An intercontinental comparison of chironomid palaeotemperature inference models: Europe vs North America

André F. Lotter[a,b,*], Ian R. Walker[a,c,d], Stephen J. Brooks[e] and Wolfgang Hofmann[f]

[a]*Geobotanical Institute, University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland*
[b]*Swiss Federal Institute of Environmental Science and Technology (EAWAG), CH-8600 Dübendorf, Switzerland*
[c]*Department of Biology, Okanagan University College, 3333 College Way, Kelowna, British Columbia, Canada V1V 1V7*
[d]*Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*
[e]*Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*
[f]*Max-Planck-Institut für Limnologie, August-Thienemann-Strasse 2, D-24302 Plön, Germany*

Abstract

Chironomid-temperature inference models based on North American, European and combined surface sediment training sets were compared to assess the overall reliability of their predictions. Between 67 and 76% of the major chironomid taxa in each data set showed a unimodal response to July temperature, whereas between 5 and 22% of the common taxa showed a sigmoidal response. July temperature optima were highly correlated among the training sets, but the correlations for other taxon parameters such as tolerances and weighted averaging partial least squares (WA-PLS) and partial least squares (PLS) regression coefficients were much weaker. PLS, weighted averaging, WA-PLS, and the Modern Analogue Technique, all provided useful and reliable temperature inferences. Although jack-knifed error statistics suggested that two-component WA-PLS models had the highest predictive power, intercontinental tests suggested that other inference models performed better. The various models were able to provide good July temperature inferences, even where neither good nor close modern analogues for the fossil chironomid assemblages existed. When the models were applied to fossil Lateglacial assemblages from North America and Europe, the inferred rates and magnitude of July temperature changes varied among models. All models, however, revealed similar patterns of Lateglacial temperature change. Depending on the model used, the inferred Younger Dryas July temperature decrease ranged between 2.5 and 6°C. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

Recently, several papers have explored the potential of fossil chironomids as Quaternary palaeotemperature indicators (Lotter *et al.*, 1997; Olander *et al.*, 1997; Walker *et al.*, 1997), and several other studies of a similar nature are currently in progress (Brooks and Birks, unpublished; Levesque *et al.*, unpublished; Porinchu *et al.*, unpublished). These analyses of the distributions of Holarctic chironomids with respect to temperature, and other studies (e.g. Rossaro, 1991; Walker and MacDonald, 1995;

Walker and Mathewes, 1989), indicate many common trends in the data, despite the great geographic separation of the study sites. For example, the abundance of the Chironomini in warm waters, and the predominance of the Orthocladiinae in cold habitats has frequently been noted. These similarities are readily apparent at both the generic and subfamily levels, but the difficulty of identifying larval chironomids to the species level has precluded any comparisons with a higher taxonomic resolution.

Present-day ecologists do not accept the assertion that all members of a chironomid genus can be considered to be ecological equivalents. Nevertheless, because of the similarities in chironomid distribution patterns among regions, Walker *et al.* (1991a) suggested that chironomids might be particularly robust palaeotemperature

---

indicators. Indeed, comparisons of the Nearctic and Palaearctic faunas (e.g. Sæther, 1979) reveal that many of the same species are present in both faunal regions. Where one indicator species is absent, it is frequently replaced by a sister species with similar environmental requirements.

It should also be noted, however, that the Nearctic and Palaearctic faunas have been greatly altered by both natural and anthropogenic environmental changes over the last 13,000 years; thus, it is possible that Lateglacial fossil assemblages may not have present-day analogues in close proximity, or even on the same continent. For example, in continental Europe, high arctic environments no longer exist, and the faunas of many temperate lakes have been greatly disrupted by eutrophication, acidification, and other perturbations to the environment. Thus, we might expect better analogues for the European Lateglacial fauna to occur in present-day North America.

It is with this perspective that we have undertaken a comparison of chironomid-temperature inference models recently developed for North America and Europe. Specifically, we compare air temperature inferences, and ecological optima and tolerances derived from the European and North American models with those obtained from a new, combined surface-sample training set. Analogue matching is used to assess the relevance of these models in reconstructing Lateglacial environments for two sites, one in North America and one in Europe.

2. Methods

In this paper, we make use of chironomid assemblage data collected in the course of earlier research. The methods used for surface sample collection, sediment processing, and identification of the subfossils have, therefore, been described in detail in earlier papers (Lotter *et al.*, 1997; Walker *et al.*, 1997). In each case, the chironomid assemblages had been collected from the uppermost lake sediments near the centre, or deepest point in each lake basin, and the sediment had been sorted and pretreated using similar procedures (Walker *et al.*, 1997; Lotter *et al.*, 1997).

To obtain better comparability between the two training sets we used mean July air temperatures as the variable to be reconstructed. Livingstone and Lotter (1998) showed a close relationship between water and air temperatures during the summer months, with the highest correlation for July temperatures. Using air temperatures may circumvent the problem of obtaining useful water temperature measurements from remote high elevation and arctic/subarctic lakes (see, e.g. Hann *et al.*, 1992, Walker *et al.*, 1992). For the Alps the mean July temperature for each site was estimated using a GIS model (Lotter *et al.*, 1997), whereas for the North American data set we estimated mean July temperatures from climate atlases (Thomas, 1953; Houde, 1978; McCalla, 1991). Hereafter, we refer to these estimates as the observed temperatures. Table 1 shows the geographical location, elevation, mean July temperature, and the maximum water depth for the sites used in this study.

Numerical analyses were performed in order to answer the following questions:

(i) What is the relationship between temperature and the response of individual chironomid taxa? To assess the statistical relationship of each individual taxon in the three training sets with regard to July temperature, a hierarchical set of taxon response models was used (see, Huisman *et al.*, 1993). The four response models used were a skewed unimodal response model, a symmetric unimodal response model, a monotonically increasing or decreasing sigmoidal response model, and a null model of no relationship with the environmental variable. Following Lotter *et al.* (1997) the simplest statistically significant response model for each taxon was found by fitting the most complex model first and progressively removing parameters from the regression model. This was done until the model could not be simplified without a significant change ($p < 0.05$) in the deviance of the model. Taxon response models were fitted by maximum likelihood with a Poisson error structure and a logarithmic link function, and were restricted to all taxa with occurrences in 20% or more of the samples in the screened training sets.

(ii) Which regression models are able to provide the most reliable inferences with regard to mean July air temperature? To answer this question the chironomid percentage data were transformed to square root values and weighted averaging partial least squares (WA-PLS), weighted averaging (WA), and partial least squares (PLS) regression models were applied using the program CALIBRATE (S. Juggins and C.J.F. ter Braak, unpubl. program). Screening for outliers in the data sets was carried out according to Jones and Juggins (1995) and Lotter *et al.* (1997).

(iii) How do optima and tolerances of European chironomid taxa compare with the optima and tolerances of the same taxa in the North American data set, and what are the optima and tolerances if the two data sets are merged? Species optima and tolerances were assessed by using WA implemented in the program WACALIB (Line *et al.*, 1994). We also compared the WA-PLS regression coefficients (Beta coefficients) for both data sets, as determined via WA-PLS and PLS regression, using the program CALIBRATE.

(iv) How well does the European model predict the mean July temperature at North American sites, and vice versa? To answer this question we employed WA-PLS, WA, and PLS calibration (ter Braak *et al.*, 1993) to the square-root transformed percentage data by means of the programs WAPLS (ter Braak and Juggins, 1993) and CALIBRATE.

(v) Does the Modern Analogue Technique (MAT) provide a better alternative for inferring temperatures than regression-based models? Can the North American chironomid assemblages be used as modern analogues for the European data set, and vice versa? Analogue matching was carried out by the program MAT (S. Juggins, unpubl. program), using Chi-squared distance as the dissimilarity coefficient. Here we also used square-root transformed chironomid percentage data, and all samples with a minimum occurrence of at least 40 head capsules. This gave 43 sites and 64 taxa for the North American data set and 58 sites with 42 taxa for the European data set. Following Jones & Juggins (1995) we used the 2nd percentile of the matrix of dissimilarities between all modern samples to define a close analogue, and the 5th percentile as a good analogue. For July temperature predictions the models that gave the highest coefficients of determination (r^2) and the lowest errors were chosen.

(vi) How similar are the temperature inferences, derived from different training sets and different modelling techniques, when hindcasting mean July air temperatures for fossil chironomid assemblages? The square-root transformed Lateglacial chironomid percentage data from Splan Pond (New Brunswick, Canada), and Whitrig Bog (Scotland) were used as fossil data sets in WA-PLS, WA and PLS calibration, and for analogue matching as carried out by the program MAT. The transfer functions resulting from question (iii) were applied to hindcast past mean July air temperatures.

3. Results and discussion

3.1. Training set merging

Modern training sets are usually developed for use in specific regions. Merging of different regional training sets may extend the environmental gradient of interest. Moreover, bigger training sets may allow better estimates of species' optima and tolerances in relation to important environmental variables and may also increase the chances for finding good modern analogues for fossil assemblages. Merging regional training sets has, nevertheless, some important disadvantages. With increasing size of the data set the heterogeneity also increases. This increase may be data-inherent and/or caused by taxonomical and methodological inconsistencies (Birks, 1994) often related to the involvement of different analysts. Furthermore, there is the problem of multiple analogues as well as biogeographical differences in species' optima. These factors may lead to a loss of analytical precision or even to misleading results. Taxonomic harmonization and analytical quality control are, therefore, an important issue (e.g. Munro *et al.*, 1990).

We have overcome differences in taxonomic treatment for the original data sets via discussions between the analysts principally responsible for the North American and European data. This taxonomic harmonization of the data sets forced the merger of some chironomid taxa, principally within the subtribe Tanytarsina, and subfamily Orthocladiinae. The original merged data set included 107 sites and 65 taxa. We used samples that contained at least 50 head capsules for regression and calibration purposes. This resulted in 39 sites with 51 different taxa for the North American data set, and 56 sites with 42 taxa for the European data set (see Table 1). With these data we could now study whether a combined, inter-regional data set would perform better than two separate, regional data sets.

3.2. Temperature response models

Detrended correspondence analysis (DCA; ter Braak, 1987) with mean July temperature as the sole constraining variable revealed a compositional gradient length of over 2 standard deviations (SD) for all data sets (Table 2), suggesting the use of unimodal methods (see, e.g. Birks, 1995). The taxa in each training set with statistically significant fits to the four response models are given in Table 3.

In the North American data set 70% of the taxa with occurrences in more that 20% of the samples had a unimodal (skewed: 44%; symmetric: 26%), 22.5% a sigmoidal response model, and 7.5% showed no relationship with July temperature (Table 3). In the European training set 67% of the major taxa had a unimodal (skewed: 19%; symmetric: 48%), 5% a sigmoidal response model, and 28% no relationship with July temperature. In the combined training set 76% of the major taxa showed a unimodal (skewed: 20%; symmetric: 56%), 12% a sigmoidal response model, and 12% no relationship with July temperature.

Visual inspection of the species abundance data suggests that many taxa showing 'no relationship' with July temperature may actually have had bimodal relationships (e.g. *Cricotopus* and *Orthocladius*, *Corynoneura* and *Thienemanniella*). All these taxonomic units incorporate several species, and it is, therefore, not surprising that they do not display simple unimodal relationships with July temperature.

Although bimodal, or more complex, relationships might be expected in other merged taxa (e.g. Tanytarsina), in many cases, the number of species included in these groups may be so large as to completely obscure the response of individual species.

