Forecasting Value-Added Tax (VAT) Revenue Using Autoregressive Integrated Moving Average (ARIMA) Box-Jenkins Method

Muchamad Irham Fathoni[a], Akbar Saputra[b],

[a] Directorate General of Taxes, Surabaya, Indonesia. Email: irham.fathoni@kemenkeu.go.id
[b] Directorate General of Taxes, Jakarta, Indonesia. Email: akbar.saputra@kemenkeu.go.id

*Corresponding Author: irham.fathoni@kemenkeu.go.id

ABSTRACT

We propose a method to forecast Value-Added Tax (VAT) revenue for Indonesia government using Autoregressive Integrated Moving Average (ARIMA) Box-Jenkins method. We experimented the ARIMA Box-Jenkins method using time-series analysis of VAT revenue data of two Tax Offices of Directorate General of Taxes (DGT) from the last five years. The result shows that it resembles the real VAT revenue more closely than when compared to the actual VAT target by Indonesia government. We then argue that this result may be used as a fail-safe tax revenue target, that can work as a tool to better measure DGT performance.

Keywords: ARIMA, Box-Jenkins, Value-Added Tax, tax revenue forecast

1. INTRODUCTION

In the first semester of 2022, Directorate General of Taxes (DGT) of the Ministry of Finance of Indonesia recorded a 55.7% tax revenue growth compared to the same period in the previous year (Kurniati, 2022). This translates to 58.5% of tax revenue target for 2022 has been collected. This seems like good news, complementary to a record that DGT finally broke in 2021 for achieving 103.9% of tax revenue targeted in the National Budget—a record that has not been broken for the last twelve years.

However, for the second semester of 2022, with the enactment of President Decree No. 98 of 2022, tax revenue target is revised to grow to IDR 1,783 billion—18.1% higher than the initial target, which was IDR 1,510 billion (Wildan, 2022). Despite

some factors that may still accelerate tax revenue collected (i.e., Income Tax revenue collected in the tax return period ended in March to April, DGT's Voluntary Disclosure Program that ran from January to June, and global commodity price boom), there is still a long way for DGT to achieve this designated target.

In the realm of tax administration in Indonesia or other countries, national budget revision—let alone tax revenue target revision—is not new. Even in Indonesia, tax revenue target has been regularly revised in every budget year due to its dynamic nature and proximity to how the economy changes. VAT, the second most significant component of tax revenue after Income Tax, is one of the most dynamically changed types of tax because it highly reflects the economy and national goods and services transactional volumes.

How to Cite:
Fathoni, M. I., & Saputra, A. Forecasting Value-Added Tax (VAT) revenue using Autoregressive Integrated Moving Average (ARIMA) Box-Jenkins method. *Scientax: Jurnal Kajian Ilmiah Perpajakan Indonesia*, *4*(2), 205–218. https://doi.org/10.52869/st.v4i2.568

While no research found adverse effects of tax revenue target revision, this might highlight a structural issue for Indonesia (especially DGT).

Currently, breaking down national tax revenue target per month and for each vertical unit in DGT is still done conventionally, meaning that no specific method is used to allocate national tax revenue target per month for each Tax Office in DGT (Hidayati, 2016). This may result in biases because, after a certain period, even tax offices in the same region may show significant differences in tax revenue collected. Regarding performance measurements, this also addresses a new issue: after the enactment of President Decree No. 96 of 2017, performance benefits are given to DGT employees based on organizational performance achievement, which factors the percentage of tax collected compared to the target. Thus, if the tax revenue target is revised and broken down without specific scientific calculation, Tax Offices may find it hard to achieve tax revenue target allocated to them and eventually fail to achieve their targeted organizational performance—resulting in a decrease in performance benefits received by employees of said Tax Offices.

More than biases affecting tax revenue target breakdown and performance measurement, a solid tax revenue target is crucial to implement an appropriate and measurable intensification strategy. With proper forecasting to determine accurate tax revenue targets, it will be helpful for DGT's vertical units to determine the direction of tax supervision strategies and potential revenue measurements in each tax period. Suppose a shortfall of tax revenue is collected from the predetermined target in one period. In that case, DGT may re-evaluate its tax intensification and supervision strategy to achieve the target in the year's remaining months.

With that said, we proposed a forecasting method called Autoregressive Integrated Moving Average (ARIMA) to determine tax revenue target for DGT. Forecasting is a conjecture or estimate of future events, and forecasting is an essential tool in effective and efficient planning (Makridakis et al.,

1999). ARIMA Box-Jenkins method—named after statisticians George Box and Gwilym Jenkins—is a time series forecasting technique based only on the behavior of observed variable data (Box & Jenkins, 1970). The ARIMA model completely ignores the independent variables because this model uses the present value and past values of the dependent variable to produce accurate short-term forecasts (Cryer & Chan, 2008).

In this research, we use the ARIMA Box-Jenkins method to determine the VAT revenue target for DGT precisely. VAT is chosen because it contributes up to 33% of total annual tax revenue. As we said in previous paragraphs, VAT revenue targets change dynamically following the economic situation and the transactional activities of goods and services. However, in this case, we propose the ARIMA Box-Jenkins method to see how VAT revenue can be projected in ceteris paribus. Our forecast may contribute to defining a more solid and less biased VAT revenue target in the future. Otherwise, from our forecast, we also elaborate on its effectiveness and accuracy, as well as its usefulness as a fail-safe VAT revenue target that goes side-by-side with the official VAT revenue target specified in the National Budget.

2. THEORETICAL FRAMEWORK
2.1 Tax Collection and Revenue Target in DGT

Generally, tax revenue target is calculated based on the tax buoyancy or tax elasticity of economic growth and tax revenue growth (Direktorat Jenderal Pajak, 2021). Calculating tax revenue targets this way is relatively simple. It only uses historical economic growth, elasticity data, and economic growth assumptions in the next fiscal year as leading indicators. Tax revenue target calculation can also be compiled from various methods on a micro-scale. On the other hand, potential tax revenue is measured based on the expansion of the tax base and the dynamics of specific economic sectors, types of taxes, and regions. Using this micro-scale approach, the

methods and indicators used to calculate tax revenue targets vary widely.

