UJM 12(1) 2023
**UNNES Journal of Mathematics**
http://journal.unnes.ac.id/sju/index.php/ujm

# Implementation of state space method for forecasting the number of patients with HIV/AIDS infectious diseases

**Wiwik Wiyanti***

Universitas Matana, Indonesia
Matana University Tower, West CBD street, Kav RW1, Curug Sangereng, Tangerang, 15810

## Article Info

Article History:
Recived January 2023
Accepted April 2023
Published May 2023

*Keywords:*
*HIV, AIDS, state space, forecasting, prediction*

## Abstract

This research concerns about the implementation of the State Space method for forecasting to the number of people with infectious diseases (a case study is the number of people with HIV/AIDS) in the Indonesia. The research was conducted from August 2022 to November 2022. The data obtained from kagle, namely new cases of HIV/AIDS infection. In this research, HIV data from 1990 to 2019 was used for forecast analysis. The forecasting method in this research is State Space method. Predict of new cases will found in 2020, 2021, 2022 and 2023 are 14095,14139, 14167 and 14184 respectively. The forecast analysis result obtained that the mean absolute percentage error is 0,4%.

## How to cite:

Wiyanti, W. 2023. Implementation of state space method for forecasting the number of patients with HIV/AIDS infectious diseases. *UNNES Journal of Mathematics*. 12(1):56-64.



*Correspondence Address
E-mail: wiwik.chae@gmail.com

p-ISSN 2252-6943
e-ISSN 2460-5859

## 1. Introduction

Based on a report from the Data and Information Center of the Indonesian Ministry of Health, it is known that the number of HIV/AIDS cases is increasing every year (Pusdatin, 2020). The number of HIV cases in Indonesia has increased for nine years from 2010 to 2019, with the highest addition of cases in 2019, amounting to 50,282 cases. This is in line with reports from several countries of an increasing trend in new HIV/AIDS infections that had previously declined. Data facts recorded at the World Health Organization (WHO) that HIV/AIDS disease has claimed 40.1 million (33.6-48.6) million lives, with continuous transmission in all countries globally (WHO, 2023).

HIV disease or called the Human Immunodeficiency Virus is a type of virus that infects white blood cells which causes a decrease in human immunity (Pusdatin, 2020). Acquired Immune Deficiency Syndrome (AIDS) is a set of symptoms that arise due to a decrease in immunity caused by infection by HIV (Pusdatin, 2020). HIV patients require treatment with Antiretroviral (ARV) to reduce the number of HIV viruses in the body so as not to enter the AIDS stage, while AIDS patients need ARV treatment to prevent opportunistic infections with various complications (Pusdatin, 2020).

HIV can be transmitted through the exchange of various body fluids such as blood, breast milk, sperm and vaginal fluids from someone who has previously been infected with the virus. HIV can also be transmitted from a pregnant woman infected with the virus to her unborn child. However, people cannot be infected through everyday contact such as skin touching, hugging, kissing, shaking hands or handling personal objects such as food and even water (WHO, 2023).

From the description above, it can be concluded that HIV and AIDS is one of the human diseases that is contagious and caused by a virus with a low level of transmission with a way of transmission that is not easy. Although the mode of transmission of HIV and AIDS is not easy, based on data found in the field, active cases of HIV/AIDS sufferers from year to year always increase.

Scientifically, the number of HIV and AIDS cases over time can be studied and predicted. By knowing the prediction of the range of HIV/AIDS sufferers in the future, it is hoped that it can be used as a basis for making policies for related parties such as the health department. For example, when the coming year is predicted that there will be an increase in the number of HIV/AIDS sufferers, the health department can make early movements to anticipate in order to minimize the predicted number.

