Earth Planets Space, **51**, 307–313, 1999

Prediction of the geomagnetic storm associated D_{st} index using an artificial neural network algorithm

Samuel Kugblenu, Satoshi Taguchi, and Takashi Okuzawa

Department of Electronic Engineering, University of Electro-Communications, Tokyo 182-8585, Japan

(Received July 5, 1997; Revised March 19, 1999; Accepted March 20, 1999)

In order to enhance the reproduction of the recovery phase D_{st} index of a geomagnetic storm which has been shown by previous studies to be poorly reproduced when compared with the initial and main phases, an artificial neural network with one hidden layer and error back-propagation learning has been developed. Three hourly D_{st} values before the minimum D_{st} in the main phase in addition to solar wind data of IMF southward-component B_s, the total strength B_t and the square root of the dynamic pressure, $\sqrt{nV^2}$, for the minimum D_{st}, i.e., information on the main phase was used to train the network. Twenty carefully selected storms from 1972–1982 were used for the training, and the performance of the trained network was then tested with three storms of different D_{st} strengths outside the training data set. Extremely good agreement between the measured D_{st} and the modeled D_{st} has been obtained for the recovery phase. The correlation coefficient between the predicted and observed D_{st} is more than 0.95. The average relative variance is 0.1 or less, which means that more than 90% of the observed D_{st} variance is predictable in our model. Our neural network model suggests that the minimum D_{st} of a storm is significant in the storm recovery process.

1. Introduction

The development of a magnetic storm is best identified at low latitudes by large decreases in the H component of the Earth's magnetic field, and consequently in the D_{st} index, a quantity derived and introduced by Sugiura (1964) as a measure of the magnetic disturbance level on the Earth.

Geomagnetic storms, illustrated by D_{st}, typically have three phases: initial, main, and recovery phases. The initial phase is caused by an increased solar wind dynamic pressure acting on the magnetosphere. The increased pressure compresses the dayside magnetosphere, forcing the magnetopause current closer to the Earth, and at the same time increasing it. The magnitude of the initial phase has been shown to be proportional to the square root of the solar wind dynamic pressure, $\sqrt{nV^2}$ (Ogilvie *et al.*, 1968; Siscoe *et al.*, 1968), where n is the solar wind density and V, the solar wind speed.

The main phase is due to an increase in energetic ions and electrons in the inner magnetosphere, where they become trapped on closed magnetic field lines and drift around the Earth, thus creating the ring current. The storm-time ring current has been shown to consist of solar wind ions and ionospheric-origin ions. H^+ ions carry the major fraction of energy in the ring current almost throughout the storm, but O^+ ions prevail near the maximum of the main phase, particularly for large storms (Gloeckler and Hamilton, 1987; Hamilton *et al.*, 1988). Recent studies (Daglis, 1997; Hamilton, 1997) have confirmed the significant contribution of O^+ ions to the minimum D_{st} value; that is, the minimum D_{st} shows a remarkable correlation with fractional O^+ concentration which can attain values as large as 70% for storms with $D_{st} \leq -300$ nT. The main phase was also found to be associated with sustained southward IMF-B_z (Rostoker and Fälthammar, 1967).

Charge exchange and Coulomb scattering have both been identified as major loss processes responsible for the decay of the ring current during the recovery phase of storms (Smith and Bewtra, 1978; Fok *et al.*, 1991). Typical H^+ and O^+ lifetimes for each of these two processes are comparable to days, characteristic of the slow recovery in D_{st} following the main phase of a storm. However, since the charge-exchange lifetime of O^+ is considerably shorter than the H^+ lifetime for ring current energies ($\geq$40 keV), Smith and Bewtra (1978) have suggested that increased O^+ influences the decay rate of the ring current.

Many different methods (Baker, 1986) have been used for predicting geomagnetic storms, such as ordinary statistical methods (which include visual correlation analyses as well), linear filtering and artificial intelligence methods. Most of these relationships are based on parameter studies in a cause-and-effect manner. Solar wind parameters constitute the causal variables, while geomagnetic indices (D_{st}, AE, AL, or AU) are the designated effect variables.

Burton *et al.* (1975) proposed an empirical linear relationship for predicting the D_{st} index from the knowledge of the solar wind velocity, density, and the southward component of IMF. In particular, Burton *et al.* developed an equation for the rate of change of pressure corrected D_{st}, showing that it was a balance between injection and decay out of the ring current. They found that decay rate for the recovery phase depends on the present strength of the ring current (D_{st}).

