International Journal of Fuzzy Mathematics and Systems.
ISSN 2248-9940 Volume 2, Number 3 (2012), pp. 217-230

http://www.ripublication.com

Temperature Prediction Using Fuzzy Time Series and Multivariate Markov Chain

V. Vamitha[1], M. Jeyanthi[2], S. Rajaram[2] and T. Revathi[2]

[1]Department of Mathematics, B.C.M.W.G. Polytehnic,
Ettayapuram, Tamil Nadu, India
E-mail: vamithav @ ymail.com
[2]Department of Mathematics ; Sri S.R.N.M. College,
Sattur; Tamilnadu; India
E-mail: rajaram_srnm@yahoo.co.in

Abstract

Since 1993 researchers proposed many methods for forecasting enrollments, Temperature prediction, stock price etc in time variant and time invariant first order, higher order, two factor and dual variables. In this paper, we propose a model to temperature prediction from correlated categorical data sequence obtained from similar source. We study a multivariate Markov chain model for categorical data sequences to fuzzy time series. The proposed method gets higher average forecasting accuracy rate than some of the existing methods on temperature prediction.

Keywords: Multivariate Markov Chain; transition frequencies; categorical data sequence; fuzzy time series.

Introduction

An ordered sequence of observed values is known as time series. If the observed values represent measured values, it is often not possible to assign precise numerical values to the observed data, they then posses data uncertainty. This paper concerns with the time series comprised of imprecise i.e., uncertain observed values. In the case of time series the uncertainty of the individual observed values as well as the interpretation of a sequence of uncertain observed values are of interest. The uncertain observed value is thus modeled as a fuzzy variable. Fuzzy sets represent concepts such as low etc are called fuzzy variables. Modeling of the individual observed values as fuzzy variables result in fuzzy time series.

In this paper, we discuss a new fuzzy time series model based on multivariate Markov chain model on categorical data sequences. Categorical data sequences have many applications in applied sciences and engineering problems, data mining, credit risk problems in finance etc. The data used for all the results that are obtained come from a single source. On the other hand, in multivariate Markov chain model categorical data sequences are taken into consideration. In other words, data sequences that have a correlation with each other are used. As a result of this, predictions which are very close to the reality can be made. By making use of the transition probability matrix a categorical data sequence of m states can be modeled by an m-state Markov chain model. The above idea can be extended to model multiple categorical data sequences. One would expect categorical data sequences generated by similar sources or same source to be correlated to each other. Therefore by exploring these relationships, one can develop better models for the categorical data sequences and hence better prediction rules.

Box and Jenkins time series models (ARIMA models) have been applied for a long time to forecasting. However, ARIMA models have several limitations such as stationary, normality etc. In fuzzy time series there are no preconditions like stationary and normality. The traditional time series forecasting methods can not be used for forecasting problems in which the historical data are linguistic values. Song and Chissom (1993, 1994) proposed time variant and time invariant fuzzy time series models and fuzzy forecasting to model and forecast processes whose observation are linguistic values. Then a number related research work have been reported. These works include enrollments [2], [3], [4], [26], [33], [34], [35], [36], length of intervals [17], [21], temperature prediction[5], [24], [25], [30], weighted method [6], [42], stock price [7], [15], [16], [20], [23], [29], [37], [39], [40], hidden Markov model [27], genetic algorithm[8], [25], neural – fuzzy system [19], [28], bulk shipping [9], [10], [11], [12], [13], [14], seasonal[1], [32], [38], heuristic models [18], [22], [31].

Wai Ki Ching., Eric S., Fung and Michael K. Ng (2002) studied multivariate Markov chain models for analyzing categorical data sequences and proposed an efficient estimation method for the model parameters. They also developed higher order Markov chain models for analyzing categorical data sequences.

Fuzzy time series

In the following, we briefly review some basic concepts of fuzzy time series from Song and Chissom (1993, 1994) and its forecasting frame work.

Definition 1: A *fuzzy set* A is defined as an uncertain subset of the fundamental set X.

A = $\{(x, \mu_A(x)) | x \epsilon X\}$

The uncertainty is assessed by the membership function $\mu_A(x)$.

Definition 2: Let Y(t) {t = 0, 1, 2, 3, …}, a subset of $\mathbb{R}$, be the universe of discourse on which fuzzy sets $f_i(t)$ (i = 1, 2, 3…) are defined and let F(t) be the collection of $f_i(t)$. Then F (t) is defined as *fuzzy time series* on Y(t).

From this definition we can see that, (1) F(t) is the 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) fi(t) (i = 1, 2, 3…) are possible linguistic values of F(t), where fi (t) (i = 1, 2, 3, ….) are represented by fuzzy sets.

Song and Chissom employed a fuzzy relational equation to develop their forecasting model under the assumption that the observations at time t are dependent only upon the accumulated results of the observation at previous times, which is defined as follows.

