Comparison between VAR, GSTAR, FFNN-VAR and FFNN-GSTAR Models for Forecasting Oil Production

[1]**Suhartono**[*], [2]**Dedy Dwi Prastyo**, [3]**Heri Kuswanto**, [4]**Muhammad Hisyam Lee**

[1,2,3]Department of Statistics, Institut Teknologi Sepuluh Nopember
Kampus ITS Sukolilo, Surabaya, Indonesia 60111

[4]Department of Mathematical Sciences
Faculty of Science, Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
[*]Corresponding author: suhartono@statistika.its.ac.id

Abstract Monthly data about oil production at several drilling wells is an example of spatio-temporal data. The aim of this research is to propose nonlinear spatio-temporal model, i.e. Feedforward Neural Network - VectorAutoregressive (FFNN-VAR) and FFNN - Generalized Space-Time Autoregressive (FFNN-GSTAR), and compare their forecast accuracy to linearspatio-temporal model, i.e. VAR and GSTAR. These spatio-temporal models are proposed and applied for forecasting monthly oil production data at three drilling wells in East Java, Indonesia. There are 60 observations that be divided to two parts, i.e. the first 50 observations for training data and the last 10 observations for testing data. The results show that FFNN-GSTAR(1_1) and FFNN-VAR(1) as nonlinear spatio-temporal models tend to give more accurate forecast than VAR(1) and GSTAR(1_1) as linear spatio-temporal models. Moreover, further research about nonlinear spatio-temporal models based on neural networks and GSTAR is needed for developing new hybrid models that could improve the forecast accuracy

Keywords Spatio-temporal; forecasting; oil production; VAR; GSTAR; FFNN-VAR; FFNN-GSTAR.

Mathematics Subject Classification 62M10, 62M45.

1 Introduction

Due to computational advances and increased computational power, nonlinear forecasting models have grown in a spectacular way and received a great deal of attention during the last decades. This fact also happens in field of time series modeling and econometrics. Neural Networks (NN) model is a prominent example of such nonlinear and nonparametric models that do not make assumptions about the parametric form of the functional relationship between the variables [1].

Vector Autoregressive (VAR) model is a quantitative forecasting method that frequently be applied for multivariate time series. This model could explain the linear relationship between observations at the same variable on different times (known as auto-correlation), and the linear relationship between observation at different variables on different times (known as cross-correlation). Furthermore, State-space model is a model that could be used for representing VAR model in forecasting problems [2,3].

Spatio-temporal problems are usually about multivariate data that have both correlation with the events at the previous times and correlation with the location or another space. Space-Time Autoregressive (STAR) model firstly proposed by Pfeifer and Deutsch [4,5] is a model that combines of these time and location dependencies in a multivariate time series data. Recently, the Generalized Space-Time Autoregressive (GSTAR)as extension of STAR is frequently used for forecasting spatio-temporal data [6]. Both GSTAR and VAR models are example of spatio-temporal linear models.

Several studies have been done relating to the application of GSTAR, such as Ruchjana [7] used GSTAR for petroleum production modeling. Deng and Athanasopoulos [8] applied Space-Time Autoregressive Integrated Moving Average (STARIMA) for forecasting domestic tourists in Australia, Wutsqa and Suhartono [9] used VAR-GSTAR model for forecasting the number of tourist arrivals, and Nurhayati *et al.* [10] applied GSTAR for forecasting GDP in Western European countries. Recently, Setiawan *et al.* [11] proposed GSTAR model for seasonal spatio-temporal and Suhartono *et al.* [12] developed GSTARX model by involving exogenous variables.

Most researches that related to GSTAR model for forecasting problems mainly focused on linear spatio-temporal data. The aims of this paper are to propose nonlinear spatio-temporal model based on NN concept and to compare the forecast accuracy to GSTAR and VAR models. The proposed models are Feedforward Neural Network-GSTAR (FFNN-GSTAR) and FFNN-VAR models. This research is mainly focused on comparison study to evaluate whether non-linear spatio-temporal model could improve the forecast accuracy of linear spatio-temporal model by applying Mean Squares Error (MSE) as evaluation criteria. Furthermore, data about monthly oil production at three drilling wells in East Java, Indonesia are used as a case study.

