



KOHONEN ARTIFICIAL NEURAL NETWORKS AND THE IndVal INDEX AS SUPPLEMENTARY TOOLS FOR THE QUANTITATIVE ANALYSIS OF PALAEOECOLOGICAL DATA

MATEUSZ PŁÓCIENNIK¹, ANDRZEJ KRUK², DANUTA J. MICHCZYŃSKA³ and H. JOHN B. BIRKS⁴

¹*Department of Invertebrate Zoology and Hydrobiology, Faculty of Biology and Environmental Protection, University of Łódź, 12/16 Banacha Str., 90-237 Łódź, Poland*

²*Department of Ecology and Vertebrate Zoology, Faculty of Biology and Environmental Protection, University of Łódź, 12/16 Banacha Str., 90-237 Łódź, Poland*

³*GADAM Centre of Excellence, Institute of Physics – CSE, Silesian University of Technology, Konarskiego 22B, 44-100 Gliwice, Poland*

⁴*Department of Biology, University of Bergen, PO Box 7803, N-5020 Bergen, Norway; and Environmental Change Research Centre, University College London, London, WC1E 6BT, UK*

Received 6 August 2014

Accepted 26 October 2015

Abstract: We applied two widely-used methods for data partitioning — constrained incremental sum-of-squares (CONISS) and Optimal Partitioning (OP) along with two supplementary methods, a Kohonen artificial neural network (self-organising map, SOM) and the indicator value (IndVal) index, for the quantitative analysis of subfossil chironomid assemblages from a palaeolake in Central Poland. The samples, taken from 79 core depths, were divided into 5–11 groups (five by SOM, seven by CONISS, 11 by OP), for which different numbers of indicator taxa were determined with the use of the IndVal index (18 for CONISS, 15 for SOM, 11 for OP). Only six indicator taxa were common to all three methods. The number of highly specific ($p < 0.001$) taxa was highest for SOM. Only the SOM analysis clearly reflected the rate of the changes in chironomid assemblages, which occurred rapidly in the Late Glacial (as a result of greater climate variability) and slowly in the Holocene (as a reflection of slow long-term changes in the local habitat, such as paludification). In summary, we recommend using SOM and the IndVal index in combination with CONISS and/or OP in order to detect different aspects of temporal variability in complex multivariate palaeoecological data.

Keywords: Late Quaternary, Chironomidae, self-organising map (SOM), CONISS, Optimal Partitioning.

1. INTRODUCTION

Diverse branches of palaeoecology, palaeolimnology, palaeogeography, palaeoclimatology and environmental archaeology are based on geological and/or biological data zonation, which is sometimes subjectively estab-

lished by a researcher. Arbitrary zonation is sufficient when few variables are being compared, but in palaeozoological and palaeobotanical data, when dozens of variables (taxa) must be considered, repeatable and explicit zonation requires multivariate methods.

A review of the main multivariate methods used in palaeoecology was made by Birks and Gordon (1985), Birks (1986), Agterberg and Gradstein (1988) and recently by Birks *et al.* (2012). Many data-analytical methods have been established for stratigraphical data zonation

Corresponding author: A. Kruk
e-mail: a.kruk@biol.uni.lodz.pl

(Birks, 2012), including constrained clustering (e.g. CONISS) and partitioning methods (e.g. Optimal Partitioning (OP)). They can be supported by supplementary data-analytical techniques which help to find ecological patterns among stratigraphical data. Methods commonly used for the latter purpose are canonical ordination (e.g. CCA, DCCA), K-means clustering or partitioning, TWINSpan (Legendre and Birks, 2012a) and other ordination techniques (Legendre and Birks, 2012b). Lately, artificial neural networks (ANNs) are also being more frequently used not only in ecology (Penczak *et al.*, 2004, 2009; Lek *et al.*, 2005; Chon, 2011) but also in palaeoecology. Malmgren and Nordlund (1997) recognised ANNs as a viable alternative to more conventional approaches to data analysis in palaeoenvironmental reconstruction.

ANNs have many advantages, which allow a researcher to apply them to “difficult” data. First, they are robust to noise in data (Lek and Guégan, 1999; Park *et al.*, 2006; Zhang *et al.*, 2011), which is especially important in palaeoecological studies. This is because the taxa abundances in samples usually do not reflect exactly the original abundances of populations (Ney, 1993). Additionally, the relationship between the real abundances of species and their abundances in samples is distorted over the thousands or millions of years separating the living populations and their sampling. The latter problem is manifested in the fragmentation caused by variation in the rate of sediment accumulation, its composition and subsequent compaction. Terrestrial sediments like peat are often exposed to desiccation, while older sediments are usually more compacted, which could cause decomposition of the fossil material.

Second, ANNs can be used for modelling complex phenomena, described by variables that are related both linearly and non-linearly, exhibiting normal or skewed distributions (Brosse *et al.*, 2001; Lek *et al.*, 2005). Abundances of organisms, especially those more rarely recorded in the samples, often have a skewed distribution, which precludes effective normalisation by any transformation, such as optimal Box and Cox or log-transformations (Quinn and Keough, 2002).

Among the various types of ANNs are Kohonen ANNs, also referred to as self-organising maps (SOMs; Kohonen, 1982, 2001). They learn features from the data themselves and they are used for classification purposes. Kohonen ANNs are often used for linking diverse biotic proxies to environmental parameter changes (Simpson and Birks, 2012). Remin (2008, 2012) used a Kohonen ANN to distinguish morphotypes in Cretaceous belemnites. Besides this application of a Kohonen ANN for taxonomic issues, this method has been used in varying ways in sediment stratigraphy analysis. For example, Sun *et al.* (2009) used it and a factor analysis based approach for geochemical data-pattern recognition, while Abreu and de Ribet (2002) used it for seismic facies mapping,

and Chang *et al.* (2002) used it for the identification of Jurassic lithofacies.

In this study we applied a Kohonen ANN to recognise coenological temporal patterns in the Late Quaternary sequence of subfossil chironomid assemblages deposited in the Rąbień palaeolake. We also assess the usefulness of the Kohonen ANN as a supplementary tool for the ecological interpretation of a sequence partitioned with methods typically used for zonation in palaeoecological studies. Additionally, we assess the efficiency of zonation obtained with the use of the indicator value index (IndVal) proposed by Dufrêne and Legendre (1997).

2. STUDY SITE

Rąbień bog is situated in central Poland, 11 km west of the City of Łódź (Fig. 1). The actual bog and former post-glacial palaeolake basin are located in a nearly flat landscape. The underlying geology of Rąbień bog comprises mid-Weichselian sand (Forysiak, 2012). The thickness of biogenic sediments reaches 6.2 m and core R2 was taken from the central part of Rąbień bog (Fig. 2). Gytja had accumulated up to a depth of 2.2 m, and was overlain by peat (Forysiak, 2012).

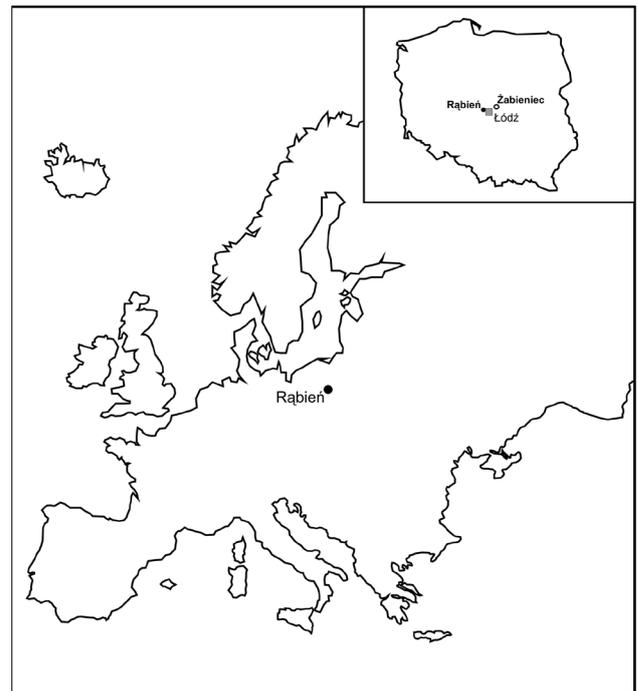


Fig. 1. Location of Rąbień bog.

