Forecasting Students' Enrolment in Fuzzy Time Series Based on Three Classes of T-Norm of Subsethood Defuzzification

Nazirah Ramli
Faculty of Computer Sciences and Mathematics
Universiti Teknologi MARA Pahang, Malaysia
Email: nazirahr@pahang.uitm.edu.my

Abu Osman Md. Tap
International Islamic University Malaysia, Malaysia
Email: abuosman@kict.iiu.edu.my

ABSTRACT

Traditional time series methods can predict the seasonal problem but fail to forecast the problems with linguistic values. An improvised forecasting method by using fuzzy time series can be applied to deal with this problems. This paper presents three classes of t-norm of subsethood defuzzification that are algebraic product, Einstein product and minimum in forecasting the students' enrolments based on fuzzy time series. The proposed method uses the historical data of students' enrolment and applies seven and ten intervals with equal length and the max-product and max-min as the composition operator in the fuzzy relations. The result shows that the t-norm of algebraic product class of subsethood defuzzification model with (10, max-product) is the best forecasting methods in terms of accuracy. The t-norm of algebraic product class with (10, max-product) also achieves higher forecasting accuracy rates compared to some of the existing methods.

Keywords: *forecasting enrolments, max-min composition, max-product composition, subsethood defuzzification, t-norm classes*

ISSN 0128-5599

Introduction

People usually deal with many forecasting activities in their daily life, such as the temperature prediction, the weather prediction, the stock prediction, the earthquake prediction, the students' enrolment prediction, etc. The traditional time series methods can predict the seasonal problem but fail to forecast the problems with linguistic data. Furthermore, the traditional time series requires more historical data and the data must obey normal distribution (Cheng, Wang & Li, 2008). In order to cope with data in linguistic values, Song and Chissom (1993a) introduced the definition of fuzzy time series, which is capable of dealing with incomplete and vague data under uncertain circumstances. Due to its better performance in some kinds of forecasting problems, fuzzy time series model has drawn much attention to many researchers. It has been employed to deal with various domain of forecasting problems, such as university enrolment forecasting (Song & Chissom, 1993b, 1994; Chen, 1996; Chen, 2002; Huarng, 2001a; Chen & Hsu, 2004; Yu, 2005; Cheng et al., 2006; Singh, 2007; Singh, 2009; etc), stock price forecasting (Huarng & Yu, 2006; Yu, 2005; Chu et al., 2009; etc), temperature forecasting (Chen, 2000), financial forecasting (Lee et al., 2006), rainfall forecasting (Li et al., 2009), etc.

Various fuzzy time series models for forecasting students' enrolment have been developed and accuracy is one the main issues discussed. As an accurate estimation of students is important for decision making and budget planning for higher education, thus, researchers have proposed many enrolment forecasting methods to improve accuracy. However, obtaining accuracy is not an easy task, as many variables have impacts on the enrolment number.

Song and Chissom (1993b, 1994) explored forecasting of fuzzy time series with enrolment data of the University of Alabama and proposed a forecasting framework composed of four major steps: partitioning fuzzy sets, fuzzification, establishing fuzzy relations and defuzzification. Although various studies may place emphasis on different steps, most studies focus on establishing fuzzy relations (Chen, 1996; Chen & Hsu, 2004; Huarng & Yu, 2006; Yu, 2005; etc). In other related studies, Huarng (2001b) focuses on the partitioning step and Nazirah and Abu Osman (2000) focus on the partitioning, establishing fuzzy relations and defuzzification steps. Findings by Nazirah and Abu Osman (2000) indicate that the number of intervals, the type of composition operator and the defuzzification techniques could influence the forecasting results. In their study, the universe of discourse is divided into seven and intervals with equal length. Instead of using the max-min composition operator alone, Nazirah and Abu Osman (2000) also applied the max-product composition operator in establishing the fuzzy relations. The defuzzification of intervals method is applied in their research. In relation to that, Nazirah and Abu Osman (2006) further propose subsethood defuzzification with t-norm of Yager class in the defuzzification step. It has improved the forecasting results by Song and Chissom (1994), Song and Leland (1996), Nazirah and Abu Osman (2000) and Yu (2005). In 2008, Nazirah and Abu Osman apply the subsethood defuzzification with t-norm of Dombi class and it has also improved the forecasting results by Song and Chissom (1994), Song and Leland (1996), Nazirah and Abu Osman (2000) and Yu (2005) but slightly less accurate compared to Nazirah and Abu Osman (2006).

