Journal of Environmental Management 110 (2012) 151–158

Contents lists available at SciVerse ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Delimiting sub-areas in water bodies using multivariate data analysis on the example of Lake Balaton (W Hungary)

József Kovács [a], Mária Nagy [b], Brigitta Czauner [a], Ilona Székely Kovács [c], Andrea K. Borsodi [d], István Gábor Hatvani [a,]*

[a] Eötvös Loránd University, Department of Physical and Applied Geology, H-1117 Budapest, Pázmány P. sétány 1/C, Hungary
[b] Middle Transdanubian Environment Inspectorate, H-8000 Székesfehérvár, Balatoni út 6, Hungary
[c] Budapest Business School, Institute of Methodology, H-1054 Budapest, Alkotmány u. 9-11, Hungary
[d] Eötvös Loránd University, Department of Microbiology, H-1117 Budapest, Pázmány P. sétány 1/C, Hungary

ARTICLE INFO

Article history:
Received 8 August 2011
Received in revised form
21 March 2012
Accepted 4 June 2012
Available online

Keywords:
Cluster analysis
Discriminant analysis
EU Water Framework Directive
Lake Balaton
Monitoring
Water body sub-areas
Wilks' lambda distribution

ABSTRACT

The main aim of the European Water Framework Directive (WFD, 2000) is to commit European Union Member states to the achievement of good qualitative and quantitative status for all water bodies by 2015. To achieve this, a reference state has to be determined and appropriate monitoring has to be carried out. Based on the fact that the WFD classifies Lake Balaton, the largest shallow freshwater lake in Central Europe, as one water body, and due to the lack of funds, the number of sampling locations on the lake was decreased. The aim of this study was to determine how many sub-areas with different WFD-related attributes (in this case, parameters) can be delimited in the so-called one water body of Lake Balaton, so that a number of representative sampling locations might be retained. To determine Lake Balaton's different water quality areas (i.e. sub-areas of water body) 23 parameters (inorganic compounds) were examined from 10 sampling locations for the time interval 1985–2004 using cluster- and discriminant analysis, and Wilks' lambda distribution. With cluster analysis we were able to determine two time intervals (1985–1997 and 1998–2004) with three patterns of sub-areas, two from the first and one from the latter interval. These patterns pointed to the fact that for the whole investigated time interval (1985–2004) a total of five sub-areas were present, changing in number and alignment. Then the results were verified using discriminant analysis, and the parameters which influenced the sub-areas the most were determined using Wilks' lambda distribution. The conclusion was that to be able to follow the changes in alignment of the sub-areas and to get a comprehensive picture of Lake Balaton, a minimum of five sampling locations should be retained, one in each sub-area. Based on this study the Water Authorities chose to keep five out of ten sampling locations so that the sub-areas could be described. We consider this a great success and the methodology as an example for setting up sub-areas in a water body.

1. Introduction

The main aim of the European Water Framework Directive (WFD) is to commit European Union Member States to the achievement of good qualitative and quantitative status for all water bodies by 2015. This is one of the biggest challenges nature conservation has ever faced; it underlines the water managements' foregoing approach to the conservation of the quality of water bodies and their biological and ecological states (WFD, 2000). Similar actions have already been undertaken in the U.S.A. as well, following ecological surveys in the interests of the restoration and management of lakes along with their watersheds and reservoirs (e.g. Cooke et al., 2005; Edmondson, 1991; USEPA, 2009; Zeng and Rasmussen, 2004).

According to the WFD's definition, a "water body" as a basic unit has to be ecologically "near homogeneous", while as a legal (administrative) term (Sánchez et al., 2009) it includes the management of the water body as well (Ferreira et al., 2006). The aim of the water bodies policy is to arrange them in such a way that they could be managed, treated, and monitored with the highest efficiency, thus the collected data could represent their current state. Since in the EU Member States the single surface waters are

* Corresponding author. Eötvös Loránd University, Faculty of Science, Department of Physical and Applied Geology, H-1117 Budapest, Pázmány Péter sétány 1/C, Hungary. Tel.: +36 70317 97 58; fax: +36 1 31 91738.
E-mail address: hatvaniig@gmail.com (I.G. Hatvani).

0301-4797/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jenvman.2012.06.002

generally considered one "water body", the practice of delimiting them based on water quality parameters is not common aside from some exceptions. For instance, Ferreira et al. (2006) raised questions regarding the monitoring of water bodies in estuaries. Muxika et al. (2007) subdivided one water body into smaller parts before determining its reference conditions according to the WFD by using parameters regarding benthic communities. Earle and Blacklocke (2008) who tried to implement the WFD in Ireland, also mentioned financial problems that every Member State had to face. Namely, a balance should be found between the financial and scientific aspects, for example by setting up enough monitoring locations. Otherwise, "it is possible that no further degradation of water body could be observed" (László et al., 2007). Eventually, based on numerical simulations Dinesh (2008) established that water bodies are dynamic objects undergoing spatial and temporal changes, such as floods and droughts.

Following of the WFD's "one lake, one water body" approach also raised significant questions in Hungary, particularly in case of Lake Balaton.

Lake Balaton (LB) is the largest (596 km^2) shallow (average depth 3.2 m) fresh water lake in Central Europe (Istvánovics et al., 2007). It is 17,000–19,000 years old (Cserny and Nagy-Bodor, 2000) and located in W Hungary, in the S-SE foreground of the Transdanubian Central Range (Fig. 1). According to paleolimnological research, trophic state of Lake Balaton was changed between meso- and eutrophy in the last Millennia (Korponai et al., 2011). It is considered as a Ramsar site; however, due to vast tourist activity it is only under the convention's protection from late summer until late spring CWIIWH, 1987.