3.3. Quantitative inference models: WA, WA-PLS and PLS

Screening for outlying samples reduced the European training set to 51 sites with a total of 40 taxa (Table 1).

Table 1a
Data pertaining to the European training set lakes used in this study (for more details see Lotter *et al.*, 1997)

Lake	Abbreviation	elevation [m asl]	Longitude [E]	Lattitude [N]	maximum depth [m]	open water area [km²]	mean July temperature [°C]	# of taxa (before merging)	# of identifiable head capsules
Hagelseewli	HAG	2339	8°02′11″	46°40′26″	18.5	0.03	7.6	6	157
Lämmerensee	LÄM	2296	7°35′13″	46°24′08″	2.5	0.07	8.7	4	717
Bachsee	BAC	2265	8°01′24″	46°40′12″	16.0	0.07	8.1	9	242
Wannisbordsee	WAN	2103	8°17′58″	46°40′58″	14.0	0.02	8.6	16	630
Iffigsee	IFF	2065	7°24′33″	46°23′16″	30.0	0.10	9.9	11	201
Flueseeli	FLU	2045	7°29′59″	46°24′38″	8.5	0.04	10.0	9	1407
Seewli See	SWL	2028	8°43′02″	46°48′45″	16.0	0.08	8.9	11	1095
Tannensee	TAN	1976	8°18′22″	46°46′30″	16.0	0.34	9.4	18	292
Sägistalsee	SÄG	1935	7°58′39″	46°40′51″	9.7	0.07	9.9	13	204
Schwellisee	SCE	1933	9°38′55″	46°45′56″	12.0	0.03	9.8	15	1705
Melchsee	MEL	1891	8°16′07″	46°46′19″	15.5	0.49	9.9	11	366
Engstlensee	ENG	1850	8°21′32″	46°46′27″	49.0	0.45	10.0	21	378
Lac Lioson	LIO	1848	7°07′46″	46°23′14″	26.0	0.07	11.1	18	449
Seebergsee	SEB	1831	7°26′38″	46°34′41″	15.5	0.06	10.8	13	201
Trüebsee	TRÜ	1764	8°23′38″	46°47′32″	7.0	0.26	10.6	14	588
Obersee	OBE	1734	9°40′53″	46°47′07″	14.5	0.08	10.9	18	347
Lutersee	LUT	1702	8°21′05″	46°50′12″	4.5	0.02	10.9	11	202
Sewenseeli	SEW	1689	8°05′15″	46°53′21″	4.5	0.03	11.8	11	78
Lac Retaud	RET	1685	7°12′00″	46°21′41″	4.5	0.01	12.1	11	267
Bannalpsee	BAN	1587	8°25′42″	46°52′12″	17.0	0.16	11.7	14	150
Lac de Nervaux	NER	1493	6°59′07″	46°22′44″	10.0	0.01	12.9	17	175
Fälensee	FÄL	1452	9°24′59″	47°15′08″	31.0	0.15	11.8	10	501
Lac Tanay	TAY	1408	6°50′33″	46°20′47″	31.0	0.18	13.6	14	180
Schwendisee	SCW	1159	9°19′55″	47°11′19″	9.5	0.04	13.7	27	301
Voralpsee	VOR	1123	9°22′43″	47°09′30″	3.3	0.15	13.9	7	165
Lac des Roussees	ROU	1058	6°05′11″	46°30′12″	11.5	0.89	14.5	19	277
Lac des Taillères	TAI	1036	6°34′07″	46°57′57″	8.5	0.44	14.4	19	125
Lag Grond	GRD	1016	9°15′29″	46°48′32″	5.0	0.02	14.9	18	219
Lac Brenet	BRE	1002	6°19′31″	46°40′29″	17.0	0.63	15.1	24	441
Lac de l'Abbaye	ABB	871	5°54′42″	46°31′48″	18.0	0.80	15.3	18	170
Seelisberg Seeli	SEL	738	8°34′16″	46°57′32″	37.5	0.18	16.3	20	194
Wilersee	WIL	730	8°37′32″	47°10′20″	21.0	0.03	16.4	24	233
Lac de Bret	BRT	674	6°46′26″	46°30′53″	18.0	0.50	17.4	11	48
Egelsee	EGE	667	8°21′39″	47°24′08″	10.0	0.02	16.7	18	87
Dittligsee	DIT	652	7°32′09″	46°45′25″	16.5	0.07	17.1	21	165
Türlersee	TÜR	643	8°30′02″	47°16′17″	21.0	0.50	17.0	20	54
Übeschisee	UEB	641	7°34′00″	46°44′07″	14.5	0.15	17.4	22	232
Burgseeli	BUG	613	7°53′11″	46°41′55″	19.0	0.09	17.3	22	168
Lac de Seedorf	SEE	609	7°02′29″	46°47′47″	7.5	0.10	17.5	17	61
Gerzensee	GER	603	7°32′53″	46°49′55″	10.0	0.27	17.6	26	255
Bichelsee	BIC	590	8°54′04″	47°27′32″	6.5	0.09	15.9	20	61
Gattiker Waldweiher	GAW	545	8°33′35″	47°16′48″	5.5	0.03	17.5	20	137
Mauensee	MAU	504	8°04′35″	47°10′20″	9.0	0.60	17.9	20	161
Burgäschisee	BUR	465	7°40′09″	47°10′10″	31.0	0.19	18.2	25	389
Inkwilersee	INK	461	7°39′50″	47°11′58″	5.0	0.12	18.1	17	139
Unterer Chatzensee	CHA	439	8°29′26″	47°25′58″	7.8	0.19	18.0	31	514
Nussbaumersee	NUS	434	8°49′58″	47°36′29″	8.2	0.25	17.9	19	130
Hüttwilersee	HÜT	434	8°50′42″	47°36′39″	15.0	0.35	17.9	18	72

Table 1a Continued

Lake	Abbrevi ation	elevation [m asl]	Longitude [E]	Lattitude [N]	maximum depth [m]	open water area [km²]	mean July temperature [°C]	# of taxa (before merging)	# of identifiable head capsules
Hasensee	HAS	434	8°49′58″	47°36′29″	5.5	0.11	17.9	20	150
Le Loclat	LOC	432	6°59′52″	47°01′13″	9.2	0.05	18.3	27	202
Rotsee	ROT	419	8°19′01″	47°04′18″	16.0	0.50	18.3	22	52
Mettmenhasler See	MET	418	8°29′35″	47°28′31″	12.5	0.03	18.1	26	240
Lago di Montorfano	MON	394	9°08′23″	45°46′59″	6.5	0.52	22.5	19	627
Lago di Segrino	SEG	374	9°16′02″	45°49′51″	8.5	0.34	20.6	27	818
Lago di Muzzano	MUZ	337	8°55′42″	45°59′53″	3.2	0.22	21.0	16	158
Lago di Endine	END	334	9°56′50″	45°46′55″	8.0	0.49	21.4	17	199
Minimum		334			2.5	0.01	7.6	4	48
Mean		1139			13.9	0.21	14.5	17	317
Median		1009			11.8	0.12	15.0	18	202
Maximum		2339			49.0	0.89	22.5	31	1705
Standard deviation		661			9.3	0.22	3.9	6	318

No outlying samples were identified in the Canadian training set.

Together with the compositional gradient length of > 2 SD units, as assessed by DCA (Table 2), the fact that over 65% of the prominent taxa show a unimodal response to July temperature suggests that inference methods based on a unimodal model may perform best with the three training sets. This has also been tested empirically by applying different inference models to the data (Table 2).

As revealed by jack-knifed coefficients of determination (r^2_{jack}) and root mean squared errors of prediction ($RMSEP_{jack}$), simple WA models consistently out-performed WA models with tolerance down-weighting. For the combined and the North American training set linear PLS models performed better than simple WA models. However, all of these models gave consistently lower r^2_{jack} and higher $RMSEP_{jack}$ than WA-PLS models using two components. WA-PLS gave a $r^2_{jack} = 0.844$ and a $RMSEP_{jack}$ of 1.47°C for the European training set, with regard to July air temperatures. The two-component WA-PLS model for the North American training set gave a $r^2_{jack} = 0.852$ with a $RMSEP_{jack} = 1.54$°C (see Table 2).

Despite taxonomic harmonization, resulting in the loss of some taxonomic resolution, comparison of the r^2_{jack} and $RMSEP_{jack}$, before and after harmonization, revealed that the new models had similar prediction capabilities to those of the original models (see Lotter et al., 1997; Walker et al., 1997). Furthermore, both models give comparable results with regard to predictive power and error estimates. The combined and screened training set includes 90 samples and 57 taxa (Table 2) and produced a somewhat lower r^2_{jack} (0.813) and higher $RMSEP_{jack}$ (1.74°C). Nevertheless, these results compare well with published results. For r^2 they range between 0.58 (Olander et al., 1997) and 0.88 (Walker et al., 1997), whereas the RMSEP lie between 1.37°C (Lotter et al., 1997) and 2.26°C (Walker et al., 1997). Note, however, that the larger errors reported by Walker et al. (1997) are to some extent related to the larger temperature gradient, and the use of water, rather than air, temperatures.

Although taxonomic harmonization, involving merging of different taxa, generally results in a loss of ecological information (Birks, 1994) this seems not to be a major problem in the case of these two data sets. The taxonomic units distinguished in fossil chironomid analysis represent already groups of species that, from a morphological

Table 1b
Data pertaining to the North American training set lakes used in this study (for more details see Walker *et al.*, 1997)

