

Prediction of intensity of moderate and intense geomagnetic storms using artificial neural network during two complete solar cycles 23 and 24

P K Singh* 

Department of Mechanical Engineering, School of Engineering, Shiv Nadar University, Tehsil Dadri, Gautam Buddha Nagar, Uttar Pradesh 201314, India

Received: 21 February 2021 / Accepted: 12 August 2021 / Published online: 3 September 2021

Abstract: This work aims to predict moderate, intense, and super geomagnetic storms during the two recent solar cycles 23 and 24 encompassing the period 1996–2018 using an artificial neural network (ANN). Optimization of the neural network includes a choice of activation function, training function, learning function, hidden layers, hidden neurons, learning rate, and momentum constant. The results obtained by the present study show the ability of the ANN model to produce an accurate estimate of the probability appearance of moderate and intense storms of about 88.9%.

Keywords: Geomagnetic storms; Dst index; Neural network; Solar cycle

1. Introduction

The disturbance caused by various solar events in the geomagnetic field can be of significant interest in the prediction of space weather. Data science and machine learning are relatively new tools used in science and mathematics that can help to improve the conventional methods used previously in space weather prediction, especially for nonlinear data as input. Burton et al. [1] first proposed an algorithm for Dst index prediction and interplanetary magnetic field (IMF) data. The algorithm also helped to detect the causes of various types of storm behavior. Kugblenu et al. [2] used an artificial neural network (ANN) with one hidden layer, a backpropagation algorithm, and three-hour Dst value before minimum Dst occurrence, and also some other inputs to predict the Dst index. Lazzús et al. [3] applied ANN to forecast the Dst index using an algorithm known as “particle swarm optimization.” Uwamahoro and Habarulema [4] used a neural network-based model to predict geomagnetic index using SW and IMF parameters as input with a correlation coefficient of 0.8. Uwamahoro et al. [5] formulated an NN-based prediction model for post-halo CME occurrence of geomagnetic conditions during 2000–2005. His model is

75% accurate for moderate storms and 86% for storm occurrence rate of halo coronal mass ejections (CMEs). Srivastava [6] implemented a logistic regression model for classifying the intense and super-intense geomagnetic storms from 1996 to 2002 utilizing solar and interplanetary properties of geo-effective CMEs. The model had 78% accuracy for the validation dataset and 85% accuracy for the training dataset. Kamide et al. [7] suggested that the NN model can prove effective for both linear and nonlinear processes. Many workers have prescribed that NN approaches are appropriate for forecasting transient solar-terrestrial phenomena. CMEs characteristics that influence magnetic storms are still a topic of consideration [8]. Therefore, selections of key solar and IP parameters are significant for the prediction of geomagnetic storms [6]. Lazzús et al. [9] employed ANN with particle swarm optimization to forecast the Dst index from 1 to 6 h onward with a correlation coefficient ranging from 0.79 to 0.86. Also, Lethy et al. [10] applied ANN with seven input parameters for 1-h to 12-h advance prediction and correlation coefficient diverging from 0.857 to 0.876.

This paper focuses on the analysis of moderate ($-100 \text{ nT} \leq \text{Dst} \leq -50 \text{ nT}$), intense ($-250 \text{ nT} \leq \text{Dst} \leq -100 \text{ nT}$), and super ($\text{Dst} \leq -250 \text{ nT}$) geomagnetic storms that have occurred during two solar cycles 23 and 24. The entire phases of two solar cycles 23 and 24 have been considered for the analysis and prediction of neural networks. It is quite significant to collect data that

*Corresponding author, E-mail: pratyush.singh997@gmail.com

contain all types and ranges of magnitude for geomagnetic storms. At the beginning of the cycle, the frequency and intensity of a geomagnetic storm are minimum. They rise to the maximum during the middle phase of the cycle, and then, they sink to a minimum again during the last phase of the cycle [11]. Hathway [11] reviewed the periodic variation of sunspot area and various other parameters that show solar activity such as a magnetic field, coronal mass ejections, cosmic ray fluxes, and geomagnetic activity. Each cycle shows a maximum and a minimum with a typical cyclic variation for these parameters. Considering the whole solar cycle gives heterogeneity in data and makes it more reliable for the prediction of the Dst index for geomagnetic storms.

