Desalination and Water Treatment
www.deswater.com

1944-3994/1944-3986 © 2012 Desalination Publications. All rights reserved
doi: 10/5004/dwt.2012.2848

Taylor & Francis
Taylor & Francis Group

40 (2012) 183–193
February

Statistical evaluation and comparison of the chemical quality of bottled water and flemish tap water

Stijn W.H. Van Hulle[a,b,*], M. Cristina Ciocci[a]

[a]*University College West Flanders, Ghent University Association, Graaf Karel de Goedelaan 5, B-8500 Kortrijk, Belgium*
[b]*BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, B-9000 Gent, Belgium*
Tel. +32 56 241 251; Fax: +32 56 241 224; email: Stijn.Van.Hulle@howest.be

Received 24 March 2011; Accepted 22 September 2011

ABSTRACT

In this study the physico-chemical composition reported on the label of 49 bottled still waters, 22 bottled sparkling waters and 13 tap waters were used to carry out a characterization study by means of multivariate pattern recognition methods such as principal components analysis (PCA) and discriminant analysis. Also, analysis of variance was used to detect statistical differences between the water types and different brands. The collected data consisted of nine major ions (calcium, magnesium, potassium, sodium, sulfate, bicarbonate, chloride, fluoride and nitrate) and pH. The different water types were represented graphically in a Piper diagram. In this Piper diagram, most of the waters were situated on the left side of the diamond. For the anions, this means that large amounts of HCO_3^- and small amounts of Cl^-, F^- and SO_4^{2-} are present. The main cations are Ca^{2+} and Mg^{2+}, rather than Na^+ or K^+. Further, it was noted that tap water had a higher chloride content, originating from chlorine dosing for disinfection. A one-way analysis of variance (ANOVA) test (with known standard deviation) confirmed the difference among different water types and water brands. Principal components analysis revealed that the differences between water types are best characterized by components that indicate saltiness, hardness and pH. The component pH allowed discriminating between sparkling water and non-sparkling water. It was not possible to divide the different water types based on saltiness or hardness, but it could be demonstrated that different types of water exist (low-mineral, oligomineral and mineral).

Keywords: Water analysis; Satistical evaluation; Chemical water quality; Principal components analysis; Linear discriminant analysis; Analysis of variance

1. Introduction

There are two common types of drinking water: bottled water and tap water. Both tap water and bottled water can have many sources. Tap water in Flanders usually originates from ground water but surface water is also used. In other parts of the world also other types of water such sea water is used [1,2]. Drinking water is also derived from natural springs, where water from underground aquifers meets the ground surface or drinking water originates from the ocean where it is pumped to the surface from depths of up to 200 m below the ocean surface layer [3]. Bottled water can also come from surface waters, such as rivers or canals. Bottled water can be naturally still water (mineral water) or naturally sparkling water (containing gaseous CO_2). Bottled sparkling

---

*Corresponding author.

water can also contain artificially added CO_2 gas [4]. Bottled water consumption has grown steadily over the past few decades. In 2003, bottled water became the second largest commercial beverage category by volume in the United States [3]. Also in other parts of the world the sale of bottled water is increasing [3,5,6]. Bottled water is considered to be a safe and readily available product and in several countries consumers prefer bottled water over tap water as drinking water [7,8]. However, bottled drinking water consumption is associated with a higher economic and ecological cost [9]. Furthermore, it can be demonstrated that the quality of tap water is not significantly different from the quality of bottled water in some developed coutries [4].

In view of this debate on the use of tap water and bottled water a statistical analysis was performed on the chemical composition of these waters in order to determine the differences in major ion content. As such this statistical analysis aims at classifying the different types of water as their chemical composition may differ. Marine water for example needs to be desalinized before it is drinkable. Natural spring water and other drinking waters can be disinfected using ozone/UV, filtered with activated carbon or treated with other techniques such as reverse osmosis [6]. These treatments change the composition of the water. Also, the different geological origins of water, the mineralogy of the soil in and around the aquifer or river from which the water is collected, the residence time of water in the aquifer, the degree of urbanization around the aquifers and the climate will introduce difference in the physico-chemical composition of the water [3,6].

The chemical composition of the waters was studied by means of one-way analysis of variance (ANOVA model with known standard deviation), principal components analysis (PCA) and linear discriminant analysis (LDA). The main goal of this study is to verify if it is indeed possible to distinguish among the different types of water and different brands on the base of the ion composition reported on the labels.

2. Materials and methods

2.1. Water database

The physico-chemical parameters of 84 different water types were collected. For most of the bottled water types the physico-chemical parameters reported on the bottle label were used, although part of the database was also obtained through the data provided [3]. The tap water composition was supplied by different Flemish water companies. The chemical parameters used in this study were calcium (Ca^{2+}), magnesium (Mg^{2+}), potassium (K^+), sodium (Na^+), sulfate (SO_4^{2-}), bicarbonate (HCO_3^-), chloride (Cl^-), fluoride (F^-) and nitrate (NO_3^-). The applied measurement method for obtaining the reported concentrations, as well as the accuracy and precision of the reported measurements was not questioned in this study in view of the legal requirements concerning tap and bottled water quality analysis. However, as an independent check on the quality of the chemical analyses in the database charge balance error was calculated [6,10]:

$$\%\text{ charge balance error} = \frac{\sum z \times m_c - \sum z \times m_a}{\sum z \times m_c + \sum z \times m_a} \times 100 \tag{1}$$

where

z = the absolute value of the ionic valence

m_c = the molarity of cationic species

m_a = the molarity of the anionic species

In accordance with Güler, only water types with a calculated charge balance error of less than ± 10% were accepted in this study [6].

The ionic strength of the different water types was calculated as a summary of the different ions in the sample with the following equation:

$$I = \frac{1}{2}\sum_{i=1}^{n} z_i^2 m_i \tag{2}$$

where n is the number of different ions.

