Development of 24 hours Dst index prediction from solar wind data and IMF Bz using NARX

F Nuraeni¹, M Ruhimat¹, M A Aris¹, E A Ratnasari¹, C Purnomo²

¹. Space Science Center - LAPAN
². Sumedang Atmospheric and Space Observatory

Email: fitrinur.fn@gmail.com

Abstract. Disturbance Storm Time (Dst) index is an index which measured the decrease in the horizontal component of the Earth’s magnetic field near the magnetic equator due to increases in the magnetospheric ring current. The index shows the strength and duration of a geomagnetic storm. The geomagnetic storms themselves could harm technological systems on earth. Dst index prediction is needed as an effort to mitigate the impact of geomagnetic storms. This paper described a 24 hours Dst index prediction using a nonlinear autoregressive exogenous (NARX) method. Solar wind parameters and interplanetary magnetic field (IMF) Bz has been used as exogenous input for the model. By using 39 geomagnetic storms data during 1997 to 2000 as training data and 7 geomagnetic storms data during 2000 to 2001 as testing data that have not been used in the training, the correlation between target and output was 0.97. The model than being used to predict Dst index for operational purposes and obtained the RMSE from 1 August to 31 October 2021 was ~5 – 20 nT during quiet days and ~18 – 45 nT during disturb days.

1. Introduction

The geomagnetic field is one of Sun-Earth environment domain that will be affected by solar and solar wind activities. Disturbance in the geomagnetic field can be hazardous to man made technologies especially space-based technologies. Primary cause of geomagnetic disturbance is associated with interplanetary structure which has intense, long duration and Southward magnetic field (Bs) [1]. Different interplanetary conditions caused different geomagnetic response [2-4]. Richardson, et.al., 2012 [5] has classified 3 types of near-Earth basic flows that are drivers to geomagnetic storms, first is high speed stream associated with coronal hole, the second is slow, inter-stream solar wind and the third is transient flows from coronal mass ejections (CME) including interplanetary CME and the associated upstream shocks and post-shock region. Interplanetary condition especially duration and intensity of Bz have significant role to control geomagnetic storm intensity [6,7].

Geomagnetic storms are characterized by depression in the horizontal component (H) of magnetic field which is caused by ring current flowing westward in the magnetosphere [8]. Disturbance Storm Time (Dst) index is often used to measure the strength of ring current and depression less than – 50 nT used to indicate a geomagnetic storm is in progress [9]. Dst is an index obtain by averaging deviation of H component of 4 ground stations around mid-latitude with its quiet days value. The index describes dynamics of ring current during quiet or disturbed days. Burton, et.al., 1975 [10] has
proposed equation to predict Dst value which is a function of solar wind speed, density and interplanetary magnetic field North-South component heading southward (negative IMF Bz).

Therefore, modeling and predicting Dst index are needed to quantify the strength of geomagnetic disturbance that will hit the Earth. As an attempt to predict geomagnetic activities many researchers have developed methods to predict Dst index as well as other global geomagnetic indexes. Modeling and prediction of Kp index has been done using artificial neural network methods and gained a good result [11,12]. Dst index and Kp prediction using artificial neural network also been done in various prediction horizon using parameter of solar wind and interplanetary magnetic field [13]. An operational forecast of geomagnetic index has also been developed by using recurrent neural network to predict one hour Dst index using only solar wind parameter as input [14].

The ability of artificial neural network methods to model and predict time series of dynamic system, makes it become a powerful tools to model dynamic of geomagnetic activities based on solar wind and interplanetary magnetic field as input. In this paper, we develop a model using NARX method to predict 24 hours Dst index, so that it can be used in operational prediction system as part of space weather information system.

2. Data
The data used to develop the NARX model as input is hourly averaged solar wind speed and density, interplanetary magnetic field North-South component (IMF-Bz) during 1997 to 2000 is obtained from https://omniweb.gsfc.nasa.gov/ and Dst index of the same year is acquired from http://wdc.kugi.kyoto-u.ac.jp/. Input data for the model that is used in the operational system are obtained from https://services.swpc.noaa.gov/products/solar-wind/mag-3-day.json for IMF Bz, https://services.swpc.noaa.gov/products/solar-wind/plasma-3-day.json for solar wind speed and density, and https://services.swpc.noaa.gov/products/kyoto-dst.json for Dst index.

