OPEN ACCESS
Indo-JC
Indonesia Journal of Computing
ISSN 2460-9056
socj.telkomuniversity.ac.id/indojc

Ind. Journal on Computing
Vol. 6, Issue. 2, September 2021. pp. 43-52
doi:10.34818/indojc.2021.6.2.576

# An Exponential Smoothing Holt-Winters Based-Approach for Estimating Extreme Values of Covid-19 Cases

Abi Rafdhi Hernandy #1, Aniq A. Rohmawati #2, P. H. Gunawan #3

*#School of Computing, Telkom University*
*Jl. Telekomunikasi 1, Terusan Buah Batu, Bandung, Indonesia*

[1] abirafdhiher@student.telkomuniversity.ac.id,
[2] aniqatiqi@ telkomuniversity.ac.id
[3] phgunawan@telkomuniversity.ac.id

**Abstract**

Covid-19 is an ongoing outbreak across the world infecting millions, having significant fatality rate, and triggering economic disruption on a large scale. The demand of healthcare facility has been significantly affected by the increased Covid-19 cases. Many countries have been forced to do lockdown and physical distancing to avoid a crucial peak of novel Covid-19 pandemic that potentially overwhelms healthcare services. Central Java is the province with the third highest population density in Indonesia and predicted to be affected significantly over a particular period of this outbreak. Our paper aims to provide a modelling to estimate extreme values of daily Covid-19 cases in Central Java, between March and April 2021. We particularly capture seasonality during this period using Exponential Smoothing Holt-Winters. We employ that Value at Risk and mean excess function based-approaches for extreme value estimation. Our simulation studies indicate that Exponential Smoothing Holt-Winters and Value at Risk provide sharp and well prediction for extreme value with zero violation. Since a number of positive cases has resulted unprecedented volatility, estimating the extreme value of daily Covid-19 cases become a crucial matter to support maintain essential health services.

**Keywords:** Covid-19, estimation, extreme value, mean excess, seasonality, value at risk

**Abstrak**

Covid-19 adalah wabah yang sedang terjadi di seluruh dunia, hingga saat ini telah menginfeksi jutaan orang, dengan tingkat kematian signifikan, dan memicu guncangan perekonimian dalam skala besar. Kebutuhan masyarakat terhadap fasilitas kesehatan dipengaruhi secara signifikan oleh peningkatan kasus Covid-19. Banyak negara melakukan *lockdown* dan *physical distancing* untuk mengantisipasi lonjakan kasus pandemi Covid-19 yang berpotensi membuat layanan kesehatan *overcapacity*. Jawa tengah merupakan provinsi dengan tingkat kepadatan penduduk tertinggi ketiga di Indonesia dan diperkirakan akan terkena dampak signifikan selama kurun waktu gelombang Covid-19 terjadi. Penelitian ini bertujuan untuk membuat model untuk memperkirakan nilai ektrim (lonjakan) kasus harian Covid-19 di Jawa Tengah, yang terjadi pada periode Maret dan April 2021, kami secara khusus menangkap pola musiman selama periode ini menggunakan *Exponential Smoothing Holt-Winters*. Kami menggunakan pendekatan *Value at Risk* dan *Mean Excess Function* untuk memperkirakan nilai ekstrim lonjakan kasus di periode tersebut. Berdasarkan simulasi, diperoleh bahwa nilai ekstrim berbasis *Value at Risk* berdasarkan model *Exponential Smoothing Holt-Winters* memberikan hasil estimasi yang baik dengan nilai pelanggaaran 0. Pergerakan jumlah harian kasus positif menunjukkan fluktuasi signifikan yang belum pernah terjadi sebelumnya, sehingga memperkirakan nilai ekstrim kasus Covid-19 harian sangat penting untuk mengantisipasi *overcapacity* fasilitas kesehatan saat terjadi lonjakan kasus Covid-19.

