20 ± 0.26	0.39 (0.05)
	RAM	0.04 ± 0.11	−0.02 (0.96)

[a] Linear trend ±95% confidence interval.
[b] Kendall tau and its significance level (between parentheses).

4.2. Multivariate trends

We perform a principal component analysis on all alkenone reconstructions. Among these reconstructions the youngest and oldest calendar ages are equal to 1460 and 9820 ka B.P., respectively. The core with minimum sampling frequency is BS7933 ($n_{\min} = 16$; Fig. 3e). Hence the time interval for which mean T values are calculated for each paleo-T record, amounts to $\overline{\Delta x} = 557$ y. For each record the mean values seem to provide a good description of the T variability on a time

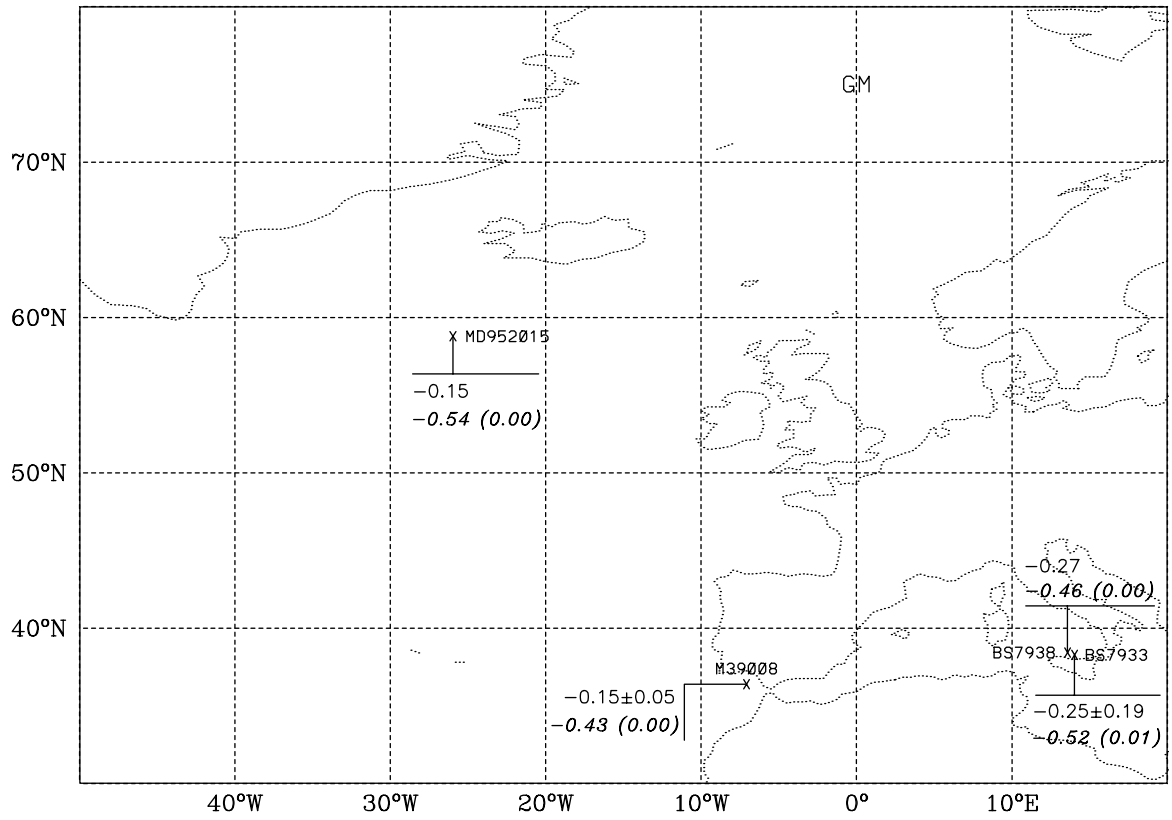

Fig. 4. Results from the univariate methods (linear regression and Mann test) for paleo-T records based on alkenones. The non-italicized values are the linear trend (b) and its 95% confidence interval reported only for cores with indication of normal regression residuals (°C kyr^{-1}). The italicized values are the rank correlation coefficient Kendall tau (t) and its significance level (p_t; between parentheses).

scale of several millennia (open and solid circles in Fig. 5). In particular the mean values for each record show, as expected, the apparent cooling over the Holocene. We note that mean T values are missing in six time intervals for BS7933 and in two intervals for M39008 (solid circles in Fig. 5). We will test the effect of removing the data from these cores on the results of the PCA.

4.2.1. Covariance and correlation matrices

The mean T values (Fig. 5) are used to calculate the covariance matrix **S** for the alkenone reconstructions (Table 3). The minimum and maximum variances in **S** differ by one order of magnitude: maximum variance is found for BS7938 and minimum variance for M39008. Although about 40% of the mean T values in BS7933 (6 of 16; Fig. 5) are "unconditional means" (Little and Rubin (1987); Appendix D.1), the variance for this core is the third in amplitude. To aid the interpretation of **S** we consider the correlation matrix **R** for the alkenone reconstructions (Table 3). Many elements in **R** are relatively high, suggesting significant correlation between several reconstructions and hence the possibility to summarize the total variation in **S** by only a few principal components. We note the relatively small correlation (0.376) between the two nearby cores BS7933 and BS7938. This likely reflects the different sampling frequencies in the two cores ($n = n_{\min} = 16$ versus $n = 32$) and the occurrence of a relatively large number of unconditional means in BS7933. The small correlation suggests that incorporating data from both BS7933 and BS7938 in the PCA would not generate much redundancy. Clearly inspection of the mean T values (Fig. 5) and of their covariance and correlation matrices (Table 3), promotes the application of a PCA to elucidate the covariation between the biomarker paleo-T records.

4.2.2. Eigenvalues

To gain a better insight into this covariation we consider the size of the eigenvalues λ_k of **S** ($k = 1, \ldots, q$). When the variables show large covariances, a plot of λ_k versus k reveals a few eigenvalues dominating markedly the remaining eigenvalues; such a plot is sometimes called a "scree graph" (Rencher, 1995). Here we plot the eigenvalue size λ_k divided by the sum of the eigenvalues $\sum_{k=1}^{q} \lambda_k$, versus k (Fig. 6). The plot has the same appearance as the corresponding scree graph but in addition displays the relative contribution of each principal component to the total variance in the

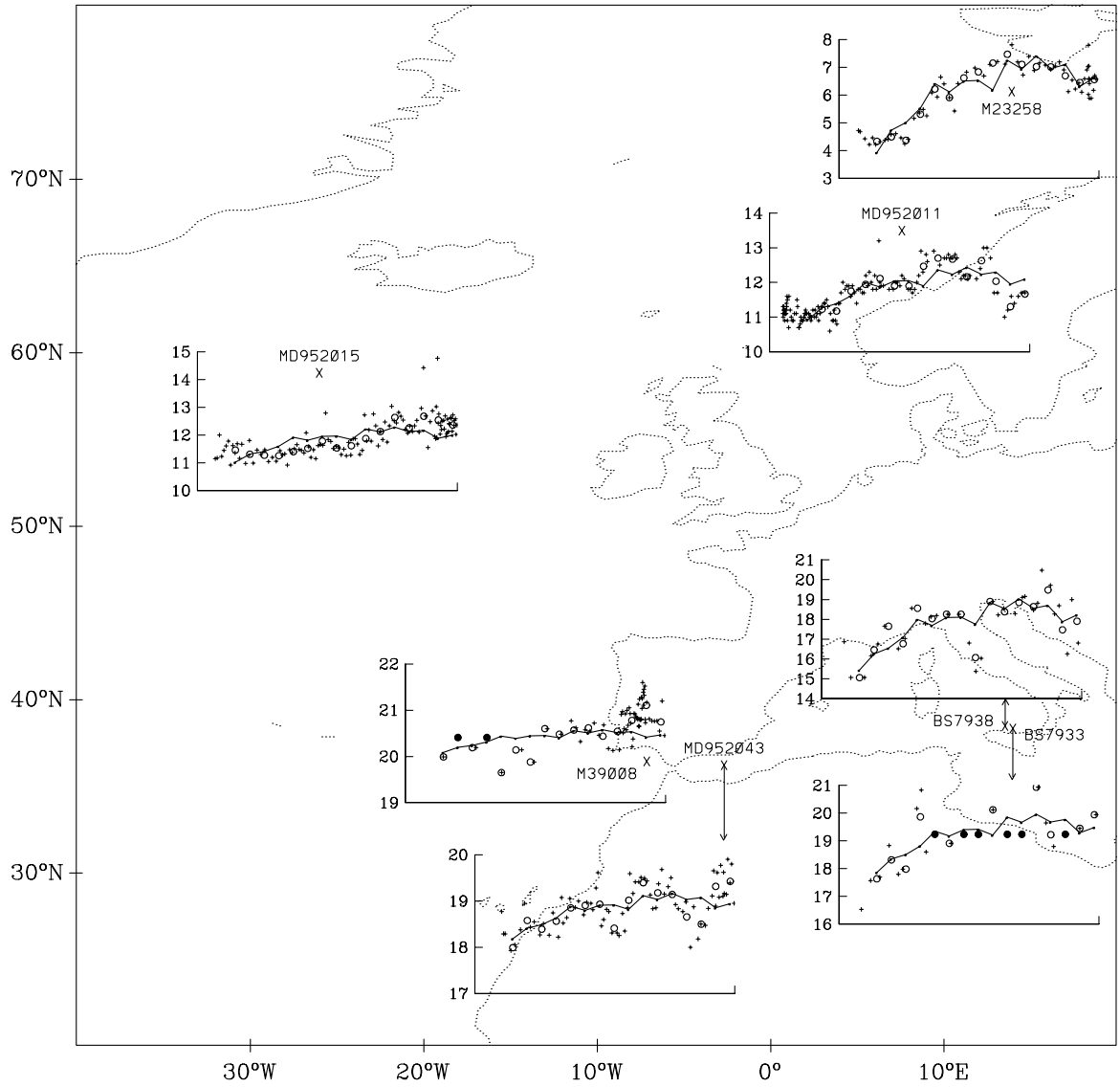

Fig. 5. Results from the principal component analysis (PCA) for paleo-T records based on alkenones. The crosses represent the original T values obtained from the alkenones. The open and solid circles are mean T in time intervals of 561 yr used in the PCA. The solid circles are unconditional means (only in cores BS7933 and M39008). The solid line is the synthesis of mean T approximated by the first principal component. The horizontal scale in each plot is linear, ranging from 0 ka B.P. (left) to 10 ka B.P. (right).

paleotemperature records (since the variance of the kth component is equal to λ_k and that $\sum_{k=1}^{q} \lambda_k$ is equal to the total variance in the records, i.e., the trace of **S** noted tr(**S**)).

The scree graph shows clearly that the first principal component dominates the alkenone reconstructions (solid circles in Fig. 6a). This component accounts for 67% of the total variance in these reconstructions ($\lambda_1 / \sum_{k=1}^{q} \lambda_k = 0.67$ tr(**S**)). The second component contributes to 17% of the total variance ($\lambda_2 / \sum_{k=1}^{q} \lambda_k = 0.17$ tr(**S**)). It separates the first component from the last five components which together accounts for only 16% of the total variance and tend to fall into a straight line in the scree graph.

Rencher (1995) reported three guidelines to determine the number of components that should be retained to effectively summarize a multivariate data set: (1) inspecting the scree graph, which may reveal a natural break between the "large" and "small" eigenvalues; (2) retaining sufficient components for a specified percentage of the total variance, say 80%; and (3) excluding components whose eigenvalues are less than the average of the eigenvalues $\sum_{k=1}^{q} \lambda_k / q$. A fourth guideline, which is a formal test requiring multivariate normality (Rencher, 1995), is not considered here. Guideline (1) is to retain those eigenvalues in the steep curve before the first one on the straight line (Rencher, 1995). Regarding guideline (3) the first, second, and third

Table 3
Covariance matrix and correlation matrix for temperature reconstructions based on alkenones

	M23258	MD952011	MD952015	BS7938	BS7933	M39008	MD952043
Covariance matrix							
M23258	1.088						
MD952011	0.471	0.293					
MD952015	0.327	0.087	0.259				
BS7938	0.776	0.351	0.326	1.383			
BS7933	0.596	0.227	0.214	0.356	0.647		
M39008	0.148	0.028	0.118	0.075	0.102	0.134	
MD952043	0.286	0.106	0.105	0.193	0.210	0.057	0.168
Correlation matrix							
M23258	1.000						
MD952011	0.835	1.000					
MD952015	0.616	0.316	1.000				
BS7938	0.633	0.552	0.545	1.000			
BS7933	0.711	0.522	0.523	0.376	1.000		
M39008	0.387	0.142	0.634	0.175	0.346	1.000	
MD952043	0.668	0.476	0.501	0.399	0.638	0.379	1.000

eigenvalues amount to 2.67, 0.67, and 0.27°C^2, respectively, which must be compared to $\sum_{k=1}^{q} \lambda_k/q = 0.57$ °C^2. Thus guidelines (1)–(3) all indicate that the first two components (contributing together to 84% of total variance) would be sufficient to summarize the total variation in the biomarker paleo-T records.

4.2.3. Eigenvectors

We now consider the eigenvectors of **S** to clarify further the covariation between these records (Table 4). All elements of the first eigenvector ($\mathbf{e}_1$) are positive (as all elements of the positive definite matrix **S** are positive; Table 3). This result implies that the first principal component (67% of total variance) is essentially a weighted average between the different paleo-T records, where the weights are function of the variances (the larger the variance, the larger the weight; compare diagonal elements of **S** in Table 3 with elements of $\mathbf{e}_1$ in Table 4). Such a component, with all elements in the corresponding eigenvector being positive, is referred to as a "size" component (Rencher, 1995). Although the records differ widely in variance, all these records contribute significantly to the first component. Different from the first component, the second component is a contrast between the reconstructions for M23258, MD952011, MD952015, BS7933, M39008, and MD952043 on the one hand, and the reconstruction for BS7938 on the other hand. Such a component, with elements in the corresponding eigenvector being sometimes positive, sometimes negative, is a "shape" component (Rencher, 1995).

