A Hybrid Generalized Space-Time Autoregressive-Elman Recurrent Neural Network Model for Forecasting Space-Time Data with Exogenous Variables

E Setyowati\textsuperscript{1,}\textsuperscript{*}, Suhartono\textsuperscript{1,}\textsuperscript{**}, D D Prastyo\textsuperscript{1,}\textsuperscript{***}

Department of Statistics, Institut Teknologi Sepuluh Nopember
Kampus ITS Sukolilo, Surabaya 60111, Indonesia

e-mail: endahsetyowati3@gmail.com, suhartono@statistika.its.ac.id, dedy-dp@statistika.its.ac.id

Abstract. This research proposes a hybrid method by combining Generalized Space-Time Autoregressive with exogenous variables and Elman Recurrent Neural Network (GSTARX-Elman RNN) to forecast space-time data. GSTAR method is used for modeling and forecasting multivariate data which including time and location factors. The modeling GSTAR with exogenous variables is to capture time series factors, i.e., trend, seasonal, and calendar variation. This method combines with Elman RNN as a nonlinear forecasting method for the data that have a nonlinear pattern. Hybrid GSTARX-Elman RNN compares with time series regression and GSTARX methods based on RMSE criteria. This research focused on simulation data that consist of a trend, seasonal, and calendar variation patterns, and using two scenarios of noise, i.e., linear and nonlinear noise. The result of these simulations showed that time series regression and GSTARX method could capture well the exogenous variables, but hybrid GSTARX-Elman RNN is a more accurate method than others. Hybrid GSTARX-Elman RNN could capture nonlinearity data pattern from these simulations. In general, the hybrid models tend to provide more accurate forecast performance than individual forecast models that it is in line with the results of the M4 forecasting competition.

Keywords: Hybrid GSTARX-Elman RNN, Time Series Regression, GSTARX, Space-Time.

1. Introduction
Modeling of spatio-temporal data or multivariate time series data containing time and location factors has grown rapidly. Forecasting method of spatio-temporal data that often used is Generalized Space-Time Autoregressive (GSTAR). GSTAR method was developing by Ruchjana in 2002 which is the development of the STAR model introduced by Pfeifer and Deutsch in 1980 [1,2,3]. GSTAR model assumes that research locations are heterogeneous, which means that by giving weights to each location can produce a space-time model with different parameters for time or location dependencies [3]. Then, the GSTAR model was developed by involving exogenous factors to capture certain patterns in data and known as the GSTARX model [4, 5].

Besides, several kinds of research combine methods (known as hybrid) between GSTAR and other forecasting methods. Similarly in univariate model, hybrid model can improve forecast accuracy, such as the ARIMA-NN model, compared to separate modeling between ARIMA and NN [6]. Previous research of spatio-temporal data using hybrid GSTARX-FFNN and GSTARX-DLNN models on simulation data containing trend, seasonal, calendar variation, and linear and nonlinear noise pattern is
also able to increase forecasting accuracy [7]. It shows that hybrid models can predict data containing linear and nonlinear patterns well. A nonlinear method that often combined with linear methods is Neural Network (NN). There are various types of models included in NN, one of which is Recurrent Neural Network (RNN).

In a recurrent neural network, neurons connect back to other neurons, and information flow is multidirectional so the activation of neurons can flow around in a loop [8]. Elman RNN is part of RNN model. This model has similarities with a multi-layer perceptron adding one or more context layers [8]. At the context layer, neurons from the hidden layer interact exclusively with other neurons on the internal network and not towards output directly. Elman RNN modeling has been done to predict short-term electricity consumption, optimizing the NN model on energy consumption data, and the hybrid EWTLSM-Elman model to predict wind speed. These results show that Elman RNN model performance is good in processing nonlinear data [9, 10, 11].

Therefore, in this research, a hybrid GSTARX-Elman RNN model was developed. GSTARX as a linear modeling involving predictor variables, such as trends, seasonal, and calendar variation, and Elman RNN model as a method to modeling nonlinear component. This research focused on the simulation data that consist of a trend, seasonal, and calendar variation patterns, by using two scenarios of noise, i.e., linear and non-linear noise series. The purpose of this research is to determine the accuracy of Hybrid GSTARX-Elman RNN to modeling data patterns that contain linear and nonlinear noise.

