ICIC Express Letters ICIC International ©2022 ISSN 1881-803X
Volume **16**, Number **3**, March **2022** pp. **327–334**

AUTOCORRELATED TIME SERIES ANALYSIS USING CLASSED TIMES SERIES AND ITS APPLICATION TO JAPANESE CONSUMER PRICE INDEX

Yoshiyuki Yabuuchi

Faculty of Economics
Shimonoseki City University
2-1-1 Daigaku-cho, Shimonoseki, Yamaguchi 751-8510, Japan
yabuuchi@shimonoseki-cu.ac.jp

Received July 2021; accepted September 2021

Abstract. *Several fuzzy time-series models have been proposed based on the Box-Jenkins approach. An autocorrelated fuzzy time-series model has been proposed in this paper, based on autoregressive (AR) modeling, to illustrate the possibilities of time-series systems while focusing on the ease of model use. Time-series models are multivariate analysis models and as such require statistical knowledge. Therefore, the proposed model was divided into classes, and the prediction was performed using the class values. The values predicted by the proposed method were more accurate than the values predicted by the AR model. In other words, the predicted values of the proposed method came closer to the check data than the AR method. This study applies the proposed method to the Japanese consumer price index to demonstrate its functionality.*
Keywords: Autoregressive model, Class interval, Class values, Consumer price index

1. **Introduction.** Various fuzzy time series models [2,6,7] have been proposed to address the ambiguity of time-series systems. Fuzzy time series models are primarily based on the rule-based model [3,5] or the Box-Jenkins model [8,10,11]. The former improves prediction accuracy when combined with clustering [1,4,9,12,13]. There are few rule-based models that can determine interval prediction values, compared to non-interval time-series models. The model proposed by Zhang and Zhu [13] uses the k-means method to adjust the clusters into intervals, to predict the interval output. The autocorrelated fuzzy time-series model (AFTSM) [11] is one of the latter Box-Jenkins models used to determine an interval output. The AFTSM is an AR model based on interval time series, such that the center value of the predicted value is the same as that of the AR model. The AFTSM thus determines the interval output from the input.

The prediction accuracy in regression analysis may be improved by creating a frequency distribution table and by increasing the class width. The ease of handling the model and prediction accuracy of the AFTSM are improved by adopting a time series of class values. In this paper, an autocorrelated fuzzy time-series model based on classified series (AFTSMC) is proposed. As a numerical example, the Japanese consumer price index (CPI) has been analyzed using the proposed model. Because the proposed model is a fuzzy AR model, it classifies the time series into classes. Thus, it is easier to handle than rule-based models with combined clustering. In the numerical example, the prediction accuracy of the proposed model is affirmed.

The remainder of this paper is organized as follows. The AFTSM, which is the original AFTSMC, is briefly explained in Section 2. Section 3 describes the proposed AFTSMC method. In Section 4, AFTSMC is used to analyze the Japanese consumer price index.

DOI: 10.24507/icicel.16.03.327

We then confirm the characteristics of the proposed model. Finally, Section 5 summarizes the study.

2. **Autocorrelated Fuzzy Time Series Model.** The AFTSM proposed by the author [11] is a fuzzy AR model with a fuzzy coefficient number. In addition, the AFTSM adopts interval-type time-series data.

The center of the values predicted by the AFTSM coincides with that of the AR model. Therefore, this study uses a triangular fuzzy number as an interval-type time series. Fuzzy time-series data $\boldsymbol{x}_t = \left[x_t^L, x_t^C, x_t^U\right]$ at time t are described by the center x_t^C and the upper and lower limits x_t^U and x_t^L, respectively. Hereafter, the center, lower, and upper limits of triangular fuzzy numbers are indicated by C, L, and U, respectively.

If the time series y_t are real numbers, they are converted to fuzzy numbers, as follows:

$$\left\{ \begin{array}{l} x_t^U = \max\{y_{t-1}, y_t, y_{t+1}\}, \\ x_t^C = y_t, \\ x_t^L = \min\{y_{t-1}, y_t, y_{t+1}\}. \end{array} \right. \tag{1}$$

Here, although the time series y_t are fuzzified using the three times $t-1$, t, and $t+1$, the conditions dictate the fuzzification.