4.2.4. Principal components

Finally, we examine the time series of the two leading principal components to estimate temporal features shared by the T reconstructions. The first component (67% of tr(**S**)) exhibits an irregular long-term decrease during the Holocene (solid circles in Fig. 7a). The second component (17% of tr(**S**)), on the other hand, does not portray a clear, long-term change during this period (solid circles in Fig. 7b). We use a synthesis formula to quantify at each coring site the T evolution associated with the first mode of variability common to all reconstructions (Appendix D.2). The evolutions associated with the first component show expected long-term decreases during the Holocene (solid line in Fig. 5). The corresponding peak-to-peak amplitude in T ranges from 0.5°C to 3.6°C, depending on the core. Our major results from the PCA are robust against (1) the inclusion or omission of data from cores for which unconditional means are imputed (BS7933 and M39008; Appendix E.1); and (2) the extraction of principal components from the covariance matrix or the correlation matrix (open circles in Figs. 6a and 7a–b; Appendix E.2).

5. Discussion

We summarize our major results. A suite of statistical analyses are used to detect trends in seven Holocene paleotemperature records based on alkenones from cores raised from the northeast Atlantic (Barents slope, off Norway, southwest of Iceland, and Gulf of Cadiz) and the Mediterranean basin (Alboran and Tyrrhenian Seas). The different statistical analyses are an attempt to check the validity and stability of these trends, given their small amplitudes compared to various potential contaminations in the records. All these records document an apparent long-term cooling during the

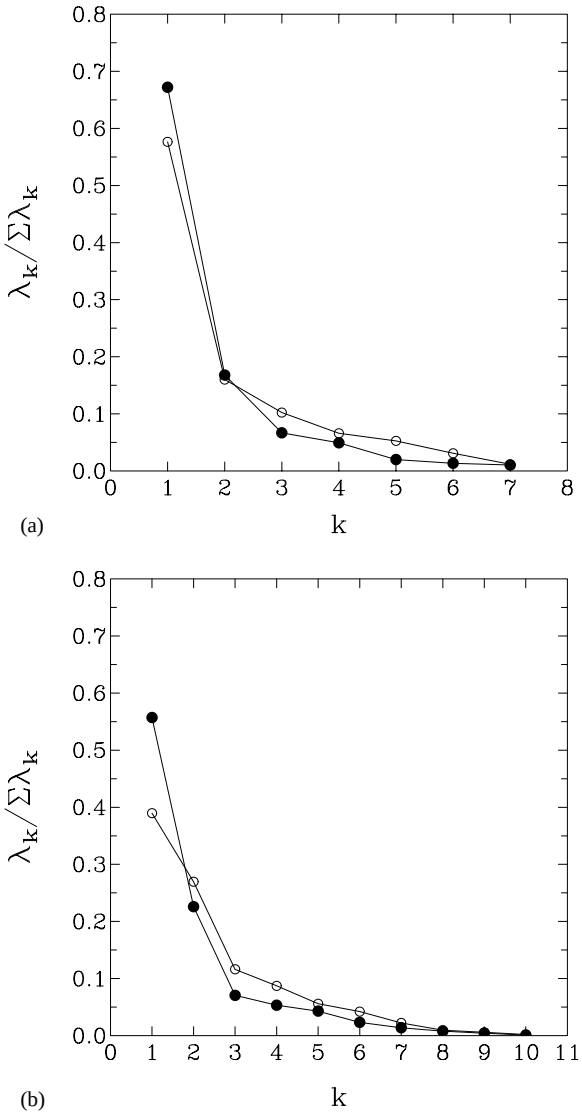

Fig. 6. Results from the PCA for (a) T records based on alkenones and (b) summer T records based on foraminiferal counts. The two panels show the eigenvalue size divided by sum of eigenvalues, versus eigenvalue number for the covariance matrix **S** (solid circles) and the correlation matrix **R** for these records (open circles).

Holocene. The linear regression analysis suggests that the apparent cooling occurred at a constant rate in four records, ranging from -0.27 to $-0.15^{\circ}C\,kyr^{-1}$. The resulting temperature changes over the Holocene (1.5–2.7°C) are comparable to those inferred for the other records on the basis of the smoothing spline (1.2–2.9°C). All these values are generally larger than the errors in the biomarker paleotemperatures considered in this paper ($<1.5^{\circ}C$). The Mann test always corroborates the conclusions from the linear regression regarding the significance of the T trends. Finally, the principal component analysis shows that a large fraction (67%) of the total variance in the biomarker paleo-T records is a spatially coherent pattern corresponding to local

Table 4
First and second eigenvectors of covariance matrices

	Eigenvector	
	$\mathbf{e}_1$	$\mathbf{e}_2$
Alkenone T		
M23258	0.59	−0.33
MD952011	0.26	−0.10
MD952015	0.22	−0.04
BS7938	0.61	0.75
BS7933	0.36	−0.53
M39008	0.08	−0.12
MD952043	0.17	−0.17
Summer Faunal T		
M23258 (MAT)	0.76	−0.23
MD952011 (WA-PLS)	0.17	0.40
NEAP4K (SIMMAX)	−0.24	0.34
MD952015 (MAT)	0.05	0.21
NEAP15K (SIMMAX)	0.29	0.65
NA8722 (RAM)	−0.15	0.10
NEAP17K (MAT)	0.29	−0.22
CH7702 (MAT)	0.36	0.27
CH6909 (RAM)	−0.13	0.23
SU8118 (RAM)	0.03	0.14

apparent toolings of 0.5–3.6°C. In the following, we discuss different possible interpretations of this pattern based on other paleodata and climate model simulations.

5.1. Widespread surface cooling

The hypothesis that the decrease in biomarker paleotemperatures reflects real cooling is consistent with an orbital control of summer temperatures at high latitudes in the northern hemisphere and with the transition from the Hypsithermal warm interval (~9–5.7 ka B.P.) to the subsequent neoglacial period (~5.7–0 ka B.P.). In this section, we first summarize evidence for a long-term cooling during the Holocene from continental records and climate model simulations. We then come back to other T proxies from marine sediment records from the same cores and from other cores from the northeast Atlantic.

5.1.1. Continental records

Several continental records from the circum North Atlantic provide evidence of a long-term cooling during the Holocene. The observational evidence come from borehole temperature measurements in the Greenland ice sheet, records of glacier advances and retreats in the northeast Atlantic area, and pollen sequences documenting vegetational changes in Europe.

The inversion of borehole temperature data at GRIP (summit of Greenland ice sheet) shows rapid warming at the beginning of the Holocene followed by a thermal

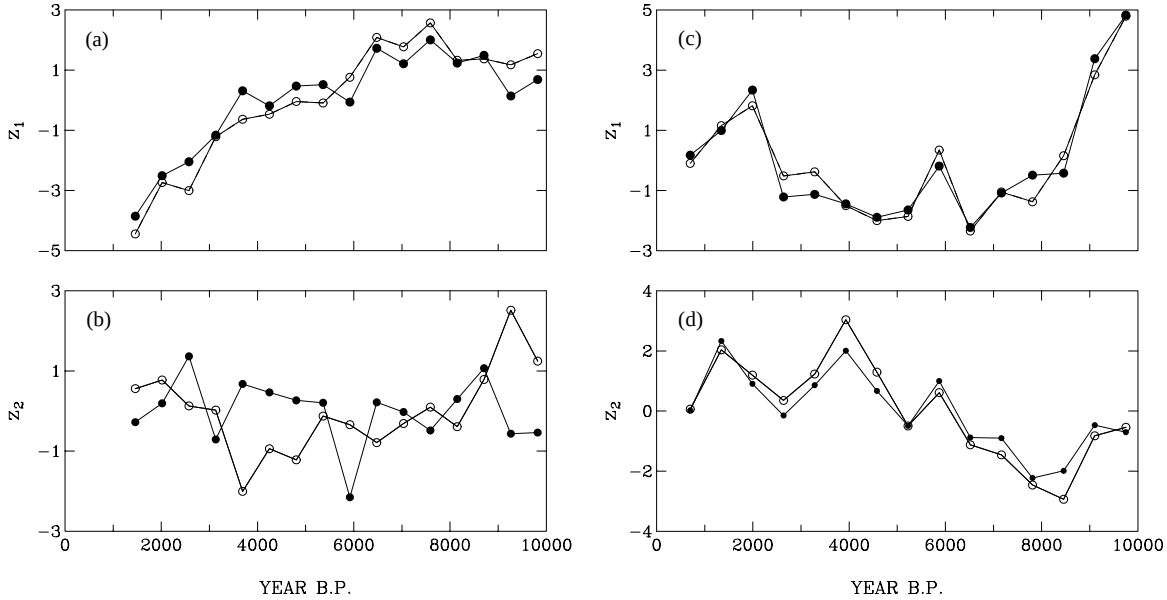

Fig. 7. Results from the PCA for ((a)–(b)) paleo-T records based on alkenones and ((c)–(d)) summer paleo-T records based on foraminiferal assemblages. Panels (a) and (c) show the first principal component and panels (b) and (d) show the second principal components of the covariance matrix **S** (solid circles) and of the correlation matrix **R** (open circles) for these records.

plateau (between 8 and 5 ka B.P.) and then general cooling (Dahl-Jensen et al., 1998). The plateau and two later thermal maxima superimposed on the gradual cooling were interpreted as the Climatic Optimum, the Medieval Warm Period, and the Little Ice Age, with temperature anomalies of $+2.5°C$, $+1°C$, and $-1°C$, respectively (Dahl-Jensen et al., 1998). The inversion of borehole temperatures at Dye 3 (southern Greenland) portrays a much shorter plateau than at GRIP and then a cooling which also initiated during the mid-Holocene (borehole temperature data at Dye 3 are only available for the 7–0 ka B.P. interval; Dahl-Jensen et al. (1998)). Interestingly the thermal features at GRIP agree qualitatively with our two northernmost biomarker T records despite differences in their timing (Fig. 3a–b). Note, on the other hand, that the ice $\delta^{18}O$ records at GRIP and GISP2 (30 km west of GRIP; Dansgaard et al. (1993); Grootes and Stuiver (1997)), and the snow accumulation record at GISP2 (Meese et al., 1994) do not reveal clear overall trends over the Holocene.

The composite records of glacier termini from Iceland (Gudmundsson, 1997), western Norway (Nesje and Dahl, 1993), Svalbard and Franz Josef Land (Lubinski et al. (1999) and references therein), show that glaciers were behind their present margins from the early to mid-Holocene and experienced several major advances during the late Holocene. Independent paleoclimate evidence from Svalbard suggest that this change in glacier mass balance resulted from a long-term decrease in summer temperature (Lubinski et al., 1999). This interpretation is in harmony with elevation changes of fossil pine wood samples in southern Norway, which was interpreted as a decrease in mean summer temperature over the last ~9 kyr (Dahl and Nesje, 1996). The changes in glacier activity in the northeast Atlantic area are consistent with the record of worldwide glacier fluctuations (including the Alps), showing less frequent expansion phases during the Hypsithermal interval than during the subsequent Neoglaciation (Grove, 1988).

The Cooperative Holocene Mapping Project (COHMAP, 1988) used pollen profiles to investigate the global and regional dynamics of climate change during the last 18 kyr. One major result of this project was a clear southward migration of the southern limit of spruce in both Europe and North America after ~6 ka B.P., which was ascribed to decreased summer temperatures (COHMAP, 1988 and references therein). The climatic anomaly patterns in Europe inferred more recently from pollen data are greater winter and summer temperatures in Scandinavia at 9 and 6 ka B.P.; lower temperatures in southern Europe were inferred for these times (Huntley and Prentice, 1993). On the other hand, a continuous pollen sequence from the Middle Atlas, Morocco, suggests that the climate was drier and warmer from ~10 to ~6.5 ka B.P. than during the period since 6.5 ka B.P. (Cheddadi et al., 1998). January and July temperatures between 10–6.5 ka B.P. would have been about 4°C higher than at present (Cheddadi et al., 1998).

5.1.2. Climate model simulations

Climate model simulations also support the hypothesis of a long-term cooling during the Holocene. The

climates at 9, 6 and 3 ka B.P. were simulated with an atmospheric general circulation model (AGCM) including Milankovitch forcing and other appropriate boundary conditions (Kutzbach and Guetter, 1986; Kutzbach et al., 1993). The simulations predict generally warmer conditions above land in summer in the northern hemisphere and less important temperature changes in winter (Kutzbach and Guetter, 1986). Positive temperature anomalies above northern Europe are predicted for both seasons at 9 and 6 ka B.P., consistent with the pollen reconstructions (Fig. 7.9 of Huntley and Prentice (1993)). More recent AGCM simulations including orbital forcing for 10 and 6 ka B.P. indicate that the Arctic region was warmer at these times, pointing to late Holocene cooling, with the warmest conditions predicted in the North Atlantic sector (Pollard et al., 2000).

Two possible limitations of the model simulations discussed above are their assumption that climate is in equilibrium with external forcing and their omission of potentially important feedbacks associated, e.g., with changes in the ocean circulation. Recently the evolution of climate over the last 9 kyr has been simulated with a coupled zonally averaged ocean–atmosphere–land biosphere model (Crucifix et al., 2001). The results show a monotonous cooling over land and ocean at high northern latitudes from 9 ka B.P. to today; north of 60°N, cooling is distributed throughout the year but is more intense for spring and summer (Crucifix et al., 2001).

Taken together, both continental records and climate model simulations support the occurrence of a long-term cooling in the circum North Atlantic during the Holocene. These evidence suggest that the interpretation of the decrease in biomarker paleotemperatures in terms of surface cooling must be considered seriously. Next, we consider whether other paleo-T records from northeast Atlantic sediments document a similar cooling. We pay attention to the fact that the comparibility between different paleo-T proxies may be confounded by possible seasonal and/or depth artifacts.

