

# Effects of within-lake variability of fossil assemblages on quantitative chironomid-inferred temperature reconstruction

O. Heiri<sup>a, b, c</sup>, H. J. B. Birks<sup>b, d</sup>, S. J. Brooks<sup>e</sup>, G. Velle<sup>c</sup> and E. Willassen<sup>c</sup>

<sup>a</sup> Botanical Palaeoecology, Laboratory of Palaeobotany and Palynology, Utrecht University, Budapestlaan 4, 3584 CD, Utrecht, The Netherlands

<sup>b</sup> Botanical Institute, University of Bergen, Allégaten 41, N-5007, Bergen, Norway

<sup>c</sup> Museum of Zoology, University of Bergen, Muséplass 3, N-5007, Bergen, Norway

<sup>d</sup> Environmental Change Research Centre, University College London, 26 Bedford Way, London WC1H 0AP, UK

<sup>e</sup> Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

Received 8 November 2002; accepted 2 June 2003. ; Available online 26 July 2003.

## Abstract

We studied the effects of within-lake variation of subfossil chironomid assemblages on chironomid-inferred temperatures in five lakes in southern Norway. In each lake six replicate surface-sediment samples in the deepest part of the lake basin and two transects of seven samples each towards the littoral were analysed for subfossil chironomid assemblages. A quantitative chironomid–July air temperature inference model applied to the subfossil assemblages inferred temperatures with a standard deviation of 0.35–0.54°C for the six mid-lake samples and of 0.38–0.59°C for all 20 sediment samples. This variation seems high when compared with the cross-validated root mean square error of prediction (RMSEP) of the model of 1.04°C. However, we show that the variability of the replicate mid-lake samples can only account for a maximum of 15% of the model prediction error if the RMSEP is decomposed into a component due to uncertainties associated with sampling the chironomid assemblage of a lake and into a residual error component. Thus, according to our results analysing multiple samples in the model lakes to obtain better estimates of their subfossil assemblage composition may only slightly reduce the model prediction error. It was difficult to detect a clear pattern of chironomid-inferred temperatures with respect to water depth in any single lake due to the high variability of inferred values. However, a comparison of all five study lakes revealed that, relative to the mid-lake samples, cooler temperatures were inferred at intermediate depths and warmer temperatures in the shallowest parts of the lakes, although two shallow-water samples with very low inferred temperatures were exceptions in this respect. This depth-related bias of inferred temperatures in our relatively shallow study lakes indicates that chironomid–temperature models calibrated on mid-lake samples should only be applied with caution to near-shore sediments and that a systematic offset in the inferred temperatures may occur.

**Author Keywords:** subfossil chironomids; temperature reconstruction; prediction error; within-lake variability; palaeolimnology

## 1. Introduction

Chironomids (Chironomidae) are a family of the insect order Diptera (true flies). Their larvae are abundant in lakes, where they inhabit the surfaces of hard substrates such as rocks, plants and wood, or burrow into the sediments. The chironomid fauna of a lake usually includes a large number of different species, many with specific ecological requirements. Therefore, chironomids have a long tradition of being used as biological indicators of water quality (e.g. [Thienemann, 1913, Sæther, 1979 and Wiederholm, 1981]). The remains of chironomid larvae, especially the strongly sclerotised head capsules, are found abundantly in lake sediments. These subfossils remain identifiable, usually to a generic or species group level, and preserve well in the sediments over thousands of years. Thus, subfossil chironomid analysis can be used to reconstruct the past chironomid fauna of lakes and the environment that the larvae lived in (e.g. [Walker, 1987 and Walker, 2001]).

Recently, subfossil chironomid assemblages in lake sediments have been increasingly used to generate reconstructions of past summer temperatures using quantitative chironomid–temperature inference models (e.g. [Walker et al., 1991, Levesque et al., 1993, Brooks and Birks, 2000a and Brooks and Birks, 2000b]). These models are based on surface-sediment samples from a large number of lakes covering the temperature gradient of interest for reconstruction. In each lake, a single sediment core is taken from the deepest part of the lake basin and the subfossil chironomid assemblage in the topmost sediment is analysed. Summer water temperatures are measured on site during field work or, alternatively, the mean summer air temperature is interpolated for each lake using temperature measurement series from nearby meteorological stations. These data are then used to develop a statistical model which predicts summer temperature based on the subfossil chironomid assemblages. Chironomid–summer air temperature inference models are now available with coefficients of determination ( $r^2$ ) of 0.74–0.94 and a leave-one-out cross-validated prediction error (root mean square error of prediction; RMSEP) of 0.80–1.54°C (e.g. [Lotter et al., 1997, Lotter et al., 1999, Vasko et al., 2000 and Brooks and Birks, 2001]). Probably due to problems associated with obtaining good summer water temperature estimates from single spot measurements, chironomid–water temperature inference models tend to perform slightly worse, with a  $r^2$  of 0.32–0.88 and a RMSEP of 1.53–2.26°C ([Walker et al., 1997, Olander et al., 1999, Brooks and Birks, 2001 and Larocque et al., 2001]).

Chironomid-based summer temperature reconstruction has become a valuable tool to assess the magnitude of the large climatic fluctuations during the Late Glacial (e.g. [Walker et al., 1991, Cwynar and Levesque, 1995, Brooks and Birks, 2000a and Brooks and Birks, 2000b]). However, recently there has been an

increasing focus of palaeoclimate research on the more stable climate during the Holocene (i.e. the past 11500 years). One of the problems with applying chironomid–temperature inference models to the Holocene is that the prediction error of the models is of the same order of magnitude as the expected temperature changes (i.e. 1–2°C). Even though chironomid-based Holocene temperature reconstructions have produced promising results (e.g. [Pellatt et al., 2000, Rosén et al., 2001 and Bigler et al., 2002]), it would greatly improve the usefulness of subfossil chironomid analysis to Holocene climate research if the prediction errors of the models could be reduced.

