



Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See

Norbert Magyar^a, István Gábor Hatvani^{a,*}, Ilona Kovácsné Székely^b, Alois Herzig^c, Mária Dinka^d, József Kovács^a

^a Eötvös Loránd University, Department of Physical and Applied Geology, Pázmány Péter stny. 1/C, H-1117 Budapest, Hungary

^b Budapest Business School, Institute of Methodology, Alkotmány u. 9–11, H-1054 Budapest, Hungary

^c Biological Station Neusiedler See, A-7142 Illmitz, Austria

^d Danube Research Institute Centre for Ecological Research of the Hungarian Academy of Sciences, Alkotmány u. 2–4, H-2163 Vácrátót, Hungary

ARTICLE INFO

Article history:

Received 16 October 2012

Received in revised form 29 January 2013

Accepted 1 February 2013

Keywords:

Cluster analysis

Neusiedler See

Principal component analysis

Spatial variability

Water quality

ABSTRACT

The aim of the present study was to distinguish areas with different chemical properties in Neusiedler See, to determine which background processes are responsible for this pattern, and to discover their spatial distribution. Uni- and multivariate data analysis was applied to the data concerning 13 mainly chemical and some biological parameters for the time period 2000–2009 from 33 sampling sites. The sampling sites were first clustered then grouped. Besides reed belt and open water areas, smaller localities, which are influenced by water inputs (the treatment plant, the river Wulka, the channels of weekend houses) were also distinguished. Using Wilks' lambda distribution it was determined that the main components (ions) have a greater effect on forming the cluster groups than those parameters which stand in close relation to biological processes. These results concurred with those obtained from the principal component analysis (PCA) conducted on the whole lake and on the groups as well. It can be stated that most of the variance in the dataset can be explained by the main components (ions). The spatial distribution of the principal component scores was visualized with isoline maps. The results of this research lead us to the view that Neusiedler See cannot be treated as one homogeneous system. This exceptional variability originates from the lake's shallow water depth, its unstable water balance, and anthropogenic activity (agriculture, tourism, sewage treatment) in the lake's vicinity.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The quality of surface water is a very sensitive issue, which transcends national boundaries (Yidana et al., 2008). One of the best examples is Lake Fertő/Neusiedler See, which lies on the Hungarian–Austrian border. Nowadays with both countries being member states of the European Union, not only does the possibility exist for them to assess the true state of the whole lake, but they are also obliged to do so (WFD, 2000/60/EC). It is recognized that water quality monitoring is one of the highest priorities in environmental protection policy (Simeonov et al., 2003), because to determine

the quality of any body of water its physical, chemical and biological parameters must be measured systematically (Sargaonkar and Deshpande, 2003). Precise and effective monitoring is, however, not enough, if the acquired dataset is not analyzed thoroughly.

In the case of the area studied in this research, because of its fragmented habitat structure and unstable water balance (Dobesch and Neuwirth, 1979b), the datasets deriving from it are complex, comprising mostly chemical and biological parameters which are often difficult to interpret and draw meaningful conclusions from. The application of different multivariate statistical methods, such as cluster analysis (CA), discriminant analysis (DA) and principal component analysis (PCA) helps in the interpretation of such complex data matrices as are found in the case of Neusiedler See in order to obtain a better understanding of water quality and the ecological state of the study area (Zhou et al., 2007). Parallel to the fragmented habitat structure the study area is daily exposed to severe anthropogenic activity, mostly on the Austrian side, in the areas used for recreation. Anthropogenic influences such as recreational use, the influence of agriculture, and changes in precipitation inputs degrade surface waters, which, because of their

* Corresponding author at: Eötvös Loránd University, Faculty of Science, Department of Physical and Applied Geology, Pázmány Péter stny. 1/C, H-1117 Budapest, Hungary. Tel.: +36 70317 97 58; fax: +36 1 31 91738.

E-mail addresses: magyarnorbert87@gmail.com (N. Magyar), hatvaniig@gmail.com (I.G. Hatvani), kovacsne.szekelyilona@kvifk.bgf.hu (I.K. Székely), alois.herzig@univie.ac.at (A. Herzig), dinka.maria@okologia.mta.hu (M. Dinka), kevesolt@geology.elte.hu (J. Kovács).

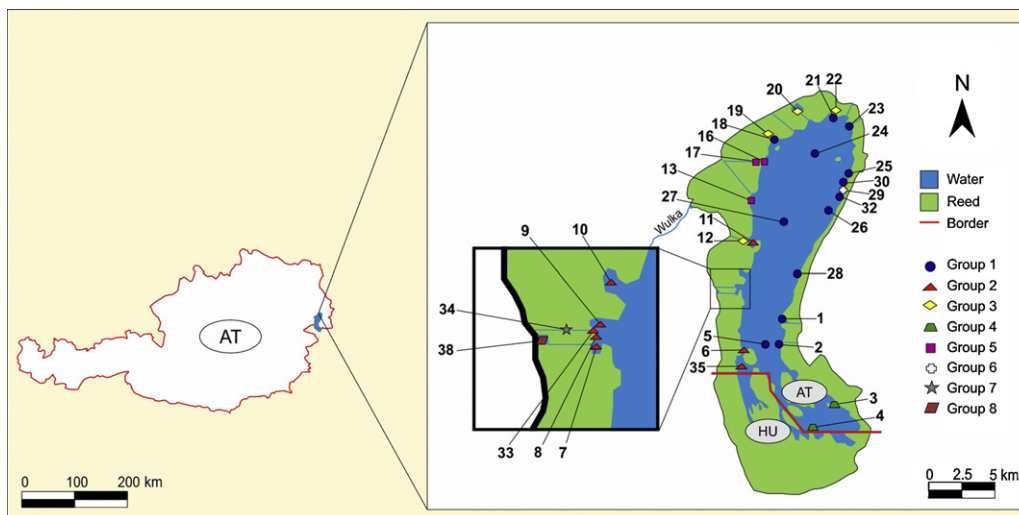


Fig. 1. Location of Neusiedler See, its sampling sites and their similarities based on the cluster results.

constantly changing vicinity, need to be monitored even more accurately. Furthermore, the area is under severe ecological control, mostly to ensure its current status.

The aim of the present study was to (1) assess Neusiedler See's (its Austrian part's) heterogeneity, i.e. distinguish areas with different chemical properties within the lake, (2) determine which parameters might give rise to this pattern, (3) point out which background variables determine the variance of the data and, what is their spatial distribution, using explorative multivariate data analysis methods.

Neusiedler See is the largest steppe lake in Europe (309 km², 244 km² of which lies in Austria; Fig. 1) (Herzig and Dokulil, 2001; Löffler, 1974). It is situated in the lowest plain of the country, at 116 m above Adriatic sea level (maA). It is a shallow (average 1.1 m), well-mixed lake with regulated outflow, and large water level fluctuations (Dokulil and Herzig, 2009). In terms of its water balance the Neusiedler See is considered to be somewhere between astatic and semiastatic (Padisák, 2005). Over the centuries the lake has dried out many times, and its water level has varied in a wide interval, sometimes resulting in enormous floods. On the one hand, as in the case of Central Europe's largest shallow freshwater lake, Balaton (3–5 m water level fluctuation up to the late 19th century), it was necessary to control the water level of the Neusiedler See as well, mainly because of succession (increase in reed coverage). To prevent this, the following measures were taken: (i) New dam management was instituted in 1965 on the Hanság Channel (outlet of the lake) and (ii) increased water level. Since 1965 the growth of the reed coverage has been quite minimal, whereas if measures had not been taken, most of the open water would probably by now be covered (Dobesch and Neuwirth, 1979b). On the other hand there is a requirement from the Austrian side to keep the summer water level at 115.8 maA and the winter water level at 115.7 maA so that the weekend houses on the Austrian side would be protected from floods. In both cases it can be said that these water management measures have exercised considerable influence on the current form of the reed belt, which – at the moment – covers about 56% of the lake. On the western and southern bank of the lake it may ever reach 3–6 km in width, however the eastern bank has more than one "reed free" areas (Fig. 1).

