Preliminary Modelling of Solar Quiet Geomagnetic Field Average using Non-Linear Autoregressive with Exogeneous Input (NARX)

M. H. Hashim, M. H. Jusoh*, K. Burhanudin
School of Electrical Engineering, College of Engineering,
Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
*huzaimy@uitm.edu.my

I. M. Yassin,
Microwave Research Institute, Universiti Teknologi MARA,
Shah Alam, Selangor, Malaysia

N. S. A Hamid
Department of Applied Physics, Faculty of Science and Technology,
University Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Z. M. Radzi
Space Exploration and Science Department,
Malaysian Space Agency (MYSA), Banting, Selangor, Malaysia

A. Yoshikawa
International Center for Space Weather Science and Education (ICSWSE)
Kyushu University, Japan

ABSTRACT

This paper discusses geomagnetic field attempt modelling using an Artificial Neural Network (ANN). The local horizontal component of geomagnetic field data was collected on April 2011 (equinox) during a solar quiet day at recent solar cycle inclination-24 using the Magnetic Data Acquisition System (MAGDAS) in Langkawi, Malaysia, in the low latitude region. The calculated average values (mean) of the H component geomagnetic field variation during Equinox 2011 characterised the dominant geomagnetic field during that particular solar cycle. The difference in amplitude of maximum and minimum values shows a regular diurnal variation of the geomagnetic field during Sq in

the low latitude region. The output training utilised these calculated mean values during the modelling attempt. Meanwhile, the input training utilised proton density, solar wind plasma speed, plasma flow pressure, and Interplanetary Magnetic Field (IMF) space data using Non-Linear Auto Regressive Input (NARX). The best modelling outputs were 3, 2, and 3 for input delay (nu), output delay (ny), and hidden layer size (h), respectively. Residual test and model fit analysis show an unbiased and high overlap between predicted and actual geomagnetic field average values, suggesting the model can potentially anticipate the geomagnetic field average during solar quiet in November (month of Equinox) on solar cycle inclination.

Keywords: *Equatorial Region; Space Weather; Geomagnetic Field; ANN; NARX*

Introduction

The Earth's magnetic field is affected by solar events like coronal holes, mass ejections, and solar flares [1]-[2]. They send out sudden magnetic impulses that change the Earth's magnetosphere and magnetic field [3]-[4] and have other effects on the atmosphere [4]-[6]. By looking at these changes in the magnetic field, scientists can learn more about how the Sun changes. During low solar-terrestrial activities, the Earth's magnetic field, also known as the geomagnetic field, is called quiet daily or solar quiet "Sq" [6]. Lots of strange events on the Sun are due to its core, which then appear in the solar cycle. Like Earth, the Sun's magnetic field has a north pole and a south pole, but it is much more chaotic and disorderly than Earth's. Every 11 years, the Sun's magnetic field changes. The Sun then goes down to its lowest point, which marks the beginning of a new cycle, and the north pole turns into the south pole and vice versa. As the cycle continues, the Sun's turbulent behaviour worsens until it reaches a peak, which flips the magnetic field.

The investigation of the Earth's magnetic field during Sq has attracted considerable interest and studied during previous solar cycles using data from several observatories at various latitudes and longitudes. Since the earliest findings by Chapman and Stagg [7] and Hibberd [8], numerous studies have investigated the variations in the Sq geomagnetic field. According to Pedatella et al. [9], the primary cause of the daily geomagnetic variation is an external current system, i.e., currents flowing above the atmosphere of the Earth.

The amplitude and variation of the geomagnetic field depend substantially on the eclipse's magnitude, space weather, local time, location of the magnetic observatory, and season [10]. In 1965, Matsushita and Maeda discovered the strong seasonal, hemisphere, and latitude dependence of the Sq field [11]. In 2011, Madeeha Talha et al. [12] used Sq field data from the South Asia Region to characterise the local geomagnetic field based on seasons.

