Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2023, 2023:21
https://doi.org/10.28919/cmbn/7882
ISSN: 2052-2541

EFFECT OF LONG VACATION ON DAILY CASES OF COVID-19 DURING PARTIAL RESTRICTION IN JAKARTA, INDONESIA

R. MAHARESI[1], I. KUSWANDARI[2,*], C.W. CHUKWU[3]

[1]Department of Information Technology, Faculty of Industrial Technology, Universitas Gunadarma, Indonesia

[2]Department of Mathematics, Faculty of Sciences and Technology, Universitas Airlangga, Indonesia

[3]Department of Mathematics Wake Forest University, Winston-Salem, NC 27109, USA

Abstract: Since the first case of COVID-19 was discovered in December 2020, the number of transmissions due to COVID-19 infection in the capital city of Jakarta has never decreased in the long term. The decrease in the number of daily positive cases through restrictions on social interaction implemented by the Government faced many obstacles during long holidays, hence it did not last long. In this research paper, we investigate the effect of long vocation on daily cases of COVID-19 by considering few factors including the level of surveillance through contact tracing and testing during social restrictions which are identified to be hampered by long holidays. Hence, based on the number of testing aspects in the time of the outbreak and serials of long holidays, a structural model was built to explain how these variables affect daily number of positive cases. The data for analysis taken at the first social restrictions period up to end of 2020 showed that daily positive cases consistently increasing. Time series structural model of the two-step approach shows that the time and number of testing variables significantly affect the number of daily positive cases, but long holidays did not directly affect the number of positive cases. In a sense, the effect of long holidays

---

[*]Corresponding author
E-mail address: inna-k@fst.unair.ac.id
Received January 09, 2023

results in an increase in the quantity of the number of tests, which in turn will affect the number of daily positive cases. The forecast function for daily new cases obtained by the model performs relatively well in terms of the mean value of the squared prediction error, which gives a positivity rate around 7%.

Keywords: social restriction; long vacation; daily cases; Jakarta province.

2020 AMS Subject Classification: 62P30.

1. INTRODUCTION

Corona Virus (COVID-19) is an infectious disease caused by Severe Acute Respiratory Syndrome Corona virus 2 (SARS-CoV-2). Humans may be infected without any symptoms while in some cases it causes severe symptoms including pneumonia, acute respiratory syndrome, kidney failure. In Indonesia, more than ten of thousands of people have been infected and more than thousands have been died from this virus up until end of 2020 (Ministry of Health RI, 2020). Due to its fast spread, World Health Organization (WHO) officially declared the COVID-19 as a pandemic on March 11, 2020 [1]. A reasonable way to deal with this pandemic is to know the best practice to implement such as correct health policies [2]. Therefore, a theoretical prediction of the virus spreading is important to make an effective health policy. Effective reproduction number (R_t) whose value describes how many persons can be infected by one who is already infected by the disease. As reported by [3], R_t define value are determined by four main factors: the duration of infection, frequency of contacts among persons within the population, number of vaccinated populations, and the implementation of health protocols to avoid the transmission within and around the environment. The spread of COVID-19 affected almost all areas in Indonesia's provinces, including capital city of Jakarta. Therefore, many studies have been conducted and reported on various aspects related to the pandemic. For instance, many questions have been raised including "how to fulfill the demand of public health facilities i.e., hospital and drug stocks required to cure the patients during the peak period of active cases?". This question may be used to address other questions, "when the pandemic will be over and how many people will be possibly infected?" Most answers to that question raised are based on a system of differential or difference equations for describing an epidemic condition, namely of the (susceptible (S), Infected (I), Recovered (R) = SIR) model. An elaboration of a SIR models give the output for the solutions in terms of a prediction of when and how many infected population would exist at the peak of the pandemic.

The system of differential equations, may be turn into a system of difference equations used for obtaining an effective number formulation, see [4] and [5] that solved numerically the differential equations system for simulating the sizes of population S, I and R under some predefined assumptions. Recently, [6] worked with an analytical solution to formulate an effective reproduction number under some imposed conditions. Since the solution for a differential equation system fall into the families of the growth function such as logistic functions [7], the model parameterization may be obtained using nonlinear estimation, which is available in some statistical packages, see [8]. In another work, [9] proposed ARIMA model for prediction the number of COVID-19 positive cases. The application of an optical sensor with surface plasmon resonance was shown to have high sensitivity and excellent detection limits to further reduce the spread of the disease [10].