5.1.3. Paleo-T records based on diatoms

We first consult our records of summer T based on the composition of another phytoplankton group, i.e., the diatoms. Comparability with the biomarker paleotemperatures might be favoured by the fact that both coccolithophorids and diatoms are photosynthetic organisms so that their habitat is restricted to the photic zone. A linear regression analysis and a smoothing spline are applied to our floral reconstructions. None of the reconstructions exhibits a constant trend (Fig. 8a–c). Smoothing the reconstructions for PS21842 and MD952011 suggests relatively rapid warming at the beginning of the Holocene, a period with high T in the first half of the interglacial, and then a gradual cooling. The two former features are also found in the reconstructions for HM796 (albeit with a different timing). On the other hand, the paleo-T record for this

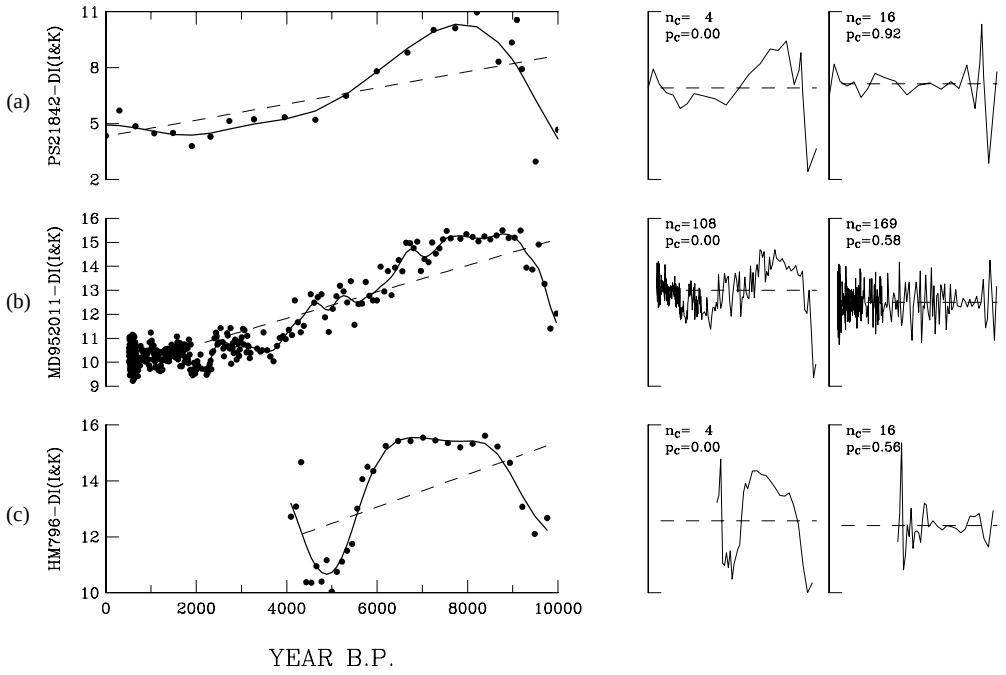

Fig. 8. Records of summer T based on diatom assemblages and results from univariate methods (linear regression and smoothing spline). Leftmost panel: Time series of T (solid circles), linear regression fit (dashed line), and cubic smoothing spline (solid line). The vertical scale represents °C. Second panel from left: Linear regression residuals versus calendar ages. n_C is the number of runs in the sequence of residuals and p_C is the significance level of this number. Third panel from left: Smoothing spline residuals versus calendar ages with the corresponding n_C and p_C values. For the second and third panels from left no labels are reported since the information is in the pattern.

core seems too short to confirm or refute the occurrence of a late cooling (T values for HM796 are available only for the period ~10–4 ka B.P.; Fig. 8c).

For a more rigorous comparison between the two proxies we consider the single core (MD952011, off Norway) for which both biomarker and floral paleotemperatures are available (Fig. 9). Both proxies document warming in the early Holocene and long-term cooling. The amplitude of the temperature changes, however, are much larger in the floral T record than in the alkenone T record. In addition, the timing of the early warming and the structure of the late cooling, differ between the two records. In spite of these differences, the first-order features in the biomarker and diatom paleotemperature records could be regarded as qualitatively consistent.

5.1.4. Planktonic $\delta^{18}O$ records

The $\delta^{18}O$ of planktonic foraminifers has been used as a proxy of T during the Holocene in the western subtropical North Atlantic (e.g., Keigwin (1996)). Bond et al. (1997), on the other hand, concluded that at least in the subpolar North Atlantic, planktonic isotopic compositions apparently are not robust indicators of temperature variability during this period. We examine the cores for which both $\delta^{18}O$ values for the planktonic foraminifer *Globigerina bulloides* and biomarker paleotemperatures are available (MD952015, BS7938, BS7933, M39008, and MD952043). The rationale for considering the isotopic changes for *G. bulloides* is that the highest abundances of this species may be coeval with peak coccolith production, thereby minimizing possible seasonal artifacts in the comparison between the two T proxies (see Chapman et al. (1996) and references therein).

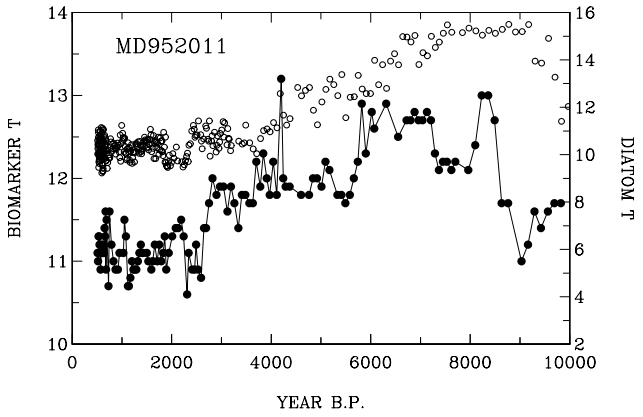

Fig. 9. Comparison between the paleo-T record based on alkenones (solid circles) and the summer paleo-T record based on diatom assemblages (open circles) for core MD9520011. Note that the scales for the two T proxies are different (left scale for biomarker T and right scale for diatom T).

Foraminiferal $\delta^{18}O$ records for the early Holocene should be contaminated by the input to the world ocean of low $\delta^{18}O$ glacial meltwater. We apply a first-order correction for this effect by assuming (1) a constant eustatic sea level rise of 40 m from 10 to 6 ka B.P., as an approximation of the $^{230}Th/^{234}U$ age–depth relationship of shallow water corals (Bard et al., 1996); and (2) that a rise of 10 m corresponds to a change in ocean mean $\delta^{18}O$ of 0.1‰ (Broecker, 1995). The differences between the

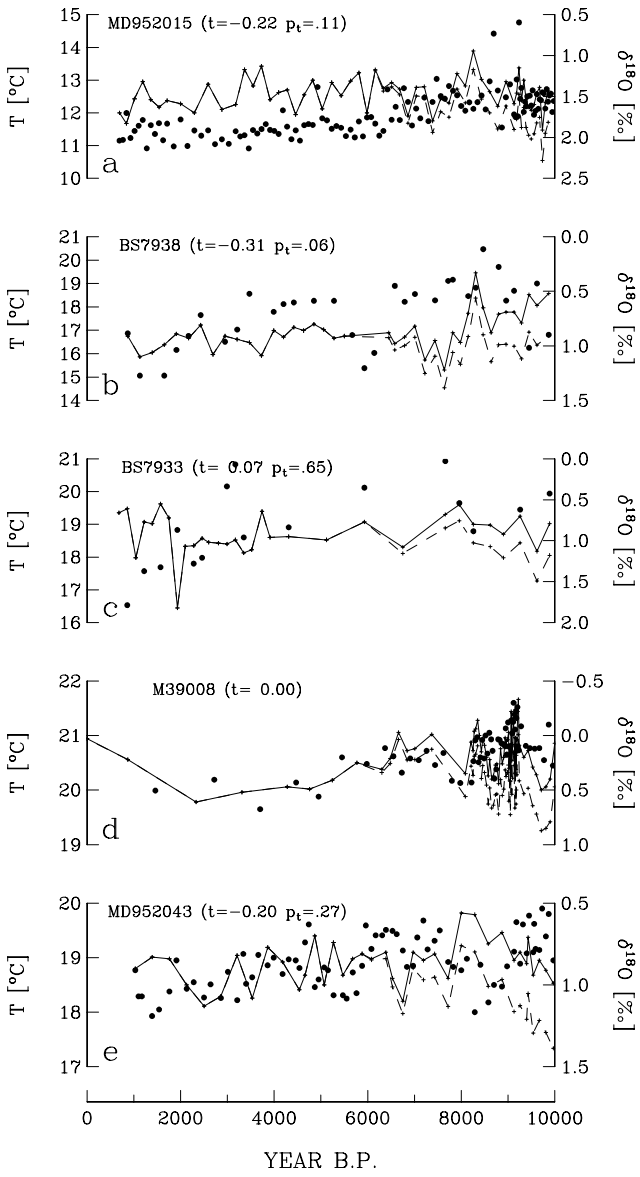

Fig. 10. Comparison between the paleo-T record based on alkenones (solid circles) and the $\delta^{18}O$ of the planktonic foraminifer *Globigerina bulloides* for five cores. Pulses connected by solid lines are $\delta^{18}O$ values uncorrected for the effect of the eustatic sea level rise in the early Holocene. Pulses connected by dashed lines are $\delta^{18}O$ values corrected for this effect (see text for explanations). t is the rank correlation coefficient Kendall tau between the $\delta^{18}O$ values and the calendar ages for the period 6–0 ka B.P. p_t is the significance level of this coefficient (not reported for core M39008 for which $t = 0.00$).

corrected and uncorrected $\delta^{18}O$ records are significant compared to the variability in the raw $\delta^{18}O$ data (compare solid line with dashed line in Fig. 10a–e). Given these differences and the uncertainties in the glacial meltwater correction, we estimate the trend in the uncorrected $\delta^{18}O$ data between 6 and0 ka B.P. We apply the Mann test to the five $\delta^{18}O$ records in order to identify possible monotonic changes in isotopic composition with time during this period. Note that applying the same test to isotopic paleotemperatures would lead to the same results since the test is based on rankings. None of the planktonic $\delta^{18}O$ records shows a monotonic change significant at 5% level between 6 and 0 ka B.P. ($p_t > 0.05$; Fig. 10a–e). If a long-term temperature change did occur, its absence in the planktonic $\delta^{18}O$ records could indicate that this change was masked by other effects. Among these effects are local variations in seawater $\delta^{18}O$, which may be relatively important in marginal seas such as the Mediterranean (where cores BS7938, BS7933, and MD952043 come from). Overall, the comparison between the biomarker and isotopic paleotemperatures is not conclusive.

5.1.5. *Paleo-T records based on foraminiferal counts*

Previously, the most comprehensive review of faunal-based T at 9 and 6 ka B.P. in the North Atlantic (north of 45°N) concluded that the paleoceanographic record was ambiguous (Ruddiman and Mix, 1993). Below, we use the univariate and multivariate methods to identify possible long-term trends in our more recent faunal records of summer paleo-T.

Consider first the results from the three univariate methods (Figs. 11 and 12; Table 2). The occurrence of a constant trend is substantiated for a relatively small number of reconstructions (6 of 16). The reconstructions for which the linear model is substantiated, exhibit an

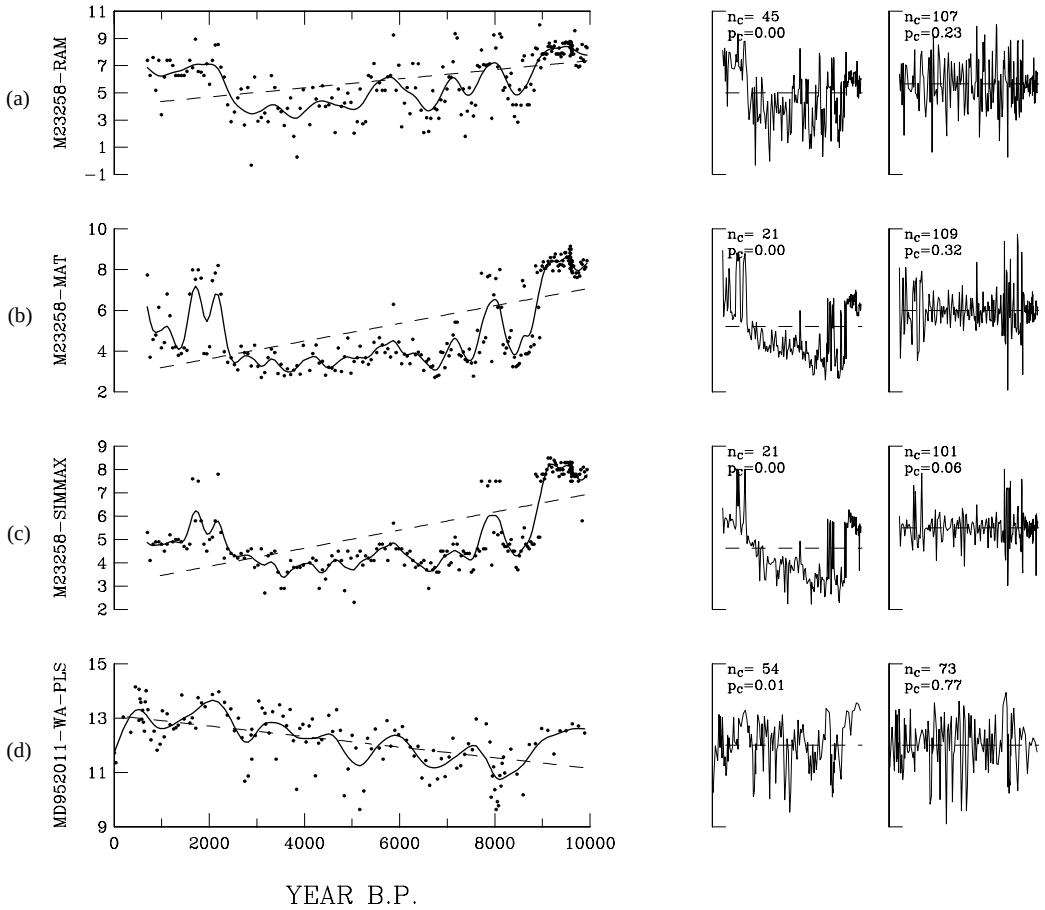

Fig. 11. Records of summer T based on foraminiferal assemblages and results from univariate methods (linear regression and smoothing spline). Leftmost panel: Time series of T (solid circles), linear regression fit (dashed line), and cubic smoothing spline (solid line). The vertical scale represents °C. Second panel from left: Linear regression residuals versus calendar ages. n_C is the number of runs in the sequence of residuals and p_C is the significance level of this number. Third panel from left (for cores for which $p_C > 0.05$): Observed versus expected standardized deletion residuals (open circles) and simulation envelopes (two solid lines). W and p_W are Shapiro–Wilk statistics and its significance level, respectively. Third panel from left (for cores for which $p_C < 0.05$): Smoothing spline residuals versus calendar ages with the corresponding n_C and p_C values. Fourth panel from left (for cores for which $p_C > 0.05$): Ratio $(b - b_{(i)})/|b|$ (in %) versus calendar ages. b is the linear paleotemperature trend and $b_{(i)}$ is the same quantity but when the ith observation is removed from the linear regression. For the second and third panels from left no labels are reported since the information is in the pattern.