Based on Lloyd's seasonal classifications, the months can be categorised into three seasons corresponding to the June Solstice (May, June, July, and August), December Solstice (November, December, January, and February), and Equinox (September, October, March, and April) [13]. This preliminary modelling chooses Sq geomagnetic data during April 2011 (equinox) as described in Figure 1.
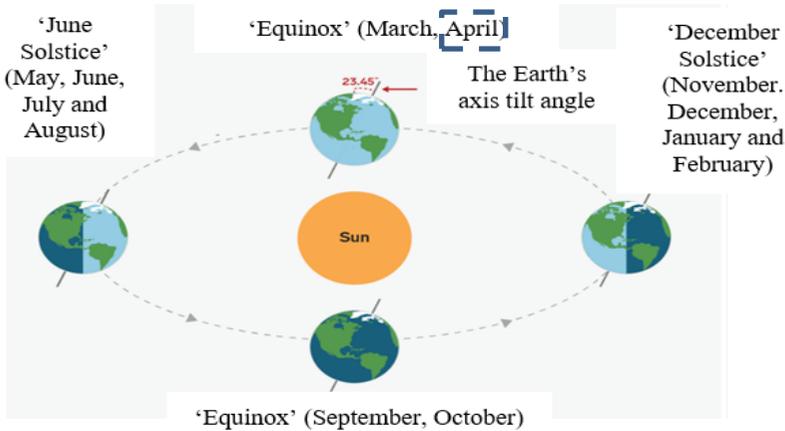
Figure 1: Earth's Equinoxes and Solstices. The 'rectangular mark' is the period (April, in equinox) of Sq geomagnetic field data taken for this study

The most recent solar cycle (Solar Cycle-24) has been especially significant for this study, as it is the weakest in the Space Era, and solar-generated space weather has been exceptionally mild [14], preceded by an extraordinarily quiet and protracted solar minimum [15].

Since 2004, the Malaysia Space Agency (MYSA) has been involved in geomagnetic field monitoring; its first geomagnetic field monitoring station was at Langkawi National Observatory (LNO) in Langkawi, Malaysia. The Magnetic Data Acquisition System (MAGDAS) monitors the daily geomagnetic field near the deep geomagnetic equatorial region [16].

The Artificial Neural Network (ANN) models investigated several novel approaches to space weather dynamics. Numerous of these models have beneficial real-world applications, allowing scientists to predict future values associated with various geophysical and solar physical systems [17]. In a previous study, Nuraeni et al. [18] utilised the NARX network to simulate conditions of the Disturbance Time Storm Index (DST) of geomagnetic fields using the interplanetary magnetic field (IMF) as the exogenous input. Therefore, this study aims to produce a model of daily averages of the *H* component of geomagnetic field values obtained from a ground geomagnetic field observatory network called the Magnetic Data Acquisition System

(MAGDAS). This attempt predicts the outline characterisation of the geomagnetic *H* component during solar quiet on solar cycle inclination based on seasonal variabilities, e.g., Equinox. At low latitudes, only the horizontal component *H* has been selected for the modelling attempt since the component has a crucial contribution to the total field, as mentioned by Rastogi and Iyer [19].

Methodology

Solar cycle incline estimation

The yearly Sunspot Number (SSN), obtained from the OMNI online data centre, was plotted to estimate the inclination period of solar cycle-24.

Mean geomagnetic field data for sun quiet (MSq) calculation

For this preliminary modelling test, the mean of geomagnetic field fluctuation during April 2011 (one month in equinox 2011) of the geomagnetic field H component, acquired from LNO during *Sq*. April 2011 was a month during the equinox of the most recent solar cycle minimum 24. Only the H component contributes significantly to the total geomagnetic field in low-latitude regions [20].

The H component provides an approximation of the strength of the Equatorial ElectroJet (EEJ) current. In addition, it integrates the impacts of all currents running across the Earth's ecosystem, which shows the daily variation in local time of the Earth's magnetic field during Sq.

Identifying the five quietest days (Sq) of the month is according to the World Data Center (WDC) for Geomagnetism. Therefore, Equation (1) calculates the mean of Sq (H) variation for the month.

$$M_{sq}(H) = \frac{\sum_{d=1}^{5} H_d}{5} \tag{1}$$

where $M_{sq}(H)$ is the mean of *H* components of the geomagnetic field for the five quietest days in April 2011, and H_d is the geomagnetic field value of each day.

