Assessment of Water Quality State Dynamics Using Adaptive Filtering Methods and Neural Networks Approaching

Case study - Danube River in Galati area

GABRIEL MURARIU[1]*, CATALINA ITICESCU[1], ADRIAN MURARIU[2], BOGDAN ROSU[3], DAN MUNTEANU[3], DANIELA L. BURUIANA[4]*

[1]Dunarea de Jos University of Galati, Faculty of Sciences and Environment, Chemistry, Physics and Environment Department, 111 Domneasca Str., 800008, Galati, Romania

[2]Imperial College London, Physics Department, Prince Consort Rd, Kensington, London SW7 2BB, UK

[3]Dunarea de Jos University of Galati, Faculty of Automation, Computer Sciences, Electronics and Electrical Engineering, 47 Domneasca Str., 800008,Galati, Romania

[4]Dunarea de Jos University of Galati, Faculty of Engineering,47 Domneasca Str., 800008, Galati, Romania

The identification of a temporal evolution model for complex systems has, since ancient times, been a subject of great interest. Whether it is mechanical systems for which it was essential knowledge of the final state or electrical systems, the problem of identifying evolution over time has always been extremely interesting. In the case of a complex system such as a river, whose condition is described by a set of physico-chemical parameters, the time description of the evolution of the state becomes a rather difficult problem. In this paper, two ways of identifying and predicting the parameters describing the state of such a system are presented. A LRS type algorithm and a process of approximating evolution over time considering neural networks was used for comparison. Recorded series of pH and carbonic acid values were used as study parameters. The data used covers the period 1990-1998 and consists of measurements of the water samples taken from the Danube River in the area of Galati City. The main result was to obtain a rapid convergence for the adaptive filter used. For comparison, a number of 6 neural network models were built. Finally, findings and discussion of the results are presented.

Keywords: RLS algorithm, Water Quality Index, neural network, Pearson correlation

The descent of surface water quality is a subject of great concern nowadays. Over time, this process was mainly attributed to a series of factors: human activities and the use of adjacent land and natural processes, including climate change [1].

The category of human activities can include the influence of municipal wastewater discharged through non-conforming sewerage systems [2, 3]. In this category, both the untreated wastewater and the water discharges treated in the sewage treatment plants can be considered [1, 3].

From this point of view, there are two important factors that can be taken into account: climate change, on the one hand, and human influence, on the other [4]. Over time, these two major issues have caused significant changes in hydrological regimes and water quality.

In the area covered by the case study, in the last decades, with the rapid development of industries and agriculture, large quantities of pollutants have been produced and discharged into rivers and lakes [4-6]. From this point of view, it has caused a severe degradation of water quality. In addition, as a direct consequence, there is a strong influence and a major restriction on the sustainable development of local economies [1, 6, 7].

Water quality investigations and sources of pollution are essential for the implementation of sustainable management strategies [8-10]. Many investigation methods and procedures have been taken into account over time. For example, using Variance Analysis (ANOVA), many authors have indicated significant spatial variability for pH and other chemical parameters that cause significant changes in water quality [1, 10, 11]. Cyclical variance analysis was further developed. Many authors have indicated that due to the seasonality of river water flows, the assessment of temporal variations in water quality could be made [1, 12-14]. These seasonal variations become an important aspect of the physical and chemical characterization of aquatic environments, and each investigation should include this research [1, 15-17].

In this paper, two ways of identifying and predicting the parameters describing the dynamics of the status parameters of such a complex aquatic systems are presented. A LRS type algorithm and a process of approximating evolution over time considering neural networks was used for comparison. In view of the previously obtained results based on the application of some methods of statistical analysis of multivariate type [1, 18, 19], it was possible to group the set of 21 parameters measured in 4 principle factors [1]. To exemplify the methods of identification and analysis of the evolution over time, the series of recorded values of *p*H and of carbonic acid [1, 20, 21] were chosen. The data used covers the period 1990-1998 for the development of a coherent dynamic model [1].

Experimental part

The study area

The Danube is the second longest river in Europe (after the Volga) and is the only European river flowing from west to east (fig. 1). The Danube course crosses 10 countries (Germany, Austria, Slovakia, Hungary, Bulgaria, Moldova and Ukraine) and collects the waters of seven tributaries before reaching the Black Sea where it created the only delta in Europe - the Danube Delta. Interest in this area is significant because Delta entered the UNESCO World Heritage Reserve for Education, Science and Culture in

* email: gabriel.murariu@ugal.ro , dnegoita@ugal.ro

1991. Because it passes through four state capitals: Vienna, Bratislava, Budapest and Belgrade, the Danube is very important for river transport and for tourism.

