Forecasting Inflow and Outflow of Currency in Central Java using ARIMAX, RBFN and Hybrid ARIMAX-RBFN

Z F Maghfiroh, Suhartono*, H Prabowo, N A Salehah, D D Prastyo, Setiawan
Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia

*E-mail: suhartono@statistika.its.ac.id

Abstract. This research aims to forecast the inflow and outflow currency in Central Java. Inflow and outflow data contained both non-linear and linear patterns with calendar variation effects. Calendar variation model based on ARIMAX as a linear model, Radial Basis Function Network (RBFN) as a non-linear model, and hybrid ARIMAX-RBFN as a combination linear and non-linear model are used to forecast inflow and outflow of currency in Central Java. The data used in this research consists of inflow and outflow of currency in Central Java from January 2010 until June 2019. The denomination used is 32 denominations of inflow and 32 denominations of outflow currency. RMSE and sMAPE values from the out-of-sample data are used to select the best model. The results show that hybrid ARIMAX-RBFN is the best model of 19 denominations of inflow currency and 22 denominations of outflow. In general, the hybrid model tends to provide a more accurate forecast than the individual forecasting model used in this research.

1. Introduction
The growth of currency circulation is also affected by the transaction needs in the related region. In the second quarter of 2018 (April-June), Central Java was one of the three regions experiencing an acceleration of economic growth, which was recorded at 5.54% (YoY), where the other two regions were Yogyakarta at 5.09% and East Java at 5.57% [1]. The economic growth of Central Java is better than national economic growth (YoY). Central Java is a destination for travelers, so the increase of currency circulation happens in a certain period. In that period, net outflow happened. Eid holidays have a significant effect on the currency circulation in Central Java, so Eid al-Fitr becomes an exogenous variable in this research. Research by Hanim and Suhartono [2] provided that the ARIMAX method generated better forecasting results. Research by Wulansari et al. [3] about forecasting netflow using the ARIMAX method and the Radial Basis Function Network (RBFN), Eid al-Fitr as an exogenous variable, showed that ARIMAX is the best model in forecasting netflow currency. The ARIMAX method has good performance with Eid al-Fitr as an exogenous variable in previous research, so the ARIMAX method with Eid al-Fitr as an exogenous variable is used in this research.

Moshiri et al. [4] researched the comparison of inflation projections between simple econometric models, namely ARIMA, Vector Autoregressive, and Bayesian Vector Autoregression with Artificial Neural Network (ANN). The research provided that ANN projection is better than projections using a simple econometric model. Therefore, RBFN is used as a non-linear model in this research. The activation function used is the Gaussian function. Learning of the Radial Basis Function (RBF) is limited to using a combination of 1 to 5 neurons in the hidden layer.
The problem of the data is non-stationary or non-linear data, so an accurate and effective tool is needed to predict the behavior of the non-stationary or non-linear data. Makridakis and Hibon [5] stated that the advantage of combining several models is to produce better forecasting with a better degree of accuracy than other single models. Combining several models is known as the hybrid model [6, 7]. The hybrid model was used in Zhang’s research [8] that combined ARIMA as a linear component and Artificial Neural Network (ANN) as a non-linear component. The results of the research indicated that the ARIMA-ANN hybrid could improve forecasting accuracy.

There are three approaches used in this research, namely linear forecasting model, non-linear forecasting model, and hybrid model. This research uses a calendar variation model based on ARIMAX as a linear model, RBFN as a non-linear model, and hybrid ARIMAX-RBFN as a combination linear and non-linear model. The results of this research are expected to assist in planning the supply of banknotes and the distribution of money from the central bank to the regions in an appropriate amount.

The rest of the paper is organized as follows. Section 2 reviews the methodology. Section 3 presents the dataset and methodology. Section 4 presents the results and discussion. The conclusion is given in section 5.

2. Research Methods

2.1. Time Series Regression

The predictor variables used are dummy variables with a categorical scale or time series variables with a numerical scale. The output or response variable series is assumed to be $Y_t$, $t=1,2,...,n$ influenced by several input variables or predictors, with the input is a fixed variable. So it can be formulated as a linear regression model [9]. In this study, the predictor variables are dummy variables (trend, seasonal and calendar variations) shown in Equation 1.

$$Y_t = \beta t + \alpha_1 M_{1,t} + \alpha_2 M_{2,t} + \ldots + \alpha_s M_{s,t} + \gamma_1 V_{1,t} + \gamma_2 V_{2,t} + \ldots + \gamma_j V_{j,t} + \epsilon_t,$$

where $t$ is trend dummy, $M_{s,t}$, $s=1,2,...,S$ is the seasonal dummy, $V_{j,t}$, $j=1,2,...,J$ is the calendar variation dummy, and $\epsilon_t$ is the residual of time series regression model that do not necessarily follow the IIDN assumptions $(0, \sigma^2)$. 

2.2. Autoregressive Integrated Moving Average (ARIMA)

The Autoregressive Integrated Average (ARIMA) model is one of the most commonly forecasting models for time series data. ARIMA model is divided into four types, namely Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) for stationary data, and Autoregressive Integrated Moving Average (ARIMA) for non-stationary data [10]. ARMA model can be expressed as Equation 2 and Equation 3.

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q},$$

or

$$\phi_p(B)Y_t = \theta_q(B)\epsilon_t,$$

where $\epsilon_t$ is the residual that fulfills the white noise assumption. ARIMA model with the seasonal effect is called SARIMA. ARIMA model that has seasonal and non-seasonal patterns is called the multiplicative SARIMA model [10]. Equation 4 is a multiplicative SARIMA model.

$$\phi_p(B)\Phi_p(B^S)(1-B)^d(1-B^S)^d Y_t = \theta_q(B)\Theta_q(B^S)\epsilon_t.$$ 

2.3. Autoregressive Integrated Moving Average with Exogeneous Variable (ARIMAX)
ARIMA model with dummy variables such as calendar variation, trend, and seasonal is called ARIMAX (Autoregressive Integrated Moving Average with Exogenous Input) model. It is a time series model used to forecast data based on seasonal patterns. ARIMAX model with dummy variables such as calendar variations, trend, and seasonal, and \( \varepsilon_t \) as residual of time series regression that does not fulfill the white noise assumption, the residual model follows this equation \( \varepsilon_t \frac{\theta_q(B)}{\phi_p(B)} \). If the residual model does not fulfill the white noise assumption, the residual is modeled with ARIMA. Equation 5 is a mathematical equation of the ARIMAX model [11]. Figure 1 shows the flowchart of the ARIMAX model.