**Kata Kunci:** Covid-19, estimasi, nilai ekstrim, *mean excess*, musiman, *value at risk*

Received on Mei, 2021. Accepted on September, 2021

## I. Introduction

COVID-19 (Corona Virus Disease 2019) is a new type of coronavirus (SARS-CoV-2) that was first discovered in the city of Wuhan, China at the end of December 2019 [1]. This virus attacks the respiratory tract and causes respiratory infections. On March 2, 2020, the first case of Covid-19 was announced in Indonesia with 2 people being tested positive for the Covid-19 virus [2]. The latest update on June 24, 2021, Covid-19 cases in Indonesia have reached 2,033,421 cases [3]. Currently, Central Java Province has confirmed that the number of positive Covid-19 cases is 232,839 and ranks third in the province with the highest number of Covid-19 cases in Indonesia. Therefore, this study was conducted to determine the prediction of extreme value of the number of Covid-19 cases in Central Java Province using time series analysis. Exponential Smoothing Holt-Winters and to find the extreme values, using scenarios from the Mean Excess Function (MEF) and Value at Risk (VAR). Estimating extreme value of the number of new cases aims to support information and maintain essential health services.

Several researchers have used the Exponential Smoothing Holt-Winters method in their research in conducting a forecast. In research [4], a study was conducted to predict the number of passengers on the Medan-Rantau Prapat train with the Exponential Smoothing Holt Winters Method. The data used in this study is data in the number of passengers on the Medan-Rantau Prapat train from 2007 until 2011. This study compares the additive seasonal model and multiplicative seasonal model. Results the additive model has a MAPE value of 0.6. Meanwhile, the multiplicative model has a MAPE value of 0.81. In research [5], This study aims to determine how the risk of stock portfolio of PT. Astra Agro Lestari Tbk (AALI) and PT. PP London Sumatra Indonesia Tbk (LSIP) use Value at Risk analysis using Varian-Covariance method at closing price of shares incorporated in Jakarta Islamic Index (JII) and Asset Value at Risk PT. Astra Agro Lestari Tbk (AALI) and PT. PP London Sumatra Indonesia Tbk (LSIP) to Value at Risk Portfolio. The results showed that if the initial fund invested to PT. Astra Agro Lestari Tbk. and PT. PP London Sumatra Indah Tbk. Rp. 10,000,000, - with a 95% confidence level obtained Value at Risk (VaR) of Rp. 369,682. this can be interpreted there is a 95% confidence that the losses received by investors will not exceed from Rp. 369,682. The result of PT. Astra Agro Lestari Tbk. against portfolio risk at 6% and PT. PP London Sumatra Indonesia Tbk. of portfolio risk is 46%.

The number of daily Covid-19 cases reveal fluctuations, several seasonal and noise patterns. As the significant severity of this pandemic, estimating the future number of extreme values of daily Covid-19 become a major concern to support information and maintain essential health services. Central Java is the province with the third highest population density in Indonesia and predicted to be affected significantly over a particular period of this outbreak. We introduce Exponential Smoothing Holt-Winters and extreme value approaches to capture high fluctuation and unprecedented volatility of daily Covid-19 cases. We propose Mean Excess Function (MEF) and Value at Risk (VaR) based variance-covariance simulation as extreme value approaches. We compare the violation rate of VaR and MEF methods. This paper is organized in the following sections: the first about the framework of Exponential Smoothing Holt-Winters and extreme value approaches are introduced. We propose Mean Excess Function (MEF) and Value at Risk (VaR) based variance-covariance simulation. Next section is detailed procedure of estimating extreme values based Exponential Smoothing Holt-Winters, we observe the forecast result of Holt-Winters with additive model and determine the extreme value of daily Covid-19 cases in Central Java corresponding to some confidence levels. In the third section, we compare the violation rate VaR and MEF methods. The last section is conclusion and proposing possibly future work.

## II. Literature Review

### *A. Exponential Smoothing Holt-Winters*

The time series method is a forecasting method using the analysis of the relationship between the variables to be estimated with the time variable [6]. Forecasting a time series data need to pay attention to the type or data pattern. In general, there are four kinds of time series data patterns, namely horizontal, trend, seasonal, and cyclical [7]. The horizontal pattern is an unexpected and random event, but its occurrence can affect time series data fluctuations data. The trend pattern is the trend in the direction of the data in the number of beds, it can be an increase or decrease. Seasonal patterns are fluctuations in data that occur periodically within one year, such as quarterly, monthly, weekly, or daily. While the cyclical pattern is a fluctuation of the data for more than one year. Holt-Winter exponential smoothing is a method that can overcome trend and seasonal factors that appear simultaneously in time series data. This method is based on three elements: seasonal, trend and residual patterns by assigning three consecutive smoothing-weightings: alpha, beta, gamma. The coefficient of α, β, dan γ lies between 0 and 1 which is determined subjectively or by minimizing the error value of forecast [8]. Holt-Winters Exponential Smoothing method is divided into two models. The Holt-Winters Exponential Smoothing time period method with the additive seasonal model is used for constant seasonal variations and with multiplicative seasonal that used for fluctuation of seasonal dataset. Holt-Winters Exponential Smoothing with seasonal additive model has the following formula [8],

