Multivariate exploration and classification applied to the chemical composition of spring waters in sanctuary forest areas

Aleksander Astel[a]; Stanisław Małek[b]; Katarzyna Krakowian[b]

[a] Environmental Chemistry Research Unit, Biology and Environmental Protection Institute, Pomeranian Academy, Słupsk, Poland [b] Department of Forest Ecology, Forest Faculty, Agricultural University of Cracow, Cracow, Poland

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

Multivariate exploration and classification applied to the chemical composition of spring waters in sanctuary forest areas

Aleksander Astel[a]*, Stanisław Małek[b] and Katarzyna Krakowian[b]

[a]*Environmental Chemistry Research Unit, Biology and Environmental Protection Institute, Pomeranian Academy, Słupsk, Poland; [b]Department of Forest Ecology, Forest Faculty, Agricultural University of Cracow, Cracow, Poland*

This study was carried out in the Zimnik and Czyrna catchments (ZC and CC, respectively) located in the Silesian Beskid Mountains of southern Poland. The data matrix with 870 observations of major inorganic elements (Cl^-, NO_3^-, SO_4^{2-}, NH_4^+, Na^+, K^+, Ca^{2+} and Mg^{2+}), pH and electrolytic conductivity of spring water samples was carried out using linear discriminant analysis, unsupervised Kohonen self-organising maps (SOM) classification and non-parametric tests. A multivariate analysis of the chemical composition of spring water was performed, taking into consideration the geological condition of the catchment area, the prevailing forest stands, slopes facing and different water levels due to the seasonal variations. For the ZC the first discriminant function (DF) reflected general geological conditions (absence of NH_4^+) while the second DF reflected the nutrient biocycle (presence of NH_4^+). In case of the CC, the first DF had a complex meaning and reflected both geological conditions and the nutrient biocycles. Based on specific chemical profiles, low water level related samples and high water level related samples were distinguished from each other in both the ZC and CC, respectively. The SOM-based classification indicated that forest types and springs location were the major factors affecting the spring water chemical profile. In general, in case of springs located above 1000 m, limited weathering connected with the effect of the highly acid spruce monoculture on leaching causes a decreasing concentration of Na^+, K^+, Mg^{2+}, Cl^- and SO_4^{2-}, while springs located between 700 and 800 m above sea level in mixed forests, showed the highest concentrations of K^+ and Cl^- and the lowest concentrations of NO_3^-, values of pH and conductivity.

Keywords: mountain catchments; Silesian Beskid; spring water; linear discriminant analysis; self-organising maps

1. Introduction

An assessment of the surface water (spring, river or lake) quality is usually based on the comparison of measured values of particular physico-chemical parameters such as pH, conductivity, salinity and the concentration of those chemical components with threshold values defined by national or international legislation. A more sophisticated approach is the application of multivariate statistics for the classification and interpretation of water reservoirs, since they consider the environment to be multivariate [1–4].

*Corresponding author. Email: astel@apsl.edu.pl

Usually, in case of surface water reservoirs, the assessment of the water quality or the optimisation of the monitoring by multivariate statistical methods includes classification of sampling locations or quality parameters, identification of possible sources of pollution and modelling the contribution of the identified source to the formation of the total concentration of the monitored chemical tracers. Up to now there have been only a few trials specified directly towards spring water reservoirs. Swanson et al. [5] applied two-way cluster analysis in order to determine the geochemical characteristics and apparent ages of sampled groundwaters to determine and thereby identify which of the two regionally extensive bedrock aquifers was the more likely source of water discharging to springs in the Nine Springs watershed. Cluster analysis, principal components analysis (PCA) and linear discriminant analysis (LDA) were applied successfully by Šnuderl et al. [6] to classify 50 samples of natural mineral waters from springs in Slovenia, Hungary, Germany, Czech Republic and other countries from former Yugoslavia. An optimised model of multivariate classification for the monitoring of 18 spring waters adjacent to Serra St. Bruno (Calabria, Italy) based on cluster analysis and partial least squares-discriminant analysis was developed by Ragno et al. [7]. However, none of these studies dealt with the spring waters located in sanctuary forest areas.

The construction of a geological profile, including the type of rock and the type of forest humus all influence the chemical profile of forests discharge waters. The appearance of varying forest sites makes unequivocal identification and explanation of the impact of a forest stand type or a geological constitution on the quality of outflow waters problematic. It was found that, in mountain area, a dominant factor responsible for the chemical profile of the outflow water relates to the washing out of minerals weathered from the rock foundation [8–11]. Additionally, a significant role is played by biological processes in the soil layer or in the stream water [12]. The chemical constitution of spring waters strongly depends on the type of soil and geological profile, as well as on the extent and nature of plant cover, the surface configuration and even the facing of slopes [13–16]. Among the factors that determine the quality and quantity of inorganic compounds in surface waters, the seasonal variability of water level in the catchment also plays a significant role [17]. Since 1997, intensive studies on mountain forest ecosystem of the Dupniański catchment of the Vistula Forest District, in the Silesian Beskid Mountains have been carried out by a research team from the Forest Ecology Department, University of Agriculture in Cracow [15,16]. In most cases, the objectives of the studies focussed on the assessment of environmental growth conditions for the Istebna spruce ecotype, the identification needs and the threats to forest sustainability. Quantitative and qualitative analyses of spring and surface water samples were carried out to produce physical and chemical balance assessment. In general, the surface waters collected from the Zimnik and Czyrna catchments (ZC and CC, respectively) indicated different chemical composition in comparison to waters collected in more remote sites of the Silesian Beskid Mountains [12,18]. However the analyses on which these conclusions were based, failed to account for the geological conditions of the catchment area, the prevailing forest stands or the seasonal variability of water level. In view of these failings, it was considered advisable to reconsider the available data using multivariate environmetrics in order to verify following hypotheses:

- the chemical composition of spring water from the ZC and CC is determined by the physico-geological constitution of the drainage areas, modified by seasonal changes of water level in the catchment,

- application of multivariate environmetrics might be useful in exploration of ways in which novel statistical methods can reveal further information on the factors influencing chemical composition of spring water (e.g. geological conditions, prevailing forest stand, spring's location).

2. Experimental

2.1 *Study area description and sampling*

The specific structure of soil and geological sediments in the southern part of the Silesian Beskid Mountains ($49°34'$N, $18°50'$E) causes differences between the ZC and CC. In the research area sandstone layers are bedded diagonally, which means the top (polygenetic conglomerates and thin bedded flinty sandstone with green and dark shales forming Haplic Podzols on ZC) and bottom layers (thin bedded glauconitic and flinty sandstone with green shales forming Haplic Podzols and Albic-Dystric Cambiosols on CC) meet the surface. In case of investigated area the geology has an impact not only on water chemistry, but also on the location of the springs. In the CC area springs are often located at the intersection of the water-bearing horizon, sometimes caused by forest road building. These waters include much more rock-mantle, which is partially reflected in higher electrolytic conductivity (COND) (average COND on CC is $7.3\,\mathrm{mS\,m^{-1}}$ while on ZC $6.7\,\mathrm{mS\,m^{-1}}$). During intensive rainfall the difference in COND between the catchments increases significantly reaching 10.4 and $7.5\,\mathrm{mS\,m^{-1}}$ for CC and ZC, respectively.

The geological structure in ZC is more homogeneous, and the spring's locations are more stabile [16]. The geographical location of the catchments investigated and the distribution of dominant soil types are presented in Figure 1(a), while the hydrological network as well as the character of the sediment's is presented in Figure 1(b). The essential difference between ZC and CC is related to the degree of weathering of surface due to different geology and soil cover. In the CC area the soils originate from dust clay and loam, while in ZC they are of sand, gravel and stones. Perhaps this is the reason why in the CC brown soil covers a much bigger surface than ZC (179 ha in comparison to 22 ha). Differences are also noticeable concerning stony soils – 14 ha in CC and 28 ha in ZC. Furthermore, the CC is more touristically developed (ski trails) than the ZC.

The vegetation close to the springs is spatial variable. The location of the springs sampled and the distribution of the dominant forest formations and their species composition is presented in Figure 2(a) and in Figure 2(b), respectively. In the area of ZC, mixed coniferous mountain forest reaches the peak. In the CC area, mixed coniferous forest reaches at 1050 m above sea level (a.s.l) and only mountain coniferous forest was observed any higher. The detailed characteristics of springs is summarised in Table 1, while in Table 2 a comparison of selected physiographical parameters of ZC and CC taken from hydrologic maps dated for 2002 and from the field research is shown.

