Fuzzy Time Series Application in Predicting the Number of Confirmation Cases of Covid-19 Patients in Indonesia

Lintang Patria

Department of Information System, Universitas Terbuka, Indonesia

Corresponding author email: lintang@ecampus.ut.ac.id

Abstract

Forecasting is a statistical method that can use historical data patterns to predict   future events. This article discusses the prediction of the number of new confirmed cases of Covid-19 patients in Indonesia. The data used is from January 1, 2021 to August 7, 2021. The methods used are Fuzzy Time Series (FTS) Chen (2014) and Cheng et al. (2008). FTS is a forecasting method that uses rules and logic on fuzzy sets. The level of prediction accuracy is then calculated based on the Mean Absolute Percentage Error (MAPE) value. The MAPE values of these two methods are then compared to know which method is more suitable in this case study. The results showed that FTS Chen produced an accuracy of 12.75% and FTS Cheng produced an accuracy of 14.27%. The results of this study indicate that FTS Chen and FTS Cheng produce good accuracy and can be used to predict new confirmed cases of Covid 19 sufferers in Indonesia.

Keywords: Chen, Cheng, Covid, Fuzzy Time Series

1. Introduction

The Covid-19 pandemic in Indonesia is part of the 2019 coronavirus disease (Covid-19) pandemic that has hit the whole world. The first positive case of Covid -19 in Indonesia was detected on March 2, 2020, when two people were confirmed to have contracted it from a Japanese citizen. On April 9, the pandemic had spread to 34 provinces with DKI Jakarta, West Java and Central Java as the provinces most exposed to SARS-CoV-2 in Indonesia (Tosepu et al., 2020; Novika and Meliyani, 2021; Munandar et al., 2020). The government then implemented a policy of large-scale restrictions and various other rules in suppressing the rate of increase in confirmed cases of Covid-19 sufferers to date. The fluctuating increase in cases per day is reported by the Ministry of Health to the public as a form of information disclosure by the government (Olivia et al., 2020; Andriyani et al., 2021; Aditantri et al., 2021; Suyatna, 2021).

The data on the increase in the number of confirmed cases of Covid-19 sufferers becomes historical data that can be analyzed and interpreted and used to predict the number of Covid-19 sufferers in the future (Anastassopoulou et al. 2020; Sarkar et al., 2020; Li et al., 2020; Yousaf et al., 2020; Bastos and Cajueiro et al., 2020). The science that studies this historical data relationship is time series. Various methods in time series have been developed by researchers, one of which is fuzzy time series (Chen, 1996; Chen, 2014; Cheng et al., 2008; Zeroual et al., 2020; Zhang et al., 2020).

Some literature for fuzzy time series can be seen in refs (Kumar and Kumar, 2021; Kumar and Susan, 2021; Xian and Cheng, 2021; Hadjira et al., 2021; Kumar and Kaur, 2021; Mishra et al., 2020). Mishra et al. (2020) investigated prediction of the new COVID-19 cases in India using Artificial Neural Network, Fuzzy Time Series and ARIMA Models. Castillo and Melin (2020) combined of fractal dimension and fuzzy logic for forecasting of COVID-19 in the world. Kumar and Kumar (2021) presented prediction of COVID-19 infected cases and deaths in India using A novel hybrid fuzzy time series. Kumar and Susan (2021) studied forecasting of COVID-19 using article swarm optimization of partitions and fuzzy order for fuzzy time series. Xian and Cheng (2021) proposed prediction of the new COVID-19 cases using Pythagorean fuzzy time series model based on Pythagorean fuzzy c-means. Hadjira et al. (2021) compared rima model, holt-winters–no seasonal and fuzzy time series for new cases of COVID-19 in Algeria. Kumar and Kaur (2021) evaluated of COVID-19 forecasting models using hybrid SOM-Fuzzy time series (SOMFTS) technique.