2. Methods

**Generalized Space-Time Autoregressive with Exogenous Variables (GSTARX)**

GSTARX model is developing of the GSTAR model, which involves exogenous or predictor factors. In general, GSTARX model uses two stages modeling, namely time series regression in series data, and GSTAR in residual of time series regression. The equation of GSTARX model with exogenous variables such as trend, seasonal, and calendar variation, is given as follows [3, 4, 12]:

\[
\begin{bmatrix}
    y_{t}^{(1)} \\
    y_{t}^{(2)} \\
    \vdots \\
    y_{t}^{(M)}
\end{bmatrix}
= \begin{bmatrix}
    \beta_{0}^{(1)} & 0 & \cdots & 0 \\
    \beta_{0}^{(2)} & \beta_{0}^{(1)} & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    \beta_{0}^{(M)} & 0 & \cdots & \beta_{0}^{(1)}
\end{bmatrix}
\begin{bmatrix}
    e_{t}^{(1)} \\
    e_{t}^{(2)} \\
    \vdots \\
    e_{t}^{(M)}
\end{bmatrix} + \begin{bmatrix}
    \alpha_{0}^{(1)} & 0 & \cdots & 0 \\
    \alpha_{0}^{(2)} & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    \alpha_{0}^{(M)} & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
    S_{t}^{(1)} \\
    S_{t}^{(2)} \\
    \vdots \\
    S_{t}^{(M)}
\end{bmatrix} + \begin{bmatrix}
    V_{t}^{(1)} \\
    V_{t}^{(2)} \\
    \vdots \\
    V_{t}^{(M)}
\end{bmatrix}
\begin{bmatrix}
    \epsilon_{t}^{(1)} \\
    \epsilon_{t}^{(2)} \\
    \vdots \\
    \epsilon_{t}^{(M)}
\end{bmatrix}
\]  

(1)

and

\[
\epsilon_{t} = \sum_{k=1}^{p} \Phi_{ik} + \sum_{j=1}^{s} \Phi_{ij} W_{ij} + \epsilon_{t}
\]

(2)

where \( i = 1, 2, \ldots, M \) refers to locations, \( t \) is trend dummy, \( S_{t,i}^{(i)} \) is seasonal dummy, and \( V_{t,i}^{(i)} \) is calendar variation dummy. Then \( \epsilon_{t} \) follows the GSTAR assumptions, \( \Phi_{ik} = \text{diag} (\varphi_{ik}^{1}, \ldots, \varphi_{ik}^{M}) \), \( \Phi_{ij} = \text{diag} (\varphi_{ij}^{1}, \ldots, \varphi_{ij}^{M}) \), \( \epsilon_{t} \) is \((M \times 1)\) noise vector following identically, independent, and multivariate normal distribution \( \mathcal{N}_{M} (\mathbf{0}, \sigma^{2} \mathbf{I}_{M}) \), and the weighted value is qualified \( w_{ij}^{(i)} = 0 \) and \( \sum_{j=1}^{s} w_{ij}^{(i)} = 1 \).

**Elman Recurrent Neural Network**

...
In general, the neurons in Feedforward Neural Network (FFNN) modeling are connecting in a layer. Likewise, Recurrent Neural Network (RNN) model has similarities to FFNN by adding one or more context layers. In the Elman RNN model, the number of neurons in the context layer is the same as in the hidden layer. Furthermore, neurons in the context layer are connecting to all neurons in the hidden layer. Neurons in the context layer used as reminders of events on the previous network by storing hidden layer neuron values then returned as additional input on the network [8].

The Elman RNN model equation using one hidden layer, $u$ hidden layer, and $v$ number of neurons is as follows:

$$
\hat{y}_t = f^0 \left[ b^0 + \sum_{j=1}^{u} w_j^h f_j^h \left( b_j^h + \sum_{i=1}^{v} w_{ji}^h z_{t-i} + \sum_{i=1}^{u} w_{ji}^h h_{i(t-i)} \right) \right]$

Where $f^0$ is activation function in the output layer, $b^0$ is bias in the output layer, $w_j^h$ is weight of the $j^{th}$ neuron in the hidden layer towards the neuron in the output layer, $f_f^j$ is activation function of the $j^{th}$ neuron in the hidden layer, $b_j^h$ is bias of the $j^{th}$ neuron in the hidden layer, $w_{ji}^h$ is the weight of the $i^{th}$ input toward the $j^{th}$ neuron in the hidden layer, $h_{i(t-k)}$ is context function, consist $h_{i(t-k)} = z_{t-k} - a_{t-k}$.