The morphologically diverse watershed of the lake is approximately 5,181 km^2 (Cserny, 1993) in which 51 inflows are located of which 20 are constant. Among the latter, the River Zala is the most important, providing 45% of LB's water supply, and 35–40% of its nutrient input (Lotz, 1988). At the mouth of the River Zala the Kis-Balaton Water Protection System (KBWPS) was established in 1985 in order to filter the nutrients entering LB with the river and thus to prevent the lake's further eutrophication. It serves a crucial role in the management of the lake. Regarding the KBWPS, Hatvani et al. (2011) published a study using methods being similar to that applied in this paper. Using cluster analysis to determine the similarity of the KBWPS' sampling site, they pointed out that, because of its water level management, the reedy areas began to decrease this way affecting the water quality of LB. On the other hand, LB's only outflow is the Sió Canal, which was put into operation in 1863 at the southeastern tip of the lake (Kovács et al., 2010). Consequently, the exterior nutrient load of LB varies systematically from the east to the west (e.g. Sagehashi et al., 2001).

However, water quality can also differ temporally and spatially in the four geographical basins that make up the lake's bed (e.g. Hajnal and Padisák, 2008; Padisák et al., 2006; Pálffy and Vörös, 2011) due to other factors as well. For instance, the lake's energetic mass cycle is basically supported by the primary production of the planktonic algae, which not only determine, but also indicate water quality (Vörös and Somogyi, 2009). Accordingly, in the winters of 1974 and 1975 algae blooms were noticeable and followed by mass fish deaths. This water quality deterioration mainly affected the western part of the lake, but the middle part becomes eutrophic as well. Subsequently, the first lake-wide algae bloom was witnessed in 1982 due to the mass reproduction of the Wolostoxin-producing, filiform, N-fixing *Cylindrospermopsis raciborskii* (Istvánovics et al., 2007; Padisák, 1997; Parpala et al., 2003; Sprőber et al., 2003). This event recurred in 1994 after a few years of drought and huge eel extinctions in 1992 and 1993, when the blue cyanobacteria turned the entire lake hypertrophic. However, thanks to a series of significant counter-measures and rainy weather LB's water quality could already been classified as "good" in 1995, and also since then it has been improving irrespective of water level changes. Thus nowadays oligotrophic trends can be observed (Hajnal and Padisák, 2008; Istvánovics et al., 2002, 2007), although it is still considered to be mesotrophic.

In spite of the spatial and temporal variability of LB's water quality described above, Hungarian typology still determines the lake as one water body (CDWE, 2010) (like every other lake in the country, and in the EU as well (WFD, 2000)). Consequently, and also due to financial constraints, in 2005 the spatial sampling of LB had to be revised, and the number of sampling locations had to be decreased. The new sampling points were settled according to this study. One sampling site was retained at the mouth of the River Zala, which is the main origin of LB's nutrient loads (Tátrai et al., 2000). Furthermore, from 2006 monitoring was retuned according to the requirements of the WFD. Thus on one hand, the biological parameters such as phytoplankton, phytobenton, macrozoobenton, fish, and macrophyton came to the fore, while on the other hand monitoring of the chemical parameters (oxygen, trophic parameters, nutrients) became secondary. In other words, if the biological parameter values are satisfactory, sampling of the chemical ones can even be discarded. Following the instructions of the WFD, monitoring of the lake can be more effective in theory.

However, already the "one lake, one water body" approach is questionable in the case of LB. It might been classified correctly by the WFD in the 16th category as a calcareous, large lake with an open water space (open area larger than 33%), because its surface is mainly consists of open water-space, while the percentage of reeds is less than 2%, and nowadays there is no significant difference in the ecological quality of the lake's western and eastern parts. However, this idea may not stand or has to be refined taking into account other facts such as the lake's geological heterogeneity, the differing nutrient loads arriving from the lake's watershed area (CDWE, 2010),

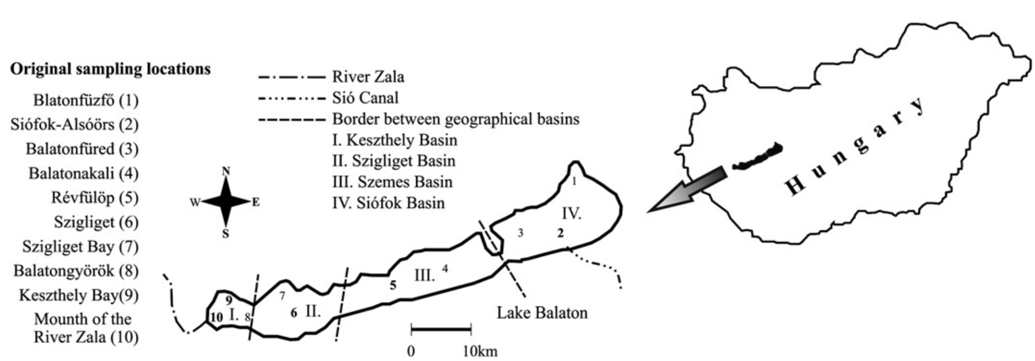

Fig. 1. Orientation of Lake Balaton with its basins and sampling locations between 1985 and 2004.

or meteorological impacts. These clearly affect the reference state, the classification, and the results of the monitoring as well.

Bearing these facts in mind our aim was to design and present an impartial method that is capable of refining the "one water body" idea by delimiting sub-areas in LB. This objective fortunately coincided with the following question on the Water Authority's part. How many sub-areas characterized by differing water quality can be delimited in the so called one water body of LB, so that a number of representative sampling locations could be retained?

2. Applied data and methods

2.1. Sampling, analysis and data handling

We acquired data concerning LB from the Middle Transdanubian Environment Inspectorate for the time interval 1985–2004. The data mainly consists of chemical parameters describing the lake's water quality. Samples were taken from ten sampling locations near the lake's centerline (Fig. 1).

