Covid-19 daily forecasting during ramadhan in countries with high muslim population

G Darmawan1,2*, D Rosadi2, and B N Ruchjana3
1Department of Statistics, Universitas Padjadjaran, Jl Raya Bandung Sumedang km 21 Jatinangor, Sumedang 45363, Indonesia
2Department of Mathematics, Universitas Gadjahmada, Sekip Utara, Yogyakarta, Indonesia
3Department of Mathematics, Universitas Padjadjaran, Jl Raya Bandung Sumedang km 21 Jatinangor, Sumedang 45363, Indonesia

*gumgum@unpad.ac.id

Abstract. For Muslims, the month of Ramadhan is a month of full ritual and non-ritual worship, such as Tarawih, Eid al-Fitr, and going home to hometown. These activities are very risky if carried out during the outbreak of Covid-19, as the spread of Covid-19 occurs within a short distance from one person to another. This study aims to look at the Covid-19 movement in Ramadhan for countries with the largest Muslim populations, such as Indonesia, Malaysia, Saudi Arabia, Egypt, Kuwait, Pakistan, Afghanistan, Oman, Qatar, Morocco, Bangladesh and India. Model the movement of Covid-19 using the Singular Spectrum Analysis Model. The method used is visual scanning and modelling using the Singular Spectrum Analysis Model. Daily forecasting is done twice as much as the first forecast for the weekly MAPE (Mean Absolute Percentage Error) from May 8th, 2020 to May 14th, while the second forecast (Data is not yet known) for the weekly MAPE from May 15th, 2020 to Eid. Based on the results of the graphical exploration, in general, the Muslim-majority countries rose in mining cases in Covid-19 during the month of Ramadhan. The states that declined in the case of Covid-19 were Oman, Qatar and Morocco. The Singular Spectrum Analysis Model with the Long Memory Time Series approach is more accurate than the Alexandrov version of the Auto grouping Model. The overall results of the MAPE.

1. Introduction
Since the Covid-19 case occurred in Wuhan City, China, many world researchers have modelled and forecasted the movements of confirmed cases, recovered cases and deaths in different countries. As in China itself [1], the real-time forecast for the Covid-19 case in Wuhan was between 5 and 24 February 2020. According to the study, the spread in the province of Hubei and surrounding areas of Covid-19 was stable.

The previous study examined the Covid-19 incubation period in Wuhan City. The results indicated that Covid-19 incubation was approximately 10-14 days, indicating that the 14-day quarantine period would be very large, ensuring no disease among individuals [2]. They were suspected of being exposed to Covid-19. If in 2018, Covid-19 modelling was carried out in Wuhan City using the Epidemiology Model and ARIMA, the results of the ARIMA Model could be used to predict the spread of different types of influenza viruses [3]. Positive influenza-positive diseases occur in children. The ARIMA model is also used to model the Covid-19 distribution data taken from the John Hopkins University website (https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html) [4].
The previous study examined the relationship between the Air Quality Index with the number of Covid-19 sufferers in China through the Poisson regression model. The result turned out to be a significant relationship between the Air Quality Index with the number of Covid-19 sufferers in an area [5].

Another study conducted work in East Asian countries such as China, Japan and South Korea on population activity in these nations [6]. I saw that while the Pandemic was international, the treatment was still local. A recent study also discusses the factors that affect the spatial distribution of Covid-19 and uses the data panel model to assess the relationship between the causal variables, including their spatial effects. As a result, the distribution of the Covid-19 population is 0.13 or 13 per 100,000 population in Wuhan between January and early March 2020 [7].

The month of Ramadhan, when Muslims observe their worship more frequently than in other months. In particular, worship in a congregation such as Tarawih and Eid prayer. Moreover, the practice of Eid, such as going home, going home amongst relatives and even shopping at the market or the Mall more frequently than other months, significantly increases the risk of Covid-19.

In this research, the spreading of Covid-19 during Ramadhan was examined in many countries with a significant Muslim population. Will the month of Ramadhan raise the number of Covid-19 patients or not? Instead, at the end of Ramadan, we modelled all Covid-19 data in 12 countries using the Singular Spectrum Analysis from two separate SSA models. The SSA model with the smallest MAPE value would be the best one for modelling the Covid-19 movement data in the countries examined.

2. Method

2.1. Singular Spectrum Analysis (SSA)

SSA is a versatile time series model that is not constrained by standard assumptions, white noise, and so on. This model is the development of a decomposition model that defines the time series data in its components. This paper uses two SSA methods of the same type, SSA Autogrouping. Auto grouping SSA on the Alexandrov model uses the Spectral frequency threshold while Alternative Autogrouping uses the Long Memory Time Series method [8, 9].