Divergence in the input values of neural networks can benefit to know the functional correlation between input and output and consequently infer how well the network has determined [12]. NN models for the prediction of geomagnetic storms employing SW data as input have been incurred by Lundstedt and Wintoft [13] who have represented the extent of geomagnetic disturbance by the Dst index. Valach et al. [14] developed a neural network-based model based on inputs such as X-ray flare, solar radio bursts, solar energetic particle flux, and high energy proton flux to predict subsequent geomagnetic storms. Dryer et al. [15] suggested that models should account for both solar and near-Earth conditions. Singh and Singh [16] discussed the development of ANN-based model to study the precursor to moderate and intense geomagnetic storms during the ascending phase of solar cycle 24 between 2009 and 2014 and reported a probability of 79% accuracy of prediction. A Gaussian process regression has been adopted for simulating the Dst index's forecasts based on probability by Chandorkar et al. [17]. Gruet et al. [18] combined the Gaussian process method with the long short-term memory (LSTM) neural network to provide a 6-h ahead forecast of the Dst index with a correlation coefficient always higher than 0.873. Recently, Xu et al. [19] implemented the bagging ensemble-learning algorithm which combines a total of three algorithms—support vector regression, artificial neural network, and long short-term memory network to forecast the Dst index 1 to 6 h before achieving an accuracy of interval prediction always higher than 90%. Out of several reported studies, till now the prediction of geomagnetic storms using a long-term analysis of the complete solar cycle using ANN is lacking. To complete the study of prediction of moderate as well as intense geomagnetic storms during two complete recent solar cycles 23 and 24, the present study has been initiated.

The present study focuses on the geomagnetic storms that occurred throughout the 23rd solar cycle from the years 1996 to 2007 and the 24th solar cycle from 2008 to 2018. Various solar and interplanetary input parameters

were used to predict the moderate, intense, and super geomagnetic storms of two recent solar cycles 23 and 24. The prediction is done using artificial neural network analysis. The number of hidden layers, hidden neurons, learning algorithm, and training algorithm have been varied and experimented with to find the optimal architecture for ANN to give the least error and highest percentage accuracy.

2. Input parameters

2.1. Solar input parameters

The solar input parameters examined in the investigation are the CME velocity, angular width (AW), and the flare class. Bruekner et al. [20] said that geomagnetic activity has been attributed to either CMEs or the interplanetary shock waves that accompany them. Hence, the CME parameters play a significant role in the intensity of geomagnetic storms. The magnitude of the acceleration of CME depends on the impulsiveness of the associated solar flares [21]. An estimate for the volume of the corona that is “blown out” can be given by the angular width [22]. Apart from AW, the CME speed is another significant property related to geo-effective CMEs. It is observed that halo CMEs obtain mean SW speed > 470 km/s and prove to be a significant parameter in the prediction of geomagnetic storms [6]. Flare class is another significant input for our neural network in the study. The logarithmic value of the X-ray flare that occurred during the CME has been considered for input. Wang et al. [8] found that geo-effective halo CMEs are generally connected with flare. Moreover, Srivastava and Venkatakrishnan [23] perceived that fast and full halo CMEs associated with massive flares lead to great geomagnetic disruptions further. The halo CME data have been obtained from the Web page: https://cdaw.gsfc.nasa.gov/CME_list/halo/halo.html.

During solar cycle 23, a total of 232 halo CMEs and during solar cycle 24, a total of 132 halo CMEs (both partial and full) have been found geo-effective. Table 1 and Table 2 represent the distribution of all halo CMEs associated with moderate, intense, and super geomagnetic storms year-wise during the whole solar cycle 24 and for solar cycle 23, respectively. There were no superstorms during solar cycle 24. Out of a total of 364 geomagnetic storms, there were 130 intense, 208 moderate, and 26 super geomagnetic storms during solar cycle 23.