Lake	Abbreviation	elevation [m asl]	Longitude [E]	Lattitude [N]	maximum depth [m]	open water area [km^2]	mean July temperature [°C]	# of taxa (before merging)	# of identifiable head capsules
Immerk Lake	IMM	30	84°34′	75°41′	6.8	1.05	5.0	10	49.5
Middle Beschel Lake	MBE	30	84°28′	75°39′	8.1	0.45	5.0	11	50.5
Phalarope Lake	PHA	30	84°37′	75°39′	5.0	1.57	5.0	7	81.5
Fish Lake	FIS	30	84°32′	75°39′	4.5	0.97	5.0	9	54.5
Lake INSTAAR	INS	10	66°16′	62°16′	10.0	0.52	6.7	12	117
Lake Mercer	MER	1	66°15′	62°16′	10.0	0.38	6.7	9	103.5
L46	L46	479	62°48′	57°20′	22.0	0.70	7.7	9	57
L68	L68	500	62°55′	57°20′	9.0	0.59	8.0	4	21.5
L45	L45	410	63°08′	57°17′	24.0	0.54	7.5	7	57.5
L48	L48	260	62°46′	57°15′	15.0	1.03	8.0	4	12
L43	L43	485	63°05′	57°08′	48.5	1.68	7.8	9	63
L49	L49	530	64°09′	56°42′	14.0	0.86	9.0	3	5
L50	L50	508	64°32′	56°39′	9.0	0.57	10.0	12	154
L52	L52	510	64°30′	56°35′	11.0	0.84	10.0	4	15
L53	L53	460	64°15′	56°22′	9.0	0.86	10.2	7	49.5
L54	L54	480	63°57′	56°17′	3.0	1.14	10.5	17	345.5
L56	L56	470	64°25′	56°10′	6.0	0.57	10.5	18	254.5
L36	L36	575	62°44′	55°12′	3.0	0.59	12.0	14	67.5
L35	L35	569	62°44′	55°06′	17.0	0.68	12.0	10	68
L64	L64	554	63°07′	55°06′	13.0	0.57	12.0	14	61.5
L61	L61	530	63°15′	55°52′	7.0	0.76	12.3	19	164.5
L58	L58	513	62°44′	54°52′	4.0	0.97	12.3	18	152
L42	L42	578	62°23′	54°49′	5.0	1.89	12.3	8	83
L41	L41	587	62°21′	54°48′	14.0	1.00	12.2	13	82
L32	L32	290	61°12′	54°28′	27.0	0.78	12.2	6	14.5
L29	L29	306	61°09′	54°22′	14.0	0.78	12.1	13	49.5
L34	L34	280	60°53′	54°18′	26.0	1.62	12.1	12	74.5
L28	L28	285	61°08′	54°18′	3.0	1.65	12.0	16	94
L23	L23	295	60°41′	53°38′	24.0	0.59	14.0	11	42
L22	L22	136	60°45′	53°37′	14.0	0.70	14.0	15	56
L24	L24	255	60°37′	53°36′	11.0	0.65	14.5	7	31
L10	L10	221	56°52′	52°52′	2.0	1.30	13.5	18	57.5
L14	L14	380	57°43′	52°22′	1.0	0.84	13.5	23	251
L20	L20	371	57°33′	52°05′	1.0	1.30	13.5	25	196
L16	L16	319	57°50′	52°05′	8.0	1.35	13.5	19	342.5
L17	L17	319	57°52′	52°05′	6.0	0.54	13.5	20	315.5
L2	L2	195	56°45′	51°35′	5.0	0.04	12.5	11	69
L3	L3	43	57°14′	51°30′	23.0	1.05	12.5	12	72.5
L1	L1	43	57°12′	51°27′	14.0	0.59	12.5	15	106
Black Lake	BLA	34	64°22′	46°06′	6.0	0.18	18.4	19	174.5
Portey Pond	POR	45	64°25′	45°21′	1.5	0.04	17.8	19	137.5
Leak Lake	LEA	0	64°21′	45°25′	14.2	0.17	18.0	25	87.5
Ritchie Lake	RIT	0	65°58′	45°25′	12.0	0.23	17.0	21	52
Long Lake	LON	0	66°04′	45°19′	9.7	0.05	16.8	18	66
Joe Lake	JOE	335	66°40′	46°45′	1.8	0.04	17.2	19	333.5
Killarneyl Lake	KIL	75	66°38′	46°01′	9.8	0.11	19.0	20	413
Pine Ridge Pond	PIN	85	67°06′	45°34′	5.6	0.01	18.5	12	131.5

Table 1b Continued

Lake	Abbreviation	elevation [m asl]	Longitude [E]	Lattitude [N]	maximum depth [m]	open water area [km^2]	mean July temperature [°C]	# of taxa (before merging)	# of identifiable head capsules
Splan Pond	SPL	106	67°20′	45°15′	10.6	0.04	18.2	29	468
Trout Pond	TRO	0	68°11′	45°05′	11.4	0.03	18.0	9	59.5
Minimum		0			1.0	0.01	5.0	3	5
Mean		276			11.0	0.72	12.1	13	119.7
Median		290			9.7	0.68	12.2	12	72.5
Maximum		587			48.5	1.89	19.0	29	468
Standard deviation		208			8.7	0.49	3.9	6	110.6

Table 2
Comparison of screened modern European and North American chironomid training sets, including performance of various chironomid-temperature regression models based upon apparent and jack-knifed statistics

	European training set	North American training set	Combined training set
Number of samples	51	39	90
Number of taxa	40	51	57
Min. July temperature (°C)	7.6	5.0	5.0
Max. July temperature (°C)	21.0	19.0	21.0
Median July temperature (°C)	15.0	12.2	12.5
Standard deviation (°C)	3.7	3.9	4.0
DCA gradient length (SD)	2.519	3.050	2.696
DCCA λ_1	0.275	0.349	0.301
λ_2	0.148	0.094	0.192
Species-environment correlation	0.914	0.934	0.902
PLS			
Number of PLS components	1	3	1
Apparent r^2	0.871	0.958	0.822
r^2_{jack}	0.832	0.842	0.794
Apparent RMSE (°C)	1.335	0.818	1.695
RMSEP$_{jack}$ (°C)	1.526	1.599	1.822
Jack-knifed mean bias	0.020	− 0.007	− 0.037
Jack-knifed maximum bias	3.194	1.545	4.916
WA (classical deshrinking)			
Apparent r^2	0.786	0.885	0.805
r^2_{jack}	0.744	0.848	0.775
Apparent RMSE (°C)	1.941	1.443	1.977
RMSE$_{jack}$ (°C)	2.082	1.623	2.101
Jack-knifed mean bias	0.049	− 0.015	− 0.001
Jack-knifed maximum bias	3.055	2.341	3.926
WA-PLS			
Number of WA-PLS components	2	2	2
Apparent r^2	0.924	0.940	0.894
r^2_{jack}	0.844	0.852	0.813
Apparent RMSE (°C)	1.028	0.979	1.308
RMSEP$_{jack}$ (°C)	1.470	1.544	1.741
Jack-knifed mean bias	0.068	0.022	− 0.053
Jack-knifed maximum bias	2.743	1.704	3.753

Table 3

Taxon response models (Huisman et al., 1993) for all taxa with occurrences in at least 20% of the samples. *Model A*: Skewed unimodal relationship between taxon abundance and July temperature. *Model B*: symmetric unimodal relationship between taxon abundance and July temperature. *Model C*: monotonically increasing or decreasing sigmoidal relationship between taxon abundance and July temperature. *Null model*: no relationship between taxon abundance and July temperature: Taxon occurs in less than 20% of the samples

Taxon	North American Training set	European Training set	Combined Training set
Tribe Pentaneurini	B	B	B
Procladius	B	B	B
Subtribe Tanytarsina	A	A	A
Stempellina	C	-	B
Stempellinella and *Zavrelia*	A	-	A
Pseudochironomus	C	B	B
Chironomus	A	B	C
Cladopelma	A	B	A
Cryptochironomus	A	-	-
Dicrotendipes	B	0	B
Endochironomus	-	B	B
Glyptotendipes	-	B	B
Lauterborniella and *Zavreliella*	A	A	B
Microtendipes	B	B	C
Pagastiella	B	-	-
Parachironomus	-	B	B
Polypedilum	A	B	B
Sergentia	A	-	A
Stictochironomus	A	-	B
Tribelos	A	-	A
Protanypus	B	-	B
Diamesa	-	A	-
Crictopus and *Orthocladius*	A	0	0
Corynoneura and *Thienemanniella*	0	0	0
Heterotanytarsus	B	-	-
Heterotrissocladius	C	-	B
Nanocladius	-	A	-
Paracladius	-	C	C
Parakiefferiella cf. *bathophila*	*A*	-	-
Other *Parakiefferiella* spp.	-	0	-
Psectrocladius	C	0	B
Zalutschia	C	-	-
Mesocricotopus thienemanii	C	-	-
Undifferentiated *Orthocladiinae*	0	0	0
n	27	21	25
Null model, no relationship	2	6	4
Model A, skewed unimodal relationship	12	4	5
Model B, symmetric unimodal relationship	7	10	14
Model C, sigmoidal relationship	6	1	3

point of view, cannot be distinguished to a lower taxonomic level. Such taxonomic units, however, very often lump several species with different environmental optima and tolerances. It is, therefore, intriguing to note how well the different quantitative inference models for chironomids generally perform. For example, harmonization resulted in the loss of 18 taxa in the European training set, but did not significantly change the predictive power of the new model in comparison to the original one.

3.4. Optima, tolerances and WA-PLS regression coefficients

Using WA, PLS, and WA-PLS regression we estimated optima, tolerances and WA-PLS regression coefficients (Beta coefficients) for the chironomid taxa on the two continents. This comparison involved a total of 34 taxa that occurred in both training sets. Twenty-three taxa (68%) had lower optima, one (3%) had the same optimum, and 10 taxa (29%) had higher optima in the
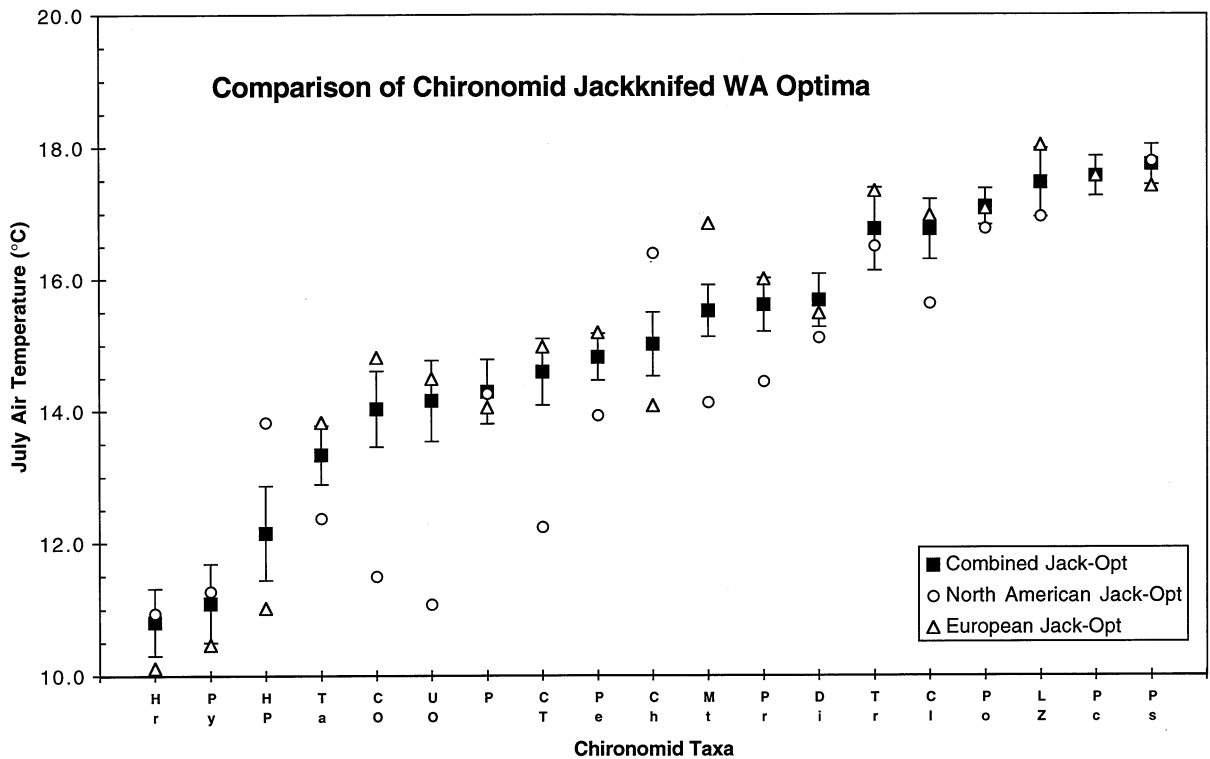
Fig. 1. Comparison of estimated WA July temperature optima for the North American, European, and combined training sets. Only taxa with a Hill's (1973) N2 > 5 in both training sets are shown. Chironomid taxa have been sorted from lowest to highest temperature optima in the combined data set. Jack-knifed optima, with bars spanning ± 1 standard error of the optima. Taxon codes (alphabetically): Ch, *Chironomus*; Cl, *Cladopelma*; CO, *Cricotopus/Orthocladius*; CT, *Corynoneura/Thienemanniella*; Di, *Dicrotendipes*; HP, *Cyphomella/Harnischia/Paracladopelma* type; Hy, *Heterotanytarsus*; L/Z, *Lauterborniella/Zavreliella*; Mt, *Microtendipes*; P, *Psectrocladius*; Pc, *Parachironomus*; Pe, Tribe Pentaneurini; Po, *Polypedilum*; Pr, *Procladius*; Ps, *Pseudochironomus*; Py, *Protanypus*; Ta, Subtribe Tanytarsina; Tr, *Tribelos*; UO, Undifferentiated Orthocladiinae.