Scientific studies on prediction with the theme of infectious diseases using forecasting methods in statistics such as HIV/AIDS have been carried out, one of which is prediction using the Exponential Smoothing method researched by Harjono & Kuntoro (2018) with the results of MAPE forecasting error of 27.87% for prediction of male gender and 47.06% for female gender. Ningsih & Jana (2018) using the ARIMA method, obtained the best model ARIMA (1,1,1) with a predicted value of 221 people with HIV/AIDS in the next four years. Still using the ARIMA method, Faqih & Jana (2022) obtained the best model ARIMA(0,0,1) with predicted values in 2019 and 2020 of 52 and 37 people respectively. Meanwhile, forecasting research studies in the health sector regarding other infectious diseases, such as dengue fever, have been researched by Achmanda (2018) using the ARIMA method with the best model results ARIMA(1,1,1). Pamungkas (2018) using the SARIMA method with the best model results ARIMA(1,1,2)(2,1,1)12 with an error rate using MAPE of 39.5% with prediction results in 2017, 2018 and 2019 of 11,490.76; 18065.73; and 21,937.28 cases, respectively. Kushartanti & Latifah (2020) have examined dengue cases in Semarang city with the results of the best model obtained is ARIMA (1,0,0) with a MAPE of 43.98%.

As for the development of science and technology, in addition to the methods that have been used above, there are several other statistical methods developed to be used as tools or tools to predict a time series type data, one of which is the State Space method. Makridakis & McGee (1997) wrote that the State Space method for predicting time series data is thought to be suitable or useful to use because first, it uses the Kalman recursive equation, which can be relatively easy to calculate. Second, the State Space model equation is easy to generalize. Third, the State Space equation can overcome missing data in time series data. In Shumway & Stoffer (2011), although the initial State Space model was introduced as a method used in aerospace research, State Space has been applied to modeling data from economics, medicine, and soil science.

Research using the State Space method in the field of medicine was conducted by Nastiti, et al (2020) using the State Space method to predict a woman's fertile period, where the results of the study obtained a prediction error of only 5%. Meanwhile, apart from the health sector, research using State Space in the transportation sector, namely predicting the number of train passengers, has also been carried out by Saraswita, Sumarjaya & Harini (2020), which in the calculation analysis results get a prediction error percentage value of 2%. Research in the economic field, namely the prediction of stock prices using State Space has also been studied by Manogari (2019), where the results of the analysis of stock price predictions are obtained very well with a mean absolute percentage error of 3.0862% for ENRGJK, 2.4146% for MEDJK and 1.9264% for ELSAJK.

Based on the description above, researchers are interested in implementing the State Space method to predict the number of infectious disease sufferers, in the case study of the number of HIV/AIDS sufferers. The purpose of this research is to find out how many predictions of the number of HIV/AIDS sufferers, as well as how big the prediction calculation error is. Furthermore, the benefits of the results of this study are expected to be used as a scientific reference for related agencies in this case is the Health Office. In addition, the results of this study can be used as a reference for academics as a reference for further research related to forecasting/prediction methods, especially for health data. As for the help of data analysis calculations in this study using the "R" software.

## 2. Method

The data source used in this study is secondary data obtained from the opensource data provider website Kaggle (www.kaggle.com). The data used is data on new cases of many people infected with HIV/AIDS from 1990 to 2019.

The method used to predict in this study is to use quantitative forecasting methods, namely the State Space method to obtain prediction numbers. According to Brockwell & Davis (2002), the State Space Model is a time series model $\{Y_t, t = 1,2,3,...\}$ which consists of two equations, namely the "observation equation" and the "state equation". The observation equation is expressed as a $w$-dimensional observation $Y_t$ as a linier function of $v$-dimensional state variables which are then added with noise. The state equation is to determine the state $X_{t+1}$ at time $t+1$ from the previous state, $X_t$ and noise where $t$ is a time variable. The mathematical equations of the State Space model can be seen in equations (3) and (4). In this study the time used is discrete, suppose that $t = 1$ means the time in the first year, namely 1990. $Y_t$ is data on the number of HIV/AIDS patients at time $t$. Suppose $Y_1 = 2.747$, meaning that the number of HIV sufferers in Indonesia in the first year, namely 1990 was 2,747.

In the writings of Makridakis & McGee (1997), it is explained that the State Space formula is assumed with the variable $Y_t$ as a forecast variable that can be expressed as a linear function of several random variables $X_{1,t}, X_{2,t}, \cdots, X_{d,t}$. The equation can be written as

$$\boldsymbol{Y_t = h_1 X_{1,t} + h_2 X_{2,t} + \cdots + h_d X_{d,t} + z_t} \quad (1)$$

(Makridakis&McGee,1997)

variabel $X_{1,t}, X_{2,t}, \cdots, X_{d,t}$ are state variables. Equation (1) is a general equation with $d$ multivariate random variables $X$. Equation (1) can be simplified to write in matrix form as equation (2).

$$\boldsymbol{Y_t = HX_t + z_t} \quad (2)$$

(Makridakis&McGee,1997)

where $X_t$ is a multivariate time series with state variables as its components and $H$ is a matrix consisting of one row with elements $h_1, ..., h_d$.