**Hybrid GSTARX-Elman RNN**

GSTARX modeling is used to capture multivariate time series data from several locations that contain several components of data pattern. This method would combine with Elman-RNN as a nonlinear forecasting method to capture nonlinear data patterns that could not be modeled by GSTARX. This method uses two stages, the first is time series regression to modeling the data with the trend, seasonal, and calendar variation, and the second is GSTAR-Elman RNN to modeling the residual of time series regression that have spatial effects. In general, hybrid GSTARX-Elman RNN model can described in Figure 1.

![Figure 1. Hybrid GSTARX-Elman RNN Modeling](image)

3. **Data**

This research is using simulation data containing trend, seasonal, calendar variation, and two scenarios noise. There are two scenarios, i.e., scenario 1 includes a linear noise pattern, and scenario 2 contains a nonlinear noise pattern. Each scenario replicated ten times. Data components combined additively with the following equation.
\[ Y_t^{(i)} = T_t^{(i)} + S_t^{(i)} + V_t^{(i)} + N_t^{(i)} \]  \hspace{1cm} (4)

where \( i = 1, 2, 3, 4 \) shows the locations.

1. Trend component

\[ T_t^{(i)} = \beta_t^{(i)} t, \]  \hspace{1cm} \text{where } \beta_t^{(i)} = 0.3, \beta_t^{(2)} = 0.25, \beta_t^{(3)} = 0.22, \text{ and } \beta_t^{(4)} = 0.18.  \]

2. Seasonal component

\[
\begin{align*}
S_t^{(1)} &= 25S_{t-1} + 28.7S_{t-2} + 30S_{t-3} + 28.7S_{t-4} + 25S_{t-5} + 20S_{t-6} + 15S_{t-7} + \ldots \\
S_t^{(2)} &= 18S_{t-1} + 21.7S_{t-2} + 23S_{t-3} + 18S_{t-4} + 13S_{t-5} + 8S_{t-6} + 4.3S_{t-7} + 3S_{t-8} + 4.3S_{t-9} + 8S_{t-10} + 13S_{t-11} + 20S_{t-12} \\
S_t^{(3)} &= 14S_{t-1} + 16.9S_{t-2} + 18S_{t-3} + 14S_{t-4} + 10S_{t-5} + 6S_{t-6} + 3.1S_{t-7} + 2S_{t-8} + 3.1S_{t-9} + 6S_{t-10} + 10S_{t-11} + 12S_{t-12} \\
S_t^{(4)} &= 13S_{t-1} + 15.2S_{t-2} + 16S_{t-3} + 15.2S_{t-4} + 13S_{t-5} + 10S_{t-6} + 7S_{t-7} + 4.8S_{t-8} + 4S_{t-9} + 4.8S_{t-10} + 7S_{t-11} + 10S_{t-12}
\end{align*}
\]

3. Calendar variation component

The calendar variation component is the data generating by the influence of certain events, for example, the Eid al-Fitr.

\[
\begin{align*}
V_t^{(1)} &= 60.22V_{t-1} + 64.78V_{t-2} + 29V_{t-3} + 19.46V_{t-4} + 27.96V_{t-5} + 32.86V_{t-6} + 32.68V_{t-7} + 32.86V_{t-8} + 60.12V_{t-9} + 83.48V_{t-10} \\
V_t^{(2)} &= 20.02V_{t-1} + 22.18V_{t-2} + 14.66V_{t-3} + 9.80V_{t-4} + 7.86V_{t-5} + 12.72V_{t-6} + 17.70V_{t-7} + 32.68V_{t-8} + 32.68V_{t-9} + 60.12V_{t-10} \\
V_t^{(3)} &= 25.08V_{t-1} + 23.52V_{t-2} + 11V_{t-3} + 8.5V_{t-4} + 8.06V_{t-5} + 12.85V_{t-6} + 16.30V_{t-7} + 32.44V_{t-8} + 32.44V_{t-9} + 60.12V_{t-10} \\
V_t^{(4)} &= 15.88V_{t-1} + 14.30V_{t-2} + 9.36V_{t-3} + 7.22V_{t-4} + 6.26V_{t-5} + 7.54V_{t-6} + 8.32V_{t-7} + 19.42V_{t-8} + 19.42V_{t-9} + 32.44V_{t-10}
\end{align*}
\]

