

## THE COMBINED APPROACH FOR SPACE WEATHER PREDICTION WITH A GUARANTEED METHOD AND EVOLUTIONARY ALGORITHM

Oleh Semenov

Space Research Institute of NASU & SSAU,  
40, Glushkov Ave., build. 4/1, Kyiv, UA-03680, Ukraine  
e-mail: oleg.semeniv@gmail.com

(Received February 03, 2015; in final form — September 20, 2015)

The problem of space weather prediction using guaranteed approach and solar wind parameters is considered. The original approach for space weather prediction based on a discrete dynamic model, evolutionary algorithm and a guaranteed interval estimation method is proposed. The problem of discrete model identification is formulated as a nonlinear programming problem with constraints. A new algorithm of *Dst*-index prediction using guaranteed interval estimation method is described. The model for the prediction of geomagnetic storms was constructed and the interaction between the solar wind parameters and the magnetosphere was discussed. The dependence between prediction accuracy and the forecasting horizon is analyzed.

**Key words:** space weather, geomagnetic activity prediction, genetic programming, guaranteed approach.

PACS number(s): 02.70.Rr, 94.05.Sd, 94.05.Sx, 94.30.Lr, 94.30.Vf

### I. INTRODUCTION

The solar energy that reaches the magnetosphere and the ionosphere makes a great impact on the technological and biological processes in the space and on the ground [1–8]. Such influence has been studied within the problem of space weather. The influence of sun on near earth space is usually characterized by geomagnetic indices (*Kp*-index, *Ap*-index, *Dst*-index and other) [5–8]. These indices describe a part of complex condition of the Solar–Earth relationships. The selection of the most relevant index for characterizing physical processes in the near-Earth space is not a trivial problem and it requires serious preliminary research [1, 2, 7, 8]. The space weather investigations can be divided into two directions. The first aspect is practical and it is affected by the prediction and reduction of space weather effects [7]. This task should be accomplished by launching a vast number of spacecraft. But, nowadays resources are insufficient to produce and maintain such a large space fleet as well as to process all the data delivered by these spacecrafts. So, another aspect is the development of forecasting techniques by reconstructing mathematical models using solar wind and magnetospheric parameters measured by ACE, WIND, SOHO, STEREO, CLUSTER and THEMIS spacecrafts, and ground-based stations observations (Intermagnet, MAGDAS, etc.). Therefore, the prediction of geomagnetic indices is an important part of space weather investigation. As was shown in [1, 2], the *Dst*-index corresponds to geomagnetic storms. So, for its prediction it is necessary to solve the problem of model structure and parameters identification. There are

many different models including a neural network, a linear regression model, and others [4, 7, 9–13] for *Dst*-index prediction. The most accurate and reliable is a nonlinear discrete dynamic model that describes a relation between spacecraft observation data and ground measurements from magnetic observatories [14]. Such a class of models was chosen in this investigation.

In the paper, according to [2, 15–20], the polynomial model for predicting *Dst*-index has been used. It is known that Burton (BMR) linear equation [21] describes *Dst*-index dynamics

$$\frac{dDst^*}{dt} = E(t) - \frac{Dst^*}{\tau}, \quad (1)$$

where  $\tau$  is *Dst*-index decay time constant,  $E(t)$  is the ring current energy injection,

$$Dst^*(t) = Dst(t) - (b \cdot \sqrt{P(t)} + a),$$

$P$  is the solar wind pressure,  $a$  and  $b$  are constants. Obviously, using equation (1) it is possible to predict *Dst*-index on the time interval  $\Delta t$

$$Dst^*(t + \Delta t) = Dst^*(t) + \left[ E(t) - \frac{Dst^*(t)}{\tau} \right] \Delta t. \quad (2)$$

Let us assumed the presence of nonlinear relation between model input and output. It is supposed that the input signal of the model can be characterized by the product of the south magnetic field component  $B_z$  and the solar wind speed  $\nu$  [2]. This permits to describe the behaviour of *Dst*-index using (2) as

