BOREAS An international journal of Quaternary research

# Chironomidae-based inference model for mean July air temperature reconstructions in the eastern Baltic area

VARVARA BAKUMENKO, ANNELI POSKA, MATEUSZ PŁÓCIENNIK, NERINGA GASTEVICIENE, BARTOSZ KOTRYS, TOMI P. LUOTO, SIMON BELLE AND SIIM VESKI

BOREAS

Bakumenko, V., Poska, A., Płóciennik, M., Gasteviciene, N., Kotrys, B., Luoto, T. P., Belle, S. & Veski, S. 2024 (July): Chironomidae-based inference model for mean July air temperature reconstructions in the eastern Baltic area. *Boreas*, Vol. 53, pp. 401–414. https://doi.org/10.1111/bor.12655. ISSN 0300-9483.

Here we present a new eastern Baltic Chironomidae training set (TS) containing 35 sites that was collected and merged with neighbouring published Finnish (82 lakes) and northern part of the Polish (nine lakes) TSs. Chironomidae, non-biting midges, are known to be strongly responsive to the July air temperature and are widely used to infer palaeotemperature. Several modern analogue-based TSs necessary for calibrating the relationships between mean July air temperature (MJAT) and chironomids are available for Europe. However, none of these is representative of the transitional climate typical for eastern Baltic (Estonia, Latvia, Lithuania). The Finno–Baltic–Polish TS contains 121 sites and covers a geographically continuous 70–50°N latitudinal and 7 °C (12.1–19.2 °C) MJAT gradient. Canonical correspondence analysis revealed that, among the tested environmental variables (pH, water depth, dissolved oxygen, MJAT), the MJAT explains the highest amount of variation, both for the eastern Baltic separately and the Finno–Baltic–Polish TSs. The weighted averaging–partial least squares-based cross-validation test reveals that the Finno–Baltic–Polish TS has a low root mean square error of prediction (0.7 °C) confirming the high reliability of the TS. The temperature optima of the taxa included in the new Finno–Baltic–Polish TS and widely used Swiss–Norwegian TS were examined. The observed dissimilarities can be attributed to the differences in the temperature ranges represented by the TS, the taxonomic identification level, the general cosmopolitan taxa distribution patterns and the influence of TS-specific geographic position, climatic or environmental conditions. The new Finno–Baltic–Polish TS adds to the knowledge on the modern distribution of Chironomidae taxa and widens the geographical area of reliable Chironomid-based MJAT reconstructions into the eastern European lowland.

*Varvara Bakumenko (varvara.bakumenko@taltech.ee) and Siim Veski, Department of Geology, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia; Anneli Poska, Department of Geology, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia and Department of Physical Geography and Ecosystem Science, Lund University, SE-221 00 Lund, Sweden; Mateusz Płóciennik, Department of Invertebrate Zoology and Hydrology, University of Lodz, Prezydenta Gabriela Narutowicza 68, 90-136 Łódź, Poland; Neringa Gasteviciene, Laboratory of Quaternary Research, Nature Research Centre, Akademijos str. 2, LT-08412 Vilnius, Lithuania; Bartosz Kotrys, Polish Geological Institute, National Research Institute, Pomeranian Branch in Szczecin, ul. Wieniawskiego 20, 71-130 Szczecin, Poland; Tomi P. Luoto, Faculty of Biological and Environmental Sciences, Ecosystems and Environment Research Programme, University of Helsinki, Yliopistonkatu 4, 00100 Helsinki, Finland; Simon Belle, Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden;* *received 17th October 2023, accepted 3rd March 2024.*

Accurate climate reconstructions are important in investigating ecosystem response to climate changes in eastern Europe. Presently available climate reconstructions are based on a variety proxy (Kaufman *et al.* 2020), where pollen-based reconstructions make the biggest contribution in eastern Europe (Feurdean *et al.* 2014). However, considerable migration delays of the terrestrial vegetation during the Lateglacial and Early Holocene and substantial anthropogenic land cover change during the Late Holocene can influence the quality of the pollen-based reconstructions, especially in northern and eastern Europe (Väliranta *et al.* 2015; Rao *et al.* 2022). Thus, there is a need for an independent proxy to increase the reliability of the climate reconstructions in the eastern Baltic (Estonia, Latvia, Lithuania).

The Chironomidae, non-biting midges from the Diptera family, are widely distributed in freshwaters, and chitinous remains of larvae that are morphologically well preserved in Quaternary lake sediments allowing their identification to the morphotype level (Brooks *et al.* 2007) are often used as an independent and reliable climate proxy. Chironomid community composition in Holarctic lakes has been reported to be strongly associated with the mean July air and water temperatures (Walker & Mathewes 1987; Barley *et al.* 2006; Heiri *et al.* 2011; Kotrys *et al.* 2020; Nazarova *et al.* 2023). Subfossils of Chironomidae larvae deposited in lake sediments are widely used to reconstruct July air temperature during the Quaternary (Eggermont & Heiri 2012; Gouw-Bouman *et al.* 2019; Plikk *et al.* 2019). However, they can be influenced by other environmental factors such as trophic state (Brooks *et al.* 2001; Luoto 2011) and deepwater oxygen concentrations (Brooks *et al.* 2001; Quinlan & Smol 2001; Verbruggen *et al.* 2011; Ursenbacher *et al.* 2020), but also pH, heavy metal concentration and lake water depth (Heiri 2004; Rees *et al.* 2008; Nazarova *et al.* 2011; Ruse *et al.* 2018; Pegler *et al.* 2020; Ni *et al.* 2023).

Relationships between chironomid assemblages and temperature can be explored using calibration training sets (TSs) that describe the distribution and abundance of

DOI 10.1111/bor.12655

chironomid taxa across a temperature gradient (Eggermont & Heiri 2012). TSs are produced by collecting sediment samples from a large number of lakes across geographic or elevation gradients and by analysing the chironomid remains contained in these sediments (Lotter *et al.* 1997), thus providing an estimate of the composition of the modern chironomid fauna of these lakes (Heiri *et al.* 2011). Numerical inference models (transfer/calibration functions) (Birks 1998; Birks *et al.* 2010) can then be applied to determine chironomid–temperature correlation and to infer July temperature values. Inference models using chironomids as a proxy are based on a range of different numerical approaches, i.e. weighted averaging-partial least square regression, maximum likelihood regression and modern analogue technique (Heiri *et al.* 2011; Self *et al.* 2011; Medeiros *et al.* 2022), and allow the estimation of inferred July temperatures with a prediction error ranging from less than 2 °C (Kaufman *et al.* 2020). Existing chironomid–July temperature calibration TSs and inference models in Europe have been developed in Switzerland (Lotter *et al.* 1997), Norway (Brooks & Birks 2000), northern Sweden (Larocque *et al.* 2001), Finland (Luoto 2009), Poland (Kotrys *et al.* 2020), Slovakia (Chamutiova *et al.* 2020) and N and NE Russia (Nazarova *et al.* 2023). While these TSs represent a wide range of biogeographical conditions, a considerable part of the sites are from highlands and none from the eastern Baltic.

The eastern Baltic area, which is a part of the European Plain, is located in the transitional zone between maritime and continental climates. The border areas separating continental-scale climatic zones exhibit significant variations in climate factors (temperature, humidity, wind speed, etc.) and climatic instability (Fu 1992). Thus, the eastern Baltic area reveals heightened sensitivity to changes in northern Europe's air circulation patterns (Giesecke *et al.* 2008; Seppä *et al.* 2009), which could result in the emergence of unique Chironomidae assemblages. The eastern Baltic Chironomidae fauna is understudied, and no local TS has been developed so far. The Swiss–Norwegian and Finnish TSs have been used for the reconstruction of summer temperatures during glacial and Early Holocene climate events in the eastern Baltic (Heiri *et al.* 2014; Šeirienė *et al.* 2021). On one hand, the Finnish TS lacks the warmer analogues that could correspond to the modern mean July air temperature (MJAT) range in the eastern Baltic region. In addition, the diverse bedrock composition, with acidic bedrock in Finland and carbonate in the Eastern Baltic, can be expected to result in varying water chemistry, which affects chironomid assemblages (Eggermont & Heiri 2012). On the other hand, the Swiss–Norwegian–Polish (Kotrys *et al.* 2020) and north European TSs (Larocque *et al.* 2001; Luoto 2009) have a spatial gap of calibration sites and lack modern analogues of Chironomidae communities reflecting the combination of the mild maritime climate and carbonate bedrock that is typical for the eastern Baltic. This complicates the use of regression-based methods for temperature reconstruction since they rely on the even spacing of samples along environmental gradients.

Another issue is that the multiple Chironomidae species with different ecological preferences belong to the same morphotype. Furthermore, the same morphotypes can represent different species in different regions, introducing additional bias into transfer functions (Velle *et al.* 2010; Juggins 2013; Medeiros *et al.* 2015). Calibration functions perform best in the geographical region in which the calibration TSs were developed, and the absence of the local TS could influence the reliability of the reconstruction (Heiri *et al.* 2011; Plikk *et al.* 2019; Płóciennik *et al.* 2022).

The objectives of this work are: (i) to compile a new chironomid TS representative of environmental conditions of the eastern Baltic; and (ii) to build a comprehensive TS representative of low-elevation areas of northern Europe and cover the temperature range expected for post-glacial conditions by merging eastern Baltic TS with samples from Finnish (Luoto 2009) and Polish TSs (Kotrys *et al.* 2020).

## Material and methods

### *Chironomid training sets*

The eastern Baltic TS consists of Chironomidae assemblage data and the corresponding MJAT for 35 lakes from Estonia, Latvia and Lithuania (Fig. 1A). Chosen lakes in the eastern Baltic TS have characteristic water chemistry for the region (Tables 1, S1, Fig. S1) and low anthropogenic impact on their surroundings. The region is located at the transition from the temperate to the boreal domains and between maritime and continental climates. The weather changes here frequently depending on the domination of the western zonal flow or continental high-pressure conditions (Johannessen 1970). The highest peak in the area is 318 m a.s.l. The bedrock in the eastern Baltic consists mainly of limestones and sandstones, thus the pH of the lakes is mostly alkaline. The surroundings of lakes chosen for sampling consist of coniferous, broad-leaved and mixed forests and low-intensity agricultural activity. Lakes were sampled in a transect along the MJAT gradient (17.2–18.4 °C).

