Journal of Atmospheric and Solar-Terrestrial Physics 71 (2009) 589–596

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar-Terrestrial Physics

journal homepage: www.elsevier.com/locate/jastp

Hybrid model for nowcasting and forecasting the K index

I. Kutiev [a,*], P. Muhtarov [a], B. Andonov [a], R. Warnant [b]

[a] Geophysical Institute, Bulgarian Academy of Sciences, Sofia, Bulgaria
[b] Royal Meteorological Institute, Brussels, Belgium

ARTICLE INFO

Article history:
Received 27 May 2008
Received in revised form
12 December 2008
Accepted 19 January 2009
Available online 29 January 2009

Keywords:
Geomagnetic K index
Empirical modeling
Solar wind-based model of K
Nowcasting and forecasting geomagnetic activity

ABSTRACT

A new empirical model nowcasting and predicting a proxy to the geomagnetic K index is developed, which is based on the combined use of solar wind parameters and ground-based magnetic data. The present approach implements the previously developed solar wind-based MAK model, calibrating its values with magnetogram-derived K index. The new model is named as Hybrid Dourbes K (HDK) model. The HDK nowcast model provides the quantity Kdf, obtained by solar wind-based Ksw and corrected with a combination of differences between several past values of Kd and Ksw. The model error of the nowcast Kdf is found to be 0.38 KU, or nearly twice less than that of the MAK model. Kdf has a good predictability. Prediction made by weighted extrapolation 6 h ahead carries an error of 1.0 KU, while for the first 1 h the error is 0.58 KU only.

1. Introduction

The planetary geomagnetic index Kp is widely used in many research areas and space weather applications for representing geomagnetic activity. Being a good measure of magnetospheric dynamics (Thomsen, 2004), Kp at the same time is available with a delay of few weeks from the current time (Wing et al., 2005). The need of nowcasting and forecasting for many on-line applications stimulates development of various models of Kp. Most of these models are based on the non-linear neural network (NN) technique (Costello, 1997; Boberg et al., 2000; Balikhin et al., 2001; Boaghe et al., 2001; Wing et al., 2005), taking advantage of the routinely available solar wind measurements of Advanced Composition Explorer (ACE) at L1 libration point.

Some models use linear regression technique to fit Kp variations to pre-defined analytical expressions containing dependences on solar wind parameters. These expressions are mostly based on the analogy with well-described physical mechanisms. Muhtarov and Andonov (2000) introduced an expression relating Kp to the Interplanetary Magnetic Field (IMF) Bz by using an analogy with voltage transfer in a electric diode rectifier circuit (DRC). The circuit used Bz as an input voltage, while the output voltage they defined as a "modified" Bz (Bzm), containing a delayed reaction to Bz. Later, Andonov et al. (2004) improved the model expression, adding terms of solar wind velocity and dynamic pressure. This model, named MAK, is a part of the hybrid model and will be described briefly below.

The official Kp index is issued by GeoForschungsZentrum (GFZ) in Potsdam, Germany (http://www.gfz-potsdam.de) with a few weeks delay. For space weather applications operated in real-time mode, a number of models provide nowcast of the Kp values by using different prediction techniques. Takahashi et al. (2001) has developed an automated procedure for near-real-time Kp estimates using on-line magnetograms from 9 midlatitude stations. They found a correlation coefficient of 0.93 between official Kp and their estimated values. These authors have also found that the accuracy of estimated Kp remains high if the number of stations is reduced to one or two. The Kp models, using solar wind parameters, provide a Kp estimate highly correlated with the magnetogram-derived Kp index. The Kp estimate can be made available in nearly real-time, as soon as ACE measurements become available. Bearing in mind that ACE spacecraft at L1 libration point (at 1.5 million km from the Earth towards Sun) measures solar wind parameters 30–60 min in advance of its arrival to Earth's magnetosphere, these models provide actually nowcast Kp values. The solar wind-based models, however, provide less accurate values; the standard deviation between the model and ground-based Kp (RMS error) is of order of one K unit (KU).

The purpose of the paper is to present a model that combines the advantages of the near-real-time Kp estimates from solar wind parameters and the higher accurate magnetogram-derived Kp values. Similar to Takahashi et al. (2001), instead of Kp, we use magnetogram-derived K index from the station of Dourbes, Belgium. In our approach, the quantity Ksw, provided by the solar wind-based MAK model, is corrected by the 3-h K index provided also in near-real-time from the ground-based magnetograms. This approach assures an immediate reaction of K to sudden changes of

* Corresponding author.
E-mail address: kutievi@geophys.bas.bg (I. Kutiev).

