Bulletin UASVM Animal Science and Biotechnologies, 68(1-2)/2011
Print ISSN 1843-5262; Electronic ISSN 1843-536X

Factor Analysis on Physical-Chemical Parameters of Wastewater from Medias's Treatment Plant

Dana M. POPA1), Letiția OPREAN2), Lorentz JÄNTSCHI3)

1)University of Agricultural Sciences and Veterinary Medicine, 3-5 Mănăştur Street,
400372 Cluj-Napoca, Romania; danampopa@gmail.com
2)"Lucian Blaga" University, 5-7 Ion Rațiu Street, 550012 Sibiu, Romania; letitia.oprean@ulbsibiu.ro
3)Technical University, 103-105 Muncii Bd., 400641 Cluj-Napoca, Romania; lori@academicdirect.org

Abstract. The physical-chemical parameters of wastewater at the entrance and exit of Medias's wastewater treatment plant during 2009-2011 were determined. Association analysis, factor analysis, multiple linear regression analysis, and t-Test (paired two samples for means) were applied in order to establish a model to explain the pollutants removal degree and the variation in treated water quality. The strongest association occurs between chemical and biochemical oxygen demand. The results indicate that ammonia also has a great influence on the efficiency of wastewater treatment process. The regression analysis leads to a statistically significant model explaining the interdependence of the parameters.

Keywords: wastewater, wastewater treatment plant (WWTP), dendrogram, factor analysis, correlation coefficient, regression analysis

INTRODUCTION

Municipal wastewater comprises of water (99.9 %) along with small concentrations of suspended and dissolved inorganic and organic substances. Among the organic substances present in sewage are carbohydrates, lignin, fats, soaps, synthetic detergents, proteins and their decomposition products, as well as various natural and synthetic organic chemicals from the process industries (Jäntschi, 2003).

The continuous monitoring of wastewater provides useful knowledge about the past and future changes in the biosphere; a major role of the water quality playing effects is on bacterial communities (Egwari *et al.*, 2002; Zhu *et al.*, 2010). The changing of the environmental parameters may change the life cycle of populations from an entire species (Woollett and Hedrick, 1970).

Then the wastewater is used for irrigation, other risks may occur when the quality of water becomes poor (Binu Kumari *et al.*, 2006).

The analysis of the water, as well as for pharmaceutical products is, from a while now, a regulated area, and should be conducted through procedures approved via legal standards (Hatton & Gibb, 1999).

A typical analysis of wastewater includes total (sometimes split into dissolved and suspended) solids, nitrogen, phosphorus, chloride, alkalinity, grease, and BOD (biochemical oxygen demand at 20°C over 5 days) water content.

In the present paper, a series of physical-chemical parameters of wastewater, before and after water treatment at Medias's purification station, recorded over two years and a half were included into analysis, in order to find relevant relationships between them.

MATERIALS AND METHODS

The recorded data at every month (for a total period of 30 months) at Medias's wastewater treatment plant (WWTP) include pH, total suspended matter (MTS), biochemical oxygen demand in five days – BOD (CBO5), chemical oxygen demand – COD (CCOCr), and ammonium (NH_4^+) and are presented in *Table 1*. Values exceeding maximum standard limits (according to NTPA 001/2002 and NTPA 002/2002, respectively) in *Table 1* are in bold.

Tab. 1

Evolution of parameters in wastewater treatment plant of Mediaş, during 2009-2011

