Statistical Analysis of Raw and Potable Water Quality from a Preparation Plant, as a Basis to Purpose the Re-schedule of the Physico-Chemical Current Routine Analysis Periodicity

DANUT GABRIEL COZMA[1], CARMEN LUIZA COSTULEANU[2*], ION SANDU[3], DOINA LUTIC[1*]

[1]Al. I. Cuza University of Iasi, Faculty of Chemistry, 11 Carol I Blvd., 700506, Iasi, Romania

[2]Ion Ionescu de la Brad University of Agricultural Sciences and Veterinary Medicine of Iasi, Department of Agrobusiness, 3, Mihail Sadoveanu Alley, 700490, Iaºi, Romania

[3]Romanian Inventors Forum, 3 Sf. P. Movila Str., L11, III/3, 700089, Iasi, Romania

The multivariate statistics analysis was used to analyze the correlations between 24 physical and chemical parameters values, for raw and potable water quality characterization, using data collected monthly from a potable water preparation plant during two years, 2015 and 2016. The t-Test (Student) and the Principal Component Analysis were used in order to analyze if there are tight dependences between the analyzed parameters, allowing eventually to purpose the analysis scheme schedule modification.

Keywords: raw water, potable water, physical parameters, chemical parameters, multivariate statistical analysis

The potable water quality in Romania is regulated by the State Standard STAS 1342-91, which is in line with the European requirements defined by Directive 98/83/CE [1]. Moreover, the detailed water parameters to be measured and the agreed methods in this respect are precisely defined by the Romanian Association for Standardization and international regulations on these topics [2]. The high number of parameters used for water characterization led to their grouping in several categories as: organoleptic, physical, chemical, bacteriological, biological and radioactive ones.

The relative importance and relevance of the physical and chemical parameters depends on their effects on the final consumer health and comfort. This is why the frequency and the detailed schedule the water analysis states some parameters to be analyzed several times a day, while others are measured once a week or even more seldom. Even after this selection, the measurements remain numerous and the global characterization of the potable water quality involves a high volume of work and costly chemicals and experimental infrastructure for the routine analysis. It is also worth to mention that the chemicals used to prepare the water analysis are sometimes intrinsic toxics and their use in analysis generates pollution of the wastewater resulted from laboratories.

The multivariate statistics analysis method has been elaborated in 1958 by Anderson [3], developed consistently and later further reviewed by Schervish (1987) [4]. The aim of the method is to allow highlighting the effects of several variables influencing a system, relating with each others, establish if the statistical results are or not representative and eventually reduce the number of variables, if there are strong dependencies in between. Several books available in Romanian [5-7] make easier for the Romanian reader the theoretical background used in the results interpretation. A big number of models can be used to perform the statistical analysis and draw results [8], each one adapted to the specific database available and aim of the data processing. In the literature, the statistical interpretation on subjects similar to our approach is widely used in data processing concerning water quality from different geographical areas and eventual toxic contamination [9-16], safe CO_2 sequestration within rocks [17] with respect of water action as a solvent for this gas, possibility of rainwater use in household applications or agriculture [18-20] or water management studies [21-22].

Experimental part

In our work, we performed a statistical analysis of the experimental results collected in two years of water analysis performed at an industrial potable water preparation plant. We were looking for tight correlations between the analyzed parameters, allowing eventually eliminating or thinning from the daily analysis schedule the ones acting as dependent variables.

The interpretation of a high number of experimental data using the statistical analysis can generate interesting, improved results by applying the multivariate statistical analysis techniques. The factorial analysis is used as a multivariate statistical technique, targeting to reduce the number of predictor variables which characterize the quality of the potable water.

In the computer era, the mathematical models for different data processing are integrated in specific software, allowing fast and complex data processing and interpretation. In our work, the SPSS 14.0 software [23] was used to process statistically the behavior of water parameters values, using the monthly average values during two years (22 parameters for 2015 and 24 parameters for 2016), for raw water from an artificial lake designed for the preliminary water clarifying procedure (denoted as *RW*) and potable water, obtained after the standard potabilization procedures applied in the plant (denoted "PW" respectively), at Darmanesti plant from Bacãu County. The t-Test (Student), Kaiser-Meyer-Olkin (KMO) statistics and Principal Component Analysis (PCA) were applied in our work.

In other words, we aim to answer if a data set of average monthly measured experimental results of analysis package (the 22 parameters listed in table 1 for 2015 and 24 parameters for 2016 from Table 2, as predictor

---

variables), defining the basic physico-chemical behavior of water, could be reduced to less. We strive to find out if the data could be *aggregated* so as to define a number of factors less than the number of months per year, depending obviously on the weather conditions, undefined now in anyway in the database. The factorial analysis purposes to establish relations between the old variables, able to generate a more correctly defined set of interactions than the empirical periodical measurement procedure (currently, the calculation of monthly average values) of the various physico-chemical water parameters. The present case provides an example of exploratory factorial analysis.

In order to reach the purposed targets, we applied:

a) The comparison of the mean values of the physico-chemical indicators between the *R* and *P* types, respectively, for each of the two years. The test is statistically significant at a standard risk factor of 5%;

b) The exploratory factorial analysis, for each year and each type of water, starting from the 22 or 24 variables (the average monthly values of the physico-chemical indicators measured for each year and type of water) in order to reduce the number of variables.

Results and discussions

The database values used in this study are displayed in table 1 and table 2.

The values from the database indicate that the potabilization procedure applied in the plant is efficient, since even when raw water has poor quality regarding some of the parameters (especially seasonal variations related to weather), the potabilization procedures brings the water within the respective parameters values in the proper ranges for potable water. Therefore, our study could just confirm the obtaining for the whole observed period the obtaining of good potable water.

