Pergamon

Phys. Chem .Earth (C), Vol. 25, No. 4, pp.275-280, 2000

1464-1917/00/$ - see front matter

PII: S1464-1917(00)00016-7

Real Time Kp Predictions from Solar Wind Data using Neural Networks

Fredrik Boberg[1], Peter Wintoft[2] and Henrik Lundstedt[2]

[1]Lund Observatory, Box 43, SE-221 00 Lund, Sweden
[2]Swedish Institute of Space Physics, Solar-Terrestrial Physics Division, Scheelevägen 17, SE-223 70 Lund, Sweden

Received 13 May 1999; accepted 20 September 1999

Abstract. Multilayer feed-forward neural network models are developed to make three-hour predictions of the planetary magnetospheric Kp index. The input parameters for the networks are the B_z-component of the interplanetary magnetic field, the solar wind density n, and the solar wind velocity V, given as three-hour averages. The networks are trained with the error back-propagation algorithm on data sequences extracted from the 21^{st} solar cycle. The result is a hybrid model consisting of two expert networks providing Kp predictions with an RMS error of 0.96 and a correlation of 0.76 in reference to the measured Kp values. This result can be compared with the linear correlation between $V(t)$ and $Kp(t + 3$ hours) which is 0.47. The hybrid model is tested on geomagnetic storm events extracted from the 22^{nd} solar cycle. The hybrid model is implemented and real time predictions of the planetary magnetospheric Kp index are available at `http://www.astro.lu.se/~fredrikb`.

1 Introduction

The interaction between the solar wind and the Earth's magnetic field generates electric currents in the ionosphere and magnetosphere. These currents produce geomagnetic disturbances that are measured at ground-based magnetic observatories recording the three magnetic field components. The mid-latitude observatories measure the amplitude of the variation of the horizontal components X and Y. The amplitude is measured in nT's and each observatory expresses the most disturbed value of a three-hourly range by the index K. The K index is given on a quasi-logarithmic scale from 0 (< 5 nT) to 9 (> 500 nT).

The planetary magnetospheric Kp index (Mayaud, 1980) is evaluated using the K indices obtained from 13 selected subauroral stations situated at geomagnetic latitudes between 48° and 63°. The three-hourly K values are first corrected for the station's geomagnetic latitude, since the geomagnetic activity is latitude-dependent, and then averaged to produce the three-hour planetary magnetospheric Kp index. The Kp index is given on a scale from 0 to 9 expressed in thirds (with a total of 28 values):

$$0_o, 0_+, 1_-, 1_o, 1_+, \ldots, 8_-, 8_o, 8_+, 9_-, 9_o$$

where 0 means very quiet conditions and 9 a very high activity. The measured Kp values obtained from the 13 ground based magnetic observatories are not available until about two months from present, while there is a preliminary Kp index based on a subset of the observatories with about a 6 hour time lag.

Because of the time resolution and the geographic positions of the observatories the Kp index can only serve as an overall measure of geomagnetic activity. The different storm phases during solar maximum (Gosling et al., 1991) and during the declining phase of the solar cycle (Tsurutani et al., 1995) are thus not resolved in detail using the Kp index. However, Kp is often used in solar-terrestrial studies.

The Lund Space Weather Model (Lundstedt, 1999) is used in the development of neural network models making real time predictions of the space weather and its effects. The model contains several different modules for the prediction of: the daily average solar wind velocity from solar magnetic field observations (Wintoft and Lundstedt, 1999); the daily Ap index from solar data and previous Ap index (Lundstedt, 1992); the hourly storm index Dst from solar wind data (Wu and Lundstedt, 1997); the 10 minute average auroral-electrojet index AE also from solar wind data (Gleisner and Lundstedt, 1997); the plasma frequency foF2 from past values of foF2 and AE (Wintoft and Cander, 1999). The construction of a neural network model to make real time three-hour predictions of the planetary magnetospheric Kp is another module in the Lund Space Weather Model. This construction is of great importance since Kp enters into several models that are intended for real time operation. Linear filters (Nagai, 1988) and neural networks (Koons and Gorney, 1991) have been developed for the prediction of energetic electron fluxes at geosynchronous orbit based on Kp. Neural

Correspondence to: F. Boberg

networks, specialized in predicting increased risks of anomalies for the Meteosat-3 and the Tele-X satellites, have also been developed using Kp as input (Wu et al., 1999). The Kp index can also be used to determine the probability of high levels of geomagnetically induced currents (GIC's) in power grids. Boteler et al. (1990) showed that there is a relation between a 3-hour GIC range value (Im) and Kp that follows the relation

$$Im = 0.319\,\mathrm{e}^{0.536\,Kp}.$$

Thus, there is a great need for real time Kp.