From 1997 to 2000 there are 39 geomagnetic disturb days that is used as model’s input in training, validating and testing phase. The geomagnetic disturb days here are defined as those with minimum Dst index less than -50 nT. For each event disturbed day, we took 36 hours before the peak of Dst minimum and 24 hours after the peak of Dst minimum. Figure 1 shows 2 events as sample of data set used in the model development. The exogenous input is also set with the same arrangement.

![Figure 1. Sample of 2 events data set used in model development. (Source: http://wdc.kugi.kyoto-u.ac.jp/)](image)

3. Model development
In this paper we use non-liner autoregressive model with exogenous input (NARX) to predict 24 hours Dst index. NARX is a dynamical neural network architecture commonly used for input-output modelling of nonlinear dynamical systems. Applied to time series prediction, the NARX network is designed as a feed-forward time delay neural network (TDNN). In general NARX model can be described as function of output which is determined by array of input u(t) and output y(t) from previous data. Mathematically described as:

\[
y(t) = f(y(t - 1), y(t - 2), ..., y(t - ny), u(t - 1), u(t - 2), ..., u(t - nu))
\]  

(1)
Data set to develop the model was divided into 3 groups, 70% for training, 15% for validating and 15% for testing. It means that 27 geomagnetic storm events used for the training, 6 events for validating and rest 6 events for testing. Each event consists of 60 hours data used as input. Figure 2 shows the model’s structure.

The exogenous input for the model are the hourly average of solar wind speed and density and also the North-South component of interplanetary magnetic field (IMF) Bz, while the hourly Dst index as the endogenous input of the model. Dst index, which is the output of the model is available during training, then we used it as feedback input instead of estimated output. Therefore the model train with true output instead of estimated output.

We used several configurations of hidden layer, delay input and output before we choose model with 10 hidden layers with 24 delay input and 6 delay output which has the best-fit regression coefficient in training, validating and testing to be used later for the prediction, the model's configuration architecture shows in figure 3. Root mean square of the model was 0.05 nT.
4. Result

Common NARX network uses previous output as feedback input, but as prediction model the previous output are not available. In that case we switch to closed loop network so the network could simulate the model using internal feedback. In order to test the model prediction with closed loop network, we use another set data consist of 7 geomagnetic storm events and 4 quiet days during 2001 to 2002, shows in table 1.

| Date             | Minimum Dst Index (nT) |
|------------------|------------------------|
| Geomagnetic storm |                        |
| 5 March 2001     | -73                    |
| 11 April 2001    | -271                   |
| 18 April 2001    | -114                   |
| 2 February 2002  | -86                    |
| 24 March 2002    | -100                   |
| 11 May 2002      | -110                   |
| 23 May 2002      | -109                   |
| Quiet days       |                        |
| 6 January 2001   | -12                    |
| 10 March 2001    | -4                     |
| 2 May 2002       | -11                    |
| 6 August 2002    | -13                    |

Using data set in table 1 we found that the model using closed loop network is working properly to predict 24 hours Dst index. Regression plot between target and output for disturb days which define as geomagnetic storm in table 1 and quiet days shows in figure 4. The RMSE for disturb days of data set.
in table 1 is from 9.05 nT to 53.63 nT while the RMSE for the quiet days of data set in table 1 is 4.75 nT to 5.46 nT.

Figure 4. Plot regression of table 1 data set on disturb days (left) and quiet days (right)

Lowest RMSE for disturbed days from data in table 1 occurs on 5 March 2001 event while the highest RMSE occurs on 23 May 2002, the result shows in figure 5. We can also see that error tend to become higher after Dst index minimum or in term of geomagnetic storm is during the recovery phase.

Figure 5. Plot of predicted Dst index of data in table 1 with lowest RMSE (left) and highest RMSE (right).