1364-6826/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jastp.2009.01.005

IMF Bz, as at the same time its variations are forced to stay close to that of ground-based K. The new hybrid model, denoted as Hybrid Dourbes K (HDK) model, provide a nowcast value Kdf, which is then used in a prediction of 10–15 h ahead. HDK model is developed in a framework of GALOCAD project (GJU/06/2423/CTR/GALOCAD), a "Galileo Local Component for nowcasting and forecasting atmospheric disturbances affecting the integrity of high precision Galileo applications". This project aims to perform a detailed study on ionospheric small- and medium-scale structures and to assess the influence of these structures on the reliability of Galileo precise positioning applications.

Further in the paper we use the K index in three different forms: the 3-h K index derived from Dourbes magnetograms, its 1-h interpolated values Kd, and Kdf values obtained from the models (Ksw from solar wind data and Kdf from the combination of Ksw and Kd). The hourly interpolated values are required by more frequent input in some applications. While the first quantity is the "authentic" K index; we consider the other two as new quantities, proxy to the K index. Indeed, the 3-h K index scales (in quasi-logarithmic scale) the difference between the lowest and maximum deviations from quiet magnetic field within 3-h intervals (Mayaud, 1980). We assign these K values to the end of the 3-h intervals they are obtained (for example, K obtained in the interval 00–03 UT is referred to 03 UT) and then digitize of the symbols "+" and "−", as the sign "−" shifts K value 0.33 below the integer number, and the sign "+" increases the value with 0.33 above this number. To check the reliability of the linear interpolation of K index on the logarithmic scale, we compared linearly interpolated K values taken over 2 years of data (2003–2004) with respective interpolated ap values (which quantify magnetic disturbances in a linear scale) and found that the mean difference is 0.066 KU, which is five times less than the K resolution of 0.33. This fact assures that the linear interpolation of K index does not introduce additional error in the model performance. As was pointed out by Thomsen (2004), K index best represents the processes in the near-Earth magnetosphere, having a comparative time scale. It is clear that the 1-h K does not have such clear physical meaning and be a measure of any identifiable geophysical process, therefore it cannot be considered as "true" geomagnetic index. Nevertheless, we use the hourly interpolated K index to increase statistical base for model quantities. We suppose that once considered as a measure of inner magnetospheric processes, the interpolated K index obtains a new legitimacy and could be used in various applications in more flexible way. To distinguish between the "true" geomagnetic K index and its model derivatives, we will further denote the latter as "K estimates".

2. MAK model

Muhtarov and Andonov (2000) found that the cross-correlation between Bz and Kp had a maximum at a time lag of about 2 h, meaning that Kp best correlates with Bz from the previous 2 h. In order to improve the dependence of Kp on Bz, they introduced a new function of Bz (denoted as Bzm), which is positive and contains a delayed reaction to Bz changes. To do this, they used an analogy with DRC circuit, which involves loading and unloading processes with different time constants. The electrical circuit shown in Fig. 1a represents such a system. The circuit includes a half-wave diode rectifier D, a smoothing capacitor C and two resistors R_1 and R_2. If the input voltage U_1 is a step-like function,

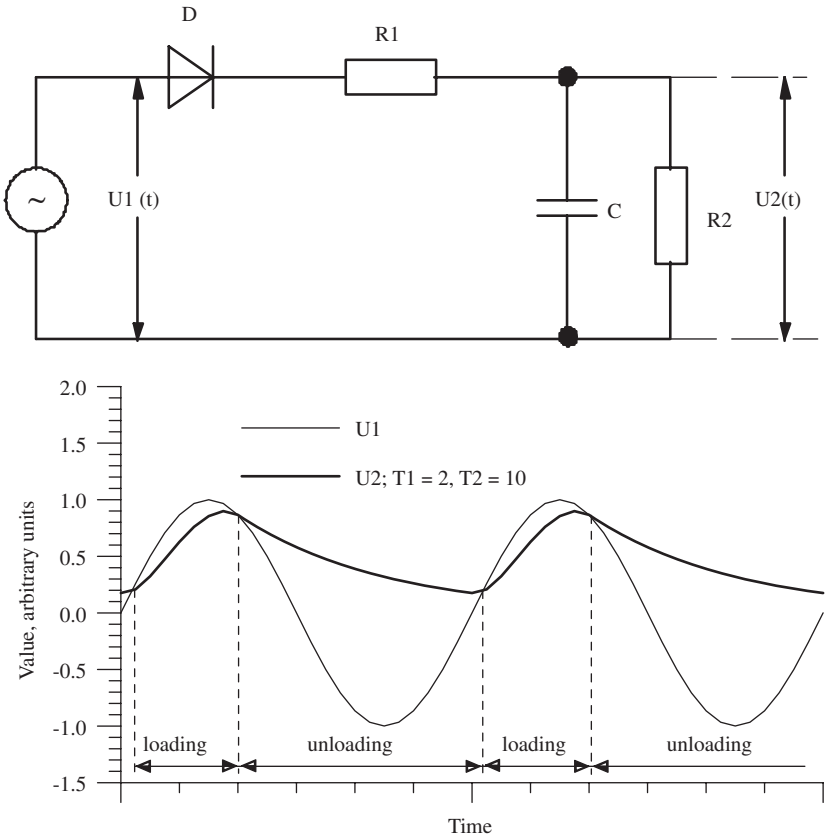

Fig. 1. (a) Equivalent electric circuit DRC, giving the delayed reaction of the output voltage U_2 to the input voltage U_1. D is a half-wave diode rectifier, C is a capacitor, R_1 and R_2 are resistors; (b) a sample of a sinusoidal input voltage U_1 and the resulting output voltage U_2. Loading and unloading phases are marked with horizontal bars. T_1 and T_2 are assured arbitrary to demonstrate the functioning of the circuit and do not relate to the values considered in the paper.