The examined components were as follows: Dissolved O_2; 5 day-Biological oxygen demand (BOD^5); Chemical oxygen demand-potassium-dichromate (COD^C); Chemical oxygen demand-potassium-permanganate (COD^P); NH_4–N, NO_2–N, NO_3–N, Organic-nitrogen (ON), Total suspended solids (TSS); CaO; Na^+; K^+; Ca^{2+}; Mg^{2+}; HCO_3^-; SO_4^{2-}; Cl^- (mg l^{-1}); Total-phosphorous (TP); PO_4–P; Chlorophyll-a (μg l^{-1}); pH; Conductivity (μS cm^{-1}); Water temperature (W_{temp}; °C); Alkalinity (mmol l^{-1}). The number of received data was around 91,000. This amount would be sufficient, however not optimal, in solving the problem. The basic requirement of multivariate exploratory techniques is that data must not be missing at any of the sampling events (Kovács et al., 2012). This situation, however, occurred many times, because during the investigated period the number of analyzed parameters changed, usually increased (e.g., in 1995 there were 25, in 1996–27, in 1997–30, and between 2000 and 2004 there were 31 investigated parameters). Thus, in order to avoid the incomparability of the results, only 23 parameters which had been continuously measured were used during the research. Moreover, the sampling methods have also changed many times. If two methods called forth similar results, then both of the parameters were retained. Another problem was that in LB the sampling was done with two different frequencies, monthly (at the sampling points of Balatonfüred, Balatonfűzfő, Balatongyörök, Révfülöp-Balatonboglár, and Szigliget Bay), and fortnightly (at the sampling points of Balatonakali, Keszthely Bay, Siófok, Szigliget and the mouth of the River Zala). Also in winter, because of the ice cover, sampling could not be done in every occasion, so the number of examined data varied between ~210 and ~420. Taking these facts into consideration, our tests were done for 85 points in time over the period 1985 to 2004, so that we were able to determine the connections between the sub-areas of the water body.

During data preparation the extreme values were handled with caution. The relative deviation of the parameters and other descriptive statistics were determined, since there may be values which seem to be extreme but according to certain parameters' variability those can be all valid. As a result, except for the obviously mistyped ones, all values were used.

2.2. Applied mathematical methods

To find the associations and the main trends in the acquired datasets, multivariate data analysis, cluster analysis (CA), discriminant analysis, and Wilks' lambda distribution were applied. Our aims could have been achieved by applying principal component analysis (PCA) and afterwards hierarchical cluster analysis (HCA), as in the work of Neilson and Stevens (1986), who, by applying T-values,[1] were able to explain 100% of the T-values' variance. Like other scientists (Cansu et al., 2008; Neilson and Stevens, 1986; Simenov et al., 2003; Singh et al., 2004; Zeng and Rasmussen, 2004) – using PCA – they too define the number of principal components based on their eigenvalues. If it is higher or equal to 1 in the case of surface water analysis it usually results in 3–5 factors, which are linear combinations of the original variables, which are uncorrelated. Using this new coordinate system, the original sampling location can be transformed into one with a smaller number of dimensions. This way the cases could be grouped based on their similarity. Based on the studies cited earlier the principal components define only 65–85% of the data's variance. This brings forward the problem that using HCA on the smaller set of independent variables can only approximate the cases in the space of the original variables. In this way the HCA result will only be representative of a part of the original variance. By using HCA on the original dataset this loss in variance could be by-passed.

"Cluster analysis classifies a set of observations into two or more mutually exclusive "unknown" groups based on combinations of interval variables. The purpose of cluster analysis is to discover a system of organizing observations, usually people, into groups where members of the groups share properties in common" (Stockburger, 2001). In our study Ward's method and squared Euclidian distance was applied. It must be stated that this tool cannot be used on time-series (because consecutive data are dependent on each other), only on distinctive sampling points in time.

During our study we used clustering on sampling locations (space) and points in time (time) as well in the order described in Fig. 2.

To test the results of the CA, discriminant analysis can be used. It shows to what extent the planes separating the groups can be distinguished. 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. The functions are generated from a sample of cases for which the group membership is known; the functions can then be applied to new cases that have measurements for the predictor variables but their group membership is as yet unknown (Afifi et al., 2004). The result of the discriminant analysis is often visualized on a surface stretched between the first two discriminating planes (Ketskeméty and Izsó, 2005).

Subsequently, the data from the sampling points were examined using Wilks' λ distribution to define which parameters played the greatest role in determining the formations of the cluster groups. During the analysis a Wilks' λ quotient was assigned to every parameter for every year, where the quotient is:

$$\lambda = \frac{\sum_i \sum_j \left(x_{ij} - \overline{x_i}\right)^2}{\sum_i \sum_j \left(x_{ij} - \overline{x}\right)^2}$$

Where x_{ij} is the jth element of group i, x_i the average of group i, and x_i the total groups average.

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. If $\lambda = 1$, then the mean of the discriminant scores is the same in all groups and there is no inter-group variability, so, in our case the parameter

[1] A T-value was formed for each parameter, containing its annual information for each sampling site.