Definition 3: Suppose F(t) is caused only by F(t-1) and is denoted by F(t-1) → F(t), then there is a fuzzy relationship between F(t) and F(t-1) and can be expressed as the fuzzy relational equation F(t) = F(t-1) ∘ R(t, t-1). Here '∘' is max- min composition operator. The relation R is called first – order model of F(t).

Further, if fuzzy relation R(t, t-1) of F(t) is independent of time t, that is to say, for different times t_1 and t_2, R(t_1, t_1-1) = R(t_2, t_2-1), then F(t) is called a *time invariant* fuzzy time series otherwise F(t) is *time variant.*

Definition 4: Suppose F(t-1) = A_i and F(t) = A_j a *fuzzy logical relationship* can be defined as $A_i \rightarrow A_j$ where A_i and A_j are called the left hand side and the right hand side of the fuzzy logical relationship respectively.

Definition 5: If F(t) is caused by more fuzzy sets F(t-n), F(t-n+1), …, F(t-1) the fuzzy relationship is represented by Ai1, Ai2, Ai3, … Ain → A_j, where F(t-n) = A_{i1}, F(t-n+1) = Ai2, …, F(t-1) = Ain. This relationship is called n^{th} order fuzzy time series model.

The Multivariate Markov Chain model

We briefly review some basic concepts of a multivariate Markov chain model to represent the behavior of multiple categorical data sequences generated by similar sources or the same source proposed by Ching et al., (2002). Consider Markov chains having finite number of states $\mathcal{M}$ = {1, 2, ..., m}. In general, a categorical data sequence $\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T$ can be logically represented by a sequence of vectors $x_1, x_2, ..., x_T$ where T is the length of the sequence, and $x_i = e_k$ (e_k is the unit vector with the k^{th} entry being one) if it is in state k. A first-order discrete-time Markov chain having m discrete states satisfies the following relationship:

$$\Pr (x_{t+1} = \boldsymbol{e}_{X_{t+1}} | x_0 = \boldsymbol{e}_{x_0}, x_1 = \boldsymbol{e}_{x_1}, ..., x_t = \boldsymbol{e}_{x_t}) = \Pr (x_{t+1} = \boldsymbol{e}_{x_{t+1}} | x_t = \boldsymbol{e}_{x_t})$$

where $\boldsymbol{x}_i \in \mathcal{M}$. The conditional probabilities $\Pr (x_{n+1} = \boldsymbol{e}_{x_{n+1}} | x_n = \boldsymbol{e}_{x_n})$ are called the single-step transition probabilities of the Markov chain. They give the conditional probability of making a transition from state i to state j when the time parameter increases from n to n+1. These probabilities are independent of n and are written as

$p_{ij} = \Pr(x_{n+1} = \boldsymbol{e_i} | x_n = \boldsymbol{e_j})$. $\forall$ i, j $\in \mathcal{M}$.

The matrix P, formed by placing p_{ij} in row i and column j for all i and j, is called the transition probability matrix. We note that the elements of the matrix P satisfy the following two properties: $0 \le p_{ij} \le 1$ $\forall$ i, j $\in \mathcal{M}$ and $\sum_{j=1}^{m} p_{ij} = 1$, $\forall$ j $\in \mathcal{M}$. We assume that p_{ij} are not all zero for each j.

Here we assume that there are s categorical sequences and each has m possible states in $\mathcal{M}$. Let $x_n^{(k)}$ be the state vector of the k^{th} sequence at time n. If the k^{th} sequence is in state j at time n then

$$x_n^{(k)} = e_j = (0,...,0, \underbrace{1}_{jthentry}, 0,...,0)^T.$$

In the multivariate Markov chain model, we assume the following relationship:

$$x_{n+1}^{(k)} = \sum_{k=1}^{s} \lambda_{jk} P^{(jk)} x_n^{(k)}, \text{ for } j = 1, 2, \ldots, s.,$$

where $\lambda_{jk} \ge 0$, $1 \le j, k \le s$ and $\sum_{k=1}^{s} \lambda_{jk} = 1$, for j = 1, 2, …, s.,

The state probability distribution of the k^{th} sequence at the $(n+1)^{th}$ step depends on the weighted average of $P^{(jk)} x_n^{(k)}$. Here $P^{(jk)}$ is a transition probability matrix from the states in the k^{th} sequence to the states in the j^{th} sequence, and $x_n^{(k)}$ is the state probability distribution of the k^{th} sequences at the n^{th} step. In matrix form we write

$$\mathbf{x}_{n+1} \equiv \begin{pmatrix} \mathbf{x}_{n+1}^{(1)} \\ \mathbf{x}_{n+1}^{(2)} \\ \vdots \\ \mathbf{x}_{n+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \cdots & \lambda_{1s}P^{(1s)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \cdots & \lambda_{2s}P^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s1}P^{(s1)} & \lambda_{s2}P^{(s2)} & \cdots & \lambda_{ss}P^{(ss)} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{n}^{(1)} \\ \mathbf{x}_{n}^{(2)} \\ \vdots \\ \mathbf{x}_{n}^{(s)} \end{pmatrix} \equiv Q\mathbf{x}_n \text{ or } x_{n+1} = Qx_n$$

Proposed Model

In this section we introduce a model to forecast the Temperature of Taipei. The historical data of daily average Temperature of Taipei from June 1996 to September 1996 are considered. The step-wise procedure of the proposed model of fuzzy time series is detailed as follows.