(1) Smoothing level

$$L_t = \alpha\,(y_t - S_{t-s}) + (1-\alpha)\,(L_{t-1} + b_{t-1}) \quad (1)$$

(2) Smoothing Trend Pattern

$$b_t = \beta\,(L_t - L_{t-1}) + (1-\beta)\,b_{t-1} \quad (2)$$

(3) Smoothing trend seasonal

$$S_t = \gamma\,(y_t - L_t) + (1-\gamma)\,S_{t-s} \quad (3)$$

(4) Forecast for the next period

$$F_{t+m} = L_t + b_t m + S_{t-s+m} \quad (4)$$

These following are Holt-Winters Exponential Smoothing parameters with seasonal multiplicative approach [8],

(1) Smoothing Level

$$L_t = \alpha\,\frac{y_t}{S_{t-s}} + (1-\alpha)\,(L_{t-1} + b_{t-1}) \quad (5)$$

(2) Smoothing Trend Pattern

$$b_t = \beta\,(L_t - L_{t-1}) + (1-\beta)\,b_{t-1} \quad (6)$$

(3) Seasonal Pattern Smoothing

$$S_t = \gamma\,\frac{y_t}{L_t} + (1-\gamma)\,S_{t-s} \quad (7)$$

(4) Forecast for the next period

$$F_{t+m} = (L_t + b_t m) + S_{t-s+m} \quad (8)$$

where

| | |
|---|---|
| $\alpha, \beta, \gamma$ | : smoothing constant, $0 < \alpha, \beta, \gamma < 1$, |
| $y_t$ | : observed value at time $t$, |
| $L_t$ | : smoothing level value at time $t$, |
| $b_t$ | : trend pattern smoothing value at time $t$, |
| $S_t$ | : seasonal pattern smoothing value at time $t$, |
| $F_{t+m}$ | : forecast for at time $t+m$, |
| $s$ | : seasonal length |

The initial parameters of Exponential Smoothing Holt-Winters are presented as follows [9],

(1) Smoothing level estimation

$$L_S = \frac{1}{S}(y_1 + y_2 + \cdots + y_s) \quad (9)$$

(2) Smoothing trend estimation

$$b_s = \frac{1}{s}\left(\frac{y_{l+1} - y_1}{s} + \frac{y_{l+2} - y_2}{s} + \cdots + \frac{y_{l+l} - y}{s}\right) \quad (10)$$

(3) Estimation of smoothing the seasonal pattern Additive model

$$Sk = (yk - Ls) \quad (11)$$

(4) Estimation of smoothing the seasonal pattern Multiplicative model

$$Sk = \frac{y_k}{L_s} \quad (12)$$

For $L_s$, $b_s$, and $S_k$ are initial value for each level, trend, and seasonal pattern respectively.

*B. Extreme Value Estimation*

*1) Value at Risk (VaR)*

There are many methods to calculate the extreme value of a forecast. Calculation of the value of this extreme value is intended to avoid the potential for loss or a very high impact on a problem. By forecasting the value of t risk, people will be more efficient and more targeted in preparing the resources needed to avoid the potential for a loss or a large impact. One method to find the extreme value of a forecast is VaR [10]. There are three main methods in calculating VaR, first is the Variance-covariance method, the Monte Carlo simulation method, and the Historical Simulation method. VaR is a method of forecasting the extreme value obtained over a certain period with the desired level of confidence. The level of confidence states how likely is that the VaR value does not exceed the maximum loss. The higher the value of the level of trust used, the higher the risk taken. The commonly confidence levels in VaR are 90%, 95%, and 99% [11]. This following is formula to find the daily extreme value forecast for a certain period using VaR,

$$VaR = \mu + k \, x \, \sigma \quad (13)$$

where μ is mean, $k$ shows quartile using standard normal distribution with a particular confidence level and $\sigma$ represent standard deviation.