When performing the t-Test for independent samples from raw water (R/RW) and potable water (P/PW), the variables were treated as independent samples. The mean-group is the average value of the group/sample, t is the value of t statistic (Student), df is the number of liberty degrees associated to t-test and p is the significance level of the test. The mean values of each group, t-value, df an p are diplayed in table 3 and table 4 for the years 2015 and 2016, respectively.

The values from the tables 3 and 4 show that the test is, respectively, statistically significant in 10 of the 22 comparisons (lines in Bold) for year 2015 and in 12 from 24 comparisons (lines in Bold) for year 2016 and is not statistically significant in the other 12 cases. The highest differences are found, respectively, in the 2015 series for the P*ermanganate Index*, *p*H and *Total Iron* and, respectively, in the 2016 series for the same indicators together with *Ammonium*. Between the pairs of values of

Table 1
DATABASE FOR WATER QUALITY PARAMETERS IN 2015

Analyzed indicators	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
R Temperature, °C	3.4	2.4	3.3	5.7	7.8	8.9	9.9	10.2	8.7	8.3	9.3	7
P Temperature, °C	3.5	2.5	3.5	5.7	8	9	10	10.5	8.6	8.4	9.8	7.4
R Turbidity, NTU	10.3	9.46	18.43	31.5	14.3	7.03	6.09	4.99	4.7	5.08	3.32	3.64
D Turbidity, NTU	4.42	3.7	4.82	1.75	1.85	2.58	4.2	4.38	3.52	3.14	1.82	2.13
R Total solids, mg/L	133	135	136	115	116	115	117	120	120	125	128	129
P Total solids, mg/L	135	135	141	119	120	113	118	120	120	126	127	130
R Total Suspended Solids, mg/L	10	9	18	29	15	9	4.75	4	4.5	6	6	4
P Total Suspended Solids, mg/L	7	7	7	2.25	2	4	3.5	3	3.2	4	2	3
R pH	7.8	7.8	7.7	7.7	7.7	7.7	7.5	7.5	7.5	7.3	7.8	7.8
P pH	7.6	7.6	7.1	7	7.1	7.2	7.3	7.4	7.4	7.4	7.6	7.7
R Total alkalinity, mval/L	1.68	1.7	1.7	1.5	1.4	1.4	1.4	1.41	1.46	1.48	1.6	1.6
P Total alkalinity, mval/L	1.63	1.7	1.4	1.1	1.15	1.26	1.36	1.4	1.43	1.47	1.6	1.6
R Acidity, mval/L	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
P Acidity, mval/L	0.14	0.1	0.29	0.3	0.25	0.17	0.12	0.1	0.12	0.14	0.1	0.1
R Conductivity, µ/cm	205	207	209	172	178	177	180	183	184	187	197	199
P Conductivity, µ/cm	207	208	217	184	184	175	181	185	185	188	196	200
R Permanganate index, mgO2/L	4	3.95	4	3.8	3.75	3.75	3.67	3.78	3.71	5.96	3.71	3.48
P Permanganate index, mgO2/L	2.92	2.79	2.55	1.89	2.24	2.58	2.71	2.83	2.87	2.76	2.86	2.74
R Dissolved Oxygen, mg/L	11.53	11.91	11.87	11.6	11.03	11.07	10.28	9.47	9.82	9.61	9.59	10.18
P Dissolved Oxygen, mg/L	11.96	12.44	12.35	11.97	11.38	11.26	10.63	9.89	10.07	10.13	10.24	10.99
R Total Hardness, °d	5.51	5.68	5.72	5.15	4.48	4.48	5.03	4.89	4.88	5.02	5.16	5.18
P Total Hardness, °d	5.51	5.68	5.7	5.13	4.82	4.86	5.04	4.83	4.87	5	5.16	5.18
R Calcium, mg/L	33.56	34.21	34.28	29.72	28.19	28.41	28.78	28.72	29.4	29.9	30.55	21.15
P Calcium, mg/L	33.56	34.21	34.21	29.72	28.19	28.25	28.8	28.72	29.42	29.87	30.55	31.15
R Magnesium, mg/L	3.58	3.92	4.05	4.35	3.93	3.9	4.39	3.83	3.3	3.66	3.68	3.7
P Magnesium, mg/L	3.58	3.92	4.1	4.35	3.84	3.89	4.74	3.81	3.32	3.63	3.69	3.6
R Total Iron, µg/L	87	67	122	187	100.5	49	48.4	46	56	60	55	50
P Total Iron, µg/L	47	52	50	16	18.6	23	33	36	36	45	37	38
R Manganese, µg/dm3	38	47	113	175	88.8	47	54.5	45	42	57	26	41
P Manganese, µg/dm3	23	38	40	17	12.3	20	37.2	38	36	44	23	22
R Ammonium, mg/L	0.016	0.018	0.033	0.04	0.028	0.018	0.015	0.014	0.015	0.017	0.014	0.014
P Ammonium, mg/L	0.014	0.015	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014
R Nitrates, mg/L	2.31	2.44	2.88	2.81	3.58	3.2	2.93	2.97	2.96	2.92	2.1	1.95
P Nitrates, mg/L	2.17	2.26	2.4	2.36	2.7	2.73	2.72	2.79	2.85	2.76	1.99	1.83
R Nitrites, mg/L	0.012	0.017	0.027	0.054	0.03	0.014	0.009	0.01	0.01	0.008	0.0014	0.012
P Nitrites, mg/L	0.01	0.013	0.008	0.004	0.003	0.006	0.005	0.006	0.006	0.005	0.004	0.004
R Chloride, mg/L	5.46	5.46	5.83	4.64	4.9	5.15	5.83	5.49	5.13	5.43	4.79	4.97
P Chloride, mg/L	5.46	5.46	5.79	4.24	4.63	4.93	5.85	5.44	5.09	5.02	4.77	4.97
R Fluorine, mg/L	0.22	0.42	0.42	0.4	0.24	0.33	0.38	0.31	0.09	0.3	0.25	0.52
P Fluorine, mg/L	0.18	0.32	0.32	0.2	0.11	0.17	0.33	0.22	0.2	0.41	0.17	0.41
R Sulfides, mg/L	0.009	0.007	0.019	0.016	0.012	0.018	0.007	0.006	0.005	0.006	0.004	0.005
P Sulfides, mg/L	0.006	0.006	0.013	0.006	0.005	0.007	0.005	0.004	0.004	0.004	0.003	0.003
R Sulfates, mg/L	19	18	20	15	17.46	17	19	18	19	16	16	20
P Sulfates, mg/L	25	20	34	32	33	26	25	22	22	20	20	20