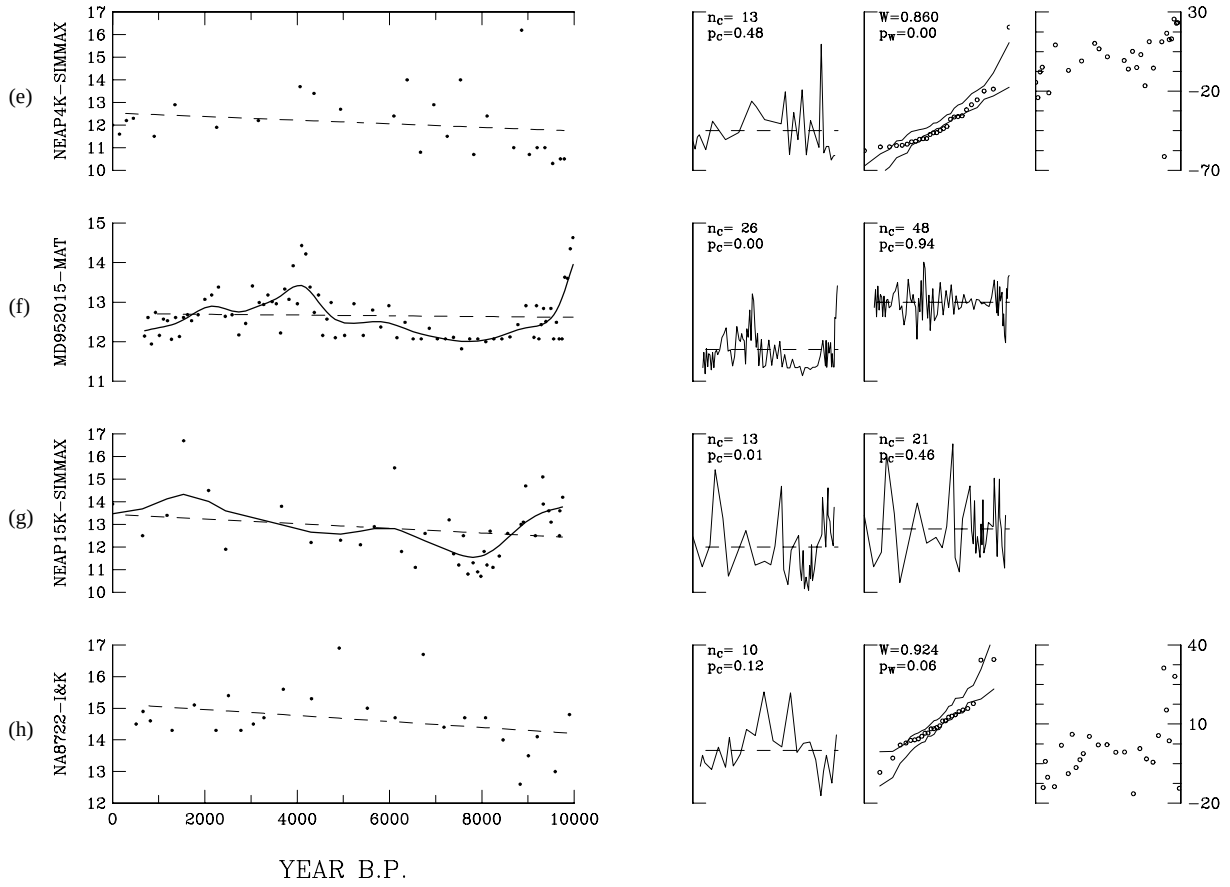

Fig. 11. (*continued*)

apparent warming rate ranging from 0.04 to $0.20^{\circ}C\,kyr^{-1}$ (Table 2). Four of these reconstructions have linear regression residuals that seem to be normally distributed; the apparent warming rate $\pm$ two standard errors is $0.09\pm0.12^{\circ}C\,kyr^{-1}$ ($0.07\pm0.07^{\circ}C\,kyr^{-1}$) in the I&K (RAM) reconstruction for NA8722 and $0.20\pm0.26^{\circ}C\,kyr^{-1}$ ($0.04\pm0.11^{\circ}C\,kyr^{-1}$) in the I&K (RAM) reconstruction for SU8118 (Table 2). None of the apparent warming in these records is therefore significant at the 5% level, which is confirmed by the Mann test (Table 2). The selected splines for the three M23258 records reveal a complex evolution of the overall trend during the Holocene, with possibly periodic fluctuations (Figs. 11a–c). The faunal reconstruction for MD952011 shows apparent warming and one of the most apparent periodicity among the paleo-T records examined in this paper (Fig. 11d). Smoothing all the other records showing indication of a time-variable trend, evidences apparent cooling in the early Holocene followed by apparent warming (Figs. 11e–f and g–j).

Consider then the results from the multivariate method. A PCA is conducted on faunal reconstructions in an attempt to detect a prominent size component similar to that identified for the alkenones. Several reconstructions based on different transfer-function techniques are available for some cores (Table 1). Since our "null hypothesis" is that the faunal T exhibit a size component corresponding to long-term cooling, we favour the hypothesis by incorporating into the PCA the records showing the "strongest" indication of such a cooling. If the PCA does not reveal a prominent size component, it would a fortiori not be the case if the other records for the same core were used in the analysis. Ten records are used in the PCA (Table l), with $n_{\min} = 15$ and $\overline{\Delta x} = 647$ yr. The analysis of **S** shows that the difference between the variance of the first component (56% of tr(**S**)) and second component (23%) is smaller than for the alkenones. That is, the first component stands up less prominently in the scree graph for the faunal T (solid circles in Fig. 6b). More important, there is an alternance of positive and negative elements in the first and second eigenvectors (Table 4). Thus, none of the two leading components is a size component. The time series of the first component is different from the one for the alkenone reconstructions (compare solid circles in Fig. 7c and a). The second component shows indication of a long-term increase (solid circles in Fig. 7d). The results above are not altered when the components are extracted from **R**

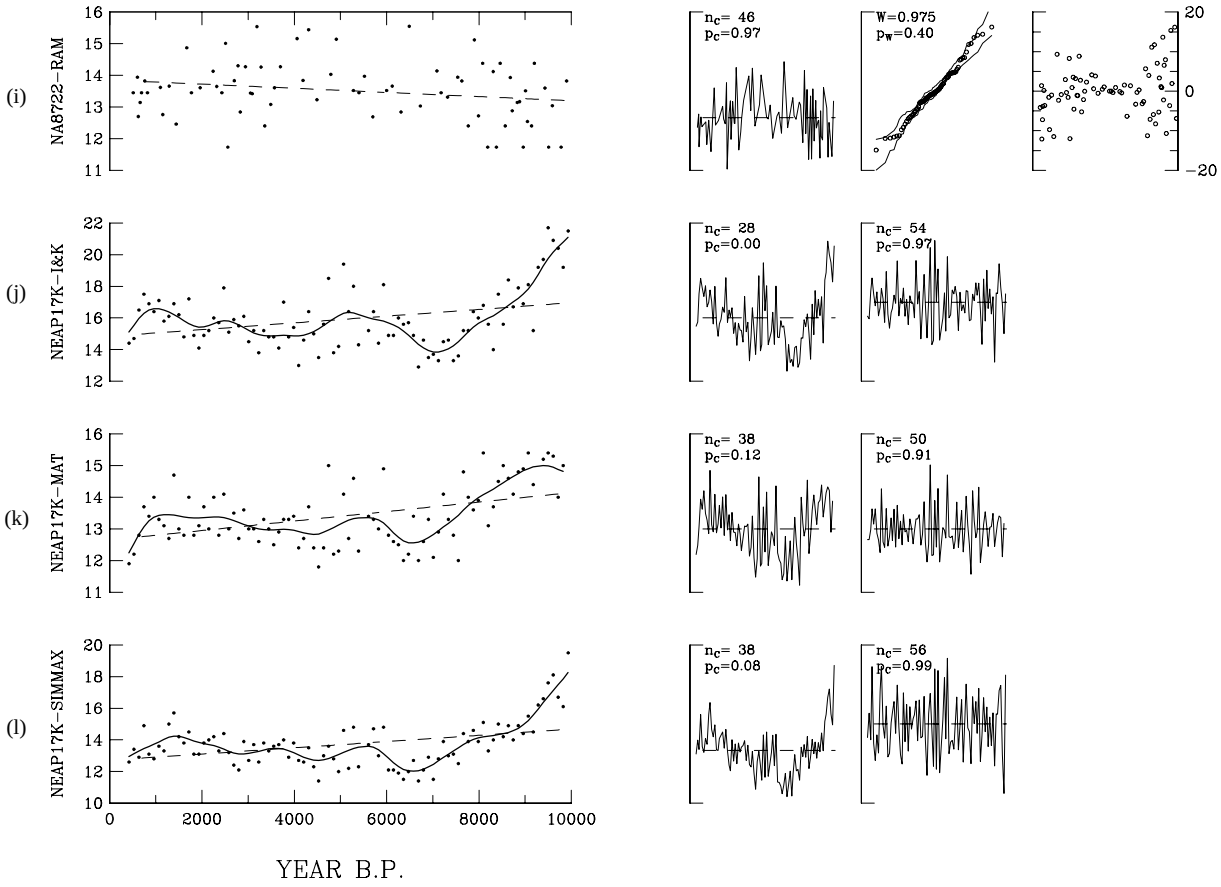

Fig. 11. (*continued*)

rather than from **S** (compare open circles with solid circles in Fig. 6b and 7c–d).

In summary, the application of the univariate and multivariate methods to our faunal T records leads to results different from those obtained for the alkenone reconstructions. The warming that is suggested in many faunal records is not significant for any of the records showing indication of a constant T trend (except for CH6909; Fig. 12). When considered in a multivariate setting, the faunal records show a covariance structure which is more complex than for the biomarker records. A possible cause of the apparent discordance between the biomarker and faunal paleotemperatures is that individual foraminiferal species could indicate a cooling trend, that would be masked by the "noise" present in the record of the other species used in transfer-function techniques. The species *Neogloboquadrina pachyderma* (left coiling) dominates planktonic foraminiferal assemblages in the modern northern North Atlantic (Prell et al., 1999). The percentage of this species in the total foraminifer assemblage was regarded as a qualitative indicator of surface temperature in the subpolar North Atlantic during the Holocene (Bond et al., 1997). To address the point above we thus consider the cores for which (1) the relative abundances of *N. pachyderma* (1) are sufficiently high; and (2) the faunal and biomarker paleotemperatures are in apparent contradiction. The cold-loving *N. pachyderma* (1) generally dominates the foraminiferal community in the Holocene section of core M23258 (Fig. 13a). In core MD952011, it averages 18% for the last 10 kyr (Fig. 13b). For both cores the evolution of *N. pachyderma* (1) over the Holocene mirrors that of the summer faunal paleotemperatures (for M23258 see Fig. 11a–c; for MD952011 see Fig. 11d). We find therefore no indication, at least in the northernmost cores, of a cooling in the records of foraminiferal counts, that would be masked by the transfer-function techniques.

5.2. Others contributions

The discordance between the biomarker and faunal paleo-T records leads us to consider other possible contributions to the apparent cooling that is consistently observed in the former records. Different contributions are discussed below, on the basis of the data in hand.

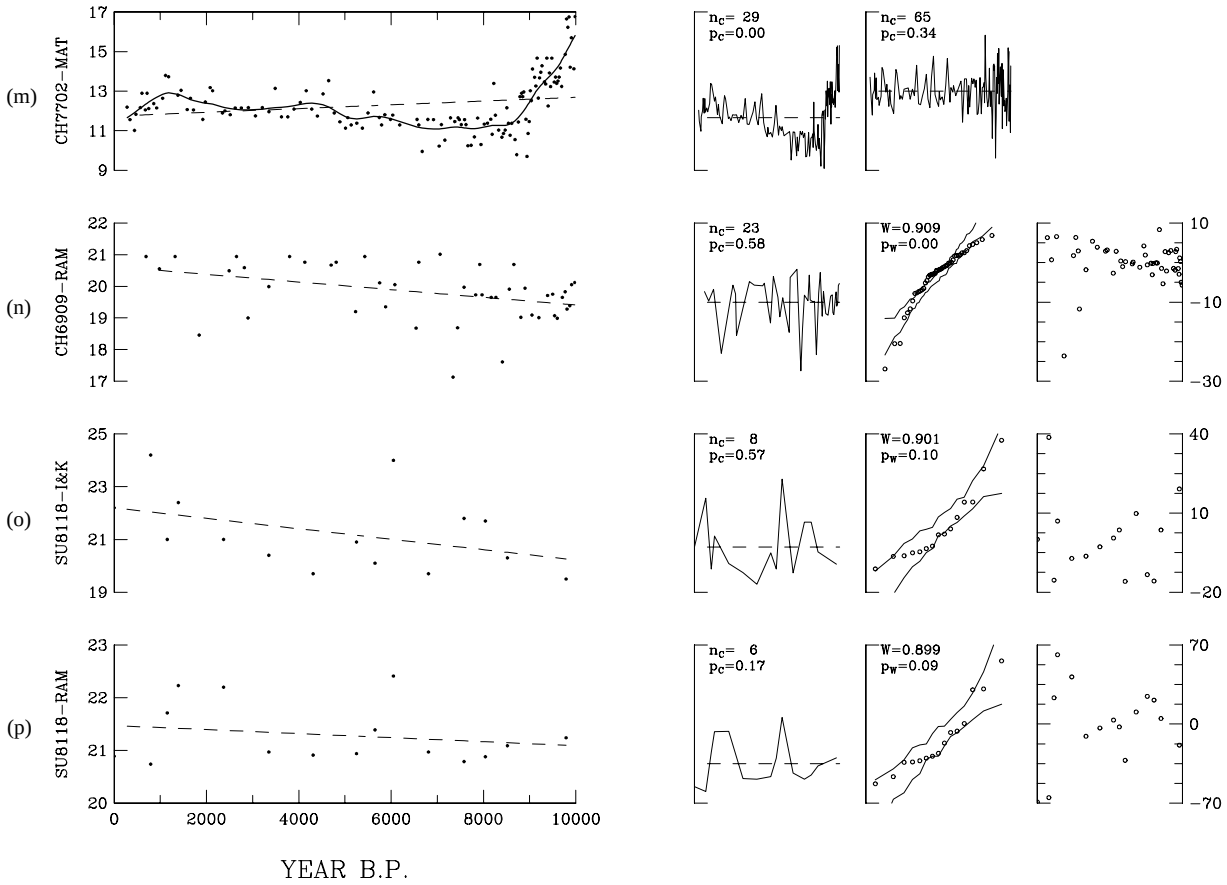

Fig. 11. (*continued*)

5.2.1. Changes in seasonal timing and/or duration of the growth period of prymnesiophyte algae

Another cause of the above discordance is that alkenones and foraminiferal tests may not be produced at the same season and/or at the same depth in the water column. The importance of these possible seasonal and depth artifacts should be linked to local hydrological and ecological factors and varies likely with core location and possibly also with time. A possible strategy to explore these artifacts is therefore to examine in detail biomarker and faunal paleotemperatures from the same core. Previous studies used this strategy to show an apparent discordance between the two T proxies in the Holocene section of North Atlantic cores (Zhao et al., 1995; Chapman et al., 1996; Sikes and Keigwin, 1996; deMenocal et al., 2000). We consider our three cores for which both biomarker and faunal paleotemperature records have been produced (Fig. 14a–c). For core M23258 (Barents slope) the biomarker paleo-T (calibrated to summer T) are higher during most of the Holocene than the faunal paleo-T for summer obtained from the three different transfer-function techniques (compare solid circles with solid lines in Fig. 14a). The biomarker paleotemperatures for cores MD952011 (culture calibration) and MD952015 (calibration to annual mean T) tend to diverge from the summer faunal paleo-T during the Holocene (Fig. 14b–c). We note that the differences between the biomarker and faunal paleo-T records in the three cores are not particularly large given the errors in these records reported in the literature. For core MD952015 for example, the error in biomarker T (from a global core-top calibration) should range from $\sim$0.5°C to 1.5°C (Section 2.2). On the other hand, the error in summer T (obtained from the MAT) would amount to 1.1°C (Waelbroeck et al., 1998). Assuming that the errors in the biomarker and faunal paleotemperatures are uncorrelated, the error in the difference between the two T proxies would be between $\sqrt{0.5^2 + 1.1^2} = 1.2$°C and $\sqrt{1.5^2 + 1.1^2} = 1.9$°C. Some differences between the biomarker and faunal T (as approximated by the spline) in the late Holocene are larger than the lower limit of 1.2°C, whereas the maximum differences are comparable to the upper limit of 1.9°C (Fig. 14c).