Step 1:

Define the universe of discourse U = [low, up], which can cover all observations of Temperature in the months of June 1996 to September 1996 of historical data set. Initially partition the universe of discourse into seven linguistic intervals u_i, i = 1, 2, …, 7. of equal length.

Step2:

Temperatures are categorized into seven possible states. Define fuzzy sets A_1 (very very low), A_2 (very low), A_3 (low), A_4 (normal), A_5 (high), A_6 (very high), A_7 (very very high). Construct the fuzzy sets A_i in accordance with the intervals in step 1. Fuzzify the historical data. For n fuzzy sets, A_1, A_2, ……A_n can be defined on U as follows:

$A_i = \sum_{j=1}^{n} \frac{\mu_{ij}}{v_j}$ where μ_{ij} is the membership degree of A_i belonging to v_j and is

defined by $\mu_{ij} = \begin{cases} 1 & if \quad j = i \\ 0.5 & if \quad j = i-1 \text{ or } i+1 \\ 0 & if \quad otherwise \end{cases}$

Then, for a given historical datum Y_t, its membership degree belonging to interval v_i is determined by the following heuristic rules.

Rule 1: if Y_t is located at v_1, the membership degrees are 1 for v_1, 0.5 for v_2 and 0 otherwise.

Rule 2: if Y_t belongs to v_i, $1 < i < n$, then the degrees are 1, 0.5 and 0.5 for v_i, v_{i-1} and v_{i+1}, respectively and 0 otherwise.

Rule 3: if Y_t is located at v_n, the membership degrees are 1 for v_n, 0.5 for v_{n-1} and 0 otherwise. Then, Y_t is fuzzified as A_j, where the membership degree in interval j is maximal.

Step 3:

Represent the subscripts of the fuzzy sets obtained from the four months of temperature data as the members of the categorical data sequences S_1, S_2, S_3, S_4.

Step 4:

Given the data sequence we count the transition frequency $f^{(jk)}_{i_j i_k}$ from the state i_k in the sequence $\{\mathbf{x}^{(k)}_n\}$ to the state i_j in the sequence $\{\mathbf{x}^{(j)}_n\}$ and therefore we construct the transition frequency matrix for the sequences as follows:

$$F^{(jk)} = \begin{pmatrix} f^{(jk)}_{11} & \cdots & \cdots & f^{(jk)}_{m1} \\ f^{(jk)}_{12} & \cdots & \cdots & f^{(jk)}_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ f^{(jk)}_{1m} & \cdots & \cdots & f^{(jk)}_{mm} \end{pmatrix}.$$

Step 5:

After the normalization, the estimates of the transition probability matrices from $F^{(jk)}$can also be obtained as follows [41]:

$$\hat{P}^{(jk)} = \begin{pmatrix} \hat{p}_{11}^{(jk)} & \cdots & \cdots & \hat{p}_{m1}^{(jk)} \\ \hat{p}_{12}^{(jk)} & \cdots & \cdots & \hat{p}_{m2}^{(jk)} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{p}_{1m}^{(jk)} & \cdots & \cdots & \hat{p}_{mm}^{(jk)} \end{pmatrix}$$

where $$\hat{p}_{i_j i_k}^{(jk)} = \begin{cases} \dfrac{f_{i_j i_k}^{(jk)}}{\sum_{i_k=1}^{m} f_{i_j i_k}^{(jk)}} & \text{if } \sum_{i_k=1}^{m} f_{i_j i_k}^{(jk)} \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Step 6:
We need to estimate the parameters λ_{jk}. The stationary vector $\hat{\mathbf{x}}$ can be estimated from the sequences by computing the proportion of the occurrence of each state in each of the sequences, and let us denote it by $\hat{\mathbf{x}} = [\,\hat{\mathbf{x}}^1, \hat{\mathbf{x}}^{(2)}, \ldots, \hat{\mathbf{x}}^{(s)}]^T$. One would expect

$$\begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \cdots & \lambda_{1s}P^{(1s)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \cdots & \lambda_{2s}P^{(2s)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{s1}P^{(s1)} & \lambda_{s2}P^{(s2)} & \cdots & \lambda_{ss}P^{(ss)} \end{pmatrix} \hat{\mathbf{x}} \approx \hat{\mathbf{x}}$$

From the above equation, it suggests one possible way to estimate the parameters $\lambda = \{\lambda_{jk}\}$ as follows. One may consider solving the following optimization problem [41]:

$$\begin{cases} \min_{\lambda} \max_{i} \left| \left[\sum_{k=1}^{s} \lambda_{jk} \hat{p}^{(jk)} \hat{x}^{(k)} - \hat{x}^{(k)} \right]_i \right| \\ \text{subject to} \sum_{k=1}^{s} \lambda_{jk} = 1 \text{ and } \lambda_{jk} \geq 0, \forall k. \end{cases}$$