The RMSEP usually used to express the uncertainty of chironomid-inferred temperatures is a summary statistic which incorporates many different error sources. It is equivalent to the standard deviation of the leave-one-out cross-validated prediction residuals calculated for all the model lakes. One of the factors which may potentially contribute to the RMSEP is the within-lake variability of subfossil chironomid assemblages in the lakes used to develop the model. Benthic macroinvertebrates commonly show a random or contagious (clumped) distribution in lakes. Thus, a large number of samples have to be analysed to obtain useful estimates of population parameters and biotic indices in surveys assessing the modern benthic fauna (e.g. [Johnson, 1995]). It is reasonable to assume that the subfossil chironomid assemblage within the deepest part of a lake basin also shows variation and that this variability will contribute to the overall prediction error of the temperature inference models.

For other biological palaeoenvironmental proxies, such as diatoms or pollen, detailed information on their distribution in lake sediments is available (e.g. [Meriläinen, 1969, Davis et al., 1971 and Anderson, 1990]). However, little is known about the within-lake variability of subfossil chironomid assemblages in the deepest parts of the small and relatively shallow lakes commonly used for organism-based quantitative climate reconstruction.

Studies of different climate proxies in the same sediment core can greatly enhance the reliability of organism-based climate reconstructions, as this will enable a more precise reconstruction of the palaeoenvironments within a lake and its catchment (e.g. [Birks et al., 2000 and Lotter et al., 2000]). Furthermore, palaeoclimatic inferences derived from biota with a range of different biologies and habitats can be compared to either support each other or point towards inconsistencies in the climate reconstructions. Within the framework of these multi-proxy studies chironomid analysis is sometimes applied to shallow-water sediment cores because remains of other proxies of interest (e.g. mites, plant macrofossils, lacustrine carbonates) may be absent or rare in the deepest part of the lake basin. It remains, however, an open question if chironomid–temperature inference models calibrated from the deepest parts of lake basins can be used to reliably reconstruct past summer temperatures from shallow-water chironomid assemblages.

We studied the distribution of chironomid subfossils in five small Norwegian lakes in order to assess the effects of within-lake variability of chironomid assemblages on quantitative chironomid–temperature reconstructions. In each lake we obtained six replicate samples from the deepest part of the lake basin, analysed the subfossil chironomid assemblages in the surface sediments and applied a quantitative chironomid–mean July air temperature inference model to the results. The variability of the inferred temperatures was then used to examine the error associated with sampling the chironomid assemblage of a lake with a single sediment core in respect to the overall model prediction error. This variability is of interest as its influence on the model could be reduced by a different sampling scheme, e.g. by obtaining several samples per lake for the model development. However, this would greatly increase the amount of work necessary to generate chironomid–temperature inference models and at present it is not clear if a significant improvement in prediction errors could be achieved. Furthermore, we analysed in each lake two transects of seven samples each, ranging from the deepest part of the lake basin towards the littoral, to assess if there exists any bias in inferred temperatures in relation to water depth.

## 2. Study sites and methods

We sampled surface sediments from five lakes in southern Norway (Table 1), Holebudalen (HOL), Histøl (HIS), Lisletjønn (LIS), Råtåsjøen (RAT), and Vestre Øykjærmyrtjern (OYK). The lakes ranged from the deciduous forest to the alpine vegetation zone, spanned a mean July air temperature gradient of 8.2–14.3°C, and were selected to be of similar size and depth as those used to develop a chironomid–mean July air temperature inference model in Norway ([Brooks and Birks, 2000b and Brooks and Birks, 2001]). After establishing a lake bathymetry using an echo sounder, six cores in the deepest part of the lake basin and two transects of seven cores ranging from the lake centre towards the littoral were taken at each site using a modified Kajak corer ([Renberg, 1991]; Fig. 1). These cores will henceforth be referred to as M-1 to M-6 for the mid-lake cores and T1-1 to T1-7 and T2-1 to T2-7 for the two transects, respectively (numbered according to decreasing water depth). In lakes with two basins of similar depth one transect was placed in each of the basins. In OYK the last sample of the second transect (T2-7) was taken close to the lake tributary instead of in line with the other transect samples (Fig. 1). From each core a sample from the top centimetre of sediment was prepared for chironomid analysis following [Walker, 2001] but excluding the acid treatment. Chironomid subfossils were identified at 400× magnification under a compound microscope. In most samples (88%) more than 80 head capsules were obtained. However, for some samples, especially from the shallowest parts of the lake basins, low chironomid abundances were a problem. The lowest count sum for a mid-lake sample was 64 and for a transect sample 41 (Table 2). Chironomid subfossils were generally identified to the taxonomic level used in the Norwegian chironomid–July air temperature inference model ([Brooks and Birks, 2000b and Brooks and Birks, 2001]; Brooks and Birks, unpublished data). Chironomid-inferred temperatures were modelled using the program CALIBRATE (S. Juggins and C.J.F. ter Braak, unpublished software) and ordination performed using CANOCO version 4.0 ([ter Braak and Smilauer, 1998]).