Field work was carried out during three measurement campaigns in 2004 at Zimnik (63 samples) and CC (51 samples) above the peasant holdings. One sample was collected from the Regional Subsurface Water Monitoring network (RMWP) located in the area of the ZC. Samples were collected at times of high (May 2–4), medium (June 18–20) and low (15–17 October 2004) water level in the catchment. The sampling days were chosen subjectively but are representative because of interannual consistency in long-term records [19]. This is why the results obtained during 2004 represent a common situation in each catchment.
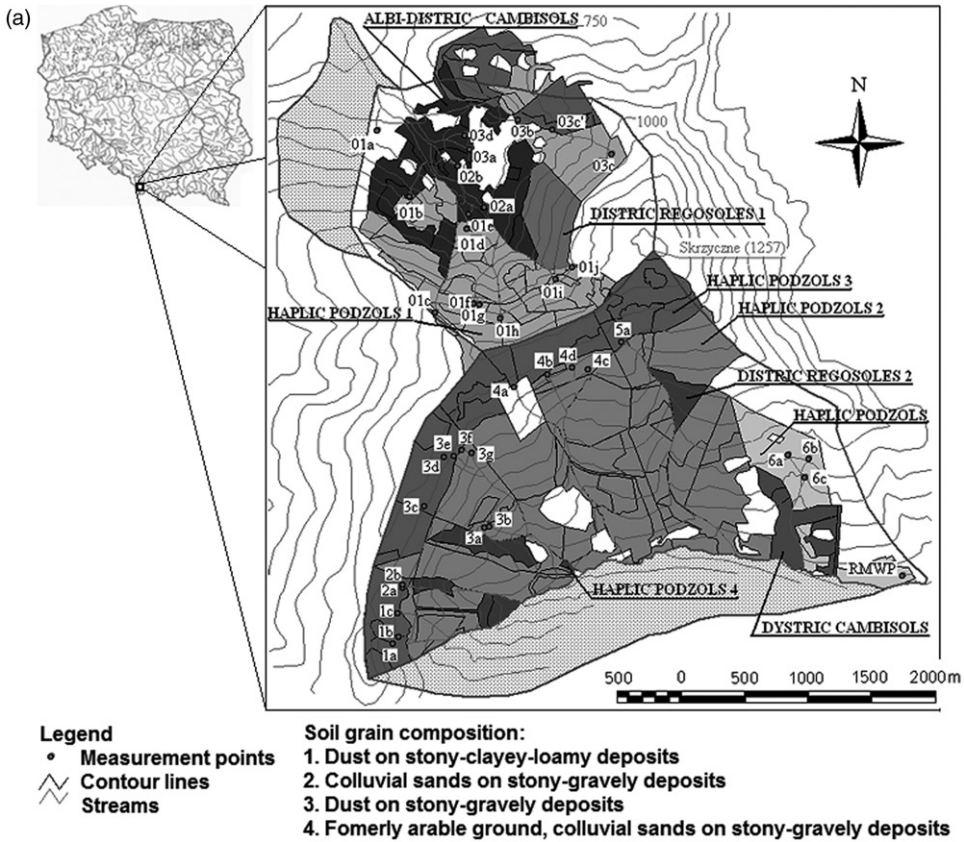
Figure 1. (a) Geographical location of the ZC and CC and the distribution of the prevailing soil types. (b) Hydrological network and types of geological sediments at the area of the ZC and CC.

As presented in Table 2, the crenologic indexes indicated a slight difference between the density of springs in the consecutive years, mainly for ZC. This is because the location of springs may vary annually. The annual variation in the location of springs corresponds to the annual variation in rainfall intensity. Twice in the case of ZC two springs changed their location in comparison to the previous sampling session. For CC it was one and two springs, respectively.

2.2 *Analytical techniques*

Water samples were analysed using ion chromatography (Dionex-320, Sunnyvale, CA, USA) to determine the concentration of: Cl^-, NO_3^-, SO_4^{2-}, NH_4^+, Na^+, K^+, Ca^{2+}, Mg^{2+} according to the current PN-ISO 10304-1:1998 [20] and PN-ISO 14911-1:1998 [21] standards. The ion chromatograph was coupled with double piston GP40 IC Pump, CD20 IC Conductivity Detector and Dionex PeakNet (ver 5.11) software.

The analysis of anions in the calibration solution and spring water samples were carried out using a Dionex IonPac AS17 Analytical Column (250×2.0 mm i.d.) with IonPac AS17 Guard Column (50×2.0 mm i.d.) and ASRS® ULTRA Anion
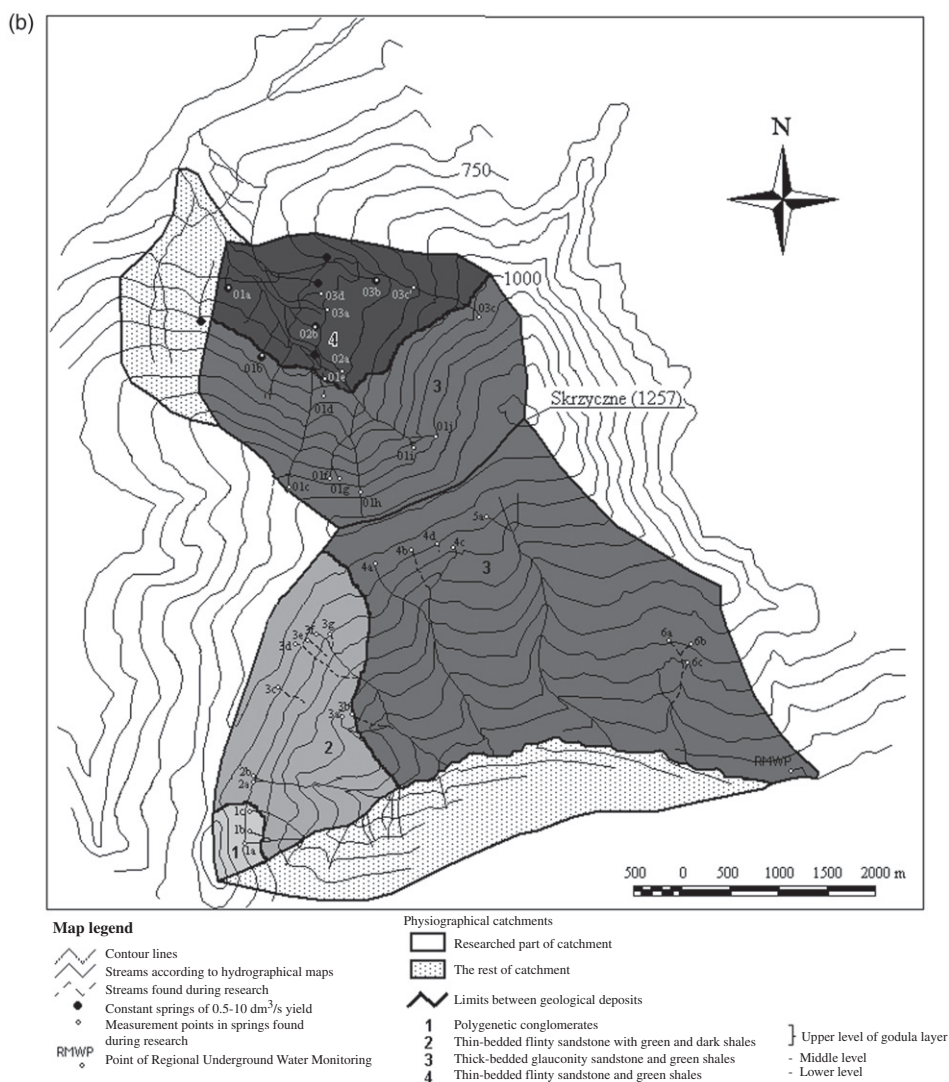
Figure 1. Continued.

Self-Regenerating Suppressor. The mobile phase was a mixture of 3.5 mmol L^{-1} Na$_2$CO$_3$ and 1.0 mmol L^{-1} NaHCO$_3$ under flow rate of 0.25 mL min^{-1}. The column temperature was 30°C and the pressure was 6.20 MPa. Both in cation and anion mode, injection was accomplished by a dose loop (15 μL) operating at room temperature.