In this study, the time series data used is one variable (data on many HIV/AIDS sufferers) which $X_t$ used is univariate time series data, so equation (1) is written as equation (3)

$$\boldsymbol{Y_t = h_1 X_{1,t} + z_t} \quad (3)$$

(Makridakis&McGee,1997)

State space models such as equation (3) can also be referred to as linear dynamic models (DLM) which have a first-order basic form (Shumway and Stoffer, 2011). If it is assumed that $X_t$ depends on the previous state, it is like the following equation (4)

$$\boldsymbol{X_t = FX_{t-1} + Ge_t} \quad (4)$$

(Makridakis&McGee,1997)

where $e_t$ is the set of calculation error values that have white noise properties and the matrices $F$ and $G$ are the measured parameters of the model. Equation (4) is referred to as the "state equation". Together, equations (2) and (3) are referred to as the "State Space" equation/model.

White noise is the prediction calculation error following a standard normal distribution. This means that the calculation error is at the mean or average around 0 and the distribution or variance of the error is around 1. In the data analysis of this study, the error in question will be referred to as residual. The mention of this residue is due to the data used in the form of samples. Determination of white noise can be done by looking at the residual plot, steam and left diagram or by the LjungBox test. All good forecasting models should meet these conditions. In Makridakis & McGee (1997) mathematically, the White Noise model is written in the form of equation (5).

$$\boldsymbol{Y_t = c + e_t} \quad (5)$$

(Makridakis&McGee,1997)

Furthermore, after the model is determined, forecasting is performed using the Akaike procedure. The use of the Akaike procedure is considered because it has been widely used in forecasting methods, on the grounds that it can handle both univariate and multivariate data types (Makridakis & McGee, 1997). The procedure has five stages, which are:

1) It is suitable for multivariate models up to 10 time lags. For each model, the AIC (Akaike Information Criterion) is calculated and the model with the smallest AIC is selected for use in the next stage.
2) Stage 1 tried to improve the model fit by further selecting the AR (Autoregressive) model by adding MA equations and eliminating some AR equations.
3) The best model that has been revised in step 2 is then approximated with a state space model by selecting the least number of parameters.
4) Parameters of the state space model are estimated.
5) Predictions from the model are obtained.

Of the five stages, the main requirement that must be met before using it is that the data must be stationary in terms of mean and variance. This means that the time series data has an average and tends to move towards the average and has a steady or unchanging variance.

In Makridakis & McGee (1997), generally, one way of checking data stationarity is by using a plot with features:

1) If the time series data on the plot does not show changes in the data against a constant average over time, it is said that the time series data is stationary to the mean.

2) If the time series data on the plot shows no change in variance over time, then the data is stationary with respect to variance.

What if the time series data is not stationary? Then the next step that must be done is to stationarize the data. In Makridakis & McGee (1997), one way to stationerize non-seasonal type time series data, such as the data in this study, is by differencing, which is to get the difference value between observation data. The formulas used are like equations 6 and 7.

$$\boldsymbol{Y'_t = Y_t - Y_{t-1}} \quad (6)$$

$$\begin{aligned} \boldsymbol{Y''_t} &= \boldsymbol{Y'_t - Y'_{t-1}} \\ &= \boldsymbol{(Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})} \\ &= \boldsymbol{Y_t - 2Y_{t-1} + Y_{t-2}} \end{aligned} \quad (7)$$

(Makridakis&McGee,1997)

$Y_t$ is the current data, $Y_{t-1}$ is the previous data with a time lag of one period. What if after differencing once but still not stationary? Analogous to the previous step, that is by differencing once again, meaning the second differencing. The formula used is like equation (7) where $Y_{t-2}$ is the previous data with a time lag of two periods.