![Flowchart of ARIMAX Model](image)

**Figure 1.** The Flowchart of ARIMAX Model
\[ Y_t = \beta t + \alpha_1 M_{1,t} + \alpha_2 M_{2,t} + \ldots + \alpha_j M_{j,t} + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \ldots + \gamma_J Y_{t-J} + \frac{\partial_y(B)}{\partial_y(B)} - a. \]  

(5)

2.4. Radial Basis Function Network (RBFN)

RBFN consists of the number of neurons in the input layer, in the hidden layer, in the output layer, and the activation function. At each node in the hidden layer, RBFN uses the Radial Basis Function (RBF) denoted by \( \phi(r) \) (non-linear activation function). The input layer comes from several nodes that connect to the network. The hidden layer contains non-linear transformations from the input layer to the hidden layer. The output layer works linearly and generates a response that comes from the activation function of the hidden layer. The flowchart of the RBFN model is shown in Figure 2.

![Figure 2. The Flowchart of RBFN Model](image)

RBFN can achieve the optimal solution by adjusting the weights with the linear optimization method. Equations 6 and Equation 7 are examples to find the output of RBFN [12].

\[ \hat{Y}_t = F(x) = \sum_{i=1}^{N} w_i \phi(||x - x_i||), \]  

or

\[ \hat{Y}_t = F(x) = \sum_{i=1}^{N} w_i \phi(||x - x_i||), i = 1, 2, \ldots, N, \]  

(7)

where \( F(x) \) is the output of RBFN, \( w_i \) is the weight from \( i^{th} \) hidden unit to output unit, \( ||x - x_i|| \) is the Euclidean norm, and \( \phi(||x - x_i||) \) is the non-linear function (radial basis function). In this research, the activation function used is Gaussian as in Equation 8.

\[ \phi(r) = e^{-\frac{r^2}{2\sigma^2}}. \]  

(8)

Equation 9 is a mathematical equation of RBFN model with 1 input, 2 hidden nodes and 1 output.

\[ F(x) = \sum_{i=0}^{2} w_i \phi(||x - x_i||) = w_0 + w_1 \phi_1(+1) + w_2 \phi_2(+1), \]  

(9)
with \( \phi_i() = \exp \left( - \frac{||x_i - y||^2}{2\sigma^2_i} \right) \), and \( \phi_2() = \exp \left( - \frac{||x_{i1} - y||^2}{2\sigma^2_{i2}} \right) \).

2.5. Hybrid Model

This research uses a hybrid ARIMAX-RBFN. ARIMAX, a linear model, is easy to interpret. In contrast, RBFN is one of several kinds of non-linear models. It is difficult to interpret. The general form of the hybrid model is shown in Equation 10 [8].

\[
Y_t = Y_t^{(i)} + Y_t^{(n)} + \varepsilon_t,
\]

where \( Y_t^{(i)} \) is the linear component, and \( Y_t^{(n)} \) is the non-linear component.

This research uses a serial hybrid model. The hybrid model consists of two stages in modeling. The first stage is using ARIMAX, and the second stage is using RBFN. The illustration of hybrid ARIMAX-RBFN can be seen in Figure 3.

### Figure 3. The Flowchart of Hybrid ARIMAX-RBFN Model

The residuals from the linear model are supposed to contain non-linear patterns so the residuals are modeled with non-linear model. \( \alpha_t \) is the residual of \( t \)th time in linear model. The equation of residual model is shown in Equation 11.

\[
\alpha_t = Y_t - \hat{Y}_t^{(i)},
\]

where \( \hat{Y}_t^{(i)} \) is the forecasting result of the linear model at \( t \)th time and \( Y_t \) is the initial value at \( t \)th time.

Hybrid ARIMAX-RBFN model is shown by Equation 12.

\[
\hat{Y}_t = \alpha_t + \hat{Y}_t^{(n)}.
\]

2.6. Best Model Selection

The best model is selected by considering the value of Root Mean Square Error (RMSE) and symmetric
Mean Absolute Percentage Error (sMAPE) of out-sample data [13]. RMSE can be expressed as Equation 13.

\[
RMSE = \sqrt{\frac{1}{L} \sum_{l=1}^{L} (\hat{Y}_{l,t} - Y_{l,t})^2}. \tag{13}
\]

sMAPE is used because it is more stable than Mean Absolute Percentage Error (MAPE). By using sMAPE, we avoid the problem of large error when the actual data values are close to zero [5]. The value of sMAPE follows Equation 14.

\[
sMAPE = \left(1 - \frac{1}{L} \sum_{l=1}^{L} \frac{2|\hat{Y}_{l,t} - Y_{l,t}|}{|Y_{l,t}| + |\hat{Y}_{l,t}|}\right) \times 100\%. \tag{14}
\]

3. Dataset and Methodology

3.1. Dataset
This research uses secondary data, namely inflow and outflow currency from each region in Central Java. Among regions are assumed to be independent. The data are monthly data from January 2010 to June 2019. The denominations of inflow and outflow used are Rp100,000.00; Rp50,000.00; Rp20,000.00; Rp10,000.00; Rp5,000.00; Rp2,000.00 and Rp100,000.00. The data is divided into in-sample data (from January 2010 to December 2017) and out-sample data (from January 2018 to June 2019).