Currently, allocating tax revenue targets to each of DGT's vertical units tends to be based on historical data adjusted with regional micro-scale parameters such as new taxation bases, business growth, and regional economic projected growth. That said, no quantitative calculations or methods are used to allocate tax revenue targets per unit or region more comprehensively and contextually. Developing a tax revenue target based on a quantitative calculation using econometric analysis in the form of time series data is necessary. By looking at how various factors affect tax revenues in certain regions, tax revenue allocated for specific DGT vertical units can be projected. The variables used may also vary, such as economic growth, income per capita, inflation, commodity prices, business growth, etc. These methods can be referred to as top-down macroeconomic data-based methods.

However, it is also necessary to set tax revenue targets based on conditions unaffected by variables (ceteris paribus). This shows the pattern that will be projected based on historical data and can be used as a fail-safe tax revenue target, aside from the official tax revenue target.

2.2 Forecasting Methods

Forecasting is crucial in policy formulation due to the time lag between when the policy is formulated and when it will be implemented. Forecasting must be performed with certain principles, i.e., forecasting involves errors, forecasting should use a benchmark for forecasting errors, and short-term forecasting is more accurate than long-term. There are also steps to follow while forecasting, i.e., defining the purpose of forecasting, creating a data plot diagram, selecting the suitable forecasting models, calculating forecasting errors, and choosing the best forecasting method with the most minor error.

There are two forecasting methods: qualitative and quantitative. The qualitative method is used where there is no mathematical model, usually because the existing data needs to be more representative to predict the future (long-term forecasting). The quantitative method is based on the availability of raw data, accompanied by a series of mathematical rules to predict future results. Quantitative methods are divided into 3 types:
a. Model Time Series Analysis (Time Series)
b. Regression Models
c. Econometric Model

One method that can be used to forecast tax revenue targets is Time Series Analysis introduced in 1970 by George E. P. Box and Gwilyam M. Jenkins through their book Time Series Analysis: Forecasting and Control. Time series can be interpreted as a series of data obtained based on observations of an event in the order in which it occurred. The time of the incident can be a period in units of seconds, minutes, hours, days, months, years, and other periods, all of which are a series of observational data based on the time of the incident with a specific time interval which is better known as a time series (Cryer & Chan, 2008).

The rationale for the Time Series is that the current observation (Z_t) depends on one or more previous observations (Z_{t-k}). In other words, time series is made because, statistically, a correlation (dependency) between a series of observations exist To see the dependencies between observations, we can perform a correlation test between observations which is often known as the autocorrelation function (ACF). In ACF, each observation is expressed as a random variable Z_t obtained on a specific time index (t_i) as a sequence of observations so that the time series data is written $Z_{t1}, Z_{t2}, Z_{t3}, ..., Z_{tn}$. Several things need to be considered in the time series method, such as the stationarity of the data, the autocorrelation function, and the partial autocorrelation function.

2.3 ARIMA

ARIMA is one of time-series models. This model consists of AR (Autoregressive), MA (Moving Average), or ARMA (Autoregressive Moving Average) components. A differencing process is carried out if the data is not stationary in its mean. ARIMA Box-Jenkins model is one of the time series model forecasting techniques that is only based on the behavior of the observed variable data.

The condition that must be met to make a forecasting model is that the data is stationary. AR model will be used to find any relations between the current and previous values by adding arbitrary values. In contrast, the MA model will be used to find relations between the current value and the previous residual value (Wei, 2006). The autoregressive process, as the name implies, is a self-regressive process. The general form of a p or AR(p) level autoregressive process is

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t \qquad (1)$$

which shows that the present value of a process can be expressed as the weighted sum of past p values plus one current random error. In this case, it is assumed that at is independent of Zt-1, Zt-2, etc. So, it can be seen that Zt is regressed on the past p-value of Z.

The identification of the ARIMA Box-Jenkins model can be used to identify the non-stationary model. If the data is not stationary in its mean, then it should be differenced; if it is not stationary in its variance, then Box-Cox can transform it. After the data is stationary in both its mean and variance, the next step is to plot the Autocorrelation Function (ACF) and Partial Auto Correlation (PACF), which are used to identify the initial ARIMA model.

3. RESEARCH METHODOLOGY

3.1 Research Methods, Variables, and Formula

This research does not try to test any hypotheses, so it uses a qualitative approach to find its conclusion. However, in the process, we use quantitative formulas using numerical data analysis and statistical models. Box-Jenkins method is used to define a suitable ARIMA model for our time-series data (Daniel, 1989), and it requires data to be stationary before getting processed. With AR and MA models integrated into our ARIMA model, we utilize the degree of AR (p), degree of differences (d), and degree of MA (q) to create a function ARIMA (p,d,q) as follows:

$$
\begin{aligned}
Yt - Yt - 1 = {}& \varphi 0 + \varphi 1 (Yt - 1 - Yt - 2) \\
& + \cdots + \varphi p(Yt - p - Yt - p - 1) \\
& + \varepsilon t - \omega 1 \varepsilon t - 1 - \omega 2 \varepsilon t - 2 \\
& - \cdots - \omega q \varepsilon t - q
\end{aligned}
$$

where

Yt	= variable of time t,
$\varphi 0$	= constant value,
φp	= p-th AR parameter,
ωq	= q-th MA parameter,
εt	= error value of time t, and
$\varepsilon t - 1, -2, ..., \varepsilon t - q$	= previous error values on related time series

The initial step of our analysis is determining a forecasting model suitable for the VAT revenue data using the time-series method with the following steps:

a. creating a time series plot on the VAT revenue data to check if there are any seasonal patterns;

b. performing the Box-Cox method to identify whether the data has stationarity problems in its variance and/or mean—if the data is not stationary in its mean, then differencing is performed, but if it is not stationary in variance, then data transformation is performed;

c. identifying data stationarity through Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, which are used to check whether the data is finally stationary after transformation and differencing, as well as to find suitable models;

d. performing initial ARIMA (*p,d,q*) model estimation, which consists of three stages—identifying model, assessment and testing, and running the model—and results in a temporary model used for later estimation and diagnostic checking;

e. performing parameter estimation based on the temporary model;

f. performing diagnostic checking;

g. selecting the best model.

