ISSN 1749-3889 (print), 1749-3897 (online)
International Journal of Nonlinear Science
Vol.17(2014) No.3,pp.234-240

Predicting Dissolved Oxygen Fluctuations in Golden Horn by Fuzzy Time Series

Xihao Sun [1], Yimin Li [2] *,
[1] Wuxi Lake Tai College,Wuxi, Jiangsu,China
[2] Faculty of Science, Jiangsu University,Jiangsu, China
(Received 7 June 2012 , accepted 10 December 2013)

Abstract: In recently years, fuzzy time series model has been applied to various applications. Fuzzy time series can solve forecasting problems in which the historical data are linguistic values. In this paper, we aim to the application of forecasting monthly dissolved oxygen (DO) amount variations by fuzzy time series. DO concentration changes in the two station located in Golden Horn at the 0.5m (upper layer) depth were modeled, and then we compare the forecasting results with TS fuzzy model and ARMA approach. Empirical analyses show that fuzzy time series model can greatly improve forecasting results; in particular, it is superior to the other two models used in previous study.

Keywords: Fuzzy time series; forecast; dissolved oxygen; Golden horn; model

1 Introduction

As we know, dissolved oxygen (DO) is a vital element for the survival of living beings in the estuatine ecosystems [1]. Thus, a sufficient supply of dissolved oxygen is vital for all higher aquatic life. DO concentration is one of the simple but practical indexes to evaluate and predict the water quality. The problems associated with low concentrations of DO in rivers have been recognized for over a century and the impacts of low DO concentrations or, at the extreme, anaerobic conditions in a normally well-oxygenated river system, are an unbalanced ecosystem with fish mortality, odours and other aesthetic nuisances [2]. When DO concentrations are reduced, aquatic animals are forced to alter their breathing patterns or lower their level of activity. Both of these actions will retard their development, and can cause reproductive problems (such as increased egg mortality and defects) and/or deformities.

Many studies have been carried out to compute DO concentration in rivers, lakes, and estuaries. [3] used a two-dimensional mathematical model for predicting the DO concentration changes in shallow estuaries. A modified approach to DO rate parameter estimation that takes temperature variation into account was studied by Butcher and Covington [4]. In epilimnion lakes, Fang and Stefan [5] analytically and numerically examined the interaction between surface gas transfer, epilimnetic diffusion and internal oxygen production on DO concentrations. Miranda et al. [6] linked a bathymetric model that estimated lake morphometry at various stage elevations, with a probabilistic risk assessment that used a simple respiration model to predict the likelihood of the lake developing unsuitable DO conditions. Cox [2] described mass-balance model, mathematical models and mechanistic models and required data for using these models to predict DO concentration in lowland river systems. Quinn [7] studied the control and management of episodes of low dissolved oxygen in DWSC. Altunkaynak et al. [1] used fuzzy logic modelling method to forecast the DO fluctuations in Golden Horn. Carlos et al. [8] describes the development of an adaptive fuzzy control strategy for tracking the DO reference trajectory applied to the Benchmark Simulation.

Fuzzy set theory was originally developed to handle problems involving human linguistic terms [9]. Because the existing statistical time series methods could not effectively analyze time series with small amounts of data and historical data, fuzzy time series methods were developed. Song and Chissom [10] proposed a first-order time-invariant model and a time-variant model of fuzzy time series. They fuzzified the enrollment at the University of Alabama in 1993 in the first application of fuzzy time series to forecasting (Song and Chissom [11]). Then, Song and Chissom [12] proposed a new fuzzy time series and compared three different defuzzification models. Chen [13] considered that the Song and

*Corresponding author: *E-mail address*: llym@ujs.edu.cn

IJNS.2014.06.15/807

Chissoms method is too complicated to apply; he therefore presented arithmetic operations instead of the logic max-min composition. The arithmetic operations have a robust specification and are superior to those applied in Song and Chissom model. Following these definitions, fuzzy time series models have been proposed for various applications, such as enrollment [10-13]stock indices [14-16] reactors [17], temperature forecasting [18], pollution concentrations [19], Taiwan export [20] and Taiwans tourism demand [21].

The goal of this study is to propose Chens fuzzy time series models to improve the forecasting of DO fluctuations in Golden Horn. The rest of this paper is organized as follows. In Section 2, the concepts of fuzzy time series are reviewed. Section 3 elaborates on the use of model in forecasting DO fluctuations in Golden Horn. Section 4 evaluates the models performance and Section 5 concludes the paper.