North American training set (Fig. 1). The largest differences in WA-optima occurred in taxa such as *Potthastia* (5.5°C), *Paracladius* (4.4°C), *Diamesa* (3.5°C), *Stempellina* (3.5°C), and the undifferentiated Orthocladiinae (3.4°C). The large discrepancies in *Diamesa* optima can be attributed to the fact that it was recorded from only a single site in North America. Similarly, *Potthastia* was recorded from only a single European site. However, "undifferentiated Orthocladiinae" and other taxa were common in both training sets. To some extent, especially with regard to the "undifferentiated Orthocladiinae", these differences may reflect differences in chironomid taxonomy that could not be addressed fully by our taxonomic harmonization process.

Generally, the optima appear very similar, and are strongly correlated ($r = 0.815$) between the two training sets (Fig. 1). Overall, 22 (65%) of the European jack-knifed optima, and 16 (47%) of the North American jack-knifed optima lie within one standard error of the mean jack-knifed optima, as calculated for the combined training set (Fig. 1).

For 31 taxa, temperature tolerance ranges could be calculated for both the North American and European data sets. The correlation ($r = 0.565$) between the European and North American tolerances was much lower than the correlation for optima. Eleven taxa (35%) had narrower tolerances, one taxon had the same tolerance, and 19 taxa (61%) had wider tolerances in the North American data set (Fig. 2). Fifteen (48%) of the European jack-knifed tolerances, and 16 (52%) of the North American jack-knifed tolerances lay within one standard error of the mean jack-knifed optima, as calculated for the combined training set (Fig. 2).

We also compared the regression (Beta) coefficients for the best WA-PLS and PLS models. This revealed correlation coefficients of 0.632, 0.736 and 0.156 between WA-PLS (two component), PLS (one component), and PLS (three component) regression coefficients, respectively.

3.5. Using one training set to predict temperatures at the other sites

Applying the North American two-component WA-PLS inference model to predict mean July temperatures at the European sites yielded sample specific standard
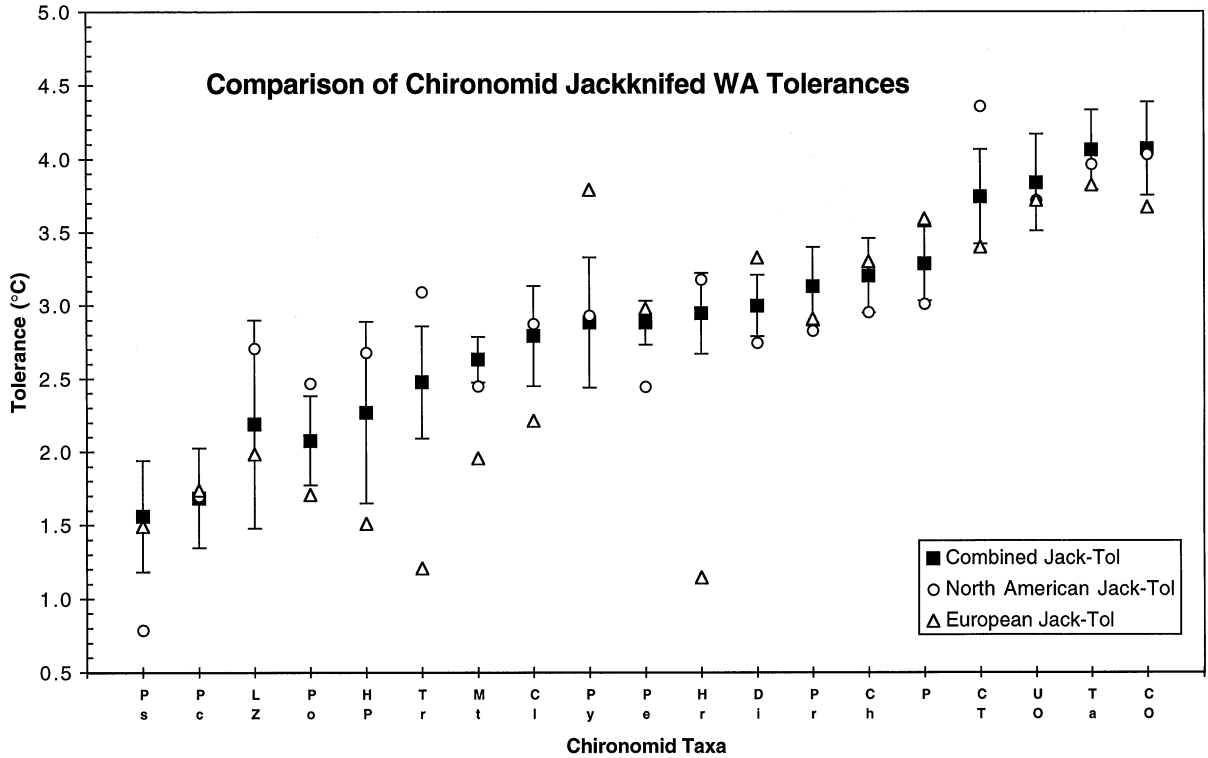
Fig. 2. Comparison of estimated WA tolerances for the North American, European, and combined training sets. Only taxa with a Hill's (1973) N2 > 5 in both training sets are shown. Chironomid taxa have been sorted from lowest to highest July temperature tolerances in the combined data set. Jack-knifed tolerances, with bars spanning ± 1 standard error of the tolerances. For taxon codes see caption for Fig. 1.

error ranges which encompassed the 'observed' temperatures for 78.4% of the sites (Fig. 3). The correlation coefficient between model predicted and observed July temperatures is 0.767 and the WA-PLS model-predictions explained 58.8% of the variance in the European July temperature data (see Table 4).

Significant deviations occurred at the lower end of the temperature gradient between 8 and 13°C, where the inference model overestimated the temperatures. At the high end of the gradient, the model underestimated the mean July temperatures for two sites from the southern Alpine foreland (Fig. 3). This phenomenon has also been observed for other organisms in connection with lakes from this so-called insubric climate region (warm, precipitation-rich summers, mild winters, see Lotter *et al.*, 1997). It is, therefore, likely that the warmer climatic conditions prevailing on the southern side of the Alps create physical properties in these lakes that are not comparable to the ones prevailing in temperate lakes on the northern side of the Alps. A separate training set focussing on lakes on the southern slope of the Alps as well as the Mediterranean region might help solve this problem.

Using the European two-component WA-PLS inference model to infer present-day temperatures at North

American sites, we noted 71.8% of the observed mean July temperatures lay within the range defined by the site-specific errors (Fig. 4). Temperatures at the lower end of the gradient, around 5°C, tended to be overestimated. Observations lying outside the site-specific errors were scattered in the middle and upper part of the temperature gradient. There is a higher correlation ($r = 0.849$, see Table 4) between model-predicted and observed July temperatures and a larger portion of the variance (72%) was explained by the European model when applied to the North American data set.

However, although the two-component WA-PLS models performed consistently superior to all other chironomid-temperature models, this result was not corroborated by the "inter-continental cross-validation" (Table 4). Our calculations indicate that the North American one-component PLS model provided much better temperature estimates ($r^2 = 0.770$, RMSE = 1.85°C) for European sites than North American WA or WA-PLS models. The European one-component WA-PLS model which is equivalent to a WA model with inverse deshrinking provided better estimates ($r^2 = 0.778$, RMSE = 1.91°C) of temperatures at North American lakes, than all other European models. This cross-validation shows that the performance of a model
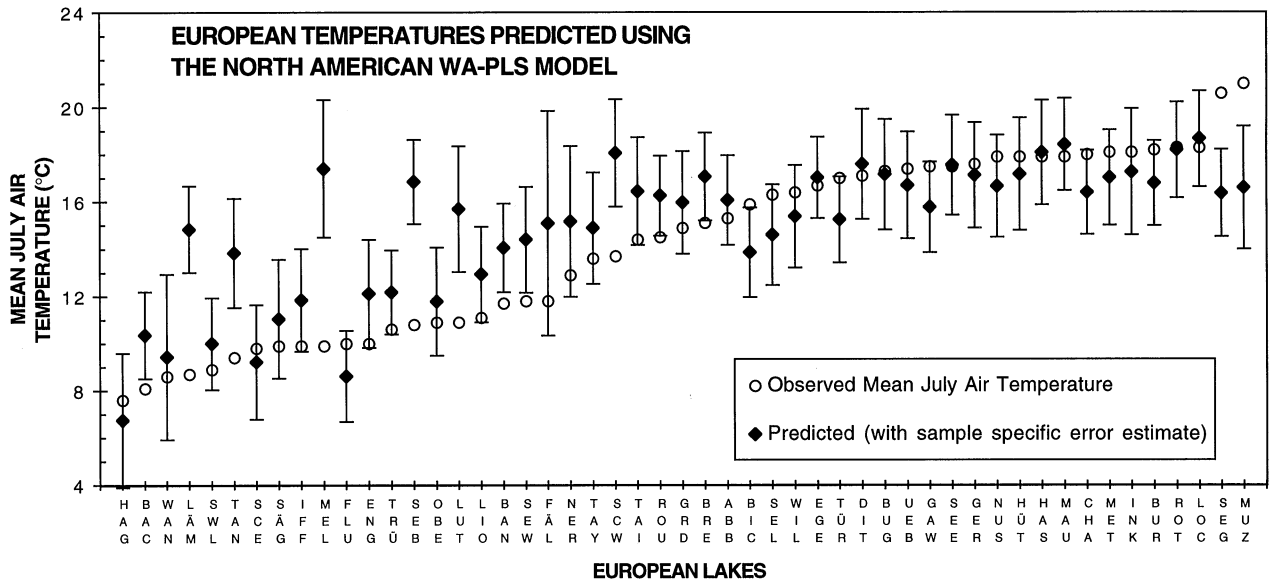
Fig. 3. Prediction of mean July temperature for the European sites using the North American two-component WA-PLS inference model. The sites are arranged according to their mean July air temperatures. Open dots signify the measured mean July air temperatures, whereas the filled diamonds show the chironomid-inferred mean July temperatures including the sample-specific error of prediction as assessed by 10,000 simulations. For site codes see Table 1a.

Table 4
Comparison of various chironomid-temperature regression models, assessed on their ability to estimate temperatures at lakes on the other continent

		North American model estimating European temperatures	European model estimating North American temperatures	Combined model estimating European temperatures	Combined model estimating North American temperatures
Number of observations		51	39	51	39
WA (Classical deshrinking)	r	0.748	0.882	0.872	0.913
	r^2	0.559	0.778	0.761	0.834
	RMSE	2.71	2.01	2.01	1.93
WA-PLS (2 components)	r	0.767	0.849	0.943	0.946
	r^2	0.588	0.721	0.889	0.896
	RMSE	2.52	2.27	1.30	1.32
PLS (1 component)	r	0.878	0.864	0.902	0.900
	r^2	0.770	0.746	0.813	0.810
	RMSE	1.85	2.19	1.65	1.75
MAT (mean of 7 analogues)	r	0.786	0.814	0.823	0.924
	r^2	0.617	0.662	0.677	0.853
	RMSE	2.46	3.02	2.19	1.78

(see, e.g. Birks, 1995) cannot be accurately assessed on the basis of the highest r^2_{jack} and the lowest $RMSEP_{jack}$ only (see Table 2). Moreover, the comparison also shows that the combined inter-regional model consistently performs better and explains more variance in the regional data sets than each regional data set alone (Table 4).