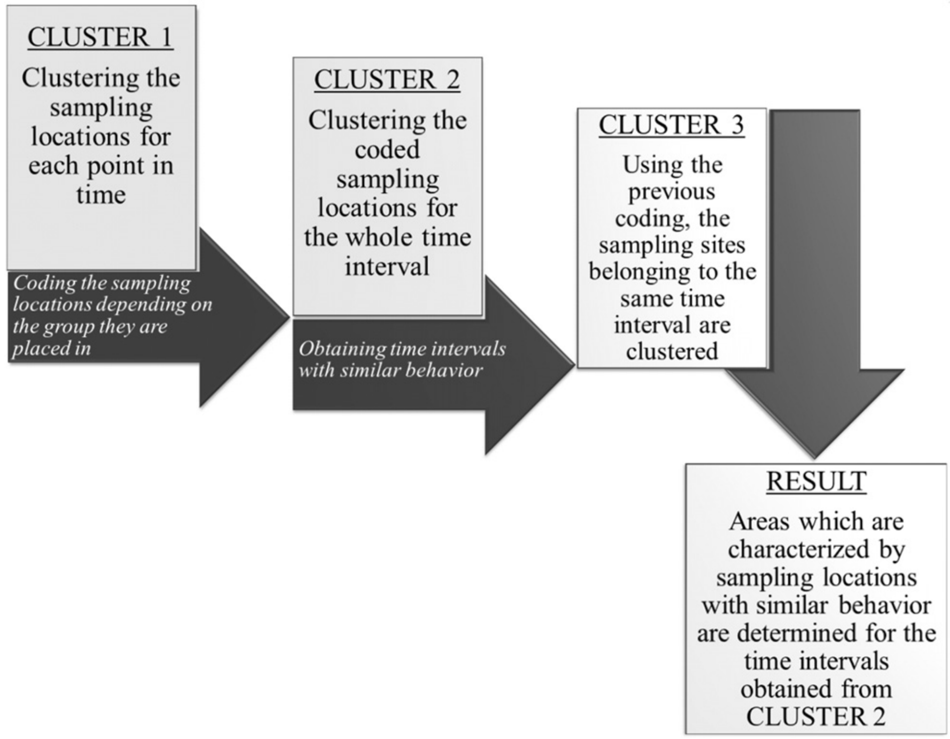

Fig. 2. Step-by-step scheme of the clustering method applied in this study.

did not affect the formation of the cluster groups (Afifi et al., 2004). On the contrary, if $\lambda = 0$, then that particular parameter affected the formation of the cluster groups the most. The smaller the quotient is, the more it determines the content of the cluster groups (Hatvani et al., 2011).

3. Results and discussion

3.1. Determining similar sampling locations for the points in time

Choosing the right method to segment the "one water body" is very important. Among the available cluster analysis seemed to be the most appropriate in selecting similar sampling locations.

During the study 85 points in time were examined employing hierarchical cluster analysis. As a result, 85 dendrograms were produced which show the groups' varying spatial distribution (examples of which may be seen in Fig. 3). Examining the composition of the dendrograms, similar sampling points were identified. The linkage distance was allocated at 20% of the linkage distance where only one group remains (every linkage distance percentage was measured in this way in the study). This method standardized the procedure and facilitated comparison. If the linkage distance had been set to a percentage less than 20%, then group number reduction would not have been possible, while if it had been higher it would have resulted in too significant a reduction in the number of groups.

At the beginning of the evaluation, it became clear that the previous, geographically-based four-basin division had changed over the years. For instance, in the most extreme cases there were six times between 1992 and 1997 when 10 sampling locations formed 6 different groups. On the contrary, there were certain points in time when only two different groups could be identified. (It has to be stated that these groups may be considered as sub-areas, but later on it will become clear that these groups are not yet able to cover the final sub-areas.)

3.2. Determining time intervals with homogenous sub-areas in LB's water body

After obtaining the 85 dendrograms, the time intervals had to be determined when homogenous sub-areas were to be found, when the results of the 85 dendrograms formed sub-areas. Because of the varying results (regarding the 85 dendrograms) – meaning that at different points in time the number of groups varied from 2 to 6 – the sampling locations belonging to the same group at a certain time have to be coded. The codes (Fig. 3) represent which group the sampling point belongs to.

As a starting point the westernmost group in LB always received -2 code, while the easternmost was $+2$. The groups the coding these two end points received integers or fractions (Fig. 3). Irrespective of whether the groups received a fractional or a whole number, the point was that the cluster analysis could identify them as different groups. By coding the groups the problem originating from the time-series character is solved as in the work of Neilson and Stevens (1986), who, instead of coding, used single (T) values.

After each group at every point in time received a code, these were clustered in time (1985–2004). The results of the cluster analysis conducted by using coding under consistent and precise execution can be considered as objective. As a result, time intervals with similar behavior were obtained when homogeneous sub-areas were to be determined (Fig. 4). There were in practice two time intervals with different behaviors, namely 1985–1997 and 1998–2004. Between 1985 and 1997 two patterns were present at the same time so we separated the interval into 1985–1997a and 1985–1997b.

3.3. Determining the similar sampling locations for the time intervals

As a final step, the points in time from the three time intervals (1985–1997a, 1985–1997b, and 1998–2004) were extracted and