2.2. Interplanetary input parameters

In the interplanetary medium, CMEs can be presented as shocks and ICME structures that connect to the

Table 1 List of all 132 CME-driven geomagnetic storm events which occurred each year during solar cycle 24

| Year | Intense storms | Moderate storms |
|-------|----------------|-----------------|
| 2008 | 0 | 0 |
| 2009 | 0 | 0 |
| 2010 | 0 | 8 |
| 2011 | 6 | 18 |
| 2012 | 11 | 20 |
| 2013 | 2 | 27 |
| 2014 | 2 | 16 |
| 2015 | 4 | 11 |
| 2016 | 2 | 1 |
| 2017 | 2 | 2 |
| 2018 | 0 | 0 |
| Total | 29 | 103 |

Table 2 List of all 232 CME-driven geomagnetic storm events which occurred each year during solar cycle 23

| Year | Intense storms | Moderate storms | Super storms |
|-------|----------------|-----------------|--------------|
| 1996 | 0 | 0 | 0 |
| 1997 | 10 | 5 | 0 |
| 1998 | 8 | 8 | 0 |
| 1999 | 3 | 1 | 0 |
| 2000 | 13 | 27 | 2 |
| 2001 | 15 | 18 | 11 |
| 2002 | 11 | 13 | 0 |
| 2003 | 19 | 7 | 4 |
| 2004 | 3 | 3 | 9 |
| 2005 | 14 | 21 | 0 |
| 2006 | 5 | 2 | 0 |
| 2007 | 0 | 0 | 0 |
| 2008 | 0 | 0 | 0 |
| Total | 101 | 105 | 26 |

magnetosphere and cause major to moderate storms. Occurrences of shock waves are recognized by a simultaneous increase in the SW speed, density, abnormal proton temperature as well as an increase in magnetic field magnitude. Gonzalez and Tsurutani [24] suggested that the intensity of the storm after the occurrence of shock-ICME structures is well linked with two parameters which are the negative IMF Bz component (Bz) and the electric field, $E_y = V B_z$, where V specifies the SW velocity. Recent findings have also established that the convective electric field has the best correlation with the Dst index [25]. Different SW data have been taken from the Web page: https://omniweb.gsfc.nasa.gov/form/omni_min.html. All

the solar and interplanetary input parameters used in the present study are listed in Table 3.

3. Principal component analysis

Principle component analysis (PCA) is a procedure that is helpful for the compression and grouping of data.

The explanation behind existing is to decrease the dimensionality of a dataset by finding another arrangement of samples, fewer than the main arrangement of samples that holds most of the sample data [26]. Dimensionality reduction can help to reduce the number of features which saves computational power and time. It removes the features which may not be essential and do not show much variation with different samples of data. Dimensionality reduction is also beneficial in reducing overfitting of data.

Steps involved in the calculation of PCA are given as follows:

1. The dataset includes all the parameters for each of the 362 CMEs which is input.
2. We will compute the arithmetic mean for each dimension of the dataset as given by Eq. (1):

$$\text{Mean} = \frac{\sum_{i=1}^n a_i}{n}. \quad (1)$$

3. We will compute the covariance matrix for the whole dataset through Eq. (2)

$$\text{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}. \quad (2)$$

4. We will calculate the eigenvectors and eigenvalues for the covariance matrix as given in Eq. (3):

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0. \quad (3)$$

5. We will arrange the eigenvalues in decreasing order and choose m eigenvectors with largest eigenvalues to form a $d \times m$ dimensional matrix W .
6. The last step is to take the transpose of the matrix W and multiply it with the original dataset.

Table 3 Input parameters considered for the ANN model

| Parameter | Unit |
|---|--------------------|
| Angular width (AW) | $> 120^\circ$ |
| Solar wind velocity (V_{sw}) | km s^{-1} |
| CME velocity (V_{cme}) | km s^{-1} |
| IMF negative magnetic component (B_z) | nT |
| Electric field component (E_y) | mV m^{-1} |
| Flare class (F_{class}) | – |

Fig. 1 Principal component analysis (PCA) plot of the dataset. The red dots represent each of the observations, i.e., the geomagnetic storms, whereas the blue lines are the six variables which are considered for input in the study