3.6. Modern analogues

The best models for inferring temperatures based on the modern analogue technique were selected on the basis of the highest r^2 and lowest RMSE. The best European model was based on a mean of the 9 closest
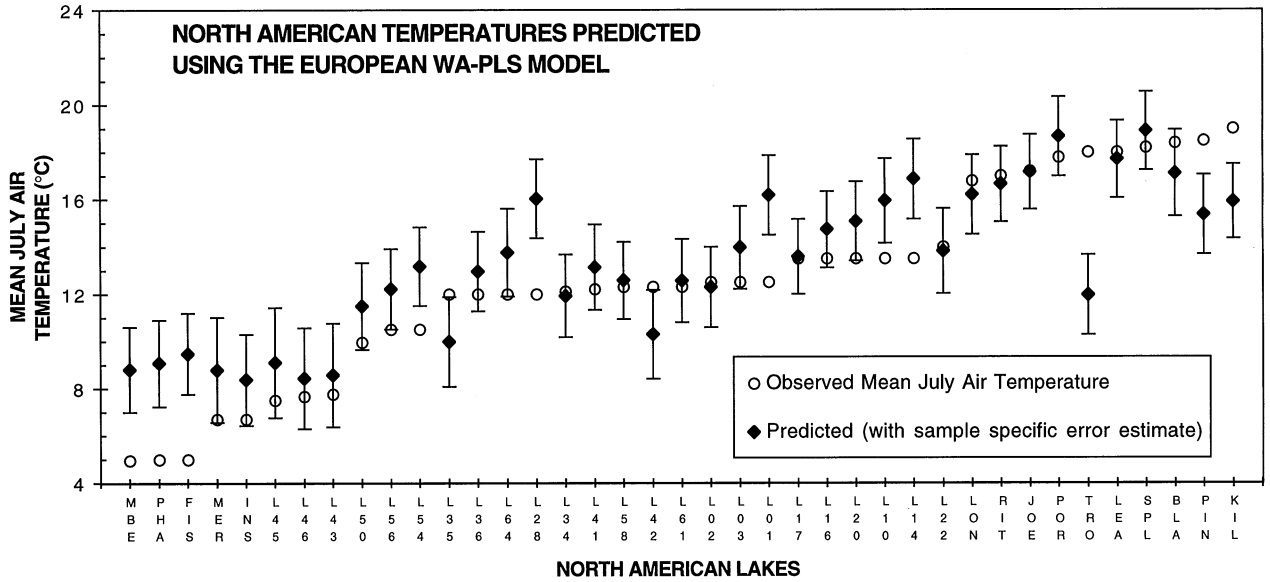
Fig. 4. Prediction of mean July temperature for the North American sites using the European two-component WA-PLS inference model. The sites are arranged according to their latitude. Open dots signify the measured mean July air temperatures, whereas the filled diamonds show the chironomid-inferred mean July temperatures including the sample specific error of prediction as assessed by 10,000 simulations. For site codes see Table 1b.

analogues ($r^2 = 0.691$; RMSE $= 2.14°$C). The best North American and combined models were based on a mean of the 3 closest analogues (North America: $r^2 = 0.930$, RMSE $= 1.08°$C; Combined: $r^2 = 0.782$, RMSE $= 1.90°$C).

When used as analogues to predict European July temperatures, the North American training set gave a lower RMSE ($2.46°$C) and a lower r^2 (0.617), than when the European training set was used to provide analogues to infer North American temperatures (RMSE $= 3.02°$C; $r^2 = 0.662$).

For the 43 North American sites neither close nor good analogues existed in the European training set (Fig. 5a). The European modern analogue model tended to overestimate the real temperatures at North American sites. Twenty observations (46.5%) lay within the range of the RMSE (Fig. 5b). For the 58 European sites there were no close analogues and only 3 good North American analogues (Fig. 6a): BUR and L17; SEL and L17; FLU and Middle Beschel Lake. The modern analogue model predicted only 21 sites (36%) within the limit of the RMSE (Fig. 6b). There was a general tendency for the model to underestimate the real temperatures. All three sites that had good analogues, according to the 5th percentile rule (Birks et al., 1990; Birks, 1995; Jones and Juggins, 1995), had chironomid-inferred July temperatures well outside the limits suggested by the RMSE, whereas the sites with good correspondence between observed and predicted temperatures had, in many cases, large Chi-squared distances (i.e. poor analogues).

The fact that despite a poor analogue situation, the models were able to reliably predict the actual temperatures, implies that temperature and other environmental reconstructions may be used in no-analogue situations, such as the Lateglacial period.

3.7. Quantitative reconstruction

We applied the different inference models to the Lateglacial chironomid stratigraphies of Splan Pond, in Atlantic Canada (Walker et al., 1991b) and Whitrig Bog in Scotland (Brooks et al., 1997a,b) to compare the performance of the different training sets, as well as the different inference models. These stratigraphies include the time between deglaciation and the Preboreal chronozone (i.e. between ca. 12,000 and 9000 conventional radiocarbon years BP). This period was marked by the Younger Dryas, a climatic cooling that took place between ca. 11,000 and 10,000 ^{14}C years BP (e.g. Ammann and Lotter, 1989; Björck et al., 1996). This cooling has been recorded in continental records on both sides of the Atlantic (e.g. Peteet, 1992) as well as in different ice-core records (e.g. Jouzel et al., 1994; Johnsen et al., 1995). Several shorter climatic oscillations have also been demonstrated in continental oxygen isotope and palynological records (see e.g. Lotter et al., 1992; Levesque et al., 1993). We concentrated our comparison on this well-established and well-defined high-amplitude Lateglacial climatic oscillation.

Because of the predominantly unimodal taxon response models (see Table 3) we have applied unimodal-based
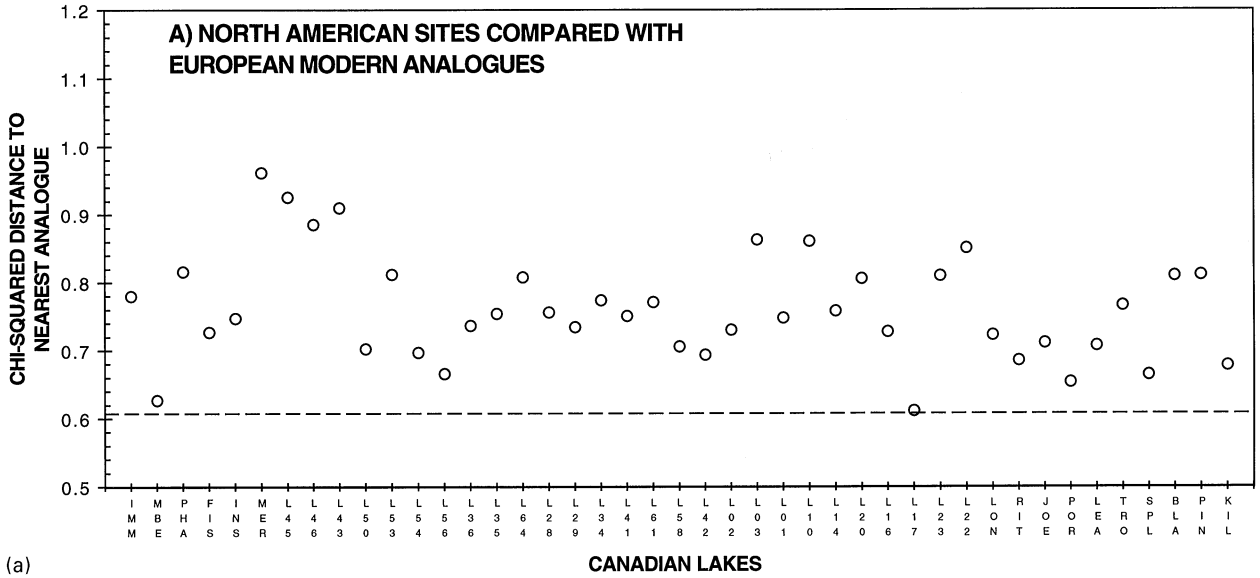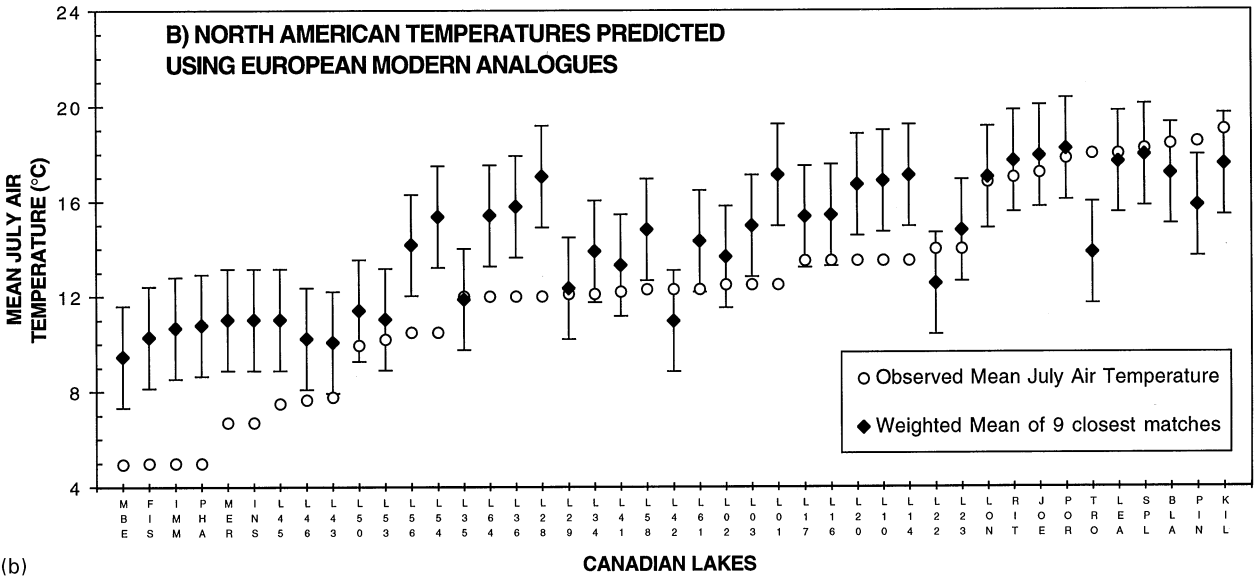
Fig. 5 (a) Chi-squared distances between the 43 North American chironomid assemblages and their closest European analogue. Values below the dashed line (5th percentile) indicate good analogues. (b) Mean July air temperatures predicted (black diamonds) by a modern chironomid analogue model using the weighted mean of the 9 closest matches in the European data set and observed mean July temperatures at 43 North American sites. The sites are ordered according to increasing mean July air temperature (see Table 1b). The error bars represent the RMSE (= ± 2.136°C).

inference models such as WA and WA-PLS to the fossil data. However, since 5–22% of the prominent taxa show a sigmoidal relationship with July temperature, we have also applied a linear-based PLS inference model. Moreover, we estimated past July temperatures using MAT. For reasons of comparison we have applied all three models to both fossil data sets. The principal focus of this comparison should, however, be the comparison of reconstructions derived from each regional data set with the reconstruction derived from the combined data set.