The reed (*Phragmites australis* (Cav.) Trin. ex. Steudel) has a unique role in the Neusiedler See ecosystem. It moderates water waves, it is the area for deposition, thereby decreasing shore

erosion, creates diverse ecological conditions, and influences water quality (biotekton). Besides its ecological functions it serves as an important industrial plant; for this reason the authorities harvest it regularly. At the Neusiedler See this harvest is carried out once a year in winter. This decreases organic succession as it facilitates nutrient removal. Winter harvesting also avoids disturbing the nesting summer visitor birdlife as well. There are, however many studies which deal with the optimal time for harvesting reeds and take into consideration many more aspects, such as nutrient removal (Hu et al., 2010), nutrient (phosphorus) retention (Kiedrzyńska et al., 2008) and P & N removal along with the vitalization of plants (Zhu et al., 2011). Naturally there are further aspects to consider, e.g. trace element removal in the case of multiple harvests in one year (Vymazal et al., 2010). Regarding the biology, ecology, growth and function in water quality melioration many studies have been published: (Kvêt, 1973; Tanner, 1996; Hu et al., 2010).

The waters of the lake can be described as salty (sodium and magnesium salts). Neusiedler See is Europe's western-most and largest steppe lake, and is considered to be a wetland (Löffler, 1974). Despite the growth in awareness of their importance as ecosystems, these wetlands are often endangered ecosystems. Their importance lies in their biological productivity, hydrological and biogeochemical cycling, and as habitats both for humans and wildlife (Mitsch and Gosselink, 1986; Dinka et al., 2010). The lake is under the protection of the following: UNESCO protection as a biosphere reserve (1977) (Löffler, 1979; Dinka et al., 2004, 2010), the Ramsar Convention; it is a part of the Fertő-Hanság National Park (1991), the Nationalpark Neusiedlersee-Seewinkel (1994) and is a part of the IUCN's¹ protection zone; furthermore, it is designated as a World Heritage site (2001).

Two main habitats can be identified in the lake, the reed belt (174 km²; 110 km² on the Austrian, 64 km² on the Hungarian part of the lake) and the open water area. A third habitat can be distinguished as well, which is the man-made channels in the reeds. The water balance of the lake is unstable, mostly because its main water input comes from precipitation and its main output is evapotranspiration. Besides precipitation, the Wulka and Rákos streams contribute to the lake's water input. The Wulka reaches the lake

¹ International Union for Conservation of Nature and Natural Resources.

Table 1
Descriptive statistics of the parameters used (2000–2009).

Statistic/parameter	pH	DO	K ⁺	Na ⁺	Ca ²⁺	Mg ²⁺	Cl [−]	SO ₄ ^{2−}	CO ₃ ^{2−}	HCO ₃ [−]	TP	NH ₄ -N	NO ₃ -N	chl-a
M	8.68	8.36	37.44	343.55	36.4	140.95	244.01	466.14	30.79	607.86	82.11	36.19	201.67	9.56
MED	8.7	8.8	38	347	34	141	245	473	30	600	68	28	152.5	7.1
SD	0.18	1.79	5.56	55.95	12.1	16.59	42.1	66.48	13.12	76.9	62.28	32.98	151.35	9.62
MIN	7.7	0.2	13	126	14	72	94	148	<LOQ	375	3	2	<LOQ	<LOQ
MAX	9.1	11.2	55	570	119	202	422	661	91	1061	552	333	872	95.3

between Purbach and Oggau (AT), while the Rákös reaches it at Fertőrákös (HU) (Gattinger, 1979). At the mouth of the Rákös, a small wetland has been constructed to decrease the loads arriving to the lake. The air is almost constantly in motion; because of orographic effects the prevailing wind direction is NW, which is followed (according to frequency) by SE (Dobesch and Neuwirth, 1979a).

It is clear that although the lake is a natural entity over recent decades, because of anthropogenic activity and “urbanization”, many measures have had to be taken and continuous management is necessary to preserve its present state.

2. Materials and methods

2.1. Data and sampling sites

The database used was acquired from the Neusiedler See Biological Station in Illmitz. This station is part of the national Water Inspectorate (AT) which monitors the lake at 38 sampling sites, both in the reed belt and on the open water. It is worth mentioning that the samples from the reed belt were taken from channels which cut through dense reed stands. These reed channels form a transition zone (ecotone) between the reed and open water habitats.

The sampling was conducted monthly, though in certain cases, weekly. The sampled parameters which suited the conditions of the methods applied (parameters have to be measured at all sampling sites at all investigated time points) are as follows: pH, dissolved O₂ (DO), K⁺, Na⁺, Ca²⁺, Mg²⁺, Cl[−], SO₄^{2−}, HCO₃[−] (mg l^{−1}), total phosphorus (TP), NH₄-N, NO₃-N, chlorophyll-a (chl-a) (μg l^{−1}). The descriptive statistics of the parameters can be seen in Table 1.

CO₃^{2−} was not used because its values are highly dependent on pH, as below pH 8.3 there is likely to be no carbonate in the water (Daniel, 1999), and at certain sampling sites was indeed below 8.3. If the zero values had been used the correlation coefficients between the parameters would have changed at those sampling sites. Because the same laboratory methods were employed in gathering the data, the results were comparable and comprehensive. The analyses were conducted for the time interval 2000–2009 because before the turn of the century the parameters in question were not measured frequently enough for these statistical methods to be applied. During the winter months and occasionally even in March most of the lake is covered with ice, so at many monitoring sites no sampling was conducted during the cold period, as a result these dates have been left out of the analyses. Specifically, 34 time points from 33 sampling sites were examined, altogether 14,586 data. The detailed description of the sampling sites can be found in A.1.

2.2. Piper diagrams

“Piper diagrams were used (Piper, 1944; see later Fig. 2) to display the relative concentrations of the major cations and anions on two separate trilinear plots, together with a central diamond plot where the points from the two trilinear plots are projected.

The central diamond-shaped field (quadrilateral field) is used to show the overall chemical character of the water (Hill, 1940; Piper, 1944).” (Güler et al., 2002).

2.3. Cluster analysis

Cluster analysis (CA) is a multivariate method which groups variables or cases based on their similarity. The most common type is hierarchical CA, where each case starts in a separate cluster and joins up to the other clusters as the linkage distance grows, and only one cluster remains (Lattin et al., 2003; McKenna, 2003). This is visualized on a dendrogram. CA was conducted on 34 time points and 13 parameters using Ward’s method (Ward, 1963), and squared Euclidean distance. A similar approach has been successfully applied to the assessment of water quality by Wunderlin et al. (2001), Simeonov et al. (2003), Singh et al. (2004), Shrestha and Kazama (2007), Zhou et al. (2007), Filik Iscen et al. (2008) and Juahir et al. (2011).

2.4. Discriminant analysis

To verify the accuracy of the cluster results, discriminant analysis (DA) was used, for the reason that the values of most of the parameters were not normally distributed. The analysis was carried out in order to show to what extent the planes separating the groups could be distinguished by building a predictive model for group membership. The model is composed of a discriminant function (for more than two groups a set of discriminant functions) based on linear combinations of the predictor variables that provide the best discrimination between the groups (Johnson and Wichern, 1992). The result of the discriminant analysis is visualized on the surface stretched between the first two discriminating planes (function 1 and function 2) (Ketskeméty and Izsó, 2005).

2.5. Wilks’ lambda distribution

After the verification of the cluster groups, the role of each parameter in determining the formation of the cluster groups was analyzed. Using Wilks’ λ distribution (Wilks, 1932) a Wilks’ λ quotient was assigned to every parameter for each sampling, where the quotient is:

$$\text{Wilks' } \lambda = \frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)^2}{\sum_i \sum_j (x_{ij} - \bar{x})^2}$$

where x_{ij} is the j th element of the i th group, $\bar{x}_i$ the i th group’s mean, and $\bar{x}$ the total mean.

The value of λ is the ratio of the within-group sum of squares to the total sum of squares. It is a number between 0 and 1. The smaller the quotient is, the more it determines the formation of the cluster groups (Hatvani et al., 2011).

