

How does taxonomic resolution affect chironomid-based temperature reconstruction?

Oliver Heiri · André F. Lotter

Received: 12 August 2009 / Accepted: 2 April 2010 / Published online: 14 April 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The resolution achievable for chironomid identifications has increased in recent years because of significant improvements in taxonomic literature. However, high taxonomic resolution requires more training for analysts. Furthermore, with greater taxonomic resolution, misidentifications and the number of rare, poorly represented taxa in chironomid calibration datasets may increase. We assessed the effects of various levels of taxonomic resolution on the performance of chironomid-based temperature inference models (transfer functions) and temperature reconstruction. A calibration dataset consisting of chironomid assemblage and temperature data from 100 lakes was examined at four levels of taxonomic detail. The coarsest taxonomic resolution primarily represented identifications to genus or suprageneric level. At the highest level of taxonomic resolution, identification to genus level was possible for 37% of taxa, and identification below genus was possible for 60% of taxa. Transfer functions were obtained using Weighted Averaging (WA) and Weighted Averaging-Partial Least Squares (WA-PLS) regression.

Cross-validated performance statistics, such as the root mean square error of prediction (RMSEP) and the coefficient of determination (r^2) between inferred and observed values improved considerably from the lowest taxonomic resolution level (WA: RMSEP 1.91°C, r^2 0.78; WA-PLS: RMSEP 1.59°C, r^2 0.86) to the highest taxonomic resolution level (WA: RMSEP 1.66°C, r^2 0.84; WA-PLS: RMSEP 1.41°C, r^2 0.89). Reconstructed July air temperatures during the Lateglacial period based on fossil chironomid assemblages from Hijkermeer (The Netherlands) were similar for all levels of taxonomic resolution, except the coarsest level. At the coarsest taxonomic level, reconstruction failed to infer one of the known Lateglacial cold episodes in the record. Also, the difference in reconstructed values based on lowest and highest taxonomic resolutions exceeded sample-specific estimated standard errors of prediction in several instances. Our results suggest that chironomid-based transfer functions at the highest taxonomic resolution outperform models based on lower-resolution calibration data. However, transfer functions of intermediate taxonomic resolution produced results very similar to models based on high-resolution taxonomic data. In studies that include analysts with different levels of expertise, inference models based on intermediate taxonomic resolution, therefore, might provide an alternative to transfer functions of maximum taxonomic detail in order to ensure taxonomic consistency between calibration datasets and down-core records produced by different analysts.

O. Heiri (✉) · A. F. Lotter
Palaeoecology, Institute of Environmental Biology,
Laboratory of Palaeobotany and Palynology,
Utrecht University, Budapestlaan 4, 3584 CD Utrecht,
The Netherlands
e-mail: O.M.Heiri@uu.nl

A. F. Lotter
e-mail: A.F.Lotter@uu.nl

Keywords Fossil chironomids · Taxonomic resolution · Transfer function · Palaeoecology · Temperature reconstruction

Introduction

In aquatic ecology, indicator-species or assemblage-based approaches are widely used to calculate indices that reflect changes in environmental conditions (Kelly 1998; Bonada et al. 2006), or to quantitatively estimate environmental variables (Hämäläinen and Karjalainen 1994; Hämäläinen and Huttunen 1996). The assumption behind these approaches is that the distribution of organisms is determined by their environment, and therefore that certain species or biotic assemblages can provide information on prevailing environmental conditions. Similar approaches have been used in palaeolimnology to reconstruct past changes in water quality or temperature based on fossil assemblages of algae and invertebrates preserved in lake sediments (Birks et al. 1990; Walker et al. 1991; Lotter et al. 1997, 1998). Usually these approaches depend on the expertise of highly trained specialists to identify indicators to a high level of taxonomic resolution, and they tend to be labor intensive. An important question is whether identification to the highest possible taxonomical resolution is necessary for organism-based environmental inferences, or whether identification to coarser taxonomic levels provides results of a comparable quality (Jones 2008). Analyses at lower taxonomic resolution tend to be less time-intensive and can, therefore, increase the efficiency of organism-based biomonitoring or reconstruction methods. Furthermore, less effort has to be invested in training analysts, an aspect which is of particular relevance for research projects involving a number of analysts or laboratories.

Fossil remains of chironomid larvae have increasingly been used in recent years to produce palaeoclimate reconstructions based on lake sediment records (Heiri and Millet 2005; Brooks 2006; Walker and Cwynar 2006; Ilyashuk et al. 2009). The distribution of chironomids in the modern environment is strongly related to summer temperature, and this relationship has been used to construct chironomid-based temperature inference models, or transfer functions (Brooks and Birks 2001; Korhola et al. 2002; Heiri

et al. 2003). Since the head capsules of chironomid larvae are preserved in lake sediments as identifiable fossils, these transfer functions, when applied to fossil assemblages, can provide quantitative estimates of past variations in summer temperatures. An interesting aspect of the chironomid-temperature relationship is that it is apparent even at a comparatively coarse taxonomic resolution. The first chironomid-temperature transfer function was based on identification to the level of the tribe, sub-tribe, genus-group, and genus and was applied successfully to reconstruct Lateglacial summer temperature changes in North America (Walker et al. 1991). Later Lotter et al. (1999) showed that at this taxonomic level the distribution of chironomid taxa with respect to temperature is similar in North America and Europe, and that a transfer function developed on one continent can be applied to fossil records from the other continent with similar results. Since then identification guides have been produced that allow a progressively higher taxonomic resolution for calibration datasets and transfer functions (e.g., Rieradevall and Brooks 2001; Brooks et al. 2007).

Fossil chironomid remains retrieved from lake sediments in well-studied regions, such as the Western Palaearctic or the Afrotropical region, can usually be identified to morphotypes below genus level, or in some cases even to species (Brooks et al. 2007; Eggermont and Verschuren 2004). However, chironomids are notoriously difficult to identify to species level even if complete, non-fossilized specimens are examined, so increased taxonomic resolution bears the potential for misidentifications (Walker 2001; Brodersen 2008). It is possible, therefore, that increased taxonomic resolution might lead to no improvement or, in extreme cases, even to a reduction of the predictive power of chironomid-based transfer functions.