Table 1. Abbreviation, geographical location, elevation, maximum lake depth and mean July air temperature for the study lakes

Lake	Abbreviation	Latitude (°N)	Longitude (°E)	Elevation (m asl)	Maximum lake depth (m)	Mean July air temperature (°C)
Holebudalen	HOL	59°50'	6°59'	1144	7.5	8.2
Histøl	HIS	58°20'	7°47'	245	9.2	14.3
Lisletjønn	LIS	59°21'	7°18'	518	7.3	12.9
Råtåsjøen	RAT	62°16'	9°50'	1169	7.5	8.6
Vestre Øykjærmyrtjern	OYK	59°49'	6°00'	570	7.8	11.0

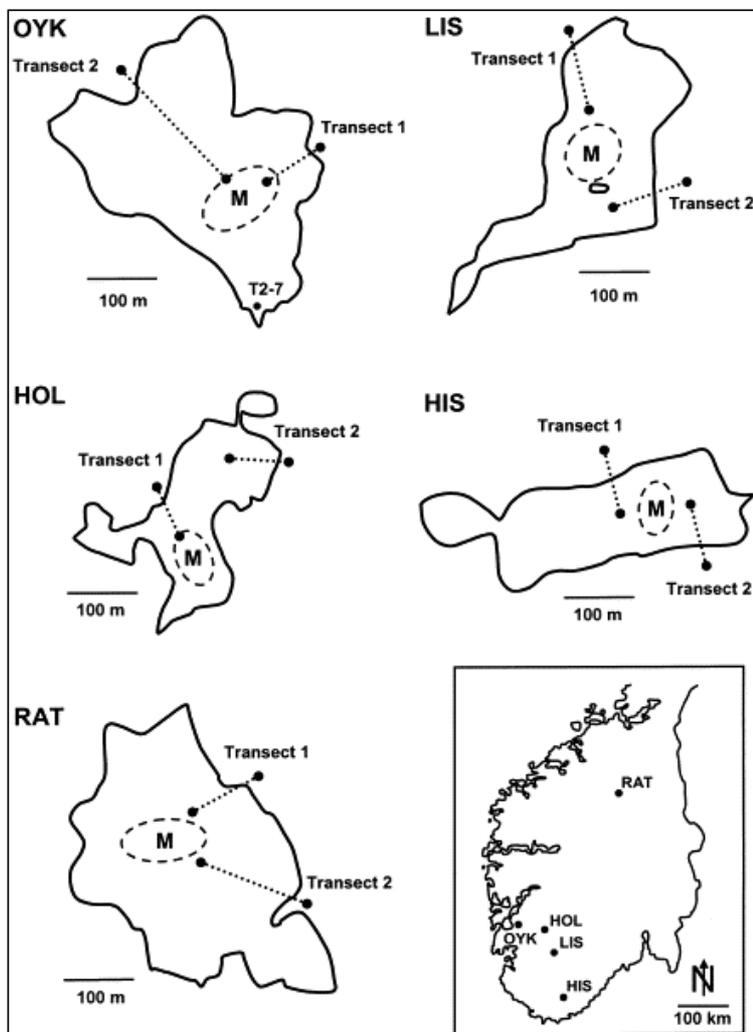


Fig. 1. Contour maps of the five study lakes. The dashed circles indicate the deepest parts of the lakes where the M-cores were obtained and the dotted lines the two transects towards the littoral. In OYK, a single near-shore sample was taken outside the transects (indicated with T2-7). The inset shows the location of the study sites in an overview map of southern Norway.

Table 2. Count size and standard deviations (S.D.) of the chironomid-inferred July air temperatures in the five study lakes

	HIS	HOL	LIS	RAT	OYK
Median (range) of counts of M-samples	93 (64–119)	182 (118–256)	132 (96–270)	153 (85–165)	123 (102–128)
Median (range) of counts of all samples	97 (49–218)	135 (66–256)	144 (82–334)	122 (52–165)	113 (41–132)
S.D. of inferred temperatures for M-samples	0.54°C	0.51°C	0.49°C	0.35°C	0.43°C
S.D. of inferred temperatures for all samples	0.56°C	0.52°C	0.59°C	0.38°C	0.56°C
S.D. of inferred temperatures for 100 simulated samples of count size 100	0.41°C	0.39°C	0.36°C	0.33°C	0.39°C

See text for details.

### 3. Results and discussion

#### 3.1. Chironomid–mean July air temperature inference model

A chironomid–mean July air temperature inference model was calculated based on subfossil assemblages in surface-sediment samples from the deepest parts of 153 small Norwegian lakes ([Brooks and Birks, 2000b and Brooks and Birks, 2001]; Brooks and Birks, unpublished data). These lakes range from southernmost Norway to Svalbard and cover a mean July air temperature gradient of 3.5–16.0°C (see [Brooks and Birks, 2000b and Brooks and Birks, 2001], for more details on air temperature interpolation procedures and criteria for lake selection). After taxonomic harmonisation with the samples obtained from our five study lakes, the model data included a total of 116 subfossil chironomid taxa. Detrended canonical correspondence analysis with July air temperature as the only explaining variable suggested a temperature gradient length of three standard deviation units in the square-root-transformed chironomid percentages and, therefore, non-linear unimodal-based regression procedures were used to develop the inference model. Following [Birks, 1998], weighted averaging partial least squares regression with three components was chosen as the ‘minimum adequate model’. The final model based on square-root-transformed chironomid percentages predicts mean July air temperature with a RMSEP of 1.04°C and a  $r^2$  of 0.91 (Fig. 2).