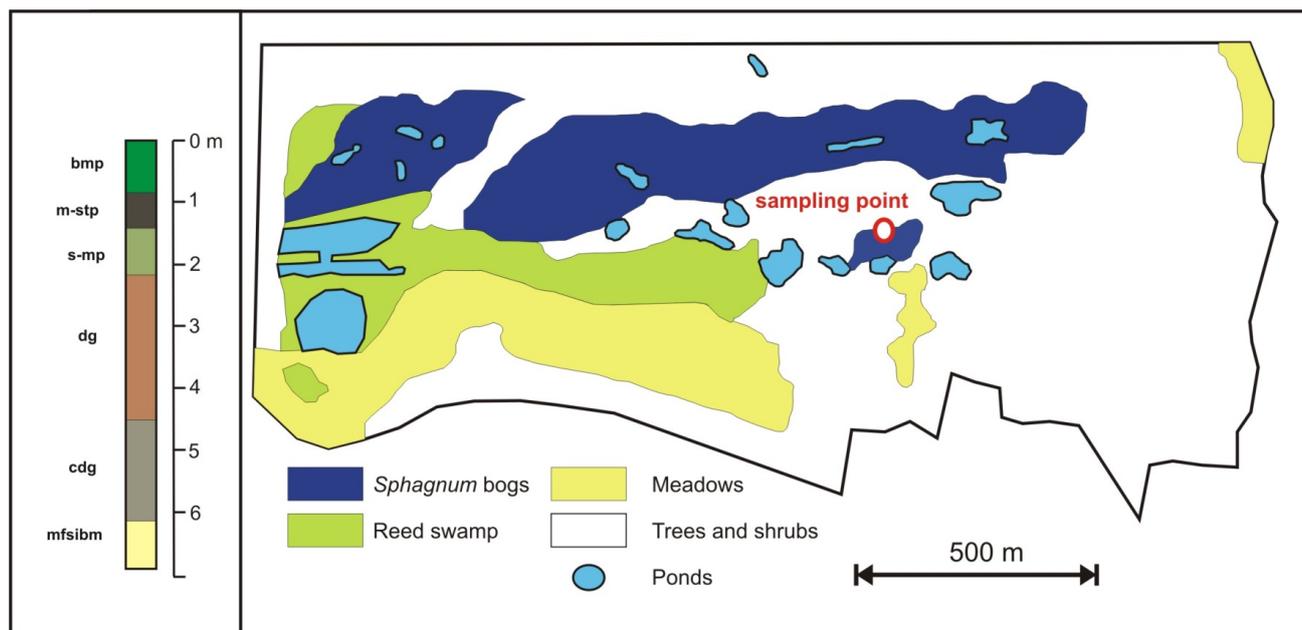


Fig. 2. Map of Rąbień bog after Kloss and Kucharski (2011) and bog lithology after Forysiak (2012): mfsibm — medium and fine-grained sand with inserts of biogenic matter, cdg — clay detrital gyttja, dg — detrital gyttja, s-mp — sedge-moss peat, m-stp — moss-sedge transition peat, bmp — bog moss peat.

3. METHODS

Field and laboratory methods

Chironomid subfossils were analysed at 10 cm (630–610 cm of core depth), 5 cm (610–430 cm) and 10 cm (430 cm to the top of the sequence) resolution. Sample volumes ranged between 3.5 cm³ and 40 cm³. Chironomid preparation methods followed Brooks *et al.* (2007). The sediments were passed through a 0.063 mm sieve. As the head capsule concentration in the upper core sequence was low, kerosene flotation was used following Rolland and Larocque (2007). Identification of chironomid head capsules followed keys by Klink and Moller Pillot (2003) and Brooks *et al.* (2007). Ecological preferences of identified taxa were based on Brooks *et al.* (2007), Moller Pillot (2009a, 2009b) and Vallenduuk and Moller Pillot (2007). The reference collection was deposited in the Department of Invertebrate Zoology and Hydrobiology, University of Łódź. Chronology of the site is based on 14 AMS radiocarbon dates and an age-depth model constructed using the P_Sequence command in OxCal programme (Bronk Ramsey, 2009) and the IntCal13 calibration curve (Reimer *et al.*, 2013). All dates cited in the paper are not results of simple calibration of radiocarbon dates, but they are estimated using the age-depth model and presented as a median and dispersion of the probability distribution of the age characteristic for each particular depth. The obtained chronology shows that the age of the palaeolake (understood as the bottom of organic sediments, gyttja) extends to the Oldest Dryas, 14820 ± 190 cal BP. The Holocene sequence (11650 ± 140 cal BP)

starts from 410 cm of the core depth. The bottom of the peat accumulation starts at 5000 ± 120 cal BP while the top of the sequence (20 cm of core depth) is dated to 2130 ± 80 cal BP (Michczyńska *et al.*, 2013, 2014, *in prep.*).

Data analysis

The dataset used for the SOM analysis consists of the abundances of 55 taxa (Table 1) sampled at 79 levels in the core. Homogeneous groups of subfossil chironomid samples were recognised using a Kohonen ANN.

ANNs are simple structural and functional models of a human brain. They are built of neurons (data-processing units) grouped into layers. Kohonen ANNs (SOMs) are built with just two (input and output) layers of neurons. The dataset is presented to the input neurons while the output neurons are responsible for data structuring and output. The input layer serves only as a flow-through layer for the input data (Lek and Guégan, 1999). Because each input neuron receives information on one taxon, the number of input neurons is equal to the number of variables in the dataset (in this study 55). The neurons in the output layer create a two-dimensional grid. We determined the final grid size of 8 × 5 neurons by applying a heuristic rule stating that the number of output neurons should be close to $5\sqrt{n}$, where n is the number of chironomid samples (compare, e.g., Vesanto and Alhoniemi, 2000).

During the SOM training, each input neuron transmits signals to all the output neurons; the intensity (weight) of a connection of each input neuron and each output neuron

Table 1. Relative abundance (A), relative frequency (F) and IndVals (I) (all in %, for five SOM sub-clusters; see Fig. 3) of all the taxa recorded in the Rabien core. IndVals highest at $p \leq 0.05$ for indicator taxa in a given SOM sub-cluster) are underlined.

SOM sub-cluster	A	F	I
<i>Lasiodiamesa</i>	1.0000		
<i>Ablabesmyia</i>	0.0051		
<i>Guttipeloplia guttipennis</i>	0.4134		
<i>Monopeloplia tenuicalcar</i>	1.0000		
<i>Procladius</i>	0.8730		
<i>Corynoneura cf. antennalis</i>	0.0251		
<i>Corynoneura edwardsi</i>	1.0000		
<i>Cricotopus type bicinctus</i>	0.6033		
<i>Cricotopus type intersectus</i>	0.1153		
<i>Cricotopus type laricomalis</i>	0.1906		
<i>Geothocladius</i>	0.1172		
<i>Limnophyes</i>	0.0794		
<i>Metriocnemus fuscipes</i>	1.0000		
<i>Orthocladius type S</i>	0.1118		
<i>Parakiefferiella type bathophila</i>	0.1172		
<i>Parametriocnemus-Paraphaenocladius</i>	0.3265		
<i>Propilocerus type N</i>	0.1145		
<i>Psectrocladius type barbatipes</i>	0.1163		
<i>Psectrocladius sordidellus</i>	0.8299		
<i>Pseudorthocladius</i>	0.1462		
<i>Pseudosmittia</i>	1.0000		
<i>Benthalia type carbonaria</i>	0.3284		
<i>Chironomus type anthracinus</i>	0.0007		
<i>Chironomus type plumosus</i>	0.0001		
<i>Cladopelma type goetghebueri</i>	0.0123		
<i>Cryptochironomus</i>	0.3373		
<i>Cryptotendipes</i>	0.4174		
<i>Demeijerea rufipes</i>	0.4195		
<i>Dicrotendipes</i>	0.0001		
<i>Endochironomus type albipennis</i>	0.0196		
<i>Glyptotendipes type barbipes</i>	0.0132		
<i>Glyptotendipes type caulicola</i>	0.4210		
<i>Glyptotendipes type imbecilis</i>	0.9929		
<i>Glyptotendipes type pallens</i>	0.0001		
<i>Lauterborniella agrayloides</i>	0.4172		
<i>Microtendipes type pedellus</i>	0.0001		
<i>Pagastiella</i>	0.1992		
<i>Polypedilum type nubeculosum</i>	0.1317		
<i>Polypedilum type sordens</i>	0.4172		
<i>Sergentia</i>	0.1118		
<i>Stictochironomus type rosenschoeldi</i>	0.0144		
<i>Stenochironomus</i>	0.0003		
<i>Synendotendipes type impar</i>	0.5040		
<i>Pseudochironomus prasinatus</i>	0.1222		
<i>Corynocera ambigua</i>	0.0041		
<i>Micropsectra type contracta</i>	0.4158		
<i>Paratanytarsus type austriacus</i>	0.4187		
<i>Paratanytarsus type penicillatus</i>	0.9979		
<i>Stempellinella-Zavrelia</i>	0.1118		
<i>Tanytarsus type glabrescens</i>	0.0001		
<i>Tanytarsus type lactescens</i>	0.0523		
<i>Tanytarsus type lugens</i>	0.1153		
<i>Tanytarsus type mendax</i>	0.0001		
<i>Tanytarsus pallidicornis type 1</i>	0.7522		
<i>Tanytarsus pallidicornis type 2</i>	0.0856		
p			
X₁			
A	100	0	0
F	4	0	0
I	4	0	0
X₂			
A	0	2	0
F	0	22	0
I	0	0	0
X₃			
A	0	53	100
F	0	67	8
I	0	<u>35</u>	8
X₄			
A	0	21	0
F	0	52	0
I	0	11	0
X₅			
A	0	24	0
F	0	67	0
I	0	16	0

becomes strengthened or weakened. On the basis of such links, a virtual chironomid sample was created in each output neuron. A similarity of any two virtual chironomid samples is reflected by their mutual positions in the output layer. Virtual chironomid samples assigned to distant neurons on the two-dimensional grid are different, while those in neighbouring neurons are similar. Additionally, the virtual chironomid samples (and thus the respective output neurons) can be clustered using hierarchical cluster analysis (Ward linkage method with Euclidean distance measure) (Ward, 1963; Vesanto and Alhoniemi, 2000). Finally, each real chironomid sample is assigned to the best matching virtual chironomid sample and the respective output neuron. Consequently, similar real chironomid samples are located nearby (in the same neuron or in adjoining neurons), and those that are considerably different are located in distant regions of the SOM (Bedoya *et al.*, 2009; Penczak, 2011; Conti *et al.*, 2012; Li *et al.*, 2013). In the above-described way, the Kohonen algorithm recognises patterns in the chironomid samples and distinguishes their groups, which may then be the subject of data interpretation (Lek *et al.*, 2005; Cheng *et al.*, 2012; Stojković *et al.*, 2013).