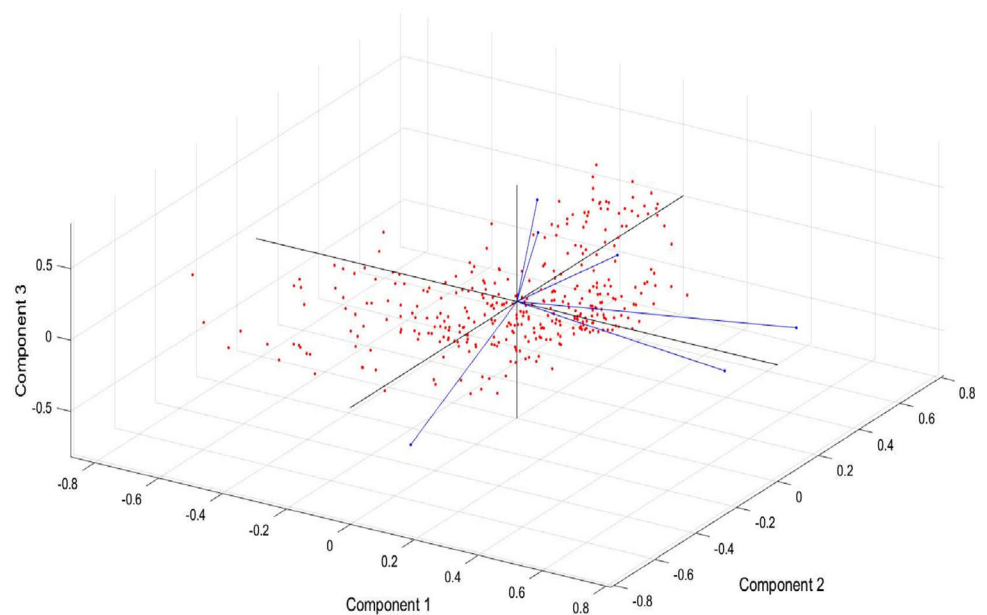


Figure 1 shows the plot for the principal component analysis; the red dots represent each of the observations, i.e., the geomagnetic storms, whereas the blue lines are the six variables that are considered for input in the study. The length of the blue vector corresponds to its contribution to the principal components, whereas the direction of the vector shows how it is correlated with each of the principal components. The major elements are shown by the axes in Fig. 1.

4. Neural networks

4.1. Feed-forward network

In this work, we practiced feeding forward networks, where the relationships in the neurons are steered along one direction from the input layer to the output layer via the hidden layer. The activation function of the output layer is linear, whereas the activation function of the hidden layer is tangent sigmoid which assists uphold and supports the network [27]. The tangent sigmoid function is used preferably because it varies from -1 to 1 ; hence, it is efficient for these data as many features attain both negative and positive values. The negative values can be mapped as strongly negative and same for positive values.

4.2. Training function

The training function used in this work is gradient descent with momentum. Training functions were implemented with the batch training technique. In batch training, modification of network weights is not done until all the

training samples are propagated to the network. The training function aims to reduce the error between the calculated and observed Dst values to their minimum using mean square error. The steps for Gradient descent with momentum are:

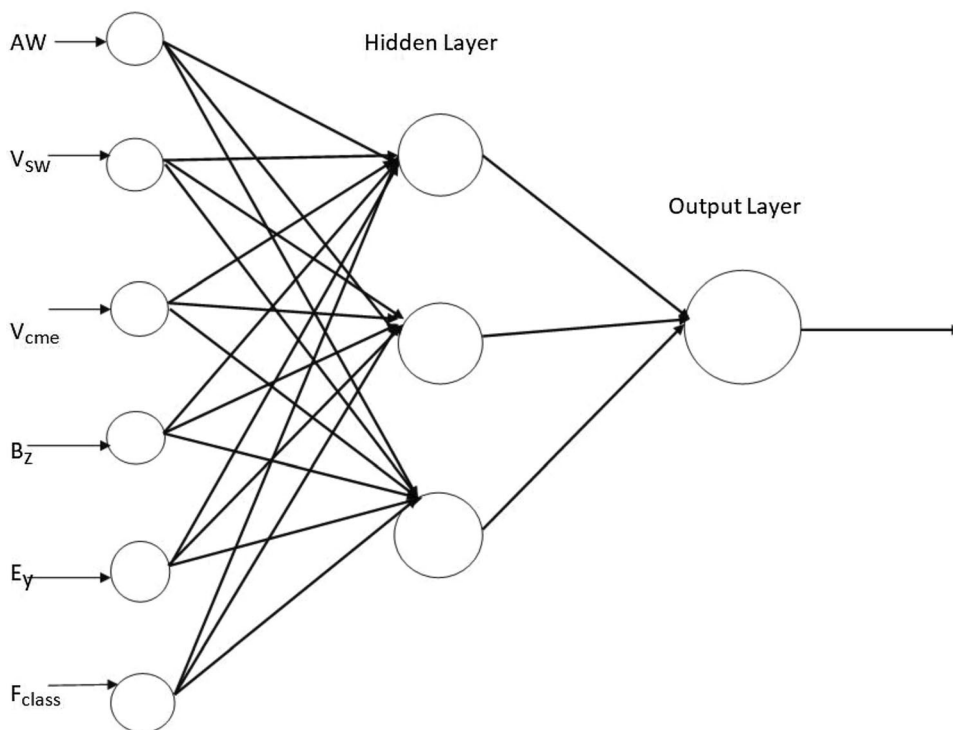
1. Random initialization of weights.
2. Reiteratively do the accompanying:
 - (a) Initialize the weight correction to zero.
 - (b) For each of the inputs:
 - (i) Calculate target data—output data.
 - (ii) Calculate the gradient of error corresponding to each weight.
 - (iii) Edit the weights as per the weight corrections.
3. Print net output and error.

