An Algorithm Forecasting Time Series Using Wavelet

Kais Ismail Ibraheem [1], Eman Bacheer Abdelahad [2]

[1] **Department of Computer, University of Mosul, Mosul, Iraq**

[2] **Department of Mathematic, University of Mosul, Mosul, Iraq**

Abstract

In this paper we used the technique of wavelets with fuzzy logic to forecast enrollment of Alabama university from 1971 to 1994 where data were taken and analyzed by using wavelets, then logic , and we used the mean square error (MSE) to compare the forecasting results with previous different forecasting methods. The results were acceptable compared with the results of previous research.

1- Introduction

Time series forecasting are widely used in many areas, Such as economics, inventory, systems, statistics, etc.

Forecasting is one of the important activities in business Enrollment, finance, etc. that helps in decision making. The Classical time series methods can not deal with forecasting problems in which the values of time series are linguistic terms represented by using wavelet and fuzzy logic.

Wavelets turned out to be very useful when applied to many Problems including analysis and synthesis of time series in both time and scale [13]. Foundations of wavelet based analysis method were laid in the beginning of the 20^{th} century. Back then, in the year 1909 Hungarian mathematician Alfred haar introduced his two-state function in appendix to his doctoral thesis published later on [6] lately a very fast development of wavelet-based data mining [18] techniques may be observed.

Fuzzy set theory is first presented by Zadeh (1965) for treatment of uncertain environment inseveral fields. Particularly fuzzy logic designs are well accepted and established for electronic devices and later fuzzy sets found a broad application potential on various studyfields [5]. Song and Chissom [15] introduced a theory for fuzzy time Series and applied fuzzy time series methods [16], [17] that modeled the enrollments of the university of Alabama, in recent years a number of techniques have been proposed for forecasting based on fuzzy set theory methods. Chen presented a method to forecast the enrollments of the university of Alabama based on fuzzy time series [1] . In [8] Huang extended Chen's work presented in [1] and used simplified calculations with the addition of heuristic rules to forecast the enrollments.

The rest of this paper is organized as follows. In section (2) we briefly review wavelet transform, in section (3) we deal with Definitions of the fuzzy time series. In section (4) we use the theory of wavelet transform and fuzzy time series to propose a new method to forecast the enrollment of the university of Alabama. In section (5) we compare the forecasting result of the forecasting result of the proposed method with the existing methods, and the conclusions are discussed in this section .

2-Wavelet Transform

According to Fourier theory, a signal can be expressed as the sum of a series of sines and cosines. This sum is also called a Fourier expansion (see Eq. (1)).However, a serious drawback of the Fourier transform is that it only has frequency resolution and no time resolution. Therefore, we can identify all the frequencies present in a signal, but we do not know when they are present. The wavelet theory is proposed [12]

$$F(w)=\int_{-\infty}^{\infty} f(t) \ e^{-jwt} \, dt$$

$$=\int_{-\infty}^{\infty} f(t) \ (\cos wt - j \sin wt) dt \qquad .....(1)$$

Mathematically , a wavelet can be defined as a function with a zero average:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \qquad ....(2)$$

and

$$\int_{-\infty}^{\infty} \psi^2(t) dt = 1 \qquad ....(3)$$

Where ψ be a real function of real variable t. Condition (2) means that for any $\in$ from an interval (0,1) there is an interval (-T,T) such that [11]

$$\int_{-T}^{T} \psi^2(t) dt = 1 - \in$$

If $\in$ is close to 0 , it may be seen that only in an interval (-T,T) Corresponding to this $\in$ values $\psi(t)$ are different than 0 . Outside of this interval they must equal 0 interval(-T,T) is small compared to an interval $(-\infty, \infty)$ on which a wale function is determined condition (2) implies that if $\psi(t)$ has some positive values , it also has to have some negative ones. If Haar function ϕ , which is a two-state function of real variable (fig. 1)

$$\phi(t) = \begin{cases} -1 & if \quad -1 \le t \le 0 \\ 1 & if \quad 0 \le t \le 1 \\ 0 & if \quad otherwise \end{cases}$$

