Hybrid SARIMA-FFNN model in forecasting cash outflow and inflow

M Monica¹*, A Suharsono², B W Otok³, A Wibisono⁴

¹,²,³Institut Teknologi Sepuluh Nopember, Indonesia
⁴Universitas Wiraraja, Indonesia

*Email: marieta.monica@gmail.com

Abstract. The monthly inflow and outflow of money from an area is one of the important concerns in the economic life of a region. This study aims to model and predict the monthly cash inflow and outflow of Kediri, East Java Province, Indonesia using the Hybrid Seasonal Autoregressive Integrated Moving Average – Feedforward Neural Network (SARIMA-FFNN) model. Seasonal time series data from monthly cash inflow and outflow of Kediri are used to test the forecasting accuracy of the proposed hybrid model. First, both variables are modeled using the SARIMA model. Then, non-linearity testing was carried out on the best SARIMA model for each variable and the results showed that only cash inflow was non-linear. Therefore, only cash inflow could be continued with the FFNN model. The best selected model was the FFNN model with the input SARIMA(0,0,0)(1,0,0)¹² with five hidden layers. The input of FFNN modeling was based on the best SARIMA model with only the autoregressive order which for non-seasonal and seasonal. The sum of hidden layers was chosen by the smallest values of MAPE and RMSE. Forecasting results with the hybrid SARIMA-FFNN model on data testing followed the actual data pattern.

1. Introduction
The economy of a country cannot be separated from the economy of each region. In everyday life, the economy cannot be separated from the need for money. Indonesia has a central bank in charge of issuing and circulating money. The bank is known as Bank Indonesia. The bank is based in Jakarta. However, each region has Representative Office of Bank Indonesia which is abbreviated as KPw BI which regulating the money needs of each region. One example is the Representative Office of Bank Indonesia in Kediri, East Java Province, Indonesia. Bank Indonesia strives to meet the needs of money in the community, both insufficient nominal and in a condition fit for circulation.

In carrying out the task of regulating and maintaining the smooth running of the payment system, Bank Indonesia is authorized to manage currency, which includes planning, spending (outflow), circulation, revocation, and withdrawal (inflow) of currency [1]. Therefore, forecasting is needed to estimate cash outflow and inflow. Bank Indonesia uses the Autoregressive Integrated Moving Average (ARIMA) method in estimating cash outflows and inflows. Not infrequently the actual data detected a seasonal pattern caused by certain events so that the method used is the Seasonal Integrated Moving Average (SARIMA). This method is often used in several studies by Elena et al. [2] which predicts tourist arrivals in Bali; Fransiska et al. modelling monthly rainfall [3]; Borhan & Arsad [4] predict
international tourism demand from the US, Japan, and South Korea to Malaysia; Martinez et al. predict the number of the cases of dengue [5]; Cong et al. predict seasonal influenza based on SARIMA model [6]; etc. In addition to the SARIMA method, there are several other development methods for forecasting cash outflows and inflows. One of them is the study by Monica & Suharsono [7] which predicts cash outflow and inflow in Jember, East Java Province, Indonesia using the ARIMAX method.

The disadvantage of these methods is that they can only be used on linear data. Tsay and Chen [8] say that the linear model is easier to use and can provide a good approximation in many applications. However, the data does not always show a linear pattern. This deficiency is overcome by the Artificial Neural Network (ANN) method which is used in non-linear data modelling. One of the ANN methods that can be used in this research is the Feedforward Neural Network (FFNN). Machmudin & Ulama [9] showed that the ARIMA and ANN methods with the FFNN approach provided a good model compared to ARIMA to predict air temperature in Surabaya. Besides that, Hidayat et al. in his research stated that the comparison between FFNN model and time series model shows that FFNN model is better than time series model [10]. Based on these results, this study uses basic methods such as ARIMA echoed by the development method, namely FFNN so that non-linear patterns in the data can be resolved. However, the data in this study indicate a seasonal pattern, so SARIMA is used as the basic method for determining the FFNN input. Therefore, the method used in this method is called hybrid SARIMA and FFNN. This research is expected to provide a better forecasting model to predict cash outflow and inflow at the Representative Office of Bank Indonesia, Kediri, West Java, Indonesia which can later be used as one of the determinants of monetary policymaking.

