

A neural network Dst index model driven by input time histories of the solar wind-magnetosphere interaction

M. Revallo*

*Geophysical Institute, Slovak Academy of Sciences,
Dúbravská cesta 9, 845 28 Bratislava, Slovak Republic*

F. Valach

*Geomagnetic Observatory, Geophysical Institute, Slovak Academy of Sciences,
Komárňanská 108, 947 01 Hurbanovo, Slovak Republic*

P. Hejda, J. Bochníček

*Institute of Geophysics, Academy of Sciences of the Czech Republic,
Boční II/1401, 141 31 Prague 4, Czech Republic*

Abstract

A model to forecast one-hour lead Dst index is proposed. Our approach is based on artificial neural networks (ANN) combined with an analytical model of the solar wind-magnetosphere interaction. Previously, the hourly solar wind parameters have been considered in the analytical model, all of them provided by registration of the ACE satellite. They were the solar wind magnetic field component B_z , velocity V , particle density n and temperature T . The solar wind parameters have been used to compute analytically the discontinuity in magnetic field across the magnetopause, denoted as $[B_t]$. This quantity has been shown to be important in connection with ground

*Corresponding author

Email addresses: geofmire@savba.sk (M. Revallo), fridrich@geomag.sk (F. Valach), ph@ig.cas.cz, jboch@ig.cas.cz (P. Hejda, J. Bochníček)

Preprint submitted to Elsevier

January 17, 2014



magnetic field variations. The method was published, in which the weighted sum of a sequence of $[B_t]$ was proposed to produce the value of Dst index. The maximum term in the sum, possessing the maximum weight, is the one denoting the contribution of the current state of the near-Earth solar wind. The role of the older states is less important – the weights exponentially decay. Moreover, the terms turn to zero if $B_z \geq 0$. In this study, we set up a more comprehensive model on the basis of the ANNs. The model is driven by input time histories of the discontinuity in magnetic field $[B_t]$, which are provided by the analytical model. At the output of such revised model, the Dst index is obtained and compared with the real data records. In this way we replaced those exponential weights in the published method with another set of weights determined by the neural networks. We retrospectively tested our models with real data from solar cycle 23. The ANN approach provided better results than a simple method based on exponentially decaying weights. Moreover, we have shown that our ANN model could be used to predict Dst one hour ahead. We assessed the predictive capability of the model with a set of independent events and found correlation coefficient $CC = 0.74 \pm 0.13$ and prediction efficiency $PE = 0.44 \pm 0.15$. We also compared our model with the so called Dst -specification models. In those models, the Dst index was derived directly through an analytic or iterative formula or a neural network-based algorithm. We showed that the performance of our model was comparable to that of Dst -specification models.

Keywords: solar wind, magnetosphere, geomagnetic storm, Dst index, artificial neural network

1. Introduction

Nowadays, there is an increasing demand to understand and predict conditions in the near-Earth space driven by the solar activity. Global magnetohydrodynamic (MHD) computational models based on first principles (Baker et al., 2004; Gombosi et al., 2001; Goodrich et al., 2004; Odstrčil et al., 2004; Siscoe et al., 2004; Tóth et al., 2012; Tsyganenko, 2013) are some examples of major thrusts in this effort. On the other hand, the use of empirical models for the purposes of forecasting has the advantage of being less computationally demanding than the MHD models. The goal is to develop short-term models which can take into account the observed features of the solar wind-magnetosphere interaction while being computationally simple and possessing real-time forecasting capability.

In Alexeev and Feldstein (2001), the dynamic paraboloid magnetospheric field model has been developed and applied for the evaluation of a variety of magnetospheric current systems and their contribution to the ground magnetic field variations during magnetic storms. In later studies by Romashets et al. (2005, 2008), a 3D paraboloid model of the solar-wind magnetosphere interaction has been proposed to evaluate the magnetic field in the near-Earth space environment. In Romashets et al. (2005), an attempt has been made to describe the magnetosheath field as given by a scalar potential, which implies a current-free approximation. In order to involve finite currents, in Romashets et al. (2008), the magnetic field has been determined by a vector potential. The model studied in Romashets et al. (2008) has been also shown to be useful for studying the solar wind-magnetosphere interaction. As a result, a function measuring the discontinuity in magnetic field across

the magnetopause, denoted as $[B_t]$, has been expressed analytically. In Romashets et al. (2008), this quantity has been pointed out to be important in connection with ground magnetic field variations.

Geomagnetic activity can be characterized by geomagnetic indices, the most common being the *Dst* index. This index serves as a good measure of the overall strength of the near-Earth global electric currents, especially the ring current, thereby providing a good measure of geomagnetic storm intensity. Correlations between the *Dst* index and possible external drivers can provide the basis for empirical prediction (Burton et al., 1975; Siscoe et al., 2005).

The use of advanced techniques such as ANNs is found to be effective in predicting *Dst*. The modeling capability of an ANN lies in its ability to learn the mappings of underlying input-output features. If the network is designed and trained properly, it can perform generalisation rather than simple fitting of the function, see Gurney (1997); Hertz et al. (1991). This approach is rather useful when information and understanding of a physical system are lacking.

The ANN can be fed with the data on solar wind or solar activity (input data) and it can be trained to provide the caused geomagnetic activity (output). The ANNs are thus usable for the forecasting of the geomagnetic activity (e.g. Andrejková et al., 1997; Valach et al., 2007, 2009). This method has been widely used for the real-time modeling of the geomagnetic responses to solar wind disturbances, e.g. Boberg et al. (2000), Lundstedt (1992), Lundstedt et al. (2002), and Wu and Lundstedt (1996). For instance, the model developed by Lundstedt et al. (2002) consists of a recurrent neural

network that requires the hourly averages of the solar wind parameters as inputs and predicts the *Dst* index in almost real-time.