The training process of the SOM was performed with the use of the SOM Toolbox (Vesanto *et al.*, 2000) developed by the Laboratory of Information and Computer Science, Helsinki University of Technology (<http://www.cis.hut.fi/projects/somtoolbox>). The length of the SOM training process was 474 epochs in the rough training phase and 1659 epochs in the fine-tuning phase. Additionally, the SOM Toolbox allows for visualisation of each taxon's occurrence and relative abundance in virtual chironomid samples (and the respective output neurons), in the form of a greyness gradient. Taxa with the same pattern of greyness over the two-dimensional grid usually co-occur in similar environmental conditions.

In order to recognise the main trends in the variation of chironomid assemblages, we also applied other data-analytical methods that are more commonly used in palaeoecology: constrained incremental sum-of-squares (CONISS) and Optimal Partitioning (OP). CONISS and optimal sum-of-squares partitioning were used to determine zones in the chironomid biostratigraphy according to Bennett (1996), Birks and Gordon (1985), and Birks (1986). The zones were then tested for statistical significance with reference to the broken-stick model (MacArthur, 1957) using the programs ZONE (Lotter and Juggins, 1991), BSTICK (J.M. Line and H.J.B. Birks, unpublished) and MVSP (Kovach, 2007). CONISS was applied to \log_{10} transformed data and it was based on Ward's method with squared (sum-of-square) Euclidean distance. The stratigraphic diagram was created in C2 (Juggins, 2007).

Because SOM, CONISS, and OP do not provide any statistical verification of associations of taxa within the distinguished groups of chironomid samples (i.e. within the distinguished stages of assemblage development), we

applied the indicator value (IndVal) approach of Dufrêne and Legendre (1997) using the untransformed data. IndVal allows identification of taxa significantly associated within each (sub)cluster of samples. The value (0–100%) is a product of (1) average taxon abundance (percentage content) in chironomid samples assigned to a given (sub)cluster over the sum of its average abundances (percentage contents) in all (sub)clusters (%), (2) the frequency of the taxon (%) in chironomid samples assigned to the (sub)cluster, and (3) the constant 100 in order to produce percentages, as follows:

$$A_{ij} = \text{abundance}_{ij} / \text{abundance}_i$$

$$F_{ij} = N \text{ chironomid samples}_{ij} / N \text{ chironomid samples}_j$$

$$\text{IndVal}_{ij} = A_{ij} \times F_{ij} \times 100 \quad (3.1)$$

The IndVals (Eq. 3.1) were calculated with PC-ORD statistical software (McCune and Mefford, 2011). The significance levels of the maximum observed IndVals in each (sub)cluster of chironomid samples were also determined in PC-ORD by a Monte Carlo randomisation test. The IndVals additionally served for making a decision regarding where to stop dividing the SOM and CONISS clusters, as this process should be stopped when the number of taxa exhibiting significant IndVals starts decreasing (Dufrêne and Legendre, 1997).

4. RESULTS

SOM partitioning

The SOM quantisation and topographic errors were 0.937 and 0.025, respectively. Two main clusters were distinguished in the output layer of SOM (Fig. 3): X and Y. The division with five sub-clusters was adopted because the number of taxa exhibiting significant ($p < 0.05$) IndVals started to decrease with further partitioning (Table 2). The decrease was even more evident when only the taxa with IndVals significant at $p < 0.001$ were concerned (Table 2). In the adopted typology, cluster X contained sub-clusters X₁ (with neurons A1-A4, B1-B3, C1, C2), X₂ (C3, C4, D1-D4, E3, E4), and X₃ (A5, B4, B5, C5, D5), while cluster Y contained sub-clusters Y₁ (E1, E2, F1, F2, G1-G3, H1-H3) and Y₂ (E5, F3-F5, G4, G5, H4, H5) (Fig. 3).

Cluster X contained both the oldest and youngest samples, i.e., from 630–595 cm and 390–010 cm of core depth, respectively. Cluster Y contained samples from 570–400 cm of core depth. The unlisted samples from 590–575 cm of core depth were alternately assigned to clusters X and Y. Among sub-clusters, the most homogeneous one was X₃, which contained exclusively samples from 320–200 cm of core depth (Fig. 3). The remaining sub-clusters contained samples from different depths, i.e., the Kohonen artificial neural network considered samples from different periods to be similar. The remaining individual sub-clusters contained samples from the following depths (with a few exceptions): X₁ – 630–620 cm, 390–

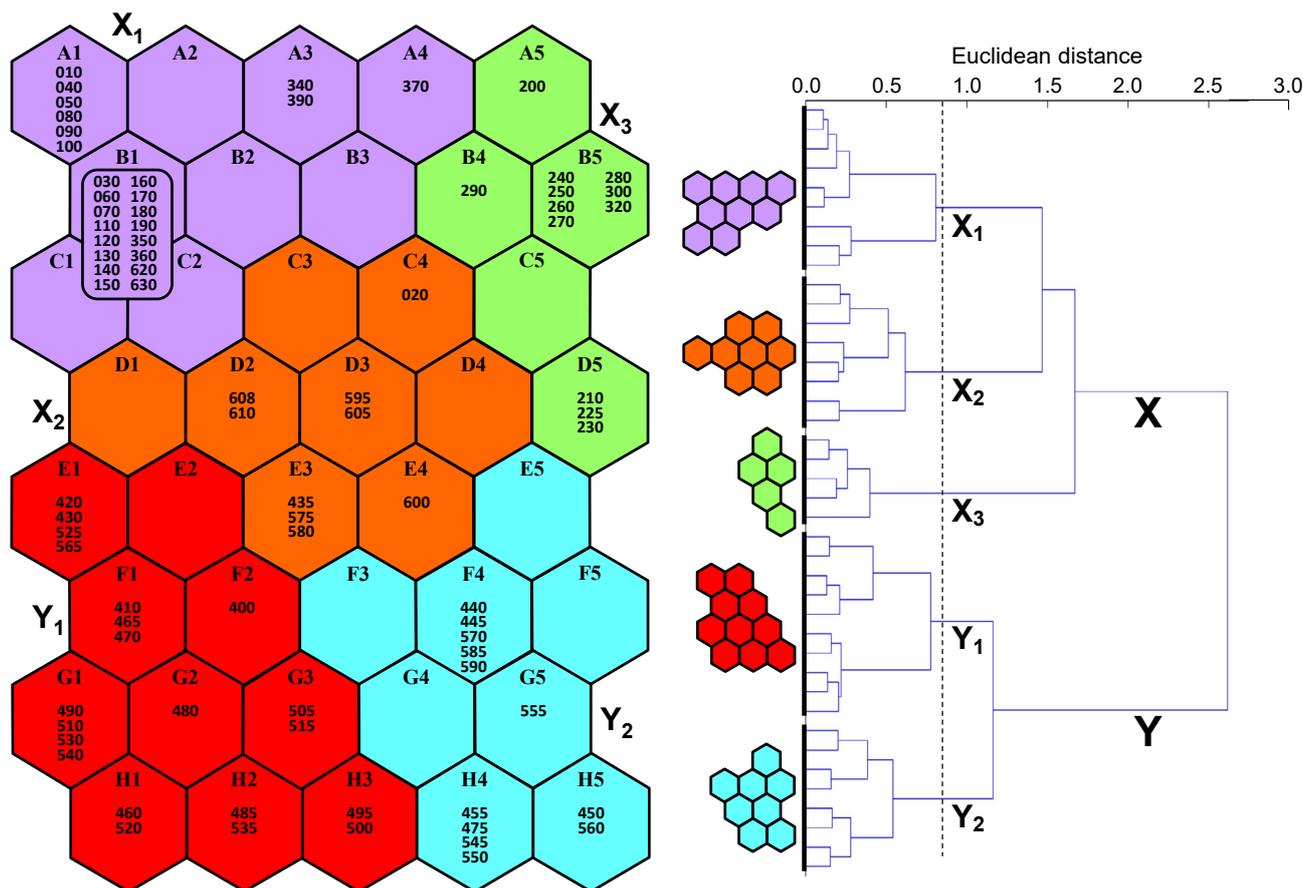


Fig. 3. Chironomid samples from 79 bog core depths (in cm) assigned to 40 self-organising map (SOM) output neurons (A1–H5) arranged into a two-dimensional lattice (8 × 5). Clusters (X, Y) and sub-clusters (X₁, X₂, X₃, Y₁ and Y₂; coloured) of neurons have been identified with the use of hierarchical cluster analysis.

340 cm and 190–010 cm of core depth, X₂ – 610–595 cm of core depth, Y₁ – 540–460 cm and 430–400 cm of core depth, and Y₂ – 570–545 cm and 455–440 cm of core depth (Fig. 3). The unlisted samples from 590–575 cm of core depth were alternately assigned to sub-clusters X₂ and Y₂. Sub-clusters X₁, Y₁ and X₃ contained samples from the largest parts of the core.

Fifteen taxa were significantly associated with certain SOM sub-clusters (Table 1 and Fig. 4). The highest numbers of such taxa were identified for sub-clusters X₂ and X₃ (five taxa in each) and the lowest numbers of taxa for X₁ and Y₂ (one taxon in each). Three taxa exhibited the significantly highest IndVals in Y₁.