With WA, WA-PLS and PLS, the European training set generally inferred lower July temperatures at Splan Pond than the North American training set (Fig. 7). When based on the combined data set, the July temperature inferences commonly lay between the European and North American estimates. This pattern was not so apparent in the Whitrig Bog reconstructions (Fig. 8). At Whitrig Bog, the European WA model tended to infer lower July temperatures than the North American WA model through the Younger Dryas, but the North American PLS model inferred lower temperatures than the
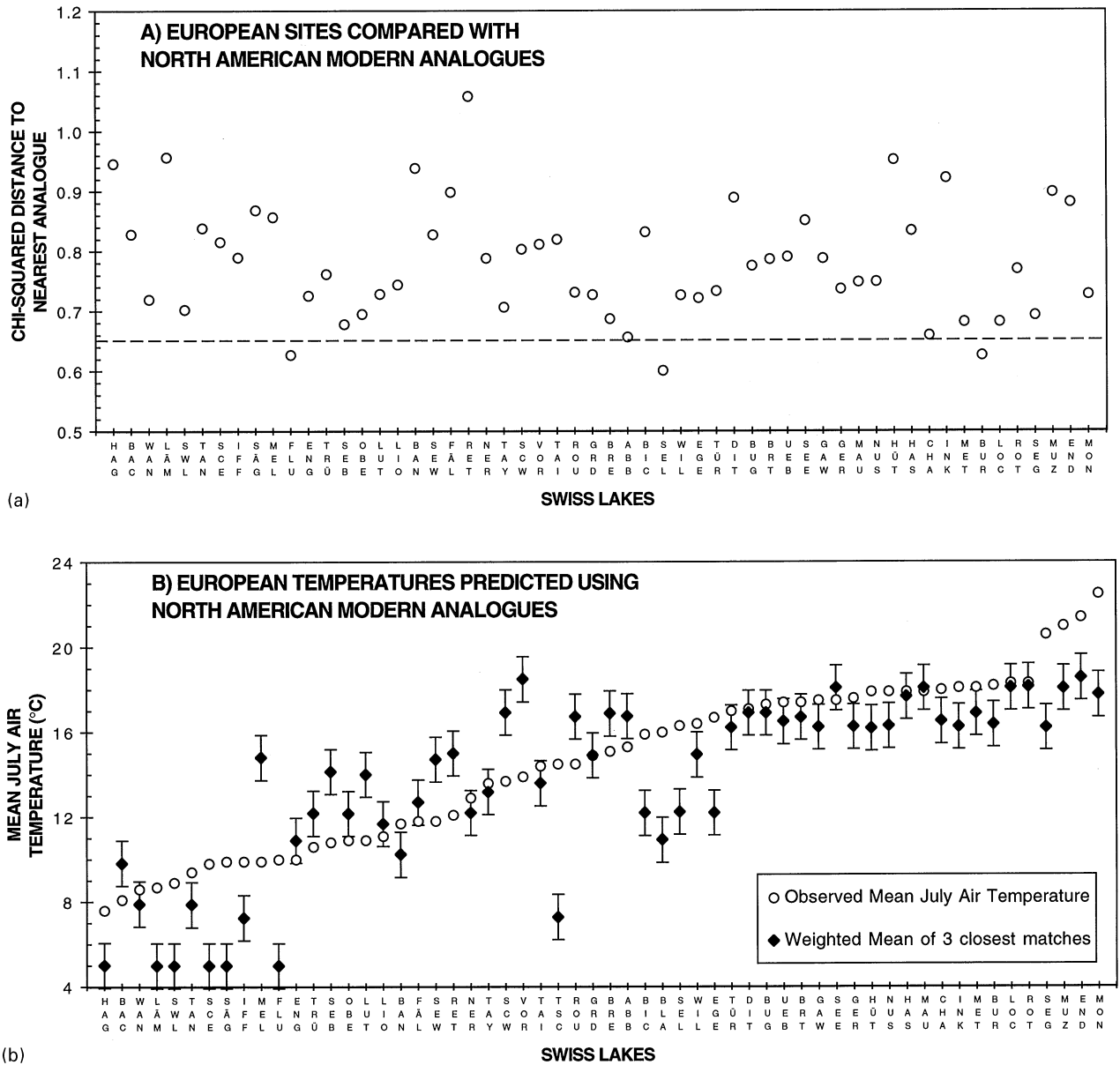
Fig. 6 (a) Chi-squared distances between the 58 European chironomid assemblages and their closest North American analogue. Values below the dashed line (5th percentile) indicate good analogues. (b) Mean July air temperatures predicted (black diamonds) by a modern chironomid analogue model using the weighted mean of the 3 closest matches in the North American data set and observed mean July temperatures at 58 European sites. The sites are ordered according to increasing mean July air temperature (see Table 1a). The error bars represent the RMSE (= ± 1.062°C).

European PLS model throughout most of the Lateglacial period.

The MAT approach produced the smoothest July temperature curves for Splan Pond (Fig. 7). The MAT reconstructions based on the North American and combined training sets were virtually identical because the method chose the same sites as closest analogues. The European data set provided one good modern analogue (CHA) for the sample at 190 cm, whereas the North American data set provided two close modern analogues (Ritchie, L35) for the samples at 95 and 105 cm and four good modern analogues (L64, L35, L17, Killarney)

for the samples at 125, 145, 183, and 190 cm. The same two close modern analogues and eight good modern analogues (twice L64, L35, L17, Leak, Portrey, CHA, MET) were identified when we used the combined training set. The closest modern samples (i.e., the samples with the lowest Chi-squared distance) remain the same, except for the sample at 190 cm (CHA instead of Killarney), but the cut-off level for good modern analogues increased from < 0.635 to 0.665 when these two training sets were combined. As a result, more good modern analogues were apparent with the combined training set.
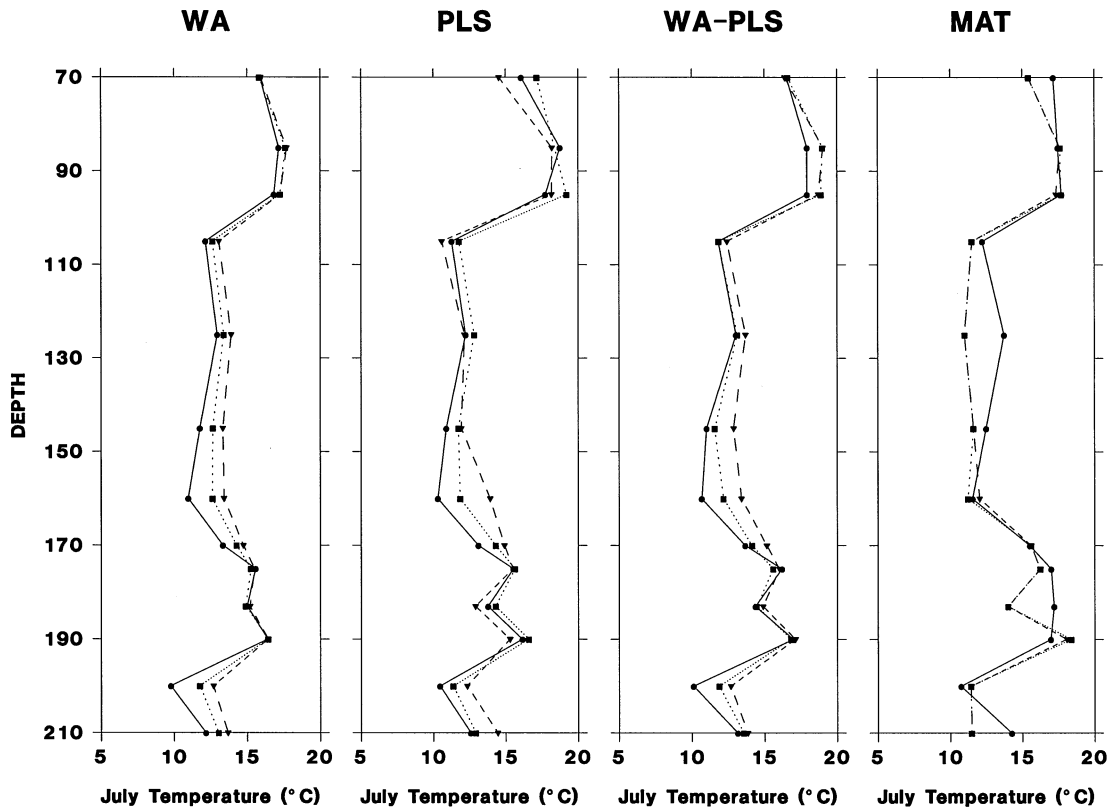
Fig. 7. Comparison of the chironomid-inferred July temperatures for the Lateglacial of Splan Pond using different quantitative inference models: Weighted Averaging with classical deshrinking (WA), Partial Least Squares (PLS), Weighted Averaging Partial Least Squares (WA-PLS), and Modern Analogue Technique (MAT) using Chi-squared distance as dissimilarity coefficient. Each model has been run with the European (dots, solid line), North American (triangles, dashed line), and combined training set (squares, dotted line).

Depending on which data set is used, different cut-off levels will result (see, e.g. Figs. 5 and 6). This highlights one problem inherent to the modern analogue approach. The method is sensitive to the presence or absence of strong taxa. The choice of close and good analogues using percentiles as cut-off levels is based on ecological and palaeoecological arguments rather than on a statistical basis. As an alternative, it may therefore be sensible to check for missing taxa in the modern training set or for taxa with low occurrences common in the modern and the fossil data set when assessing and evaluating the reliability of inference models (see Birks, in press).

The three MAT reconstructions for Whitrig Bog all appear very similar (Fig. 8), except for the anomalously low values inferred by the Canadian and combined data sets at 106 cm. Fish Lake, in the Canadian high arctic, was identified as the only site providing a good analogue for this sample. In applying the European model, only four samples were identified as having close modern analogues, with nine additional samples having good modern analogues. Similarly, using the North American model, eight Whitrig Bog samples had close analogues, and 13 additional samples had good analogues.

When we combined the training sets, 19 (25%) of the Whitrig Bog samples were identified as having close modern analogues. Good modern analogues were apparent for an additional 21 (28%) samples. In total, 29 close analogues and 63 good analogues were identified with the combined training set. North American sites contributed 52% of the close analogues, but European sites provided 65% of the good analogues. In general, North American sites provided better analogues for Younger Dryas samples, but European sites provided better analogues for the Allerød period. This may well be related to the fact, that the North American training set includes arctic/subarctic sites with lower temperatures than the European training set (see Tables 1 and 2). Therefore, this training set may provide better modern analogues for past low-temperature assemblages. This is consistent with the observation that Europe currently has no high arctic landscapes.