The underlying study is a contribution towards the *Dst* index modeling on the basis of the model proposed in Romashets et al. (2008). We employ the method of ANN to develop a revised version of the model, hereafter referred to as the revised RPV model. Unlike the approach by Lundstedt et al. (2002), where the solar wind parameters are used directly as the ANN input, here we feed the ANN with past hourly means of the function $[B_t]$ known from Romashets et al. (2008). As such, the presented ANN model can be thought as driven by input time histories of the solar wind-magnetosphere interaction. We obtain the *Dst* index series as the ANN output and compare it with the real data records. We evaluate the model for the set of intense geomagnetic storms of the 23-rd solar activity cycle. This study concerns strong geomagnetic storms because the intense events and their impacts on the terrestrial environment interest the space weather community (e.g. Echer et al. (2010), Gopalswamy et al. (2005), Siscoe et al. (2006), Srivastava (2005b), Srivastava and Venkatakrishnan (2004), Szajko et al. (2013) and Zhang et al. (2003)). Nevertheless, we must also admit a disadvantage of such a treatment: The downside of focusing on extreme storms is the limited number of the observed events, which partly reduces the potency of our arguments when drawing conclusions.

The paper is organised as follows. In Section 2, the data resources are specified. Development of the revised model for the *Dst* index and the results are presented in Section 3. The main findings are summarized in Section 4.

2. Data used

In Romashets et al. (2008), the Dst index for the so-called Bastille day event, on 14-15 July 2000, has been computed using the hourly solar wind parameters: the solar wind magnetic field component B_z , velocity V , particle density n and temperature T ; all of them provided by registration of the ACE satellite operating at the libration point L1.

In this study, 16 major geomagnetic storms from solar cycle 23 are considered, as listed in Tab. 1 (according to Table 1 in Tripathi and Mishra (2006)). Note, that two successive storms of November 2004 are treated as a single event. For each of the events considered, the series of the model Dst index is computed. The observed true Dst values, required to compare with the model values, were obtained from the World Data Center for Geomagnetism, Kyoto.

3. Models and results

In what follows, the revised model for the Dst index will be developed by combining the original analytical model by Romashets et al. (2008) with the approach of ANN. First, the analytical expression for the jump in magnetic field $[B_i]$ will be shown. The original model by Romashets et al. (2008) will be referred to as *the primal RPV model*, with its output denoted as Dst^p . This model will be evaluated for the set of geomagnetic storms considered and the need for its revision will be argued. As a preliminary step, a neural network model without hidden neurons will be presented, referred to as *the preliminary revised RPV model*, with its output denoted as Dst^n . The final revised version of this model will be constructed involving the neural network

possessing hidden layer and will be referred to as *the revised RPV model*, with its output denoted as Dst^{nn} . Hereafter, the notation Dst will stand for the observational data record.

3.1. Analytical expression for the jump in magnetic field

Considering the magnetopause as paraboloidal in shape, Romashets et al. (2008) constructed an analytical representation of magnetic fields in the region where the solar wind interacts with the Earth’s magnetosphere. The paraboloidal coordinates (σ, τ, φ) were adopted, defined by

$$x = \frac{1}{2}(\sigma^2 - \tau^2), \quad (1)$$

$$y = \sigma\tau \cos \varphi, \quad (2)$$

$$z = \sigma\tau \sin \varphi, \quad (3)$$

where x, y, z are solar ecliptic coordinates, with axis x pointing to the Sun. The components of magnetic field have been expressed analytically in paraboloidal coordinates. The full development of this analytical model can be found in Romashets et al. (2008) and will therefore not be reproduced here.

Magnetic field exhibits a discontinuity in tangential component when moving across the magnetopause from the solar wind (the IMF) to the magnetosphere (the internal field). There is no normal component of the magnetic field at the magnetopause. Romashets et al. (2008) used the notation $[B_t]$ for the magnitude of the jump in magnetic field across the magnetopause and argued for the relevance of this quantity for modeling the geomagnetic activity. We refer to the final expression for $[B_t]$ (Section 5 in Romashets et



al. (2008)) which reads

$$[B_t] = B_z \left[4.2629 \left(\frac{V_\infty}{500} \right) \left(\frac{10^6}{T_\infty} \right)^{1/2} - 1 \right] - 34.2109 \left(\frac{n_\infty}{5} \right)^{1/2} \left(\frac{V_\infty}{500} \right). \quad (4)$$

Here, the subscript ∞ stands for the undisturbed solar wind parameters far before the interaction with the magnetosphere, V_∞ is the velocity measured in km s^{-1} , n_∞ is the particle density measured in cm^{-3} , T_∞ is the temperature measured in K, B_z represents the z -component of the interplanetary magnetic field (IMF) measured in nT and the model Dst^p index is measured in nT.

3.2. Primal RPV model

In Romashets et al. (2008), the time histories of the jump in magnetic field $[B_t]$ have been related to the geomagnetic activity measured in terms of the Dst^p index as a weighted sum

$$Dst^p = \xi \sum_{k=0}^{12} x_k \exp(-k/12). \quad (5)$$

Here, *the entry functions* x_k are defined as

$$x_k = \frac{1}{2} [1 - \text{sgn}(B_{zk})] [B_t]_k \quad (6)$$

and ξ is the free parameter to be determined. The index k denotes the contribution to the geomagnetic disturbance of flows that came in k hours ago. In (6), the sign function is involved because the south orientation of the IMF is known to be crucial for the development of strong geomagnetic activity. The free parameter ξ in (5) can be computed to minimize the normalized mean square error defined by

$$NMSE^p = \frac{1}{M^2} \sum_{s=1}^M (Dst_s^p - Dst_s)^2, \quad (7)$$



where M is the length of the record (in hours).

