Earth, Planets and Space

FULL PAPER Open Access

Investigation of the relationship between geomagnetic activity and solar wind parameters based on a novel neural network (potential learning)

Ryozo Kitajima[1*†], Motoharu Nowada[2*†] and Ryotaro Kamimura[3]

Abstract

Predicting geomagnetic conditions based on in-situ solar wind observations allows us to evade disasters caused by large electromagnetic disturbances originating from the Sun to save lives and protect economic activity. In this study, we aimed to examine the relationship between the K_p index, representing global magnetospheric activity level, and solar wind conditions using an interpretable neural network known as potential learning (PL). Data analyses based on neural networks are often difficult to interpret; however, PL learns by focusing on the "potentiality of input neurons" and can identify which inputs are significantly utilized by the network. Using the full advantage of PL, we extracted the influential solar wind parameters that disturb the magnetosphere under southward Interplanetary magnetic field (IMF) conditions. The input parameters of PL were the three components of the IMF (B_x, B_y, B_z), solar wind flow speed (V_x), and proton number density (N_p) in geocentric solar magnetospheric (GSM) coordinates obtained from the OMNI solar wind database between 1998 and 2019. Furthermore, we classified these input parameters into two groups (targets), depending on the K_p level: $K_p = 6$–9 (positive target) and $K_p = 0$ to 1 + (negative target). Negative target samples were randomly selected to ensure that numbers of positive and negative targets were equal. The PL results revealed that solar wind flow speed is an influential parameter for increasing K_p under southward IMF conditions, which was in good agreement with previous reports on the statistical relationship between the K_p index and solar wind velocity, and the K_p formulation based on the IMF and solar wind plasma parameters. Based on this new neural network, we aim to construct a more correct and parameter-dependent space weather forecasting model.

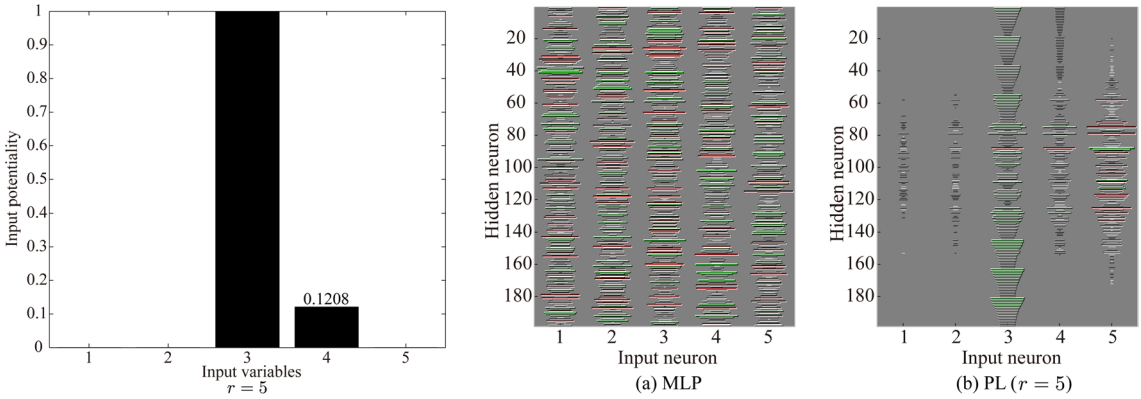
Keywords: Space weather modeling, Solar wind conditions, Geomagnetic activity, Neural network, Data classification

†Ryozo Kitajima and Motoharu Nowada have equally contributed to this paper

*Correspondence: r.kitajima@eng.t-kougei.ac.jp; moto.nowada@sdu.edu.cn

[1] Department of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan
[2] Shandong Provincial Key Laboratory of Optical Astronomy and Solar-Terrestrial Environment, Institute of Space Sciences, Shandong University, 180 Wen-Hua West Road, Weihai, Shandong 264209, People's Republic of China
Full list of author information is available at the end of the article

Graphical Abstract



Introduction

The terrestrial magnetosphere protects life from the harmful radiation effects associated with the high-speed plasma streams (solar wind) and is constantly undergoing dynamic changes due to interactions with solar wind and the interplanetary magnetic field (IMF) originating from the Sun (e.g., Black, 1967; Glassmeier et al. 2009; Glassmeier and Vogt, 2010). Drastic changes from quiet to active geomagnetic conditions start from a violation of the "frozen-in-condition" of the geomagnetic field caused by reconnecting the geomagnetic field with solar wind field lines, known as magnetic reconnection. Substorms, magnetic storms and auroral signatures are phenomena observed in the magnetosphere that occur due to reconnection-associated transfers of solar wind energy into the magnetosphere, and the resultant magnetospheric activity is of a high level.