Next are the possible models to be used in the State Space, the criteria are based on Akaike's (1976) procedure that can overcome the problem of univariate and multivariate data types, namely stages 1 and 2, namely by getting the smallest Akaike's Information Criterion (AIC) value and adding and subtracting Autoregressive (AR) and Moving Average (MA) models by considering choosing the least AR parameters.

As for estimating the best model in stage 2 or ARIMA($p$,$d$,$q$) by looking at the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) correlograms with conditions such as Table 1. The notation $p, d, q$ in ARIMA itself is $p$ is the order of the AR model, $d$ is the degree of differencing performed, and $q$ is the order of the MA model. To better understand, suppose the ARIMA(2,3,1) model means an AR model with order 2 and MA order is 1 where the original data has been differenced 3 times to get stationary data.

Table 1. Model selection

| **Process** | **ACF** | **PACF** |
|---|---|---|
| AR(1) | Decays exponentially. | Tapered sharply at lag 1, then immediately breaks/cuts off to zero. |
| AR($p$) | Decays exponentially or damped sine wave. | Tapered sharply at lag 1 then immediately breaks off/cuts off to zero. |
| MA(1) | Tapering at lag 1 then immediately cuts off to zero. | Decays exponentially. |
| MA($q$) | Tapering sharply at lags 1 to q, then cuts off to zero. | Decays exponentially or damped sine wave. |

Furthermore, after the data is stationary, a State Space model will be determined to obtain the predicted value. Some reasons why using a state space model is useful in forecasting are:

1) Kalman (1960) and Kalman and Bucy (1961) developed a recursive equation for forecasting. It is commonly known as the "Kalman recursive equation" or "Kalman filter". This equation is also quite easy to use as it calculates each step by calculating the prediction error and its likelihood. So, as long as a model can be written in a State Space model, then all calculations can be done using the "Kalman filter".
2) State space models are easy to generalize. For example, it is possible to use a regression model for an ARIMA error model where the parameters of the ARIMA error model change over time.
3) The state space formulation makes it easy to handle missing values in time series.

Hubbard (2023) describes the Kalman filter model itself as a two-stage forward procedure which is a linear optimization filter based on multivariate distribution. The next filter only uses the $t$ data to make inference on data/components that are not observed at time $t$ and will be optimum for the next value. The stages for the Kalman filter itself are starting the prediction stage, then updating, then smoothing and estimation. The implementation of Kalman filter in R software is with the sourcecode *kf=kalman_filter(ssm,yt,Xo,Xs,w,smooth)* with *ssm* is a matrix-matrix object that describes the State Space, *yt* is the multiplication of many variables with many time periods, *Xo* is an optional matrix in the observation equation, *w* is a weight matrix, and *smooth* is a *boolean* that will determine whether the step goes forward or backward more smoothly or not. For a more detailed explanation, you can learn from Hubbard (2023).

After the estimation/prediction value stage is obtained, the next step is to calculate the amount of prediction error. In this study, the calculation of prediction error uses Mean Absolute Percentage Error (MAPE). The MAPE mathematical model is as in equation (8).

$$\boldsymbol{MAPE} = \frac{1}{n}\sum_{t=1}^{n}|\boldsymbol{PE_t}| \quad (8)$$

where $n$ is the amount of data used/analyzed, $PE_t$ is the ratio of the difference between predicted data ($F_t$) and actual data ($Y_t$) divided by actual data ($Y_t$) multiplied by 100. Mathematically like equation (9)

$$\boldsymbol{PE_t} = \frac{(Y_t - F_t)}{Y_t} \times \mathbf{100} \quad (9)$$

(Makridakis&McGee,1997)

The outline carried out in this study from the description above in detail, schematically, can be seen in the summary of Figure 1 to facilitate understanding.