Would be transformed into

$$\psi^{(H)}(t) = \begin{cases} -\frac{1}{\sqrt{2}} & if \quad -1 \le t \le 0 \\ \frac{1}{\sqrt{2}} & if \quad 0 \le t \le 1 \\ 0 & if \quad otherwise \end{cases}$$

Then the resulting function $\psi^{(H)}$ satisfies condition (2) and (3) called Haar basic wavelet function (fig. 2) [11]

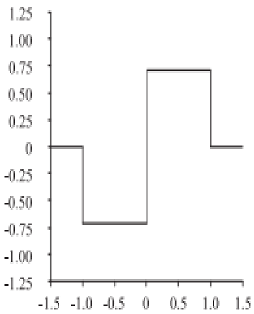

Fig 1 Haar funcation

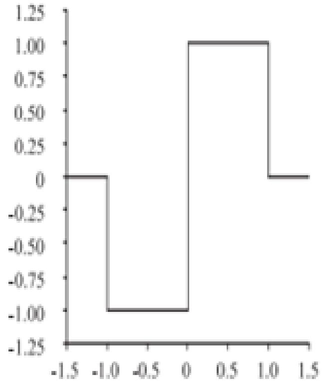

Fig. 2 Modified function-the first basic wavelet function

3-Fuzzy time series definition

In this section briefly summarizes basic fuzzy time series.

Definition (1)

Assume $Y(t) \subset R \quad (t = ...0,1,2,.....)$ to be a universe of discourse defined by the fuzzy set $f_i(t)$. $F(t)$ consisting of $f_i(t)$, $i = 1,2,...$, is defined as a fuzzy time series on $Y(t)$ .At that $F(t)$ can be understood as a linguistic variable , where $f_i(t)$, $i = 1,2,.....$ are possible linguistic values of $F(t)$. [14]

Definition (2)

If there exists a fuzzy relationship $R(t\,,t-1)$ such that $F(t) = F(t-1) \times R(t-1\,,t)$ where symbol ($\times$) is an operator , then $F(t)$ is said to be caused by $F(t-1)$. The existing relationship between $F(t)$ and $F(t-1)$ can be denoted by the expression $F(t-1) \rightarrow F(t)$. [14]

Definition (3)

Let $R(t\,,t-1)$ be a first-order model of $F(t)$ .If for any t, $R(t\,,t-1) = R(t-1\,,t-2)$, then $F(t)$ is called a time-invariant fuzzy time series. Otherwise, it is called a time-variant fuzzy time series.[4]

Definition (4)

Suppose $F(t-1) = A_i$ and $F(t) = A_j$, thefirst order univariate fuzzy logical relationship can be defined as $A_i \rightarrow A_j$, where A_i and A_j are called the left-hand side (LHS) and right-hand side (RHS) of the fuzzy logical relationship respectively.[5]

4- Anew method proposed

In this section , we present a new method to forecast the enrollments of the university of Alabama by using wavelets transform and fuzzy. This method is described in three partitions .

Partition one .

The original data were taken from the university of Alabama, they were applying the Haar wavelet by using algorithm A which is described below and we get the following data which is explained in table (1)

Algorithm -A-

Procedure Standard Decomposition

(c: array $[1....2^j,1.....2^k]$ of reals)

for $row \leftarrow 1 \quad to \quad 2^j \quad do$

Decomposition($c[row,1.....2^k]$)

end for

for $col \leftarrow 1 \quad to \quad 2^k \quad do$

Decomposition ($c[1.....2^j, col]$)

end for

end procedure

Table -1- The original data and wavelet

Year	*enrollment*	*wavelet*
1971	13055	1782
1972	13563	1804
1973	13867	-259
1974	14696	-41
1975	15460	3683
1976	15311	-217
1977	15603	3819
1978	15861	-145
1979	16807	2019
1980	16919	2004
1981	16388	192
1982	15433	-29
1983	15497	3943
1984	15145	153
1985	15163	3927
1986	15984	-68
1987	16859	2262
1988	18150	2314
1989	18970	-200
1990	19328	-109
1991	19337	4734
1992	18876	7
1993	18909	4754
1994	18707	77

Partition two.