In case the pH was not mentioned on the bottle label, it was measured using a HI 9023, Prominent® – PHEX 112 SE pH-electrode.

2.2. Piper diagram

The Piper diagram is a relatively old but still widely used method to graphically represent the composition of waters in a single diagram [11]. The Piper diagram only represents the presence of cations (Ca^{2+}, Mg^{2+}, Na^+ and K^+) and anions (HCO_3^-, CO_3^{2-}, SO_4^{2-}, Cl^- and F^-) in waters. The Piper plot consists of two trilinear diagrams, one representing the relative concentrations of the cations, the other representing the relative concentrations of the anions. The two trilinear diagrams are then combined in a diamond shaped field. This field, commonly referred to as simply the "diamond," allows to classify a water into different zones. For example a zone can be distinguished where the chloride + sulfate share is greater than 90%. These different zones can thus group different waters with the same properties together (Fig. 1) [12,13]. A template implemented in Excel supplied by the U.S. Geological Survey

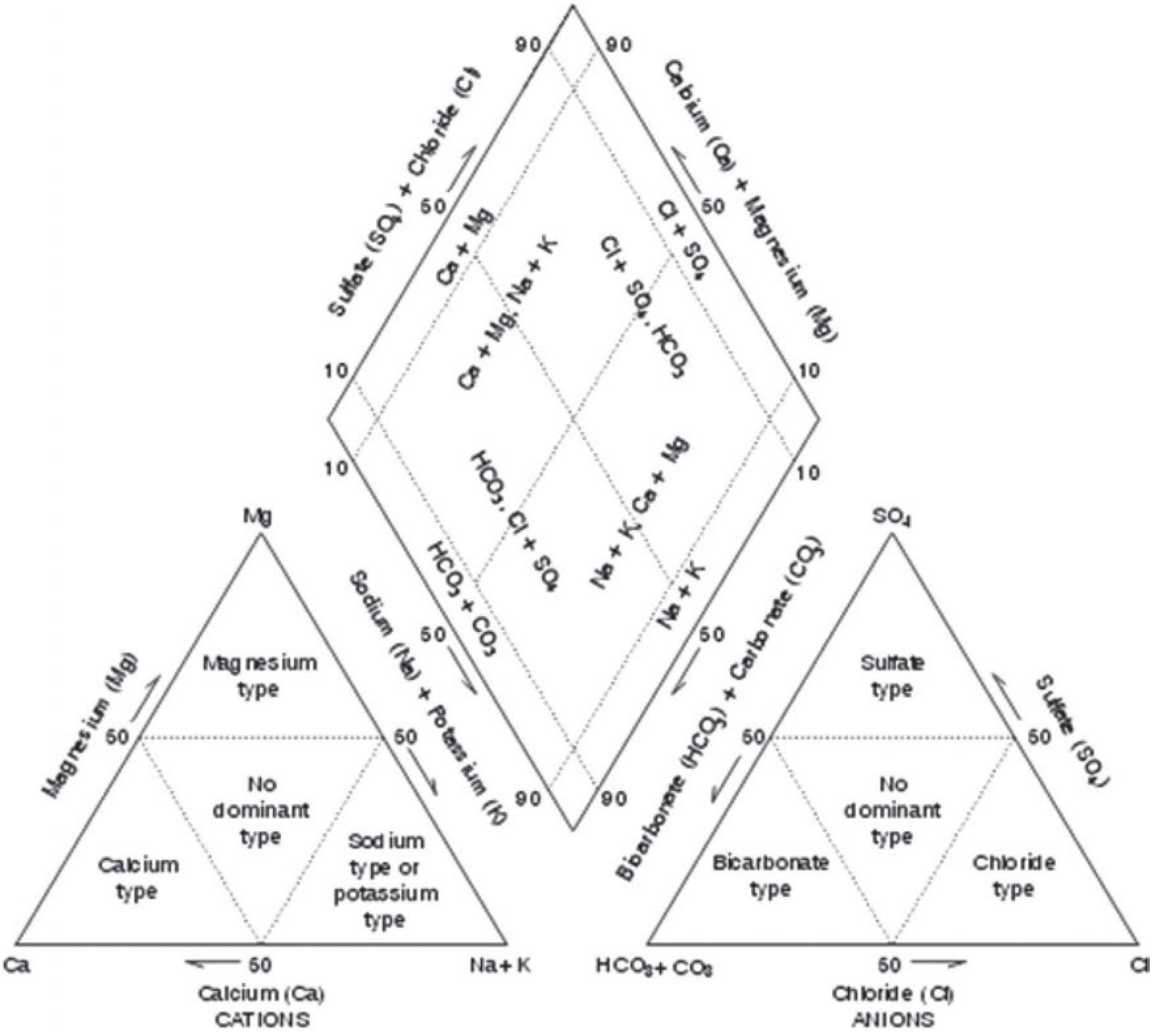

Fig. 1. The sectors of the Piper diagram which represent a type of water, based on dominant presence of ions [13] (http://nevada.usgs.gov/tech/excelforhydrology).

(http://nevada.usgs.gov/tech/excelforhydrology) was used for the construction of the Piper diagram. In this study fluoride was combined with chloride. Also, there was no data on the carbonate concentration available and only the bicarbonate concentration was taken into account. This will not strongly affect the accuracy of the study, as the amount of the carbonate ion (CO_3^{2-}) is only relevant, with respect to the amount of bicarbonate ion (HCO_3^-), in strong alkaline solution which is not the case for drinking waters.