The K index, defined as the value representing the level of geomagnetic disturbances driven by the solar wind based on the perturbations in the Earth's magnetic field, was used to determine geomagnetic conditions. Moreover, it defines geomagnetic disturbances using an integer with signs of '+' and '−' in the range 0–9 values, indicating a threshold to judge whether or not the Earth's magnetospheric activity is under disturbed or quiet level. This index was first introduced by Bartels (1939) and is derived from the maximum fluctuations of horizontal components observed on a magnetometer with a temporal resolution of 3 h. Today, the K_p index, derived from the weighted average of the K indices of 13 geomagnetic observatories around the world (Bartels 1949), is known as one of parameters for measuring geomagnetic activity. Other examples of indices representing geomagnetic activity use the current intensity associated with aurora (AL and AU indices) and the strength of looped currents (ring current) flowing around the magnetic equator region built up by magnetic storms (D_{st} index). Since their inception, K_p values have been used as an important and reliable index to representations of global geomagnetic activity. However, the time resolution of K_p (3 h) is lower than those of the other geomagnetic indices, such as AL, AU, PC (Polar Cap), and D_{st}, which have time resolutions ranging from 1 min. to 1 h (see Rangarajan 1987).

Based on solar wind parameters, these indices to show global or specific region geomagnetic activity have been obtained. In particular, K_p has been derived from solar wind parameters at Lagrange point 1 (L1) obtained by satellites (e.g., Wing et al. 2005; Wintoft et al. 2017; Zhelavskaya et al. 2019; Shprits et al. 2019). Newell et al. (2008) formulated the K_p index based on the IMF and solar wind–magnetosphere coupling functions, which are equations that quantitatively evaluate the amount of solar wind energy inputs to the magnetosphere based on IMF and solar wind plasma parameters (Newell et al. 2007). The equations are as follows:

$$K_p = 0.05 + 2.244 \times 10^{-4} \left(\frac{\mathrm{d}\Phi_{MP}}{\mathrm{d}t} \right) + 2.844 \times 10^{-6} N_p^{\frac{1}{2}} V_{\mathrm{sw}}^2 \tag{1}$$

$$\frac{\mathrm{d}\Phi_{MP}}{\mathrm{d}t} = V_{sw}^{\frac{4}{3}} B_t^{\frac{2}{3}} \sin^{\frac{8}{3}} \left(\frac{\theta_{\mathrm{clock}}}{2} \right) \tag{2}$$

According to Eq. (1), K_p can be represented by the solar wind proton number density (N_{p}); velocity (V_{sw}); IMF clock angle, defined as the angle between the IMF-B_{y} and -B_{z} components ($\theta_{\mathrm{clock}} = \arctan(\text{IMF-}B_{\mathrm{y}}/$

IMF-B_z)); and IMF intensity (B_t), which is included in Newell's solar wind–magnetosphere coupling function ($d\Phi_{MP}/dt$), as calculated using Eq. (2). Equations (1) and (2) show that the K_p index can be described as the function of solar wind velocity. The closed relation between K_p and solar wind velocity is also advocated by studies by Snyder et al. (1963) and Elliott et al. (2013). Nevertheless, it is difficult to uniquely determine the geomagnetic disturbance level based on solar wind conditions.

Recently, machine learning (or deep learning) approaches have been used to predict K_p. The artificial neural network (NN) is one of the most popular algorithms for forecasting the K_p, D_{st} and polar cap (PC) indices (e.g., Nagai 1994; Costello 1998). Later, Boberg et al. (2000) and Wing et al. (2005) developed a prediction model based on NN using IMF, solar wind plasma density, and solar wind velocity as input parameters. Boberg et al. (2000) sequentially built a multi-layer feed-forward network using IMF-B_z component, solar wind plasma density (N_p), and velocity (V_{sw}) as the input parameters, and evaluated the developed algorithm in terms of "training", "validation", and "test" based on the correlation and root-mean-square error (RMSE). Furthermore, an NN was developed by Bala and Reiff (2012) to forecast three indices: K_p, D_{st}, and AE (as defined by AU–AL). They obtained and compared several forecasting patterns of the K_p index with various solar wind input parameters and found significant differences in the RMSE and correlation between the obtained models. They also evaluated the prediction time for forecasting performance and concluded that RMSE tends to become larger as K_p prediction time increases.

Following these NNs, Ji et al. (2013) introduced a support vector machine (SVM) to build a K_p forecasting model and evaluated the forecasting results from SVM by comparing the K_p prediction results with those from an NN. They constructed a forecasting model under high magnetic activity conditions with the K_p level higher than 6 ($K_p \geq 6$). Tan et al. (2018) constructed and evaluated a K_p forecasting model using the solar energy input function (a coupling function) and the associated viscous term as inputs (Newell et al. 2008). Their models can also consider the K_p forecasting error and were built based on long short-term memory (LSTM), which was developed from recurrent NNs (RNNs) (Hochreiter and Schmidhuber 1997).

In this study, we developed the K_p classification model based on potential learning (PL) to extract the solar wind parameters that significantly effect geomagnetic disturbances. PL has been used to conduct analyses, where high model performance and high interpretability are required. For example, in a study that applied PL to supermarket data (ID-POS) by Kitajima et al. (2016a), a model was developed that used the "consumer's purchase behavior in the past three months" as an input parameter to determine the "customer's probability to visit the store in two months in the future". They determined that the model based on PL performed better than the conventional method and succeeded to extract an important variable. In addition, PL has been applied to data in various fields, such as the messages displayed in Twitter at the time of a disaster (Kitajima et al. 2016b) and the words involved in the president messages of the Japanese companies (Kitajima et al. 2019). Since PL has been used for data analysis in various fields, we aim to identify the most significant parameters that disturb the magnetosphere based on PL. Furthermore, we will run several PLs by changing the parameters which are included in an essential equation in the PL algorithm and evaluate the performance of the application of PL to space physics data.