This paper attempts to approach the forecasting issue by proposing another three classes of t-norm; algebraic product, Einstein product and minimum in the subsethood defuzzification as each t-norm offers a good approach and it sees things under a different perspective (Iliadis et al., 2007). The proposed method also uses seven and ten intervals with equal length and the max-product and max-min as the composition operator. The analysis shows that the t-norm of algebraic product class of subsethood defuzzification with (10, max-product) increases the forecasting accuracy rate. The t-norm of algebraic product class with (10, max-product) also achieves higher forecasting accuracy rates compared to some of the existing methods.

Preliminaries

In this section, the definition and characteristic of t-norm are briefly reviewed.

T-Norms

Triangular norms (T-norms) are the functions used for intersection of fuzzy sets. A t-norm is a binary operation t on the interval $[0,1]$ satisfying the following properties:

(1) Commutativity : $t(a,b) = t(b,a)$
(2) Associativity : $t[t(s,b),c] = t[a,t(b,c)]$
(3) Monotonicity : If $a \leq a'$ and $b \leq b'$ then $t(a,b) \leq t(a',b')$

(4) Boundary conditions : $t(0,0) = 0$ and $t(a,1) = a$

Some of the t-norms classes are as follows:

(1) Minimum : $t_m(a,b) = \min(a,b)$
(2) Algebraic product : $t_{ap}(a,b) = ab$
(3) Bounded sum : $t_b(a,b) = \max(0, a + b - 1)$

(4) Drastic product :
$$t_{dp}(a,b) = \begin{cases} a & ,b = 1 \\ b & ,a = 1 \\ 0 & ,\text{otherwise} \end{cases}$$

(5) Dombi class :
$$t_\lambda(a,b) = \frac{1}{1 + \left[\left(\frac{1}{a} - 1\right)^\lambda + \left(\frac{1}{b} - 1\right)^\lambda\right]^{\frac{1}{\lambda}}},$$

$\lambda \in (0, \infty)$

(6) Yager class :
$$t_w(a,b) = 1 - \min\left[1, \left((1-a)^w + (1-b)^w\right)^{\frac{1}{w}}\right],$$

$w \in (0, \infty)$

(7) Einstein product :
$$t_{ep}(a,b) = \frac{ab}{2 - (a + b - ab)}$$

Theorem (Wang, 1997, p. 42)

For any t-norm which satisfies the commutative, associative, monotonic and boundary conditions properties, the following inequality holds for any $a,b \in [0,1]$, $t_{dp}(a,b) \le t(a,b) \le \min(a,b)$.

Fuzzy Time Series

In this section, the basic concept of fuzzy sets (Zadeh, 1965) and fuzzy time series (Song & Chissom, 1993b, 1994) are briefly overviewed. Let U be the universe of discourse with $U = \{x_1, x_2, \ldots, x_n\}$. A fuzzy set A_i of U is defined as:

$A_i = \mu_{A_i}(x_1)/x_1 + \mu_{A_i}(x_2)/x_2 + \mu_{A_i}(x_3)/x_3 + \ldots + \mu_{A_i}(x_n)/x_n$, where μ_{Ai} is the membership function of the fuzzy set A_i, such that $\mu_{A_i} : U \rightarrow [0,1]$ and $\mu_{Ai}(x_j)$ represents the grade of membership of x_j in A_i, $\mu_{Ai}(x_j) \in [0,1]$.

Definition 1

Let $Y(t)$ (t = ..., 0, 1, 2, ...), is a subset of $\Re$, be the universe discourse defined by the fuzzy set $f_i(t)$. If $F(t)$ is the collection of $f_i(t)$ ($i = 1, 2, \ldots$), then $F(t)$ is defined as a fuzzy time series on $Y(t)$ ($t = \ldots, 0, 1, 2, \ldots$).

Definition 2

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

Definition 3

Supposed $F(t)$ is caused by $F(t-1)$ only, and $F(t) = F(t-1) \times R(t,t-1)$. For any t, if $R(t,t-1)$ is independent of t, then $F(t)$ is named a time-invariant fuzzy time series, otherwise it is a time-variant fuzzy time series.

Subsethood Defuzzification

The subsethood defuzzification (SD) is based on sigma-count measurement and mean value theorem (Oliveira, 1995). The SD coincides with the unifying structure of fuzzy set theory, that is, with the Kosko's subsethood theorem (Kosko, 1992).