The Finnish 110-site TS (Luoto 2009) and nine sites from the Polish TS (Kotrys *et al.* 2020) were merged (see below, section 'Chironomidae identification and TSs harmonisation') with the eastern Baltic TS. The resulting Finno–Baltic–Polish TS (Fig. 1B, Table 1) covers an MJAT range of 12.1–19.2 °C and a latitudinal range of 69.44–53.9°N.

Detailed information about Finnish and Polish sampling sites and their Chironomidae assemblages can be found in the original publications (Luoto 2009; Kotrys *et al.* 2020).

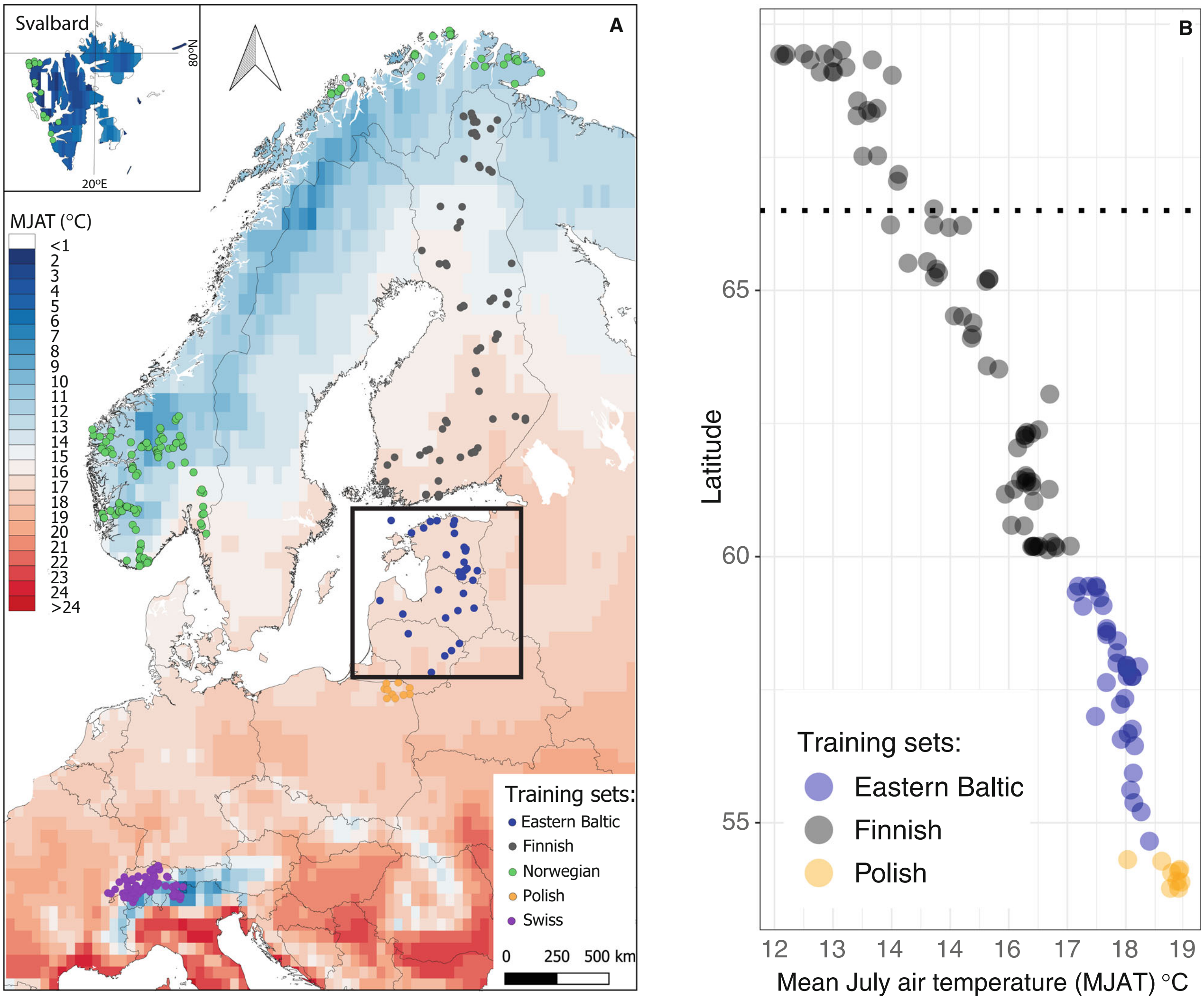

*Fig. 1.* Location of the sites (A) included in the Finno–Baltic–Polish and Swiss–Norwegian training sets (TSs) and the distribution of the Finno–Baltic–Polish TS sites, where the grey frame indicates new eastern Baltic sites (B) along the mean July air temperature (MJAT) and latitudinal gradient. The dotted black line represents the polar circle 66.5°N boundary.

The Swiss–Norwegian TS (Heiri *et al.* 2011) and supplementary environmental data were downloaded from the National Oceanic and Atmospheric Administration website. It covers a broad MJAT and latitudinal range (Table 1) in central and northern Europe. The Swiss–Norwegian TS was used without excluding any outliers.

### *Sampling and laboratory procedures*

The eastern Baltic TS is based on surface sediment samples of 35 medium-sized natural lakes. The samples were collected from the deepest part of the analysed lakes using a gravity corer. During the sampling, the topmost 2 cm of sediment was taken and later analysed in the laboratory. Sampling in Estonia was carried out in 2021 (February), in Latvia in 2019–2021 (February) and in Lithuania in 2022 (February) (Table S1). The water chemistry data were collected during the summers of 2021 and 2022. The sediment samples were water-sieved with a 100 μm mesh size sieve to remove fine sediment. The deflocculation in 10% KOH, which had been applied in samples processing of Finnish (Luoto 2009) and Polish (Kotrys *et al.* 2020) TSs, was skipped owing to the generally loose and watery sediment structure. Each sample was thereafter transferred into a Petri dish, and Chironomidae head capsules were collected with fine forceps under a stereomicroscope at 25× magnification. The obtained head capsules were air-dried and mounted in Aquatex® mounting medium. Taxonomic identification was conducted under the microscope at 100–400× magnification. The average number of head capsules collected per sample was 65, with counts ranging from 42 to 125.

*Table 1.* Environmental data for different subsets of the Finno–Baltic–Polish and Swiss–Norwegian training sets (TSs).

| Training set | Finnish | Eastern Baltic | Polish | Swiss–Norwegian |
|---|---|---|---|---|
| Number of lakes | 82 | 35 | 9 | 265 |
| Latitude (°N) | 60.13–69.44 | 54.65–59.45 | 53.9–54.3 | 46.15–79.8 |
| Longitude (°E) | 22.0–30.13 | 24.2–27.34 | 22.06–23.48 | 5–31.04 |
| Sampling water depth range (m) | 0.5–9.0 | 1.4–21.8 | 3.3–15.0 | – |
| July air temperature (°C) | 12.1–17.5 | 17.2–18.4 | 18.4–19.2 | 3.5–18.4 |
| pH | 4.4–9.3 | 4.48–9.5 | 7.81–8.73 | – |
| Dissolved oxygen (g $L^{-1}$) | 0.5–11.8 | 6.4–16.6 | – | – |
| Water phosphorus content (g $mL^{-1}$) | – | 0.007–0.169 | – | – |

The laboratory procedures for the Finnish, Swiss–Norwegian, and Polish TSs samples are described in the original papers (Luoto 2009; Heiri *et al*. 2011; Kotrys *et al*. 2020).

### *Chironomidae identification and TSs harmonization*

Chironomid head capsules from the Eastern Baltic TS were identified according to Schmid (1993), Klink and Pillot (2003), Brooks *et al*. (2007), Larocque-Tobler (2014) and Andersen *et al*. (2013). The eastern Baltic TS Chironomidae collection is stored and available in the Department of Geology at Tallinn University of Technology, Tallinn, Estonia.

The taxonomic resolution of Eastern Baltic TS and all of the used published TSs (Finnish (Luoto 2009), Polish (Kotrys *et al*. 2020) and Swiss–Norwegian (Heiri *et al*. 2011)) is based on that described in Brooks *et al*. (2007). The taxonomic harmonization between TSs was done on the count data. The few apparent taxonomic differences were handled as follows:

- Head capsules that lacked identifying features and were only identified to the genus or subfamily taxonomic level (Tanytarsini, *Tanytarsus* spp., *Paratanytarsus* spp., Tanypodinae, Chironomini, Orthocladinae) were excluded from the Finno–Baltic–Polish TS owing to the possibility of introducing bias into the inference model by including groups containing taxa representing different spectra of ecological conditions.
- The *Cricotopus intersectus* type was merged with the *Cricotopus laricomalis* type into one type owing to the high chance of misidentification of these morphotypes.
- *Ablabesmyia* spp. was merged with *Ablabesmyia monilis* type, *Ablabesmyia longistyla* type and *Ablabesmyia phatta* type owing to the absence of identification features in the eastern Baltic and Polish TSs.
- The *Corynoneura scutellata* type was merged with *Corynoneura edwardsi* type and *Corynoneura arctica* type following the identifications used in the Finnish TS (Luoto 2009).

Following the above-described taxonomic harmonization, the relative abundances were recalculated.

### *Water chemistry, lake depth and climate data*

For the eastern Baltic TS, lake depth was measured during lake sampling using a mechanical tape; the pH and oxygen concentration were measured in the field with the YSI ProDSS probe at 30–40 cm above the sediment surface of the lake; and the water phosphorus was determined in the laboratory from the top water layer samples using a HACH LKC349 analysis kit and Hach Lange DR 2800 spectrophotometer by Anna Lanka (unpublished data). Some additional water chemistry information was obtained from lake monitoring centres (Latvian Environment, Geology and Meteorology Center and the Lithuanian Environmental Protection Agency under the Ministry of the Environment). The pH data for Eastern Baltic TS are available for 33 lakes and dissolved oxygen and water phosphorus data are available for 29 lakes from the eastern Baltic TS. The dissolved oxygen for eastern Baltic sites was measured in July and August, whereas for Finnish lakes it was measured between February and April; thus, joining these data series could be unrepresentative of environmental conditions. For the Finnish, Polish and Swiss–Norwegian TSs the water chemistry and lake water depth were derived from the original publications (Luoto 2009; Heiri *et al*. 2011; Kotrys *et al*. 2020). The MJAT for each lake in Eastern Baltic TS was estimated using 30-year gridded (0.1 × 0.1°) observational data (1991–2020), obtained from the E-OBS TS (Cornes *et al*. 2018) downloaded from the Copernicus Climate Data Store. The same approach interpolating the meteorological stations' observations and calculating 30 year means was used for the Finnish, Polish and Swiss–Norwegian TSs (Luoto 2009; Heiri *et al*. 2011; Kotrys *et al*. 2020).