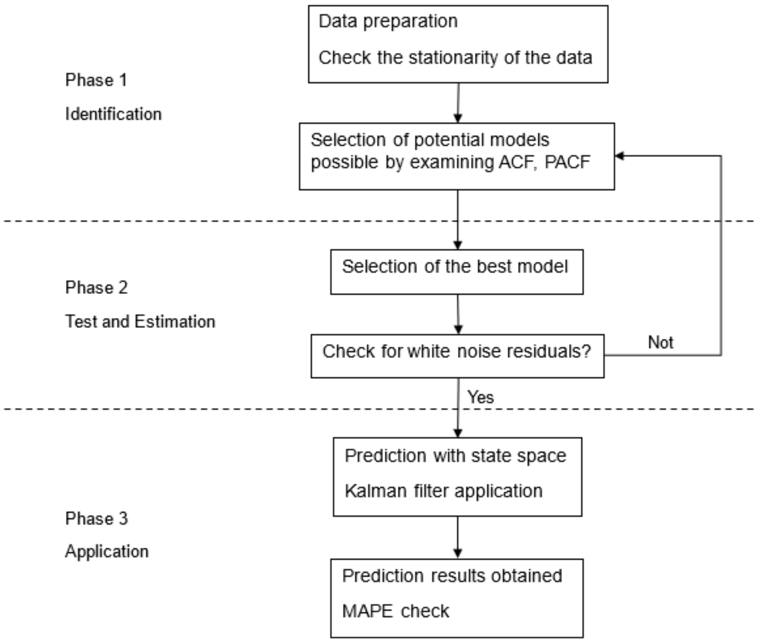



Figure 1. Schematic of research data analysis

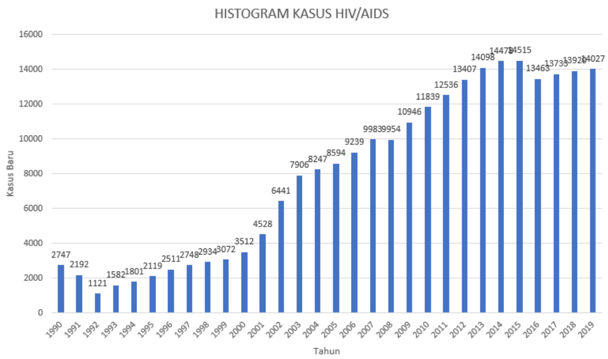
As for the need for calculation assistance when analyzing the data from this study using R software which can be downloaded for free opensource so that it can be used by researchers for free. Although it is free, R software is software developed by statisticians around the world for statistical research purposes.

The R software used for data analysis in this study is version RStudio-2022.07.2-576 for Windows. For more details on learning R software, regarding the use, shortcomings and applications of R for forecasting time series data from basic to advanced levels, the author recommends reading the book Rosadi (2016) for Indonesian-language references, and Derryberry (2014) for English-language references.

The R package library source code used in this study includes (readxl, tsm,lm, tidyverse, tseries, dplyr, lubridate, ggplot2, MLmetrics, forecast, KFAS, KalmanForecast, kf). For a more detailed explanation of the source code functions, please refer to the help tools in Software R.

## 3. Result and discussions

The data used in this study is data on new cases of many people infected with HIV/AIDS from 1990 to 2019. Data obtained from Kaggle web.




Data Source: Hrterhrter, kaggle.com (2022)

Figure 2. Histogram of new HIV/AIDS cases in Indonesia

From the data obtained, it is divided into two parts, the first is called "*data latih* (training data)" which is used for model selection, and the second is called "*data uji* (test data)" which is used to check the error of prediction results with real data. The " *data latih* (training data)" used is from 1990 to 2010, while the data used for " *data uji* (test data)" is data from 2011 to 2019. The prediction/estimation/forecast data are for the months of 2020, 2021, 2022 and 2023.

The first stage for prediction with the state space method as described in the research method section, is to ensure that the time series data obtained is stationary. To see stationary data, one of them can be done by looking at the stationarity plot, which can be seen as in Figure 3.

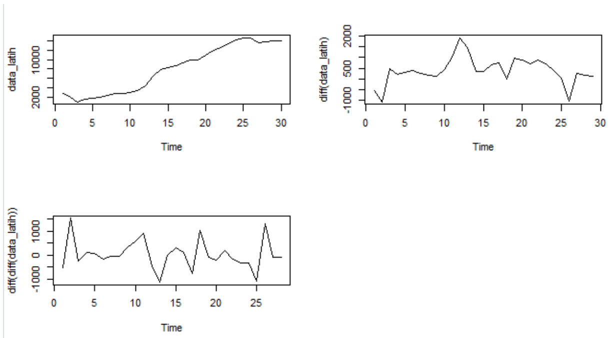



Figure 3. Stationarity plot of HIV/AIDS data

From Figure 2, HIV data without differencing (data_latih) namely the first order image looks non-stationary to the mean but stationary to the variance. After integrating or differencing, it can be seen that HIV/AIDS data is more stationary to the mean and variance. To ensure this, it is then seen using the Augmented Dickey Fuller Test (ADF Test). From the ADF results obtained HIV data without differencing produces a p-value (0.598), by differencing once obtained p-value (0.5367) and differencing twice obtained p-value (0.07585). From the acquisition of the p-value, successively it can be concluded that, by using ADF-test analysis, the data is not stationary at 5% alpha.