Space data

In a statistical prediction of the disturbance storm time of geomagnetic fluctuations, Caswell [20] suggested deploying several solar wind parameters (proton density in Ncm^{-3}, solar wind plasma speed in kms^{-1}, and plasma flow pressure in nPa) and Interplanetary Magnetic Field (IMF) parameters (Bx, By, and Bz in nT) as independent variables. All data collected from Goddard Space Flight Center, Space Physics Data Facility, and OmniWeb is for input training.

This preliminary experimental modelling utilised the solar wind parameters during the five quietest days in April 2011, during the equinox.

The multilayer perceptron (MLP) – NARX model

The modelling network utilised nonlinear processes. The mean of the geomagnetic field acted as the target output. MATLAB 2016a built an ANN to produce a nonlinear autoregressive network with exogenous input (NARX). The geomagnetic field means $M_{Sq}(H)$ from II (B) served as the target variable specified for the auto-regressive function. Meanwhile, the solar wind and IMF data obtained from II (C) functioned as the input training.

Overall, 70% of input data was allocated for training, 15% for validation, and 15% for testing trials. A previous study suggested a maximum of 1,000 epochs as the limit of neural network learning [20]. This modelling experiment employed Levenberg Marquardt (LMANN) as a training algorithm, which set the ratio of input delay (nu), output delay (ny), and hidden units (h) to 1:1:3 for the end loop. A closed-loop series-parallel architecture was used for the NARX network, as shown in Figure 2.

In a closed loop NARX network, the output of the current time step serves as the input for the subsequent time step, which is slightly different from an open loop architecture. Closed loop systems return the predicted results of the model as input, in contrast to the open loop architecture, which uses actual outputs as feedback to the network.
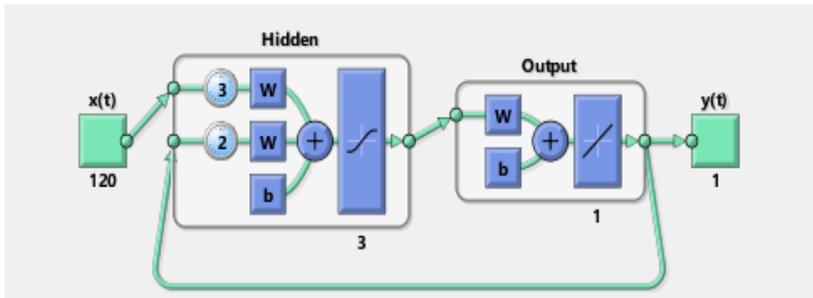
Figure 2: One of the networks was designed with a hidden unit, *h* of 3, input delay (*nu*) of 3, and output delay (*ny*) of 2. Input layer *x*, consisting of 120 exogenous input time series with six solar wind and IMF different parameters. The average of a geomagnetic field is the targeted output

Validation of model

Model validation is required to ensure that the model adequately reflects system behaviour. Model fit and residual tests are two types of tests for validating the model. The One Step Ahead (OSA) was used for the model fit analysis, while the residual test analysis applied Mean Squared Error (MSE),

error histogram, and regression tests. Equation (2) shows the calculation for the MSE.

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{2}$$

where *MSE* is the MSE value, n is the number of data points, y is the actual output, and $\hat{y}$ is the predicted output. The three smallest MSE values indicated the best modelling outputs; however, this paper only discusses the best value.

Result and Discussion

Solar cycle inclination
Figure 3 presents the yearly SSN. It shows the solar cycle started to incline in 2009 before finally reaching its peak in 2014.
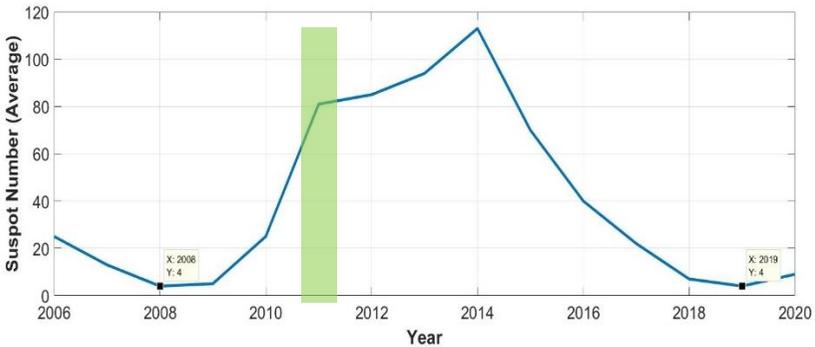
Figure 3: SSN for Solar Cycle-24. The grey region represents the period of this study, which is 2011