AFTSM using a fuzzy autoregressive coefficient $\boldsymbol{a}_j$ $(j = 1, 2, \dots, p)$ can be formulated as

$$\left. \begin{array}{l} \boldsymbol{x}_t = \boldsymbol{a}_1 \boldsymbol{x}_{t-1} + \boldsymbol{a}_2 \boldsymbol{x}_{t-2} + \cdots + \boldsymbol{a}_p \boldsymbol{x}_{t-p}, \\ \boldsymbol{a}_j = \left[a_j^L, a_j^C, a_j^U\right], \quad j = 1, 2, \dots, p. \end{array} \right\} \tag{2}$$

The AFTSM is a fuzzified AR model, and its coefficients are determined by autocorrelation. The autocovariance $\boldsymbol{v}_k$, and autocorrelation coefficient $\boldsymbol{r}_k$ of lag k were obtained using the fuzzy time series $\boldsymbol{x}_t$. Therefore, fuzzy autocovariance $\boldsymbol{v}_k$, and fuzzy autocorrelation $\boldsymbol{r}_k$ are also interval-type fuzzy numbers, $\boldsymbol{v}_k = \left[v_k^L, v_k^C, v_k^U\right]$ and $\boldsymbol{r}_k = \left[r_k^L, r_k^C, r_k^U\right]$, respectively. The fuzzy autocovariance $\boldsymbol{v}_k$, and fuzzy autocorrelation coefficient $\boldsymbol{r}_k$ of lag k are given by

$$\boldsymbol{v}_k = Cov[\boldsymbol{x}_t, \boldsymbol{x}_{t-k}], \quad \boldsymbol{r}_k = \frac{\boldsymbol{v}_k}{\boldsymbol{v}_0}, \quad \boldsymbol{r}_0 := [1, 1, 1]. \tag{3}$$

The fuzzy autocorrelation coefficient of lag 0 is defined as $\boldsymbol{r}_0 = [1, 1, 1] = 1$.

Generally, an autoregressive coefficient is obtained using the Yule-Walker equation; however, both the autocorrelation and autoregressive coefficients of AFTSM are fuzzy numbers. Therefore, the Yule-Walker equation does not apply, but it can be extended as follows:

$$\begin{bmatrix} \boldsymbol{r}_0 & \boldsymbol{r}_1 & \cdots & \boldsymbol{r}_{p-1} \\ \boldsymbol{r}_1 & \boldsymbol{r}_0 & \cdots & \boldsymbol{r}_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{r}_{p-1} & \boldsymbol{r}_{p-2} & \cdots & \boldsymbol{r}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_p \end{bmatrix} \supseteq \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_p \end{bmatrix}, \tag{4}$$

where $\boldsymbol{r}$ and $\boldsymbol{a}$ are fuzzy numbers. Therefore, the fuzzy Yule-Walker equation, as indicated in Equation (4), is not an equation. In addition, a fuzzy autocorrelation coefficient has a constraint $\boldsymbol{r} \subseteq [-1, 1]$. Therefore, a fuzzy autoregressive coefficient is obtained using the following procedure.

Step 1: Adjustment of fuzzy autocorrelation coefficient

Fuzzy autocorrelation coefficients are adjusted by α-cut, thereby becoming $\boldsymbol{r} \subseteq [-1, 1]$.

Step 2: Calculation of fuzzy autoregressive coefficient

The widths of the values predicted by the interval model are vague. Equation (4), using the fuzzy autoregressive coefficients $\boldsymbol{a}$, minimizes the sum of vagueness, $\sum_{j=1}^{p} \left(\rho_j^U - \rho_j^L\right)$, of AFTSM, which is rewritten in linear programming as follows:

$$\left.\begin{array}{c}\min\limits_{\boldsymbol{a}} \sum\limits_{j=1}^{p} \left(\rho_j^U - \rho_j^L\right) \\ \text{s.t. } \boldsymbol{Ra} \supseteq \boldsymbol{r},\ \rho_j^C = r_j^C,\ j = 1, 2, \ldots, p.\end{array}\right\} \tag{5}$$

It is evident from Equation (5) that the center of the predicted value of the AFTSM coincides with the AR model.