Step 7:
Formulate s linear programming problems from the above optimization problem as follows:

For each j: $$\begin{cases} \min_{\lambda} w_j \\ \text{subject to} \begin{pmatrix} w_j \\ w_j \\ \vdots \\ w_j \end{pmatrix} \geq \hat{X}^{(j)} - B \begin{pmatrix} \lambda_{j1} \\ \lambda_{j2} \\ \vdots \\ \lambda_{js} \end{pmatrix}, \begin{pmatrix} w_j \\ w_j \\ \vdots \\ w_j \end{pmatrix} \geq -\hat{X}^{(j)} + B \begin{pmatrix} \lambda_{j1} \\ \lambda_{j2} \\ \vdots \\ \lambda_{js} \end{pmatrix}, \\ w_j \geq 0, \\ \sum_{k=1}^{s} \lambda_{jk} = 1, \lambda_{jk} \geq 0, \forall k. \end{cases}$$

where B = [$\widehat{P}^{(j1)}\,\hat{\mathbf{x}}^{(1)}$ | $\widehat{P}^{(j2)}\,\hat{\mathbf{x}}^{(2)}$ |...| $\widehat{P}^{(js)}\,\hat{\mathbf{x}}^{(s)}$].

Step 8: Construct the models using the values obtained in steps 6 and 7 and forecast the vector to the historical data. Compare with the actual vector by taking minimum among the vector elements. Calculate the forecasted value as follows:

Forecasted value = $\frac{\sum_{i=1}^{7} v_i\, m_i}{\sum_{i=1}^{7} v_i}$ where v_i are entries of the resulting vector and m_i are midpoints of the corresponding u_i.

Step 9: Choose α in (0, 1). Make an error analysis for every month as follows:

New forecasted value = $\frac{Actual\ value + forecasted\ value - \alpha}{2}$

Compute Root Mean Square Error (RMSE) for different α values on new forecasted values. Fix α corresponding to minimum RMSE value through the graph. The forecasted values with respect to this α are the expected forecasted values.

Step 10: Compare this model with some of the existing models.
In the next section, we give an example to demonstrate the construction of a multivariate Markov model using fuzzy time series from four categorical data sequences.

Performance evaluation of the model:

The four categorical data sequences obtained from the temperature data as follows:

S_1 = {3, 4, 5, 7, 6, 6, 6, 6, 5, 6, 6, 5, 5, 4, 6, 5, 5, 6, 6, 7, 7, 5, 4, 4, 4, 4, 5, 4, 5, 6}

S_2 = {6, 5, 5, 6, 6, 6, 6, 6, 5, 5, 5, 6, 4, 5, 4, 5, 5, 6, 7, 7, 7, 7, 7, 7, 5, 4, 5, 4, 6, 4}

S_3 = {4, 5, 5, 6, 5, 5, 5, 5, 4, 5, 5, 5, 6, 4, 3, 4, 4, 5, 5, 6, 6, 6, 4, 5, 5, 5, 5, 4, 3, 3} and

S_4 = {4, 3, 3, 4, 3, 5, 5, 5, 6, 6, 6, 7, 6, 6, 6, 5, 5, 5, 5, 5, 3, 3, 2, 4, 3, 3, 2, 1, 1, 1}

By counting the intra transition frequencies

S_1:3⟶4⟶5⟶7⟶6⟶6⟶6⟶6⟶5⟶6⟶6⟶5⟶5⟶4⟶6⟶5⟶5⟶6
⟶6⟶7⟶7⟶5⟶4⟶4⟶4⟶4⟶5⟶4⟶5⟶6

S_2:6⟶5⟶5⟶6⟶6⟶6⟶6⟶6⟶5⟶5⟶5⟶6⟶4⟶5⟶4⟶5⟶5⟶6
⟶7⟶7⟶7⟶7⟶7⟶7⟶5⟶4⟶5⟶4⟶6⟶4

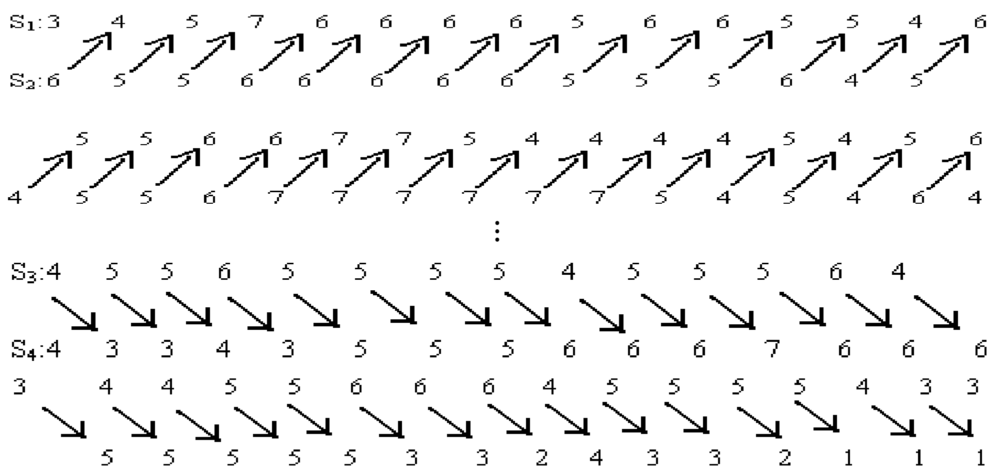
S_3:4⟶5⟶5⟶6⟶5⟶5⟶5⟶5⟶4⟶5⟶5⟶5⟶6⟶4⟶3⟶4⟶4⟶5
⟶5⟶6⟶6⟶6⟶4⟶5⟶5⟶5⟶5⟶4⟶3⟶3