This study examines the effect of taxonomic resolution on the performance of chironomid-based transfer functions for temperature, both in modern environments, and when the transfer functions are used to reconstruct temperatures from fossil sequences. We utilize a calibration dataset from Central Europe that has a high level of taxonomic resolution. Based on these data we construct and evaluate inference models with maximum taxonomic detail but also with lower levels of taxonomic resolution. We then apply these transfer functions to a fossil chironomid record from

the Lateglacial period (ca. 11,000–15,000 calibrated ^{14}C years BP; cal. BP) to assess whether reconstructions are affected by differences in taxonomic resolution, and to determine how inferences agree with the known climatic development during this period.

Methods

Calibration dataset

Subfossil chironomid assemblages in the surface sediments of 114 lakes in and around the Swiss Alps and July air temperature estimates for these sites were available for calculating chironomid-based transfer functions for temperature (Lotter et al. 1997; Heiri and Lotter 2005; Bigler et al. 2006). Assemblage data were based on subfossil chironomid remains isolated from the top 1–2 cm of sediment obtained from the deepest part of the lake basins. Samples were sieved with a 100- μm sieve, the sieve residue was examined under a stereomicroscope at about 35 \times magnification, and chironomid fossils were mounted on permanent microscope slides. This calibration dataset previously has been used to develop a transfer function to reconstruct past temperature changes based on fossil chironomid assemblages (Heiri et al. 2003, 2007). For the present study microscope slides were re-examined as necessary to reach an identical, high taxonomic resolution for all assemblages. In most instances this taxonomic level is identical to the identification scheme described by Brooks et al. (2007) with morphotypes representing chironomid species, species groups or genera. A total of 14 assemblages in the calibration dataset were excluded from further numerical analyses because of ecological reasons (von Gunten et al. 2008). The final dataset, therefore, consisted of 100 assemblages from lakes spanning an altitudinal range of 418–2,815 m asl, and a July air temperature gradient of 5.0–18.4 $^{\circ}\text{C}$.

Inference models

Transfer functions were constructed based on this calibration dataset at four different taxonomic resolutions, which approximate levels of taxonomic detail used in published chironomid-based temperature inference models in the past. The calibration dataset at the coarsest resolution is based on assemblage data at the

genus, genus-group, tribe, and subfamily level as applied by Walker et al. (1991, 1997) (Table 1). At the second lowest taxonomic resolution, chironomids were identified to genus level in most cases and to morphotypes below the genus level in some instances. This dataset represents the taxonomic level achieved by Lotter et al. (1997). At the third lowest taxonomic resolution, chironomids were identified to generic level in 49% of instances, and to below genus level in 46% of instances. This dataset approximates the taxonomic level achieved by Brooks and Birks (2001). The calibration dataset at the highest taxonomic resolution was based on chironomids identified to below genus level for 60% of identified taxa. These calibration datasets will henceforth be referred to as TS1991, TS1997, TS2001, and TS2009, respectively (Table 1).

Most available chironomid-based transfer functions for temperature are based on weighted averaging (WA) or weighted averaging-partial least squares regression (WA-PLS) (e.g., Walker et al. 1991; Lotter et al. 1997; Brooks and Birks 2001; Larocque et al. 2001; Luoto 2009). To assess the influence of taxonomic resolution on transfer-function performance, we calculated both inference models based on simple WA with inverse deshrinking as well as on WA-PLS, with the number of useful components assessed following Birks (1998), and this approach was applied to each level of taxonomic resolution. Cross-validated error and performance statistics, such as the root mean square error of prediction (RMSEP) and the coefficient of determination (r^2), were calculated using bootstrapping with 999 cycles. All transfer functions were constructed based on square-root transformed percentage abundances using the program C2 version 1.4.2 (Juggins 2003). Ordinations and diversity estimates (Hill's N_2 values) were calculated with the program CANOCO 4.51 (ter Braak and Šmilauer 1998), and the statistical significance of ordination axes was assessed using 9,999 unrestricted permutations.

Fossil data

The transfer functions at four taxonomic resolutions were applied to the fossil chironomid record from Hijkermeer, the Netherlands (Heiri et al. 2007). This sequence covers the Lateglacial period from the early interstadial onwards. Chironomid-inferred temperatures from Hijkermeer indicate a very similar temperature development as reconstructed for Central

Table 1 Numerical properties of chironomid assemblages and taxa in the calibration dataset at the four different levels of taxonomic resolution

	TS1991	TS1997	TS2001	TS2009
Calibration dataset				
Number of sites	100	100	100	100
Total number of taxa	46	75	104	121
Number of taxa at supra-generic level	8 (17%)	6 (8%)	5 (5%)	4 (3%)
Number of taxa at generic level	37 (80%)	53 (71%)	51 (49%)	45 (37%)
Number of taxa below generic level	1 (2%)	16 (21%)	48 (46%)	72 (60%)
Number of taxa per site	3–23	3–35	3–38	3–40
Number of occurrences per taxon	1–100	1–76	1–68	1–68
Hill's N2 per site	2–17	2–25	2–28	2–30
Hill's N2 per taxon	1–90	1–62	1–57	1–56
Rare taxa (Hill's N2 <5)	14 (30%)	24 (32%)	42 (40%)	51 (42%)
DCA axis 1 gradient length (SD units)	2.5	3.8	4.6	5.5
DCCA axis 1 gradient length (SD units)	2.3	2.8	3.2	3.5
Variance explained by July air temperature (%)	15.7	15.0	13.5	13.3
Significance of chironomid-temperature relationship (<i>P</i> value)	0.0001	0.0001	0.0001	0.0001
Fossil dataset				
DCA axis 1 gradient length (SD units)	2.2	2.4	2.7	2.7
Total number of taxa	35	55	58	72
Number of taxa not in transfer function	3	7	11	11
Mean abundance of taxa not in transfer function	0.4%	2.5%	4.1%	4.1%
Maximum abundance of taxa not in transfer function	7.7%	20.3%	20.9%	20.9%

DCCA was calculated with July air temperature as the only constraining variable

Greenland by $\delta^{18}\text{O}$ in the Greenland ice cores. A number of well described centennial- to millennial-scale temperature oscillations are prominent in this chironomid record, such as the Younger Dryas (YD), the Gerzensee Oscillation (GO) (or Greenland interstadial event GI-1d) and the Aegelsee Oscillation (AO) (or GI-1d). This sequence, therefore, provides an opportunity to examine how taxonomic resolution affects the outcome of chironomid-based temperature reconstruction when the approach is applied to chironomid assemblages deposited during a period with several known shifts in temperature.