2.3. Statistical analysis

Initially a one-way ANOVA test with known standard deviation was performed manually to demonstrate the possible differences in composition amongst the different types of water. For this calculation an analytical uncertainty of 5% on the measured values was assumed [3]. From this uncertainty, the standard deviation on an individual ion concentration measurement was calculated to be 5% of the measured value. Further, by definition, the variance on a measurement is the squared value of the standard deviation. With these variances, a mean of sample variances (MSV) can be calculated [14]. Further, the variance of the sample means (VSM) can be calculated as the variance on the individual ion concentration measurements. The F-value in the ANOVA test can then be calculated as:

$$F = \frac{n * \text{VSM}}{\text{MSV}} \tag{3}$$

where n is the number of measurements, arbitrarily chosen to be 10.

For further statistical analysis, the software program SPSS version 17 was used (www.spss.com). As an initial step PCA and LDA were used to select the most discriminating parameters and to investigate the overall variation of the data. The pH was not measured for all the water types and therefore PCA and LDA were performed considering the 57 water types for which the pH was known.

PCA was used as pattern recognition method and aims at reducing a large number of variables to a smaller number of representative variables (principal components or PC's) [6,15]. Varimax normalized rotation of principal components was carried out in order to reduce the contribution of variables with minor significance and increase interpretability of the components [15]. The number of PCs extracted (to explain the underlying data structure) is defined by using the "Kaiser criterion" where only the PCs with eigenvalues greater than unity are retained [16].

Initially 10 factors (all ion concentrations and the pH) were used. A first PCA analysis revealed that the correlation of the NO_3^- concentration with the other factors (i.e., concentrations) values is too low (<0.4) [15]. This can probably be explained by the fact that only about 10% of all recorded NO_3^- concentrations differ with at least 2 standard deviations from the average NO_3^- concentration. As such the NO_3^- concentration was not included in further statistical analysis and only 9 factors are considered further. The rule of thumb that the ratio "Cases to Factors" should be at least 5/1 is met with 57 cases and 9 factors.

LDA was used to achieve maximum discrimination among known groups. Groups are forced to be as statistically different as possible by forming a weighted linear combination of the discriminating variables (i.e., the ion concentrations and the pH), the weights or canonical discriminant function coefficients (β_i) are estimated so that they result in an optimal separation between the groups. Also, LDA provided better insight into the relationship between group membership and the variables used to predict group membership [17,18]. The Wilk's Lambda was used to catch the importance of the discriminating function. The Wilk's Lambda is an inverse measure of the functions discriminating power; the smaller the value of Wilk's Lambda the better the discriminating power of the function [15].

As LDA gives unstable solutions when the variables are highly correlated and when the assumption of multivariate normal distribution is not met, we also performed LDA using the PC-scores obtained by previous PCA instead of the original variables. As similar results were obtained with the PC-scores as with the original values, only the results from the original values will be discussed further.

3. Results and discussion

3.1. Collected data and Piper diagram

In Table 1, the collected data of 84 different water types is presented. Of these 84 different water types, 4 are considered to be low mineral waters with a total mineral content lower than 50 mg l^{-1}. Another 47 water types are considered to be oligomineral with a total mineral content between 50 and 500 mg l^{-1}. The remaining 37 water types are considered to be mineral (TDS > 500 mg l^{-1}).

In Fig. 2 the data from Table 1 is summarized in a Piper diagram. In the diamond of the Piper plot it can be seen that most of the waters are situated on the left side. For the anions, this means that large amounts of HCO_3^- and small amounts of Cl^-, F^- and SO_4^{2-} are present. The main cations are Ca^{2+} and Mg^{2+}, rather than Na^+ or K^+. It is remarkable that the different tap waters are situated more in the center of the diamond than the other water types. This means that there is more Cl^-, F^- and SO_4^{2-} and less HCO_3^- present in the water. The anion triangle reveals that mainly the Cl^- concentration, rather than the SO_4^{2-} concentration, is higher in tap water. This is probably because of Cl^- is formed after the addition chlorine [19]. This chlorine is added to tap water to inactivate micro-organisms and to prevent micro-organism regrowth during water flow in the pipes [9]. Most of the sparkling waters are situated in the left corner of the anion triangle. This indicates that (relative) large amounts of HCO_3^- are present because of the addition of CO_2 gas to sparkling water. The still waters are scattered but in general most of the still waters are located in the Ca^{2+} + HCO_3^- zone. The data presented in the Piper diagram are very similar to data originating from other studies as most of the waters are situated on the left side of the diamond in the Piper diagram [3,6]. Also, little difference can be seen between the different water types. As such, further statistical analyses are necessary.

3.2. ANOVA

As a first step, the data was submitted to a one-way ANOVA test. The resulting calculated F-values for the effect of the different water type brands (i.e., the different rows in Table 1) and the effect of ion composition (i.e., the different columns in Table 1) are compared to the tabulated critical F-value with a significance level of 0.05 in Table 2. The ANOVA test was performed on the total data set, as well as on the individual water types (bottled still water, bottled sparkling water and tap water). It can be seen that the calculated F-values are always larger than the critical F-values, indicating that there is a difference in ion composition between the different water types.

