

Operational forecasts of the geomagnetic *Dst* index

H. Lundstedt

Swedish Institute of Space Physics, Lund, Sweden

H. Gleisner

Danish Meteorological Institute, Copenhagen, Denmark

P. Wintoft

Swedish Institute of Space Physics, Lund, Sweden

Received 22 August 2002; revised 4 October 2002; accepted 29 October 2002; published 24 December 2002.

[1] We here present a model for real time forecasting of the geomagnetic index *Dst*. The model consists of a recurrent neural network that has been optimized to be as small as possible without degrading the accuracy. It is driven solely by hourly averages of the solar wind magnetic field component B_z , particle density n , and velocity V , which means that the model does not rely on observed *Dst*. In an evaluation based on more than 40,000 hours of solar wind and *Dst* data, it is shown that this model has smaller errors than other models currently in operational use. A complete description of the model is given in an appendix. **INDEX TERMS:** 2722 Magnetospheric Physics: Forecasting; 2784 Magnetospheric Physics: Solar wind/magnetosphere interactions; 3220 Mathematical Geophysics: Nonlinear dynamics; 2788 Magnetospheric Physics: Storms and substorms. **Citation:** Lundstedt, H., H. Gleisner, and P. Wintoft, Operational forecasts of the geomagnetic *Dst* index, *Geophys. Res. Lett.*, 29(24), 2181, doi:10.1029/2002GL016151, 2002.

1. Introduction

[2] The *Dst* index is the longitudinally averaged magnetic field depression at low latitudes. It is the primary measure of the magnitude of magnetic storms, and provides a convenient way to monitor the magnetospheric ring current. The first predictions of *Dst* based on neural networks were presented in [Lundstedt, 1991, 1993] and in [Freeman and Nagai, 1993]. In both cases multi-layer feed-forward neural networks were used. The neural networks were trained to predict *Dst* one hour ahead. The models were later improved using more data and studied in detail statistically [Lundstedt and Wintoft, 1994]. In 1996 the first Elman recurrent neural networks were applied [Wu and Lundstedt, 1996], leading to a more accurate description of the recovery phase of geomagnetic storms.

[3] With the launch of the ACE solar wind probe in 1997, and the release of real-time solar wind data measured near the Sun-Earth libration point L1, real-time forecasts became possible. At the same time the Lund Space Weather Model was introduced [Lundstedt, 1998], in which the *Dst* forecast model was included as a module in an intelligent hybrid system. The aim of this hybrid system is to provide tailored forecasts for different users, e.g., power system operators and satellite operators.

[4] ESA initiated the Space Weather Programme Study in 1999. We participated in the consortium led by Alcatel Space, where we developed a prototype forecast service of space weather and effects, using real-time knowledge-based neurocomputing [Lundstedt, 2002]. The Lund operational *Dst* model is included in the prototype. The *Dst* forecast is also offered as a service through the RWC-Sweden of ISES (www.lund.irf.se/rwc/dst) and the forecast is regularly accessed by different users. Several research groups use the forecasted *Dst* values as inputs to their models, e.g., forecasts of the radiation doses astronauts are exposed to [Wuest et al., 2002].

[5] In the next section current operational *Dst* forecast models are described and compared to the Lund *Dst* model. In section 3 we discuss some operational aspects, and section 4 contains the main conclusions. Finally, we give a complete description of the Lund *Dst* model in Appendix A.

2. Current Operational Models

2.1. Representing the Dynamics of the *Dst* Index

[6] The *Dst* index can be regarded as an output signal from a dynamic system driven by an external input, namely the solar wind. A number of empirical models have been suggested as valid representations of the *Dst* dynamics [Burton et al., 1975; Fenrich and Luhmann, 1998; O'Brien and McPherron, 2000; Lundstedt and Wintoft, 1994; Wu and Lundstedt, 1996]. These models represent different approaches to the problem of finding a description of the system dynamics, but they share the property that the model coefficients are empirically determined rather than derived. Although most of these models could, in principle, be used to make forecasts, only a few of them have actually been converted to operational use. We here consider those models that are currently used in real time operations: models based on first-order differential equations and the Lund *Dst* model which is based on a recurrent neural network.

