

3-2 Prediction of the Dst Index from Solar Wind Parameters by a Neural Network Method

WATANABE Shigeaki, SAGAWA Eiichi, OHTAKA Kazuhiro,
and SHIMAZU Hironori

Using the Elman-type neural network technique, operational models are constructed that predict the Dst index two hours in advance. The input data consist of real-time solar wind velocity, density and magnetic field data obtained by the Advanced Composition Explorer (ACE) spacecraft since May 1998 (<http://www2.crl.go.jp/uk/uk223/service/nnw/index.html>). During the period from February to October 1998, eleven storms occurred with minimum Dst values below -80nT . For ten of these storms the differences between the predicted minimum Dst and the minimum Dst calculated from ground-based magnetometer data were less than 23%. For the remaining one storm (beginning on 19 October 1998) the difference was 48%. The discrepancy is likely to stem from a imperfect correlation between the solar wind parameters near ACE and those near the earth. While the IMF Bz remains to be the most important parameter, other parameters do have their effects. For instance, Dst appears to be enhanced when the azimuthal direction of IMF is toward the sun. A trapezoid-shaped increase in the solar wind density enhances the main phase Dst by almost 10% compared with the case of no density increase. Velocity effects appear to be stronger than the density effects. Our operational models have, in principle, no limitations in applicability with respect to storm intensity.

Keywords

Dst prediction, Space weather, Neural network, Ring current, Solar wind

1 Introduction

Recent advances in space science have resulted in the use of an increasing number of satellites as part of the infrastructure supporting activities in space, as well as in communications, broadcasting, and meteorological applications. Furthermore, it has now become evident that disturbances of the geomagnetic field have adverse effects on many elements of this infrastructure, both on the ground and in the middle to high latitudes. Geomagnetic storms can induce abnormal charge build-up and discharge, interfering with satellite functions. The effects of such disturbances on satellites, such as abnormal changes to satel-

lite attitude caused by an increase of friction due to atmospheric expansion, may even cause a satellite to fall from orbit. Ground-based electric transmission lines may also be affected by abnormal increases in induced currents. To avoid these adverse effects and to ensure reliable operation of satellites and transmission lines, it is important to make precise predictions of the geomagnetic field disturbance index, even if such predictions are short-term.

The advantage of time-series analysis lies in its ability to enable prediction even when the relevant physical processes in the solar wind plasma, the corresponding magnetic field, and the Earth's magnetosphere remain unresolved. If, for instance, a linear correla-

tion can be found, predictions will prove extremely simple, and a response should be determined using only the uniquely defined impulse response function and the convolution of the input. Iemori et al. (1979) analyzed the impulse response function assuming a linear response, and found that there was a time lag of approximately 40 minutes in the response time between the onset of changes in the solar wind parameter on the sunward side of the Earth and the time of peak disturbance of the substorm (AL, AE). Blanchard and McPherron (1992) were able to express accurately the AL index using two impulse response functions, and explained that the causal factor existed at the noon side of the Earth and in the magnetotail. However, the magnitude of and time interval between the two impulses varied widely for different storms, and the resultant findings could not be used in predictions. Numerous other attempts were made to create a linear prediction model using indices of geomagnetic field disturbances, such as Kp, AE, and AL, but none succeeded in producing a viable model.

The *Dst* and ring current *Dst** during disturbances were expressed in the form of an ordinary differential equation by Burton et al. (1975), and their expression has frequently been used as a classical means of short-term prediction. Klimas et al. (1997, 1998) developed a non-linear prediction method by introducing a filter coefficient. Ebihara and Ejiri (1998, 2000) developed a ring current prediction method based on particle tracking of the plasma sheet and simulations. Various neural network methods (hereafter referred to as "NN methods") have been introduced as methods of prediction in space science, mainly by Lundstedt and Wintoft (1994) and others. Kugblenu et al. (1999) carefully selected 20 geomagnetic storms to produce a few target time series, successfully developing models that exceeded expectations. Wu and Lundstedt (1997a, b) developed an Elman-type NN method (1990) with feedback connection to the input layer from the hidden layer. Mainly by using storm data as the learning time series, they succeeded

ed in improving the coefficient of correlation of *Dst* prediction to 0.9.

Operational models must function properly under any conditions (e.g., in cases of missing data, storms, or quiet periods). **2** presents a model developed by the authors and discusses the performance of the model in 1998, while **3** introduces the process of selecting an operational model. **4** discusses the performance of the model during the period from Feb.-Oct. 1998, and presents some interesting features observed during this period. The solar wind density, velocity, and the *Dst* response of the magnetic field are analyzed in **5**. In **6**, our model is compared to the Burton model and modified models. **7** gives our conclusions.

2 Development of the Elman-type Feedback NN Method

An operational model using the NN technique was constructed based on the procedure established by Wu and Lundstedt (1997a), with novel features in the learning time series and the method of in-process changes to iteration. Fig.1 shows the architecture of our model. The six input components ($R = 6$

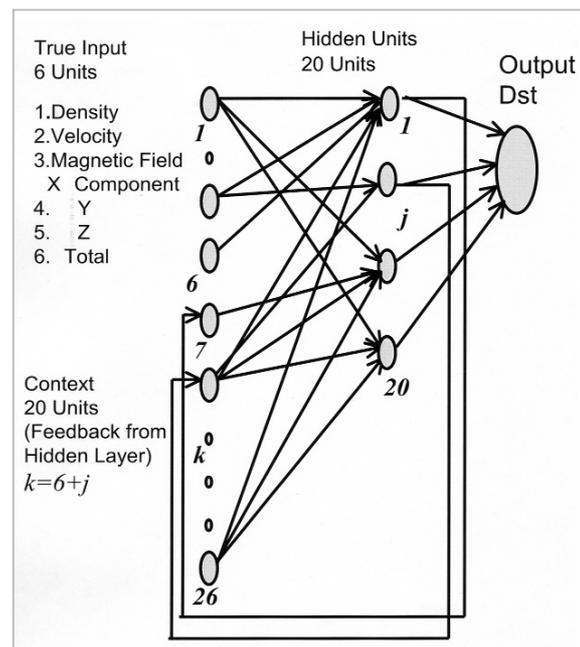


Fig. 1 A schematic diagram of the Elman-type neural network

units) are connected to the hidden layer ($S = 20$ units) through the connection strengths of w_{kj} , and the output of the hidden layer is connected to the Dst through the connection strengths of W_j . Furthermore, after a delay time of one hour, the hidden layer is fed back to units 7 to 26 of the input layer, under the guise of normal input signals, and is reconnected to the hidden layer. These processes can be expressed by equations (1) and (2) below. The components of the normal input I and V , which are returned from the hidden layer, are each connected to the hidden layer by w , and the hidden layer output V having the hyperbolic transfer function of $Tanh$, is connected directly to output Dst . The output is also fed back to the input layer (the first step of the time series) after a delay time of one hour.

$$V_j^t = \tanh\left[\sum_{k=1}^R w_{kj} I_k^t + \sum_{K=R+1}^{R+S} w_{kj} V_{k-R}^{t-1}\right] \quad (1)$$

$$Dst = \sum_{J=1}^S W_j V_j^t \quad (2)$$

The coefficient t on the upper right of I and V in Eqs. (1) and (2) represents time, expressed in units of one hour in the time series. Therefore, the construction of an NN prediction model is equivalent to determining the connection strengths w_{kj} and W_j between each unit. The important point, or our final goal, will then be to find connection strengths that best reflect the natural phenomena (Dst). The current model (Model C) has been used since Dec. 1998, and it has a hidden layer of 20 units, an input layer of six units (solar wind velocity, density, three components of the interplanetary magnetic field (IMF), and their absolute value).

Therefore, a total of 540 connection strengths (520 (20~26) for w_{kj} and 20 for W_j) have been calculated, and the connection strength set giving the best results are used in the model. In previous models, the input components consisted of only three parameters—solar wind velocity, density, and the southward component of the IMF—and so the total

number of connection strengths was 480 (460 (20~23) for w_{kj} and 20 for W_j). This was operated from Apr.–Nov. 1998. The coordinate system used in this paper is the GSM system. The X-axis is in the sunward direction. The Z-axis lies in the plane defined by the X-axis and the northward axis of the dipole magnetic field, and is perpendicular to the X-axis and points toward the north. The Y-axis is set so that the X, Y, and Z-axes form a Cartesian coordinate structure.