2. Method

2.1. Cash Outflow and Inflow
Cash Outflow is money that comes out of Bank Indonesia through withdrawals by the public and commercial banks either directly, mobile cash, or cash deposit [11]. Meanwhile, cash inflow is money that enters Bank Indonesia through deposits by the public and commercial banks either directly, mobile cash, or cash deposit.

2.2. Seasonal Autoregressive Integrated Moving Average (SARIMA)
Unlike the ARIMA model, the SARIMA model has six components which consist of autoregressive, integrated, moving average, seasonal autoregressive, seasonal integrated, and seasonal moving average. Model identification can be done by identifying Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) graphs. In general, the SARIMA model can be written in equation (1) [12].

$$\Phi_p(B^s)\phi_p(B)(1-B)^d(1-B^s)^\delta Y_t = \theta_q(B)\Theta_q(B^s)a_t$$

where:

$$\Phi_p(B^s) = 1-\Phi_1B^s-\Phi_2B^{2s}-...-\Phi_pB^{ps}$$

$$\phi_p(B) = 1-\phi_1B-\phi_2B^2-...-\phi_pB^p$$

$$\Theta_q(B^s) = 1-\Theta_1B^s-\Theta_2B^{2s}-...-\Theta_qB^{qs}$$

$$\theta_q(B) = 1-\theta_1B-\theta_2B^2-...-\theta_qB^q$$
For convenience, we often call \( \phi_j(B) \) and \( \theta_j(B) \) the regular autoregressive and moving average factors (polynomials) and \( \Phi_j(B^s) \) and \( \Theta_j(B^s) \) the seasonal autoregressive and moving average factors (or polynomials), respectively. Subindex \( s \) refers to the seasonal period.

### 2.3. Feedforward Neural Network (FFNN)

Lazzeri said that deep learning neural networks can automatically learn arbitrary complex mappings from input to output and support multiple inputs and outputs [13]. This powerful feature can solve non-linear time series problems for forecasting both with multivalent-input and multi-step forecasting. There are several types of artificial neural networks such as Radial-Basis Function Networks, Generalized Regression Neural Networks, Feedforward Neural Networks (FFNN), etc. Based on Awchi’s research, it is said that FFNN gives better results than GRNN and RBFN [14]. Feedforward Neural Network (FFNN) is a model that can be used to predict time series data. This method is also a type of Artificial Neural Network method where the process goes forward from the input layer to the output with a directed acyclic graph [15]. Figure 1 shows the FFNN network architecture which consists of an input layer, a hidden layer, and an output layer.

![Figure 1. FFNN Network Architecture.](image)

In general, the equation of the FFNN model can be seen in equation (2).

\[
Y_t = \psi_0 + \sum_{j=1}^{J} \psi_j f \left( \tau_{j,0} + \sum_{k=1}^{K} \tau_{j,k} X_{t,k} \right) + a_t
\]

(2)

where \( K \) is the number of input nodes, \( J \) is the number of hidden nodes, \( \{ \psi_j, j = 0, 1, ..., J \} \) is the weight vector from hidden layer to output, \( \{ \tau_{j,k}, k = 0, 1, ..., K; j = 0, 1, ..., J \} \) is the weight vector from the input layer to the hidden layer.

The data used in this study are cash outflows and inflows from the Representative Office of Bank Indonesia, Kediri, East Java Province, Indonesia with a monthly period starting from January 2015 to February 2021. Sixty data from January 2015 to December 2019 were used as training data and the rest as testing data. The division of training-testing was carried out to see whether the modelling that has been carried out has followed the actual data or not. The research variables are presented in Table 1.