Months

Parameter												
R Temperature, °C	4.7	4.6	4.7	5.8	7.3	9.8	12.5	12.6	11.7	9.8	9.1	6.6
P Temperature, °C	4.7	4.7	4.9	5.9	7.4	9.7	11.8	12.3	11.5	9.9	9.2	7.1
R Turbidity, NTU	3.38	4.28	3.76	13.3	18.3	74.6	27.07	22.7	16.9	21.3	16.2	13.3
P Turbidity, NTU	2.41	2.71	2.58	2.82	1.18	1.15	1.23	1.24	1.23	1.08	0.99	0.89
R Total solids, mg/L	131	140	141	130	128	104	104	103	107	112	123	125
P Total solids, mg/L	132	140	136	143	136	117	125	118	114	118	129	131
R Total Suspended Solids, mg/L	4	6	6	12	18	74	21	14	13	19	19	13
P Total Suspended Solids, mg/L	3	3	3	4	4	2	2	2	2	2	2	2
R pH	7.8	7.9	7.9	7.9	7.9	7.8	7.8	7.7	7.7	7.6	7.8	7.9
P pH	7.7	7.8	7.8	7.5	7	6.8	6.7	6.7	7.1	7	7.2	7.2
R Total alkalinity, mval/L	1.62	1.69	1.7	1.7	1.59	1.3	1.2	1.2	1.2	1.32	1.5	1.52
P Total alkalinity, mval/L	1.62	1.69	1.7	1.47	1.05	0.72	0.7	0.7	0.82	0.96	1.2	1.22
R Acidity, mval/L	0.1	0.09	0.09	0.09	0.09	0.08	0.08	0.09	0.1	0.1	0.1	0.1
P Acidity, mval/L	0.1	0.09	0.09	0.19	0.28	0.31	0.3	0.3	0.27	0.25	0.25	0.25
R Conductivity, µ/cm	202	215	209	214	197	160	159	159	164	172	190	192
P Conductivity, µ/cm	204	216	209	219	209	163	181	182	176	182	199	201
R Permanganate index, mgO2/L	3.26	3.53	3.42	4.02	3.94	4.74	4.7	5.1	4.85	4.6	4.4	3.92
P Permanganate index, mgO2/L	2.47	2.56	2.63	2.69	2.16	1.94	1.77	1.76	2.19	2.17	2.11	2.03
R Dissolved Oxygen, mg/L	11.23	12.21	12.1	11.75	11.65	10.37	8.95	8.65	8.23	8.09	9.57	10.9
P Dissolved Oxygen, mg/L	11.78	12.51	12.3	11.64	11.57	10.38	10.25	10.19	9.78	9.87	10.61	11.4
R Total Hardness, °d	5.27	5.41	5.47	5.51	5.22	4.33	4.3	4.4	4.43	4.67	5.12	5.3
P Total Hardness, °d	5.25	5.41	5.47	5.5	5.21	4.46	4.3	4.4	4.43	4.67	5.12	5.3
R Calcium, mg/L	32.73	33.01	32.76	32.35	31.25	25.23	24.3	24.6	24.7	25.8	28.7	28.6
P Calcium, mg/L	32.73	33.01	32.76	33.32	31.27	26.18	24.3	24.6	24.7	25.8	28.7	28.6
R Magnesium, mg/L	3.07	3.42	4.36	4.3	3.7	3.65	3.77	4.25	4.25	4.6	4.8	5.1
P Magnesium, mg/L	3.07	3.42	4.36	4.18	3.7	3.42	3.89	4.28	4.25	4.6	4.8	5.1
R Total Iron, µg/L	51	94	93	235	241	504	518	522	423	586	529	337
P Total Iron, µg/L	42	79	71	88	110	110	184	189	101	27	22	17
R Manganese, µg/dm3	33	44	38	122	122	507	272	199	153	196	191	116
P Manganese, µg/dm3	25	30	31	91	40	19	23	12	12	12	12	8
R Ammonium, mg/L	0.014	0.014	0.014	0.023	0.025	0.087	0.04	0.026	0.021	0.026	0.023	0.018
P Ammonium, mg/L	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014
R Nitrates, mg/L	2.15	2.01	2.33	2.63	2.89	3.21	3.16	2.93	3.09	2.99	2.57	2.61
P Nitrates, mg/L	2.01	1.93	2.2	2.48	2.52	2.55	2.47	2.55	2.68	2.4	2.22	2.24
R Nitrites, mg/L	0.011	0.012	0.012	0.027	0.03	0.11	0.039	0.033	0.027	0.033	0.026	0.02
P Nitrites, mg/L	0.003	0.005	0.004	0.004	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003
R Chloride, mg/L	4.93	4.58	4.38	4.38	4.38	4.38	4.38	4.38	4.38	4.38	4.38	4.38
P Chloride, mg/L	4.91	4.57	4.38	4.38	20.45	24.59	24.76	23.65	9.18	4.38	4.38	4.38
R Fluorine, mg/L	0.31	0.33	0.34	0.26	0.29	0.13	0.23	0.26	0.15	0.32	0.41	0.27
P Fluorine, mg/L	0.25	0.32	0.18	0.42	0.25	0.13	0.15	0.31	0.23	0.26	0.44	0.21
R Sulfides, mg/L	0.011	0.011	0.011	0.015	0.011	0.083	0.094	0.017	0.019	0.022	0.013	0.011
P Sulfides, mg/L	0.011	0.011	0.011	0.015	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
R Sulfates, mg/L	19	21	18	22	22	27	19	23	21	21	22	22
P Sulfates, mg/L	21	22	20	34	22	19	12	14	19	35	31	31
R Aluminium, µg/L	8	8	9	27	26	30	17	15	22	36	31	26
P Aluminium, µg/L	17	12	14	79	31	11	9	7	54	74	81	69