The method of linear filtering prediction was used by

Iyemori *et al.* (1979) to predict geomagnetic-storm indices. This technique allows one to empirically determine the most general linear relationship between a solar wind input function and a geomagnetic disturbance output function, taking time delays and frequency response into consideration (Clauer, 1986). Iyemori *et al.* (1979), and Iyemori and Maeda (1980) used IMF-B_z as an input function, and their output function was one of the geomagnetic indices, D_{st}, AE, AL, and AU. They concluded that using a single input solar-wind function is not enough to predict completely a geomagnetic disturbance based on the assumption of a linear system. They attributed this inability to a possible nonlinear response of the magnetosphere.

Artificial neural network (ANN) models enjoyed a resurgence in popularity as a prediction tool during the late 1980s, as a consequence of the discovery of the backpropagation of errors learning algorithm. One of the most unique property of these ANNs is their ability to generalize to new situations after having been trained on a number of examples of a relationship. They can then induce a complete relationship that interpolates and extrapolates from the examples. ANNs therefore offer the possibility to study large complex nonlinear systems of highly inter-correlated data.

Several ANN models (e.g., Freeman *et al.*, 1993; Lundstedt and Wintoft, 1994; Gleisner *et al.*, 1996) have shown good performance for the reproduction of D_{st}. Although the reproduction of the initial and main phases was excellent, there was somewhat difficulty in reproducing the recovery phase well. In this study, in order to enhance the recovery phase reproduction, we use a feed forward multilayer neural network with error-backpropagation learning algorithm. This network is trained using three hourly values before the minimum D_{st} in the main phase designated as $D_{st}(-1)$, $D_{st}(-2)$, and $D_{st}(-3)$, in addition to solar wind parameters for the minimum D_{st}, i.e., the magnetic field strength B_t, the southward IMF component B_s and the solar wind dynamic pressure $\sqrt{nV^2}$. In other words, we used information only on the main phase for the network training. This trained network is able to reproduce the recovery phase with high accuracy.

2. Artificial Neural Networks

ANNs make up a new approach to the computation that involves developing mathematical structures with learning ability. This approach is the result of academic investigations to model the human nervous system learning. ANN is essentially a group of interconnected computing elements (i.e., neurons).

The multi-layer feed-forward error backpropagation algorithm (Hertz *et al.*, 1991, for detail) was used in this study. This network belongs to the class of supervised networks, i.e., it learns from known answers. Typically, the network is arranged in layers of neurons (nodes), where every neuron in a layer computes the sum of its inputs and passes this sum through a nonlinear function (an activation function) as its output. Each neuron has only one output, but this output is multiplied by a weighting factor if it is to be used as an input to another neuron (in a next higher layer). There are no connections among neurons in the same layer.

Figure 1 shows a typical three layer network structure. The

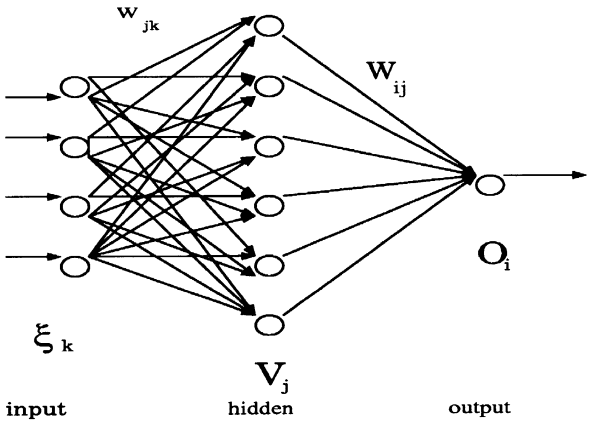

Fig. 1. A three-layered neural network.

input, hidden, and output layers are denoted by k, j, and i, respectively. For a given pattern, μ, a neuron j in the hidden layer receives, from a neuron k in the input layer, a net input signal given by

$$x_j^\mu = \sum_k w_{jk}\xi_k^\mu + b_k \tag{1}$$

where ξ_k^μ is the input signal fed to neuron k in the input layer, and w_{jk} is the connection strength between neuron j in the hidden layer and neuron k in the input layer, while b_k is a bias connected to the input layer. The bias is an additional input to a neuron that serves to normalize its output, and normally has a constant activation of 1.