Year	Month	pH		MTS		CBO_5		CCOCr		NH_4^+	
		inflow	outflow	inflow	outflow	inflow	outflow	inflow	outflow	inflow	outflow
2009	Jan	7.17	7.29	110.0	4.6	89	**37**	240.3	106.4	**55.7**	**39.0**
	Feb	7.36	7.56	140.6	14.0	118	18	322.8	62.4	**80.7**	**30.0**
	Mar	7.39	7.34	204.0	16.6	108	22	287.6	72.1	**74.9**	**30.7**
	Apr	7.25	7.29	104.0	19.3	69.0	22	296.5	68.6	**76.9**	**50.0**
	May	7.21	7.43	113.0	6.0	129	**38**	396.6	118.2	**70.2**	**32.0**
	Jun	7.31	7.28	53.2	12.3	72	**30**	210.6	92.7	**56.0**	**27.3**
	Jul	7.36	7.34	60.4	29.6	107	**34**	286.2	111.4	**61.1**	**33.4**
	Aug	7.27	7.33	124.4	**58.0**	118	**38**	316.1	108.6	**65.6**	**37.0**
	Sep	7.26	7.27	122.0	26.4	110	**41**	320.2	118.6	**63.0**	**33.0**
	Oct	7.17	7.29	179.0	21.3	208	**30**	426.6	107.7	**74.8**	**31.0**
	Nov	7.32	7.22	105.0	**43.2**	112	**33**	320.2	118.0	**63.1**	**31.3**
	Dec	7.34	7.25	90.4	29.6	92	**39**	323.0	102.4	**68.3**	**33.3**
2010	Jan	7.35	7.28	126.0	10.4	154.0	11.0	387.3	52.4	**37.40**	**22.60**
	Feb	7.35	7.07	100.8	16.4	87.2	10.0	343.4	49.0	**31.60**	**24.50**
	Mar	7.41	7.07	136.4	17.6	104.0	9.0	324.2	51.6	**37.20**	**28.70**
	Apr	7.47	7.01	110.8	13.2	94.0	5.5	263.0	39.8	**43.70**	**32.30**
	May	7.47	7.23	118.0	20.7	127.0	6.0	315.6	28.2	25.80	**25.30**
	Jun	7.41	7.28	102.8	4.0	124.0	7.0	276.5	47.7	28.40	**25.00**
	Jul	7.29	7.06	102.8	10.0	92.0	5.0	248.4	36.6	29.30	**26.70**
	Aug	7.41	7.31	115.6	13.3	86.0	11.0	340.3	49.6	**33.30**	**30.50**
	Sep	7.23	7.13	119.2	6.0	92.0	15.0	329.6	61.4	**31.10**	**29.20**
	Oct	7.36	7.02	120.5	16.8	76.0	**26.0**	348.0	82.8	**32.70**	**26.40**
	Nov	7.36	7.05	109.0	18.0	110.0	8.0	272.5	28.8	**32.65**	**26.35**
	Dec	7.50	7.07	62.0	30.8	64.0	4.0	273.0	49.8	27.25	**25.95**
2011	Jan	7.17	7.27	128.0	30.0	90	**26**	282.3	94.2	**35.5**	**27.4**
	Feb	7.37	7.30	122.0	30.8	126	**28**	280.4	91.3	**33.0**	**29.2**
	Mar	7.28	7.25	109.2	32.4	95	20	283.5	94.4	**32.0**	**29.1**
	Apr	7.31	7.25	86.8	35.5	80	18	235.0	72.6	29.2	**27.5**
	May	7.29	7.33	76.4	34.4	78	15	226.6	64.4	**30.8**	**27.7**
	Jun	7.32	7.03	153.2	6.8	176	16	426.5	53.4	**39.4**	**29.8**

The conducted analysis consisted from an association analysis (in which a single linkage tree dendrogram was constructed) (Johnson, 1967; Milligan, 1980), a factor analysis (in which principal component analysis was involved) (Browne, 1968; Schönemann and Steiger, 1976), a multiple linear regression analysis (involving the strongest observed association of variables) (Kvalseth, 1985; Jäntschi and Bolboacă, 2007), and a mean based comparison (involving Student's t test) (Student, 1908; Hassainia *et al.*, 1994).

RESULTS AND DISCUSSION

The association analysis was conducted with 1-r as classification measure, where r represents the Pearson's correlation coefficient (Pearson, 1900). The results are depicted in *Figure 1*.