Without the loss of generality, the universe discourse is normalized to $X \in [0,1]$. For a given fuzzy set A, the averaging defuzzification in terms of set measures is given by $\hat{x} = \frac{M_p(A \cap M_A)}{M_p(A)}$ where the sigma-count measure of a set denoted as $M_p(X) = \sqrt[p]{\mu^p{}_x(x_1) + \ldots + \mu^p{}_n(x_n)}$, the mirror set of support A is expressed as $M_A = \sum_{i=1}^{n} \mu_M(x_i)/x_i$ and $\mu_M(x_i) \cong x_i$. The intersection of fuzzy sets A and M_A denoted as $A \cap M_A$, is defined as $\mu_{A \cap M_A}(x) = \mu_A(x)\, t\, x_i$ where t is a triangular norm. From the averaging defuzzification, the SD method can be written as

$$x_{SD} = \frac{\sqrt[p]{\sum_{i=1}^{n} (\mu_A(x_i)t\, x_i)^p}}{\sqrt[p]{\sum_{i=1}^{n} \mu^p{}_A(x_i)}}$$

where for $i = 1, 2, \ldots n$, $\mu_A(x_i)$ is the degree of the membership of the i-th element in the support of A and $p > 0$ is the sigma count parameter.

Enrolment Forecasting

The enrolment forecasting using three t-norm classes of subsethood defuzzification; algebraic product, Einstein product and minimum is implemented on a well known historical data that is the students' enrolment at University of Alabama. Seven and ten equal intervals with the max-min and max-product composition operator from Nazirah and Abu Osman (2000) are applied. The enrolment forecasting procedure using fuzzy time series which is basically from Song and Chissom (1994) with some modifications on the interval length, type of operator and defuzzification method is described step by step as follows:

Step 1: Define the universe of discourse U within the historical data. Let U= [13000, 2000] and for seven and ten equal intervals, the length is 1000 and 700 respectively.

Step 2: Seven or ten linguistic values must be determined to define fuzzy sets on the universe U. The A_i (i = 1, 2, 3, ...) are the possible linguistic values for *enrolment*. For seven intervals, $A_1, A_2, A_3, \ldots, A_7$ represent poor enrolment, below average enrolment, average enrolment, good enrolment, very good enrolment, excellent enrolment and extraordinary enrolment respectively. Each A_i is defined by the intervals $u_1, u_2, u_3, \ldots, u_7$ as follows:

$$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + \ldots + 0/u_6 + 0/u_7$$
$$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + \ldots + 0/u_6 + 0/u_7$$
$$A_3 = 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + \ldots + 0/u_6 + 0/u_7$$
$$\ldots$$
$$A_6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$$
$$A_7 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7$$

Table 1: The Fuzzified Historical Enrolments for Seven Intervals

Year	Actual enrolment	Fuzzified enrolment
1971	13055	A_1
1972	13563	A_1
1973	13867	A_1
1974	14696	A_2
1975	15460	A_3
1976	15311	A_3
1977	15603	A_3
1978	15861	A_3
1979	16807	A_4
1980	16919	A_4
1981	16388	A_4
1982	15433	A_3
1983	15497	A_3
1984	15145	A_3
1985	15163	A_3
1986	15984	A_3
1987	16859	A_4
1988	18150	A_6
1989	18970	A_6
1990	19328	A_7
1991	19337	A_7
1992	18876	A_6

Step 3: Choose a model basis $w = 4$ and at a given time t, calculate the fuzzy relation $R^{w}(t,t-1) = f^{T}(t-2) \times f(t-1) \cup f^{T}(t-3) \times f(t-2) \cup \ldots f^{T}(t-w) \times f(t-w+1)$ and fuzzy forecast $F(t) = F(t-1) \circ R(t,t-1)$ where $\circ$ is the max-min and max-product composition.