2.2. Models Based on Differential Equations

[7] Burton et al. [1975] described the evolution of the ring current by the first-order differential equation $\frac{dD_{st}^*}{dt} = Q - \lambda D_{st}^*$ where $D_{st}^* = Dst - b\sqrt{P_{dyn}} + c$ is the pressure corrected *Dst* index from which the effects of the magnetopause currents have been removed. This effect is assumed to depend solely on the solar-wind dynamic pressure P_{dyn} , and is described by the constant coefficients b and c . In the differential equation above, Q represents the injection of particles to the ring current, and λD_{st}^* repre-

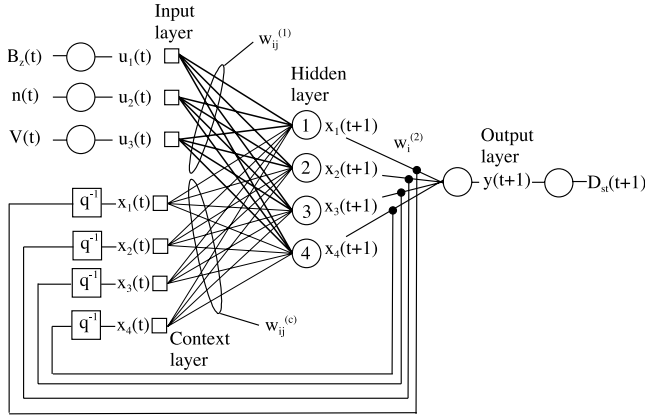


Figure 1. The figure shows the Elman neural network architecture. At the top left we have the B_z , n , and V inputs that are normalized to u_1 , u_2 , and u_3 . The weights connecting the neurons are illustrated with lines. The q^{-1} is the unit delay. The weights connecting from the hidden layer to the unit delays are all set to one (1).

sents the loss of particles with an e -folding time $1/\lambda$. Several Dst models have been derived from this model. In the original paper by *Burton et al.* [1975] the decay time is assumed to be constant ($1/\lambda = 7.72$ hours), while the source term is a linear function of VB_s when the interplanetary electric field E_y exceeds 0.5 mV/m. For smaller values of E_y , the source term vanishes. *Fenrich and Luhmann* [1998] presented a slightly modified version of the *Burton et al.* model in which the source term depends on both VB_s and P_{dyn} , and the decay time is reduced to 3.0 hours when E_y exceeds 4.0 mV/m. Compared to the original model, these modifications give some improvements, particularly for high geomagnetic activity levels. Other modifications to the original model were proposed by *O'Brien and McPherron* [2000], the most important being a decay time continuously varying from 2.4 hours for high activity levels to 19 hours at low geomagnetic activities.

[8] The differential equations are integrated using a forward difference method [*O'Brien and McPherron, 2000*], preceded and followed by magnetopause-current corrections. In discrete time t , and with the time step set to 1 hour, the equations can be written

$$Dst^*(t) = Dst(t) - b\sqrt{P_{dyn}(t)} + c \quad (1)$$

$$Dst^*(t+1) = Dst^*(t) + [Q(t) - \lambda(t)Dst^*(t)] \quad (2)$$

$$Dst(t+1) = Dst^*(t+1) + b\sqrt{P_{dyn}(t+1)} - c \quad (3)$$

According to equations (1) and (2), we can compute $Dst^*(t+1)$ from solar-wind data at time t . To apply the final magnetopause-current correction we would need solar-wind observations at time $t+1$. Modifying equation (5) to

$$Dst(t+1) = Dst^*(t+1) + b\sqrt{P_{dyn}(t)} - c$$

we arrive at a set of equations - equations (3), (4), and (6) - that allows us to forecast $Dst(t+1)$ from solar-wind data observed at time t . However, we use $P_{dyn}(t)$ in equation 3 instead of the more correct $P_{dyn}(t+1)$. The evolution of Dst can thus be predicted solely from solar-wind data with no

dependence on observed Dst . The substitution of $P_{dyn}(t+1)$ with $P_{dyn}(t)$ leads to an increase of the RMS prediction errors, particularly for strong positive Dst disturbances. However, the overall effects of this approximation are small and do not change the general conclusions.