End.

The gradient descent with momentum helps to perform better than gradient descent. Bishop [28] said that gradient descent with momentum can lead to faster convergence toward the minimum cost without causing any divergent oscillations.

Figure 2 represents a feed-forward network with the inputs used in the study and a single hidden layer with three hidden neurons. Each of the inputs will have a weight and a bias which will be used to send a value to the hidden neurons, and then, the values will be used as an input for the activation function which will process it and send the calculated value to the output layer.

Fig. 2 Feed-forward neural network used in the present study with used input parameters, one hidden layer, and three hidden neurons



4.3. Hidden neurons

There is no specific guideline or hypothesis for finding the specific number of neurons in the hidden layer. Nevertheless, Huang and Huang [29] derived a constraint on the number of hidden neurons. Huang and Babri [30] showed that the upper limit of the number of neurons required to replicate the desired outputs of the training samples is approximately close to the number of training samples in the training set. Mirko and Christain [31] showed that ten neurons are thought to be a good initial choice in practice. The best configuration can be found by experimental and trial, varying the number of neurons from lower to higher number and observing the output performance.

5. Result and discussion

The input of the ANN was such that the columns contained the number of samples whereas the rows contained the number of attributes. The dataset was split randomly in training, testing, and validation in the ratio of 70:15:15, respectively. The random shuffling meant that all kinds of data occurring throughout the cycle are included in training, testing and validation dataset. It is advised that the three datasets contain data from the entire domain of data to ensure the better performance of the neural network model. The activation function used for the neural network in all layers is tangent sigmoid. The number of hidden

layers and number of hidden neurons for the optimal neural network are found by experimenting with various configurations. The algorithm used is a feed-forward backpropagation algorithm and the training function is gradient descent with momentum. Table 4 represents the mean squared error (MSE) performance for all neural networks with tangent sigmoid as the activation function. The hyper parameters were tuned for each neural network architecture to obtain the best accuracy. The error could vary due to the nature of input hence only the least error for all the computations has been included.

The best performance result achieved was with 26 hidden layers, 15 hidden neurons with tangent sigmoid as the activation function. The mean squared error performance for the validation dataset observed was 1669. The overall accuracy is 88.9% with a validation performance of 92.70% as shown in Fig. 3. The parameters for the ANN

Table 4 Mean squared error (MSE) performance of all neural networks with tangent sigmoid as activation function

| Hidden layers | 15 neurons | 16 neurons | 17 neurons |
|---------------|------------|------------|------------|
| 25 layers | 2210 | 2106 | 2451 |
| 26 layers | 1669 | 1956 | 2065 |
| 27 layers | 1909 | 1975 | 2265 |
| 28 layers | 1820 | 2071 | 2107 |
| 29 layers | 1875 | 2591 | 1887 |

Fig. 3 Percentage accuracy for training, testing, validation, and overall dataset comprising of geomagnetic storms