Because it was not possible to decide which inference model was the most appropriate for the fossil data sets we have used a combined reconstruction, utilizing the results of the four different inference models (Figs. 9 and 10). All four models gave slightly different reconstructions
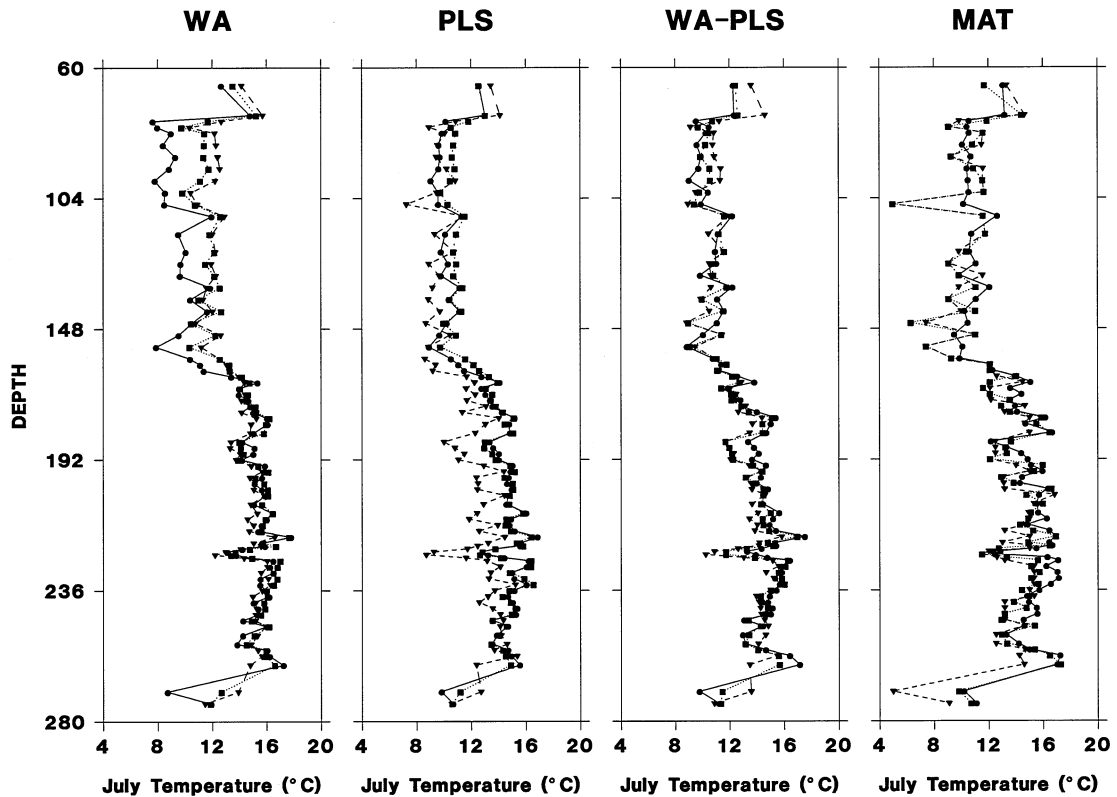
Fig. 8. Comparison of the chironomid-inferred July temperatures for the Lateglacial of Whitrig Bog using different quantitative inference models: Weighted Averaging with classical deshrinking (WA), Partial Least Squares (PLS), Weighted Averaging Partial Least Squares (WA-PLS), and Modern Analogue Technique (MAT) using Chi-squared distance as dissimilarity coefficient. Each model has been run with the European (dots, solid line), North American (triangles, dashed line), and combined training set (squares, dotted line).

because they emphasise different features of the data. Given RMSEP of between 1.5 and 2°C (Table 2) most of the inferred July temperatures lay within this range. The smoothed curves presented in Figs. 9 and 10, therefore, represent estimates of past temperature change that are encompassed by these inherent errors. Applying uniquely the inference model with the best performance (i.e. highest r^2 and lowest RMSEP) to the fossil data would ignore the cross-validation results (Table 4) as well as the different proportions of unimodal and sigmoidal taxon-environment response models (Table 3).

Each training set gave a curve of similar shape for the Splan Pond and Whitrig Bog data sets. The European training set, however, inferred a more rapid temperature change than the other data sets for the Younger Dryas cooling, but not for the Preboreal warming. At Splan Pond, this feature is mainly due to the fact that the European models infer a stronger cooling already at 170 cm, whereas the other models suggest that cooling only began at 160 cm of sediment depth. This holds true for all models applied to the Splan Pond data (Figs. 7 and 9), except the MAT model. With the MAT models, the temperatures inferred from all three training sets suggest a gradual cooling at the onset of the Younger Dryas.

For Whitrig Bog, the European WA model infers the most rapid temperature changes, at both the beginning and end of the Younger Dryas, but the North American PLS model infers more rapid temperature changes than the European PLS model (Fig. 8). With close interval sampling at Whitrig Bog, several brief climatic oscillations, perhaps correlative with the Older Dryas, Aegelsee and/or Amphi-Atlantic (Killarney and Gerzensee) Oscillations (Brooks et al., 1997b), are revealed within the Lateglacial interstadial. (Figs. 8 and 10).

Both sites and the reconstructions with all different models suggest a gradual decrease in summer temperatures at the onset of the Younger Dryas, whereas the warming at its end is more abrupt. These results are in good agreement with ice-core data that also suggest a gradual cooling at the onset of the Younger Dryas stadial and a rapid warming at the transition to the Holocene (e.g. Dansgaard et al., 1989; Taylor et al., 1993). The inferred cooling at the Allerød-Younger Dryas transition seems to be more pronounced at Whitrig Bog. Depending on the modern training set applied to the fossil chironomid data the temperature decrease at the Scottish site ranges between 4 and 5.5°C, whereas at Splan Pond a cooling of 2.5–4.5°C in July temperature
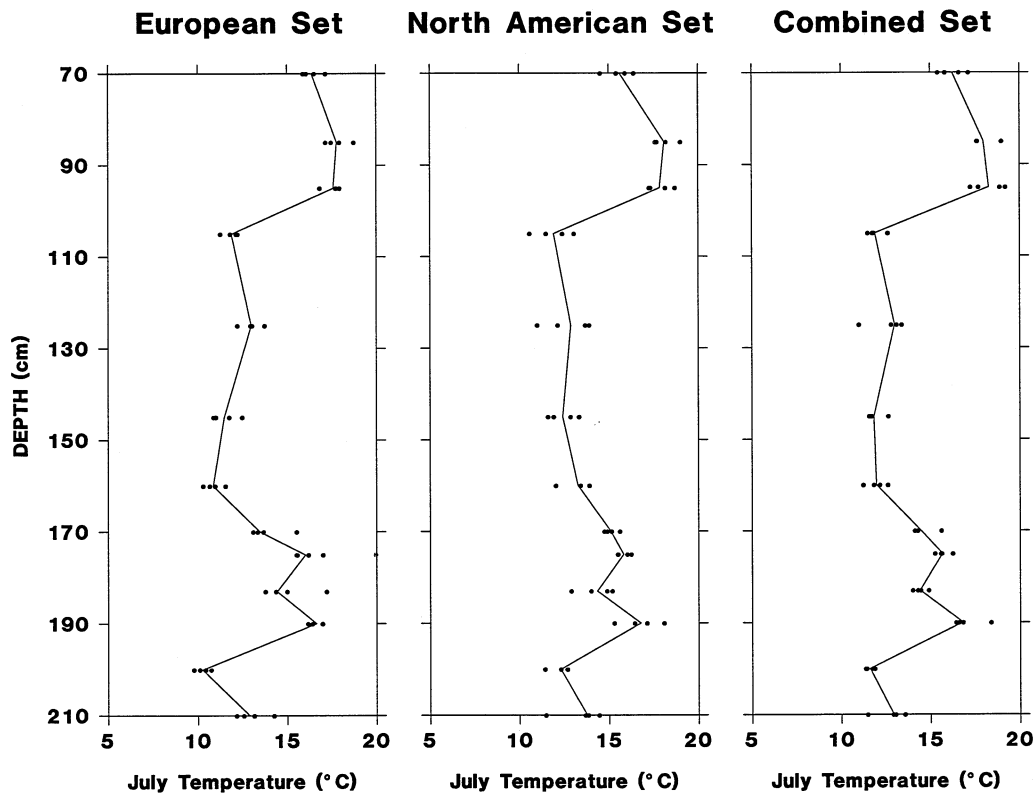
European Set North American Set Combined Set

Fig. 9. Comparison of the chironomid inferred July temperature for the Lateglacial of Splan Pond. For each training set a combined reconstruction has been carried out based on the inferences of the four models shown in Fig. 7. A LOWESS smoother was fitted through the data points.

was inferred. The warming at the Younger Dryas-Preboreal transition seems to be stronger at the North American site. At Splan Pond an increase in July temperatures between 6 and 6.5°C was inferred, whereas the European site suggested a warming in the range of 2.5–3°C. There are only very few studies using comparable quantitative climate reconstructions as of yet. Pollen and cladoceran-based inference models at Gerzensee on the Swiss Plateau reconstructed a gradual decrease in mean summer temperature of between 2 and 3°C at the onset of the Younger Dryas stadial and a more rapid warming in the order of 2°C at the Younger Dryas-Preboreal transition (Lotter *et al.*, in press). Brooks and Birks (1997) and Duigan and Birks (1997) used chironomid and cladoceran assemblages to reconstruct Younger Dryas mean summer air temperatures at Kråkenes, a site in western Norway (Birks *et al.*, 1996), where the Younger Dryas cooling was inferred to be about 2°C.

4. Conclusions

(1) Most of the chironomid taxa (67–76%) show unimodal responses to temperature, but a number of taxonomic units have sigmoidal responses (5–22%) or showed no relationship (7–28%) to July temperature.

(2) The July temperature optima estimated on the basis of North American, European, and combined training sets were similar and highly correlated. Other species parameters (i.e. tolerances and regression coefficients) tended also to be similar, but much less highly correlated.

(3) The processes of taxonomic harmonization and merger of the two training sets did not greatly affect the predictive power of the models. The combined, inter-regional model consistently predicted the temperatures at the European and North American sites better than each single, regional temperature inference model did for lakes on the other continent.

(4) Two-component WA-PLS models were consistently identified by $r^2_{(jack)}$ and $RMSEP_{(jack)}$ as being superior to other WA, WA-PLS, PLS, and MAT models. However, when the European training set was used to infer North American temperatures, and vice versa, other models provided better temperature inferences.

(5) The various inference models were able to provide good July temperature inferences, even for sites on

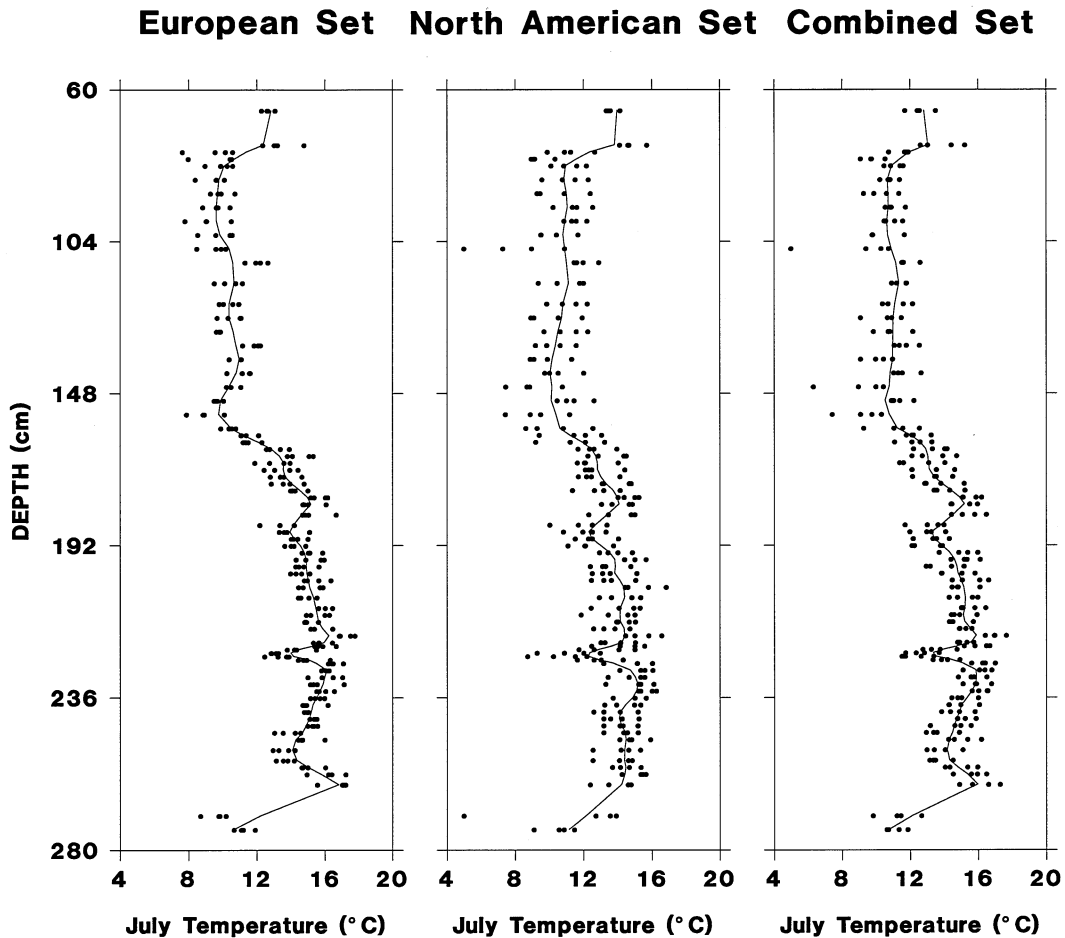## European Set   North American Set   Combined Set