S_4:4⟶3⟶3⟶4⟶3⟶5⟶5⟶5⟶6⟶6⟶6⟶7⟶6⟶6⟶6⟶5⟶5⟶5
⟶5⟶5⟶3⟶3⟶2⟶4⟶3⟶3⟶2⟶1⟶1⟶1

$$F^{(11)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 & 3 & 1 \\ 0 & 0 & 0 & 1 & 3 & 5 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \cdots \quad F^{(44)} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 6 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

After normalization we have the transition probability matrices

$$P^{(11)} = \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & 1 & \frac{3}{7} & \frac{1}{3} & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & 0 & \frac{3}{7} & \frac{2}{9} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{7} & \frac{1}{7} & 0 & \frac{1}{7} & \frac{1}{3} & \frac{5}{9} & \frac{1}{3} \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} \end{pmatrix} \cdots \; P^{(44)} = \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{7} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{7} & 1 & \frac{1}{8} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{7} & 0 & \frac{3}{4} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{2}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} & 0 \end{pmatrix}$$



Moreover, by counting the inter-transition frequencies we have

$$F^{(12)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 3 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 4 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 \end{pmatrix} \cdots \quad F^{(43)} = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 3 & 3 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

After normalization we have the transition probability matrices:

$$P^{(12)} = \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{4} & \frac{1}{5} & \frac{1}{9} & \frac{1}{2} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{3}{4} & \frac{3}{10} & \frac{2}{9} & \frac{1}{6} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 & \frac{2}{5} & \frac{2}{3} & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & 0 & \frac{1}{10} & 0 & \frac{1}{3} \end{pmatrix} \ldots \quad P^{(43)} = \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & \frac{1}{2} & \frac{1}{7} & \frac{1}{15} & 0 & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & \frac{1}{15} & \frac{1}{5} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & 0 & \frac{1}{7} & \frac{1}{5} & \frac{3}{5} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & 0 & \frac{1}{7} & \frac{1}{15} & 0 & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{2} & \frac{2}{7} & \frac{1}{3} & 0 & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & 0 & \frac{2}{7} & \frac{1}{5} & \frac{1}{5} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & 0 & 0 & \frac{1}{15} & 0 & \frac{1}{7} \end{pmatrix}$$

By solving the corresponding minimization problems, through linear programming

$$\omega_4 \ \geq \ \frac{1}{10} - \frac{1}{10}\lambda_{41} - \frac{13}{120}\lambda_{42} - \frac{13}{60}\lambda_{43} - \frac{2}{15}\lambda_{44}$$

$$\vdots$$

$$\omega_4 \ \geq \ \frac{1}{30} - \frac{1}{27}\lambda_{41} - \frac{1}{30}\lambda_{42} - \frac{1}{30}\lambda_{43} - \frac{1}{30}\lambda_{44}$$

and

$$\omega_4 \ \geq \ -\frac{1}{10} + \frac{1}{10}\lambda_{41} + \frac{13}{120}\lambda_{42} + \frac{13}{60}\lambda_{43} + \frac{2}{15}\lambda_{44}$$

$$\vdots$$

$$\omega_4 \ \geq \ -\frac{1}{30} + \frac{1}{27}\lambda_{41} + \frac{1}{30}\lambda_{42} + \frac{1}{30}\lambda_{43} + \frac{1}{30}\lambda_{44}$$

we obtain the optimal solution by step 7

$$\Lambda = [\lambda_{jk}] = \begin{pmatrix} 0.0000 & 0.2683 & 0.4541 & 0.2777 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.5112 & 0.4848 & 0.0000 \\ 0.4529 & 0.1824 & 0.1861 & 0.1786 \end{pmatrix}$$

and by step 8, the multivariate Markov model for four categorical data sequences is as follows:

$$\begin{cases} \mathbf{x}_{n+1}^{(1)} = 0.0000\,\hat{P}^{(11)}\mathbf{x}_n^{(1)} + 0.2683\,\hat{P}^{(12)}\mathbf{x}_n^{(2)} + 0.4541\hat{P}^{(13)}\mathbf{x}_n^{(3)} + 0.2777\hat{P}^{(14)}\mathbf{x}_n^{(4)} \\ \mathbf{x}_{n+1}^{(2)} = 0.0000\,\hat{P}^{(21)}\mathbf{x}_n^{(1)} + 0.0000\,\hat{P}^{(22)}\mathbf{x}_n^{(2)} + 0.0000\,\hat{P}^{(23)}\mathbf{x}_n^{(3)} + 1.0000\,\hat{P}^{(24)}\mathbf{x}_n^{(4)} \\ \mathbf{x}_{n+1}^{(3)} = 0.0000\,\hat{P}^{(31)}\mathbf{x}_n^{(1)} + 0.5112\,\hat{P}^{(32)}\mathbf{x}_n^{(2)} + 0.4888\hat{P}^{(33)}\mathbf{x}_n^{(3)} + 0.0000\,\hat{P}^{(14)}\mathbf{x}_n^{(4)} \\ \mathbf{x}_{n+1}^{(4)} = 0.4529\,\hat{P}^{(41)}\mathbf{x}_n^{(1)} + 0.1824\,\hat{P}^{(42)}\mathbf{x}_n^{(2)} + 0.1861\hat{P}^{(43)}\mathbf{x}_n^{(3)} + 0.1786\hat{P}^{(44)}\mathbf{x}_n^{(4)} \end{cases}$$

By step 8 as an example when n = 1, the forecasting vector of the second position in the fourth sequence is $\mathbf{x}_2^{(4)} = (0.05 \quad 0 \quad 0.70 \quad 0.03 \quad 0.13 \quad 0.09 \quad 0)^T$ and the actual fuzzy set of the second position in the fourth sequence is A3 (i.e in the second day of the September temperature). The corresponding vector is $(0 \quad 0.5 \quad 1 \quad 0.5 \quad 0 \quad 0 \quad 0)^T$. Compare the actual and forecasting vectors by taking minimum, we get $(0 \quad 0 \quad 0.70 \quad 0.03 \quad 0 \quad 0 \quad 0)^T$.

The forecasting value corresponding to second position in the fourth sequence by step 9 is $\frac{(0\times m1)+(0\times m2)+(0.70\times m3)+(0.03\times m4)+(0\times m5)+(0\times m6)+(0\times m7)}{0.70+0.03} = 26.3.$

By step 9, we find $\alpha = 0.06$ and the new forecasted value = 26.52.