2 Neural Network Model

The neural network model used is a multilayer feed-forward network consisting of one input layer, one hidden layer, and one output layer. Time sequences of the three solar wind parameters (B_z, n, and V) are given as input data. The activation function for the hidden nodes is a hyperbolic tangent function whereas a linear combiner is used for the output node. The resulting network output is a prediction of the Kp index for the following three-hour period.

The network training utilizes the error back-propagation algorithm where the actual response $a(n)$ of the network, when presenting the n^{th} training example, move closer to the applied desired response $d(n)$ in a statistical sense by minimizing the summed square error

$$\mathcal{E} = \sum_{n=1}^{N} [d(n) - a(n)]^2, \tag{1}$$

where N denotes the total number of examples contained in the training set.

The objective of the learning process is to adjust the free parameters (i.e. the synaptic weights and biases) connecting the different layers of the network to minimize $\mathcal{E}$. The weight update is performed after the presentation of all the training examples that constitute an epoch. The weight correction applied to the synaptic weight connecting neuron i to neuron j at epoch s is defined by the generalized delta rule:

$$\Delta w_{ji}(s) = -\eta \frac{\partial \mathcal{E}(s)}{\partial w_{ji}(s)} + \alpha \Delta w_{ji}(s-1). \tag{2}$$

The gradient $-\frac{\partial \mathcal{E}(s)}{\partial w_{ji}(s)}$ determines the direction of search in weight space and η is the learning rate parameter. To accelerate descent in steady downhill directions and stabilize the descent in directions that oscillate in sign, a momentum term is included consisting of a momentum constant α and the weight update from the previous epoch (Haykin, 1999).

One measure describing the network accuracy is the root-mean-square error defined as

$$RMSE = \sqrt{\frac{\mathcal{E}}{N}}. \tag{3}$$

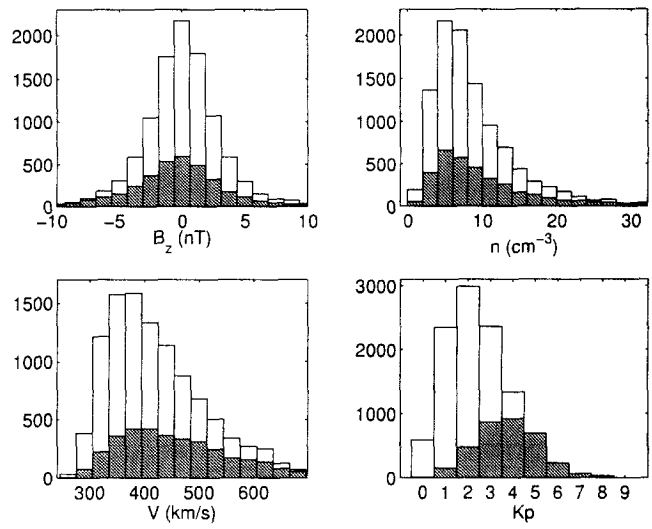

Fig. 1. Comparison between the full set (10460 examples represented by white bars) and the storm period set (3458 examples given as grey bars) extracted from the 21st solar cycle. The storm period set contain the majority of representative solar wind examples causing geomagnetic storms, i.e. examples with negative B_z and high velocities.

Another measure is the correlation coefficient C given by

$$C = \frac{1}{N\sigma_a\sigma_d} \sum_{n=1}^{N} [(a(n) - \langle a \rangle)(d(n) - \langle d \rangle)], \tag{4}$$

where $\langle a \rangle$ and $\langle d \rangle$ are the actual and desired network output averages; σ_a and σ_d are their respective standard deviation. We thus obtain a low $RMSE$ and a correlation close to 1 when making accurate predictions.