Table 1
Overview of different water data used

Bottled still water

Brand	Ca^{2+} (mg l^{-1})	Mg^{2+} (mg l^{-1})	K^{+} (mg l^{-1})	Na^{+} (mg l^{-1})	SO_4^{2-} (mg l^{-1})	HCO_3^{-} (mg l^{-1})	Cl^{-} (mg l^{-1})	F^{-} (mg l^{-1})	NO_3^{-} (mg l^{-1})	pH [–]	Ionic strength (eq l^{-1})	References
Evian	80	26	1	6.5	12.6	360	6.8	0	3.7	7.2	0.0096	www.evian.be
Acqua Panna	30.2	6.9	0.9	6.5	21.4	100	7.1	0.1	5.7	8.2	0.0036	www.sanpellegrino.com
Contrex	486	84	3.2	9.1	1187	403	10	0	2.7	nd	0.0596	www.contrex.com
Saint-Amand	176	46	5	28	372	312	37	1.3	0	nd	0.0241	www.saint-amand.com
Spa Reine	4.5	1.3	0.5	3	4	15	5	0	1.9	6	0.0007	www.spa.be
Valvert	67.6	2	0.2	1.9	18	204	4	0	3.5	7.7	0.0057	www.valvert.fr
Chaudfontaine	65	18	2.5	44	40	305	35	0.4	0	7	0.0095	www.chaudfontaine.com
Duke	10	6	0.5	5	33	30	7	0	0	nd	0.0021	www.inbev.be
Spontin	38	15	2.5	14	28	180	19	0	6.5	7.3	0.0058	www.spontin.eu
Top Bronnen	1.3	0.8	12	170	60	300	71.5	0	0.9	nd	0.0087	www.topbronnen.be
Romy 365	104	3.7	0	3.7	52	280	4	0	2	nd	0.0090	www.roman.be
Source de Ciseau	72	8	1.6	4	20	242	5	0	0.5	7.3	0.0068	www.softdrinksinternational.com
Louise	110	19	11	32	79	379	32	0	2	nd	0.0131	www.cristaline.tm.fr
L'oiselle	164	79	22	86	430	312	144	1	1	7.2	0.0304	www.saint-amand.com
Fiji	17	13	5	18	2	140	9	0	0.2	nd	0.0037	www.fijiwater.com
1 litre	52	16	0.04	2	20.5	176	2	0.1	0.2	nd	0.0059	www.1litre.com
Aurele	106	3.8	1.8	3.5	58.9	272	3.8	0.9	2	7.2	0.0092	www.cristaline.tm.fr
Orée du Bois	234	70	9	43	635	292	62	1.3	1	7.2	0.0350	www.saint-amand.com
Freu Hé	17	3.5	1	11	18	70	5	0	0.2	nd	0.0024	www.delhaize.be
Léberg	120	14	4	11	80	300	50	0	0	nd	0.0123	www.leberg.be
Ordal	143	7	5	17	91	317	42	0.1	1	nd	0.0132	www.ordal.be
Hépar	549	119	0	14.2	1530	383.7	0	0.1	4.3	7.2	0.0725	www.mont-roucous.com
Extreme H_2O	96.5	6.5	1	10	25.5	230	24.5	0	2.9	nd	0.0084	www.springwater.nl
Finé	9.7	4.7	1.7	8.7	5.7	52.8	5.3	0	0	7.8	0.0017	www.finejapon.com
Montcalm	3	0.6	0.4	1.5	8.7	5.2	0.6	0	0.9	nd	0.0005	www.saint-amand.com
Serra de estrella	2.71	0	0	5.5	0	8.7	7.9	0	0	6.2	0.0004	www.aguaserradaestrela.pt
Earth water	58.7	6.3	6.5	1.9	48	145	13.6	0.09	0.3	nd	0.0060	www.earth-water.org
Nestlé Aquarel	70	2.1	0	2	0	210	0	0	4	nd	0.0055	www.nestle.com
Swiss Alpina	159.7	48.7	4.4	65.5	90.4	770.1	9.3	0.23	1.3	nd	0.0218	www.coop.nl
San Benedetto	128.9	17.9	2.4	11.7	43.3	359.9	35.3	0	0.1	7.52	0.0125	www.sanbenedetto.it
Roche des Ecrins	63	10.2	0.4	1.4	51.3	173.2	1	0	2	7.6	0.0065	www.carrefour.eu

(Continued)

Table 1 (*Continued*)

Brand	Ca^{2+} (mg l^{-1})	Mg^{2+} (mg l^{-1})	K^+ (mg l^{-1})	Na^+ (mg l^{-1})	SO_4^{2-} (mg l^{-1})	HCO_3^- (mg l^{-1})	Cl^- (mg l^{-1})	F^- (mg l^{-1})	NO_3^- (mg l^{-1})	pH [–]	Ionic strength (eq l^{-1})	Reference
Vauban	230	66	8	40	620	280	58	1.3	1	7.2	0.0339	www.saint-amand.com
Emma (Pierval)	38	15	15	11	32	190	10.4	0.15	2	7.2	0.0060	www.carrefour.eu
Montfras	96	6.1	3.7	10.6	9.3	297	22.6	0	2	7.2	0.0085	www.springwater.nl
Kildevæld	76	14	5.9	30	39	244	52	0	1	nd	0.0092	www.carlsbergdanmark.dk
Albert Heijn	102	13	2.7	5	32	360	5	0.2	0	7.5	0.0100	www.ah.nl
Waterline	22.97	4.86	0.91	4.3	10.37	95.8	3.72	0	4.72	7.33	0.0027	Bong et al. (2009)
Soonsoo	21.24	0.74	0.25	13.2	29.2	66.69	5.57	0	1.18	7.45	0.0027	Bong et al. (2009)
Seoksu	22.32	2.82	1.25	3.93	6.8	79.93	5.35	0	6.76	7.23	0.0024	Bong et al. (2009)
Icis	12.46	4.69	0.63	1.51	4.55	58.88	2.71	0	2.05	7.08	0.0017	Bong et al. (2009)
Volvic	13.39	9.21	5.67	12.44	8.61	82.5	15.67	0	7.44	7.78	0.0029	www.volvic.fr
Yaksan	12.02	1.34	0.31	5.82	3.74	48.94	2.17	0	4.03	6.98	0.0014	Bong et al. (2009)
Dongwon	31.38	3.79	1.18	6.3	34.39	80.42	5.32	0	2.93	7.33	0.0035	Bong et al. (2009)
Elumgol	7.78	0.71	0.63	23.38	12.3	59.98	6.52	0	5.59	7.37	0.0018	Bong et al. (2009)
Sammool	13.43	5.11	0.63	1.54	4.74	60.59	2.84	0	1.78	7.3	0.0018	Bong et al. (2009)
Samdasoo	3.04	2.55	1.71	5.63	1.92	28.62	6.06	0	1.76	6.89	0.0009	Bong et al. (2009)
Power O_2	43.61	31.31	1.11	12.62	26.99	279.15	22.48	0	1.93	7.54	0.0082	Bong et al. (2009)
Whistler	12.46	1.31	0.57	2.51	7.08	33.93	5.89	0	0.84	7.01	0.0013	Bong et al. (2009)
Tynant	23.49	13.44	0.9	21.36	7.32	165.05	13.93	0	0.73	7.41	0.0045	Bong et al. (2009)
Average	82.05	17.45	3.20	17.40	121.54	200.80	18.47	0.15	2.01	6.99[1]	0.0104	