4. Noise

Noise component in these simulations consists of linear and nonlinear noise. For linear noise patterns following GSTAR(1) model as follows:

\[ N_t^{(i)} = \phi_t^{(i)} N_{t-1} + \phi_t^{(i)} N_{t-2} + \phi_t^{(i)} N_{t-3} + \phi_t^{(i)} N_{t-4} + \phi_t^{(i)} N_{t-1} + \alpha_t^{(i)} \]  \hspace{1cm} (5)

with the parameter coefficient used in the GSTAR(1) model following the stationarity requirements of the GSTAR parameter, i.e., the eigenvalue of the parameter is less than 1.

\[
\begin{align*}
N_t^{(1)} &= 0.36N_{t-1} + 0.19N_{t-2} + 0.19N_{t-3} + 0.19N_{t-4} + a_t^{(1)} \\
N_t^{(2)} &= 0.20N_{t-1} + 0.33N_{t-2} + 0.20N_{t-3} + 0.20N_{t-4} + a_t^{(2)} \\
N_t^{(3)} &= 0.12N_{t-1} + 0.12N_{t-2} + 0.12N_{t-3} + 0.12N_{t-4} + a_t^{(3)} \\
N_t^{(4)} &= 0.16N_{t-1} + 0.16N_{t-2} + 0.16N_{t-3} + 0.37N_{t-4} + a_t^{(4)}
\end{align*}
\]

Then for nonlinear noise following the equation as follows:

\[
\begin{align*}
N_t^{(1)} &= 3N_{t-1} \times \exp(-0.25N_{t-1}^{2}) + 0.8N_{t-2} \times \exp(-0.25N_{t-2}^{2}) + 1.1N_{t-3} \times \exp(-0.25N_{t-3}^{2}) + 0.9N_{t-4} \times \exp(-0.25N_{t-4}^{2}) + a_t^{(1)} \\
N_t^{(2)} &= 1.3N_{t-1} \times \exp(-0.25N_{t-1}^{2}) + 3.5N_{t-2} \times \exp(-0.25N_{t-2}^{2}) + 0.9N_{t-3} \times \exp(-0.25N_{t-3}^{2}) + 1.1N_{t-4} \times \exp(-0.25N_{t-4}^{2}) + a_t^{(2)} \\
N_t^{(3)} &= 0.9N_{t-1} \times \exp(-0.25N_{t-1}^{2}) + 0.6N_{t-2} \times \exp(-0.25N_{t-2}^{2}) + 3N_{t-3} \times \exp(-0.25N_{t-3}^{2}) + 1.1N_{t-4} \times \exp(-0.25N_{t-4}^{2}) + a_t^{(3)} \\
N_t^{(4)} &= 1.1N_{t-1} \times \exp(-0.25N_{t-1}^{2}) + 0.8N_{t-2} \times \exp(-0.25N_{t-2}^{2}) + 0.8N_{t-3} \times \exp(-0.25N_{t-3}^{2}) + 3N_{t-4} \times \exp(-0.25N_{t-4}^{2}) + a_t^{(4)}
\end{align*}
\]

where \( \alpha_t \sim N(0, \Sigma) \) and between locations correlate with

\[
\Sigma = \begin{bmatrix}
1.00 & 0.65 & 0.70 & 0.55 \\
0.65 & 1.00 & 0.50 & 0.60 \\
0.70 & 0.50 & 1.00 & 0.75 \\
0.55 & 0.60 & 0.75 & 1.00 
\end{bmatrix}
\]
Figure 2 and Figure 3 show time series plots of simulation data of replication 1 at both scenarios. Data patterns show an increasing trend and tend to be high in certain months due to the influence of seasonal components. In these figures, the high data values are in the month of the Eid al-Fitr and the months after. It shows that the calendar variation of Eid al-Fitr affects the data pattern.