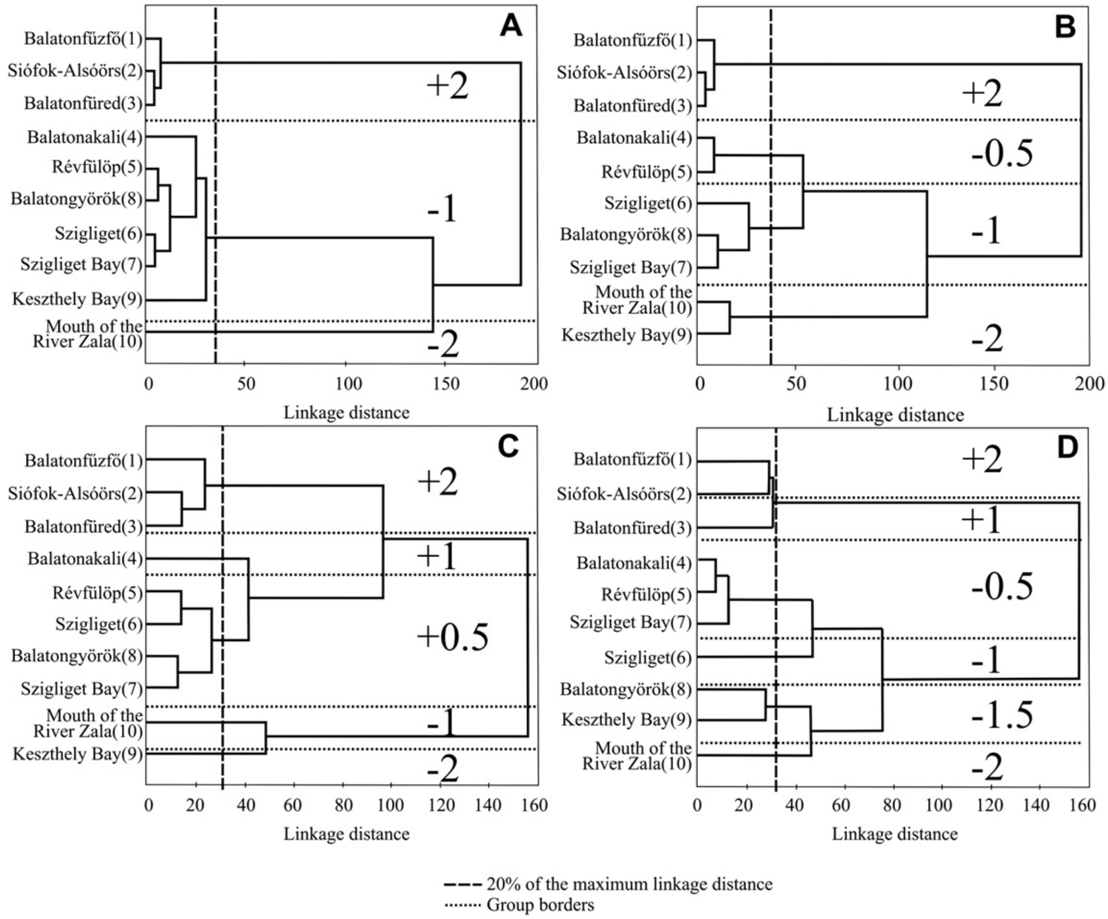

Fig. 3. Examples on the 85 clustered points in time and the coding applied to the sampling locations, 07.12.1987 A); 10.10.1988 B); 02.04.1991 C); 03.07.1995 D).

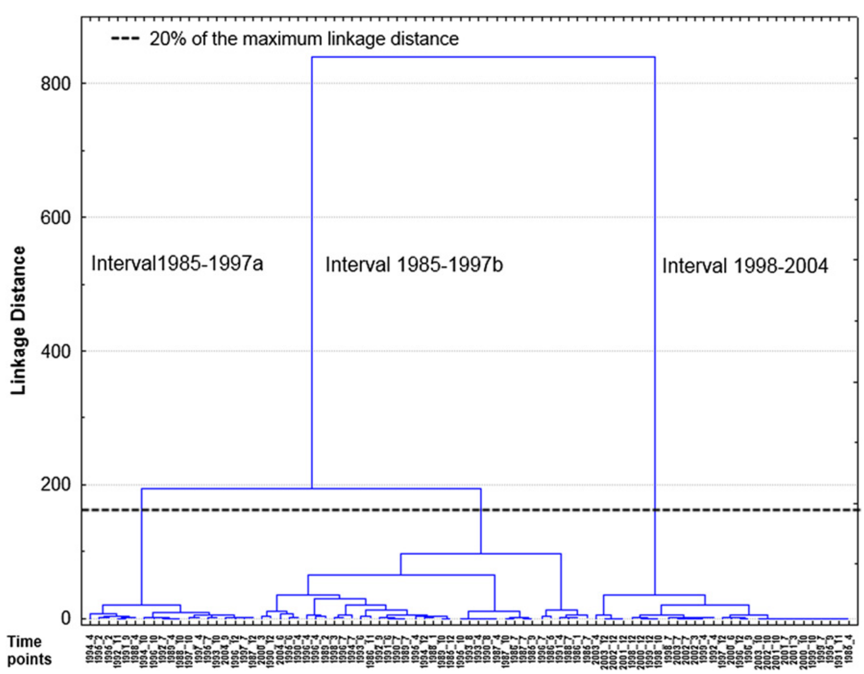

Fig. 4. Dendrogram of the three time intervals with similar behavior (1985–1997a; 1985–1997b and 1998–2004).

clustered. This time the cluster analysis was conducted on the sampling locations from the points in time belonging to the same time intervals. This led us to the final results. In this case, because of the coding, a "new dataset" was obtained and this permitted the use of a linkage distance which, using discriminant analysis proved to be correctly classified in 81.8% of the cases.

Two patterns were identifiable from 1985 to 1997 and one from 1998 to 2004 at a linkage distance of 10%, resulting in four sub-areas from 1985 to 1997, and three from 1998 to 2004 (Fig. 5).

Over the course of 1985–2004 five distinctive sub-areas were identifiable in LB's water body.

Regarding the time interval 1985–1997 the ecological crisis that the lake suffered can easily be recognized in the data. The effects of the 1991 fish extinction followed by the hypertrophy seemed to ease up until 1995 (Istvánovics et al., 2002, 2007). The most significant difference between Fig. 5A and B is the question of whether the two sampling locations in the middle of the lake connect to the eastern or western basin. Bearing these facts in mind, in order to be able to follow the eventual changes between sub-areas of the water body, we suggest a minimum of five sampling locations, each one in a different area (Fig. 6).