Fig. 10. Comparison of the chironomid inferred July temperature for the Lateglacial of Whitrig Bog. For each training set a combined reconstruction has been carried out based on the inferences of the four models shown in Fig. 8. A LOWESS smoother was fitted through the data points.

the other continent, and where neither good nor close analogues existed. Inferences for no analogue samples were as good as those for samples having close or good analogues.

(6) When the models were applied to fossil Lateglacial data from North America and Europe, the inferred rates and magnitude of July temperature changes varied among models; however, all models revealed similar patterns of Lateglacial temperature change. This suggests that the chironomid temperature inference models provide robust means for inferring past climatic change.

Acknowledgements

We would like to thank S. Juggins for allowing us to use his unpublished inference software, and H. J. B. Birks and S. Juggins for many discussions on inference models and quantitative reconstruction of past environments. S. Juggins and an anonymous reviewer made valuable comments on an earlier version of the manuscript.

References

Ammann, B., & Lotter, A. F. (1989). Late-Glacial radiocarbon- and palynostratigraphy on the Swiss Plateau. *Boreas, 18,* 109–126.

Birks, H. J. B. (1994). The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions. *Review of Palaeobotany and Palynology, 83,* 107–117.

Birks, H. J. B. (1995). Quantitative palaeoenvironmental reconstructions. *In* D. Maddy, & J.S. Brew (Eds.), *Statistical modelling of Quaternary science data* (pp. 161–254). Cambridge: Quaternary Research Association.

Birks, H. J. B. Numerical tools in fine-resolution palaeolimnology — progress, potentialities, and problems. *Journal of Paleolimnology* (in press).

Birks, H. J. B., Line, J. M., Juggins, S., Stevenson, A. C., & ter Braak, C. J. F. (1990). Diatoms and pH reconstruction. *Philosophical Transactions Royal Society London, B 327,* 263–278.

Björck, S., Kromer, B., Johnsen, S., Bennike, O., Hammarlund, D., Lemdahl, G., Possnert, G., Rasmussen, T.L., Wohlfarth, B., Hammer, C.U., & Spurk, M. (1996). Synchronized terrestrial-atmospheric deglacial records around the North Atlantic. *Science*, *274*, 1155–1160.

Brooks, S. J., Mayle, F. E., & Lowe, J. J. (1997a). Chironomid-based Lateglacial climatic reconstruction for southeast Scotland. *Journal of Quaternary Science*, *12*, 161–167.

Brooks, S. J., Mayle, F. E., & Lowe, J. J. (1997b). The Late Devensian Lateglacial palaeoenvironmental record from Whitrig Bog, SE Scotland. 2. Chironomidae (Insecta: Diptera). *Boreas*, *18*, 297–308.

Brooks, S. J., & Birks, H. J. B. (1997). Chironomidae from a laminated lake in Norway: a sensitive archive of Lateglacial palaeoenviron-ments. *Würzburger Geographische Manuskripte*, *41*, 43–44.

Dansgaard, W., White, J. W. C., & Johnsen, S. J. (1989). The abrupt termination of the Younger Dryas climate event. *Nature*, *339*, 532–534.

Duigan, C. A., & Birks, H. H. (1997). Reconstructing temperature changes from cladoceran microfossil assemblages at the late-glacial site of Kråkenes, Western Norway. *Würzburger Geographische Manuskripte*, *41*, 63–64.

Hann, B. J., Warner, B. G., & Warwick, W. F. (1992). Aquatic invert-ebrates and climate change: a comment on Walker *et al.* (1991). *Canadian Journal of Fisheries & Aquatic Sciences*, *49*, 1274–1276.

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, *54*, 427–432.

Houde, A. (1978) *Atlas climatologique du Québec*: *température-precipita-tion*. Québec: Service de la Météorologie, Ministère des richesses naturelles,.

Huisman, J., Olff, H., & Fresno, L. F. M. (1993). A hierarchical set of models for species response analysis. *Journal of Vegetation Science*, *4*, 37–46.

Johnsen, S. J., Dahl-Jensen, D., Dansgaard, W., & Gundestrup, N. (1995). Greenland palaeotemperatures derived from GRIP bore hole temperature and ice core isotope profiles. *Tellus*, *47B*, 624–629.

Jones, V. J., & Juggins, S. (1995). The construction of a diatom-based chlorophyll *a* transfer function and its application at three lakes on Signy Island (maritime Antarctic), subject to differing degrees of nutrient enrichment. *Freshwater Biology*, *34*, 433–445.

Jouzel, J., Lorius, C., Johnsen, S., & Grootes, P. (1994). Climate instabi-lities: Greenland and Antarctic records. *C.R. Acad. Sci. Paris*, *II*, 65–77.

Levesque, A. J., Mayle, F. E., Walker, I. R., & Cwynar, L. C. (1993). The Amphi-Atlantic Oscillation: a proposed late-glacial climatic event. *Quaternary Science Reviews*, *12*, 629–643.

Line, J. M., ter Braak, C. J. F., & Birks, H. J. B. (1994). WACALIB version 3.3 — a computer program to reconstruct environmental variables from fossil assemblages by weighted averaging and to derive sample-specific errors of prediction. *Journal of Paleolimnol-ogy*, *10*, 147–152.

Livingstone, D. M., & Lotter, A. F. (1998). The relationship between air and water temperatures in lakes of the Swiss Plateau: a case study with palaeolimnological implications. *Journal of Paleolimnology*, *19*, 181–198.

Lotter, A. F., Eicher, U., Birks, H. J. B., & Siegenthaler, U. (1992). Late-glacial climatic oscillations as recorded in Swiss lake sediments. *Journal of Quaternary Science*, *7*, 187–204.

Lotter, A. F., Birks, H. J. B., Hofmann, W., & Marchetto, A. (1997). Modern diatom, cladocera, chironomids, and chrysophyte cyst as-semblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. I. Climate. *Journal of Paleolimnology*, *18*, 395–420.

Lotter, A. F., Birks, H. J. B., Eicher, U., Hofmann, W., Schwander, J., & Wick, L. Younger Dryas and Allerød summer temperatures

at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeography*, *Palaeoclimatology*, *Palaeoecology* (in press).

McCalla, R. J. (1991). *The Maritime Provinces Atlas*. Maritext, Halifax. 96 pp.

Munro, M. A. R., Kreiser, A. M., Battarbee, R. W., Juggins, S., Steven-son, A. C., Anderson, N. J., Berge, F., Davis, R. B., Flower, R. J, Fritz, S. C., Haworth, E. Y., Jones, V. J., Kingston, J. C., & Renberg, I. (1990). Diatom quality control and data handling. *Phil. Trans. R. Soc. Lond. B*, *327*, 257–261.

Olander, H., Korhola, A., & Blom, T. (1997). Surface sediment Chironomidae (Insecta: Diptera) distributions along an ecotonal transect in subarctic Fennoscandia: developing a tool for palaeotemperature reconstructions. *Journal of Paleolimnology*, *18*, 45–59.

Peteet, D. M. (1992). The palynological expression and timing of the Younger Dryas event - Europe versus eastern North America. *In*: E. Bard, & W. S. Broecker, (Eds.), *The last deglaciation* (pp. 327–344). NATO ASI Series, *I2*. Springer: Berlin.

Rossaro, B. (1991). Chironomids and water temperature. *Aquatic In-sects*, *13*, 87–98.

Sæther, O. A. (1979). Chironomid communities as water quality indi-cators. *Holarctic Ecology*, *2*, 65–74.

Taylor, K. C., Lamorey, G. W., Doyle, G. A., Alley, R. B., Grootes, P. M., Mayewski, P. A., White, J. W. C., & Barlow, L. K. (1993). The 'flickering switch' of late Pleistocene climate change. *Nature*, *361*, 432–436.

ter Braak, C. J. F. (1987). Ordination. *In*. R. H. G. Jongman, C. J. F. ter Braak, & O. F. R. van Tongeren (Eds.), *Data analysis in community and landscape ecology* (pp. 91–173). Wageningen, Pudoc.

ter Braak, C. J. F., & Juggins, S. (1993). Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstruct-ing environmental variables from species assemblages. *Hy-drobiologia*, *269/270*, 485–502.

ter Braak, C. J. F., Juggins, S., Birks, H. J. B., & van der Voet, H. (1993). Weighted averaging partial least squares regression (WA-PLS): definition and comparison with other methods for species-environ-ment calibration. *In*: G. P. Patil, & C. R. Rao (Eds.), *Multivariate Environmental Statistics* (pp. 525–560). Elsevier, Amsterdam.

Thomas, M. K. (1953) *Climatological Atlas of Canada*. Ottawa: National Research Council of Canada.

Walker, I. R., & Mathewes, R. W. (1989). Chironomid (Diptera) remains in surficial lake sediments from the Canadian Cordillera: analysis of the fauna across an altitudinal gradient. *Journal of Paleolimnology*, *2*, 61–80.

Walker, I. R., Smol, J. P., Engstrom, D. R., & Birks, H. J. B. (1991a). An assessment of Chironomidae as quantitative indicators of past cli-matic change. *Canadian Journal of Fisheries and Aquatic Sciences*, *48*, 975–987.

Walker, I. R., Mott, R. J., & Smol, J. P. (1991b). Allerød-Younger Dryas lake temperatures from midge fossils in Atlantic Canada. *Science*, *255*, 1010–1012.

Walker, I. R., Smol, J. P., Engstrom, D. R., & Birks, H. J. B. (1992). Aquatic invertebrates, climate, scale, and statistical hypothesis test-ing: a response to Hann, Warner, and Warwick. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*, 1276–1280.

Walker, I. R., & MacDonald, G. M. (1995). Distributions of Chironomidae (Insecta: Diptera) and other freshwater midges with respect to treeline, Northwest Territories, Canada. *Arctic and Alpine Research*, *27*, 258–263.

Walker, I. R., Levesque, A. J., Cwynar, L. C., & Lotter, A. F. (1997). An expanded surface-water palaeotemperature inference model for use with fossil midges from eastern Canada. *Journal of Paleolimnology*, *18*, 165–178.