Table 2
DATABASE FOR WATER QUALITY PARAMETERS IN 2016

Table 3
t-TEST (STUDENT) COMPARING THE VALUES OF THE PARAMETERS OF PHYSICO-CHEMICAL ANALYSIS BETWEEN R AND P WATER (2015)

R vs. P Compared values, 2015	Mean – Group 1	Mean – Group 2	t-value	df	p
Temperature, °C	7.075	7.242	-0.148	22	0.883
Turbidity, NTU	**9.903**	**3.193**	**2.810**	**22**	**0.010**
Total solids, mg/L	124.083	125.333	-0.373	22	0.713
Total Susp. Solids, mg/L	**9.938**	**3.996**	**2.668**	**22**	**0.014**
pH, units	**7.650**	**7.367**	**3.477**	**22**	**0.002**
Total alkalinity, mval/L	1.527	1.425	1.576	22	0.129
Acidity, mval/L	**0.100**	**0.161**	**-2.783**	**22**	**0.011**
Conductivity, µ/cm	189.833	192.500	-0.506	22	0.618
Permanganate index, mgO2/L	**3.963**	**2.645**	**6.406**	**22**	**0.000**
Dissolved Oxygen, mg/L	10.663	11.109	-1.172	22	0.254
Oxygen saturation, %	87.483	90.667	-1.757	22	0.093
Total Hardness, °d	5.098	5.148	-0.338	22	0.739
Calcium, mg/L	29.739	30.554	-0.676	22	0.506
Magnesium, mg/L	3.858	3.872	-0.106	22	0.917
Total Iron, µg/L	**77.325**	**35.967**	**3.285**	**22**	**0.003**
Manganese, µg/dm3	**64.525**	**29.208**	**2.814**	**22**	**0.010**

Nitrates, mg/L	2.754	2.463	1.732	22	0.097
Nitrites, mg/L	**0.017**	**0.006**	**2.618**	**22**	**0.016**
Chloride, mg/L	5.257	5.138	0.670	22	0.510
Fluorine, mg/L	0.323	0.253	1.592	22	0.126
Sulfides, mg/L	**0.009**	**0.006**	**2.304**	**22**	**0.031**
Sulfates, mg/L	**17.872**	**24.917**	**-4.377**	**22**	**0.000**

Table 4

t-TEST (STUDENT) COMPARING THE VALUES OF THE PARAMETERS OF PHYSICO-CHEMICAL ANALYSIS   BETWEEN R AND P WATER (2016)

R vs. P Compared values, 2016	Mean – Group 1	Mean – Group 2	t-value	df	p
Temperature, °C	8.267	8.258	0.01	24	0.995
Turbidity, NTU	**19.591**	**1.626**	**3.28**	**24**	**0.003**
Total solids. mg/L	120.500	128.250	-1.55	24	0.136
Total Suspended Solids. mg/L	18.250	43.833	-1.65	24	0.113
pH	**7.808**	**7.208**	**4.98**	**24**	**0.000**
Total alkalinity. mval/L	**1.462**	**1.154**	**2.42**	**24**	**0.024**
Acidity. mval/L	**0.092**	**0.223**	**-5.34**	**24**	**0.000**
Conductivity. µ/cm	186.083	195.083	-1.10	24	0.284
Permanganate index. mgO2/L	**4.207**	**2.207**	**10.07**	**24**	**0.000**
Dissolved Oxygen. mg/L	10.308	11.023	-1.36	24	0.186
Oxygen saturation. %	**85.333**	**92.833**	**-3.12**	**24**	**0.005**
Total Hardness. °d	4.952	4.960	-0.04	24	0.970
Calcium. mg/L	28.669	28.831	-0.11	24	0.914
Magnesium. mg/L	4.106	4.089	0.07	24	0.946
Total Iron. µg/L	**344.417**	**87.000**	**4.38**	**24**	**0.000**
Manganese. µg/dm3	**166.083**	**26.250**	**3.67**	**24**	**0.001**
Ammonium. mg/L	**105.000**	**26.274**	**5.74**	**24**	**0.000**
Nitrates. mg/L	**2.354**	**2.714**	**-2.71**	**24**	**0.013**
Nitrites. mg/L	**51.002**	**0.032**	**3.31**	**24**	**0.003**
Chloride. mg/L	89.126	53.509	2.02	24	0.056
Fluorine. mg/L	0.275	0.263	0.35	24	0.732
Sulfides. mg/L	**94.418**	**42.939**	**2.93**	**24**	**0.008**
Sulfates. mg/L	21.417	23.333	-0.83	24	0.414
Aluminium. µg/L	21.250	38.167	-1.82	24	0.082