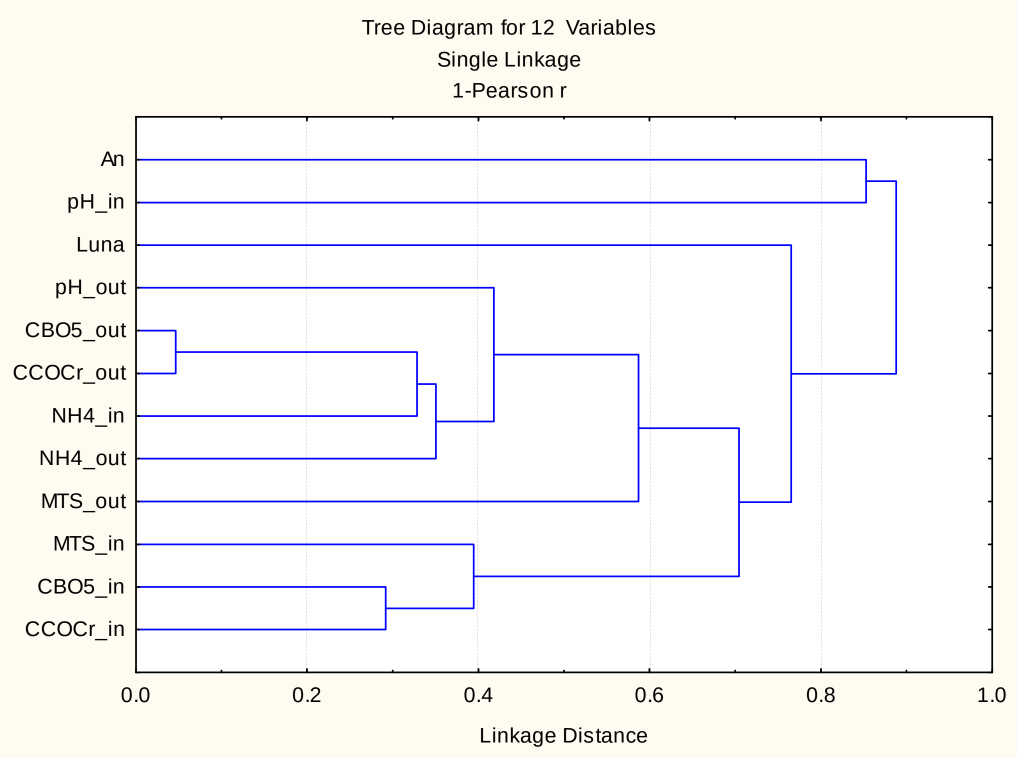

Fig. 1. Single linkage between input (*_in) and output (*_out) values of the observed variables

As can be seen in *Figure 1* (in which year (An) and month (Luna) variables were also included in order to reveal if periodicity is the main factor in the given observable) the strongest association is between CBO5 (BOD–biochemical oxygen demand at 20°C in 5 days) and CCOCr (COD–chemical oxygen demand) values at the output of WWTP (being below 0.1 degree of dissimilarity). The next link is at about 0.3 degree of dissimilarity (about 0.7 correlations) and is established between the same observable, but at the input of the purification station. An interesting association is established between the values of CBO5 and CCOCr at the output of the station and the value of NH_4^+ at the input of the station (the third by relevance association in Fig. 1), suggesting that somehow the efficiency of the purification in the treatment station is strongly affected by the amount of the ammonium in treated water. Total suspended matter at input (MTS_in) of WWTP also connects with the cluster formed by CBO5_in and CCOCr_in (almost 0.4 degree of dissimilarity), suggesting the influence of season on the quality of inflow. That means that quantity of total suspended matter, which increases along with increase of rain amount, affects oxygen exchange within surface water.

The factor analysis was constructed with year and month as supplementary classifiers variables and with the remaining ones (pH, ammonium, biochemical oxygen demand) as active variables. The next figure (Fig. 2) depicts the explanatory degree of the principal components in the values of the variables depending on the number of the components.