This article will discuss the fuzzy time series developed by Chen and Cheng. Forecasting using fuzzy time series can explore past data patterns to predict future events by using various properties, principles and logic in fuzzy sets. Fuzzy time series is a forecasting method introduced by Song and Chisom. The difference between the Chen and

Cheng methods lies in the defuzzification process. The Chen method does not use weighting, while the Cheng method uses weighting.

The purpose of this article is to predict the number of confirmed cases of Covid-19 sufferers using the Chen and Cheng Fuzzy Time Series method. The prediction accuracy of the method is then calculated using MAPE. MAPE results from both methods were then compared. This article is a manifestation of the application of mathematics and statistics in the world of health in Indonesia.

2. Methodology

The data on the increase in confirmed cases of Covid-19 sufferers per day are observed from January 1, 2021 to August 7, 2021can be seen in Table 1.

Table 1. Data on additional confirmed cases of Covid-19 sufferers

No	Date	New Case
1	01-Jan	8072
2	02-Jan	7203
3	03-Jan	6877
…	…	…
…	…	…
217	05-Aug	35764
218	06-Aug	39532
219	07-Aug	31753

The calculation scheme for the Fuzzy Time Series Chen and Cheng can be seen in Figure 1.
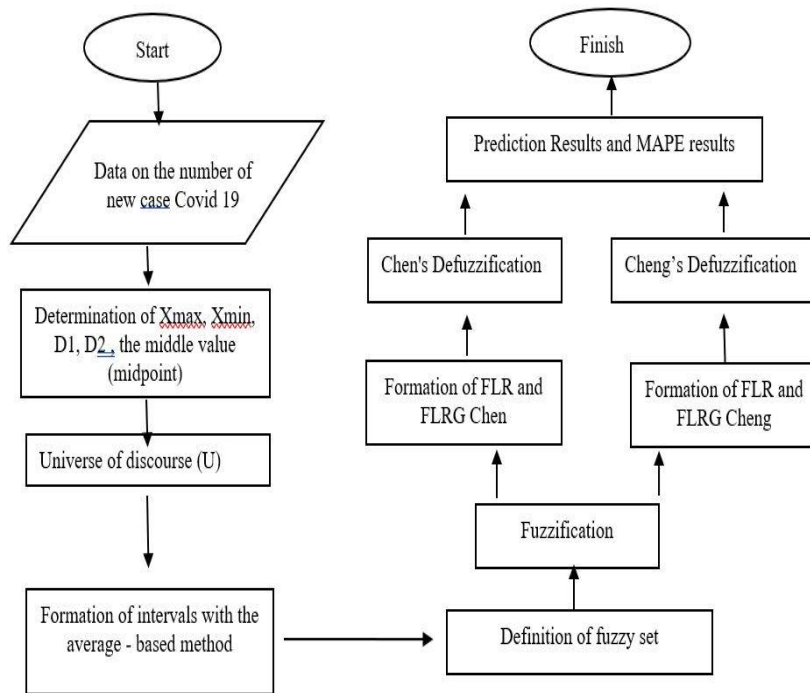

Figure 1. Schematic of calculation of average - based fuzzy time series

The stages of research on average based fuzzy time series (Chen 1996; Cheng et al., 2008; Sukono et al., 2020)

1. Data on the increase in confirmed cases of Covid-19 sufferers is used as input in Chen and Cheng's Fuzzy Time Series method
2. Determine Xmax, Xmin, $D1$, $D2$ and middle point
   Xmin = Minimum data
   Xmax = Maximum data

$D1$ and $D2$ are arbitrary positive numbers and are determined by the researcher. Usually, the values of D1 and D2 are numbers that make it easy to calculate the universe set from the historical data formed.

3. Determining U is the Universe of discourse or the set of historical data universes, namely,

$$U = \{Xmin\text{-}D1, Xmax\text{+}D2]U = \{Xmin\text{-}D1, Xmax\text{+}D2] \tag{1}$$

4. The length of the interval is determined according to the justification of the researcher. The number of intervals is adjusted to the Universe of discourse or the universe of historical data sets.