Handling the COVID-19 pandemic in an attempt for reducing the frequencies of contacts among people in Jakarta was done through the implementation of great social restriction (PSBB). Study in [11] shows that restriction for at least 50% from the total contacts between individual within a certain region was able to suppress the pandemic. Contact tracing for COVID-19 involves three processes namely, identifying, assessing, and managing people who already infected to prevent wider spread in order to break the chains of transmission of this infectious disease. The identified persons who may have been exposed to COVID-19 through the appropriate test will be supervised up daily for 14 days within the region where last time they found [12]. As WHO noted that one of the critical elements of the implementation of contact tracing is a real-time data analysis. The data analysis should include some potentially factors that have significant impacts either directly or indirectly towards the correct information needed to tackle the pandemic. For the direct impact example is an evaluation testing and contact tracing, see [13]. While for the indirect impact, e.g. to do evaluation of government's interventions see [14], [15] to break the chains of human-to-human transmission for ensuring that the number of new cases generated by each confirmed maintained below.

Different to many countries which already implemented fully level of social restriction through a lock down policy to the area being isolated, Indonesia applies some partials of social restrictions due to the level of emergency. The social restriction sometimes must be rearranged due to some moments of vacation which are considered as routine traditions. As an example a long periods of vacation is during Christmas up to New Year occasion. When people usually go to visit and gather their families. The people mobility may hence increase unexpected mobility and crowds of people during the pandemic. In this occasion based on the number of cases and tests each day together with the length day of vacations, we wish to test whether the daily number of cases may be positively affected by some serial's moments of vacations along end of year time during the social restriction periods. For this purpose, a structural model for time series data is chosen as our model. By implementing a structural model, at least three benefits to gain, firstly, daily fluctuations number of positive cases in the capital city of Jakarta may be described using more than one nonlinear functional forms e.g. logarithmic function to describe the trend components according to the effects of the number of tests as well as long holiday moments. Secondly, the use transformed logarithmic form into the linear functional provides an interval confidence for the estimate. Thirdly, as already observed by [16], that in India country whose demographic characteristics are not much different to Indonesia's, the use of model formulation based on SIR nor simulation were not good enough to predict the figure of the daily cumulative COVID-19 positive cases for certain periods of time.

2. METHODS

This study used secondary data from the website of the capital city of Jakarta (DKI Jakarta) Provincial Government's Covid-19 Task Force Provide Citation. The time series data was taken starting from 10th April 2020 to 20th December 2021 after the implementation of the first large-scale social restrictions. The data is arranged in the form of a table containing data of daily date, number of people tested per day, number of positive cases per day. At the early stage, the collected purposive data sampling was processed with MS Excel to yield the required the data modeling. The processed data then were imported to a statistical program to perform model parameters estimation. The estimation of the structural model parameters consists of two steps followed by statistical testing for the feasibility of the model to be used as a predictor for the number of daily new positive cases.

The training data considered was obtained from April 10 to December 21, 2020 while the sample of predicted data was calculated starting from December 22, 2020. The level of accuracy was evaluated using Mean Absolute Percent Error (MAPE). The structural model in the study is illustrated by Figure 1.

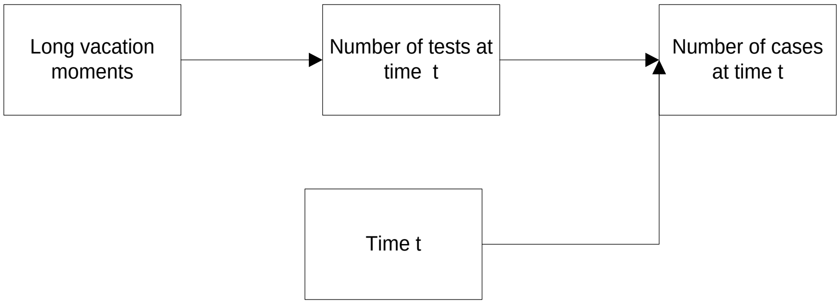

Figure 1. Two-stages time series data structural model

The historical data revealed for any long vacation periods during the pandemic, mobility people for meeting each other tend to increase. Consequently, the more risk of human to human virus transmissions due to the coincidence of long vacation and viral incubation periods. To find out the effect of long holidays on the number of daily cases, the model constructed, must stated the relationship between the number of daily tests as a respond variable and a predictor variable measuring the effect of long vacation D_L. For this intention, we used binary dummy variable for representing all of regular dates (assigned value = 0) and vacation dates (assigned value = 1) in our collected data sample of size $n = 261$ containing $m = 8$ vacation dates. To take account the risk of holiday in the pandemic situation towards the respond variable, the dummy variable then were accumulated and this value later on be added with *1*, to take account of the incubation period, as given by Equation (1):

$$D_L = \sum_{t=1}^{m} D_{Lt} + 1, L = 1, 2, 3, \ldots, 261; \; t = 1, 2, 3, \ldots, m \; D_{Lt} = \begin{cases} 1, & L = \text{holiday date} \\ 0, & L \neq \text{holiday date} \end{cases} \quad (1)$$