formed by discrete values at arbitrary moments of time and $R_2 \gg R_1$, the output voltage U_2 within the time-step $[t_i, t_i{+}1]$ is given by the well-known relationship:

$$U_2(t[i+1]) = \begin{cases} (U_2(t_i) - U_1(t_i))\ \exp(-t[i+1] - t_i T_1) + U_1(t_i), & U_2(t_i) < U_1(t_i), \\ U_2(t_i)\ \exp\left(-\frac{t[i+1]-t_i}{T_2}\right), & U_2(t_i) > U_1(t_i) \end{cases} \quad (1)$$

where $T_1 = R_1C$ and $T_2 = R_2C$. The expressions (1) have recurrent feedback: the voltage U_2 obtained from the previous step $[t_{i-1}, t_i]$ is placed on the right-hand side of the equations for obtaining U_2 in the next step $[t_i, t_{i+1}]$. The first expression in (1) represents a process of loading the capacitor C with a time constant T_1, while the second expression represents the unloading process with a time constant T_2. The loading takes place while U_1 is higher than U_2 and the diode is open. If U_1 becomes lower than U_2, the diode is closed and the capacitor starts discharging through the resistor R_2. The whole process is schematically presented in Fig. 1b. The input voltage U_1 is represented as a simple sinusoid (thin line) and the output voltage U_2 is given by the solid line. The loading takes place when $U_1 > U_2$. The output voltage U_2 accepts now only positive values, gradually decreasing when U_1 is lower. Taking use of the analogy with this electrical scheme, Muhtarov and Andonov (2000) defined Bzm through Eq. (1) with a replacement of U_1 with $-$Bz and U_2 with Bzm. Using 27 years of IMF data (1973–1999), Muhtarov and Andonov (2000) estimated that Bzm improved the correlation between Bz and Kp from -0.4 to 0.7. Later, Andonov et al. (2004) further improved the model, denoted as MAK, adding dependences on solar wind dynamic pressure and velocity.

Andonov et al. (2004) calculated the cross-correlation between Kp and all available solar wind parameters. Ranking the parameters by their cross-correlation with Kp, they included in the model, besides Bzm, the dependences on solar wind velocity V and dynamic pressure P. Based on the form of the average dependences, the MAK model was given as follows:

$$\mathrm{Ksw} = a_o + a_1\mathrm{Bzm} + a_2P + a_3V + a_4\mathrm{Bzm}^2 + a_5P^2 \quad (2)$$

Ksw represented the model value of (2), which presumably was very close to Kp. Coefficients a_i were obtained by fitting the expression on the right side of (2) to the ground-based Kp values. Solar wind parameters were taken this time from Advanced Composition Explorer space platform for the years 1998–2004. Because the ACE database was compiled from hourly values of Bz, V and P, the 3-h Kp values for the same period were interpolated to obtain also hourly values. To make this, they converted the Kp grades around each unit into decimal numbers by adding -0.33, 0 and +0.33. For example, 2^-, 2^0, 2^+ were converted to 1.66, 2.0 and 2.33. Performing the fitting of expression (2), Andonov et al. (2004) first obtained the time constants T_1 and T_2 in Bzm (expression (1)). The time constants were obtained in the following way. The right-hand side of (2) was repeatably fitted to Kp values from the whole database by using grid of pairs (T_1, T_2). Corresponding root mean square (RMS) deviations of model from data were calculated for each pair and then compared. Andonov et al. (2004) defined the time constants as the pair (T_1, T_2) having the minimum RMS deviation. Using ACE database in period 1998–2004, MAK found $T_1 = 3$ and $T_2 = 7$ h. Therefore, the time constants of the delayed reaction of Kp to Bz forcing were obtained as average over the whole database. Andonov et al. (2004) estimated the average root mean square deviation between the observed Kp and model Ksw (overall model error). MAK model error was estimated at 0.63 KU, while the model error by using Bzm only was 0.96 KU.