Results of Optimal Partitioning and CONISS analysis

Two divisions could be adopted for CONISS, i.e., into: 1) three significant zones (CNB 1–3) indicated with the broken-stick model, or 2) seven clusters (CN 1–7) based on the maximum number of taxa exhibiting significant IndVals (Table 2 and Fig. 5). The first variant was rejected from further consideration because of a lower number of such taxa (15 vs 18), and a lack of such taxa

for CNB3 (465–010 cm, i.e., 74% of core depth) (Table 2). The OP method allowed for determination of 11 statistically significant zones (Fig. 5). Zone OP10 contained only one sample from the core depth of 50 cm, where only one head capsule (*C. cf. antennalis*) was found.

The total number of taxa significantly associated with any stage distinguished with SOM (in the final version with 5 sub-clusters), CONISS (in the final version with 7 sub-clusters) and OP is 22, but only six taxa (*Chironomus* type *anthracinus*, *Cladopelma* type *goetghebuerei*, *Dicrotendipes*, *Glyptotendipes* type *pallens*, *Corynocera* type *ambigua*, *Tanytarsus* type *glabrescens*) are common to all three methods (Table 2). The number of indicator taxa for CONISS (18) is higher than for SOM (15) and OP (11) (Table 2). However, the number of highly specific ($p < 0.001$) taxa for groups of chironomid communities distinguished with SOM and OP is similar (eight and seven taxa, respectively), and much higher than for CONISS (three taxa). The number of common indicator taxa for SOM and CONISS is 12; for SOM and OP is eight; and for CONISS and OP also eight (out of 22 in total) (Table 2). This indicates that the differences

Table 2. Indicator taxa, i.e. taxa exhibiting a significant ($p < 0.05$) *IndVal* (*l*) in any (sub)cluster (*C*) distinguished with the SOM (divisions into both 5 and 6 sub-clusters), CONISS (divisions into 3, 7 and 8 clusters) and Optimal Partitioning methods. Explanations: significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; SOM sub-cluster X_{1B} contains the following neurons of X_1 : A3, A4 and B3 (see Fig. 3).

Taxon	SOM with 5 sub-clusters (final version)			SOM with 6 sub-clusters			CONISS with 3 clusters (broken stick variant)			CONISS with 7 clusters (final version)			CONISS with 8 clusters			Optimal Partitioning (final version)		
	C	l	p	C	l	p	C	l	p	C	l	p	C	l	p	C	l	p
<i>Ablabesmyia</i>	X ₃	35	**	X ₃	35	*	CNB2	39	**	CN6	39	*						
<i>Procladius</i>				X _{1B}	88	***												
<i>Corynoneura cf. antennalis</i>	X ₁	20	*													OP8	96	***
<i>Parametricnemus-Paraphaenocladus</i>							CNB1	13	*	CN1	100	***	CN1	100	***	OP1	100	***
<i>Psectrocladius type barbatipes</i>										CN6	41	*	CN6	37	*			
<i>Pseudorthocladus</i>																OP9	97	***
<i>Benthalia type carbonaria</i>										CN6	41	*						
<i>Chironomus type anthracinus</i>	X ₂	40	***	X ₂	40	*	CNB1	46	***	CN2	73	**	CN2	73	**	OP2	42	*
<i>Chironomus type plumosus</i>	X ₂	61	***	X ₂	61	**	CNB1	41	**	CN2	36	*	CN2	36	*			
<i>Cladopelma type goetghebueri</i>	Y ₁	33	*	Y ₁	33	*	CNB2	60	***	CN4	39	*				OP3	42	*
<i>Dicrotendipes</i>	X ₃	69	***	X ₃	69	***				CN6	61	***	CN6	54	**	OP6	71	***
<i>Endochironomus type albipennis</i>	Y ₁	24	*				CNB2	31	***									
<i>Glyptotendipes type barbipes</i>	X ₂	27	*				CNB1	22	*	CN2	32	*	CN2	32	*			
<i>Glyptotendipes type pallens</i>	X ₃	75	***	X ₃	47	*				CN6	37	*				OP6	58	*
<i>Microtendipes type pedellus</i>	X ₃	68	***	X ₃	68	***										OP6	68	***
<i>Pagastiella</i>							CNB2	14	*									
<i>Polypedilum type nubeculosum</i>							CNB1	38	**									
<i>Stictochironomus type rosenschoeldi</i>	X ₂	22	*	X ₂	22	*	CNB1	13	*	CN2	40	*	CN2	40	*			
<i>Stenochironomus</i>	X ₃	42	***	X ₃	42	**				CN6	83	**	CN6	83	**			
<i>Pseudochironomus prasinatus</i>				X _{1B}	57	**												
<i>Corynocera ambigua</i>	X ₂	34	**	X ₂	34	*	CNB1	45	**	CN3	36	*	CN3	30	*	OP4	42	***
<i>Paratanytarsus type penicillatus</i>							CNB1	27	***	CN1	48	*	CN1	48	*	OP1	49	**
<i>Tanytarsus type glabrescens</i>	Y ₁	72	***	Y ₁	72	***	CNB2	91	***	CN4	82	***	CN4	69	***	OP3	63	***
<i>Tanytarsus type lactescens</i>							CNB1	13	*	CN3	25	*	CN3	25	*			
<i>Tanytarsus type mendax</i>	Y ₂	60	***	Y ₂	60	**				CN3	38	*						
<i>Tanytarsus pallidicornis type 2</i>							CNB1	27	***	CN3	50	**	CN3	50	**			
Number of indicator taxa at $p < 0.001$		8			4			6			3			2			7	
Total number of indicator taxa ($p < 0.05$)		15			14			15			18			13			11	

between the three methods in indicator taxa were at least moderate. The highest number of different indicator taxa recorded for CONISS compared to OP is ten, for SOM compared to OP it is seven, and for CONISS compared to SOM it is six. The lowest number of different indicator taxa (three) was recorded in SOM compared to CONISS, and in OP compared to SOM and to CONISS (Table 2).

5. DISCUSSION

Ecological interpretation

The SOM analysis distinguishes five stages of chironomid assemblages development: X_1 , X_2 , X_3 , Y_1 and Y_2 (Figs. 3, 5 and 6). The stages are understood here as chironomid subfossil assemblages reflecting specific ecological states of the lake habitat occurring in neighbouring or distant periods. This is why they cannot be treated as zones defined as stratigraphically unique intervals. Such intervals can be distinguished with OP and/or CONISS methods, usually applied to derive a Quaternary sequence zonation. The SOM analysis is complementary to these methods.

Stages Y_1 , Y_2 and X_2 fall mainly into the Late Glacial core section (14.7–10.6 kyr cal BP). Stage X_2 aggregates mainly samples from 14.7–14.2 kyr cal BP, Y_2 — 14.0–13.8 kyr cal BP, and 12.8–12.45 kyr cal BP, while Y_1 ranges between 14.1–10.6 kyr cal BP. Stage X_3 falls into the mid-Holocene (7.75–4.5 kyr cal BP) and X_1 mainly into the early-Holocene (10.6–7.75 kyr cal BP) and late-Holocene section (4.5–2.1 kyr cal BP) (Figs. 5, 6).

Stage X_2 is characterised by the significantly highest abundance and/or occurrence of five Chironominae species: *Chironomus type plumosus*, *Chironomus type anthracinus*, *Corynocera ambigua*, *Glyptotendipes type barbipes* and *Stictochironomus type rosenschoeldi* (Fig. 4). *C. type anthracinus*, *C. ambigua* and *S. type rosenschoeldi* are often recorded from Late Glacial sediments and were found in the Bølling-Allerød Interstadial and Younger Dryas sediments in the nearby Żabieniec palaeolake (Płóciennik *et al.*, 2011) during its cold, oligotrophic phase. *C. ambigua* and *S. type rosenschoeldi* are considered as oligostenotherm taxa (Brooks *et al.*, 2007). *C. type anthracinus* is related to intermediate climatic conditions and summer temperatures rather below 16°C (Heiri *et al.*, 2011). Both *C. type anthracinus* and *C. type*