2 Fuzzy time series revisited

In this section, we briefly review some basic concepts of fuzzy time series. [10] first proposed the definitions of fuzzy time series. The concepts of fuzzy time series are described as follows.

Let U be the universe of discourse, where $U = \{u_1, u_2, \cdots, u_n\}$. A fuzzy set A_i of U is defined by

$$A_i = \frac{f_{A_i}(u_1)}{u_1} + \frac{f_{A_i}(u_2)}{u_2} + \cdots + \frac{f_{A_i}(u_n)}{u_n}, \tag{1}$$

where f_{A_i} is the membership function of the fuzzy set A_i, $f_{A_i} : U \to [0, 1]$. u_k is an element of fuzzy set A_i, and $f_{A_i}(u_k)$ is the degree of belongingness of u_k to A_i. $f_{A_i}(u_k) \in [0, 1]$ and $1 \le k \le n$.

Definition 1 $Y(t)(t = \cdots, 0, 1, 2, \cdots)$, *is a subset of R. Let* $Y(t)$ *be the universe of discourse defined by fuzzy set* $f_i(t)$*. If* $F(t)$ *consisted of* $f_1(t), f_2(t), \cdots$, $F(t)$ *is defined as a fuzzy time series on* $Y(t)(t = \cdots, 0, 1, 2, \cdots)$.

From definition 1, we can see that:

1. $F(t)$ is a function of time.
2. $F(t)$ can be regarded as a linguistic variable, which is a variable whose values are linguistic values represented by fuzzy sets.
3. $f_i(t)(i = 1, 2, \cdots)$ are possible linguistic values of $F(t)$, where $f_i(t)(i = 1, 2, \cdots)$ are represented by fuzzy sets.

Definition 2 *If there exists a fuzzy relationship* $R(t-1, t)$*, such that* $F(t) = F(t-1) \times R(t-1, t)$ *where* $\times$ *represents an operator, then* $F(t)$ *is said to be caused by* $F(t-1)$*. Let* $F(t-1) = A_i$ *and* $F(t) = A_j$*. The relationship between* $F(t)$ *and* $F(t-1)$*(referred to as a fuzzy logical relationship, FLR) can be denoted by* $A_i \to A_j$*; where* A_i *is called the left-hand side (LHS) and* A_j *the right-hand side (RHS) of the FLR.*

Definition 3 *Given two FLRs with the same fuzzy sets on the LHSs* $A_i \to A_{j1}$*,* $A_i \to A_{j2}$*. Both FLRs can be grouped together into fuzzy logical relationship groups (FLRG)* $A_i \to A_{j1}, A_{j2}$*.*

Different fuzzy time-series models have been proposed by following Song and Chissoms definition of fuzzy time-series. For example, there have been Song and Chissoms time-invariant model [11], Song and Chissoms time-variant model, the Markov model [22], Chens model [13], weighted model [16], etc.

It is common for these models to include the following steps: (1) define the universe of discourse and the intervals for the observations; (2) partition the universe based on the intervals; (3) define the fuzzy sets for the observations; (4) fuzzify the observations; (5) establish the fuzzy logic relationship and fuzzy logic relationship group; (6) perform the forecast; and (7) defuzzify the forecasting results.

3 DO fluctuation Forecasting

In the following, the aforementioned methodology is applied to the data for Golden Horn. Data of the two stations are used in this study. Next, we follow the forecast procedure described in section 2 to forecast the DO fluctuation. For simple, we take data for Galata Bridge station for illustration.

Step1: Defining the universe of discourse and intervals.

The universe of discourse for observations, U, is defined as $[D_{min} - D_1, D_{max} + D_2]$, where D_{min} and D_{max} is the minimum and maximum of known historical data, D_1, D_2 are two proper positive numbers. According to the data, we can see that $D_{min} = 6.04$ and $D_{max} = 10.04$. Thus, the universe of discourse is defined as $U = [6, 10.2]$.

Then U can be partitioned into equal-length intervals $u_1, \cdots, u_{14}$, and the midpoints of these intervals are $m_1, \cdots, m_{14}$, respectively, where $u_1 = [6, 6.3], \cdots, u_{14} = [9.9, 10.2]$.

Step2: Defining fuzzy sets for observations.