The output V_j produced by the jth neuron in the hidden layer is related to the activation value for that neuron, by a transfer function $g_h(x)$ which will be defined later, in such a way as

$$V_j^\mu = g_h(x_j^\mu) = g_h\left(\sum_k w_{jk}\xi_k^\mu + b_k\right). \tag{2}$$

A neuron i in the output layer receives this signal from the neuron j in the hidden layer as input, and similarly produces an output, O_i^μ, which is related to the activation value for that neuron, by a transfer function $g_o(x)$ (which will be assumed to be the same form of $g_h(x)$ in this paper)

$$O_i^\mu = g_o(x_i^\mu) = g_o\left(\sum_j W_{ij}V_j^\mu + b_j\right). \tag{3}$$

The biases in the equations, b_k and b_j, can be omitted as they can be considered as an extra input of unit value connected to all units in the network. The final net output for an input pattern μ, ($\mu = 1, \ldots, p$) can be then described by

$$O_i^\mu = g_o\left(\sum W_{ij}g_h\left(\sum w_{jk}\xi_k^\mu\right)\right). \tag{4}$$

Activation functions for the hidden layers are needed to introduce nonlinearity into the network. Without nonlinearity, hidden layers would not make networks more powerful than just simple perceptrons (which do not have any hidden layers, only input and output layers). A composition of

linear functions is again a linear function. It is the nonlinearity (i.e., the capability to represent nonlinear functions) that makes multi-layer networks so powerful. Almost any nonlinear function is applicable. For backpropagation learning, however, it must be differentiable and saturating at both extremes. Sigmoid functions such as the logistic and hyperbolic tangent functions, and the Gaussian function are the most common choices. If the transfer functions were chosen to be linear, then the network would become identical to a linear filter (Iyemori *et al.*, 1979). A backpropagation net with nonlinear transfer functions could be thus regarded as a nonlinear generalization of a linear filter (see Gleisner and Lundstedt, 1997).

The steepness of the logistic sigmoid can be modified by a slope parameter σ. The more general sigmoid function (with range between 0 and 1) is given by

$$g(x) = g_{\mathrm{h}}(x) = g_{\mathrm{o}}(x) = \frac{1}{1 + \exp(-\sigma x)} \tag{5}$$

with its derivative as

$$g\prime(x) = \sigma g(x)[1 - g(x)]. \tag{6}$$

The slope may be determined such that the sigmoid function achieves a particular desired value for a given value of x, the input (Fausett, 1994). In the present study σ has been set to 4 for nonlinear multiregression analysis (Ichikawa, 1993).

The training of a network by backpropagation involves three stages: the feedforward of the input training pattern, the calculation and backpropagation of the associated error, and the adjustment of the weights. After training, application of the net involves only the computations of the feedforward phase. In order to train the network, input is shown to the net together with the corresponding known output, and if there exists a relation between the input, ξ_k^μ, and the output, O_i^μ, the net *learns* by adjusting the weights until an optimum set of weights that minimizes the network error is found and the network then converges.

The network error, E, which is defined as the sum of the individual errors over a number of examples, is given by

$$E = \frac{1}{2} \sum_{\mu=1}^{p} \sum_{i=1}^{N_{\mathrm{out}}} (T_i^\mu - O_i^\mu)^2 \tag{7}$$

where T_i^μ is the target or desired output ($i = 1, \ldots, N_{\mathrm{out}}$). Substituting (4) for the net output, O_i^μ, we have

$$E = \frac{1}{2} \sum_{\mu=1}^{p} \sum_{i=1}^{N_{\mathrm{out}}} \left[T_i^\mu - g\left(\sum_j W_{ij} g\left(\sum_k w_{jk} \xi_k^\mu \right) \right) \right]^2 . \tag{8}$$

At the completion of a pass through the entire data set, all the nodes change their weights based on the accumulated derivatives of the error with respect to each weight. These weight changes move the weights in such a direction that the error declines most quickly. The standard learning algorithm which updates the weights can be expressed by the gradient descent rule, which means that each weight, say, w_{pq}, changes by an amount Δw_{pq} which is proportional to the gradient of the error E at the present location.

For the hidden-to-output connections, the gradient descent rule gives the change in weight as

$$W_{ij}^{\mathrm{new}} = W_{ij}^{\mathrm{old}} + \Delta W_{ij} \tag{9}$$

where

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} \tag{10}$$

and η stands for the learning rate. Similarly, the weight change for the input-to-hidden connections, is given by

$$w_{jk}^{\mathrm{new}} = w_{jk}^{\mathrm{old}} + \Delta w_{jk}, \tag{11}$$

where

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \sum_\mu \frac{\partial E}{\partial V_j^\mu} \frac{\partial V_j^\mu}{\partial w_{jk}}. \tag{12}$$

3. Data and Training

The neural network software used for this work was originally developed by Ichikawa (1993). The network structure is flexible to configure since all parameters are commutative. Our model is a feed forward multi-layer network with error-backpropagation learning, and applied as a nonlinear, multi-variable least squares algorithm.