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.503	95.859	95.859	11.503	95.859	95.859
2	0.459	3.822	99.681			
3	0.028	0.229	99.910			
4	0.007	0.060	99.970			
5	0.002	0.014	99.984			
6	0.001	0.010	99.994			
7	0.000	0.004	99.997			
8	0.000	0.001	99.999			
9	8.32 E-005	0.001	100.000			
10	2.65 E-005	0.000	100.000			
11	9.00 E-006	7.50 E-005	100.000			
12	4.68 E-006	3.90 E-005	100.000			

Table 5

TOTAL VARIANCE EXPLAINED FOR R (2015) SAMPLES

the above mentioned indicators, the values of the t-test are the most significant, since in these cases, the risk coefficient used for the comparison had been decreased from 0.05 to 0.01 or even 0.001, which means t-test statistically significant even at 0.001 ($p < 0.001$).

In the beginning, the opportunities of applying Principal Component Analysis (PCA) should be assessed [5]. The application of PCA should fulfill a series of conditions: Coakes' criteria - at least 5 subjects per variable, a sample of 100 subjects (acceptable) or 200 subjects (preferably). On another hand, Steiner and Gorsuch [6] recommend 10 cases per variable when the sample is less than 100 cases. In our case, we have 23 subjects/12 variables, both for *P* and *R* series and both years). Howitt and Cramer [7] recommend 2-3 times more subjects (cases) than the number of variables. In the present study, the variables are values of the physico-chemical indicators, measured on each of the two types of water. Thus, certain grounds do not recommend the use of PCA. Total Variance Explained for P (2015) and R (2015) samples is presented in table 5 and table 6, respectively.

The value of the correlation matrix determinants (1.55 E-035 for the R series and 4.98 E-040 for P series) is lower than E-005, indicating that the variables correlate very strongly and the risk of multi-colinearity is a contraindication for the method. According to tables 5 and 6, in the case of R (2015) and P (2015), only one factor is extracted, therefore the method is rejected.

Testing the independence of the investigated variables is necessary; therefore, the following hypotheses are formulated: H0= the hypothesis of independence, when the matrix of correlations is a unit matrix; H1= the addiction hypothesis.

The choice of one of the two hypotheses that exclude each other is based on the KMO and Bartlett test (table 7).

The result of the Bartlett sphericity test with a significance level $p < 0.001$ shows that the correlation matrix differs significantly from the unit matrix (in which the variables would not correlate with each other) and the variables are suitable for factorial analysis; the data set confirm the choice of the hypothesis H1 against the H0.

Another possibility for the adequacy verification of the factorial analysis for the data set of sample R (2016) is applying the Kaiser-Meyer-Olkin (KMO) statistics as a measure to select a representative sample, the value should be higher than 0.5 (in our study the value is 0.770). Thus, the existence of significant statistical connections is proven and the PCA can be applied.

However, the value of the correlation matrix determinant (3.26 E-033) is less than E-005, which bring the risk of multicollinearity for the sample R (2016) and is a contraindication for applying this method. Discussions on the Bartlett and KMO tests are, however, encouraging to perform the PCA.

According to table 8, in the case of 2016-R, two factors are extracted. The way of grouping (rallying) the 12 initial variables (months of 2016) in two new derived variables, referred further on to as factors, is discussed subsequently.

Figure 1 shows the values of the own vectors (eigenvalues) associated to each of the 12 former variables. The existence of only two clearly defined factors is observed, with eigenvalues of 9.83 and 1.84, respectively.

In table 9 are presented the variance values of the new variables, since after the application of PCA standardized variables, mean zero and variance one are obtained.

The values in table 9 are the square values of the multiple correlation coefficients between the variable and the factors (more precisely, R^2 from the multiple regression where the studied variable plays the role of the dependent variable and the extracted factors play the role of the independent variable).

Based on the result of figure 2, the results from 12 months 2016-R are grouped in two unequal groups conventionally named *hot* and *cold*, depending on the result of the factorial analysis. The purpose of this procedure is to assign to the temperature the role of *predictor*, which is

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.932	99.432	99.432	11.932	99.432	99.432
2	0.049	0.404	99.836			
3	0.014	0.113	99.949			
4	0.004	0.030	99.979			
5	0.002	0.015	99.993			
6	0.000	0.004	99.998			
7	0.000	0.002	99.999			
8	3.73E-005	0.000	100.000			
9	2.69E-005	0.000	100.000			
10	1.24E-005	0.000	100.000			
11	6.06E-006	5.05E-005	100.000			
12	1.28E-006	1.06E-005	100.000			

Table 6
TOTAL VARIANCE EXPLAINED FOR P (2015) SAMPLES

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.770
Bartlett's Test of Sphericity	Approx. Chi-Square	1358.954
	df	66
	Sig.	0.000

Table 7
KMO AND BARTLETT'S TEST FOR R (2016) SERIES OF SAMPLES

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.830	81.920	81.920	9.830	81.920	81.920
2	1.942	16.183	98.102	1.942	16.183	98.102
3	0.187	1.555	99.657			
4	0.037	0.311	99.969			
5	0.002	0.015	99.984			
6	0.001	0.011	99.995			
7	0.000	0.003	99.998			
8	0.000	0.001	99.999			
9	0.000	0.001	100.000			
10	2.98E-005	0.000	100.000			
11	1.20E-005	0.000	100.000			
12	5.26E-006	4.39E-005	100.000			