Results

At the highest taxonomic resolution (TS2009), the Alpine chironomid-temperature calibration dataset consists of 121 taxa with a taxonomic diversity of 3–40 taxa per site (Table 1). The distribution of most of these taxa shows a strong relationship with

temperature (Fig. 1). Detrended Correspondence Analysis (DCA) reveals a very long unconstrained compositional gradient (5.5 standard deviation units) and Detrended Canonical Correspondence Analysis (DCCA) indicates a long compositional gradient in chironomid assemblages along the range of temperature (3.5 standard deviations) (Table 1), indicating that the dataset is well suited for analysis with numerical methods assuming a unimodal response model (Birks 1998). At progressively coarser taxonomic resolution (TS2001, TS1997, TS1991) the total number of taxa, the taxa per site, the unrestricted compositional gradient length, and the compositional gradient length along the temperature gradient all decrease. However, even at the coarsest resolution, the dataset still features gradient lengths of 2.3–2.5 standard deviation units. The relationship between July air temperature and chironomid assemblages is highly significant at all levels of taxonomic resolution (Table 1). The percentage of variance explained in the chironomid assemblages by July air temperature

decreases with higher levels of taxonomic resolution. However, this can be explained by the greater overall variance in the chironomid data caused by an increase in the number of chironomid taxa from 46 (TS1991) to 121 (TS2009).

Taxonomic diversity of biotic assemblages can be assessed by examining the number of taxa per sample. Alternatively, Hill's N_2 (Hill 1973), the effective number of occurrences in a sample (Birks 1998), can be calculated. Hill's N_2 per sample decreases with coarser taxonomic resolution from 3 to 40 in TS2009 to 3–24 in TS1991. In samples with low diversity, the number of taxa per sample and Hill's N_2 are very similar at all four levels of taxonomic resolution (Fig. 2). However, in samples with a high diversity, larger differences become apparent, with both the number of taxa per sample and Hill's N_2 distinctly lower in TS1991 than in datasets with a greater taxonomic resolution (Fig. 2).

Chironomid-based temperature inference models calculated with WA and inverse deshrinking are characterized by a successively stronger relationship between chironomid-inferred and observed July air temperatures when the inference models are tested in the modern environment, with the cross-validated r^2 increasing from 0.78 for TS1991 to 0.84 for TS2009. Both the cross-validated maximum bias and the root mean square error of prediction (RMSEP) tend to be lower at greater taxonomic resolution. However, TS1997 slightly outperforms TS2001 if the r^2 or maximum bias is examined (Table 2). WAPLS inference models calculated for the calibration data at different resolution show similar results. For all four taxonomic levels, WAPLS with two components is identified as the most appropriate model following the criteria outlined by Birks (1998). Both r^2 and RMSEP improve with increasing taxonomic resolution, with the model calculated based on TS1991 featuring an r^2 of 0.86 and a RMSEP of 1.59°C and the model based on TS2009 an r^2 of 0.89 and a RMSEP of 1.41°C. In contrast, the maximum bias increases with taxonomic resolution from a value of 1.34°C in the model based on TS1991 to 1.44°C in the model based on TS2009.

If the relationship between chironomid-inferred and observed July air temperature is examined in detail for the different models, it becomes apparent that all WA-based transfer functions are characterized by a very uneven distribution of the prediction

residuals (Fig. 3a). All of the models have problems reconstructing extreme values at both ends of the temperature gradient, leading to a noticeable bias in the prediction residuals below a July air temperature of ~6°C and above ~16°C. This bias is typical for WA-based inference models and usually referred to as the 'edge effect' (Birks 1998). In addition, the WA-based models show a distinctly higher variability of the prediction residuals in the cooler half of the temperature gradient than in the warmer sections. In the WA-PLS based models (Fig. 3b), both the edge effect and the changing variability of residuals along the temperature gradient are noticeably reduced. However, on the cold end of the temperature gradient, all WA-PLS models still have a clear tendency to overestimate July air temperature.

When the different transfer functions are applied to the fossil chironomid record from Hijkermeer, a similar range of July air temperatures of ~12–17°C is inferred by all models and at all taxonomic resolutions (Fig. 4). Reconstructions based on WA and TS1997, TS2001, and TS2009 all show a very similar pattern of temperature changes, with pronounced cold phases inferred during the AO, GO, and the YD, with the latter episode characterized by the most pronounced cooling. In the WA-based reconstruction based on TS1991, cooler temperatures also are reconstructed during both the AO and YD, whereas no cooling is apparent during the GO. Furthermore, an additional, short-lived cooling is inferred at 1,105 cm, which is not apparent in the records based on higher taxonomic resolution. The results using WA-PLS are very similar to those using WA, at least for reconstructions based on TS1997, TS2001, and TS2009. However, the cooling during the GO is more pronounced than is suggested by WA-based models. Furthermore, the WA-PLS based reconstruction using the TS1991 data features very high between-sample variability, suggesting that noise in the Hijkermeer record may be affecting the reconstruction. The GO is not very distinct in this reconstruction, and a centennial-scale temperature oscillation is again inferred by this model, which is not apparent in the other reconstructions.