In Romashets et al. (2008), the model Dst^p index given by (5) has been compared with the real Dst index for the so-called Bastille day event. In this study, we attempted to use the same technique for the set of intense geomagnetic storms considered, i.e. the 15 events (Tab. 1). We evaluated the primal RPV model, computed the common free parameter ξ using (7) over all the event records and obtained the value $\xi = 0.0058$. With the single free parameter ξ to be computed, poor agreement was achieved between the primal RPV model output and the set of observational data, as shown in Figs. 4a,b (green line). That is why, we intended to revise the model given by (5) and suggested a more general approach based on the method of ANN.

3.3. Preliminary revised RPV model

In the primal RPV model given by (5), the consecutive hourly means of $[B_t]$ given by (4) were multiplied by an exponential weighting function. In subsequent analysis, the idea was to drop the exponential term $\exp(-k/12)$ in (5) and thus leaving the weights unspecified. To determine the weights corresponding to particular hourly means of $[B_t]$, we adopted the ANN without hidden neurons. We designed the ANN to consist of 13 input neurons and one output neuron; 13 values of the entry function x_k given by (6) were used to feed the ANN model, with the Dst^n index obtained at the output.

Mathematically, the ANN model without hidden neurons can be expressed as

$$Dst^n = \frac{1}{1 + \exp[-(\sum_{j=0}^{12} w_j x_j - \Theta)]}. \quad (8)$$



In the above formula, the input vector x_j (for $j = 0$ to 12) consists of the entries given by (6) and w_j (for $j = 0$ to 12) is the vector of weights to be found. Function Θ characterizes the sensitivity threshold of the output neuron. Note, that the model of ANN expressed by (8) resembles the logistic regression model adopted in Srivastava (2005a, 2006) for predicting the occurrence of intense geomagnetic storms.

The vector of weights w_j (for $j = 0$ to 12) can be determined by the training process. We used the standard backward propagation algorithm (e.g. Gurney, 1997; Hertz et al., 1991) to train the neural networks. The algorithm is based on minimizing the normalized mean square error defined as

$$NMSE^n = \frac{1}{M^2} \sum_{s=1}^M (Dst_s^n - Dst_s)^2, \quad (9)$$

where M is the number of training patterns, i.e., the sets of data consisting of 13 input and 1 output values. The training process is usually initiated by random shooting of the components of the vector of weights, sensitivity thresholds of the neurons and some other control parameters, with fixed number of the iteration steps.

Here, the ANN was trained repeatedly with as many as 101 runs. We performed such a high number of runs because the result of training, i.e. the resulting vector of weights, slightly depended on the initial values of the vector of weights, which were set as a vector of small random numbers. In order to eliminate this inaccuracy, the median vector was computed from 101 resulting vectors¹. Knowing the vector of weights enables us to determine

¹Each element of the final vector was computed as the median from the corresponding

the importance of particular contributions of the ANN input given by (6) for each of the 13 time intervals. In Fig. 1, the weights at particular one-hour intervals are shown. We observe that the components of the weighting vector decrease with backward time intervals. However, for the current time interval, when the Dst^n index is evaluated, the contribution to the weight is negligible. This is in contrast to the primal RPV model, where the current solar wind input is considered with maximum weight. This finding is the main reason why the preliminary revised RPV model is presented in this paper.

3.4. Revised RPV model

Consequently, we included a hidden layer of neurons into the ANN model thereby setting up a more general model. In the case of ANN possessing hidden layer, there were more variables to be adjusted and therefore we expected to achieve more accuracy with observational data.

In Section 3.3 we found that the contribution of the current state of the near-Earth solar wind was small in the ANN model (see the value of the weight for step 0 in Fig. 1), hence it could be omitted. We considered contributions of only past time intervals as the ANN entries. The values of $[B_t]$ and $\text{sgn}(B_z)$, appearing in the entry function defined by (6), were taken as the ANN inputs. Leaving out the components for the current time interval, the input vector consisted of 24 components, with 2 components for each time interval. The model values of index Dst^{nn} were expected as the ANN output.

elements of 101 vectors of weights.

Mathematically, the ANN with a hidden layer consisting of H neurons can be expressed as

$$Dst^{nn} = \frac{1}{1 + \exp[-\sum_{i=1}^H (W_i \frac{1}{1 + \exp[-(\sum_{j=1}^{24} W_{i,j} X_j - \Theta_i)])]} . \quad (10)$$

In the above formula, for $i = 1$ to H and $j = 1$ to 24, X_j is the input vector consisting of 12 values of $[B_t]$ and 12 values of $\text{sgn}(B_z)$, $W_{i,j}$ is the weighting matrix characterizing the strength of the connection between the j -th input and the i -th hidden neuron, W_i stands for the weight between the i -th hidden neuron and the output neuron and vector Θ_i characterizes the sensitivity threshold of the hidden neurons. For the output neuron, the sensitivity threshold is set to 0. The number of the hidden neurons H is a free parameter to be determined.