3.2. Research Variable
The research variables are divided into two variable, i.e. data inflow and data outflow. The variables can be seen in Table 1, where \(i = 1,2,3,4\) for each region in Central Java, namely Semarang, Solo, Purwokerto, and Tegal.

| Variable | Inflow    | Variable | Outflow   |
|----------|-----------|----------|-----------|
| \(Y_{1,t}^{(i)}\) | Rp1,000.00 | \(Y_{5,t}^{(i)}\) | Rp1,000.00 |
| \(Y_{2,t}^{(i)}\) | Rp2,000.00 | \(Y_{6,t}^{(i)}\) | Rp2,000.00 |
| \(Y_{3,t}^{(i)}\) | Rp5,000.00 | \(Y_{7,t}^{(i)}\) | Rp5,000.00 |
| \(Y_{4,t}^{(i)}\) | Rp10,000.00 | \(Y_{8,t}^{(i)}\) | Rp10,000.00 |
| \(Y_{5,t}^{(i)}\) | Rp20,000.00 | \(Y_{9,t}^{(i)}\) | Rp20,000.00 |
| \(Y_{6,t}^{(i)}\) | Rp50,000.00 | \(Y_{10,t}^{(i)}\) | Rp50,000.00 |
| \(Y_{7,t}^{(i)}\) | Rp100,000.00 | \(Y_{11,t}^{(i)}\) | Rp100,000.00 |
| \(Y_{8,t}^{(i)}\) Total inflow | | \(Y_{12,t}^{(i)}\) Total outflow |

The pattern of calendar variations follows the dates of Eid al-Fitr in a certain month and year. The dates of Eid al-Fitr from 2010 to 2019 can be seen in Table 2.
Table 2. List of Eid Al-Fitr 2010-2019

| The Dates of Eid al-Fitr | Week $j^{th}$ | Dummy Variable |
|--------------------------|---------------|----------------|
| 10-11 September 2010     | 2             | August         |
|                          |               | September      |
|                          |               | October        |
| 30-31 August 2011        | 4             | July           |
|                          |               | August         |
|                          |               | September      |
| 19-20 August 2012        | 3             | July           |
|                          |               | August         |
|                          |               | September      |
| 08-09 August 2013        | 2             | July           |
|                          |               | August         |
|                          |               | September      |
| 28-29 July 2014          | 4             | June           |
|                          |               | July           |
|                          |               | August         |
| 17-18 July 2015          | 3             | June           |
|                          |               | July           |
|                          |               | August         |
| 06-07 July 2016          | 1             | June           |
|                          |               | July           |
|                          |               | August         |
| 26-27 June 2017          | 4             | May            |
|                          |               | June           |
|                          |               | July           |
| 15-16 June 2018          | 3             | May            |
|                          |               | June           |
|                          |               | July           |
| 5-6 June 2019            | 1             | May            |
|                          |               | June           |
|                          |               | July           |

3.3. Step of Analysis

In this research, there are three methods used, i.e. ARIMAX, RBFN, and hybrid ARIMAX-RBFN. The step of analysis in this research is shown below, and the flowchart can be seen in Figure 4.

1. Analyze the characteristics of inflow and outflow in each region of Central Java and identify time series plots for each denomination.

2. Divide the data into in-sample data and out-sample data. Modeling in-sample data of each denomination in each region uses the ARIMAX method with the following steps:
   a. Modeling in-sample data of each denomination uses time series regression.
   b. Check diagnostics of the time series regression residual. If the residuals fulfill the white noise assumption, forecasting can be continued using the time series regression model, and the analysis step is complete. If the residuals do not fulfill the assumptions, continue to step (c).
   c. Estimating the order of the ARIMA model based on the ACF and PACF plots of the residuals.
   d. Check the diagnostic of ARIMAX residuals. If the residuals fulfill the white noise assumption, continue to step (e). If the residuals do not fulfill the assumptions, return to step (c).
   e. Determine the RMSE and sMAPE of out-sample data.

3. Modeling in-sample data of each denomination uses the RBFN method with the following steps:
   a. Determine the input, i.e. the significant PACF lag.
   b. Specify the centroid and width,
   c. Determine the number of neurons in the hidden layer,
   d. Calculate the weight,
   e. Obtain the RBFN model,
   f. Determine the RMSE and sMAPE of out-sample data for each denomination (the best neuron combination).

4. Modeling in-sample data of each denomination uses the hybrid ARIMAX-RBFN method with the following steps:
   a. Identify the significant lags of the PACF plot from residual ARIMAX. The lag is used as input of RBFN.
   b. Combine ARIMAX and RBFN models into hybrid ARIMAX-RBFN.
   c. Determine RMSE and sMAPE of out-sample data (the best neuron combination).

5. Compare RMSE and sMAPE of each model to get the best model.

6. Make a conclusion and suggestion.
4. Result and Discussion

4.1. Characteristics of Inflow and Outflow in Central Java
Identification characteristics of the inflow and outflow data are necessary because it aims to provide an overview of data patterns. The characteristics of inflow and outflow in each region can be seen in Table 3.

| Data   | Region         | Mean | Standard Deviation | Min  | Max   |
|--------|----------------|------|--------------------|------|-------|
| Inflow | Semarang       | 2407 | 1546               | 521  | 9396  |
|        | Solo           | 1396.8 | 802.6           | 200.4 | 4461.2 |
|        | Purwokerto     | 825.3 | 643.7             | 184.4 | 3555.3 |
|        | Tegal          | 416.5 | 327.3             | 22.3  | 1846  |
| Outflow| Semarang       | 1605 | 1448              | 27    | 9070  |
|        | Solo           | 745.2 | 737.4             | 3     | 4073.3 |
|        | Purwokerto     | 639.6 | 616.2             | 5.2   | 3463.6 |
|        | Tegal          | 580.3 | 604.2             | 1.3   | 3678.7 |

Figure 4. The Flowchart of Analysis Step
The highest average inflow in Central Java is 2407 billion rupiahs in Semarang. It is the same as the highest average outflow. It shows that the flow of banknotes is the most prevalent in Semarang. Semarang is a regional cash depot in Region V that has an important role in the money distribution in Central Java. The highest inflow occurs in the month of Eid al-Fitr and one month after Eid al-Fitr, whereas the highest outflow occurs in the month of Eid al-Fitr and one month before Eid al-Fitr as in Table 4. The increase of inflow during the month of Eid al-Fitr and one month after Eid al-Fitr is influenced by people's tendency to deposit money after Eid al-Fitr. The behavior to withdraw money for buying necessities for the holidays caused an increase in the outflow.