To determine the connection strength, Eqs. (1) and (2) are initially calculated using small values. The results are compared to existing Dst index values calculated from ground-based values at the C2 Center of Kyoto University (hereafter referred to simply as "ground-based Dst values"), with nearly 100,000 iterations to minimize error. It can easily be seen that this trial-and-error technique of determining the connection strength set is time-consuming. However, once a set has been determined, the prediction calculation can be completed in less than one second using workstations that have been available since 1998.

2.1 Creation of a Target Time Series and Solar Wind Learning Data

Target time series: a Dst time series determined from past ground-based geomagnetic field observations. An expected time series.

Learning time series: a time series of solar wind parameters corresponding to past ground-based Dst values. By using this time series to predict past Dst , the connection strengths, or the model, is determined through automatic iteration.

Numerous models can be constructed by varying the length and components of the time series and the geomagnetic storm period ratio in the target time series. The various models were evaluated, and two were selected as final candidates for the six-component input model.

All data for learning were taken from the solar wind OMNI data of NASA's National Space Science Data Center (NSSDC) for the period from 1978-1982, which offers relative-

ly continuous solar wind data. The target time series for prediction was created based on the final ground-based *Dst* data. The learning data used in creating the prediction model consisted of observed values near Earth or values that have been converted into such observed values. Solar wind observed by an ACE satellite positioned at the Lagrangian point (L1, located approximately 1.5 million km away in the sunward direction) arrives at the Earth approximately one hour later, although this delay time depends on solar wind velocity.

The parameters observed by the ACE satellite are corrected for variations in elapsed time and used as input data for the prediction models. In other words, the time required for prediction under the operational model is the sum of the time it takes for the solar wind to travel from the L1 point to the Earth and the NN prediction time (one hour at present). The prediction time of the current operational model is 2 hours.

Model A has been constructed using a 5,077-hour learning time series consisting of solar wind density, velocity, and the southward component of the IMF for over 50 disturbance events (including at least one sub-storm) that occurred between 1978-1982. This model was first used when the Real Time Solar Wind (RTSW) data for space weather became available, with the launch of the ACE satellite. Although model A is an initial operational model, its performance is good. The coefficient of correlation between the *Dst* actually observed and the *Dst* (A) predicted by model A based on scientific data from the ACE satellite was approximately 0.89. The error in prediction results mainly from imperfections in the model and from the poor correlation between solar wind plasma near the ACE satellite and near Earth. Given that we have been aware of this poor correlation through previous reports (Paularena et al., 1998; Richardson et al., 1998), the performance of model A is considered to be fairly good.

The Elman-type prediction model used in

this study will be introduced below. The *Dst* time series given by Eq. (2) is continually compared to ground-based values in each step of the iteration and during processing through the use of absolute differences and coefficients of correlation. A successful model is assumed to have been achieved when the difference reaches a minimum, when the model is saturated at the limit, and when various indices (such as the coefficient of correlation) indicate the appropriateness of the model. Therefore, a model with good results can ultimately be obtained, even when the method involved is somewhat inappropriate, since the iteration proceeds by comparison with ground-based values.

This paper will present a method of *Dst* prediction using the Elman-type NN technique.

Some caution must be taken when creating a target time series and a learning time series by connecting many time series sections. High-priority conditions for the section of the time series near the connections are:

- i) The absolute value of the *Dst* must be small at the end of one section and at the beginning of the next section, regardless of whether the portion corresponds to a quiet period or a disturbance period. The worst connection would result from placing a part of the storm recovery phase at the end of a section, where B_z is already positive and *Dst* has a large negative value. When calculations are initiated in the recovery phase, the storm will fail to be recognized, and the whole section will be regarded as a quiet period.
- ii) The difference in B_z values should be small enough to allow a smooth connection. (However, B_z must be positive and have a value near 0.)
- iii) The difference in velocity should be small enough to allow a smooth connection.

A long time series should be created by connecting each section, bearing the above conditions in mind. Of course, under ideal conditions, there would also be only small discontinuities in density, but when priorities are given to the above considerations, discontinu-

ities will to some extent be inevitable. However, this is not considered a serious setback, as there are many cases in nature where only density displays a sudden increase. In any case, the above conditions should be given the highest priority.

Empirically, it has been found that the models with learning time series having higher proportions of quiet periods performed more stably during disturbances relative to model B, which was created by learning only disturbances with large absolute *Dst* values. Model C was constructed using learning time series containing 44 % quiet period data.

2.2 Procedures in Iteration

2.2.1 Start of Iteration

The values of w_{kj} and W_j in Eqs. (1) and (2) are assumed to be small. The authors have used small random numbers. V also has a small value. The increment from W (ΔW) to the next step in the iteration is also small or 0. In the following equations, values of α near 1 and η near 0 are recommended. Wu and Lundstedt (1997a) made major contributions regarding the method for getting these values, save for construction of the learning time series.

2.2.2 During the Iteration Process

The absolute mean of the difference between the predicted value and the expected value can be expressed as:

$$E = \frac{1}{2} \sum_{t=1}^N (T(t) - Dst(t))^2 \quad (3)$$

Here, T is the expected ground-based *Dst* value, and *Dst* is the predicted value at a certain point in the iteration. The variable t represents time, or the sequential number of the sampling within the time series.

In this way, E is calculated at each step in the iteration to judge the increase in the precision of the model. The coefficient of correlation is also calculated to evaluate performance.

First, the connection strengths W_j between

the hidden layer and the output *Dst* in Eq. (2) are determined. The iteration increment ΔW for time τ can be expressed as:

$$\Delta W_j(\tau) = (1 - \alpha)\eta \sum_{t=1}^N \delta^t V_j^t + \alpha \Delta W_j(\tau - 1), \quad (4)$$

$$\delta^t = T^t - Dst^t, \quad (5)$$

Both V and W are values at a given point in the iteration. Next, the formula for the connection strength w between the input layer and the hidden layer is given. The value of w_{kj} for one hour earlier is used as the rate of change to w at time τ .

$$\Delta w_{jk}(\tau) = (1 - \alpha)\eta \sum_{t=1}^N \delta_k^t I_k^t + \alpha \Delta w_{jk}(\tau - 1), \quad (6)$$

In the $k = 1, \dots, R,$

$$\Delta w_{jk}(\tau) = (1 - \alpha)\eta \sum_{t=1}^N \delta_j^t V_{k-R}^{t-1} + \alpha \Delta w_{jk}(\tau - 1), \quad (7)$$

$$k = R + 1, \dots, R + S$$

above equation, δ is:

$$\delta_j^t = (1 - (V_j^t)^2) W_j \delta^t, \quad j = 1, \dots, S \quad (8)$$

In Eqs. (6) and (7), $0 < \eta < 1$ and $0 < \alpha < 1$. These values determine the speed of convergence and whether a good model can be obtained. They are adjusted automatically in the iteration. The rate of change for η during the calculation process is calculated according to the following equation.

$$\Delta \eta = \begin{cases} +a\eta & \Delta E < 0 \\ -b\eta & \Delta E > 0 \end{cases} \quad (9)$$

Here, a and b are arbitrary positive constants. When ΔE is positive, the iteration diverges and η becomes smaller; appropriate measures are thus taken, such as changing the constants.