The SSA stage consists of Embedding, Singular Value Decomposition, Grouping and Diagonal Averaging. Embedding is the method of translating univariate data to \( L \times K \) dimension trajectory matrices, with \( K = N - L + 1 \). \( L \) is the length of the window of the SSA parameters obtained by trial and error, while \( N \) is the sum of time series data. During the second step, the singular value Decomposition is calculated; at this step, the triple eigent value of the matrix where \( X \) is the trajectory vector. The third stage is grouping, which is based on the same pattern (trend, seasonal). Precise is Diagonal Averaging; at this point, matrix data is transformed back to new univariate data. The SSA phases can be seen in [10, 11].

2.2. Auto grouping SSA Alexandrov

In this section, SSA automatic grouping based on trend extraction based on Alexandrov method will be explained. The first, we introduce the periodogram of a time series. Let us consider the Fourier representation of the elements of a time series of length \( N \).

\[
x_N = c_0 + \sum_{\substack{k \in \mathbb{Z} \setminus \{-N/2, -N/2 + 1, \ldots, N/2 - 1\} \quad \text{if} \quad N \text{ is odd} \quad \text{or} \quad \{0\} \quad \text{if} \quad N \text{ is even}}} (c_k \cos(2\pi nk/N) + s_k \cos(2\pi nk/N) + (-1)^n c_{N/2})
\]

(1)

where \( k \in \mathbb{Z} \), \( 0 \leq n \leq N - 1 \) and \( c_{N/2} = 0 \) if \( N \) is an odd number. The periodogram of \( X \) at the frequencies \( \omega \in \{k/N\}_{k=0}^{\lfloor N/2 \rfloor} \) is defined as

\[
I^x_N(\omega) = \frac{N}{2} \begin{cases} 
2c_0^2, & k = 0 \\
2s_0^2 + 2s_1^2, & 0 < k < N/2 \\
2c_{N/2}^2, & k = N/2, N \text{ even} 
\end{cases}.
\]

(2)
Define the $k$-th element of the discrete Fourier transform of $X$ as

$$F_k(X) = \sum_{n=0}^{N-1} e^{-i2\pi nk/N} x_n,$$

The periodogram $I_X^N(\omega)$ at the frequencies $\omega \in \{k/N\}_{k=0}^{N/2}$ can be simplified as

$$I_X^N(k/N) = \frac{1}{N} \left\{ \begin{array}{ll} 2|F_k(X)|^2 & \text{if } 0 < k < N/2 \\ |F_k(X)|^2 & \text{if } k = 0 \text{ or } N \text{ is even and } k = N/2. \end{array} \right.$$

All frequency in the interval $(0,0.5)$ is multiplied by two. This equation is done to ensure the following property:

$$\|X\|_2^2 = \sum_{n=0}^{N-1} x_n^2 = \sum_{k=0}^{N/2} I_X^N(k/N)$$

where $\pi_X^N(\omega) = \sum_{k:0 \leq k \leq N/2} I_X^N(k/N), \omega \in [0,0.5]$ is the cumulative contribution of the frequencies $[0,\omega]$. Then, for given $\omega_0 \in (0,0.5)$, define the contribution of low frequencies from the interval $[0,\omega_0]$ to $X \in \mathbb{R}^N$ as

$$C(X,\omega_0) = \pi_X^N(\omega_0)/\pi_X^N(0.5).$$

Then, given parameters $\omega_0 \in (0,0.5)$ and $\omega_0 \in [0,1]$, [3] propose to select those SVD components whose eigenvectors satisfy the following criteria:

$$C(U_j,\omega_0) \geq C_0$$

where $U_j$ is a corresponding $j$th eigenvector.