For both network training and prediction, we used the OMNI data set from the database of the National Space Science Data Center of NASA and the WDC-2 Data Center of Kyoto University. These data sets consist of hourly averages of the solar wind plasma and IMF data from various spacecraft. Hourly average of D_{st} is also available in these databases. For the storm selection, we required that the storm be preceded by a relatively quiet period of D_{st} activity. Storms with large and sustained southward IMF were given priority in the selection process, as well as sudden increases in the velocity of solar wind, mostly above 450 km/s. Selection of storm events was, however, limited owing to incompleteness of solar wind data. Twenty storms of different intensities with $D_{st} \leq -30$ nT were selected from the years of 1972, 1973, 1974, 1975, 1977, 1978, 1979, 1981, and 1982. These are shown in Table 1.

Table 1 also shows the hour and values of the minimum D_{st} (peak value) for the 20 storms (column 2). When the solar wind data set corresponding to the minimum value of D_{st} is not available in the database, the nearest hour having a complete data set was taken. This hour is indicated in column 3, where the hour is different from that of column 2 for November 1, 1972, April 25, 1979, and September 26, 1982 storm events.

The minimum D_{st} peak value has been used in ring current energization studies. Pudovkin *et al.* (1985) used the hourly D_{st} peak value to study the relationship of the ring current energy to solar wind functions and to calculate the characteristics decay time. For our network training, therefore, we made a data set consisting of minimum D_{st} value in the main phase for each storm, which we designate $D_{st}(0)$, the corresponding solar wind data (IMF total stength B_{t}, IMF southward component B_{s}, and dynamic pressure $\sqrt{nV^2}$) for that hour, and 3 consecutive preceding hourly D_{st} values in the main phase. These D_{st} values are designated by $D_{st}(-1)$, $D_{st}(-2)$, and $D_{st}(-3)$. The $D_{st}(0)$ value is shown to the net

Table 1. List of storms used for network training.

Date	Hour of D_{st} min.	Hour used for training
November 1, 1972	08 (−199 nT)	02 (−103 nT)
March 19, 1973	18 (−84 nT)	18
April 1, 1973	22 (−211 nT)	22
July 6, 1974	06 (−204 nT)	06
November 17, 1975	20 (−89 nT)	20
January 30, 1977	22 (−95 nT)	22
November 26, 1977	00 (−87 nT)	00
December 11, 1977	11 (−112 nT)	11
August 4, 1978	11 (−32 nT)	11
August 28, 1978	09 (−226 nT)	09
September 6, 1978	09 (−56 nT)	09
September 29, 1978	10 (−224 nT)	10
April 4, 1979	03 (−202 nT)	03
April 25, 1979	14 (−149 nT)	13 (−136 nT)
June 7, 1979	06 (−40 nT)	06
September 18, 1979	15 (−158 nT)	15
April 13, 1981	06 (−311 nT)	06
July 25, 1981	20 (−226 nT)	20
March 2, 1982	05 (−211 nT)	05
September 26, 1982	18 (−187 nT)	15 (−164 nT)

as the known output parameter. A data set for one storm thus consists of a six parameter input and a corresponding one parameter output. This is presented to the net during training. For prediction, we used the six parameters with no corresponding output parameter. This includes three preceeding previous D_{st} values and the solar wind data set for each hour of a storm event.

Determination of the number of hidden nodes is a difficult task. Although some techniques have been used by different authors (e.g., Lundstedt and Wintoft, 1994), no established methods exist. We have therefore applied different network architectures with various number of hidden nodes and learning rates to the network during training in order to determine the best such parameters. A compromise for the best fit has been found to be 7 nodes for the hidden layer and 0.15 for the learning rate η. This learning rate is the best among an applicable range of 0.05 to 0.2.

4. Network Prediction

4.1 December 19, 1980 storm event

Our purpose is to reproduce the D_{st} index for the recovery phase using data from the main phase, nevertheless we apply our trained ANN to the period starting with several hours before the beginning of the initial phase. A sudden increase in the D_{st} plot at 29th hour (05:00 UT, December 19) defines the beginning of the initial phase of the storm. The initial phase lasted for roughly about 10 hours. This was followed by a very rapid decrease in D_{st} reaching its minimum approximately in the 42nd hour (18:00 UT, December 19). This constitutes the main phase. The D_{st} index then began a rapid recovery at first, followed by a long and slow one until the 119th hour. The D_{st} minimum marks the beginning of the recovery phase during which the ring current decays. This event has 119 data points.

Figure 2 shows the prediction result for the storm of December 19, 1980 compared with observed D_{st}. The solid and broken lines represent the observation and prediction, respectively. Good agreement between the measured D_{st} and the predicted one exists in the recovery phase.