### *Data analysis*

For the statistical analysis, chironomid assemblage data of Finno–Baltic–Polish TS were transformed into relative abundances (%). Only taxa with an abundance of at least 2% in one sample were considered during numerical analysis and square root transformation was applied to equalize variances among taxa.

Chironomidae assemblages were analysed using detrended correspondence analysis (DCA; Hill & Gauch 1980) to examine the general distribution of communities and the compositional gradient lengths along the first two DCA axes. The length of DCA Ax1 and Ax2 (3.1 and 2.4 standard deviation (SD) units, respectively) falls into the intermediate category, for which the use of both redundancy analysis (RDA) and canonical correspondence analysis (CCA) has been recommended (Birks 1998; Lepš & Šmilauer 2003). We have used RDA and CCA for Finno–Baltic–Polish TS to produce results comparable with earlier studies (Luoto 2009; Kotrys *et al*. 2020). The CCA was applied to evaluate the significance of the environmental variables that explain significant variation in the chironomid data. Owing to a different number of observations for each variable the CCAs were run with only one environmental variable at a time (Table 1). The statistical significance of each selected variable was tested by a Monte Carlo permutation test (9999 unrestricted permutations) (Ter Braak 1992; Ter Braak & Verdonschot 1995). To choose the most relevant explanatory variables, $\lambda1$:$\lambda2$ ratios were also calculated, where $\lambda1$ is the eigenvalue of the first constrained CCA axis and $\lambda2$ is the eigenvalue of the second unconstrained axis. RDA was applied on eastern Baltic TS owing to the shorter length of DCA Ax1 and Ax2 (2.4 and 2.0 SD units, respectively).

The MJAT optima for individual taxa across the Finno–Baltic–Polish TS were estimated based on weighted-average regression with inverse deshrinking (Ter Braak & Juggins 1993). The generalized additive models (GAMs) (Wood 2011) in the whole Finno–Baltic–Polish TS were applied to estimate the taxon-specific MJAT–abundance relationships. For taxa for which a significant relationship with MJAT was revealed, the variance was calculated. To study the temperature-related distribution of taxa, the TS was divided into three biogeographic zones based on MJAT intervals: northern boreal (12.1–15.0 °C), southern boreal (15.0–17.0 °C) and temperate (17.0–19.2 °C) (Fig. 1B). Linear regression was applied to the taxa presented within the biogeographic zones to ensure that they reveal significant dependency in the same biogeographical zone where their weighted averaging-based optima are situated. For some understudied taxa, additional GAMs for estimating the dependency between their abundances and dissolved oxygen level, water phosphorus, sampling depth and pH values were calculated.

### *The Chironomid-inferred MJAT model building*

The Finno–Baltic–Polish inference model was built using cross-correlated weighted averaging–partial least squares (WA-PLS) regression and calibration (Ter Braak & Juggins 1993). The best transfer functions were selected as those producing the lowest cross-validated root mean squared error of prediction (RMSEP). The components were accepted as statistically significant at the $p \leq 0.05$ level. Bootstrapping techniques (9999 permutations; Birks *et al*. 1990; Birks 1998) were used to calculate cross-validated error and performance statistics for the WA-PLS inference model, such as the RMSEP, the maximum bias, the mean bias and the coefficient of determination ($R^2$) between inferred and predicted values within the individual calibration TSs (Finnish, Finno–Baltic–Polish and Swiss–Norwegian TSs). July air optima and tolerance of the individual taxa were estimated based on the WA regression (Ter Braak & Juggins 1993).

Numerical analysis and plots were performed with the free software program R version 4.1.1. using the following packages: 'tidyverse' for data restructuring and visualizing (Wickham *et al*. 2019), 'dplyr' for data restructuring and basic calculations (Wickham *et al*. 2022), 'vegan' for ordination (Oksanen *et al*. 2022), 'rioja' for performing the WA-PLS analysis and plotting the stratigraphic diagram (Juggins 2022) and 'mgcv' for performing the GAMs (Wood 2017).

## Results

### *Composition of Chironomidae assemblages*

A total of 112 Chironomidae morphotypes were identified in 35 lakes from the eastern Baltic region, of which 30 morphotypes were not present at the Finnish and Polish sites. The most abundant taxa (Fig. 2) in the eastern Baltic TS are *Psectrocladius sordidellus* type (8.6%), *Chironomus plumosus* type (8.6%), *Dicrotendipes nervosus* type (6.8%), *Neozavrelia* (6.8%), *Polypedilum nubeculosum* type (5.28%), *Glyptotendipes pallens* type (4.8%), *Microtendipes pedellus* type (4.6%) and *Corynoneura ambigua* type (4.1%).

The merged Finno–Baltic–Polish TS after deleting species with abundances less than 2% contains 106 taxa and 121 sites. Following Luoto (2009), five sites from the Finnish TS were considered outliers and excluded from the MJAT inference model. The most abundant taxa in the final TS were *P. sordidellus* type (average abundance 15.6%), *C. plumosus* type (7.3%), *Ablabesmyia* (7.2%), *Procladius* (7%), *M. pedellus* type (6.8%), *Cladotanytarsus mancus* type (6.1%) and *Zalutschia zalutschicola* type (5.7%).

### *Ordination analysis*

The variance gradient of DCA calculated for eastern Baltic Chironomidae assemblages equals 2.4 SD and 2.0 SD for Ax1 and Ax2, respectively, whereas the Finno–Baltic–Polish TS DCA variance gradient equals 3.1 SD and 2.4 SD, respectively (Fig. 3).

The RDAs of the 35 eastern Baltic sites revealed that MJAT and water phosphorus content were the only environmental variables that explain the Chironomidae community distribution significantly (Table 2). In contrast, the RDAs and CCAs of the Finno–Baltic–Polish TS revealed that each of the examined

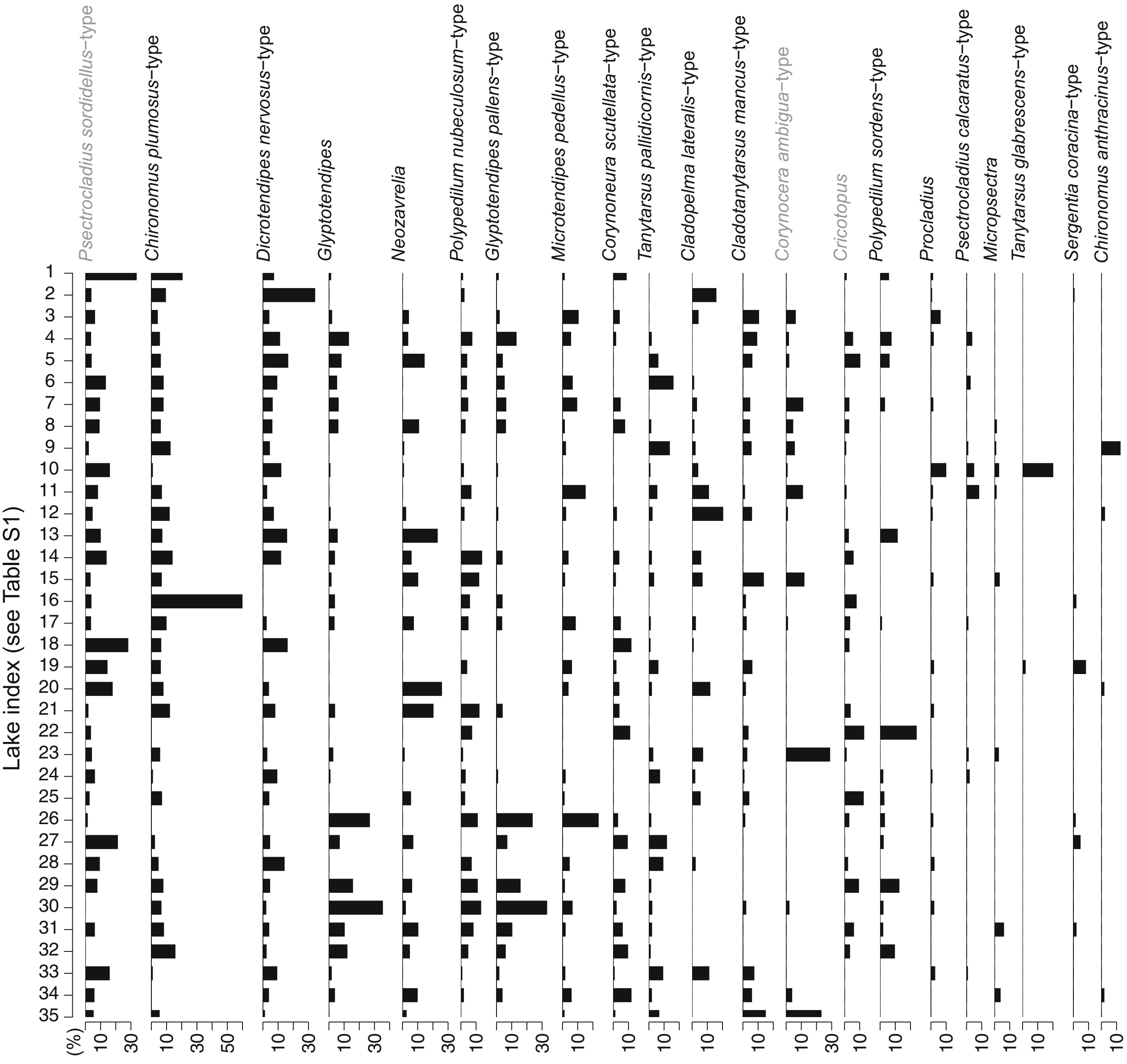

*Fig. 2.* Selected eastern Baltic chironomid morphotypes from the lake surface sediment layer ((0–2)–2 cm) included in the Finno–Baltic–Polish TS. Lake indices are ordered according to decreasing latitude (Table S1). Species are organized according to the decreasing abundances in the eastern Baltic TS; taxa with MJAT dependency are marked with grey.

environmental variables explains the Chironomidae assemblages significantly. The MJAT explains the 9.1% CCA-based (14.4% RDA-based) Chironomidae assemblage variance followed by pH (7.2% (CCA), 11.5 (RDA)), dissolved oxygen (4.5% (CCA), 7.5% (RDA)) and sampling depth (3.0% (CCA), 3.9% (RDA)) (Table 3, Fig. 4). The strongest explanatory variable for the Chironomidae community according to the $\lambda 1{:}\lambda 2$ ratio was MJAT (1.4 CCA-based, 1.3 RDA-based; Table 3). Training sets with ratios $\lambda 1{:}\lambda 2 > 1$ are characterized by strong relationships between the examined environmental variable and the assemblage data relative to the remaining variance in the TS and are therefore suitable for developing inference models (Goldenberg Vilar *et al.* 2018). However, it must be kept in mind that the CCA plot demonstrating the dependency of chironomid assemblages from MJAT exhibits the horseshoe effect, which could affect the results.