After that the data which we obtained from the previous partition were applied in fuzzy logic algorithm B which is described below , and we get these data in table -3-

Algorithm –B-

1- Define the universe of discourse U = $[a.b]$ where a is the minimum value little less than it of the wavelet data obtained by partition one and b is the maximum value or little more than it of the wavelet data obtained by partition one.

2- Divide U = $[a.b]$ into several equal-length intervals . $u_1, u_2, ..., u_n$ where n is the positive integer number .

3- Define ench fuzzy set X_i based on the re – divided intervals .

4- Determine fuzzy logical relationships $X_i \to X_j$ where X_i is fuzzified enrollment wavelet of the year $(n-1)$ and X_j is the fuzzified enrollments wavelet of the year (n),table (2) shows fuzzy logical relationship .

Table -2- fuzzy logical relationships

$X_7 \to X_7$	$X_1 \to X_6$	$X_{10} \to X_1$	$X_2 \to X_{10}$
$X_7 \to X_1$	$X_6 \to X_6$	$X_1 \to X_6$	$X_{10} \to X_2$
$X_1 \to X_2$	$X_6 \to X_2$	$X_6 \to X_6$	
$X_2 \to X_{11}$	$X_2 \to X_1$	$X_6 \to X_1$	
$X_{11} \to X_1$	$X_1 \to X_{10}$	$X_1 \to X_1$	
$X_1 \to X_{11}$	$X_{10} \to X_2$	$X_1 \to X_{10}$	
$X_{11} \to X_1$	$X_2 \to X_{10}$	$X_{10} \to X_2$	

5-Divide each interval derived in step (2) into four subintervals of equal length where the 0.25 point and 0.75 point of each interval are used as the upward and downward forecasting points of the forecasting . We use the following rules to determine whether the trend of the forecasting goes up, down or middle .

1-If $j \geq i$, $j \leq i$ and

$$((W_{n-1} - W_{n-2}) - (W_{n-2} - W_{n-3})) \rangle 0 \Rightarrow$$

the trend of the forecasting will go up and we use rule (1) to compute the forecast value , 2- If $j \geq i$, $j \leq i$ and

$((W_{n-1} - W_{n-2}) - (W_{n-2} - W_{n-3})) \langle 0 \Rightarrow$ the trend of the forecasting will go down and we use rule (2) to compute the forecast value

Rule (1)

Let $|(W_{n-1} - W_{n-2}) - (W_{n-2} - W_{n-3})| = W^*$

1- If $(W^* * 2 + W_{n-1}) \in X_j$

or $(W_{n-1} - W^* / 2) \in X_j$

then the trend of the forecasting of the interval will be upward (0.75 point) .

2- If $(W^* / 2 + W_{n-1}) \in X_j$

or $(W_{n-1} - W^* / 2) \in X_j$

then the trend of the forecasting of this interval will be downward (0.25 point) , if neither of these cases , then the trend of the forecasting of this interval will be middle .

Rule (2)

Let $|(W_{n-1} - W_{n-2}) - (W_{n-2} - W_{n-3})| = W^*$

1- if $(W^* / 2 + W_{n-1}) \in X_j$

or $(W_{n-1} - W^* / 2) \in X_j$

then the trend of the forecasting of this interval will be downward(0.25 point)

2- If $(W^* * 2 + W_{n-1}) \in X_j$

or $(W_{n-1} - W^* * 2) \in X_j$

then the trend of the forecasting of this interval will be upward (0.75 point)

If neither of these cases , then the trend of the forecasting of this interval will be middle . [3]

Partition three .

After that the data which taken from the second step were be back by using algorithm inverse wavelet C .