---

$$Dst(t + \Delta t) = F \left[ Dst(t), Dst(t - \Delta t), \dots, Dst(t - n\Delta t), \nu(t) \cdot B_z(t), \nu(t - \Delta t) \cdot B_z(t - \Delta t), \dots, \nu(t - n\Delta t) \cdot B_z(t - n\Delta t) \right], \quad (3)$$

where  $F$  is a polynomial function;  $n\Delta t$  is an associated maximal lag,  $t$  is time [2]. Now, expressing  $t = k$  and  $n\Delta t$  as  $n_y, n_u, n_\xi$ , equation (3) can be represented as

$$y(k+1) = F^l \left[ y(k), y(k-1), \dots, y(k-n_y), u(k), u(k-1), \dots, u(k-n_u), \xi(k), \xi(k-1), \dots, \xi(k-n_\xi) \right], \quad (4)$$

where  $F^l[\bullet]$  is a polynomial function of the order  $l$  with variables  $y(k)$ ,  $u(k)$  and  $\xi(k)$ ;  $u(k) = -(\nu(k) \cdot Bz(k)) \cdot 10^{-3}$ ;  $y(k)$  is *Dst*-index values at time  $k$ ;  $\xi(k)$  is a quantity that accounts a possible noise and uncertainties at the time  $k$  [22, 23],  $n_y, n_u, n_\xi$  are their associated maximal lags. As a result the problem of *Dst*-index prediction, can be reduced to the model (4) identification. In previous studies [19, 20, 24] an original algorithm of model structure and parameters identification has been developed. It is based on the selection of the most significant linear and nonlinear terms for the most accurate prediction of *Dst*-index. The model is presented as “input-output” system [9, 22]. Its reconstruction is stopped when the model error meets the requirements and model complexity reaches the minimum value on the training data set.

The time series that characterize the solar wind consists of multidimensional components. These components increase the dimensionality of the model input; as a result the complexity of the identification problem significantly increases. It would be shown that the identification problem can be reduced to the solution of the corresponding mathematical programming problem with constraints. Such an approach substantially differs from

the widely used statistical methods [10, 11, 25, 26] where the great number of input variables is exploited. This paper focuses on a novel approach to reconstruction of the guaranteed prediction interval, the so called “prediction tube”. Also, the problem of *Dst*-index model identification with an evolutionary algorithm application is discussed in details.

## II. A GUARANTEED PREDICTION INTERVAL ESTIMATION

According to [24], made in this chapter are some improvements in solving the problem of guaranteed interval estimation of  $\tilde{y}(k+q)$  of the  $y(k+q)$  and polynomial model identification. It can be seen that it is more convenient to predict not a consequence  $y(k+q)$ , but the subjects, i.e.  $u(k), u(k-1), \dots, u(k-n_u), \dots, y(k), y(k-1), \dots, y(k-n_y)$ .

Let us now estimate the interval  $\tilde{y}(k+q)$  of  $y(k+q)$  when  $q = 1$  with mathematical model (4). It is necessary to obtain  $\tilde{u}(k+1)$  of  $u(k+1)$ . Let us have the change rate of  $u(k)$  as

$$-\delta \leq \Delta u(k) \leq \delta, \quad (5)$$

$$\delta = \max (|u(k-1) - u(k)|, |u(k-2) - u(k-1)|, \dots, |u(k-N_\delta) - u(k-N_\delta-1)|),$$

where  $N_\delta$  is the lag of time interval in  $\delta$  is calculated. From (5) it is estimated  $u(k+1)$

$$u(k) - \delta \leq u(k+1) \leq u(k) + \delta. \quad (6)$$

Rewriting (4) in view (5), so that

$$y(k+1) = F^l \left[ y(k), y(k-1), \dots, y(k-n_y), \dots, u(k) + \Delta u(k), u(k), u(k-1), \dots, u(k-n_u) \right]. \quad (7)$$

From (7) the interval  $\tilde{y}(k+1)$  can be found as

$$\underline{y}(k+1) \leq \tilde{y}(k+1) \leq \overline{y}(k+1), \quad (8)$$

where

$$\underline{y}(k+1) = \min F^l \left[ y(k), y(k-1), \dots, y(k-n_y), u(k) + \Delta u(k), u(k), u(k-1), \dots, u(k-n_u) \right], \quad (9)$$

$$\overline{y}(k+1) = \max F^l \left[ y(k), y(k-1), \dots, y(k-n_y), u(k) + \Delta u(k), u(k), u(k-1), \dots, u(k-n_u) \right]. \quad (10)$$

If  $F^l(u(k) + \Delta u(k))$  is a multi-extreme function, the interval  $\{-\delta; \delta\}$  can be presented by a discrete data set  $\Delta u_j(k)$

$$\Delta u_j(k) = \left( \left( \frac{2\delta}{J} \right) \cdot (j-1) - \delta \right), \quad (11)$$

$j = \overline{1; J+1}, J$  is integer constant,

then the problems (9) and (10) can be reduced to standard combinatorial problems, which can be solved using the brute force direct search. All the above related to (9) and (10) and according to [24] can be established for predicting in  $q$  steps,  $q \geq 2$ .