Table 2: The Fuzzy Forecast for Seven Intervals and Max-product Composition Operator

Year	Fuzzy output						
1975	0.5	0.5	0.5	0	0	0	0
1976	0.25	0.5	0.5	0.5	0	0	0
1977	0.25	0.5	1	0.5	0	0	0
1978	0	0.5	1	0.5	0	0	0
1979	0	0.5	1	0.5	0	0	0
1980	0	0.5	0.5	0.5	0.5	0	0
1981	0	0.5	0.5	1	0.5	0	0
1982	0	0	0.5	1	0.5	0	0
1983	0	0.5	0.5	0.5	0.5	0	0
1984	0	0.5	1	0.5	0.5	0	0
1985	0	0.5	1	0.5	0	0	0
1986	0	0.5	1	0.5	0	0	0
1987	0	0.5	1	0.5	0	0	0
1988	0	0.5	0.5	0.5	0.5	0	0
1989	0	0	0	0	0.25	0.25	0.25
1990	0	0	0	0	0.5	1	0.5
1991	0	0	0	0	0.5	0.5	0.5
1992	0	0	0	0	0.5	0.5	1
1993	0	0	0	0	0.5	0.5	1

Step 4: Interpret the forecast outputs. By using the SD with t-norm of algebraic product $t_{ap}(\mu(x_i),x_i) = \mu(x_i)x_i$, Einstein product $t_{ep}(\mu(x_i), x_i) = \dfrac{\mu(x_i).x_i}{2-(\mu(x_i)+x_i-\mu(x_i).x_i)}$ and minimum class $t_{min}(\mu(x_i),x_i) = \min(\mu(x_i),x_i)$, the forecast enrolment can be written as

$$x_{SD_{ap}} = \frac{\sqrt[p]{\sum_{i=1}^{n}[\mu(x_i).x_i]^p}}{\sqrt[p]{\sum_{i=1}^{n}\mu^p(x_i)}} \text{ for algebraic product,}$$

$$x_{SD_{ep}} = \frac{\sqrt[p]{\sum_{i=1}^{n}\left[\frac{\mu(x_i).x_i}{2-(\mu(x_i)+x_i-\mu(x_i).x_i)}\right]^p}}{\sqrt[p]{\sum_{i=1}^{n}\mu^p(x_i)}} \text{ for Einstein product and}$$

$$x_{SD_{\min}} = \frac{\sqrt[p]{\sum_{i=1}^{n}[\min(\mu(x_i),x_i)]^p}}{\sqrt[p]{\sum_{i=1}^{n}\mu^p(x_i)}} \text{ for minimum class.}$$

Results and Discussion

Figures 1-4 show the forecasting results of t-norm classes of algebraic product, Einstein product and minimum with four different pairs of intervals and composition operator.

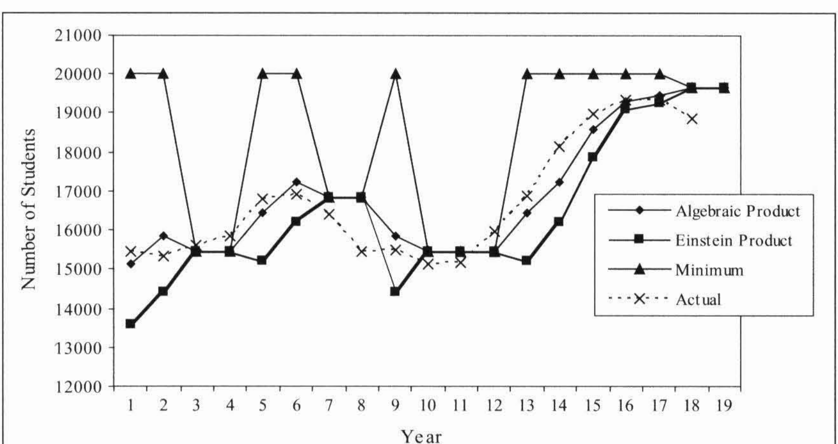

Figure 1: Forecast and Actual Enrolments with (10, max-product)

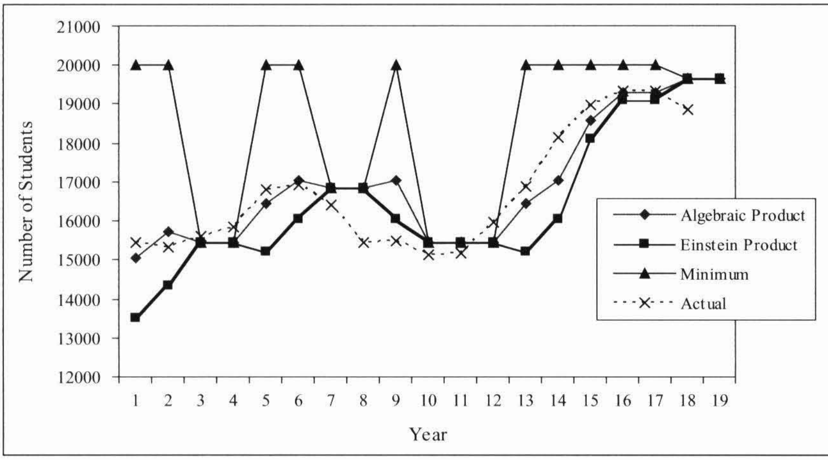

Figure 2: Forecast and Actual Enrolments with (10, max-min)