3. HDK model development

The flow chart of HDK model, as designed for the GALOCAD project, is shown in Fig. 2. HDK model consists of two branches, which are combined in the last stage to produce the final product, e.g. the nowcast Kdf. The upper branch represents MAK model with re-calculated coefficients for Dourbes K data. The lower branch represents the production of the Kd values by sliding the 3-h time window with 1-h step. At the end of the chart, both branches merge to obtain Kdf. Production of the hourly Kd values from magnetograms is a separate task and will not be included in this paper. For the present analysis, the routinely issued 3-h K index at Dourbes will be linearly interpolated to obtain the hourly Kd values. Further, we denote the 3-h K index as "Dourbes K" and the hourly interpolated values as "Kd". Therefore, all derivatives of the Dourbes K index used by HDK model are K estimates, not necessarily connected with magnetic disturbances, from where the 3-h K index originate. The model development and error estimates will be based on retrospective K data. The errors will be evaluated separately for the interpolated hourly Kd and the 3-h K values.

3.1. Adoption of MAK model to Dourbes K index

To adapt MAK model to Dourbes K data, we first obtained the cross-correlation coefficient between the 3-h Kp and Dourbes K over the whole database, in order to check whether the average dependences of K on solar wind velocity and pressure remain the same. Indeed, the maximum cross-correlation of all time-intervals was placed at the time lag = 0, which means that there is no delay between the variations of the two indices. Fig. 3 shows the cross-correlation between Dourbes K and Kp at time lag = 0. It has a maximum of 0.929 in time interval (00–03) h and minimum of 0.869 in (09–12) h interval. The local time difference is expected, because of the asymmetry of current system in the auroral oval. The high cross-correlation between Dourbes K and Kp assures that the MAK definitions are applicable to Dourbes K index. The respective coefficients in Eq. (2) are given in Table 1. The time constants T_1 and T_2 were also re-calculated. Their average values remain the same as for the Kp-based model: $T_1 = 3$ h and $T_2 = 7$ h, which is reasonable in view of the high correlation between Kp and Dourbes K.

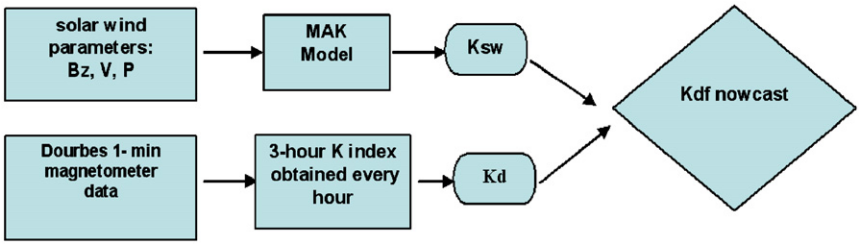

Fig. 2. Flow chart of HDK model development. Upper branch represents adoption of MAK-derived Ksw to Dourbes K index. Lower branch represents production of the K index in hourly basis.

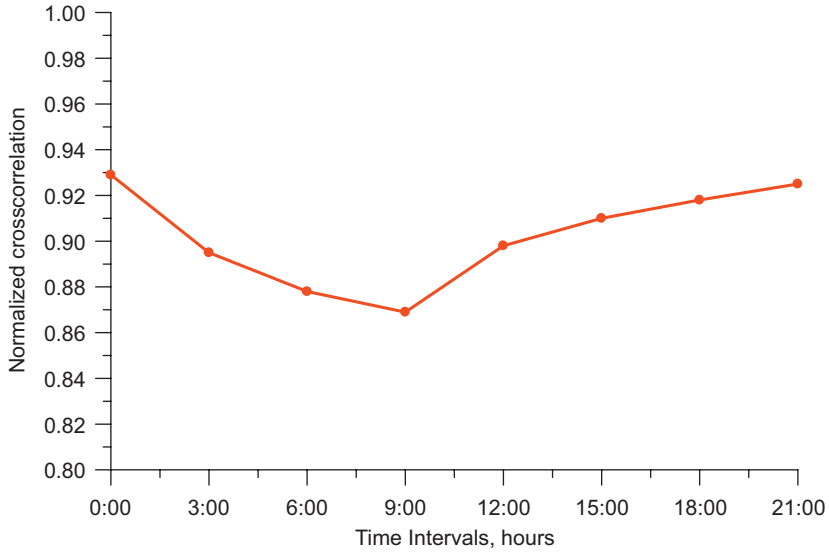

Fig. 3. Local time variation of the cross-correlation between 3-h Kp and Dourbes K indices at time lag = 0.

Table 1
Coefficients in Eq. (2) for Dourbes K data.

a0	a1	a2	a3	a4	a5
−0.687897	0.350934	0.188879	0.004763	−0.009347	−0.005151

3.2. HDK nowcasting algorithm

The main idea of combined use of Ksw and Kd is that at any given moment Ksw is corrected by the differences between Ksw and Kd (named Delta) obtained at some past moments. We use the fact that at the current moment t, Ksw is a measure of the current conditions in the solar wind (within the last hour, if we deal with its hourly values), while Kd represents the averaged conditions of the Earth's magnetic field in the last 3 h. So, at the current moment t, HDK provides a value Kdf, which is based on MAK Ksw at the same moment, corrected by a combination of weighted Delta values at some previous hours. While Ksw is more variable, Kd stays more conservative; therefore, the nowcast Kdf should be tightly connected to Kd. In the nowcast algorithm this is assured by the double action of Delta and its mean value MDelta.