In the above procedure, the coefficients of the AR model interval output (AFTSM) are obtained by fuzzification. However, the values predicted by the AFTSM can vary significantly compared with the measured values [11]. In addition, when the equation is verified without using the original series, its forecasting accuracy occasionally decreases. This may be due to inadequate preprocessing in model building or in determining the characteristics of the model. In addition, because AFTSM uses fuzzy numbers, its fuzzy autoregressive coefficients may be affected by the vagueness of the adopted time series and by the fuzzy operations. Therefore, the predicted widths are likely to be large. To address these limitations, this study proposes the utilization of classes. When a time series is divided into classes, it is interval-valued, and there is no need for fuzzy operations. The prediction accuracy can also be improved by appropriately setting the class width. Accordingly, AFTSM adopts an interval-type time series. To improve prediction accuracy and ease of use, an AFTSM with a class time series (AFTSMC) is proposed.

3. **Class Time Series.** The classes used in this study are presented in Table 1. Initially, set a class value e_i, and a class width $2w$, such that time series $\boldsymbol{x}_t$ can be properly classified into class I_i. A class value e_i is assigned to a central value z_t^C of a time series $\boldsymbol{z}_t = \left(z_t^C, w\right)$, classified as class I_i. The time series adopted in AFTSMC are real-valued. Therefore, the autoregressive coefficients of the AFTSMC have real values. However, because the time series to be predicted, $\boldsymbol{z}_t$, is a fuzzy number, an interval-type fuzzy number and the width w of an interval are applied to the predicted values.

Table 1. Class

i	Interval	Class value
$\vdots$	$\vdots$	$\vdots$
$i-1$	$(e_{i-1} - w, e_{i-1} + w]$	e_{i-1}
i	$(e_i - w, e_i + w]$	e_i
$i+1$	$(e_{i+1} - w, e_{i+1} + w]$	e_{i+1}
$\vdots$	$\vdots$	$\vdots$

A fuzzy time series model AFTSMC, which adopts class time series, can be described as follows:

$$z_t^C = b_1 z_{t-1}^C + b_2 z_{t-2}^C + \cdots + b_p z_{t-p}^C. \tag{6}$$

As mentioned above, the time series $\boldsymbol{z}_t$, and the autoregressive coefficient b_i of Equation (6) are real values.

$$\begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & \rho_0 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & \rho_0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}. \tag{7}$$

ρ_k is the autocorrelation coefficient of lag k in time series $\boldsymbol{z}_t$:

$$\rho_k = \frac{Cov\left[z_t^C, z_{t-k}^C\right]}{Cov\left[z_t^C, z_t^C\right]}. \tag{8}$$

Here, the autocorrelation coefficient at lag 0 is clearly $\rho_0 = 1$, as in the AR model.

The procedure to build AFTSMC is as follows.

Step 1: Transformation of an original series to a class time series

A time series y_t is transformed into a stationary time series, and the classes indicated in Table 1 are defined. The stationary time series is then divided into classes and transformed into a class time series $\boldsymbol{z}_t$.

Step 2: Obtaining autoregressive coefficients

The autocorrelation coefficient ρ_k is set to the center z_t^C of the class time series. Using the autocorrelation coefficient ρ_k, the autoregressive coefficient b_i of AFTSMC is determined using the Yule-Walker equation, shown as Equation (7).

Step 3: Prediction

Classify the predictions of z_t^C obtained from the autoregressive Equation (6) into classes. The predicted value of an original series is determined by transforming the class into an original series.

The three steps outlined above are adopted to predict class time series by AFTSMC.