Table 8
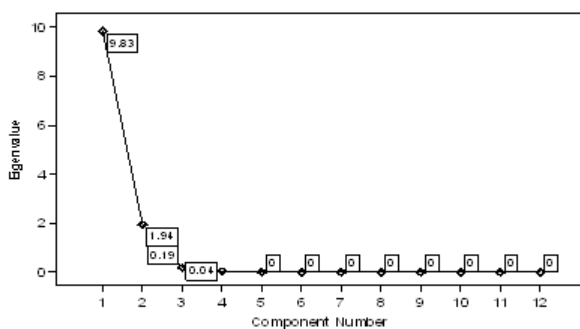TOTAL VARIANCE EXPLAINED
FOR R (2016)

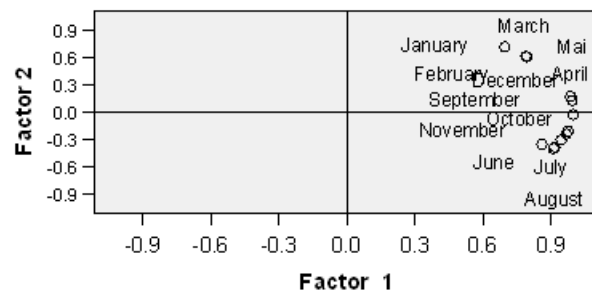Fig. 1. Extracted factors and corresponding eigenvalues

Fig. 2. Positions of former variables on the first two factorial axes

Table 9
VARIANCES OF THE STANDARDIZED VARIABLES

Initial Variable	Component	
	1	2
January	0.694	0.718
February	0.792	0.610
March	0.787	0.616
April	0.983	0.175
Mai	0.991	0.129
June	0.859	-0.354
July	0.909	-0.400
August	0.914	-0.383
September	0.973	-0.208
October	0.943	-0.305
November	0.965	-0.240
December	0.994	

grouping all the physical-chemical analysis results from the 2016-R data set into two large categories, the temperature averages being responsible for a statistically significant difference.

Table 10 is evidencing the *hot* and *cold* labels for 12 months obtained for raw water set, based on factorial analysis.

When performing the t-test between the two temperature *subgroups*, the obtained results are presented in table 11.

The t test is statistically significant at a standard risk factor of 1% (t = -3.288, p <0.01). Therefore, it is justified to attribute the *hot* and *cold* characteristics for the 12 months from the set 2016-R; the temperature is, in this case, only a predictor of the data packet separation in two distinct entities.

The results obtained for Total Variance Explained for potable water (2016 samples) are included in table 12.

As in the cases of R (2015) and P (2015) sample series, the value of the correlation matrix 1.71 E-032 is lower than E -005 indicates that the variables correlate strongly between each other, meaning that there is risk of multi-colinearity. According to Table 12, in the series P (2016)

Table 10
ASSIGNING *HOT* AND *COLD* LABELS FOR 12 MONTHS (DATA SET R (2016)) BASED ON THE FACTORIAL ANALYSIS

Months	Factor	Months	Factor	Months	Factor	Months	Factor
January	cold	April	cold	July	hot	October	hot
February	cold	May	cold	August	hot	November	hot
March	cold	June	hot	September	cold	December	cold

Compared values, „cold" vs „hot", 2016-R	Mean – cold	Mean – hot	t-value	df	p
R Temperature, °C	6.486	10.760	-3.288	10	0.008

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.054	92.121	92.121	11.054	92.121	92.121
2	0.776	6.469	98.590			
3	0.088	0.731	99.321			
4	0.068	0.564	99.885			
5	0.009	0.073	99.958			
6	0.002	0.019	99.977			
7	0.002	0.018	99.995			
8	0.000	0.004	99.999			
9	5.47E-005	0.000	100.000			
10	4.13E-005	0.000	100.000			
11	1.48E-005	0.000	100.000			
12	5.42E-007	4.52E-006	100.000			

only one factor is extracted, therefore the method is rejected.

Such type of extended and deepened statistical analysis is used in many fields and is absolutely necessary for data processing [24-43], related also to environmental issues [44-57]. Resolving the environmental issues induced by pollution will increase the quality of human life, as well as of the environment [58].

Conclusions

The databases used for this study confirms in a synthetic and easy-to-follow manner the proper efficiency of the treatment procedure for good quality potable water preparation in the potabilization plant.

The statistic significant differences between the water types R and P, investigated by applying the t-test (Student), appear in just over half of analyzes. Among the subsets of *R* and *P*, there exist statistically significant differences in about half of the tested physical and chemical indicators, in terms of the result of the t (Student) test, both for 2015 and for 2016.

For samples/data sets 2015-R, 2015-P and 2016-P, the performing of Principal Component Analysis (PCA) show that the values for the investigated indicators do not group with each others.

For the 2016-R data set sample, the values determined for the water quality indicators values could be grouped in two categories named *hot* and *cold*, labeled in general as connected to climatic differences (temperature). The assignment of October and November in the *hot* range, and of April, May and September in the *cold* area, respectively, is related to the temperature characteristics of 2016 and falls into the theoretical predictions provided by the factorial analysis.

In conclusion, the processing of the water parameters values from the database highlighted that the correlations between the analyzed parameters were not strong enough to allow the re-scheduling of some analyses more seldom than regulated now.