To perform nowcasting, the HDK model uses the method of weighted extrapolation of Wienner–Hopf (Korn and Korn, 1968). This method was first applied to ionospheric short-term predictions by Muhtarov and Kutiev (1999). In this method, the weighting coefficients, assigned to the past data, are obtained by the autocorrelation function of respective extrapolated quantity. If the autocorrelation function can be approximated by an exponential function, as is the case in most ionospheric processes, the weighting factors are presented by simple exponents with increment $(t-\tau)/T$; τ is the hours in the past. For this type of autocorrelation functions, the weighted extrapolation is based on nearest (first) past value only. The time constant T of the approximated autocorrelation function is defined as the time at which the tangent drawn from the time lag = 0 crosses the abscissa. At this time lag, the autocorrelation function decreases e times. Fig. 4 shows the empirical autocorrelation function of the difference Delta = Ksw–Kd (solid blue line) and its exponential approximation (dashed line). The difference Delta is not so well autocorrelated; its time constant of 6 h is much less compared with 20-h time constant of Kd. We explain this with the larger variability of Ksw. The autocorrelation of Delta deviates from exponential form at time lags above 6 h. Fortunately, for time lags shorter than the time constant, Wienner–Hopf extrapolation is well applicable. In the nowcasting procedure, we use not only differences Delta, but also deviations of individual Delta values from their average in the last 6 h, named MDelta. This double deviation (Delta–MDelta) assures stronger control over the nowcasting Kdf during the large excursions of Ksw. The nowcasting is now performed by a simple extrapolation of the double deviation (Delta–MDelta) from previous hour $(t-1)$ to the current hour t. Note that the deviation of Delta from its average has the same (normalized) autocorrelation function as that of Delta. The Kdf at the current moment t is defined by the following expression:

$$\mathrm{Kdf}(t) = \mathrm{Ksw}(t) - \mathrm{MDelta} - [\mathrm{Ksw}(t-1) - \mathrm{Kd}(t-1) - \mathrm{MDelta}]\exp(-1/6) \quad (3)$$

Double deviation from previous hour (in brackets) is weighted by exp(1/6), in which $(t-\tau) = 1$ and $T = 6$. HDK nowcasting works in the following way. At the current moment t, the model calculates (predicts) Kdf by using Ksw and the average deviation MDelta. Kd at the time t is not taken into account. Kd actually goes one step behind and contributes to the term in brackets. Indeed, at the current moment t, Kd represents geomagnetic disturbances happened in the previous 3 h, while Ksw represents current conditions or that, which will happen in the near future. The nowcast timing will be discussed later in the paper. The term in brackets in (3) is usually small during quiet conditions. If at time t Ksw suddenly increases, Kdf will increase proportionally, bearing in mind that MDelta does not change. At next step, t+1, the term in brackets will be positive ($\mathrm{Ksw}(t) > \mathrm{Kd}(t)$) and tends to oppose an increase of Kdf(t+1) due to increased Ksw value. This effect will be illustrated later in Fig. 5.

If one or two recent Kd values are missing, the model (3) can still calculate Kdf, because instead of Kd from the previous hour, values from earlier hours may be used, with their respective weights. This is important property of the model, because it makes possible using magnetogram-derived K with larger sampling rate. For example, it is possible to use 3-h Dourbes K values to obtain MDelta. Then MDelta will be an average of two 3-h K values only.

Fig. 5 shows examples of HDK nowcast Kdf, along with Ksw and Kd for 2 periods with different geomagnetic activity. Fig. 5a

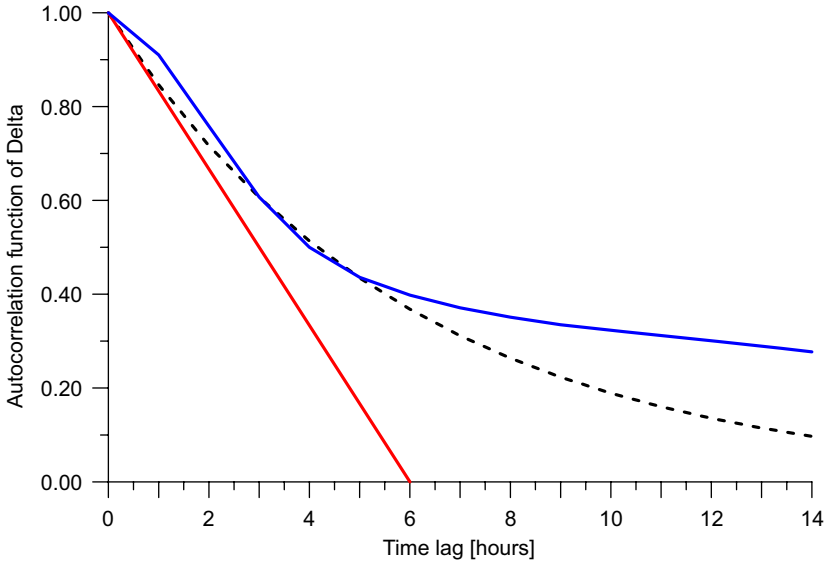

Fig. 4. Empirical autocorrelation function of Delta (solid line) and its exponential approximation (dashed line). The time constant of Delta is 6 h.