Water samples were collected near the Galati City area, from the Danube River border, between January 1990 and December 1998.

Water analyses

All the 20 indicators for all water samples were investigated, each month during the monitoring period [1]. The used methods are described in the table 1 [1].

All water samples were analysed according to the Romanian standard procedures. Water Standards permissible values used in this study were taken from Romanian Legislation, Order 161/2006.

Statistical approach

In general, each statistical study must be based primarily on recorded data series, which represent the body of water explored to be able to discover the laws governing related phenomena. Previous statistical studies [1] show the existence of several groups of correlated factors. In this paper, methods of temporal evolution analysis were used using identification methods using RLS adaptive filters, the optimization criterion being considered the sum of squares of errors.

Dynamical approach

The recursive least-squares (RLS) procedure is one of the most well-known method used in adaptive filtering and system identification [17, 21- 23] The main advantage is principally due to its fast convergence quickness, which is considered to be ideal in practice. The RLS technique is typically used to minimize a weighted linear least squares function relating to the input parameters [21- 23].

In our case studies on the water quality parameters dynamics determination the time series of measurements could be used as input or output signals.

Identification of dynamic systems using RLS adaptive filters

For this paper, it could be included a briefly identification method presentation of adaptive filters usage by considering the system input *x* and the system output *y*.

Table 1

PARAMETERS AND METHODS OF ANALYSIS

Methods	*Parameter*	*Measurement unit*
Electrometrical	pH	upH
Volumetric	COD-Mn	mg $O_2 \cdot L^{-1}$
	Ca^{2+}	$mg \cdot L^{-1}$
	Mg^{2+}	$mg \cdot L^{-1}$
	Cl^-	$mg \cdot L^{-1}$
	CO_3^{2-}	$mg \cdot L^{-1}$
	HCO_3^-	$mg \cdot L^{-1}$
	Alkalinity	$mval \cdot L^{-1}$
	W. hardness	^0dH
Mathematic	Ca^{2+}/Mg^{2+}	-
Spectrophotometric	Fe_{total}	$mg \cdot L^{-1}$
	NO_2^-	$mg \cdot L^{-1}$
	NO_3^-	$mg \cdot L^{-1}$
	PO_4^{3-}	$mg \cdot L^{-1}$
	NH_3	$mg \cdot L^{-1}$
	NH_4^+	$mg \cdot L^{-1}$
Gravimetric	SO_4^{2-}	$mg \cdot L^{-1}$
	Rf_{105}	$mg \cdot L^{-1}$
	TSM	$mg \cdot L^{-1}$

If we denote by $h(t)$ the impulse response of the parameter that will be recognised, where t is the discrete time ($t = 0, 1, 2....$) and we denote by T_s the sampler period, the trials of the response $h(t)$ could be written as [23, 24]:

$$h(0.T_s) \equiv h_0;\ h(1.T_s) \equiv h_1;$$
$$h(2.T_s) \equiv h_2; \ldots ; h((n-1).T_s) \equiv h_{n-1} \qquad (1)$$

If the process is significant irregular, the output responses of the system at moment $t = kT_s$ will be denoted by $h_i[k]$, $i=\overline{1, n-1}$, and the values vectors of the output magnitudes could be written as [23, 24]:

$$\boldsymbol{h}[k] = \left[h_0[k] \quad h_1[k] \ldots h_{n-1}[k]\right]^T \qquad (2)$$