This paper is organized as follows. "Data and methodology" section presents the data used, and methodology in this study. We also describe the details of the PL structure in this section. The evaluation of the performance of PL, the results obtained based on PL and the differences between PL and MLP are shown in "Results" section. In "Summary and discussion" section, finally, we present the discussion and our conclusions of this study.

Data and methodology

Database compiling

In this study, we used the three components of magnetic field vectors of the interplanetary magnetic field (IMF; B_x, B_y, B_z), solar wind plasma density (N_p), solar wind velocity (V_x) in geocentric solar magnetospheric (GSM) coordinates from January 1 1998 to December 31 2019, as input parameters for PL. To make the classification model of PL, we used the global geomagnetic activity index (K_p index) for the same interval as the input parameters. Detailed information on the parameters of

Table 1 Detailed information on the parameters in compiled database used in this neural network

No.	Input Parameters	Unit
1	B_X [IMF GSM-X component]	nT
2	B_Y [IMF GSM-Y component]	nT
3	V_X [Solar wind velocity in GSM]	km/s
4	N_P [Ion number density]	/cm^3
5	$B_s^{\dagger}$ [Southward IMF GSM-Z component]	nT

$$B_s^{\dagger} = \begin{cases} 0 \ (B_z > 0), \\ B_z \ (B_z \leq 0). \end{cases}$$

These input solar wind parameters are normalized in the range between 0 and 1 using the equation of $x' = (x - \min(x))/(\max(x) - \min(x))$

the solar wind and geomagnetic activity index is summarized in Table 1. Note that we normalized these input solar wind parameters in the range between 0 and 1 using the equation: $x' = (x - \min(x))/(\max(x) - \min(x))$, where x' is the normalized parameter whose range is 0 to 1, x is input solar wind parameter, min(x) and max(x) are minimum and maximum values of the input solar wind parameter, respectively. The solar wind parameters with temporal resolution of 1 min. obtained from the OMNI database and the K_p index whose time resolution is 3 h were utilized. We calculated the 3 h average of the solar wind data to give these parameters the same temporal resolution as K_p. If the parameter had a data gap larger than 40%, the averages were not computed. To further exclude the observation data in the magnetosphere from the database, we established a threshold, where the satellite GSM-X component (sun-earthward) was larger than the nominal nose point (~ 15 R_E) of the model bow shock, proposed by Farris and Russell (1994), that is, the database used completely comprised the observation values in interplanetary space. Therefore, we can reveal the relationship between the geomagnetic activity and solar wind conditions without the significant influences originating from the Earth's bow shock.

In this study, we considered only magnetospheric activity under the southward (negative) IMF-B_z case and excluded the northward (positive) IMF-B_z component as this NN input parameter, identified with "B_s" in Table 1. There were two main reasons for this criterion for the IMF-B_z component. First, we considered that geomagnetic conditions are favorable to be disturbed, because high occurrences of magnetic reconnection can be expected in dayside magnetosphere under southward IMF conditions (e.g., Dungey 1961). Second, PL learns by focusing on the highest variance of the parameters [see Eqs. (3) and (4) in "Details of the potential learning (PL)" section] and extracts the focused parameter as the most significant factor driving the magnetospheric disturbances. Therefore, we excluded cases of the IMF-B_z component highly fluctuating between positive and negative around 0 nT.

Before inputting the solar wind conditions to the PL, we classified the K_p values into two groups (targets) of "positive" and "negative" targets. K_p index with values from 6- to 9, and the associated solar wind data were labeled as "positive target (group)". Whereas K_p values ranging from 0 to 1 + and the associated solar wind parameters were labelled as "negative target (group)". The positive (active geomagnetic condition) and negative (quiet geomagnetic condition) targets are categorized, depending on whether or not geomagnetic conditions were highly disturbed or very quiet. This categorization was made according to the geomagnetic activity level, referred to the space weather scales provided from the National Oceanic and Atmospheric Administration (NOAA, https://www.swpc.noaa.gov/noaa-scales-explanation), and that by Ji et al. (2013). The total number of compiled (averaged) data points was 27,493 with the positive (negative) target number being 803 (26,690). To equalize the number of data between the positive and negative targets, we randomly chose and extracted 803 points out of the 26,690 negative target data points. Finally, we analyzed 1606 positive and negative data points.

Methodology of database analysis

By adopting a new NN (PL) to take the 3 h average solar wind parameters as "input parameters" and classify whether or not the associated K_p index belongs to "positive" or "negative" targets, we investigated the relationship between geomagnetic activity levels and solar wind parameters.

Recently, NNs have frequently been used to build forecasting models of geomagnetic indices. However, it is difficult to interpret which solar wind parameters are the most important in disturbing the magnetosphere. In this study, we demonstrate the PL method in the field of space science (space weather) research. The PL algorithm has been previously utilized to resolve the problem on the consumer behavior by Kitajima et al. (2016a) and developed based on two NNs; selective potentiality maximization, proposed by Kamimura and Kitajima (2015), and self-organizing selective potentiality learning (Kamimura, 2015).

We trained the PL by setting the number of neurons (see the details on the manner to train PL are described in "Details of the potential learning (PL)" section), as listed in Table 2. The number of hidden neurons was automatically determined by the software of "SOM Toolbox v2.1" which was developed by Vatanen et al. (2015).