The accuracy and effectiveness of the forecasting method can be seen from the difference between the forecast value and the actual value. Two methods are used to measure forecast accuracy:

(2)

a. MAPE (Mean Absolute Percentage Error) measures the accuracy of the estimated model value expressed in the form of the average absolute percentage error and can be expressed as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \quad (3)$$

where
Y_t = actual value on year t
$\hat{Y}_t$ = forecast value on year t
n = number of data

b. MSE (Mean Squared Error) measures the accuracy of the estimated value of the model expressed in the average square of the error and can be expressed as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \quad (4)$$

where
Y_t = actual value on year t
$\hat{Y}_t$ = forecast value on year t
n = number of data

In addition to MAPE and MSE, we used Root-Mean-Square Error (RMSE) and Ljung-Box test to measure accuracy. RMSE, as the name suggests, is the root of MSE and is often preferred to MSE as it is on the same scale as the data (Hyndman & Koehler, 2006). Ljung-Box test is a statistical measure to test if there are some autocorrelation groups in a time-series data that value different from zero (Ljung & Box, 1978).

3.2 Research Scope

Due to limitations and unavailability of broader data, we use VAT revenue data from only two Primary Tax Offices in DGT: Surabaya Karangpilang Tax Office and Jakarta Setiabudi Dua Tax Office. These two tax offices are located in two big cities in Indonesia, which are Surabaya and Jakarta, respectively. These two tax offices run the same responsibilities, meaning that there are no special responsibilities or functions given to these tax

offices that may make them different from other tax offices in DGT.

We take monthly VAT revenue data from 2017 to 2022 from both offices. To better represent the actual VAT revenue, we only consider regular VAT payments and omit VAT payments due to fines or sanctions or VAT payments withheld by the government, treasurers, and other VAT collectors.

4. RESULT AND DISCUSSION

In this study, forecasting will be carried out using VAT revenue data from Surabaya Karangpilang Tax Office and Jakarta Setiabudi Dua Tax Office. To assist us, we utilized Minitab for data visualization, as well as EViews and SPSS for data processing.

4.1 Forecasting VAT Revenue in Surabaya Karangpilang Tax Office

We observed Surabaya Karangpilang Tax Office's VAT revenue from January 2017 to December 2021, consisting of 60 observations. The first step of this research is to determine the forecasting model by the data on the amount of Value Added Tax receipts analyzed using the time series method. Value Added Tax revenue data is included in time series data that can be predicted using the ARIMA method with the following stages of analysis.

a. Identifying Models
1) Data Pilotting
We plotted VAT revenue data from Surabaya Karangpilang Tax Office, as shown in Figure 1. Figure 1 indicates that the data was
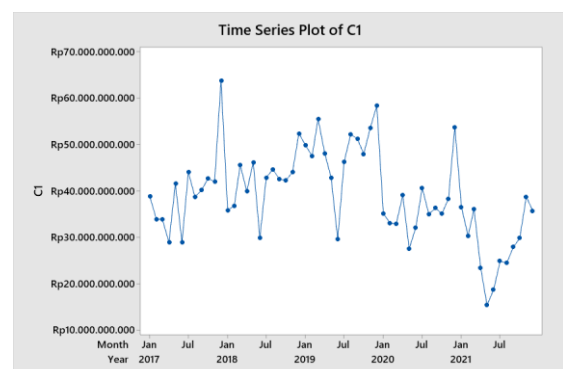

Figure 1 Surabaya Karangpilang Tax Office VAT Revenue Data Plot
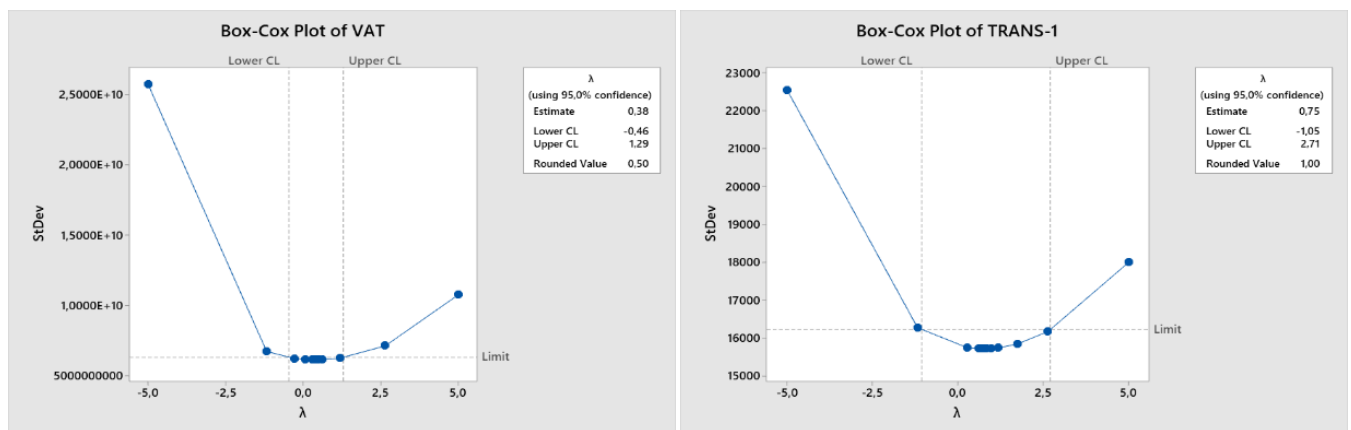Source: Processed data by the author

Figure 2 Box-Cox Plot Before and After Transformation
Source: Processed data by the author

not stationary because a trend appears in the plot. Non-stationary data in variance can be seen from the series plot, where points are not spread out evenly because they increase or decrease with time. We have to transform non-stationary data to be stationary. Diagnostic checking using the Box-Cox method showed a rounded value of 0.50 and data interval between -0.46 and 1.29, or less than 1. We transformed this non-stationary data until we got a rounded value of 1. The result is shown in Figure 2.