After obtaining the model and architecture network that can be used for prediction, the model is implemented into an operational prediction system. From August 2021 to October 2021 the system has been operated to test whether the system runs smoothly and obtains good prediction results. During August to October 2021 there are 6 disturbed days which results are summarized in table 2, while the RMSE for days other than the disturbed date listed in table 2 are in the range of 5 to 20 nT. Disturbed days in table 2, classified according to Loewe and Prölls,1997 [15] which are -30 nT ≥ Dst_{min} > -50 nT consider as weak, -50 nT ≥ Dst_{min} > -100 nT as moderate, -100 nT ≥ Dst_{min} > -200 nT as strong geomagnetic storms. Table 2 shows that the operational 24 hours prediction Dst index system gives a lower value of predicted Dst index at 7 out of 8 geomagnetic storm peaks time, the highest and lowest RMSE shows in figure 6. The minimum value of predicted Dst index does not always occur at precise time as the peak of geomagnetic storm. The operational system calculates the prediction everyday on 00 UT, so that the input used in the calculation may not show any pattern of disturbance as already been recognized by the model in learning process. This may be the cause of the overestimate prediction value and inaccuracy of the peak time.
Table 2. Result of operational Dst index prediction system

| Geomagnetic Disturbed days (date) | Dst\textsubscript{min} peak time (UT) | Dst\textsubscript{min} value (nT) | Predicted Dst\textsubscript{min} on peak time (nT) | ΔDst\textsubscript{min} = Obs - Predicted (nT) | RMSE (nT) |
|----------------------------------|---------------------------------------|----------------------------------|-----------------------------------------------|-----------------------------------------------|----------|
| 6-7 August 2021                  | 7                                    | -38                              | -69                                           | 31                                            | 25.41    |
| 27-28 August 2021                | 24                                   | -82                              | -105                                          | 23                                            | 18.45    |
| 17-18 September 2021             | 22                                   | -64                              | -144                                          | 80                                            | 44.99    |
| 12-13 October 2021               | 7 (1\textsuperscript{st} peak) 15 (2\textsuperscript{nd} peak) | -57                              | -70                                           | 13                                            | 18.96    |
| 17-18 October 2021               | 24                                   | -55                              | -136                                          | 81                                            | 37.17    |
| 31 October 2021                  | 6 (1\textsuperscript{st} peak) 15 (2\textsuperscript{nd} peak) | -33                              | -24                                           | -7                                            | 19.48    |

Figure 6. Dst index prediction of disturbed day with highest RMSE occurred on 17-18 September 2021 event (left) while the lowest RMSE occurred on 27-28 August 2021 event (right). Dst index predicted plotted as black line while the Dst index plotted in blue line.

From the result, solar wind parameter together with interplanetary magnetic field North-South component (Bz) can be used as exogenous input of NARX network to predict 24 hours Dst index. During training, we had an almost fitted model with very low RMSE while at testing and operational system tends to have larger RMSE, might be caused by data selection. For training, we purposely select geomagnetic disturbed days that had a clear track such as single event geomagnetic storms without double peak. While in testing and operational systems, we predict Dst index using all kind of known or real time input without targeting type of geomagnetic disturbance. Still, model accuracy needs to be improved especially for the operational purpose.

5. Summary
To summarize, our results are:
- Solar wind parameters and interplanetary magnetic field (IMF) Bz has been used as exogenous input for the model.
39 geomagnetic storms data during 1997 to 2000 as training data and 9 geomagnetic storms data during 2000 to 2001 testing data. 60 hours input used to predict 24 hours Dst index ahead.

Model correlation between target and output was 0.97.

The model than being used to predict Dst index for operational purposes and obtained the RMSE ~ 18 to 45 nT during disturb days and ~ 5 to 20 nT during quiet days.

Acknowledgments
Authors acknowledge the OMNIweb, WDC Kyoto, and SWPC team that have provide solar wind and interplanetary magnetic data and also geomagnetic indices. Authors also would like to thank reviewer and editorial team that has given constructive comments and suggestions.

FN contributed in research design, study literature, methodology development, neural network training, and results analysis. MR contributed in research design and neural network training. MAA, EAR and CP contributed in data preparation, and assist in the training process of the neural network.