Since D_L value must represents a risk degree while maintaining the consistency of its dimensions with the number of tested persons per day so the dimension must be in unit day. The unit time is taken to ease the interpretation of the analysis results. Besides it seems to be a common sense, the longer the exposure time is the bigger infectious chance to happen. And also, the number of people tested per day is a kind of a controlled variable as its value may be adjusted or determined by the regulator according to a certain degree of emergency in handling the pandemic. In the literature of SIR model, [7] stated that number of positive cases per unit time and the infectious period changes exponentially with a mean one over the recovery rate.

$$J_t = 1.2^{\alpha_0 + \alpha_1 D_L + \varepsilon_t};$$

$$y_t = 1.2^{\beta_0 + \beta_1 t + \beta_2 J_t + \varepsilon_{1t}} \quad (2)$$

Applying the logarithmic of based 1.2 for the daily number of persons tested J_t to yield J_{1t} hence a linear regression model describing relationship between the dependent variable J_t and the independent variable D_L written in the Equation (3)

$$J_{1t} = \log_{1.2} J_t = \alpha_0 + \alpha_1 D_L + \varepsilon_t \quad (3)$$

Constant α_0 denotes the number of persons being tested per day. Coefficient α_1 measures the effect of vacation moments on the number of persons being tested per day. For making the model fitted result easily interpreted, the dimension for this coefficient should be persons per day^2. This represents the acceleration for number of persons being tested due to a unit risk of long vacation D_L. The random variable ε_t represents error model of stage one which is assumed to be $n \sim (\mu = 0, \sigma = \text{constant})$. By applying the logarithm of 1.2 as a base, the second equation of the model is obtained as follows:

$$y_{1t} = \log_{1.2} y_t = \beta_0 + \beta_1 t + \beta_2 J_t + \varepsilon_{1t} \quad (4)$$

The term y_{1t} is in a base 1.2 logarithm of the number of new positive cases per day. The term β_0 is a constant denoting the number of new positive cases per day. Hence the dimension is persons per day. The trend of new positive cases β_1 with periods *t* rescaled to *t*/4 has dimension persons per day^2, has a meaning an acceleration caused by duration time *t*. For the purpose of interpretation, dimension less coefficient β_2 measures the effect of the number of people tested per day to the daily number of new positive cases during the long vacation moments which usually considered as positivity rate. The random variable ε_t^1 represents error model of stage two which is assumed to be $n \sim (\mu = 0, \sigma = \text{constant})$.

3. RESULTS AND DISCUSSION

Parameter estimation for the structural model is obtained by exchanging the random error ε_t with the estimated error e_t as given in equations (5) below:

$$J_t = 1.2^{a_0+a_1D_L+e_t}$$

$$y_t = 1.2^{b_0+b_1t+b_2J_t+e_{1t}} \qquad (5)$$

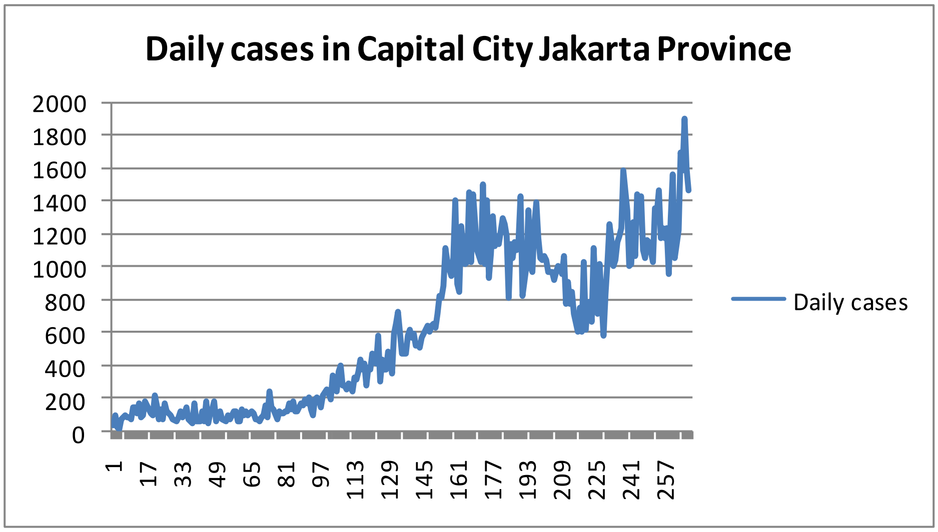

Figure 2. Time series data plot of daily new positive cases in DKI Jakarta province.
Source: author's research data.