### III. A NONLINEAR MODEL IDENTIFICATION ALGORITHM

Now, let us discuss the problem of model identification for the geomagnetic index  $y(k)$  approximation. Let us represent the model (4) as

$$y(k) = \sum_{m=1}^M c_m F_m(\Psi(k)), \quad (12)$$

where  $\Psi(k)$  is a vector which is formed on the basis of observation data by combining of "input-output" variables

$$\Psi(k) = (u(k-1), \dots, u(k-n_u), y(k-1), \dots, y(k-n_y)),$$

where  $F_1(\psi_1(k)), \dots, F_M(\psi_M(k))$  are regressors;  $m = \overline{1; M}$ ,  $M \in \mathbb{R}^n$  is the number of regressors;  $\mathbb{R}^n$  is the integer set;  $c_1, \dots, c_M$  are model parameters. The identification problem consists in the model search with the minimum number of the regressors  $M$ , which yields the most accurate calculation of  $y(k)$  to the observations  $\hat{y}(k)$ . For providing maximal fitting of the model result to the real data the statistical parameter  $\chi^2$  was used

$$\min \chi^2 = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2, \quad (13)$$

where  $N$  is the size of the training data set. Also, the next constraints were taken into account: the constraints on the model parameters  $\mathbf{c} = \{c_1, \dots, c_M\}$ ,  $c_m \in \mathbb{R}^d$ ,  $\mathbb{R}^d$  is the constraint set of real numbers, the constraints on polynomial function complexity  $l \leq l^*$ , where  $l^*$  is the maximal value of the polynomial function order, and the constraints on the number of the regressors  $M \leq M^*$ , where  $M^*$  is the maximal number of the regressors.

For the nonlinear model structure selection the Genetic Programming (GP) method [27–30] was used. The model identification involves some procedures similar to biological processes in nature [30]. First of all, two data sets are chosen from the measurements: the training and the testing once. Also, the lags are introduced for finding the proper set of the regressors to the time  $k$ . The set of individuals (regressors) is randomly selected in the algorithm. These individuals form the first generated population [27, 28, 30]. The process of regressors

selection according to evolution principal is similar to the biological process of organisms survival [30]. In the same way as in nature, genetic algorithms carried out the search of the most perfect individuals without using of information about them. Every individual has a fitness value that expresses efficiency of the solution for solving the problem. Better solutions are assigned higher values of fitness than worse solutions. The fitness function also determines how successful the individual will be at propagating its genes to subsequent generations in the next population. It permits to select the most adapted individuals according to the evolution survival principal [30]. After that the fitness functions values are calculated; the most enduring individuals are chosen to generate the new population. The individuals in genetic algorithm are randomly initialized by genetic operators: mutation, recombination and crossover. In our algorithm the genetic procedure based on roulette wheel selection method and elitism principle [30] has been used. The modeling of the evolution process is applied in a numerical algorithm (Fig. 1). In the end of the algorithm the most significant regressors are chosen.

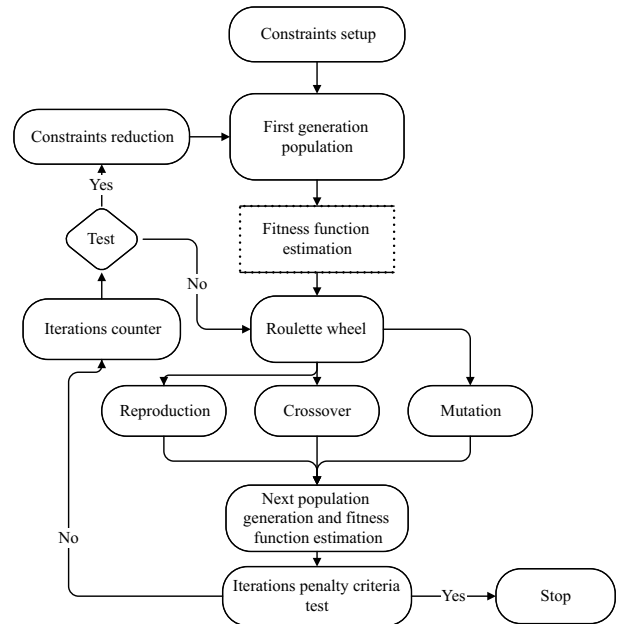


Fig. 1. Block diagram of numerical algorithm for regressors selection.