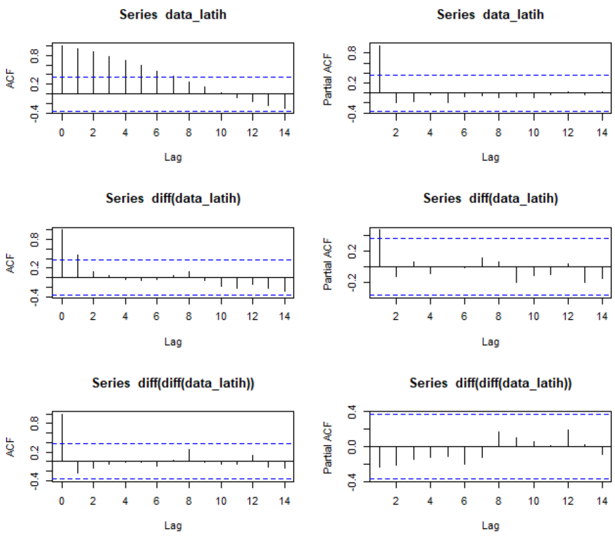



Figure 4. Correlogram of HIV/AIDS data

In addition to the two steps above, stationarity can also be seen from the ACF and PACF correlograms (Figure 4), which is when the correlogram diagram exponentially decays towards zero. From the correlogram diagram, it is obtained that the data after the first differencing is stationary because the ACF corellogram decays rapidly towards zero. Likewise for the second differencing, the first lag in the ACF tapers sharply, while the other lags are within the limit line (dashed). From the three stationarity check steps above, it can be concluded that the HIV/AIDS data is stationary after differencing one and two times.

Furthermore, after the data stationary assumption is met, the second stage is the selection of models that can be considered as State Space models. The selection of the best model can be seen from the smallest Akaike's Information Criterion (AIC) value in accordance with the provisions described in the previous research method. Before getting the AIC value, we will first collect possible ARIMA (p,d,q) models that can be used or considered to be a State Space model. ARIMA(p,d,q) models that can be considered to be State Space models can be selected from ARIMA(p,d,q) models with significant Pr(>|z|) values during the ARIMA(p,d,q) coeftest analysis.

As for the possibility of ARIMA(p,d,q) models seen from the correlogram of ACF and PACF according to the rules in Table 1. From Table 1, the *correlogram* diff(data_latih) obtained the possible models are ARIMA(2,1,0), ARIMA(2,1,2), ARIMA(1,1,0), ARIMA(0,1,1), ARIMA(0,1,2). While for diff(diff(data_latih)) the possibility of ARIMA(0,2,1) model. The results of the ARIMA(p,d,q) coeftest analysis are summarized in table 2. The value of Pr|>z| is compared with the alpha value of 0.05. If Pr|>z|< alpha then the ARIMA(p,d,q) model is said to be significant, meaning it can be considered for use as a State Space model. From Table 2, the model that can be considered or feasible to be used as a State Space model is ARIMA(1,1,0). This model selection consideration is in accordance with the stages in the previous research method chapter. From the summary of the coeftest results obtained, the only model that can be used for the State Space model in the case of this research data is ARIMA(1,1,0).

After the best model is determined, the third stage is to check whether the calculation error of the ARIMA (1,1,0) model is white noise. As for seeing white noise, you can see in addition to the plot of the residuals, it can also be seen from the steam and left diagram or the LjungBox test. The results of the scatter plot diagram using the LjungBox test, it is concluded that all residual data are above the dotted line as shown in Figure 5. From the results of the p-value for LjungBox statistics, it is concluded that the residues of the ARIMA (1,1,0) model are white noise.