The plot of the number of new COVID-19 cases against time on Figure 2 allows for assuming the exponential functional for the data, in this study, the base value of the logarithm = 1.2. The seemingly plot of daily new cases showed an exponential growth also met in [9] who used a logarithmic transformation to eliminate seasonal and cyclical variations in the data used for their

ARIMA model. Determination of the basis of logarithms is done by trial and error with some criteria, including the graph visually does not cut the horizontal axis which represents time and shows exponential curvature as well the descriptive statistics which presents significant positive correlation between the number of new positive cases per day with time and the number of people tested per day. Hence following [7] in this study the time t is multiplied by 1/4 to represent the recovery rate of 4 days. The classical assumptions must fulfill namely the normal distribution of the residuals, non-co linearity of the independent variables, the best fit criteria and the significance of the influence of the independent variables to the dependent variable.

The four variables, namely: i) Time (t): In this data, the count of time begins at the same time as the first Large-Scale Social Restrictions (PSBB) in Jakarta, from April 10, and ended to 2020 to December 21, 2020. So the values of t inputted into the model is $t_1 = t/4$, where $t = 1$ to $t = 261$. ii) Number of new positive cases per day (y_t) considered to swab tests with positive results amongst all people tested per day. In the data table, y_t values transformed into base 1.2 logarithm to use as y_{1t} values. iii) The number of tests (J_t) is the number of people included in the contact tracing who underwent a swab test, with positive or negative results and J_{1t} are its corresponding base 1.2 logarithm values. iv) Holiday moments (D_L) were calculated based on Equation (1), which in the output SPSS denoted by D_L'.

The following is a list of long holidays, with duration of at least three days considered in this paper: April 10 (Friday), May 1 (Friday), May 7 (Thursday), May 21 (Thursday), May 22 (Friday), 24-25 May (Sunday & Monday), 26-29 May (Tuesday to Friday), June 1 (Monday), July 31 (Friday), August 17 (Monday), August 20 (Thursday), August 21 (Friday), October 29 (Thursday), October 30 (Friday). Meanwhile, data on holiday moments that were not included in the research sample were: Birthday of the Prophet Muhammad SAW 24 December (Thursday) followed by Christmas Holiday 25 December (Friday), with a total time of 11 days. Table 1 presents data on long holiday moments used in this study.

Table 1. Variables date involved in the data of long holiday moments $D_L= D_L'$ (Source: author)

Date	t	D_L'	Date	t	D_L'	Date	t	D_L'
4/10/2020	15	1	5/23/2020	58	4	8/1/2020	128	5
4/11/2020	16	2	5/24/2020	59	4	8/2/2020	129	5
4/12/2020	17	2	5/25/2020	60	4		147	…
	36	..	5/26/2020	61	4	8/20/2020		…
5/1/2020		..	5/27/2020	62	4	8/21/2020	148	6
5/2/2020	37	3	5/28/2020	63	4	8/22/2020	149	6
5/3/2020	38	3	5/29/2020	64	4	8/23/2020	150	6
	56	..	5/30/2020	65	4		217	...
5/21/2020		…	5/31/2020	66	4	10/29/2020		..
5/22/2020	57	4	6/1/2020	67	4	10/30/2020	218	7
			…	….	..	10/31/2020	219	7
			7/31/2020	127	4	11/1/2020	220	7

As parametric inferential statistical method, regression analysis requires normality assumption of residuals, i.e., the difference between the value of the dependent variable and the estimated results. The Least Squares method used to obtain the estimates of the model parameters was based on sample data obtained from the URL address https://public.tableau.com/profile/jsc.data#!/vizhome/LandingPageCovid-19Jakarta2/TablePositivityRate. Meanwhile, data on the start and end of the PSBB in DKI Jakarta were obtained from the URL https://ppid.jakarta.go.id/siaran-pers.