2) Data Stationarity and Differentiation

Using data plotting, we then performed a stationarity test in the mean for this data. Again, we found a trend in the data, so we have to transform it using differentiation. Figure 3 shows the time-series data plot after transformation and differentiation. From here, we found that the data has already been stationary.

3) ACF and PACF plot

We continue our testing by further performing stationarity tests in the mean using ACF and PACF plots. The result is shown in Figure 4. From that figure, we can
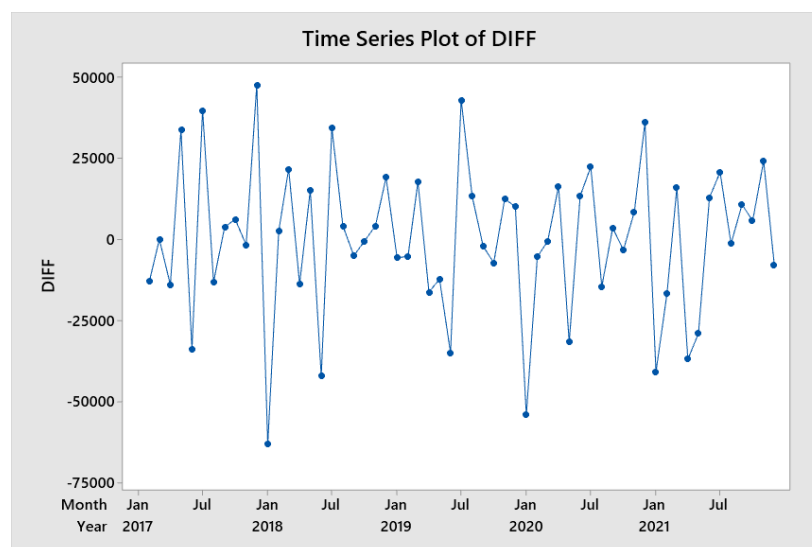

Figure 3 Time-series Data Plot After Differentiation 1
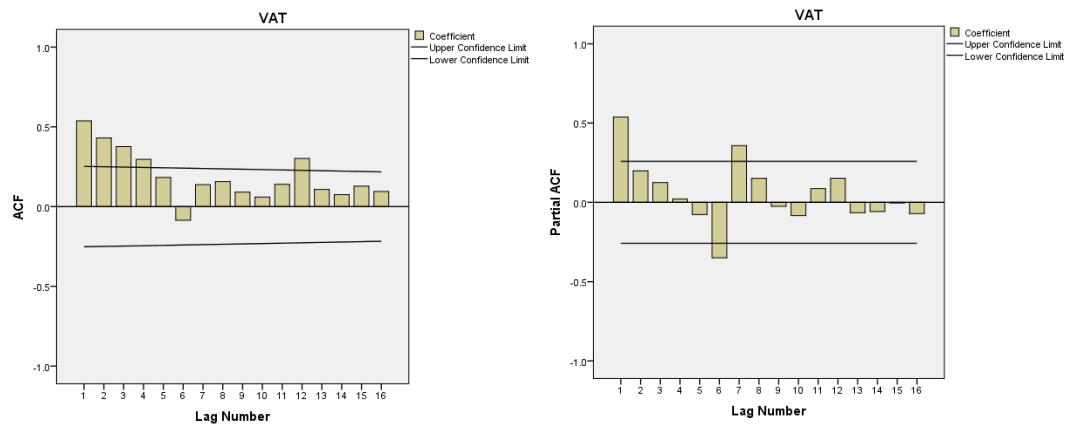Source: Processed data by the author

Figure 4 ACF and PACF Plot
Source: Processed data by the author

visually analyze that the data is not stationary in mean because some ACF plots appeared to die down. We can also tell from Figure 4 that the plot does not constantly fluctuate around parallel lines. Thus, it is necessary for us to perform differentiation. The result of the data after differentiation is shown in Figure 5.

Figure 5 exhibits that the data plot constantly fluctuates along the mean line. Thus, we conclude that the data is now stationary in the mean. We can move forward by finding the best and most feasible model to do forecasting. Figure 5 exhibits that the ACF bar crosses the line on lag 1 and PACF bars show a sinusoidal pattern but crosses the line on lag 1. From this,

we find several possible models, i.e. ARIMA (1,1,1), ARI (1,1), and IMA (1,1). However, considering the time-series nature of VAT revenue data, we predict that the more feasible models are ARIMA (0,1,1), ARIMA (1,1,1), and ARIMA (1,1,0). From here, we can continue to estimate the parameters.

b. Testing Model Feasibility

After predicting three feasible forecasting models, we need to calculate the criteria of the best model among those three. The output of this calculation is shown in Table 1.

We can find the best model from three outputs by comparing the statistical value of RMSE, MAPE, MAE, and Ljung-Box of three feasible ARIMA models. The smaller the
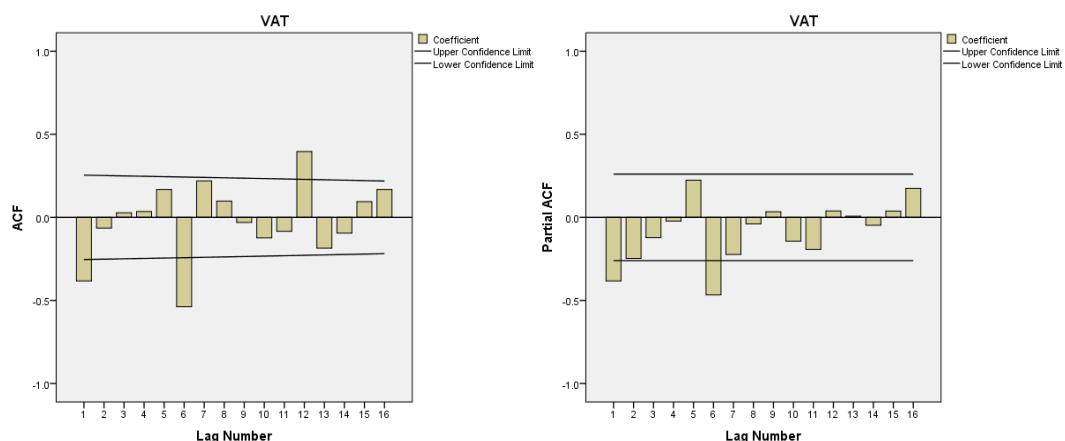