According to the conventional routine, the 15 events considered (Tab. 1) were divided into three sets; each of them intended to be employed for a different purpose: 7 storms as training patterns, 4 storms as validation patterns and 4 storms for independent test. Random selection was used to assign the events to the three categories. The training patterns were used for running several ANNs with different number of hidden neurons. Then, we employed validation tests to search for the proper ANN architecture. The value of H has been varied from $H = 1$ to 36. The validation patterns were used for performance tests of each type of the ANN. We used quantitative measures to compare the performance in terms of the correlation coefficient CC_{ANN} and the root mean square error $RMSE_{ANN}$, see Figs. 2,3. We plotted the results of all independently accomplished trainings as well as the medians. For the neural networks with small number of hidden neurons, which showed better

results than those with more hidden neurons, the trainings were repeated several times. We did so because the results of the training of a neural network depended slightly on the initialization of the inter-neuron connections (weights) and the sensitivity thresholds - the initial values were set to small random numbers. The dispersion of the data visible in the graphs was caused by this effect. We have found the ANNs possessing the least number of hidden neurons to provide best results, namely for $H = 1$ to 4, we evaluated $CC_{ANN} = 64.9 \pm 8.6$ and $RMSE_{ANN} = (301 \pm 13)$ nT. We fixed $H = 3$ for the optimum ANN with hidden neurons. By common practice, the learning rate and the momentum term were set to 0.1 and 0.6, respectively. We checked that varying these values within a reasonable range caused practically no change in the result of the training. The maximum number of the learning epochs was set to 100. However, the trainings were mostly terminated before reaching 100 epochs under the following rule: After achieving the local maximum of the correlation between the ANN outputs and the desired outputs, ten more learning epochs were performed. The training was stopped when the results of the local maximum remained unsurpassed.

3.5. Results and skill scores

The series of the Dst^{nn} index as the revised RPV model output are shown in Figs. 4a,b (red line). Much closer agreement with the observational Dst index is now obtained. Motivated by the study of Rastätter et al. (2013), we employed two types of skill scores, the correlation coefficient CC and the prediction efficiency PE , to quantitatively assess the revised RPV model performance. The correlation coefficient CC between the series of Dst^{nn} and



Dst is defined as

$$CC = \frac{\sum_{s=1}^M (Dst_s - \overline{Dst})(Dst_s^{nn} - \overline{Dst^{nn}})}{\sqrt{\sum_{s=1}^M (Dst_s - \overline{Dst})^2} \sqrt{\sum_{s=1}^M (Dst_s^{nn} - \overline{Dst^{nn}})^2}}, \quad (11)$$

where $\overline{Dst^{nn}}$ and \overline{Dst} stand for the arithmetic means of the series of Dst^{nn} and Dst , respectively, and M is the length of the record. Using the same notation, the prediction efficiency PE for a discrete time series is defined as

$$PE = 1 - \frac{\sum_{s=1}^M (Dst_s - Dst_s^{nn})^2}{\sum_{s=1}^M (Dst_s - \overline{Dst})^2}. \quad (12)$$

The value $PE = 1$ indicates perfect model performance and $PE = 0$ indicates performance comparable to predicting the arithmetic mean of the observed signal. PE can reach unlimited negative values. The values $PE < 0$ occur when the observed mean is a better predictor than the model. In contrast to the correlation coefficient, PE includes the amplitude of the modeled signal in addition to the shape of the time series. Signals with good correlation but incorrect amplitudes may result in negative PE scores.

The skill scores for each particular event are shown in Tab. 1. We run the revised RPV model with the set of independent test events listed in Tab. 1 in order to assess the model accuracy. For the event of May 1998, the model run scored poorly; we obtained the skill scores as low as $CC = 0.33$ and $PE = -0.64$. The event of May 1998 is a long duration storm with rather complex Dst index record and therefore appears to be difficult to capture by the model. As such, we excluded this particular test event from the statistics and considered only the remaining independent test events, i.e. the events of Oct 1999, Aug 2000 and Nov 2004. Similarly as in Table 3 in Rastätter et al. (2013), we evaluated the averages and standard deviations for the skill scores and found $CC = 0.74 \pm 0.13$ and $PE = 0.44 \pm 0.15$.



In Rastätter et al. (2013), various types of the *Dst* prediction models were subject to ranking. Among others, the class of so called *Dst*-specification models was assessed, in which the *Dst* index was derived directly through an analytic or iterative formula or a neural network based-algorithm. The revised RPV model falls into the class of the *Dst*-specification models. As shown in Table 3 in Rastätter et al. (2013), the skill scores for the *Dst*-specification models range from $CC = 0.58 \pm 0.35$ and $PE = 0.18 \pm 0.56$ to $CC = 0.92 \pm 0.07$ and $PE = 0.78 \pm 0.21$. The performance of the revised RPV model quantified in terms of the skill scores CC and PE can be compared with the performance of the *Dst*-specification models.

4. Discussion and conclusions

In this study, the value of the discontinuity in magnetic field [B_t] across the magnetopause known from the model by Romashets et al. (2008) was treated in connection with ground magnetic field variations. We have revised the model using the method of ANN in order to assess the weights corresponding to time histories of the function [B_t]. A new time prediction model to represent the *Dst* index has come out of our study. We named the new model the revised RPV model.

We have found, that the most relevant input to the revised RPV model was the one which represented the value of [B_t] one hour prior to the period in which the *Dst* index was computed. On the other hand, we have shown that the current contribution from the function [B_t] to the *Dst* index value was negligible (Fig. 1). As such, the model for producing the series of the *Dst* index can be considered to be a one-step ahead forecast instead of nowcast.

On physical grounds, the revised RPV model responds to the fact that there is a time delay of the order of one hour between the solar wind-magnetosphere interaction and development of geomagnetic activity.