Using a convolution-type procedure of the process, the system output can be written as [23, 24]:

$$y[k] = \sum_{i=0}^{n-1} h_{i-1}[k].x[k-i] \qquad (3)$$

and thus the process model can be written as [23, 24]

$$y[k] = \boldsymbol{h}^T[k]x[k] \qquad (4)$$

A classical tuning of the state magnitude vector, h[k] could be achieved by the familiar algorithm RLS [22- 24]:

$$\boldsymbol{h}[k] = \boldsymbol{h}[k-1] + \boldsymbol{g}[k].\left(y[k] - \boldsymbol{h}^T[k-1]\boldsymbol{x}[k]\right) \qquad (5)$$

with the adapting improvement g[k] is computed as [22-24]:

$$g[k] = \frac{\boldsymbol{C}_{xx}^{-1}[k-1]\,\boldsymbol{x}[k]}{1 + \boldsymbol{x}^T[k]\boldsymbol{C}_{xx}^{-1}[k-1]\boldsymbol{x}[k]} \qquad (6)$$

where the autocorrelation matrix inverse, c_{xx}^{-1} is updated iteratively:

$$\boldsymbol{C}_{xx}^{-1}(k) = \boldsymbol{C}_{xx}^{-1}(k-1) - \boldsymbol{g}(k).\boldsymbol{x}^T(k)\boldsymbol{C}_{xx}^{-1}(k-1) \qquad (7)$$

For comparison, it was considered useful to build an approximation model using neural networks [3, 4, 11].

Results and discussions

In figures 2a and 2b the monthly evolutions of the two status parameters considered to be analysed - *p*H and HCO_3, respectively, are presented. The results of the ANOVA analyses are included to highlight the seasonal dependence of these parameters. The statistical analysis reveals the existence of an insignificant variation for these two parameters, which must be treated with caution, given the multitude of factors that influence the values and processes.

In order to obtain fittings with acceptable results, it was considered useful to construct neural network models of

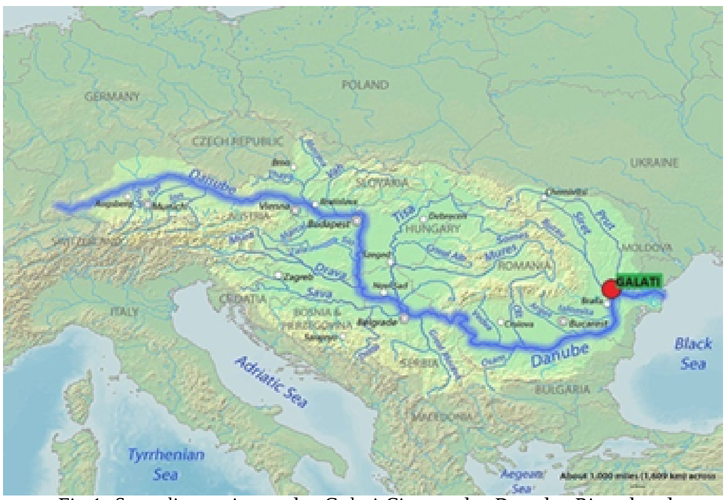

Fig.1. Sampling point at the Galati City on the Danube River border

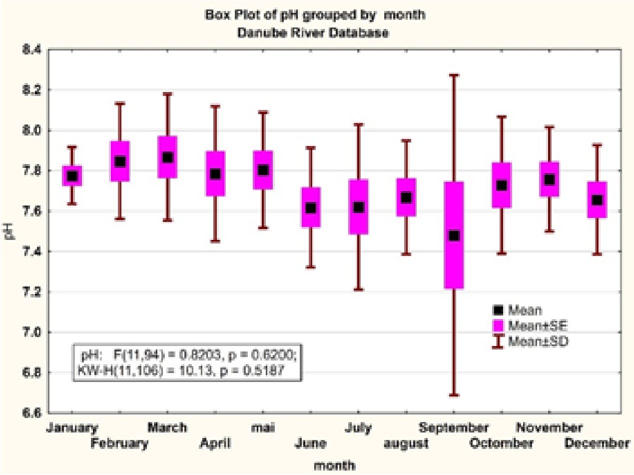

Fig. 2.a Monthly evolution for chemical parameters *p*H

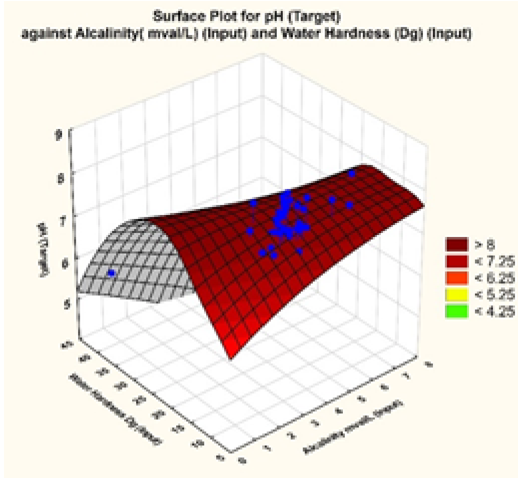

Fig.3.a The *p*H values as target values according to the values of alkalinity and water hardness