4. **Analysis of the Japanese National Consumer Price Index.** To check the characteristics of the AFTSMC, the Consumer Price Index for Japan, based on the year 2010, was used. There were 525 monthly data sets from January 1970 to September 2013. In this study, the period from January 1970 to October 2012 was used for model building, and the ensuing 12-month period from November 2012 to September 2013 was used to verify the prediction accuracy.

When the original series y_t represents the consumer price index, the value range of the stationary time series $\Delta\Delta_6 y_t$ is $[-3.4, 4.1]$, and the width of the class is set to 0.2. The autocorrelation coefficients of the difference series $\Delta\Delta_6 y_t$ and center z_t^C of the class time series are presented in Table 2 and Figure 1.

As illustrated in Table 2, the difference in the autocorrelation coefficients of the difference series $\Delta\Delta_6 y_t$, and its class value z_t^C is negligible. In addition, as illustrated in Figure 1, $\Delta\Delta_6 y_{t-1}$ and $\Delta\Delta_6 y_{t-6}$ are highly correlated. Hence, the second-order AFTSMC can be expressed as

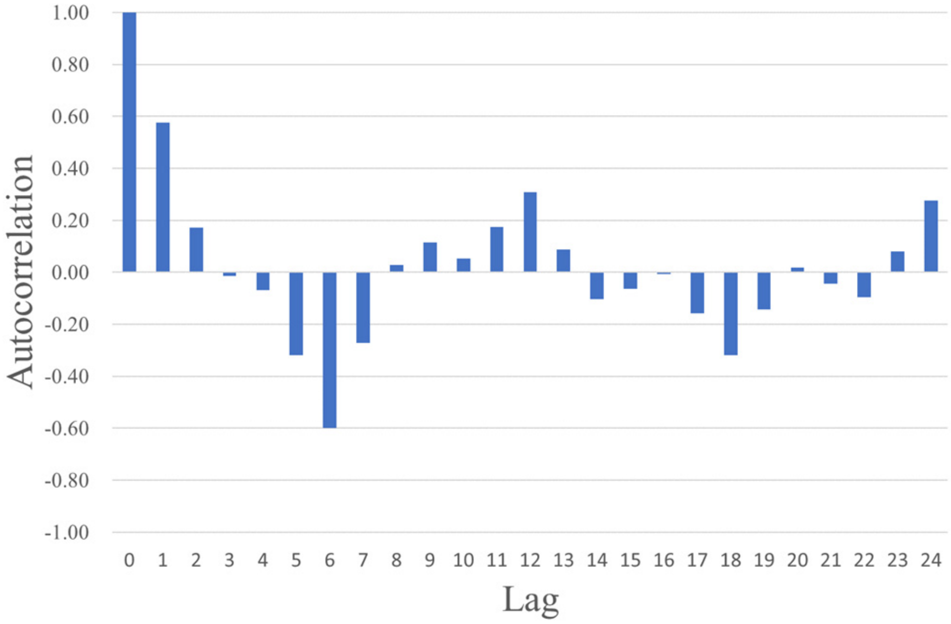
$$z_t^C = 1.373 z_{t-1}^C - 1.388 z_{t-6}^C. \tag{9}$$