As regards the time lapse between solar wind parameters and geomagnetic effect, in Borovsky and Funsten (2003) the solar wind-magnetosphere coupling has been investigated with the effect of MHD turbulence. In Borovsky and Funsten (2003), the correlation between the amplitude of the MHD turbulence in the upstream solar wind and the amplitude of the Earth's geomagnetic activity indices has been explored. The correlations have been performed both with and without a one-hour time lag between the solar wind parameters and the geomagnetic indices. As shown in Borovsky and Funsten (2003), the correlations between the turbulence amplitudes and the *Dst* index were weak. However, the correlations improved if a several-hour time lag was introduced between the solar wind measurements and the measurement of the *Dst* index.

To measure the performance of models for the *Dst* index prediction, various types of skill scores have been employed (Ji et al., 2012; Rastätter et al., 2013). In Table 3 in Rastätter et al. (2013), three classes of the *Dst* index prediction models were subject to ranking. They were the 3D magnetosphere models, the ring current kinetic models and the *Dst*-specification models. As pointed out in Rastätter et al. (2013), model ranking varied widely by skill score used and none of the models consistently performed best for all events.

The revised RPV model developed in this study, falls into the class of the *Dst*-specification models. We quantified its performance in terms of two types of skill scores, the correlation coefficient *CC* and the prediction effi-



ciency PE (Tab. 1). In case of long-duration geomagnetic storms or those with a more complex Dst index record, the runs of the revised RPV model scored poorly. For these events low values of the correlation coefficient and negative values of the prediction efficiency were obtained. On average, performance of the revised RPV model in terms of the skill scores is comparable to that of the Dst -specification models, according to the model ranking chart shown in Table 3 in Rastätter et al. (2013). However, exact comparison and inclusion to the ranking chart in Rastätter et al. (2013) is not possible, due to different selection of events and different lengths of data records.

In this study, we pointed out the relevance of the quantity $[B_t]$ as a measure of the solar wind-magnetosphere interaction in connection with the geomagnetic response modeling. We have also demonstrated, that relying on the ANN approach it is possible to enhance modeling capability of the primal model proposed by Romashets et al. (2008) and improve its response to observational data. As we bounded our study to strong geomagnetic storms, the restricted number of observed events has to be considered to be a limiting factor to our conclusions. In future studies, it could be valuable to modify the model to produce the so-called dynamic pressure corrected Dst index (known as Dst^*), which takes into account the motion of the magnetopause currents Burton et al. (1975).

Acknowledgments

The support from the VEGA grant 2/0030/14 of the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences, by the Slovak Research and



Development Agency under the contract No. APVV-0662-12 and by the Ministry of Education, Youth and Sports of the Czech Republic under project OC09070 is acknowledged. The data on solar wind were provided by OMNIWeb Space Physics Data Facility at omniweb.gsfc.nasa.gov. The *Dst* data were obtained from the World Data Center (WDC) for Geomagnetism, Kyoto (wdc.kugi.kyoto-u.ac.jp/dst/dir/). The authors also gratefully acknowledge the constructive comments and suggestions of the anonymous referees.

References

- Alexeev, I.I., Feldstein, Ya.I., 2001. Modeling of geomagnetic field during magnetic storms and comparison with observations, *J. Atmos. Sol. Terr. Phys.*, 63, 431–440.
- Andrejková, G., Tóth, H., Kudela, K., 1997. Fuzzy neural networks in the prediction of geomagnetic storms, *Proceedings of "Artificial Intelligence in Solar-Terrestrial Physics"*, Publisher European Space Agency, Lund, pp. 173–179.
- Baker, D.N., Weigel, R.S., Rigler, E.J., McPherron, R.L., Vassiliadis, D., Arge, C.N., Siscoe, G.L., Spence, H.E., 2004. Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical methods, *J. Atmos. Sol. Terr. Phys.*, 66 (15-16 SPEC. ISS.), 1491–1497.
- Boberg, F., Wintoft, P., Lundstedt, H., 2000. Real Time Kp Predictions from Solar Wind Data using Neural Networks, *Phys. Chem. Earth (C)*, Vol. 25, No. 4, pp. 275–280.

- Borovsky, J.E., Funsten, H.O., 2003. Role of solar wind turbulence in the coupling of the solar wind to the Earth's magnetosphere, *J. Geophys. Res.*, 108(A6), 1246, doi:10.1029/2002JA009601.
- Burton, R.K., McPherron, R.L., Russell, C.T., 1975. An Empirical Relationship Between Interplanetary Conditions and *Dst*, *J. Geophys. Res.*, 80, No. 31, pp. 4204–4214.
- Echer, E., Tsurutani, B.T., Guarnieri, F.L., 2010. Interplanetary origins of November 2004 superstorms, *J. Atmos. Sol. Terr. Phys.*, 72, 280–284.
- Gombosi, T.I., DeZeeuw, D.L., Groth, C.P.T., Powell, K.G., Clauer, C.R., Song P., 2001. From Sun to Earth: Multiscale MHD Simulations of Space Weather, in *Space Weather*, *Geophys. Monogr. Ser.*, vol. 125, edited by Song, P., Singer, H.J., Siscoe, G., pp. 169–176, AGU, Washington, D.C.
- Goodrich, C.C., Sussman, A.L., Lyon, J.G., Shay, M.A., Cassak, P.A., 2004. The CISM code coupling strategy, *J. Atmos. Sol. Terr. Phys.*, 66 (15-16 SPEC. ISS.), 1469–1479.
- Gopalswamy, N., Yashiro S., Michalek, G., Xie, H., Lepping, R.P., Howard, R.A., 2005. Solar source of largest geomagnetic storm of cycle 23, *Geophys. Res. Lett.*, vol. 32, doi:10.1029/2004GL021639.
- Gurney, K., 1997. *An Introduction to Neural Networks*, 234 pp., UCL Press, London.
- Hertz, J., Krogh, A., Palmer, R.G., 1991. *Introduction to the Theory of Neural Computation*, Addison-Wesley, Reading, Mass.