Bottled sparkling water

Brand	Ca^{2+} (mg l^{-1})	Mg^{2+} (mg l^{-1})	K^+ (mg l^{-1})	Na^+ (mg l^{-1})	SO_4^{2-} (mg l^{-1})	HCO_3^- (mg l^{-1})	Cl^- (mg l^{-1})	F^- (mg l^{-1})	NO_3^- (mg l^{-1})	pH [–]	Ionic strength (eq l^{-1})	Reference
Koningsbronnen	2.8	0.1	8.7	180	50	335	71	4.5	1	nd	0.0091	www.topbronnen.be
SPA Barisart	5.5	1.5	0.5	5	7.5	18	5.5	0	1.5	nd	0.0009	www.spa.be
Ordal	143	7	5	17	91	317	42	0.1	1	4.98	0.0132	www.ordal.be
Tönissteiner	157	126	16.1	106	30	1312	31.3	26	0	nd	0.0332	www.toenissteiner.de
San Pellegrino	185.6	52.5	0	35	443.8	238	59	0	0	6.08	0.0264	www.sanpellegrino.com
Appolinaris classic	90	120	30	470	100	1800	130	0	0	5.8	0.0436	www.apollinaris.de
Vichy	103	10	66	1172	138	2989	235	5	0	6.8	0.0631	www.vichy-celestins.com
Rogaska Donat Mg	393	1070	15	1544	2250	7750	80	0.17	0.9	nd	0.2529	www.donatmg.net
Ramlösa	3	0.5	2	210	6	520	21	2.7	0	nd	0.0095	www.ramlosa.se

Kaiserbrunnen	62	9	69	1295	277	876	1486	0	0	nd	0.0668	www.kaiserbrunnen.com
Badiot	190	85	10	150	40	1300	40	0	0	6	0.0319	www.badoit.be
River	2.8	0.1	8.7	180	50	335	71	4.5	1	nd	0.0091	www.topbronnen.be
Source de Ciseau	72	8	1.6	4	20	242	5	0	0.5	nd	0.0068	www.softdrinksinternational.com
Saint-Léger-du-Nord	64	26	21	93	70	478	32	0	2	nd	0.0135	www.aldi.be
Gerolsteiner	326.9	98.6	17.32	99.0	39.1	1728	37.5	0	5.38	6.63	0.0424	www.gerolsteiner.de
Solé	112.2	33.62	0.43	3.51	16.86	445.8	4.64	0	10.94	6.31	0.0126	Bong et al. (2009)
Montes	114.8	54.9	4.11	29.54	306.7	202.6	18.47	0	0.73	5.98	0.0193	Bong et al. (2009)
Perrier	181.9	6.18	1.07	12.3	40.73	435.3	26.48	0	6.39	6.12	0.0147	Bong et al. (2009)
Heidiland	117.9	39.92	0.56	4.35	257.6	85.73	1.5	0	2.07	5.65	0.0154	Bong et al. (2009)
San Benedetto	53.47	31.92	1.04	7.47	4.76	293.8	3.24	0	8.62	6.16	0.0081	Bong et al. (2009)
Tynant Bruis	24.09	13.4	0.85	21.79	7.22	124.0	13.83	0	0	5.81	0.0042	Bong et al. (2009)
Tau	23.41	13.15	0.85	21.98	6.93	125.8	13.65	0	0	5.86	0.0041	Bong et al. (2009)
Average	11.38	82.15	12.72	257.32	193.33	997.77	110.37	1.95	1.91	5.76	0.0319	

Tap water

Brand	Ca^{2+} (mg l^{-1})	Mg^{2+} (mg l^{-1})	K^{+} (mg l^{-1})	Na^{+} (mg l^{-1})	SO$_4^{2-}$ (mg l^{-1})	HCO$_3^{-}$ (mg l^{-1})	Cl^{-} (mg l^{-1})	F^{-} (mg l^{-1})	NO$_3^{-}$ (mg l^{-1})	pH [–]	Ionic strength (eq l^{-1}]	Reference
WPC Kooigem	118.92	35.5	11.61	33.15	109.44	441.53	35	0.82	3	nd	0.0162	www.vmw.be
WPC Zillebeke	59.38	11.91	12.46	42.94	67.85	139.54	82.95	0.25	4.77	nd	0.0088	www.pidpa.be
Ghent area	77	8.4	2.8	24.8	62.1	190.31	35.7	0.22	15.7	7.8	0.0086	www.vmw.be
WPC De Gavers	84.47	12.12	11.06	45.61	93.95	180.97	85.58	0.25	8.05	nd	0.0111	www.vmw.be
WPC St-leger	139.46	30.54	8.48	35.15	164.5	378.45	53.63	0.55	3	nd	0.0177	www.vmw.be
WPC De Blankaart	124.94	16.25	24.47	74	190.53	172.66	151.45	0.25	20.73	nd	0.0172	www.vmw.be
WPC Lommel	34.91	3.46	2.29	6.91	3.22	121.06	9.78	0.2	3	nd	0.0034	www.vmw.be
WPC Herent	152.9	12.79	2.27	46.49	131.33	302.09	107	0.2	3	nd	0.0165	www.vmw.be
WPC Holsbeek	58.78	4.95	4.76	8.94	39.43	139.74	24.14	0.2	3	nd	0.0059	www.vmw.be
Wuustwezel area	92	6.2	10.5	12.7	81.6	216	18.4	0.16	2	7.47	0.0093	www.pidpa.be
WPC kluizen	93.1	10.84	11.79	50.51	81.94	219.66	91.78	0.2	8.11	nd	0.0117	www.vmw.be
WPC Hasselt	68.64	6.6	3.2	34.25	70.17	169.64	30.8	0.2	26	7.83	0.0083	www.vmw.be
Mol area	18.3	2.6	3.2	22.2	7.1	108	10.1	0.086	1.3	8.38	0.0028	www.pidpa.be
Average	86.37	12.47	8.38	33.67	84.86	213.82	56.64	0.28	7.82	7.77	0.0106	

nd, not determined; WPC, Water Production Centre.