The analyses of cations were carried out using an IonPac CS12A (250 × 2.0 mm i.d.) analytical column with an IonPac CG12A Guard Column (50 × 2.0 mm i.d.) and CSRS® ULTRA Auto Suppressor working in recycle mode. The mobile phase was 18 mM methanosulphonic acid at a flow rate of 1.0 mL min^{-1}. The column temperature was 30°C and its pressure was 11.37 MPa. A low-pH acid rain sample from southern Ontario (Canada), RAIN.97 and No 409 served as certified reference materials (CRM) were
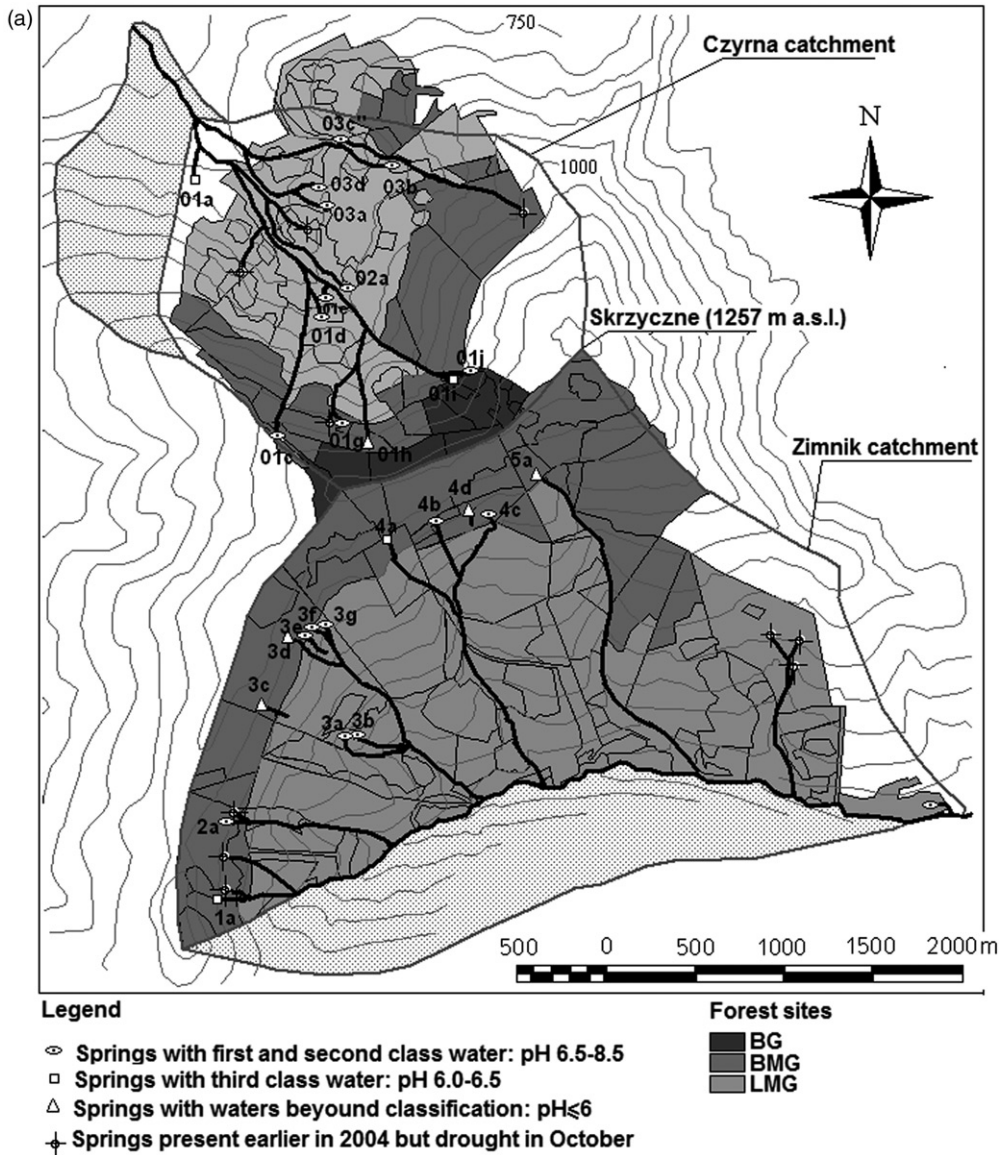
Figure 2. (a) Geographical location of springs under investigation, and the distribution of prevailing forest formations (BG – mountain coniferous forest (spruce monoculture); BMG – mixed mountain coniferous forest; LMG – mixed mountain forest). (b) Composition of forest stands ranked according to the species variability in the ZC and CC.

analysed simultaneously. A detailed description of the analytical technique, its calibration and validation based on CRM can be found elsewhere [13].

Both, relative uncertainties and expanded uncertainties were calculated according to the formulae presented by Konieczka *et al.* [22], as well as with the scheme presented by Tepuš and Simonič [23]. The following relative uncertainties for the average elemental content were obtained: $Cl^- - 0.4\%$, $NO_3^- - 0.4\%$, $SO_4^{2-} - 0.5\%$, $NH_4^+ - 0.3\%$, $Na^+ - 0.4\%$,
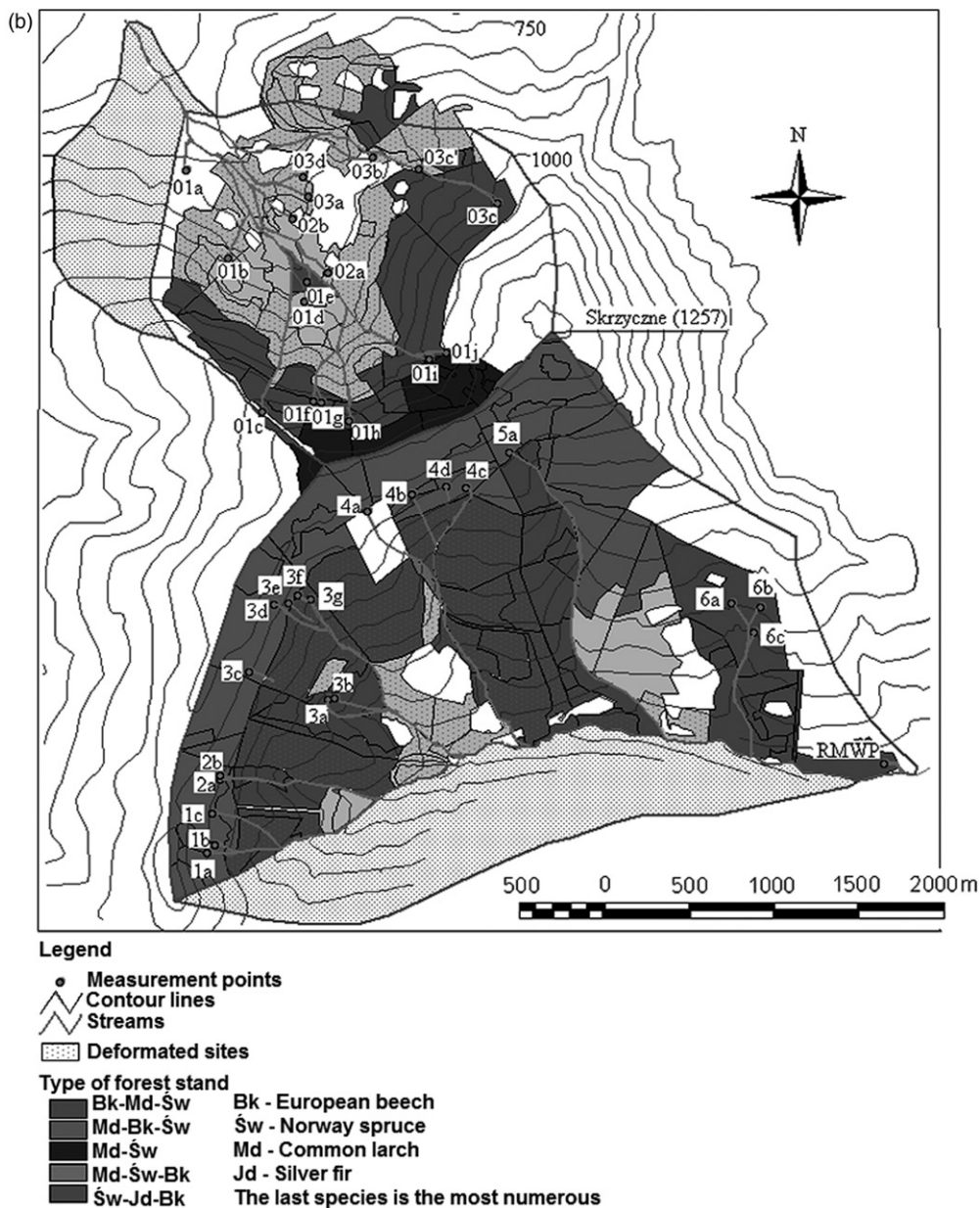
Figure 2. Continued.

$K^+ - 0.5\%$, $Ca^{2+} - 0.4\%$, $Mg^{2+} - 0.3\%$ and were comparable with relative uncertainties calculated earlier by Röhker *et al.* [24] and Miskaki *et al.* [25] for the determination of routine ions in potable water. The calculated expanded uncertainties were equal to 4.8, 6.2, 5.5, 5.8, 4.3, 8.5, 6.2 and 6.4% for Cl^-, NO_3^-, SO_4^{2-}, NH_4^+, Na^+, K^+, Ca^{2+} and Mg^{2+} ions, respectively. Determined values were comparable with values published before by others [23,25].

Table 1. Detailed characteristics of springs concerning elevation, facing of slopes, forest stand composition and age of the dominant tree species.

Spring	Elevation [m a.s.l.]	Exposition	Main species composition	Age of the main species [years]	Other characteristics
Zimnik catchment					
1a	1050	E	Md–Bk–Św	130	Bog-spring located near to the boundary of an old spruce stand with forest crops
1b	1020	E	Md–Bk–Św	120	
1c	1020	E	Md–Bk–Św	120	
2b	1020	E	Md–Bk–Św	120	
2a	1050	E	Md–Bk–Św	120	
3a	880	SE	Św–Jd–Bk	110	A beech stick stand with maple admixture surrounds
3b	880	SE	Św–Jd–Bk	110	
3c	1100	SE	Md–Bk–Św	105	Peatmoss-grown
3d	1120	SE	Md–Bk–Św	105	
3e	1080	SE	Md–Bk–Św	105	
3f	1075	SE	Św–Jd–Bk	110	An outflow from a stone rubble
3g	1070	SE	Św–Jd–Bk	110	
4a	1130	SE	Md–Bk–Św	100	
4b	1110	SE	Md–Bk–Św	5	Located within 5-year-old spruce and pine crops limits
4d	1100	SE	Md–Bk–Św	5	
4c	1060	SE	Md–Bk–Św	115	A vast bog-spring
5a	1110	SW	Św–Jd–Bk	190	
6a	810	S	Św–Jd–Bk	65	
6b	810	S	Św–Jd–Bk	65	
6c	740	S	Św–Jd–Bk	65	
RMWP	585	S	Św–Jd–Bk	75	

Czyrna catchment

01a	730	NE	Md–Św–Bk	95	Discovered in multispecies plantings near to pastures, herbaceous vegetation strongly developed
03d	755	NW	Md–Św–Bk	70	
01b	775	NE	Md–Św–Bk	95	
01c	1120	NW	n.a.	n.d.	A muddy pool in the middle of the ski trail
01d	825	N	n.a.	n.d.	A very vast bog-spring located along the boundary between a pasture and a spruce forest
01e	775	NE	Bk–Md–Św	45	A bog-spring
01f	1075	N	Bk–Md–Św	3	
01g	1125	N	Bk–Md–Św	3	Located on vast spruce and mountain pine crops, trickled slowly through a hygrophilous plants patch
01h	1075	N	Md–Św	3	Located on vast spruce and mountain pine crops, an outflow from stone rubble
01i	1070	NW	Bk–Md–Św	60	An outflow straight from a bedrock (about 5 metre high wall of sandstone)
01j	1100	NW	Md–Św	55/140	
02a	795	W	Md–Św–Bk	130	The highest flow rate, an outflow from beneath a big stone
02b	730	W	n.a.	n.d.	A bog-spring
03a	750	NW	Md–Św–Bk	70	
03b	775	NW	Bk–Md–Św	100	A water intake, hydrophilous vegetation with a stand about 30-years old-beech and maple mixture
03c	1010	NW	Bk–Md–Św	10/100	A part of a streambed drought
03c′	845	SW	Md–Św–Bk	15	

Notes: Bk – European Beech – *Fagus sylvatica* L.; Md – European Larch - *Larix decidua* Mill; Św – Norway Spruce – *Picea abies* Karst; Jd – Fir– *Abies alba* Mill; RMWP – sampling site belonging to regional subsurface water monitoring network; n.a. – not afforested; n.d. – not determined.