2.6. Principal component analysis (PCA)

The final method was PCA, which was used to determine the reduced number of latent variables (principal components) which

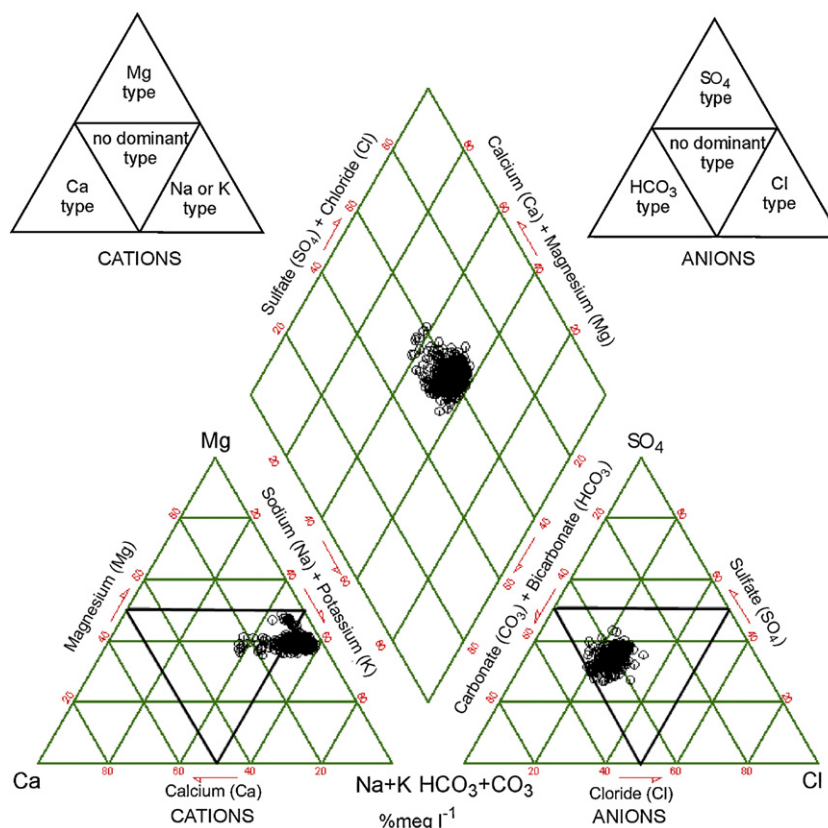


Fig. 2. Piper diagram for anion and cation facies derived from major ion percentages.

explain most of the variance in the experimental dataset in the groups obtained by cluster analysis, and for the whole lake as well. In other words, it is a powerful pattern recognition technique that attempts to explain the variance of a large dataset of intercorrelated variables with a smaller set of independent variables principal components (Simeonov et al., 2003). The reason why PCA was applied instead of factor analysis (FA) is that, while the PCA uses a correlation matrix, the FA uses an adjusted correlation matrix. In this correlation matrix the elements on the diagonal (commonalities) can be estimated in different ways, and as a result different solutions (of the FA) may be acceptable. The principal component (PC) can be expressed as

$$Z_{ij} = a_{i1}X_{1j} + a_{i2}X_{2j} + \dots + a_{im}X_{mj}$$

where z is the component score, a is the component loading, x is the measured value of the variable, i is the component number, j is the sample number, and m is the total number of variables (Juahir et al., 2011). A similar approach has been successfully applied to the assessment of water quality by Simeonov et al. (2003), Shrestha and Kazama (2007), Yidana et al. (2008), Filik Iscen et al. (2008), Kazi et al. (2009) and Hatvani et al. (2012). For PCA it is vital that there is enough data to work with, otherwise the matrix will become ill-conditioned.

2.7. Data representation techniques

As discussed before, the Piper diagram in general represents the whole system; however information concerning the inter-group differences is also needed. The most common means of representing these differences is the box-and-whiskers plot. More than one statistic can be visualized on such a plot, where the boxes show the interquartile intervals and the black lines in the boxes the median.

Two vertical lines can be drawn from the boxes if the database requires them. The data between 1.5 and 3 times the interquartile range are indicated with a circle (outliers), and the ones with higher values than 3 times the interquartile range are considered to be extreme values indicated with an asterisk (Norušis and SPSS, 1993; Vega et al., 1998).

2.8. Software used and data handling

All maps in the study were prepared using Golden Software Surfer 9 and kriging as a linear estimation method. Interpolation was applied between the sampling sites, while beyond the sampling sites (near the shore of the lake) extrapolation was applied. The statistical evaluations were made using IBM SPSS Statistics 19. Using cluster analysis the data were z-scale standardized.

3. Results and discussion

At first all the measured values for the ions were presented on Piper diagrams (Fig. 2) (including CO_3^{2-}),² in order to determine which facies the lake belongs to, based on the dominant ions. However, in the case of cations no dominant type could be determined in about 40% of the samples, while in about 60% of the samples only the sodium type could be identified. In the case of anions, again no dominant type was determined. This is based on the works of Morgan and Winner (1962) and Back (1966), quoted in Freeze and Cherry (1979).

It is clear that no groups can be identified on the basis of the Piper diagram, because the samples were located close to each other,

² A Piper diagram cannot be produced without carbonate.

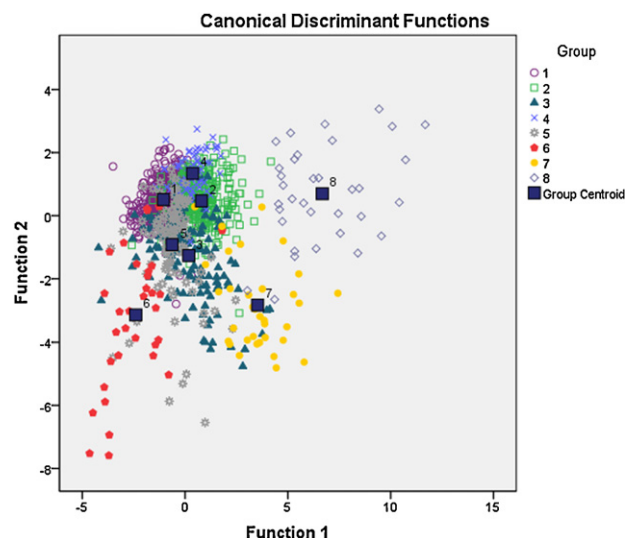


Fig. 3. Cross-validated discriminant results.

unlike in the study by Güler et al. (2002). Despite the lack of groups in the Piper diagram, the lake is known to be highly variable (Dinka et al., 2004) so it was expedient to use multivariate methods for further investigation.

In order to determine the spatial variations in water quality clustering was applied to the 34 time points, eight representative groups were identified. Hereinafter these are considered to be relatively similar water quality areas. Their sampling sites and spatial locations can be seen in Fig. 1.

Along with the three large areas, two smaller could be distinguished in the NW and SE regions, and three sampling sites (29, 34, 38) formed separate “groups” in themselves. The characteristics of the groups will only be interpreted after their statistics have been discussed.

Discriminant analysis allowed us to see that the original groups were correctly identified in 71% of the cases (Fig. 3).

As the next step in assessing the different water quality areas a Wilks' lambda quotient was calculated for each of the 34 time points, for every parameter examined, and their averages were determined. In increasing order, the smallest quotient (0.2–0.3) belongs to the main components (ions); these are followed by pH (0.38), dissolved oxygen (0.42), while even higher values follow: TP (0.53), $\text{NO}_3\text{-N}$ (0.53), $\text{NH}_4\text{-N}$ (0.53) and chlorophyll-a (0.62). The next step was to cluster the obtained Wilks' lambda quotients ($34 \times$ the number of parameters).

It is clear that the parameters can be distributed into two distinct groups according to their Wilks' lambda quotients on the dendrogram (Fig. 4). The main components (anions and cations) with pH and dissolved oxygen (average 0.29) influenced the formation of the cluster groups the most, despite the fact that on the Piper diagram (Fig. 2) no groups were distinguishable. Nutrients and chl-a, however, had a smaller influence (average 0.55).