![Figure 2. Time Series Plot of Simulation Data Replication 1 of Scenario 1](image1)

![Figure 3. Time Series Plot of Simulation Data Replication 1 of Scenario 2](image2)

4. Result

4.1. Individual Model

In this subsection, both scenarios of simulation data were done modeling with individual model. The methods applied to individual model consist of time series regression, GSTAR, and Elman RNN.

4.2.1. Time Series Regression Model

Time series regression model in both scenarios follows equation (1) with applied trend, seasonal, and calendar variation. The following equation of time series regression model from scenario 1 is:
Based on the result of parameter estimation, it's shown that the estimated parameter has approached the parameter value of data generating.

4.2.2. GSTAR Model

GSTAR order used in this model is GSTAR (1) adjusted to the generating of linear noise components that follow GSTAR (1). The weight of this model is uniform weights which assume that the distances between four locations are considered the same. The result of this model in scenario 1 is the following equation:

\[
\begin{align*}
Y^{(1)}_t &= 0.29 0 0 0 \quad \varepsilon^{(1)}_t = 25.63 0 0 0 \\
Y^{(2)}_t &= 0 0.24 0 0 \quad \varepsilon^{(2)}_t = 18.41 0 0 \\
Y^{(3)}_t &= 0 0 0.21 0 \quad \varepsilon^{(3)}_t = 14.33 0 0 \\
Y^{(4)}_t &= 0 0 0 0.18 \quad \varepsilon^{(4)}_t = 12.82 0 0 \\
&+ \ldots + 20.40 0 0 0 \quad S^{(1)}_{t-1} \\
&+ 0 13.34 0 0 \quad S^{(2)}_{t-1} \\
&+ 0 0 10.08 0 \quad S^{(3)}_{t-1} \\
&+ 0 0 0 9.73 \quad S^{(4)}_{t-1} \\
\end{align*}
\]

(6)

4.2.3. Elman RNN Model

The input of Elman RNN model is using lag 1 following the previous GSTAR model. The activation function used is hyperbolic tangent in the hidden layer and linear in the output layer. The number of neurons consists of 1, 2, 3, 4, 5, 10, and 15, and selected using the smallest RMSE value. Elman RNN model equation in scenario 1 using the best number of neurons, 4, is as follows.

\[
\hat{Y}^{*} = 1.81f^{(1)}_1(z) - 1.46f^{(2)}_1(z) - 0.75f^{(3)}_1(z) - 1.39f^{(4)}_1(z)
\]

(8)

where,

\[
\begin{align*}
f^{(1)}_1(z) &= \text{tanh}(6.97Y^{(1)}_{i-1} + \ldots + 2.04Y^{(4)}_{i-1} + 2.46(Y^{(1)}_{i-1} - a^{(1)}_{i-1}) + \ldots + 1.35(Y^{(4)}_{i-1} - a^{(4)}_{i-1})) \\
f^{(2)}_1(z) &= \text{tanh}(-3.52Y^{(1)}_{i-1} + \ldots + 3.63Y^{(4)}_{i-1} - 2.26(Y^{(1)}_{i-1} - a^{(1)}_{i-1}) + \ldots + 0.72(Y^{(4)}_{i-1} - a^{(4)}_{i-1})) \\
f^{(3)}_1(z) &= \text{tanh}(6.93Y^{(1)}_{i-1} + \ldots + 1.28Y^{(4)}_{i-1} - 4.43(Y^{(1)}_{i-1} - a^{(1)}_{i-1}) + \ldots - 10.96(Y^{(4)}_{i-1} - a^{(4)}_{i-1})) \\
f^{(4)}_1(z) &= \text{tanh}(-0.70Y^{(1)}_{i-1} + \ldots - 4.68Y^{(4)}_{i-1} + 6.44(Y^{(1)}_{i-1} - a^{(1)}_{i-1}) + \ldots - 4.91(Y^{(4)}_{i-1} - a^{(4)}_{i-1}))
\end{align*}
\]