In the knowledge utilization step, hyperbolic tangent and softmax functions were used for the activation functions of the hidden and output neurons, respectively. We searched for the most suitable "r" value by varying the value of the parameter (r) from 1 to 10 with a step of 1. A total of 1,126 (70%) of the 1,606 samples were used for training. Half of the remaining 240 samples (15%) were utilized to prevent training from overfitting (early

Table 2 List of numbers of neurons in potential learning (PL)

Setup of PL	
# of input neurons	5
# of output neurons at Knowledge Accumulation step	198
# of hidden neurons at Knowledge Utilization step	198
# of output neurons at Knowledge Utilization step	2

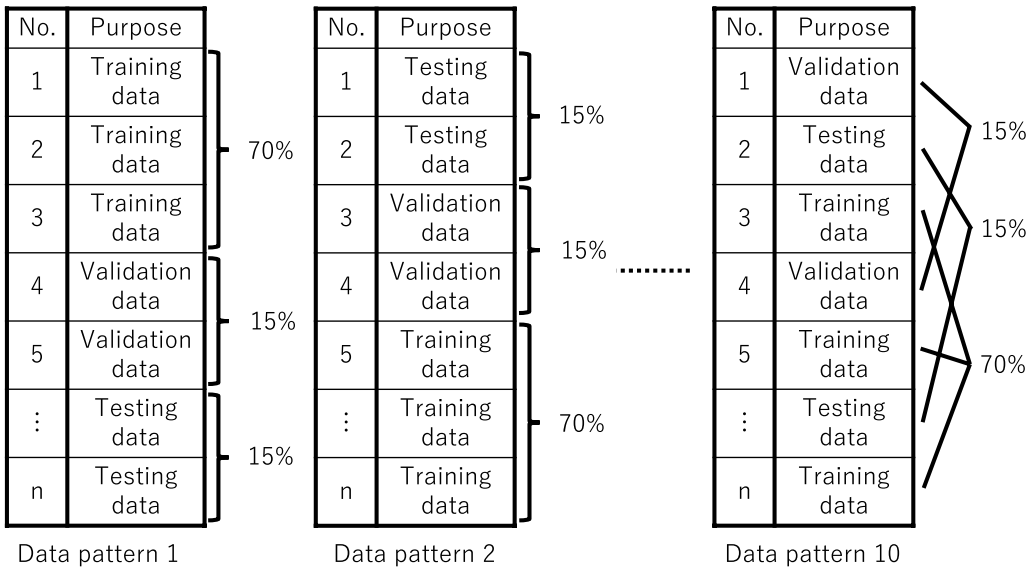


Fig. 1 Block diagrams of details of 10 potential learning (PL) models. In each model, the data for training (Training data), the data to prevent training from overfitting (early stopping) (Validation data), and the data for testing the model (Testing data) are included. The percentages for the three kinds of data are 70%, 15%, and 15%, respectively

stopping) and the other half (15%) were used for testing. We maintained these allocation rates during these PL runs, and made 10 different models in a random choice manner as shown in Fig. 1. Furthermore, we evaluated the performance of each model by calculating the average values of the 10 models.

Details of the potential learning (PL)

PL consists of two steps: knowledge accumulation, based on self-organizing maps (SOM), the concept of which is shown in Fig. 2a; and knowledge utilization, originating from multi-layer perceptron (MLP), the details of which are shown in Fig. 2b. During knowledge accumulation, the potentiality of the input neuron is calculated, and knowledge is acquired (training). Here, we define "potentiality" as ability to respond to various conditions of neuron. In case of "neuron with high potentiality", it indicates the neuron which can play an important role in training. In general, NNs are referred to as "black box," but, in the PL, we can interpret which input parameters are important by interpreting the potentiality after training. If assigning the number $k(k = 1, 2, \ldots, K)$ to the input neuron, we can derive the potentiality of the kth input neuron (Φ_k^r) between 0 and 1, using the following equation:

$$\Phi_k^r = \left(\frac{V_k}{\max\limits_{k=1,\ldots,K} V_k} \right)^r . \tag{3}$$

where V_k is the variance of the kth input neuron, which is computed based on "weight" ($w_{j,k}$) connected to the kth input neuron from jth $\left(j = 1, 2, \ldots, J\right)$ output neuron and r is the parameter that controls the potentiality calculated using the algorithm. The larger the "r" value becomes, the input neuron with larger variance can have larger potentiality.

After the potentiality was calculated, PL was trained based on self-organizing maps (SOM), in which the

Table 3 Summary of accuracy, precision, recall and F-measure values

	PL										MLP
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$	$r=10$	
Accuracy	0.9858	0.9858	0.9858	0.9858	**0.9875**	0.9863	0.9863	0.9858	0.9850	0.9850	**0.9904**
Precision	**0.9875**	0.9859	0.9867	**0.9875**	0.9867	0.9859	0.9867	**0.9875**	0.9850	0.9842	**0.9925**
Recall	0.9842	0.9858	0.9850	0.9842	**0.9883**	0.9867	0.9858	0.9842	0.9850	0.9858	**0.9883**
F-measure	0.9858	0.9858	0.9858	0.9858	**0.9875**	0.9862	0.9862	0.9858	0.9850	0.9850	**0.9904**