*2) Mean Excess Function*

Mean Excess Function (MEF) is also one of methods considered to assess an extreme value [11]. Forecasting the extreme value using the MEF works by finding the average value of the actual value, which is greater than the conditional value, in this case it means the forecast value.

$$\text{MEF} = E\,[X \mid X > u] \quad (14)$$

where $u$ is the expected value of Holt-Winters Exponential Smoothing forecast and $X$ presents the actual (observed) data sets.

*3) Extreme Value Violation*

In calculating the extreme value forecast, there is the potential for a violation. What is meant by violation is when the result of the forecast extreme value is lower than actual forecasting data in the daily period. So, to overcome this problem, it is necessary to check the violation.

$$\text{Violation Check} = \frac{m}{n} \text{ x } 100\% \quad (15)$$

with $n$ is a number of forecasting data, $m$ presents a number of the forecast values is less than the actual data.

## III. RESEARCH METHOD

We present an overview of design system to estimate extreme values of daily Covid-19 cases in Central Java between March and April 2021 in Figure 1. We introduce Exponential Smoothing Holt-Winters and extreme value approaches to capture high fluctuation and unprecedented volatility of daily Covid-19 cases to support maintain essential health services. We propose Mean Excess Function (MEF) and Value at Risk (VaR) based variance-covariance simulation as extreme value approaches. We compare the violation rate of VaR and MEF methods.

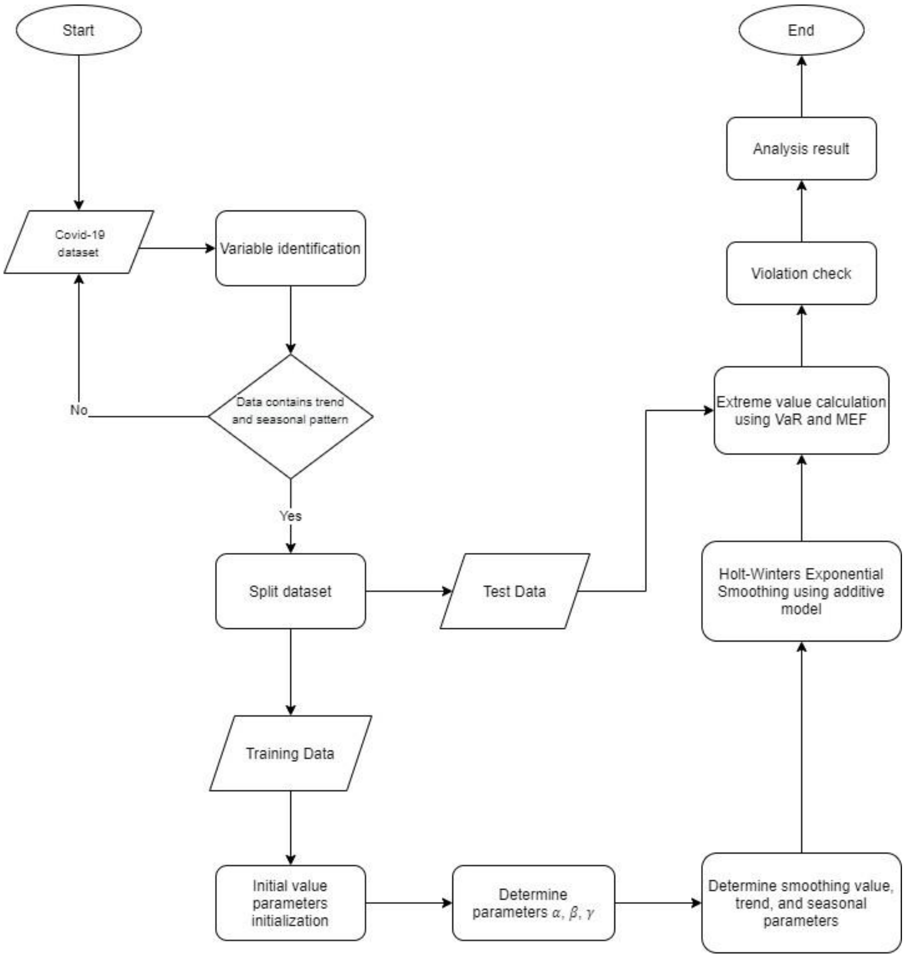


Figure 1. System Flowchart of Estimating Extreme Value Based-Approach Exponential Smoothing Holt Winters

### *A. Dataset*

We employ the Exponential Smoothing Holt-Winters on daily new cases Covid-19 in Central Java Province from March 14 to April 17, 2021 extracted from kawalcovid19.id website. We observed seasonal and trend pattern during the particular periods. Holt-Winters Exponential Smoothing captures trend and seasonal patterns from time series data in which data from March 14 to April 17 2021 to capture trend and seasonal patterns. The data has a seasonal length of 7 and a seasonal periodicity that repeats every 7 days. Thus, we consider to apply the seasonal periodicity in the testing-our splitting data which is around 7.