2.3. Model Based on Elman Neural Network

[9] The Elman neural network [*Elman, 1990*] is one type of network that belongs to the class of recurrent neural networks. The network architecture has been found through a standard training and optimization procedure using two separate data sets: 1) training set and 2) validation set [*Haykin, 1994*]. Each set contains approximately 8000 hours of data. A large number of networks, with different architectures, were trained on data from the training set. Then the root-mean-square errors were computed for each network using data from the validation set, and the network with the smallest validation error was selected as the optimal network (Figure 1). The input layer consists of three units that receive the solar wind data. The hidden layer contain four neurons that receive weighted data from the input layer and the context layer. The context layer keeps a copy of the output from the hidden layer at the previous time step. Finally, the output layer, which only has one neuron, is the weighted sum of the hidden layer output.

[10] The input and output are the normalized solar wind data and Dst according to App. A. The output from the network is described by the following equations

$$x_i(t+1) = \tanh\left(\sum_{j=1}^{n_1} w_{ij}^{(1)} u_j(t) + \sum_{j=1}^{n_c} w_{ij}^{(c)} x_j(t) + b_i^{(1)}\right) \quad (4)$$

$$y(t+1) = \sum_{i=1}^{n_2} w_i^{(2)} x_i(t+1) + b^{(2)}. \quad (5)$$

The first sum of equation (4) goes over the three ($n_1 = 3$) external inputs $u_j(t)$ with weights $w_{ij}^{(1)}$. The second sum goes over the four ($n_c = 4$) context inputs $x_j(t)$ with weights $w_{ij}^{(c)}$, where the context input is the output from hidden neuron j at the previous time step. A constant bias $b_i^{(1)}$ is also added at each hidden neuron. The output from the network $y(t+1)$ is the weighted sum of the hidden outputs with the weights $w_i^{(2)}$, bias $b^{(2)}$, and $n_2 = 4$.

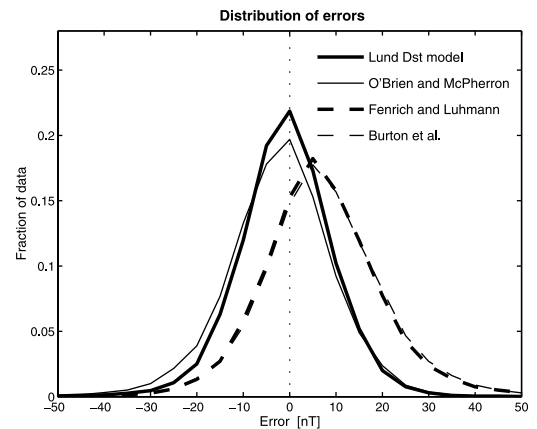


Figure 2. The error distribution of the Lund Dst model, compared to three differential-equation based models.

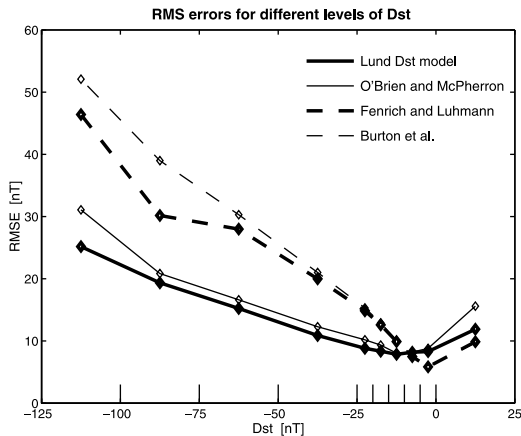


Figure 3. The RMS errors for different levels of *Dst* for the Lund *Dst* model, compared to three differential-equation based models.