### Table 4. The Pattern of Inflow and Outflow Data (Billion Rupiah) in Central Java

| Data | Region   | Trend | Seasonal | Calendar Variation (Eid al-Fitr) | Before | Eid al-Fitr | After |
|------|----------|-------|----------|----------------------------------|--------|-------------|-------|
| Inflow | Semarang | Yes   | Yes      | No                               | Increase | Increase    |       |
|       | Solo     | Yes   | Yes      | No                               | Increase | Increase    |       |
|       | Purwokerto | Yes | Yes      | No                               | Increase | Increase    |       |
|       | Tegal    | Yes   | Yes      | No                               | Increase | Increase    |       |
| Outflow | Semarang | Yes   | Yes      | Increase                         | Increase | No          |       |
|        | Solo     | Yes   | Yes      | Increase                         | Increase | No          |       |
|        | Purwokerto | Yes | Yes      | Increase                         | Increase | No          |       |
|        | Tegal    | Yes   | Yes      | Increase                         | Increase | No          |       |

4.2. Inflow and Outflow Modeling in Central Java with Calendar Variation Model Based on ARIMAX

Modeling in-sample data with time series regression needs to be done first then the residual model is used to check the white noise assumption. If the residual does not comply with the assumption, the residual is modeled with ARIMA. The mathematical equation of the time series regression for inflow denomination Rp100,000.00 in Semarang is defined by Equation 15.

\[
\hat{Y}_t^{(1)} = 15.508 + 1044.3M_{1,t} + 290.423M_{2,t} + 332.733M_{3,t} + 231.245M_{4,t} + 164.812M_{5,t} + 32.68M_{6,t} + 589.525M_{7,t} + 362.605M_{8,t} - 299.379M_{9,t} + 251.053M_{10,t} + 257.305M_{11,t} - 96.044M_{12,t} + 297.745V_{1,t} + 297.745V_{2,t} + 1972.7V_{3,t} + 145.789V_{4,t} + 124.5V_{5,t} + 727.85V_{6,t} - 257.305V_{7,t} + 252.955V_{8,t} + 617.64V_{9,t} + 3011.3V_{10,t} + \]

The test results show that the residual does not fulfill the white noise assumptions, so it needs to be modeled with ARIMA. The mathematical equation of the ARIMAX model is shown in Equation 16.

\[
\hat{Y}_t^{(1)} = 15.617t + 1040.7M_{1,t} + 265.154M_{2,t} + 328.398M_{3,t} + 226.405M_{4,t} + 170.088M_{5,t} + 25.95M_{6,t} + 578.042M_{7,t} + 348.813M_{8,t} - 343.103M_{9,t} + 225.739M_{10,t} + 285.239M_{11,t} - 107.186M_{12,t} + 3017.5V_{1,t} + 2054.9V_{2,t} + 1250.6V_{3,t} - 705.508V_{4,t} - 308.716V_{5,t} + 382.517V_{6,t} + 675.419V_{7,t} + 3040.4V_{8,t} + \]

\[
\frac{1}{1 + 0.242B^t} \]

The residuals of the ARIMAX model have fulfilled the white noise and normal distribution assumption. The same steps are also done for other denominations. ARIMAX and time series regression (TSR) model for inflow data in each region can be seen in Table 5. The ARIMAX residuals from some denominations do not fulfill the normal distribution assumptions, but all residual models meet the white noise assumption. The residual models that do not comply with the normal distribution assumption are
the residual model of small denominations, such as Rp1,000.00 and Rp2,000.00. Outlier data influenced the fulfillment of the normal distribution assumption. Fulfillment assumptions do not always affect the forecasting results. Even though the model residuals do not fulfill the normal distribution assumptions, the forecasting results are not necessarily bad. The model evaluation can be done by calculating RMSE and sMAPE of out-sample data. ARIMAX models for outflow data are shown in Table 6. To select the best model for each data can use RMSE and sMAPE.

**Table 5. ARIMAX Model of Inflow Data in Central Java**

| Denomination | Region      | ARIMAX/TSR | RMSE   | sMAPE  |
|--------------|-------------|------------|--------|--------|
| Rp100,000.00 | Semarang    | ([5],0,0)  | 788.106| 25.702 |
|              | Solo        | ([2,4],0,0)| 441.762| 26.411 |
|              | Purwokerto  | TSR        | 289.084| 27.994 |
|              | Tegal       | (2,0,0)    | 204.680| 58.698 |
| Rp50,000.00  | Semarang    | TSR        | 243.341| 13.162 |
|              | Solo        | ([1,2,22],0,0)| 201.508| 22.246 |
|              | Purwokerto  | ([2,36],0,0)| 159.027| 23.805 |
|              | Tegal       | (2,0,0)    | 94.292 | 47.283 |
| Rp20,000.00  | Semarang    | TSR        | 34.278 | 36.567 |
|              | Solo        | ([1,3],0,0)| 22.123 | 44.038 |
|              | Purwokerto  | (0,[1,12,24])| 9.128  | 27.478 |
|              | Tegal       | ([12,36],0,0)| 11.683| 55.285 |
| Rp10,000.00  | Semarang    | TSR        | 31.859 | 58.264 |
|              | Solo        | ([1,3],0,0)| 21.964 | 53.519 |
|              | Purwokerto  | ([12,34,36],0,0)| 7.952  | 36.228 |
|              | Tegal       | ([1,36],0,0)| 10.968| 69.840 |
| Rp5,000.00   | Semarang    | ([3,12,23],0,0)| 11.552| 26.764 |
|              | Solo        | ([1,11],0,0)| 10.362| 42.068 |
|              | Purwokerto  | ([36],0,0) | 11.510 | 54.294 |
|              | Tegal       | TSR        | 6.122  | 53.689 |
| Rp2,000.00   | Semarang    | (0,[1,5])  | 7.2472 | 45.153 |
|              | Solo        | ([1,4],0,0)| 4.222  | 50.174 |
|              | Purwokerto  | ([1,36],0,0)| 4.236  | 59.541 |
|              | Tegal       | (1,0,0)    | 2.426  | 51.656 |
| Rp1,000.00   | Semarang    | ([3,5],1,0)| 1.194  | 192.815|
|              | Solo        | ([1,3,10],0,0)| 0.189| 83.727 |
|              | Purwokerto  | ([1,7],0,0)| 0.419  | 168.928|
|              | Tegal       | ([1,9],0,0) | 0.055  | 74.216 |
| Total inflow | Semarang    | TSR        | 797.583| 15.156 |
|              | Solo        | (0,[2,8])  | 459.479| 18.768 |
|              | Purwokerto  | TSR        | 369.289| 20.448 |
|              | Tegal       | (0,[1,6])  | 285.470| 52.108 |
Table 6. ARIMAX Model of Outflow Data in Central Java

