Prediction of wastewater organic matter with time series

Facundo Cortés-Martínez, T. Espinoza-Fraire, Agustín Sáenz López, Julio Roberto Betancourt Chávez

***Abstract*— The implementation of methods for waste water quality prediction, in the treatment systems, is a great tool. These provide valuable information for decision-making. In the present research, we propose the application of time series with the criteria of moving average with different conditions, as well as the exponential smoothing. The above with the purpose of predicting the Biochemical Oxygen Demand in the effluent of a treatment system. The indicated criteria were applied and the measurements were carried out in 17 months in the system's effluent. The obtained results show that the smoothing with moving average of order 2 resulted in the best prognosis: 79 mg/L., and the minimum error were obtained.**

***Index Terms*— moving median, exponential smoothing, time series, operating organisms, biochemical oxygen demand.**

I. Introduction

A time series is a sequence of observations of a variable taken at several instants of time. The literature indicates that the time series includes four theoretical components: trend, seasonal variations, cyclical variations and residual variations. The trend lines are used to observe the meaning of the data analyzed, in addition to carrying out the prediction. The seasonal variations refer to oscillations that are generated in the series with a time equal to or less than 12 months; while the cyclical variations are variations that are generated in a period greater than one year: repetition of the behavior of the series, ie cycles. Residual variations relate to fluctuations showing a series, but without a definite trend (Pérez, 2002).

Scientific papers have been published for the prediction of costs and quality of treated wastewater, for example, Zhu et al (1998) applied simulation model to predict expenditure in the effluent treatment system with time series and neural networks; Kusiak et al. (2012) published the results of their research: prediction of treated wastewater quality, for which they were supported in time series and data mining. Then Verma et al. (2013) conducted a study in order to predict the concentration of suspended solids, used time series and data mining. Other many published studies for predicting the quality of wastewater using neural networks, some are: Mjalli et al. (2007); Atasoy et al. (2013); Pai et al. (2011). The studies indicated above have contributed to different criteria in the prediction of pollutants in wastewater treatment systems.

Facundo Cortés-Martínez. Faculty of Engineering, Science and Architecture Juarez University of Durango State, Mexico. (FICA-UJED).

T. Espinoza-Fraire[1]Faculty of Engineering, Science and Architecture Juarez University of Durango State, Mexico. (FICA-UJED).

Agustín Sáenz López. Faculty of Engineering, Science and Architecture Juárez University of Durango State, Mexico. (FICA-UJED).

Julio Roberto Betancourt Chávez. Faculty of Engineering, Science and Architecture Juárez University of Durango State, Mexico. (FICA-UJED).

The objective of the present study was to calculate the forecasting of the concentration of biochemical oxygen demand in mg/L at the effluent of a wastewater treatment plant from stabilization ponds.

The contribution of the study is the prediction of organic matter from wastewater with time series by applying the criteria of moving average analysis. In the literature review were not found research jobs with the previously mentioned application.

According to the Comisión Nacional del Agua y el Instituto Mexicano de Tecnología del Agua (CNA and IMTA) the Biochemical Oxygen Demand (BOD) or organic matter, refers to the dissolved oxygen in the wastewater required by living organisms, in order to carry out the decomposition of organic matter. This parameter is used for both the design of treatment systems, to mediate the quality of treated wastewater (CNA and IMTA, 2000).

II. Methodology

According to Perez (2012) smoothing by the moving average is used to reduce fluctuations in a time series, the indicated in order to obtain an adequate line, free of seasonal variations and thus determine a correct prediction. If you have a time series where t = 1, 2, ...T. For example for a moving average of order 5 we have:

$$T_s = (X_{s-2} + X_{s-1} + X_s + X_{s+1} + X_{s+2})/5 \tag{1}$$

$$s = 3,4, \dots T-2$$

One of the disadvantages of this method is the loss of initial and final data of the series, so that as the order number grows the amount of data is reduced: if there are no previous reference values.

Exponential smoothing: the method consists of a self-correction where the forecasts are adjusted in the opposite direction to the previous errors. This method is used both to smooth time series to carry out forecasts (Suárez, 2011).

$$Y_{t+1} = \propto . X_t + (1-\propto). Y_t \tag{2}$$

Where:

Y_{t+1} = Forecast for any future period

$\propto$ = Smoothing constant, which is assigned a value between 0 and 1

X_t = Actual value for the period

Y_t = Forecast made previously for the time period

The measurements belong to a stabilization lagoon treatment system located in northern Mexico. Then the concentration data of BOD_5 mg/L., per month at the outlet of the processing system: 96.8, 86, 57.9, 88.6, 86.4, 77, 53, 82, 71, 61, 97, 71, 80; 95, 111, 76 and 81. It is important to indicate that all these values are below the maximum allowed limit that indicates the regulation. According to Mexican Official Standard NOM-001-SEMARNAT-1996 (DOF 1997). The monthly average concentration of BOD_5 for agricultural irrigation is 150 mg/L.

