Performance of backpropagation artificial neural network to predict el nino southern oscillation using several indexes as onset indicators

B Aprilia, Marzuki* and I Taufiq

Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Andalas, Padang, Indonesia

marzuki@sci.unand.ac.id

Abstract. El Nino Southern Oscillation (ENSO) is a world’s climate anomaly that occurs repeatedly, unavoidable, has significant natural disaster impact for countries around the Pacific Ocean include Indonesia. ENSO has a time series of predictors, so it can be predicted the Artificial Neural Network (ANN). ANN has several important advantages over the more statistical models traditionally used. ANNs can accommodate non-linear relationship and the flexibility in testing multiple inputs. This research aims to predict the onset of ENSO using the ANN-backpropagation method of learning rate and momentum variation. The prediction is based on several indexes during 1979-2018, i.e., Sea Surface Temperature (Nino 1.2, Nino 3, Nino 3.4, and Nino 4), Southern Oscillaiton Index (SOI), Multivariate ENSO and then verified with prediction data from the International Research Institute (IRI). The results of this prediction stated that in the period JAS (July-August-September) and DJF (December-January-February) 2020/2021 world climate conditions are in normal ENSO in which there are no El Nino and La Nina phenomena. Thus, the ANN-backpropagation method is an appropriate method to predict ENSO.

1. Introduction

El Nino Southern Oscillation (ENSO) is an atmospheric phenomenon with a sea of warm periods (El Nino) and cold periods (La Nina) as well as one of the most significant inter-annual variations of the Sea Surface Temperature (SST) Pacific Ocean [1]. ENSO phenomenon modulated from decades to decades or even from century to century [2]. The El Nino and La Nina periods lasted several months with varying intensity [3].

ENSO phenomenon is one of the main factors affecting the climate variability of the world, so it can lead to drought disasters, floods, storms, and fires in various parts of the world that are mainly in the regions surrounding the Pacific Ocean [4]. One of the areas that feel the bad impact of ENSO is Indonesia. Climate anomaly characterized by ENSO’s symptoms of El Nino have occurred in the Indonesian region at the end of the dry season in 2002 and early in 2003. The results of NOAA-14 satellite imagery show a tendency to increase the number, intensity, and spread of fire spots (hot spots) on Sumatra and Kalimantan [5]. Weather changes occurring in Indonesia when El Nino is reduced in the intensity of rainfall under the normal limit of reaching 50-300 mm/month in August-October 2015 especially in the region of southern Indonesia. On the other hand, in the period of La Nina, the weather...
Natural disasters due to ENSO are inevitable, but early mitigation by predicting the phenomena can be done to prevent any adverse impact that will occur. ENSO is a periodic (recurring) event but has an oscillating mechanism that is difficult to explain, causing ENSO prediction to be an interesting problem for meteorological researches [7]. ENSO can be predicted with statistical models [8] and dynamic [9]. ENSO predictions with dynamic models and statistical models are always updated at all times by International Research Institute (IRI) Columbia University [10], so research on ENSO predictions with other methods of comparing with the IRI model [11, 12].

ENSO has complex and time-series data characteristics, so the Artificial Neural Network (ANN) method is the right method to predict ENSO [13]. ANN is computing systems consisting of some simple processing elements (neurons), which process information through an optimization function from an input variable to an output variable [14]. ANN has several advantages over other statistical models that can accommodate non-linear relationships and flexibility in testing multiple inputs [15]. One type of ANN that has succeeded in predicting the data of the time-series climatology, meteorological, and oceanography is the backpropagation [16].

ENSO predictions with the Southern Oscillation Index (SOI) and Nino 3 using the ANN-backpropagation by [17] conducted to only know how accurate the performance of the ANN method is without any time prediction of ENSO. This study resulted in the performance accuracy of ANN-backpropagation is 75 %. Long Short-Term Memory (LSTM) is an ANN variant architecture with a new non-linear machine learning algorithm capable of studying long-term temporal dependency data from complex phenomena [18]. ENSO prediction with SST Nino 3.4 index using LSTM by [19] stated that this method is so complex that it still needs to be developed again. ENSO verification of the IRI models of the Nino 3.4 index by [20] stated that the IRI models are comprehensive enough to use high-spec computer capabilities.

The ANN-backpropagation method is ANN simple but has high predictive accuracy capability if done with a lot of input data [21]. Therefore, ongoing model training with more data is indispensable. This final task study predicts ENSO using ANN-backpropagation by multiplying the input data. The data used are Sea Surface Temperature (Nino 4, Nino 3.4, Nino 3, Nino 1.2, SOI (Southern Oscillation Index), and Multivariate ENSO Index version 2 (MEI.v2) taken from the year 1979-2018. ENSO prediction results are expected to be an early mitigation step to reduce the harm of natural disasters caused by ENSO.