The evolutionary algorithm has some advantages in comparison to other optimization techniques. Let us present the result of model identification in the analytical form. The selection of the optimal model (12) is provided by solving a mix optimization problem

$$\min \chi^2 = \frac{1}{N} \sum_{k=1}^N \left( \hat{y}(k) - \left( \sum_{m=1}^M c_m F_m^l(\Psi(k)) \right) \right)^2, \quad (14)$$

where  $N = N_a^* + N_b^* + N_c^*$ . The training data set consists of  $N_a^*$  data set for model reconstruction;  $N_a^* + N_b^*$

data set for model parameters estimation, and  $N$  data set for the accuracy test [20]. The binary trees are used to represent the results of identification and this representation consists of the most significant  $M$  regressors  $\{\psi_1(k), \dots, \psi_M(k)\}$ . The system expression can be generated through the trees preorder traversal based on the above-mentioned rules and putting some intuitive as-

$$y(k) = c_1 F_1^{l_1} \psi_1(k) + c_2 F_2^{l_2} \psi_2(k) + c_3 F_3^{l_3} \psi_3(k) + \dots + c_M F_M^{l_M} \psi_M(k).$$

The penalty term of the fitness function for regressor selection was suggested

$$0.25 < \phi_m < 0.99, \quad \phi_m = \frac{|r_m|}{1 + \exp(\beta(l_m - l_{max}))}, \quad (15)$$

where  $r_m$  is the correlation coefficient between regressor  $m$  and observed output  $\hat{y}$ , which is calculated as

$$r_m = \frac{\sum_{k=1}^N (\hat{y}(k) - \bar{y}) \cdot (\psi_m(k) - \bar{\psi})}{\sqrt{\left(\sum_{k=1}^N (\hat{y}(k) - \bar{y})\right)^2 \cdot \left(\sum_{k=1}^N (\psi_m(k) - \bar{\psi})\right)^2}},$$

$$\bar{y} = \frac{1}{N} \sum_{k=1}^N \hat{y}(k), \quad \bar{\psi} = \frac{1}{N} \sum_{k=1}^N \psi(k),$$

$l_m$  is the order of the regressor with index  $m$ ,  $l_{max}$  is the maximal possible order of the regressors,  $\beta$  is control parameter for regressor complexity optimization (Fig. 2). As seen on Fig. 2, parameter  $\beta$ ,  $0 \leq \beta \leq 1$  can give preference to some class of the regressors complexity. For  $\beta = 1$  the preference will be given to the selection of linear regressors, and for  $\beta = 0$  the selection will be made without notification. In the first population the fitness function (15) is calculated for all regressors and then the most significant are selected for generating new population of individuals (regressors). It is made by using the inequality

$$\Phi_M \geq \Phi^*, \quad \Phi_M = \frac{\sum_{m=1}^M \phi_m}{\sum_{\tilde{m}=1}^{\tilde{M}} \phi_{\tilde{m}}}, \quad (16)$$

$$\frac{M}{\tilde{M}} \leq \alpha, \quad (17)$$

where  $\Phi^*$  is a constant which characterizes the integer significance of the selected regressors to the sum of all reconstructed regressors with fitness function ( $\Phi^* \geq 0.8$ ),  $\tilde{M}$  is the total number of reconstructed regressors,  $M$  is

assumptions. Let  $i$  be the number of iterations (in each circle it is equal to the number of the most significant regressors  $i = \overline{1, M}$ ), and  $i = 1$  correspond to the traversal result  $F_1^{l_1}$ , and continue traversal  $i \rightarrow i + 1$ . Suppose that the traversal result is  $F_1^{l_1}, F_2^{l_2}, \dots, F_M^{l_M}$ , so the corresponding model for  $y(k)$  is

the number of selected regressors,  $\alpha$  is a constant value for constraint of the selected regressors number according to (16),  $0 < \alpha \ll 1$ . The process of regressors selection lasts until conditions (16) and (17) are met or the maximal number of iterations (10000) is reached. If the condition (16) is not satisfied, it is possible to weaken the constraint (17). After the most significant regressors are selected the model parameters  $\mathbf{c}$  on the data set  $N_a^* + N_b^*$  are estimated using the global search genetic algorithm. Then the model accuracy is tested on the data set  $N$ . If the result is unsatisfactory for us the identification process is repeated with training data set redistributing and weakens constraints.

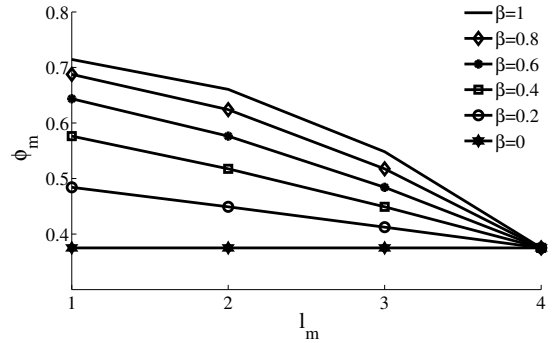


Fig. 2. The dependence of the fitness function penalty term  $\phi_m$  from the complexity control parameter  $\beta$  ( $r_m = 0.75$ ).