The recovery phase, which is more dependent on the internal processes of the magnetosphere, is bound to be represented by the D_{st} history. The minimum D_{st} should also be well reproduced since main phase data were used in training the network. The initial phase appears to be well reproduced, but this is not the general trend as is shown later by the poor fitting of the initial phase in Fig. 5. However, it is interesting to reproduce somewhat the D_{st} jump at the beginning of the initial phase. Considering that there is often a gradual decrease in the dynamic pressure during the main phase where D_{st} decreases, we speculate that the network can learn some relation between the variations of dynamic pressure and D_{st} and produce the D_{st} increase in accordance with a sharp in-

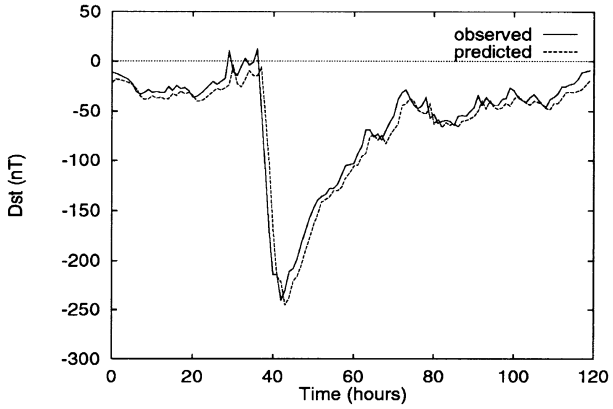

Fig. 2. Observed and predicted D_{st} plot of December 19, 1980 storm. The horizontal axis represents time from the beginning of December 18, 1980.

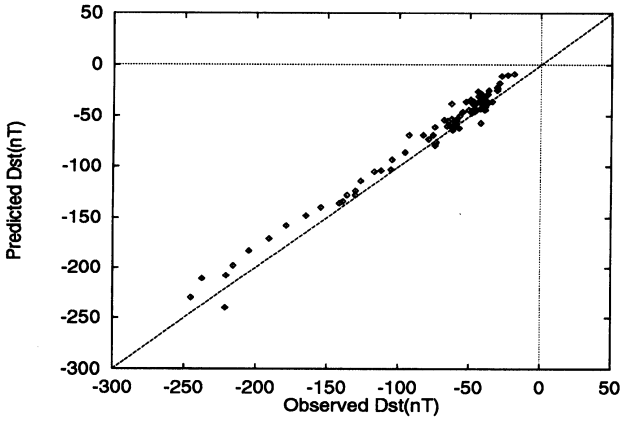

Fig. 3. Correlation plot of predicted D_{st} and observed D_{st} of December 19, 1980 storm.

crease in the dynamic pressure for the initial phase.

Figure 3 shows correlation between the network prediction output and the target for the recovery phase, i.e., a period of time from the 18th hour of December 19 to the 23th hour of December 22. The correlation coefficient for this interval is 0.98. We also calculated the average relative variance (*ARV*), i.e., the mean squares error normalized by the variance of the data,

$$ARV = \frac{\sum_{t=1}^{N} (D(t) - O(t))^2}{\sum_{t=1}^{N} (D(t) - \langle D \rangle)^2} \tag{13}$$

and the root mean squares error (*RMSE*):

$$RMSE = \left[\frac{1}{N} \sum_{t=1}^{N} (D(t) - O(t))^2 \right]^{1/2} \tag{14}$$

for the same interval, where $D(t)$, $\langle D \rangle$ and $O(t)$ denote observed, its averaged, and predicted D_{st}, respectively. *ARV* is 0.04, which means that 96 percent of the observed D_{st} variance is predictable from both solar wind and D_{st} history. The value of *RMSE* is 11 nT, which is very small compared with the lowest peak value of D_{st}. This storm event has been also

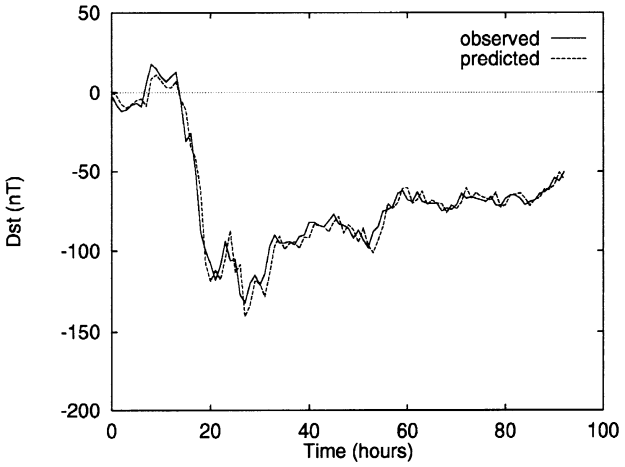

Fig. 4. Observed and predicted D_{st} plot of April 25–26, 1989 storm. The horizontal axis represents time from the 3rd hour of April 25, 1989.