Table 2. Model possibilities

| Model Name ARIMA(p,d,q) | Pr\|>z\| | Description | AIC | MAPE |
|---|---|---|---|---|
| ARIMA(2,1,0) | 0,0008644<br>0,7709374 | Significant.<br>Not significant. | 454,51 | 9,262 |
| ARIMA(2,1,2) | 0,1804<br>0,8457<br>0,5509<br>0,6509 | Not significant.<br>Not significant.<br>Not significant.<br>Not significant. | 457,75 | 9,439 |
| ARIMA(1,1,0) | 2,969e-06 | Significant | 452,6 | 9,4 |
| ARIMA(1,1,1) | 0,001616<br>0,724486 | Significant<br>Not significant. | 454,47 | 9,26 |
| ARIMA(2,1,0) | 0,0008644<br>0,7709374 | Significant<br>Not Significant | 454,51 | 9,4 |
| ARIMA(0,2,1) | 0,1205 | Not significant | 440,06 | 9,13 |

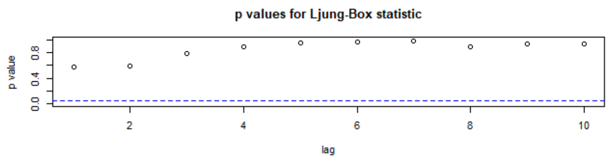



Figure 5. ARIMA(1,1,0) residual test

Next, after the model is obtained is the formation of the State Space model of ARIMA(1,1,0). In the ARIMA(1,1,0) model, the matrix parameters in the State Space model are obtained which are then used to obtain prediction/forecasting values. The parameters that have been obtained or the matrix output obtained are summarized in Figure 6.

The matrices in Figure 6 are the basic matrices used in the Kalman filter to obtain the estimated/forecast value. From the matrix obtained in Figure 6, it is then included in the Kalman algorithm to get a prediction value.

By using the Kalman filter algorithm, the prediction/estimation value of the number of HIV/AIDS sufferers in 2020, 2021, 2022 and 2023 is 14095.35; 14139.00; 14166.88; and 14184.70, respectively. Because the number of HIV/AIDS sufferers is a non-negative integer, the prediction results are rounded to 14095, 14139, 14167, and 14184, respectively. The MAPE result obtained using the State Space method is 0.4%.

```
> fit3$model
$phi
[1] 0.6387414

$tetha
numeric(0)

$Delta
[1] 1

$Z
[1] 1 1

$a
[1] 107    13920

$P
                [,1]            [,2]
[1,] 0.000000e+00  4.440954e-21
[2,] 4.440954e-21  -4.440954e-21

$T
           [,1]  [,2]
[1,] 0.6387414    0
[2,] 1.0000000    1

$V
     [,1]  [,2]
[1,]    1     0
[2,]    0     0

$h
[1]  0

$Pn
                [,1]            [,2]
[1,] 1.000000e+00  -2.836621e-21
[2,] -2.836621e-21  -4.440954e-21
```

Figure 6. Output State Space Matrix

## 4. Conclusion

The conclusion obtained from the results of data analysis and discussion is that by using the State Space method, the prediction of the number of HIV/AIDS sufferers in Indonesia in 2020, 2021, 2022 and 2023 is 14095, 14139, 14167, and 14184, respectively. The error of the prediction calculation results using the State Space method is 0.4%.

The suggestion for future researchers is the application of time series data with the type of seasonal time series data or seasonal time series data. Furthermore, the data used in this study is data with a single variable (univariate), then it is recommended to try state space applications for bivariate or multivariate data.

## References


Achmanda, D., A. 2018. Aplikasi Model ARIMA untuk Meramalkan Jumlah Penderita Penyakit Demam Berdarah Dengue (DBD). *Skripsi*. Medan: Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sumatera Utara.

Brockwell, P. J., dan Davis, R. A. 2002. *Introduction to Time Series and Forecasting second Edition*. USA: Springer-Verlag New York, Inc. ISBN 0-387-95351-5

Derryberry, D.R. (2014). *Basic Data Analysis for Time Series with R*. New Jersey: John Wiley & Sons, Inc.