### Table 1. Research Variable*

| Variable         | Unit       |
|------------------|------------|
| Cash Inflow (Y_1) | Trillion Rupiah |
The steps of SARIMA and FFNN hybrid modelling are explained as follows.

a. Sharing training-testing data and identifying data characteristics using descriptive statistics,
b. Testing data stationarity, model identification, parameter estimation, testing white noise assumption,
c. Choosing the best SARIMA model,
d. Testing for linearity, if it is not linear, then proceed with the FFNN method where the input used comes from the best SARIMA model,
e. Determining the number of neurons in the hidden layer for modelling,
f. Network learning to get the FFNN model, and
g. Forecasting cash outflow and inflow.

3. Results and Discussion

3.1. Characteristic of Cash Outflow and Inflow

| Table 2. Descriptive Statistics |
|--------------------------------|
| **Cash Inflow**               |
| Mean (Trillion Rupiah)        | 1.453 |
| Standard deviation            | 0.662 |
| Minimum                       | 0.587 |
| Maximum                       | 3.623 |
| **Cash Outflow**              |
| Mean (Trillion Rupiah)        | 1.561 |
| Standard deviation            | 1.012 |
| Minimum                       | 0.199 |
| Maximum                       | 5.746 |

Based on Table 2, most of the monthly cash inflows in Kediri amounted to 1.453 trillion Rupiah from 2015 to 2021. Meanwhile, the cash outflow was 1.561 trillion Rupiah, so it can be said that the money that came out of Bank Indonesia was more than the money that came in either through direct withdrawals, mobile cash, and cash deposits by the public and commercial banks. In terms of the size of the data center, which is the average, there was not much different between the cash outflow and inflow. However, when viewed from the size of the data spread, cash outflows have more diverse values than cash inflows. This is evidenced by the value of the standard deviation of the outflow is greater than the inflow. The smallest cash inflow was 587 billion Rupiah in December 2015 and 199 billion Rupiah for cash outflow in January 2018. On the other hand, the largest cash inflow was 3.623 trillion Rupiah in June 2019 and cash outflow was 5.746 trillion Rupiah in June 2017. If you pay attention, the lowest and highest values for both inflow and outflow occur in certain months such as Christmas in December, New Year in January, and Eid al-Fitr in June 2017 and June 2019. These patterns can be seen in Figure 2 where there are spikes in certain months. These indicates that there are seasonal patterns that occurred.
3.2. Cash Outflow and Inflow with SARIMA Modelling

The stationarity of the data needs to be tested first before carrying out time series analysis. Data is said to be stationary if the mean and variance of the data do not change systematically over time. Two things need to be tested in the stationarity of the data, namely the variance and the average. Based on the results of stationarity test of variance, both variables in this study were transformed by the Box-Cox transformation first so that they are stationary for variance. The results of transformation are shown in Figure 3. The data is said to be stationary to a variance if the rounded value is 1. While the average stationarity is carried out by Augmented Dickey-Fuller test was accompanied by an alternative hypothesis, namely stationary data, and obtained a p-value of 1% (less than the 5% level of significance) so that it can be said that the data is stationary to the average.

The identification of the ARIMA model is carried out using ACF and PACF plots as shown in Figure 4. Based on Figure 4, it can be seen that at lag 12, both ACF and PACF are significant. This indicates a seasonal pattern in outflow and inflow data. After that, seasonal differentiation was carried out to obtain a model that was formed for each variable. The identification of the two models has the same form as a tentative model. Testing the white noise assumption was conducted before determining the best model. The result has shown that all tentative models for both variables meet the white noise assumption. The best model obtained for the same cash outflow and inflow was SARIMA (0,0,0) (1,0,0)_{12}. The selection of the best model was based on the smallest Akaike Information Criterion (AIC) value from several tentative models. Table 3 presents a description of the AIC values of each tentative model for the two variables.
The general form of the SARIMA model for the two variables can be seen in equation (3).

\[ Y_t = Y_{t-12} + a_t \]
\[ Y_t = Y_{t-12} + a_t \]

where \( Y_t \) and \( Y_{t-12} \) are the value of cash inflows and outflows at the time \( t \); \( Y_{t-1} \) and \( Y_{t-12} \) are the value of cash inflows and outflows at the time \( (t-12) \); \( a_t \) and \( a_{t-12} \) are the residuals cash inflow and outflow value at the time \( t \).