### 2.3 Alternative Method of Autogrouping SSA
In this section, we propose our new automatic algorithm on Singular Spectrum Analysis. The procedure of our automatic Grouping Algorithm can be described using the following steps:

1) Denote the original series as $X = (x_1,x_2,\ldots,x_N)$, the univariate time series data with length $N$.
2) Define Window Length ($L$).
3) Determine the shape of the trajectory matrix based on $L$ as in equation.
4) Find the Eigen Triple, $i = 1,2,\ldots,d$ ($d$ is the number of positive eigenvalues) of the $XX^T$ matrix, where $X$ is the Trajectory matrix.
5) Identify the order difference of each series $U_i$.
6) Make a differencing for all $U_i$ based on the value of in order to obtain the UD$_i$ series. Using this step, then we obtain Data wich stationary pattern.
7) Specify the maximum value of seasonal order of Pi of UD$_i$; Here, the $P_i$ values are between 0 and 24.
8) Group UD$_i$ series based on $P_i$ values obtained from the previous step 7, a grouping of UD$_i$ is based on periods, 3,6,12 and 24, UD$_i$ with $P_i=0$ is grouped as non-seasonal data, where when $P_i>24$ is considered as cycle pattern.
9) Create diagonal averaging for grouping component from step 8.
10) Forecasting (Linear Recurrent Method).
11) Determine MAPE values based on the number of our sample data.
The MAPE of the $h$-step ahead forecast is defined as

$$
    MAPE = \left( \frac{1}{n} \sum_{i=1}^{n} \frac{y_{ni} - \hat{y}_{ni}}{\hat{y}_{ni}} \right) \times 100\%
$$

where $n$ denotes the last data used in the in-sample data and represents the forecast values.

12) Go to step 2 until all of the L choices have been calculated. The best model will give the smallest MAPE value.

3. Result and Discussion

The Covid-19 data source used in this paper was drawn from John Hopkins University. In the first part of the descriptive study, data on the rise in Covid-19 cases were plotted both before and during Ramadhan. Data are plotted on the increase in daily cases in 12 countries with the largest Muslim population.

In the second part, the modelling is done using the two version SSA model, namely SSA Alexandrov and SSA Alternative. The selection of the best models of the two models is based on MAPE for the last six days of the month of Ramadan. The smallest MAPE results show that the SSA model is better than any other SSA model.

The program used in this review is R 3.6.1 Program with many packages designed to run R macros. Package R needed is Rssa which functions to perform SSA analysis, both of which are used for grouping in alternative SSA.

3.1. Descriptive Analysis

In the descriptive analysis, confirmed cases of 12 countries were plotted from the first month to May. All figures are created by R 3.6.1.

![Descriptive Analysis](image)

Figure 1. Covid-19 movements in four countries (Indonesia, Malaysia, Bangladesh and India)
Figure 1 shows the pattern of Covid-19 in four countries (Indonesia, Malaysia, Bangladesh and India). Of the four countries, three (Indonesia, Bangladesh, India) have the same pattern of ascending both before the month of Ramadan (not yet the dotted black vertical line) and after the month of Ramadhan. The Covid-19 movement in Malaysia passed the decline about the thirtieth day, but at the beginning of the month of Ramadan, it increased again and gradually declined towards Eid. It has happened because many Muslim people move from the city to their village for gathering with their family. The end of Ramadhan, the spreading of Covid-19 decreased because the people improve their immune because of fasting. The biggest surge occurred in India (3000), second Bangladesh (1200), followed by Indonesia (600) and the smallest Malaysia (200).

Figure 2. Covid-19 movements in four countries (Saudi Arabia, Egypt, Kuwait and Morocco)

Figure 2 demonstrates the trend of the introduction of cases to Covid-19 in four countries (Saudi Arabia, Egypt, Kuwait and Morocco). Of the four nations, three (Saudi Arabia, Egypt, Kuwait) have the same pattern of ascending both before the month of Ramadan (not yet the dotted black vertical line) and after the month of Ramadhan. The Covid-19 campaign in Morocco passed its decline about forty-five days, reaching the month of Ramadan, gradually returning to the Eid. The highest rise occurred in Saudi Arabia by 1,000 seconds in Kuwait, followed by Egypt by 500 and the smallest of Morocco About 250 patients.
Figure 3 shows the pattern of adding cases of Covid-19 in four countries (Afghanistan, Pakistan, Oman and Qatar). All four countries have the same pattern of rising both before the month of Ramadan (not yet the dotted black vertical line) or after the month of Ramadhan. The Covid-19 movement in Oman sloped into the month of Ramadhan, but nearing Eid there was a significant increase nearing Eid. The biggest surges occurred in Qatar and Pakistan by 1500 patients and second in Afghanistan and Oman by 300 patients.

3.2. Covid-19 Daily Forecasting Results With SSA
In this section, the two SSA approaches, namely Alexandrov and Alternative SSA, are compared for the Covid-19 movement in 12 countries with the largest Muslim population. Comparisons were made using the actual MAPE for six days along with the average MAPE. The first day will begin on May 15th, 2020 and the sixth day on May 21st, 2020.