Mean geomagnetic field data for sun quiet *(MSq)*
The variance of the seasonal *MSq* for the H component, *MSq(H)*, during the April Equinox 2011 is revealed in Figure 4.
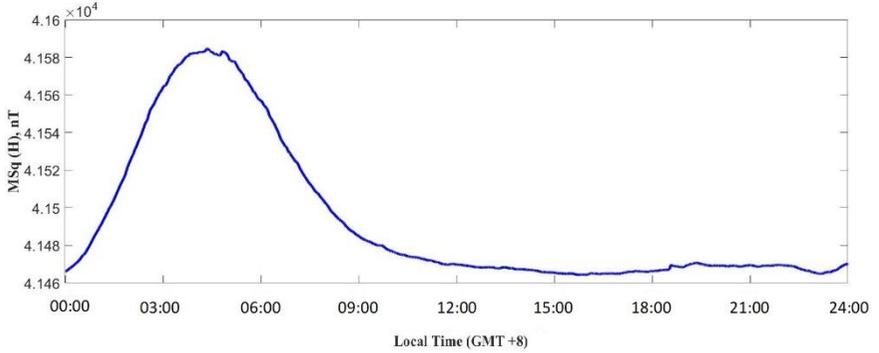
Figure 4: Mean Sq of Geomagnetic *H* Component on April Equinox 2011 at local time (GMT+8)

The mean Sq pattern of the geomagnetic field H component on the April Equinox 2011 is similar to local diurnal variation monitored for 24 hours [21]-[23]. The maximum value of M_{sq} is 4.159 x 10^4nT during the daytime, while the lowest value is 4.146 x 10^4nT during the nighttime. These few tens of nanoteslas of amplitude show a regular diurnal variation of the geomagnetic field during Sq in low and middle latitudes, as mentioned by Wu et al. [24].

Mean square error (MSE)
The best three output models are selected based on the three smallest MSE values, as shown in Table 1.

Table 1: Input Delays, Feedback Delays, Size of Hidden Layer, MSE, and Square Root MSE based on the three lowest MSE

	Input Delays (nu)	Feedback Delays (ny)	Hidden Layer Size (h)	Mean Square Error (MSE)	Square Root MSE
1.	3	2	3	4.997	2.235
2.	1	2	2	19.898	4.460
3.	1	3	3	40.901	6.395

NARX closed loop training performance
Figure 5 shows the effectiveness of NARX closed-loop training performance. The training is terminated at epoch 4 with a final best validation performance MSE score of 16.741 as a result of Early Stopping (ES) as automatically activated in MATLAB to avoid the NARX loop overtrained.

Best Validation Performance is 16.7413 at epoch 4
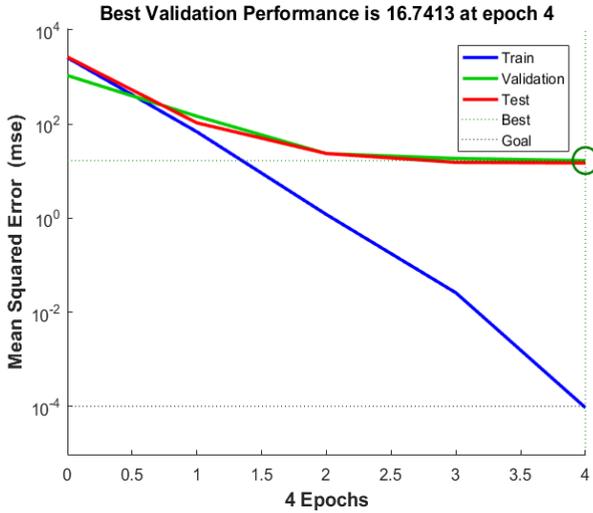
Figure 5: MLP Training Performance for $nu = 3$, $ny = 2$, and h = 3

Analysis of regression analysis

Regression analysis evaluates how well a machine learning model fits the data relative to its training data. A regression model is well-fitting when the Pearson correlation coefficient is greater than 0.9718, which indicates that the desired output closely matches the target output, as shown in Figure 5. With a few exceptions, the desired output lies on or near the fitted target output line, signifying a good fit with the original data.