We examined differences among the reconstructions calculated for Hijkermeer at different taxonomic resolutions in the context of the sample-specific estimated standard errors of prediction (eSEP). Differences between inferences based on WA and either

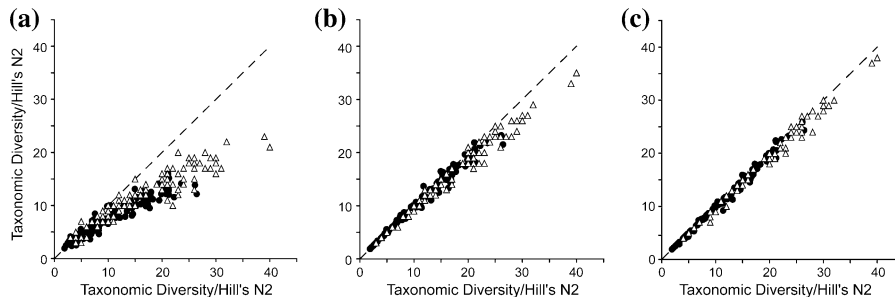


Fig. 2 Total species richness (*triangles*) and Hill's N2 values (*circles*) of assemblages in the Alpine chironomid-temperature calibration dataset at highest taxonomic resolution (TS2009) plotted against species richness and Hill's N2 of lower resolution datasets: **a** TS1991, **b** TS1997, **c** TS2001

Table 2 Performance statistics of Weighted Averaging (WA) and Weighted Averaging—Partial Least Squares (WA-PLS) transfer functions developed based on calibration datasets of different taxonomic resolution

Dataset	Method	Components	r^2	RMSEP (°C)	Maximum bias (°C)
TS1991	WA	—	0.78	1.91	2.01
TS1991	WA-PLS	1	0.78	1.92	1.99
TS1991	WA-PLS	2	0.86*	1.59*	1.34*
TS1991	WA-PLS	3	0.87	1.61	1.06
TS1997	WA	—	0.81	1.77	1.79
TS1997	WA-PLS	1	0.81	1.77	1.84
TS1997	WA-PLS	2	0.87*	1.50*	1.32*
TS1997	WA-PLS	3	0.88	1.51	1.11
TS2001	WA	—	0.80	1.81	1.85
TS2001	WA-PLS	1	0.80	1.82	1.90
TS2001	WA-PLS	2	0.87*	1.50*	1.49*
TS2001	WA-PLS	3	0.89	1.46	1.20
TS2009	WA	—	0.84	1.66	1.52
TS2009	WA-PLS	1	0.83	1.67	1.56
TS2009	WA-PLS	2	0.89*	1.41*	1.44*
TS2009	WA-PLS	3	0.90	1.37	1.10

The asterisk (*) indicates the optimal WA-PLS model selected based on the criteria outlined in Birks (1998)

Discussion

Chironomid-based temperature reconstruction relies on the strong relationship between summer temperature and the composition of chironomid assemblages in the modern environment. The method assumes that a similar relationship between temperature and chironomid taxa (species, morphotypes, genera, or higher taxonomic units) existed in the past and that chironomid assemblages responded to changing temperatures by shifts in taxonomic composition. For chironomids, it has been shown that closely related species often colonize habitats with similar environmental conditions. For example, species of *Diamesa* typically are restricted to cool, running water habitats (Rossaro et al.

2006) and many species of the genus *Micropsectra* to cold lakes or cool stream and spring habitats (Säwedal 1982; Brooks and Birks 2001). Such restricted distributions are also apparent at higher taxonomic levels of the Chironomidae. Most members of the subfamily Diamesinae, for instance, are restricted to cold habitats, whereas the tribe Chironomini is usually found at higher diversity in warmer lakes, ponds, streams, and rivers (Lindegaard 1995; Boggero et al. 2006). However, closely related species in the Chironomidae can also show distinct differences in their distribution. In Scandinavian lakes, for example, different species of *Heterotrissocladius* show a marked difference in distribution relative to temperature, with *H. marcidus*, *H. grimshawi*, *H. maeaeri*, and *H. subpilosus* generally