[11] Equation (4) describes a non-linear dynamic system where the state $\mathbf{x}(t + 1)$ is determined from the external input $\mathbf{u}(t)$ and past states $\mathbf{x}(t)$, $\mathbf{x}(t - 1)$, \dots . The weights connecting the context layer to the hidden layer ($w_{ij}^{(c)}$ or $\mathbf{W}^{(c)}$ in Appendix A) determine the dynamics of the model. However, the weights $\mathbf{W}^{(c)}$ are not easily interpreted because of the non-linear tanh function. To overcome this, equation (4) may be linearized using a Taylor expansion and rewritten into a set of difference equations. At each time step the coefficients of the linear difference equations change. The coefficients derived from the diagonal elements in $\mathbf{W}^{(c)}$ can now be interpreted as decay terms similar to λ in equation (1). The diagonal element accounting for most of the variance in the output corresponds to a decay rate varying between 2 and 8 hours, with the low values at storm main phase. The interpretation of the network will be further explored in a future paper.

2.4. Evaluation of Model Performances

[12] The statistical performances of the models were evaluated on more than 40,000 hours of solar wind-*Dst* data obtained from the OMNI data set. This large number of data, in relation to the few parameters of the models, ensures that the models' generalization abilities are thoroughly tested. The evaluation is based on 68 periods longer than 10 days, with no data gaps longer than 2 hours, with all periods starting at low *Dst* levels ($-15 \text{ nT} < Dst < 10 \text{ nT}$), and with *Dst* falling below -50 nT at least once during each period. All four models are driven solely by observed solar-wind data with no dependence on observed *Dst*.

[13] From the statistical evaluation we see that the error distributions (Figure 2) are roughly Gaussian. The Lund *Dst* model and the model by O'Brien and McPherron [2000] have a mean close to zero, whereas the other two models tend to over-estimate *Dst*. The RMS errors for different *Dst* levels (Figure 3) demonstrate a similar division between the older Burton et al. type models on the one hand, and the O'Brien-McPherron and Lund *Dst* models on the other hand. The largest differences between the models occur for enhanced ring-current conditions, and become particularly significant during magnetic storms at times when the *Dst* dynamics is dominated by ring-current decay. For small, or positive, *Dst* levels all models tend to underestimate *Dst*.

The linear correlation between predicted and observed *Dst* computed over the more than 40,000 hours of data are 0.88 for the Lund *Dst* model, and 0.84, 0.79, and 0.78, respectively, for the other three models.

3. Operational Aspects

3.1. Simulation vs. Extrapolation

[14] For a model to be regarded as a valid representation of the *Dst* dynamics, we normally require the predicted *Dst* to closely follow the observed *Dst* when any remaining influences from the model initialization have become negligible. This is sometimes referred to as out-of-sample prediction. However, an alternative in real-time operations is to actually let the models accommodate observed *Dst* data, and to use the model for extrapolation. We have to make a distinction between these two modes of operation: simulation (when the model is driven solely by observed solar-wind data, and with negligible influence from any observed *Dst* data) and *extrapolation* (when the model also accommodates observed *Dst*). Only the simulation mode of operation is dealt with here.

3.2. Data Averaging and Forecast Lead Time

[15] We have used 1 h averaged data to maintain consistency with earlier studies. A consequence of the 1-hour averaging is that a forecast of *Dst* valid for a certain time interval $[t, t + 1.0]$ can not be issued until time t , at the end of the preceding averaging interval. However, even though observed 1-hour averaged data are not fully known until the end of the corresponding averaging intervals, they are partially known before that. By using approximate averages of the solar-wind data, we could increase the forecast lead time so that at some point during the time interval $[t - 1.0, t]$ we can begin to issue preliminary forecasts of *Dst* valid for the interval $[t, t + 1.0]$.

[16] In operational forecasting we get an additional forecast lead time due to the solar-wind propagation from L1 to the Earth. Depending on the solar-wind speed the forecast lead time increases with 25 to 75 min.