3 Data Preparation

The three solar wind parameters B_z, n, and V together with the planetary magnetospheric Kp index for solar cycles 21 and 22 were downloaded from the *National Space Science Data Center* (NSSDC). The 21st solar cycle contain 10 years of data (1976 to 1985) and the 22nd cycle contain 11 years of data (1986 to 1996). Data gaps in the hourly solar wind parameters of up to 3 consecutive hours were filled in with linear interpolation. The network input parameters were then averaged into three-hour values to attain uniformity with the network output parameter Kp.

3.1 Training And Validation Sets

A network trained with the back-propagation algorithm will bias its performance so that the most common input structures will be accurately predicted while predictions for other structures will be poor. This means that storm events, which are far outnumbered by quiet periods, will not be accurately predicted. Therefore, two different sets are extracted from the 21st solar cycle. The two sets are compared in Fig. 1.

The first set, represented by white bars, contains all examples for solar cycle 21. This set contains a majority of low

Kp examples, with a peak at $Kp = 2$ as shown in the lower right panel, and will thus be used in the construction of an expert network A specialized in making predictions during geomagnetically quiet periods.

The second set, given by grey bars in Fig. 1, contains storm periods. We define a storm period as a continuous period with one or more events with $Kp \geq 5_-$ where the spacing between individual events is 12 hours or less. The 12 hours before the first event and the 12 hours after the last event are also added to the period. This means that one storm period contains at least 27 hours of data, with at least one three-hour period with $Kp \geq 5_-$. Typically, a storm period contains several events with $Kp \geq 5_-$ and extends over 45 hours. With this definition of a storm period we generally capture the complete evolution of individual storms. This second set contains a majority of high Kp examples with a peak at $Kp = 4$. This set will subsequently be used when constructing an expert network B specialized in making predictions during geomagnetic storm periods.

Each set is further divided into two equal subsets: a training set and validation set. The training set is used as network input/output during the training session and the validation set is used during the subsequent validation session to find the optimal network architecture.

3.2 Test Set

The averaged data for the 22nd solar cycle will be used to test the final optimized network chosen. In this test geomagnetic storm events will be compared with predictions. The test set is hence kept as a temporal sequence.

4 Previous Work

Before we proceed with the network training it is interesting to study previous work related to Kp predictions. A number of empirical studies have been made attempting to find correlations between solar wind parameters and magnetospheric activity represented by the Kp index. The studies have been focused mainly on nowcasting, in contrast to predictions, and a frequently used relation is of the form

$$\sum Kp = \frac{\langle V \rangle - C_1}{C_2}, \tag{5}$$

where $\sum Kp$ is the sum of the 8 daily Kp values and $\langle V \rangle$ is the daily average of the solar wind speed given in km/s. Different values of the constants C_1 and C_2 have been presented. Snyder et al. (1968) found that $C_1 = 262$ km/s and $C_2 = 6.3$; Hargreaves (1992) presented $C_1 = 330 \pm 17$ km/s and $C_2 = 8.44 \pm 0.74$.

Pudovkin et al. (1980) included daily means of the IMF parameters B_z and σ_B, given in nT's, together with the daily mean of V given in 10^2 km/s. They evaluated the relationship

$$\sum Kp = 5.1 + (1.03 \pm 0.03)(\langle V \rangle \langle \sigma_B \rangle) - (0.51 \pm 0.03)(\langle V \rangle \langle B_z \rangle). \tag{6}$$