III. RESULTS AND DISCUSSION

Table 1 shows the results of the three analyzes

Table 1. Results of the three time series analyzes with Excel

Moving average of orden 2		Moving average of orden 3		Exponential Smoothing	
Forecast	Typical error	Forecast	Typical error	Forecast	Typical error
#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
91	#N/A	#N/A	#N/A	97	#N/A
72	11	80	#N/A	91	#N/A
73	15	78	#N/A	75	#N/A
88	11	78	15	82	22
82	3	84	9	84	21
65	9	72	13	81	9
68	13	71	13	67	17
77	11	69	13	74	19
66	5	71	9	73	18
79	13	76	13	67	11
84	16	76	14	82	19
76	10	83	12	76	20
88	6	82	8	78	19
103	8	95	12	87	12
94	14	94	16	99	17
79	13	89	15	87	22
	157		162		225

The numbers appearing in green in Table 1 are the forecasts for the 18th month; while the yellow ones refer to the sum of the errors. According to table 1 the analysis with the moving average of order 2 was the best prognostics: 79 mg/L BOD_5 followed by analysis of exponential smoothing with a difference of 8 mg/L., organic matter. The 10.12 %. Finally the analysis of moving average of order 3. The difference was of 10 mg/L representing 12.65% approximately. Nonetheless, the three analyzes perfectly comply with the permitted concentration limits for the discharge of wastewater to the receiving bodies: 79, 89 and 87 mg/L. Moving average of order 2 and 3 as well as exponential smoothing respectively.

Similarly the sum of the lowest error was studying the moving average of order 2. According to Suárez (2011) the error is an indicator of forecast accuracy. Figures 1, 2 and 3 show the time series with the criteria of the moving average of order 2 and 3 as well as the exponential smoothing.

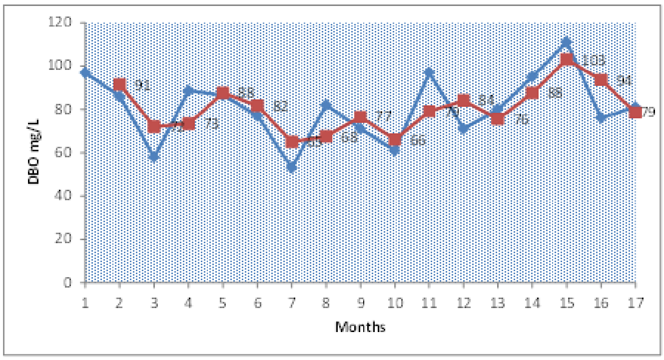

Figure 1. Forecast of wastewater quality with moving average of order 2.

In the analysis of the series with order 3 that is indicated in figure 2, it is observed with smaller fluctuations than that of order 2. Nevertheless, the forecast was lower in the series with order 2 as well as the sum of the errors.

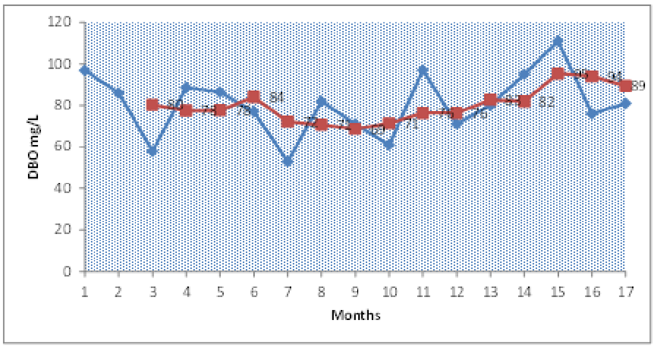

Figure 2. Residual water quality forecast with moving average of order 3.