Based on Table 3, according to [17] the Kolmogorov Smirnov's statistics test on the residual model showed normality assumption because all values in the sig column indicating the z statistic values were less than 1.97 with significant level of 5%.

Table 2. One-Sample Kolmogorov Smirnov Test of Normality

		Unstandardized Residual
N		261
Normal Parameters[a,,b]	Mean	.0000000
	Std. Deviation	2.25257268
Most Extreme Differences	Absolute	.026
	Positive	.026
	Negative	-.020
Kolmogorov-Smirnov Z		.424
Asymp. Sig. (2-tailed)		.994

Table 3. Estimation results of the two-stage regression model with log 1.2

Stage 1

Model Summary

Model	R	Adjusted R Square	Std. Error of the Estimate
1	.875[a]	.765	2.1752517

a. Predictors: (Constant), DL'

Coefficients[a]

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		83.581	.000
	DL'	.875	29.111	.000

Stage 2

Model Summary

Model	R	Adjusted R Square	Std. Error of the Estimate
1	.920[a]	.845	2.2612867

Coefficients[a]

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		4.193	.000
	j1t	.239	4.600	.000
	t1	.702	13.479	.000

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7265.535	2	3632.768	710.438	.000[a]
	Residual	1319.262	258	5.113		
	Total	8584.797	260			

a. Predictors: (Constant), t1, j1t
b. Dependent Variable: y1t

The SPSS outputs presented in Table 3 depicts the results of regression modeling with two stages. Concerning the classical assumption test of the regression model based on model residual, the test reveal that the first stage model still has not passed the fulfillment of the normal distribution assumption for the residual value, even though the mean value and standard deviation value are close to the normal distribution parameter values if normality assumption the residuals met, this was due to the formation of binary variable to account for variable D_L. But for the second stage model, the residuals normality assumption fulfilled as shown by the Kolmogorov-Smirnov Test, which gives a significant value of more than 0.05, i.e., 0.994. Co-linearity between independent variables is not justified; this is tested using VIF statistics. The independent variables, namely the number of daily testing and time, provide a statistical value of VIF 4.549 (< 5) so that the model can be said to be free from co-linearity problems.

The structural model equations for prediction purposes presented by Equations (6) by first replacing $D_L = D_L'$:

$$\log_{1.2} J_t = a_0 + a_1 D_L \quad \text{for } D_L = 7 + l, \quad l \geq 1$$
$$\log_{1.2} y_t = b_0 + b_1 t + b_2 J_t \tag{6}$$

where l = the counting of vacation moments after December 22, 2020. Note, when using the prediction function, the estimate for the number of positive cases after the Christmas holiday, e.g., December 27, the variable describing of long vacation value can be assigned with $D_L = 8$. The steps below show simplified terms involving in the equations for calculating the effect of changes in each unit independent variable's values namely $x_1 = t/4$ and $x_2 = J_{1t} = {}^{1.2}\log J_t$ on the value of the dependent variable value y_t representing the daily number of new positive cases: If $y_t = 1.2^{b_0 + b_1 x_1 + b_2 x_2}$ and $\Delta y_t = yx_1 + \Delta x_1, x_2 + \Delta x_2) - y_t$, then we obtain $\Delta y_t = (1.2^{b_1 \Delta x_1} 1.2^{b_2 \Delta x_2} - 1) y_t$. The last equation becomes $\Delta y_t = (1.2^{b_1 + b_2} - 1) y_t$ after assigning $\Delta x_1 = \Delta x_2 = 1$.

SPSS software showed all coefficient estimate were significant at a significant level $\alpha < 1\%$. Hence each unit changes in the value of the independent variables significantly contributes to the change for the dependent variable value y_t. The SPSS output in Table 3 showed both the independent variables simultaneously contribute in the variability of daily number of new positive COVID-19 for about cases by 84% for level of significant 1%, and about 16% supplied by other factors. By recalling the independent variables mentioned above, hence the daily predicting new cases given by Equation (7):

$$y_t = 1.2^{(11.392 + 0.0535\, t + .307 J_{1t})} \quad \text{for } J_{1t} = 35.640 + 2.385 D_L. \tag{7}$$

The predicted daily increment case obtained using regression coefficients in Table 4, i.e., gives $\Delta y_t = (1.2^{b_1 + b_2} - 1) y_t$ to yield ${}^{\Delta y_t}/_{y_t} = 0.068 \approx 7\%$ considered as positivity rate estimate since the unit is dimensionless.