| Denomination | Region   | ARIMAX/TSR | RMSE  | sMAPE  |
|--------------|----------|------------|-------|--------|
| Rp100,000.00 | Semarang | TSR        | 843.822 | 43.666 |
|              | Solo     | (0.0,[2,3,12]) | 331.821 | 47.509 |
|              | Purwokerto | (0.0,[12]) | 306.418 | 39.140 |
|              | Tegal    | ([3,12,15],0,0) | 401.838 | 28.617 |
| Rp50,000.00  | Semarang | ([2],0,0)  | 525.342 | 43.436 |
|              | Solo     | ([3],0,0)  | 215.628 | 49.375 |
|              | Purwokerto | ([2,3,9,27],0,0) | 159.253 | 64.166 |
|              | Tegal    | TSR        | 195.965 | 53.110 |
| Rp20,000.00  | Semarang | ([1,2,35],0,0) | 54.256 | 68.913 |
|              | Solo     | TSR        | 41.162 | 103.576 |
|              | Purwokerto | TSR       | 18.422 | 82.045 |
|              | Tegal    | TSR        | 12.025 | 108.666 |
| Rp10,000.00  | Semarang | ([1,35],0,0) | 60.220 | 72.599 |
|              | Solo     | TSR        | 44.855 | 100.635 |
|              | Purwokerto | TSR       | 18.195 | 83.581 |
|              | Tegal    | TSR        | 10.957 | 103.784 |
| Rp5,000.00   | Semarang | (0.0,[2,11]) | 36.124 | 77.586 |
|              | Solo     | TSR        | 24.578 | 82.460 |
|              | Purwokerto | TSR       | 13.241 | 89.025 |
|              | Tegal    | TSR        | 8.188 | 103.784 |
| Rp2,000.00   | Semarang | (0.0,[35]) | 12.479 | 79.303 |
|              | Solo     | (1,0,0)    | 9.177 | 97.947 |
|              | Purwokerto | ([11,12,36],0,0) | 2.754 | 76.717 |
|              | Tegal    | ([36],0,0) | 2.235 | 106.129 |
| Rp1,000.00   | Semarang | ([1,11,12],0,0) | 2.970 | 150.672 |
|              | Solo     | (0.0,[11]) | 1.872 | 178.999 |
|              | Purwokerto | (0.0,[11]) | 1.284 | 169.735 |
|              | Tegal    | ([1,11,35],0,0) | 0.498 | 147.258 |
| Total outflow| Semarang | TSR        | 1386.482 | 40.207 |
|              | Solo     | (0.0,[3])  | 664.105 | 42.268 |
|              | Purwokerto | (0.0, [1,4,8]) | 434.749 | 52.857 |
|              | Tegal    | ([1,9,12],0,0) | 523.800 | 30.443 |

4.3. Inflow and Outflow Modeling in Central Java with RBFN Model

Determination of significant lags on the PACF plot is the first step to do modeling in-sample data using RBFN. Lag 1,2,3,12,13,14,15 is the input for modeling inflow denomination the Rp100,000.00 in Semarang. The best model is a model with a combination of 3 neurons in the hidden layer. The mathematical equation of the RBFN model is shown in Equation 17. The architecture of the RBFN model can be seen in Figure 5.

\[
\hat{Y}_{t+s}^{(1)} = F(x) = w_0 + w_1\phi_1(\cdot) + w_2\phi_2(\cdot) + w_3\phi_3(\cdot) = 1666.892 - 829.184\phi_1(\cdot) - 651.94\phi_2(\cdot) + 58.731\phi_3(\cdot)
\] (17)
\[ \phi_1() = \exp \left( \frac{\|y_{i+1} - \bar{y}_{i+1}\|}{2\sigma_{i+1}} + \frac{\|y_{i+2} - \bar{y}_{i+2}\|}{2\sigma_{i+2}} + \ldots + \frac{\|y_{i+96} - \bar{y}_{i+96}\|}{2\sigma_{i+96}} \right) \quad \text{, for } i = 1, 2, \ldots, 96 \]

\[ \phi_2() = \exp \left( \frac{\|y_{i+1} - \bar{y}_{i+1}\|}{2\sigma_{i+1}} + \frac{\|y_{i+2} - \bar{y}_{i+2}\|}{2\sigma_{i+2}} + \ldots + \frac{\|y_{i+96} - \bar{y}_{i+96}\|}{2\sigma_{i+96}} \right) \quad \text{, for } i = 1, 2, \ldots, 96 \]

\[ \phi_3() = \exp \left( \frac{\|y_{i+1} - \bar{y}_{i+1}\|}{2\sigma_{i+1}} + \frac{\|y_{i+2} - \bar{y}_{i+2}\|}{2\sigma_{i+2}} + \ldots + \frac{\|y_{i+96} - \bar{y}_{i+96}\|}{2\sigma_{i+96}} \right) \quad \text{, for } i = 1, 2, \ldots, 96 \]

**Figure 5.** The Architecture of RBFN Model for Inflow of Denomination Rp100,000.00 in Semarang

RBFN model for other denominations is also carried out with the same steps. RBFN model from each denomination can be seen in Table 7. Forecasting results of the RBFN model tend to produce small values, so the RBFN model is suitable for modeling small denominations of inflow and outflow, such as Rp1,000.00 and Rp2,000.00. It is in contrast with the forecasting result of denomination Rp100,000.00 in Tegal. RBFN model is suitable to model that denomination. RBFN tends to provide a constant value of forecasting results. It can cause forecasting results not to follow the pattern of actual data. The forecasting result shows the different values from the actual data. There is a large difference between the forecasting result and actual data.
The same steps are also carried out for modeling outflow data. The combinations of neurons used are 1, 2, 3, 4, and 5 neurons in the hidden layer. Table 8 shows the RBFN model of outflow data. The evaluation model is done by calculating RMSE and sMAPE values.