2. Methodology
The research uses six ENSO indexes obtained from the official National Oceanic and Atmospheric Administration (NOAA) website. The data are Nino 4, Nino 3.4, Nino 3, Nino 1.2, SOI (Southern Oscillation Index), and Multivariate ENSO Index version 2 (MEI.v2) taken from the year 1979-2018.

2.1. Data Normalization
The ANN-backpropagation method uses the sigmoid activation function which will carry inputs with an unlimited range of values to limited output value (not reaching 0 or 1). Data normalization is performed so that unlimited input values can be carried (adjusted) with a limited output value. The transformation (normalization) of data is carried out at smaller intervals i.e. (0,1 : 0,9), as stated by [22] as in the following equation:

$$x' = \frac{0.8(x - b)}{(a - b)} + 0.1 \quad (1)$$

with $x'$ is the result of normalization, $x$ is the initial data, $b$ is the minimum value of the initial data and $a$ is the maximum value of the initial data.

2.2. Data Training
The process of training ENSO prediction data with backpropagation is carried out with the following steps:

1. Input training data. The data used is 80% of the total number of years of the data.
2. Architectural creation. The initial architecture of the backpropagation neural network artificial used was 12-10-1 (input layer-hidden layer-output layer).
3. Pattern recognition (training). The training process is done by adjusting the weight value. In this research, the weight value will be determined randomly using an additional learning rate as well as momentum. Here is the initial value that will be used from each parameter:
   - Epoch : 1000
   - Learning rate : 0.1 - 0.9 with multiples of 0.1
   - Target error : 0.001
   - Momentum : 0 - 0.40 with multiples of 0.05
   - Activation function : sigmoid binary
4. Data validation is done by entering new data that has never been trained before to find out the resulting Mean Square Error (MSE) value. Predicted data with the smallest MSE value will be used in the testing process.

2.3. Testing Data
The process of testing ENSO prediction data with backpropagation is carried out with the following steps:

1. Calling the results of the training process
2. Input testing data. The data used is 20% of the total number of years of the data.
3. Data validation. Validate the data by entering new data that has never been tested before to find out the resulting Mean Square Error (MSE) value. Predicted data with the smallest MSE value will be used as a clue to predict ENSO.

2.4. Data Analysis
The predicted value that will be used as an ENSO prediction due is the prediction data obtained from the backpropagation test results with the smallest MSE value. The smallest MSE value indicates high predictive accuracy which will then be used as a due to analyzing ENSO events through the characteristic (anomalous values) of each index and then verified with prediction data from the International Research Institute (IRI).

3. Results
ENSO’s prediction was also made by Columbia University’s International Research Institute (IRI). IRI made the prediction using several models for the Nino 3.4 index from May 2020 to March 2021, as shown in Figure 1. Table 1 and Table 2 showing ENSO predictions based on models from IRI with Nino 3.4 index and ANN methods from this research for the period May 2020 to March 2021. Of the eight models used IRI, there was a considerable difference in value from each model.

Differences are also noticeable when compared to the ANN method of researching this final task, as shown in Table 3. Relatively small differences were observed in the JAS period (July-August-September) and DJF (December –January-February) 2020/2021. The average predicted value of ENSO IRI models in the JAS period was 0.28 and in the DJF period was 0.21; whereas from the ANN method, the value is 0.27 for JAS and 0.23 for DJF. The correlation value between ANN and the CPC CA/IRI method is the largest of 0.89; so the ANN method has a result equation that will be relatively similar to the CPC CA method for predicting ENSO. It also states that it cannot be said which model is most accurate until the modeled period is reached, as this is a predictive topic. The accuracy of each model can be tested at the time index for the period already exists.
Figure 1. ENSO IRI model prediction chart from May [23]

Table 1. ENSO prediction from IRI/CPC based on Nino 3.4 for May 2020 – March 2021 [23]

| Model             | MJJ  | JJA  | JAS  | ASO  | SON  | OND  | NDJ  | DJF  | JFM  |
|-------------------|------|------|------|------|------|------|------|------|------|
| NTU_CODA          | 0.61 | 0.58 | 0.54 | 0.41 | 0.34 | 0.36 | 0.36 | 0.37 | 0.59 |
| BCC_RZDM          | 0.14 | -0.04| -0.18| -0.28| -0.35| -0.43| -0.52| -0.55| -0.52|
| CPC_MRKOV         | 0.17 | 0.15 | 0.16 | 0.21 | 0.26 | 0.33 | 0.42 | 0.48 | 0.46 |
| CPC_CA            | 0.26 | 0.24 | 0.12 | 0.05 | 0.03 | 0.08 | 0.07 | 0.11 | 0.16 |
| CSU_CLIPR         | 0.62 | 0.64 | 0.66 | 0.68 | 0.81 | 0.93 | 1.06 | 0.88 | 0.70 |
| IAP-NN            | 0.28 | 0.12 | -0.04| -0.17| -0.26| -0.32| -0.36| -0.37| -0.34|
| FSU_REGR          | 0.31 | 0.22 | 0.17 | 0.15 | 0.16 | 0.18 | 0.18 | 0.13 | 0.08 |
| UCLA-TCD          | 0.72 | 0.77 | 0.77 | 0.74 | 0.71 | 0.68 | 0.66 | 0.62 | 0.57 |
| Average, Statistical Models | 0.39 | 0.33 | 0.28 | 0.22 | 0.21 | 0.23 | 0.23 | 0.21 | 0.21 |