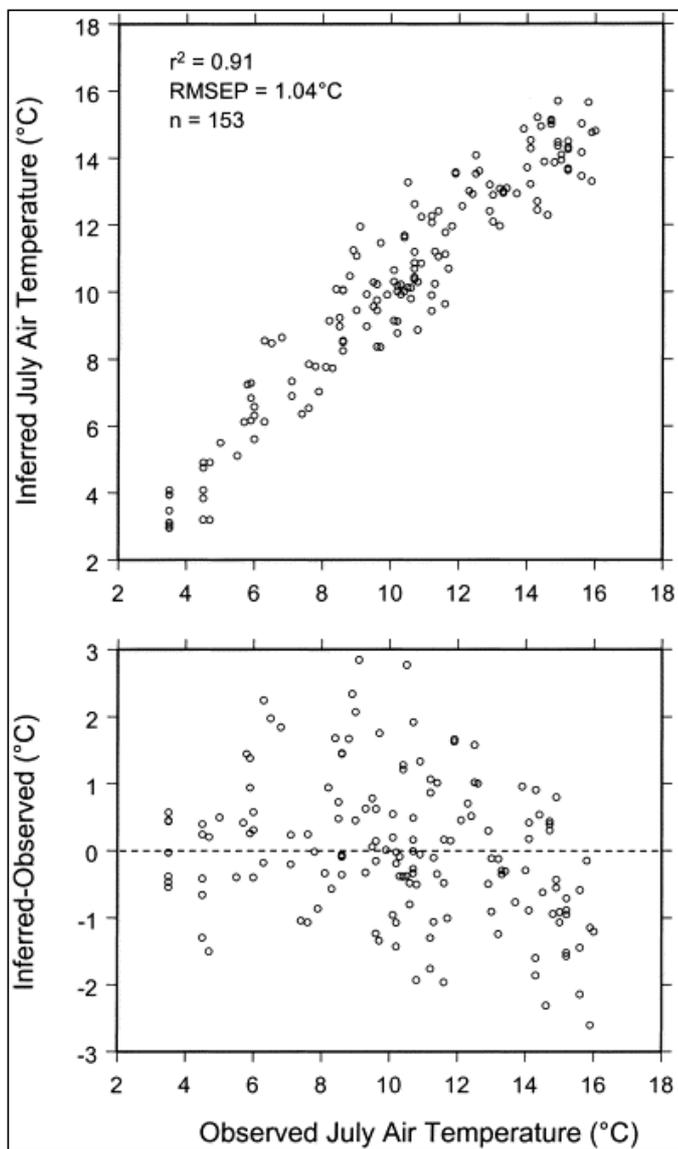


Fig. 2. Comparison of observed mean July air temperature with the chironomid-inferred temperatures and the prediction residuals for the Norwegian chironomid–July air temperature inference model. The model is based on three-component weighted averaging partial least squares regression. RMSEP indicates the leave-one-out cross-validated root mean square error of prediction,  $r^2$  the coefficient of determination and  $n$  the number of samples in the model.

### 3.2. Chironomid-inferred temperatures within the lake basins

Chironomid-inferred mean July air temperatures were fairly variable across a lake basin and generally spanned a temperature range of about 2°C (Fig. 3). The standard deviation of the inferred temperatures for the six mid-lake samples ranged from 0.35 to 0.54°C (Table 2). For all 20 samples analysed within a lake basin the standard deviation was slightly higher in all five study lakes (0.38–0.59°C). The standard deviations of the six mid-lake samples can be interpreted as the variability of chironomid-inferred temperatures due to replicate sampling. Thus, they provide an estimate of the variation in our model due to the error associated with assessing the chironomid assemblages in the deepest parts of the lakes with a single sediment sample. Six is a low sample number to estimate this variability from and ideally a higher number of samples per lake should have been obtained to calculate this variation. However, four of the five standard deviations for the mid-lake cores lie within the narrow range of 0.43–0.54°C and the standard deviations for the mid-lake samples and for all 20 samples from within a lake are very similar (Table 2). This suggests that we have obtained realistic estimates of the variation of chironomid-inferred temperatures due to replicate sampling in the deepest parts of the lake basins and that a standard deviation of 0.43–0.54°C may be a typical value for this variability in small Norwegian lakes.

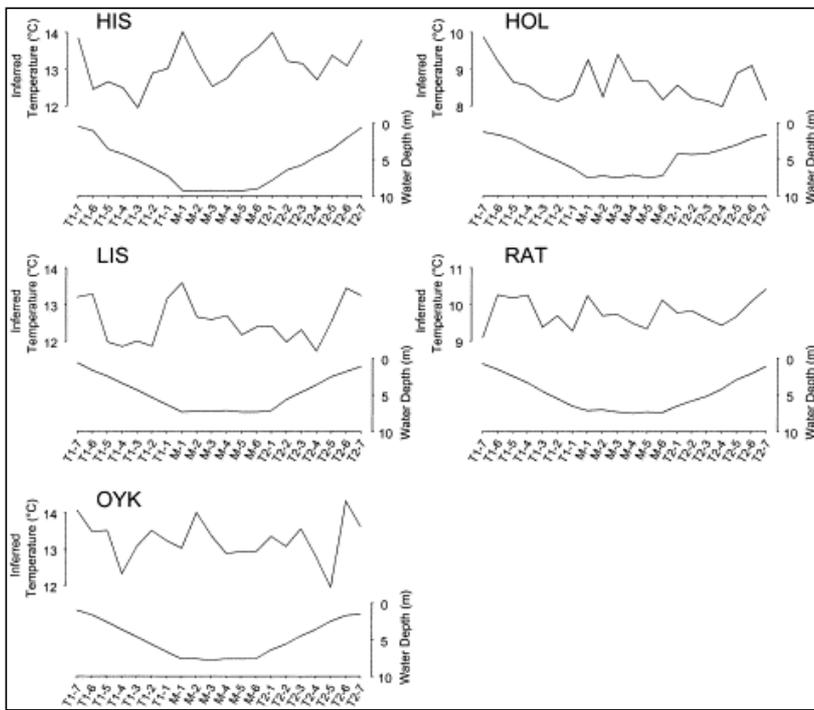


Fig. 3. Chironomid-inferred mean July air temperatures and water depth plotted for the 20 samples from each study lake. M refers to the samples taken from the deepest part of the lakes whereas T1 and T2 refer to the two sampling transects.