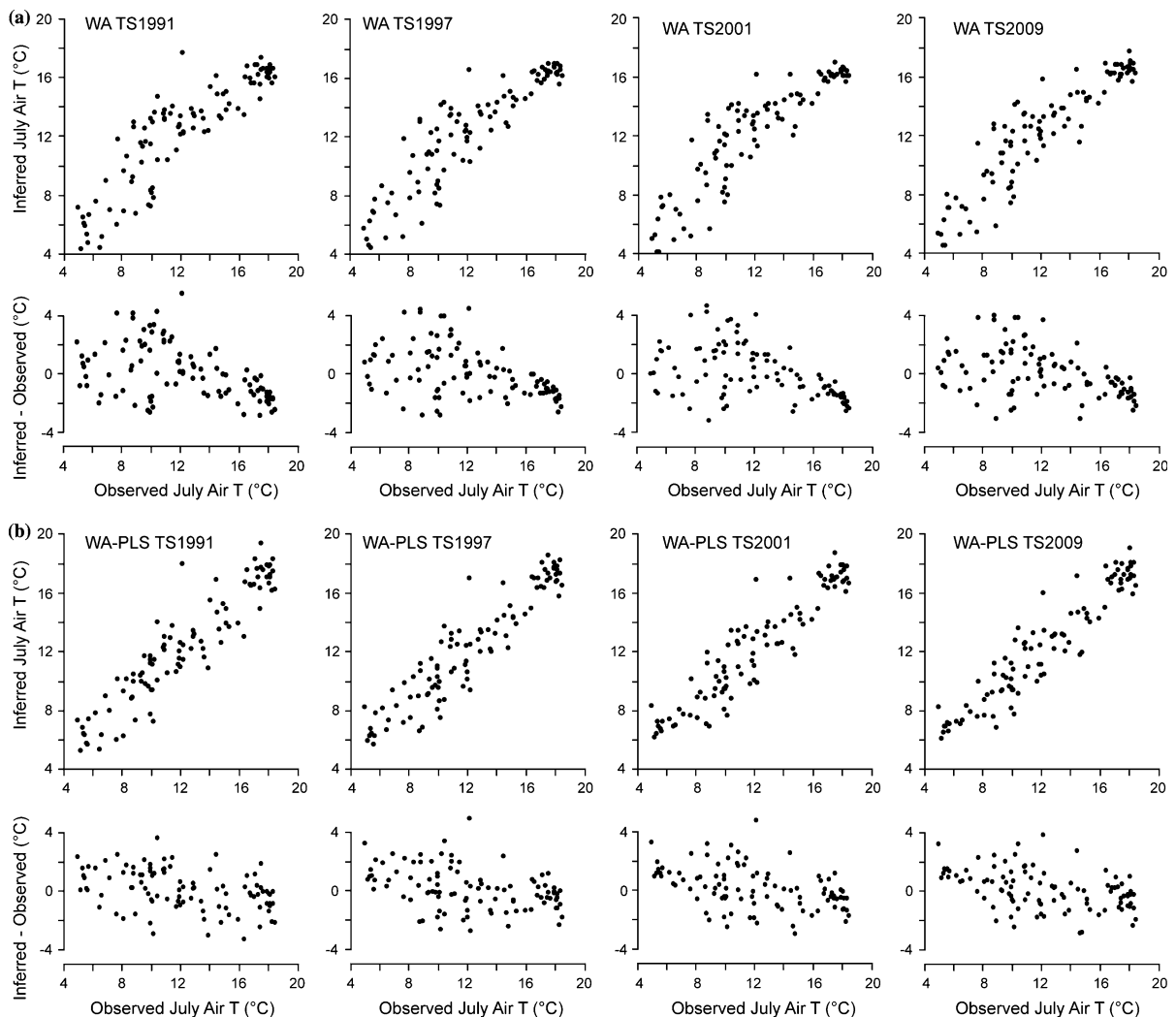


Fig. 3 Inferred July air temperatures and prediction residuals plotted against observed July air temperatures for transfer functions based on weighted averaging (WA) (a) and weighted averaging partial least squares (WA-PLS) (b) and the Alpine

chironomid-temperature calibration dataset at four different taxonomic resolutions (TS1991, TS1997, TS2001, TS2009). Inferred values and prediction residuals are cross-validated by bootstrapping (999 cycles)

found in progressively cooler lakes (Brooks and Birks 2001). Different species of *Tanytarsus* or *Chironomus* also are known to have distinctly different requirements with respect to nutrient conditions in lakes (Saether 1979). One might expect, therefore, that higher taxonomic resolution would allow differences in distributions of closely related chironomids to be resolved in calibration datasets and lead to improved performance of chironomid-based transfer functions. However, identification of chironomid larvae can be difficult, even for experienced analysts (Boggero et al. 2006). For fossil chironomid remains, identification is

further complicated by damaged specimens and the absence of morphological traits used for identifying modern chironomids (Brooks et al. 2007). Walker (2001) has drawn attention to this problem and indicated that with increasing taxonomic resolution, the chance of misidentification may increase. As a consequence, an increase in taxonomic effort does not necessarily lead to increased performance of chironomid-based transfer functions. Recently, Brodersen (2008) reviewed problems associated with misidentifications of fossil chironomids and indicated that very high taxonomic resolution may also lead to the separation of

Fig. 4 Chironomid-inferred July air temperatures for the Lateglacial sediment sequence from Hijkermeer, the Netherlands (Heiri et al. 2007) reconstructed with transfer functions based on WA (a) and WA-PLS (b) and the Alpine chironomid-temperature calibration dataset at four different levels of taxonomic resolution (TS1991, TS1997, TS2001, TS2009). AO, GO, and YD indicate Lateglacial cold oscillations apparent in the Hijkermeer record (see text for details)

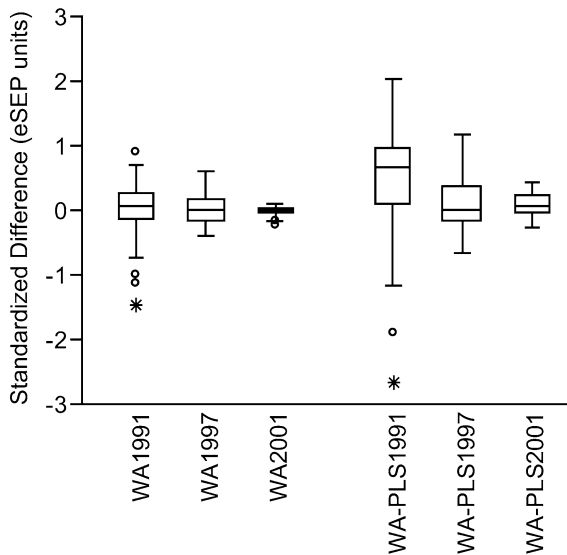
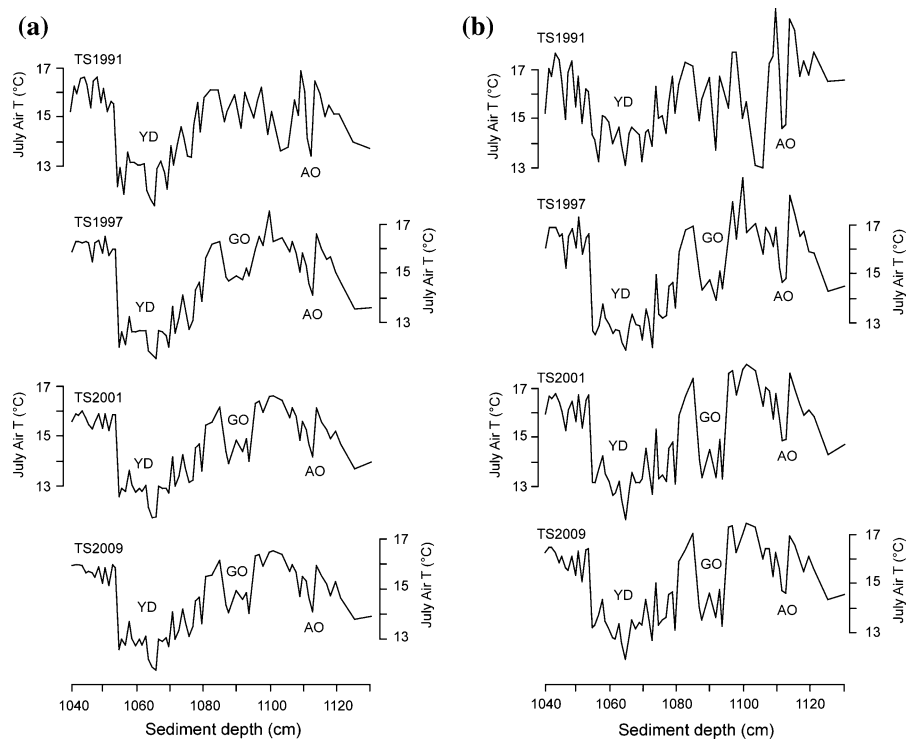