2 Methods

In this section, the linear and nonlinear spatio-temporal models that are used for forecasting oil production are presented.

2.1 VAR (Vector Autoregressive) Model

Let $Z_i(t)$ with $t \in T$, $t = \{1, 2, \ldots, T\}$ and $i = \{1, 2, \ldots, N\}$ are index oftime and variables, then the VAR model can be expressed as follows [10]:

$$\boldsymbol{\Phi_p}(B)\,\mathbf{Z}(t) = \mathbf{e}(t) \qquad (1)$$

where

$$\boldsymbol{\Phi_p}(B) = \boldsymbol{\Phi}_0 - \boldsymbol{\Phi}_1 B - \boldsymbol{\Phi}_2 B^2 - \cdots - \boldsymbol{\Phi}_p B^p$$

is the autoregressive matrix polynomials of order p, $\boldsymbol{\Phi}_0$ is nonsingular $m \times m$ matrix, and $\mathbf{e}(t)$ is m-dimensional white noise random vector mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$.

There are four iterative steps for building VAR model as the Box-Jenkins procedure, i.e. identification, parameter estimation, diagnostic checks, and forecasting [3]. The objective of identification step is to determine the order of VAR model and usually is done by using a time series plot, MCCF (Matrix Cross Correlation Function), MPCCF (Matrix Partial Cross Correlation Function), and the AIC (Akaike's Information Criterion). Parameter estimation step is performed by the method of Least Square or Maximum Likelihood. Then, diagnostic check is employed for evaluating whether the residuals of the model satisfy the assumptions. Finally, the calculation of the forecast values, both point and interval prediction, is done by using the best model.

2.2 GSTAR Model

Let $\{Z(t) : t = 0, \pm 1, \pm 2, \cdots\}$ is a space time data of N locations, then GSTAR with time order p and spatial order $\lambda_1, \lambda_1, \cdots, \lambda_p$, i.e. GSTAR $(p; \lambda_1, \lambda_1, \cdots, \lambda_p)$, in matrix notation can be written as follows [6]:

$$\mathbf{Z}(t) = \sum_{s=1}^{p} \left(\mathbf{\Phi}_{s0} + \sum_{k=1}^{\lambda_s} \mathbf{\Phi}_{sk} \mathbf{W}^{(k)} \right) \mathbf{Z}(t-s) + \mathbf{e}(t) \tag{2}$$

where $\mathbf{\Phi}_{s0} = \text{diag}\left(\phi_{s0}^1, \ldots, \phi_{s0}^N\right)$, $\mathbf{\Phi}_{sk} = \text{diag}\left(\phi_{sk}^1, \ldots, \phi_{sk}^N\right)$, $\mathbf{e}(t)$ is residual model that satisfies identically, independent, distributed with mean and covariance Σ. For example, GSTAR model with time and spatial order one for three locations is as follows:

$$Z(t) = \Phi_{10}Z(t-1) + \Phi_{11}W^{(1)}Z(t-1) + e(t) \tag{3}$$

and in matrix form, equation(3) can be written as follows:

$$\begin{pmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \end{pmatrix} = \begin{pmatrix} \phi_{10} & 0 & 0 \\ 0 & \phi_{20} & 0 \\ 0 & 0 & \phi_{30} \end{pmatrix} \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix}$$
$$+ \begin{pmatrix} \phi_{11} & 0 & 0 \\ 0 & \phi_{21} & 0 \\ 0 & 0 & \phi_{31} \end{pmatrix} \begin{pmatrix} 0 & w_{12} & w_{13} \\ w_{21} & 0 & w_{23} \\ w_{31} & w_{32} & 0 \end{pmatrix} \begin{pmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{pmatrix}. \tag{4}$$

Borovkova *et al.* [6] stated that several matrices of spatial weights or $\mathbf{W}$ are usually used in GSTAR model, i.e. uniform weight, weight based on inverse of distance between locations, weight based on normalization of cross correlation inference, and weight based on normalization of partial cross correlation inference. In general, the number of parameters in VAR is greater than in GSTAR model.