Algorithm –C-

Procedure Standard Redecom Position

(c:array[$1..2^j, 1..2^k$] of reals)

for $col \leftarrow 1\ to\ 2^k\ do$

Red composition(c[$1..2^j, col$])

end for

for $row \leftarrow 1\ \ to\ \ 2^j\ do$

Red composition(c[$row, 1..2^k$])

end for

end procedure.

Table-3- Actual enrollment and forecasting enrollment by fuzzy and wavelet of the University of Alabama

Year	*Enrollment*	*Fuzzy Set*	*Forecast by Fuzzy and Wavelet*	e_i
1971	13055	X_7	13350	295
1972	13563	X_7	13669	106
1973	13867	X_1	14155	288
1974	14696	X_2	14155	541
1975	15460	X_{11}	15162	298
1976	15311	X_1	14835	476
1977	15603	X_{11}	15644	41
1978	15861	X_1	15644	217
1979	16807	X_6	16754	53
1980	16919	X_6	16073	846
1981	16388	X_2	15978	411
1982	15433	X_1	15405	29
1983	15497	X_{10}	15550	53
1984	15145	X_2	14927	218
1985	15163	X_{10}	14803	360
1986	15984	X_1	15376	609
1987	16859	X_6	17204	345
1988	18150	X_6	18224	74
1989	18970	X_1	19050	80
1990	19328	X_1	18814	514
1991	19337	X_{10}	18988	349
1992	18876	X_2	18224	652
1993	18909	X_{10}	18526	383
1994	18707	X_2	18290	417

Table -4- A comparison of the forecasting results of different forecasting methods

Year	*Enrollment*	*Song[16]*	*Song[17]*	*Chen[1]*	*Hwang[8]*	*Huarng[7]*	*Chen[2]*	*Jilani[10]*	*Proposed method*
1971	13055	-----	-----	-----	-----	-----	-----	14464	13350
1972	13563	14000	-----	14000	-----	14000	-----	14464	13669
1973	13867	14000	-----	14000	-----	14000	-----	14464	14155
1974	14696	14000	-----	14000	-----	14000	14500	14710	14155
1975	15460	15500	14700	15500	-----	15500	15500	15606	15162
1976	15311	16000	14800	16000	16260	15500	15500	15606	14835
1977	15603	16000	15400	16000	15511	16000	15500	15606	15644
1978	15861	16000	15500	16000	16003	16000	15500	15606	15644
1979	16807	16000	15500	16000	16261	16000	16500	16470	16754
1980	16919	16813	16800	16833	17407	17500	16500	16470	16073
1981	16388	16813	16200	16833	17119	16000	16500	16470	15978
1982	15433	16789	16400	16833	16188	16000	15500	15606	15405
1983	15497	16000	16800	16000	14833	16000	15500	15606	15550
1984	15145	16000	16400	16000	15497	15500	15500	15606	14927
1985	15163	16000	15500	16000	14745	16000	15500	15606	14803
1986	15984	16000	15500	16000	15163	16000	15500	15606	15376
1987	16859	16000	15500	16000	16384	16000	16500	16470	17204
1988	18150	16813	16800	16833	17659	17500	18500	18473	18224
1989	18970	19000	19300	19000	19150	19000	18500	18473	19050
1990	19328	19000	17800	19000	19770	19000	19500	19155	18814
1991	19337	19000	19300	19000	19928	19500	19500	19155	18988
1992	18876	-----	19600	19000	15837	19500	18500	18473	18224
1993	18909	-----	-----	-----	-----	-----	-----	-----	18526
1994	18707	-----	-----	-----	-----	-----	-----	-----	18290
MSE		775687	407507	321418	226611	86694	86694	227194	147488

5- Conclusion :

In this paper we proposed a new algorithm to forecast enrollment of Alabama university from 1971 to 1994 , where we used wavelet with logic and applied this method to predict these data. Table -3- contains registration data real values forecasted by using the method described previously . Table -4- included a comparison of predictions of the proposed method with previous methods. Researchers formerly used mean square error (MSE) to compare the forecasting results of different methods . Results obtained using this method were good compared with the methods used previously ; except for three values where error rate appeared to be of fairly large compared with other values . Each value appeared in the range of three groups into which these data are .