Based on the plot of rescaled time corresponding to the daily cases, it was noticeable that a structural change present mainly due to the implementation of large-scale social restrictions, confirmed with a decrease in daily new positive cases within some passages of time. However, due to several factors, such as discipline in implementing social restrictions, public obedience in carrying out health procedures, and also the 11-day long holiday during Christmas up to the New Year, an exponential spike was seen with new parameters greater than the original parameter values. This similar pattern also met in the study by [14] who found nearly 44.06% increase in number of cases due to untimely end of lockdown scenarios in a certain region in India. As agreed with it, [11] reported in their simulation that of at least 50% social restriction were quite effective to decrease the cases. Moreover, as long as the social restrictions were carried out using appropriate procedures, [15] reported that the number of COVID-19 cases in Kuwait can be reduced. They use a modified ARIMA model to evaluate the impact of government's interventions through the partial and total lockdown interventions in Kuwait. The study showed that both interventions were very effective in lowering the number COVID-19 cases in Kuwait. The similar study was also conducted by [18] which integrate some aspects relating with the entities of demographic in the Boston metropolitan. It was found that a combination of a period of strict social distancing and a robust level of testing, contact-tracing and household quarantine could keep the disease within the capacity of the healthcare system while enabling the reopening of economic activities.

Based on the SPSS output in Table 4, where the statistical value of F is significant at the < 1% level of significance, it is possible to determine the predictive model for the periodic data structural model as follows. By applying the base 1.2 logarithm gives the result: $y_t = 1.2\ y_{it} = 1.2^{11.392+0.214t_1+0.307J_{1t}}$ where $J_{1t} = 35.640 + 2.385D_L$. Thus, to make a prediction of the future time involving Christmas and New Year holiday moments, an equation is used to predict the number of people who are tested daily, namely:

$$J_{1t} = {}^{1.2}log\ J_t = 35{,}640 + 2{,}385\ D_L = 54.72 \qquad \text{for } D_{\mathrm{L}} = 8, 9, 10, \ldots.$$

For example, assigning a value of $D_L = 8$ for a long vacation after December 21, to obtain prediction of new positive cases at January 16, 2020, we get $t = 278$ or $t_1 = 278/4 = 69.5$. So, from stage 1, it produces ${}^{1.2}log\ J_{1t} = 54.72$. Based on these results, it is obtained that $y_t = 1.2^{11{,}392+0.214(69.5)+0.307(54.72)} = 2778$, so that the predicted value for the number of new positive cases on January

16, 2021 is 2778 persons. Table 4 shows the predicted value starting on January 8, 2021 to January 19, 2021 along with the prediction error using MAPE, which measures forecast accuracy by concentrating on a rescaled version of a measure.

Table 4. Prediction with logarithmic model

Date	Actual number of Positive cases	Predicted value	Percentage error of predicted value
8/1/2021	2774	2570	0.0795484
9/1/2021	2892	2595	0.1145454
10/1/2021	2346	2620	0.1046532
11/1/2021	2869	2646	0.0843204
12/1/2021	3088	2856	0.0811354
13/1/2021	3214	2884	0.1143266
14/1/2021	2509	2913	0.138548
15/1/2021	3217	2751	0.1693199
16/1/2021	2559	2778	0.0788797
17/1/2021	2361	2805	0.1583997
18/1/2021	2563	2833	0.0952631
19/1/2021	2901	3058	0.0513693
MAPE			1.2703091

Table 4 provides prediction results using the value of the D_L variable, namely the moment of long holiday with a value range of 8 - 8.5, considering that according to epidemiologists, the severity of outbreak transmission follows a normal distribution, in this case the data of the moment of long vacation in the period from 8 to 19 January 2021 were used. The total difference in the prediction value and the actual values MAPE is relatively less than 1.3%. As stated by [19], a forecast can be categorized accurate if MAPE is less than 5%.

The COVID-19 pandemic has resulted in many adverse outcomes and challenges including problem involved in the testing which may subjects to false positive or false negative apart from the problems involved in contact tracing managements as well as application for social restriction to the government policy and individual commitment as a member of the community to obey the health protocol in order to finish the pandemic. The false-positive COVID-19 results have caused significant consequences on both the individual and the society level. The misleading COVID-19 diagnosis can result in significant consequences such as delaying surgeries, unnecessary quarantine and treatments, transplant lists omission, and unnecessary sick leaves as stated by [20], [21]. Despite of these issues which may lead to substantial impact, there is a scarcity in the literature of its prevalence or impact, as admitted by [21], hence those issues must be considered to our result and then based on the gathered information relevant to our research papers we make some comparisons towards our findings as follows.