Bold letters indicate their maxima

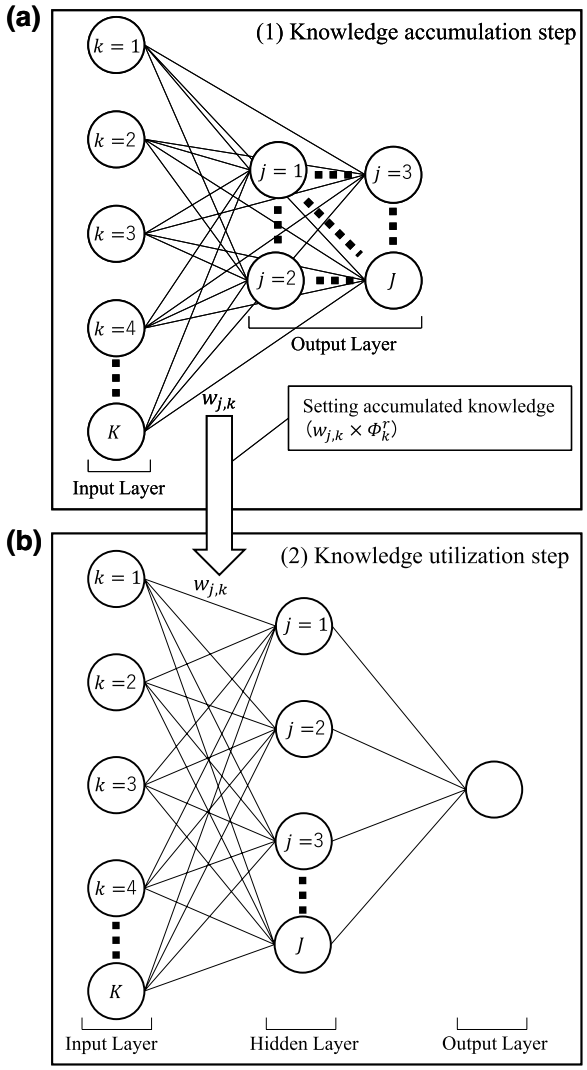


Fig. 2 The concept of potential learning (PL). PL comprises two important steps: (**a**) knowledge accumulation and (**b**) knowledge utilization

potentiality was used to calculate the distance (d_j) between the input neuron (the input from the k^{th} input neuron is denoted by x_k) and the jth output neuron with the following formula:

$$\boldsymbol{d_j} = \sqrt{\sum_{\boldsymbol{k}=1}^{\boldsymbol{K}} \varnothing_{\boldsymbol{k}}^{\boldsymbol{r}} \left(\boldsymbol{x_k} - \boldsymbol{w_{j,k}}\right)^2}. \quad (4)$$

Equation (4) means that the "distance," weighted by the potentiality of the input neuron, was used in the training process. The training logic was the same as those of the SOM. Through the knowledge accumulation step, PL starts to conduct the training at the step of knowledge utilization, based on MLP. In this step, the weight obtained in the knowledge accumulation step was multiplied by the potentiality and set as the initial weight between the input and hidden layers for learning. In general, the results of the training based on MLP depend on the initial weights. However, PL is expected to provide more precise training based on the knowledge obtained from the input parameters (data).

Results

Evaluation of model performances

To evaluate first the K_p classification model based on PL, we calculated the values of four measures (accuracy, precision, recall and F-measure) with changing value of parameter "r" from 1 to 10. Table 3 shows the calculation results of the four measures, indicating the extent to which the model successfully predicted the test data. When "r" was 5, the value of "accuracy (0.9875)" was the highest. The main purpose of creating the 10 models was to extract of the variables that play essential roles in classifying the K_p index into two targets: negative and positive. Because we found that most of measures (accuracy, recall and F-measure) as shown in Table 3 were the highest with $r=5$, we applied the best model with $r=5$ in this study.

We also compared the PL results with another algorithm called multi-layer perceptron (MLP), a basic NN. As shown in Table 3, all four measure values in the various PL iterations ($r=1$ to 10) were close to those of MLP. In particular, the difference in accuracy between PL and MLP was only 0.0029. MLP can be better for classifying K_p into two targets than PL, if the main purpose is only the prediction of geomagnetic activity. However, PL actively selects the input values to be utilized for classification, while MLP does not.

When $r=5$, the values of the three measures (accuracy, recall, and F-measure) of evaluating model performance in PL reached their maxima but were slightly lower than those in, or the same as MLP. PL, however, has a strong advantage in extracting the most influential solar wind parameters that cause geomagnetic disturbances. Therefore, we applied the PL model with $r=5$. The values of all four measures (accuracy, precision, recall, and F-measure), their averages and numbers of true positives, false positives, true negatives, and false negatives in 10 patterns (models) at $r=5$, and all equations to derive the four measures are summarized in the supplementary information.