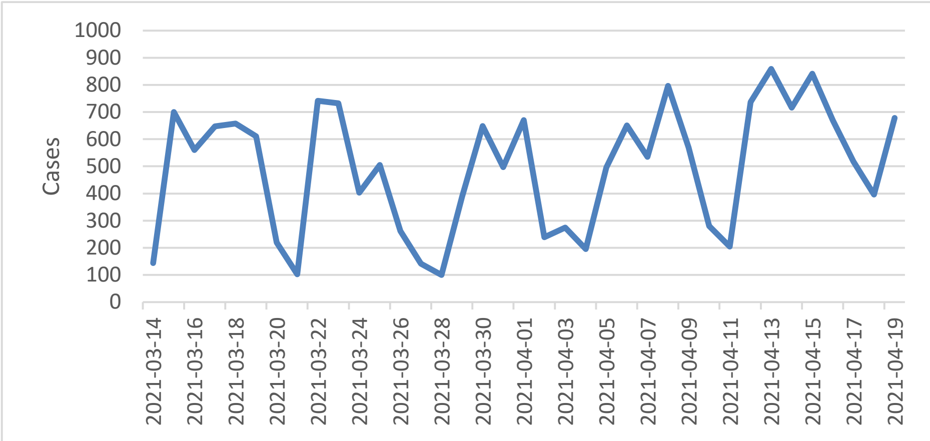


Figure 2. Plot of New Covid-19 Cases in Central Java Province March 14 to April 17, 2021

### *B. Parameters Estimation*

The initial value parameters are carried out as a first step in the Holt-Winters Exponential model. To calculate the initial value parameters, we consider using Equation (9)-(12). To get the initial value of the seasonal period, it is necessary to have at least one complete season data. To determine the initial parameters (α, β, γ), we observed by looking the smallest error value as the optimal value for each parameter by using random values between 0 and 1. We applied GRG-Nonlinear Solving Method to obtain the optimal parameters. Indeed, for each smoothing patterns we employ Equation (1)-(3).

### *C. Extreme Value Estimation*

To anticipate a high volatility of Covid-19 new cases, we employ an estimation of extreme value based on two approaches. We follow Equation (13)-(14) to perform VaR and MEF as extreme value estimation. Furthermore, to measure the performance of VaR and MEF based Holt-Winters, we perform violation check using Equation (15).

## IV. Results And Discussion

### *A. Trend and Seasonal Pattern Identification*

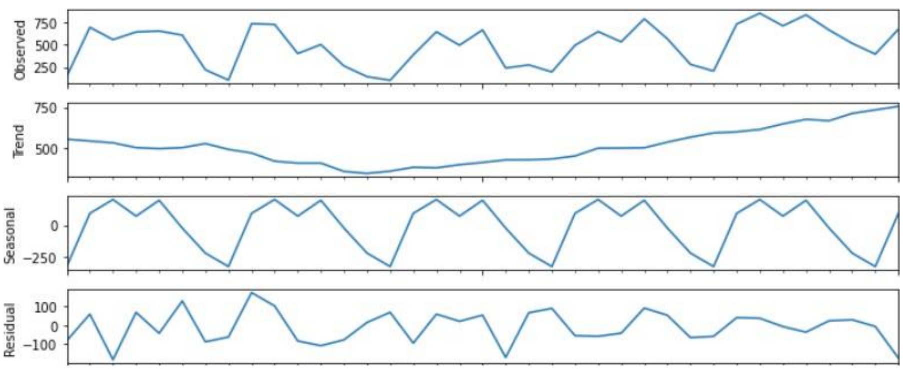


Figure 3. The Trend, Seasonal and Residual Pattern Observed in Data Sets of Daily Covid-19 in Central Java

At first, we analyze the presence or absence of trend pattern and seasonal pattern in the data sets. Indications of trend and seasonal patterns show that these data have important parameters required by Holt-Winters

Exponential Smoothing model. The trend pattern and seasonal pattern contained in the data are shown in Figure 3. Figure 3 shows the actual data for new Covid-19 cases has an upward trend, and a seasonality pattern as requirements of Holt-Winters Exponential Smoothing model. Based on the trend and seasonality pattern, we consider to apply Exponential Smoothing Holt-Winters with additive model.