Then the AR of the difference series $\Delta\Delta_6 y_t$ is

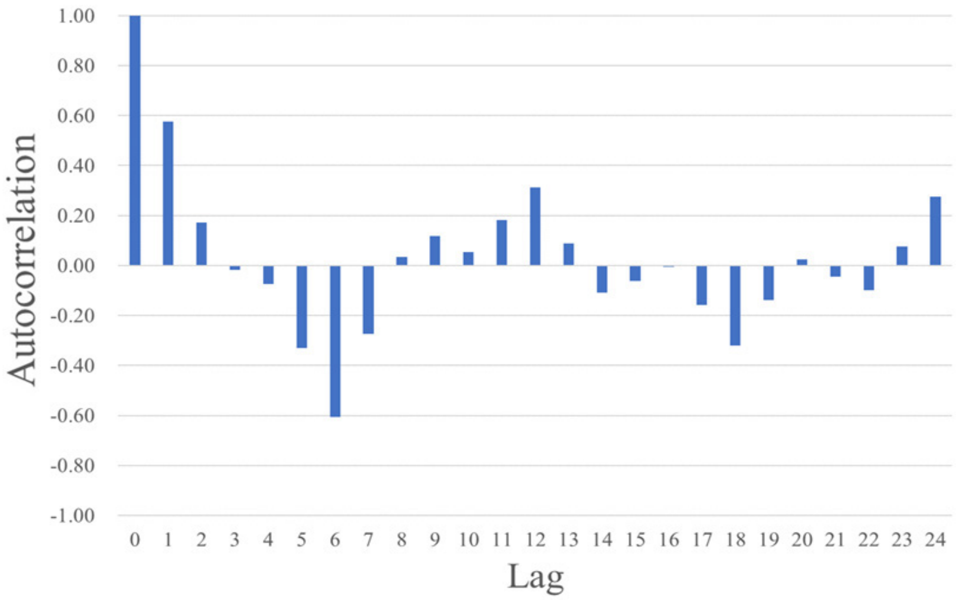
$$x_t = 1.386 x_{t-1} - 1.405 x_{t-6}. \tag{10}$$

TABLE 2. Autocorrelation coefficients of $\Delta\Delta_6 y_t$ and z_t^C

Lag	$\Delta\Delta_6 y_t$	z_t^C	Lag	$\Delta\Delta_6 y_t$	z_t^C
1	0.576	0.575	13	0.088	0.086
2	0.172	0.171	14	−0.108	−0.102
3	−0.018	−0.013	15	−0.062	−0.062
4	−0.074	−0.070	16	−0.005	−0.006
5	−0.331	−0.320	17	−0.157	−0.157
6	−0.606	−0.598	18	−0.320	−0.320
7	−0.273	−0.272	19	−0.138	−0.144
8	0.034	0.027	20	0.024	0.017
9	0.119	0.114	21	−0.045	−0.044
10	0.054	0.052	22	−0.099	−0.095
11	0.183	0.174	23	0.075	0.080
12	0.313	0.307	24	0.276	0.275



(a) Difference series



(b) Class value

FIGURE 1. Correlograms

The predicted values and original series of Equations (9) and (10) are illustrated in Figure 2. The model was constructed using the time series of the vertical line in Figure 2 until September 2012. The model was then validated. Since October 2012, forecasts have been made by adopting predicted values. In the model building period, it is evident that both AR and AFTSMC have large forecast errors. In the validation of prediction accuracies, the original series displayed stable values; however, both predictions displayed large fluctuations. In addition, it can be observed that at the end of the validation period, the predictions of AR are more disordered than those of AFTSMC.

Because increasing the width of the class may improve the prediction accuracy, the change in the prediction is checked by varying the width of the class w by approximately 1.0. The resulting autoregressive coefficients are listed in Table 3, and the prediction accuracies are listed in Table 4. The correlation coefficients of the original series and the