Based on Table 1, the three most suitable countries to be modeled by Alexandrov SSA are Malaysia (MAPE = 0.24%), Egypt (MAPE = 0.58%) and Morocco (MAPE = 0.71%). The three countries with the highest MAPE are Oman (MAPE = 0.59%), Pakistan (MAPE = 2.72%) and Kuwait (MAPE = 2.17%). The smallest daily MAPE (MAPE = 0%) occurred in Malaysia at day-1, while the highest MAPE (MAPE = 5.56%) occurred in Oman at day-4.
Table 1. MAPE values for Alexandrov SSA models.

| Negara         | MAPE SSA Alexandrov (%) | Average |
|----------------|--------------------------|---------|
|                | Day-1 | Day-2 | Day-3 | Day-4 | Day-5 | Day-6 |       |
| Indonesia      | 1.42  | 0.56  | 1.92  | 0.29  | 0.17  | 0.44  | 0.80  |
| Malaysia       | 0.00  | 0.31  | 0.25  | 0.41  | 0.36  | 0.10  | 0.24  |
| Bangladesh     | 0.75  | 1.47  | 0.24  | 2.01  | 1.51  | 0.40  | 1.06  |
| India          | 0.34  | 1.21  | 0.20  | 1.24  | 0.90  | 0.88  | 0.80  |
| Saudi Arabia   | 0.55  | 1.18  | 0.28  | 0.89  | 0.77  | 0.79  | 0.74  |
| Egypt          | 0.20  | 0.66  | 0.26  | 0.69  | 1.25  | 0.39  | 0.58  |
| Kuwait         | 3.56  | 0.89  | 3.09  | 1.94  | 1.84  | 1.72  | 2.17  |
| Morocco        | 0.72  | 1.10  | 1.22  | 0.46  | 0.53  | 0.25  | 0.71  |
| Afghanistan    | 1.59  | 0.84  | 1.74  | 0.23  | 1.71  | 0.36  | 1.08  |
| Pakistan       | 1.12  | 4.42  | 4.13  | 4.51  | 0.61  | 1.51  | 2.72  |
| Oman           | 2.79  | 0.56  | 4.94  | 5.56  | 3.65  | 4.01  | 3.59  |
| Qatar          | 1.58  | 0.80  | 1.43  | 1.06  | 1.29  | 0.01  | 1.03  |

Based on Table 2, Malaysia (MAPE = 0.21), India (MAPE = 0.41%) and Morocco (MAPE = 0.69%) are the three most appropriate countries modeled on this Alternative SSA. The three countries with the highest MAPE are OMAN (MAPE = 3.60), Pakistan (MAPE = 3.26%) and Kuwait (MAPE = 1.65%). The smallest daily MAPE (MAPE = 0.01%) occurred in Malaysia at day-1, while the largest MAPE (MAPE = 5.29%) occurred in Pakistan at day-5.

Table 2. MAPE values for alternative SSA models.

| Negara         | MAPE SSA Alternative (%) | Average |
|----------------|--------------------------|---------|
|                | Day-1 | Day-2 | Day-3 | Day-4 | Day-5 | Day-6 |       |
| Indonesia      | 1.36  | 0.09  | 1.83  | 0.77  | 0.29  | 0.30  | 0.77  |
| Malaysia       | 0.01  | 0.26  | 0.15  | 0.42  | 0.24  | 0.16  | 0.21  |
| Bangladesh     | 1.38  | 1.05  | 0.23  | 1.02  | 1.78  | 0.61  | 1.01  |
| India          | 0.41  | 0.46  | 0.20  | 1.00  | 0.37  | 0.04  | 0.41  |
| Saudi Arabia   | 0.51  | 1.21  | 0.23  | 1.00  | 0.67  | 0.76  | 0.73  |
| Egypt          | 0.97  | 0.59  | 0.80  | 0.89  | 0.66  | 0.53  | 0.74  |
| Kuwait         | 3.58  | 0.79  | 2.98  | 0.01  | 2.11  | 0.41  | 1.65  |
| Morocco        | 0.72  | 0.8   | 1.30  | 0.09  | 0.54  | 0.67  | 0.69  |
| Afghanistan    | 0.17  | 0.73  | 1.68  | 0.31  | 2.29  | 0.86  | 1.01  |
| Pakistan       | 0.44  | 6.04  | 0.59  | 5.5   | 5.29  | 1.7   | 3.26  |
| Oman           | 2.81  | 0.50  | 5.21  | 5.02  | 3.37  | 4.67  | 3.60  |
| Qatar          | 1.76  | 0.69  | 1.11  | 0.69  | 0.67  | 0.40  | 0.89  |