To assess the relationship between MJAT and chironomid assemblages in different parts of the Finno–Baltic–Polish TS, CCA with MJAT as an explanatory variable for different subsets was performed (Table 4). The variance explained by the MJAT in Finno–Baltic–Polish TS is 9.1% (CCA-based), which is almost the same as in only the Finnish TS (9.6%). The ratio of the unique taxa in the TS (Table 4) was calculated as a percentage of taxa

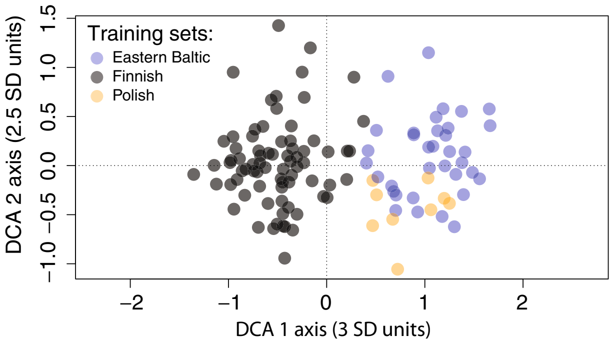

*Fig. 3.* Detrended correspondence analysis (DCA) ordination diagram for the lakes from the Finno–Baltic–Polish TS.

observed only in a particular TS. This ratio is almost equal for the Polish–Baltic area and the Finnish one.

### *Species–MJAT relations analysis*

The linear regression analyses (Table S2) indicated that the most MJAT-dependent taxa in the different biogeographic zones of the Finno–Baltic–Polish TS were *Micropsectra insignilobus* type (35.5% of variance explained), *Zalutschia mucronata* type (16%), *Psectrocladius septentrionalis* type (15.3%), *Tanytarsus lugens* type (14.8%) and *Psectrocladius calcaratus* type (12.3%) for the northern boreal part of the TS; *Procladius* (23.1%), *Tanytasus mendax* type (21.6%), *C. scutellata* type (20.8%), *Chironomus anthracinus* type (18.9%) and *Ablabesmyia* (11.9%) for the southern boreal part; and *D. nervosus* type (35.5%), *Ablabesmyia* (33.13%), *C. arctica* type (25.6%), *C. plumosus* type (21.2%) and *Neozavrelia* (20.2%) in the temperate part of the gradient. The taxa whose distribution revealed the strongest MJAT dependence based on GAM over the entire Finno–Baltic–Polish TS (Fig. 5) were *P. septentrionalis* type (51.6% of variation explained by the MJAT), *P. calcaratus* type (47.6%), *M. insignilobus* type (43.7%), *T. mendax* type (39.4%), *Ablabesmyia* (35.6%), *Procladius* (33.7%), *Dicrotendipes pulsus* type (32%) and *C. plumosus* type (28.3%). In contrast, *P. sordidellus* type, *M. pedellus* type and *Corynoneura lobata* type did not reveal any MJAT dependency at all. *Limnophyes* and *C. ambigua* type morphotypes had MJAT dependency only within one biogeographical zone but not across the entire gradient. The average variation explained by MJAT for the different taxa was 9.2% less within individual zones than in the GAM-based estimations for the whole Finno–Baltic–Polish TS. The July air temperature optima for each taxon are situated in the same biogeographic zones, where they reveal significant distribution (Table S2).

### *Inference model for Chironomid-inferred July air temperature reconstructions*

The Finnish and Finno–Baltic–Polish TSs WA-PLS cross-validation models both demonstrated the smallest RMSEP value with the two-component model with 0.8 and 0.7 °C, respectively (Table 5). The bootstrapped maximum bias was 1.1 and 0.4 °C, and $R^2$ was 0.9 for Finnish and Finno–Baltic–Polish TSs. Thus, the performance of the new TS is comparably similar to that of the Finnish TS. Scatterplots of cross-validated predicted vs. observed MJATs in the Finno–Baltic–Polish cross-validation model generally follow a 1:1 relationship (Fig. 6). However, observed MJATs below 13 °C are consistently overestimated by the model, suggesting a minor bias on the colder end of the temperature gradient.

## Discussion

The DCA revealed that eastern Baltic and Polish TSs form a separate group from Finnish TS sites, whereas samples from Poland are closer and partially embedded into the eastern Baltic sample space (Fig. 3). Even though 30 taxa were different from the Finnish or Polish Chironomidae communities, only seven of them reached an abundance of 2% at least in one sample and were included in the Finno–Baltic–Polish TS. These seven unique taxa account for 9.1% of the whole Finno–Baltic–Polish TS Chironomidae taxa composition (Table 4). Thus, the eastern Baltic and Polish samples provide new information about species communities and give new modern analogues to the Finno–Baltic–Polish TS.

### *Chironomidae assemblages' sensitivity to the environmental parameters*

Most MJAT relationships were studied on a wide regional or climatological scale with a gradient length greater than

*Table 2.* The results of RDA of the eastern Baltic TS using only one environmental variable. The number of sites with the respective environmental variable measured, the amount of variability explained, *p*-values and the $\lambda1:\lambda2$ ratio are provided.

| Name of the measured variable | Mean July air temperature | Water depth | Dissolved oxygen | pH | Water phosphorus content |
|---|---|---|---|---|---|
| Number of sites | 35 | 35 | 29 | 31 | 29 |
| Percentage of variability explained | 5.4 | – | – | – | 5.6 |
| *p*-Value | 0.048 | 0.1 | 0.2 | 0.1 | 0.047 |
| $\lambda1:\lambda2$ | 0.37 | – | – | – | 0.34 |

*Table 3.* The results of canonical correspondence analysis (CCA) and RDA of the Finno–Baltic–Polish TS using only one environmental variable. The number of sites with the respective environmental variable measured, the amount of variability explained, *p*-values and the $\lambda 1:\lambda 2$ ratio are provided.

| Name of the measured variable | Mean July air temperature | Water depth | Dissolved oxygen | pH |
|---|---|---|---|---|
| Number of sites | 121 | 121 | 56 | 79 |
| Ordination method | CCA/RDA | CCA/RDA | CCA/RDA | CCA/RDA |
| Percentage of variability explained | 9.1/14.4 | 3.0/3.9 | 4.5/7.5 | 7.2/11.5 |
| *p*-Value | 0.001/0.001 | 0.001/0.001 | 0.001/0.001 | 0.001/0.001 |
| $\lambda 1:\lambda 2$ | 1.4/1.3 | 0.3/0.23 | 0.5/0.4 | 1/1 |

5 °C (Larocque *et al.* 2001; Luoto 2009); however, the minimal gradient length requirements are unknown. In the eastern Baltic TS, the MJAT explains a significant proportion (5.4% RDA-based) of the Chironomidae assemblage composition, suggesting that a MJAT range of approximately 2 °C as in the eastern Baltic TS is already enough to find significant MJAT–Chironomidae relationships. The RDA and CCA of the Finno–Baltic–Polish TS revealed that all observed environmental variables had a significant influence on the chironomid communities, with the MJAT explaining the highest amount of distribution (9.1% CCA-based) and having the highest $\lambda 1:\lambda 2$ ratio (1.4) of the studied ones. This agrees with other chironomid calibration TSs from Europe: 6.2% for the Swiss–Norwegian TS (Heiri *et al.* 2011), 7.7% for the northern Sweden TS (Larocque *et al.* 2001) and 9.6% for the Swiss–Norwegian–Polish TS (Kotrys *et al.* 2020) TSs.

The pH, water depth and oxygen level in the eastern Baltic TS (Table 2) were non-significant and this might be explained by the short gradient range and generally stable environmental conditions. The water phosphorus content has been recorded to have a significant influence on the Chironomidae communities (Brooks *et al.* 2001; Luoto 2011). In contrast to the eastern Baltic TS the statistically significant secondary factors in the Finno–Baltic–Polish TS are pH gradient (7.2% (CCA), 11.5% (RDA) of distribution explained), dissolved oxygen (4.5% (CCA), 7.5% (RDA)) and sampling depth (3.0% (CCA), 3.9% (RDA)). The pH is known to be an influential factor in the limnology of aquatic environments and the distribution of chironomids (Orendt 1999;