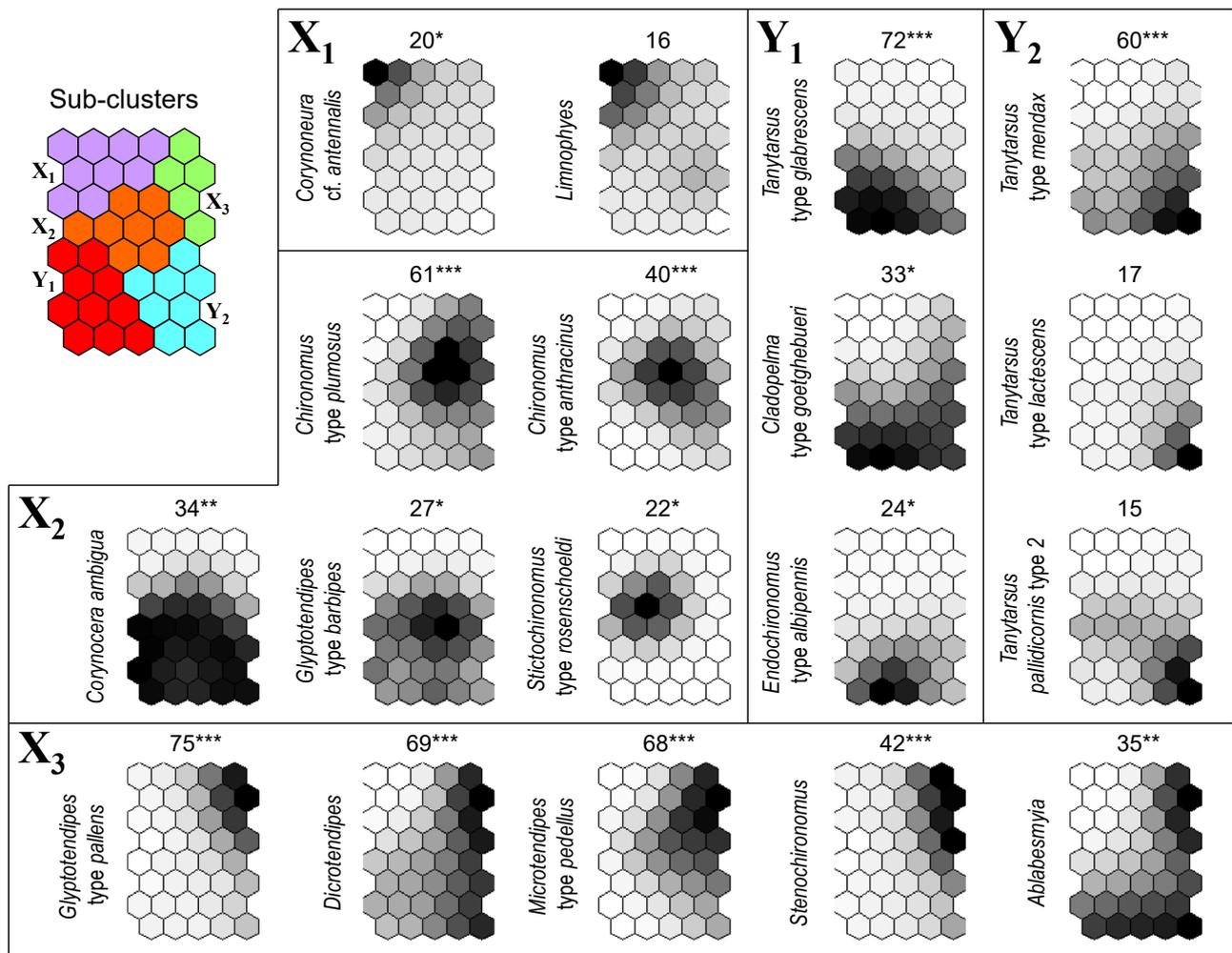


Fig. 4. Eighteen taxa associated with any self-organising map (SOM) sub-cluster at $p < 0.1$. The strength of associations with SOM regions is visualised with shading (darker for stronger associations), which is scaled independently for each taxon. Taxa with similar patterns of greyness over SOM usually have similar environmental requirements. The highest indicator value (IndVal) calculated for a given taxon and the respective significance level (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$) are shown above each taxon plane.

plumosus are often reported from lakes of higher trophy, being found also in a profundal zone and anoxic conditions, but they include many *Chironomus* species of diverse ecological preferences. The appearance of other oligostenotherm, Late Glacial taxa such as *Sergentia* and *T. type lugens* only in this stage, is remarkable (Table 1).

The typical taxa for stage Y₂ are *Tanytarsus* morphotypes: the indicator *Tanytarsus* type *mendax*, accompanied by *Tanytarsus* type *lactescens* and *Tanytarsus pallidicornis* type 2 (Fig. 4). These three morphotypes, especially *T. type mendax*, aggregate many taxa that are difficult to separate in the subfossil material. Most of the *Tanytarsus* species live in relatively warm, productive lakes (Brooks *et al.*, 2007). Gilka (2011) classifies most of the European *Tanytarsus*, including *T. mendax*, *T. lactescens* and *T. pallidicornis*, as specific to temperate zones of intermediate climatic conditions. He showed that *T. pallidicornis* mostly occurs in flowing waters, but is also found in stagnant waters. In the nearby Żabieniec

palaeolake, *T. pallidicornis* type 2 sensu Brooks *et al.* (2007) was abundant from the Younger Dryas to the late Holocene.

Stage Y₁ falls mainly into the mid-Interstadial and second stage of the Younger Dryas. The indicator species for Y₁ are *Tanytarsus* type *glabrescens*, *Cladopelma* type *goetghebueri* and *Endochironomus* type *albipennis* (Fig. 4). *T. type glabrescens* and *C. ambigua* are the most abundant species of the Late Glacial core section (Fig. 6). The share of *C. ambigua* is highest in X₂, and in Y₁ it is accompanied by *T. type glabrescens*. Gilka (2011) classifies *T. type glabrescens* as occurring mainly in temperate climates but appearing also in warmer regions. In the Swiss–Norwegian training set (Heiri *et al.*, 2011) *Cladopelma* and *Endochironomus* type *albipennis* appear in warmer temperate regions, the latter species mainly between 13°C and 18°C. Brooks *et al.* (2007) mention *C. type goetghebueri* as a polystenotherm. Both species dwell on the bottom and on plants in meso/eutrophic

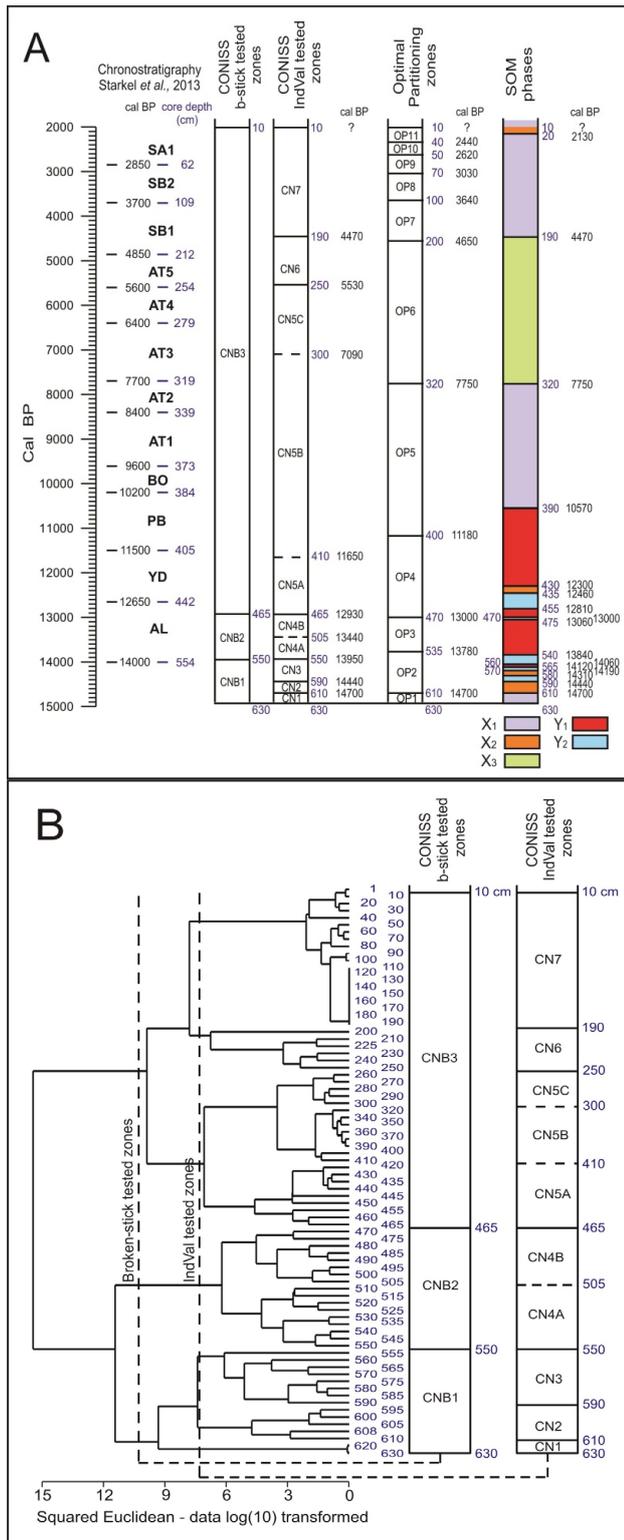


Fig. 5. Comparison of core zones determined with constrained incremental sum-of-squares (CONISS) and Optimal Partitioning (OP) with the stages of chironomid assemblages development distinguished with the self-organising map (SOM) method with age stratigraphy after Michczyńska (2013, 2014 and *in prep.*) (A). Dendrogram used for the CONISS analysis (B).

conditions and they often graze on algae (Moller Pillot, 2009a, 2009b). *E. type albipennis* is commonly found mining macrophyte tissues (Moller Pillot, 2009a).

For stage X₃ there are five indicator morphotypes: *Glyptotendipes type pallens*, *Dicrotendipes*, *Microtendipes type pedellus*, *Stenochironomus* and *Ablabesmyia* (Fig. 4). All these taxa are associated with a warm or intermediate temperate climate and meso/eutrophic waters with ample vegetation (Szczerkowska-Majchrzak *et al.*, 2010; Głowacki *et al.*, 2011). Larvae of *Glyptotendipes type pallens* are often found to be macrophyte miners. *Stenochironomus* mine living and dead plant stems including wood. *Ablabesmyia* and *Dicrotendipes* are also often found on aquatic plants (Brooks *et al.*, 2007; Valleduuk and Moller Pillot, 2007; Moller Pillot, 2009b). *Microtendipes type pedellus* sensu Brooks *et al.* (2007) in fact aggregates mostly *M. chloris* agg. species sensu Moller Pillot (2009a, 2009b). While *M. pedellus* agg. sensu Moller Pillot (2009a, 2009b) inhabits nearly exclusively small streams, *M. chloris* agg. is abundant in stagnant waters. Both aggregations fit the *M. type pedellus* sensu Brooks *et al.* (2007) and cannot be separated in subfossil material. Young larvae from the *M. chloris* agg. are often found on plants and filamentous algae, and as they grow they move to the bottom. Representatives of *M. pedellus* agg. are more restricted to the bottom (Moller Pillot, 2009a, 2009b).