The divergences between the biomarker and faunal paleo-T in the three cores are non-random, however, suggesting the possibility of a systematic difference between the two proxy records. Chapman et al. (1996) produced two paleotemperature records based on the $U^{K'}_{37}$ index and foraminifer assemblage

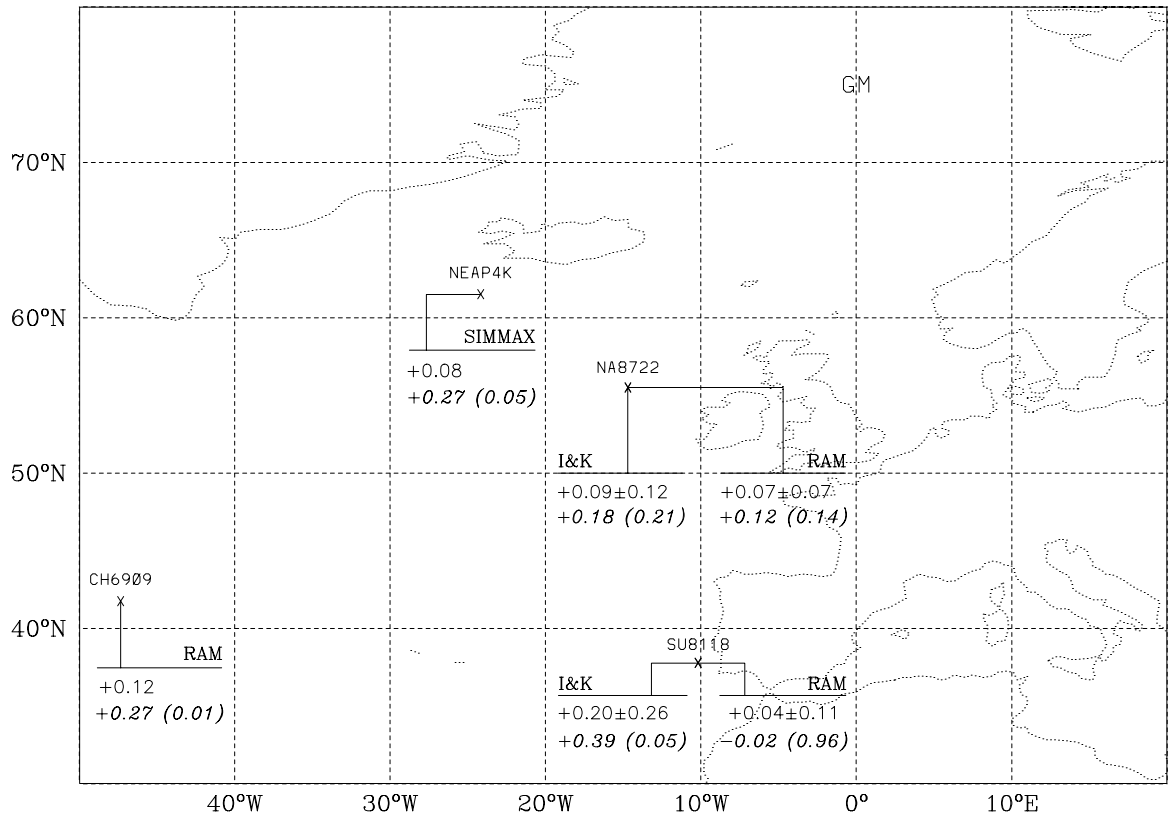

Fig. 12. Results from the univariate methods (linear regression and Mann test) for summer paleo-T based on foraminiferal assemblages. The non-italicized values are the linear trend (b) and its 95% confidence interval (both in $^{\circ}C\,kyr^{-1}$). The italicized values are the rank correlation coefficient Kendall tau (t) and its significance level (p_t; between parentheses).

for the same core from the eastern subtropical Atlantic. They noted that the faunal summer T diverge by 3°C between 8–6 ka B.P., an offset which persists through the late Holocene. They ascribed the offset to a shift in the seasonal timing of maximum coccolith production from summer in the glacial ocean to late spring–early summer in the modern ocean. Likewise, it is conceivable that the divergence between the biomarker and faunal paleotemperatures observed in MD952011 (Fig. 14b) and MD952015 (Fig. 14c) reflects the fact that coccolithoporids flourished later in the year and/or more abundantly during the warmest months in the early Holocene than in the late Holocene. Alternatively the possibility that significant changes in the seasonal timing and/or depth habitat of foraminifer species occurred during this period cannot be ruled out. Clearly, if both biomarker and fauna1 paleo-T are accurate, our data suggest that the seasonal and/or depth artifacts and their possible changes with time strongly compromise the comparability between the two T proxies.

5.2.2. *Changes in the composition of the prymnesiophyte community*

We examine the possibility that the apparent cooling in the biomarker paleo-T records reflects changes in the composition of the coccolithophorid community. Indeed culture experiments showed that different genera, species, and even strains of alkenone-producing organisms can be characterized by different $U^{K'}_{37}$ index at the same temperature (for reviews see Brassell (1993); Müller et al. (1998)). We consider the two cores for which both a biomarker paleo-T record and percentages of dominant species in the total coccolith assemblage are available (Fig. 15a–b). Drastic changes in the proportion of two dominant species are found in MD952015: the relative abundance of *Emiliania huxleyi* rised by ~40% from ~6.5 to 3.4 ka B.P., whereas the proportion of *Gephyrocapsa muellerae* dropped by ~50% over this period (Giraudeau et al., 2000). None of these changes is clearly reflected in the biomarker paleotemperature record (compare pluses and open circles with solid circles in Fig. 15a). Similarly in BS7933 there is no clear relationship between the changes in *E. huxleyi* and *Gephyrocapsa* spp. on the one hand and the changes in biomarker paleo-T on the other hand (Fig. 15b) (Sbaffi et al., 2001). These results are consistent with previous studies that showed that changes in the composition of the coccolithophore population do not dominate marine sediment records of the $U^{K'}_{37}$ index (Müller et al. (1997); Herbert et al. (1998); Villanueva et al., "A detailed comparison of the $U^{K'}_{37}$ and coccolithophorid records

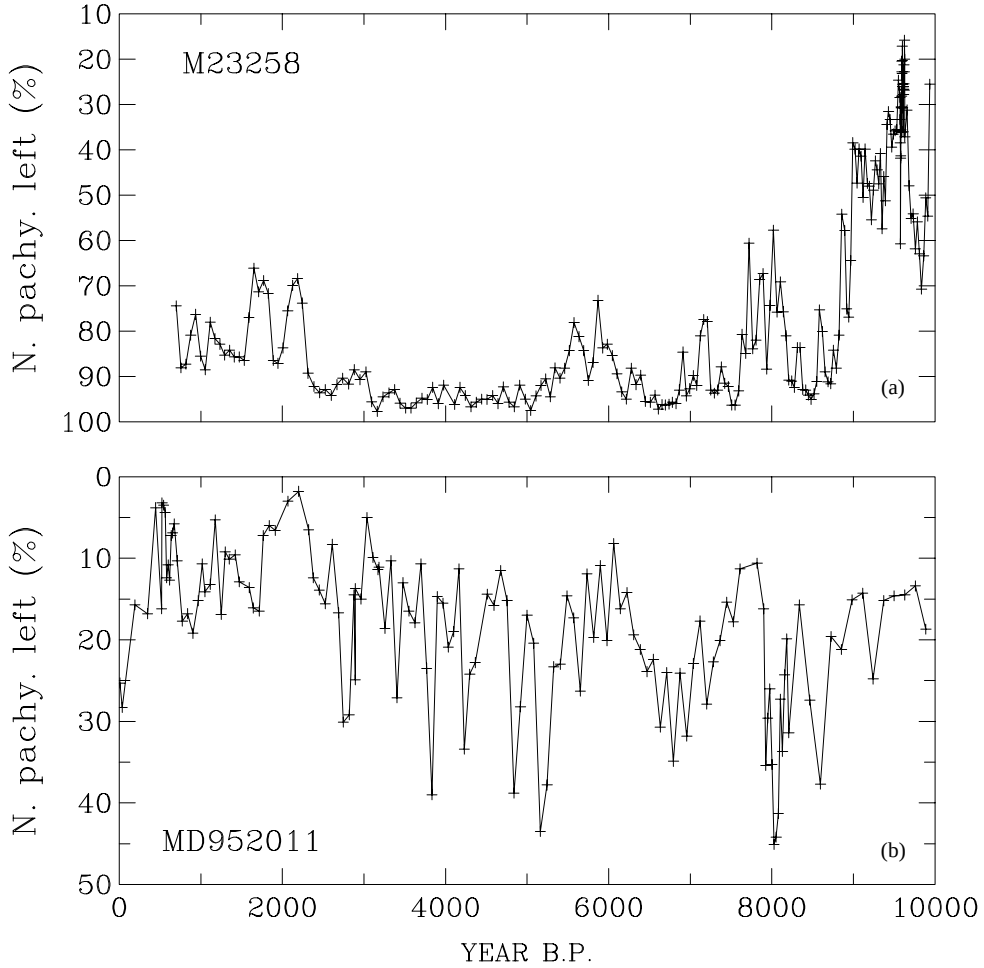

Fig. 13. Relative abundance of *Neogloboquadrina pachyderma* (left coiling) in total foraminiferal assemblage in cores M23258 and MD952011.

over the last 290 kyr: Implications for the alkenone paleotemperature record", submitted).

5.2.3. Reworking of alkenones

Finally, we address the possible effect of alkenone reworking in the sediment. One major difficulty in Holocene paleoceanographic reconstructions is to obtain core top sections with sufficiently high accumulation, and some of our cores come from drift deposits for which the above effect may be particularly important (e.g., M23258, MD952011, and M39008). We consider core M39008 (Gulf of Cadiz) to explore the potential impact of alkenone reworking on the biomarker paleo-T record. This core is located in the present pathway of the Mediterranean Outflow Water, which in this area is a relatively strong current and produces a winnowing of the sediments. The sedimentation rates for this core were much larger before $\sim$8 ka B.P. (reaching 323 cm kyr^{-1} in one interval), as reflected in the change of resolution in the biomarker T record (Fig. 3f). It is thus likely that before $\sim$8 ka B.P. winnowing of sediments did not occur and that the current was rather providing some drift material which would have increased the sedimentation rates. To assess the potential effect of the change in sediment regime at $\sim$8 ka B.P., we recalculate the linear trend b and the "monotonic" trend t for core M39008 by excluding the T data between 10 and 8 ka B.P. We find $b = -0.10$°C kyr^{-1} and $t = -0.22$, with the first value significant and the second value insignificant at 5% level. These results do not differ markedly from the ones obtained from the full record ($b = -0.15$°C kyr^{-1} and $t = -0.43$, both significant at 5% level). Thus, at least for core M39008, the probable reworking of alkenones does not largely affect the paleo-T record.

6. Conclusions

Seven records of surface paleotemperature during the Holocene were obtained by the alkenone method from high accumulation sediment cores raised from the Barents slope, off Norway, southwest of Iceland, the

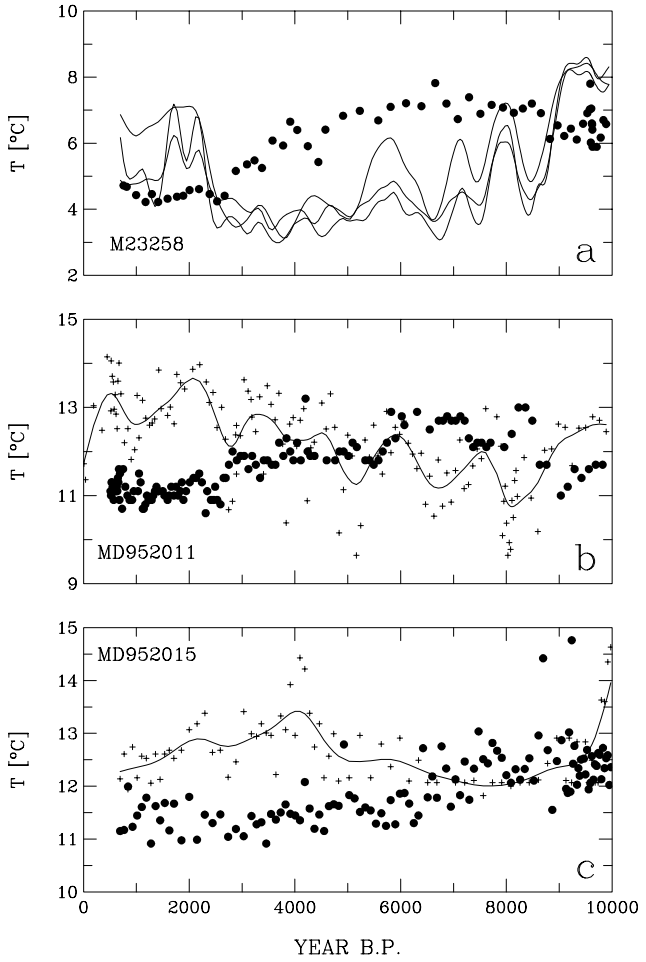

Fig. 14. Comparison between the paleo-T record based on alkenones (solid circles) and the summer paleo-T record based on foraminiferal assemblages (pluses) for cores (a) M23258, (b) MD952011, and (c) MD952015. The solid lines are the splines for the faunal reconstructions. For M23258 the original data are not reported (for a better clarity of the Figure) and the splines for the three summer reconstructions are shown.