#### IV. RESULTS

Two models, linear and nonlinear, were reconstructed with the constraints on lags  $n_u = 25$ ,  $n_y = 24$ , for  $Dst$ -index guaranteed prediction. The first model was identified only from most significant linear regressors  $\beta = 1$

$$y(k) = 1.25y(k-1) - 0.35y(k-2) + 2.5u(k) + 0.15y(k-3) + 0.01y(k-6) - 2.74u(k-1) + 0.95u(k-2) - 0.2u(k-4) + 0.14u(k-8), \quad (18)$$

and the second one was identified using linear and nonlinear nodes  $\beta = 0$

$$\begin{aligned}
 y(k) = & 2.5u(k) - 3.13u(k-1) + 0.65u(k-2) - 0.13u(k-4) + 0.19u(k-8) \\
 & + 0.11u(k-12) - 0.03u(k-14) + 0.07u(k-25) + 1.17y(k-1) \\
 & - 0.41y(k-2) + 0.12y(k-3) + 0.04y(k-6) + 0.02y(k-12) \\
 & - 0.004y(k-17) - 0.03u(k-3)y(k-9) + 0.01u(k-7)y(k-12) \\
 & + 0.01y(k-15)u(k-1) + 0.03u(k-15)u(k-6) \\
 & - 0.003y(k-13)y(k-17) + 0.03y(k-20)u(k-4) \\
 & - 0.05y(k-21)u(k-4) - 0.35u(k-1)u(k-1) \\
 & + 0.03u(k-16)y(k-2) - 0.022y(k-12)u(k-2) \\
 & + 0.002y(k-15)y(k-12) - 0.2u(k-1)u(k-13).
 \end{aligned} \tag{19}$$

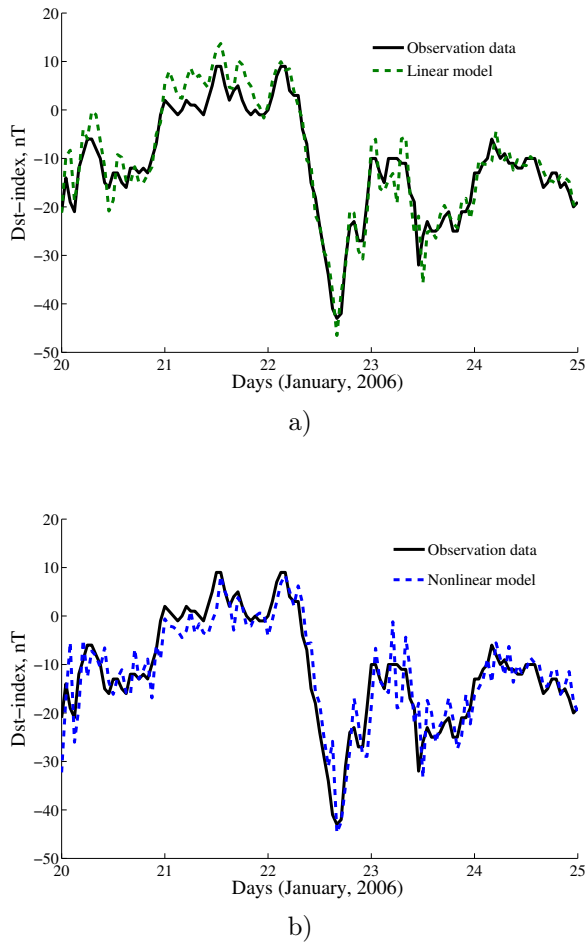


Fig. 3. Illustration of  $Dst$ -index modelling: a) — linear model (18); b) — nonlinear model (19).

The structure and parameters of models (18) and (19) were numerically determined using the identification algorithm and training data set of measurements for  $N = 2000$  hours. The statistical characteristics of these models are the linear correlation coefficient (LC) and the residual mean square (RMS): model (18) LC = 0.96, RMS = 3.36; model (19) LC = 0.97, RMS = 3.19. Parameter (13) was used for testing models accuracy. On Fig. 3 the results of  $Dst$ -index behaviour modeling using

models (18) and (19) are shown together with observations. The time scaled window of 150 hours was chosen for the analysis of  $u(k)$  behaviour and determination of maximal changes of the  $u(k)$  and  $\delta$  calculation. On Fig. 4 the results of  $Dst$ -index prediction interval estimation with models (18) and (19) are shown.