2.3 FFNN-GSTAR and FFNN-VAR Model

In this paper, two types of nonlinear spatio-temporal model are proposed for forecasting spatio-temporal data, i.e. FFNN-GSTAR and FFNN-VAR.The architecture of FFNN-GSTAR and FFNN-VAR that be used for forecasting oil production data could be seen at Figure 1 and 2, respectively.
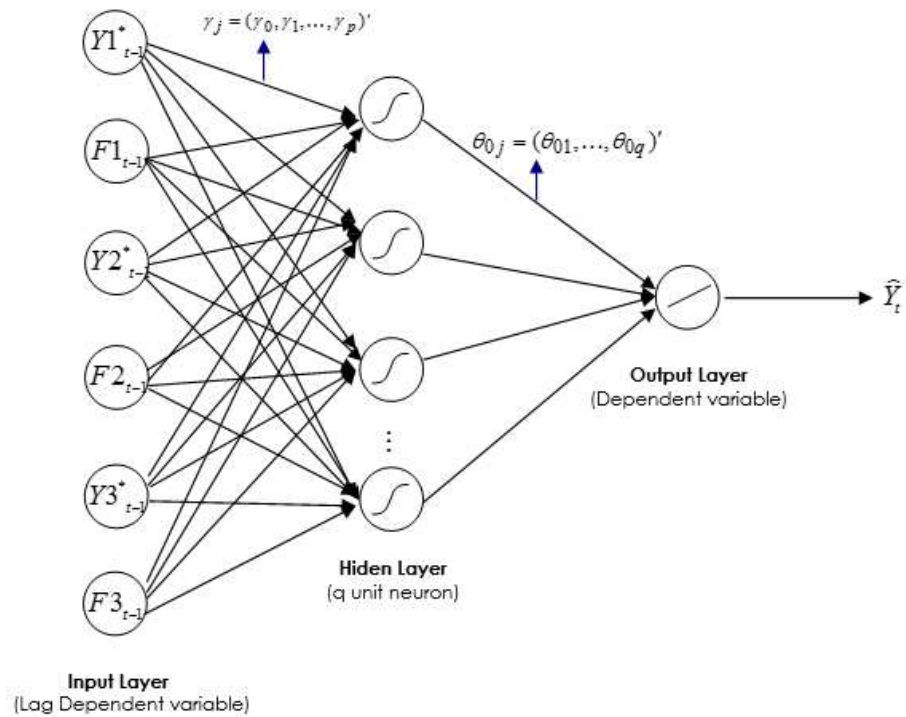
Figure 1: Architecture of FFNN-GSTAR(1_1) with one Hidden Layer, 6 Input Neurons, q Neurons in Hidden Layer, and 1 Neuron in Output Layer
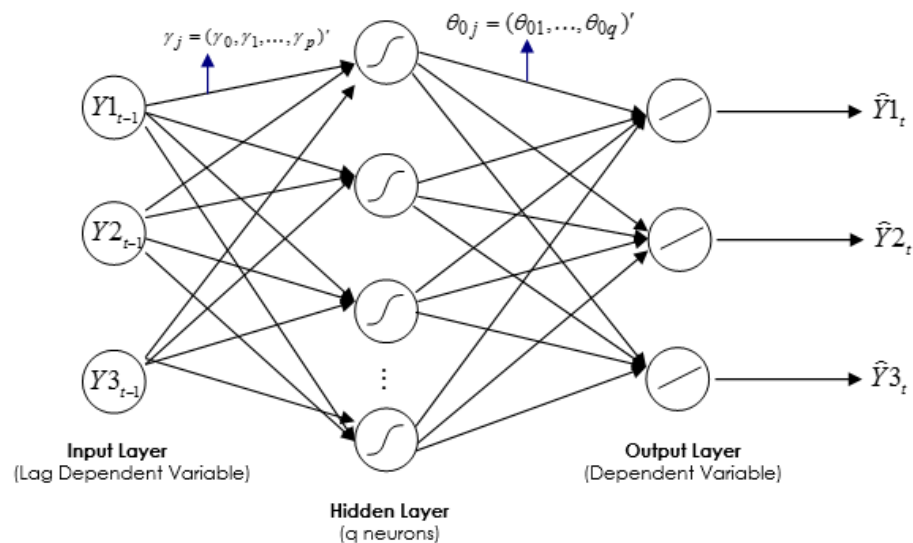