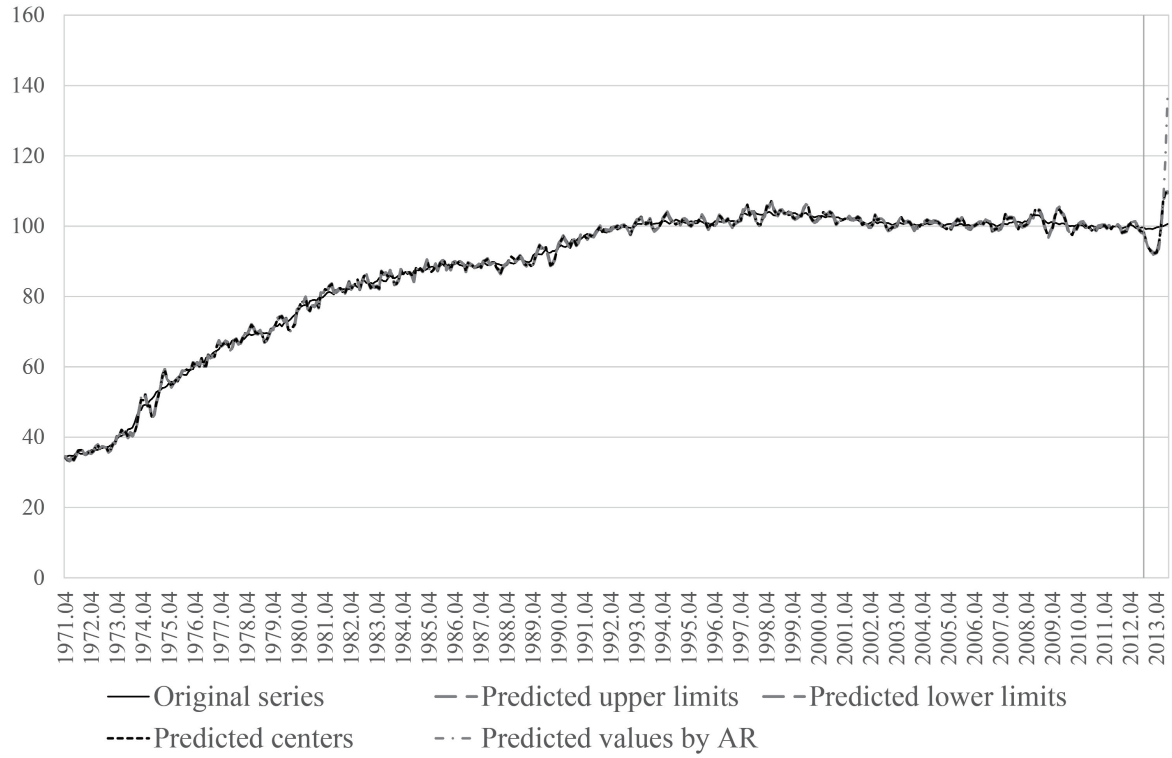


FIGURE 2. Original series and predicted values of AFTSMC ($w = 0.2$) and AR

TABLE 3. Autoregression coefficients and widths of class intervals

Lag	$w = 0.2$	$w = 0.4$	$w = 0.8$	$w = 1.0$	AR
1	1.373	1.292	1.066	0.948	1.386
6	−1.388	−1.318	−1.101	−0.970	−1.405

TABLE 4. Prediction accuracies and widths of class intervals

	$w = 0.2$	$w = 0.4$	$w = 0.8$	$w = 1.0$
Correlation coefficient between original series and the predicted values by AFTSMC	0.9975	0.9976	0.9986	0.9986
Correlation coefficient between original series and the predicted values by AR	0.9973			
Residual sum of squares between original series and the predicted values by AFTSMC	932.36	894.54	534.70	500.10
Residual sum of squares between original series and the predicted values by AR	997.1204			
Sum of squares of distances between original series outside the predicted interval, and the nearest boarder of the interval	497.8	411.4	173.0	114.9

predicted values are similar for AR and AFTSMC. However, when the class width value is increased from $w = 0.2$ to $w = 1.0$, the correlation coefficient approaches 1. The residual sum of squares calculated from the values predicted by the AFTSMC and the original series is smaller than that of AR. As the width of the class value increases, the residual sum of squares of the AFTSMC decreases. The values are largest at $w = 0.2$, 932.36, and smallest at $w = 1.0$, 500.10. In other words, the prediction accuracy can be improved by increasing the width of the class.