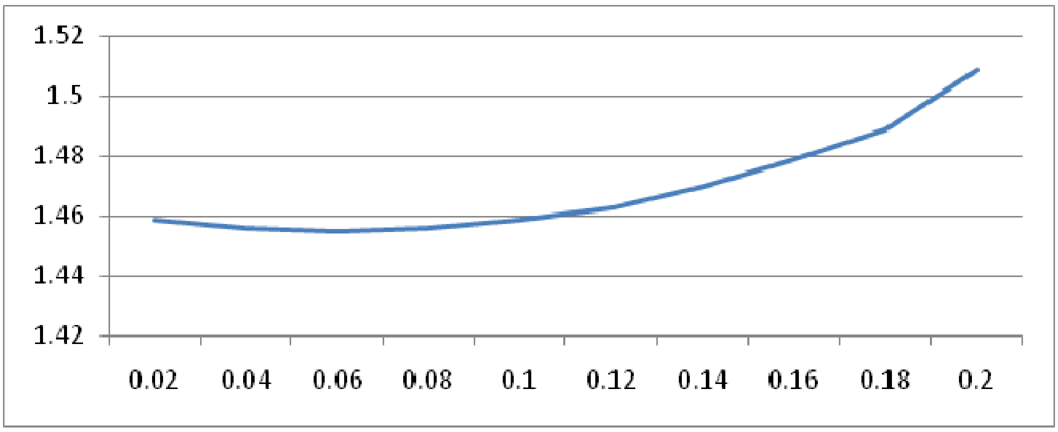


Figure 1: RMSE results for different α values to September 1996

Table 1: Actual and forecasted values of temperature

Day	June		July		August		September	
	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast
1	26.1	-	29.9	-	27.1	-	27.5	-
2	27.6	27.74	28.4	28.74	28.9	28.75	26.8	26.52
3	29.0	28.89	29.2	29.04	28.9	28.75	26.4	26.32
4	30.5	30.34	29.4	29.64	29.3	29.10	27.5	27.52
5	30.0	29.84	29.9	29.49	28.8	28.75	26.6	26.37
6	29.5	29.49	29.6	29.74	28.7	28.75	28.2	28.52
7	29.7	29.64	30.1	30.04	29.0	28.85	29.2	29.02
8	29.4	29.49	29.3	29.64	28.2	28.45	29.0	28.92
9	28.8	28.69	28.1	28.74	27.0	27.60	30.3	29.72
10	29.4	29.39	28.9	28.79	28.3	28.40	29.9	29.67
11	29.3	29.39	28.4	28.54	28.9	28.75	29.9	29.62
12	28.5	28.69	29.6	29.29	28.1	28.35	30.5	30.32
13	28.7	28.84	27.8	27.64	29.9	29.40	30.2	29.77
14	27.5	27.74	29.1	28.89	27.6	27.60	30.3	29.87
15	29.5	29.44	27.7	28.04	26.8	26.80	29.5	29.62
16	28.8	28.89	28.1	28.39	27.6	27.40	28.3	28.67
17	29.0	28.99	28.7	29.04	27.9	27.90	28.6	28.77
18	30.3	29.94	29.9	29.94	29.0	28.75	28.1	28.52
19	30.2	29.89	30.8	30.59	29.2	28.95	28.4	28.62
20	30.9	30.64	31.6	30.99	29.8	29.50	28.3	28.57
21	30.8	30.64	31.4	30.89	29.6	29.50	26.4	26.27
22	28.7	28.49	31.3	30.84	29.3	29.35	25.7	26.07
23	27.8	27.84	31.3	30.84	28.0	28.05	25.0	25.22
24	27.4	27.59	31.3	31.14	28.3	28.55	27.0	26.97
25	27.7	27.74	28.9	28.99	28.6	28.75	25.8	26.02
26	27.1	27.64	28.0	28.14	28.7	28.65	26.4	26.32
27	28.4	28.44	28.6	28.74	29.0	28.65	25.6	25.52
28	27.8	27.84	28.0	27.74	27.7	27.95	24.2	24.17
29	29.0	28.89	29.3	29.59	26.2	26.55	23.3	23.72
30	30.2	29.74	27.9	27.69	26.0	26.45	23.5	23.72
31	-	-	26.9	-	27.7	-	-	-
RMSE		1.1489		2.6575		1.7475		2.1171

Table 2: (AFER) (In percentage)

Month	Lee et.al(2006)	Lee et.al(2007)	Proposed Method
June	1.44	1.24	0.54
July	1.59	1.23	0.87
Aug	1.26	1.09	0.69
Sept	1.89	1.28	0.83

Conclusion

In this chapter, we have proposed a new forecasting model based on multivariate Markov chain on categorical sequences for forecasting the daily average temperature of the Taipei, Taiwan. Through the fuzzification of the temperature data we obtained four categorical data sequences and on which multivariate Markov chain on categorical sequences is applied. From the experimental results the proposed method provides the smallest AFER (see Table 2) and improves on other methods using fuzzy times series forecasting methods. We may obtain further accuracy by applying higher order multivariate Markov chain model on categorical data sequences.

References

[1] Chang, P.T. (1997). Fuzzy seasonality forecasting. Fuzzy Sets and Systems, 90, l-10.

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

[3] Chen, S.M. (2002). Forecasting enrollments based on high-order fuzzy time series. Cybern. Syst., Int. J., 33(1), 1–16.

[4] Chen, S.M. and Hsu, C.C. (2004). A new method to forecast enrollments using fuzzy time series. Int. J. Appl. Sci. Eng., 2(3), 234–244.

[5] Chen, S.M. and Hwang, J.R. (2000). Temperature prediction using fuzzy time series. IEEE Trans. Syst., Man, Cybern. B, Cybern., 30(2), 263–275.

[6] Cheng, C. H., Chen, T. L., & Chiang, C. H. (2006). Trend-weighted fuzzy time-series model for TAIEX forecasting. Lecture Notes in Computer Science, 4234, 469–477.

[7] 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 (1), 165–171.

[8] Day, M. Y. (1995). Research of Applying Genetic Algorithms to Fuzzy Forecasting- Focus on Sales Forecasting. M.S. thesis, Tamkang Univ., Taipei, Taiwan, R.O.C.

[9] Duru, O., & Yoshida, S. (2008a). Composite forecast: A new approach for forecasting shipping markets. In Proceedings for the international association of maritime economists conference, Dalian, China, April 2– 4.

[10] Duru, O., & Yoshida, S. (2008b). Market psychology. Lloyd's Shipping Economist, 30, 30–31.

[11] Duru, O., & Yoshida, S. (2009). Composite forecasting of dry bulk shipping index: Judgments and statistics. KAIUN Japanese Shipping Journal, 976, 107–109.

[12] Duru, O., (2010).A fuzzy integrated logical forecasting model for dry bulk shipping index forecasting:An improved fuzzy time series approach. Expert Syst. Appl., 37, 5372-5380.

[13] Duru, O., (2012).A multivariate model of fuzzy integrated logical forecasting method (M-FILF) and multiplicative time series clustering: A model of time-varying volatility for dry cargo freight market. Expert Syst. Appl., 39, 4135-4142.

[14] Hampton, M. J. (1991). Long and short shipping cycles. Cambridge: Cambridge Academy of Transport.

[15] Hodgson, A., Masih, A. M. M., & Masih, R. (2006). Futures trading volume as a determinant of prices in different momentum phases. International Review of Financial Analysis, 15, 68–85.