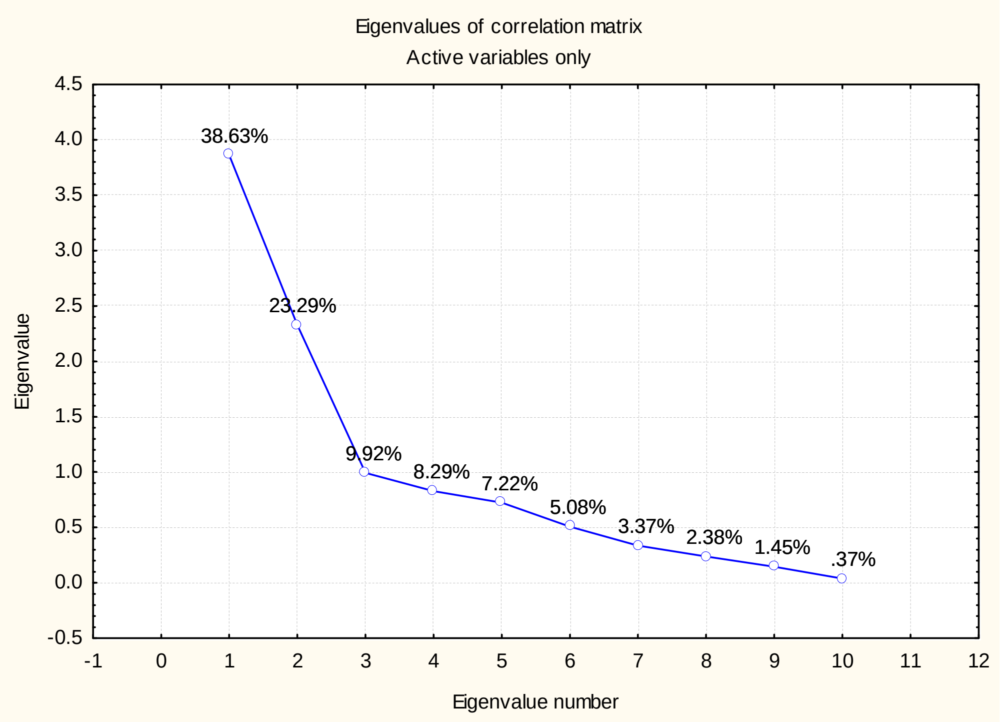

Fig. 2. Explained variance by each consecutive component in principal component analysis