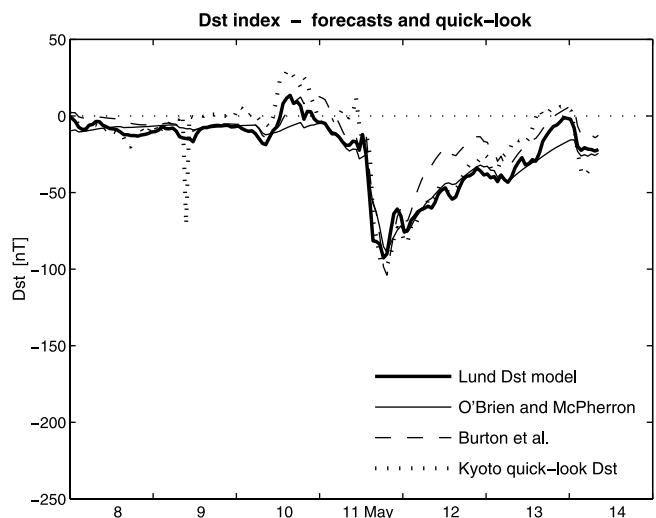


Figure 4. Operational forecasts issued in the evening of May 14, 2002. Predicted *Dst* from the four models are shown along with the Kyoto quick-look data.

3.3. Operational Forecasts and Quick-Look *Dst*

[17] Figure 4 shows operational forecasts issued in the evening of May 14, 2002. The *Dst* index as predicted by the four models are shown together with the latest quick-look *Dst* obtained from the World Data Center in Kyoto. Such quick-look data can be used for rapid, qualitative assessments of the forecast accuracy. They can also be used to drive an operational model in extrapolation mode (see section 3.1). However, if quick-look *Dst* is used to initiate an operational *Dst* model, errors like those that occur on May 9 (Figure 4) have to be identified and dealt with. In the present study all *Dst* models are run in simulation mode and, as evident in Figure 4, the quick-look *Dst* does not influence the *Dst* forecasts. Note that the similarities between the Lund and the O'Brien-McPherron models that appeared in the statistical evaluation are evident also here.

4. Conclusions

[18] The model presented here makes one hour predictions of the *Dst* index using only solar wind data as input. Being primarily an operational model, emphasis has been placed on simplicity and ease of implementation, which are also important properties when interpreting model coefficients in terms of system dynamics. Still, the model is close to optimal as judged by more than 40,000 hours of solar wind and *Dst* data. It has smaller prediction errors than other models currently in operational use, particularly during magnetic storms and for enhanced ring-current conditions.

Appendix A: The Lund *Dst* Model

[19] Here we describe the algorithm to compute the one hour forecast of *Dst* from solar wind data. We assume that we have contiguous observations of B_z , n , and V with one hour cadence. The algorithm is described using vectors and matrices, and T denotes the transpose. When the model is applied the first time the internal state $\mathbf{x} = (x_1, x_2, x_3, x_4)^T$ is not known. We therefore set it to the average value computed from our 40,000 hour data set. Depending on where in the storm the model is first initiated it takes from less than an hour up to about 30 hours for the model to converge. Matlab and Java implementations can be found at www.lund.irf.se/rwc/dst together with some examples.

1. Set $\mathbf{x} = (0.39, 0.22, -0.83, 0.14)^T$.
2. Get the first observation of B_z , n , and V .
3. Compute $\mathbf{u} = (u_1, u_2, u_3)^T = \left(\frac{B_z}{30}, \frac{n}{80} - 1, \frac{v}{400} - 1.5\right)^T$.
4. Compute $\mathbf{x} \leftarrow \tanh(\mathbf{W}^{(1)}\mathbf{u} + \mathbf{W}^{(c)}\mathbf{x} + \mathbf{b}^{(1)})$.
5. Compute $y \leftarrow \mathbf{w}^{(2)}\mathbf{x} + b^{(2)}$.
6. Compute *Dst* from $Dst = 150y - 100$.
7. Get the next observation of B_z , n , and V .
8. Go back to step 3.