5. After getting the number of intervals, we define the fuzzy set according to the number of classes. Example $A_1, A_2, ...A_p$ is a fuzzy set that has a linguistic value, the definition of the set $A_1, A_2, ...A_p$ in the universe of U is as follows

$$A_1 = 1/u_1 + 0{,}5/u_2 + 0/u_3 + ... + 0/u_p$$
$$A_2 = 0{,}5/u_1 + 1/u_2 + 0{,}5/u_3 + ... + 0/u_p$$
$$A_3 = 0/u_1 + 0{,}5/u_2 + 1/u_3 + 0{,}5/u_4 ... + 0/u_p$$

....

....

....

$$A_p = 0/u_1 + 0/u_2 + 0/u_3 + ...0{,}5/u_{p\ 1} + 1/u_p$$

Where $u_i (i = 1, 2, 3, ..., p)$ is elements of the universal set U and dan number marked with "/ " state the degree of membership $\mu_{A_i}(u_i)$ to $A_i (i = 1, 2, 3, ..., p)$ where the value is 0.05 or 1.

6. The historical data is classified according to the fuzzy set-in step 5.

7. After getting the fuzzy set, we form the FLR. FLR is a fuzzy relation between data in year i and data in year j. If $(t - 1) = Ai$ and $(t) = Aj$, then the FLR formed is $A_i \rightarrow A_j$. FLR is a fuzzy relation between data in year i and data in year j. After all historical data are classified into FLR, FLRG is formed. FLRG is a group obtained from FLR which has the same left side. The difference between the Chen and Cheng methods is in the weighting of the FLRG. The Chen method does not use weighting and the Cheng method uses weighting.

8. The following are the forecasting rules for the defuzzification process, namely:

    Rule 1

    If the result of fuzzification in year t is A_i and there is no fuzzy logic relation. Example, if $A_i \rightarrow \varnothing\square$, then the forecasting result F_{t+1} is m_i and m_i is the middle value u_i .

    Rule 2.

    If the result of fuzzification in year t is A_i and there is only one FLR on the FLRG, Example, if $A_i \rightarrow A_j$, where A_i and A_j is a fuzzy set, then the forecasting result F_{t+1} is m_j .

    Rule 3.

    If the result of fuzzification in year t is A_i and A_i has multiple FLR on the FLRG, Example, if $A_i \rightarrow A_{j1}, A_{j2}...A_{jk}$, where $A_i, A_{j1}, A_{j2}, ...A_{jk}$ is a fuzzy set and the maximum value of the membership function of $A_{j1}, A_{j2}, ...A_{jk}$ are in the interval $u_{j1}, u_{j2}, ...u_{jk}$ and the middle value from $u_{j1}, u_{j2}, ...u_{jk}$ are $m_{j1}, m_{j2}, ..., m_{jk}$, then the forecasting result

    $$F_{t+1} \text{ is } \frac{m_{j1} + m_{j2} + ... + m_{jk}}{k} \tag{5}$$

    where k is the number of middle values. To find the middle value m_i the fuzzy set interval can be used the following equation:

    $$m_i = \frac{\text{upper limit } + \text{lower limit}}{2} \tag{6}$$

9. The error rate of the prediction or MAPE can be calculated by the equation (7)

$$MAPE = \sum_{i=1}^{n} \left| \frac{x_i - p_i}{x_i} \right| \times 100\%$$

(7)

where;

x_i = actual data in year i

p_i = forecast value in year i

n = amount of data

3. Results and Discussion

Based on these data, we get Xmin = 2383 and Xmax = 56757, the selected D1 and D2 are D1 = 85 and D2 = 1543, so the value obtained is U = [2300,58300]. According to the researcher's justification, the length of the interval is determined to be 2000. According to U, the number of intervals (fuzzy numbers) is 28. The obtained intervals are 28 intervals, namely [2300,4300], [4300,6300], [6300,8300], … , [54300,56300], [56300,58300]. The fuzzification stage is carried out according to the number of intervals formed. The fuzzification results can be seen in the following Table 2.