Refrains

[1] Chen, S.N. "Forecasting enrollments based on fuzzy time series", Set and System, Vol. 81, 1996, pp. 311-319.

[2] Chen, S.M. "Forecasting enrollments based on high-order fuzzy timeseries". Cybernetics and Systems: An International Journal , Vol. 33, 2002, pp. 1-16.

[3] Chen, S.M. and Hsu, C.C."A new method to forecast enrollmentsusing fuzzy time series". International Journal of applied science and engineering, Vol. 3, No. 2, 2004, pp. 234-244.

[4] Chen, S. and Chung, N. " Forecasting enrollments of students by using fuzzy time series and genetic algorithms", information and management sciences, Vol. 17, No. 3, 2006, pp. 1-17.

[5] Duru, O. and Yoshida, S. "Fuzzy time series methods for shippingfreight markets". Lisbon, portual, vol. 11, no. 15, 2010, pp. 1-13.

[6] Harr, A. "Zur theorie der orthogonalen funkion systeme" , Math.Ann. Vol. LXIX. 1910, pp. 331-371.

[7] Huarng, K. "Heuristic models of fuzzy time series for forecasting",Fuzzy sets and systems, 2001, 123: pp. 369-386.

[8] Hwang, J.R., Chen, S.M. and Lee, C.H. "Handling forecasting proble- ms using fuzzy time series". Fuzzy sets and systems, Vol. 100,pp. 217-228.

[9] Jasim, H.T., Ibraheem, K.I. and Salim, A.G.J. "A hybrid algorithms toforecast enrollment based on genetic algorithms and fuzzy timeseries", Iajit First Online Publication, 2013.

[10] Jilani, T.A. , Burney, S.M.A. and Ardil, C. "Fuzzy metric approachfor fuzzy time series forecasting based on frequency density based partitioning proceedings of world academy of science, engineeringand technology". Vol. 23, 2007, pp. 333-338.

[11] Kozlowski, B."Time series denoising with wavelet transform", Journal of telecommu-nication and information technology, Vol. 3, 2005, pp. 91-95.

[12] Norgaard, C. "Financial time series forecasting using improvedwavelet neural network", Pedersen, 2009.

[13] Percival, D.B. and Walden, A.T. "Wavelet methods for time series analysis", Cambridge: Cambridge university press, 2000.

[14] Sah, M. and Degtiarev, K.Y. "Forecasting enrollment model based on first-order fuzzy time series" Proceedings of world academy of science, engineering and technology . Vol. 1, 2005, pp. 375-378.

[15] Song, Q. and Chissom, B.S. "Fuzzy time series and its models", Fuzzy set and systems.Vol. 54,1993, pp. 269-277.

[16] Song, Q. and Chissom, B.S. "Forecasting enrollments with fuzzy time series ", Part 1, fuzzy sets and systems ,Vol. 54, 1993, pp. 1-9.

[17]Song, Q. and Chissom, B.S., " Forecasting enrollments with fuzzytime series", Part 11, fuzzy sets and systems. Vol. 62, 1994, pp. 1-8 .

[18] Zhu, S. , Li, Q. and Ogihara, M. , "Survey on wavelet applications in data mining", Vol. 4, no. 2, 2003, pp. 49-68.

First Author: Kais Ismail Ibraheem received the Msc in Applied Mathemtics and Phd degree in Electrical Engineering from Gadjah Mada University Indonesia, in 1998 and 2003, respectively.

He is currently a head of department of Computer Sciences faculty of education, Mosul University, Iraq. His research interests include artificial intelligence,wavelet, numerical analysis, etc.

Second Author: Eman Bacheer Abdelahad received her Bsc degree in Mathematics, Department of Mathmatics, College of Education, Mosul University, Iraq, 1985.

Her Msc Degree in Time Series from Department of Mathmatics, College of Sciences, 1990. Her research interests include forecasting Time Series by using wavelet and Fuzzy Time Series.She is author of 6 research articles.