Figure 2: Architecture of FFNN-VAR(1) with One Hidden Layer, 3 Input Neurons, q Neurons in Hidden Layer, and 3 Neurons in Output Layer

The graph at Figure 1 shows that FFNN-GSTAR only uses one neuron in output layer, i.e. $\hat{\mathbf{Y}}_{\mathbf{t}}$. In general,the FFNN-GSTAR model with one hidden layer,p inputs neurons, q neurons in hidden layer,and one neuron in output layer, can be written as follows:

$$\hat{\mathbf{Y}}_{\mathbf{t}} = f(\mathbf{x_t}, \gamma, \theta) = g_2 \left\{ \sum_{\mathbf{j=0}}^{\mathbf{q}} \theta_{\mathbf{0j}}\, g_1 \left[\sum_{\mathbf{i=0}}^{\mathbf{p}} \gamma_{\mathbf{ij}} \mathbf{x_{it}} \right] \right\} \tag{5}$$

where $\mathbf{x_t} = \left\{ \mathbf{Y1^*_{t-1}, Y2^*_{t-1}, Y3^*_{t-1}, F1^*_{t-1}, F2^*_{t-1}, F3^*_{t-1}} \right\}$ are the inputs,

$$\gamma = \left\{ \gamma_{ij} \right\}_{j=1,2,...,q;\ i=1,2,...,p}$$

are the weights that connect input layer to hidden layer, $\theta = \{\theta_{0j}\}_{j=1,2,...,q}$ are the weights that connect the hidden layer to the output layer, $g_1(\cdot)$ and $g_2(\cdot)$ are the activation function in hidden layer and output layer, respectively. In this research, the logistic sigmoid and linear (identity) function are used as the activation function in hidden layer and output layer, respectively.

This output is a vector that consist of oil production data at three drilling wells and could be written as follows:

$$\hat{\mathbf{Y}}_{\mathbf{t}} = \left[\begin{array}{c} \hat{\mathbf{Y}}\mathbf{1_t} \\ \hat{\mathbf{Y}}\mathbf{2_t} \\ \hat{\mathbf{Y}}\mathbf{3_t} \end{array} \right].$$

The explanations about notations at Figure 1 are as follows:

$$\widehat{\mathbf{Y}}_t = \left[\begin{array}{c} \widehat{\mathbf{Y}}\mathbf{1_t} \\ \widehat{\mathbf{Y}}\mathbf{2_t} \\ \widehat{\mathbf{Y}}\mathbf{3_t} \end{array} \right], \quad \mathbf{Y1^*_{t-1}} = \left[\begin{array}{c} \widehat{\mathbf{Y}}\mathbf{1_{t-1}} \\ \mathbf{0} \\ \mathbf{0} \end{array} \right], \quad \mathbf{Y2^*_{t-1}} = \left[\begin{array}{c} \mathbf{0} \\ \widehat{\mathbf{Y}}\mathbf{2_{t-1}} \\ \mathbf{0} \end{array} \right], \quad \mathbf{Y3^*_{t-1}} = \left[\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ \widehat{\mathbf{Y}}\mathbf{3_{t-1}} \end{array} \right],$$