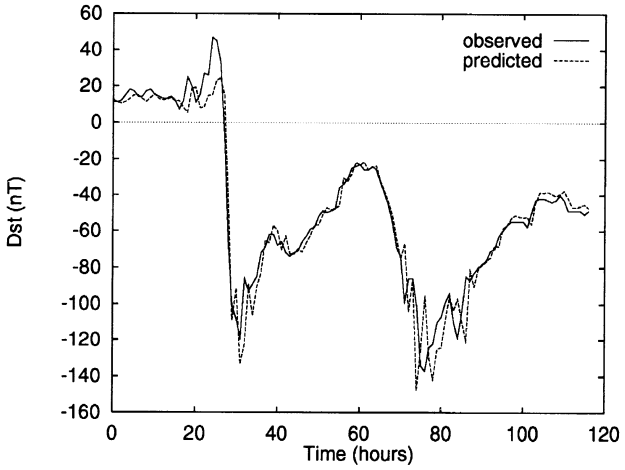

Fig. 5. Observed and predicted D_{st} plot of May 8–12, 1981 storm. The horizontal axis represents time from the biginning of May 8, 1981.

studied by Lundstedt and Wintoft (1994) for D_{st} network prediction. Our model shows much better performance in the recovery phase.

4.2 April 25–26, 1989 storm event

April 25–26, 1989 storm event (Fig. 4) began at the 7th hour of the 25th day of April, 1989. From a value of +13 nT, D_{st} decreased sharply to a low value of −118 nT to comprise the first stage of a two-stage main phase (Kamide *et al.*, 1998). The D_{st} index then increased briefly to a value of −94 nT (due to a rise in IMF-B_z at that hour) before finally falling to the minimum value of −132 nT. The D_{st} index began a disturbed, long-lasting and slow recovery until the 85th hour, which we defined as the end of the recovery phase. In this case, we selected 93 data points, and obtained good agreement between measured and predicted D_{st}. The correlation coefficient between the measured and predicted D_{st} over the recovery phase duration is 0.96. The diagnostic parameter *ARV* has been calculated to be 0.1, indicating that

90 percent of the observed D_{st} variance is predictable. The value of *RMSE* is 5.5 nT.

4.3 May 8–12, 1981 storm event

As shown in Fig. 5, this storm event has two distinct minima. The main phase started around the 25th hour of the event, and fell rapidly to the first minimum of −120 nT. The recovery phase started with the northward turning of the IMF until about the 62th hour, when the IMF turned southward again to produce another main phase. This was again followed by another northward turning of the IMF to produce the consequent recovery phase.

We used 117 data points to reproduce this storm. Figure 5 illustrates our result. The coefficient of correlation has been calculated to be 0.96. The *RMSE* becomes 11 nT and the *ARV* has been found to be 0.07; that is, 93% of the storm is predictable.

5. Discussion

Recently, Gleisner *et al.* (1996) have shown that the solar wind history for 18–24 hours, which is a longer time window than our 3 hours, is needed as input to a trained network in order to reproduce all phases of a geomagnetic storm. They developed a time-delay feed forward neural network based on a temporal sequence of solar wind data. Their networks showed better performance with larger temporal size of the input data sequence, and were able to reproduce 85% of the D_{st} variance, which is quite an improvement on the earlier work of Lundstedt and Wintoft (1994). In our model 90% of D_{st} have been reproduced, at least, for the recovery phase of the three storm events.

The use of previous D_{st} values as input has been suggested by Lundstedt and Wintoft (1994) as another way to model all phases of the magnetic storm. However, as the measured D_{st} is not instantaneously available, they proposed the use of the predicted D_{st} a few hours back in time as input to the net. It is noted that Freeman *et al.* (1993) had already used previous D_{st} data together with solar wind inputs to predict 1-hour ahead D_{st} for input to their Magnetospheric Specification and Forecast Model (MSFM). For input to their network, they used the hourly averages of the IMF-B, B_z, the solar wind pressure, and D_{st} values for the four previous hours. Their output is the 1-hour ahead D_{st} value.