Firstly, [14] found doubling testing levels may results in identifying more cases about 13.84% in part of India. Also stated by [13], good and proper contact tracing system prior to the testing procedures were associated with a reduction in subsequent new infections of 63%. Those findings are critical elements toward the implementation of contact tracing in order to be able to work with a real-time data analysis which supplies valid information for monitoring the pandemic.

Secondly, according to [20], false positive results in the low prevalence level can have several adverse effects as already mentioned above, a particular one is a risk of subsequent increased exposure due to the believing that having been infected and being placed with other patients with COVID-19 and to more exposed to the virus, which may bring to more severe health condition. Beside this negative consequently individually, the reported daily positive case results in bias information caused by unnecessarily addition of positive cases. Thirdly, [22], stated that 58% of COVID-19 patients may have initial false-negative RT-PCR, this figure was obtained based on their 95% CI finding for the overall false-negative rate 0.12 (0.10 to 0.14). To overcome this problem, [23] suggested a screening strategy in order to lower false-negative rate using group Mix with less than six members. In which it was claimed to be up to 63% saving in the number of tests compared to individual testing. By combining the two above mentioned methods, it is worth to expect the amount of tests can be increased more to reduce false-negative rate up to 58%. Fourthly, for the question of what will happen if large numbers of metropolitan people visit a less populated province during the long vacation? As qualitatively argued by [24] that minimizing events gathering both in the province and metropolitan groups and/or reducing the number of short-term visitors could substantially decrease spreading as could measure to lower the fraction initially infectious upon arrival. So based it is understandable during the long vacation moments the government control towards the citizen mobility still being relevant.

In the absence of quantitative measure toward the social restrictions at the time of the collected data being analyzed, a prediction function for the daily cases which measures the effect of long vacation moments through the number of daily testing done in the end of year of 2020 through the early of year 2021 was obtained. If we bring the result of [14] that doubling testing levels may results in identifying more cases about 13.84%, in terms of our predicted model increment as stated in Equation (7), doubling the number of testing may increase the positivity rate for about 12.9%, which is almost less than *2* times greater apart from the possibility of both negative and or positive false in the reported data as well as the mathematical functional being used.

CONCLUSION

The results of secondary data analysis using a two-step structural time series model with a base 1.2 exponential function formulation were successfully reported. They have successfully provided answers to the research objectives. For instance, the first hypothesis for the number of daily tests, which affects the number of new positive cases per day is significant at a significant level of 5%. It indicates that the moment of long holiday indirectly affects the number of new positive cases per day. Effect of long vacation moment has a direct impact on increasing the number of daily testing. The evaluation performance results of the prediction model for the number of daily positive cases in DKI Jakarta showed good results by using a two-step nonlinear structural time series model for the mean value of the squared prediction error. Cases also showed an estimated positivity rate of about 7%.

ACKNOWLEDGMENTS

The authors acknowledge Professor S. Wadi Harun and Professor Moh. Yasin for some valuable discussions and supports during the preparation of this manuscript. This work was supported by Universitas Airlangga Grant (No. 1913/UN3.1.8/LT/2018).

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] A. Qazi, J. Qazi, K. Naseer, et al. Analyzing situational awareness through public opinion to predict adoption of social distancing amid pandemic COVID-19, J. Med. Virol. 92 (2020), 849-855. https://doi.org/10.1002/jmv.25840.

[2] I. Hossain, A.R. Mullick, M.H. Khan, et al. Epidemiology of corona virus disease: past, present, future prospects and its journey towards Bangladesh, Epidemiology, 25 (2020), 2517-2529.

[3] H. Weiss, The SIR model and the foundations of public health, MATerials MATemàtics, 17 (2013), 1887-1097. https://ddd.uab.cat/record/108432.

[4] H. Susanto, V.R. Tjahjono, A. Hasan, et al. How many can you infect? Simple (and naive) methods of estimating the reproduction number, Commun. Biomath. Sci. 3 (2020), 28–36. https://doi.org/10.5614/cbms.2020.3.1.4.

[5] B.A. Wahyudi, I. Palupi, Prediction of the peak Covid-19 pandemic in Indonesia using SIR model, J. Teknol. Dan Sist. Komput. 9 (2020), 49–55. https://doi.org/10.14710/jtsiskom.2020.13877.