Figure 5 ACF and PACF After Data Differentiation
Source: Processed data by the author

Table 1 Comparison of Statistical Value of RMSE, MAPE, MAE, and Ljung-Box of Three ARIMA Models
(row in bold shows the best model)
Source: Processed Data by the Author

	RMSE	MAPE	MAE	Ljung-Box
ARIMA (0,1,1)	$8,403 \times 10^9$	18,178	$6,409 \times 10^9$	54,736
ARIMA (1,1,1)	**$8,359 \times 10^9$**	**17,755**	**$6,248 \times 10^9$**	**46,467**
ARIMA (1,1,0)	$8,750 \times 10^9$	18,265	$6,623 \times 10^9$	74,236

Table 2 ARIMA (1,1,1) Model Parameters
Source: Processed Data by the Author

		Estimate	SE	t	Sig.
VAT-Model_1	Constant	-52,302,636.085	283,614,207.478	-0.184	0.854
No Trans-formation	AR Lag 1	0.308	0.182	1.690	0.097
	Difference	1.000			
	MA Lag 1	0.835	0.110	7.604	0.000

statistical value, the more feasible and suitable the model with our observation data. In Table 1, we can see that the smallest statistical value came from ARIMA (1,1,1). Thus, ARIMA (1,1,1) is the best model for our time-series data.

Further feasibility testing on ARIMA (1,1,1) shows MAPE of 17.755, which is sufficient considering it is still in the range between 10 to 20. We then estimate ARIMA (1,1,1) mode parameters, as shown in Table 2.

Using the following time-series equation of ARIMA (1,1,1): (5)

$$Z_t = \mu + \varphi_1 Z_{t-1} + \cdots + \varphi_p Z_{t-p} + a_t - \psi_1 a_{t-1} - \cdots - \psi_q a_{t-q}$$

we can substitute the estimated value from Table 2, with φ_p as the coefficient of AR(p) and ψ_q as the coefficient of MA(q), as follows:

$$Z_t = -52.3(10^6) + 0.308 Z_{t-1} + a_t - 0.835 a_{t-1}$$ (6)

However, because the model is differentiated to lag 1, we expand Z_t=Y_t-Y_(t-1), as follows:

$$Y_t - Y_{t-1} = -52.3(10^6) + 0.308(Y_{t-1} - Y_{t-2}) + a_t - 0.835 a_{t-1}$$ (7)

$$\boldsymbol{Y_t = -52.3(10^6) + 1.308\, Y_{t-1}}$$

$$\boldsymbol{-0.308\, Y_{t-2} + a_t - 0.835 a_{t-1}}$$

c. Forecasting
Using our ARIMA (1,1,1) equation, we can now start to forecast VAT revenue of Surabaya Karangpilang Tax Office throughout 2022. The result of our forecast can be seen in Table 3.

Table 3 VAT Forecasted Revenue of
Surabaya Karangpilang Tax Office
Source: Processed Data by the Author

Month		VAT Forecasted Revenue
2022	1	Rp32,749,473,639
	2	Rp31,859,834,811
	3	Rp31,593,374,692
	4	Rp31,513,565,906
	5	Rp31,489,661,983
	6	Rp31,482,502,401
	7	Rp31,480,357,999
	8	Rp31,479,715,718
	9	Rp31,479,523,346
	10	Rp31,479,465,727
	11	Rp31,479,448,470
	12	Rp31,479,443,301
Total		Rp379,566,367,993

4.2 **Forecasting VAT Revenue in Jakarta Setiabudi Dua Tax Office**

We can repeat the same process to forecast VAT revenue in Jakarta Setiabudi Dua Tax Office. Observation data still ranges for the same period, which is January 2017 until December 2021, resulting in 60 observations.

a. Identifying Models

1) Data Pilotting and Stationarity

We plotted VAT revenue data from Jakarta Setiabudi Dua Tax Office to detect its stationarity, as shown in Figure 6. From
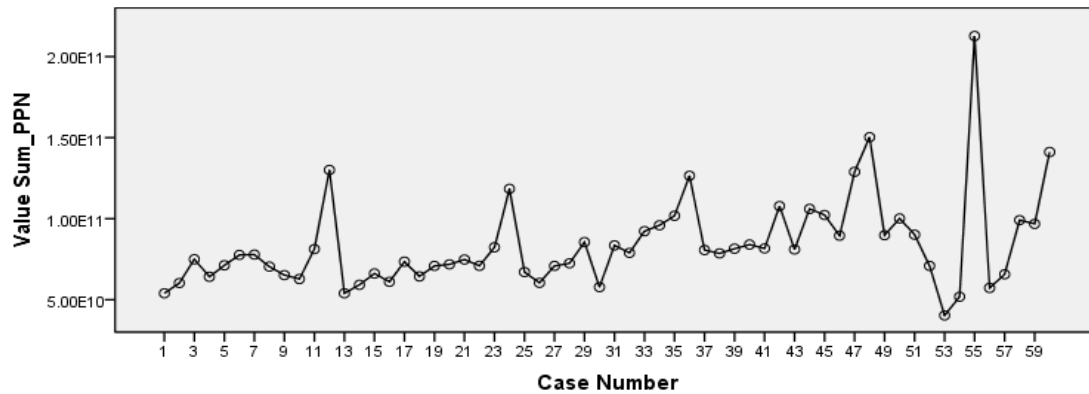