Each linguistic observation, A_i , can be defined by the interval $u_1, \cdots, u_{14}$, as follows:

$$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + \cdots + 0/u_{14};$$
$$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + \cdots + 0/u_{14};$$
$$A_3 = 0 + 0.5/u_2 + 1 + 0.5/u_4 + \cdots + 0/u_{14};$$
$$\cdots$$
$$A_{13} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + \cdots + 0.5/u_{12} + 1/u_{13} + 0.5/u_{14};$$
$$A_{14} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + \cdots + 0/u_{12} + 0.5/u_{13} + 1/u_{14}.$$

Step3: Fuzzifying observations.

Each data can be fuzzified into a fuzzy set. Each data and their corresponding fuzzy set A_i are listed in Table 1.

Table 1: Fuzzified DO fluctuation

Month	DO fluctuation	Fuzzy Sets
January	6.91	A_4
February	8.31	A_8
March	9.53	A_{12}
April	7.77	A_6
May	10.04	A_{14}
June	7.78	A_6
July	7.31	A_5
August	6.04	A_1
September	7.41	A_5
October	7.76	A_6
November	9.26	A_{11}
December	8.94	A_{10}

Step4: Establishing FLRs and FLRGs.

Based on the fuzzy set in step 3, the FLRs are established, as in Table 2. The FLRs can be rearranged to FLRGs, as in Table 3.

Table 2: Fuzzy Logic relationships

$A_4 \to A_8, A_8 \to A_{12}, A_{12} \to A_6, A_6 \to A_{14}, A_{14} \to A_6, A_6 \to A_5, A_5 \to A_1$
$A_1 \to A_5, A_5 \to A_6, A_6 \to A_{11}, A_{11} \to A_{10}$

Step5: Forecasting.

Forecasting is conducted by the following rules:

Rule 1: If the current fuzzy set is A_i, and the fuzzy logical relationship group of A_i is empty, i.e.,$A_i \to$, the forecast is m_i, the midpoint of u_i.

$$Forecasting = m_i$$

Rule 2: If the current fuzzy set is A_i, and the fuzzy logical relationship group of A_i is one-to-one, i.e., $A_i \to A_j$, the forecast is m_j, the midpoint of u_j.

$$Forecasting = m_j$$

Table 3: Fuzzy Logic relationships

$A_4 \to A_8$
$A_8 \to A_{12}$
$A_{12} \to A_6$
$A_6 \to A_{14}, A_5, A_{11}$
$A_{14} \to A_6$
$A_5 \to A_6, A_1$
$A_1 \to A_5$
$A_{11} \to A_{10}$

Rule 3: If the current fuzzy set is A_i, and the fuzzy logical relationship group of A_i is one-to-many, i.e., $A_i \to A_{j1}, A_{j2}, \cdots, A_{jn}$, the forecast is equal to the average of $m_{j1}, m_{j2}, \cdots, m_{jn}$, the midpoints of $u_{j1}, u_{j2}, \cdots, u_{jn}$, respectively.

$$Forecasting = \frac{\sum_{i=1}^{n} m_{ji}}{n}$$

Defuzzifying the forecast In the following, we only illustrate the process of some months. The similar procedure can be applied to other months.

[March] Because the fuzzified data of February shown in Table 1 is A_8, and from Table 3, we can see that the corresponding FLRGs to A_8 is as follows:

$$A_8 \to A_{12}$$

Following the forecasting rules 2, the forecasted data of March is equal to $m_{12} = 9.45$.

[July] Because the fuzzified data of June shown in Table 1 is A_6, and from Table 3, we can see that the corresponding FLRGs to A_6 is as follows:

$$A_5 \to A_{14}, A_5, A_{11}$$

Following the forecasting rules 3, the forecasted data of July is equal to

$$\frac{1}{3}(m_{14} + m_{11} + m_5) = \frac{1}{3}(10.5 + 7.35 + 9.15) = 8.85$$

[October] Because the fuzzified data of September shown in Table 1 is A_5, and from Table 3, we can see that the corresponding FLRGs to A_5 is as follows:

$$A_5 \to A_1, A_6$$

Following the forecasting rules 3, the forecasted data of October is equal to

$$\frac{1}{2}(m_1 + m_6) = \frac{1}{2}(6.15 + 7.65) = 6.9$$

4 Empirical Analyses

Following section 3, the out-of-sample defuzzified forecasts for Chens model are listed in Table 4 and 5 as an illustration.