From Table 3, they are comparing the total MAPE of the two Alexandrov SSA methods and Alternative SSA for the 12 countries with large Muslim populations. The asterisk (*) is the smallest MAPE value compared to the MAPE value of the form next to it. From the table above, it appears that the Nine Countries (Indonesia, Malaysia, Bangladesh, India, Saudi Arabia, Kuwait, Morocco, Afghanistan and Qatar) are better suited with the Alternative SSA system than the Alexandrov SSA.
The MAPE values modelled by the SSA Alternatives smaller than the Alexandrov MAPE grades. The better countries to use the Alexandrov SSA Method are Egypt, Pakistan and Oman because they have a smaller MAPE value compared to the MAPE value produced by the Alternative SSA Method.

Table 3. Comparison of MAPE both methods.

| Country    | SSA Alexandrov | SSA Alternative |
|------------|----------------|-----------------|
| Indonesia  | 0.80           | 0.77*           |
| Malaysia   | 0.24           | 0.21*           |
| Bangladesh | 1.06           | 1.01*           |
| India      | 0.80           | 0.41*           |
| Saudi Arabia | 0.74         | 0.73*           |
| Egypt      | 0.58*          | 0.74           |
| Kuwait     | 2.17           | 1.65*           |
| Morocco    | 0.71           | 0.69*           |
| Afghanistan | 1.08          | 1.01*           |
| Pakistan   | 2.72*          | 3.26           |
| Oman       | 3.59*          | 3.60           |
| Qatar      | 1.03           | 0.89*           |

4. Conclusion
Based on descriptive analysis, the country which experienced a decline when Ramadhan was Morocco, while Malaysia rose again when Ramadhan was Ramadhan. Many ten countries (Indonesia, India, Bangladesh, Kuwait, Egypt, Saudi Arabia, Oman, Qatar, Pakistan and Afghanistan) are experiencing a growing pattern in addition to Covid-19 patients both before and during Ramadhan.

For the 12 countries based on the SSA model, nine countries (75%) are better off using the Alternative SSA method than the Alexandrov SSA method. The three most relevant countries are Malaysia (MAPE = 0.21%), India (MAPE = 0.41%) and Morocco (MAPE = 0.69%). In general, the two SSA models (both Alexandrov and Alternative) are still suitable for modelling in the 12 countries, as can be seen from the average MAPE of the 12 countries, which is still below 5%.

Acknowledgments
The authors thank Rector, the Directorate of Research, Community Service, and Innovation (DRPMI) of Padjadjaran University who provided financial assistance for the cost of publishing this article and supported the research through Academic Leadership Grant 2020 with contract number: 1427/UN6.3.1/LT/2020. This support is given in the context of increasing the quantity of research and publications for lecturers at Padjadjaran University.

References
[1] Roosa K, Yee L, Luo R, Kirpich A, Rothernberg R, Hyman J M, Yan P and Chowell G 2020 Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th 2020, *Infectious Disease Modelling* 5 256–263
[2] Linton N 2020 Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data *Journal of Clinical Medicine* 9 2 538
[3] He Zand, Tao, H 2018 Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan , China : A nine-year retrospective study,” *International Journal.*
Infection Diseases 74 61–70

[4] Benvenuto D, Giovanetti M, Vassallo L, Angeletti S and Ciccozzi M 2020 Application of the ARIMA model on the COVID-2019 epidemic dataset Data in Brief 29 105340

[5] Xu H, et al 2020 Possible environmental effects on the spread of COVID-19 in China Science Total Environment 731 139211

[6] Shaw R, Kim Y, and Hua J 2020 Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia Progress in. Disaster Science 6 100090.

[7] Guliyev H 2020 Determining the spatial effects of COVID-19 using the spatial panel data model Spatial Statistics 38 100443

[8] Alexandrov T 2009 A Method of Trend Extraction Using Singular Spectrum Analysis arXiv 3 1–23

[9] Alexandrov T and Golyandina N 2005 Automatic extraction and forecast of time series cyclic components within the framework of SS in Proceedings of the Fifth Workshop on Simulation 45–50

[10] Golyandina N, Nekrutkin V and Zhigljavsky A 2001 Analysis of Time Series Structure SSA and Related Techniques (London: hapman & Hall/CRC)

[11] Hassani H and Mahmoudvand R 2018 Singular Spectrum Analysis Using R (London: Palgrave Macmillan UK)