Table 2. Fuzzification

No	Date	New Case	Fuzzification
1	01-Jan	8072	A3
2	02-Jan	7203	A3
3	03-Jan	6877	A3
…	…	…	…
…	…	…	…
…	…	…	…
217	05-Aug	35764	A17
218	06-Aug	39532	A19
219	07-Aug	31753	A15

1. FLR process

   FLR formed by fuzzy relationship A_i from year to year for $1 \le i \le 28$ can be seen in Table 3.

Table 3. FLR

No	Date	FLR
1	01 Jan -> 02 Jan	A3->A3
2	02 Jan -> 03 Jan	A3->A3
3	03 Jan -> 04 Jan	A3->A3
…	…	…
…	…	…
…	…	…
216	04 Aug->05 Aug	A17->A17
217	05 Aug->06 Aug	A17->A19
218	06 Aug->07 Aug	A19->A15

2. Forming FLRG

   FLRG is done by grouping fuzzy sets that have the same current state, then grouped into one group in the next state based on the FLR table. Chen's FLRG method can be seen in Table 4 and Cheng's FLRG method can be seen in Table 5.

Table 4. Chen Method using FLRG

No	FLRG	No	FLRG
1	A1->A1,A2	15	A16->A17,A19,A24
2	A2->A1,A2,A3	16	A17->A17,A19,A20
3	A3->A1,A2,A3,A4	17	A18->A15,A17
4	A4->A3,A4,A5,A6	18	A19->A13,A15,A16,A18
5	A5-> A3,A4,A5,A6	19	A20->A18,A23
6	A6->A5,A6,A7	20	A21->A20,A23
7	A7->A5,A6,A10	21	A22->A16,A19,A223
8	A9->A10	22	A23->A21,A27
9	A10->A9,A10,A12	23	A24->A22,A24
10	A11->A16	24	A25->A22
11	A12->A12,A13	25	A26->A25
12	A13->A13,A14,A22	26	A27->A28
13	A14->A15	27	A28->A26
14	A15->A11,A17		

Table 5. Cheng Method using FLRG

No	FLRG	No	FLRG
1	A1->5A1,5A2	15	A16->A17,A19,A24
2	A2->4A1,62A2,11A3	16	A17->2A17,2A19,A20
3	A3->A1,10A2,13A3,8A4	17	A18->A15,A17
4	A4->5A3,11A4,4A5,2A6	18	A19->A13,A15,A16,A18
5	A5-> 2A3,3A4,9A5,4A6	19	A20->A18,A23
6	A6->4A5,5A6,3A7	20	A21->A20,A23
7	A7->A5,A6,A10	21	A22->A16,A19,A223
8	A9->A10	22	A23->A21,A27
9	A10->A9,4A10,A12	23	A24->A22,A24
10	A11->A16	24	A25->A22
11	A12->A12,A13	25	A26->A25
12	A13->A13,A14,A22	26	A27->A28
13	A14->A15	27	A28->A26
14	A15->A11,A17		

3. The process of defuzzification of forecasting values, namely by finding the middle value of each interval based on equation (5), and calculating the forecast value according to the defuzzification rules that have been discussed. The results of defuzzification for Chen's method can be seen in Table 6 and Cheng's method can be seen in Table 7.