Figure 6 Jakarta Setiabudi Dua Tax Office VAT Revenue Data Plot
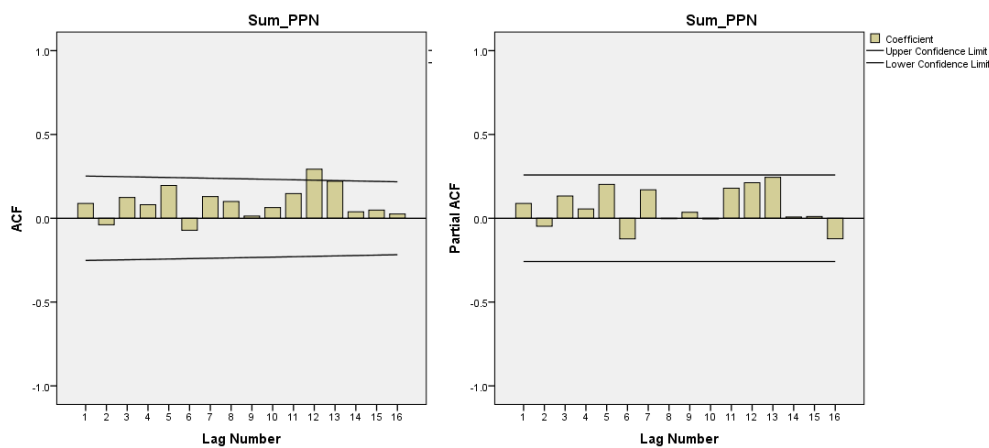Source: Processed data by the author

Figure 7 ACF and PACF Plot Before Differentiation
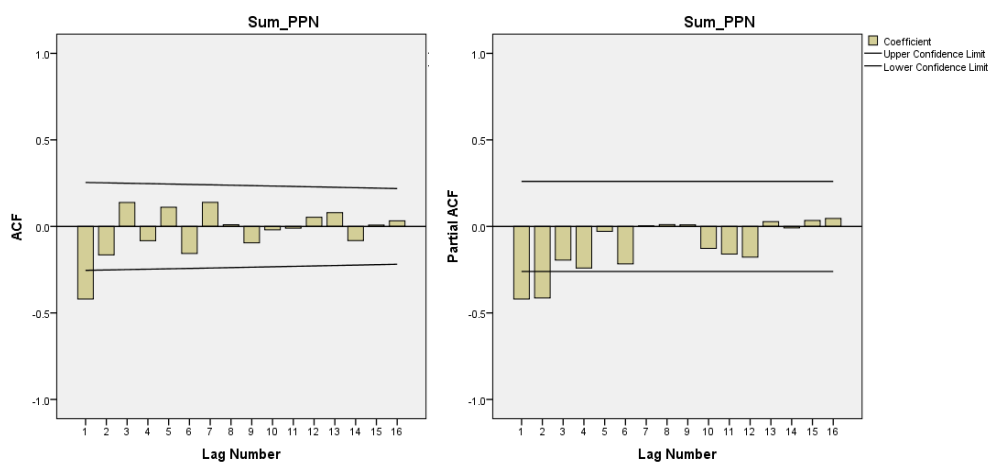Source: Processed data by the author

Figure 8 ACF and PACF After Differentiation
Source: Processed data by the author

213

Figure 6, we can tell that—different from Surabaya Karangpilang VAT revenue data—Jakarta Setiabudi Dua data is relatively more stationary in variance. Thus, we skipped differentiation and moved forward to ACF and PACF plots.

2) ACF and PACF Plot

Figure 7 depicts ACF and PACF plots for our data. We can see that dying down occurs, so our data tends to be non-stationary in the mean. We then perform order-1 differentiation on the data, resulting in a more significant down value that indicates stationarity (see Figure 8). Figure 8 exhibits that ACF shows a sinusoidal pattern and crosses the confidence limit line on lag 1, while PACF crosses the line on lag 2. Thus, we can predict possible ARIMA models, which are ARIMA (1,1,2), ARI (1,1), and IMA (1,2).

b. Testing Model Feasibility

We then compare the statistical value of the three possible ARIMA models. The comparison is shown in Table 4.

The smallest statistical value of RMSE, MAPE, MAE, and Ljung-Box is found on ARIMA (1,1,2). Thus, ARIMA (1,1,2) is the best model to forecast our data. Model estimation of ARIMA (1,1,2) is shown in Table 5.

Using the following time-series equation of ARIMA (1,1,2):

$$Z_t = \mu + \varphi_1 Z_{t-1} + \cdots + \varphi_p Z_{t-p} + a_t - \psi_1 a_{t-1} - \cdots - \psi_q a_{t-q}$$

we can substitute the estimate value from our model parameters, with φ_p as the coefficient of AR(p) and ψ_q as the coefficient of MA(q), as follows:

$$\boldsymbol{Z_t = 1.477(10^9) - 0.818\, Z_{t-1} + a_t - 0.117\, a_{t-1} + 0.883\, a_{t-2}}$$

Table 4 Comparison of Statistical Value of RMSE, MAPE, MAE, and Ljung-Box
(row in bold shows the best model)
Source: Processed Data by the Author

	RMSE	MAPE	MAE	Ljung-Box
ARIMA (1,1,2)	**2.765 x 10^{10}**	**24.358**	**1.909 x 10^{10}**	**10.071**
ARIMA (1,1,0)	3.410 x 10^{10}	24.962	2.085 x 10^{10}	19.816
ARIMA (0,1,2)	2.850 x 10^{10}	25.211	1.974 x 10^{10}	13.192

Table 5 ARIMA (1,1,2) Model Parameters
Source: Processed Data by the Author

		Estimate	SE	t	Sig.
VAT-Model_1	Constant	1,477,258,589.000	893,046,818.543	1.654	0.104
No Trans-formation	AR Lag 1	-0.818	0.121	-6.770	0.000
	Difference	1.000			
	MA Lag 1	-0.117	10.438	-0.011	0.991
	MA Lag 2	0.883	9.204	0.096	0.924

Table 6 VAT Forecasted Revenue of
Jakarta Setiabudi Tax Office
Source: Processed Data by the Author