### 3.3. Cash Outflow and Inflow with FFNN Modelling

Before doing FFNN modelling, a non-linear detection test was first performed. The test was conducted to determine whether the relationship between \( Y \) and the input variable was non-linear. The non-linear detection test used was the Terasvirta test with an alternative hypothesis, namely a non-linear model. Table 4 shows the results of the Terasvirta test for each variable. The input was derived from the SARIMA modelling in sub-chapter 4.2.

| Variable            | P-Value |
|---------------------|---------|
| Cash Inflow (Y_{11})| 0.00035 |
| Cash Outflow (Y_{12})| 0.1077 |

* Row marked in grey is subject to reject the null hypothesis

Based on Table 4, it can be concluded that the cash inflow variable can be continued with FFNN modelling while cash outflow continues to use the SARIMA model for modelling and forecasting. In FFNN modelling, it is necessary to first identify many hidden layers that are used for the learning algorithm process. In this study, the number of hidden layers is limited to 10 neurons with the selection of many hidden layers based on the smallest Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. Table 5 shows the MAPE and RMSE values for each hidden layer and 5 hidden layers were selected for FFNN modelling in this study.

| Sum of Hidden Layer | MAPE   | RMSE   |
|---------------------|--------|--------|
| 1                   | 0.146965 | 0.160488 |
| 2                   | 0.114188 | 0.14001 |
| 3                   | 0.11667  | 0.140713 |
| 4                   | 0.114802 | 0.127887 |
| 5                   | 0.096346 | 0.103827 |
| 6                   | 0.102253 | 0.132121 |
The general form of the FFNN model for the inflow variables can be seen in equation (4).

\[
\tilde{Y}_t = f^a((-1.693f_{1}^a(-1.074\tilde{Y}_{t+1,12} + 0.224a_t - 0.021) + 2.540f_{2}^b(1.341\tilde{Y}_{t-1,12} - 0.725a_t - 0.157) + \\
1.613f_{3}^b(1.191\tilde{Y}_{t+1,12} - 1.109a_t - 0.631) - 1.925f_{4}^b(-0.490\tilde{Y}_{t-1,12} - 7.702a_t + 1.109) + \\
4.235f_{5}^b(0.620\tilde{Y}_{t+1,12} + 6.487a_t - 0.899) + 4.235)
\]

The FFNN diagram can be shown in Figure 5 with input from the SARIMA model and 5 neurons hidden layer.

![FFNN Diagram](image)

**Figure 5.** FFNN Plot for Cash Inflow.

**Figure 6.** Forecasting Testing Data Cash Inflow with FFNN.

Basically, the FFNN model cannot be interpreted but is used for forecasting. Figure 6 is the result of the training-testing data forecast. Based on Figure 6, the forecast follows the actual data pattern even though it does not detect a high spike.

| Variable            | MSE   |
|---------------------|-------|
| Cash Inflow (Y_{11})| 0.015 |
| Cash Outflow (Y_{12})| 0.164 |

Looking at the MSE values of the forecasting results, it can be determined that the model can be used for forecasting. MSE values range from 0 to 1, the smaller the MSE value, the better the model. Based in Table 6, it can be said that the FFNN model for cash inflow is good enough to be used for forecasting, or in other words the FFNN model followed the actual patterns of data. The same thing was also carried out for cash outflows. Figure 7 shows the data testing forecast using SARIMA (0,0,0) (1,0,0)_{12} and the forecast followed the actual patterns of data. Based on Table 6, since the resulting MSE value for cash outflows was 0.164, it can be said that the SARIMA model followed the actual patterns of outflow data.
Furthermore, forecasting cash outflows and inflows for the next six months using SARIMA and FFNN are shown in Table 7. From Table 7, it can be concluded that the forecast for the highest inflow value occurred in May 2021, presumably due to the Eid Al-Fitr holiday occurring in that month. It is suspected that due to the Covid-19 pandemic, people tend to save rather than deposit their money in the bank. Meanwhile, the highest outflow occurred in August 2021. The cause of the spikes in these months is required further analysis, particularly related to other economic indicators such as in terms of inflation or can be detailed by the increase or decrease in prices of basic commodities and other necessities.