For our purposes it would be of interest to quantify how much of the overall model prediction error such variability could account for. We can assume that the variability associated with assessing the subfossil assemblage composition of the model lakes with a single sediment sample is superimposed on the model prediction error we would obtain if we could calibrate it on the actual chironomid assemblages. Following [Birks et al., 1990], it is then possible to decompose the overall prediction error of the model into the two separate components using:

$$\text{MSEP} = \text{MSE}_{\text{replicate}} + \text{MSE}_{\text{residual}} \quad (1)$$

where MSEP is the leave-one-out cross-validated mean square prediction error of the chironomid–temperature model,  $\text{MSE}_{\text{replicate}}$  is the mean square error associated with using a single sample to assess the chironomid assemblage composition of the model lakes, and  $\text{MSE}_{\text{residual}}$  is the residual mean square error component of our model, an estimate of the model prediction error we would obtain if it were possible to sample perfectly the chironomid assemblages. A number of statistical assumptions are implicit to this approach, amongst them the assumption that (1) the three error components have a Gaussian distribution, (2)  $\text{MSE}_{\text{replicate}}$  and  $\text{MSE}_{\text{residual}}$  are independent and (3) the three error components are of the same size over the whole temperature gradient. Using the RMSEP of the chironomid–temperature model to estimate MSEP and the squared observed standard deviations of the six mid-lake samples as an estimate of  $\text{MSE}_{\text{replicate}}$ , the residual root mean square error ( $\text{RMSE}_{\text{residual}}$ ) can thus be estimated to 0.89–0.98°C or 85–94% of the RMSEP of the Norwegian chironomid–July air temperature inference model. The prediction residuals of the chironomid–temperature inference model tend to be smaller at the cold end of the model temperature gradient (Fig. 2), suggesting that our assumption (3) may not be valid for the overall model prediction error. Therefore, we recalculated the model, the RMSEP and  $\text{RMSE}_{\text{residual}}$  excluding sites with observed temperatures below 6°C. However, the results were very similar, suggesting a  $\text{RMSE}_{\text{residual}}$  of 87–95% of the model prediction error.

The pattern of inferred temperatures within a given lake basin differed from site to site and between transects and it is difficult to detect an overall pattern for all five study lakes (Fig. 3). In some transects there seems to be a tendency towards warmer inferred temperatures in near-shore samples (e.g. HIS, LIS, HOL T-1, OYK T-1). However, other shallow-water samples infer exceptionally cool temperatures (e.g. HOL T2-7, RAT T1-7). In some lakes, samples of intermediate water depth gave relatively cool temperatures (HOL, LIS, OYK), although in others the between-sample variation in inferred temperatures is too high to detect any trend (e.g. OYK). The relationship between chironomid-inferred temperatures and water depth becomes more obvious if the inferred temperatures are plotted relative to the average temperature of the six mid-lake samples and versus standardised water depth (Fig. 4). At intermediate depths of about 0.4–0.6 times the maximum lake depth the inferred air temperatures are, on average, about 0.4–0.5°C cooler than the mean inferred temperature for the mid-lake samples. At depths shallower than 0.2 times the maximum lake depth the model tends to infer above-average temperatures. However, two samples inferring exceptionally cool temperatures are inconsistent with this trend.

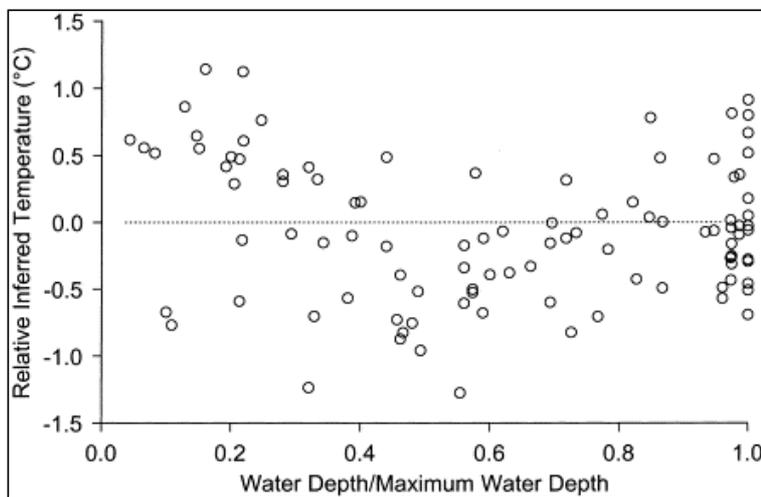


Fig. 4. Chironomid-inferred July air temperatures and water depth for all 100 samples from the five study lakes. The inferred temperatures are plotted relative to the mean of the six mid-lake samples (relative inferred temperature) and water depth values as depth divided by maximum water depth.