SOM stage X₁ is composed of samples with very low subfossil amounts. The stages in the Rąbień palaeolake and bog development, from which there are only singular head capsules of diverse taxa, are as follows: initial stage of palaeolake development of 630–620 cm, early Holocene of 390–340 cm and late Holocene of 190–10 cm (Fig. 5). There are no characteristic species for the Late Glacial and early Holocene phases of X₁, whereas the late Holocene phase of X₁ is characterised by the dominance of Orthoclaadiinae, mainly *Corynoneura cf. antennalis* with co-occurrences of *Limnophyes*, *Pseudorthocladius* and *Parametriocnemus-Paraphaenocladius*. The last three taxa occur mainly in the late X₁ phase (Table 1 and Fig. 6) but, mostly because of their low stability of occurrence, their associations with X₁ are not significant. *Corynoneura cf. antennalis*, *Limnophyes* and *Pseudorthocladius* are recorded in the final stage of the Żabieniec bog (Płóciennik *et al.*, 2011). Similar to the 190–10 cm of Rąbień core section, it is a stage of semi-terrestrial conditions of mire and raised-bog habitat. *Corynoneura cf. antennalis* appears also in lowland streams and spring marshes (Moller Pillot, 2009b). *Parametriocnemus-Paraphaenocladius* contains many terrestrial species (Brooks *et al.*, 2007; Moller Pillot, 2009b) and is recorded in central Poland from the Late Glacial/Holocene fen sediments with groundwater supply (Dzieduszyńska *et al.*, 2014; Płóciennik *et al.* 2015). *Limnophyes* and *Pseudorthocladius* are typically terrestrial taxa often found on mosses (Brooks *et al.*, 2007; Moller Pillot, 2009b).

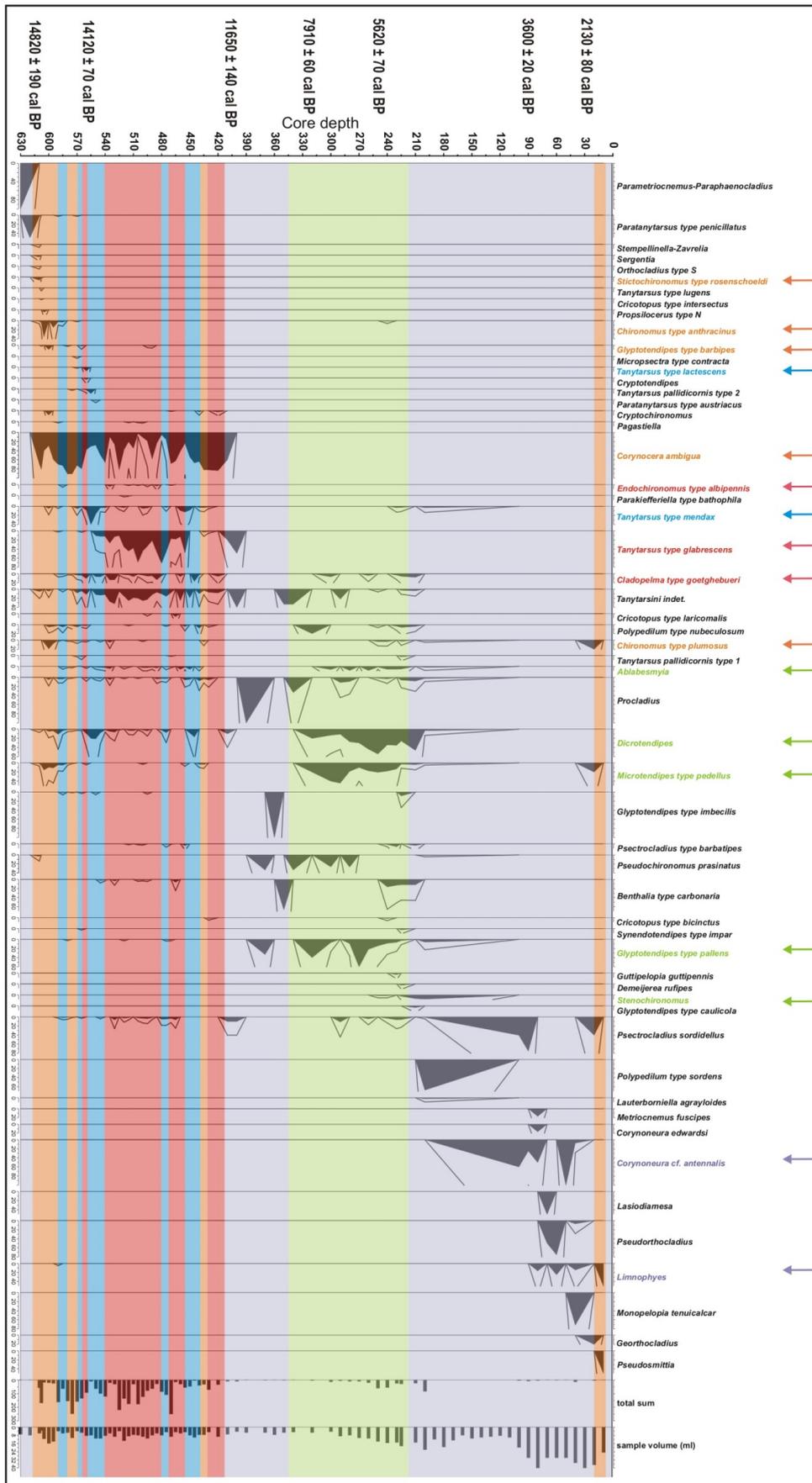


Fig. 6. Percentage diagram of the main chironomid taxa in the Rabień core vs the self-organising map (SOM) stages (compare with Figs. 4 and 5). Colours reflect the SOM sub-clusters. Modelled ages are presented next to a few chosen core depths.

Chronological reconstruction

The age-depth model (Michczyńska *et al.*, 2013, 2014, *in prep.*) and palynological data (Obremska, unpublished data) indicate that the initial phase of the Rąbień palaeolake (630–620 cm of core depth, X₁) represents the cold climatic and ultra-oligotrophic conditions of the Oldest Dryas. In the succeeding, younger deposits (610–590 and 580–570 cm, X₂) the chironomid assemblages and palynomorphs exhibit preferences for the warmer, but still cool, climate of the early Bølling Interstadial. According to the age–depth model, the samples 590–585, 570 and 560–545 cm (Y₂) of the core represent the Older Dryas and short intermediate oscillations before it. The indicator taxa for Y₂ are typical of intermediate climatic conditions or they may represent cooling in the warmer stages of the Bølling/Allerød Interstadial. The chironomid assemblage from 565 cm (Y₁) corresponds to a short warm oscillation before the Older Dryas, and at 540–460 cm it is associated with the Allerød Interstadial. The chironomid assemblages at 475 cm are assigned to Y₂ and may be linked with the Gerzensee oscillation (Michczyńska *et al.*, 2013, 2014, *in prep.*; Obremska, unpublished data). Chironomid assemblages at 455–440 cm (Y₂) and at 435 cm (X₂) fall into the first stage of the Younger Dryas. Interestingly, the youngest samples of Y₁ (430–400 cm) fall into the second Younger Dryas phase and the onset of the Holocene. On the one hand, this may suggest that the second phase of the Younger Dryas after 12.3 kyr cal BP could have been associated with warmer conditions. On the other hand, this is a time of high dominance of *C. ambigua*, a boreal species, which would indicate cold climatic conditions (Fig. 6). The early Holocene (from Preboreal to Early Atlantic periods), attributable to the period of 10.6–7.75 kyr cal BP (X₁), was a time of recession in the chironomid communities, when their abundance fell to singular specimens in the studied samples. During the mid Holocene (Atlantic Period), i.e., 7.74–4.5 kyr cal BP (X₃), Rąbień was a shallow palaeolake of higher trophic, overgrown by macrophytes. Finally, in the Subboreal–Early Subatlantic Periods (4.5–2.1 kyr cal BP; X₁) Rąbień palaeolake became a mire with semi-terrestrial conditions. The short episode of ca. 2.25–2.1 kyr cal BP (X₂), may be a reflection of more humid conditions with a shallow pool on the bog surface.

Comparison of the methods

The classical methods of dataset partitioning allowed the sequence to be divided into clearly defined chronologically constrained zones, unlike SOM. The latter method showed a few main stages which are not chronologically constrained. They alternately partition the sequence into longer or shorter phases and thus the SOM analysis clearly reflects the rate of the changes in chironomid assemblages, which occurred much faster in the Late Glacial (as a result of higher climate variability) and

more slowly in the Holocene (as a reflection of the local paludification) (Fig. 5). This is an advantage not available with classical partitioning methods, which do not show that certain stages in the composition of assemblages may be repeated. In the well-defined partitioning of the Late Glacial section, the phases distinguished by SOM correspond well to the chronological and palynological zonation (Michczyńska *et al.*, 2013, 2014, *in prep.*).