\(Y_{i-1}^{(1)}\) is standardize of \(Y_{i-1}^{(i)}\) in the \(i\)th locations

\[
\begin{align*}
\bar{Y}_{i-1}^{(1)} &= \begin{bmatrix} \bar{Y}_{i-1}^{(1)} \\ \bar{Y}_{i-1}^{(2)} \\ \bar{Y}_{i-1}^{(3)} \\ \bar{Y}_{i-1}^{(4)} \end{bmatrix}, \quad Y_{i-1}^{(1)} = \begin{bmatrix} Y_{i-1}^{(1)} \\ Y_{i-1}^{(2)} \\ Y_{i-1}^{(3)} \\ Y_{i-1}^{(4)} \end{bmatrix}, \quad Y^{(1)}_{i-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Y^{(2)}_{i-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad Y^{(3)}_{i-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Y^{(4)}_{i-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\end{align*}
\]

(9)

4.2.4. Comparisons of RMSE among Individual Model

RMSE from both scenarios shown in Table 1 shows a high enough value at time series regression model. This model is only able to capture the trend, seasonal, and calendar variation. While the noise patterns, both in scenario 1 and 2, could not be modeled only by using time series regression model, so the error value of this model is also quite high in both scenarios.

| Table 1. Average RMSE Value of Individual Model of Both Scenarios |
|------------------|------------------|------------------|
| Location         | Scenario 1       | Scenario 2       |
|                  |                  |                  |
|                  |                  |                  |
|                  |                  |                  |

\[ \text{RMSE} \]
GSTAR model shows that RMSE value is large enough because GSTAR(1,1) model used could not yet model well because it is possible can be modeled using a higher GSTAR order. The RMSE value of the Elman RNN model in Table 1 shows that the RMSE value tends to be large because the data patterns, especially calendar variation, could not be captured by the model. In this modeling, the input use is lag 1, while the lag calendar variation from the data do not include as input.

4.2. Hybrid Forecast Model

Hybrid forecast model is combine some models to correct errors in individual models. This modeling used GSTARX model and GSTARX-NN model, especially GSTARX-Elman RNN model.

4.2.1. GSTARX Model

The data patterns of a trend, seasonal, and calendar variation can be captured well in the first stage of modeling, time series regression, follows equation (6). Then residual time series regression model, \( \epsilon_t \), that contain location dependencies are modeled with GSTAR(1) following this equation:

\[
\begin{bmatrix}
\varepsilon_t^{(0)} \\
\varepsilon_t^{(1)} \\
\varepsilon_t^{(2)} \\
\varepsilon_t^{(3)} \\
\varepsilon_t^{(4)} \\
\end{bmatrix} =
\begin{bmatrix}
0.34 & 0.19 & 0.19 & 0.19 \\
0.19 & 0.15 & 0.19 & 0.19 \\
0.12 & 0.12 & 0.38 & 0.12 \\
0.15 & 0.15 & 0.15 & 0.27 \\
\end{bmatrix}
\begin{bmatrix}
\varepsilon_{t-1}^{(0)} \\
\varepsilon_{t-1}^{(1)} \\
\varepsilon_{t-1}^{(2)} \\
\varepsilon_{t-1}^{(3)} \\
\varepsilon_{t-1}^{(4)} \\
\end{bmatrix}
+ \begin{bmatrix}
a_{t-1}^{(0)} \\
a_{t-1}^{(1)} \\
a_{t-1}^{(2)} \\
a_{t-1}^{(3)} \\
a_{t-1}^{(4)} \\
\end{bmatrix}
\]

(10)

4.2.2. Hybrid GSTARX-Elman RNN Model

In this method, data patterns that contain trends, seasonal, and calendar variation are modeled through time series regression first, and it’s following equation (6). Furthermore, residuals (\( \varepsilon_t \)) are modeled using hybrid GSTAR-NN, especially GSTAR-Elman RNN. Figure 4 shows the architecture of GSTARX-Elman RNN model in scenario 1 using the best number of neurons.