### 3.3. Implications for quantitative chironomid-inferred temperature reconstruction

The chironomid-inferred mean July air temperatures were rather variable in the deepest parts of the five studied lake basins and it seems reasonable to assume a similar variability exists in all the lakes used to develop the Norwegian chironomid–temperature inference model. This variation (i.e.  $RMSE_{\text{replicate}}$ ) has a direct influence on the model prediction error, as in each leave-one-out cross-validation sequence used to calculate the RMSEP, the observed temperature of a lake will be compared to the temperature predicted from the chironomid assemblage of a single sediment sample only.  $RMSE_{\text{replicate}}$  is not exclusively determined by the within-lake variability of fossil assemblages and includes other error sources such as the sample count size and coring artefacts. Low counts can be a problem in chironomid analysis (e.g. [Larocque, 2001 and Quinlan and Smol, 2001]) and increase the variability of chironomid-inferred temperatures ([Heiri and Lotter, 2001]). At a given count sum, this variability can differ from lake to lake and is ultimately dependent on the assemblage composition of a sample as well as on the distribution of the chironomid taxa in the applied chironomid–temperature inference model. In our five study lakes the number of chironomid head capsules isolated per sample was very variable, ranging from 41 to 334 (Table 2), although 77% of the mid-lake samples had counts of 100 head capsules or more. In order to obtain an idea of the variation in inferred temperatures that could be caused in the lakes by count problems we produced for each lake 100 simulated samples of 100 counts each (see [Heiri and Lotter, 2001], for more information on the basic approach). The simulations were based on random sampling with replacement from a population with a distribution following the average percentage composition of the six mid-lake samples. They were used to provide an estimate of the variability in chironomid-inferred temperatures we would expect in a sample of 100 counts, if the fossils are distributed randomly in the sediments. The standard deviations of the chironomid-inferred temperatures for these simulated samples ranged from 0.33 to 0.41°C and were only slightly smaller than our estimates for  $RMSE_{\text{replicate}}$  (Table 2).

Our study was not designed to separate between the effects of within-lake variability of fossil assemblages and count size on chironomid-inferred temperatures, and these two sets of standard deviations should therefore not be directly compared. However, the large variation in inferred temperatures of the simulated samples suggests that in our study lakes  $RMSE_{\text{replicate}}$  is strongly influenced by the count size of the samples.

A clumped distribution of subfossils in the deepest parts of the lakes would be a further source of variability contributing to  $RMSE_{\text{replicate}}$ . In principle this variability could be reduced by sampling several cores per lake for the model development. The average subfossil assemblage composition could then be calculated for each lake prior to calibrating the model. However, subfossil chironomid analysis is very time-consuming and includes hand-sorting of chironomid remains from sieved sediments under a stereomicroscope ([Walker, 2001]). Depending on the sediment type and concentrations of the subfossils, the analysis of a single sample can take from several hours up to two days. Thus, the effort needed to analyse several samples per lake would vastly increase the time used for developing chironomid–temperature inference models. However, since  $RMSE_{\text{replicate}}$  in Norwegian lakes seems to be strongly influenced by count size, a considerable amount of the model improvement possible by a ‘multi-core’ approach might actually be attained by taking multiple cores only in lakes with low count problems to increase the minimum count sum in the model.

Extant chironomid assemblages are known to show strong changes in species composition with water depth, especially in deep, stratified lakes. This bathymetric zonation has been attributed to some extent to the water temperature gradient during summer stratification. For example, chironomid species such as *Micropsectra radialis*, *Sergentia coracina* or *Stictochironomus rosenschoeldi*, which are solely found in the cool profundal of deep stratified lakes in temperate regions, occur in littoral regions in colder subarctic

and mountain lakes ([Brundin, 1949](#) and [Hofmann, 1971](#)). Within each of our five study lakes all the 20 chironomid samples were generally dominated by the same set of taxa, although there was a shift in subfossil chironomid assemblages towards the littoral. Therefore, it is not surprising that the chironomid-inferred temperatures show a trend with decreasing water depth. However, the finding that cooler July air temperatures are inferred for intermediate water depth is unexpected. At the bottom of stratified lakes summer water temperatures increase with decreasing water depth. Hence, eurythermic chironomid species tend to dominate in littoral regions of temperate lakes and one would have expected increasingly warmer temperatures to be inferred along the transects towards the littoral. We have examined the chironomid data to try to determine which taxa are responsible for these cooler inferred temperatures at intermediate water depth. However, it was not possible to link major changes in the chironomid assemblages with this trend. The reasons for the decrease in inferred summer temperatures at intermediate water depth therefore remain unknown. It can be stated, however, that even though this trend is only weakly expressed in some of our study lakes ([Fig. 3](#)) the overall picture in all five lakes is surprisingly consistent ([Fig. 4](#)) and suggests that on average cooler temperatures of about 0.4–0.5°C are inferred from samples taken at about half the maximum lake depth. The rise in chironomid-inferred temperatures in samples close to the lake shore is simpler to explain. Summer temperatures of the surface water layer in stratified lakes can be significantly warmer than in the deepest parts, even in alpine and subarctic lakes (e.g. [Barbieri et al., 1999](#) and [Sorvari et al., 2000](#)). Consequently, chironomid larvae living in the shallowest parts of the lake basin need to be able to tolerate these warmer temperatures and cold stenothermic species are often absent. Although the inferred temperatures in the shallowest parts of the lakes tended to be more than 0.5°C warmer than the average inferred temperature of the mid-lake samples, exceptionally cool temperatures were inferred in two shallow-water samples ([Fig. 4](#)), indicating that there are exceptions to this trend.