The IndVal index provides a statistical verification of taxa associations with the clusters of samples distinguished with different methods, which considerably supports the ecological interpretation of the resulting core divisions. Moreover, it is an objective tool, based on unambiguous criteria, that supports decisions regarding where to stop dividing the clusters distinguished using SOM and CONISS. The division obtained with the broken-stick model seemed to be less satisfactory because of the lower number of indicator taxa and the lack of such taxa for almost $\frac{3}{4}$ (465–010 cm) of the studied core.

Despite the fact that the lists of indicator taxa for SOM, CONISS, and OP are quite different, they result in comparable ecological interpretations. The important difference between SOM and the zonation methods is the indicator taxa for the lowest and upper core sections. The assemblage-based stratigraphy (Fig. 6) reveals *Paratanytarsus-Paraphaenocladus* and *Pseudorthocladus* to be strongly associated, respectively, with the first (CN1, OP1) and last (OP9) stages of the lake development. These taxa are typical of semi-terrestrial conditions, and describe well the habitat character of the initial and final phases at the Rąbień palaeolake. The other indicator morphotypes for CONISS and OP, but not for SOM, (e.g. *Paratanytarsus* type *penicillatus*, *Tanytarsus* type *lactescens* or *Tanytarsus pallidicornis* type 2) are more eurytopic and play a less important role in ecological interpretation. Interestingly, some of these are also associated with certain SOM sub-clusters at significance levels slightly higher than 0.05 (Table 1 and Fig. 4). The number of highly specific ($p < 0.001$) taxa is highest for SOM. Of these, *Chironomus* type *plumosus* clearly separates a wet episode at 20 cm of core depth from the late-Holocene sequence represented by semi-terrestrial taxa. The increase in the abundance of *Tanytarsus* type *mendax* at 455 cm indicates intermediate conditions at the onset of the Younger Dryas. Finally, *Microtendipes* type *pedellus* and *Stenochironomus* clearly indicate ample macrophyte vegetation in the lake during X₃ stage. Only *Dicrotendipes* and *Tanytarsus* type *glebrescens* are associated at $p < 0.001$ with groups of core samples distinguished by all three methods. *Tanytarsus* type *glebrescens* is one of the most important species for climate reconstruction in the Late Glacial. In contrast to the also abundant *Coryncera ambigua*, it represents warm climatic conditions and it allows a researcher to separate interstadial from cool stadial phases. *Dicrotendipes* represents a group of taxa that are typical of the warm, eutrophic and shallow mid-Holocene Rąbień lake.

6. CONCLUSIONS

In summary, the Kohonen artificial neural network and the IndVal index appear to be valuable supplementary methods to aid the ecological explanation of sequence partitioning in palaeoecological analyses. SOM revealed the pace of changes and the main trends in the palaeolake history, which, combined with a well-developed visualisation of the intensity of biotic variables in the SOM output layer, provides an opportunity for a more detailed interpretation of the results obtained with other methods. The comparison of SOM, CONISS, and OP clearly shows that they are complementary. Consequently, we recommend using SOM in combination with CONISS and/or OP in order to reveal different aspects of temporal variability in palaeoecological data. Moreover, the IndVal index considerably extends the potential of analyses with the use of any of the three methods.

ACKNOWLEDGEMENTS

The manuscript has greatly benefited from constructive comments from journal reviewers including Dr Zbigniew Remin. The authors are also grateful to Ewa Janowska for valuable assistance in laboratory analysis of the samples, and to Dominik Pawłowski for computing the CONISS zonation. The study was carried out under the projects financed from the funds of the Polish Ministry of Science and Higher Education No. N N306 034040 and No. N N306 276735.

REFERENCES

- Abreu CE and de Ribet B, 2002. Trace shape and multi-attribute seismic facies analysis applied to Paleocene/Eocene reservoirs on deep-water Campos Basin. *Revista Brasileira de Geofísica* 20: 89–95.
- Agterberg FP and Gradstein FM, 1988. Recent developments in quantitative stratigraphy. *Earth-Science Reviews* 25: 1–73, DOI 10.1016/0012-8252(88)90099-2.
- Bedoya D, Novotny V and Manolagos ES, 2009. Instream and off-stream environmental conditions and stream biotic integrity: importance of scale and site similarities for learning and prediction. *Ecological Modelling* 220: 2393–2406, DOI 10.1016/j.ecolmodel.2009.06.017.
- Bennett KD, 1996. Determination of the number of zones in a biostratigraphical sequence. *New Phytologist* 132: 155–170, DOI 10.1111/j.1469-8137.1996.tb04521.x.
- Birks HJB, 1986. Numerical zonation, comparison and correlation of Quaternary pollen-stratigraphical data. In: Berglund BE, ed., *Handbook of Holocene Palaeoecology and Palaeohydrology*. Chichester, New York, J. Wiley & Sons Ltd.: 743–774.
- Birks HJB, 2012. Analysis of stratigraphical data. In: Birks HJB, Lotter AF, Juggins S and Smol JP, eds., *Tracking Environmental Change Using Lake Sediments. Volume 5: Data Handling and Numerical Techniques*. Dordrecht, Springer: 355–378.
- Birks HJB and Gordon AD, 1985. *Numerical Methods in Quaternary Pollen Analysis*. London, Academic Press: 317 pp.
- Birks HJB, Lotter AF, Juggins S and Smol JP (eds), 2012. *Tracking Environmental Change Using Lake Sediments. Data Handling and Numerical Techniques*. Dordrecht, Springer: 745 pp.
- Bronk Ramsey C, 2009. Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1): 337–360.
- Brooks SJ, Langdon PG and Heiri O, 2007. The Identification and Use of Palaearctic Chironomidae Larvae in Palaeoecology. *QRA Technical Guide* No. 10. London, Quaternary Research Association: 276 pp.
- Brosse S, Girardel JL and Lek S, 2001. Utilisation of non-supervised neural networks and principal component analysis to study fish assemblages. *Ecological Modelling* 146: 159–166, DOI 10.1016/S0304-3800(01)00303-9.
- Chang H-C, Kopaska-Merkel DC and Chen H-C, 2002. Identification of lithofacies using Kohonen self-organizing maps. *Computers and Geosciences* 28: 223–229, DOI 10.1016/S0098-3004(01)00067-X.
- Cheng L, Lek S, Lek-Ang S and Li Z, 2012. Predicting fish assemblages and diversity in shallow lakes in the Yangtze River basin. *Limnologia* 42: 127–136, DOI 10.1016/j.limno.2011.09.007.
- Chon TS, 2011. Self-Organizing Maps applied to ecological sciences. *Ecological Informatics* 6: 50–61, DOI 10.1016/j.ecoinf.2010.11.002.
- Conti L, Grenouillet G, Lek S and Scardi M, 2012. Long-term changes and recurrent patterns in fisheries landings from Large Marine Ecosystems (1950–2004). *Fisheries Research* 119–120: 1–12, DOI 10.1016/j.fishres.2011.12.002.
- Dufrène M and Legendre P, 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs* 67: 345–366, DOI 10.1890/0012-9615(1997)067[0345:SAAI]2.0.CO;2.
- Dzieduszyńska DA, Kittel P, Petera-Zganiak J, Brooks SJ, Korzeń K, Krąpiec M, Pawłowski D, Płaza DK, Plóciennik M, Stachowicz-Rybka R and Twardy J, 2014. Environmental influence on forest development and decline in the Warta River valley (Central Poland) during the Late Weichselian. *Quaternary International* 324: 99–114, DOI 10.1016/j.quaint.2013.07.017.
- Forysiak J, 2012. Record of changes in the natural environment of the Late Weichselian and Holocene preserved in the sediments of peatlands of the Łódź region. *Acta Geographica Lodziana* 99: 1–164 (in Polish).
- Giłka W, 2011. Analysis of faunistic diversity in chironomids of the tribe Tanytarsini in Europe (Diptera: Chironomidae). *Dipteron* 27: 11–31 (in Polish).
- Głowacki Ł, Grzybkowska M, Dukowska M and Penczak T, 2011. Effects of damming a large lowland river on chironomids and fish assessed with the (multiplicative partitioning of) true/Hill biodiversity measure. *River Research and Applications* 27: 612–629, DOI 10.1002/rra.1380.
- Heiri O, Brooks SJ, Birks HJB and Lotter AF, 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, DOI 10.1016/j.quascirev.2011.09.006.
- Juggins S, 2007. *C2 Version 1.5 User Guide. Software for Ecological and Palaeoecological Data Analysis and Visualisation*. Newcastle University: 73 pp.
- Klink AG and Moller Pillot HKM, 2003. *Chironomidae Larvae. Key to Higher Taxa and Species of the Lowlands of Northwestern Europe*. Amsterdam, ETI. CDROM.
- Kloss M and Kucharski L, 2011. History of vegetation of the "Rąbień Bog" reserve based on interdisciplinary research. In: Zieliński A, ed., *Interdisciplinary researches in natural sciences*. Kielce, Institute of Geography, Jan Kochanowski University: 47–58.
- Kohonen T, 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 43: 59–69, DOI 10.1007/BF00337288.
- Kohonen T, 2001. *Self-Organizing Maps*. Third Extended Edition. Berlin, Springer: 501 pp.
- Kovach WL, 2007. *MVSP-A Multivariate Statistical Package for Windows, ver. 3.1*. Pentraeth, Kovach Computing Services.
- Legendre P and Birks HJB, 2012a. Clustering and partitioning. In: Birks HJB, Lotter AF, Juggins S and Smol JP, eds., *Tracking Environmental Change Using Lake Sediments. Volume 5: Data Handling and Numerical Techniques*. Dordrecht, Springer: 167–200.
- Legendre P and Birks HJB, 2012b. From classical to canonical ordination. In: Birks HJB, Lotter AF, Juggins S and Smol JP, eds., *Tracking Environmental Change Using Lake Sediments. Volume 5: Data Handling and Numerical Techniques*. Dordrecht, Springer: 201–248.