3 Dependency on Learning Data

Table 1 shows the model and corresponding learning data used. The data acquisition period for model B is the same as that for model A, but instead of the southward compo-

Table 1 Specifications of the Learning Time Series in Creating Models

Model	Condition of Learning Data	Duration of Learning	Parameters in Learning	Operational Period
Model A	Mainly during storms	5,077 hours	Velocity, density, and IMF Bs	Model #1 From Apr. 1998 to Nov. 1998
Model B	Mainly during storms	5,077 hours	Velocity, density, IMF Bx, By, Bz, Bt	
Model C	During storms and quiet times	9,058 hours	Velocity, density, IMF Bx, By, Bz, Bt	Model #2 From Dec. 1998 to present

ment of IMF, model B uses B_x , B_y , B_z , and B_t (the total IMF) from the GSM coordinate system, for a total of six input parameters. Furthermore, model C was created through learning using time series data (9,058 hours) with a significantly higher quiet-period ratio, representing 44 % of the whole period. The coefficients of correlation between the ground-based and predicted values for both models B and C were 0.9 or higher. Using this defined coefficient of correlation in the operational NN model, a difference of 10 nT for $Dst = -10$ nT will correspond to an error of 100 %. A large part of the learning time series is composed of quiet times featuring such large error components. A model providing a coefficient of correlation of 0.9 over the whole period is considered to be an extremely satisfactory model. Furthermore, the mean square of error (\sqrt{E}) was generally lower than 10 nT.

The model was evaluated based on a predictive test in 1978 and on the results of predictions since 1998. A comparison of typical prediction results for models B and C is introduced for the 140-hour period beginning at 12 UT, Aug. 2, 1978. The top panel in Fig.2 shows the ground-based Dst values (black), values predicted by model B (green), and those predicted by model C (blue). The mid-

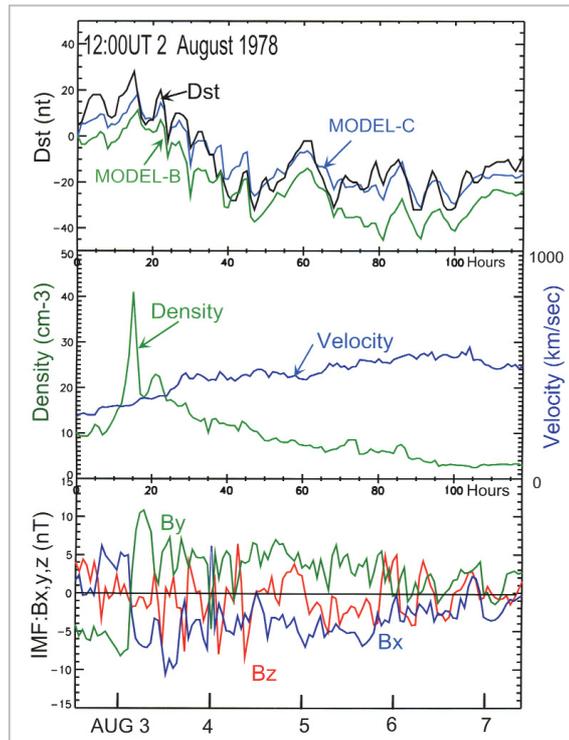


Fig.2 Comparison of the ground-based Dst values (black), values predicted by model B (green), model C (blue) for the 140-hour period starting Aug. 2, 1978

The second panel shows the solar wind density (green) and velocity (blue) used in the Dst prediction based on the neural network. The third panel shows the components of the IMF: B_x (blue), B_y (green), and B_z (red).

dle panel shows the solar wind velocity within the range of 0-1,000 km/s (blue) and particle density for the range of 0-50 (cm^{-3}) for the corresponding period. The solar wind velocity increases from 11 UT Aug. 2 to 20 UT Aug.6, from 274 km/s to 578 km/s, which is higher than normal. Thereafter, it maintains a steady, high-speed plateau, then gradually declines. Density peaks at 41 cm^{-3} at 03UT Aug.3, then rapidly decreases. The bottom panel shows the strength of the IMF in the range from -15 nT to $+15$ nT. It can be seen that B_z (red) neither falls below -10 nT nor displays prolonged periods below -5 nT.

Therefore, there is no significant growth in Dst , and the quiet period continues. The results of prediction under model C are clearly closer to ground-based Dst value, and this is believed to be the effect of the large propor-

tion of quiet periods contained in the learning time series for model C. The prediction results of model C also show stable consistency with ground-based values, even during geomagnetic storms. Model C replaced model A as the operational *Dst* prediction model as of Dec. 1998.

4 Precision of Predictions Under Model C

Model C was developed using solar wind and *Dst* data obtained during the solar maximum of 1978-1982. However, its effectiveness was demonstrated in 1998, when solar activity began to increase from a previous solar minimum. Fig.3 shows the results of predictions under model C (green line) using continuous ground-based *Dst* observation data (black) and ACE satellite observation data for

the period from Feb.-Oct. 1998. However, a period for which data is missing (due to satellite malfunction) is excluded. The solar wind data used consisted of the final calibration values, i.e., level-2 or scientific data, and not the "quick-look" values. These ACE data values are distributed through the Web by NSSDC/NASA and the ACE Science Center. Fig.3 is drawn with the prediction time shifted by 1-2 hours to account for the maximum time required for the solar wind to pass from the ACE satellite to the Earth's magnetosphere. At the left of each panel is the date on which the section begins and the Bartels number of the Sun. During this period (Feb.-Oct. 1998), there were 11 major geomagnetic storm events with minimum *Dst* below -80 nT. Of the 11 events, 10 had prediction errors of less than 23 %. The remaining event had a large error of 48 %. This large error is thought to be due to

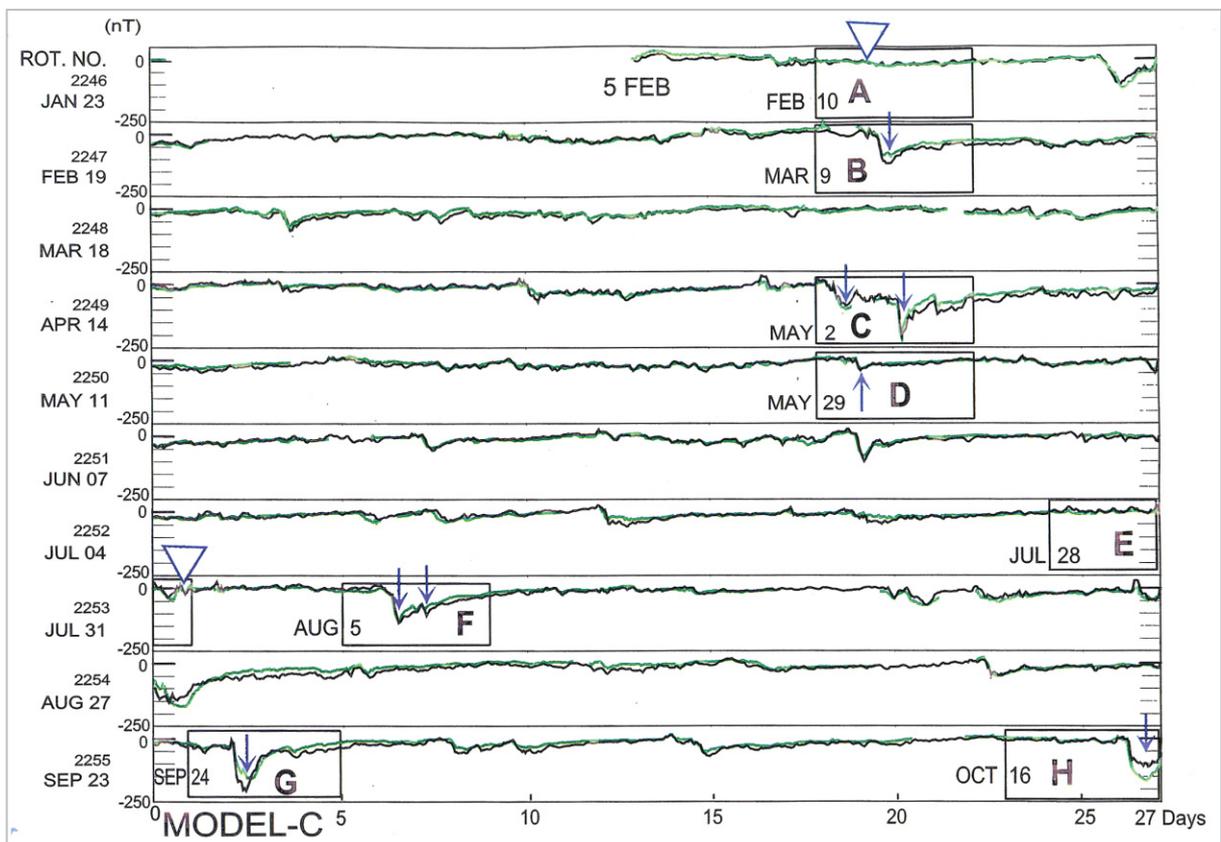


Fig.3 The curve of ground-based *Dst* values (black) and values predicted by model C (green) for the period from Feb. through the end of Oct. 1998

The Bartels number of the sun's rotation and the date of the beginning of observation are shown at the left of each section. The solar wind parameters in boxes labeled from A to H with dates will be explained in detail in Fig. 4.

the poor correlation of solar wind parameters between the Lagrangian point and the near-Earth region. If this one event is disregarded, we can conclude that the model created using data for the 1978-1982 solar maximum is also effective for predicting *Dst* in 1998, during which solar activity began to increase from a previous solar minimum. The rectangular box in Fig.3 highlights several periods of interest (4 days each), which will be investigated in detail in Fig.4.