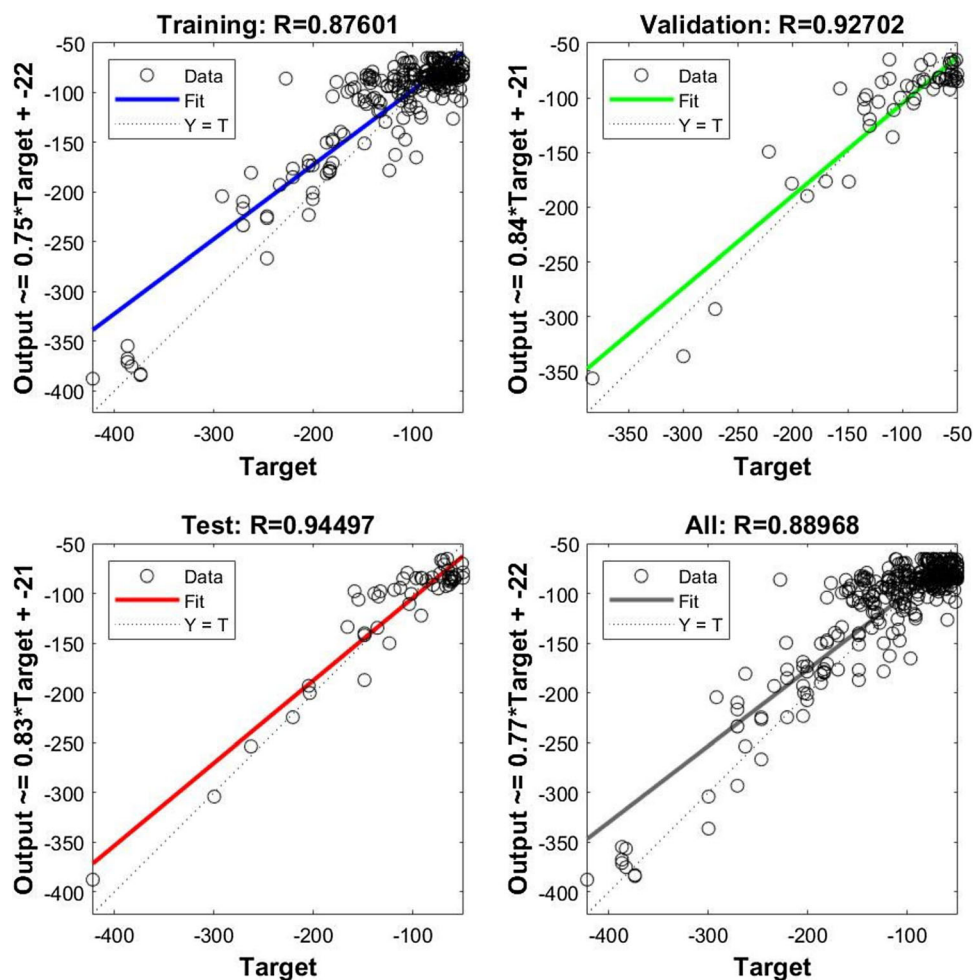


Table 5 Parameters for the best ANN model

| Parameters | Description |
|----------------------------|---|
| Network type | Feed-forward backpropagation |
| Training function | Gradient descent with momentum and adaptive learning rate |
| Adaption learning function | Gradient descent with momentum |
| Activation function | Tangent sigmoid |
| Momentum constant | 0.9 |
| Learning rate | 0.008 |

model with the best performance are given in Table 5. The least mean squared error for the validation dataset is as shown in the plot in Fig. 4. The best performance achieved with tan sigmoid was with 26 hidden layers and 15 hidden neurons. Uwamahoro et al. [5] gave the best performance of 75%. Singh and Singh [16] gave an accuracy of 79%. Lazzus et al. [9] ANN predicts 1–6 h and Lethy et al. [10] predict up to 12 h in advance with a significant correlation coefficient. Our model had an overall accuracy of 88.9% with a validation performance of 92.70% which is close to

the other ANN models considered above. Xu et al. [19] used the Bagging Ensemble-learning algorithm which had an accuracy of above 90% and is better than our ANN model. Gruet et al. [18] combined the Gaussian process method with the long short-term memory (LSTM) neural network and able to predict in 6-h advance with a correlation coefficient greater than 0.87. Thus, our ANN model accuracy is consistent with the results of Gruet et al. [18]. Our study takes into input all the moderate, intense, and super geomagnetic storms that occurred throughout the