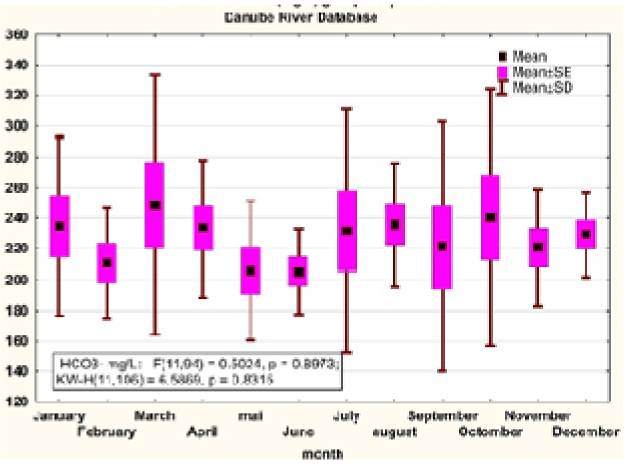

Fig. 2.b Monthly evolution for chemical parameters HCO_3

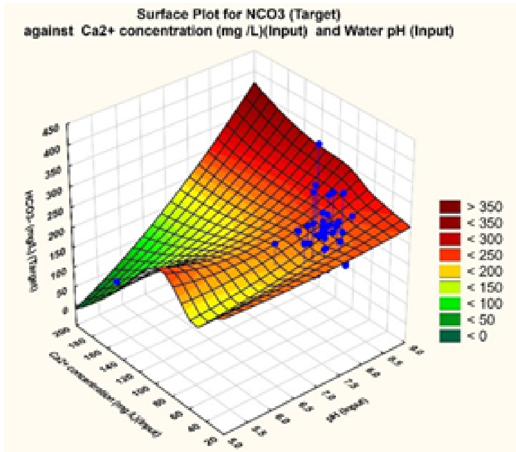

Fig.3.b The carbonic acid HCO_3 concentration according to *p*H values and calcium ion concentration

Table 2

SUMMARY OF ACTIVE NETWORKS (DANUBE RIVER DATABASE)

Index	Net. name	Training perf.	Test perf.	Validation perf.	Training error	Test error
1	MLP 2-3-1	0.662321	0.306912	0.646648	0.041084	0.054057
2	MLP 2-3-1	0.501109	0.290074	0.370834	0.054811	0.056442
3	MLP 2-7-1	0.705567	0.174060	0.316346	0.036698	0.058290
4	MLP 2-6-1	0.668858	0.128722	0.443897	0.040507	0.058151
5	MLP 2-4-1	0.722812	0.516059	0.378775	0.034897	0.045718
6	MLP 2-9-1	0.649012	0.122470	0.486531	0.043684	0.057849
Index	Net. name	Training perf.	Test perf.	Validation perf.	Training error	Test error

type ANN. A set of 20 different models has been built using Multi-Layer Perceptions (MLP) and Radial Basis Function, often used in literature to get the most accurate results. Figure 3a presents the result obtained using *p*H values as target values according to the values of alkalinity and water hardness [1, 3, 5]. Respectively, figure 3b presents the result obtained using the HCO3 carbonic acid values according to *p*H values and calcium ion concentration. The analysis of the group of these parameters was presented in previous papers [1, 3, 5]. Table 2 presents the characteristics and performance of these models and in table 3 the values of the correlation coefficients. The results of the network number 5 are observed.