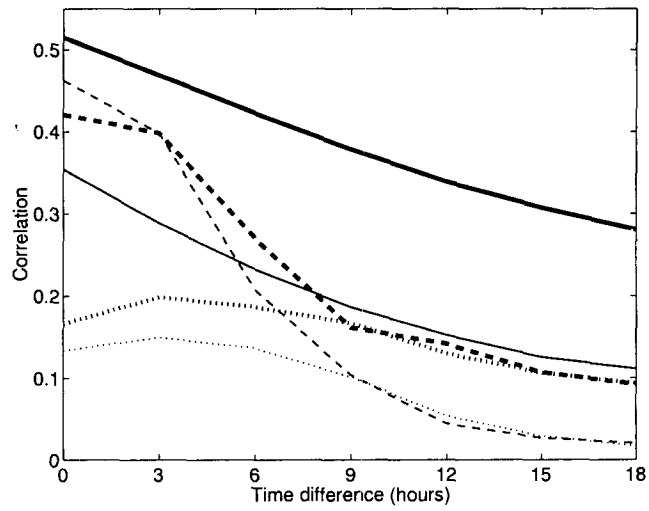

Fig. 2. Linear correlation between the Kp index and the three solar wind parameters as a function of time difference. The $[Kp, V]$ correlations are given as a solid lines, the $[Kp, -B_z]$ correlation as dashed lines, and the $[Kp, n]$ correlations with dotted lines. The thick lines mark the correlations for the full set and the thin lines represent correlations using only the storm period set (cf. Fig. 1).

A different approach was carried out by Ballif et al. (1969). They found, using data observed by Mariner 4, the empirical relationship

$$\langle Kp \rangle = 9 \left[1 - \mathrm{e}^{-(\langle \sigma B_{T,N} \rangle - 0.35)/7.70} \right], \tag{7}$$

where $\langle Kp \rangle$ is the daily average and $\langle \sigma B_{T,N} \rangle$ is a measure of the daily average transverse fluctuations of the interplanetary field.

Finally, we can also examine the linear correlations between the solar wind parameters and Kp for the two sets described in the previous section (see Fig. 2). The correlations between $[B_z(t), n(t), V(t)]$ at t hours and $Kp(t+3)$ are $[-0.40, 0.20, 0.47]$ for the full set and $[-0.40, 0.15, 0.29]$ for the storm period set. The relatively high correlation between the solar wind velocity and Kp for the full set is due to the majority of low activity periods where V and Kp are both relatively low and stable. These correlations are purely linear and the aim of the neural network training procedure is to extract non-linear correlations between the solar wind parameter signatures and the Kp index.

5 Kp Prediction

5.1 Network Training And Optimization

The input layer of the neural network consists of solar wind parameter units and the output layer consists of only one unit, the predicted Kp value. To construct an optimized neural network model a training procedure is performed. In this procedure neural networks with different lengths of input data sequences, i.e. different number of input units per solar wind parameter, and different number of hidden units are

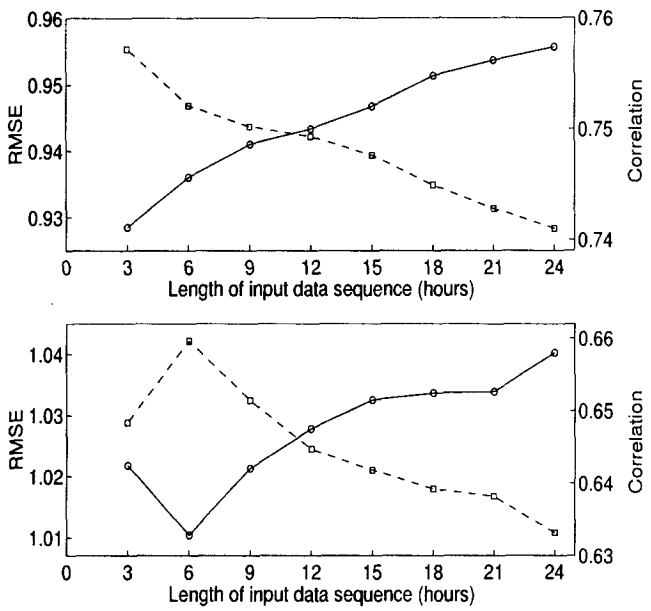

Fig. 3. Validation RMS errors (solid lines) and correlation coefficients (dashed lines) as a function of input data sequence length for neural networks trained and validated on data from the 21st solar cycle. The top panel displays the results using the full validation set and the bottom panel displays the results using the storm period validation set.