Table 6. The results of defuzzification using the Chen

No	Group	Forecasting	No	Group	Forecasting
1	A1	4300	15	A16	41300
2	A2	5300	16	A17	38633
3	A3	6300	17	A18	33300
4	A4	10300	18	A19	32300
5	A5	10300	19	A20	42300
6	A6	13300	20	A21	41300
7	A7	15300	21	A22	39967
8	A9	21300	22	A23	49300
9	A10	21967	23	A24	47300
10	A11	33300	24	A25	45300
11	A12	26300	25	A26	51300

12	A13	33967	26	A27	57300
13	A14	31300	27	A28	53300
14	A15	29300			

Table 7. The results of defuzzification using the Cheng

No	Group	Forecasting	No	Group	Forecasting
1	A1	4300	15	A16	41300
2	A2	5482	16	A17	38100
3	A3	7050	17	A18	33300
4	A4	9573	18	A19	32300
5	A5	10967	19	A20	42300
6	A6	17511	20	A21	41300
7	A7	15300	21	A22	39967
8	A9	21300	22	A23	49300
9	A10	17417	23	A24	47300
10	A11	33300	24	A25	45300
11	A12	26300	25	A26	51300
12	A13	33967	26	A27	57300
13	A14	31300	27	A28	53300
14	A15	29300			

4. Forecasting results for positive confirmed case of Covid-19 can be seen in Table 8. Also, forecasting plots using the Chen and Cheng method can be seen in Figure 2.

Table 8. Forecasting results using Chen dan Cheng methods

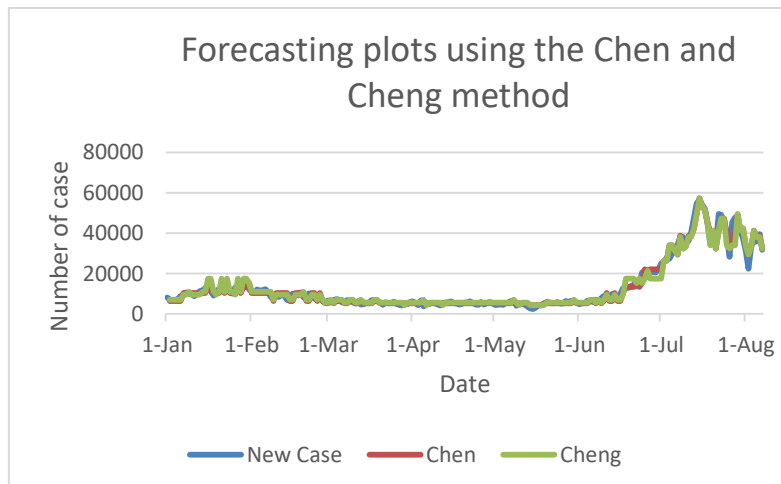| No	Date	New Case	Chen	Cheng
1	01-Jan	8072		
2	02-Jan	7203	6300	7050
3	03-Jan	6877	6300	7050
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
217	05-Aug	35764	38633	38100
218	06-Aug	39532	38633	38100
219	07-Aug	31753	32300	32300

Figure 2. Forecasting plots using the Chen and Cheng method

4. Conclussion

In this paper, we have predicted the number of positive confirmed case of Covid-19 using the fuzzy time series method The Chen and Cheng's Fuzzy Time Series can be used to predict the number of confirmed cases of Covid-19 sufferers. The MAPE of the Chen method is 12.75% and the MAPE of the Cheng method is 14.27%. It can be seen that the Fuzzy Time Series can be used to predict the number of confirmed cases of Covid-19 sufferers.

References

Aditantri, R., Mahliza, F., & Wibisono, A. D. (2021). Urban Planning and E-Commerce: Understanding the Impact During Pandemic Covid-19 in Jakarta. *International Journal of Business, Economics, and Social Development*, *2*(3), 135-142.

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PloS one*, *15*(3), e0230405.

Andriyani, D., Nailufar, F., Yurina, Y., Ratna, R., & Rahmah, M. (2021). Analyzing the Sustainability of Micro, Small and Medium Enterprises during Covid-19 Pandemic in Bireuen Regency, Indonesia. *International Journal of Business, Economics, and Social Development*, *2*(3), 119-126.

Bastos, S. B., & Cajueiro, D. O. (2020). Modeling and forecasting the early evolution of the Covid-19 pandemic in Brazil. *Scientific Reports*, *10*(1), 1-10.