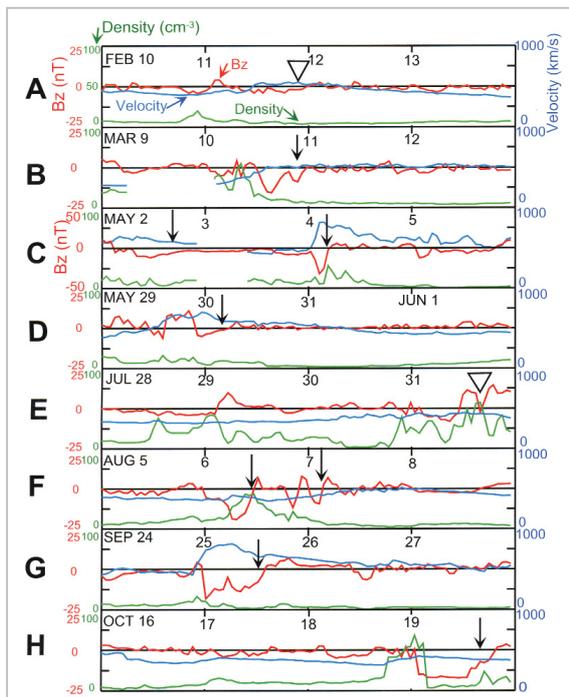


Fig.4 The solar wind parameters of significance in A-H for four days

solar wind velocity (blue), density (green), and *Bz* (red). The date of the beginning of observation is shown in the upper-left hand corner of each panel. The arrows show the minimum *Dst*. The inverted triangle indicates a notable point.

Fig.4 shows the *Bz* component of the IMF (under the GSM coordinate system) in red in each panel. The full scale of the panels are from -25 nT to $+25$ nT, except for panel C, which is set from -50 nT to $+50$ nT for full display of the large *Bz*. The blue lines represent solar wind velocity shown within the range of 0-1,000 km/s. Density is shown in green for a range between 0-100 particles/cm³. The gaps in density and velocity in panels B

and C indicate missing data. The blue arrows in Fig.3 indicate the most intense periods of the geomagnetic storm (end of the main phase), and the periods correspond to those indicated by the black arrows in Fig.4. These periods roughly correspond to the time at which *Bz* switches from negative to positive. Each event will be discussed in the following section.

Panel 4A (Feb. 10-13): Velocity exceeds 500 km/s on Feb. 11, and remains high. However, *Bz* has a value higher than -3 nT, and so, as indicated by the inverted triangle, the ground-based *Dst* has a moderate value of -29 nT. There is almost no prediction error in this case.

Panel 4B (March 9-12): After a period of missing data, velocity begins to increase at 03 UT March 10 from 300 km/s, as shown in panel 4A, and remains above 500 km/s after 15UT on March 10. This solar wind was ejected from the same coronal hole as the solar wind in 4A after a single solar rotation from 4A. The density peak exceeding 40 cm⁻³ at 04 UT and 09 UT on March 10 forms the positive *Dst* peak of the initial phase. The *Bz* component falls to -15 nT at 15 UT on March 10 and remains below -11 nT for three hours. This causes a geomagnetic storm of -116 nT at 20 UT on March 10. The prediction error was 22%.

Panel 4C (May 2-5): This was one of the largest geomagnetic storms in 1998. The velocity has two peaks: 646 km/s at 06 UT on May 2 and 825 km/s at 04 UT on May 4. *Bz* also displays two valleys: -12 nT at 12 UT on May 2 and -32 nT at 03 UT on May 4. The corresponding *Dst* valleys are at 18 UT on May 2 (-85 nT) and at 05 UT on May 4 (-205 nT). The prediction errors were 22 % for the former and 21 % for the latter. These valleys are believed to correspond to magnetic clouds formed at coronal holes.

Panel 4D (May 29-June 1): During this period, an extremely fast solar wind with maximum speed of 705 km/s associated with a solar flare was observed at 00 UT on May 30. However, the fast region shown in the panel

coincides with a positive B_z peak, and only a shallow valley of about -5 nT is present. Therefore, as in period D in Fig.3, only a moderate disturbance in Dst was induced. The prediction error is extremely small. Model C is concluded to be effective for this event.

Panel 4E (July 28-31): This is among the periods of highest density observed in 1998. The peak density was 57.3 cm^{-3} at 16 UT on July 31. There is a positive B_z peak immediately before and after this density peak. The period corresponding to the inverted triangle does not have a large Dst valley, even though the dynamic pressure is large. The prediction error is also small.

Panel 4F (Aug.5-8): The velocity displays a gradual bump between 360-500 km/s. There is a density peak at 43 cm^{-3} at 10UT Aug.6, but otherwise only gradual variations are seen. On the other hand, there are three B_z valleys from Aug.6-7. The main valley reaches -19.2 nT at 08 UT, and there are smaller valleys of -11.7 nT at 21 UT and -9.2 nT at 03 UT on Aug.7. In contrast, Dst shows a major storm of -138 nT at 11 UT on Aug.6, and another storm of -108 nT at 05 UT on Aug.7 during the recovery phase of the former. The two storms are indicated by blue arrows in Fig.3. The prediction error for the main storm is only 6%.

Panel 4G (Sept.24-27): A large velocity peak with maximum velocity of 826 km/s is observed accompanied by a large B_z valley (-18 nT); these phenomena are precursors to a large storm. The minimum Dst of -169 nT was recorded at 09 UT on Sept.25, and the prediction error was 18%.

Panel 4H (Oct.16-19): Velocity remained relatively constant during this period. There was a large density peak of 68 cm^{-3} at 01 UT Oct.19, but otherwise, density also remained relatively constant. However, B_z switched to the negative and fell significantly at the density peak. B_z slowly recovered from -11 nT to a positive value from 02-15 UT on Oct.19. Therefore, a large negative Dst of -166 nT was predicted, but the actual ground-based Dst was only -112 nT, resulting in a large predic-

tion error (48 %). However, the WIND satellite (positioned at $X_{se} = 95$ Re, $Y_{se} = 32$ Re, $Z_{se} = 6$ Re) had also observed a large negative value for B_z of below -15 nT during the period of 02-15 UT. This indicates that a poor correlation between the solar wind parameters at the ACE satellite position and the near-Earth region may have caused the inconsistency, as opposed to imperfections in the model C.

5 Dst Response to Solar Wind Parameters

The NN model has no difficulty reflecting phenomena that have been included in the learning, but it is less capable of extrapolating phenomena from analysis data. This problem is currently the focus of discussion among specialists. In this study, this problem is addressed through attempts to identify the role of each solar wind parameter by inputting data that is artificially generated, but nevertheless highly likely to be produced under natural conditions. If, for two phenomena in nature, only one parameter changes independently and all others respond in the same manner, then the role of the single parameter can be extracted empirically from the observed data. However, the system of responses to the solar wind consists of many basic input parameters such as velocity, density, and the components of the IMF, making it difficult to isolate two such phenomena from the observation data. Furthermore, only the Z component of the IMF has a strong effect on observed phenomena. It is almost impossible to identify the characteristics of the other parameters, since their effects are masked by the effect of the dominant parameter. Therefore, we believe that it will prove significant to study the effects of various parameters by feeding artificial input data into the NN prediction model. In this process, the parameters are defined as being near the Earth and not in the vicinity of the Lagrangian point, so that the effect of distortion by propagation can be avoided. Model C is considered to be appropriate for testing with the NN method, since the coefficient of

correlation between the predicted and observed data for the period from 1978-1982 is higher than 0.91, and excellent prediction results have been obtained for ACE data after 1998. Furthermore, subtle effects of minor parameters were evaluated using NN models produced using different learning time series. The results of these examinations generally displayed the same trend as model C, and so will be included in this paper.