Because Neusiedler See is a sodic lake, the input waters, which are low in ions, may produce great differences between the areas. This may be the reason why ions have a much more determining role in forming the cluster groups than the other parameters. Aside from the ions, pH and dissolved oxygen were among the more determining parameters. Their concentrations show big differences in the two main habitats of the lake (reed belt, open water). Another explanation behind the nutrients' small determining role is the retaining function of the reeds in relation to them. The water being filtered by the biotekton decreases in runoff in the concentration of its parameters, while by facilitating the sedimentation

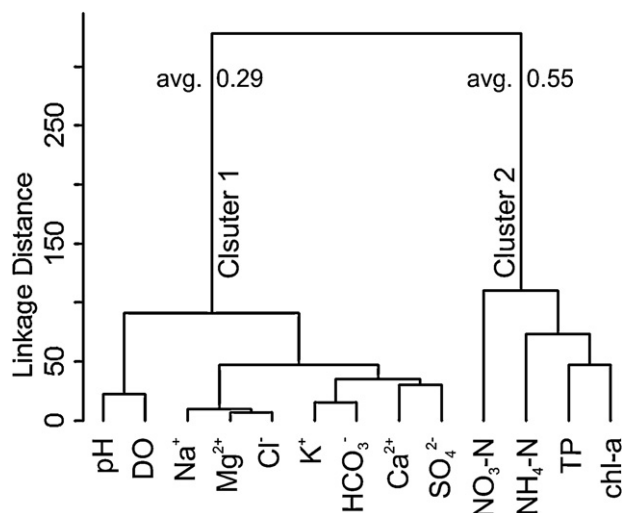


Fig. 4. Dendrogram of Wilks' lambda quotients (2000–2009).

of suspended solids increases in opacity. Using its aquatic roots, reed takes up nutrients directly from the water. There have been many studies published which elaborate on the melioration effect of reeds on natural waters, mostly by decreasing its nutrient content (Dinka, 1986; Raghi-Atri and Bornkamm, 1979; Köhl and Köhl, 1992). The previously discussed filtrating effect of reed is exploited in constructed wetlands (Brix, 1987; Vymazal et al., 2010; Tanner, 1996; Lu et al., 2009).

The groups obtained from the 34 time points were compared using boxplots. The median of pH (Fig. 5a) in Groups 1 and 4 is higher than in the others, the smallest values can be observed in Groups 7 (Station³ 34) and 8 (Station 38). In Group 6 (Station 29) the interquartile interval of pH was the widest; this means that this is the location where the values varied between the widest boundaries. As in the case of pH the concentrations of dissolved oxygen (Fig. 5b) were higher in the eastern part of the lake (Groups 1 and 4) than elsewhere. The lowest values were observable in Group 7 (Station 34).

The ions, except for calcium, behave similarly, so sodium was chosen to represent these ions (Fig. 5c). The smallest concentrations were observable in Groups 5 and 6, while the highest values were in Groups 4, 7 and 8. The median of total phosphorus (Fig. 5d) is higher in Groups 5, 6, and 8 than in the others. In Groups 1, 2 and 4, which cover most of the eastern and central part of the lake, 45 extreme values were recorded, 27 of them in April and May 2004. The chlorophyll-a concentrations in Groups 1, 2, 4 and 6 (the middle and eastern parts of the lake) are lower than in the western, Groups 3, 5 and 7. In Group 8 (Station 38) values clearly higher than the average were registered on many occasions (Fig. 5e).

Differences became obvious between the groups based on coefficients of variation (CV). These CVs were determined for the whole lake and the distinct groups. The groups' CVs were compared to each other and the whole lake as well (Fig. 6). In the following, the CVs of a few exceptional parameters will be discussed. The CV of pH is the highest in Group 6, while its variability is still considerably high in Groups 3, 5, 7 and 8, and is the smallest in Groups 1, 2 and 4. The CV of dissolved oxygen is similar to that of the pH, which is lower in Groups 1, 2 and 4; on the contrary the CV of Group 7 (Station 34) is three times the average of the whole lake. The CVs of the ions are highest in Groups 5 and 6. In the

³ In Austria sampling sites are officially named Station.

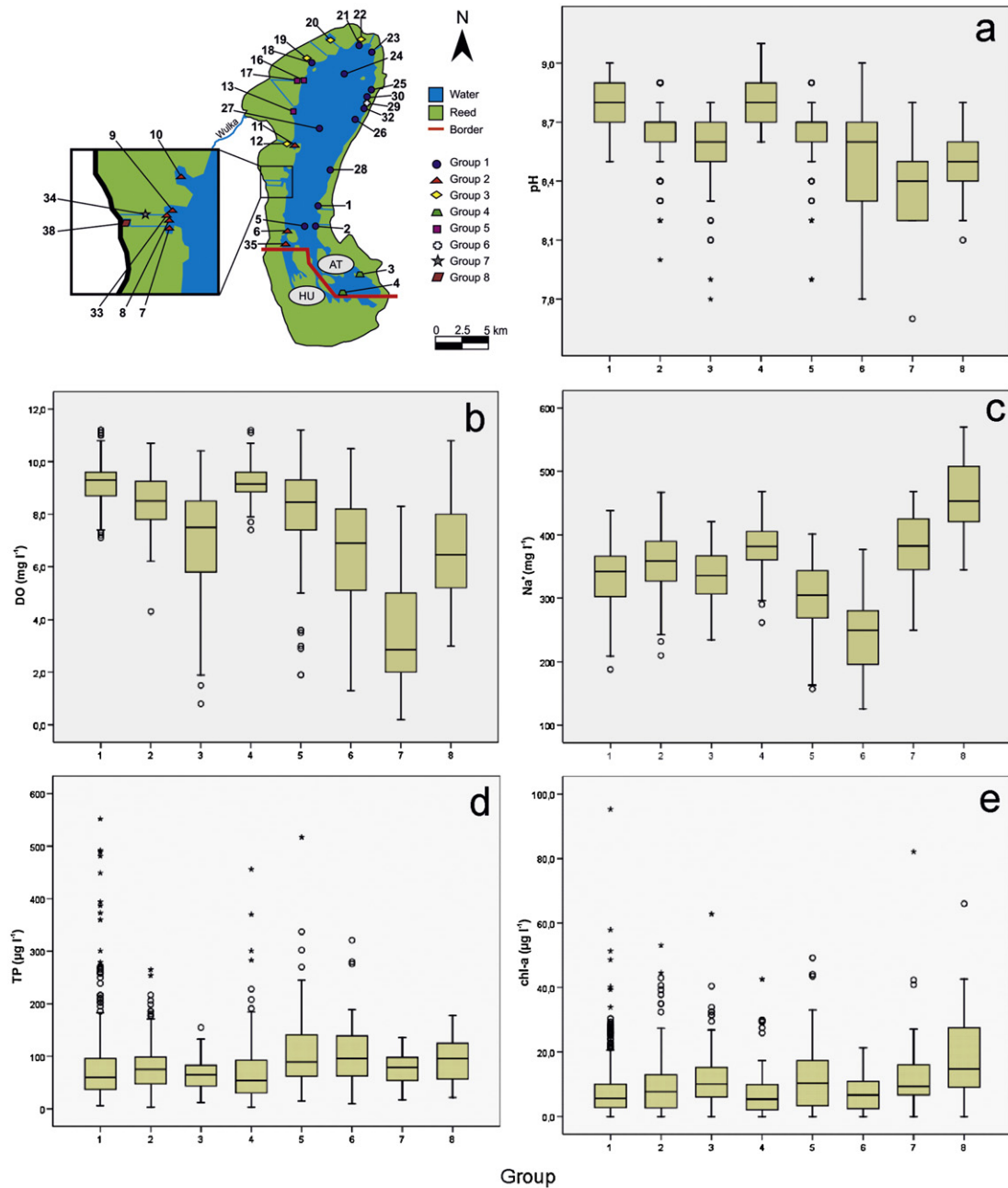


Fig. 5. Box-and-whiskers plots of pH (a), DO (b), sodium (c), TP (d) and chl-a (e).

case of total phosphorus there are again great spatial differences between the groups. In the western groups (3, 7 and 8, but not Group 5), the CV of TP is half the average for the lake and one-third of that of Groups 1 and 4. Last but not least the variability of chlorophyll-a is higher in Groups 1, 4 and 7 than the average for the lake.