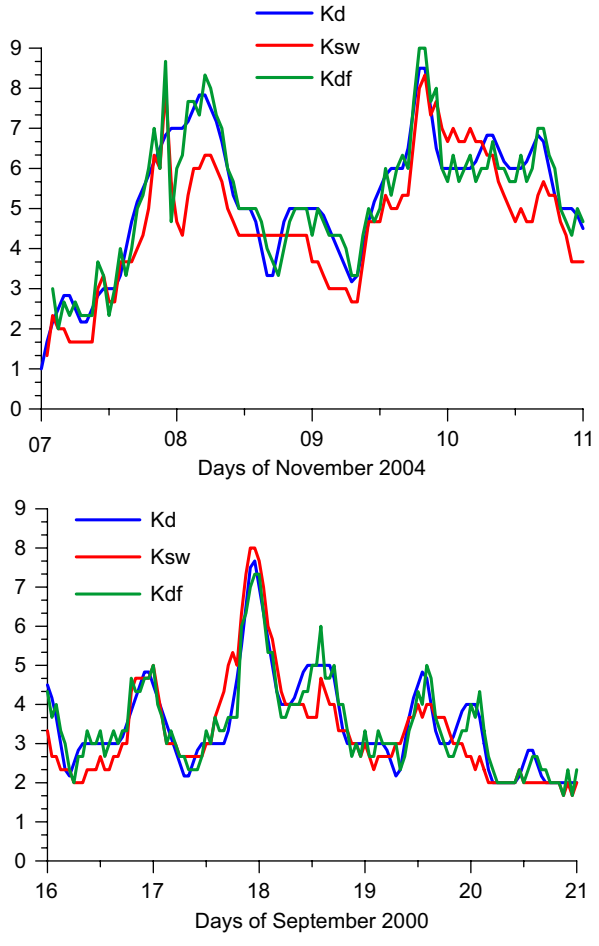

Fig. 5. Hourly interpolated Dourbes K (Kd), solar wind-derived Ksw, and the nowcasted Kdf during the period 7–10 November 2004 (a), and 16–20 September 2000 (b). Date labels are placed at the beginning of the day.

represents the period 7–10 November 2004, when two severe geomagnetic storms occurred on 7–8 and 10–11 November. Kdf (green line) follows much closer the ground-based Kd (blue line), than the solar wind Ksw (red line) does. It is interesting to follow the behavior of Kdf and Ksw during the late hours of 7 November. First Kd increases faster than Ksw and Kdf follows closer this increase. Few hours later Ksw exhibits a spike-like increase, followed closely by Kdf, while Kd values increase smoother. The back slope of Ksw cannot be compensated by Delta and for a short-time Kdf followed Ksw well below Kd values. Compensation from Delta takes effect few hours after sharp changes of Ksw and for the most of the time keeps Kdf closer to Kd. Fig. 5b shows a period with several modest storms (Kd not exceeding 5) and one major storm around 18 September 2000. Around noon on 17th September, Ksw sharply increases towards its maximum value of 8.3 at midnight. In this case Kdf stays closer to Kd during the whole storm. The small peak of Kdf in the afternoon hours on 18 September is provoked by a small pulse of Ksw on top of relatively high Kd values. The agreement between Kdf and Kd is excellent even in the periods when Ksw deviate form Kd.

3.3. *Error assessment of the nowcast model*

The HDK nowcast model performs surprisingly well. The model error (standard deviation between Kdf and Kd), estimated over the whole database, is 0.38 KU. Cross-correlation between Kdf and Kd at time lag = 0 is 0.97. Fig. 6 compares MAK and HDK nowcast model errors as depending on K magnitude. For the whole range of K, the reduction of model error is between 0.2 and 0.5 KU. The error reduction is especially large at higher K values, when predictions are most important for space weather applications.

As was mentioned above, the use of the 3-h K values in MDelta yields larger model error. Fig. 7 illustrates this fact. The red curve represents HDK nowcast error by using hourly Kd values and the red curve is the error by using the 3-h K values only. The difference here is of order of 0.4–0.6 KU in the whole K range. We have to note that interpolated hourly Kd values have lower dispersion than the 3-h K values, because the interpolated values triple the number of data points without contributing with their own scatter.