$$\mathbf{F1^*_{t-1}} = \left[\begin{array}{c} w_{12}\widehat{\mathbf{Y}}\mathbf{2_{t-1}} + w_{13}\widehat{\mathbf{Y}}\mathbf{3_{t-1}} \\ \mathbf{0} \\ \mathbf{0} \end{array} \right], \quad \mathbf{F2^*_{t-1}} = \left[\begin{array}{c} \mathbf{0} \\ w_{21}\widehat{\mathbf{Y}}\mathbf{1_{t-1}} + w_{23}\widehat{\mathbf{Y}}\mathbf{3_{t-1}} \\ \mathbf{0} \end{array} \right], \text{ and}$$

$$\mathbf{F3^*_{t-1}} = \left[\begin{array}{c} \mathbf{0} \\ \mathbf{0} \\ w_{31}\widehat{\mathbf{Y}}\mathbf{1_{t-1}} + w_{32}\widehat{\mathbf{Y}}\mathbf{2_{t-1}} \end{array} \right].$$

Moreover, it is different with Figure 2 that consist of three neurons in output layer. Neuron 1, 2, and 3 at output layer in this FFNN-VAR show about the oil production at drilling wells 1, 2 and 3, respectively. In general, the FFNN-VAR model with one hidden layer, p input neurons, q neurons in hidden layer, and 3 neurons in output layer, can be written as follows:

$$\hat{\mathbf{Y}}_{kt} = f(\mathbf{x_t}, \gamma, \theta) = g_2 \left\{ \sum_{j=0}^{q} \theta_{k\mathbf{j}}\, g_1 \left[\sum_{i=0}^{p} \gamma_{\mathbf{ij}} \mathbf{x_{it}} \right] \right\} \tag{6}$$

where $\mathbf{x_t} = \{\mathbf{Y1_{t-1}, Y2_{t-1}, Y3_{t-1}}\}$ are the inputs, $\gamma = \{\gamma_{ij}\}_{j=1,2,...,q;\ i=1,2,...,p}$ are the weights that connect input layer to hidden layer, $\theta = \{\theta_{kj}\}_{j=1,2,...,q}$ are the weights that connect the

hidden layer to the k-th neuron in output layer, $g_1(\cdot)$ and $g_2(\cdot)$ are the activation function in hidden layer and output layer, respectively.

Hence,the main differences between FFNN-GSTAR and FFNN-VAR architectures are the number of inputs and outputs. For example, the FFNN-GSTAR(1_1) model has 6 inputs that represent lags observation in the same location (time effect) and difference location (space effect) as GSTAR representation in Equation (4). Otherwise, the FFNN-VAR(1) model consists of 3 inputs as in Figure 2. Thus, determination of inputs in both FFNN-GSTAR and FFNN-VAR is done by decomposing the model and representing it as in Figure 1 and Figure 2.

The training process to estimate the optimal weights are done by the back propagation algorithm. In this research, **nnet** package in R that employ BFGS (Broyden, Fletcher, Goldfard and Shanno) training algorithm is used for back propagation optimization to find the optimal weights in both FFNN-GSTAR and FFNN-VAR.

3 Results

In this section, we firstly present the data description, then the results of comparison study.