- Ji, E.-Y., Moon, Y.-J., Gopalswamy, N., Lee, D.-H., 2012. Comparison of Dst forecast models for intense geomagnetic storms, *J. Geophys. Res.*, 117, A03209, doi:10.1029/2011JA016872.
- Lundstedt, H., 1992. Neural networks and predictions of solar-terrestrial effects, *Planet. Space. Sci.*, 40, 457–464.
- Lundstedt, H., Gleisner, H., Wintoft, P., 2002. Operational forecasts of the geomagnetic *Dst* index, *Geophys. Res. Lett.*, 29(24), 2181, doi:10.1029/2002GL016151.
- Odstrčil, D., Riley, P., Zhao, X.P., 2004. Numerical simulation of the 12 May 1997 interplanetary CME event, *J. Geophys. Res.*, 109, A02116, doi:10.1029/2003JA010135.
- Rastätter, L., et al., 2013. Geospace environment modeling 2008-2009 challenge: Dst index, *Space Weather*, 11, 187-205, doi:10.1002/swe.20036.
- Romashets, E.P., Poedts, S., Vandas M., 2008. Modeling of the magnetic field in the magnetosheath region, *J. Geophys. Res.*, 113, A02203, doi:10.1029/2006JA012072.
- Romashets, E.P., Vandas, M., Nagatsuma, T., 2005. Modelling of magnetic field near the magnetopause, *Planet. Space Sci.*, 53, 127–131.
- Siscoe, G., Baker, D., Weigel, R., Hughes, J., Spence, H., 2004. Roles of empirical modeling within CISM, *J. Atmos. Sol. Terr. Phys.*, 66 (15-16 SPEC. ISS.), 1481–1489.



- Siscoe, G., Crooker, N.U., Clauer, C.R., 2006. Dst of the Carrington storm of 1859, *Adv. Space Res.*, 38, 173–179.
- Siscoe, G., McPerron, R.L., Liemohn, M.W., Ridley, A.J., Lu, G., 2005. Reconciling prediction algorithms for *Dst*, *J. Geophys. Res.*, 110, A02215, doi:10.1029/2004JA010465.
- Srivastava, N., 2005a. A logistic regression model for predicting the occurrence of intense geomagnetic storms, *Annales Geophysicae*, 23, 2969–2974.
- Srivastava, N., 2005b. Predicting the occurrence of super-storms, *Annales Geophysicae*, 23, 2989–2995.
- Srivastava, N., 2006. The Challenge of Predicting the Occurrence of Intense Storms, *J. Astrophys. Astr.*, 27, 237–242.
- Srivastava, N., Venkatakrishnan, P., 2004. Solar and interplanetary sources of major geomagnetic storms during 1996–2002, *J. Geophys. Res.*, 109, A10103, doi:10.1029/2003JA010175.
- Szajko, N.S., Cristiani, G., Mandrini, C.H., Dal Lago, A., 2013. Very intense geomagnetic storms and their relation to interplanetary and solar active phenomena, *Adv. Space Res.*, 51, 1842–1856.
- Tóth, G., van der Holst, B., Sokolov, I.V., De Zeeuw, D.L., Gombosi, T.I., Fang, F., Manchester, W.B., Meng, X., Najib, D., Powell, K.G., Stout, Q.F., Glocer, A., Ma, Y.-J., Opher, M., 2012. Adaptive numerical algorithms in space weather modeling, *J. Comput. Phys.*, 231 (3), 870–903.



- Tripathi, R., Mishra, A.P., 2006. Occurrence of severe geomagnetic storms and their association with solar-interplanetary features, ILWS Workshop, preprint.
- Tsyganenko, N.A., 2013. Data-based modelling of the Earth's dynamic magnetosphere: A review, *Annales Geophysicae*, 31 (10), 1745–1772.
- Valach, F., Hejda, P., Bochníček, J., 2007. Geoeffectiveness of XRA events associated with RSP II and/or RSP IV estimated using the artificial neural network, *Stud. Geophys. Geod.*, 51, 551–562.
- Valach, F., Revallo, M., Bochníček, J., Hejda, P., 2009. Solar energetic particle flux enhancement as a predictor of geomagnetic activity in a neural network-based model, *Space Weather*, 7, S04004, doi:10.1029/2008SW000421.
- Wu, J.-G., Lundstedt, H., 1996. Prediction of geomagnetic storms from solar wind data using Elman recurrent neural networks, *Geophys. Res. Lett.*, 23, 319–322.
- Zhang, J., Dere, K.P., Howard, R.A., Bothmer, V., 2003. Identification of solar sources of major geomagnetic storms between 1996 and 2000, *Astrophys. J.*, vol. 82, pp. 520–533.

Tables

This is the author's version of a manuscript that was accepted for publication in *Journal of Atmospheric and Solar-Terrestrial Physics*. The definitive version was subsequently published in: Miloš Revallo, Fridrich Valach, Pavel Hejda and Josef Bochníček: A neural network Dst index model driven by input time histories of the solar wind-magnetosphere interaction.