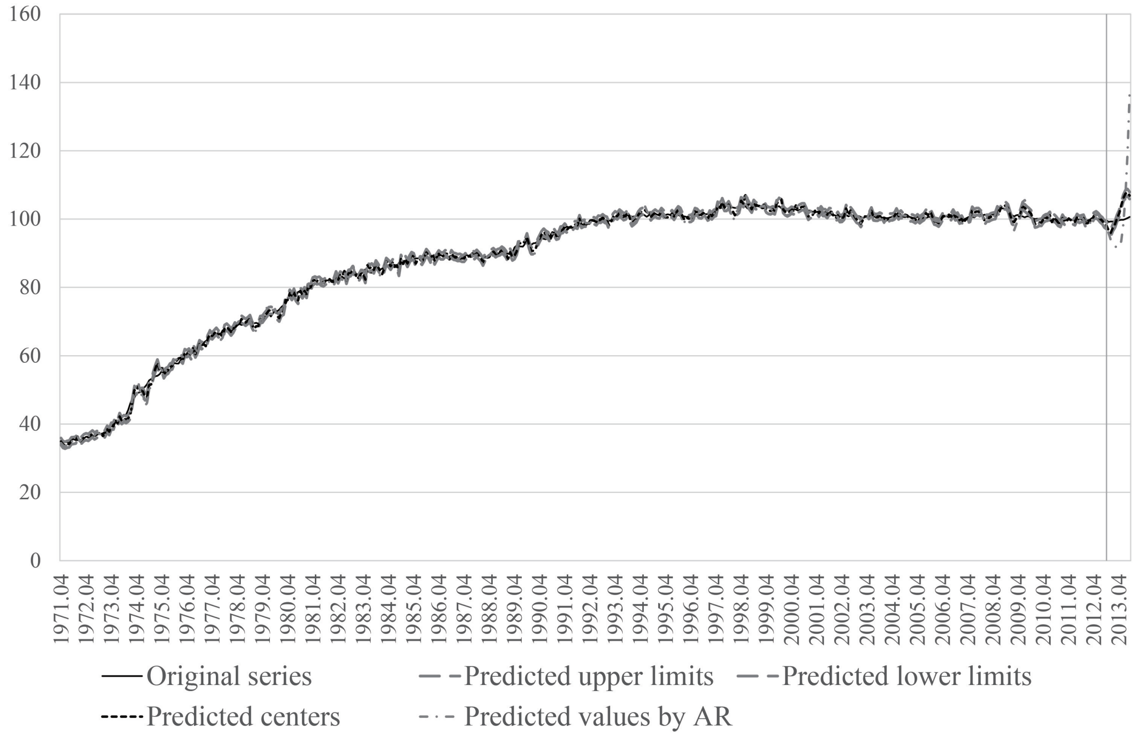


FIGURE 3. Original series and predicted values of AFTSMC ($w = 1.0$) and AR

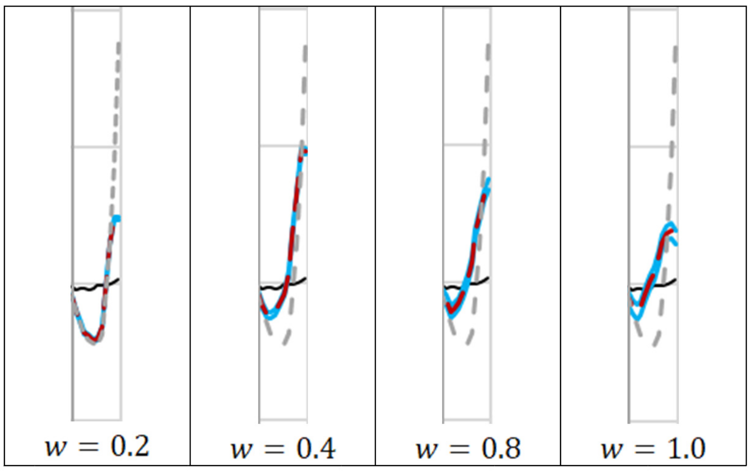


FIGURE 4. Widths of class and prediction accuracies

The sum of squares of the distances between the original series outside the predicted interval and the nearest border of the interval also decreases as w increases. The distances are largest at $w = 0.2$, 497.8, and smallest at $w = 1.0$, 114.9. However, the distances to these intervals cannot be directly compared in terms of prediction accuracy because the intervals become larger as w increases.

The original series and AFTSMC predictions for a class width of $w = 1.0$ are illustrated in Figure 3. In Figure 3, the width of the class does not just increase; the center of the predicted values seems to be closer to the original series. Furthermore, in the verification of AFTSMC, the predicted values are closer to the original series, and the irregularity of the fluctuations decreases.