3.1 Data

Spatio-temporal data about oil production at three drilling wells in West Java, Indonesia are used as case study. These data are monthly data that consist of 60 observations and then be divided become two parts, i.e. 50 and 10 observations for training and testing dataset, respectively. The time series plot of these monthly oil production data after centering, i.e. subtracted by the mean, are illustrated at Figure 3.
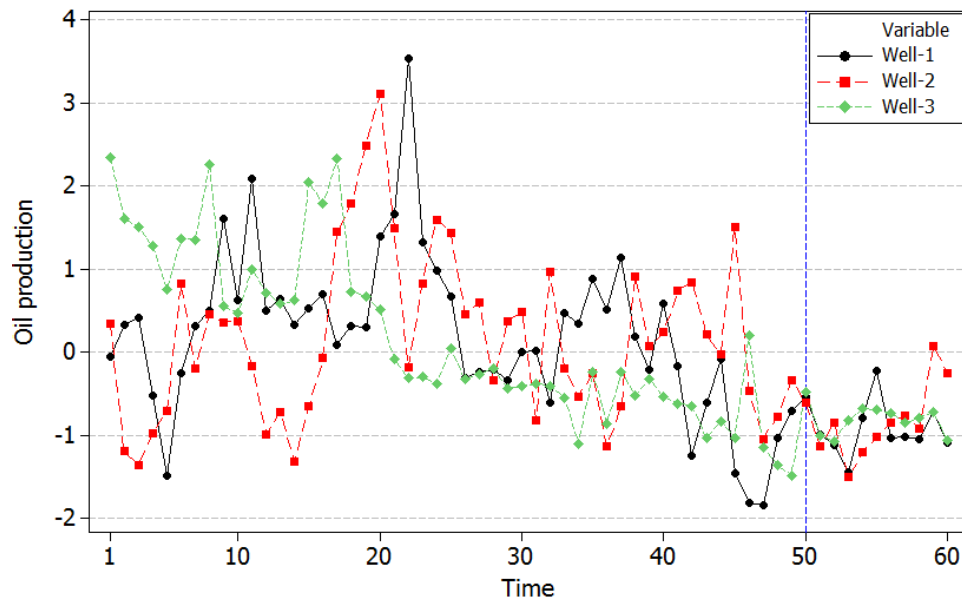

Figure 3: Time Series Plot of Oil Production Data at Three Drilling Wells after Centering by the Mean

3.2 The Comparison of Forecast Accuracy

The identification, estimation and diagnostic check steps show that the best model for forecasting these oil production data are VAR(1) and GSTAR(1_1). Due to the location of three drilling wells are at the same area with relative similar distance, the uniform weight is used as a matrix of spatial weights or **W** in this GSTAR(1_1) model. Furthermore, the FFNN-VAR(1) and FFNN-GSTAR(1_1) architectures are developed based on these VAR and GSTAR models, respectively.

The results of forecast accuracy comparison on these four spatio-temporal models could be seen at Table 1.

Table 1: The Comparison Results of Forecast Accuracy between GSTAR, VAR, FFNN-GSTAR, and FFNN-VAR at Oil Production Data

Model	MSE at Training Data			MSE at Testing Data		
	Y1	**Y2**	**Y3**	**Y1**	**Y2**	**Y3**
1. GSTAR(1_1)	0.497	0.649	0.310	0.177	0.255	0.092
2. VAR(1)	0.498	0.6339	0.402	0.207	0.258	0.097
3. FFNN-GSTAR(1_1)						
- 1 neuron	0.487	0.651	0.247	0.179	0.240	0.063
- 2 neurons	0.376	0.541	0.226	0.251	0.314	0.042
- 3 neurons	0.333	0.447	0.174	0.505	0.335	0.501
- 4 neurons	**0.271**	**0.361**	**0.180**	**0.143**	**0.348**	**0.036**
- 5 neurons	0.198	0.325	0.138	0.271	0.627	0.079
4. FFNN-VAR(1)						
- 1 neuron	0.710	0.966	0.400	0.278	0.789	0.044
- 2 neurons	0.617	0.596	0.219	0.399	0.797	0.036
- 3 neurons	0.456	0.435	0.208	0.178	0.618	0.046
- 4 neurons	**0.485**	**0.212**	**0.265**	**0.140**	**0.258**	**0.074**
- 5 neurons	0.395	0.221	0.150	0.330	0.731	0.088