Figure 5.a shows the result obtained using the RLS algorithm for identifying the temporal evolution of the *p*H

Table 3

CORRELATIONS - PREDICTIONS SPREADSHEET FOR *p*H (DANUBE RIVER DATABASE) THE NOTED VALUES ARE SIGNIFICANT

	pH target	pH - Output NR-1	pH - Output NR-2	pH - Output NR-3	pH - Output NR-4	pH - Output NR-5	pH - Output NR-6
pH target	1.0000	.6623	.5011	.7056	.6689	.7228	.6490
	p= ---	p=.000	p=.000	p=.000	p=.000	p=.000	p=.000
pH - Output NR-1	.6623	1.0000	.8926	.9361	.9911	.9122	.9824
	p=.000	p= ---	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00
pH - Output NR-2	.5011	.8926	1.0000	.7081	.8596	.6979	.8697
	p=.000	p=0.00	p= ---	p=.000	p=0.00	p=.000	p=0.00
pH - Output NR-3	.7056	.9361	.7081	1.0000	.9465	.9573	.9186
	p=.000	p=0.00	p=.000	p= ---	p=0.00	p=0.00	p=0.00
pH - Output NR-4	.6689	.9911	.8596	.9465	1.0000	.9177	.9951
	p=.000	p=0.00	p=0.00	p=0.00	p= ---	p=0.00	p=0.00
pH - Output NR-5	.7228	.9122	.6979	.9573	.9177	1.0000	.8972
	p=.000	p=0.00	p=.000	p=0.00	p=0.00	p= ---	p=0.00
pH - Output NR-6	.6490	.9824	.8697	.9186	.9951	.8972	1.0000
	p=.000	p=0.00	p=0.00	p=0.00	p=0.00	p=0.00	p= ---

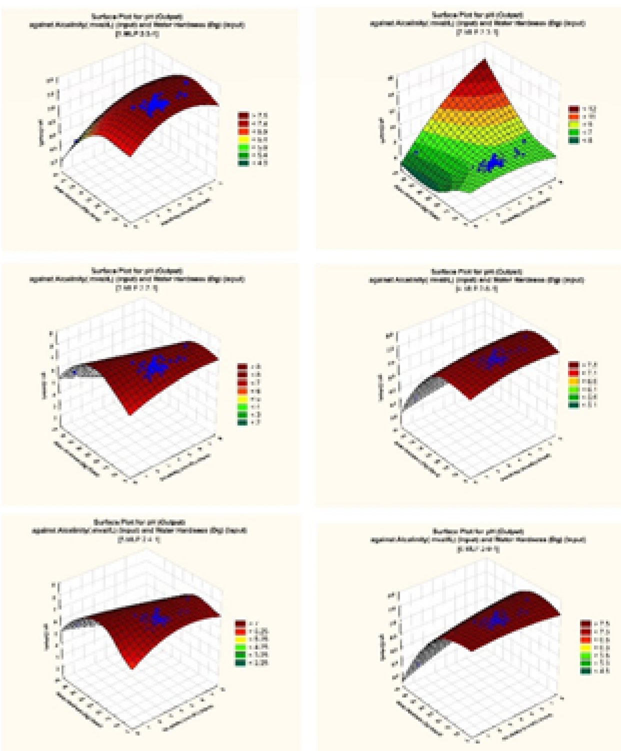

Fig.4.a The *p*H values as output values according to the values of alkalinity and water hardness- RNN results

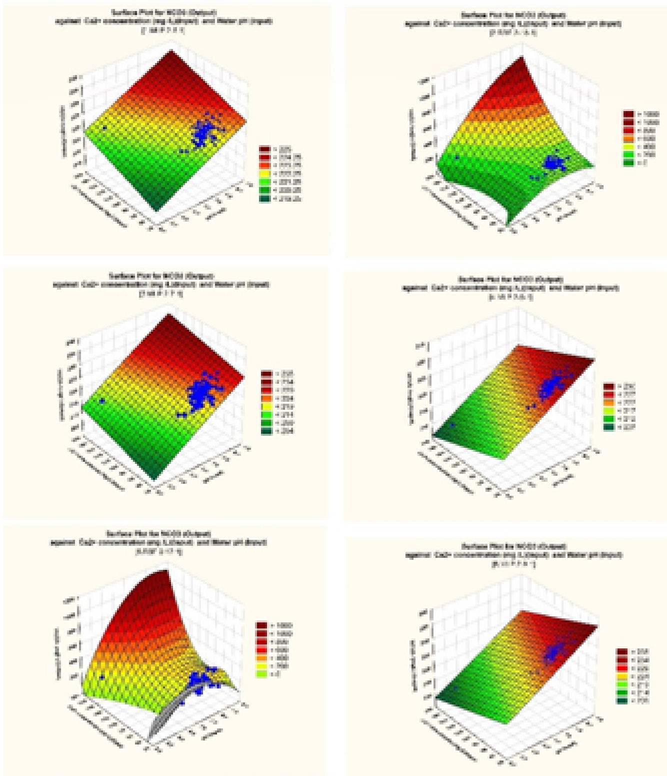

Fig.4.b The carbonic acid HCO_3 concentration as output according to *p*H values and calcium ion concentration - RNN results