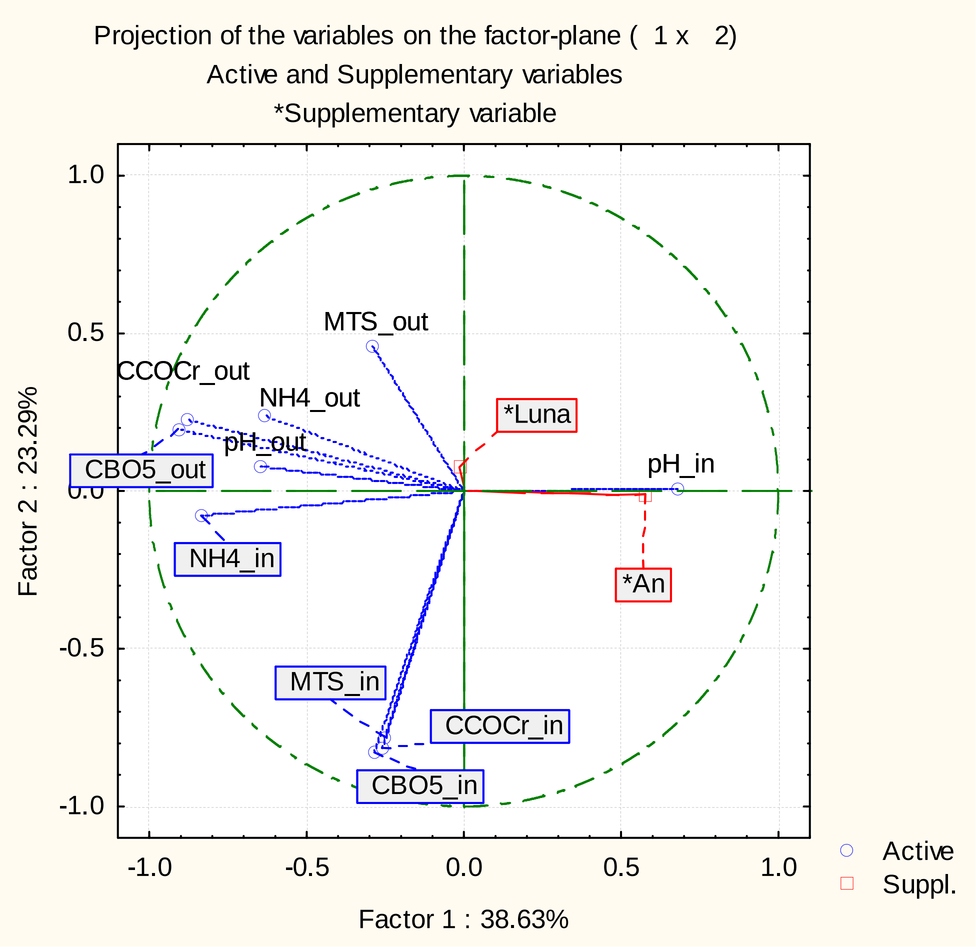

Fig. 3. Projection of the variables in the plane of the first two principal components

The values depicted in *Figure 2* reveals (as other authors noted too: Condit *et al.,* 1996; Chamberlain *et al.,* 1999) that no more than three principal components should be assigned to the deterministic factors in monitored parameters. It is easy to see how the factors starting from factor 3 (with explained variance of 9.92 %) till factor 10 (with explained variance of 0.37 %) fit to a straight line and their explanatory variance is only an expression of overfitting (Hawkins, 2004).

The next figure (Fig. 3) depicts the projection of the variables on the plane of the first two principal components (explaining over 60 % of the variance present in the observed values of the active variables).

The analysis of the projections given in *Figure 3* reveals a strong association of the input value of the pH at the input of the WWTP with the year (An) factor as well as that both are along the first component of the total variability. This result suggests that the main factor giving variation in the observed variables is the value of pH, and this factor is a slow changing one, varying in approximately good agreement with the year of the observation. Excepting CBO5 and CCOCr at the input of the purification station and the MTS (total suspended matter) at input and output of the station, the rest of the variables have a strong negative linear relationship (revealed by their projection along the first component in its negative part of the axis) with the pH value at the input of the purification station.

By entering the input CBO5 and CCOCr active variables and month supplementary classifier variable into a multiple linear regression analysis (given in Tab. 2), a relationship between CBO5 and CCOCr can be established (Tab. 2).

Tab. 2

Regression analysis of input values for biochemical oxygen demand

Model statistic	CCOCr_in = Linear(month,MTS_in,CBO5_in)	CCOCr_in = Linear(CBO5_in)
Multiple R	0.75004	0.708387425
R Square	0.56256	0.501812744
Adjusted R Square	0.51208	0.484020341
Standard Error	37.9980	39.07542249
Observations	30	30
F	11.14542	28.20370
pF	6.923E-05	1.187E-05
Intercept	135.46	176.087
t(intercept)	4.123	6.87337
month	3.185	-
t(month)	1.510*	-
MTS_in	0.439	-
t(MTS_in)	1.499*	-
CBO5_in	0.967	1.230
t(CBO5_in)	3.376	5.3107
* the corresponding parameter is not statistically significant and was deleted (see right model)		

The analysis from *Table 2* reveals a statistically significant relationship established between the CBO5 and CCOCr input variables. As statistical analysis of the obtained model (CCOCr_in=176(±53)+1.23(±0.57)·CBO5_in) based on the adjusted value of the determination coefficient (Giraitis *et al.,* 2003) gives (Adjusted R Square=0.48) about 48% from the variance in CCOCr_in variable is explained by the variance in CBO5_in variable.

Comparison of paired values (whenever this comparison has a meaning) is given in Table 3. In most cases, the comparison reveals differences, which were statistically significant (see Table 2). But in a series of comparisons, it reveals that the values did not distinguish at 5% risk being in error. Thus, for pH values, no significant differences (which can be put into account of not enough observed data or not different sampled population) were observed between input and output values of the pH in year 2009, between input values from years 2009 and 2011, and between output in 2010 and 2011 and input in year 2011 values. This pattern of the pH values suggests that the remained ones (including input values from years 2010 and 2011) were somehow affected by external factors (such as a major pollution).