## 4. Conclusions

Our results indicate that the variation of chironomid-inferred July air temperatures of replicate samples from the deepest parts of lakes can be quite significant when compared to the overall prediction error of the Norwegian chironomid–temperature inference model. However, this variability, which can be interpreted as an estimate of the model error due to uncertainties of assessing the chironomid assemblages from a single sample only, can at best account for 15% of the model RMSEP. Clearly, investing a large amount of time and effort in reducing this component of the prediction error does not seem to be a promising approach to increasing the performance of chironomid–temperature inference models. Based on [Eq. 1](#) and our estimates of the different error components, the overall prediction error of the model would still be 92% its original size, even if  $RMSE_{\text{replicate}}$  could be reduced by as much as 75%. Given the large amount of the prediction error that remains unexplained (i.e.  $RMSE_{\text{residual}}$ ) it seems more encouraging to focus future efforts on minimising other possible sources of variability in the chironomid–temperature inference models. Potential error sources which come to mind are the quality of the temperature data used to calibrate the models, the ability of the applied statistical models to realistically capture the relationship between the chironomid assemblages and temperature, and environmental factors other than temperature affecting the limnology of the study lakes and their chironomid assemblages (e.g. nutrient loading, bottom water oxygen concentrations, or pH).

In our transects we found clear indications of a bias in inferred temperatures associated with the water depth in which sediments were sampled. This error ranged from on average 0.4–0.5°C cooler temperatures inferred from intermediate water depths to generally more than 0.5°C warmer temperatures inferred from shallow-water samples. It is to be expected that this bias will be even more pronounced in deeper lakes with stronger bathymetric water temperature gradients than the ones we studied. Chironomid analysts should be aware of a possible systematic bias when applying chironomid temperature inference models calibrated on deep-water samples to sediment cores obtained from shallower parts of the lake basin. This may be of less concern if the focus of the temperature reconstruction is on assessing the overall pattern of past temperature development or on the magnitude of past temperature fluctuations. However, if the main interest is on the absolute temperature values, a water depth-related bias will add to the prediction uncertainty of the model.

## Acknowledgements

We would like to thank Jan Berge, Aina Dahlø, Lapager Duoje, and Jorunn Larsen for assistance with the field work, and Klaus P. Brodersen and Ian R. Walker for helpful comments on the manuscript. The stay of O.H. at the University of Bergen was financed by a Swiss National Science Foundation fellowship for prospective researchers (Fellowship 81BE-66224) and field work was supported by the Norwegian Research Council Strategic University Project ‘Norwegian Palaeoenvironments and Climate’ (NORPEC).

## References

- [Anderson, 1990](#). N.J. Anderson, Variability of diatom concentrations and accumulation rates in sediments of a small lake basin. *Limnol. Oceanogr.* **35** (1990), pp. 497–508.
- [Barbieri et al., 1999](#). A. Barbieri, M. Veronesi, M. Simona, S. Malusardi and V. Straškrabová, Limnological survey in eight high mountain lakes located in Lago Maggiore watershed (Switzerland). *J.*

Bigler et al., 2002. C. Bigler, I. Larocque, S.M. Peglar, H.J.B. Birks and R.I. Hall, Quantitative multiproxy assessment of long-term patterns of Holocene environmental change from a small lake near Abisko, northern Sweden. *Holocene* **12** (2002), pp. 481–496.

Birks et al., 2000. H.H. Birks, R.W. Battarbee and H.J.B. Birks, The development of the aquatic ecosystem at Kråkenes Lake, western Norway, during the late-glacial and early-Holocene – a synthesis. *J. Paleolimnol.* **23** (2000), pp. 91–114.

Birks, 1998. H.J.B. Birks, Numerical tools in palaeolimnology – progress, potentialities, and problems. *J. Paleolimnol.* **20** (1998), pp. 307–332.

Birks et al., 1990. H.J.B. Birks, J.M. Line, S. Juggins, A.C. Steveson and C.J.F. ter Braak, Diatoms and pH reconstruction. *Phil. Trans. R. Soc. London B* **327** (1990), pp. 263–278.

Brooks and Birks, 2000a. S.J. Brooks and H.J.B. Birks, Chironomid-inferred Late-glacial air temperatures at Whitrig Bog, southeast Scotland. *J. Quat. Sci.* **15** (2000), pp. 759–764.

Brooks and Birks, 2000b. S.J. Brooks and H.J.B. Birks, Chironomid-inferred late-glacial and early-Holocene mean July air temperatures for Kråkenes Lake, western Norway. *J. Paleolimnol.* **23** (2000), pp. 77–89.

Brooks and Birks, 2001. S.J. Brooks and H.J.B. Birks, Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. *Quat. Sci. Rev.* **20** (2001), pp. 1723–1741.

Brundin, 1949. L. Brundin, Chironomiden und andere Bodentiere der südschwedischen Urgebirgsseen. *Rep. Inst. Freshwater Res. Drottningholm* **30** (1949), pp. 1–914.

Cwynar and Levesque, 1995. L.C. Cwynar and A.J. Levesque, Chironomid evidence for late glacial climate reversals in Maine. *Quat. Res.* **43** (1995), pp. 405–413.

Davis et al., 1971. M.B. Davis, L.B. Brubaker and J.M. Beiswenger, Pollen grains in lake sediments: Pollen percentages in surface sediments from southern Michigan. *Quat. Res.* **1** (1971), pp. 450–467.

Heiri and Lotter, 2001. O. Heiri and A.F. Lotter, Effect of low count sums on quantitative environmental reconstructions: an example using subfossil chironomids. *J. Paleolimnol.* **26** (2001), pp. 343–350.