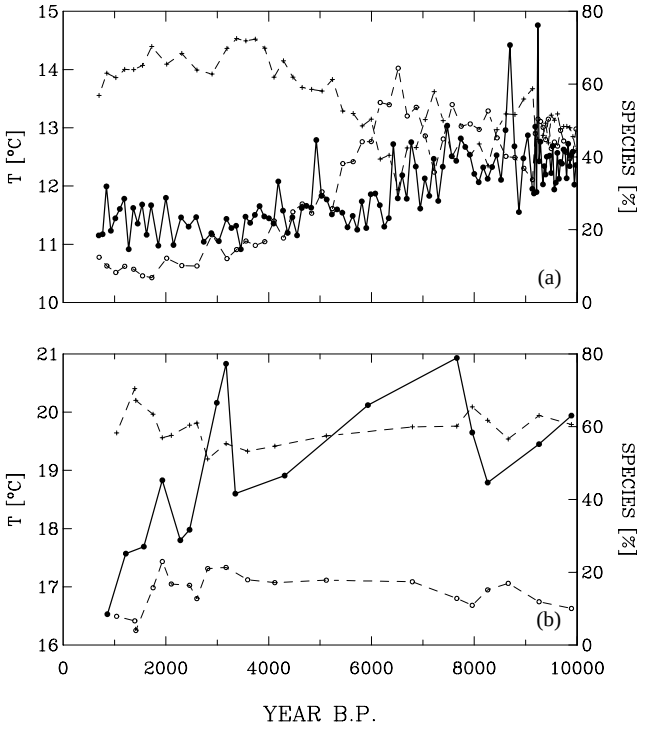

Fig. 15. Comparison between the paleo-T record based on alkenones (solid circles) and the percentages of coccolith species in total coccolith assemblage for cores (a) MD952015 and (b) BS7933. In panel (a) the pluses are for *Emiliania huxleyi* and the open circles for *Gephyrocapsa muellerae* (data from Giraudeau et al. (2000)). In panel (b) the pluses are for *E. huxleyi* and the open circles for *Gephyrocapsa* spp. (*G. muellerae*, *G. ericssoni*, and *G. protohuxleyi*; data from Sbaffi et al. (2001)).

Gulf of Cadiz, the Alboran Sea, and the Tyrrhenian Sea. All these records show an apparent long-term cooling during the Holocene. The apparent cooling occurred at a constant rate in four records, ranging from -0.27 to $-0.15^{\circ}\mathrm{C\,kyr^{-1}}$. The resulting temperature changes are comparable to those inferred from the three paleo-T records showing a time-variable trend (1.2–2.9°C). The biomarker T records are strongly dominated by a single factor (67% of total variance in these records) that is spatially coherent and corresponds to local apparent coolings of 0.5–3.6°C. Two possible contributions to this spatial pattern are a widespread surface cooling and a change in the seasonal timing and/or duration of the growth period of alkenone producers during the Holocene. The first contribution is consistent with many climate proxy records and with climate model simulations including Milankovitch forcing. A recent reconstruction of northern hemisphere temperature shows that the 20th century warming counters a cooling trend during the last millennium (Mann et al., 1999). It could be speculated that this trend is the tail of a long-term cooling which occurred over a much longer portion of the present interglaciation. The second contribution is suggested by the divergence between the biomarker and fauna1 paleo-T in two cores. It is worth mentioning that the apparent discrepancy between biomarker and fauna1 T observed here for the relatively stable Holocene period, is much reduced when the records encompass the last deglaciation (Chapman et al., 1996). More work is necessary, and in particular the above discrepancy must be understood, to better constrain the physical and ecological contributions to the consistent apparent cooling observed in biomarker T records.

Acknowledgements

We thank Richard Norris for information about the genetic diversity in marine plankton, Joël Hirschi, Lloyd

Keigwin, and two anonymous reviewers for comments on the manuscript, and René Fehlmann for discussions on the eigenvalue–eigenvector problem. This study was funded by the European project "Holocene Ocean Instability".

Appendix A. Linear regression

If y is the "dependent variable" (T in our case) and x is the "independent" variable (calendar age) the linear regression model reads

$$y_i = \alpha + \beta \cdot x_i + \varepsilon_i, \tag{A.1}$$

where α and β are the *parameters* of the model, ε is an *error*, and $i = 1, \ldots, n$, where n is the sample size (number of values in a given reconstruction). The goal is to estimate β for each reconstruction.

A.1. Gauss–Markov conditions and normality assumption

We use the ordinary least squares technique to estimate α and β for each reconstruction (the least square estimate of β is noted b hereafter). Statistical theory defines the conditions, called Gauss–Markov (G–M) conditions, for which the prediction made by a least square fitted equation is "good" (e.g., Sen and Srivastasa (1990)),

$$E[\varepsilon_i] = 0 \quad \text{for all } i, \tag{A.2}$$

$$V[\varepsilon_i] = \sigma^2 \quad \text{for all } i, \tag{A.3}$$

$$C[\varepsilon_i, \varepsilon_j] = 0 \quad \text{for all } i \neq j, \tag{A.4}$$

where E[·] is the expectation (mean), V[·] is the variance, and C[·, ·] is the covariance. Condition (A.2) states that the mean of the error ε_i is zero for all i, i.e., that the deviations of y_i from the linear model (A.1) are random. Condition (A.3) states that the variance of the error ε_i is constant for all i, a condition termed homoscedasticity (unequal variance is termed heteroscedasticity). Finally, condition (A.4) states that the covariance between the errors is zero.

The construction of a confidence interval for the regression parameters α and β requires, in addition to the G–M conditions, that the ε_i are normally distributed. If the errors have zero mean and variance σ^2 the normality assumption reads

$$\varepsilon_i \sim N(0, \sigma^2), \tag{A.5}$$

where the symbol "$\sim$"stands here for "has the distribution". Various diagnostics have been proposed in the statistical literature to examine a posteriori whether conditions (A.2)–(A.5) are met for a given linear regression analysis.

A.2. Regression diagnostics

The regression residuals $e_i = y_i - \hat{y}_i$, where $\hat{y}_i$ is the value on the fitted line, are the least square estimates of the regression errors ε_i. The e_i are thus natural diagnostics to assess the quality of the linear model (A.1). Other types of residuals have been proposed as diagnostics for this model, i.e., the standardized residuals and the standardized deletion residuals (e.g., Sen and Srivastasa (1990); Ryan (1997)). We have calculated these two latter types of residuals and found that they do not provide a better insight into the validity of (A.2–A.4) than the usual residuals e_i in our regressions. We thus focus on the statistical properties of the e_i to assess the G–M conditions and on the distribution of the standardized deletion residuals to evaluate the normality assumption.

A.2.1. Randomness

Traditionally "randomness" is loosely defined as a pattern intermediate between "clustering" and "mixing". In this paper randomness is meant to represent also mixing since mixing of residuals (when considered versus calendar ages) can also be taken as a support for the validity of a regression model. Inspection of the e_i is useful to detect a nonrandom pattern in the regression errors ε_i, as well as heteroscedasticity and non-normality of the ε_i. We plot the e_i versus the x_i; a plot versus the $\hat{y}_i$ would reveal the same pattern since the x_i and $\hat{y}_i$ are related through a perfect linear relationship in a simple linear regression setting. We use the runs test to assess whether the regression errors ε_i tend to cluster (Draper and Smith, 1998). The test is based on the number of sign changes in the sequence of e_i when considered versus x_i; this number is the number of runs n_C. A small n_C points to a tendency of the residuals to cluster, whereas a large n_C points to a tendency to mix. If n_C is lower than its expected value in random sequences of the same length n, we suspect a clustering of the residuals. We thus consider the value of n_C (comparatively to n) and the lower-tail p-value (noted p_C) of the corresponding test statistics (a low p_C indicates clustering). A plot of e_i versus x_i, and the values of n_C and p_C are reported in Figs. 3, 8, and 11 (second panel from left).

A.2.2. Homoscedasticity

An association between the absolute values of the $e_i(|e_i|)$ and the $\hat{y}_i$ can indicate whether the variance in the regression errors changes with time, i.e., whether these errors are heteroscedastic (e.g., Madansky (1988); Sen and Srivastasa (1990)). We have found that the plots of $|e_i|$ versus $\hat{y}_i$ do not provide a real insight into the possible violation of G–M condition (A.3). These plots are therefore not included here.

A.2.3. Normality

To examine whether the distribution of regression errors is non-normal we produce normal plots for the standardized deletion residuals $e_{(i)}$ (Sen and Srivastasa, 1990; Ryan, 1997),

$$e_{(i)} = \frac{e_i}{s_{(i)}\sqrt{1-h_i}}, \tag{6}$$

where $s_{(i)}$ is the residual variance $s = (n-2)^{-1}\sum_i (y_i - \hat{y}_i)^2$ that results when the ith observation is not used in the regression, and $h_i = n^{-1} + (x_i - \bar{x})^2 / \sum_i (x_i - \bar{x})^2$ is the leverage for the ith observation, where $\bar{x} = n^{-1}\sum_i x_i \left(\sum_i \equiv \sum_{i=1}^n\right)$. The normal plots show the observed $e_{(i)}$ as a function of the $e_{(i)}$ expected as if the standardized deletion residuals were normally distributed. Thus a linear pattern in such plots indicates normality of the ε_i. To aid the interpretation of normal plots we calculate "simulation envelopes" following the method of Atkinson (1981). These envelopes are estimates of the 5th and 95th percentiles of the distribution of the ith-order statistic $e_{(i)}$, given $N(0, \sigma^2)$ errors. We calculate the Shapiro–Wilk statistics (W, based on the e_i) and its significance level (p_W) to assist further the interpretation of the normal plots (Madansky, 1988; Sen and Srivastasa, 1990). The values of W and p_W are determined from the algorithm described by Royston (1982). Low values for both quantities indicate non normality. A normal plot, and the values of W and p_W are reported for reconstructions showing indication of a random pattern in the regression residuals (third panel from left in Figs. 3 and 11).

A.2.4. Influential observations

Outliers are observations which are characterized by relatively large residuals but which do not have necessarily a large effect in the regression analysis. *Influential observations*, by contrast, are not removed from the regression line, but they have a disproportionate effect on the least square estimate of the regression parameters (e.g. Sen and Srivastasa (1990)). To detect observations that have a large influence on the least square estimate of β we calculate for each observation i the following quantity

$$\frac{b - b_{(i)}}{|b|},$$

where $b_{(i)}$ is the regression coefficient obtained when the ith observation is removed from the analysis ($b - b_{(i)}$ is termed DFBETA in the statistical literature; Sen and Srivastasa (1990); Ryan (1997)). A plot of $(b - b_{(i)})/|b|$ versus x_i is reported for regressions with a random pattern in regression residuals (rightmost panel in Figs. 3 and 11).

Appendix B. Mann test

B.1. Kendall Tau

The Kendall tau (noted t) is a usual nonparametric measure of the association between two random variables (Kendall and Gibbons, 1990; Gibbons and Chakraborti, 1992; Gibbons, 1997). Let $x_1, \ldots, x_n$, be the values of one variable (e.g., calendar ages) and $y_1, \ldots, y_n$ the values of the other variable (T values), ranked in any order. A score a_{ij} is allotted to each pair (i, j) in the x-ranking and a score b_{ij} is allotted to each pair (i, j) in the y-ranking. In the x-ranking, $a_{ij} = +1$ if $x_i < x_j$, $a_{ij} = 0$ if $x_i = x_j$ ("tie"), and $a_{ij} = -1$ if $x_i > x_j$. The same rules are applied for b_{ij} in the y-ranking. If $a_{ij}b_{ij} > 0$ the pair is said to be "concordant", whereas if $a_{ij}b_{ij} < 0$ the pair is "discordant". The Kendall tau is the proportion of concordant pairs minus the proportion of discordant pairs (Kendall and Gibbons, 1990),

$$t \equiv \frac{\sum_{i,j} a_{ij}b_{ij}}{\sqrt{\sum_{i,j} a_{ij}^2 \sum_{i,j} b_{ij}^2}}, \tag{B.1}$$

where a_{ij} and b_{ij} are regarded as zero if $i = j$ and $\sum_{i,j} \equiv \sum_{i=1}^n \sum_{j=1}^n$. It can be shown that $-1 \leqslant t \leqslant 1$, with $t = -1$, 0, and $+1$ indicating, respectively, perfect discordance, equal proportion of concordant and discordant pairs, and perfect concordance between the two variables. The Kendall tau measures therefore the intensity of the *monotone* relationship between two variables. Besides its simple and intuitive interpretation the Kendall tau has the other advantage, compared to other nonparametric measures of monotonicity, that its sampling distribution reaches rapidly normality as the sample size increases, a useful property in significance tests (Kendall and Gibbons, 1990).