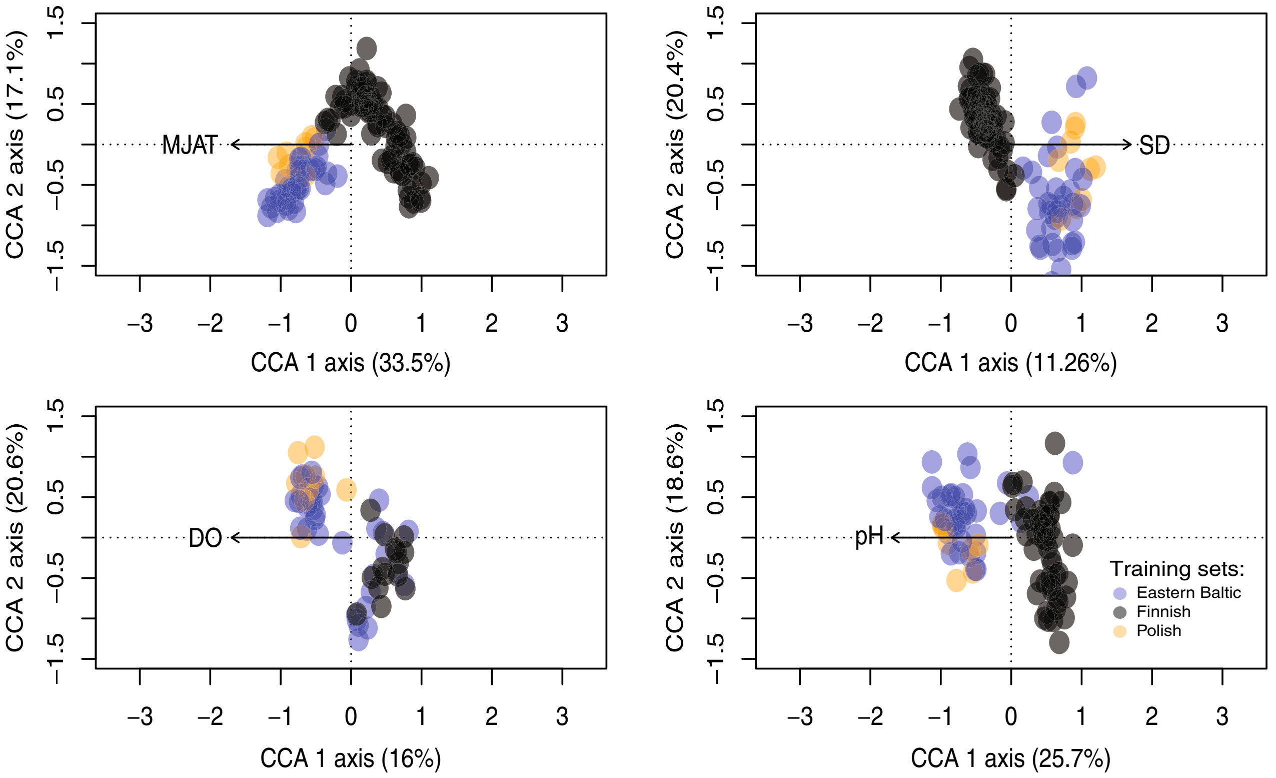

*Fig. 4.* The canonical correspondence analysis (CCA) biplots for the lakes from the Finno–Baltic–Polish TS with MJAT, sampling depth (SD), dissolved oxygen (DO) and pH as explanatory variables. The eigenvalues of the CCA1 and CCA2 axes are 0.3191 and 0.0569, respectively.

*Table 4*. Results of CCAs of the different parts of the Finno–Baltic–Polish and Swiss–Norwegian full dataset (Heiri *et al.* 2011) TSs set with mean July air temperature as an explanatory variable. The number of sites, variability explained by the environmental factor, $p$-values and $\lambda1{:}\lambda2$ ratio are provided for each subset and combinations of these. MJAT = mean July air temperature.

| | Training set | | | |
|---|---|---|---|---|
| | Polish–Baltic (Polish; Baltic) | Finnish | Finno–Baltic–Polish | Swiss–Norwegian |
| Number of sites | 44 (9; 35) | 77 | 121 | 265 |
| Total taxa included | 82 (53; 77) | 79 | 101 | 154 |
| Unique taxa for the component (%) | 26.8 (0; 9.1) | 24.0 | – | 100 |
| Variance explained by MJAT (°C) (%) | 6.3 (–; 5.4)[1] | 9.6 | 9.1 | 5.6 |
| $p$-Value | 0.001 (0.45; 0.048) | 0.001 | 0.001 | 0.001 |
| $\lambda1{:}\lambda2$ | 0.39 (–; 0.37) | 1.37 | 1.37 | 1.13 |
| MJAT (°C) | 17.2–19.2 (18.4–19.2; 17.2–18.4) | 12.1–17.5 | 12.1–19.2 | 3.5–18.4 |

[1] RDA-based.

Brooks *et al.* 2007). For instance, in the Swiss–Norwegian–Polish TS (Kotrys *et al.* 2020), it explains 3.85%, and in the TS from Russia (Self *et al.* 2011), it explains 3.8%. In the Finno–Baltic–Polish TS, pH is the second most significant factor, although in the eastern Baltic part of the TS, pH does not explain the

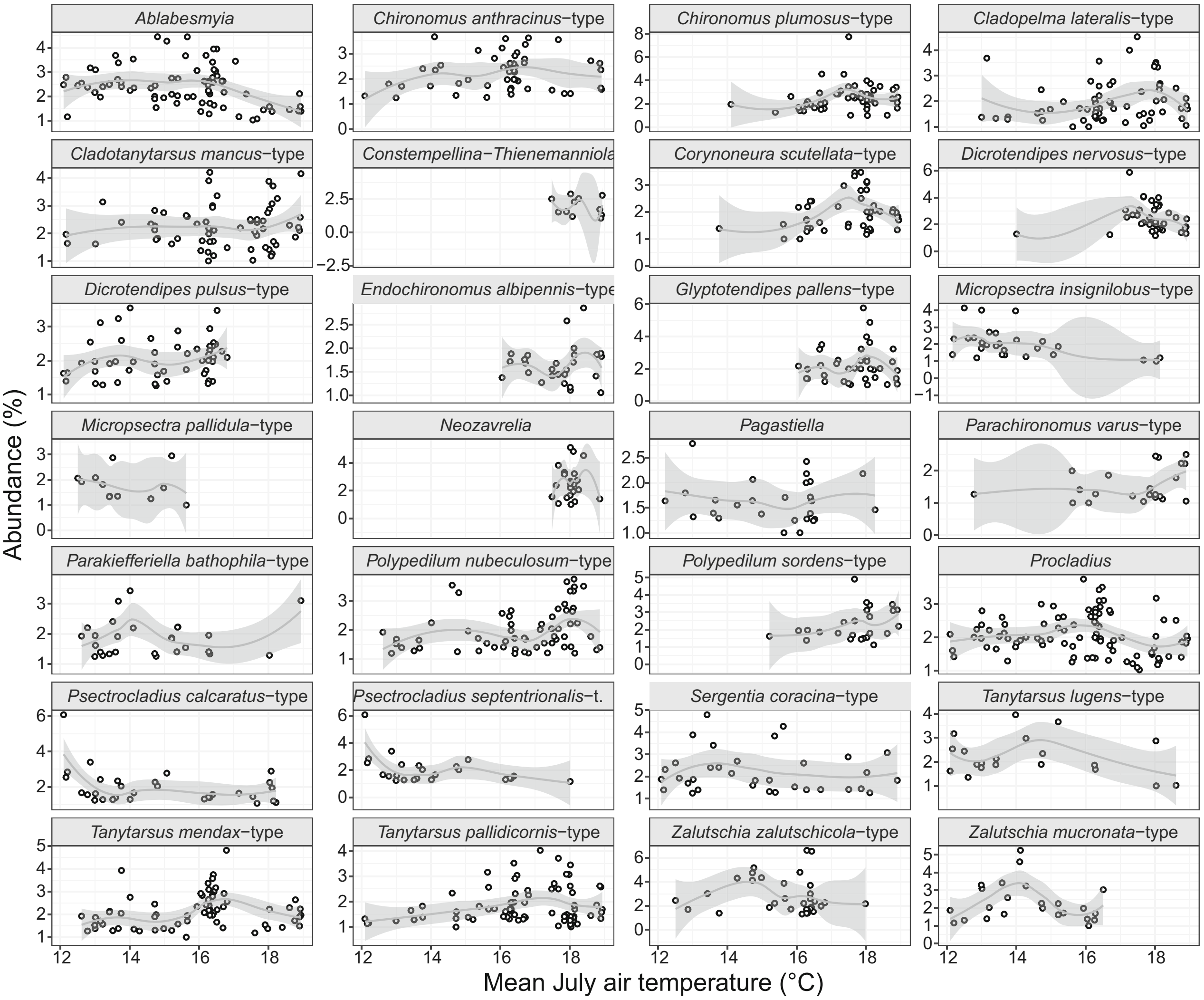

*Fig. 5.* Distribution of taxa with a significant relationship with temperature and abundances >2% in the Finno–Baltic–Polish TS. The lines are calculated by loess smoothing over the whole TS.

*Table 5.* Weighted averaging–partial least squares cross-validated models based on the Finnish and Finno–Baltic–Polish TS error statistics. Values are based on bootstrapping with 9999 bootstrap cycles except for the RMSE, which represents non-cross-validated values. Outliers and the least abundant taxa (max abundance under 2%) were excluded before calculating the inference model. RMSE = root mean square error; RMSEP = root mean square error of prediction; WA-PLS = weighted averaging–partial least squares.

| Training set | Finnish | | | Finno–Baltic–Polish | | |
|---|---|---|---|---|---|---|
| WA-PLS components | 1 | 2 | 3 | 1 | 2 | 3 |
| RMSE (°C) | 0.6 | 0.5 | 0.4 | 0.8 | 0.6 | 0.4 |
| RMSEP (°C) | 0.8 | 0.8 | 0.8 | 0.9 | 0.7 | 0.7 |
| Maximum bias (°C) | 0.8 | 1.1 | 0.4 | 1.2 | 0.4 | 0.3 |
| Average bias (°C) | 0.02 | 0.01 | 0.00 | 0.00 | 0.01 | −0.01 |
| $R^2$ | 0.8 | 0.9 | 0.9 | 0.8 | 0.9 | 0.9 |

chironomids distribution significantly. The pH ranges between 4.5 and 8.9 in the eastern Baltic lakes, with only three lakes having pH values lower than 7 and 32 lakes having pH values of 7–8.9. Thus, eastern Baltic Chironomidae communities reflect the changes within 1.9 pH units; however, adding the alkaline eastern Baltic sites to the more acidic Finnish ones increases the influence of the pH variable from 2.4% (CCA-based; Luoto 2009) to 7.2% (CCA-based). Sampling depth explained 2.4% of the variation in the Finnish TS and increased to 3.0% after merging with eastern Baltic and Polish lakes. The increase in the explanatory power is the result of adding lakes with depths of 10–21 m, whereas the deepest lake in the Finnish TS is 9 m.

Based on the GAM, among 44 taxa that had a mean abundance higher than 2% in the northern boreal, southern boreal or temperate parts of the TS (Table S2), 30 morphotypes revealed significant MJAT dependency (Fig. 5). Thus, the estimated taxa groups and their proportions can be considered characteristic features for the corresponding biogeographical zones.