Tab. 3

Paired Two Sample for Means Test comparing variables during 2009 - 2011 years

pH	pH_in_2009	pH_out_2009	pH_in_2010	pH_out_2010	pH_in_2011	pH_out_2011
pH_in_2009		0.040(2.2E-1)	0.100(8.2E-3)		0.008(7.7E-1)	
pH_out_2009				0.193(6.3E-4)		0.127(3.1E-2)
pH_in_2010				0.256(3.5E-5)	0.120(1.2E-2)	
pH_out_2010						0.0817(3.3E-1)
pH_in_2011						0.0517(3.9E-1)
pH_out_2011						
NH_4^+	NH4_in2009	NH4_out2009	NH4_in2010	NH4_out2010	NH4_in2011	NH4_out2011
NH4_in2009		33.53(8.19E-8)	34.992(1.6E-8)		35.75(1.5E-3)	
NH4_out2009				7.042(4.7E-4)		6.383(1.5E-1)
NH4_in2010				5.575(1.0E-3)	0.7(8.5E-1)	
NH4_out2010						2.05(2.4E-1)
NH4_in2011						4.867(1.4E-2)
NH4_out2011						
CBO5	CBO5_in2009	CBO5_out2009	CBO5_in2010	CBO5_out2010	CBO5_in2011	CBO5_out2011
CBO5_in2009		79.167(1.1E-5)	10.15(4.9E-1)		3.35(9.5E-1)	
CBO5_out2009				22.042(7.0E-6)		7.3333 (1.7E-1)
CBO5_in2010				91.058(1.3E-7)	29.283(4.7E-1)	
CBO5_out2010						33.6(2.2E-3)
CBO5_in2011						87(2.6E-3)
CBO5_out2011						
CCOCr	CCOCr_in2009	CCOCr_out2009	CCOCr_in2010	CCOCr-out2010	CCOCr_in2011	CCOCr_out2011
CCOCr_in2009		213.3(4.74E-8)	2.075(9.2E-1)		3.35(9.5E-1)	
CCOCr_out2009				50.78(2.69E-5)		8.35(5.7E-1)
CCOCr_in2010				262.0(5.72E-11)	29.283(4.7E-1)	
CCOCr_out2010						33.6(2.2E-3)
CCOCr_in2011						210.67(1.4E-3)
CCOCr_out2011						
MTS	MTS_in2009	MTS_out2009	MTS_in2010	MTS_out2010	MTS_in2011	MTS_out2011
MTS_in2009		93.758(2.4E-5)	6.842(5.3E-1)		8.2(7.7E-1)	
MTS_out2009				8.642(9.3E-2)		16.183(2.1E-2)
MTS_in2010				95.56(1.5E-8)	3.2(8.3E-1)	
MTS_out2010						14.6(3.1E-3)
MTS_in2011						84.283(2.7E-3)
MTS_out2011						

Values expressed as the mean of the difference, in parenthesis being given the probability to be different (from Student t)

The yearly output values of NH_4^+ do not significantly differ, implying either a poor efficiency of wastewater treatment process for ammonium, or an intense bacterial activity continuously producing new amounts of ammonium.

The comparison of the paired (by month) values for CBO5 shows that there is no significant difference between input values of the biochemical oxygen demand, and thus no major change in the microbiological content of the water should be suspected during 2009-2011 timeframe.

Chemical oxygen demand (CCOCr) represents the mass concentration of oxygen equivalent to the amount of potassium dichromate consumed by dissolved and suspended materials (MTS) when a water sample is treated with this oxidant under specified conditions. This direct proportionality also reflects in the results with no significant difference obtained by pairing values for both CCOCr and MTS for the inflow in WWTP.

CONCLUSIONS

All applied statistical analyses revealed a strong relationship between chemical and biochemical demand, especially for the outflow of WWTP, which explains why the efficiency of CBO5 and CCOCr reduction represents a criterion for assessing wastewater purification degree. These results correlated with the exceeding values reported during the monitored period (2009-2011) show the influence of ammonium on the efficiency of wastewater treatment process.