[1] Determined based on the average proton concentration.

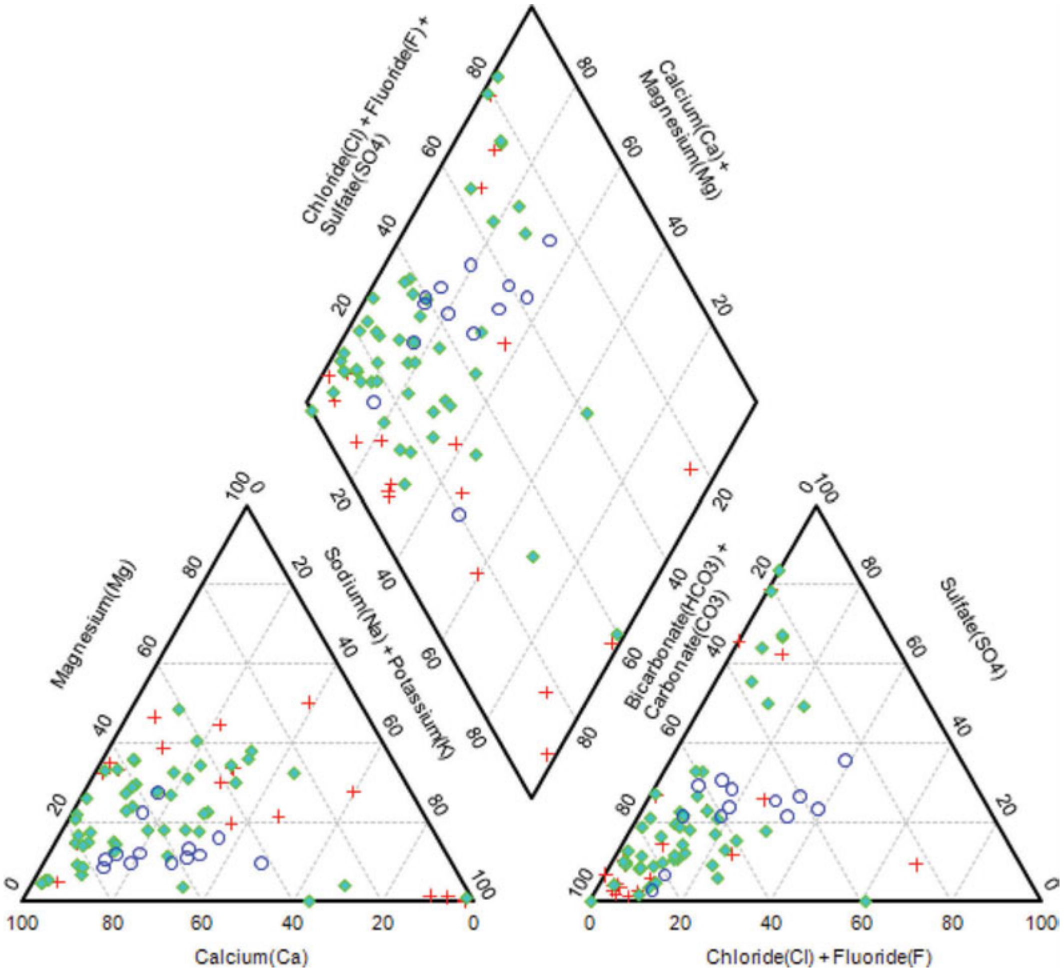

Fig. 2. Piper plot for bottled still waters (♦), bottled sparkling water (+) and tap water (○).

Table 2
Resulting F-values for the one-way ANOVA test without replication investigating the effect of water type brands and the effect of ion composition

Brand	Tabulated F value	Ca^{2+}	Mg^{2+}	K^{+}	Na^{+}	SO_4^{2-}	HCO_3^{-}	Cl^{-}	F^{-}	NO_3^{-}	pH	Ionic strength
Complete dataset	1.29	2202	3741	3050	3653	3466	3368	3719	3864	2774	1805	3150
Bottled still waters only	1.39	2578	2764	2636	2928	3483	1400	2676	3472	1973	1425	2644
Bottled sparkling water only	1.61	1877	3673	2899	3142	3570	3056	3704	3721	2984	1922	3031
Tap water only	1.84	711	1613	1497	976	1208	742	1530	1362	2120	3334	759

3.3. *Principal components analysis of water data*

In Table 3 the resulting rotated pattern of principal components (after varimax rotation) is presented. It was decided to use the three principal components with an eigenvalue larger than 1 (Kaiser criterion [16]) for further analysis. These three principal components can explain 84.8 % of the total variance (Table 3). Further, the screeplot (not shown here) had a clear elbow justifying retaining the first three components.