3.1. Detailed group description

The eight groups individually cover certain (smaller or larger) areas in the Austrian part of the lake. With reference to the areas covered, each group received a name.

Group 1 ("open water") contains most of the lake's sampling sites. Here the pH and dissolved oxygen values are higher, while

the nutrients and chl-a show lower concentrations in comparison to most of the other groups.

Group 2 ("transient") is considered to be a transition zone between the reed belt and the open water area; it covers the SW part of the lake and consists mostly of the sampling sites located within a 100 m range of the reed belt's border with the open water.

Group 3 ("the reed belt") sampling sites are mainly the ones located in the channels that are used to gain access to the reeds for harvesting and fishing. In these 5–6 m wide channels pH and dissolved oxygen values are lower, while chlorophyll-a shows higher concentrations than at most of the sampling sites.

Group 4 ("the bay") consists of the two sampling sites (Stations 3 and 4) located in the SE part of the lake. This area often "behaves" differently from the open water. The open water has its own general current system operating clockwise, mostly independently of

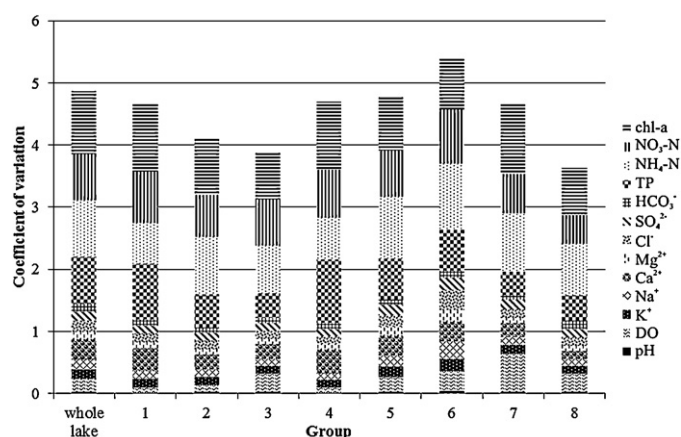


Fig. 6. CVs of the parameters in each cluster group (2000–2009).

the prevailing wind directions (NW & SE). The connection between Group 4 and the open water with the general current is ensured by a couple of narrow channels. However if the wind is strong enough, intensive exchange can be observed between the “separated” bay and the open water. In bays like Group 4 counterclockwise currents can develop if there is a NW wind. If the before wind conditions are strong enough, it may result in the mixing of local water (Jungwirth, 1979). This may be the reason for the separation of sampling sites 3 and 4. In this particular group the concentration of dissolved oxygen and the main components (anions and cations except calcium) is noticeably high and the chlorophyll-a is low.

Group 5 (“the Wulka”) is influenced by the inflow of the river Wulka. It enters the lake in a diffuse way, through a 5 km long reed belt. The low ion concentrations are caused by the diluting effect of the Wulka (originating from the Leitha Mountains) (Grosina, 1990). From the high phosphorus values it became obvious that this input source may be responsible for a considerable part of the lake’s phosphorus loads; it transports the phosphorus washed into its water from agricultural land. This frequently happens during heavy rains.

Group 6 (“the sewage treatment plant”) consist of only one sampling site (Station 29). It is located near the sewage treatment plant situated in the vicinity of Podersdorf am See. The ion concentrations at sampling site 29 are lower than in its surroundings and the interquartile interval of pH is quite wide. This may occur because of the chemical composition of the treated sewage reaching the lake causing a local effect (dilution). It must be stated that the local currents carry these extra inputs toward the inner parts of the lake.

Group 7 (“the weekend houses”) consists of sampling site 34 alone. In the group’s vicinity many weekend houses can be found, all connected to a communal sewage system. Despite this fact, extremely low dissolved oxygen values were occasionally recorded. The reason behind this is the oxygen consumption of the organic substances’ oxidation.

Last but not least Group 8 (“the hotel”) again contains only one sampling site (Station 38), which is located near a hotel. The water in front of the hotel has good light conditions because it is shielded from the wind by the nearby buildings. Furthermore it has a high nutrient content (like all the areas near the shore). The conditions are optimal for algae to reproduce. This may be the reason behind the high chlorophyll-a content of the water along with its low CV. The concentrations of pH and dissolved oxygen were low, while the values for ions showed high concentrations. This sampling site separated from its neighboring ones because of the high ion concentration.

Table 2

1st and 2nd principal components of the PCA results (2000–2009; except 2004).

Parameter/PC	1st PC	2nd PC
pH	0.29	0.81
DO	−0.07	0.78
K ⁺	0.85	0.03
Na ⁺	0.96	−0.06
Ca ²⁺	−0.31	−0.55
Mg ²⁺	0.94	−0.05
Cl [−]	0.95	−0.06
SO ₄ ^{2−}	0.92	0.04
HCO ₃ [−]	0.73	−0.51
TP	−0.08	−0.37
NH ₄ -N	−0.12	−0.47
NO ₃ -N	−0.39	−0.17
Chl-a	0.19	−0.16
Explained variance	5.23	2.25
Proportion of total	0.403	0.173

The differences between the groups can be seen in the box-plots. For example the chemical, biological processes in the reed belt produce humic substances which originate mostly from the decomposition of reed detritus. This, and the diluting effect of the inflowing water, results in lower pH in Groups 3, 5, 7 and 8 (Fig. 5a) in comparison to the other groups. These inflowing waters can be considered to be truly representative of the lake only after flowing slowly through the reed belt. In the open water, because of the frequent mixing, and low nutrient input, the oxygen conditions are good. On the contrary, in the reed belt oxygen concentrations are low (Fig. 5b) or in certain cases even anaerobic conditions may occur (Neuhuber and Hammer, 1979; Dinka and Szeglet, 2001). The lack of mixing and the oxidation of organic substances decreases dissolved oxygen content.

In the reed channels (Groups 3, 7 and 8, but not Group 5) the CV of total phosphorus is lower than average (Fig. 6), because the reed stands have a role in the exchange of nutrients, absorbing nutrients from the sediment and water, and returning these through the processes of excretion and mineralization (Dykyjová and Úlehlová, 1998), thereby forming a crucial link in the detritus food chain (Wetzel, 2001; Dinka et al., 2010). The reed belt functions as a sink for a significant percentage of the phosphorus coming from diffuse sources in the vegetation period (agriculture, livestock, etc.), dampening its effect in the channels. In the process of harvesting the retained and consumed nutrient may be extracted from the water, thereby decreasing succession.

With regard to the chlorophyll-a parameter, an interesting situation arises which should be discussed, as in the open water (Groups 1, 2, 4 and 6) the chlorophyll-a content is much lower than in the channels and on the lakeward side of the reed belt in the western part of the lake (Groups 3, 5, 7 and 8) (Fig. 5e). This happens because the reeds protect the water from the wind and as a result mixing is less frequent, while the width of the channels enables good light conditions, and the area is rich in nutrients, providing good conditions for algae growth. On the contrary, in the open water wind frequently stirs up the sediment and light conditions limit algae growth, in this way prohibiting their multiplication (Dokulil and Padisák, 1993).

3.2. Background processes in Neusiedler See

In determining the latent variables present over a long time period in the lake and the distinct groups, PCA was applied to the data for each year 2000–2009. All of these years, except for the year with extreme droughts (2004) seemed homogenous. Therefore 2004 was analyzed separately and all the other years together (Table 2). From this point on the time interval 2000–2009 only

Table 3
1st and 2nd principal components (2004).

Parameter/PC	1st PC	2nd PC
pH	0.54	−0.34
DO	0.17	−0.70
K ⁺	0.96	0.07
Na ⁺	0.98	0.06
Ca ²⁺	−0.42	−0.73
Mg ²⁺	0.91	−0.23
Cl [−]	0.95	0.22
SO ₄ ^{2−}	0.90	0.32
HCO ₃ [−]	0.67	−0.15
TP	0.07	−0.87
NH ₄ -N	−0.35	0.39
NO ₃ -N	−0.34	0.03
Chl-a	0.29	−0.36
Explained variance	5.69	2.42
Proportion of total	0.438	0.186

refers to those years with the experience of 2004. Only component loadings with an absolute value higher than 0.70 were considered to be a part of further investigations (in bold). The decision regarding the margin where a component score is considered to be relevant or not was made based on the study of Cloutier et al. (2008). Only the first two principal components were evaluated because these were the only ones explaining a significant part of the data's variance.