Table 4: Comparison of fuzzy time series (FTS) model and TS fuzzy,ARMA model results for Galata Bridge station

Year	Month	Observation	Prediction			Relative error (%)		
			FTS	TS fussy	ARMA	FTS	TS fuzzy	ARMA
2001	January	6.91	7.05	7.50	6.91	2.27	7.91	0.00
	February	8.31	8.25	7.47	7.75	0.59	10.13	6.70
	March	9.53	9.45	8.47	7.25	0.82	11.07	23.90
	April	7.77	7.65	8.03	6.54	1.31	3.21	15.83
	May	10.04	8.85	7.25	7.84	13.75	27.77	21.86
	June	7.78	7.65	8.75	7.93	1.22	11.07	1.86
	July	7.31	8.85	6.73	6.54	20.13	8.00	10.50
	August	6.04	6.90	8.08	5.96	12.93	25.26	1.35
	September	7.41	7.35	6.84	5.99	0.87	7.76	19.22
	October	7.76	6.90	7.53	6.25	12.46	2.98	19.46
	November	9.26	8.85	8.70	6.97	4.74	6.04	24.71
	December	8.94	8.85	8.63	7.74	23.4	3.50	13.44
	Average	8.09	8.05	7.83	6.97	5.99	10.39	13.23

Table 5: Comparison of fuzzy time series (FTS) model and TS fuzzy, ARMA model results for Kasimpasa station

Year	Month	Observation	Prediction			Relative error (%)		
			FTS	TS fussy	ARMA	FTS	TS fuzzy	ARMA
2001	January	5.99	6.150	5.56	6.62	2.67	11.52	9.52
	February	10.31	10.15	6.60	6.86	1.55	31.33	33.42
	March	9.41	9.650	8.65	9.37	2.55	11.28	0.45
	April	9.05	9.150	9.45	6.79	1.10	2.290	25.01
	May	8.80	8.650	9.63	8.59	1.70	7.530	2.35
	June	12.77	10.65	9.80	8.72	16.6	24.40	31.70
	July	7.63	7.650	13.94	8.77	0.26	28.95	13.03
	August	6.43	6.650	6.73	8.04	3.42	11.71	20.07
	September	6.85	6.900	7.06	8.04	0.73	0.410	14.79
	October	7.01	6.900	7.52	7.46	1.57	7.160	6.02
	November	8.44	8.650	7.73	8.39	2.49	8.160	0.59
	December	8.63	10.65	9.41	8.96	23.4	12.84	3.66
	Average	8.44	8.470	8.51	8.05	4.84	13.13	13.38

As can be seen in Table 4 and 5, fuzzy time series model predict the observed values better than the other two models. Moreover, the superiority of fuzzy time series model also reflected in the mean of DO amounts prediction for 2001.

The observed value is 8.09; while the FTS model forecasted it as 8.05, TS fuzzy model predicted it as 7.83 and the ARMA model as 6.97 for Galata Bridge station. Also it is seen that FTS fuzzy model exhibits more accurate results when evaluation is considered in terms of average relative error (Table 4). For Kasimpasa station the mean observed value 8.44 is predicted as 8.51 and 8.05 for TS fuzzy model and ARMA model, respectively. Under FTS model, the predicted results is 8.47 and the mean relative error plays best among three models (Table 5).

Furthermore, fuzzy time series modeling has advantages over the other two models. There are amounts of computation in TS model, while some basic assumptions are needed for the ARMA type modeling such as at least 50 and preferably more than 100 are required to apply ARMA models (Chen, 2000), etc. However, in fuzzy time series model there is no need of considering these assumptions. And it can be applied to the case while historical data are linguistic values.

In the following, we also present the comparison results in figures.

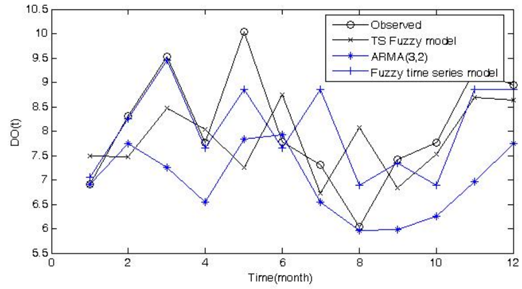

Figure 1: Comparisons of DO amounts forecasts for Galata Bridge Stations

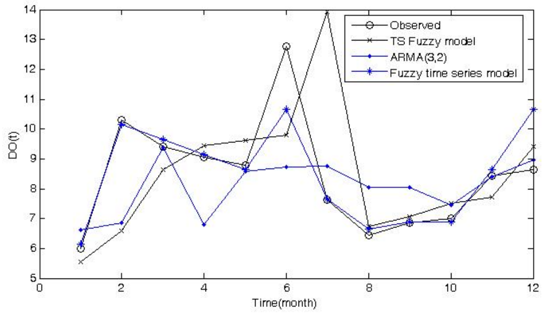

Figure 2: Comparisons of DO amounts forecasts for Kasimpasa Station

In figures, we can see that fuzzy time series model follows the general trend more closely than the other two approaches.