The gray dashed line in Figure 4 indicates the AR prediction. The wide red dashed line indicates the center of AFRSMC, and the solid light blue line indicates the boundary of the section. As illustrated in Figure 4, the prediction accuracy of the proposed model improves as the width of the class increases. In summary, it was confirmed that the use of the rank time series improves the forecasting accuracy of the proposed model.

5. **Conclusion.** In this study, a fuzzy time-series model with class values was proposed. The proposed model can predict with high accuracy, when an appropriate class range is set.

Using the consumer price index of Japan as a numerical example, it was confirmed that the proposed model was able to forecast with high accuracy. In addition, it was confirmed that the proposed model was easy to handle.

In a previous study [11], the predicted values behaved unnaturally, and the width of the interval output also increased. In this study, by classifying the time series into classes, the prediction accuracy was improved. However, the width of the interval output does not fully express the vagueness of the time-series system. In future work, the proposed model will be further developed to describe the ambiguity of the time-series system.

REFERENCES

[1] M. E. A. Alqatqat and M. T. Feng, Methods in fuzzy time series prediction with applications in production and consumption electric, *American Journal of Mathematics and Statistics*, vol.10, no.3, pp.79-95, 2020.

[2] A. Inagaki, Fuzzy time-series of qualitative data, *J. Japan Society for Fuzzy Theory and Systems*, vol.14, no.6, pp.656-663, 2002 (in Japanese).

[3] Z. Ismail and R. Efendi, Enrollment forecasting based on modified weight fuzzy time series, *J. Artificial Intelligence*, vol.4, no.1, pp.110-118, 2011.

[4] D. Shan, W. Lu and J. Yang, The data-driven fuzzy cognitive map model and its application to prediction of time series, *International Journal of Innovative Computing, Information and Control*, vol.14, no.5, pp.1583-1602, 2018.

[5] Q. Song and B. S. Chissom, Forecasting enrollment with fuzzy time series – Part 1, *Fuzzy Sets and Systems*, vol.54, no.3, pp.269-277, 1993.

[6] H. J. Teoh, C. H. Cheng, H. H. Chu and J. S. Chen, Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets, *Data & Knowledge Engineering*, vol.67, pp.103-117, 2008.

[7] H. Tian, J. Li, F. Zhang, Y. Xu, C. Cui, Y. Deng and S. Xiao, Entropy analysis on intuitionistic fuzzy sets and interval-valued intuitionistic fuzzy sets and its applications in mode assessment on open communities, *J. Adv. Comput. Intell. Intell. Inform.*, vol.22, no.1, pp.147-155, 2018.

[8] F. M. Tseng, G. H. Tzeng, H. C. Yu and B. J. C. Yuan, Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems*, vol.118, no.1, pp.9-19, 2001.

[9] W. Qiu, P. Zhang and Y. Wang, Fuzzy time series forecasting model based on automatic clustering techniques and generalized fuzzy logical relationship, *Math. Probl. Eng.*, Article ID 962597, 2015.

[10] Y. Yabuuchi, T. Kawaura and J. Watada, Fuzzy autocorrelation model with fuzzy confidence intervals and its evaluation, *J. Adv. Comput. Intell. Intell. Inform.*, vol.20, no.4, pp.512-520, 2016.

[11] Y. Yabuuchi, An interval-type autocorrelated fuzzy time-series model used other fuzzification and its sequential reconstruction, *Proc. of Int'l Conf. Knowledge Based and Intelligent Information and Engineering Systems (KES2018)*, pp.996-1005, 2018.

[12] Y. Zhang, H. Qu, W. Wang and J. Zhao, A novel fuzzy time series forecasting model based on multiple linear regression and time series clustering, *Math. Probl. Eng.*, Article ID 9546792, 2020.

[13] Z. Zhang and Q. Zhu, Fuzzy time series forecasting based on K-means clustering, *Proc. of 2012 World Congress on Engineering and Technology*, pp.100-103, 2012.