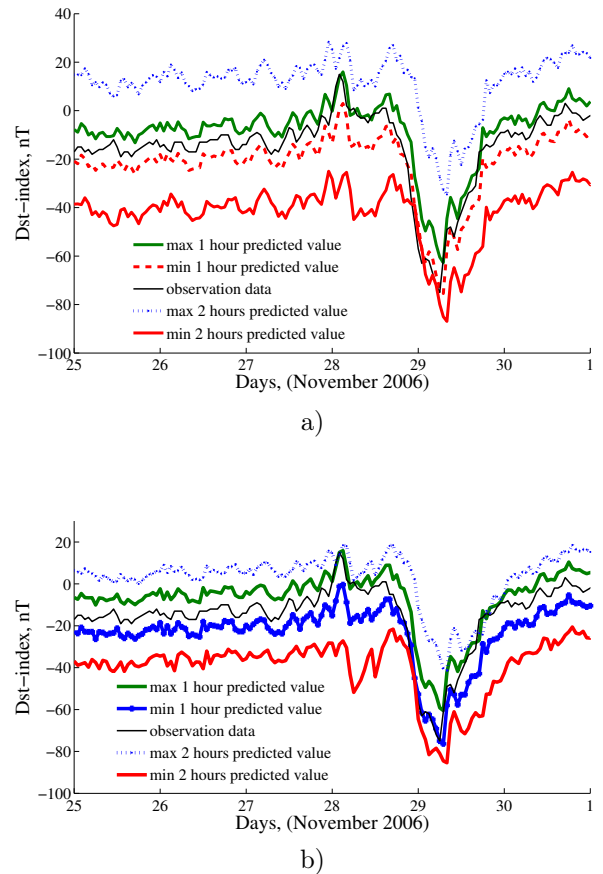


Fig. 4.  $Dst$ -index observation data and results of guaranteed prediction for one and two hours: a) — linear model (18); b) — nonlinear model (19).

Let us analyze the interval  $d_q$  when  $q = 1$  for the simplest case, on the example of the linear model (18). Let us substitute (18) in (9) and (10)

$$\begin{aligned} \underline{y}(k+1) = \min & \left[ 1.255y(k) - 0.355y(k-1) + 0.153y(k-2) + 0.013y(k-5) \right. \\ & + 2.5(u(k) - \Delta u(k)) - 2.74u(k) + 0.95u(k-1) - 0.2u(k-3) \\ & \left. + 0.14u(k-7) \right], \end{aligned} \tag{20}$$

$$\begin{aligned} \overline{y}(k+1) = \max & \left[ 1.255y(k) - 0.355y(k-1) + 0.153y(k-2) + 0.013y(k-5) \right. \\ & + 2.5(u(k) + \Delta u(k)) - 2.74u(k) + 0.95u(k-1) - 0.2u(k-3) \\ & \left. + 0.14u(k-7) \right], \end{aligned} \tag{21}$$

and then conduct the numerical modeling of the  $\{\overline{y}(k+1); \underline{y}(k+1)\}$  values subject to  $\Delta u_j(k)$ . The value  $\overline{y}(k+1)$  reaches its maximum at  $\Delta u_j(k) = \delta$ , and  $\underline{y}(k+1)$  its the minimum at  $\Delta u_j(k) = -\delta$ . Then

$$\underline{y}(k+1) = f(y, u) - 2.5\delta,$$

$$\overline{y}(k+1) = f(y, u) + 2.5\delta,$$

where

$$\begin{aligned} f(y, u) = & 1.255y(k) - 0.355y(k-1) + 0.153y(k-2) + 0.013y(k-5) \\ & - 0.24u(k) + 0.95u(k-1) - 0.2u(k-3) + 0.14u(k-7). \end{aligned}$$

Let us represent the change of the interval  $d_1$  as

$$d_1 = |\underline{y}(k+1) - \overline{y}(k+1)| \leq 5\delta + 2\varepsilon,$$

where  $\varepsilon$  is the model error.