- Lek S and Guégan JF, 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling* 120: 65–73, DOI 10.1016/S0304-3800(99)00092-7.
- Lek S, Scardi M, Verdonshot PFM, Descy JP and Park YS, 2005. *Modelling Community Structure in Freshwater Ecosystems*. Berlin, Springer: 518 pp.
- Li F, Bae MJ, Kwon YS, Chung N, Hwang SJ, Park SJ, Park HK, Kong DS and Park YS, 2013. Ecological exergy as an indicator of land-use impacts on functional guilds in river ecosystems. *Ecological Modelling* 252: 53–62, DOI 10.1016/j.ecolmodel.2012.09.006.
- Lotter AF and Juggins S, 1991. POLPROF, TRAN and ZONE. Programs for Plotting, Editing and Zoning of Pollen and Diatom Data. *INQUA Commission for the study of Holocene. Working Group on Data-Handling Methods, Newsletter* 6.
- MacArthur RH, 1957. On the relative abundance of bird species. *Proceedings of the National Academy of Sciences USA* 43: 293–295.
- Malmgren BA and Nordlund U, 1997. Application of artificial neural networks to paleoceanographic data. *Palaeogeography, Palaeoclimatology, Palaeoecology* 136: 359–373, DOI 10.1016/S0031-0182(97)00031-X.
- McCune B and Mefford MS, 2011. *PcOrd Multivariate Analysis of Ecological Data, Version 6.06*. Gleneden Beach, Oregon, MjM Software Design.
- Michczyńska DJ, Forsytek J, Pawłowski D, Płóciennik M, Borówka RK, Witkowski A, Obremaska M, Słowiński M, Żurek S, Brooks SJ and Michczyński A, 2013. *The environment changes and chronology of the Late Vistulian (Weichselian) sediments in the Rąbień mire*. In: Piotrowska N, ed., 11th International Conference "Methods of Absolute Chronology", 15–18th May 2013, Podlesice, Poland. Abstracts and Programme: 79.
- Michczyńska DJ, Borówka RK, Okupny D, Obremaska M, Forsytek J, Pawłowski D, Płóciennik M, Słowiński M, Żurek S, Brooks SJ, Michczyński A and Witkowski A, 2014. *The environment changes and chronology of the Late Vistulian (Weichselian) and Early Holocene sediments in the Rąbień mire, Central Poland*. In: Abstract Volume and Programme, INTIMATE Open Workshop and COST Action ES0907 Final Event, Zaragoza, 15th–21st June 2014: 53.
- Michczyńska DJ, Forsytek J, Borówka RK, Brooks SJ, Luoto TP, Michczyński A, Nevalainen L, Obremaska M, Okupny D, Pawłowski D, Payron O, Płóciennik M, Self A, Słowiński M, Witkowski A, Żurek S, in preparation. Response of bio- and geochemical proxies to Late Glacial and early-Holocene climatic fluctuations recorded in Rąbień paleolake sediment (Poland).
- Moller Pillot HKM, 2009a. *Chironomidae Larvae. Biology and Ecology of the Chironomint*. Zeist, KNNV Publishing: 270 pp.
- Moller Pillot HKM, 2009b. *A Key to the Larvae of the Aquatic Chironomidae of the North-west European Lowlands*. Private print, not published: 77 pp.
- Ney JJ, 1993. Practical use of biological statistics. In: Kohler CC and Hubert WA, eds., *Inland Fisheries Management in North America*. Bethesda, MD, American Fisheries Society: 137–158.
- Park YS, Tison J, Lek S, Giraudel JL, Coste M and Delmas F, 2006. Application of a self-organizing map to select representative species in multivariate analysis: A case study determining diatom distribution patterns across France. *Ecological Informatics* 1: 247–257, DOI 10.1016/j.ecoinf.2006.03.005.
- Penczak T, 2011. Usefulness of the SOM algorithm for estimation of species distribution and significance in comparing habitats. *Journal of Applied Ichthyology* 27: 1371–1374, DOI 10.1111/j.1439-0426.2011.01867.x.
- Penczak T, Lek S, Godinho F and Agostinho AA, 2004. Patterns of fish assemblages in tropical streamlets using SOM algorithm and conventional statistical methods. *Ecology and Hydrobiology* 4: 139–146.
- Penczak T, Agostinho AA, Gomes LC and Latini JD, 2009. Impacts of a reservoir on fish assemblages of small tributaries of the Corumbá River, Brazil. *River Research and Applications* 25: 1013–1024, DOI 10.1002/rra.1200.
- Płóciennik M, Self A, Birks HJB and Brooks SJ, 2011. Chironomidae (Insecta: Diptera) succession in Żabieniec bog and its palaeo-lake (Central Poland) through the Late Weichselian and Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 307: 150–167, DOI 10.1016/j.palaeo.2011.05.010.
- Płóciennik M, Kruk A, Forsytek J, Pawłowski D, Mianowicz K, Elias S, Borówka RK, Kloss M, Obremaska M, Coope R, Krąpiec M, Kittel P and Żurek S, 2015. Fen ecosystem responses to water-level fluctuations during the early and middle Holocene in central Europe: a case study from Wilczków, Poland. *Boreas* 44: 721–740, DOI 10.1111/bor.12129.
- Quinn GP and Keough MJ, 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge, Cambridge University Press: 537 pp.
- Reimer PJ, Bard E, Bayliss A, Beck JW, Blackwell PG, Bronk Ramsey C, Buck CE, Cheng H, Edwards RL, Friedrich M, Grootes PM, Guilderson TP, Haflidason H, Hajdas I, Hatté C, Heaton TJ, Hoffmann DL, Hogg AG, Hughen KA, Kaiser KF, Kromer B, Manning SW, Niu M, Reimer RW, Richards DA, Scott EM, Southon JR, Staff RA, Turney CSM and van der Plicht J, 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* 55(4): 1869–1887.
- Remin Z, 2008. Artificial Kohonen neural networks as a tool in paleontological taxonomy – An introduction and application to Late Cretaceous belemnites. *Przegląd Geologiczny* 56: 58–66 (in Polish with English abstract).
- Remin Z, 2012. The *Belemnella* stratigraphy of the Campanian–Maastrichtian boundary; a new methodological and taxonomic approach. *Acta Geologica Polonica* 62: 495–533.
- Rolland N and Larocque I, 2007. The efficiency of kerosene flotation for extraction of chironomid head capsules from lake sediments samples. *Journal of Paleolimnology* 37: 565–572, DOI 10.1007/s10933-006-9037-2.
- Simpson GL and Birks HJB, 2012. Statistical learning in palaeolimnology. In: Birks HJB, Lotter AF, Juggins S and Smol JP, eds., *Tracking Environmental Change Using Lake Sediments. Volume 5: Data Handling and Numerical Techniques*. Dordrecht, Springer: 249–327.
- Starkel L, Michczyńska D, Krąpiec M, Margielewski W, Nalepka D and Pazdur A, 2013. Progress in the holocene chronostratigraphy of Polish territory. *Geochronometria* 40(1): 1–21, DOI 10.2478/s13386-012-0024-2.
- Stojković M, Simić V, Milošević D, Mancev D and Penczak T, 2013. Visualization of fish community distribution patterns using the self-organizing map: A case study of the Great Morava River system (Serbia). *Ecological Modelling* 248: 20–29, DOI 10.1016/j.ecolmodel.2012.09.014.
- Sun X, Deng J, Gong Q, Wang Q, Yang L and Zhao Z, 2009. Kohonen neural network and factor analysis based approach to geochemical data pattern recognition. *Journal of Geochemical Exploration* 103: 6–16, DOI 10.1016/j.gexplo.2009.04.002.
- Szczerkowska-Majchrzak E, Grzybkowska M and Dukowska M, 2010. Effect of flow fluctuations on patch dynamics and chironomid distribution in a medium-sized lowland river. *Journal of Freshwater Ecology* 25: 437–448, DOI 10.1080/02705060.2010.9664387.
- Vallenduuk HJ and Moller Pillot HKM, 2007. *Chironomidae larvae of the Netherlands and adjacent lowlands. General ecology and Tanyptodinae*. Zeist, KNNV Publishing: 143 pp.
- Vesanto J and Alhoniemi E, 2000. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks* 11: 586–600, DOI 10.1109/72.846731.
- Vesanto J, Himberg J, Alhoniemi E and Parhankangas J, 2000. *SOM Toolbox for Matlab 5. Report A57*. Helsinki, Helsinki University of Technology. WEB site: <<http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf>>. Accessed 2015 March 5.
- Ward Jr JH, 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–244.
- Zhang Q, Zhang JS, Zhang B, Cheng J and Tian S, 2011. Self-organizing feature map classification and ordination of *Larix principis-rupprechtii* forest in Pangquangou Nature Reserve. *Acta Ecologica Sinica* 31: 2990–2998 (in Chinese with English summary).