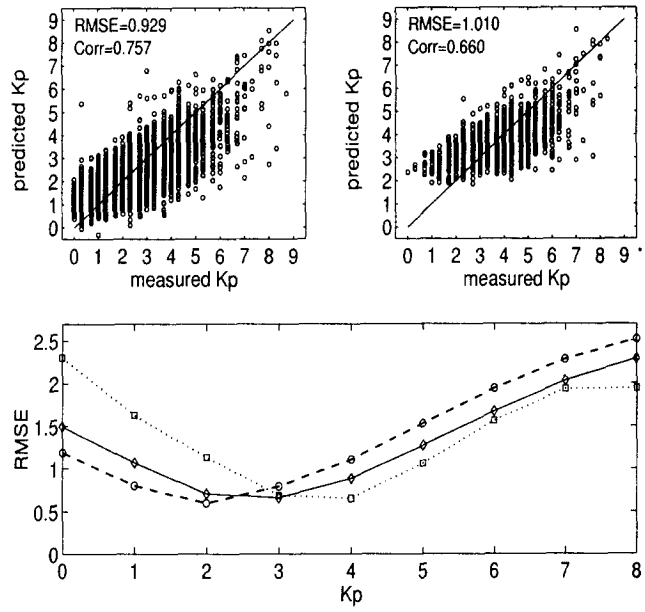

Fig. 4. Comparison between the predictions made with the optimized network A using the full validation set (top left panel and dashed line in bottom panel) and the predictions made with the optimized network B using the storm period validation set (top right panel and dotted line in bottom panel). The solid line in the bottom panel shows the final prediction accuracy when presenting the full validation set to the constructed hybrid model C.

trained using the error back-propagation algorithm. Each trained network is then validated using the validation set. Based on these validation results, an optimized network is obtained. This procedure is first executed using the full set in an attempt to find a network A specialized in making predictions during geomagnetically quiet periods. The procedure is then executed using the geomagnetic storm period set in an attempt to find a network B specialized in making predictions during geomagnetic storm periods.

The optimum number of hidden units is found to be 10 for both network A and B. Figure 3 presents the $RMSE$'s (solid lines) and correlation coefficients (dashed lines) as a function of input data sequence length obtained using 10 hidden units. The top panel displays the validation results using the full set showing that the most accurate three hour prediction of the Kp index is obtained with only the present solar wind data as input. The optimal network A is thus unable to correlate solar wind signatures of the past with the present to attain higher accuracy in the Kp predictions. This result is explained by the fact that the training set contains a majority of low activity examples where the solar wind parameters only influence the present Kp index and not the future indices.

The bottom panel in Fig. 3 displays the validation results using the storm period set and the most accurate predictions are here obtained using a network with a six hour time delay, i.e. with two inputs per parameter. The ability to correlate solar wind signatures of the past with the present is now possible for the optimal network B. This is due to a training set containing mainly storm examples where the solar wind parameters influence both present and future Kp indices.

5.2 Network Construction And Testing

The predictions made with the two optimized neural networks are presented in Fig. 4. As indicated in the bottom panel, network A makes the most accurate low Kp predictions and network B makes the most accurate high Kp predictions. A hybrid model C is constructed by combining these two optimized networks. The final output from model C is the weighted average

$$Kp = \frac{Kp_A/e_A^2 + Kp_B/e_B^2}{1/e_A^2 + 1/e_B^2}, \tag{8}$$

where Kp_A and Kp_B are the predicted Kp from networks A and B, respectively, and the errors e_A and e_B are taken as the respective $RMSE$'s interpolated from the curves in Fig. 4. The resulting prediction using this hybrid model is indicated by the solid line in the bottom panel in Fig. 4.

Finally, the hybrid model is tested on geomagnetic storm events extracted from the 22nd solar cycle. The results are displayed in Fig.s 5 and 6: The discrepancy between measured and predicted Kp at hour 39 in Fig. 5 might be explained by an influence of the negative (-5 nT) interplanetary B_y component (Gleisner and Lundstedt, 1997) for this period. The discrepancy for the storm period between 51 and 60 hours is more difficult to explain. An interesting remark though is that a similar storm course appears 24 hours later as well as 48 hours later (not shown).

The final results from the neural network training/validation session and the test session are summarized in Table 1. Although the $RMSE$ is optimal using network A, we are in-

Table 1. Final results obtained from the training/validation session and the test session. All results are obtained using the full data set except for the training/validation results for network B where the storm data set is used. The results for network model C given in brackets are obtained presenting the training/validation sets to the existing model.