![Figure 4](image-url)

**Figure 4.** The architecture of GSTARX-Elman RNN in Scenario 1

The architecture in Figure 4 explained by the following equation:

\[
\hat{\varepsilon}_t = -1.16 \times \tanh \left( -2.33e_t^{(0)} - 1.79F_t^{(1)} - 1.35e_t^{(2)} - 3.35F_t^{(2)} - 2.10e_t^{(3)} - 2.14F_t^{(3)} - 0.82e_t^{(4)} - 3.62F_t^{(4)} - 0.62(\varepsilon_{t-1} - a_t^{(0)}) \right)
\]

where the \( \hat{\varepsilon}_t \) adjusts to equation (9), and
4.3. Best Model Selection

The average RMSE ratio is used to determine the error rate of the reduced time series regression method when compared to others. If RMSE ratio is less than 1, then the method is better than time series regression. This comparison does not include the GSTAR and RNN from individual model because the RMSE values of both are very high. Figure 10 shows that hybrid GSTARX model combine with neural network, i.e., FFNN, DLNN, or Elman RNN for linear noise is similarly with GSTARX model, but those methods can model well on data containing nonlinear patterns. It is in line with the results of the M4-Competition: Results, Findings, Conclusions, and Way Forward, that hybrid forecasting models tend to be more accurate than individual forecasting models [15]. According to average RMSE ratio from scenario 2 in Figure 10, hybrid GSTARX-FFNN and hybrid GSTARX-DLNN models in location 1 can reduce RMSE of time series regression model by 99.6%, while hybrid GSTARX-Elman RNN model can reduce it by 98.8%.

![Figure 5. The Average RMSE Ratio to Time Series Regression Model in Scenario 1 and 2](image)

5. Conclusion

This research examined the performance of several forecasting methods using individual models, i.e. time series regression, GSTAR, and Elman RNN, and hybrid models, i.e. GSTARX, GSTARX-FFNN, GSTARX-DLNN, and GSTARX-Elman RNN. Hybrid GSTARX model with a neural network performs similarly with GSTARX in simulation data with a linear noise. Moreover, the result of simulation data with nonlinear noise shows that hybrid GSTARX-FFNN, GSTARX-DLNN, and GSTARX-Elman RNN give more accurate forecast than others. Hence, these results indicate that the hybrid forecasting model tends to be better than the individual model as the conclusion from the M4-Competition [15].

6. Acknowledgement

This research was funded by DRPM under the scheme of “Penelitian Tesis Magister”, project No. 793/PKS/ITS/2019. The authors thank to the General Director of DIKTI for funding and anonymous referees for their useful suggestions.

References

[1] Pfeifer P E and Deutsch S J 1980 Technometrics 22(1) 35-47.
[2] Pfeifer P E and Deutsch S J 1980 Technometrics 22(1) 397-408.
[3] Ruchjana B N 2002 Forum Statistika dan Komputasi ISSN 0853-8115 1-6.
[4] Suhartono, Wahyuningrum S R, Setiawan, & Akbar M S 2017 Malaysian Journal of Mathematical Sciences (10) 91-103.
[5] Astuti D, Ruchjana B N, and Soematini 2017 Journal of Physics: Conf. Series 893.
[6] Zhang G P 2003 *Neurocomputing* **50** 159-175.

[7] Dana I M, Suhartono, and Rahayu S P 2018 Model Hybrid GSTARX-ANN Untuk Peramalan Data Space-Time Dengan Efek Variasi Kalender (Studi Kasus: Data Inflow dan Outflow Uang Kartal di Bank Indonesia Wilayah Jawa Timur). Tesis Magister: Statistika ITS.

[8] Lewis N D 2016 *Deep Learning Made Easy with R*. South Carolina: Createspace Independent Publishing Platform.

[9] Suhartono and Endharta A J 2009 Peramalan Konsumsi Listrik Jangka Pendek dengan ARIMA Musiman dan Elman-Recurrent Neural Network. Surabaya: ITS Surabaya.

[10] Ruiz L G B, Rueda R, Cuéllar M P, and Pegalajar M C 2018 *Expert Systems with Applications*, **92** 380–389.

[11] Liu H, Mi X, and Li Y 2018 *Energy Conversion and Management* **156** 498-514.

[12] Lee M H, Suhartono, and Hamzah N A 2010 *Regional Conference on Statistical Sciences* 30-41.

[13] Suhartono 2007 *Feedforward Neural Network untuk Pemodelan Runtun Waktu* Yogyakarta: Disertasi Doktoral, Universitas Gajah Mada.

[14] Ahmad I S, Setiawan, Suhartono, and Masun N H 2015 *AIP Conference Proceedings* 1-8.

[15] Makridakis S, Spiliotis E, and Assimakopoulos V 2018 *International Journal of Forecasting* **34** 802-808.