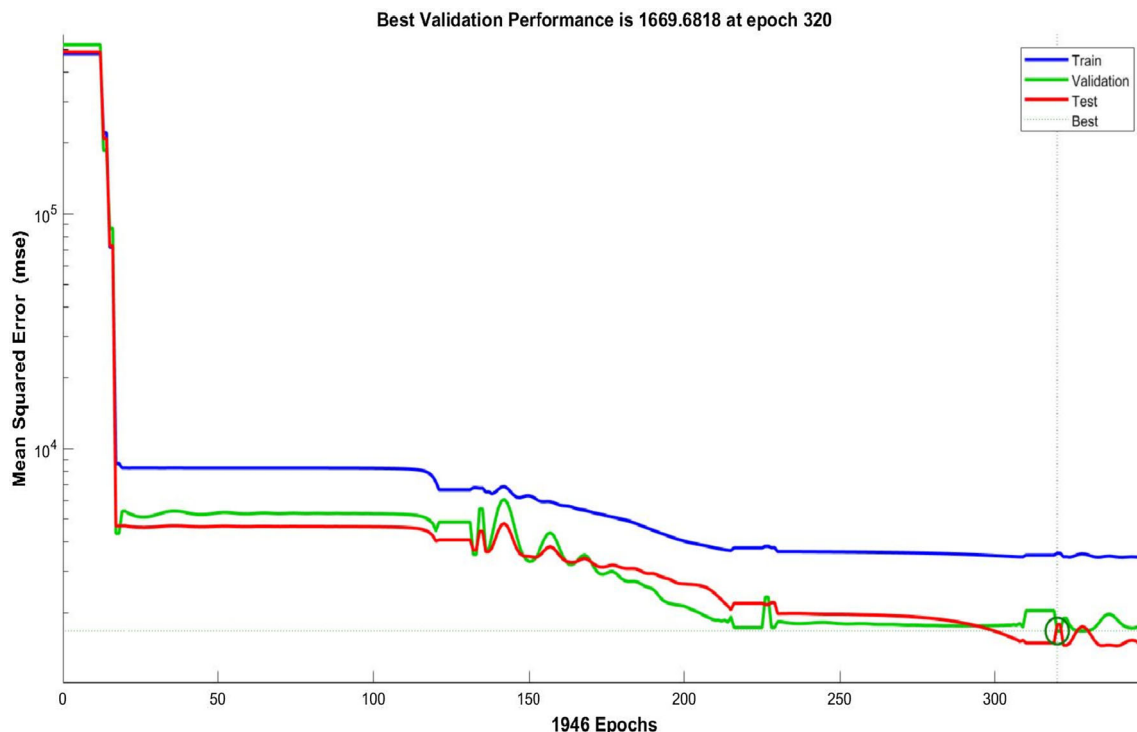


Fig. 4 Mean squared error (MSE) performance for training, testing, and validation dataset for best performance

23rd and 24th solar cycles during 22 years of data that have not been done previously.

6. Conclusions

The present study presents the prediction of moderate, intense, and super geomagnetic storms using inputs from solar and interplanetary parameters over the whole phases of two recent solar cycles 23 and 24 using an artificial neural network (ANN). The complete dataset used in the study was from 1996 to 2018 about 22 years and covers a large period so that the prediction of the Dst index can be reliable. The best performance for the neural network gave an accuracy of 88.9% which provides accuracy that is comparable to other studies that used ANN to predict geomagnetic storms. As the PCA revealed that the data vary without following a recognizable pattern and our ANN model was able to give outputs at such high accuracy, it shows that ANN can be used to predict geomagnetic storms with good effect. Thus, the ANN model can be used to predict geomagnetic storms with an accuracy of 88.9% and can be used in applications such as satellite communication and weather forecast.

Future work should be aimed at including more parameters that can assist with improving the expectation of geomagnetic storms. Various algorithms can be applied

for the prediction apart from ANN which may be more capable of predicting the Dst index and can obtain higher accuracy.