according to the two input parameters. We see the rapid convergence of the algorithm used and the end result performance. Figure 5b shows the result obtained using the RLS algorithm to identify the temporal evolution of HCO_3 according to the two input parameters. In this case, the rapid convergence of the algorithm used and the performance of the result is also observed. In figure 5, a satisfactory correspondence between the identified models outputs and the emulated process can be seen.

Figure 6.a shows the variation of the Cxx convergence matrix for the values of the *p*H dynamics. Figure 6.b shows the variation of the Cxx convergence matrix for the identified values of carbonic acid dynamics.

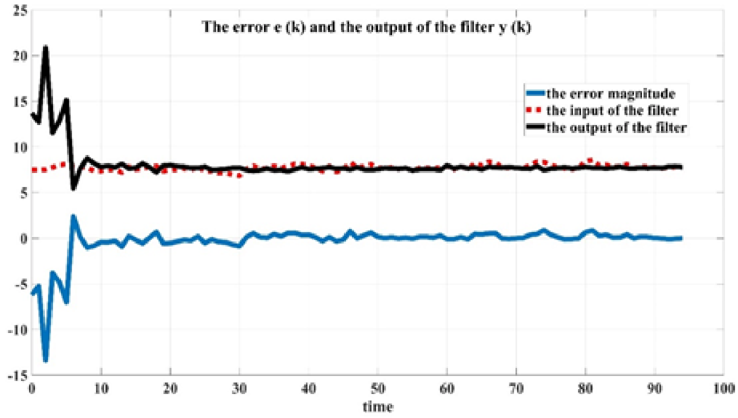

Fig.5.a The obtained result using the RLS algorithm for identifying the temporal evolution of the *p*H

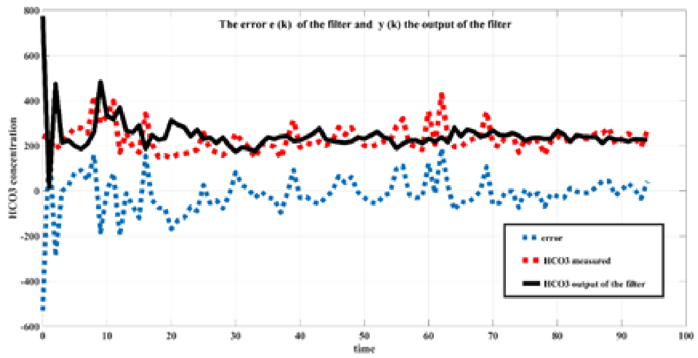

Fig.5.b The obtained result using the RLS algorithm for identifying the temporal evolution of the HCO_3