References
1. *** DIRECTIVE 98/83/CE CONCERNING POTABLE WATER (available at: http://eur-lex. europa.eu/legal-content/RO/TXT/PDF/?uri= CELEX:31998L0083&from=RO, accessed November 9th, 2017).
2. *** ROMANIAN ASSOCIATION FOR STANDARDIZATION, List of Standards referring to water quality (available at http://www.asro.ro/ ?p=3864, accessed November 9th, 2017).
3. ANDERSON, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons Inc., New York, 1958.
4. SCHERVISH, M.J., Stat. Sci., **2**, no. 4, 1987, p. 396.
5. POPA, M., Multivariate Statistics Applied in Psychology (Statistici multivariate aplicate in psihologie), Polirom Editing House, Iasi, 2010, p.291.
6. PINTILESCU, C., Multivariate Statistical Analysis (Analiza statistica multivariata), Al. I. Cuza University Editing House, Iasi, 2007, p. 53.
7. LABAR, A.V., SPSS for Education Sciences (SPSS pentru stiintele educatiei), Polirom Editing House, Iasi, 2008, p. 306.
8. HOWITT, D., CRAMER, D., Introduction in SPSS for Psychology (Introducere in SPSS pentru psihologie), Polirom Editing House, Iasi, 2006, p. 191.
9. https://en.wikipedia.org/wiki/Multivariate_statistics
10. MAGESH, N.S., CHANDRASEKAR, N., ELANGO, L., Chemosphere, **185,** 2017, p. 468.
11. VINCE, F., AOUSTIN, E., BRÉANT, P., MARECHAL, F., Desalination, **220**, 2008, p. 37.
12. AYOKO, G.A., SINGH, K., BALEREA, S., KOKOT, S., J. Hydrol., **336**, 2007, p.115.
13. MHLONGO, S., MATIVENGA, P. T., MARNEWICK, A., J. Clean. Prod., **171**, 2018, p. 446.
14. KARKRA, R., KUMAR, P., BANSOD, B. K. S., KRISHNA, C. R., Procedia Computer Science, Proceedings of the 6th International Conference on Advances in Computing and Communications, Kochi, India, **93,** 2016, p. 988.
15. CAVALCANTE, Y.L., HAUSER-DAVIS, R.A., SARAIVA, A.C.F., BRANDAO, I.L.S., OLIVEIRA, T.F., SILVEIRA, A.M., Sci. Total Environ., **442**, 2013, p. 509.
16. MATIATOS, I., Sci. Total Environ., **541**, 2016, p. 802.

17. GIMENEZ-FORCADA, E., VEGA-ALEGRE, M., TIMON-SANCHEZ, S., Sci. Total Environ., **593–594,** 2017, p. 211.

18. IRANMANESH, A., LOCKE, R.A., WIMMER, B.T., Energy Procedia, 12th International Conference on Greenhouse Gas Control Technologies, (GHGT), Austin, TX, USA, **63,** 2014, p. 3182.

19. SANTOS SILVA, A., GHISI, E., J. Environ. Manage., **180,** 2016, p. 82.

20. VIALLE, C., SABLAYROLLES, C., LOVERA, M., JACOB, S., HUAU, M.C., MONTREJAUD-VIGNOLES, M., Water Res., **45,** 2011, p. 3765.

21. KELLY MARINOSKI, A., FORGIARINI RUPP, R., GHISI, E., J. Environ. Manage., **206,** 2018, p. 28.

22. WEN-SHYONG LIN, T., LEE, M., HUANG, Y.C., DEN, W., Conserv. Recycl., **94,** 2015, p. 35.

23. SPSS BASE 14.0 USER'S GUIDE, Copyright © 2005 by SPSS Inc.

24. LUTIC, D., COROMELCI, C.G., JUZSAKOWA, T., CRETESCU, I., Environ. Eng. Manag. J., **16,** no. 4, 2017, p. 801.

25. DELETZE, E., ANTONIADIS A., KITSIOU, V., KOSTOPOULOU, E., LUTIC, D., CRETESCU, I., POULIOS, I., Desalin. Water Treat., **57,** no. 40, 2016, p. 18897.

26. SIMION, V.A., CRETESCU, I., LUTIC, D., LUCA, C., POULIOS, I., Environ. Eng. Manag. J., **14,** no. 3, 2015, p. 595.

27. HRISTODOR, C., COPCIA, C., LUTIC, D., POPOVICI, E., Rev. Chim. (Bucharest), **61,** no. 3, 2010, p. 285.

28. LUTIC, D., CRETESCU, I., Rev. Chim. (Bucharest), **67,** no. 1, 2016, p. 134

29. LUTIC, D., COROMELCI-PASTRAVANU, C., CRETESCU, I., POULIOS, I., STAN, C.D., Int. J. Photoenergy, 2012, art. no. 475131, DOI: 10.1155/2012/475131.

30. SEFTEL, E.M., COOL, P., LLOYD-SPETZ, A., LUTIC, D., CR Chim., **17,** no. 7-8, 2014, p. 717.

31. SEFTEL, E.M., COOL, P., LUTIC, D., Mater. Sci. Eng. C-Master. Biol. Appl., **33,** no. 8, 2013, p. 5071.

32. SANDU, I.C.A., HAYASHI, M., VASILACHE, V., COZMA, D.G., PRUTEANU, S., URMA, M., SANDU, I., Rev. Chim. (Bucharest), **66,** no. 4, 2015, p. 587.