	Month	VAT Forecasted Revenue
2022	1	Rp95,494,519,785
	2	Rp98,079,903,747
	3	Rp95,964,946,856
	4	Rp97,695,073,876
	5	Rp96,279,754,482
	6	Rp97,437,547,502
	7	Rp96,490,422,293
	8	Rp97,265,212,041
	9	Rp96,631,400,220
	10	Rp97,149,885,945
	11	Rp96,725,741,999
	12	Rp97,072,710,254
Total		Rp1,162,287,119,000

However, because the model is differentiated to lag 1, we expand $Z_t = Y_t - Y_{t-1}$, to create our ARIMA (1,1,2) equation as follows:

$$Y_t - Y_{t-1} = 1.477(10^9) - 0.818\,(Y_{t-1} - Y_{t-2}) + a_t - 0.117\,a_{t-1} + 0.883\,a_{t-2}$$
$$Y_t = \mathbf{1.477(10^9) + 0.182\,Y_{t-1} + 0.818\,Y_{t-2} + a_t - 0.117\,a_{t-1} + 0.883\,a_{t-2}}$$

c. Forecasting
The final result of Jakarta Setiabudi Dua Tax Office's forecasted VAT revenue for 2022 is shown in Table 6.

4.3 Discussion

From the forecasting results of both Surabaya Karangpilang and Jakarta Setiabudi Dua VAT revenue, we found that forecasted revenue tends to be lower than the official VAT target. For example, Surabaya Karangpilang's official VAT revenue target for 2022 is Rp403,781,162,000. Nevertheless, its forecasted VAT revenue is Rp379,566,367,993—Rp24,214,794,007, or 5.99% lower. This is also the case for Jakarta Setiabudi Dua's official and forecasted VAT revenue: the official VAT revenue target of Jakarta Setiabudi Dua Tax Office is Rp1,349,070,583,000, but the forecasted VAT revenue for the same year is Rp1,162,287,119,000—which is Rp186,783,464,000 or 13.84% lower than the official target. Official VAT targets in both tax offices refer to the revised tax revenue target after the enactment of President Decree No. 98 of 2022.

Compared to both Tax Offices' net VAT revenue (at least in the first semester of 2022), their deviations from their forecasted VAT revenues are smaller than those from their official target. For instance, Surabaya Karangpilang's net VAT revenue for the first semester of 2022 is Rp186,415,317,564. This is roughly 92.33% of half of Surabaya Karangpilang's official VAT revenue target but 97.76% compared to the forecast result of January to June 2022. The same thing happens for Jakarta Setiabudi Dua's data: VAT revenue of Jakarta Setiabudi Dua in January to June 2022 is Rp829,734,161,466, translating to 123.01% of half of its official VAT revenue target or 109.30% of forecasted VAT revenue of January to June 2022.

Seeing that the differences between net VAT revenue and forecasted revenue is closer to zero than those of net VAT revenue and official VAT target, one should not necessarily think that forecasted revenue describes actual VAT revenue pretty well. This is because the official target might have taken other external factors, while forecasted revenue from the ARIMA method only uses historical data.

Another way to interpret Figure 9 is that there might be problems with tax compliance. In Surabaya Karangpilang, there is an under-compliance situation in which actual VAT revenue falls way below the official target. Meanwhile, Jakarta Setiabudi Dua might be over-compliant, so VAT revenue is higher than the official target. Both are still valid interpretations, albeit further elaboration with more than one-year case is needed to make them even more valid. That said, this might not be about the forecasting at all: ARIMA or other forecasting methods will work just fine, but the situation is rooted in a more profound discrepancy: taxpayer's compliance and behavior.

However, while we leave room for further research to delve deeper into the compliance area, Figure 9 depicts another potential aspect to
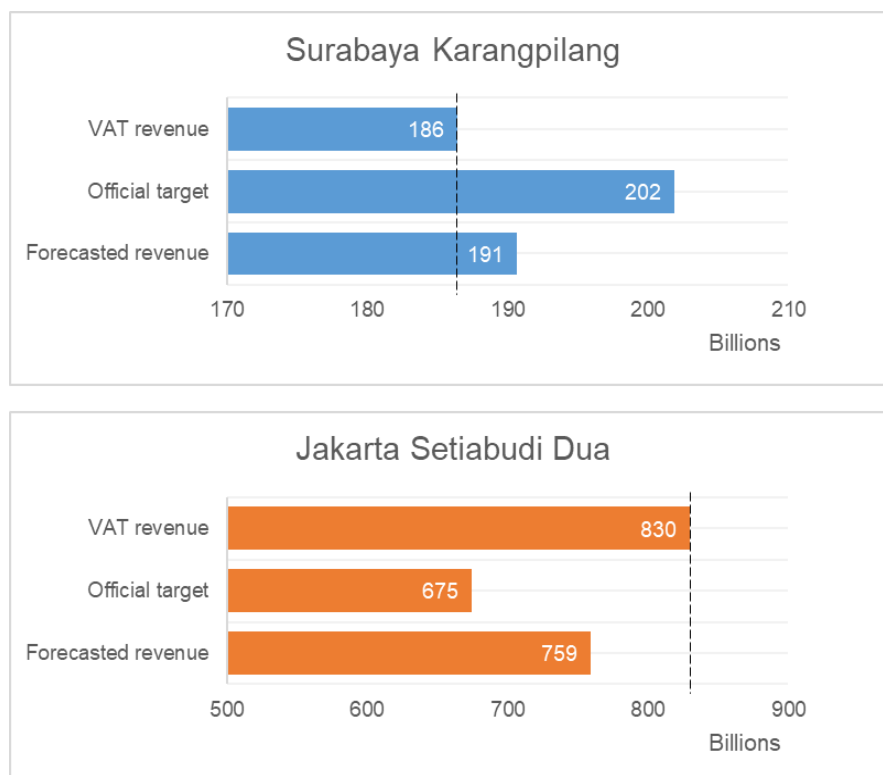
Figure 9 Comparison of Net VAT Revenue, Official VAT Target, and
Forecasted VAT Revenue of Both Tax Offices for the first half of 2022.
Numbers in billion of rupiahs
Source: Processed data by the author

elaborate on, especially regarding tax revenue target definition. We can think of forecasted revenue as tax revenue that the office should generally achieve when everything—taxpayers' compliance, tax regulations, economic growth, and geopolitical situation, among other things—goes the same as in previous years.