*B. Initialization of Holt-Winters Exponential Smoothing with Additive Model*

These following are steps to taken in determining the initial value parameters of the additive model. We apply Equation (9)-(11) to determine the initial value of smoothing level, trend, and seasonal. This following is the initial value of smoothing level by averaging of the actual data in the first season period.

$L_s = \frac{1}{S}(y_1 + y_2 + \cdots + y_s)$

$L_s = \frac{1}{7}(144 + 700 + 560 + 647 + 658 + 611 + 219)$

$L_s = 505.57$

The initial value of smoothing trend,

$b_s = \frac{1}{s}\left(\frac{y_{l+1} - y_1}{s} + \frac{y_{l+2} - y_2}{s} + \cdots + \frac{y_{l+l} - y_l}{s}\right)$

$b_s = \frac{1}{7}\left(\frac{102 - 144}{7} + \frac{741 - 700}{7} + \frac{732 - 560}{7} + \frac{741 - 700}{7} + \cdots + \frac{732 - 560}{7}\right.$

$b_s = -13.33$

The initial value of smoothing seasonal pattern,

$S_k = (y_k - L_s)$

$S_1 = -361{,}57, S_2 = 194{,}43, S_3 = 54{,}43, S_4 = 141{,}43, S_5 = 152{,}43, S_6 = 105{,}43, S_7 = -286{,}57$

*C. Parameters of Holt-Winters Exponential Smoothing*

In the Holt-Winters Exponential Smoothing model, the first process of additive model is to determine the values of α, β, γ, with range of values between 0 and 1. As for the forecasting experiment, the values of α, β, γ, used are $\alpha = 0.19$, $\beta = 0.35$, $\gamma = 0.36$. The next step is to calculate the smoothing level value using Equation (1). Then, calculating the trend pattern smoothing value using Equation (2) and smoothing the seasonal pattern involves Equation (3).

TABLE I
THE SMOOTHING WEIGHTINGS

| Date | Smoothing Level Value (Lt) | Smoothing Trend Pattern Value (bt) | Smoothing Seasonal Pattern Value (St) |
|---|---|---|---|
| March 21, 2021 | 486.78 | -15.27 | -370.02 |
| March 22, 2021 | 485.82 | -10.19 | 216.53 |
| March 23, 2021 | 514.15 | 3.48 | 113.90 |
| … | … | … | … |
| April 9, 2021 | 541.75 | 24.07 | -223.81 |
| April 11, 2021 | 551.74 | 19.08 | -308.74 |

*D. Forecasting Exponential Smoothing Holt-Winters Using Additive Model*

We employ Holt-Winters based additive model to forecast new cases of Covid-19 corresponding to the testing data set from April 12 to April 17, 2021.

TABLE II
ADDITIVE MODEL FORECASTING RESULT

| Date | Actual Data | Forecasting |
|---|---|---|
| April 12, 2021 | 737 | 728 |
| April 13, 2021 | 859 | 811 |
| April 14, 2021 | 716 | 726 |
| April 15, 2021 | 841 | 879 |
| April 16, 2021 | 669 | 669 |
| April 17, 2021 | 518 | 461 |

Table 2 shows the result of forecasting using the Holt-Winters Exponential Smoothing additive method have a high accuracy with a MAPE value of 3.94. From Table 2, we can see that the highest new cases happened in April 15, 2021 and the lowest was on April 17, 2021, with a number of new cases was 879 and 461 cases respectively.