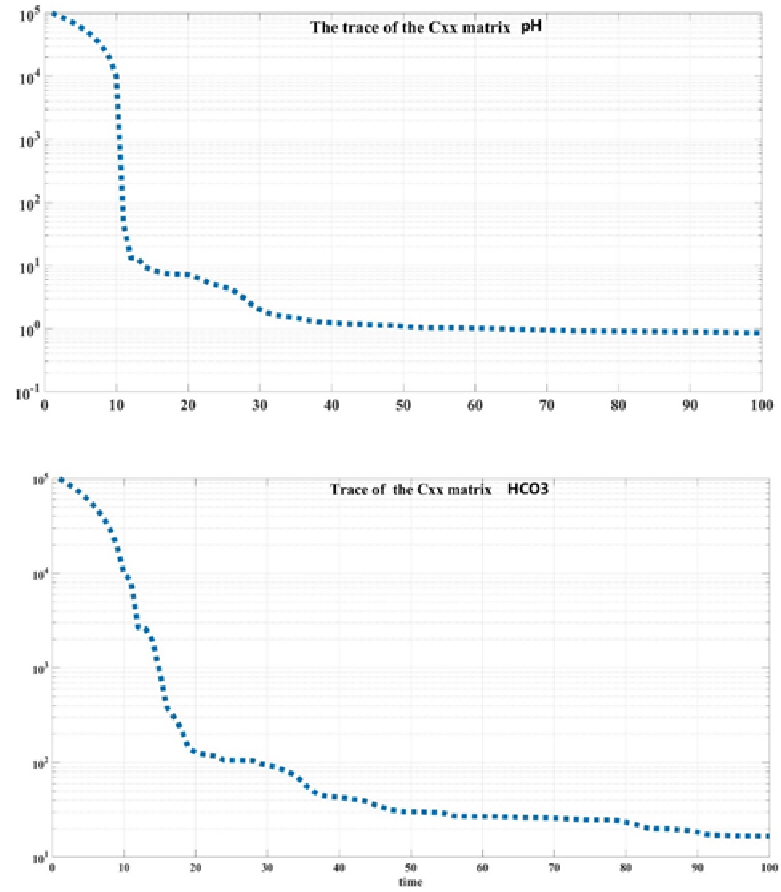

Fig.6.The variation of the Cxx convergence matrix for the values of the *p*H dynamics and respectively the HCO_3 dynamics

Conclusions

This paper's aim was to obtain a dynamic model using adaptive filters for two reasons: firstly, to highlight the correlation between pH alkalinity and water hardness - as already described in the literature on one hand and the correlation between HCO3, *p*H and concentration alkaline ions on the other hand. Secondly, the study was conducted to allow an analysis of the influence of the two inputs on the output channel. The promising results will allow for the development of the models considered to be included in the numerical approach of an extended set of status parameters of the studied aquatic system.

Acknowledgements. The authors are grateful for the financial support granted through the 4PS/2017 Research to support the development of the capacity to monitor, evaluate and capitalize on natural resources provided by wetlands of international importance in Romania and the Black Sea coastal zone. This work was supported by the *project Strategy and actions for preparing the national participation in the DANUBIUS-RI Project, acronym DANS, financed by the Romanian Ministry of Research and Innovation.*

References

1. POPA P., MURARIU G., TIMOFTI M., GEORGESCU L. P., Multivariate statistical analyses of water quality of Danube River at Galati, Romania, Environmental Engineering and Management Journal, 17, 5, 1249-1266, (2018).

2. ROMANESCU G., TIRNOVAN A., COJOC G. M., SANDU L. G., Temporal variability of minimum liquid discharge in Suha basin. Secure water resources and preservation possibilities International Journal of Conservation Science, 7, 4, 1135-1144, (2016).

3. DRASOVEAN R., MURARIU G., CONDURACHE-BOTA S., CONSTANTINESCU G., Studies on the water quality of the Siret river, near Galati city, 18th International Multidisciplinary Scientific Geoconference SGEM 2018, 18(3.1),687-695, (2018).

4. RADU, V. M., IVANOV, A. A., IONESCU, P., DEAK, G., DIACU, E., Overall assessment of water quality on Lower Danube River using multi-parametric Quality Index, Rev. Chim (Bucharest), **67**, no. 3, 2016, p. 391-395

5. ITICESCU C., GEORGESCU LP., TOPA C., MURARIU G., Monitoring the Danube water quality near the Galati City, Journal of Environmental Protection and Ecology, 15, 1, 30-38, (2014).

6. DRASOVEAN, R., MURARIU, G., CONSTANTINESCU, G. CIRCIUMARU, A., Assessment of surface water quality of Danube River in terms of usual parameters and correlation analyses, Rev. Chim. (Bucharest), **70**, no. 2, 2019, 392-406

7. BURADA A., TEODOROF L., DESPINA C., SECELEANU-ODOR D., TUDOR M., IBRAM O., NÃVODARU I., MURARIU G., ÞOPA C. M., TUDOR M., Trace elements in fish tissue with commercial value, of The Danube Delta Biosphere Reserve, Environmental Engineering and Management Journal, 16, 3, 731-738, (2017).

8. IONESCU, P., RADU, V.-M., DIACU, E., Sediments as indicators of heavy metal contamination in the Lower Danube River,Rev. Chim. (Bucharest),**66**, no. 11, 2015, p. 1725-1727

9. CROSA G., FROEBRICH J., NIKOLAYENKO V., STEFANI F., GALLID P., CALAMARI D., Spatial and seasonal variations in the water quality of the Amu Darya River (Central Asia), Water Research, 40, 2237-2245, (2006).