Based on the analysis results, the RBFN model is suitable for modeling small denominations. It happens because small denominations tend to have little inflow and outflow values. The forecasting result of outflow data is the same as the forecasting result of inflow data. RBFN model tends to generate a small value of forecasting results. It can be proved by the value of RMSE and sMAPE. The value is not the smallest of other models. It can be found in small denomination, such as Rp10,000.00; Rp2,000.00 and Rp1,000.00.

| Denomination | Region | Neuron | RMSE  | sMAPE |
|--------------|--------|--------|-------|-------|
| Rp100,000.00 | Semarang| 3      | 1143.058 | 37.456 |
|              | Solo   | 4      | 497.216  | 33.725 |
|              | Purwokerto | 3   | 456.019  | 42.725 |
|              | Tegal  | 4      | 160.822  | 51.795 |
| Rp50,000.00  | Semarang| 1      | 607.371  | 27.624 |
|              | Solo   | 4      | 265.904  | 27.041 |
|              | Purwokerto | 5   | 275.126  | 34.267 |
|              | Tegal  | 3      | 108.389  | 43.232 |
| Rp20,000.00  | Semarang| 3      | 41.827   | 35.631 |
|              | Solo   | 4      | 28.097   | 44.511 |
|              | Purwokerto | 1   | 15.568   | 32.930 |
|              | Tegal  | 1      | 13.202   | 39.116 |
| Rp10,000.00  | Semarang| 5      | 36.030   | 50.088 |
|              | Solo   | 3      | 27.273   | 48.423 |
|              | Purwokerto | 1   | 14.675   | 42.437 |
|              | Tegal  | 1      | 10.118   | 51.177 |
| Rp5,000.00   | Semarang| 4      | 17.521   | 45.016 |
|              | Solo   | 2      | 15.194   | 51.227 |
|              | Purwokerto | 2   | 9.800    | 41.032 |
|              | Tegal  | 1      | 6.554    | 40.004 |
| Rp2,000.00   | Semarang| 2      | 6.558    | 35.441 |
|              | Solo   | 2      | 3.808    | 36.339 |
|              | Purwokerto | 3   | 2.327    | 31.371 |
|              | Tegal  | 1      | 1.484    | 30.205 |
| Rp1,000.00   | Semarang| 5      | 0.953    | 117.325 |
|              | Solo   | 5      | 0.390    | 90.025 |
|              | Purwokerto | 5   | 0.227    | 106.203 |
|              | Tegal  | 4      | 0.068    | 71.653 |
| Total inflow | Semarang| 1      | 2030.251 | 30.443 |
|              | Solo   | 3      | 967.540  | 29.844 |
|              | Purwokerto | 2   | 887.801  | 38.188 |
|              | Tegal  | 4      | 355.824  | 47.054 |

The same steps are also carried out for modeling outflow data. The combinations of neurons used are 1, 2, 3, 4, and 5 neurons in the hidden layer. Table 8 shows the RBFN model of outflow data. The evaluation model is done by calculating RMSE and sMAPE values.

Based on the analysis results, the RBFN model is suitable for modeling small denominations. It happens because small denominations tend to have little inflow and outflow values. The forecasting result of outflow data is the same as the forecasting result of inflow data. RBFN model tends to generate a small value of forecasting results. It can be proved by the value of RMSE and sMAPE. The value is not the smallest of other models. It can be found in small denomination, such as Rp10,000.00; Rp2,000.00 and Rp1,000.00.
### Table 8. RBFN Model of Outflow Data in Central Java

| Denomination | Region   | Neuron | RMSE   | sMAPE  |
|--------------|----------|--------|--------|--------|
| Rp100,000.00 | Semarang | 2      | 1414.045 | 62.108 |
|              | Solo     | 3      | 593.066  | 63.703 |
|              | Purwokerto | 2    | 461.785  | 52.907 |
|              | Tegal    | 1      | 637.495  | 63.616 |
| Rp50,000.00  | Semarang | 4      | 713.047  | 58.496 |
|              | Solo     | 2      | 314.997  | 54.449 |
|              | Purwokerto | 2   | 250.439  | 63.626 |
|              | Tegal    | 4      | 255.647  | 68.734 |
| Rp20,000.00  | Semarang | 5      | 103.162  | 102.393|
|              | Solo     | 5      | 58.084   | 111.746|
|              | Purwokerto | 5  | 28.677   | 99.569 |
|              | Tegal    | 5      | 20.139   | 116.339|
| Rp10,000.00  | Semarang | 5      | 101.092  | 103.790|
|              | Solo     | 5      | 70.150   | 124.910|
|              | Purwokerto | 3 | 35.399   | 101.246|
|              | Tegal    | 2      | 21.672   | 100.754|
| Rp5,000.00   | Semarang | 5      | 70.165   | 105.088|
|              | Solo     | 5      | 47.936   | 129.573|
|              | Purwokerto | 3 | 24.708   | 116.987|
|              | Tegal    | 2      | 14.949   | 112.596|
| Rp2,000.00   | Semarang | 2      | 28.019   | 93.167 |
|              | Solo     | 4      | 13.670   | 122.928|
|              | Purwokerto | 3 | 6.703    | 103.582|
|              | Tegal    | 1      | 4.275    | 97.986 |
| Rp1,000.00   | Semarang | 4      | 2.335    | 107.584|
|              | Solo     | 5      | 0.634    | 99.373 |
|              | Purwokerto | 4 | 0.821    | 113.786|
|              | Tegal    | 2      | 0.301    | 107.774|
| Total outflow | Semarang | 2      | 59.2612  | 2394.87|
|              | Solo     | 5      | 954.201  | 59.938 |
|              | Purwokerto | 5 | 776.422  | 55.690 |
|              | Tegal    | 4      | 856.742  | 54.061 |

### 4.4. Inflow and Outflow Modeling in Central Java with Hybrid ARIMAX-RBFN Model

Identification of the PACF plot needs to be done first for modeling data using hybrid ARIMAX-RBFN. The lag is an input for modeling residual using RBFN. In this research, the input used is the lag of the AR component of the ARIMAX model and certain lags (1, 12 and 35) as justification for trends, seasonal and calendar variations. For inflow denomination Rp 100,000 in Semarang, there are two kinds of inputs that will be used. The first input is a certain lag (1, 12 and 35), and the second input is the AR component of the ARIMAX model ([5], 0, 0).
\begin{equation}
\tilde{a}_{i,t}^{(1)} = F(x) = \omega_0 + \omega_1 \varphi(t) + \omega_2 \varphi(x) = -78.567 + 131.989 \varphi(t) - 42.189 \varphi(x)
\end{equation}

(18)

In this research, \( \tilde{a}_{i,t}^{(1)} \) is the same with \( \tilde{Y}_{i,t}^{(1,n)} \) as non-linear component so the equation for hybrid model is defined by Equation 19.

\begin{equation}
\tilde{Y}_{i,t}^{(1)} = \tilde{Y}_{i,t}^{(1,l)} + \tilde{Y}_{i,t}^{(1,n)}
\end{equation}

(19)

\( \hat{Y}_{i,t}^{(1,l)} \) is the linear component and \( \hat{Y}_{i,t}^{(1,n)} \) is the forecast result of hybrid model. The best hybrid model for inflow data can be seen in Table 9. It uses the best combination of neurons in the hidden layer. Moreover, RMSE and sMAPE from each model are shown in Table 9.