Now let us consider the dependence of  $d_q$  from  $\delta$  during the estimation of  $\tilde{y}(k+q)$ ,  $q = 2$ . For this purpose let us rewrite (20) and (21) for  $k+2$

$$\underline{y}(k+2) = f'(y, u) - 5.3975\delta,$$

$$\overline{y}(k+2) = f'(y, u) + 5.3975\delta,$$

where

$$\begin{aligned} f'(y, u) = & 1.22y(k) - 0.29y(k-1) + 0.19y(k-2) + 0.01y(k-4) + 0.02y(k-5) \\ & + 0.89u(k) + 1.19u(k-1) - 0.2u(k-2) - 0.25u(k-3) \\ & + 0.14u(k-6) + 0.18u(k-7). \end{aligned}$$

As a result the interval  $d_2$  of the value  $\tilde{y}(k+2)$  can be represented as

$$d_2 = |\underline{y}(k+2) - \overline{y}(k+2)| \leq 10.795\delta + 4.51\varepsilon.$$

Obviously, when  $q$  increases, then  $d_q$  value will also increase. Also it can be seen that the use of nonlinear model in the case  $q > 2$  leads to a quicker increase of  $d_q$  value. However, by using the optimization methods it is possible to conduct the adaptive calculation of value  $\delta$  when interval  $d_q$  is minimal. In further research, there would be examined the problem of prediction interval estimation solving the min-max mathematical programming problem.

---

Our predictions were also compared with results of other authors, [10, 11, 25, 31], and [26]. Our method provides a much more precise guaranteed forecast than most empirical and typical neural network models. It is hard to calculate the statistical characteristics of the models accuracy as in our method the bounds of the interval are predicted. Earlier models for *Dst*-index prediction suffered from time shifting. Our nonlinear model is almost free from this effect, because the minimal value of the in-

terval will be always lower than the original value ( $LC = 0.86$  and  $RMS = 10.75$ ). So, that is why the guaranteed prediction method is more accurate for strong storms prediction, which are of greater interest. The huge advantages of the guaranteed prediction method are: firstly, it gives the opportunity to predict the interval “tube” of the possible geomagnetic index values; at secondly, the identification of the mode structure and parameters should be performed only once. The prediction itself takes about 10 seconds on an average PC, which allows creating fully automated operational online space weather forecast services.

## V. CONCLUSION

A new guaranteed approach to geomagnetic activity prediction using satellite observations of solar wind parameters has been proposed. It is based on maximal and minimal index values forming “interval tube” and on a nonlinear discrete “input–output” dynamical system. The model structure and parameters are chosen by solving a nonlinear mathematical programming problem with constraints. The advantages of this approach are as follows: 1) automatic selection of regressors for nonlinear models; 2) solution of the mathematical programming problem by genetic algorithm for selecting the optimal

model structure and unknown parameters. The nonlinear dependence between input model parameters and prediction interval has been discovered. It was shown that the prediction error of  $Dst$ -index is nonlinearly rising with the increase of the prediction time. The numerical calculation shown that time horizon for  $Dst$ -index forecast can be increased by using an optimization approach. A new model for the prediction of  $Dst$ -index on the basis of the solar wind has been proposed. A prospective real time dynamical prediction of  $Dst$ -index has been discussed

## ACKNOWLEDGMENT

We thank the ACE MAG and SWEPAM instruments teams and the ACE Science Center for providing the ACE data. The OMNI-data were obtained from the GSFC/SPDF OMNIWeb interface at <http://omniweb.gsfc.nasa.gov>. We acknowledge the use of the geomagnetic data from Kyoto WDC for Geomagnetism (<http://swdcd.kugi.kyoto-u.ac.jp>), Geoforschung Zentrum Potsdam (<http://www-app3.gfz-potsdam.de>). The research leading to these results has received funding from the European Commission’s Seventh Framework Programme (FP7/2007-2013) under the grant agreement #263506 (AFFECTS project, [www.affects-fp7.eu](http://www.affects-fp7.eu)).