Dominant MJAT-dependent taxa in the eastern Baltic and Polish sites (17.5–19.2 °C) (*C. plumosus* type, *D. nervosus* type, *Tanytarsus pallidicornis* type, *Polypedilum sordens* type, *Cladopelma lateralis* type, *G. pallens* type, *P. nubeculosum* type, *C. mancus* type, *C. lateralis* type and *Constempellina–Thienemanniola*) were also considered warm-related in the Polish, Swiss–Norwegian, Canadian and northern America TSs (Walker *et al.* 1997; Heiri *et al.* 2011; Kotrys *et al.* 2020; Medeiros *et al.* 2022). *Chironomus plumosus* type, *C. lateralis* type, *P. sordens* type and *G. pallens* type have rare findings in the northern boreal part of the TS; however, they increase their abundance when the MJAT is warmer than 16 °C. *Dicrotendipes nervosus* increases its abundances above 17 °C. *Tanytarsus pallidicornis* type as well as *C. mancus* type appear in all parts of the Finno–Baltic–Polish TS, although they tend to increase in abundance from 16 to 19.1 °C. *Polypedilum nubeculosum* type, which has always been reported as a warm-related taxon (Heiri *et al.* 2011; Medeiros *et al.* 2022), has two peaks of distribution according to GAM (Fig. 5)–one at approximately 14.5 °C and another at approximately 18 °C – which could be explained by the occurrence of different species included in the type or by the presence of compensatory ecological factors in the northern boreal zone of the Finno–Baltic–Polish TS. Also, the *P. nubeculosum* type includes many species, whose subfossils cannot be separated from each other; it is possible that peaks represent two different species – one dominating in the northern boreal zone of MJAT gradient and another dominating in the temperate Baltic–Polish range of the TS. The *Constempellina–Thienemanniola* morphotype revealed MJAT optima of approximately 18.3 °C, with 16.2% variance explained, thus the type can be considered as warm-adapted in the Finno–Baltic–Polish TS. This is confirmed by the Swiss–Norwegian–Polish TS when it appears in many warm lakes of northern Poland. *Neozavrelia* has been described as a stenotherm taxon distributed in calcareous waters

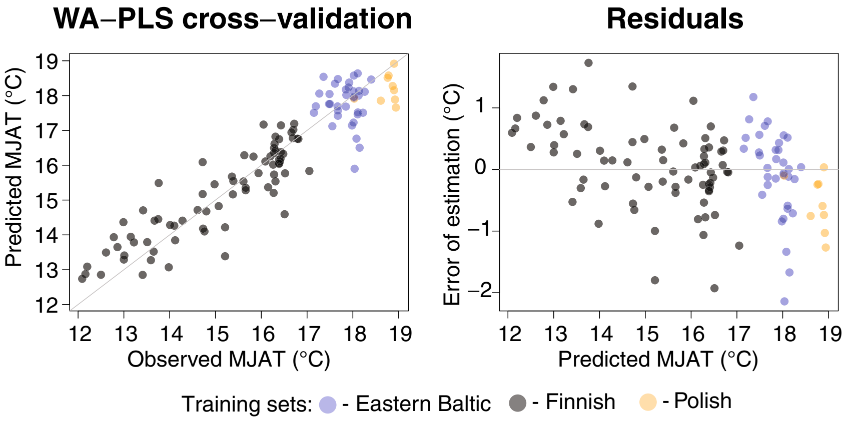

*Fig. 6.* Diagnostic plots of cross-validated estimates and prediction residuals compared with observed values of the Finno–Baltic–Polish TS calculated with a WA-PLS-based model based on two components.

(Ekrem 2006; Brooks *et al.* 2007). In the Finno–Baltic–Polish TS, *Neozavrelia* shows a strong positive relationship with MJAT and is mainly found in calcareous waters, which is in good agreement with Brooks *et al.* (2007). *Procladius* appears over the whole Finno–Baltic–Polish TS, with a broad distribution, which can be explained by the low identification resolution and the existence of regional morphotypes with different climatic optimums. The same observation was made within the study from Russia (Self *et al.* 2011), where *Procladius* had been considered a cosmopolitan taxon. *Chironomus plumosus* type, *C. lateralis* type and *T. glabrescens* type, although present in the temperate and southern boreal biogeographical zones, have significant positive correlations with a MJAT only in the temperate part of the Finno–Baltic–Polish TS. These taxa also seem to be related to secondary environmental variables along their distribution. For example, *C. plumosus* type has a strong negative relationship with dissolved oxygen concentrations, which agrees with studies by Little & Smol (2000) with ecological data that indicate that the taxon is related to low oxygen concentrations. *Limnophyes* and *C. ambigua* type revealed MJAT dependency in the southern and northern boreal zones respectively but not on the scale of the whole Finno–Baltic–Polish TS. *Limnophyes* has a significant MJAT relationship in the southern boreal zone of the TS and does not have any dependencies with secondary environmental gradients, so the pattern of its distribution must be clarified in future research. Previously, *Limnophyes* was described as a MJAT-related species and a possible indicator of water level fluctuations (Brooks *et al.* 2007), and a negative correlation with water depth was found (Self *et al.* 2011). The distribution of *C. ambigua* type revealed two MJAT optima in the Swiss–Norwegian–Polish TS (Kotrys *et al.* 2020), which together with findings in the current study, can be the basis for considering the presence of different cryptic species inside the type.

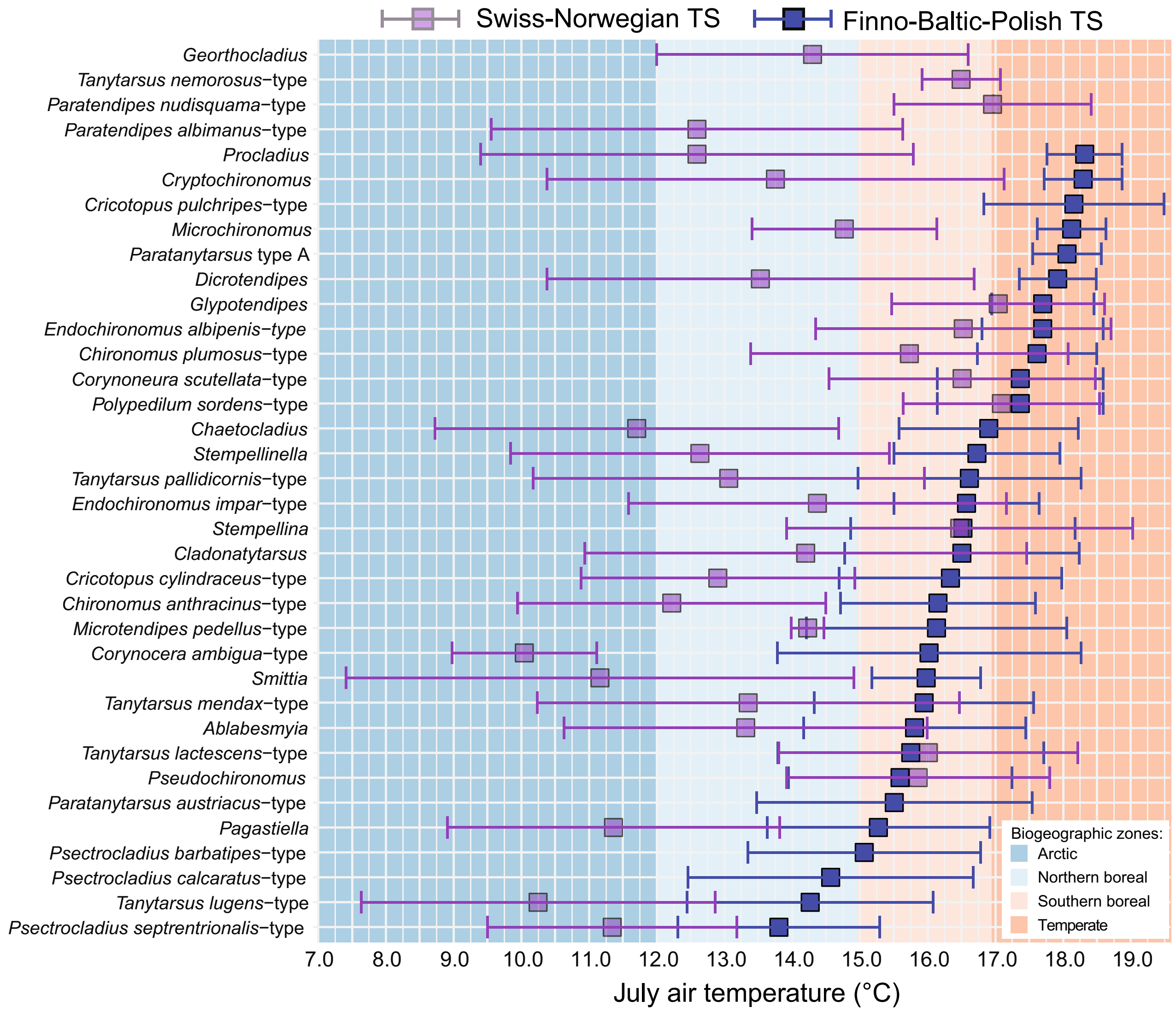

*Fig. 7.* The WA-based July taxa optima calculated in Finno–Baltic–Polish and Swiss–Norwegian TSs.

### *Inference model cross-validation and taxa MJAT optima estimations*

The cross-validation performance of the Finno–Baltic–Polish TS is slightly better compared with the Finnish TS (Table 5; 0.7 and 0.8 of RMSEP ( °C) respectively). Furthermore, the RMSEP, obtained from the Finno–Baltic–Polish WA-PLS model is comparably smaller than some of the other widely used TSs. For instance, the North American TS exhibited an RMSEP of 1.93 (Fortin *et al*. 2015) and the Swiss–Norwegian TS revealed an RMSEP of 1.4 (Heiri *et al*. 2011).