Analysis of one-step ahead (OSA)

The OSA test evaluates the predictive ability of a model as it attempts to predict one step using historical data. Figure 7 depicts the OSA results for the average H-component geomagnetic field (actual) versus the modelling output produced by NARX. A high degree of overlap between actual and predicted values indicates a good agreement.
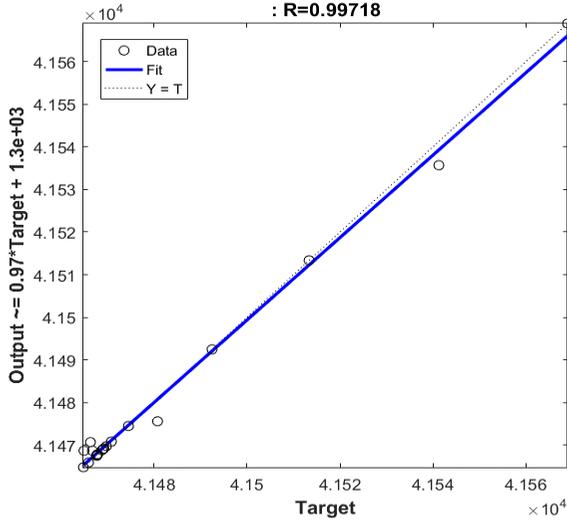
Figure 6: Regression analysis for $nu = 3$, $ny = 2$, and h = 3
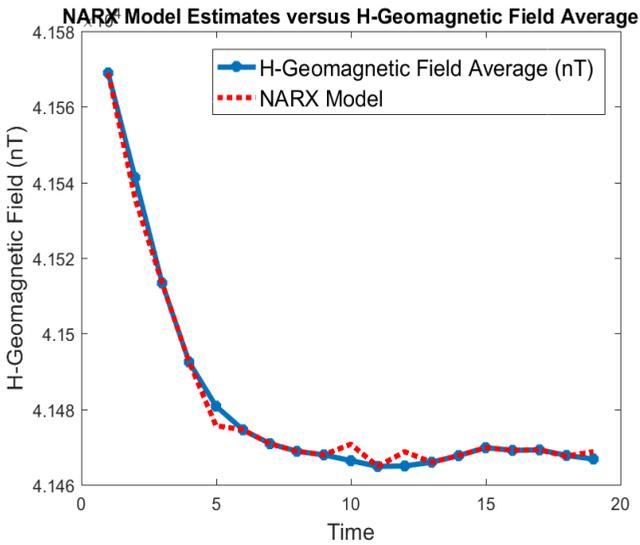
Figure 7. OSA for a) $nu = 3$, $ny = 2$, and h = 3

Analysis of histogram error

The histogram of errors shown in Figure 8 also supports the conclusion of randomly distributed residuals due to the appearing distribution pattern as a Gaussian 'bell-shaped like' with most of the frequency counts grouped in the middle, approaching zero error and tapering off at left and right shoulders. Therefore, we consider the NARX model to be valid and acceptable as a result.

Error Histogram with 20 Bins
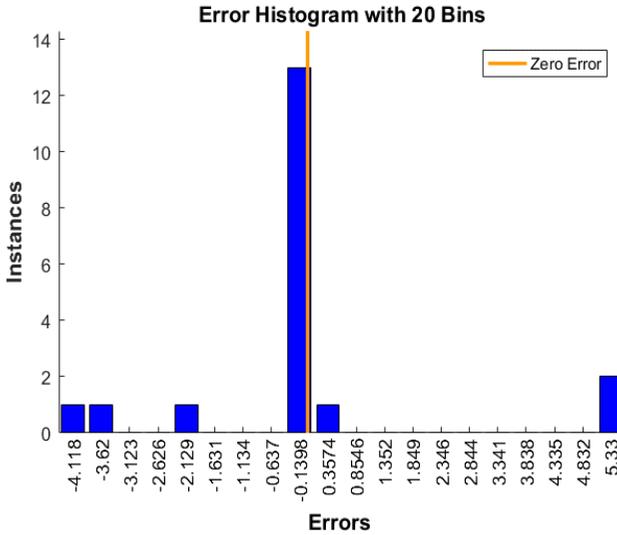
Figure 8: Histogram of MLP residual for a) $nu = 3$, $ny = 2$, and $h = 3$