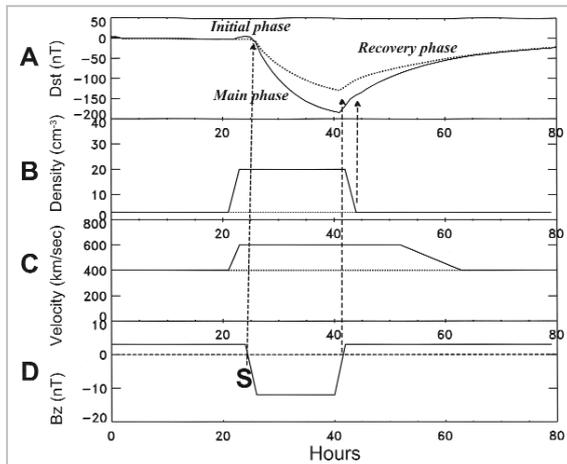


Fig.5 The result of a storm calculated for artificial input parameters in model C.

Panels A, B, C, and D show the calculated (predicted) *Dst*, solar wind density, solar wind velocity, and the B_z component of the IMF, respectively. The dotted lines in panels A, B, and C represent the results for a constant solar wind density and velocity. Solid lines indicate a non-steady state of density and velocity.

To begin, we input simple variations in parameters that are normally observed in actual solar wind behavior. In Fig.5, the X and Y components of the IMF are given as $B_x = 10$ nT and $B_y = 0$ nT. B_z is initially fixed at +3 nT, and at point S (where the B_z component becomes negative, or first turns southward), it is decreased to -12 nT in two hours. This condition lasts for 15 hours, after which B_z returns to +3 nT in two hours. In the beginning, there is no change in density or velocity, set to 3 cm^3 and 400 km/s , respectively. The density and velocity are shown by the dotted lines in panels B and C, respectively. The B_z component is shown by the solid line in panel D. The output calculated for this input is

shown by the dotted line in panel A. This output corresponds to the *Dst* calculated by model C. Since there is no change in the dynamic pressure, the positive peak normally seen in the initial phase is absent. The minimum *Dst* for the main phase reaches -130 nT immediately before the start of the recovery phase, which begins when the B_z switches from negative to positive. In actual geomagnetic storms, the *Dst* also takes a minimum value near the end of the B_z valley, as indicated by arrows in panels B, C, D, F, G, and H in Figs.3 and 4.

In the next step, the behavior of B_z is assumed to be the same as for the above case, shown by the solid line in panel D. In this case, the velocity is assumed to increase from 400 km/s to 600 km/s from S-3 hours in two hours. The velocity then remains at 600 km/s for the next 30 hours, and takes 10 hours to return to the original value of 400 km/s . The density is assumed to take two hours to increase from 3 cm^3 to 20 cm^3 beginning at S-3 hours, to remain at 20 cm^3 for 20 hours, and then to return to 3 cm^3 . The change in density and velocity are shown by solid lines in panels B and C. Since this case involves a change in dynamic pressure, a positive peak of approximately +7 nT is present in the initial phase. The change in velocity results not only in a change in dynamic pressure, but will also accompany the added effect of the $V \times B_z$ electric field. As can be seen from the solid line in panel A, a large valley is formed in the main phase, and the minimum *Dst* (maximum value of disturbance) reaches -185 nT. The recovery phase begins at the point at which B_z switches from negative to positive, indicated by an arrow. Furthermore, when the density plateau ends during the recovery phase, distortion appears in the *Dst* curve.

5.1 The Effect of Density

Fig.6 shows the collective change in *Dst* for five geomagnetic storm events, with the duration of the B_z valley increased from five to a maximum of 25 hours in five-hour increments. B_x and B_y have the same values as in

the case shown in Fig.5. To evaluate the effect of density, velocity is fixed at 400 km/s (dotted line in Fig.5C), and two separate conditions were assumed: in one case without a density plateau and in the other with a plateau. The time series for the two conditions are shown by the dotted and solid lines in Fig.5B. The third storm, shown by a dotted line in Fig.6 and having a 15-hour B_z plateau, coincides with the case indicated by the dotted line in Fig.5A. When there is no change in velocity, the dynamic pressure is expected to be smaller than in the case shown in Fig.5., but the positive peak in the initial phase seen in Fig.6 is slightly higher than for Fig.5, at +10 nT. (However, this difference may fall within the range of error.) The change in dynamic pressure affects the Chapman-Ferraro current, resulting in a positive shift of Dst . The case shown by the solid line in Fig.5A accompanies both density and velocity changes and involves a larger dynamic pressure. However, the B_z has not yet turned negative. Thus, the $V \times B_z$ electric field has the effect of strongly pushing back the plasma from the magnetotail, which would otherwise act as the seed for the ring current. This may be the reason for the higher positive peak in the initial phase seen

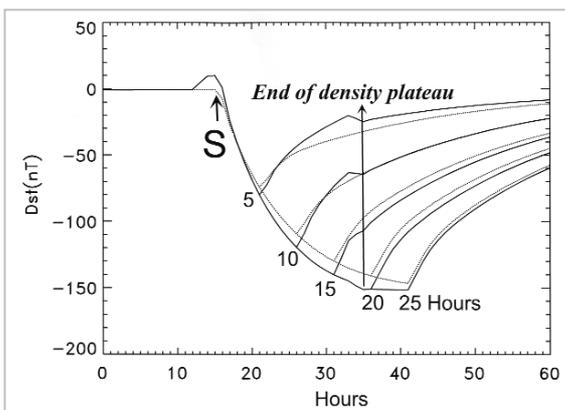


Fig.6 A plot of Dst values calculated for increasing durations in five-hour increments, assuming that the valley for B_z is at $-12nT$

Solid and dotted lines correspond to cases with and without density plateaus, respectively. When the density is constant and the duration of the B_z valley is 15 hours, the results will be the same as shown by the dotted line in Fig. 5.

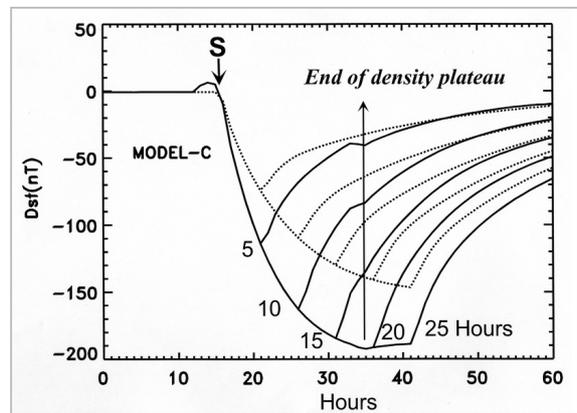


Fig.7 Dst calculated for artificial input

Solid and dotted lines correspond to cases with and without density (20 cm^{-3}) and velocity (600 km/s) plateaus, respectively.

in Fig.6 compared to Fig.5A.

It can be seen from Fig.6 that the main phases of cases with density plateaus (solid lines) have peaks shifted toward negative values compared to those without changes in density (dotted lines). Since B_z is southward and points to the negative, the plasma essentially flows from the plasma sheet into the ring current. Furthermore, there have been reports that the solar wind plasma is supplied to the plasma sheet, with a certain time lag (Ebihara and Ejiri, 1998, 2000). It is believed that the Dst is enhanced toward the negative side when an increase in density results from this penetration of plasma and a plateau is formed, as seen in Fig.6. In the case of the geomagnetic storms shown in Fig.5, the duration of negative B_z is 15 hours, and the minimum Dst is pushed lower by approximately 10 % relative to the case in which no density change occurs.