Table 2. A comparison of selected physiogeographical parameters of the ZC and CC on hydrological maps dated for 2002, and those used for the research.

Feature	ZC	CC
Catchment length (km)	4.73	2.32
Location of the highest spring (m a.s.l)	1130	1210
Hydrographical maps 2002	1130	1125
The research		
River network density according to Neuman (km km^{-2})		
Hydrographical maps 2002	1.99	2.91
The research – maximum	2.07	2.57
The research – minimum	1.85	2.08
Springs density index (crenologic index) (amount km^{-2})		
Hydrographical maps 2002	1.73	3.87
The research – all springs	2.79	3.86
The research – springs occurring in all seasons	1.86	3.14

2.3 *Environmetric techniques*

Advanced multivariate statistical approaches contribute to a better understanding of the data collected by long-term monitoring. Nevertheless, careful specification of the objectives of the particular research should be done before selecting a suitable data mining technique. To find a relatively small number of factors from a data set of many correlated variables or interpret hidden, complex and casually determined relationships between features and objects a PCA is often applied [26–29]. On the other hand, when statistical assessment of the differences between mean values in *a priori* defined groups is the objective, an analysis of variance (ANOVA) offers many advantages.

Keeping in mind the above mentioned, and due to the specific PCA and ANOVA constrains, a linear discriminant function (DF) analysis (LDA) was chosen to test if the chemical composition of springs located in ZC and CC depends on the physico-geological constitution of their drainage areas and catchments, considering seasonal variability of water level in the catchment. LDA enables the determination of the variables responsible for discriminating between two or more expected, naturally occurring groups in the data [30,31]. There are three different modes of LDA: standard, forward stepwise and backward stepwise. In standard LDA, the discrimination model is built using the complete set of variables. In forward stepwise LDA analysis, the discrimination model is built step-by-step. Specifically, at each step all variables are reviewed and evaluated to determine which one will contribute most to the discrimination between groups. That variable will then be included in the model, and the process starts again. In the backward stepwise LDA all variables are included in the model and then, at each step, the variable that contributes least to the prediction of group membership is eliminated.

Thus, as the result of a successful LDA analysis, one would only keep the most 'important' variables in the model, that is, those variables that contribute the most to the discrimination between groups. Usually, both forward and backward stepwise modes of LDA are used to determine if groups differ with regard to the mean of the most differentiating variables, and then use these selected variables to predict group member-ship (e.g. in new samples). The investigation presented here is not focussed on the development of prediction rules and hence standard LDA was applied. LDA has been

widely described before elsewhere and so for more details readers are referred to the literature [27,32–42].

It can be concluded that as with PCA, LDA enables, the compression of a data set dimensionality in the features' space, by the attribution of scientific explanation to the identified DFs. On the other hand, like ANOVA, it enables statistical assessment of the contribution of independent variables to the discrimination of *a priori* defined groups. LDA complements PCA and ANOVA, and hence was chosen as the most suitable tool for testing the first hypothesis presented above. Moreover, the main reason for our choice is the successful application of LDA by others [5–7,43].

The second major hypothesis, which assumes that application of advanced multivariate exploration techniques (especially unsupervised), enables the discovery of impacts of other factors on chemical profile of spring water was tested by an application of self-organising map (SOM)-based classification.

A SOM algorithm, proposed by Kohonen [44], is a neural-network model that implements a characteristic non-linear projection from the high-dimensional space of sensory or other objects onto a low-dimensional array of neurons [45]. The term 'self-organising' refers to the ability of learning and organising information without being given the associated-dependent output values for the input pattern [46]. Similarly, as LDA SOM has been widely described before elsewhere and so for more details readers are referred to the literature [44–49,51]. For the SOM algorithm, there are no precise rules for choosing the various parameters (e.g. global topology type (sheet, cylinder, toroid), local topology type (hexagonal, rectangular), map dimension, neighbour functions, etc.) [47]. In this study, the Kohonen map was chosen as a rectangular grid with the number of nodes (n) determined from the following formula: $n = 5 \cdot \sqrt{\text{number of springs}}$ [48]. Furthermore, a hexagonal lattice of the map was preferred since it does not favour horizontal or vertical directions [47]. Subsequently, in the ZC the dimensionality of Kohonen's map was determined as 4×9 ($n = 5 \cdot \sqrt{43} = 32.8$) while in CC the dimensionality was determined as 5×7 ($n = 5 \cdot \sqrt{44} = 33.2$). Basically, the two largest eigenvalues of the training data were calculated, and the ratio between the side lengths of the map grid was set to the ratio between the two maximum eigenvalues. The actual side lengths were then set so that their product was close to the determined number of map units. Similar conditions were applied in the study of Park *et al.* [49]. In this study, in clustering step, the non-hierarchical K-means algorithm was applied [50]. Different numbers of predefined clusters were tried, and finally, the best classification with the lowest Davies–Bouldin (DB) index was chosen [51].

Prior to chemometric analysis, all samples were checked in order to fulfil the criteria of ionic balances and similar salinity. Twenty samples collected on the area of ZC and seven samples collected in the area of CC did not fulfil rules mentioned above and were excluded from analysis leaving a final data matrix of 44 samples for ZC and 43 for CC for evaluation.

Standard LDA based on sequential Wilk's Lambda as well as SOM analysis were performed on a complete (10 variables × 44 samples = 440 results for ZC; 10 variables × 43 samples = 430 results for CC), raw data set tested successfully for normality by the use of Shapiro–Wilk's test [52]. LDA was independently applied to determine variables, which distinguish possible groups corresponding to water level (high, medium or low) in the catchments of Zimnik and Czyrna. Both in case of ZC and CC two significant DFs were obtained.

Throughout the study the commercial statistics software packages SPSS 16.0 Evaluation, Matlab® and Statistica® 8.0 (StatSoft, Inc.) running on Windows VISTA platform were used for calculation and visualisation purposes.

3. Results and discussion

Descriptive statistics calculated for the inorganic content of spring water samples collected at Zimnik and Czyrna research catchments are depicted in Table 3. The summary of the standard LDA for Zimnik ($\lambda = 0.03986$, $F = 12.426$, $p < 0.05$) as well as for Czyrna ($\lambda = 0.02078$, $F = 17.218$, $p < 0.05$) is shown in Table 4. The statistics, also presented in Table 4, illustrate the contribution of variables to the LDA model.

Decreasing the contribution of the variables to the separation of the spring water samples separation at ZC is given by following sequence: $Cl^- > NO_3^- > NH_4^+ > COND > pH > Ca^{2+} > Mg^{2+} > Na^+ > K^+ > SO_4^{2-}$. The value of tolerance (R^2) and $1 - R^2$, respectively, deliver information of the correlation of the respective variable, with all other variables included in the model. The most correlated ($R^2 > 0.5$) variables were Ca^{2+} and pH, while the least correlated ($R^2 < 0.5$) variables were K^+, Mg^{2+}, Na^+, NH_4^+, SO_4^{2-}, NO_3^- and COND. Similarly to ZC, for CC the highest variable contributions to the model were delivered by Ca^{2+}, NH_4^+, pH, K^+, COND and NO_3^-, while all other variables have similar contributions. SO_4^{2-}, NH_4^+ and pH were included in the group of containing the most redundant variables ($1 - R^2 > 0.5$). The standardised canonical DF coefficients (variables ordered by absolute value of correlation with the first DF) for the catchments investigated are presented in Table 5.

For ZC the first DF was mostly associated with Cl^-, NO_3^-, Na^+, K^+, Mg^{2+}, Ca^{2+} and pH, and was characterised by an eigenvalue of 6.52, a canonical correlation of 0.93 and explaining 74% of the variance. The second DF was characterised by an eigenvalue of 2.33, and a canonical correlation of 0.84. It explained 26% of the total variance being mostly associated with NH_4^+, Mg^{2+} and COND, as well as with ions as NO_3^- and K^+, previously associated with the first DF.