The Kaiser–Meyer–Olkim (KMO) criterium for sampling adequacy (KMO = 0.358) was not taken into account. However, visual inspection of the correlation

Table 3
The resulting rotated patterns of factors after varimax rotation

Variable	Factor		
	1	2	3
Ca^{2+}	0.425	0.845*	.094
Mg^{2+}	0.507	0.748*	–0.102
K^+	0.920*	–0.267	0.131
Na^+	0.902*	–0.343	0.019
SO_4^{2-}	0.321	0.826*	0.173
HCO_3^-	0.897*	–0.086	–0.058
Cl^-	0.866*	–0.176	0.129
F^-	0.715*	–0.339	0.003
pH	–0.256	–0.091	0.951*
Eigenvalue	4.33	2.31	1
% of explained variance	48.1	25.6	11.0

*Significant loadings.

matrix and the Barlett's test of sphericity ($\chi^2(36) \approx 6.48$; $p < 0.05$), both indicated that correlations between items were sufficiently large for PCA [17].

The first principal component explains 48.1% of the total variance and contains most of the information on the monovalent ions K^+, Na^+, HCO_3^-, Cl^- and F^-. The first principal component (PC1) can therefore be considered as a measure for the water saltiness [4]. The second principal component (PC2) explains 25.6% of the total variance and contains most of the information on the bivalent ions Ca^{2+}, Mg^{2+}, and SO_4^{2-}. As such, this principal component can be considered as a measure of the water hardness. The third principal component (PC3) explains 11% of the total variance and mostly contains information on the pH. As such, the following relation can be derived between the three principal components and the original variables:

$$\begin{cases} PC1 = 0.962\,[K^+] + 0.961\,[Na^+] + 0.853\,[HCO_3^-] \\ \qquad + 0.878\,[Cl^-] + 0.785\,[F^-] \\ PC2 = 0.946\,[Ca^{2+}] + 0.9\,[Mg^{2+}] + 0.865\,[SO_4^{2-}] \\ PC3 = pH \end{cases} \tag{4}$$

In Fig. 3 the resulting score plots are presented. It can be seen that bottled sparkling water can be distinguished from bottled still water and tap water by its pH value (PC3). Saltiness (PC1) and hardness (PC2) cannot be used to discriminate between different water types as all points in Fig. 2 (top) are situated near the origin. This is remarkable as PC3 only explains 11% of the variance and should therefore be not as discriminating as PC1 and PC2.

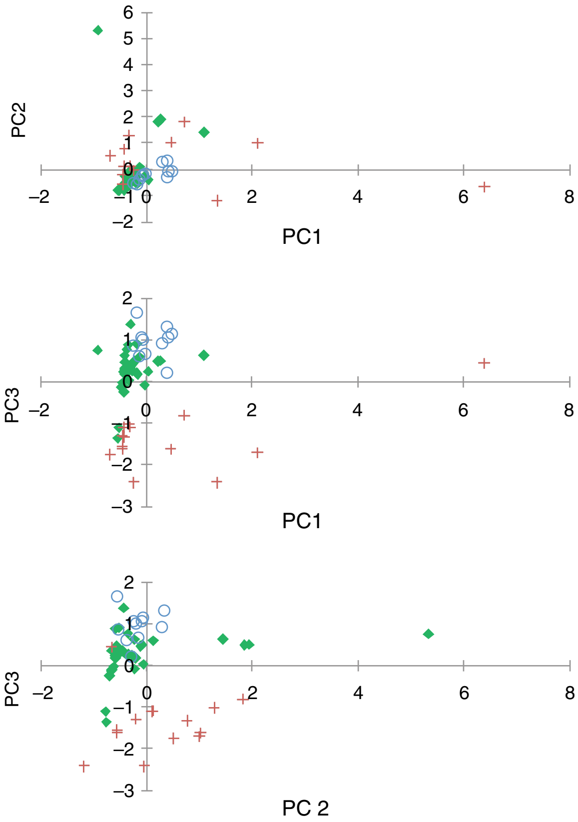

Fig. 3. Score plots of the PCA analysis with the water data (top: PC1 vs. PC2; middle: PC1 vs. PC3; bottom: PC2 vs. PC3): bottled still waters (♦), bottled sparkling water (+) and tap water (○).

It should further be noted that saltiness (PC1) and hardness (PC2) are two factors that can be used to discriminate between different brands of the same water type, as also suggested by Versari et al. [2]. A separation can be made between low-mineral, oligomineral and mineral still waters as it can be seen in Fig. 4 where the saltiness (PC1) and hardness (PC2) of bottled still water are represented. It can be seen that a positive linear correlation exists between saltiness and hardness.

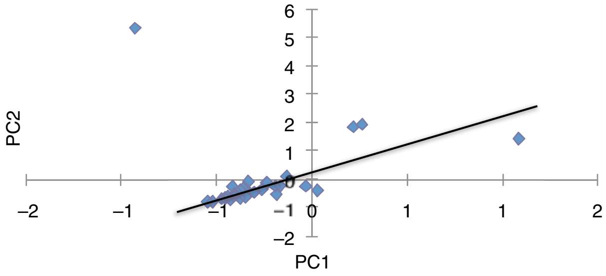

Fig. 4. Comparison of saltiness (PC1) and hardness (PC2) for bottled mineral water, demonstrating the positive correlation between these two components.

It can also be noted that on average (*see* Table 1) bottled sparkling water has a higher mineral content than non-sparkling water.

3.4. Discriminant analysis of water data

In Table 4 the canonical discriminant function coefficients for the different original variables are presented for the performed discriminant analysis. The eigenvalues of the discriminant functions are presented as well. In total 57 different water types (with measured pH) were used. In Fig. 5 the canonical discriminant functions for the different water types are presented.

It can be seen that there are two discriminant functions. Both discriminating functions are discriminating as the Wilk's Lambda is statistically significant for both functions at a significance level $\alpha = 0.05$. The first discriminant function (DF1) accounts for 94.5% of the differences among the three groups, while the second discriminant function (DF2) accounts for 5.5%. Note that DF2 differentiates between bottled still water and tap water. The variable pH is strongly correlated with DF1, while DF2 is mainly correlated with the chloride ion, confirming previous findings: chloride and pH are the main discriminating factors between the types of water.