Table 9. Hybrid Model of Inflow Data in Central Java

| Denomination | Region | Neuron | RMSE  | sMAPE |
|--------------|--------|--------|-------|-------|
| Rp100,000.00 | Semarang | 5      | 785.073 | 26.428 |
|              | Solo   | 5      | 432.238 | 26.883 |
|              | Purwokerto | 1  | 289.176 | 28.091 |
|              | Tegal  | 1      | 196.447 | 57.119 |
| Rp50,000.00 | Semarang | 2      | 237.128 | 12.497 |
|              | Solo   | 3      | 180.434 | 20.435 |
|              | Purwokerto | 1  | 153.559 | 22.460 |
|              | Tegal  | 4      | 88.228  | 44.741 |
| Rp20,000.00 | Semarang | 3      | 28.656  | 38.255 |
|              | Solo   | 2      | 21.368  | 44.012 |
|              | Purwokerto | 3  | 8.542   | 26.866 |
|              | Tegal  | 5      | 11.496  | 49.830 |
| Rp10,000.00 | Semarang | 5      | 28.713  | 39.465 |
|              | Solo   | 3      | 21.663  | 51.860 |
|              | Purwokerto | 4  | 6.713   | 26.952 |
|              | Tegal  | 1      | 11.199  | 72.006 |
| Rp5,000.00  | Semarang | 5      | 10.881  | 25.339 |
|              | Solo   | 4      | 10.171  | 39.514 |
|              | Purwokerto | 5  | 10.707  | 45.717 |
|              | Tegal  | 1      | 6.424   | 59.216 |
| Rp2,000.00  | Semarang | 1      | 7.459   | 45.742 |
|              | Solo   | 3      | 3.827   | 48.472 |
|              | Purwokerto | 3  | 3.584   | 51.423 |
|              | Tegal  | 2      | 2.242   | 48.664 |
| Rp1,000.00  | Semarang | 2      | 1.023   | 168.895 |
|              | Solo   | 1      | 0.183   | 79.358 |
|              | Purwokerto | 2  | 0.425   | 163.878 |
|              | Tegal  | 4      | 0.055   | 83.002 |
| Total inflow | Semarang | 1      | 803.047 | 15.301 |
|              | Solo   | 2      | 453.163 | 18.044 |
|              | Purwokerto | 3  | 351.601 | 19.038 |
|              | Tegal  | 5      | 269.697 | 48.918 |
RMSE (billion rupiahs) and sMAPE of out-sample data can be seen in Table 9 and Table 10. These values will be compared with the values of linear and non-linear models so the best model can be obtained. The best hybrid model of outflow data can be seen in Table 10.

Table 10. Hybrid Model of Outflow Data in Central Java

| Denomination | Region | Neuron | RMSE   | sMAPE  |
|--------------|--------|--------|--------|--------|
| Rp100,000.00 | Semarang | 1      | 844.647| 43.684 |
|              | Solo    | 5      | 326.621| 46.608 |
|              | Purwokerto | 1    | 302.017| 39.405 |
|              | Tegal   | 5      | 397.897| 29.117 |
| Rp50,000.00 | Semarang | 5      | 517.445| 43.422 |
|              | Solo    | 4      | 213.496| 48.548 |
|              | Purwokerto | 2    | 165.351| 63.676 |
|              | Tegal   | 1      | 194.699| 52.640 |
| Rp20,000.00 | Semarang | 5      | 53.455 | 70.826 |
|              | Solo    | 4      | 40.892 | 101.859|
|              | Purwokerto | 5    | 18.153 | 79.503 |
|              | Tegal   | 3      | 12.119 | 97.770 |
| Rp10,000.00 | Semarang | 4      | 59.818 | 74.915 |
|              | Solo    | 3      | 44.726 | 89.469 |
|              | Purwokerto | 1   | 18.218 | 83.110 |
|              | Tegal   | 5      | 10.890 | 92.327 |
| Rp5,000.00  | Semarang | 2      | 34.931 | 83.964 |
|              | Solo    | 1      | 24.494 | 84.507 |
|              | Purwokerto | 1   | 13.080 | 90.237 |
|              | Tegal   | 5      | 7.778  | 101.670|
| Rp2,000.00  | Semarang | 2      | 12.252 | 75.325 |
|              | Solo    | 2      | 9.020  | 97.214 |
|              | Purwokerto | 4   | 2.777  | 74.176 |
|              | Tegal   | 5      | 1.980  | 97.190 |
| Rp1,000.00  | Semarang | 5      | 3.0177 | 142.437|
|              | Solo    | 4      | 1.725  | 180.857|
|              | Purwokerto | 4   | 1.291  | 174.870|
|              | Tegal   | 1      | 0.463  | 146.322|
| Total outflow| Semarang | 2      | 1369.677| 40.808 |
|              | Solo    | 3      | 670.214| 43.026 |
|              | Purwokerto | 1   | 413.401| 50.992 |
|              | Tegal   | 5      | 521.265| 30.384 |