33. ANTOCI, V., MANTU, D., COZMA, D.G., USRU, C., MANGALAGIU, L.I., Med. Hypotheses, **82,** no. 1, 2014, p. 11.

34. SANDU, I., POPESCU-CIOCIRLIE, I., GRUIA, I., CALANCIA, O., COSMA D.G., SANDU, I.G., COZMA, D.G., Rev. Chim. (Bucharest), **49,** no. 11, 1998, p. 756.

35. COZMA, D.G., CRUCEANU, A., COJOC, G.M., MUNTELE, I., MIHU-PINTILIE, A., International Multidisciplinary Scientific GeoConference-SGEM; Water Resources, Forest, Marine and Ocean Ecosystems, SGEM 2015, vol. 1, 2015, p. 625.

36. CRUCEANU, A., COJOC, G.M., COZMA, D.G., MUNTELE, I., MIHU-PINTILIE, A., International Multidisciplinary Scientific GeoConference-SGEM; Water Resources, Forest, Marine and Ocean Ecosystems, SGEM 2015, vol. 1, 2015, p. 159.

37. PUI, A., TANASE, C., COZMA, D.G., BALAES, T., Environ. Eng. Manag. J., **12,** no. 3, 2013, p. 527.

38. TANASE, C., PUI, A., OLARIU, R., COZMA D.G., Rev. Chim. (Bucharest), **59,** no. 5, 1998, p. 479.

39. CECAL, A., COLISNIC, D., POPA, K., PARASCHIVESCU, A.O., BILBA, N., COZMA, D.G., Cent. Eur. J. Chem., **2,** no. 1, 2004, p. 247-U1.

40. GHERCA, D., CIOCARLAN, R.G., COZMA, D.G., CORNEI, N., NICA, V., SANDU, I., PUI, A., Rev. Chim. (Bucharest), **64,** no. 8, 2013, p. 848.

41. CAPLANUS, I., SUNEL, V., BABAN, C.I., COZMA, D.G., PUI, A., RUSU, G.I., Rev. Chim. (Bucharest), **60,** no. 12, 2009, p. 1247.

42. PUI, A., COZMA, D.G., PUI, M., Mat. Plast., **42,** no. 2, 2005, p. 138.

43. HUMELNICU, D., COZMA, D.G., HUMELNICU, I., Rev. Chim. (Bucharest), **54,** no. 9, 2003, p. 759.

44. COSTULEANU, C.L., VINTU, C.R., ROBU, D.A., IGNAT, G., BREZULEANU, S., Rev. Chim. (Bucharest), **66,** no. 5, 2015, p. 743.

45. COSTULEANU, C.L., BREZULEANU S., IGNAT, G., BOLDUREANU, G., TOMA, M.C., ANDRUSEAC, G.G., Rev. Chim. (Bucharest), **67,** no. 10, 2016, p. 1990.

46. COSTULEANU, C.L., BOLDUREANU, G., ANDRUSEAC, G.G., Rev. Chim. (Bucharest), **68,** no. 5, 2017, p. 1058.

47. PAHONIE, R.C., STEFAN, A., COSTULEANU, C.L., BOLDUREANU, D., ANDRUSEAC, G.G., Mat. Plast., **54,** no. 1, 2017, p. 155.

48. PAHONIE, R.C., STEFAN, A., ADOCHIEI, I.R., COSTULEANU, C.L., ANDRUSEAC, G.G., UNGUREANU, G., SARDARU, D.P., Mat. Plast., **54,** no. 2, 2017, p. 396.

49. IGNAT, G., COLIBABA, C., COSTULEANU, C.L., MORARU, I., VINTU, C.R., UNGUREANU, G., BEJINARIU, C., Mat. Plast., **54,** no. 3, 2017, p. 586.

50. ANDRUSEAC, G.G., PASARICA, A., BREZULEANU, C.O., IGNAT, G., BREZULEANU, S., COSTULEANU, C.L., Rev. Chim. (Bucharest), **68,** no. 6, 2017, p. 1357.

51. IGNAT, G., COLIBABA, C., COSTULEANU, C.L., BALAN, A., ROTARU, L., SANDU, I.C.A., Mat. Plast., **53,** no. 3, 2016, p. 367.

52. IGNAT, G., BALAN, G., SANDU, I., COSTULEANU, C.L., VILLE, S.T.S., Rev. Chim. (Bucharest), **67,** no. 8, 2016, p. 1560.

53. COSTULEANU, C.L., BOBITAN, N., DUMITRESCU, D., Environ. Eng. Manag. J., **14,** no. 6, 2015, p. 1423.

54. UNGUREANU, G., IGNAT, G., BOGHITA, E., COSTULEANU, C.L., VINTU, C.R., BODESCU, D., BEJINARIU, C., Rev. Chim. (Bucharest), **68,** no. 10, 2017, p. 2350.

55. COSTULEANU, C.L., IGNAT, G., BREZULEANU, O., UNGUREANU, G., ROBU, D., VINTU, C.R., BOGHITA, G., BREZULEANU, S., Rev. Chim. (Bucharest), **68,** no. 11, 2017, p. 2597.

56. UNGUREANU, G., BOGHITA, G., IGNAT, G., COSTULEANU, C.L., SANDU, A.V, BEJINARIU, C., VINTU, C.R., Rev. Chim. (Bucharest), **68,** no. 11, 2017, p. 2662.

57. PAHONIE, C.R., LARCO, C., MIHAILA ANDRES, M., NASTASESCU, V., BARBU, C., COSTULEANU, C.L., Mat. Plast., **54,** no. 4, 2017, p. 768.

58. COSTULEANU, C.L., DUMITRESCU, D., BREZULEANU, S., BOBITAN, N., Rev. Cercet. Interv. Soc., **48,** 2015, p.50.