Fig. 5 Differences between chironomid-inferred July air temperatures for the Hijkermeer record based on the calibration dataset at highest taxonomic resolution and values reconstructed based on the calibration data at lower taxonomic resolution (TS1991, TS1997, TS2001). The differences are standardized by the sample-specific estimated standard error of prediction (eSEP) of the individual inferences

morphotypes that do not actually represent distinct species, but instead represent variation in morphological traits within a species. In addition to problems associated with misidentifications and increasing effort necessary for training of analysts, higher taxonomic resolution also will result in a higher proportion of rare taxa in calibration datasets. For example, at the coarsest taxonomic resolution (TS1991), the Alpine chironomid-temperature dataset contains 30% of chironomid taxa that can be considered rare if a threshold of $N_2 < 5$ is used to define rare taxa (Table 1). At the highest taxonomic resolution (TS2009), 42% of taxa are rare, according to this definition. Since the relationship to the environment is less well constrained for rare taxa than for abundant taxa, this is an additional source of uncertainty that may reduce the performance of chironomid-based transfer functions if taxonomic resolution is increased.

The transfer functions based on various taxonomic resolutions of the Alpine chironomid-temperature calibration dataset have been evaluated using cross-validation (bootstrapping). The cross-validated performance statistics (r^2 , RMSEP) indicate that the

transfer functions with the highest taxonomic resolution outperform models based on coarser resolution for both WA and WA-PLS (Table 2). It seems, therefore, that confounding effects of higher taxonomic resolution did not outweigh the benefits of more information on the distribution of chironomids along the temperature gradient. For all four levels of taxonomic resolution, WA-PLS outperformed WA, with a decrease in the RMSEP of 15–17% (Table 2). A noticeable decrease in RMSEP of both WA and WA-PLS models was apparent if taxonomic resolution increased from the level of TS1991 to TS1997, and from the level of TS2001 to TS2009. Interestingly though, the RMSEP and other performance statistics remain almost unchanged if the resolution is increased from the level of TS1997 to TS2001.

When applied to the fossil chironomid assemblages from Hijkermeer, the transfer functions based on TS1997, TS2001, and TS2009 produce reconstructions that clearly record all three known cold events (GO, AO, and YD) in the sequence. Minor differences, nevertheless, are apparent in the reconstructions. However, the maximum difference between inferred values for any given sample is 1.04 and 1.96°C for WA and WA-PLS based reconstructions, respectively, and 2.44°C if both WA and WA-PLS based inferences are compared for the three different taxonomic resolutions. In contrast, there are marked differences between reconstructions based on TS1991 and the other three datasets. The GO is not recognized as a cold oscillation if transfer functions based on TS1991 are applied, whereas a temperature oscillation is apparent in the record that is not shown by other reconstructions based on greater taxonomic detail (Fig. 4). If reconstructions at all four taxonomic resolutions are compared, the maximum difference between inferences for any given fossil sample are 3.23 and 4.56°C for WA and WA-PLS, respectively. If differences between the reconstructions are examined relative to sample-specific errors calculated for fossil samples, it becomes apparent that differences between WA- and WA-PLS-based reconstructions at the resolution of TS1997, TS2001, and TS2009 are mostly within the sample specific eSEP (Fig. 5). However, differences between inferences based on TS1991 and TS2009 are larger and clearly exceed the eSEP in a number of instances.

The effects of taxonomic resolution on palaeoenvironmental inferences have rarely been examined. Birks (1994) examined the effects of taxonomic precision on quantitative palaeoenvironmental reconstructions based on diatom and pollen assemblages. Using two modern calibration datasets he examined the consequences of deleting rare taxa from the datasets for the prediction error of transfer functions based on WA. Birks (1994) concluded that the transfer functions performed best if all taxa were included unless WA was calculated with tolerance down-weighting. In a second study, Finkelstein et al. (2006) indicated that species-level identification of fossil pollen allowed more complete reconstruction of the vegetation history of eastern North America than records identified to genus-level.

The effect of taxonomic resolution on environmental assessments has been a major focus in biomonitoring, where communities of aquatic organisms provide the basis for assessing the quality of streams, rivers, and lakes. Heino and Soininen (2007) examined how identification efforts affect assessments of community structure and taxonomic richness based on stream macroinvertebrates and diatoms from Finland. They indicated that very similar results were produced by analyses at species, genus, and family level, although for diatoms, the number of families is probably too low to adequately capture biodiversity patterns. Lane (2007) examined the performance of autecological indices based on diatoms identified to genus, species, and subspecies level in assessing the biotic integrity of isolated herbaceous wetlands in Florida. The author concluded that all three taxonomic levels provided very similar results. Jones (2008) reviewed how taxonomic resolution of benthic macroinvertebrate data affects bioassessment of freshwater ecosystems. He indicated that, despite the many studies that examined this problem for various datasets and biotic indices, it is difficult to find consensus on whether identification to species or to higher taxonomic levels produce comparable results. He concluded that only species-level identification ensures that maximum information content is available for interpreting bioassessment results, although it may sometimes be necessary to resort to higher taxonomic levels ('taxonomic minimalism') if finances, time, equipment, or expertise are limiting factors.

The effect of increasing taxonomic resolution on the Alpine calibration dataset we used is perhaps best illustrated for the tribe Tanytarsini. In the first chironomid-based temperature transfer function (Walker et al. 1991), chironomid head capsules belonging to this tribe were separated into the three categories Tanytarsina, *Stempellinella/Zavrelia*, and *Stempellina*. At this initial stage, more detailed identification of the subtribe Tanytarsina was not possible because specialized keys for subfossil Tanytarsini were not available in the early 1990s. Since then, descriptions for a number of morphotypes within the Tanytarsina have been produced, allowing more detailed identification of this group. If in the Alpine calibration dataset the Tanytarsini are identified to the three taxonomic groups mentioned above, the Tanytarsina should be considered a taxon that represents intermediate temperatures, with a July air temperature optimum of ~12°C as estimated by WA, compared with the overall July temperature gradient of 5.0–18.4°C. *Stempellinella/Zavrelia* and *Stempellina* are both more common in warmer lakes of the temperature gradient (Fig. 6). However, if the Tanytarsina are identified to the highest taxonomic resolution, it becomes apparent that different morphotypes exist that are typical for different sections of the temperature gradient. For example, *Micropsectra radialis*-type and *Paratanytarsus austriacus*-type feature WA-based July air temperature optima of ~7 and 9°C, whereas *Tanytarsus lactescens*-type and *T. glabrescens*-type are both assigned optima of ~17°C.