Season represents one of the main variation factor, especially concerning the change in suspended matter (MTS) content.

Projection of the variables on the factor plane is able to point out in detail around which observables gather the identified variation factors.

The obtained statistically significant model allows an analytical expression of intrinsic link between the physico-chemical parameters.

Comparing the values observed in different observation points (input and output) in different times (month, year) are able to highlight on the one hand the extent that the treatment process caused differences in values of environmental parameters controlled and monitored in WWTP and how much systematic changes appeared year to year to influence the water quality if untreated and treated, respectively.

REFERENCES

1. Binu, Kumari, S., A. Kavitha Kirubavathy, R. Thirumalnesan (2006). Suitability and water quality criteria of an open drainage municipal sewage water at Coimbatore, used for irrigation. J. Environ. Biol. 27(4):709-712.
2. Browne, M.W. (1968). A comparison of factor analytic techniques. Psychometrika, 33:267-334.
3. Chamberlain, D.E., J.D. Wilson, R.J. Fuller (1999). A comparison of bird populations on organic and conventional farm systems in southern Britain. Biological Conservation. 88(3):307-320.
4. Condit, R., S.P. Hubbell, R.B. Foster (1996). Assessing the response of plant functional types to climatic change in tropical forests. J. Veg. Sci. 7(3):405-416.
5. Egwari, L., O.O. Aboaba (2002). Environmental impact on the bacteriological quality of domestic water supplies in Lagos, Nigeria (in French). Revista de Saude Publica 36(4):513-520.
6. Giraitis, L., P. Kokoszka, R. Leipus, G. Teyssière (2003). Rescaled variance and related tests for long memory in volatility and levels. J. Econometrics 112(2):265-294.
7. Hassainia, F., D. Petit, J. Montplaisir (1994). Significance Probability Mapping: The final touch in t-statistic mapping. Brain Topography 7(1):3-8.
8. Hatton, A.D., S.W. Gibb (1999). A Technique for the Determination of Trimethylamine-N-oxide in Natural Waters and Biological Media, Analytical Chemistry, 71(21):4886-4891.
9. Hawkins, D.M. (2004). The Problem of Overfitting. Journal of Chemical Information and Computer Sciences 44(1):1-12.

10. Jäntschi, L. (2003). Physical Chemistry. Chemical and Instrumental Analysis (in Romanian), AcademicDirect, Internet, 64 p.

11. Jäntschi, L., S.D. Bolboacă (2007). The Jungle of Linear Regression Revisited, Leonardo Electronic Journal of Practices and Technologies, 6(10):169-187.

12. Johnson, S.C. (1967). Hierarchical clustering schemes. Psychometrika, 32:241-254.

13. Kvalseth, T.O. (1985). Cautionary Note about R2, The American Statistician, 39(4):279-285.

14. Milligan, G.W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika, 45:325-342.

15. Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine 5th Ser 50:157-175.

16. Schönemann, P.H., J.H. Steiger (1976). Regression component analysis. British Journal of Mathematical and Statistical Psychology, 29:175-189.

17. Student (1908). The probable error of a mean. Biometrika, 6:1-25.

18. Woollett, L.L., L.R. Hedrick (1970). Ecology of yeasts in polluted water. Antonie van Leeuwenhoek, 36(1):427-435.

19. Zhu, J.-N., X.-D. Wang, Z.-H. Zhai, W.-L. Ma, R.-Q. Li, X.-L. Wang, Y.-H. Li (2010). Changes of bacterial community structure on reusing domestic sewage of Daoxianghujing hotel to landscape water Huanjing Kexue, 31(5):1299-1306.

20. NTPA 001/2002: Regulation issued by Romanian Governmental Decision 188/2002. Standard for establishing the limits of pollutants charging of industrial and municipal wastewater at discharge into natural receivers. Monitorul Oficial I 187, March 20, 2002.

21. NTPA 002/2002: Regulation issued by Romanian Governmental Decision 188/2002. Standard for establishing conditions of wastewater discharge into the local sewerage networks and directly into treatment plants. Monitorul Oficial I 187, March 20, 2002.