Table 2. ENSO prediction data for ANN-backpropagation method from May 2020 – March 2021

| Index  | MJJ  | JJA  | JAS  | ASO  | SON  | OND  | NDJ  | DJF  | JFM  |
|--------|------|------|------|------|------|------|------|------|------|
| Nino 1.2 | 0.68 | 0.53 | 0.31 | 0.11 | -0.08| 0.06 | 0.01 | -0.10| -0.21|
| Nino 3  | 0.10 | -0.04| 0.01 | 0.11 | 0.13 | 0.19 | 0.13 | 0.00 | -0.15|
| Nino 3.4| 0.01 | 0.05 | 0.27 | 0.47 | 0.50 | 0.43 | 0.38 | 0.23 | 0.05 |
| Nino 4  | -0.08| -0.07| 0.02 | 0.20 | 0.32 | 0.40 | 0.33 | 0.33 | 0.19 |
| SOI     | -0.60| -0.36| -0.83| -0.48| -0.38| -0.66| -0.75| -0.69| -0.58|
| ME1.v2  | -0.26| -0.31| -0.25| -0.02| -0.04| 0.03 | 0.22 | 0.43 | 0.53 |

Table 3. ANN-backpropagation Nino 3.4 regression values with IRI model methods

| Regression | NTU_CODA | BCC_RZDM | CPC_MRKOV | CPC_CA | CSU_CLIPR | IAP-NN | FSU_REGR | UCLA-TCD |
|------------|----------|----------|-----------|--------|-----------|--------|----------|----------|
| ANN        | 0.77     | 0.20     | 0.01      | 0.89   | 0.29      | 0.31   | 0.09     | 0.02     |
As mentioned above that the difference between IRI and ANN predictions from this final task is relatively small for JAS and DJF 2020/2021. Predictions of the ENSO phenomenon from ANN state that these two periods are normal periods that are not the case of El Nino or La Nina. This is the same as that obtained by the IRI model wherein JAS and DJF 2020/2021 the state of the world climate is in normal condition with a predicted probability of 58% and 44% (Table 4). Thus, the ANN-backpropagation method developed in this research has the potential to be used in predicting ENSO although there are some differences with the IRI model as mentioned above.

| Table 4. ENSO prediction probability [23] |
|-----------------------------------------|
| Season  | La Nina | Neutral | El Nino |
| MJJ 2020 | 2%      | 89%     | 9%      |
| JJA 2020 | 15%     | 71%     | 14%     |
| JAS 2020 | 25%     | 58%     | 17%     |
| ASO 2020 | 30%     | 50%     | 20%     |
| SON 2020 | 31%     | 44%     | 25%     |
| OND 2020 | 31%     | 40%     | 29%     |
| NDJ 2020 | 30%     | 41%     | 29%     |
| DJF 2020 | 25%     | 44%     | 31%     |
| JFM 2020 | 17%     | 49%     | 34%     |

4. Conclusion
Predictive results with the ANN-backpropagation method can be said to be precise if verified by the International Research Institute (IRI) model. The analysis of ENSO predictions from the ANN model of this research for the Nino 3.4 index in the JAS period (July-August-September) and DJF (December-January-February) 2020/2021 is under normal circumstance i.e. the absence of El Nino or La Nina phenomena. This analysis was also stated similarly by the IRI model which is in the JAS and DJF period 2020/2021 the state of the world's climate is in normal condition with a predicted probability of 58% and 44%. This states that the ANN-backpropagation is the correct method and provides high accuracy for predicting ENSO time series data.

Acknowledgments
This study was supported by the 2020 Basic Research grant from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia.