Thus, the forecasted tax revenue target from the ARIMA Box-Jenkins method may work as a fail-safe target, which is a target that some tax offices should normally achieve if external factors are omitted, and taxpayers pay taxes with the same abilities. Therefore, if some tax offices achieve less than their fail-safe target, these offices perform less than they historically did. On the other hand, if some tax offices achieve more than their fail-safe target, it is indicated that they perform better than they normally do. In this case, Surabaya Karangpilang's VAT revenue in the first half of 2022 still falls short of its fail-safe target—since, historically, it can achieve at least Rp5 billion more than what it achieved until June 2022. Likewise,

Jakarta Setiabudi Dua went stronger than it normally did because it achieved around Rp71 billion higher than what ARIMA predicted it could.

This fail-safe target can further be used to measure DGT's tax collection performance better. Looking back at Figure 9, if Surabaya Karangpilang achieves about Rp190 billion of VAT until June 2022, it is a little unfair to say that this tax office still falls away from its designated target. This is because compared to what it historically can achieve, it is only Rp1 billion shy. The same logic goes for Jakarta Setiabudi Dua's performance: if it achieves Rp700 billion of VAT in the first semester of 2022, we cannot say it has achieved more than its target. This is because, historically speaking, Jakarta Setiabudi Dua can still achieve Rp59 billion more.

This fail-safe target can also work as a starting point for decision-making. Heads of Tax Offices of DGT can use this fail-safe target to see how much they fall below their expected tax collection performances based on previous years

and what strategies they have to do to fulfill the gap. This should be their first priority. Meanwhile, the differences between the official tax revenue target and the forecasted tax revenue can be seen as the impact of external factors—thus, a second priority. In case of VAT, if this difference is positive, it means that the current economic situation is better than how it was in previous years. If the difference is negative, the previous years' situation was worse.

5. CONCLUSION

We forecast VAT revenue of two Tax Offices—Surabaya Karangpilang Tax Office and Jakarta Setiabudi Dua Tax Office—of DGT using the ARIMA Box-Jenkins method. Using historical VAT revenue from 2017 until 2021, after several tests, we found that ARIMA (1,1,1) is the best model for Surabaya Karangpilang, while ARIMA (1,1,2) is the best model for Jakarta Setiabudi Dua. The results show that forecasted VAT revenue resembles VAT revenue of both offices, meaning that differences between forecasted VAT revenue and actual net VAT revenue is closer than zero compared to the deviance of forecasted revenue to the official VAT revenue target defined by the Indonesian government.

We further argue that these results can be used as a fail-safe target to measure DGT's performance better. Since until now, there is no specific method to calculate and allocate tax revenue target, the result of our ARIMA Box-Jenkins forecasting can be used as an initial tool to find if certain tax offices achieve more or less than what they historically can. Especially for VAT which, by nature, highly relies on the economic situation, this fail-safe target can work as a first-priority target that should be the foundation of tax intensification strategies for tax offices in DGT.

6. LIMITATION

This research is, to many extents, limited due to many aspects, one of which is unavailability of comprehensive tax revenue data, both in terms of the scope and the details of the data. Further research is eagerly welcomed to see how ARIMA can forecast nationwide tax revenue targets for all types of taxes—not just VAT.

While we intended this research to be more experimental, further research should also compare the effectiveness and efficiency of using other forecasting methods (i.e., trend analysis) instead of ARIMA to forecast tax revenue targets.

As was also described in the 4.3 Discussion section, in terms of finding the root cause of the gap between actual tax revenue and tax revenue target, further research can delve more profound into the compliance and behavioral area. We understand that interpreting Figure 9 based on the ceteris paribus assumptions is highly flawed, and further defining it as a fail-safe target is pretty premature. That said, with more research in the tax compliance and behavioral area, we can have a more comprehensive view of a better tax revenue target definition methodology

ACKNOWLEDGEMENTS

REFERENCES

[1] Box, G., & Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. Holden-Day.

[2] Cryer, J., & Chan, K. (2008). Time Series Analysis. Springer.

[3] Daniel. (1989). Statistik Nonparametrik Terapan. Penerbit Gramedia.

[4] Direktorat Jenderal Pajak. (2021). Laporan Tahunan 2020.

[5] Hidayati, Nur. (2016). Analisis Metode Distribusi Rencana Penerimaan Pajak Per Kantor Wilayah Direktorat Jenderal Pajak. [Diploma thesis, Politeknik Keuangan Negara STAN]. Repositori Digital Politeknik Keuangan Negara STAN. http://eprints.pknstan.ac.id/id/eprint/170

[5] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4).

[6] Kurniati, D. (2022, July 27). Penerimaan Pajak Tumbuh 55,7%, Sri Mulyani: Kenaikan Luar Biasa

Kuat. DDTC News. https://news.ddtc.co.id/penerimaan-pajak-tumbuh-557-sri-mulyani-kenaikan-luar-biasa-kuat-40839

[7] Ljung, G. M., & Box, G. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2).

[8] Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1999). Metode dan Aplikasi Peramalan Jilid 1. Penerbit Erlangga.

[9] President Decree No. 96 of 2017 concerning Amendment to President Decree No. 37 of 2015 concerning Employee Performance Benefits in the Directorate General of Taxes, (2017) (testimony of President of the Republic of Indonesia).

[10] President Decree No. 98 of 2022 concerning Amendment to President Decree No. 104 of 2021 concerning Details on the State Budget for Budget Year of 2022, (2022) (testimony of President of the Republic of Indonesia).

[9] Wei, W. W. S. (2006). Time Series Analysis Univariate and Mutivariate Methods. Pearson.

[9] Wildan, M. (2022, June 30). Perpres Revisi APBN 2022 Terbit, Target Perpajakan Resmi Naik. DDTC News. Perpres Revisi APBN 2022 Terbit, Target Perpajakan Resmi Naik