3.4. Verification of cluster results and determination of parameters which mostly influenced these groups

Verification of the cluster results by using discriminant analysis showed that the determined sub-areas in the time intervals 1985–1997a and b were in 91% and 84% correctly classified, respectively, while 77.6% of the classification for the time interval 1998–2004 is correct.

To find out which parameters determine the cluster groups the most, a Wilks' λ quotient was assigned to every parameter for every point in time. Grouping the c. 1900 quotients, parameters were separated by using CA, and two different groups were obtained according to how much they determine the original spatial cluster groups.

The parameters in *Group 1* were Alkalinity, pH, COD^C, COD^P, TP, Cond., Na^+; K^+; Ca^{2+}; Mg^{2+}; HCO_3^-; SO_4^{2-}; and Cl^-, which had an

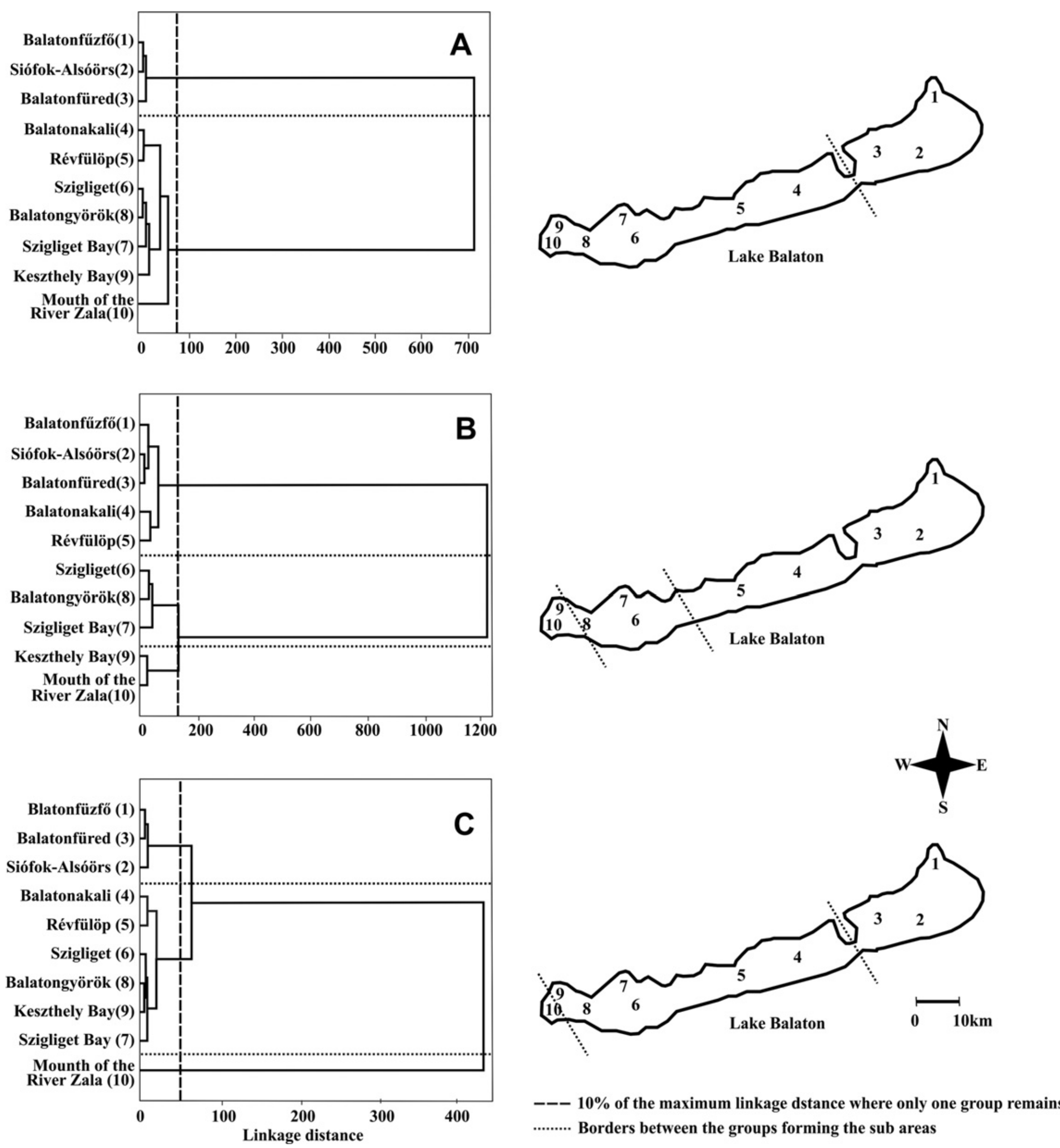

Fig. 5. Division of sub-areas for the time intervals with similar behavior, 1985–1997a A); 1985–1997b B); 1998–2004 C).

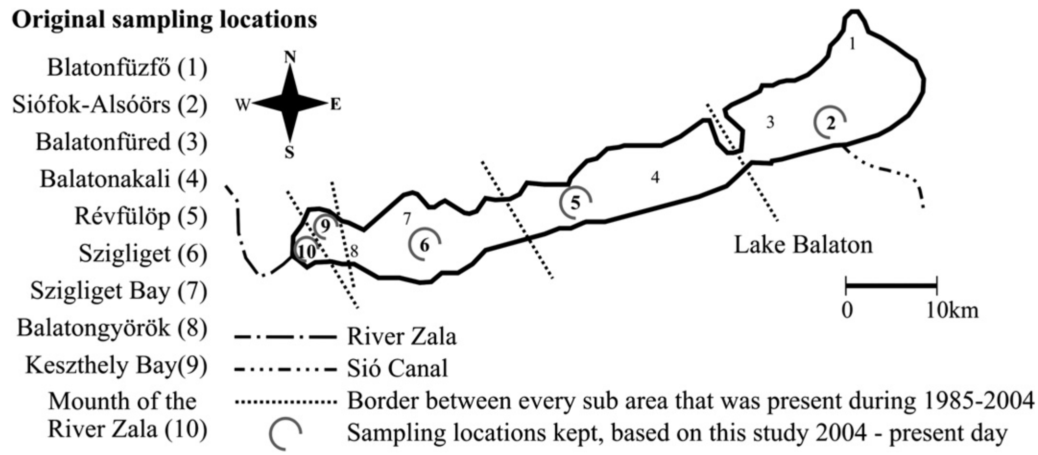

Fig. 6. Sub areas discoverable during 1985–2004 in Lake Balaton, and the five sampling locations retained by the Water Authority based on our study.