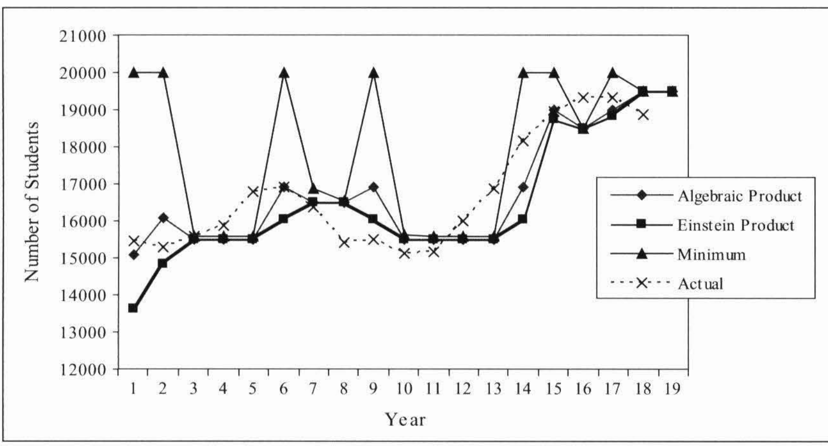

Figure 3: Forecast and Actual Enrolments with (7, max-product)

For t-norm class of algebraic product with (7, max-product) and (7, max-min), the best forecast enrolments (with the smallest errors) occurs at sigma count parameter $p = 38$ and 37 respectively, whereas for (10, max-product) and (10, max-min), the best forecasting occurs at the same sigma count parameter ($p = 50$). Similar with Einstein product class, the best forecasting occurs at $p = 50$ for both intervals. However, for the minimum class, the best forecasting occurs at different value of p where for (10, max-product) occurs at $p = 14$, (10, max-min) at $p = 16$, (7, max-product) at $p = 8$ and (7, max-min) at $p = 9$.

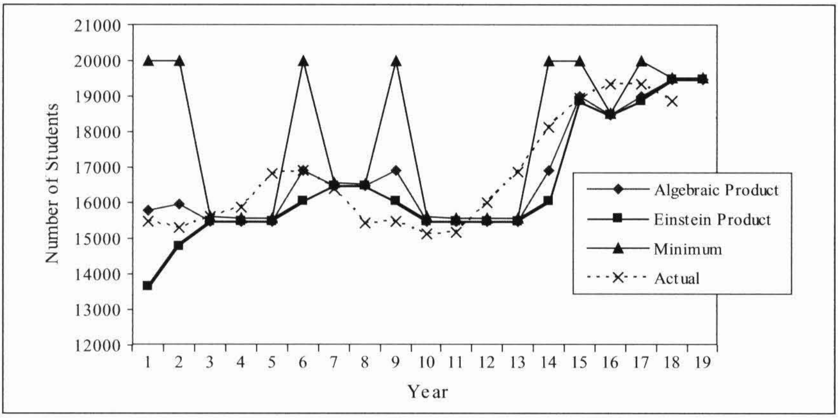

Figure 4: Forecast and Actual Enrolments with (7, max-min)

Table 3: Comparison of Performance Indicator for Three Classes of T-norms

T-norm	Type of interval & composition operator	MAD	RMSE	MAPE
Algebraic product	(10, max-product)	448.11	547.52	2.72
	(10, max-min)	512	663.24	3.13
	(7, max-product)	616.06	772.90	3.71
	(7, max-min)	607.44	765.57	3.67
Einstein product	(10, max-product)	864.56	1050.35	5.21
	(10, max-min)	857.44	1053.25	5.15
	(7, max-product)	750.11	937.89	4.48
	(7, max-min)	748.22	939.17	4.47
Minimum	(10, max-product)	1761.28	2372.04	10.80
	(10, max-min)	1761.39	2372.05	10.80
	(7, max-product)	1524.94	2154.32	9.29
	(7, max-min)	1510.67	2153.33	9.30

The accuracy of the forecasting results are measured by using three criteria, namely, mean absolute deviation (MAD), root mean square error (RMSE) and mean absolute percentage error (MAPE) as in Table 3.