References
[1] Gonzalez, W.D., Joselyn, J.A., Kamide, Y., Kroehl, K.W., Rostoker, G., Tsurutani, B.T., and Vasyliunas, V.M.: What is a geomagnetic storm?, *Journal of Geophysical Research*, Vol. 99, No. A4, pages 5771 -5792, April 1, 1994
[2] Clauer, C.R.: The geomagnetic storm-time response to different solar wind driving conditions, *ILWS WORKSHOP 2006, GOA, FEBRUARY 19-24, 2006*
[3] Borovsky, J.E., and Denton, M.H.: Differences between CME-driven storms and CIR-driven storms, *Journal of Geophysical Research*, Vol. 111, A07S08, doi:10.1029/2005JA011447, 2006
[4] Guarnieri, F.L., Tsurutani, B.T., Gonzalez, W.D., Gonzalez, A.L.C., Grande, M., Soraas, F., and Echer, E.: ICME and CIR storms with particular emphases on HILDCAA events, *ILWS WORKSHOP 2006, GOA, FEBRUARY 19-20, 2006*
[5] Richardson, I.G., and Cane, H.V.: Solar wind drivers of geomagnetic storms during more than four solar cycles, *J.Space Weather Space Clim. 2* (2012) A01, doi:10.1051/SWS/2012001
[6] Lakhina, G.S., Alex, S., Tsurutani, B.T., & Gonzalez, W.D.: Supermagnetic storms: Hazard to society. Extreme Events and Natural Hazards: The Complexity Perspective Geophysical Monograph Series 196 © 2012. American Geophysical Union. All Rights Reserved. 10.1029/2011GM001073
[7] Alex, S., Mukherjee, S., & Lakhina, G.S.: Geomagnetic signatures during the intense geomagnetic storms of 29 October and 20 November 2003. *Journal of Atmospheric and Solar-Terrestrial Physics* 68 (2006) 769–780. doi:10.1016/j.jastp.2006.01.003.
[8] Kamide, Y., Baumjohann, W., Daglis, I.A., Gonzalez, W.D., Grande, M., Joselyn, J.A., McPherron, R.L., Phillips, J.L., Reeves, E.G.D., Rostoker, G., Sharma, A.S., Singer, H.J., Tsurutani, B.T., and Vasyliunas, V.M., Current understanding of magnetic storms: Storm/substorm relationships, *Journal of Geophysical Research: Space Physics* August 1998. DOI: 10.1029/98JA01426
[9] Temerin, M., and Li, Xinlin, A New Model for the Prediction of Dst on the Basis of the Solar Wind, *Journal of Geophysical Research Atmospheres*: December 2002. DOI: 10.1029/2001JA007532
[10] Burton, R.K., McPherron, R.L., and Russel, C.T.: An empirical relationship between interplanetary conditions and Dst, *Journal of Geophysical Research*, Vol. 80, NO. 31, 1975
[11] Ayala Solares, J. R., H.-L. Wei, R. J. Boynton, S. N. Walker, and S. A. Billings (2016), Modeling and prediction of global magnetic disturbance in near-Earth space: A case study for Kp index using NARX models, *Space Weather*, 14, 899–916, doi:10.1002/2016SW001463
[12] Uwamahoro, J., and Habarulema, J.B., Empirical modeling of the storm time geomagnetic
indices: a comparison between the local K and global Kp indices, *Earth, Planets and Space* 2014, 66:95.

[13] Myagkova, I., Shiroky, V., and Dolenko, S., Prediction of geomagnetic indexes with the help of artificial neural networks, Solar-Terrestrial Relations and Physics of Earthquake Precursors, E3S Web of Conferences 20, 02011 (2017). DOI: 10.1051/e3sconf/20172002011

[14] Lundstedt, H., Gleisner, H., Wintoft, P., Operational forecasts of the geomagnetic Dst index, *Geophysical Research Letters*, VOL. 29, NO. 24, 2181, doi:10.1029/2002GL016151, 2002

[15] Loewe, C.A., and Prölss, G.W., Classification and mean behaviour of magnetic storms, *Journal Of Geophysical Research*, Vol. 102, No. A7, Pages 14,209-14,213, July 1, 1997