average quotient of 0.125. In *Group 2* the following parameters were found: TSS, BOD^5, NH_4–N, NO_2–N, Dissolved O_2, CaO, Chl-a, ON, PO_4–P, and W_{temp}, which had a quotient of 0.243.

Regarding the effect of the parameters on the result stated above, it can be said that the parameters which stand in close relation to eutrophication play a smaller role in determining the sub-areas in LB's water body than the inorganic ones. This should come as no surprise, because LB's water surface is in the main lacking in coverage and the opposite is witnessed in eutrophic and macrophyte covered waters (Hatvani et al., 2011).

4. Conclusions

According to the WFD, Lake Balaton currently consists of only one water body. During our research we were able to point out that preconceived sub-areas of the water body change dynamically over time, and on different scales this movement stands in close connection with the "recent" ecological events in the lake (e.g. mass fish deaths (Molnár et al., 1991), oligotrophic trends (Istvánovics et al., 2002, 2007; Hajnal and Padisák, 2008) etc.). In the first time period (1985–1997) two patterns were noticeable: one of two and one of three sub-areas. However, from 1998 only one pattern, namely that of three sub-areas, becomes apparent, thus the lake presents a uniform picture. Considering that the intensive changes in the lake's water quality described in the introduction ceased in 1998, the coincidence is conspicuous.

Data analysis for the 20 year time interval explicitly showed that the borders between the five sub-areas of LB's water body can move any time. At a given point in time the division of the whole water body into sub-areas is a key question in planning the monitoring of the lake. Namely, where the borders are expected to move, the spatial frequency of sampling should be increased.

The uniqueness of the study lies in the application of cluster analysis, because if no transformation is made on the dataset, then no variance loss is expected from the original data. As a result to what extent the general pattern obtained exists in the identified time intervals can be determined precisely.

Based on this study the Water Authority chose to keep five out of ten sampling locations so that the sub-areas could be described. We consider this as a great success and the methodology as an example for the setting up of sub-areas in a water body.

Acknowledgments

We the authors would like to thank Paul Thatcher for his work on our English version and say thanks for the help of Peter Tanos on the artwork. The work of I.G. Hatvani was funded by the "Habilitas" scholarship of the Hungarian Development Bank Plc. (MFB Zrt.) which we are very grateful for.

References

Afifi, A., Clark, V.A., May, S., Raton, B., 2004. Computer-aided Multivariate Analysis, fourth ed. Chapman & Hall/CRC, Boca Raton, FL.

Cansu, F.I., Özgür, E., Semra, I., Naime, A., Veyesel, Y., Seyhan, A., 2008. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environmental Monitoring and Assessment 144, 269–276.

CDWE, Central Directorate for Water and Environment, 2010. Watershed Management Plan, Domestic Implementation of the Water Framework Directive Watershed Management Plan. Hungarian Ministry of Rural Development, Budapest. http://www.vizeink.hu/files/Reszvizgyujto_VGT_Balaton_13.pdf.

Cooke, G.D., Welch, E.B., Peterson, S., Nichols, S.A., 2005. Restoration and Management of Lakes and Reservoirs, third ed. CRC Press, Boca Raton.

Cserny, T., Nagy-Bodor, E., 2000. Limnogeological investigations on Lake Balaton. In: Gierlowski-Kordesch, E., Kelts, K. (Eds.), Lake Basins Through Space and Time. AAPG Studies in Geology, vol. 46, pp. 605–618.

Cserny, T., 1993. Lake Balaton, Hungary. In: Gierlowski-Kordesch, E., Kelts, K. (Eds.), A Global Geological Record of Lake Basins. Cambridge University Press, Cambridge, pp. 397–401.

CWIIWH, 1987 Convention of International Importance especially as Waterflow Habitat. Ramsar (Iran), 2 February 1971. UN Treaty Series No. 14583. As Amended by the Paris Protocol, 3 December 1982, and Regina Amendments, 28 May 1987.

Dinesh, S., 2008. Characterization of influence zones of simulated droughts and floods of water bodies in a floodplain. Computer and Geosciences 34, 1791–1797.

Earle, R., Blacklocke, S., 2008. Master plan for water framework directive activities in Ireland leading to River Basin Management Plans. Desalination 226, 134–142.

Edmondson, W.T., 1991. The Uses of Ecology; Lake Washington and Beyond. University of Washington Press, Seattle.

Ferreira, J.G., Nobre, A.M., Simas, T.C., Silva, M.C., Newton, A., Bricker, S.B., Wolff, W.J., Stacey, P.E., Sequeira, A., 2006. A methodology for defining homogenous water bodies in estuaries – Application to the transitional systems of the EU Water Framework Directive. Estuarine, Coastal and Shelf Science 66, 468–482.

Hajnal, E., Padisák, J., 2008. Analysis of long-term ecological status of Lake Balaton based on the ALMOBAL phytoplankton database. Hydrobiologia 599, 259–276.

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. Ecological Engineering 37, 629–635.

Istvánovics, V., Somlyódi, L., Clement, A., 2002. Cyanobacteria-mediated internal eutrophication in shallow Lake Balaton after load reduction. Water Research 36, 3314–3322.