Extraction of significant solar wind parameters that cause magnetospheric disturbances

Figure 3 shows the result of PL for the input neurons at $r=5$. PL extracted the solar wind velocity (V_x) as the parameter with the highest "input potentiality" (~1.0), suggesting that PL at $r=5$ judged solar wind velocity to

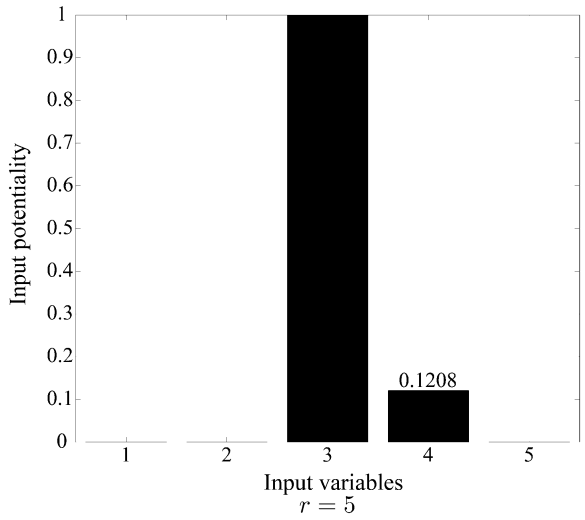


Fig. 3 Results of the application of PL at $r = 5$. The five OMNI solar wind parameters (IMF-B_x, IMF-B_y, V_x, N_p, and B_s) are chosen as the input data to PL. The horizontal and vertical axes give input potentiality and the numbers of input five solar wind parameters, respectively. The potentialities of the solar wind velocity and density are 1.0 and 0.1208, respectively

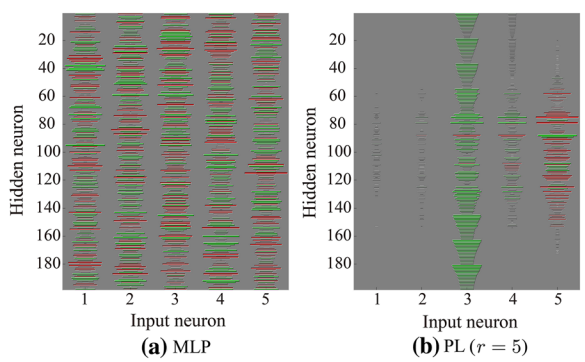


Fig. 4 Weights in the input–hidden layers in the networks of MLP (**a**) and PL (**b**). Horizontal and vertical axes give the number of five input variables (neurons) and number of neurons in hidden layer, respectively. The length of the bar shows the weight in PL, and the signs of the weight values are indicated with red (plus) and green (minus), respectively

be the most significant parameter causing geomagnetic disturbances under the B_s (southward IMF) condition. The parameter with the next highest potentiality was the solar wind density (N_p) at 0.1208; however, this can almost be ignored when compared with the potentiality of the solar wind velocity.

Figure 4 shows the weights of the input and hidden layers used in the PL and MLP for comparison. The length of the bar shows the magnitude of the weight, and the signs of the weight values are indicated with red (plus) and green (minus), respectively. The panel (a) in Fig. 4 shows that MLP has various plus and minus values for weights at each input neuron (parameter), indicating that it is difficult to identify which input neuron (parameter) was used in the network. However, in the PL network with $r = 5$ (panel b), most of the weight was concentrated on the third variable (solar wind velocity); however, the fourth (ion number density) and fifth (southward IMF) variables also had some weight and were thus also used in the PL network. PL uses potentiality to set up the initial weight in the knowledge utilization step (see Fig. 2b). Although three variables (third, fourth, and fifth) had high weight values, we judged the parameter with the highest weight value, the third variable (solar wind velocity), as having the most significant potentiality among them.

Summary and discussion

We reported the results of benchmarks of the application of a new neural network (Potential Learning) for the prediction of geomagnetic activity, driven by solar wind, and the successful extraction of the most significant solar wind parameter in causing geomagnetic field disturbances. This study is the first attempt for applying the PL to the numerical data analyses in space plasma. We also used 22 years of OMNI solar wind data and K_p indices as input neurons but only used the data when the IMF B_z was southward. This was because geomagnetic activity is favorable to be disturbed by dayside magnetic reconnection under southward IMF-B_z conditions (e.g., Dungey 1961), and it is thus easier to extract the crucial solar wind parameter(s) that drive the geomagnetic disturbances.

We excluded the solar wind data under northward IMF conditions due to an inherent disadvantage of the current PL algorithm; PL identifies the largest variance value with the highest potentiality. Therefore, if data under northward IMF conditions were included in the database, the stable (non-excursive) but intensive southward IMF-B_z component would not be chosen by PL as the solar wind parameter with the highest potentiality. Furthermore, the fluctuating IMF-B_z around 0 nT may be chosen as the most significant parameter that cause magnetospheric disturbance. To avoid these cases, we utilized only solar wind data during the southward IMF intervals as input neuron. In future studies, we will try to change the method of calculating the potentiality and improve the associated algorithm of PL so that some other input parameters, except the variance value, can be identified as the highest potentiality.

Based on a large solar wind database, PL extracted the solar wind velocity as the parameter with the highest potentiality when $r=5$ (see Fig. 3), suggesting that solar wind speed (V_x) is an important parameter in disturbing geomagnetic conditions. However, it is required to note that this result can be obtained under the conditions that our PL analysis was conducted 1) only with the solar wind parameter data under southward IMF conditions ($B_z \leq 0$), and 2) based on classification of the two geomagnetic activity levels; "highly" disturbed ($K_p=6$–9) or "very" quiet ($K_p=0$ to $1+$) geomagnetic conditions.