To calculate t we use an operational definition equivalent to (7) but which clarifies the treatment of ties in the x-ranking (i.e., identical calendar ages) and/or in the y-ranking (identical T values). The total score S is defined as the number of positive scores minus the number of negative scores, i.e. $S = \frac{1}{2}\sum_{i,j} a_{ij}b_{ij}$ (the fraction $\frac{1}{2}$ arises as the sum $\sum_{i,j}$ is carried out twice for each pair (i, j)). In the absence of ties, any term a_{ij}^2 is unity, so that $\sum_{i,j} a_{ij}^2$ is equal to the number of possible terms, i.e., $\sum_{i,j} a_{ij}^2 = n(n-1)$. The same rule holds for b_{ij}^2 so that $\frac{1}{2}\sqrt{\sum_{i,j} a_{ij}^2 \sum_{i,j} b_{ij}^2}$ is equal to $\frac{1}{2}n(n-1)$. If there is a tie of u consecutive members in the x-ranking, all the scores arising from any pair chosen from them is zero. There are $u(u-1)$ such pairs. Consequently the sum $\sum_{i,j} a_{ij}^2$ is equal to $n(n-1) - \sum u(u-1)$, where the summation is over all sets of u tied scores. We define $U = \frac{1}{2}\sum u(u-1)$ for ties in the x-ranking and $V = \frac{1}{2}\sum v(v-1)$ for ties in the y-ranking. The Kendall tau is

then calculated from (e.g., Kendall and Gibbons (1990))

$$t = \frac{S}{\sqrt{[\frac{1}{2}n(n-1) - U]}\sqrt{[\frac{1}{2}n(n-1) - V]}} \quad \text{(B.2)}$$

B.2. Significance level

The Mann test for a monotone trend is based on the significance level of the Kendall tau. Consider that a random sample has been drawn from a population and that the value of t between two random variables in the sample has been calculated. The significance level of the sample t is then the proportion of values $> |t|$ observed in a population where the two variables are not correlated. We determine this level as follows (Kendall and Gibbons, 1990). Since $S = \frac{1}{2}\sum_{i,j} a_{ij}b_{ij}$ testing the significance of t is equivalent to testing the significance of S. It can be shown that the sampling distribution of S for a population where the total score $= 0$ has the following variance,

$$\begin{aligned} V[S] = {} & \frac{1}{18}\Big[n(n-1)(2n+5) - \sum u(u-1)(2u+5) \\ & - \sum v(v-1)(2v+5)\Big] \\ & + \frac{1}{9n(n-1)(n-2)}\Big[\sum u(u-1)(u-2)\Big] \\ & \times \Big[\sum v(v-1)(v-2)\Big] \\ & + \frac{1}{2n(n-1)}\Big[\sum u(u-1)\Big]\Big[\sum v(v-1)\Big]. \end{aligned} \quad \text{(B.3)}$$

The test statistic $z = S/\sqrt{V[S]}$ approaches the standard normal distribution for large sample sizes. There is probably little important error involved in using the normal approximation for $n \geqslant 10$, unless the ties are very extensive or very numerous (Kendall and Gibbons, 1990).

To compare with the standard normal we apply a continuity correction, i.e. $z = (S+1)/\sqrt{V[S]}$ if $S<0$ and $z = (S-1)/\sqrt{V[S]}$ if $S>0$. The significance level of S (and thus of t) is twice the left-tail probability of z (if $z<0$) or right-tail probability of z (if $z>0$) as determined from the standard normal. A low value of this level indicates association between the two variables (i.e., a T trend in our case).

Appendix C. Smoothing spline

A model for which the smoothing spline is applicable is (e.g., Hutchinson and De Hoog (1985))

$$y_i = f(x_i) + \varepsilon_i, \quad \text{(C.1)}$$

where f is a suitable smooth, but unknown, function ("signal"), and the ε_i are random errors satisfying (A.2–A.4) ("noise"). It is assumed that f is smooth in the sense that, for some positive integer m, f admits $m-1$ continuous derivatives and has a square integrable mth derivative. Under these restrictions a natural estimator is the function $\hat{f}$ which minimizes

$$\sum_i^n [y_i - f(x_i)]^2 + \rho \int_{t_0}^{t_1} \left[\frac{\mathrm{d}^m f}{\mathrm{d}x^m}\right]^2 \mathrm{d}x.$$

The solution to this problem is a polynomial spline of order $2m$ with knots at the x_i that is usually referred to as a smoothing spline (Eubank, 1988). The parameter ρ governs the balance between fidelity to the data (as measured by $\sum_i [y_i - f(x_i)]^2$, and smoothness (as gauged by $\int_{t_0}^{t_1} \left[\frac{\mathrm{d}^m f}{\mathrm{d}x^m}\right]^2 dx$. If $\rho \to 0$ the smoothing approaches an interpolating spline, which is of no interest in our case. If $\rho \to \infty$ it approaches the least squares line.

The values of m and ρ must be prescribed for each smoothing spline. We use for all reconstructions $m = 2$, a typical choice, which produces *cubic* smoothing splines. Different approaches are available to constrain the smoothing parameter ρ (Eubank, 1988). In this paper we try various values of ρ and select a posteriori a "good" value based on diagnostics suggested for smoothing splines.

C.1. Regression diagnostics

The diagnostic methodology for smoothing splines is at a much less advanced level of development than for the linear model (1) (Eubank, 1988). Residuals, standardized residuals, and standardized deletion residuals were proposed to assess the quality of a smoothing spline (e.g., Eubank (1984, 1986)). We focus below on the residuals e_i. A runs test is performed on the e_i to assess whether the ε_i associated with the spline are random (i.e., whether they do not tend to cluster; Eubank (1988)). A plot of e_i versus x_i, and the corresponding values of n_C and p_C are reported for reconstructions with a nonrandom pattern in the linear regression residuals (rightmost panel in Figs. 3, 8, and 11).

Appendix D. Principal component analysis

The PCA is a standard procedure to reduce the dimensionality of a multivariate data set by a few combinations, the *principal components*. We give below a brief description of the PCA (for details see, e.g., Jolliffe (1986); Preisendorfer (1988)). In the following q is the number of *variables* measured on a number n of *subjects*.

D.1. Temperature values used in PCA

The number and the values of the calendar ages from different reconstructions, i.e., the number of *subjects* and the *subjects* themselves, are not the same (this would occur exceptionally when considering different marine sediment records). The application of PCA to the original reconstructions is therefore not possible.

We use the following approach to apply the PCA to a number q of T reconstructions. We consider that the characteristic time scale of covariation that can be detected by the PCA, is limited by the core with minimum sampling frequency. We note $n_{\min}$ the number of T values for this core. The youngest and oldest calendar ages among the q reconstructions are identified to find the time period (between 10 and 0 ka B.P.) for which T values are available in all reconstructions; these ages are noted $x_{\min}$ and $x_{\max}$, respectively. We then calculate a mean sampling interval $\overline{\Delta x} = (x_{\max} - x_{\min})/(n_{\min} - 1)$. This allows us to define a number $n_{\min}$ of time intervals between $x_{\min} - (\overline{\Delta x}/2)$ and $x_{\max} + (\overline{\Delta x}/2)$, the boundaries between the intervals being calculated as $x_j = (x_{\min} - \overline{\Delta x}/2) + (j-1)\overline{\Delta x}$ with $j = 1, \ldots, n_{\min} + 1$. The T values used in the PCA are then the mean T in each time interval.

We have to deal with the possible situation of missing T values in some time intervals. A general and flexible method for handling missing-data problems in multivariate analysis is *imputation,* i.e., the prediction of missing values on the basis of available observations (Little and Rubin, 1987). We use a particularly simple form of imputation by estimating the missing value(s) in a reconstruction from the mean of available values for that reconstruction; this mean is termed the "unconditional mean" (Little and Rubin, 1987). Using this form in PCA was suggested by Jolliffe (1986). The variances and covariances between variables are underestimated when imputing unconditional means (Little and Rubin, 1987). We keep this aspect in mind when interpreting the outputs from the PCA (Appendix E.1).

D.2. Extraction of principal components

The principal components are obtained from the covariance matrix $\mathbf{S}$ for a number q of T reconstructions. $\mathbf{S}$ can be expressed in terms of the $q \times 1$ observations vectors $\boldsymbol{y_1}, \ldots, \boldsymbol{y_n}$ and the $q \times 1$ mean vector $\bar{\boldsymbol{y}}$, i.e., $\mathbf{S} = (n-1)^{-1} \sum_i^n (\boldsymbol{y_i} - \bar{\boldsymbol{y}})(\boldsymbol{y_i} - \bar{\boldsymbol{y}})^{\mathrm{T}}$, where "T" denotes the transpose. The diagonal elements of $\mathbf{S}$ are the variances of the reconstructions; the total variance in the data set is thus the sum of these elements, i.e., the trace of $\mathbf{S}$, noted tr($\mathbf{S}$). The off-diagonal elements of $\mathbf{S}$ are the covariances between the reconstructions. $\mathbf{S}$ is thus a symmetric, positive definite matrix of size $q \times q$. To obtain the principal components we solve the eigenvalue–eigenvector problem for $\mathbf{S}$,

$$\mathbf{Se} = \lambda \mathbf{e}, \tag{D.1}$$

where $\mathbf{e}$ is an eigenvector and λ is the corresponding eigenvalue. The solution of (D.1) yields the q eigenvectors and the q eigenvalues of $\mathbf{S}$. The sum of eigenvalues is equal to the total variance in the data set, i.e. $\sum_{k=1}^{q} \lambda_k = \mathrm{tr}(\mathbf{S})$. Since $\mathbf{S}$ is symmetric its eigenvectors are mutually orthogonal, i.e., they portray uncorrelated "modes of variability" in the data set. Since $\mathbf{S}$ is positive definite all its eigenvalues are positive.

The principal component z_{ij} for the ith time interval and jth mode is given by projecting the centred observation vector $\boldsymbol{y_i} - \bar{\boldsymbol{y}}$ onto the eigenvector $\mathbf{e}_j$,

$$z_{ij} = \mathbf{e}_j^{\mathrm{T}}(\boldsymbol{y_i} - \bar{\boldsymbol{y}}) = \sum_{k=1}^{q} \mathrm{e}_{jk}(\boldsymbol{y_{ik}} - \bar{\boldsymbol{y}}_{\boldsymbol{k}}). \tag{D.2}$$

The latter equation was termed the "analysis" of y_{ik} (Preisendorfer, 1988). The time series $\{z_{ij}\}$ with $i = 1, \ldots, n$ depicts the evolution of the jth mode of variability in the data set. We follow a common practice by normalizing all eigenvectors so that they have unit length, i.e., $\mathbf{e}_j^{\mathrm{T}}\mathbf{e}_j = 1$, with $j = 1, \ldots, q$. In this case the variance of the jth principal component is equal to the jth eigenvalue of $\mathbf{S}$. To find a common mode of variability between several reconstructions we will examine the leading principal components, i.e., the components with the largest variances.

Alternately, the original (centred variables) can be exactly represented in the form

$$y_{ik} - \bar{y}_k = \sum_{j=1}^{q} e_{jk} z_{ij}. \tag{D.3}$$

The latter equation was termed the "synthesis" of y_{ik} (Preisendorfer, 1988). Here we use this equation to quantify T changes at a given location associated with the leading mode of variability common to all reconstructions.

D.3. Physical interpretation and inference in PCA

It is noteworthy that principal components are statistical constructions, which may not correspond to a distinct process in a physical system. The reason is that principal components depict uncorrelated modes of variability in a multivariate data set, whereas dominant processes in a physical system are likely to be correlated. In technical terms interpreting principal components can be delicate owing to the constraint of orthogonality between the eigenvectors from the same covariance matrix (von Storch, 1995).

On the other hand, formal inference procedures in PCA rely on the independence of the observation vectors $\boldsymbol{y_1}, \ldots, \boldsymbol{y_n}$, as well as on multivariate normality.

They cannot therefore be used if more than very weak dependence is present between $y_1, \ldots, y_n$. Such a dependence is likely to occur in our case, where the observation vectors correspond to spatial fields at different times. Multivariate normality, however, is not required for estimation of principal components. As our main objective with the PCA is descriptive, not inferential, complications such as non-independence should not seriously affect this objective (Jolliffe, 1986).

Appendix E. Robustness of PCA results

E.1. Effect of unconditional means

A substantial portion of mean T values for cores BS7933 and M39008 are unconditional means (Fig. 5). We thus explore the sensitivity of our major results for the alkenones to the omission in the PCA of data from these cores. When data from BS7933 are omitted the first and second principal components contribute to 62% and 23% to the total variance in the paleo-T records, respectively. When data from M39008 are removed these percentages amount to 70% and 15%, respectively. When data from both M39008 and BS7933 are discounted they amount to 65% and 24%, respectively. In the three cases the first component exhibits a long-term decrease during the Holocene (not shown here). In the three cases all elements of the first eigenvector have the same sign, in contrast to elements of the second eigenvector. We thus conclude that the major results from the PCA are robust against the inclusion or omission of data from cores for which unconditional means are imputed.

E.2. Extraction from correlation matrix

Principal components may be extracted from the correlation matrix rather than from the covariance matrix. Extraction of the components from the correlation matrix (or equivalently from the covariance matrix for the standardized variables) is justified when the measurement units of the variables are not commensurate or when the variance of one variable is particularly large (in the latter case the first component essentially replicates this variable; Rencher (1995)).

In our application of the PCA all variables have the same physical unit (°C), so that extracting the principal components from **S** is justified. However, inspection of **S** for the alkenones shows that the variances differ widely among the alkenone reconstructions (Table 3), which may bias the leading component(s) towards the reconstruction(s) with the largest variance(s). We find that most of the total variation in **R** (74%) is summarized by the two leading components (open circles in Fig. 6a). The first component of **R** exhibits an evolution over the Holocene, which is similar to that of the first component of **S** (compare open circles with solid circles in Fig. 7a). The second component of **R** does not show a clear long-term trend during this period (open circles in Fig. 7b). Thus the extraction of the principal components from **S** and **R** leads essentially to the same results.

References

Atkinson, A.C., 1981. Two graphical displays for outlying and influential observations in regression. Biometrika 68, 13–20.

Bard, E., Hamelin, B., Arnold, M., Montaggioni, L., Cabioch, G., Faure, G., Rougerie, F., 1996. Deglacial sea-level record from Tahiti corals and the timing of global meltwater discharge. Nature 382, 241–244.

Berger, A., 1978. Long-term variations of daily insolation and Quaternary climatic changes. Journal of Atmospheric Sciences 35, 2362–2367.

Bhattacharyya, G.K., 1984. Tests of randomness against trend or serial correlation. Vol. 4. Elsevier Science Publishers, North-Holland, Amsterdam, pp. 89–111.

Bond, G., Showers, W., Cheseby, M., Lotti, R., Almasi, P., DeMenocal, P., Priore, P., Cullen, H., Hajdas, I., Bonani, G., 1997. A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. Science 278, 1257–1266.

Brassell, S.C., 1993. Applications of biomarkers for delineating marine paleoclimatic fluctuations during the Pleistocene. In: Engel, M.H., Macko, S.A. (Eds.), Organic Geochemistry. Plenum, New York, pp. 699–738.

Broecker, W.S., 1995. The glacial World According to Wally. Lamont–Doherty Geological Observatory, Palisades. Eldigio Press, New York.

Broecker, W.S., 1998. The end of the present interglaciation: How and When? Quaternary Science Reviews 17, 689–694.

Cacho, I., Grimalt, J.O., Canals, M., Sbaffi, L., Shackleton, N.J., Schönfeld, J., Zahn, R., 2001. Variability of the Western Mediterranean Sea surface temperatures during the last 25,000 years and its connection with the northern hemisphere climatic changes. Paleoceanography 16, 40–52.

Cacho, I., Grimalt, J.O., Pelejero, C., Canals, M., Sierro, F.J., Flores, J.A., Shackleton, N.J., 1999. Dansgaard–Oeschger and Heinrich event imprints in Alboran Sea paleotemperatures. Paleoceanography 14, 698–705.