4. Development of the HDK forecasting technique

Kdf values exhibit a good reproducibility. Fig. 8 represents its autocorrelation function (blue line) and the exponential

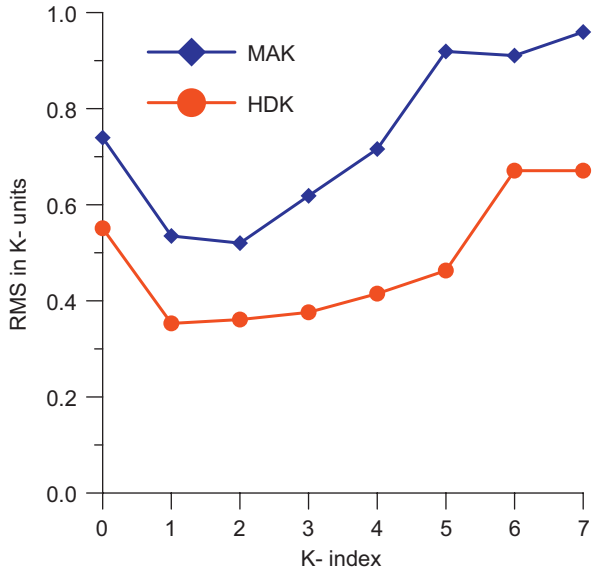

Fig. 6. Model errors of MAK (blue diamonds) and HDK nowcast (red dots) models. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

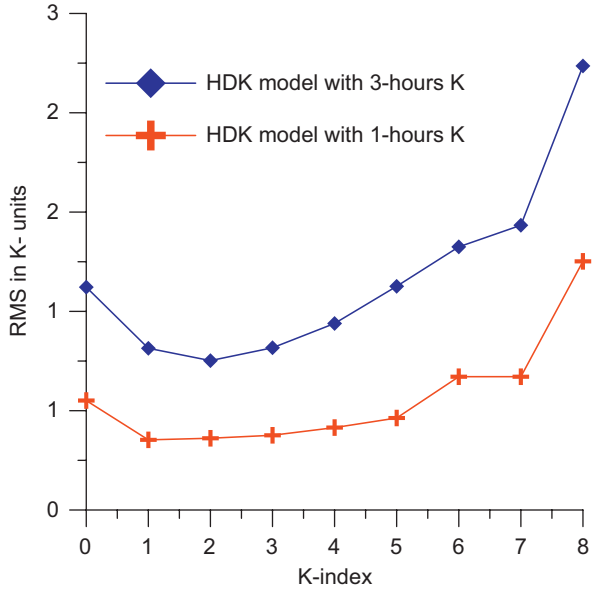

Fig. 7. Model errors of HDK using 3-h Dourbes K index (blue line) and hourly sampled Kd values. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

approximation (dashed line) with a time constant $T = 13$ h. To make prediction of Kdf, we use the same Wienner–Hopf method, as we did for the nowcasting. If we have the last obtained Kdf value at moment t, the predicted value Kdfp at the future moment $(t+\tau)$ is defined as

$$\mathrm{Kdfp}(t+\tau) = \mathrm{Kdfmean} + [\mathrm{Kdf}(t) - \mathrm{Kdfmean}]\exp(-\tau/13) \qquad (4)$$

Here Kdfmean, in analogy with MDelta, is the average of Kdf values from the past 13 h; τ is the lead time of prediction. Expression (4) clearly shows the main property of the weighted extrapolation: with increasing the lead time τ, Kdfp asymptotically approaches the average Kdfmean value.

Prediction error is defined as the standard deviation of predicted Kdfp from Kd values and estimated again over the whole database. In Fig. 9, the blue line represents the prediction error obtained by using 3-h K values and red line representing prediction error obtained by using the hourly interpolated Kd values. The expression (4) assures that predictions error up to 6 h ahead does not exceed 1.0 KU. This is, of course, a statistically averaged error; for individual cases the error could exceed the theoretical value.

5. Comparison with NN nowcasting and forecasting models

To our knowledge, HDK is the only empirical (linear regression) model, which combines solar wind and ground-based magnetic data for nowcasting and forecasting K index by using analytical expressions. There exist, however, models based on NN technique, which involve Kp along with solar wind parameters as input for their calculations. Wing et al. (2005) developed a prediction model, named APL and compared their results with those of three other NN models: Costello (Costello, 1997), NARMAX (Balikhin et al., 2001), and Lund (Boberg et al., 2000). In this comparison, Wing et al. (2005) have estimated model error by using correlation coefficient r, not RMS error. To make comparison, we also calculated HDK correlation coefficient between Kd and Kdf values. Table 2 presents a summary of the nowcast model errors, compiled from Wing et al. (2005) paper. First column shows the name of the models; the second column indicates their types. Third column gives the correlation coefficients and the fourth column shows the input parameters of the models; $|\mathbf{B}|$ is the absolute value of IMF. It is clear that APL and HDK models show higher correlation coefficients r, due to the fact that they combine ground-based K with solar wind parameters.