3.2.1. Latent variables present in the whole lake for the time interval 2000–2009

Two principal components were identified as responsible for 57.6% of the data's variance (Table 3). The first one explained 40.3% of the total variance of the data and in it only the ions with the exception of calcium had significant component loadings, while the second explained 17.3% of total variance, where only the pH and dissolved oxygen parameters had high

component loadings. The component scores were determined for each sampling site; as a result, we were able to produce isoline maps describing their spatial changes. Three areas were distinguishable on this map regarding the first PC (Fig. 7a): two areas with smaller component scores in the NE and NW parts of the lake, and one with high component scores on the western side. The second PC's component scores gradually increase from NW to SE (Fig. 7b), suddenly reaching an explicit boundary (open water), indicating a change in habitat. The anomaly (absence of lower values forming an ellipse in the Wulka delta) seen in the NW part of the lake in the case of the first PC cannot be observed in the case of the second PC.

Regarding the exceptional year 2004, the first PC explained 43.8% and the second 18.6% of the total variance in the data. As in the case of the PCA conducted on the longer datasets (2000–2009) the ions were present in the first PC except for calcium. The difference was that in the case of 2004 HCO₃[−] component loadings did not reach the critical 0.70 value. In the second PC pH lost its determining role, while TP and calcium gained their position next to dissolved oxygen as a determining parameter. In the first PC's map (Fig. 8a) there is no noticeable difference between 2004 and years 2000–2009, however in the second PC's map (Fig. 8b) the component scores show a gradual decrease from their maximum in the NW part of the lake toward its center. This picture is highly different from the one presented in Fig. 7b.

Given that Neusiedler See is a salt lake, during PCA in the first PC only the inorganic parameters were dominant (except for calcium). It can be stated that ions determine the variance of the dataset the most, while calcium and biogenic carbonate precipitation play a less significant role. In the second PC the reason behind the determining role of pH and dissolved oxygen is the presence of humic substances and the oxidation of organic matter respectively. These are responsible for the water's brown color in the reed belt (Dokulil and Herzig, 2009).

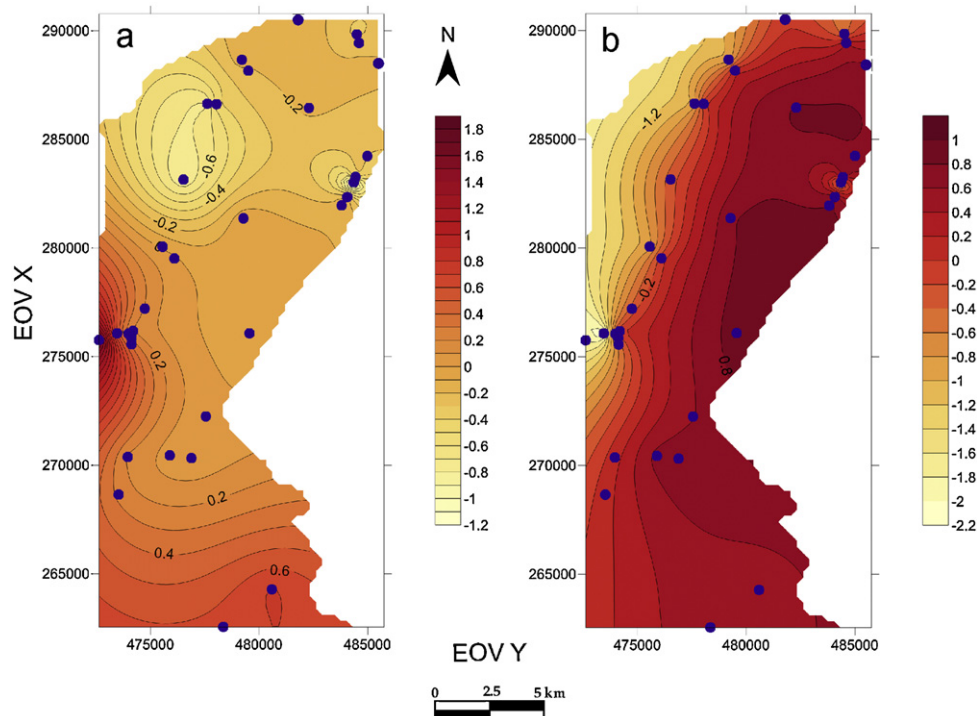


Fig. 7. Spatial distribution of the 1st (a), and the 2nd PC scores (b) (2000–2009; except 2004).

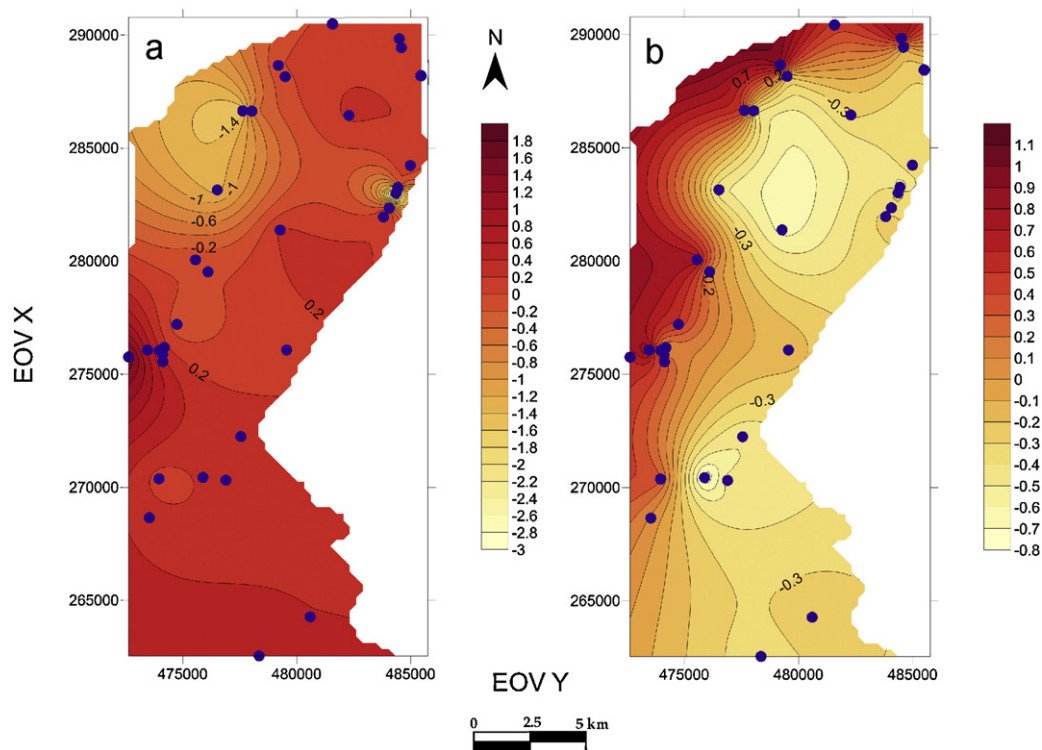


Fig. 8. Spatial distribution of the 1st (a), and the 2nd PC scores (b) (2004).

From the component loading maps (2000–2009) (Figs. 7 and 8) it became clear that the Wulka, the sewage treatment plant, the hotel and the weekend houses may have a pronounced local impact on the processes evolving in Neusiedler See. Because all of these sampling sites exercise influence on the ion content of the water, all of their locations can clearly be seen on the first PC's maps (Figs. 7a and 8a). On the maps with the values of the second PC (Fig. 7b) the effect of the Wulka cannot be seen because it does not have a serious effect on the pH and oxygen conditions, unlike the other three sampling sites (Stations 29, 34, 38). In 2003 when there was a devastating drought and the water in the reed belt almost completely evaporated, the reed belt's role as a significant sink factor for TP decreased. This event can clearly be seen in the data for the year 2004, when the water level was still low, so the unusually strong wind conditions stirred up the lake's sediment in this way increasing the suspended solids, and as a result TP as well.