Chapman, M., Shackleton, N.J., Zhao, M., Eglinton, G., 1996. Faunal and alkenone reconstructions of subtropical North Atlantic surface hydrography and paleotemperature over the last 28 kyr. Paleoceanography 11, 343–357.

Cheddadi, R., Lamb, H.F., Guiot, J., van der Kaars, S., 1998. Holocene climate change in Morocco: a quantitative reconstruction from pollen data. Climate Dynamics 14, 883–890.

COHMAP, 1988. Climatic changes of the last 18,000 years: observations and model simulations. Science 241, 1043–1052.

Conte, M.H., Eglinton, G., Madureira, L.A.S., 1995. Origin and fate of organic biomarker compounds in the water column and sediments of the eastern North Atlantic. Philosophical Transactions of the Royal Society, London. B 348, 169–178.

Crucifix, M., Loutre, M.-F., Tulkens, P., Fichefet, T., Berger, A., 2001. Climate evolution during the Holocene: A study with an Earth system model of intermediate complexity, Climate Dynamics, in press.

Dahl, S.O., Nesje, A., 1996. A new approach to calculating Holocene winter precipitation by combining glacier equilibrium line altitudes

and pine-tree limits: a case study from Hardangerjøkulen central south Norway. The Holocene 6, 381–398.
Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G.D., Johnsen, S.J., Hansen, A.W., Balling, N., 1998. Past temperatures directly from the Greenland ice sheet. Science 282, 268–271.
Dansgaard, W., Johnsen, S.J., Clausen, H.B., Dahl-Jensen, D., Gundestrup, N.S., Hammer, C.U., Hvidberg, C.S., Steffensen, J.P., Sveinbjornsdottir, A.E., Jouzel, J., Bond, G., 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. Nature 364, 218–220.
deMenocal, P., Ortiz, J., Guilderson, T., Sarnthein, M., 2000. Coherent high- and low-latitude climate variability during the Holocene warm period. Science 288, 2198–2202.
Denton, G.H., Karlen, W., 1973. Holocene climatic variations—their pattern and possible cause. Quaternary Research 3, 155–205.
Draper, N.R., Smith, H., 1998. Applied Regression Analysis. Wiley Series in Probability and Statistics. Wiley, New York, 706pp.
Duplessy, J.-C., Labeyrie, L., Arnold, M., Paterne, M., Duprat, J., Van Weering, T.C.E., 1992. Changes in surface salinity of the North Atlantic Ocean during the last deglaciation. Nature 358, 485–488.
Eubank, R.L., 1984. The hat matrix for smoothing splines. Statistics and Probability Letters 2, 9–14.
Eubank, R.L., 1986. Diagnostics for smoothing splines. Journal of the Royal Statistical Society, B 47 (2), 332–341.
Eubank, R.L., 1988. Spline Smoothing and Nonparametric Regression, Vol. 90, Statistics: Text-books and Monographs. Marcel Dekker, New York, 438 pp.
Gibbons, J.D., 1997. Nonparametric Methods for Quantitative Analysis, 3rd Edition, American Series in Mathematical and Management Sciences. American Science Press, Columbus, 537pp.
Gibbons, J.D., Chakraborti, S., 1992. Nonparametric Statistical Inference, Vol. 131, Statistics: Textbooks and Monographs. Marcel Dekker, New York, 544pp.
Giraudeau, J., Cremer, M., Manthé, S., Labeyrie, L., Bond, G., 2000. Coccolith evidence for instabilities in surface circulation south of Iceland during Holocene times. Earth and Planetary Science Letters 179, 257–268.
Grootes, P.M., Stuiver, M., 1997. Oxygen 18/16 variability in Greenland snow and ice with 10^{-3}- to 10^{5}-year time resolution. Journal of Geophysical Research 102, 26,455–26,470.
Grove, J.M., 1988. The Little Ice Age. Methuen, London, 498pp.
Gudmundsson, H.J., 1997. A review of the Holocene environmental history of Iceland. Quaternary Science Reviews 16, 81–92.
Herbert, T.D., Schuffert, J.D., Thomas, D., Lange, C., Weinheimer, A., Peleo-Alampay, A., Herguera, J.-C., 1998. Depth and seasonality of alkenone production along the California margin inferred from a core top transect. Paleoceanography 13, 263–271.
Huntley, B., Prentice, I.C., 1993. Holocene vegetation and climates of Europe. In: Wright, H.E., Kutzbach, J.E., Webb, T., Ruddiman, W.F., Street-Perrott, F.A., Bartlein, P.J. (Eds.), Global Climates since the Last Glacial Maximum, University of Minnesota Press. Minneapolis, pp. 136–168.
Hutchinson, M.F., De Hoog, F.R., 1985. Smoothing noisy data with spline functions. Numerische Mathematik 47, 99–106.
Imbrie, J., Kipp, N.G., 1971. A new micropaleontological method for quantitative paleoclimatology: Application to a late Pleistocene Carribean core. In: Turekian, K.K. (Ed.), The Late Cenozoic Glacial Ages. Yale University Press, New Haven, CT, pp. 71–181.
Indermühle, A., Stocker, T.F., Joos, F., Fisher, H., Smith, H.J., Wahlen, M., Deck, B., Mastroianni, D., Tschumi, J., Blunier, T., Meyer, R., Stauffer, B., 1999. Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antartica. Nature 398, 121–126.
Jolliffe, I.T., 1986. Principal Component Analysis. Springer, New York, 271pp.
Keigwin, L.D., 1996. The Little Ice Age and Medieval Warm Period in the Sargasso Sea. Science 274, 1504–1508.
Kendall, M.G., Gibbons, J.D., 1990. Rank Correlation Methods, 5th Edition. Oxford University Press, New York, 237pp.
Kerwin, M.W., Overpeck, J.T., Webb, R.S., DeVernal, A., Rind, D.H., Healy, R.J., 1999. The role of oceanic forcing in mid-Holocene Northern Hemisphere climate change. Paleoceanography 14, 200–210.
Koç, N., Jansen, E., Haflidason, H., 1993. Paleoceanographic reconstructions of surface ocean conditions in the Greenland, Iceland and Norwegian Seas through the last 14 ka based on diatoms. Quaternary Science Reviews 12, 115–140.
Koç, N., Jansen, E., 1992. A high resolution diatom record of the last deglaciation from the SE Norwegian Sea: Documentation of rapid climatic changes. Paleoceanography 7, 499–520.
Koç, N., Jansen, E., 1994. Response of the high latitude Northern Hemisphere to orbital climate forcing: Evidence from the GIN-Seas. Geology 22, 523–526.
Koç, N., Schrader, H., 1990. Surface sediment diatom distribution and Holocene paleotemperature variations in the Greenland, Iceland and Norwegian Sea. Paleoceanography 5, 557–580.
Kutzbach, J.E., Guetter, P.J., 1986. The influence of changing orbital parameters and surface boundary conditions on climate simulations for the past 18,000 years. Journal of Atmospheric Sciences 43, 1726–1759.
Kutzbach, J.E., Guetter, P.J., Behling, P.J., Selin, R., 1993. Simulated climatic changes: results of the COHMAP climate-model experiments. In: Wright, H.E., Kutzbach, J.E., Webb, T., Ruddiman, W.F., Street-Perrott, F.A., Bartlein, P.J. (Eds.), Global Climates since the Last Glacial Maximum. University of Minnesota Press, Minneapolis, pp. 24–93.
Labeyrie, L., Leclaire, H., Waelbroeck, C., Cortijo, E., Duplessy, J.-C., Vidal, L., Elliot, M., LeCoat, B., Auffret, G., 1999. Temporal variability of the surface and deep waters of the North West Atlantic Ocean at orbital and millennial scales. In: Webb, R., Clark, P.U., Keigwin, L.D. (Eds.), Mechanisms of Global Climate Change at Millennial Time Scales, Vol. 112, Geophysical Monograph, American Geophysical Union, pp. 77–98.
Little, R.J.A., Rubin, D.B., 1987. Statistical Analysis with Missing Data. Wiley, New York, 278pp.
Lubinski, D.J., Forman, S.L., Miller, G.H., 1999. Holocene glacier and climate fluctuations on Franz Jozef Land, Artic Russia, 80°N. Quaternary Science Reviews 18, 85–108.
Madansky, A., 1988. Prescriptions for Working Statisticians. Springer Texts in Statistics. Springer, New York, 295pp.
Mangerud, J., Andersen, S.T., Berglund, B.E., Donner, J.J., 1974. Quaternary stratigraphy of Norden, a proposal for terminology and classification. Boreas 3, 109–127.
Mann, M., Bradley, R.S., Hughes, M.K., 1999. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. Geophysical Research Letters 26, 759–762.
McManus, J.F., Oppo, D.W., Cullen, J.L., 1999. A 0.5-million-year record of millennial-scale climate variability in the North Atlantic. Science 283, 971–975.
Meese, D.A., Gow, A.J., Grootes, P., Mayewski, P.A., Ram, M., Stuiver, M., Taylor, K.C., Wadding-ton, E.D., Zielinski, G.A., 1994. The accumulation record from the GISP2 core as an indicator of climate change throughout the Holocene. Science 266, 1680–1682.
Mix, A., Bard, E., Schneider, R., 2001. Environmental processes of the Ice Age: land, oceans, and glaciers (EPILOG). Quaternary Science Reviews 20, 627–657.
Müller, P.J., Cepek, M., Ruhland, G., Schneider, R., 1997. Alkenone and coccolithophorid species changes in Late Quaternary sedi-

ments from the Walvis Ridge: implications for the alkenone paleotemperature method. Paleogeography Paleoclimatology Paleocology 135, 71–96.

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

Nesje, A., Dahl, S.O., 1993. Lateglacial and Holocene glacier fluctuations and climate variations in western Norway: a review. Quaternary Science Reviews 12, 255–261.

O'Brien, S.R., Mayewski, P.A., Meeker, L.D., Meese, D.A., Twickler, M.S., Whitlow, S.I., 1995. Complexity of Holocene climate as reconstructed from a Greenland ice core. Science 270, 1962–1964.

Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: Dissimilarity coefficients and the method of modern analogs. Quaternary Research 23, 87–108.

Pflaumann, U., Duprat, J., Pujol, C., Labeyrie, L., 1996. SIMMAX: A modern analog technique to deduce Atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments. Paleoceanography 11, 15–35.

Pollard, D., Bergengren, J.C., Stillwell-Soller, L.M., Felzer, B.S., Thompson, S.L., 2000. Climate simulations for 10,000 and 6000 years BP using the GENESIS Global Climate Model. Paleoclimates. Data and Modelling 2, 183–218.

Prahl, F.G., Wakeham, S.G., 1987. Calibration of unsaturation patterns in long-chain ketone compositions for paleotemperature assessment. Nature 330, 367–369.

Preisendorfer, R.W., 1988. Principal Components Analysis in Meteorology and Oceanography. Elsevier, Amsterdam, Oxford, New York, Tokyo, 422pp.

Prell, W., Martin, A., Cullen, J., Trend, M., 1999. NOAA-NGDC Paleoclimatology Program, Boulder CO, USA: IGBP PAGES-World Data Center-A for Paleoclimatology, Data Contribution Series # 1999–027.

Rencher, A.C., 1995. Methods of Multivariate Analysis. Wiley-Interscience, New York, 627pp.

Rosell-Melé, A., Eglinton, G., Pflaumann, U., Sarnthein, M., 1995. Atlantic core-top calibration of the $U_{37}^{K'}$ index as a sea-surface palaeotemperature indicator. Geochimica and Cosmochimica Acta 59, 3099–3107.

Royston, J.P., 1982. Algorithm AS181: the *W* test for normality. Applied Statistics 31, 176–180.

Ruddiman, W.F., Mix, A.C., 1993. The North and Equatorial Atlantic at 9000 and 6000 yr B.P. In: Wright, H.E., Kutzbach, J.E., Webb, T., Ruddiman, W.F., Street-Perrott, F.A., Bartlein, P.J. (Eds.), Global Climates since the Last Glacial Maximum. University of Minnesota Press, Minneapolis, pp. 94–124.

Ryan, T.P., 1997. Modern Regression Methods. Wiley Series in Probability and Statistics. Wiley, New York, 515pp.

Sbaffi, L., Wezel, F.C., Kallel, N., Paterne, M., Cacho, I., Ziveri, P., Shackleton, N.J., 2001. Response of the pelagic environment to palaeoclimatic changes in the central Mediterranean Sea during the Late Quaternary. Marine Geology 178, 39–62.

Sen, A., Srivastasa, M., 1990. Regression Analysis: Theory, Methods, and Applications. Springer Texts in Statistics. Springer, New York, 347pp.

Sikes, E.L., Keigwin, L.D., 1996. A reexamiation of northeast Atlantic sea surface temperature and salinity over the last 16 kyr. Paleoceanography 11, 327–342.

Stuiver, M., Braziunas, T., Becker, B., Kromer, B., 1991. Climatic, solar, oceanic, and geomagnetic influences on late-glacial and Holocene atmospheric $^{14}C/^{12}C$ change. Quaternary Research 35, 1–24.

Stuiver, M., Reimer, P.J., 1993. Extended C-14 data base and revised Calib 3.0 age calibration program. Radiocarbon 35, 215–230.

Stuiver, M., Reimer, P.J., Bard, E., Beck, W., Burr, G.S., Hughen, K.A., Kromer, B., McCormac, G., van der Plicht, J., Spurk, M., 1998. Intcal98 radiocarbon age calibration, 24,000-0 cal B.P. Radiocarbon 40, 1041–1083.

Ter Braak, C.J.F., Juggins, S., 1993. Weighting averaging partial least square regression (WA-PLS): An improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269, 485–502.

Villanueva, J., Flores, J.A., Grimalt. J.O., A detailed comparison of the $U_{37}^{K'}$ and coccolithophorid records over the past 290 kyrs: Implications to the alkenone paleotemperature method. Submitted.

von Storch, H., 1995. Spatial patterns: EOFs and CCA. In: von Storch, H., Navarra, A. (Eds.), Analysis of Climate Variability. Springer, New York, pp. 227–257.

Waelbroeck, C., Labeyrie, L., Duplessy, J.-C., Guiot, J., Labracherie, M., Leclaire, H., Duprat, J., 1998. Improving past sea surface temperature estimates based on planktonic fossil faunas. Paleoceanography 13, 272–283.

Zhao, M., Beveridge, N.A.S., Shackleton, N.J., Sarnthein, M., Eglinton, G., 1995. Molecular stratigraphy of cores off northwest Africa: sea surface temperature history over the last 80 ka. Paleoceanography 3, 661–675.