For most WA-based morphotypes, their MJAT optima are situated in the biogeographic zones, where their distribution is significant. Thus the calculated optima values are expected to be reliable. Generally, optima for the same taxa, calculated based on the Finno–Baltic–Polish TS, are warmer than those calculated using the Swiss–Norwegian TS (Fig. 7). Some taxa revealed the same MJAT optima regardless of TS origin (*Glyptotendipes*, *P. sordens* type, *Stempellina*, *Tanytarsus lactescens* type, *Pseudochironomus*). *Endochironomus albipennis* type, *C. plumosus* type and *C. scutellata* type MJAT optima in Finno–Baltic–Polish TS are within the error bars of those in the Swiss–Norwegian TS, e.g. exhibiting similar but narrower distribution areas. These morphotypes have a very low abundance or are absent in the Finno–Baltic–Polish TS sites with MJAT <16 °C. However, the same taxa revealed a more cold-adapted distribution pattern, appearing in sites with <14 °C MJAT in the Swiss–Norwegian TS, especially in the Swiss part of it (Heiri *et al*. 2011). The broader difference can be related to geographic position-, climate- or environment-driven differences (elevation, bedrock acidity, climate continentality, etc.) in the Swiss Alps and Finno–Baltic–Polish areas. *Endochironomus impar* type, *P. septentrionalis* type, *T. lugens* type, *Pagastiella* and *Ablabesmyia* exhibited optima with the intersecting error bars in the tested TSs. Considering that in the Finno–Baltic–Polish TS these taxa revealed quite a cosmopolitical distribution pattern (Fig. 5), we conclude that the same pattern is apparent in the Swiss–Norwegian TS. The most pronounced differences in optima are observed for taxa that were identified on a lower taxonomic level (e.g. *Procladius*, *Dicrotendipens*, *Chaetocladius*, *Microchironomus*, *Stempellinella*, *Smittia*) and can include species with different climatic preferences, thus leading to the mismatch of the optima in two tested TSs even within the error bars. This aligns well with the study of Heiri & Lotter (2010), which stated that the chironomid-based TSs with the high taxonomic resolution have the smallest RMSEP.

The absence of the cold (<12 °C) part of the MJAT gradient in the Finno–Baltic–Polish TS makes it impossible to reconstruct the colder phases of the Lateglacial climate. This could be handled by further adding colder sites, e.g. northern Norwegian ones. The advantage of the new Fino–Baltic–Polish TS is in providing warmer modern analogues, which fixes the issue of the warm end of the gradient in the Finnish TS. This statistical issue mentioned previously (Heiri *et al.* 2011) appears because of the lack of warm analogues and results in inferred July temperature overestimation. Having adequate warm analogues from the eastern Baltic is beneficial in reconstructions of the warm periods of the Middle and Late Holocene along with estimations of the previous July temperature change rates, which can be compared and discussed with modern ones.

## Conclusions

The eastern Baltic TS contributes new information about Chironomidae taxa distribution patterns in the understudied Baltic region. The modern assemblages improve considerably the statistical performance of the inference model of July temperature when merged with other TSs from the region. Thus, building the Chironomidae TS representative of the environmental conditions of the study area is an important prerequisite for high-quality chironomid-based palaeoenvironmental reconstructions. The advantages of the Finno–Baltic–Polish TS are geographic and climatic continuity and the presence of warm modern analogues.

The MJAT and water phosphorus content were the only statistically significant environmental variables among the other tested ones (pH, water depth and dissolved oxygen level) which explained a Chironomidae distribution in the eastern Baltic TS. Thus, the MJAT range of approximately 2 °C as in the eastern Baltic TS is already sufficient to find significant MJAT–Chironomidae relationships. The MJAT had the highest explanatory power in the regional Finno–Baltic–Polish TS.

The differences in estimated Chironomidae distribution–MJAT optima in the Finno–Baltic–Polish and Swiss–Norwegian TSs highlight the importance of the TS selection. The differences can be attributed to the temperature range represented by the TS, taxonomic identification level, general cosmopolitan taxa distribution pattern and the influence of geographic position-, climate- or environment-driven differences. These results suggest that further tests are needed to investigate the impact of Chironomidae TS taxonomic composition and biogeographic origin on palaeotemperature reconstructions.

*Acknowledgements*. – This study was financially supported by the Estonian Research Council grant PRG323 and PRG1993. Varvara Bakumenko was supported by the Doctoral School of Earth Sciences and Ecology, supported by the European Union, European Regional Development Fund (ASTRA 'TTÜ arenguprogramm aastateks 2016–2022'). The water chemistry data were collected by Anna Lanka (Tallinn University of Technology). Special gratitude goes to Oliver Heiri for chironomid identification help and scientific guidance. The three journal reviewers are gratefully acknowledged for their valuable suggestions. We express our gratitude to Jan A. Piotrowski, the Editorin-Chief, for his diligent corrections and editorial management. None of the authors has any competing interests.

*Author contributions.* – Conceptualization: VB, SV, AP; Laboratory analysis/sample collection: VB, SV, MP, BK, TPL; Supervision – SV, SB; Writing – original draft: VB; Writing – review and editing: VB, AP, SV, SB, MP, NG.

*Data availability statement.* – Data available on request.

## References

Andersen, T., Sæther, O. A., Cranston, P. S. & Epler, J. H. 2013: The larvae of Orthocladiinae (Diptera: Chironomidae) of the Holarctic region-keys and diagnoses. *Insect Systematics & Evolution 66*, 189–385.

Barley, E. M., Walker, I. R., Kurek, J., Cwynar, L. C., Mathewes, R. W., Gajewski, K. & Finney, B. P. 2006: A northwest North American training set: distribution of freshwater midges in relation to air temperature and lake depth. *Journal of Paleolimnology 36*, 295–314.

Birks, H. J. B. 1998: D.G. Frey and E.S. Deevey review 1: numerical tools in palaeolimnology – progress, potentialities, and problems. *Journal of Paleolimnology 20*, 307–332.

Birks, H. J. B., Braak, C. T., Line, J. M., Juggins, S. & Stevenson, A. C. 1990: Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 327*, 263–278.

Birks, H. J. B., Heiri, O., Seppä, H. & Bjune, A. E. 2010: Strengths and weaknesses of quantitative climate reconstructions based on Late-Quaternary. *The Open Ecology Journal 3*, 68–100.

Brooks, S. J. & Birks, H. J. B. 2000: Chironomid-inferred late-glacial and early-Holocene mean July air temperatures for Kråkenes Lake, western Norway. *Journal of Paleolimnology 23*, 77–89.

Brooks, S. J., Bennion, H. & Birks, H. J. B. 2001: Tracing lake trophic history with a chironomid–total phosphorus inference model. *Freshwater Biology 46*, 513–533.

Brooks, S. J., Langdon, P. G. & Heiri, O. 2007: The identification and use of Palaearctic Chironomidae larvae in palaeoecology. *Quaternary Research Association Technical Guide 10*, i–vi.

Chamutiova, T., Hamerlik, L. & BituŠÍk, P. 2020: Subfossil chironomids (Diptera, Chironomidae) of lakes in the Tatra Mountains: an illustrated guide. *Zootaxa 4819*, 216–264.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. & Jones, P. D. 2018: An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres 123*, 9391–9409.

Eggermont, H. & Heiri, O. 2012: The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. *Biological Reviews 87*, 430–456.

Ekrem, T. 2006: A redescription of *Neozavrelia cuneipennis* (Edwards) comb. nov., with a checklist of *Neozavrelia* species of the world (Diptera: Chironomidae). *Zootaxa 1153*, 1–16.

Feurdean, A., Perşoiu, A., Tanţău, I., Stevens, T., Magyari, E. K., Onac, B. P., Marković, S., Andrič, M., Connor, S., Fărcaş, S. & Gałka, M. 2014: Climate variability and associated vegetation response throughout Central and Eastern Europe (CEE) between 60 and 8 ka. *Quaternary Science Reviews 106*, 206–224.

Fortin, M. C., Medeiros, A. S., Gajewski, K., Barley, E. M., Larocque-Tobler, I., Porinchu, D. F. & Wilson, S. E. 2015: Chironomid-environment relations in northern North America. *Journal of Paleolimnology 54*, 223–237.

Fu, C. 1992: Transitional climate zones and biome boundaries: a case study from China. *In* Hansen, A. J. & di Castri, F. (eds.): *Landscape Boundaries. Ecological Studies 92*, 394–402. Springer, New York.

Giesecke, T. A. E. B., Bjune, A. E., Chiverrell, R. C., Seppä, H., Ojala, A. E. K. & Birks, H. J. B. 2008: Exploring Holocene continentality changes in Fennoscandia using present and past tree distributions. *Quaternary Science Reviews 27*, 1296–1308.

Goldenberg Vilar, A., Donders, T., Cvetkoska, A. & Wagner-Cremer, F. 2018: Seasonality modulates the predictive skills of diatom based salinity transfer functions. *PLoS One 13*, e0199343, https://doi.org/10.1371/journal.pone.0199343.

Gouw-Bouman, M. T. I. J., Van Asch, N., Engels, S. & Hoek, W. Z. 2019: Late Holocene ecological shifts and chironomid-inferred summer temperature changes reconstructed from Lake Uddelermeer, the Netherlands. *Palaeogeography, Palaeoclimatology, Palaeoecology 535*, 109366, https://doi.org/10.1016/j.palaeo.2019.109366.

Heiri, O. 2004: Within-lake variability of subfossil chironomid assemblages in shallow Norwegian lakes. *Journal of Paleolimnology 32*, 67–84.

Heiri, O. & Lotter, A. F. 2010: How does taxonomic resolution affect chironomid-based temperature reconstruction? *Journal of Paleolimnology 44*, 589–601.

Heiri, O., Brooks, S. J., Birks, H. J. B. & Lotter, A. F. 2011: A 274-lake calibration data-set and inference model for chironomid-based summer air temperature reconstruction in Europe. *Quaternary Science Reviews 30*, 3445–3456.

Heiri, O., Brooks, S. J., Renssen, H., Bedford, A., Hazekamp, M., Ilyashuk, B., Jeffers, E. S., Lang, B., Kirilova, E., Kuiper, S. & Millet, L. 2014: Validation of climate model-inferred regional temperature change for late-glacial Europe. *Nature Communications 5*, 4914, https://doi.org/10.1038/ncomms5914.

Hill, M. O. & Gauch, H. G. 1980: Detrended correspondence analysis: an improved ordination technique. *Vegetatio 42*, 47–58.

Johannessen, O. M. 1970: Note on some vertical profiles below ice floes in the Gulf of St. Lawrence and near the North Pole. *Journal of Geophysical Research 75*, 2857–2861.