The general explanation of individual DFs was proposed according to the pattern of elements' association. The first DF reflected general geological conditions (absence of NH_4^+) and allowed the separation of low water level related samples from others, while the second DF reflected the nutrient biocycle (presence of NH_4^+), and separated out the high water level related samples. It has to be emphasised that the NO_3^-, K^+ and Mg^{2+} variation were probably partially related to the geological conditions and their involvement in nutrient biocycles (especially in the case of NO_3^-) [12–16]. Standardised canonical DF coefficients (Table 5) indicated that together with increasing water level, the concentrations of Cl^-, Na^+, Ca^{2+} as well as pH increased. Such phenomena are in opposition to the observations of Maciaszek and Zwydak for lower Istebna sandstone leaching dynamics [53], and could be explained by spring-time thawing being supreme of weathering processes. Standardised canonical DF coefficients also indicated that, in comparison to low and medium water samples, the NH_4^+, NO_3^- and Mg^{2+} concentrations increased greatly in the case of high water level related samples. High water levels corresponded to initial phases of the vegetation succession, and such observations are in agreement with annual nitrogen changes in surface water [54], and with the results published by Szczęsny and Zięba [17]. In the middle and at the end of the vegetation

Table 3. Descriptive statistics of concentrations of Na^+, NH_4^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, NO_3^-, SO_4^{2-} ($\mu g\,L^{-1}$), pH and electrolytic conductivity ($mS\,m^{-1}$) in spring water samples collected at Zimnik and Czyrna research catchments, split on low, medium and high water levels.

	N	Water level	ZC					N	CC				
			Mean	Median	Min	Max	SD		Mean	Median	Min	Max	SD
Na^+	18	Low	39^{ab}	37	8	69	14	15	43^{ab}	37	19	91	19
NH_4^+			3^b	2	2	11	2		8^b	2	2	23	7
K^+			10^{ab}	10	6	17	3		14^b	12	6	28	7
Ca^{2+}			305^a	308	48	581	130		323^b	326	82	550	152
Mg^{2+}			61^b	57	24	104	23		70^b	61	14	125	38
Cl^-			22^{ab}	21	14	37	6		27^{ab}	30	13	39	8
NO_3^-			82^a	88	27	129	25		92	92	19	147	42
SO_4^{2-}			238^{ab}	235	175	324	37		258^b	270	124	346	57
pH			6.4^{ab}	6.4	6.2	6.6	0.1		6.2^{ab}	6.3	5.4	6.4	0.2
COND			69.5	66.9	41.9	109.0	17.8		66.8^a	61.8	26.3	96.4	21.5
Na^+	16	Medium	63	61	32	105	25	17	61^c	58	37	124	24
NH_4^+			3^c	2	2	11	3		9^c	2	2	62	17
K^+			14	13	6	26	6		15^c	16	6	24	5
Ca^{2+}			429	421	192	773	172		434^c	410	215	691	143
Mg^{2+}			79	77	36	127	31		93^c	87	37	211	45
Cl^-			52	53	27	68	11		48	57	20	73	18
NO_3^-			49	52	4	94	22		113	62	5	425	138
SO_4^{2-}			268^c	257	223	357	34		254^c	260	95	334	60
pH			6.1^c	6.2	5.1	6.9	0.4		6.6	6.6	6.4	6.9	0.1
COND			74.9	72.5	41.2	117.0	22.3		105.0^c	110.0	50.1	173.0	32.8

(*Continued*)

Table 3. Continued.

		ZC						CC				
N	Water level	Mean	Median	Min	Max	SD	N	Mean	Median	Min	Max	SD
9	High						12					
Na^+		80	75	60	132	22		89	84	57	133	25
NH_4^+		63	33	18	271	80		51	44	13	114	31
K^+		20	17	12	41	9		24	22	11	37	7
Ca^{2+}		378	410	74	652	238		121	130	2	229	65
Mg^{2+}		117	95	57	186	46		150	149	61	251	65
Cl^-		51	48	42	68	9		55	57	29	81	15
NO_3^-		64	67	9	96	25		56	68	6	87	27
SO_4^{2-}		306	300	273	364	34		308	307	226	391	51
pH		6.6	6.7	5.5	7.2	0.5		6.9	7.0	5.9	7.4	0.5
COND		72.0	64.2	41.9	122.0	25.2		72.7	77.5	40.1	95.8	19.5

Notes: COND – Electrolytic conductivity.
[a]Statistical difference between 'low' and 'medium' (U Mann Whitney test, $p < 0.05$).
[b]Statistical difference between 'low' and 'high' (U Mann Whitney test, $p < 0.05$).
[c]Statistical difference between 'medium' and 'high' (U Mann Whitney test, $p < 0.05$).

Table 4. A contribution of variables to the LDA model for the ZC and CC.

	Wilk's Lambda	Partial	*F*-remove	*p*-level	Tolerance (R^2)	$1 - R^2$
			ZC springs			
Na^+	0.044	0.908	1.56	0.225	0.297	0.703
NH_4^+	0.058	0.684	7.17	0.003	0.298	0.702
K^+	0.042	0.939	1.01	0.375	0.116	0.884
Ca^{2+}	0.045	0.885	2.00	0.152	0.622	0.378
Mg^{2+}	0.045	0.887	1.98	0.156	0.119	0.881
Cl^-	0.088	0.455	18.56	0.000	0.491	0.509
NO_3^-	0.068	0.590	10.77	0.000	0.465	0.535
SO_4^{2-}	0.040	0.989	0.17	0.848	0.364	0.636
pH	0.050	0.805	3.75	0.035	0.543	0.457
COND	0.056	0.715	6.17	0.006	0.486	0.514
			CC springs			
Na^+	0.022	0.965	0.52	0.601	0.419	0.581
NH_4^+	0.033	0.633	8.40	0.001	0.548	0.452
K^+	0.029	0.718	5.68	0.008	0.157	0.843
Ca^{2+}	0.049	0.427	19.42	0.000	0.254	0.746
Mg^{2+}	0.022	0.936	1.00	0.382	0.155	0.845
Cl^-	0.022	0.952	0.74	0.487	0.455	0.545
NO_3^-	0.023	0.886	1.86	0.173	0.444	0.556
SO_4^{2-}	0.022	0.958	0.64	0.535	0.591	0.409
pH	0.030	0.692	6.45	0.005	0.547	0.453
COND	0.025	0.836	2.85	0.074	0.243	0.757

Note: COND – Electrolytic conductivity.

Table 5. Standardised canonical DF coefficients for LDA of the ZC and CC.

Standardised canonical DF coefficients

ZC	1st DF	2nd DF	CC	1st DF	2nd DF
Cl^-	1.13	−0.05	Ca^{2+}	1.54	0.06
NO_3^-	−0.85	−0.61	K^+	−1.36	0.22
Na^+	0.59	−0.07	pH	−0.69	−0.42
K^+	−0.58	0.58	NH_4^+	−0.84	−0.02
Mg^{2+}	−0.46	−1.05	NO_3^-	0.52	0.07
Ca^{2+}	0.44	0.14	Mg^{2+}	−0.50	0.54
pH	0.41	−0.55	Na^+	−0.10	−0.35
SO_4^{2-}	0.17	0.07	Cl^-	0.09	−0.40
COND	0.12	0.90	SO_4^{2-}	−0.06	0.34
NH_4^+	−0.01	−1.23	COND	0.02	−1.06
Eigenvalue	6.52	2.33	Eigenvalue	18.19	1.51
Cumulative proportion	0.74	1.00	Cumulative proportion	0.92	1.00

Note: COND – Electrolytic conductivity.

succession the leaching of NO_3^-, Mg^{2+} and NH_4^+ may be limited by the increasing demands of forest photosynthesis [55,56].

For CC the first DF, associated mostly with Ca^{2+}, K^+, pH, NH_4^+, NO_3^- and Mg^{2+}, was characterised by an eigenvalue of 18.19, a canonical correlation of 0.97 and explained 92% of the variance. The second DF was characterised by an eigenvalue of 1.51 and a canonical correlation of 0.78. It explained 8% of the total variance being mostly associated with COND, pH and Cl^- as well as with Mg^{2+}, also associated with the first DF. The first DF had a complex meaning and reflected both geological conditions and the nutrient biocycles. It allowed the distinction of high water level related samples from others, while the second DF enabled medium and low water level related samples to be isolated. Similarly, standardised canonical DF coefficients indicated that the concentration of NH_4^+, Mg^{2+} and pH increased together with increasing water level, while the concentration of Ca^{2+} decreased. In general, the observed variation agreed with explanations proposed above for ZC.

In the final step of the LDA model, calculation classification rules were tested on the basis of the derived DFs. For ZC, an almost excellent separation of the spring water samples was achieved. All the samples were correctly classified as being either low or high water level related, while concerning medium water level related samples the classification correctness was reduced to 94% as a single sample was misclassified. For CC, the classification success was slightly less effective in the case of low water level samples. For these samples 86% (12) of samples were classified correctly, although two samples were erroneously included in the low water level group. In general, DFs indicated a high efficiency in the classification process connected to the investigation of statistically significant discriminant parameters responsible for distinguishing between chemical profiles of samples related to various water levels on the area of investigated catchments. This statement was well supported by two-dimensional scatter plots using DFs on samples along DF1 and DF2 (Figure 3a and b, respectively). In summary, LDA delivered satisfactory, although very general, information related to surface water chemical composition variability as a function of seasonal changes in water level in these mountain catchments. This generality was why the application of other, especially unsupervised classification techniques, became advisable.