Faqih., A., dan Jana, P. 2021. Pemodelan ARIMA Penderita HIV/AIDS dengan Minitab. *Jurnal Euclid*, p-ISSN 2355-1712, e-ISSN 2541-4453, Vol.8, No.2, pp.83-173. DOI: *http://dx.doi.org/10.33603/e.v8i2.3312*

Harjono, A., E. dan Kuntoro, K. 2018. Perbandingan Hasil Peramalan Jumlah Kasus HIV Berdasarkan Jenis Kelamin di Kota Malang dengan metode *Exponential Smoothing*. *Jurnal Biometrika dan Kependudukan (Journal of Biometrics and Population)*, Vol.6, No.1, Juli 2018: 9-16.

Hubbard, A. 2023. *Kalman Filter for State Space Models*. Diunduh di https://cran.r-project.org/web/packages/kalmanfilter/vignettes/kalmanfilter_vignette.html pada tanggal 28 februari 2023

Indaru, D.F.B. (2021). *Kelebihan dan Kekurangan R*. Diakses di *https://rpubs.com/Farrel44/822334* pada tanggal 1 Januari 2022.

Hrterhrter. (2022). *HIV/AIDS*. Diunduh di https://www.kaggle.com/datasets/programmerrdai/hiv-aids?select=new-cases-of-hiv-infection.csv tanggal 3 April 2023

Kushartanti, R., dan Latifah, M. 2020. *Autoregressive Integrated Moving Average* (ARIMA) Sebagai Model Peramalan Kasus Demam Berdarah Dengue. *Jurnal Kesehatan Lingkungan* Vol.10, No.2, Oktober 2020, pp.76-80. ISSN 2615-188X(Online), ISSN 2089-0451(Print). DOI: 10.47718/jkl.v10i2.1165. Jurnal Kesehatan Lingkungan (poltekkes-manado.ac.id)

Makridakis, S.,G., Whellright, S., C., dan McGee, V., E. 1997. *Forecasting Methods and Applications/2nd Ed*. Wiley: ISBN 0471532339, 9780471532330

Manogari, N. 2019. Analisis Deret Waktu Menggunakan Metode *State Space* dan Aplikasinya pada Harga Saham. *Skripsi*. Bandar lampung: Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

Nastiti, V.R.S, Risqiwati, D dan Siswoyo, A.N. 2020. Prediksi masa subur pada wanita menggunakan pemodelan *State Space* berbasis suhu basal tubuh. *Jurnal Tekno Kompak* Vol.14, No.1., 2020, 1-4, P-ISSN 1412-9663, E-ISSN 2656-3525

Ningsih, S., C., dan Jana, P. 2018. Pemodelan Penderita HIV/AIDS dengan metode ARIMA. *Jurnal Mercumatematika: Jurnal Penelitian Matematika dan Pendidikan Matematika* Vol. 3, No. 1, Oktober 2018, pp. 23-28. ISSN: 2548-1819. DOI : *http://dx.doi.org/10.26486/jm.v3i1.648*

Pamungkas, M., B. 2018. Peramalan Kasus Demam Berdarah Dengue dengan Metode SARIMA di Provinsi Jawa Timur. *Thesis*. Surabaya: Fakultas Kesehatan Masyarakat Universitas Airlangga. Diunduh dari https://repository.unair.ac.id/70153/ pada tanggal 10 November 2022.

Pusdatin. 2020. *Infodatin Pusat Data dan Informasi Kementerian Kedehatan RI*. Diunduh di https://pusdatin.kemkes.go.id/resources/download/pusdatin/infodatin/infodatin-2020-HIV.pdf tanggal 15 Juli 2022

Rosadi, D. 2016. *Analisis Runtun Waktu dan Aplikasinya dengan R*. Yogyakarta: Gadjah Mada University Press.

Saraswita, F.A.D, Sumarjaya, I.W dan Harini, L.P,I. 2020. “Metode *State Space* dalam Meramalkan Jumlah Penumpang Kereta Api di Pulau Jawa”. *E-Jurnal Matematika* Vol.9(1), Januari 2020, pp.8-13.

Shumway, R. H., and Stoffer, D. S. 2011. *Time Series Analysis and Its Applications with R Examples Third Edition*. Springer Texts in Statistics. Springer Science+Business Media, LLC 2011.

WHO. 2023. Diunduh di *https://www.who.int/news-room/fact-sheets/detail/hiv-aids* tanggal 25 April 2023.