Acknowledgements Acknowledgment has to be given to authors of the LASCO/SOHO catalog list of CMEs, available online at <http://cdaw.gsfc.nasa.gov/CMElist> which was used for the study. The work was also supported by the National Aeronautics and Space Administration for providing us the necessary data of the most significant flare, available online at <http://hesperia.gsfc.nasa.gov/goes/goes> event listings/goes Xray event list 2014.txt.

References

- [1] R K Burton, R L McPherron and C T Russell *J. Geophys. Res.* **80** 4204 (1975)
- [2] S Kugblenu, S Taguchi and T Okuzawa *Earth Planets Space* **51** 307 (1999)
- [3] J A Lazzús, C H López-Caraballo, P Rojas, I Salfate, M Rivera and L Palma-Chilla *J. Phys. Conf. Ser.* **720** 012001 (2016)
- [4] J Uwamahoro and J B Habarulema *Earth Planets Space* **66** 95 (2014)
- [5] J Uwamahoro, L A McKinnell and J B Habarulema *Ann. Geophys.* **30** 963 (2012)
- [6] N Srivastava *Ann. Geophys.* **23** 2969 (2005)
- [7] Y Kamide et al. *J. Geophys. Res. Space Phys.* **103** 17705 (1998)
- [8] Y M Wang, P Z Ye, S Wang, G P Zhou and J X Wang *J. Geophys. Res.* **107** 1340 (2002)
- [9] J A Lazzús, P Vega, P Rojas and I Salfate *Space Weather* **15** 1068 (2017)

- [10] A Lethy, M A El-Eraki, A Samy and H Adeebes *Space Weather* **16** 1277 (2018)
- [11] D H Hathaway *Living Rev. Sol. Phys.* **7** 1 (2010)
- [12] H Lundstedt, H Gleisner and P Wintoft *Geophys. Res. Lett.* **29** 34 (2002)
- [13] H Lundstedt and P Wintoft *Ann. Geophys.* **12** 19 (1994)
- [14] F Valach, M Revallo, J Bochnicek and P Hejda *Space Weather* **7** S04004 (2009)
- [15] M Dryer, Z Smith, C D Fry, W Sun, C S Deehr and S I Akasofu *Space Weather* **2** S09001 (2004)
- [16] G Singh and A K Singh *J. Earth Syst. Sci.* **125** 899 (2016)
- [17] M Chandorkar, E Camporeale and S Wing *Space Weather* **15** 1004 (2017)
- [18] M A Gruet, M Chandorkar, A Sicard and E Camporeale *Space Weather* **16** 1882 (2018)
- [19] S B Xu, S Y Huang, Z G Yuan, X H Deng and K Jiang *Astrophys J Suppl S* **248** 14 (2020)
- [20] G E Brueckner et al. *Geophys. Res. Lett.* **25** 3019 (1998)
- [21] B Vrsnak, D Maricic, A L Stanger, A M Veronig, M Temmer and D Rosa *Sol. Phys.* **241** 85 (2007)
- [22] E Robbrecht, D Berghmans and R A M Van der Linden *Astrophys. J.* **691** 1222 (2009)
- [23] N Srivastava and P Venkatakrishnan *J. Geophys. Res.* **109** A10103 (2004)
- [24] W D Gonzalez and B T Tsurutani *Planet. Space Sci.* **35** 1101 (1987)
- [25] E Echer, W D Gonzalez, B T Tsurutani and A L C Gonzalez *J. Geophys. Res.* **113** A05221 (2008)
- [26] I T Jolliffe *Principal Component Analysis*. (New York: Springer-Verlag) (2002)
- [27] M M Poulton *Geophysics* **67** 979 (2002)
- [28] C Bishop *Pattern Recognition and Machine Learning*. (New York: Springer-Verlag) (2006)
- [29] S C Huang and Y F Huang *IEEE Trans. Neural Netw. Learn. Syst* **2** 47 (1991)
- [30] G B Huang and H A Babri *IEEE Trans. Neural Netw. Learn. Syst* **9** 224 (1998)
- [31] B Mirko and J Christian *Geophysics* **65** 1032 (2000)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.