In order to discover the impact of other factors on spring water chemistry, a SOM-based exploration and classification was applied. After the net has been trained, the spring water samples were clustered based on the K-means clustering mode. Different values of k (predefined number of clusters) were tried, and the sum of squares for each run was calculated. The best classification with the lowest DB index (a function of the ratio of the sum of within-cluster scatter and between-cluster separation) values (presented for ZC and CC in Figure 4a and b) was chosen. Cluster borders were set based on the obtained unified distance matrix [51]. A unified distance matrix is a form of a summary map reflecting a simultaneous contribution of each input parameter to self-organisation of the map. In the form of an intermediate SOM procedure step, the unified distance matrix helps to identify the cluster structures of the final map where high values of the U-matrix indicate a cluster border, while uniform areas of low values indicate clusters themselves. A four cluster-configuration had the lowest index for ZC, and a three clusters-configuration for CC. Due to the dimensions of the Kohonen's map for ZC and CC defined above, it was clear that more than one case from the initial data set ($n = 43$ and $n = 44$ for ZC and CC, respectively) was related to a particular hexagon. Samples included into each hexagon were grouped in agreement with the cluster borders. Contribution of particular spring water
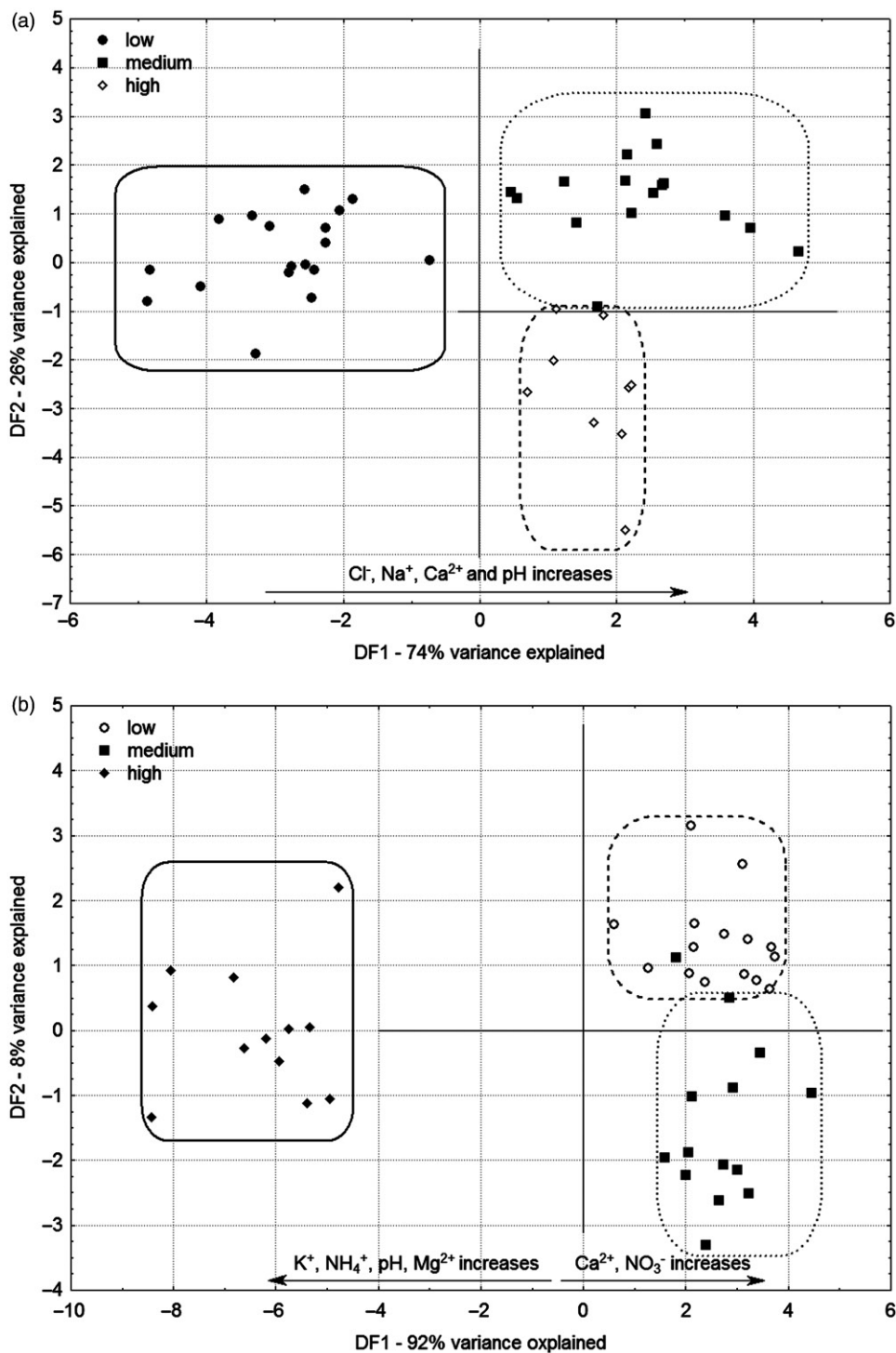
Figure 3. Scatterplot of canonical scores on the plan described by DF1 and DF2 for the (a) Zimnik and (b) Czyrna catchments.
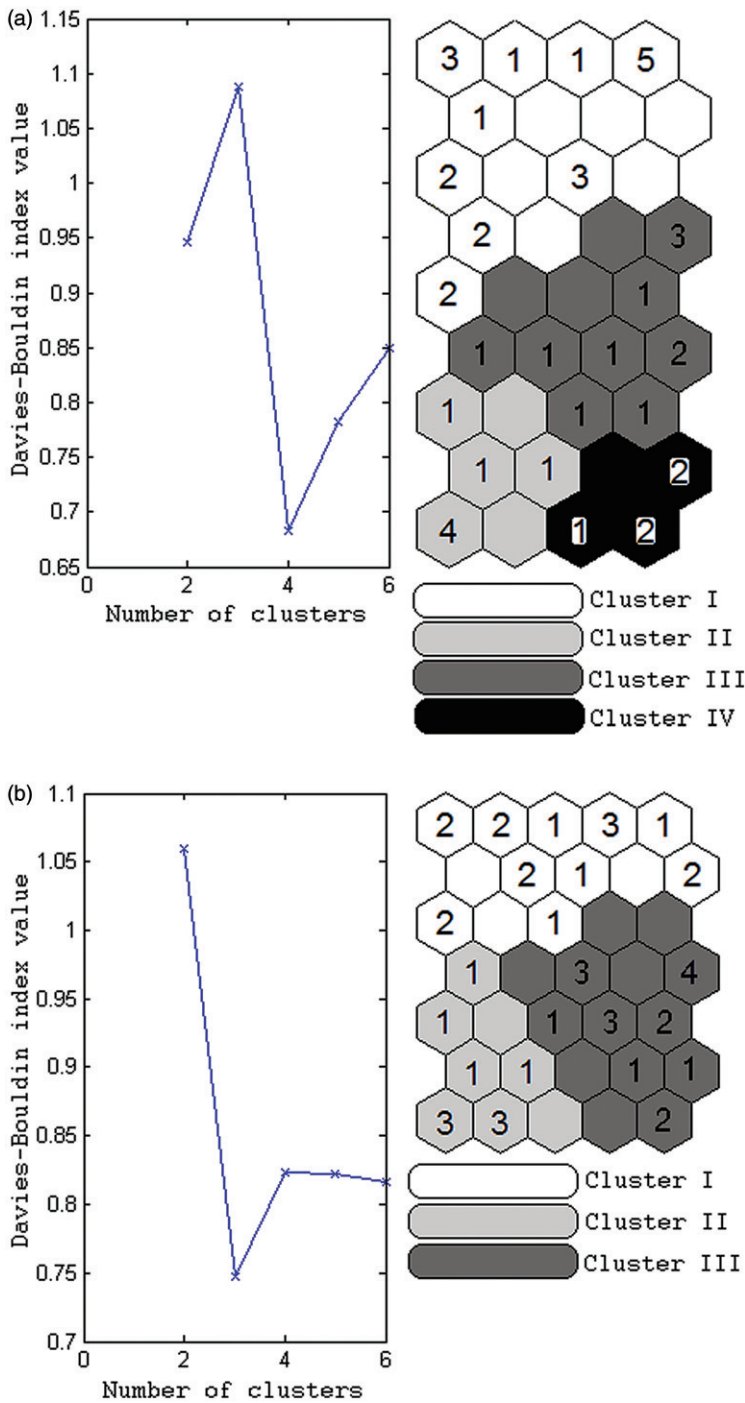
A. Astel et al.

Figure 4. SOM based clustering pattern according to the Davies–Bouldin index minimum value for the (a) Zimnik and (b) Czyrna catchments. The digits placed in hexagonal SOM units correspond to the number of spring water samples belonging to a particular cluster.

samples to the clusters derived by SOM clustering are shown in Table 6. Setting up initially determined values of physiochemical features along with the classification results allowed for obtaining a clustering pattern associated with the impact of forest stands and elevation on the chemical profile of samples in the ZC and CC. Table 7 shows the results of the application of Kruskal–Wallis and Dunn tests on the differences between the levels of spring water quality indicators for ZC (clusters I–IV) and CC (clusters I–III) obtained by SOM-based classification. For ZC, the clusters I–IV included various quantities of a total 43 cases as follows: I – 20, II – 7, III – 11 and IV – 5.