| Month   | Outflow | Inflow |
|---------|---------|--------|
| March 2021 | 1.281   | 1.570  |
| April 2021 | 1.203   | 1.457  |
| May 2021   | 0.976   | 2.087  |
| June 2021  | 1.724   | 1.072  |
| July 2021  | 0.886   | 1.928  |
| August 2021| 2.061   | 1.460  |

4. Conclusion
Hybrid SARIMA and FFNN models were used for modelling and forecasting the cash inflows in Kediri, East Java Province, Indonesia. Meanwhile, the cash outflow was used the SARIMA (0,0,0) (1,0,0)_{12} model since, there was no non-linear pattern. The predictions can be used by the Representative Office of Bank Indonesia and the local government in Kediri to make evidence-based policies that are effective to ensure the smooth running of the payment system, the task of regulating and maintaining, as well as the planning, spending (outflow), circulation, revocation, and withdrawal (inflow) of currency in the future. Suggestions for further research are to link the cash outflow and inflow with other variables that may have significant contribution to the flows so that the increase or decrease in the cash outflow and inflow can be further analysed.

References
[1] Bank Indonesia 2012 Peraturan BI 14-7-PBI-2012 tentang Pengelolaan Uang Rupiah Lembaran Negara RI Tahun 2012, 138 (Jakarta: DPU)
[2] Maria E, Lee M H, Suharto, Hossein, Haizum N, and Bazilah N A 2012 Fuzzy Time Series and SARIMA Model for Forecasting Tourist Arrivals to Bali J. Teklogi 57
[3] Fransiska H, vianti P, and Agustina D 2019 Modelling Monthly Rainfall in Bengkulu City Using Seasonal Autoregressive Integrated Moving Average (Sarima) Case Study: Climatology Station In Bengkulu 390–395

[4] Borhan N and Arsad Z 2014 Forecasting international tourism demand from the US, Japan, and South Korea to Malaysia: A SARIMA approach in AIP Conference Proceedings 955–60

[5] Zangiacomi E, Martinez, Soares Da Silva E A, Lelis A, and Fabbro D 2011 Artigo A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of São Paulo, Brazil Um modelo SARIMA para previsão do número de casos de dengue em Campinas, Estado de São Paulo, Revista da Sociedade Brasileira de Medicina Tropical 44 436–40

[6] Cong J, Ren M, Xie S, and Wang P 2019 Predicting seasonal influenza based on SARIMA model, in mainland China from 2005 to 2018 International Journal of Environmental Research and Public Health 16

[7] Monica M and Suharso A 2021 Forecasting cash outflow and inflow in Jember with ARIMAX calendar variation effect, in AIP Conference Proceedings 2326

[8] Tsay R S and Chen R 2019 Nlinear Time Series Analysis (New York: John Wiley & Sons)

[9] Machmudin A and Ulama B S S 2012 Peramalan Temperatur Udara di Kota Surabaya dengan Menggunakan ARIMA dan Artificial Neural Network Jurnal Sains dan Seni ITS 1 118–23

[10] Hidayat Y, Sutijo B, Talib Bin Bon A, and Supian S 2016 Indonesian Financial Data Modeling and Forecasting by Using Ecometrics Time Series and Neural Network Global Journal of Pure and Applied Mathematics 12 3745-57

[11] Bank Indonesia 2019 Peraturan Bank Indonesia mor 21/10/PBI/2019 tentang Pengelolaan Uang Rupiah (Jakarta: Bank Indonesia)

[12] Wei W W S 2006 Time Ser ies Analysis (United States: Pearson Education, Inc)

[13] Lazzeri F 2021 Machine learning for time series forecasting with Python (UK: Wiley)

[14] Awchi T A 2014 River discharges forecasting in rthern Iraq using different ANN techniques Water Resources Management 28 801–14

[15] Fine T L 1999 Feedforward Neural Network Methodology (New York: Springer)