### *E. Extreme Value Based on Mean Excess Function (MEF) and Value at Risk (VaR)*

To estimate extreme value of Covid-19 new cases, we consider using MEF and VaR with confidence level of 90% and 95%. We observe the forecasting results of Holt-Winters model as the input to construct MEF and VaR following Equation (13)-(14). We also present the violation check to measure MEF and VaR performances using Equation (15). To maintain a huge necessity of extreme number of Covid-19, we estimate the optimal extreme value that based forecast result Holt-Winters and the violation rate. The result is, in the period April 12 to April 17 2021, extreme value of the number of Covid-19 cases using MEF was 705 cases with a violation rate is 67%. The violation rate from the MEF here is still quite large, meaning that in that period if there is a spike in Covid-19 cases, the probability is only 33% that the spike in cases will reach 705 cases. In contrast, the extreme values based on VaR with 90% and 95% confidence levels have zero violation rates. Moreover, this VaR results more well-performed than MEF because we consider confidence levels to calculate extreme values. Thus, we can control the estimation considering the data violation. the higher significance level, then the lower violation we get.

## V. CONCLUSION

The number of daily Covid-19 cases reveal fluctuations, several seasonal and noise patterns. As the significant severity of this pandemic, estimating the future number of extreme values of daily Covid-19 become a major concern to support information and maintain essential health services. Central Java is the province with the third highest population density in Indonesia and predicted to be affected significantly over a particular period of this outbreak. We present a design system to estimate extreme values of daily Covid-19 cases in Central Java between March and April 2021. We introduce Exponential Smoothing Holt-Winters and extreme value approaches to capture high fluctuation and unprecedented volatility of daily Covid-19 cases to support maintain essential health services. We propose Mean Excess Function (MEF) and Value at Risk (VaR) based variance-covariance simulation as extreme value approaches. We compare the violation rate of VaR and MEF methods. The violation rate from the MEF here is still quite large, meaning that in that period if there is a spike in Covid-19 cases. Our simulation studies indicate that Exponential Smoothing Holt-Winters and Value at Risk with a MAPE value of 3.94 provide sharp and well prediction for extreme value with zero violation. Since a number of positive cases has resulted unprecedented volatility, estimating the extreme value of daily Covid-19 cases become a crucial matter to support maintain essential health services.

## REFERENCES

[1] Yuliana, “Corona virus diseases (Covid-19); Sebuah tinjauan literatur,” *Wellness and Healthy Magazine.* Lampung, 2020.

[2] Public Health of Indonesia, “The First Confirmed Cases Of Covid-19 In Indonesian Citizens,” 18 May 2020. Online Available: http://stikbar.org/ycabpublisher/index.php/PHI/article/view/337/pdf/. [Accessed 1 June 2021].

[3] KawalCOVID19, “Data Update Covid-19,” 1 March 2020. Online. Available: https://kawalcovid19.id/. [Accessed 24 June 2021].

[4] L. Sinay, T. Pentury and D. Anakotta, “’Peramalan Curah Hujan di Kota Ambon Menggunakan Metode Holt-Winters Exponential Smoothing,” *Jurnal Ilmu Matematika dan Terapan*, Ambon, 2017.

[5] N. Wahidah, Y. Nasution and N. Rizky, “Analisis *Value at Risk* Portofolio Saham Menggunakan Varian-Kovarian (Studi Kasus : Harga Penutupan Saham Harian PT. Astra Agro Lestari Tbk dan PT. PP London Sumatra Indonesia Tbk bulan Juli-Desember 2016 ,” *Jurnal EKSPONENSIAL*, Samarinda, 2018.

[6] S. Hansun, “Peramalan Data IHSG Menggunakan Fuzzy Time Series.” *Indonesian Journal of Computing and Cybernetics Systems*, Tangerang, 2012.

[7] A. Mahfud, Widiarti, D. Kurniasari and M. Usman, “Peramalan Data *Time Series* Menggunakan Metode Analisis Spektral,” *Jurnal Siger Matematika*, Lampung, 2020.

[8] S. Makridakis, S. C. Wheelwright and V.E McGree, Metode dan Aplikasi Peramalan, Jakarta: Erlangga, 1999.

[9] D. Rosadi, Ekonometrika dan Analisis Runtun Waktu Dengan Eviews, Yogyakarta: Andi, 2012.

[10] D. Marudanni and A. Purbowati, “Pengukuran *Value at Risk* Pada Aset Tunggal dan Portofolio Dengan Simulasi Monte Carlo,” *Media Statistika*, Semarang, 2009.

[11] A. Rejeb, O. Salha and J. Rejeb, “Value at Risk Analysis for The Tunisian Currency Market: A Comparative Study,” *International Journal of Economics and Financial Issues*, Tunisia, 2012.