- 
- [1] S. I. Akasofu, S. Chapman, *An Account of the Wave and Particle Radiations from the Quiet and the Active Sun, and of the Consequent Terrestrial Phenomena* (Clarendon Press, Oxford, 1972).
  - [2] M. Balikhin, I. Bates, S. Walker, *Adv. Space Res.* **28**, 1123 (2001).
  - [3] A. Galeev, V. Krasnoselskikh, V. Lobzin, *Sov. J. Plasma Phys.* **14**, 697 (1988).
  - [4] H. Gleisner, H. Lundstedt, P. Wintoft, *Ann. Geophys.* **14**, 679 (1996).
  - [5] K. Schindler, *Physics of Space Plasma Activity* (Cambridge University Press, Cambridge, 2007).
  - [6] J. Goodman, *Space Weather, Telecommunication* (The Springer International Series in Engineering and Computer Science, Vol. 782, 2005).
  - [7] V. Bothmer, I. A. Daglis, *Space Weather – Physics and Effects* (Springer-Verlag, 2010).
  - [8] M. Moldwin, *An Introduction to Space Weather* (Cambridge University Press, Cambridge, 2008).
  - [9] S. Billings, Q. Zhu, *Int. J. Control* **60**, 1107 (1994).
  - [10] A. Parnowski, *J. Phys. Stud.* **12**, 4003 (2008).
  - [11] A. Parnowski, *Astrophys. Space Sci.* **323**, 169 (2009).
  - [12] J. Sharifi, B. N. Araabi, C. Lucas, *Earth Planets Space* **58**, 331 (2006).
  - [13] M. Temerin, X. Li, *J. Geophys. Res.* **111**, A04221 (2006).
  - [14] OMNI2 database. National Space Science Data Center / Space Physics Data Facility. <http://nssdc.gsfc.nasa.gov/omniweb/> (Cited 24 Sep 2014).
  - [15] S. Chen, S. Billings, *Int. J. Control* **49**, 1013 (1989).
  - [16] I. Leontaritis, S. Billings, *Int. J. Control* **41**, 303 (1985).
  - [17] I. Leontaritis, S. Billings, *Int. J. Control* **45**, 311 (1987).
  - [18] O. Cheremnykh, V. Yatsenko, O. Semeniv, Iu. Shatokhina, *Phys. J.* **53**, 504 (2008).
  - [19] O. Semeniv, V. Sidorenko, Yu. Shatokhina, O. Cheremnykh, V. Yatsenko, *J. Automat. Inf. Sci.* **40**, 8 (2008).
  - [20] O. Semeniv, in *UK–Ukraine Meeting on Solar Physics and Space Science joint with Topical Advanced Summer School SPSS/TASS* (2011); [http://swat.group.shef.ac.uk/Conferences/Ukraine\\_UK\\_2011/Participations](http://swat.group.shef.ac.uk/Conferences/Ukraine_UK_2011/Participations).
  - [21] R. Burton, R. McPherron, C. Russell, *J. Geophys. Res.* **80**, 4204 (1975).
  - [22] S. Billings, S. Chen, *Int. J. Control* **50**, 1873 (1989).
  - [23] D. McCaffrey *et al.*, *Adv. Space Res.* **25**, 1571 (2000).
  - [24] V. Yatsenko, O. Cheremnykh, V. Kuntsevich, O. Semeniv, *J. Automat. Inf. Sci.* **41**, 12 (2009).
  - [25] T. P. O’Brien, R. L. McPherron, *J. Atm. Sol.-Terr. Phys.* **62**, 1295 (2000).
  - [26] G. Palocchia *et al.*, *Mem. Soc. Astron. It. Suppl.* **9**, 120 (2006).
  - [27] J. Koza, *Genetic Programming: On the programming of Computers by Means of Natural Evolution* (The MIT Press, 1998).
  - [28] J. Koza, *Genetic Programming IV: Routine Human Competitive Intelligence* (Kluwer Academic Publishers, 2003).
  - [29] K. Rodríguez-Vázquez, C. M. Fonseca, P. J. Fleming, *IEEE Trans. Systems Man Cybernet.* **34**, 531 (2004).
  - [30] M. Afenzeller, S. Winkler, S. Wagner, A. Beham, *Genetic*

*Algorithms and Genetic Programming: Modern Concepts and Practical Applications* (Chapman and Hall/CRC, 2009).

[31] C. Cid *et al.*, in *Solar Wind 11/SOHO-16: Connecting Sun and Heliosphere*, June 13–17, 2005, p. 116.

**КОМБІНОВАНИЙ ПІДХІД ПРОГНОЗУВАННЯ КОСМІЧНОЇ ПОГОДИ  
НА ОСНОВІ МЕТОДУ ГАРАНТОВАНОГО ОЦІНЮВАННЯ  
ТА ЕВОЛЮЦІЙНОГО АЛГОРИТМУ**

О. В. Семенів

*Інститут космічних досліджень НАНУ та ДКАУ  
просп. Глушкова, 40, корп. 4/1, Київ, 03680, Україна*

Розглянуто проблему прогнозування космічної погоди з використанням методу гарантованого оцінювання та параметрів сонячного вітру. Запропоновано оригінальний алгоритм прогнозування геомагнітного *Dst*-індексу на основі дискретних динамічних моделей, еволюційного алгоритму та гарантованого інтервального оцінювання. Представлено проблему ідентифікації дискретної моделі у вигляді розв'язання задачі нелінійного програмування з обмеженнями. Досліджено залежність точності прогнозування від горизонту прогнозування та побудовано модель прогнозування геомагнітних буревіїв.