The weights and biases are

$$\mathbf{W}^{(1)} = \begin{pmatrix} -0.5946 & 0.0904 & 0.0428 \\ 0.6215 & 0.4485 & 0.2516 \\ -1.4217 & 0.4988 & 0.1409 \\ 0.3372 & -0.5065 & 0.7190 \end{pmatrix}$$

$$\mathbf{W}^{(c)} = \begin{pmatrix} 0.6099 & 0.2223 & -0.4379 & 0.0053 \\ 0.7935 & 0.5000 & -1.2654 & -0.0428 \\ -0.3496 & 0.1024 & 0.3605 & 0.1790 \\ -0.2067 & 0.2270 & 0.5446 & 0.1019 \end{pmatrix}$$

$$\mathbf{b}^{(1)} = \begin{pmatrix} -0.1081 \\ -0.7252 \\ -0.2307 \\ 0.5406 \end{pmatrix}$$

$$\mathbf{w}^{(2)} = (0.2535 \quad 0.3172 \quad -0.4860 \quad -0.2346)$$

$$\mathbf{b}^{(2)} = \mathbf{0.0712}$$

[20] **Acknowledgments.** We would like to thank the ACE spacecraft team and SEC for making the real-time solar wind data available, and the NSSDC for making the OMNI dataset available. This work was supported by an ESA Research Fellowship grant (Hans Gleisner, DMI), and by the Swedish National Space Board (Swedish Institute of Space Physics, Lund).

References

- 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.
- Elman, J. L., Finding structure in time, *Cognitive Sci.*, *14*, 179–211, 1990.
- Fenrich, F. R., and J. G. Luhmann, Geomagnetic response to magnetic clouds of different polarity, *Geophys. Res. Lett.*, *25*, 2999–3002, 1998.
- Freeman, J., and A. Nagai, The Magnetospheric Specification and Forecast Model: Moving from Real-Time to Prediction, *Solar-Terrestrial Predictions IV, Proc. of a Workshop at Ottawa, Canada, May 18–22, 1992*, vol. 2, edited by J. Hruska, M. A. Shea, D. F. Smart, and G. Heckman, NOAA, Boulder, Colorado, 524–539, 1993.
- Haykin, S., *Neural Networks - A Comprehensive Foundation*, Macmillan College Publ. Comp., Inc., New York, 1994.
- Lundstedt, H., Neural network predictions of geomagnetic activity, in *AGA Programs and Abstracts*, XX General Assembly, IUGG, Vienna, 1991.
- Lundstedt, H., A Trained Neural Network, Geomagnetic Activity and Solar Wind Variation, *Solar-Terrestrial Predictions IV, Proc. of a Workshop at Ottawa, Canada, May 18–22, 1992*, vol. 2, NOAA, Boulder, Colorado, 607–610, 1993.
- Lundstedt, H., Lund Space Weather Model: Status and Future Plans, in *Proceedings of the 2nd Workshop on AI Applications in Solar-Terrestrial Physics, July 29–31, 1997, Lund, Sweden*, ESA WPP-148, Paris, 107–112, 1998.
- Lundstedt, H., Forecasting Space Weather and Effects Using Knowledge-Based Neurocomputing, in *Proceedings of ESA Workshop on Space Weather: Looking Towards a European Space Weather Programme, Dec 17–19, 2001, ESTEC, Noordwijk, the Netherlands*, ESA, in press, 2002.
- 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.
- O'Brien, T. P., and R. L. McPherron, Forecasting the ring current index *Dst* in real time, *J. Atmos. Solar-Terr. Phys.*, *62*, 1295–1299, 2000.
- 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.
- Wüest, M., R. A. Frahm, J. K. Jennings, and J. R. Shaber, Predicting electron precipitation based on geomagnetic condition, presented at Space Weather Week, SEC, Boulder, April 16–19 2002.

H. Lundstedt, Swedish Institute of Space Physics, Scheelevägen 17, SE-22370 Lund, Sweden. (henrik@lund.irf.se)

H. Gleisner, Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, Denmark. (hgl@dmu.dk)

P. Wintoft, Swedish Institute of Space Physics, Scheelevägen 17, SE-22370 Lund, Sweden. (peter@lund.irf.se)