Conclusion

This paper presents NARX preliminary modelling for the Solar Quiet Geomagnetic Field Average during the April 2011 Equinox solar cycle 24 inclination. The data from Malaysia's low latitude region, which represents part of Southeast Asia, is utilised for these modelling attempts. The solar cycle 24 inclination period is identified between 2009 and 2014. The average value of the Sq geomagnetic field H component is 4.158×10^4nT during the daytime, while the lowest value is 4.146×10^4nT during the nighttime. The best modelling outputs are 3, 2, and 3 for input delay (nu), output delay (ny), and hidden layer size (h), respectively. However, the parameter values might be optimised using the parameter optimisation platform to reduce MSE values. The residual analysis using histograms and regression confirms that the model is unbiased. Based on validation and fitting tests, the preliminary model successfully shows high overlaps between predicted and actual values of the

geomagnetic field average. Thus, this suggests that the model can potentially predict the geomagnetic field average during solar quiet in November (the month of Equinox) on solar cycle inclination.

Acknowledgement

References

[1] S. Kahler, D. Reames, and E. Cliver, "Coronal sources of impulsive fe-rich solar energetic particle events," in *Proceedings of Science*, 2015, The Hague, Netherlands, vol. 30-July-20, doi: 10.22323/1.236.0049.

[2] N. Gopalswamy, "Large-Scale Solar Eruptions," *Astrophys. Sp. Sci. Proc.*, vol. 0, no. 202549, pp. 53–71, 2010, doi: 10.1007/978-3-642-11341-3_4.

[3] A. V. Moiseev, D. G. Baishev, E. S. Barkova, A. Du, and K. Yumoto, "Specific features of the generation of long-periodic geomagnetic pulsations in the event on June 25, 2008," *Cosm. Res.*, vol. 53, no. 2, pp. 111–118, Mar. 2015, doi: 10.1134/S0010952515020057.

[4] R. S. Selvaraj, S. T. Selvi, and and S. P. V. Priya, "Association between surface ozone and solar activity," *Indian J. Sci. Technol.*, vol. 3, no. 3, pp. 332–334, Mar. 2010, doi: 10.17485/IJST/2010/V3I3.18.

[5] R. Samuel Selvaraj, T. Gopinath, and K. Jayalakshmi, "Statistical relationship between surface ozone and solar activity in a tropical rural coastal site, India," *Indian J. Sci. Technol.*, vol. 3, no. 7, pp. 792–794, 2010, doi: 10.17485/IJST/2010/V3I7/29816.

[6] S. Matsushita, "Solar quiet and lunar daily variation fields," in *International Geophysics*, 1967, vol. 11, no. P1, pp. 301–424, doi: 10.1016/B978-0-12-480301-5.50013-6.

[7] S. Chapman and J. Bartels, *Geomagnetism: Geomagnetic and Related Phenomena*, 1st ed., vol. 1. 1940.

[8] Y. Yamazaki, K. Häusler, and J. A. Wild, "Day-to-day variability of midlatitude ionospheric currents due to magnetospheric and lower

atmospheric forcing," *J. Geophys. Res. Sp. Phys. Day*, vol. 121, no. 7, pp. 7067–7086, 2016, doi: 10.1002/2016JA022817.

[9] N. M. Pedatella, J. M. Forbes, and A. D. Richmond, "Seasonal and longitudinal variations of the solar quiet (Sq) current system during solar minimum determined by CHAMP satellite magnetic field observations," *J. Geophys. Res. Sp. Phys.*, vol. 116, no. A4, p. 4317, Apr. 2011, doi: 10.1029/2010JA016289.