Follow the above numerical and graphical comparisons, we can conclude that the fuzzy time series model plays best in the forecasting of DO amount among three models.

5 Conclusions

In this paper, we introduce fuzzy time series model and focus on the dissolved oxygen measurement prediction in Golden Horn. Fuzzy time series would be efficient to cooperating historical data with linguistic value, and would improve forecast results. We then applied to the fuzzy time series model to the DO prediction in Golden Horn in test. Empirical analyses indicate that the fuzzy time series plays best in the predication among three models. Therefore, it would be an efficient way to improve forecasting.

Acknowledgments

The authors would like to acknowledge the supported by the National Science Fundation: Bionic intelligent control method of chaotic ecosystems under slow disturbance (11072090).

References

[1] A. Altunkaynak Fuzzy logic modeling of the dissolved oxygen fluctuations in Golden Horn. *Ecological Modelling*, 189(2005):436-446.

[2] B.A.Cox, A review of dissolved oxygen modelling techniques for lowland rivers. *Sci.Total Environ*, 314-316(2003):303C334.

[3] A. Garcia and J.A. Revilla and R. Medina and C. Alvarez and J.A. Juanes, A model for predicting the temporal evolution of dissolved oxygen concentration in shallow estuaries. *Hydrobiologia* , 475-476 (2002):205C211.

[4] J.B. Butcher and S. Covington Dissolved-oxygen analysis with temperature dependence. *J.Environ.Eng.* 121(1995):756C759.

[5] X. Fang and H.G. Stefan, Interaction between oxygen transfer mechanisms in lake models. *J. Environ. Eng.*, 121(1995):447C454.

[6] L.E. Miranda and J.A. Hargreaves and S.W. Raborn Predicting and managing risk of unsuitable dissolved oxygen in a eutrophic lake. *Hydrobiologia*, 457(2001):177C185.

[7] N. Quinn and K. Jacobs and C.M. Chen and W.T. Stringfellow, Elements of a decision support system for real-time management of dissolved oxygen in the San Joaquin River Deep Water Ship Channel. *Environment Modelling & Software*, 20(2005):1495-1504.

[8] A. Carlos and A. Rui and A. Jorge, Dissolved oxygen control of the activated sludge wastewater treatment process using stable adaptive fuzzy control. *Computers & Chemical Engineering*, 37(2012):156-172.

[9] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8(1975): 199-249.

[10] Q. Song and B.S. Chissom, Fuzzy time series and its models. *Fuzzy sets and Systems* , 54(1993):269-277.

[11] Q. Song and B.S. Chissom, Forecasting enrollments with fuzzy time series-part 1. *Fuzzy sets and Systems* , 54(1993):1-9.

[12] Q. Song and B.S. Chissom, Forecasting enrollments with fuzzy time series-part 2. *Fuzzy sets and Systems* , 62(1994):1-8.

[13] S.M.Chen, Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, 81(1996):311-319.

[14] K.H. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems*, 123(2001):387-394.

[15] H.K. Yu, A refined fuzzy time-series model for forecasting. *Physica A*, 346(2005): 657-681.

[16] H.K. Yu, Weighted fuzzy time series models for TAIEX forecasting. *Physica A*, 349(2005): 609-624.

[17] M. Versaci and F.C. Morabito, Fuzzy time series approach for disruption prediction in Tokamak reactors. *IEEE Trans. Magnetics*, 39(2003):1503-1506.

[18] S.M.Chen and J.R.Hwang, Temperature prediction using fuzzy time series. *IEEE Trans Syst Man Cybern*, 30(2000):263-275.

[19] D. Domańska and M. Wojtylak Application of fuzzy time series models for forecasting pollution concentrations. *Expert Systems with Applications*, 39(9)(2012):7673-7679.

[20] C.C.Wang, A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. *Expert Systems with Applications*, 38(8)(2011):9296-9304.

[21] C. Ruey and C. Ting, The adaptive fuzzy time series model with an application to Taiwans tourism demand. *Expert Systems with Applications*, 38(8)(2011):9164-9171.

[22] J. Sullivan and W.H. Woodall, A comparison of fuzzy forecasting and Markov modeling. *Fuzzy Sets Systems*, 64(1994):279-293.