As shown in Figs.6 and 7, the absolute value of Dst generally increases with longer durations of southward B_z . However, the increase becomes less dramatic and eventually appears not to increase. Burton et al. (1875) expressed the behavior of Dst using an ordinary differential equation. The second term of the solution is a logarithmic function, and so this equation may be considered to be qualitatively consistent with the above trend. In other words, since the term of ring current generation balances with the decay term, the

change in Dst will become smaller and finally be reversed. Details of further analysis will be given in 6.

In the recovery phase, the B_z is positive, as in the initial phase. The dynamic pressure will work to push the Dst in the positive direction. After 20 hours, the density plateau will become non-existent, as explained above. The sudden drop in dynamic pressure will cause a dramatic decrease in the Chapman-Ferraro current, and the magnitude of Dst will begin to decrease at the point indicated by the arrow.

5.2 The Effect of Velocity

In the calculation for Fig.7, velocity as well as density is assumed to change, following the curve shown in Fig.5C. The velocity increase in the main phase intensifies the westward electric field, and plasma is supplied from the plasma sheet of the magnetotail by the $E \times B$ drift, and the seed of the ring current is carried toward the Earth. The dotted lines, the same as in Fig.6, are plotted here for comparison. The third storm, with negative B_z duration of 15 hours, is the same as in the case shown by the solid line in Fig.5. The minimum Dst calculated for the case involving only a velocity increase (a 50 % increase, from 400 to 600 km/s) was -170 nT, and for the case involving only a density increase (a 570 % increase, from 3 to 20 cm^{-3}), was -140 nT. Thus, it can be concluded that the effect of velocity is greater than that of density.

5.3 The Effect of the Azimuthal Angle Component of the Interplanetary Magnetic Field (IMF)

All calculations up to this point contained only IMF components in the X-direction within the XY-plane. In Fig.8, the absolute value of the IMF component in the XY plane is fixed, and rotated from the X-axis for each case to examine the azimuthal-angle dependency. An example with a 10-nT component in the XY plane is analyzed. The density, velocity, and the southward IMF component are assumed to follow the curve shown by the dotted and solid lines in Fig.5. The results of cal-

culations using the prediction model show that the minimum Dst (maximum storm intensity) is distributed around 10 - 30° in the clockwise direction from the X-axis direction, whether or not there are increases in velocity and density. The difference between the maximum and minimum values in Fig.8 is approximately 20 nT.

Furthermore, when the absolute value of the IMF in the XY-plane is assumed to be 15 nT, the difference is approximately 30 nT. This is qualitatively consistent with model B. This suggests that the effect of the azimuthal angle is significant. Calculations were performed using the density, velocity, and B_z of an actual geomagnetic storm, with the component in the XY-plane having, again, an absolute value of 10 nT. Fig.9 shows a geomagnetic storm frequently referenced in the literature. The geomagnetic storm began on Sept. 29, 1978, and the minimum ground-based Dst was -224 nT. The solid and dotted lines in the top panel of Fig.9 represent ground-based and predicted values, respectively. It can be seen that the prediction is

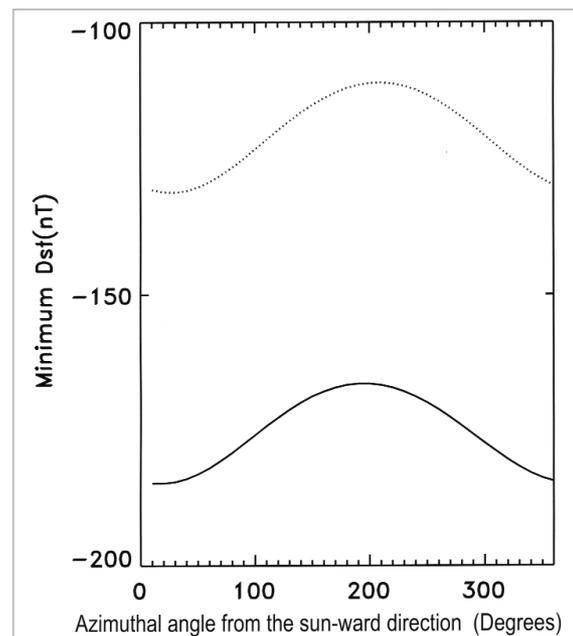


Fig.8 The dependency of Dst on azimuthal angle (in the counterclockwise direction relative to the sunward direction) of the IMF

Solid and dotted lines correspond to cases with and without density and velocity plateaus, respectively.

extremely consistent with observation. The middle panel shows how a shock is formed by extremely high velocity (solid line) and an increase in density (dotted line). As shown in the bottom panel, the southward component of B_z (solid line) is strengthened and causes a strong storm.

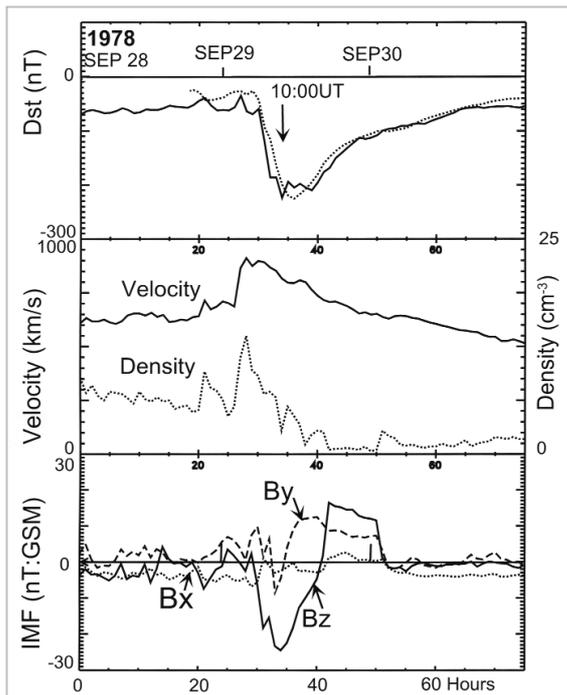


Fig.9 Upper panel: The ground-based Dst values (solid line) and predicted values from actual solar wind parameters (dotted line) for the three-day period beginning Sept. 28, 1978
Middle panel: Velocity (solid line) and density (dotted line) of solar wind
Lower panel: The X, Y, and Z components of the IMF in the GSM system: X, Y, and Z are represented by dotted, dashed, and solid lines, respectively

Fig.10 represents the same storm as in Fig.9. Except for the X-Y component of the IMF, all parameters are the observed values shown in Fig.9. The minimum Dst is calculated for each geomagnetic storm by rotating the IMF component in the XY plane with the absolute value $\sqrt{B_x^2 + B_y^2}$ set at 10 nT. According to Fig.10, the storm has maximum intensity when the X-Y component is pointing in the sunward direction. In this example, the difference between the maximum and minimum values resulting from azimuthal-angle dependency is approximately 20 nT. Fig.11 shows

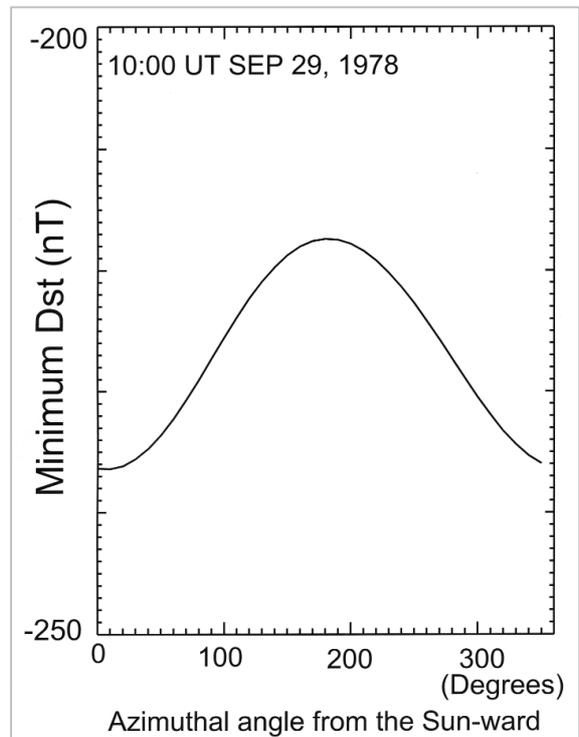


Fig.10 The azimuthal-angle dependency of Dst calculated from actual solar wind parameters
The magnitude in the XY-plane is fixed at 10 nT.