Journal of Atmospheric and Solar-Terrestrial Physics, Volumes 110-111, April 2014, Pages 9-14. DOI: 10.1016/j.jastp.2014.01.011

© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>



Table 1: Intense geomagnetic storms during solar cycle 23, according to Table 1 in Tripathi and Mishra (2006), their usage in the revised RPV model and the values of correlation coefficient CC and prediction efficiency PE . The final values of Dst_{min} are listed, according to the Kyoto WDC. Note: In Table 1 in Tripathi and Mishra (2006) the provisional values of Dst_{min} have been considered.

Event	Date	Onset	Dst_{min}	Usage	CC	PE
May 1998	04/05/1998	6 UT	-205	test	0.33	-0.64
Sep 1998	25/09/1998	10 UT	-207	validation	0.71	0.42
Oct 1999	22/10/1999	6 UT	-237	test	0.59	0.31
Apr 2000	07/04/2000	1 UT	-288	training	0.61	0.19
July 2000	15/07/2000	22 UT	-301	training	0.68	0.21
Aug 2000	12/08/2000	10 UT	-235	test	0.82	0.40
Mar 2001	31/03/2001	9 UT	-387	training	0.79	0.48
Apr 2001	11/04/2001	23 UT	-271	training	0.38	-0.01
Nov 2001a	06/11/2001	6 UT	-292	training	0.90	0.80
Nov 2001b	24/11/2001	15 UT	-221	training	0.34	-0.01
Oct 2003	30/10/2003	23 UT	-383	validation	0.76	0.54
Nov 2003	20/11/2003	20 UT	-422	validation	0.84	0.64
Nov 2004	08/11/2004	7 UT	-373	test*	0.81	0.60
Nov 2004	10/11/2004	10 UT	-289	test*	0.81	0.60
May 2005	15/05/2005	8 UT	-263	validation	0.58	0.28
Aug 2005	24/08/2005	11 UT	-216	training	0.80	0.31

Note *: The two storms of November 2004 were treated as one test event.



Figure captions

Fig. 1. The plot of weighting vector w_j components. For each backward time interval, $j = 0$ to 12, whisker boxes are shown, with upper and lower quartiles and medians (thick bars). In total, 101 values at each time interval were considered, as a result of 101 runs of the preliminary revised RPV model without hidden neurons.

Fig. 2. Correlation coefficient CC_{ANN} for the ANN with hidden layer of neurons in dependence on the number of hidden neurons.

Fig. 3. Root mean square error $RMSE_{ANN}$ for the ANN with hidden layer of neurons in dependence on the number of hidden neurons.

Fig. 4a. Observational and modeled series of the Dst index for the sets of training data.

Fig. 4b. Observational and modeled series of the Dst index for the sets of validation data and independent test data. Legend explaining the meaning of individual lines is shown in Fig. 4a.