[10] L. F. Chernogor, "Geomagnetic Effect of the Solar Eclipse of June 10, 2021," *Kinemat. Phys. Celest. Bodies*, vol. 38, no. 1, pp. 11–24, 2022, doi: 10.3103/S0884591322010020.

[11] S. Matsushita and H. Maeda, "On the geomagnetic lunar daily variation field," *J. Geophys. Res.*, vol. 70, no. 11, pp. 2559–2578, 1965, doi: 10.1029/JZ070i011p02559.

[12] M. Talha, G. Murtaza, J. L. Rasson, N. Ahmed, and M. Peerzada, "Sq(H) field variations at Sonmiani geomagnetic observatory, Pakistan for solar cycle 24," *Adv. Sp. Res.*, vol. 67, no. 1, pp. 66–74, 2021, doi: 10.1016/j.asr.2020.09.013.

[13] E. O. Elemo, M. O. Ehigiator, and R. Ehigiator-Irughe, "Seasonal variations of the Vertical Total Electron Content (VTEC) of the ionosphere at the GNSS cor station (SEERL) UNIBEN and three other cors stations in Nigeria," *Niger. J. Technol.*, vol. 37, no. 2, p. 286, 2018, doi: 10.4314/njt.v37i2.1.

[14] N. Gopalswamy, S. Akiyama, S. Yashiro, H. Xie, P. Makela, and G. Michalek, "The Mild Space Weather in Solar Cycle 24," 2015, [Online]. Available: http://arxiv.org/abs/1508.01603.

[15] S. Basu, "The peculiar solar cycle 24-where do we stand?," *J. Phys. Conf. Ser.*, vol. 440, no. 1, p. 12001, 2013, doi: 10.1088/1742-6596/440/1/012001.

[16] S. Law, S. Communication, S. Engineering, R. Sensing, and S. Science, "ASM Science Malaysia in Space," vol. 12, no. 2, 2019.

[17] E. Camporeale, "The Challenge of Machine Learning in Space Weather: Nowcasting and Forecasting," *Sp. Weather*, vol. 17, no. 8, pp. 1166–1207, Aug. 2019, doi: 10.1029/2018SW002061.

[18] F. Nuraeni, M. Ruhimat, M. A. Aris, E. A. Ratnasari, and C. Purnomo, "Development of 24 hours Dst index prediction from solar wind data and IMF Bz using NARX," *J. Phys. Conf. Ser.*, vol. 2214, no. 1, 2022, doi: 10.1088/1742-6596/2214/1/012024.

[19] R. G. Rastogi and K. N. Iyer, "Quiet Day Variation of Geomagnetic H-Field at Low Latitudes," *J. Geomagn. Geoelectr.*, vol. 28, no. 6, pp. 461–479, 1976, doi: 10.5636/jgg.28.461.

[20] J. M. Caswell, "A Nonlinear Autoregressive Approach to Statistical Prediction of Disturbance Storm Time Geomagnetic Fluctuations Using Solar Data," *J. Signal Inf. Process.*, vol. 2014, no. 02, pp. 42–53, May 2014, doi: 10.4236/JSIP.2014.52007.

[21] R. Umar *et al.*, "Magnetic Data Acquisition System (MAGDAS) Malaysia: installation and preliminary data analysis at ESERI, UNISZA," *Indian J. Phys.*, vol. 93, no. 5, pp. 553–564, 2019, doi: 10.1007/s12648-018-1318-x.

[22] Z. Iffah Abd Latiff *et al.*, "The first solar-powered Magdas-9 installation and possible geomagnetically induced currents study at Johor, Malaysia," in *Journal of Physics: Conference Series*, Mar. 2019, vol. 1152, no. 1, doi: 10.1088/1742-6596/1152/1/012030.

[23] N. A. Zakaria *et al.*, "Installation and preliminary data analysis of Penang magnetic data Acquisition system (MAGDAS) in Malaysia," *Adv. Sp. Res.*, vol. 67, no. 7, pp. 2199–2206, Apr. 2021, doi: 10.1016/j.asr.2021.01.009.

[24] Y. Wu, L. Liu, and Z. Ren, "Equinoctial asymmetry in solar quiet fields along the 120◦ e meridian chain," *Appl. Sci.*, vol. 11, no. 19, 2021, doi: 10.3390/app11199150.