[16] Hsu, Y.Y., Tse, S.M. and Wu, B. (2003). A new approach of bivariate fuzzy time series analysis to the forecasting of a stock index. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., 11(6), 671–690.

[17] Huarng, K. (2001a). Effective lengths of intervals to improve forecasting in fuzzy time-series. Fuzzy Sets and System, 123 (3), 387–394.

[18] Huarng, K. (2001b). Heuristic models of fuzzy time-series for forecasting. Fuzzy Sets and System, 123, 369–386.

[19] Huarng, K., & Yu, H. K. (2006). The application of neural networks to forecast fuzzy time series. Physica A, 363, 481–491.

[20] Huarng, K., & Yu, H. K. (2005). A type 2 fuzzy time-series model for stock index forecasting. Physica A, 353, 445–462.

[21] Huarng, T.H.K. & Yu, T.H.K., (2006). Ratio-Based Lengths of Intervals to Improve Fuzzy Time Series Forecasting, IEEE transactions on systems, man, and cybernetics-part b: cybernetics, 36(2), 328-340.

[22] Huarng, T.H.K. Yu, T.H.K. and Hsu, Y.W (2007). A multivariate heuristic model for fuzzy time-series forecasting. IEEE Trans. Syst., Man, Cybern. B, Cybern., 37(4), 836–846.

[23] Kim, M. J., Min, S. H., & Han, I. G. (2006). An evolutionary approach to the combination of multiple classifiers to predict a stock price index. Expert Systems with Applications, 31, 241–247.

[24] Lee, L.W. Wang, L.H. Chen, S.M. and Leu, Y.H. (2006). Handling forecasting problems based on two-factor high-order fuzzy time series. IEEE Trans. Fuzzy Syst., 14(3), 468–477.

[25] Lee, L.W. Wang, L.H. and Chen, S.M. (2007). Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. Expert Syst. Appl., 33(3), 539–550.

[26] Li, S.T. and Cheng, Y.C. (2007). Deterministic fuzzy time series model for forecasting enrollments. Comput. Math. Appl., 53(12), 1904–1920.

[27] Li, Y.(2005). Hidden Markov models with states depending on observations. Pattern Recognit. Lett., 26(7), 977–984.

[28] Li Bo, Zhang shiying and Wang Xiuf Eng, (2003). Research on fuzzy modeling forecasting method based on stochastic neural network. Taylor and Francis, Systems Analysis Modelling Simulation, 43(12), 1709–1715.

[29] Marc, B., & Hiraki, T. (1999). Volume and individual security returns on the Tokyo stock exchange. Pacific-Basin Finance Journal, 7, 351–370.

[30] Nai - Yi Wang, Shyi - Ming Chen. (2009). Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two – factors high - order Fuzzy time series. Expert systems with Applications, 36, 2143 - 2154.

[31] Own, C.M. and Yu, P.T. (2005). Forecasting fuzzy time series on a heuristic high order model. Cybernetics and systems: An International Journal, 36: 705-717.

[32] Phillip G. Gould, Anne B. Koehler, J. Keith Ord, Ralph D. Snyder, Rob J. Hyndman, Farshid Vahid-Araghi, (2008). Stochastics and Statistics, Forecasting time series with multiple seasonal patterns. European Journal of Operational Research 191, 207–222.

[33] Rajaram, S. & Sakthisree, P. (2010). Forecasting enrollments using interval length, weightage factor and fuzzy logical relationships. International journal of statistics and systems, 5(2).173-182.

[34] Song, Q., & Chissom, B. S. (1993a). Fuzzy forecasting enrollments with fuzzy timeseries – Part 1. Fuzzy Sets and Systems, 54, 1–9.

[35] Song, Q., & Chissom, B. S. (1993b). Fuzzy time series and its models. Fuzzy Sets and Systems, 54, 269–277.

[36] Song, Q., & Chissom, B. S. (1994). Fuzzy forecasting enrollments with fuzzy time series – Part 2. Fuzzy Sets and Systems, 62, 1–8.

[37] Tahseen Ahmed Jilani, & Syed Muhammad Aqil Burney, (2008). A refined fuzzy time series model for stock market forecasting. Physica A, 387, 2857–2862.

[38] Tasha R. Inniss, (2006). Seasonal clustering technique for time series data, Stochastics and Statistics. European Journal of Operational Research, 175, 376–384.

[39] Ting, J. J. L. (2003). Causalities of the Taiwan stock market. Physica A, 324, 285–295.

[40] Wang, C. Y., & Chin, S. T. (2004). Profitability of return and volume-based investment strategies in China's stock market. Pacific-Basin Finance Journal, 12, 541–564.

[41] Wai Ki Ching., Eric S., Fung and Michael K. Ng (2002). A Multivariate Markov Chain Model for Categorical Data Sequences and its Applications in Demand Predictions. IMA Journal of Management Mathematics, 13, 187-199.

[42] Yu, H. K. (2005). Weighted fuzzy time series models for TAIEX forecasting. Physica A, 349, 609–624.