Figure 1

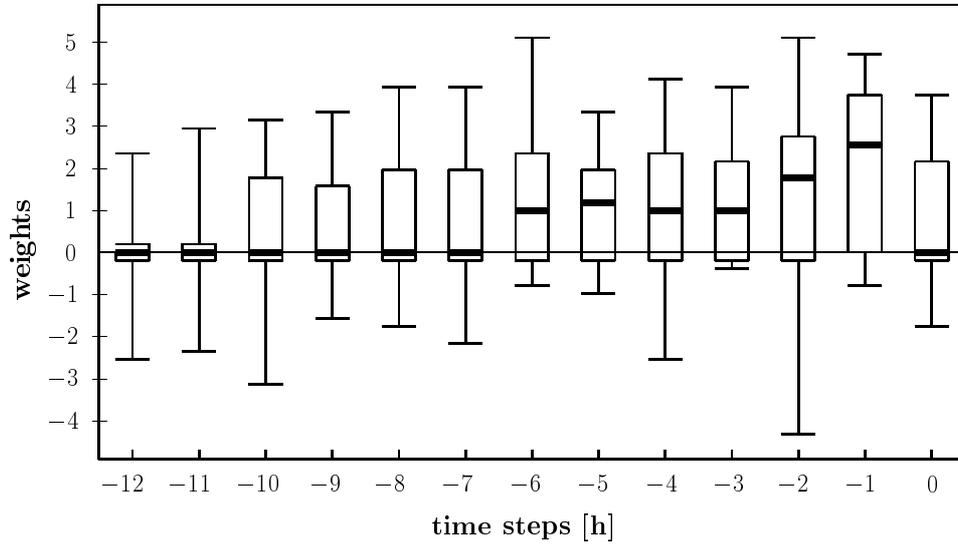


Figure 2

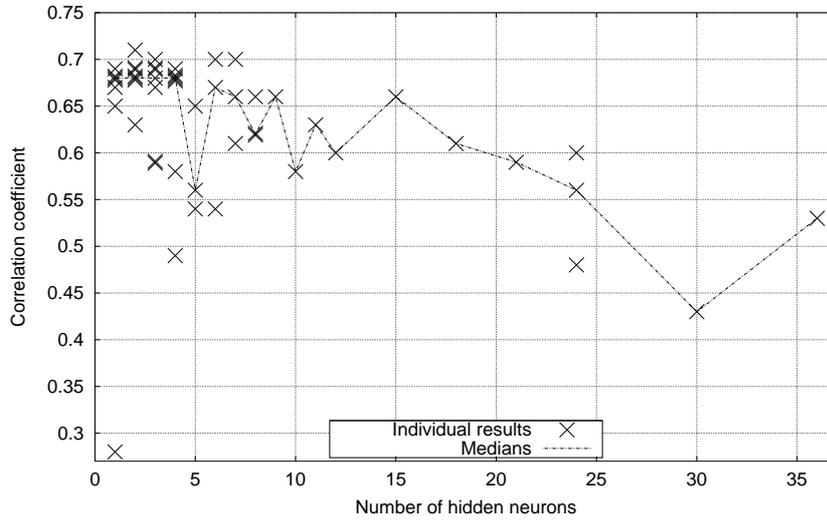


Figure 3

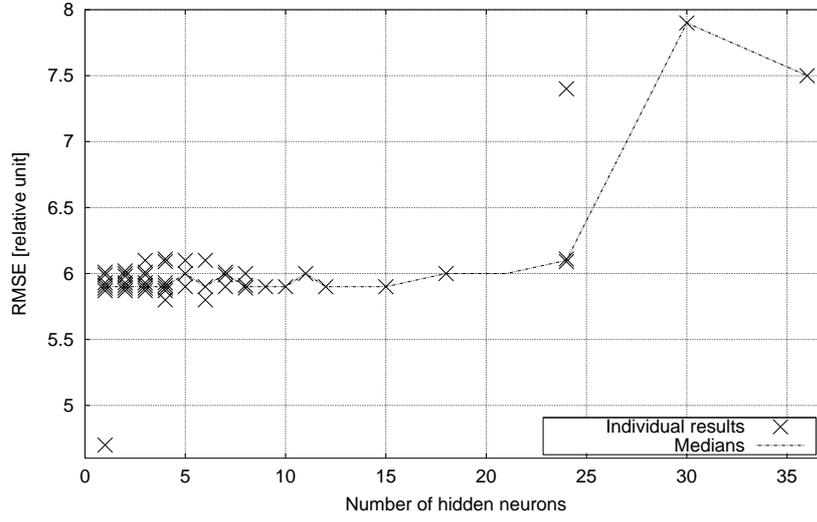


Figure 4a

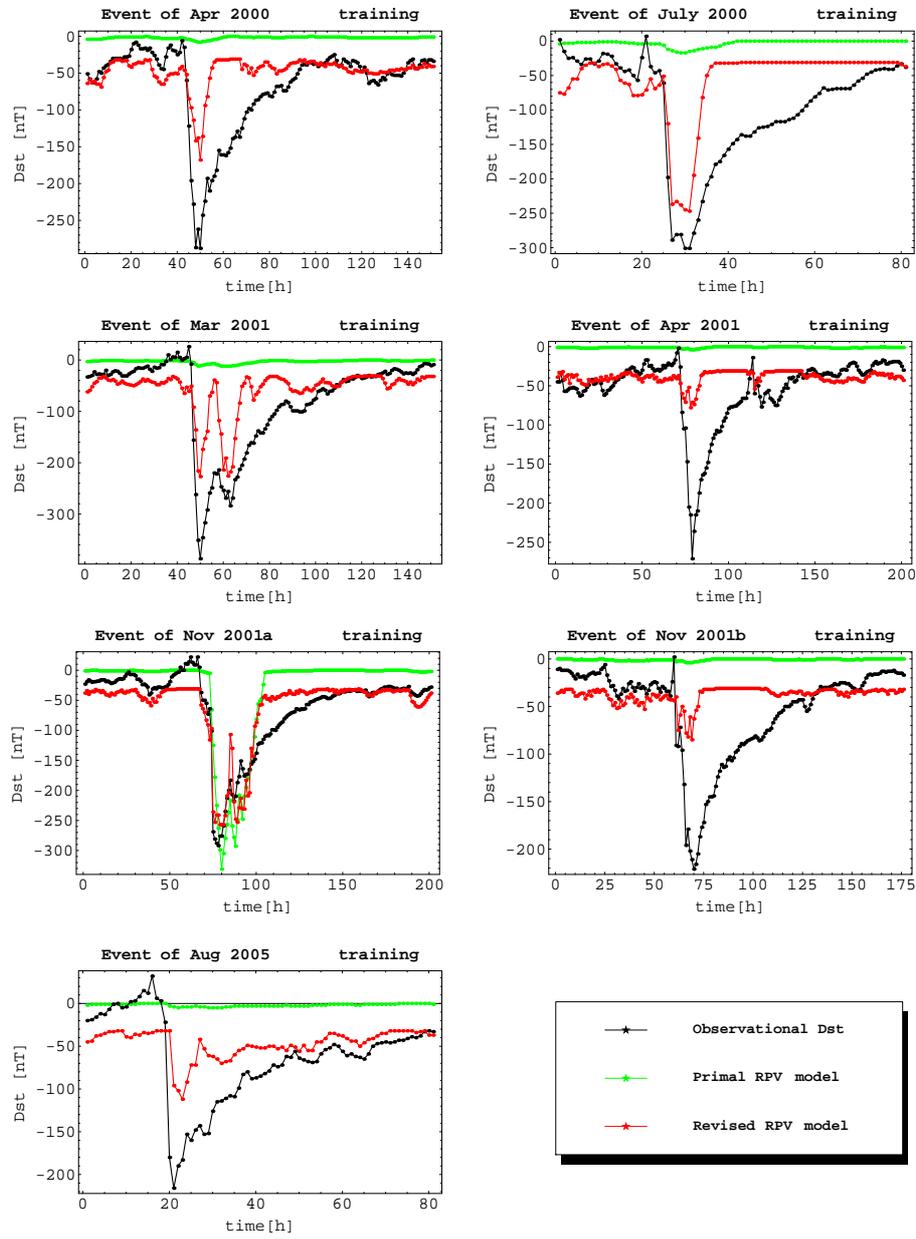


Figure 4b