4.5. The Best Model of Each Inflow and Outflow Data

The best model is needed to forecast inflow and outflow values for the next eighteen months (July 2019-December 2020). RMSE and sMAPE values of out-sample data are used to determine the best model. The model with the smallest RMSE and sMAPE values will be chosen to be the best model. The best model of inflow and outflow data can be found in Table 11.
Table 11. The Best Model of Each Inflow dan Outflow Data

| Denomination | Region   | The Best Model |
|--------------|----------|----------------|
|              |          | Inflow         | Outflow       |
| Rp100,000.00 | Semarang | Hybrid Model   | TSR           |
|              | Solo     | Hybrid Model   | Hybrid Model  |
|              | Purwokerto | TSR   | Hybrid Model  |
|              | Tegal    | RBFN          | Hybrid Model  |
| Rp50,000.00  | Semarang | Hybrid Model   | Hybrid Model  |
|              | Solo     | Hybrid Model   | Hybrid Model  |
|              | Purwokerto | Hybrid Model | ARIMAX        |
|              | Tegal    | Hybrid Model   | Hybrid Model  |
| Rp20,000.00  | Semarang | Hybrid Model   | Hybrid Model  |
|              | Solo     | Hybrid Model   | Hybrid Model  |
|              | Purwokerto | Hybrid Model | Hybrid Model  |
|              | Tegal    | Hybrid Model   | TSR           |
| Rp10,000.00  | Semarang | Hybrid Model   | Hybrid Model  |
|              | Solo     | Hybrid Model   | Hybrid Model  |
|              | Purwokerto | Hybrid Model | TSR           |
|              | Tegal    | RBFN          | Hybrid Model  |
| Rp5,000.00   | Semarang | Hybrid Model   | Hybrid Model  |
|              | Solo     | Hybrid Model   | Hybrid Model  |
|              | Purwokerto | RBFN  | Hybrid Model  |
|              | Tegal    | RBFN          | Hybrid Model  |
| Rp2,000.00   | Semarang | RBFN          | Hybrid Model  |
|              | Solo     | RBFN          | Hybrid Model  |
|              | Purwokerto | RBFN  | ARIMAX        |
|              | Tegal    | RBFN          | Hybrid Model  |
| Rp1,000.00   | Semarang | RBFN          | RBFN          |
|              | Solo     | Hybrid Model   | RBFN          |
|              | Purwokerto | RBFN  | RBFN          |
|              | Tegal    | ARIMAX        | RBFN          |
| Total        | Semarang | TSR           | Hybrid Model  |
|              | Solo     | Hybrid Model   | ARIMAX        |
|              | Purwokerto | Hybrid Model | Hybrid Model  |
|              | Tegal    | Hybrid Model   | Hybrid Model  |

The hybrid model is the best model for most of the inflow and outflow data. Comparison of the best models for each denomination of the inflow and outflow in Central Java. RBFN model is the best model for small denominations, such as Rp1,000.00; Rp2,000.00; and Rp5,000.00. RBFN is also the best model of inflow Rp100,000.00 in Tegal. ARIMAX and time series regression are the best models of only a few denominations shown in Table 11. Figure 6 shows that the hybrid model is generally superior to the individual models used in this research. It is in line with the result of the M4 forecasting competition, i.e. the hybrid method can improve the forecasting accuracy of the individual method [14].
5. Conclusion
In this study, hybrid ARIMAX-RBFN will be compared with TSR, ARIMAX, and RBFN as individual methods. Inflow and outflow data in Central Java have seasonal patterns, trends, and calendar variation effects (Eid al-Fitr). The hybrid model is the best model of 19 denominations of inflow and 22 denominations of outflow. ARIMAX and TSR are the best models of only a few denominations. RBFN is the best model of nine denominations of inflow, and four denominations of outflow. Based on the analysis, hybrid models are generally superior to linear and non-linear models. It shows that the hybrid model can capture linear and non-linear patterns. For further research, other hybrid linear-nonlinear models can be proposed. Outlier detection needs to be done in linear modeling. It can affect the fulfillment of the normal distribution assumptions. A spatial-temporal model can also be used in further research to forecast inflow and outflow data simultaneously in Central Java.

Acknowledgments
This research was supported by DRPM under the scheme of “Penelitian Unggulan Perguruan Tinggi”, project No. 1226/PKS/ITS/2020. The authors thank the support from General Director of DIKTI for funding and anonymous referees for their useful suggestions.

References
[1] Bank Indonesia 2018 Regional Economic and Financial Studies of Central Java Semarang: Bank Indonesia.
[2] Hanim Y and Suhartono 2015 Application of Time Series an ARIMAX for Forecasting Inflow and Outflow of Currency in East Java, DKI Jakarta and Indonesia Surabaya: Statistics Department FMIPA ITS.
[3] Wulansari E, and Suhartono 2014 Forecasting Netflow of Currency using ARIMAX and RBFN (Bank Indonesia Case Study) POMITS Journal of Science and Arts 3 (2) 73-78.
[4] Moshiri S and Cameron N 2000 Neural Networks Versus Econometric Models in Forecasting Inflation Journal of Forecasting 19 201-217.
[5] Makridakis S and Hibon M 2000 The M3-Competition: Results, Conclusions and Implications International Journal of Forecasting 16 451-476.
[6] Hajirahimi Z and Khashei M 2019 Hybrid Structures in Time Series Modeling and Forecasting: A Review Engineering Applications of Artificial Intelligence 86 83-106.
[7] Suhartono, Prabowo H, Prastyo D D and Lee M H 2019 New Hybrid Statistical Method and Machine Learning for PM_{10} Prediction Communications in Computer and Information Science 1100 142-155.
[8] Zhang G P 2003 Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model Neurocomputing 50 159-175.
[9] Suhartono, Lee M H and Hamzah NA 2010 Calendar Variation Model Based on ARIMAX for Forecasting Sales Data with Ramadhan Effect Proceedings of the Regional Conference on Statistical Sciences 10 30-41.
[10] Wei W W S 2006 Time Series Analysis: Univariate and Multivariate Methods, 2nd Edition New York: Pearson.
[11] Suhartono, Lee M H and Prastyo D D 2015 Two Levels ARIMAX and Regression Models for Forecasting Time Series Data with Calendar Variation Effects AIP Conference Proceedings 1691 150-164.
[12] Haykin S 2009 Neural Networks and Learning Machines Third Edition New York: Pearson Education, Inc.
[13] Hyndman R J and Koehler A B 2006 Another Look At Measures of Forecast Accuracy International Journal of Forecasting 22 679-688.
[14] Makridakis S, Spiliotis E and Assimakopoulos V 2018 The M4 Competition: Results, Findings, Conclusion and Way Forward International Journal of Forecasting 34 802-808