Among the forecast enrolments for t-norm of algebraic product, Einstein product and minimum, the algebraic product has the best accuracy since its MAD, RMSE and MAPE are the lowest for all types of intervals and composition operator. Besides that, the t-norm of algebraic product with (10, max-product) is also the best forecasting in terms of accuracy compared to other types of forecasting methods.

The performance (in terms of accuracy) of the forecasting result for t-norm of algebraic product (10, max-product) is compared with other forecasting models (Table 4). The RMSE for algebraic product class (10, max-product) is better than the other previous forecasting models. It shows that algebraic product class with (10, max-product) outperforms the models by Song and Chissom (1993, 1994), Sullivan and Woodall (1994), Song and Leland (1996), Chen (1996), Yu (2005) and Nazirah and Abu Osman (2000, 2006, 2008) in forecasting the students' enrolment in fuzzy time series.

Table 4: Comparison of RMSEs with Various Models

Method	RMSE
Song and Chissom (1993b)	642.26
Song & Chissom (1994)	880.73
Sullivan and Woodall (1994)	621.33
Chen (1996)	638.36
Song & Leland (1996)	796.36
Nazirah & Abu Osman (2000)	883.25
Yu (2005)	1020.38
Nazirah & Abu Osman (2006)	557.91
Nazirah & Abu Osman (2008)	563.93
Algebraic product (10, max-product)	547.52

Conclusion

This paper presents three classes of t-norm of subsethood defuzzification; that is, algebraic product, Einstein product and minimum, in forecasting the university students' enrolment based on fuzzy time series. The algebraic product class with (10, max-product) is the best forecasting method compared to Einstein product and minimum classes with two types of length and composition operators. Furthermore, the algebraic product class with (10, max-product) has also improved the forecasting results by Song and Chissom (1993, 1994), Sullivan and Woodall (1994), Song and Leland (1996), Chen (1996), Yu (2005) and Nazirah and Abu Osman (2000, 2006, 2008).

References

Chen, S.M. (1996). Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems, 81, 311-319. Chen, S.M. (2000). Temperature prediction using fuzzy time series. IEEE Transactions on Systems, Man and Cybernetics, 30, 263-275.

Chen, S.M. (2000). Temperature prediction using fuzzy time series. IEEE Transactions on Systems, Man and Cybernetics, 30, 263-275.

Chen, S.M. (2002). Forecasting enrollments based on high-order fuzzy time series. *Cybernetics and Systems*, 33, 1-16.

Chen, S.M. & Hsu, C.C. (2004). A new method to forecast enrollment using fuzzy time series. *International Journal of Applied Science and Engineering*, 3, 234-244.

Cheng, C.H., Chang, J.R. & Yeh, C.A. (2006). Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost. Technological Forecasting & Social Change, 73, 524-542.

Cheng, C.H., Wang, J.W. & Li, C.H. (2008). Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix. *Expert Systems with Applications*, 34, 2568-2575.

Chu, H.H., Chen, T.L., Cheng, C.H. & Huang, C.C. (2009). Fuzzy dual factor time series for stock index forecasting. *Expert Systems with Applications*, 36, 165-171.

Huarng, K. (2001a). Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems*, 123, 369-386.

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

Huarng, K. & Yu, T.H.K. (2006). The application of neural networks to forecast fuzzy time series. *Physica A*, 363, 481-491.

Iliadis, L.S., Spartalis, S. & Tachos, S.I. (2007, August 29-31). An innovative artificial neural network evaluation model: Application in industry. Proceedings of the 10th International Conference on Engineering Applications of Neural Networks (EANN-2007), Thessaloniki, Hellas.

Kosko, S. (1992). *Neural networks and fuzzy systems*. Prentice Hall: New Jersey.

Lee, C.H.L., Liu, A. & Chen, W.S. (2006). Pattern discovery of fuzzy time series for financial prediction. IEEE Transactions on Knowledge and Data Engineering, 18, 613-625.

Li, S.H., Kuo, S.H., Cheng, Y.C. & Chen, C.C. (2009). Deterministic vector long-term forecasting for fuzzy time series. Fuzzy Sets and System, doi: 10.1016/j.fss.2009.10.028. In Press.

Nazirah Ramli & Abu Osman Md. Tap. (2000, April 1-2). Ramalan kabur kemasukan pelajar suatu alternatif. Paper presented at the Simposium Kebangsaan Sains Matematik Ke-8, Kuala Terengganu, Terengganu, Malaysia.