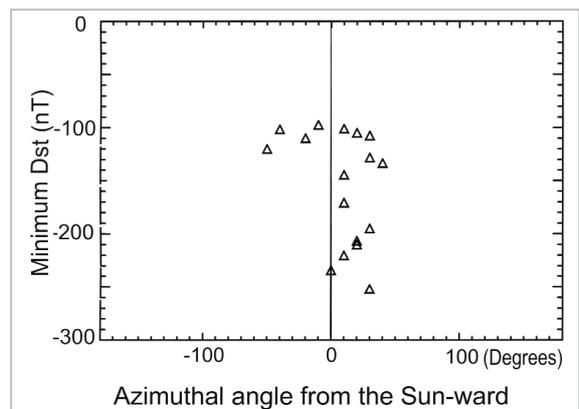


Fig.11 The azimuthal-angle dependency of Dst for 17 apparent cases from 1978 to 1982, giving the minimum Dst during storms

the results of similar attempts made for 17 geomagnetic storms between 1978 and 1982, all having relatively simple structures. Storms with complex structures were excluded from analysis since the time of the minimum Dst (maximum storm intensity) differed for different azimuthal angles, complicating analysis. Fig.11 shows the relationship between the

magnitude of the storm and the azimuthal angle that gives the minimum *Dst* for each storm. According to this plot, the storm is most intensified when the X-Y component points sunward, regardless of the magnitude of the storm.

6 Discussion

This chapter will discuss the classical equations used in short-term predictions, the Burton model, and the NN model. The Burton model (Burton et al., 1975) is not outstanding in terms of prediction, but is helpful in understanding the physical process. Below are the equations used in the Burton model.

$$\frac{dDst^*}{dt} = F(E) - \frac{Dst^*}{\tau} \quad (10)$$

$$Dst^* = Dst - b\sqrt{P} + c \quad (11)$$

$$F(E) = 0 \quad Ey < 0.5mV/m$$

$$F(E) = d(Ey - 0.5) \quad Ey > +0.5mV/m \quad (12)$$

$$b = 16nT(nPa)^{-1/2}, \quad c = 20nT$$

$$d = -5.4 \times 10^{-3}nT(mV/m)/s$$

$$\tau = 7.7hours$$

$$Ey = VBz \quad 10^{-3}mV/m$$

$$P = nV^2 \quad 10^{-2}eV/cm^3$$

Here, *P* is the solar wind dynamic pressure, *Ey* is the westward electric field, and *Bz* is the southward component of the IMF.

In Eq. (10), *Dst** represents the increase in ring current during geomagnetic storms, which is equivalent to the total of the ring current during quiet periods and the effect of the Chapman-Ferraro current subtracted from the *Dst*. Here *F*(*E*), *Dst**/ τ , and τ represent the term of ring current generation, the term of ring current decay, and lifetime, respectively. In the Burton model, τ is fixed at 7.7 hours based on the results of statistical analysis.

This has been used by many researchers as a reference model. However, it has been subsequently noted that τ does not seem to be an appropriate parameter in reflecting the phenomena, and many have proposed models that adopt *Dst*, ϵ (a parameter for the influx of solar wind energy into the magnetosphere),

and *U* (rate of energy supply to ring current) as parameters.

One of the authors of this classical Burton model (McPherron) later proposed a modified model known as the O'Brien and McPherron (OM) model (O'Brien and McPherron, 2000). The primary feature of the OM model is that τ is a function of the duskward electric field *Ey*, and is expressed as:

$$\tau = 2.40exp[9.74/(4.69+VBs)] , \quad (13)$$

Other coefficients are also slightly modified, and

$$F(E)=0 \quad Ey < Ec=0.49mV/m , \quad (14)$$

$$F(E)=-4.4(VBs-Ec) \quad Ey > Ec=0.49mV/m ,$$

$$Dst^*=Dst-7.26P^{1/2} + 11nT. \quad (15)$$

Here, *Bs* is the southward component of *Bz*.

When *Bz* is northward, τ has a maximum value of 19.2 hours in this model. As the southward component of *Bz* increases, τ becomes smaller. Fig.12 shows a comparison of the results of prediction under the NN model, Burton model, and OM model for a geomagnetic storm that began on Oct. 10, 1980 and for one that began on March 2, 1982. These storms were selected because they were used in analysis by O'Brien and McPherron (2000), with the results presented in Figs. 8 a and 8 b of their paper. The minimum *Dst* values observed on the ground for the two storms are -104 nT and -211 nT.

Panel A in Fig.12 compares the values predicted by each model. Ground-based *Dst* is shown by a thick gray line, while the predicted values for the NN model, the Burton model, and the OM model are shown in red, green, and blue, respectively. The predicted values under the OM model are the line referred to as the "Multi-Step" in O'Brien and McPherron (2000). The Adams-Bashforth-Moulton method is used to solve the differential equation. Since this method requires four known points at initial steps, the actual ground-based values were used in the calculation. Since the

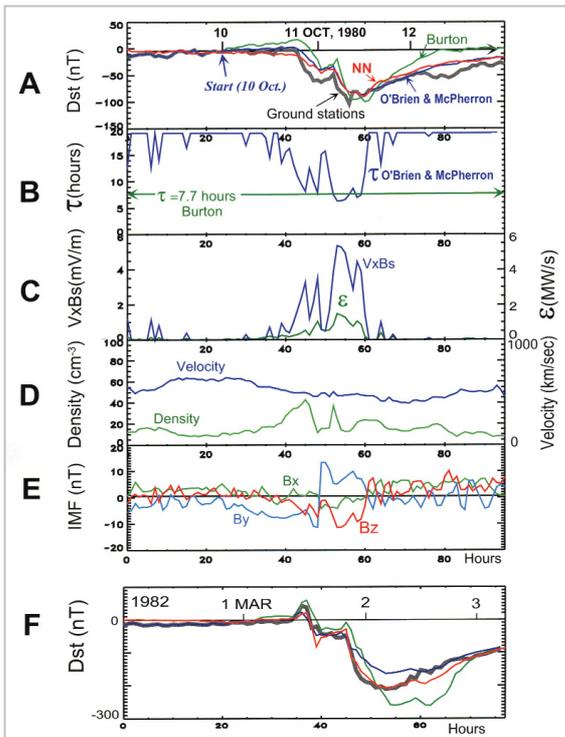


Fig. 12 Comparison of the NN model, the Burton model, and the O'Brien and McPherron (modified Burton) model

Panel A: Predicted values for a storm on Oct. 9-12, 1980. Ground-based values (thick gray), and values predicted by the NN model (red), Burton (green), and O'Brien and McPherron (blue) models.

Panel B: Decay time τ for the Burton (green) and the O'Brien and McPherron (blue) models.

Panel C: VBs and ϵ of solar wind.

Panel D: Solar wind velocity (km/s) and density (cm^{-3}).

Panel E: The three components of the IMF in the GSM system.

Panel F: Comparison of predicted values calculated according to the three models for a storm occurring Feb. 28-March 3, 1982. Line colors are the same as for panel A.

NN model has a lasting time dependency, calculations were initiated one day before the other methods. However, in actual use of the NN method, the run is not terminated unless a system error occurs and continues for several months to years. Thus, the effect of the initial values can be ignored.