Table 4

Component loadings in Groups 1, 2, 3.

Parameter/PC	Group 1		Group 2		Group 3	
	1st PC	2nd PC	1st PC	2nd PC	1st PC	2nd PC
pH	0.59	0.44	0.37	0.64	0.33	0.78
DO	–0.22	0.00	–0.16	0.75	–0.03	0.88
K ⁺	0.85	0.08	0.84	–0.02	0.88	0.04
Na ⁺	0.96	0.02	0.93	0.01	0.97	–0.02
Ca ²⁺	–0.05	–0.80	–0.07	–0.55	–0.07	0.02
Mg ²⁺	0.93	–0.05	0.93	0.06	0.92	0.02
Cl [–]	0.95	0.06	0.94	0.01	0.90	–0.08
SO ₄ ^{2–}	0.95	0.00	0.94	–0.02	0.89	0.14
HCO ₃ [–]	0.78	–0.43	0.69	–0.36	0.60	–0.66
TP	–0.02	–0.75	0.14	–0.37	0.08	–0.07
NH ₄ -N	–0.02	0.04	–0.13	–0.56	–0.01	–0.39
NO ₃ -N	–0.50	0.06	–0.48	0.05	–0.29	0.06
Chl-a	0.17	–0.37	0.14	0.48	0.27	0.36
Explained variance	5.58	1.74	5.15	2.08	4.80	2.14
Proportion of total	0.429	0.134	0.396	0.160	0.369	0.164

3.2.2. Latent variables present in the different cluster groups

Because cluster analysis pointed out that there is more than one distinguishable area in the lake, PCA was conducted on the individual groups as well. Only those groups (Groups 1, 2, 3) were part of the analysis which had enough samples to conduct PCA on.

The first PC explained 37–43% of the total variance of the data (Table 4) and the second PC 13–16%. As in the case of the previous analyses (2000–2009) the ions were present in the first PC again except for calcium. It is worth mentioning that the component loading of HCO₃[–] in Groups 2 and 3 was lower than 0.70. The second PC does not show homogeneity, as the first PC did. In the second PC of Group 1 calcium and TP, in Group 2 dissolved oxygen and in Group 3 pH and dissolved oxygen had component loadings higher than 0.70.

Regarding the results obtained from the PCA conducted on the groups it can be said that the differences in the second PC originate from the different habitats' processes.

4. Conclusions

This study demonstrated that uni- and multivariate data analysis methods are excellent tools in interpreting complex datasets. The sampling sites of Neusiedler See were grouped and the grouping verified by discriminant analysis. At the different time points the sites showed different group patterns which indicate the variability of the lake's habitats and processes (just as the CVs did). Eight different areas were distinguished. Using Wilks' lambda statistics it was verified that the ions are most responsible for the formation of the previously discussed cluster groups. The nutrients and chl-a were less relevant in this question. Based on the boxplots derived from the separated groups it was proven that the anthropogenic activity in the lake's surroundings has a great localized effect on the lake's water chemistry. The Wulka and the sewage treatment

plant both have a clear effect on ion variability. The local (treated sewage) and diffuse (materials with high nutrient content washed in from agricultural areas) phosphorus loads are both considerable. The reed belt had a great contribution in buffering the arriving nutrient loads at almost all of the inlets.

The final approach was to identify the parameters accounting for most of the variance of the dataset. It became clear that the ions are most determining. The component score maps proved the suspicion that the inflows and reeds have a great effect on the pH and oxygen conditions along with concentrations of the main components (anions and cations) in the area. The determining component loadings proved to be stable throughout 2000–2009 except the year 2004. In space, however, regarding the different groups the first principal component showed again a homogenous pattern while in the second principal component the component loadings of the variables differ to a great extent between the reed belt and the open water.

In summary it can be said that Neusiedler See cannot be treated as one homogeneous system. The parameters are highly variable in space and time, to a greater degree than was suspected. This exceptional variability originates from the lake's shallow water depth, unstable water balance, and anthropogenic activity (agriculture, tourism, sewage treatment, etc.) in the lake's surroundings which can be dampened with optimal management of the reed belt.

The lake is aging. In order to preserve the present state, joint management measures have to be taken each year, such as reed harvesting, managing the runoff and water quality of the inlets and outlets and last but not least managing the water level, which has a widespread effect on all the previously mentioned factors. Based on the information obtained it is possible to optimize the current management even further, in this way increasing its cost efficiency, which is a yet greater question in this period of financial crisis. If in the not too distant future crucial changes have to be made (improving the reed belt's water supply, maintenance of channels, etc.) in the lake's management the present study (by assessing the data series of many decades) can be of great help.

Acknowledgments

We the authors would like to thank the Biological Station Neusiedler See in Illmitz for providing the datasets; furthermore we would like to thank Georg Wolfram, Martin Dokulil and the colleagues working at the Department of Physical and Applied Geology for the stimulating discussions. The work of J. Kovács was funded by the Hungarian National Science Foundation, OTKA-K 101065, while the work of I.G. Hatvani was realized in the frames of TÁMOP 4.2.4. A/1-11-1-2012-0001 "National Excellence Program – Elaborating and operating an inland student and researcher personal support system". The project was subsidized by the European Union and co-financed by the European Social Fund and the "Habilitas" scholarship of the Hungarian Development Bank Plc. (MFB Zrt.) which we are very grateful for.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ecoleng.2013.02.005>.

References

Back, W., 1966. Hydrochemical facies and ground-water flow patterns in northern part of Atlantic Coastal Plain. U.S. Geol. Surv. Prof. Paper 424-D. United States Government Printing Office, Washington, DC.

Brix, H., 1987. Treatment of wastewater in rhizosphere of wetland plants: the root-zone method. Water Sci. Technol. 19, 107–108.

Cloutier, V., Lefebvre, R., Therrien, R., Savard, M.M., 2008. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. J. Hydrol. 353 (3–4), 294–313.

Daniel, J.J., 1999. Introduction to Atmospheric Chemistry. Princeton University Press, New Jersey.

Dinka, M., 1986. The effect of mineral nutrient enrichment of Lake Balaton on the common reed (*Phragmites australis*). Folia Geobot. Phytotax. 21, 65–84.

Dinka, M., Szeglet, P., 2001. Some characteristics of reed (*Phragmites australis*/Cav./Trin ex Steudel) that indicate different health between vigorous and die-back stands. Verh. Internat. Ver. Limnol. 27, 3364–3369.

Dinka, M., Ágoston-Szabó, E., Szeglet, P., 2010. Comparison between biomass and C, N, P, S contents of vigorous and die-back reed stands of Lake Fertő/Neusiedler See. Biologica 65 (2), 237–247.

Dinka, M., Ágoston-Szabó, E., Berczik, Á., Kutrucz, Gy., 2004. Influence of water level fluctuation on the spatial dynamic of the water chemistry at Lake Fertő/Neusiedler See. Limnologica 34 (1–2), 48–56.

Dobesch, H., Neuwirth, F., 1979a. Climatic conditions. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Hague, pp. 47–64.

Dobesch, H., Neuwirth, F., 1979b. Water balance. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Hague, pp. 79–84.

Dokulil, M., Padisák, J., 1993. Langfristige (1968–1990) und jahreszeitliche Dynamik der planktischen Diatomeen im Neusiedlersee. BfB-Bericht 79, 5–11.

Dokulil, M.T., Herzig, A., 2009. An analysis of long-term winter data on phytoplankton and zooplankton in Neusiedler See, a shallow temperate lake, Austria. Aquat. Ecol. <http://dx.doi.org/10.1007/s10452-009-9282-3>.

Dykyjová, D., Úlehlová, B., 1998. Mineral economy and cycling of minerals in wetlands. In: Westlake, D.F., Kvet, J., Szczepansky, A. (Eds.), The Production Ecology of Wetlands. Cambridge University Press, Cambridge, pp. 319–366.