Cluster I grouped spring water samples collected above 1000 m a.s.l. on south-east slopes dominated by spruce forest. In comparison to other groups, these samples were characterised by the lowest concentration of Na^+, K^+, Mg^{2+}, Cl^- and SO_4^{2-} due to the low weathering rate at a high elevations in a poor soil formed on a sandstone bedrock [53,54], and the increasing effect of spruce monoculture on leaching [13].

Cluster II grouped samples collected between 710 and 810 m a.s.l. on south facing slopes. In the lower samples mixed mountain forest prevails with beech as a dominant. These samples were characterised by the highest contents of K^+ and Cl^- as well as the lowest contents of NO_3^-, pH and COND. These characteristics relate to the higher weathering rate (especially of K^+) at this elevation, and a richer soils with a higher amounts of clay particles that was formed on richer bedrock [53,54], as well as lower nitrification rates and an increased impact of beech trees [55,56].

Cluster III had a mixed character. It grouped spring water samples located on easterly slopes that are characterised by a so-called 'precipitation shadow' and hence lower rainfall. These samples were characterised by the highest concentrations of Ca^{2+}, and also of NO_3^- due to more intensive nitrification processes [57,58].

Cluster IV consisted of only five samples. For three of them, sediments were characterised as polygenic conglomerate. These samples were characterised by the highest contents of Na^+, NH_4^+, Mg^{2+} and SO_4^{2-} and the lowest of Ca^{2+}. In these samples also the highest pH and COND were related to denitrification processes [57,58].

Similarly to ZC the clusters I–III for CC consisted of various numbers of a total of 44 cases as follows: I – 17, II – 10, III – 17. Cluster I grouped samples collected between 1050 and 1070 m a.s.l. on north facing slopes dominated by young spruce monocultures. These samples were characterised by the lowest concentrations of Na^+, K^+, Mg^{2+}, Cl^-, NO_3^-, SO_4^{2-}, as well as of low pH and COND as a consequence of low weathering rate and poor soils formation on poor bedrock [53,54], and also because of an increased acidification effect of the spruce monocultures causing increased leaching of elements [57,58].

Cluster II contained samples collected from springs located between 750 and 800 m a.s.l. on north facing slopes, where mostly European beech prevailed in major forest stands. These samples were characterised by the highest contents of all ions excluding Ca^{2+} and NO_3^-, resulting from higher weathering at this elevation and the rich clay soils formed on the bedrock [30,31], as well as lower rates of nitrification [56,57].

Cluster III contained samples collected between 730 and 850 m a.s.l. with highly modified European beech. These samples were characterised by the lowest concentrations of NH_4^+ and the highest of NO_3^- and Ca^{2+}, resulting from intensive nitrification processes [57,58]. Also high leaching from soils with elevated amount of clay particles formed on reach bedrock [53,54], and an impact of beech can explain this cluster [55,56].

Table 6. A distribution of cluster episodes.

ZC								CC					
Cluster I		Cluster II		Cluster III		Cluster IV		Cluster I		Cluster II		Cluster III	
Water level	Sample i.d.	Water level	Sample i.d.	Water level	Sample i.d.	Water level	Sample i.d.	Water level	Sample i.d.	Water level	Sample i.d.	Water level	Sample i.d.
Low	3b	Medium	5a	Medium	3e	High	1a	Low	01f	High	01i	Medium	01i
Low	4a	Medium	1a	Medium	3b	Medium	2a	Low	01h	High	01j	Medium	02b
Low	4b	Medium	6a	Medium	4b	High	3b	High	01c	High	01e	Medium	03c
Low	3e	Medium	6b	Medium	4c	High	2a	High	01h	High	01a	Medium	03a
Medium	3d	Medium	6c	Low	1c	High	RMWP	Low	01c	High	01g	Medium	01j
Medium	4a	High	3d	Low	2b			Medium	01f	High	03d	Medium	03c'
Medium	3a	Medium	RMWP	Low	RMWP			Medium	01c	High	02a	Medium	03d
Medium	3f			Low	2a			Medium	01h	High	01d	Medium	02a
Medium	3c			High	3a			Low	01d	High	03a	Low	01e
High	3e			High	3f			Low	01a	High	03b	Low	02a
Low	6a			High	3g			Medium	01e			Low	03a
Low	1a							Low	01b			Medium	01d
Low	1b							Low	01i			Medium	01g
Low	4c							Low	01j			Medium	03b
Low	4d							Low	01g			Low	03d
Low	3a							Low	03b			Medium	01a
Low	3c							Low	03c'			Medium	01b
Low	3d												
Low	3f												
Low	5a												
Low	70.0%	Low	0%	Low	36.4%	Low	0%	Low	65.0%	Low	0%	Low	24.0%
Medium	25.0%	Medium	85.7%	Medium	36.4%	Medium	20.0%	Medium	23.5%	Medium	0%	Medium	76.0%
High	5%	High	14.3%	High	27.3%	High	80.0%	High	12.5%	High	100%	High	0%

Table 7. A statistical assessment (Kruskal–Wallis and Dunn tests) of differences between median values of chemical indicators of spring water quality for ZC (clusters I–IV) and CC (clusters I–III) obtained by the SOM algorithm.

Feature		Median values — Clusters				Kruskal–Wallis' test	Dunn's multiple comparison test					
		I	II	III	IV		I–II	I–III	I–IV	II–III	II–IV	III–IV
ZC	Na^+ [$\mu g\,L^{-1}$]	37	83	62	91	KW = 29.17; $p < 0.001$	***	*	***	n.s.	n.s.	n.s.
	NH_4^+ [$\mu g\,L^{-1}$]	2	2	2	26	KW = 8.867; $p < 0.05$	n.s.	n.s.	*	n.s.	n.s.	n.s.
	K^+ [$\mu g\,L^{-1}$]	9	21	12	18	KW = 27.5; $p < 0.001$	***	*	***	n.s.	n.s.	n.s.
	Ca^{2+} [$\mu g\,L^{-1}$]	254	458	511	174	KW = 16.39; $p < 0.001$	n.s.	***	n.s.	n.s.	n.s.	n.s.
	Mg^{2+} [$\mu g\,L^{-1}$]	51	103	89	151	KW = 30.84; $p < 0.001$	**	**	***	n.s.	n.s.	n.s.
	Cl^- [$\mu g\,L^{-1}$]	26	61	36	50	KW = 17.79. $p < 0.001$	***	n.s.	n.s.	n.s.	n.s.	n.s.
	NO_3^- [$\mu g\,L^{-1}$]	76	46	81	61	KW = 13.15. $p < 0.01$	*	n.s.	n.s.	**	n.s.	n.s.
	SO_4^{2-} [$\mu g\,L^{-1}$]	238	247	276	348	KW = 18.43; $p < 0.001$	n.s.	n.s.	***	n.s.	*	n.s.
	PH	6.4	5.8	6.4	7.0	KW = 19.40; $p < 0.001$	**	n.s.	n.s.	*	***	n.s.
	COND [$mS\,m^{-1}$]	67.0	54.1	79.7	97.0	KW = 11.67; $p < 0.01$	n.s.	n.s.	n.s.	n.s.	*	n.s.
CC	Na^+ [$\mu g\,L^{-1}$]	37	91	64		KW = 29.844; $p < 0.001$	***	***		n.s.		
	NH_4^+ [$\mu g\,L^{-1}$]	6	38	2		KW = 18.39; $p < 0.001$	**	n.s.		***		
	K^+ [$\mu g\,L^{-1}$]	10	22	17		KW = 23.50; $p < 0.001$	***	**		n.s.		
	Ca^{2+} [$\mu g\,L^{-1}$]	231	140	488		KW = 30.78; $p < 0.001$	n.s.	***		***		
	Mg^{2+} [$\mu g\,L^{-1}$]	45	159	109		KW = 27.47; $p < 0.001$	***	***		n.s.		
	Cl^- [$\mu g\,L$]	27	59	47		KW = 17.76; $p < 0.001$	***	**		n.s.		
	NO_3^- [$\mu g\,L^{-1}$]	49	73	77		KW = 3.719; $p > 0.05$	n.s.	n.s.		n.s.		
	SO_4^{2-} [$\mu g\,L^{-1}$]	230	327	270		KW = 11.79; $p < 0.01$	**	n.s.		*		
	PH	6.3	7.0	6.6		KW = 14.83; $p < 0.001$	***	n.s.		n.s.		
	COND [$mS\,m^{-1}$]	55.5	82.3	110.0		KW = 25.0; $p < 0.01$	n.s.	***		n.s.		

Notes: n.s. – not significant statistically.
Statistical level of significance (p) lower than 0.001 marked as '***', in the range of $0.001 \leq p < 0.01$ marked as '**', while in the range of $0.01 \leq p < 0.05$ marked as '*'.