In total 52 out of 57 water types were classified correctly. This corresponds with 91.2% of the water types. This indicates that a considerable amount of discriminatory information is accounted for by the discriminatory variables.

Table 4
Canonical discriminant function coefficients (β_i) for the different original variables and eigenvalues of the discriminant functions

Variable	Function	
	1	2
Ca^{2+}	0.077	0.063
Mg^{2+}	0.144	0.070
K^+	−0.056	0.057
Na^+	0.068	0.040
SO_4^{2-}	−0.034	−0.025
HCO_3^-	−0.022	−0.017
Cl^-	−0.054	0.000
F^-	0.512	0.002
pH	−2.242	0.335
Constant value	14.86	−3.480
Eigenvalue	6.9	0.4
% of explained variance	94.5	5.5

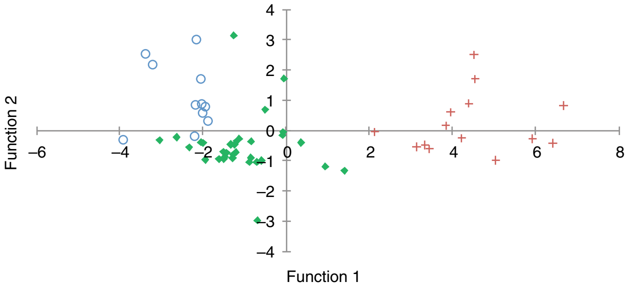

Fig. 5. The canonical discriminant functions for the different water types): bottled still waters (♦), bottled sparkling water (+) and tap water (○).

4. Conclusions

This study attempted at separating different water types (bottled still water, bottled sparkling water and tap water) based on their chemical composition. Different techniques such as ANOVA, hierarchical cluster analysis, PCA and discriminant analysis were used.

Based on the ANOVA it could be concluded that there exists a difference between the different water types and that there also is exists a difference between different water brands. Principal component analysis revealed that the three main differences between different water types are saltiness, hardness and pH. Based on pH it was possible to discriminate between sparkling water and non-sparkling water, although only 11% of the variance is explained by this component. It was not possible to divide the different water types based on saltiness or hardness, but it could be demonstrated that different types of water exist (low-mineral, oligomineral and mineral).

The Piper diagram and discriminant analysis revealed that the main discriminating factor for tap water is the slightly increased chloride content. It was otherwise not possible to differentiate tap water from bottled still water, indicating the fact that the chemical composition is not statistically different.

References

[1] J.K. Edzwald, Water Quality and Treatment, 6th Ed. McGraw-Hill, New York (2010).
[2] H. Tanaka and Y.S. Ho, Global trends and performances of desalination research, Desalin. Water Treat., 25 (2011) 1–12.
[3] Y.-S. Bong, J.-S. Ryu and K.-S. Lee, Characterizing the origins of bottled water on the South Korean market, Anal. Chim. Acta, 631 (2009) 189–195.
[4] A. Versari, G.P. Parpinello and S. Galassi, Chemometric survey of Italian bottled mineral waters by means of their labelled physico-chemical and chemical composition, J. Food Compos. Anal., 15 (2001) 251–264.
[5] A. Ikem, S. Odueyungbo, N.O. Egiebor and K. Nyavor, Chemical quality of bottled waters from three cities in eastern Alabama, Sci. Total Environ., 285 (2002) 165–175.

[6] C. Güler, Characterization of Turkish bottled waters using pattern recognition methods, Chemom. Intell. Lab. Syst., 86 (2007) 86–94.
[7] A. Azoulay, P. Garzon and M.J. Eisenberg, Comparison of the mineral content of tap water and bottled waters, J. Gen. Intern. Med., 16 (2001) 168–175.
[8] A. Misund, B. Frengstad, U. Siewers and C. Reimann, Variation of 66 elements in European bottled mineral waters, Sci. Total Environ., 243–244 (1999) 21–41.
[9] S. Puget, N. Beno, C. Chabanet, E. Guichard and T. Thomas-Danguin, Tap water consumers differ from non-consumers in chlorine flavor acceptability but not sensitivity, Water Res., 44 (2009) 956–964
[10] R.A. Freeze and J.A. Cherry, Groundwater, Prentice-Hall, Englewood Cliffs, NJ (1979).
[11] A.M. Piper, A graphic procedure in the geochemical interpretation of water analyses, Trans. Am. Geophys. Union, 25 (1944) 914–923.
[12] N.S. Rao, MHPT.BAS: a computer program for modified Hill–Piper diagram for classification of ground water, Comput. Geosci., 24 (1998) 991–1008
[13] W. Back, Hydrochemical facies and ground-water flow patterns in northern part of Atlantic Coastal Plain. U.S. Geological Survey Professional Paper 498-A, 42 (1966).
[14] G. Currell and A. Dowman, Essential Mathematics and Statistics for Science. Wiley, New York (2005).
[15] R.G. Brereton, Chemometrics, Data Analysis for the Laboratory and Chemical Plant. Wiley, UK (2003) 487p.
[16] H.F. Kaiser, The application of electronic computers to factor analysis, Educ. Psychol. Meas., 20 (1960) 141–151.
[17] J.F. Hair, R.E. Anderson, R.L. Tatham and W.C. Black, Multivariate Data Analysis, 5th Ed. Prentice-Hall, Upper Saddle River (1998).
[18] E.M. Smeti, N.C. Thanasoulias, E.S. Lytras, P.C. Tzoumerkas and S.K. Golfinopoulos, Treated water quality assurance and description of distribution networks by multivariate chemometrics, Water Res., 43 (2009) 4676–4684.
[19] G. Tchobananoglous, F.L. Burton and H.D. Stensel, Wastewater Engineering: Treatment and Reuse, 4th Ed. Metcalf and Eddy Inc. and McGraw-Hill, New York (2003).