References
[1] Feng Y and Tung K 2019 ENSO Modulation: Real and Apparent; Implications for Decadal Prediction Springer-Verlag GmbH Germany
[2] Li J, Xie S, Cook E R, Morales MS, Christie D A, Johnson N C, Chen F, D’Arrigo R, Fowler A M, Gou X and Fang K 2013 El Nino Modulations Over The Past Seven Centuries Nature Climate Change 3 822 – 826
[3] Philander S G 2013 El Nino, La Nina, and the Southern Oscillation Academic Press Inc, San Diego
[4] Ludescher J, Gozolchiianib A, Bogacheva M I, Bundea A, Havlinb S and Schellnhuber H J 2014 Very early warning of next El Nino Proceedings of the National Academy of Sciences 111 2064 - 2066
[5] Sudibyakto 2003 Anomali Iklim Dan Mitigasi Kebakaran Hutan Di Indonesia Majalah Geografi Indonesia 17 1 71 – 80.
[6] Athoillah I, Sibarani R M and Doloksaribu D E 2017 Analisi Spasial El Nino Kuat Tahun 2015 Dan La Nina Lemah Tahun 2016 (Pengaruhnya Terhadap Kelembapan, Angin Dan Curah Hujan Di Indonesia Jurnal Sains & Teknologi Modifikasi Cuaca 18 1 33 – 41
[7] Santos A, Mchpaden M J and Cai W 2017 The Defining Characteristics of ENSO Extremes and the Strong 2015/2016 El Nino Reviews of Geophysics 55 4 1079 -1129
[8] Barnston A G, Tippett M K, Ranganathan M and LHeureux M L 2019 Deterministic skill of
ENSO predictions from the North American Multimodel Ensemble *Clim Dyn* **53** 12 7215 – 7234

[9] Petersik P J and Dijkstra H A 2020 Probabilistic forecasting of El Nino using neural network models *American Geophysical Union*

[10] Barnston A G and Tippett M K 2014 Climate Information, Outlooks, and Understanding—where Does the IRI Stand? *Earth Perspectives* **1** - 20

[11] Andika Halide H and Hasanah N 2017 Verifikasi Prediksi Enso Musiman Model Dinamik Operasional IRI (International Research Institute) *Jurnal Geofisika Universitas Hasanuddin*

[12] Barnston A G, LI S, Mason S J, Dewitt D G, Goddard L and Gong X 2010 Verification of the First 11 Years of IRI’s Seasonal Climate Forecasts *Journal of Applied Meteorology and Climatology* **59** 493 - 520

[13] Mu B, Peng C, Yuan S and Chein L 2019 ENSO Forecasting over Multiple Time Horizons Using ConvLSTM Network and Rolling Mechanism *International Joint Conference on Neural Networks Budapest*

[14] Nootooboop P D, Feng Q Y, López C, ‘Hernández-Garcia E and Dijkstra H A Using Network Theory and Machine Learning to Predict El Nino 2018 *Earth System Dynamics* **9** 969 - 983

[15] Abbot J and Marohasy J 2017 Application Of Artificial Neural Networks To Forecasting Monthly Rainfall One Year In Advance For Locations Within The Murray Darling Basin, *Australia Int. J. Sus. Dev. Plann* **12** 8 1282 – 1298

[16] Hasieh W W and Tang B 1998 Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography *Bulletin of the American Meteorological Society*

[17] Baawain M S, Nour M H, El-Din A G and El-Din M G 2005 El Nino Southern-Oscillation Prediction Using Selatanern Oscillation Index and Niño3 as Onset Indicators: Application of Artificial Neural Networks *NRC Research Press* 13 – 21

[18] Gers F A, Schmidhuber J and Cummins F 2000 Learning to Forget: Continual Prediction with LSTM *Neural Comput* **12** 10 2451 – 2471

[19] Broni-Bedaiko C, Katsriku F A, Unemi T, Shinomiya N, Jamal-Deen Abdulai J and Atsumi M 2018 El Nino-Southern Oscillation Forecasting Using Complex Networks Analysis of LSTM Neural Networks *The Twenty-Third International Symposium on Artificial Life and Robotics Bappu*

[20] Barnston A G, Tippet M K, L’Heureux M L, Li S and Dewitt, D G 2012 Skill Of Real-Time Seasonal Enso Model Predictions During 2002–11 *American Meteorological Society*

[21] Dijkstra H A, Petersik P, Hernández-Garcia E and Lopez C 2019 The Application of Machine Learning Techniques to Improve El Niño Prediction Skill *Frontiers In Physics* **7** 1 - 13

[22] Hutabarat M A P, Julham M and Wanto A 2018 Penerapan Algoritma Backpropagation dalam Memprediksi Produksi Tanaman Padi Sawah Menurut Kabupaten/Kota di Sumatera Utara *semanTIK* **4** 1

[23] International Research Institute (IRI) 2020 IRI/CPC ENSO Predictions Plume *https://iri.columbia.edu/ourexpertise/climate/forecasts/enso/current/?enso_tab=enso-sst_table*, accessed Mei 2020