4. Conclusions

Spring water chemistry is highly related with water level in mountain catchments. Application of LDA enables the identification of the most discriminating parameters responsible for the chemical profile of springs. For both the ZC and CC, the most important factor is related to geological conditions (expressed by variability of Cl^-, Na^+, K^+, Mg^{2+}, Ca^{2+} and pH), while the second relates to the nutrient biocycle (mainly NH_4^+ and NO_3^-). Due to a poor soil composition in ZC the spring-time thawing dominated over weathering processes, and limited the impact of Cl^-, Na^+ and Ca^{2+} in spring water. The nutrient biocycle varies seasonally and is dependent on the photosynthesis activity of the forest. Thus, in the initial phase of the vegetation succession, with high water level in the catchment, the concentrations of NH_4^+ and NO_3^- increased, but towards the middle and at the end of the vegetation succession the leaching, and hence the concentration of NO_3^-, Mg^{2+} and NH_4^+ was limited by the increasing demand related to the photosynthesis activity of the forest ecosystems. Unsupervised, Kohonen SOM-based classification enables logical clustering of objects of interest, and hence shows that location of the springs and the prevailing forest stand are the major influences on the chemical profile of spring water. In case of springs located above 1000 m, limited weathering connected with the effect of the highly acid spruce monoculture on leaching causes a decreasing concentration of Na^+, K^+, Mg^{2+}, Cl^- and SO_4^{2-}. Limited leaching caused by the effect of spruce monocultures dominated during the decadent succession of vegetation, along with low water levels. Springs located between 700 and 800 m a.s.l. in mixed forests, showed the highest concentrations of K^+ and Cl^- and the lowest concentrations of NO_3^-, values of pH and COND. Such phenomenona were an effect of high weathering and low nitrification rates, as well as the impact of the beech stand.

Acknowledgements

References

[1] C. Cun and R. Vilagines, Sci. Total Environ. **208**, 59 (1997).
[2] V. Simeonov, J.W. Einax, I. Stanimirova, and I. Kraft, Anal. Bioanal. Chem. **374**, 898 (2002).
[3] V. Simeonov, C. Sarbu, D.L. Massart, and S. Tsakovski, Microchim. Acta **137**, 243 (2001).
[4] A. Astel, S. Tsakovski, V. Simeonov, E. Reisenhofer, S. Piselli, and P. Barbieri, Anal. Bioanal. Chem. **390**, 1283 (2008).
[5] S.K. Swanson, J.M. Bahr, M.T. Schwar, and K.W. Potter, Chem. Geol. **179**, 73 (2001).
[6] K. Šnuderl, M. Simonič, J. Mocak, and D. Brodnjak-Vončina, Acta Chim. Slov. **33–39**, 54 (2007).
[7] G. Ragno, M. De Luca, and G. Ioele, Microchem. J. **87**, 119 (2007).
[8] L.A. Munk and G. Faure, Appl. Geochem **19**, 1065 (2004).
[9] S.T. Trudgill, A.M. Pickles, T.P. Burt, and R.W. Crabtree, Eur. J. Soil. Sci. **32**, 433 (2006).

[10] G.C. Lacey and R.B. Grayson, J. Hydrol. **204**, 231 (1998).

[11] D.W. Clow, R.G. Striegl, L. Nanus, M.A. Mast, D.H. Campbell, and D.P. Krabbenhoft, Water Air Soil Pollut. **2**, 139 (2002).

[12] S. Wróbel, in *Environmental Degradation in the Czarna Wisełka and Biała Wisełka Catchments. Western Carpathians*, edited by S. Wróbel (Studia Naturae, Cracow, 1998), Vol. 44, pp. 81–100.

[13] S. Małek and A. Astel, Environ. Pollut. **155**, 517 (2008).

[14] S. Małek and T. Gawęda, Zielona Góra Univeristy Scientific Papers: Environmental Engineering, **131**, 257 (2004).

[15] S. Małek and T. Gawęda, Sylwan **3**, 39 (2006).

[16] S. Małek and K. Krakowian, in *The Functioning of Geoecosystems of River Catchment 4 – Extreme Processes in the Geographical Environment*, edited by A. Kostrzewski and J. Szpikowski (Bogucki Scientific Press, Poznań, 2007), pp. 331–344.

[17] B. Szczęsny and D. Zięba, J. Nat. Conservat. **58**, 109 (2001).

[18] J. Pawlik-Dobrowolski, Jagiellonian University Scientific Papers, **34**, (1965).

[19] E. Feliksik and G. Durło, Dendrobiology **51**, 47 (2004).

[20] PN-ISO 10304-1:1998: Water quality – Determination of dissolved fluoride, chloride, nitrite, orthophosphate, bromide, nitrate and sulphate ions using liquid chromatography of ions – Part 1, Method for water with low contamination.

[21] PN-ISO 14911-1:1998: Water quality – Determination of dissolved Li^+, Na^+, NH_4^+, K^+, Mn^{2+}, Ca^{2+}, Mg^{2+}, Sr^{2+} and Ba^{2+} using ion chromatography method.

[22] P. Konieczka, J. Namieśnik, B. Zygmunt, E. Bulska, A. Świtaj-Zawadka, A. Naganowska, E. Kremer, and M. Rompa, *Quality Assessment and Quality Control of Analytical Results–QA/QC* (Jagiellonìan University Cracow, Poland, 2004).

[23] B. Tepuš and M. Simonič, Cent. Eur. J. Chem. **5**, 557 (2007).

[24] K. Röhker, O. Rienitz, and D. Schel, Accred Qual. Assur. **9**, 671 (2004).

[25] P. Miskaki, E. Lytras, L. Kousouris, and P. Tzoumerkas, Desalination **213**, 182 (2007).

[26] J. Einax, D. Truckenbrodt, and O. Kampe, Microchem. J. **58**, 315 (1998).

[27] J. Einax, H. Zwanziger, and S. Geiß, *Chemometrics in Environmental Analysis* (John Wiley & Sons Ltd, Chichester, 1997).

[28] V. Simeonov, C. Sarbu, D.L. Massart, and S. Tsakovski, Microchim. Acta **137**, 243 (2001).

[29] V. Simeonov, J.W. Einax, I. Stanimirova, and I. Kraft, Anal. Bioanal. Chem. **374**, 898 (2002).

[30] M.M. Jordán, J. Mateu, and A. Boix, Water Air Soil Pollut. **107**, 91 (1998).

[31] K.P. Singh, A. Malik, and V.K. Singh, Water Air Soil Pollut. **170**, 383 (2006).

[32] B.F.J. Manly, *Multivariate Statistical Methods* (Chapman and Hall, London, 1986).

[33] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, and L. Kaufman, *Chemometrics: A Textbook* (Elsevier, Amsterdam, 1980).

[34] R.G. Brereton, *Chemometrics: Applications of Mathematics and Statistics to the Laboratory* (Ellis Horwood, Chichester, 1990).

[35] G. Scarponi, I. Moret, G. Capodaglio, and P. Cescon, J. Agric. Food Chem. **30**, 1135 (1982).

[36] J.F. Diáz-Flores, E.F. Diáz-Flores, C. Hernández Calzadilla, E.M. Rodríguez, C. Diáz Romero, and L. Serra-Majem, Eur. J. Clin. Nut. **58**, 449 (2004).

[37] S. Mikkonen, K.E.J. Lehtinen, A. Hamed, J. Joutsensaari, M.C. Facchini, and A. Laaksonen, Atmos. Chem. Phys. Disc. **6**, 8485 (2006).

[38] W.R. Klecka, *Discriminant Analysis* (Sage Publications Inc., Chester, USA, 1970).

[39] P.A. Lachenbruch, *Discriminant Analysis* (Hafner Press, New York, USA, 1975).

[40] M.M. Tatsuoka, *Discriminant Analysis: The Study of Group Differences* (Institute for Personality and Ability Testing, Champaign, 1970).

[41] C.J. Huberty, *Applied Discriminant Analysis* (Willey, New York, USA, 1994).

[42] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition* (Willey, New York, USA, 1992).

[43] A. Astel, S. Małek, and S. Makowska, Water Air Soil Pollut. **195**, 137 (2008).

[44] T. Kohonen, Biol. Cybern. **43**, 59 (1982).

[45] T. Kohonen, E. Oja, O. Simula, and A. Visa, J. Proc. IEEE **84** (10), 1358 (1999).

[46] Introduction to Chemometrics and Statistics < http://dissertations.ub.rug.nl/FILES/faculties/ science/1998/j.nilsson/c2.pdf > .

[47] T. Kohonen, *Self-Organizing Maps*, 3rd ed. (Berlin, Springer, 2001).

[48] J. Vesanto, in Proceedings of Symposium on Tool Environments and Development Methods for Intelligent Systems (TOOL-MET2000), Oulu, Finland, 184 (2000).

[49] Y.S. Park, J. Tison, S. Lek, J.L. Giraudel, J. Coste, and F. Delmas, Ecol. Inform. **1**, 247 (2001).

[50] J. MacQueen, in Proceedings of the Fifty Berkeley Symposium on Mathematic, Statistics and Probability, **1**, 281-296, (1967).

[51] D.L. Davies and B.W. Bouldin, IEEE Transactions on Pattern Recognition and Machine Intelligence, **1**, 224 (1979).

[52] S.S. Shapiro and M.B. Wilk, Biometrika **52**, 591 (1965).

[53] W. Maciaszek and M. Zwydak, in *Environmental Degradation in the Czarna Wiselka and Biala Wiselka Catchments. Western Carpathians*, edited by S. Wróbel (Studia Naturae, Cracow, 1998), Vol. 44, pp. 27–52.

[54] R.J. Dojlido, *Surface Water Chemistry* (Economics and Environment, Białystok, 1995).

[55] S. Małek, Dendrobiology **51**, 65 (2004).

[56] S. Małek and M. Kizior, in *Cycles of Elements in the Environment, T II*, edited by B. Gworek and J. Misiak (Environmental Protection Institute, Warsaw, 2003) pp. 88–92.

[57] A. Rothe, C. Huber, K. Kreutzer, and W. Weis, Plant Soil **240**, 33 (2002).

[58] J.H. Jussy, M. Colin-Belgrand, E. Dambrine, J. Ranger, B. Zeller, and S. Bienaime, Biogeochem. **69**, 241 (2004).