Conclusions

The Alpine chironomid July air temperature transfer function has the smallest prediction error at the highest taxonomic resolution. At intermediate taxonomic resolution (TS1997, TS2001), both the RMSEP and other performance statistics are slightly, but noticeably less favorable. However, our example based on the Hijkermeer record suggests that down-core reconstructions with the calibration datasets at all three taxonomic resolutions are still very similar. In contrast, reconstructions using the coarsest taxonomic resolution show significant differences compared to the records with more taxonomic detail, and the transfer functions based on TS1991 feature the highest RMSEP. A conclusion of our study, therefore, is that

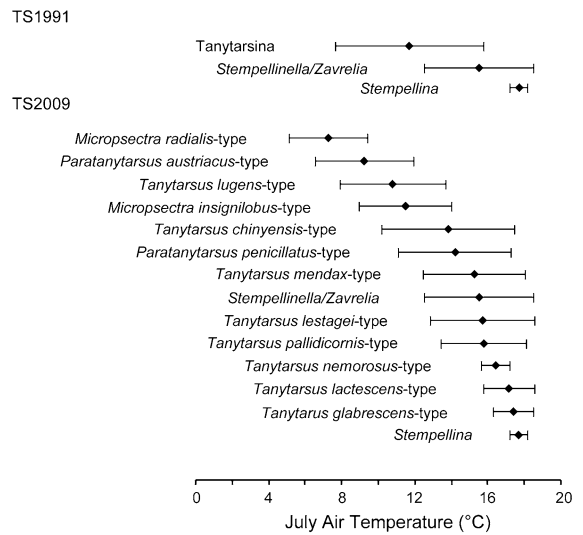


Fig. 6 Weighted averaging optima (diamonds) and tolerances (bars) for chironomid taxa within the tribe Tanytarsini calculated for the calibration dataset at coarsest (TS1991) and at highest taxonomic resolution (TS2009)

whenever possible, the highest achievable taxonomic resolution should be used for chironomid-based temperature transfer functions and down-core reconstructions. In our study all samples were analyzed by the same analyst. Increasingly, chironomid-based temperature inference models are applied to down-core records that have not been identified by the same analyst (Heiri and Millet 2005; Larocque et al. 2009; Larocque-Tobler et al. 2009), or regional calibration datasets by various analysts are combined to cover a broader environmental range (Barley et al. 2006). At very high taxonomic resolution, the possibility of taxonomic inconsistencies among analysts increases. Therefore, it might be feasible to decrease taxonomic resolution to intermediate levels in studies that involve analysts with a range of skill levels. This would decrease the possibility of misidentifications and inconsistencies while allowing chironomid-based temperature reconstructions to be consistent with inferences based on higher taxonomic resolution.

Acknowledgments We thank the participants of the fossil chironomid workshop 2007 in Reykjavik, Iceland, and especially Klaus Brodersen and Ian Walker, for stimulating discussions on the effects of taxonomic resolution on palaeoecological inferences based on fossil chironomids. We also thank Isabelle Larocque, Tom Whitmore and two anonymous reviewers for valuable comments and suggestions on earlier versions of the manuscript. The research presented in