[6] A. Fitriyah, Rumus Bilangan Reproduksi Dasar Covid-19 dengan Adanya Vaksinasi Dosis 1 dan 2, Limits: J. Math. Appl. 19 (2022), 87-100. https://doi.org/10.12962/limits.v19i1.12249.

[7] J. Ma, Estimating epidemic exponential growth rate and basic reproduction number, Infect. Dis. Model. 5 (2020), 129–141. https://doi.org/10.1016/j.idm.2019.12.009.

[8] F.R. Pratikto, Prediksi Akhir Pandemi COVID-19 di Indonesia dengan Simulasi Berbasis Model Pertumbuhan Parametrik, Jurnal Rekayasa Sistem Industri, 9 (2020) 63–68. https://doi.org/10.26593/jrsi.v9i2.4018.63-68.

[9] D. Benvenuto, M. Giovanetti, L. Vassallo, et al. Application of the ARIMA model on the COVID-2019 epidemic dataset, Data Brief. 29 (2020), 105340. https://doi.org/10.1016/j.dib.2020.105340.

[10] H.A. Zain, M. Batumalay, H.R.A. Rahim, et al. Surface plasmon resonance optical sensor for COVID-19 detection, Nanosystems: Phys. Chem. Math. 12 (2021) 575–582. https://doi.org/10.17586/2220-8054-2021-12-5-575-582.

[11] H.N. Fadhilah, E. Apriliani, D.K. Arif, Prediksi Penyebaran Covid-19 di Indonesia dan Jawa Timur dengan

Metode Extended Kalman Filter, Limits: J. Math. Appl. 18 (2021) 1. https://doi.org/10.12962/limits.v18i1.6933.

[12] World Health Organization, Contact tracing in the context of COVID-19: interim guidance, 10 May 2020. World Health Organization. https://apps.who.int/iris/handle/10665/332049. (2020).

[13] T. Fetzer, T. Graeber, Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment, Proc. Natl. Acad. Sci. U.S.A. 118 (2021), e2100814118. https://doi.org/10.1073/pnas.2100814118.

[14] A. Dixit, S. Vishnoi, S. Bikas Paul, Adding structure to statistics: A study on COVID-19 dynamics in India, medRxiv, 2020. https://doi.org/10.1101/2020.05.26.20113522.

[15] S.S. BuHamra, J.A. Dallal, Exploring the impact of government interventions on COVID-19 pandemic spread in Kuwait, Int. J. Healthcare Inform. Syst. Inform. 16 (2021), 93. https://doi.org/10.4018/ijhisi.288893.

[16] S. Ghosh, Predictive model with analysis of the initial spread of COVID-19 in India, Int. J. Med. Inform. 143 (2020), 104262. https://doi.org/10.1016/j.ijmedinf.2020.104262.

[17] R.E. Walpole, R.H. Myers, S.L. Myers, Probability & statistics for engineers & scientists, 9th Edition, Prentice Hall, 2012.

[18] A. Aleta, D. Martín-Corral, A. Pastore y Piontti, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, Nat. Human Behav. 4 (2020) 964–971. https://doi.org/10.1038/s41562-020-0931-9.

[19] D. Swanson, On the relationship among values of the same summary measure of error when it is used across multiple characteristics at the same point in time: An examination of MALPE and MAPE, Rev. Econ. Finance, 5 (2015), 1–14.

[20] B. Healy, A. Khan, H. Metezai, et al. The impact of false positive COVID-19 results in an area of low prevalence, Clin. Med. 21 (2020), e54–e56. https://doi.org/10.7861/clinmed.2020-0839.

[21] S.H. Alsheikh, K. Ibrahim, D. AlFaraj, The impact of false positive COVID-19 result, Cureus. 13 (2021), e20375. https://doi.org/10.7759/cureus.20375.

[22] V. Pecoraro, A. Negro, T. Pirotti, et al. Estimate false-negative RT-PCR rates for SARS-CoV-2. A systematic review and meta-analysis, Eur. J. Clin. Investigation. 52 (2021), e13706. https://doi.org/10.1111/eci.13706.

[23] A.R. Alizad-Rahvar, S. Vafadar, M. Totonchi, et al. False negative mitigation in group testing for COVID-19 screening, Front. Med. 8 (2021), 661277. https://doi.org/10.3389/fmed.2021.661277.

[24] T. Britton, F. Ball, Summer vacation and COVID-19: effects of metropolitan people going to summer provinces, arXiv. (2020). https://doi.org/10.48550/ARXIV.2006.00579.