Juggins, S. 2013: Quantitative reconstructions in palaeolimnology: new paradigm or sick science? *Quaternary Science Reviews 64*, 20–32.

Juggins, S. 2022: *rioja: analysis of quaternary science data, R package version (1.0-5)*. Available at: https://cran.r-project.org/web/packages/rioja/rioja.pdf.

Kaufman, D., McKay, N., Routson, C., Erb, M., Davis, B., Heiri, O., Jaccard, S., Tierney, J., Dätwyler, C., Axford, Y. & Brussel, T. 2020: A global database of Holocene paleotemperature records. *Scientific Data 7*, 115, https://doi.org/10.1038/s41597-020-0445-3.

Klink, A. G. & Pillot, H. K. M. 2003: *Chironomidae Larvae: Key to the Higher Taxa and Species of the Lowlands of Northwestern Europe*. ETI 6338, Expert Center for Taxonomic Identification, University of Amsterdam, Amsterdam.

Kotrys, B., Płóciennik, M., Sydor, P. & Brooks, S. J. 2020: Expanding the Swiss–Norwegian chironomid training set with Polish data. *Boreas 49*, 89–107.

Larocque, I., Hall, R. I. & Grahn, E. 2001: Chironomids as indicators of climate change: a 100-lake training set from a subarctic region of northern Sweden (Lapland). *Journal of Paleolimnology 26*, 307–322.

Larocque-Tobler, I. 2014: The Polish sub-fossil chironomids. *Palaeontologia Electronica 17*, 1–28.

Lepš, J. & Šmilauer, P. 2003: *Multivariate Analysis of Ecological Data Using CANOCO*. 280 pp. Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511615146.

Little, J. L. & Smol, J. P. 2000: Changes in fossil midge (Chironomidae) assemblages in response to cultural activities in a shallow, polymictic lake. *Journal of Paleolimnology 23*, 207–212.

Lotter, A. F., Birks, H. J. B., 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. *Journal of Paleolimnology 18*, 395–420.

Luoto, T. P. 2009: Subfossil Chironomidae (Insecta: Diptera) along a latitudinal gradient in Finland: development of a new temperature inference model. *Journal of Quaternary Science 24*, 150–158.

Luoto, T. P. 2011: The relationship between water quality and chironomid distribution in Finland – a new assemblage-based tool for assessments of long-term nutrient dynamics. *Ecological Indicators 1*, 255–262.

Medeiros, A. S., Chipman, M. L., Francis, D. R., Hamerlík, L., Langdon, P., Puleo, P. J., Schellinger, G., Steigleder, R., Walker, I. R., Woodroffe, S. & Axford, Y. 2022: A continental-scale chironomid training set for reconstructing Arctic temperatures. *Quaternary Science Reviews 294*, 107728, https://doi.org/10.1016/j.quascirev.2022.107728.

Medeiros, A. S., Gajewski, K., Porinchu, D. F., Vermaire, J. C. & Wolfe, B. B. 2015: Detecting the influence of secondary environmental gradients on chironomid-inferred paleotemperature reconstructions in northern North America. *Quaternary Science Reviews 124*, 265–274.

Nazarova, L., Herzschuh, U., Wetterich, S., Kumke, T. & Pestryakova, L. 2011: Chironomid-based inference models for estimating mean July air temperature and water depth from lakes in Yakutia, northeastern Russia. *Journal of Paleolimnology 45*, 57–71.

Nazarova, L., Syrykh, L., Grekov, I., Sapelko, T., Krasheninnikov, A. B. & Solovieva, N. 2023: Chironomid-based modern summer temperature data set and inference model for the northwest European part of Russia. *Water 15*, 976, https://doi.org/10.3390/w15050976.

Ni, Z., Zhang, E., Meng, X., Sun, W. & Ning, D. 2023: Chironomid-based reconstruction of 500-year water-level changes in Daihai Lake, northern China. *Catena 227*, 107122, https://doi.org/10.1016/j.catena.2023.107122.

Oksanen, J. and 32 others 2022: *vegan: community ecology package. R package version 2.6-4*. Available at: https://CRAN.R-project.org/package=vegan.

Orendt, C. 1999: Chironomids as bioindicators in acidified streams: a contribution to the acidity tolerance of chironomid species with a classification in sensitivity classes. *International Review of Hydrobiology 84*, 439–449.

Pegler, S., Simmatis, B., Labaj, A. L., Meyer-Jacob, C. & Smol, J. P. 2020: Long-term changes in chironomid assemblages linked to lake liming and fertilization in previously acidified middle lake (Sudbury, Canada). *Water, Air, & Soil Pollution 231*, 410, https://doi.org/10.1007/s11270-020-04780-y.

Plikk, A., Engels, S., Luoto, T. P., Nazarova, L., Salonen, J. S. & Helmens, K. F. 2019: Chironomid-based temperature reconstruction for the Eemian interglacial (MIS 5e) at Sokli, northeast Finland. *Journal of Paleolimnology 61*, 355–371.

Płóciennik, M., Mroczkowska, A., Pawłowski, D., Wieckowska-Lüth, M., Kurzawska, A., Rzodkiewicz, M., Okupny, D., Szmańda, J., Mazurkevich, A., Dolbunova, E. & Luoto, T. P. 2022: Summer temperature drives the lake ecosystem during the Late Weichselian and Holocene in Eastern Europe: a case study from East European Plain. *Catena 214*, 106206, https://doi.org/10.1016/j.catena.2022.106206.

Quinlan, R. & Smol, J. P. 2001: Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. *Journal of Paleolimnology 26*, 327–342.

Rao, Z., Tian, Y., Guang, K., Wei, S., Guo, H., Feng, Z., Zhao, L. & Li, Y. 2022: Pollen data as a temperature indicator in the late Holocene: a review of results on regional, continental and global scales. *Frontiers in Earth Science 10*, 84565, https://doi.org/10.3389/feart.2022.845650.

Rees, A. B., Cwynar, L. C. & Cranston, P. S. 2008: Midges (Chironomidae, Ceratopogonidae, Chaoboridae) as a temperature proxy: a training set from Tasmania, Australia. *Journal of Paleolimnology 40*, 1159–1178.

Ruse, L. P., Greaves, H. M., Sayer, C. D. & Axmacher, J. C. 2018: Consequences of pond management for chironomid assemblages and diversity in English farmland ponds. *Journal of Limnology 77*, 160–168, https://doi.org/10.4081/jlimnol.2018.1789.

Schmid, P. E. 1993: *A Key to the Larval Chironomidae and Their Instars from Austrian Danube Region Streams and Rivers. Part 1: Diamesinae, Prodiamesinae and Orthocladiinae*. 512 pp. Federal Institute for Water Quality, Vienna.

Šeirienė, V., Gastevičienė, N., Luoto, T. P., Gedminienė, L. & Stančikaitė, M. 2021: The Lateglacial and early Holocene climate variability and vegetation dynamics derived from chironomid and pollen records of Lieporiai palaeolake, North Lithuania. *Quaternary International 605*, 55–64.

Self, A. E., Brooks, S. J., Birks, H. J. B., Nazarova, L., Porinchu, D., Odland, A., Yang, H. & Jones, V. J. 2011: The distribution and abundance of chironomids in high-latitude Eurasian lakes with respect to temperature and continentality: development and application of new chironomid-based climate-inference models in northern Russia. *Quaternary Science Reviews 30*, 1122–1141.

Seppä, H., Bjune, A. E., Telford, R. J., Birks, H. J. B. & Veski, S. 2009: Last nine-thousand years of temperature variability in Northern Europe. *Climate of the Past 5*, 523–535.

Ter Braak, C. J. 1992: Permutation versus bootstrap significance tests in multiple regression and ANOVA. *In Bootstrapping and Related Techniques: Proceedings of an International Conference, Held in Trier, FRG, June 4–8, 1990*, 79–85. Springer, Berlin.

Ter Braak, C. J. & Juggins, S. 1993: Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *In Proceedings of the Twelfth International Diatom Symposium, Renesse, the Netherlands*, 485–502. Springer, Dordrecht.

Ter Braak, C. J. & Verdonschot, P. F. 1995: Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences 57*, 255–289.

Ursenbacher, S., Stötter, T. & Heiri, O. 2020: Chitinous aquatic invertebrate assemblages in Quaternary lake sediments as indicators of past deepwater oxygen concentration. *Quaternary Science Reviews 231*, 106203, https://doi.org/10.1016/j.quascirev.2020.106203.

Väliranta, M., Salonen, J. S., Heikkilä, M., Amon, L., Helmens, K., Klimaschewski, A., Kuhry, P., Kultti, S., Poska, A., Shala, S. & Veski, S. 2015: Plant macrofossil evidence for an early onset of the Holocene summer thermal maximum in northernmost Europe. *Nature Communications 6*, 6809, https://doi.org/10.1038/ncomms7809.

Velle, G., Brodersen, K. P., Birks, H. J. B. & Willassen, E. 2010: Midges as quantitative temperature indicator species: lessons for palaeoecology. *The Holocene 20*, 989–1002.

Verbruggen, F., Heiri, O., Meriläinen, J. J. & Lotter, A. F. 2011: Subfossil chironomid assemblages in deep, stratified European lakes: relationships with temperature, trophic state and oxygen. *Freshwater Biology 56*, 407–423.

Walker, I. R. & Mathewes, R. W. 1987: Chironomids, lake trophic status, and climate. *Quaternary Research 28*, 431–437.

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.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. & Yutani, H. 2019: Welcome to the tidyverse. *Journal of Open Source Software 4*, 1686, https://doi.org/10.21105/joss.01686.

Wickham, H., François, R., Henry, L. & Müller, K. 2022: *Dplyr: a grammar of data manipulation. R package version 1.0.10*. Available at: https://CRAN.R-project.org/package=dplyr.

Wood, S. N. 2011: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology 73*, 3–36.

Wood, S. N. 2017: *Generalized Additive Models: An Introduction with R*. 496 pp. Chapman and Hall/CRC, Boca Raton.

## Supporting Information

Additional Supporting Information to this article is available at http://www.boreas.dk.

*Fig. S1*. Distribution of environmental variables in lakes of the Eastern Baltic training set.

*Table S1*. Environmental and geographical data of Eastern Baltic lakes.

*Table S2*. Linear regression and general additive models of species with mean abundance >1% with the mean July air temperature (MJAT) as an explanatory variable.