Istvánovics, V., Clement, A., Somlyody, L., Specziar, A., G.-Toth, L., Padisak, J., 2007. Updating water quality targets for shallow Lake Balaton (Hungary), recovering from eutrophication. Hydrobiologia 581, 305–318.

Ketskeméty, L., Izsó, L., 2005. Introduction to the SPSS Software. Methodical Guide and Task Collection for Statistical Analysis. Eötvös Kiadó, Budapest.

Korponai, J., Varga, K.A., Lengré, T., Papp, I., Tóth, A., Braun, M., 2011. Paleolimnological reconstruction of the trophic state in Lake Balaton (Hungary) using Cladocera remains. Hydrobiologia 676, 237–248.

Kovács, J., Hatvani, I.G., Korponai, J., Kovácsné, Sz. I., 2010. Morlet wavelet and autocorrelation analysis of long term data series of the Kis-Balaton Water Protection System (KBWPS). Ecological Engineering 36, 1469–1477.
Kovács, J., Tanos, P., Korponai, J., Kovácsné, Sz.I., Gondár, K., Gondár-Sőregi, K., Hatvani, I.G., 2012. Analysis of water quality data for scientists. In: Voudouris, V., Voutsa, D. (Eds.), Water Quality and Water Pollution: Evaluation of Water Quality Data. In Tech Open Access Publisher, Rijeka, pp. 65–94.
László, B., Szilágyi, F., Szilágyi, E., Heltai, Gy., Licskó,, I., 2007. Implementation of the EU Water Framework Directive in monitoring of small water bodies in Hungary, I. Establishment of surveillance monitoring system for physical and chemical characteristics for small mountain watercourses. Microchemical Journal 85, 65–71.
Lotz, Gy, 1988. A Kis-Balaton Vizvedelmi Rendszer. Hidrologiai Tajekoztato (The Kis-Balaton Water Protection System. Hidrologycal Bulettin). Oktober 20–22.
Molnár, K., Székely, Cs., Baska, F., 1991. Mass mortality of eel in Lake Balaton due to *Anguillicola crassus* infection. Bulletin of the European Association of Fish Pathologists 11, 211.
Muxika, I., Borja, Á., Bald, J., 2007. Using historical data, expert judgment and multivariate analysis in assessing reference conditions and benthic ecological status, according to the European Framework Directive. Marine Pollution Bulletin 55, 16–29.
Neilson, M.A., Stevens, R.J.J., 1986. Determination of water quality zonation in Lake Ontario. Developments in Water Science 27, 99–116.
Padisák, J., Borics, G., Grigorszky, I., Soroczki-Pintér, E., 2006. Use of phytoplankton assemblages for monitoring ecological status of lakes within the Water Framework Directive: the assemblage index. Hydrobiologia 553, 1–14.
Padisák, J., 1997. *Cylindrospermopsis raciborskii* (Wołoszynska) Seenayya et Subba Raju, an expanding, highly adaptive cyanobacterium: worldwide distribution and review of its ecology. Arch. Hydrobiologia 107, 563–593.
Pálffy, K., Vörös, L., 2011. A fitoplankton diverzitása a Balatonban: a fajösszetételtől a funkcionális diverzitásig (Fitoplankton diversity in lake Balaton). Ecology of lake Balaton/A Balaton Ökológiája 1, 61–75.
Parpala, L., G.-Tóth, L., Zinevici, V., Németh, P., Szalontai, K., 2003. Structure and production of the metazoan zooplankton in Lake Balaton (Hungary) in summer. Hydrobiologia 506, 347–351.
Sagehashi, M., Sakoda, A., Suzuki, M., 2001. A mathematical model of a shallow and Eutrophic Lake (The Keszthely Basin, Lake Balaton) and simulation of restorative manipulations. Water Research 35, 1675–1686.
Sánchez, D., Carrasco, F., Andreo, B., 2009. Proposed methodology to delineate bodies of groundwater according to the European Water framework directive. Application in a pilot Mediterranean river basin (Málaga, Spain). Journal of Environmental Management 90, 1523–1533.
Simenov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M., Kouimtzis, Th, 2003. Assessment of the surface water quality in Northern Greece. Water Research 37, 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 Research 38, 3980–3992.
Sprőber, P., Hesham, M.S., Présing, M., Kovács, A.W., Herodek, S., 2003. Nitrogen uptake and fixation in the cyanobacterium *Cylindrospermopsis raciborkii* under different nitrogen conditions. Hydrobiologia 506-509, 169–174.
Stockburger, W.D., 2001. Multivariatestatistics: Concepts, Models, and Applications. Missouri State University.
Tátrai, I., Kálmán, M., Korponai, J., Paulovits, G., Pomogyi, P., 2000. The role of the Kis-Balaton water protection system in the control of water quality of lake Balaton. Ecological Engineering 16, 73–78.
U.S. Environmental Protection Agency (USEPA), 2009. National Lakes Assessment: a Collaborative Survey of the Nation's Lakes. EPA 841-R-09–001. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C.
Vörös, L., Somogyi, B., 2009. A Balaton algaegyüteseinek szerepe a tó vízminőségének alakításában (The role of Lake Balaton's algae composition in water quality). In: Bíró, P., Banczerowsky, J. (Eds.), A Balaton kutatásának 2008. évi eredményei (The results of the research of Lake Balaton in the year 2008). Hungarian Academy of Sciences, Budapest, pp. 7–16.
WFD, 2000. 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. PE-CONS 3639/1/00 REV 1.
Zeng, X., Rasmussen, T.C., 2004. Multivariate statistical characterization of water quality in lake Lanier, Georgia. Journal of Environmental Quality 34, 1980–1991.