this article was supported by the Netherlands Organization for Scientific Research (NWO)/Aard- en Levenswetenschappen (ALW) (Grant no. 818.01.001), by the European Commission's Research Infrastructure Action via the SYNTHESYS Project (GB-TAF-114), and the project "European climate change at the end of the Last Glaciation (EUCLIM)". This is Netherlands Research School of Sedimentary Geology (NSG) publication no. 20100202.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Barley EM, Walker IR, Kurek J, Cwynar LC, Mathewes RW, Gajewski K, Finney B (2006) A northwest North American training set: distribution of freshwater midges in relation to air temperature and lake depth. *J Paleolimnol* 36:295–314
- Bigler C, Heiri O, Krskova R, Lotter AF, Sturm M (2006) Distribution of diatoms, chironomids and cladocera in surface sediments of thirty mountain lakes in south-eastern Switzerland. *Aquat Sci* 68:154–171
- Birks HJB (1994) The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions. *Rev Palaeobot Palynol* 83:107–117
- Birks HJB (1998) Numerical tools in palaeolimnology—progress, potentialities, and problems. *J Paleolimnol* 20:307–332
- Birks HJB, Line JM, Juggins S, Steveson AC, ter Braak CJF (1990) Diatoms and pH reconstruction. *Philos Trans Roy Soc Lond B* 327:263–278
- Boggero A, Fureder L, Lencioni V, Simcic T, Thaler B, Ferrarese U, Lotter A, Ettinger R (2006) Littoral chironomid communities of Alpine lakes in relation to environmental factors. *Hydrobiologia* 562:145–165
- Bonada N, Prat N, Resh VH, Statzner B (2006) Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Ann Rev Entomol* 51:495–523
- Brodersen KP (2008) Book review: S. J. Brooks, P. G. Langdon and O. Heiri, The identification and use of palaeartic Chironomidae larvae in palaeoecology. QRA technical guide No. 10. *J Paleolimnol* 40:751–753
- Brooks SJ (2006) Fossil midges (Diptera: Chironomidae) as palaeoclimatic indicators for the Eurasian region. *Quat Sci Rev* 25:1894–1910
- Brooks SJ, Birks HJB (2001) Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. *Quat Sci Rev* 20:1723–1741
- Brooks SJ, Langdon PG, Heiri O (2007) The identification and use of Palaeartic Chironomidae larvae in palaeoecology. *Quat Res Assoc Tech Guid* 10:1–276
- Eggermont H, Verschuren D (2004) Sub-fossil Chironomidae from East Africa. 2. Chironominae (Chironomini and Tanytarsini). *J Paleolimnol* 32:413–455
- Finkelstein SA, Gajewski K, Viau AE (2006) Improved resolution of pollen taxonomy allows better biogeographical interpretation of post-glacial forest development: analyses from the North American Pollen Database. *J Ecol* 94:415–430
- Hämäläinen H, Huttunen P (1996) Inferring the minimum pH of streams from macroinvertebrates using weighted averaging regression and calibration. *Freshw Biol* 36:697–709
- Hämäläinen H, Karjalainen J (1994) An assessment of Harpacticoida and resting stages of Cyclopoida as trophic indicators in Finnish lakes. *Ann Zool Fenn* 31:377–388
- Heino J, Soiminen J (2007) Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organisms? *Biol Conserv* 137:78–89
- Heiri O, Lotter AF (2005) Holocene and Lateglacial summer temperature reconstruction in the Swiss Alps based on fossil assemblages of aquatic organisms: a review. *Boreas* 34:506–516
- Heiri O, Millet L (2005) Reconstruction of Late Glacial summer temperatures from chironomid assemblages in Lac Lautrey (Jura, France). *J Quat Sci* 20:33–44
- Heiri O, Lotter AF, Hausmann S, Kienast F (2003) A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. *Holocene* 13:477–484
- Heiri O, Cremer H, Engels S, Hoek WZ, Peeters W, Lotter AF (2007) Lateglacial summer temperatures in the Northwest European lowlands: a chironomid record from Hijkermeer, the Netherlands. *Quat Sci Rev* 26:2420–2437
- Hill MO (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–432
- Ilyashuk B, Gobet E, Heiri O, Lotter AF, van Leeuwen JFN, van der Knaap WO, Ilyashuk E, Oberli F, Ammann B (2009) Lateglacial environmental and climatic changes at the Maloja Pass, Central Swiss Alps, as recorded by chironomids and pollen. *Quat Sci Rev* 28:1340–1353
- Jones FC (2008) Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environ Rev* 16:45–69
- Juggins S (2003) C2 User guide. Software for ecological and palaeoecological data analysis and visualization. University of Newcastle, Newcastle upon Tyne
- Kelly MG (1998) Use of the trophic diatom index to monitor eutrophication in rivers. *Water Res* 32:236–242
- Korhola A, Vasko K, Toivonen H, Olander H (2002) Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modeling. *Quat Sci Rev* 21:1841–1860
- Lane CR (2007) Assessment of isolated wetland condition in Florida using epiphytic diatoms at genus, species, and subspecies taxonomic resolution. *EcoHealth* 4:219–230
- Larocque I, Hall RI, Grahm E (2001) Chironomids as indicators of climate change: a 100-lake training set from a subarctic region of northern Sweden (Lapland). *J Paleolimnol* 26:307–322
- Larocque I, Grosjean M, Heiri O, Bigler C, Blass A (2009) Comparison between chironomid-inferred July temperatures and meteorological data AD 1850–2001 from varved Lake Silvaplana, Switzerland. *J Paleolimnol* 41:329–342
- Larocque-Tobler I, Heiri O, Wehrli M (2009) Late Glacial and Holocene temperature changes at Egelsee, Switzerland,

- reconstructed using subfossil chironomids. *J Paleolimnol*. doi:10.1007/s10933-009-9358-z
- Lindgaard C (1995) Classification of water-bodies and pollution. In: Armitage PD, Cranston PS, Pinder LCV (eds) *The Chironomidae—the biology and ecology of non-biting midges*. Chapman & Hall, London, pp 385–404
- Lotter AF, Birks HJB, Hofmann W, Marchetto A (1997) Modern diatom, cladocera, chironomid, and chrysophyte cyst assemblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. I. Climate. *J Paleolimnol* 18:395–420
- Lotter AF, Birks HJB, Hofmann W, Marchetto A (1998) Modern diatom, cladocera, chironomid, and chrysophyte cyst assemblages as quantitative indicators for the reconstruction of past environmental conditions in the Alps. II. Nutrients. *J Paleolimnol* 19:443–463
- Lotter AF, Walker IR, Brooks SJ, Hofmann W (1999) An intercontinental comparison of chironomid paleotemperature inference models: Europe vs. North America. *Quat Sci Rev* 18:717–735
- Luoto TP (2009) Subfossil Chironomidae (Insecta: Diptera) along a latitudinal gradient in Finland: development of a new temperature inference model. *J Quat Sci* 24:150–158
- Rieradevall M, Brooks SJ (2001) An identification guide to subfossil Tanyptodinae larvae (Insecta: Diptera: Chironomidae) based on cephalic setation. *J Paleolimnol* 25: 81–99
- Rossaro B, Lencioni V, Boggero A, Marziali L (2006) Chironomids from Southern Alpine running waters: ecology, biogeography. *Hydrobiologia* 562:231–246
- Saether OA (1979) Chironomid communities as water quality indicators. *Holarct Ecol* 2:65–74
- Säwedäl L (1982) Taxonomy, morphology, phylogenetic relationships and distribution of *Micropsectra* Kieffer, 1909 (Diptera: Chironomidae). *Ent scand* 13:371–400
- ter Braak CJF, Šmilauer P (1998) *CANOCO reference manual and user's guide to Canoco for windows*. Centre for Biometry, Wageningen
- von Gunten L, Heiri O, Bigler C, van Leeuwen J, Casty C, Lotter AF, Sturm M (2008) Seasonal temperatures for the past 400 years reconstructed from diatom and chironomid assemblages in a high-altitude lake (Lej da la Tscheppa, Switzerland). *J Paleolimnol* 39:283–299
- Walker IR (2001) Midges: Chironomidae and related Diptera. In: Smol JP, Birks HJB, Last WM (eds) *Tracking environmental change using lake sediments*. Volume 4. Zoological indicators. Kluwer, Dordrecht, pp 43–66
- Walker IR, Cwynar LC (2006) Midges and palaeotemperature reconstruction—the North American experience. *Quat Sci Rev* 25:1911–1925
- Walker IR, Smol JP, Engstrom DR, Birks HJB (1991) An assessment of Chironomidae as quantitative indicators of past climatic change. *Can J Fish Aquat Sci* 48:975–987
- Walker IR, Levesque AJ, Cwynar LC, Lotter AF (1997) An expanded surface-water paleotemperature inference model for use with fossil midges from eastern Canada. *J Paleolimnol* 18:165–178