Furthermore, it could be concluded that in general FFNN-GSTAR(1_1) and FFNN-VAR(1) as nonlinear spatio-temporal models give more accurate forecast than GSTAR(1_1) and VAR(1) as linear spatio-temporal models. These results are supported by the forecast of FFNN-GSTAR(1_1) and FFNN-VAR(1) with four neurons in hidden layer that tend to yield more accurate forecast at all three locations in training dataset, and two locations in testing dataset. Hence, the results also give new opportunity to do further research about NN for forecasting

spatio-temporal that be developed from both GSTAR and VAR models. Moreover, these results also in line with the third results and conclusion of M3 competition, i.e. the accuracy when various methods are being combined outperforms, on average, the individual methods being combined and does very well in comparison to other methods [13]. Additionally, the results also showed that FFNN-GSTAR with smaller number of weights (parameters) than FFNN-VAR yield mostly similar forecast with FFNN-VAR results. This result also in line with the first results of M3 competition, i.e. sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones.

4 Conclusion

The present study was designed to propose nonlinear spatio-temporal model, i.e. FFNN-VAR and FFNN-GSTAR, for forecasting oil production data and to compare their forecast accuracy to linear spatio-temporal model, i.e. VAR and GSTAR. Based on the empirical results, it could be concluded that the proposed nonlinear spatio-temporal models, i.e. FFNN-GSTAR(1_1) and FFNN-VAR(1), could work well for forecasting oil data production. Moreover, the results also showed that FFNN-GSTAR(1_1) and FFNN-VAR(1) tend to give more accurate forecast than VAR(1) and GSTAR(1_1) both at training and testing datasets. Additionally, the results showed that FFNN-GSTAR with smaller number of parameters (weights) than FFNN-VAR yield similar accurate forecast as FFNN-VAR results. Furthermore, more research is needed for validating these initial results, particularly by performing simulation study and comparing the results to more datasets.

Acknowledgement

References

[1] Suhartono. *Feedforward Neural Network for Time Series Forecasting.* Gadjah Mada University: Ph.D. Thesis. 2007.

[2] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. *Time Series Analysis: Forecasting and Control.* 3rd edition, Englewood Cliffs: Prentice Hall. 1994.

[3] Wei, W. W. S. *Time Series Analysis: Univariate and Multivariate Methods.* Addison-Wesley Publishing Co., USA. 2006.

[4] Pfeifer, P. E. and Deutsch, S. J. A Three stage iterative procedure for space-time modeling. *Technometrics.* 1980a. 22(1): 35-47.

[5] Pfeifer, P. E. and Deutsch, S. J. Identification and interpretation of first order space-time ARMA models. *Technometrics.* 1980b. 22(1): 397-408.

[6] Borovkova, S. A., Lopuhaa, H. P., and Nurani, B. Generalized STAR model with experimental weights. In M Stasinopoulos & G Touloumi (Eds.). *Proceedings of the 17th International Workshop on Statistical Modelling.* 2002. 139-147.

[7] Ruchjana, B. N. *A Generalized Space-Time Autoregressive Model and Its Application to Oil Production.* InstitutTeknologi Bandung: Ph.D. Thesis,

[8] Deng, M. and Athanasopoulos, M. Modeling Australian domestic and international inbound travel: a spatial-temporal approach. *Tourism Management.* 2011. 32(5): 1075-1084.

[9] Wutsqa, D. U. and Suhartono. Seasonal multivariate time series forecasting on tourism data by using VAR-GSTAR mode. *Jurnal ILMU DASAR.* 2010. 11(1): 101-109.

[10] Nurhayati, N., Pasaribu, U. S. and Neswan, O. Application of generalized space-time autoregressive model on GDP data in West European countries. *Journal of Probability and Statistics.* 2012. 1-16.

[11] Setiawan, Suhartono, and Prastuti, M. S-GSTAR-SUR model for seasonal spatio temporal data forecasting. *Malaysian Journal of Mathematical Sciences.+* 2016. 10(S) March: 53-65.

[12] Suhartono, Wahyuningrum, S. R., Setiawan, and Akbar, M. S. GSTARX-GLS model for spatio-temporal data forecasting. *Malaysian Journal of Mathematical Sciences.* 2016. 10(S) March: 91-103.

[13] Makridakis, S. and Hibon, M. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting.* 2000. 16: 451-476.