Our model is consistent with the linear autoregressive moving average (ARMA) filter assumption that geomagnetic activity O, can be described as a function of a time series of both solar activity I, and the previous geomagnetic activity,

$$O_t = F(I_{t-1}, I_{t-2}, \ldots, I_{t-T_I}, O_{t-1}, O_{t-2}, \ldots, O_{t-T_S}) \tag{15}$$

where T_I is the system memory for solar wind inputs, and T_S is the memory for previous magnetospheric states. Nonlinear generalizations of this filter have been referred to as state-input space models (Vassiliadis *et al.*, 1995). The linear autoregressive moving average (ARMA) filter is extensively discussed by Detman and Vassiliadis (1997). A direct ANN equivalent is the Elman recurrent network.

Wu and Lundstedt (1996), attempting to reproduce the recovery phase which was difficult to model with the networks of Lundstedt and Wintoft (1994), used Elman recurrent networks. The Elman network is an extension of the multi-layer backpropagation networks with an addition of a feedback connection from hidden to input layers which allows the network to both detect and generate time-varying patterns. Their Elman network model obtained a relatively high correlation coefficient, i.e., 0.91, for a very long period of D_{st} (900 hours). For each storm period, however, their model does not always give a good fit to the observations. Our model uses three hours of observed D_{st} history, whilst their model depends on the feedback element to generate the decay rate; our network is updated with previously measured data, and in their case, it is updated with previously predicted data.

Vassiliadis *et al.* (1996) presented a method that converts an ARMA model to a physical model, namely, a nonlinear damped oscillator. This is expressed as a second order differential equation in terms of the D_{st} index as

$$\frac{d^2 D_{st}}{dt^2} + \nu \frac{d D_{st}}{dt} + \Omega^2 D_{st} = \alpha E_y \tag{16}$$

where the model parameters ν, Ω and α are determined locally for many different levels of activity of D_{st} and E_y, the input of solar-wind electric field. Our model can be considered as a neural net generalization of such category of second-order differential equation.

The number of previous hourly D_{st} values included beside the solar wind data as input to the network may be significant in the modeling of the current D_{st}. Training sequences with one and two D_{st} values did not produce good fits to the observed data (not shown here). The use of one D_{st} value as input produces overfitting (more positive) in the recovery phase, while the inclusion of two D_{st} values produces underfitting. The length of D_{st} history may be an indication of the timescale of the dissipation processes that determine the magnetospheric system memory.

We have used IMF-B_t as input to our net, and investigated its effect on the network output. We found that deficiency of the B_t parameter in the input leads to poor modeling of the initial stage of the recovery phase, implying that this stage is associated with IMF-B_t. Prediction is therefore improved with B_t as input.

6. Conclusion

To reproduce the recovery phase D_{st} index we have developed a model by training the ANN with the D_{st} minimum, three hourly D_{st} values before the minimum in the main phase, and the solar wind data for the D_{st} minimum. This model has greatly enhanced the recovery phase reproduction. This suggests that a process producing minimum D_{st} of a storm is significant in the storm recovery process. This may be related to the suggestion that the increased O^+ which dominates the ring current during storm maximum influences the decay rate of the ring current (Smith and Bewtra, 1978). Further work would be needed for the decisive reason for the significance of the D_{st} minimum for the storm recovery process.

Acknowledgments. The authors thank T. Shibata, University of Electro-Communications, Y. Kamide, STE Lab., Nagoya University, and T. Iyemori, Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University, for their valuable discussions.

The data supply from World Data Center C2 of Kyoto University and NSSDC/NASA are gratefully acknowledged.

References

Baker, D. N., Statistical analyses in the study of solar wind-magnetosphere coupling, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 17–38, Terra Scientific Pub., Tokyo, 1986.

Burton, R. K., R. L. McPherron, and C. T. Russell, An empirical relationship between interplanetary conditions and D_{st}, *J. Geophys. Res.*, **80**, 4204–4214, 1975.

Clauer, C. R., The technique of linear prediction filters applied to studies of solar wind-magnetosphere coupling, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 39–57, Terra Scientific Pub., Tokyo, 1986.

Daglis, I. A., The role of magnetosphere-ionosphere coupling in magnetic storm dynamics, in *Magnetic Storms*, edited by B. T. Tsurutani, W. D. Gonzalez, Y. Kamide, and J. K. Arballo, pp. 107–116, AGU, Washington, D.C., 1997.

Detman, T. R. and D. Vassiliadis, Review of techniques for magnetic storm forecasting, in *Magnetic Storms*, edited by B. T. Tsurutani, W. D. Gonzalez, Y. Kamide, and J. K. Arballo, pp. 253–266, AGU, Washington, D.C., 1997.

Fausett, L., *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, 461 pp., *Prentice Hall*, Englewood Cliffs, NJ 07632, 1994.

Fok, M.-C., J. U. Kozyra, A. F. Nagy, and T. E. Cravens, Lifetime of ring current particles due to Coulomb collisions in the plasmasphere, *J. Geophys. Res.*, **96**, 7861–7867, 1991.