		Training		Validation		Test	
Network	Time delay (hours)	*RMSE*	Correlation	*RMSE*	Correlation	*RMSE*	Correlation
A	3	0.920	0.740	0.929	0.757	0.900	0.765
B	6	1.058	0.641	1.010	0.660	1.220	0.764
C	3 and 6	(0.947)	(0.745)	(0.956)	(0.764)	0.985	0.768

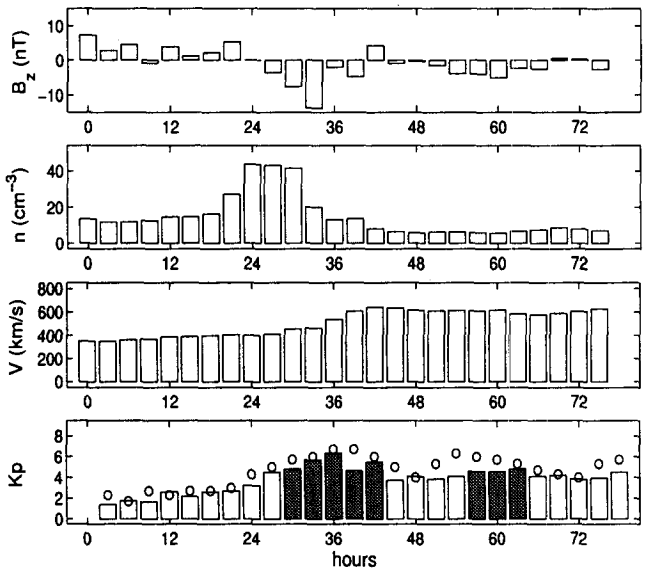

Fig. 5. Solar wind input data and Kp indices for a geomagnetic storm event recorded by the IMP-8 satellite the 26^{th} of April, 1989. The predicted Kp indices are given by bars, and these are shaded when the geomagnetic activity is predicted to be high ($Kp \geq 5_-$). The measured Kp indices are given as circles.

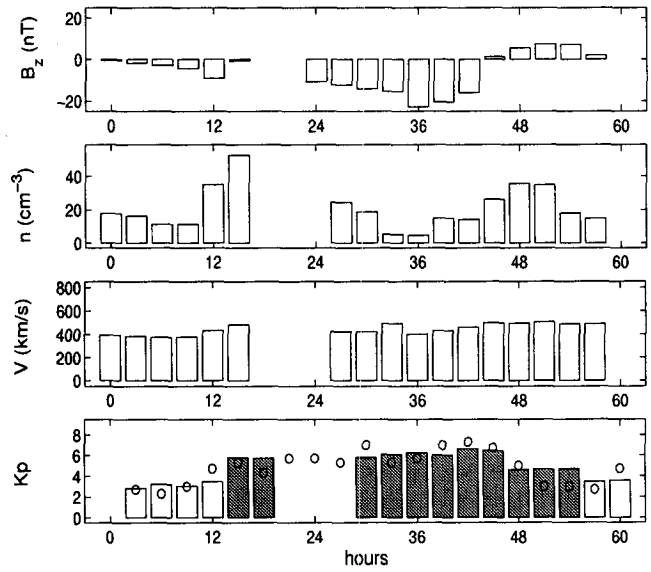

Fig. 6. Solar wind input data and Kp indices for a geomagnetic storm event recorded by the IMP-8 satellite the 21^{st} of February, 1992. The predicted Kp indices are given by bars, and these are shaded when the geomagnetic activity is predicted to be high ($Kp \geq 5_-$). The measured Kp indices are given as circles.

terested in making accurate geomagnetic storm predictions and this is achieved using the hybrid model C. Note that the training and validation results for model C in Table 1 are obtained presenting the data sets to the existing hybrid model.

5.3 Real Time Predictions

The final step is the real time implementation of the hybrid model C. The latest solar wind data are first downloaded from the NOAA's Space Environment Center in Boulder. These data are then presented to the hybrid model. The resulting network output obtained using Eq. (8), together with adherent solar wind data, are presented at `http://www.astro.lu.se/~fredrikb`. This real time prediction is updated every three hours. To illustrate the accuracy of these predictions, the latest preliminary Kp values are downloaded and compared with the predictions. These preliminary values are not final and not up to date, but still a reliable measure of our predictions.