Significant enhancements of the geomagnetic activity level due to increases in the solar wind speed (V_x) have been suggested by previous research (e.g., Snyder et al. 1963; Vasyliunus et al. 1982; Borovsky et al. 1998; Gholipour et al. 2004; Newell et al. 2007, 2008; Elliott et al. 2013, and references therein). Taking a look at the two empirical equations; Eqs. (1) and (2) proposed by Newell et al. (2008), the term of solar wind convection electric field, which can be calculated using the solar wind velocity (V_{sw}) and IMF (B_{IMF}), was involved. Furthermore, according to these equations, K_p can be described by the two velocity terms with square and 3/4 powers of "$V_{sw}{}^2$" and "$V_{sw}{}^{3/4}$", indicating that solar wind velocity is the most important parameter in controlling geomagnetic activity (K_p). Therefore, the most significant parameter extracted by PL (solar wind velocity) was determined to be the most significant parameter that causes disturbances to the Earth's magnetosphere, being the same tendency as the previous statistical results using large solar wind parameter and geomagnetic index databases (Gholipour et al. 2004; Newell et al. 2008; Elliott et al. 2013, and references therein).

Comparing the MLP results with those of PL, the accuracies were not significantly different. However, unlike MLP, PL was able to extract the most significant solar wind parameter to cause geomagnetic disturbances. This benchmark for the application of PL to extraction of the specific parameters verified its effectiveness in predicting the solar wind driving geomagnetic activity and significant solar wind parameters that cause geomagnetic disturbances.

In this study, we showed that PL can extract the most significant solar wind parameter which causes geomagnetic disturbances under certain conditions. Therefore, we can aim to construct a more correct and parameter-dependent forecasting model for geomagnetic activity based on PL.

Abbreviations

PL: Potential learning; IMF: Interplanetary Magnetic Field; RMSE: Root-Mean-Square Error; NN: Artificial neural network; MLP: Multi-layer perceptron; GSM coordinates: Geocentric Solar Magnetospheric coordinates; SOM: Self-organizing maps.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-022-01697-0.

Additional file 1: Table S1. Values of all four measures (accuracy, precision, recall, and F-measure), their averages and numbers of true positives, false positives, true negatives, and false negatives in 10 patterns (models) at r = 5 and all equations to the four measures are summarized.

Acknowledgements

We would like to thank Editage (www.editage.com) for English language editing.

Author contributions

MN conceived the research project. RK performed all data analyses, made all the figures, and tuned the PL codes. MN and RK wrote the paper and edited the manuscript. RK developed the main engine of the PL program and edited the draft. RK and MN equally contributed to this work and paper. All authors critically reviewed and revised the manuscript draft and approved the final version for submission.

Funding

M.N. was supported by a grant from the National Natural Science Foundation of China (NSFC 42074194).

Availability of data and materials

Solar wind OMNI data were obtained from the Coordinated Data Analysis Web (https://cdaweb.sci.gsfc.nasa.gov/index.html/), provided by GSFC/NASA. K_p index data were provided by the World Data Center for Geomagnetism, Kyoto (http://swdcdb.kugi.kyoto-u.ac.jp/).

Declarations

Competing interests

The authors declare that they have no competing interest.

Author details

[1]Department of Engineering, Tokyo Polytechnic University, 1583 Iiyama, Atsugi, Kanagawa 243-0297, Japan. [2]Shandong Provincial Key Laboratory of Optical Astronomy and Solar-Terrestrial Environment, Institute of Space Sciences, Shandong University, 180 Wen-Hua West Road, Weihai, Shandong 264209, People's Republic of China. [3]IT Education Center, Tokai University, 4-1-1 Kitakaname, Hiratsuka, Kanagawa 259-1292, Japan.

Received: 3 April 2022 Accepted: 28 August 2022
Published online: 28 September 2022

References

Bala R, Reiff P (2012) Improvements in short-term forecasting of geomagnetic activity. Space Weather 10:S06001. https://doi.org/10.1029/2012SW000779

Bartels J (1939) Potsdamer Erdmagnetische Kennziffern, 4. Mitteilung Zeitschrift Für Geophysik 15:214–221. https://doi.org/10.23689/fidgeo-3179

Bartels J (1949) The standardized index, Ks, and the planetary index, Kp. IATME Bull 12b: 97–120.

Black DI (1967) Cosmic ray effects and faunal extinctions at geomagnetic field reversals. Earth Planet Sci Lett 3:225–236. https://doi.org/10.1016/0012-821X(67)90042-8

Boberg F, Wintoft P, Lundstedt H (2000) Real time Kp predictions from solar wind data using neural networks. Phys Chem Earth Part C Solar Terr Planet Sci 25(4):275–280. https://doi.org/10.1016/S1464-1917(00)00016-7

Borovsky JE, Thomsen MF, Elphic RC (1998) The driving of the plasma sheet by the solar wind. J Geophys Res Space Phys 103(A8):17617–17639. https://doi.org/10.1029/97JA02986