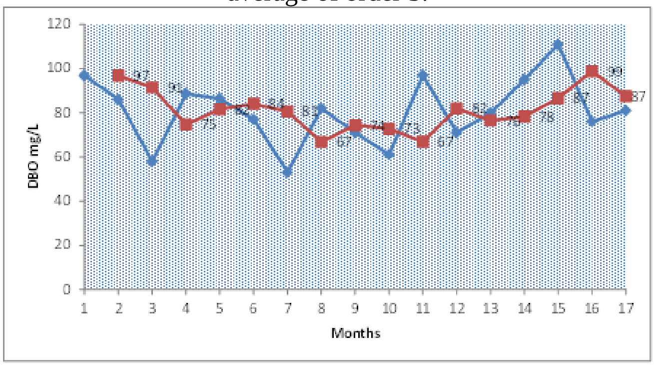

Figure 3. Residual water quality forecast with exponential smoothing

Although in the present study only one month of forecast in the concentration of the organic matter was considered, it is prudent to take into account this result, since it will serve as the basis to make decisions in both the influent and the effluent of the treatment system. For example control the level of concentration of BOD_5 in the influent: This is possible through a program of verification of industrial processes that generate organic matter (EPA, 1987; CNA; and IMTA, 2000, UNAM, 2010; Cortes, et., 2009; Martínez et al., 2010). About the effluent carefully review the operation of the treatment system. In the present case, the forecast, as already indicated, was below that mentioned by the regulation. For this reason it will not be necessary to carry out any emergent program, reason why the costs of operation will not be increased. Nevertheless, it is advisable to monitor the organic matter constantly.

IV. CONCLUSION

The prognosis of the organic matter was carried out in the effluent of the stabilization lagoon system.

The analysis criteria and the statistical tool used are simple and easy to execute, so that they can be performed by technicians of the operating agencies of drinking water in the operation of wastewater treatment systems.

ACKNOWLEDGE

The author would like to thank the strengthening program for quality in educational institutions (PROFOCIE) 2015 for its support in the realization of this research.

REFERENCES

[1] Atasoy, A.D., Babar, B., Sahinkaya, E., (2013). Artificial neural network prediction of the performance of upflow and downflow fluidized bed reactors treating acidic mine drainage water. *Mine Water Environ*. 32 (3), 222–228

[2] Comisión Nacional del Agua e Instituto Mexicano de Tecnología del Agua (2000). "Guía para el control de descargas a los sistemas de alcantarillado urbano o municipal". México.

[3] Cortes Martínez, F., Martínez García, I., Betancourt Hernández, J. y Duke Herrera, J. (2009). Cálculo del índice de incumplimiento y costo por pago de derechos en la descarga de agua residual de procesos industriales. (Caso Comarca Lagunera). Ingeniería, Revista Académica de la FI-UADY, 13-2, pp. 33-39, ISSN: 1665-529X.

[4] Diario Oficial de la Federación. (1997). *Norma Oficial Mexicana NOM-001-SEMARNAT-1996:* que establece los límites máximos permisibles de contaminantes en las descargas de aguas residuales a los sistemas en aguas y bienes nacionales. México.

[5] EPA, (1987), Environmental Protection Agency "Manual guía para prevenir la interferencia en las POTWs". USA.

[6] Kusiak, A., Verma, A., & Wei, X. (2012). A data-mining approach to predict influent quality. *Environmental monitoring and assessment*, *185*(3), 2197-2210.

[7] Martínez, F. C., Silva, F. B., & Chávez, J. F. M. (2010). Control Inicial en la Descarga de Aguas Residuales Industriales y Comerciales. *Conciencia Tecnológica*, (39), 43-49.

[8] Mjalli, F. S., Al-Asheh, S., & Alfadala, H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, *83*(3), 329-338.

[9] Pai, T. Y., Yang, P. Y., Wang, S. C., Lo, M. H., Chiang, C. F., Kuo, J. L., & Chang, Y. H. (2011). Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Applied Mathematical Modelling*, *35*(8), 3674-3684

[10] Pérez L.C. (2002). Estadística aplicada a través de Excel. Universidad Complutense de Madrid. España.

[11] SUÁREZ Mario, (2011), Métodos de suavizamiento y pronóstico para series de tiempo. Consultado el 29 enero 2017. Disponible en: http://www.monografias.com/trabajos87/metodos-suavizamiento-y-pronostico-series-tiempo/metodos-suavizamiento-y-pronostico-series-tiempo.shtml

[12] Universidad Nacional Autónoma de México. (2000), "Curso uso eficiente del agua y control de calidad de las descargas de aguas residuales en la industria". México.

[13] Verma, A., Wei, X., & Kusiak, A. (2013). Predicting the total suspended solids in wastewater: a data-mining approach. *Engineering Applications of Artificial Intelligence*, *26*(4), 1366-1372.

[14] Zhu, J., Zurcher, J., Rao, M., Meng, M.Q.H., (1998). An on-line wastewater quality prediction system based on a time delay neural network. *Eng. Appl. Artif. Intell*. 11 (2), 747–758.