6 Conclusions

The results obtained in Sect. 5 indicate that it is difficult to evaluate a single network making accurate Kp predictions over the whole index interval from 0 to 9. This difficulty is resolved by merging two expert networks into one hybrid model. The first network is specialized on making low Kp predictions and the second network on making high Kp predictions. The resulting hybrid model makes Kp predictions with an accuracy indicated by the solid line in the bottom panel in Fig. 4 and by the bottom line in Table 1. The network testing examples in Fig. 5 and Fig. 6 illustrate this prediction accuracy. The hybrid model is implemented and real time predictions of the planetary magnetospheric Kp index are available at `http://www.astro.lu.se/~fredrikb`.

Acknowledgements. The staff at NOAA's Space Environment Center in Boulder and the ACE project teams are gratefully acknowledged for making the solar wind data available in real time.

References

Ballif J. R., Jones D. E., and Coleman Jr P. J., Further evidence on the correlation between transverse fluctuations in the interplanetary magnetic field and *Kp*, *J. Geophys. Res., 74,* 2289-2300, 1969.

Boteler D. H., Watanabe T., and Butler D. B., Prediction of geomagnetically induced current levels in the B.C. Hydro 500 kV system, *Solar-Terrestrial Predictions:* Proc. of a Workshop at Leura, Australia, October 16-20, 1989, Vol. 2, 30 - 41, 1990.

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

Gosling J. T., McComas D. J., Philips J. L., and Bame S. J., Geomagnetic activity associated with earth passage of interplanetary shock disturbances and coronal mass ejections, *J. Geophys. Res., 96,* 7831-7839, 1991.

Hargreaves J. K., *The solar-terrestrial environment*, ch. 8, Cambridge University Press, 1992.

Haykin S., *Neural Networks - a comprehensive foundation*, ch. 4, Prentice-Hall, 1999.

Koons H. D. and Gorney D. J., A neural network model of the relativistic electron flux at geosynchronous orbit, *J. Geophys. Res., 96,* 5549-5556, 1991.

Lundstedt H., Neural networks and predictions of solar-terrestrial effects, *Planet. Space Sci. 40,* 457-464, 1992.

Lundstedt H., The Swedish Space Weather Initiatives, in the Proceedings of 'Workshop on Space Weather', November 11-13, 1998 in Noordwijk, Netherlands, *ESA WPP-155*, 1999

Mayaud P. N., *Derivation, Meaning, and Use of Geomagnetic Indices*, ch. 5, American Geophysical Union, 1980.

Nagai T., Space weather forecast: Predictions of relativistic electron intensity at synchronous orbit, *Geophys. Res. Lett., 15,* 425-428, 1988.

Pudovkin M. I., Shukhtina M. A., Ponyavin D. I., Zaitseva S. A., and Ivanov O. U., On the geoefficiency of the solar wind parameters, *Ann. Géophys., 36,* 549-553, 1980.

Snyder C. W., Neugebauer M., and Rao U. R., The solar wind velocity and its correlation with cosmic-ray variations and with solar and geomagnetic activity. *J. Geophys. Res., 68,* 6361, 1963.

Tsurutani B. T., Gonzales W. D., Gonzales A. L. C., Tang F., Arballo J. K., and Okada M., Interplanetary origin of geomagnetic activity in the declining phase of the solar cycle, *J. Geophys. Res., 100,* 717-733, 1995.

Wintoft P., and Lundstedt H., A neural network study of the mapping from solar magnetic fields to the daily average solar wind velocity, *J. Geophys. Res. 104,* 6729 - 6736, 1999.

Wintoft P. and Cander Lj. R., Ionospheric foF2 storm forecasting using neural networks, This issue, 1999.

Wu J.-G. and Lundstedt H., Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks, *J. Geophys. Res., 102,* 14255-14268, 1997.

Wu J.-G., Lundstedt H., Eliasson L., Andersson L., and Norberg O., Spacecraft anomaly forecasting using non-local environment data, Study of plasma and energetic electron environment and effects, *ESA Tech. Note*, ESTEC/Contract No. 11974/96/NL/JG(SC), 1999.