Costello KA (1998) Moving the Rice MSFM into a real-time forecast mode using solar wind driven forecast modules, Doctoral dissertation, Rice University.
Dungey JW (1961) Interplanetary magnetic field and the auroral zones. Phys Rev Lett 6(2):47–48. https://doi.org/10.1103/PhysRevLett.6.47
Elliott HA, Jahn JM, McComas DJ (2013) The Kp index and solar wind speed relationship: insights for improving space weather forecasts. Space Weather 11:339–349. https://doi.org/10.1002/swe.20053
Farris MH, Russell CT (1994) Determining the standoff distance of the bow shock: Mach number dependence and use of models. J Geophys Res Space Phys 99(A9):17681–17689. https://doi.org/10.1029/94JA01020
Gholipour A, Lucas C, Araabi BN (2004) Black box modeling of magnetospheric dynamics to forecast geomagnetic activity. Space Weather 2:S07001. https://doi.org/10.1029/2003SW000039
Glassmeier KH, Vogt J (2010) Magnetic polarity transitions and biospheric effects. Space Sci Rev 155(1–4):387–410. https://doi.org/10.1007/s11214-010-9659-6
Glassmeier KH, Richter O, Vogt J, Möbus P, Schwalb A (2009) The Sun, geomagnetic polarity transitions, and possible biospheric effects: review and illustrating model. Int J Astrobiol 8(3):147–159. https://doi.org/10.1017/S1473550409990073
Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
Ji EY, Moon YJ, Park J, Lee JY, Lee DH (2013) Comparison of neural network and support vector machine methods for Kp forecasting. J Geophys Res Space Phys 118:5109–5117. https://doi.org/10.1002/jgra.50500
Kamimura R, Kitajima R (2015) Selective potentiality maximization for input neuron selection in self-organizing maps. In: Proceedings of 2015 international joint conference on neural networks (IJCNN), pp 1–8. doi:https://doi.org/10.1109/ijcnn.2015.7280541.
Kamimura R (2015) Self-organizing selective potentiality learning to detect important input neurons. In: Proceedings of 2015 IEEE international conference on systems, man and cybernetics (SMC), pp 1619–1626. doi:https://doi.org/10.1109/SMC.2015.286.
Kitajima R, Endou K, Kamimura R (2016a) Creating a model for detecting non-continuous customers in retail stores by focusing on the potentiality of input neurons. Commun Operat Res Soc Japan 61(2):88–96 (**In Japanese**)
Kitajima R, Sakai H, Kamimura R (2019) Analysis of relationships between top messages and profitability by potential learning. J Japan Soc Fuzzy Theory Intell Inform 31(2):636–644. https://doi.org/10.3156/jsoft.31.2_636(InJapanese)
Kitajima R, Kamimura R, Uchida O, Toriumi F (2016b) Identifying important tweets by considering the potentiality of neurons. IEICE Trans Fundam Electron Commun Comput Sci E99-A:8:1555–1559. https://doi.org/10.1587/transfun.e99.a.1555
Nagai A (1994) Prediction of magnetospheric parameters using artificial neural networks, Doctoral dissertation, Rice University.
Newell PT, Sotirelis T, Liou K, Meng CI, Rich FJ (2007) A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. J Geophys Res Space Phys. https://doi.org/10.1029/2006JA012015
Newell PT, Sotirelis T, Liou K, Rich FJ (2008) Pairs of solar wind-magnetosphere coupling functions: combining a merging term with a viscous term works best. J Geophys Res Space Phys 113:A04218. https://doi.org/10.1029/2007JA012825
Rangarajan GK (1987) Indices of geomagnetic activity. In: Jacobs JA (ed) Geomagnetism, vol 3. Academic Press, pp 323–384
Shprits YY, Vasile R, Zhelavskaya IS (2019) Nowcasting and predicting the Kp index using historical values and real-time observations. Space Weather 17:1219–1229. https://doi.org/10.1029/2018SW002141
Snyder CW, Neugebauer M, Rao UR (1963) The solar wind velocity and its correlation with cosmic-ray variations and with solar and geomagnetic activity. J Geophys Res 68:6361. https://doi.org/10.1029/JZ068i024p06361
Tan Y, Hu Q, Wang Z, Zhong Q (2018) Geomagnetic index Kp forecasting with LSTM. Space Weather. https://doi.org/10.1002/2017SW001764
Vasyliunas VM, Kan JR, Siscoe GL, Akasofu S-I (1982) Scaling relations governing magnetospheric energy transfer. Planet Space Sci 30:359. https://doi.org/10.1016/0032-0633(82)90041-1
Vatanen T, Osmala M, Raiko T, Lagus K, Sysi-Aho M, Orešič M, Honkela T, Lähdesmäki H (2015) Self-organization and missing values in SOM and GTM. Neurocomputing 147(5):60–70. https://doi.org/10.1016/j.neucom.2014.02.061
Wing S, Johnson JR, Jen J, Meng CI, Sibeck DG, Bechtold K, Takahashi K (2005) Kp forecast models. J Geophys Res. https://doi.org/10.1029/2004JA010500
Wintoft P, Wik M, Matzka J, Shprits Y (2017) Forecasting Kp from solar wind data: input parameter study using 3-hour averages and 3-hour range values. J Space Weather Space Clim 7:A29. https://doi.org/10.1051/swsc/2017027
Zhelavskaya IS, Vasile R, Shprits YY, Stolle C, Matzka J (2019) Systematic analysis of machine learning and feature selection techniques for prediction of the Kp index. Space Weather 17:1461–1486. https://doi.org/10.1029/2019SW002271

Publisher's Note