Freeman, J., A. Nagai, P. Reiff, W. Denig, S. Gussenhoven, M. A. Shea, M. Heinemann, F. Rich, and M. Hairston, The use of neural networks to predict magnetospheric parameters for input to a magnetospheric forecast model, in *Proceedings of the International Workshop on Artificial Intelligence Applications in Solar Terrestrial Physics*, edited by J. Joselyn, H. Lundstedt, and J. Trolinger, pp. 167–182, Lund, Sweden, 1993.

Gleisner, H. and H. Lundstedt, Response of the auroral electrojets to the solar wind modeled with neural networks, *J. Geophys. Res.*, **102**, 14269–14278, 1997.

Gleisner, H., H. Lundstedt, and P. Wintoft, Predicting geomagnetic storms from solar wind data using time-delay neural networks, *Ann. Geophys.*, **14**, 679–686, 1996.

Gloeckler, G. and D. C. Hamilton, AMPTEE ion composition results, *Phys. Scr.*, T18, 73–84, 1987.

Hamilton, D. C., 5. Storm dynamics/ring current, in *Magnetic Storms*, edited by B. T. Tsurutani, W. D. Gonzalez, Y. Kamide, and J. K. Arballo, pp. 6–7, AGU, Washington, D.C., 1997.

Hamilton, D. C., G. Gloeckler, F. M. Ipavich, W. Studemann, B. Wilken, and G. Kremser, Ring current development during during the great geomagnetic storm of February 1986, *J. Geophys. Res.*, **93**, 14343–14355, 1988.

Hertz, J., A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, lecture notes vol. 1, Santa Fe Institute Studies in the sciences of complexity, 327 pp., Addison-Wesley, Redwood City, CA 94065, 1991.

Ichikawa, H., *Layered Neural Networks*, 184 pp., Kyoritu Pub., Tokyo, 108, 1993.

Iyemori, T. and H. Maeda, Prediction of geomagnetic activities from solar wind parameters based on the linear prediction theory, in *Solar-Terrestrial Predictions Proceedings*, Vol. 4, U.S. Dept. of Commerce, Boulder, CO, A-1–A-7, 1980.

Iyemori, T., H. Maeda, and T. Kamei, Impulse response of geomagnetic indices to interplanetary magnetic fields, *J. Geomag. Geoelectr.*, **31**, 1–9, 1979.

Kamide, Y., N. Yokoyama, W. Gonzalez, B. T. Tsurutani, I. A. Daglis, A. Brekke, and S. Masuda, Two step development of geomagnetic storms, *J. Geophys. Res.*, **103**, 6917–6921, 1998.

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.

Ogilvie, K. W., L. F. Burlaga, and T. D. Wilkerson, Plasma observations on Explorer 34, *J. Geophys. Res.*, **73**, 6809–6824, 1968.

Pudovkin, M. I., S. A. Zaitseva, and L. Z. Sizova, Growth and decay of magnetospheric ring current, *Planet. Space Sci.*, **33**, 1097–1102, 1985.

Rostoker, G. and C.-G. Fälthammar, Relationships between changes in the interplanetary magnetic field and the variations in the magnetic field at the earth's surface, *J. Geophys. Res.*, **72**, 5853–5863, 1967.

Siscoe, G. L., V. Formisano, and A. J. Lazarus, Relation between geomagnetic sudden impulses and solar wind pressure changes—an experimental investigation, *J. Geophys. Res.*, **73**, 4869–4874, 1968.

Smith, P. H. and N. K. Bewtra, Charge exchange lifetimes for ring current ions, *Space Sci. Rev.*, **22**, 301–318, 1978.

Sugiura, M., Hourly values of equatorial D_{st} for the IGY, *Ann. Int. Geophys. Year*, **35**, 49, 1964.

Vassiliadis, D., A. J. Klimas, D. N. Baker, and D. A. Roberts, A description of solar wind magnetosphere coupling based on nonlinear filters, *J. Geophys. Res.*, **100**, 3495–3512, 1995.

Vassiliadis, D., A. J. Klimas, and D. N. Baker, Nonlinear ARMA models for the D_{st} index and their physical interpretation, paper presented at the Third International Conference on Substorms (ICS-3), Versailles, France, May 12–17, 1996.

Wu, J.-G. and H. Lundstedt, Prediction of geomagnetic storms from solar wind data using Elman recurrent neural networks, *Geophys. Res. Lett.*, **23**, 319–322, 1996.

S. Kugblenu (e-mail: samuel@flare.ee.uec.ac.jp), S. Taguchi, and T. Okuzawa