Filik Iscen, C., Emiroglu, Ö., İlhan, S., Arslan, N., Yilmaz, V., Ahiska, S., 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environ. Monit. Assess. 144, 269–276.

Freeze, R.A., Cherry, J.A., 1979. Groundwater. Prentice-Hall, New Jersey.

Gattinger, T., 1979. The hydrogeology of Neusiedlersee and its catchment area. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Hague, pp. 85–88.

Grosina, H., 1990. Forschungsbericht 1987–1989. Rötzer-Druck, Eisenstadt.

Güler, C., Thyne, G.D., McCray, J.E., Turner, A.K., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeol. J. 10, 455–474.

Hatvani, I.G., Kovács, J., Kovácsné Székely, I., Jakusch, P., Korponai, J., 2011. Analysis of long term water quality changes in the Kis-Balaton Water Protection System with time series-, cluster analysis and Wilks' lambda distribution. Ecol. Eng. 37 (4), 629–635.

Hatvani, I.G., Magyar, N., Tanos, P., Korponai, J., Székely, I.K., Herzig, A., Kovács, J., 2012. Determining anthropogenic effects using principal component analysis on a fluvial (E Hungary) and two lake ecosystems (W Hungary, E Austria). In: Visco, G. (Ed.), CMA4HC: Use of Multivariate Analysis and Chemometrics in Cultural Heritage and Environment. Rome, Italy, 27–30 May 2012, pp. 30–31 (ISBN: 9788875473327).

Herzig, A., Dokulil, M., 2001. Neusiedler See – ein Steppensee in Europa. In: Dokulil, M., Hamm, A., Kohl, J.G. (Eds.), Ökologie und Schutz von Seen. Fakultät-Universität, Wien, pp. 401–415.

Hill, R.A., 1940. Geochemical patterns in Coachella Valley, California. Trans. Am. Geophys. Union 21, 46–49.

Hu, L., Hu, W., Deng, J., Li, Q., Gao, F., Zhu, J., Han, T., 2010. Nutrient removal in wetlands with different macrophyte structures in eastern Lake Taihu, China. Ecol. Eng. 36 (12), 1725–1732.

Johnson, R.A., Wichern, D.W., 1992. Applied Multivariate Statistical Analysis, 3rd ed. Prentice Hall, New Jersey.

Juahir, H., Zain, S.M., Yusoff, M.K., Hanidza, T.I.T., Armi, A.S.M., Toriman, M.E., Mokhtar, M., 2011. Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. Environ. Monit. Assess. 173 (1–4), 625–641.

Jungwirth, M., 1979. Currents. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Håga, pp. 85–88.

Kazi, T.G., Arain, M.B., Jamali, M.K., Jalbani, N., Afridi, H.I., Sarfraz, R.A., Baig, J.A., Shah, A.Q., 2009. Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotoxicol. Environ. Saf. 72, 301–309.

Ketskeméty, L., Izsó, L., 2005. Introduction into the SPSS System (Bevezetés az SPSS programrendszerbe). ELTE Eötvös Kiadó, Budapest, Hungary (in Hungarian).

Kiedrzyńska, E., Wagner, I., Zalewski, M., 2008. Quantification of phosphorus retention efficiency by floodplain vegetation and a management strategy for a eutrophic reservoir restoration. Ecol. Eng. 33 (1), 15–25.

Kühl, H., Kohl, J.G., 1992. Nitrogen accumulation, productivity and stability of reed stands (*Phragmites australis* (Cav.) Trin. ex Steud.) at different lakes and sites of the lake district of Uckermark and Mark Brandenburg (Germany). Int. Rev. Gesamten Hydrobiol. Hydrogr. 77 (1), 85–107.

Květ, J., 1973. Mineral nutrients in shoots of reed (*Phragmites communis* Trin.). Pol. Arch. Hydrobiol. 20, 137–147.

Lattin, J.M., Carroll, J.D., Green, P.E., 2003. Analyzing Multivariate Data. Duxbury Press, New York.

- Löffler, H., 1974. Der Neusiedlersee. Naturgeschichte eines steppensee. Molden Verlag, Wien, Austria.
- Löffler, H., 1979. Foreword. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Håga, pp. 47–64.
- Lu, S., Hu, H., Sun, Y., Yang, J., 2009. Effect of carbon source on the denitrification in constructed wetlands. *J. Environ. Sci.* 21 (8), 1036–1043.
- McKenna, J., 2003. An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environ. Modell. Softw.* 18, 205–220.
- Mitsch, W.J., Gosselink, J.G., 1986. Wetlands. Van Nostrand Reinhold Company, New York.
- Morgan, C.O., Winner Jr., M.D., 1962. Hydrochemical facies in the 400 foot and 600 foot sands of the Baton Rouge Area, Louisiana. U.S. Geol. Surv. Prof. Paper 450-B. United States Government Printing Office, Washington, DC.
- Neuhuber, F., Hammer, L., 1979. Oxygen conditions. In: Löffler, H. (Ed.), Neusiedlersee. The Limnology of a Shallow Lake in Central Europe. Dr. W. Junk B.V. Publishers, Håga, pp. 121–130.
- Norušis, M.J., 1993. SPSS for Windows Professional Statistics Release 6.0. Prentice Hall, Englewood Cliffs, NJ.
- Padisák, J., 2005. General Limnology (Általános limnológia). ELTE Eötvös Kiadó, Budapest (in Hungarian).
- Piper, A.M., 1944. A graphic procedure in the geochemical interpretation of water-analyses. *Trans. Am. Geophys. Union* 25, 914–923.
- Raghi-Atri, F., Bornkamm, R., 1979. Wachstum und chemische Zusammensetzung von Schilf (*Phragmites australis*) in Abhängigkeit von der Gewässereutrophierung. *Arch. Hydrobiol.* 85 (2), 192–228.
- Sargaonkar, A., Deshpande, V., 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environ. Monit. Assess.* 89 (1), 43–67.
- Shrestha, S., Kazama, F., 2007. Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ. Modell. Softw.* 22 (4), 464–475.
- Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsas, D., Anthemidis, A., Sofoniou, M., Kouimtzi, Th., 2003. Assessment of the surface water quality in Northern Greece. *Water Res.* 37 (17), 4119–4124.
- Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): a case study. *Water Res.* 38 (18), 3980–3992.
- Tanner, C.C., 1996. Plants for constructed wetlands treatment systems – a comparison of growth and nutrient uptake of eight emergent species. *Ecol. Eng.* 7, 59–83.
- Vega, M., Pardo, R., Barrado, E., Debán, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res.* 32 (12), 3581–3592.
- Vymazal, J., Kröpfelová, L., Švehla, J., Štrichová, J., 2010. Can multiple harvest of above-ground biomass enhance removal of trace elements in constructed wetlands receiving municipal sewage? *Ecol. Eng.* 36 (7), 939–945.
- Ward Jr., J.H., 1963. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58 (301), 236–244.
- Wetzel, R.G., 2001. Limnology: Lake and River Ecosystems, 3rd ed. Academic Press, San Diego.
- WFD (Water Framework Directive), Directive of the European Parliament and of the Council 2000/60/EC Establishing a Framework for Community Action in the Field of Water Policy, European Union, Luxembourg, 2000, PE-CONS 3639/1/00 REV 1.
- Wilks, S.S., 1932. Certain generalizations in the analysis of variance. *Biometrika* 24, 471–494.
- Wunderlin, D.A., Diaz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C., Bistoni, M.A., 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba-Argentina). *Water Res.* 35 (12), 2881–2894.
- Yidana, S.M., Ophori, D., Banoeng-Yakubo, B., 2008. A multivariate statistical analysis of surface water chemistry data – the Ankobra Basin, Ghana. *J. Environ. Manage.* 86 (1), 80–87.
- Zhou, F., Liu, Y., Guo, H., 2007. Application of multivariate statistical methods to water quality assessment of the watercourses in Northwestern New Territories, Hong Kong. *Environ. Monit. Assess.* 132 (1–3), 1–13.
- Zhu, L., Li, Z., Ketola, T., 2011. Biomass accumulations and nutrient uptake of plants cultivated on artificial floating beds in China's rural area. *Ecol. Eng.* 37 (10), 1460–1466.