Hofmann, 1971. W. Hofmann, Zur Taxonomie und Palökologie subfossiler Chironomiden (Dipt.) in Seesedimenten. *Arch. Hydrobiol. Beih.* **6** (1971), pp. 1–50.

Johnson, 1995. Johnson, R.K., 1995. The indicator concept in freshwater biomonitoring. In: Cranston, P. S. (Ed.), Chironomids – From Genes to Ecosystems. CSIRO, Melbourne, Vic., pp. 11–27.

Larocque, 2001. I. Larocque, How many chironomid head capsules are enough? A statistical approach to determine sample size for palaeoclimatic reconstructions. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **172** (2001), pp. 133–142.

Larocque et al., 2001. I. Larocque, R.I. Hall and E. Grahn, Chironomids as indicators of climate change: a 100-lake training set from a subarctic region of northern Sweden (Lapland). *J. Paleolimnol.* **26** (2001), pp. 307–322.

Levesque et al., 1993. A.J. Levesque, F.E. Mayle, I.R. Walker and L.C. Cwynar, A previously unrecognized late-glacial cold event in eastern North America. *Nature* **361** (1993), pp. 623–626.

Lotter et al., 1997. A.F. Lotter, H.J.B. Birks, W. Hofmann and A. Marchetto, 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** (1997), pp. 395–420.

Lotter et al., 1999. A.F. Lotter, I.R. Walker, S.J. Brooks and W. Hofmann, An intercontinental comparison of chironomid paleotemperature inference models: Europe vs. North America. *Quat. Sci. Rev.* **18** (1999), pp. 717–735.

Lotter et al., 2000. A.F. Lotter, H.J.B. Birks, U. Eicher, W. Hofmann, J. Schwander and L. Wick, Younger Dryas and Allerød summer temperatures at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **159** (2000), pp. 349–361.

Meriläinen, 1969. J. Meriläinen, Distribution of diatom frustules in recent sediments of some meromictic lakes. *Mitt. Int. Verein. Limnol.* **17** (1969), pp. 186–192.

- Olander et al., 1999. H. Olander, H.J.B. Birks, A. Korhola and T. Blom, An expanded calibration model for inferring lakewater and air temperatures from fossil chironomid assemblages in northern Fennoscandia. *Holocene* **9** (1999), pp. 279–294.
- Pellatt et al., 2000. M.G. Pellatt, M.J. Smith, R.W. Mathewes, I.R. Walker and S.L. Palmer, Holocene treeline and climate change in the subalpine zone near Stoyoma Mountain, Cascade Mountains, southwestern British Columbia, Canada. *Arct. Antarct. Alp. Res.* **32** (2000), pp. 73–83.
- Quinlan and Smol, 2001. R. Quinlan and J.P. Smol, Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. *J. Paleolimnol.* **26** (2001), pp. 327–342.
- Renberg, 1991. I. Renberg, The HON-Kajak sediment corer. *J. Paleolimnol.* **6** (1991), pp. 167–170.
- Rosén et al., 2001. P. Rosén, U. Segerström, L. Eriksson, I. Renberg and H.J.B. Birks, Holocene climatic change reconstructed from diatoms, chironomids, pollen and near-infrared spectroscopy at an alpine lake (Sjuodjijaure) in northern Sweden. *Holocene* **11** (2001), pp. 551–562.
- Sæther, 1979. O.A. Sæther, Chironomid communities as water quality indicators. *Holarct. Ecol.* **2** (1979), pp. 65–74.
- Sorvari et al., 2000. S. Sorvari, M. Rautio and A. Korhola, Seasonal dynamics of the subarctic Lake Saanajärvi in Finnish Lapland. *Verh. Int. Verein. Limnol.* **27** (2000), pp. 507–512.
- ter Braak and Smilauer, 1998. ter Braak, C.J.F., Smilauer, P., 1998. CANOCO Reference Manual and User's Guide to Canoco for Windows. Centre for Biometry Wageningen, Wageningen, 351 pp.
- Thienemann, 1913. A. Thienemann, Der Zusammenhang zwischen dem Sauerstoffgehalt des Tiefenwassers und der Zusammensetzung der Tierfauna unserer Seen. *Int. Rev. Hydrobiol.* **12** (1913), pp. 1–65.
- Vasko et al., 2000. K. Vasko, H.T.T. Toivonen and A. Korhola, A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. *J. Paleolimnol.* **24** (2000), pp. 243–250.
- Walker, 1987. I.R. Walker, Chironomidae (Diptera) in paleolimnology. *Quat. Sci. Rev.* **6** (1987), pp. 29–40.
- Walker, 2001. Walker, I.R., 2001. Midges: Chironomidae and related Diptera. In: Smol, J.P., Birks, H.J. B., Last, W.M. (Eds.), Tracking Environmental Change Using Lake Sediments. Zoological Indicators. Kluwer Academic, Dordrecht, pp. 43–66.
- Walker et al., 1991. I.R. Walker, R.J. Mott and J.P. Smol, Allerød-Younger Dryas lake temperatures from midge fossils in Atlantic Canada. *Science* **253** (1991), pp. 1010–1012.
- Walker et al., 1997. I.R. Walker, A.J. Levesque, L.C. Cwynar and A.F. Lotter, An expanded surface-water paleotemperature inference model for use with fossil midges from eastern Canada. *J. Paleolimnol.* **18** (1997), pp. 165–178.
- Wiederholm, 1981. T. Wiederholm, Associations of lake-living Chironomidae. A cluster analysis of Brundin's and recent data from Swedish lakes. *Schweiz. Z. Hydrol.* **43** (1981), pp. 140–150.