10. ZHOU F., GUO H., LIU Y., JIANG Y., Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong, Marine pollution bulletin, 54, 745-756, (2007),

10. IONESCU, P., RADU, V.-M, DEAK, G., DIACU, Rev. Chim. (Bucharest),**65**, no. 9, 2014, p. 1092.

11. LI S., GU S., LIUA W., HAN H., QUANFA Z, Water quality in relation to land use and land cover in the upper Han River Basin, China, CATENA, 75, 216-222, (2008).

12. AKTER T., TUZ JHOHURA F., AKTER F., ROY CHOWDHURY T., MISTRY S. K., DEY D., MILAN, BARUA K., ISLAM MD. A., RAHMAN M., Water quality index for measuring drinking water quality in rural Bangladesh: a cross-sectional study, J Health Popul Nutr., 35(4), (2016), doi: 10.1186/s41043-016-0041-5.

13. OUYANG Y., NKEDI-KIZZA P., WU Q.T., SHINDE D., HUANG C.H., Assessment of seasonal variations in surface water quality, Water Research, 40, 3800-3810, (2006).

14. SUNDARAY S.K., PANDA U.C., NAYAK B.B., BHATTA D., Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of the Mahanadi river–estuarine system (India) - a case study, Environmental Geochemistry Health, 28, 317-330, (2006).

15. WANG J., DA L., SONG K., LI B.L., Temporal variations of surface water quality in urban, suburban and rural areas during rapid urbanization in Shanghai, China, Environmental Pollution, 152, 387-393, (2008).

16. VASILINIUC I., PATRICHE C.V., PIRNAU R., ROSCA B., Statistical spatial models of soil parameters. An approach using different methods at different scales, Environmental Engineering and Management Journal, 12, 417-616, (2013).

17. MURARIU G., GEORGESCU L., TIMOFTI M., ITICESCU C., TOPA C., POPA P., MOCANU I.R., Comparison between pristine lagoon water systems in Romania and Italy. Case study - Leahova and Sinoe lagoons (Romania) and Cesine lagoon (Italy), Current Opinion in Biotechnology, 24, S66, (2013).

18. MOCANU, G., A Correlation Analysis on the coordination tests of Physical Education and Sports Faculty's students, from Galati, the Annals of Dunarea de Jos University of Galati 2013, Physical Education and Sport Management, 2013, no.2, fascicle XV, ISSN 1434-9832, www.ann.ugal.ro/efms.

19.*** http://www.statisticssolutions.com/correlation-pearson-kendall-spearman/

20. POPA, P., MOCANU, R., SARBU, C., Data multidimensional analysis regarding Danube water quality over the period 1990-1996, Rev. Chim. (Bucharest), **49**, 1998, p. 846-854

21. MURARIU G., GEORGESCU L., DOBRE M., ITICESCU C., TOPA C., DOBREA (POPESCU) A.A., DOBREA M., Using the physico-chemical and biological parameters in the state functions definition. Case study: River Candelaro (Italy), Current Opinion in Biotechnology, 24, S79-S80, (2013),

22. XU X., HE H.-G., HU D., Journal of Artificial Intelligence Research 16 259-292, (2002)

23.KOH C., SEE L., Identification and Uncertainty Estimation of Structural Parameters., J. Eng. Mech., 120(6), 1219–1236, (1994).

24.TIMOFTI, M., POPA P., MURARIU, G.; ET AL, Complementary approach for numerical modelling of physicochemical parameters of the Prut River aquatic system, Journal of Environmental Protection and Ecology, 17,1, 53-63, (2016)

25. ELENA L. LISA, GABRIEL B. CARP, DANIELA L. BURUIANA, MARIUS BODOR, CLAUDIA SIMONA STEFAN, The Necessity of Improving the Quality of Water for Usage and Drainage in Environment, In The Metropolitan Area Galati-Braila, SGEM 2017, Bulgaria, Section Hydrology and water resources, volume 17, pp 717, ISBN 978-619-7408-04-1.

26. ITICESCU, C., MURARIU, G., GEORGESCU, L., BURADA, A, TOPA, M. C., Seasonal variation of the physico-chemical parameters and Water Quality Index (WQI) of Danube water in the transborder Lower Danube area. Rev. Chim. (Bucharest), **67**, no. 9, 2016, p. 1843 -1849.

Manuscript received: 16.01.2019