Panel B is a plot of the calculation results for τ . The τ for the OM model is longer in the early initial phase and the recovery phase than the fixed value of 7.7 hours in the Burton model. Panel A shows that the shorter τ in the

recovery phase of the Burton model results in a faster recovery from a geomagnetic storm. Panel C shows the magnitude of the $E_y = VB_s$ westward electric field that controls the lifetime τ for the OM model and the total energy ϵ flowing into the magnetosphere from the solar wind during geomagnetic storms. Akasofu (2001) grouped geomagnetic storms by magnitude according to the following classification:

Weak geomagnetic storm	$\epsilon \leq 0.25 \text{ MW/s}$
Moderate geomagnetic storm	$\epsilon \leq 1.4 \text{ MW/s}$
Extremely strong geomagnetic storm	$\epsilon \leq 8.0 \text{ MW/s}$

According to this classification, the geomagnetic storm on Oct. 10, 1980 is a moderate storm, since the peak ϵ (green) is approximately 1.5 MW/s, as seen from panel C.

Panel D shows the velocity (km/s; blue) and density (cm^{-3} , green). The change in velocity is gradual, and the geomagnetic storm is a non-shock type lacking a distinct initial phase. The B_z curve (red) shown in panel E controls the progress of the storm. Like the NN model, the OM model more neatly traces the ground-based Dst than the Burton model.

The OM model is a good model, but as explicitly set forth by those who proposed it, it cannot be applied to strong geomagnetic storms with Dst below -150 nT . However, such storms are not rare events; this makes the OM model unsuitable as an operational model. In contrast, the NN model produces impressive prediction results and has almost no limits on potential applications. The results of analysis in Fig. 8 b of O'Brien and McPherron (2000) is also shown in panel F for comparison. The ground-based Dst for this case is shown by a thick, gray line, and the minimum Dst was -211 nT . The results for the Burton model are shown by a green line. It can be seen that the curve deviates significantly from the ground-based Dst from the main phase to the recovery phase. This results from the overestimation of τ , which should have been given a value shorter than 7.7 hours. In the OM model, the absolute value of the Dst is

smaller than the actual value, but as stated previously, this results from the application of the model beyond its limits, a shortcoming that has been clearly recognized and pointed out by the authors of the model. On the other hand, the curve of the values predicted by the NN model (red) coincides extremely well with the ground-based values, demonstrating the effectiveness of the model.

As stated previously, the prediction error of ground-based values contains the error of correlation between the solar wind environment at the L1 point and near Earth. Dividing the plasma near the L1 point into blocks of six hours, Paularena et al. (1998) compared the plasma conditions to near-Earth conditions and found that the coefficient of correlation is only around 0.7. Richardson et al. (1998) have also reported a higher coefficient of correlation for the IMF than for plasma conditions. Crooker et al. (1982) have also stated that there are limits to the correlation between the solar wind parameters at the L1 point and near Earth.

7 Conclusions

Using actual time series data observed by the ACE satellite, we constructed a *Dst* prediction model that has been in operation since May 1998 (<http://www2.crl.go.jp/uk/uk223/service/nnw/index.html>). During the period from February to the end of October 1998, there were 11 major storms with minimum *Dst* below -80 nT. Predicted values for 10 out of

the 11 storms had errors within 23 %, while one had a prediction error of 43 %. The large error is believed to result from poor correlation between the state of the solar wind plasma near the ACE satellite (positioned 1.5 million km away from the Earth) and near Earth.

It is widely known that the southward turning of the IMF is the fundamental cause of geomagnetic disturbances. It has now been determined that the positive *B_x* (sunward) component works to increase the magnitude of the disturbance.

Changes in the density of solar plasma increase dynamic pressure. Therefore, the enhancement of the Chapman-Ferraro current has the effect of shifting the *Dst* in the positive direction. However, when *B_z* is negative, the main phase will lower the *Dst*, and so the density increase will intensify a storm.

It can be concluded from the evaluation of the NN models that the effects of velocity are greater than the density effect.

Our operational *Dst* prediction model is essentially applicable to almost any type of storm.

Acknowledgements

We are grateful to the staff of the ACE/SWEPAM instrumentation group and the ACE/MAG instrumentation group, as well as the staff at the ACE Science Center, which provides ACE data. We also thank Prof. Masahisa Sugiura for his review of the manuscript.

References

- 1 Akasofu, S.-I., "Predicting Geomagnetic Storms as a Space Weather Project", Geophys. Monogr. Ser., 125, edited by P. Song, H. J. Singer, G. L. Siscoe, AGU Washington DC, 329-337, 2001.
- 2 Blanchard, G. T., and R. L. McPherron, "A bimodal representation of the response function relating the solar wind electric field to the AL index", in Solar-Terrestrial Prediction Proc., vol. 2, pp. 479-486, U. S. Dept. of Commerce, Boulder, Colo., 1992.
- 3 Burton, R. K., R. L. McPherron, and C. T. Russell, "An empirical relationship between interplanetary conditions and *Dst*", J. Geophys. Res. 80, 4204-4214, 1975.
- 4 Crooker, N. U., G. L. Siscoe, C. T. Russell, and E. J. Smith, "Factors controlling degree of correlation between ISEE 1 and ISEE 3 interplanetary magnetic field measurements", J. Geophys. Res. 87, 2224-2230,

1982.

- 5 Ebihara, Y., and M. Ejiri, "Modeling of solar wind control of the ring current buildup: A case study of the magnetic storms in April 1997", *Geophys. Res. Lett.* 25, 3751-3754, 1998.
- 6 Ebihara, Y., and M. Ejiri, "Simulation study on fundamental properties of the storm-time ring current", *J. Geophys. Res.* 105, 15843-15859, 2000.
- 7 Elman, J. L., "Finding structure in time", *Cognitive Sci.* 14, 179, 1990.
- 8 Feldstein, Y. I., "Modeling of the magnetic field of magnetospheric ring current as a function of interplanetary medium parameters", *Space Sci. Rev.*, 59, 83-165, 1992.
- 9 Klimas, A. J., D. Vassiliadis, and D. N. Baker, "Data-derived analogues of the magnetospheric dynamics", *J. Geophys. Res.* 102, 26993-27009, 1997.
- 10 Klimas, A. J., D. Vassiliadis, and D. N. Baker, "Dst index prediction using data-derived analogues of the magnetospheric dynamics", *J. Geophys. Res.* 103, 20435-20447, 1998.
- 11 Kugblenu, S., S. Taguchi, and T. Okuzawa, "Prediction of the geomagnetic storm associated Dst index using an artificial NN algorithm", *Earth Planet Sci.* 51, 307-313, 1999.
- 12 Iyemori, T., H. Maeda, and T. Kamei, "Impulse response of geomagnetic indices to interplanetary magnetic field", *J. Geomagn. Geoelectr.* 31, 1-9, 1979.
- 13 Lundstedt, H., and P. Wintoft, "Prediction of geomagnetic storms from solar wind data with the use of a neural network", *Ann. Geophys.* 12, 19-24, 1994.
- 14 O'Brien T. P., and R. L. McPherron, "An empirical phase space analysis of ring current dynamics: Solar wind control of injection and decay", *J. Geophys. Res.* 105, 7707-7719, 2000.
- 15 Paularena, K. I., G. N. Zastenker, A. J. Lazarus, and P. A. Dalin, "Solar wind plasma correlations between IMP 8, INTERBALL-1, and WIND", *J. Geophys. Res.* 103, 14601-14617, 1998.
- 16 Richardson, J. D., F. Dashevskiy, and K. I. Paularena, "Solar wind plasma correlations between L1 and Earth", *J. Geophys. Res.* 103, 14619-14629, 1998.
- 17 Wu, J.-G., and H. Lundstedt, "Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks", *J. Geophys. Res.* 102, 14255-14268, 1997a.
- 18 Wu, J.-G., and H. Lundstedt, "Neural network modeling of solar wind- magnetosphere interaction", *J. Geophys. Res.* 102, 14457-14466, 1997b.



WATANABE Shigeaki, Dr. Sci.
Senior Researcher, Space Weather Group, Applied Research and Standards Division
Space Weather Forecast, Plasma Wave

SAGAWA Eiichi, Dr. Sci.
Senior Researcher, Space Weather Group, Applied Research and Standards Division
Space Weather



OHTAKA Kazuhiro
Researcher, Space Weather Group, Applied Research and Standards Division
Space Weather



SHIMAZU Hironori, Dr. Sci.
Senior Researcher, Space Simulation Group, Applied Research and Standards Division
Space Plasma Physics