Castillo, O., & Melin, P. (2020). Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic. *Chaos, Solitons & Fractals*, *140*, 110242.

Chen, S. M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems, 81*(3), 311-319.

Chen, M. Y. (2014). A high-order fuzzy time series forecasting model for internet stock trading. *Future Generation Computer Systems, 37*, 461-467.

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*(4), 2568-2575.

Hadjira, A., Salhi, H., & El Hafa, F. (2021). A Comparative Study between ARIMA Model, Holt-Winters–No Seasonal and Fuzzy Time Series for New Cases of COVID-19 in Algeria. *American Journal of Public Health*, *9*(6), 248-256.

Kumar, A., & Kaur, K. (2021, February). A hybrid SOM-Fuzzy time series (SOMFTS) technique for future forecasting of COVID-19 cases and MCDM based evaluation of COVID-19 forecasting models. In *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 612-617). IEEE.

Kumar, N., & Kumar, H. (2021). A novel hybrid fuzzy time series model for prediction of COVID-19 infected cases and deaths in India. *ISA transactions*. In Press

Kumar, N., & Susan, S. (2021). Particle swarm optimization of partitions and fuzzy order for fuzzy time series forecasting of COVID-19. *Applied Soft Computing*, 107611.

Li, Q., Feng, W., & Quan, Y. H. (2020). Trend and forecasting of the COVID-19 outbreak in China. *Journal of Infection*, *80*(4), 469-496.

Mishra, P., Fatih, C., Rawat, D., Sahu, S., Pandey, S. A., Ray, M., ... & Sanusi, O. M. (2020). Trajectory of COVID-19 Data in India: Investigation and Project Using Artificial Neural Network, Fuzzy Time Series and ARIMA Models. *Annual Research & Review in Biology*, 46-54.

Munandar, D., Supian, S., & Subiyanto, S. (2020). Probability distributions of COVID-19 tweet posted trends uses a nonhomogeneous Poisson process. *International Journal of Quantitative Research and Modeling*, *1*(4), 229-238.

Novika, F., & Meliyani, R. (2021). Procrustes Analysis of Indonesian Mortality Table Iv and Indonesia's Death Rate During Covid-19 Pandemic. *International Journal of Quantitative Research and Modeling*, *2*(2), 91-96.

Olivia, S., Gibson, J., & Nasrudin, R. A. (2020). Indonesia in the Time of Covid-19. *Bulletin of Indonesian Economic Studies*, *56*(2), 143-174.

Sarkar, K., Khajanchi, S., & Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in India. *Chaos, Solitons & Fractals*, *139*, 110049.

Sukono, Sambas, A., He, S., Liu, H., Vaidyanathan, S., Hidayat, Y., & Saputra, J. (2020). Dynamical analysis and adaptive fuzzy control for the fractional-order financial risk chaotic system. *Advances in Difference Equations*, *674*(1), 1-12.

Suyatna, N. (2021). The Effect of Stimulus Policies and Risk-Based Capital Levels on Solvency at PT. Indonesian Takaful using Sharia Principles during the Majeur Force Period (Covid-19 Pandemic) in Indonesia. *International Journal of Business, Economics, and Social Development*, *2*(1), 24-31.

Tosepu, R., Gunawan, J., Effendy, D. S., Lestari, H., Bahar, H., & Asfian, P. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Science of the total environment*, *725*, 138436.

Xian, S., & Cheng, Y. (2021). Pythagorean fuzzy time series model based on Pythagorean fuzzy c-means and improved Markov weighted in the prediction of the new COVID-19 cases. *Soft Computing*, *25*(22), 13881-13896.

Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos, Solitons & Fractals*, *140*, 110121.

Zhang, H., Song, H., Wen, L., & Liu, C. (2021). Forecasting tourism recovery amid COVID-19. *Annals of Tourism Research*, *87*, 103149.