Wing et al. (2005) developed two models for predicting Kp for 1 and 4 h ahead. Unfortunately, NN methods cannot provide prediction formula, function of the lead time as that of (4). Therefore, each prediction time needs separate model. For 1 and 4 h lead time, Wing et al. (2005) models provide $r = 0.92$ and 0.79, respectively. The red bars in Fig. 10 represent correlation coefficient between measured Kd and predicted Kdfp values for 6 h of prediction. Numbers at the top axis show correlation coefficient of each prediction. The bar at lead time = 0 shows the correlation coefficient between Kd and Kdf. For prediction 1 and 4 h lead time, the HDK correlations are lower comparing with APL: 0.88 and 0.68 against 0.92 and 0.79. Unfortunately, we do not have enough information to discuss this comparison: what is database used, how they obtain the nowcast value of Kp, etc.

6. Prediction and forecast

The meaning of the terms "prediction" and "forecast" used in this paper should be clearly distinguished. Prediction is a feature of the model, its capacity to guess a future state. Forecast is a service. Usually the prediction model is a part of the forecasting software, which has additionally: data collection part, I/O modules, data adjustment (correction) part, etc. In some forecasting software, as those developed in the framework of EU COST 251 and 247 actions (see for example, Kutiev et al., 1999; Muhtarov et al., 2001), data adjustment is a part separated from the models. The data adjustment technique is important for forecasting, because it keeps model prediction close to the current data. In the above cited papers, data adjustment is performed by correcting the model prediction by the currently obtained difference between prediction and data. In the HDK model this adjustment is part of the algorithm. Using the analogy with the above cited works, we can regard MAK model as prediction model

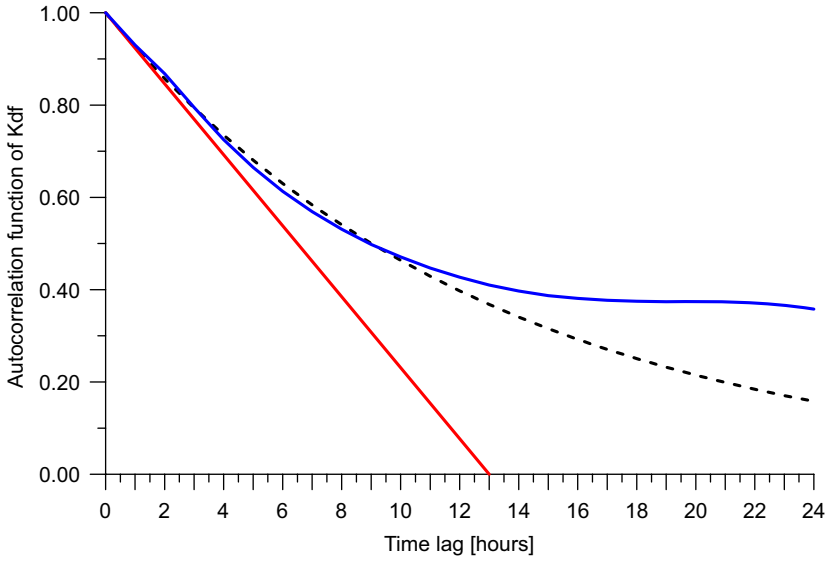

Fig. 8. Empirical autocorrelation function of Kdf (solid blue line) and its exponential approximation (dashed line). The red line marks the time constant of 13 h. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

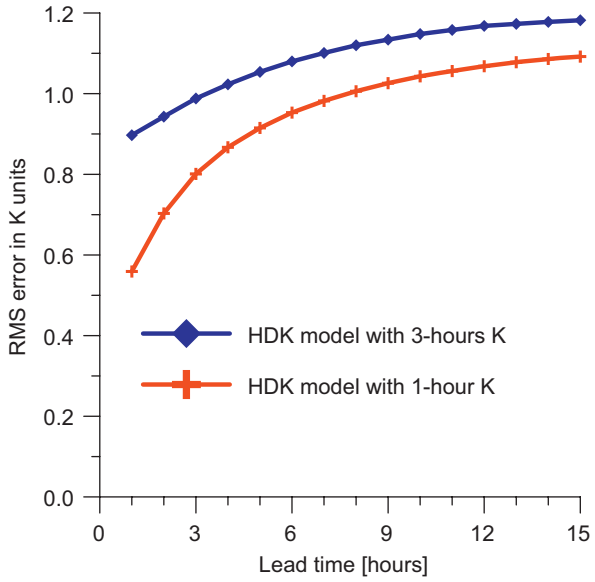

Fig. 9. Prediction error (standard deviation between predicted and nowcasted Kdf for 1-h (red line) and 3-h (blue line) sampling rate. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

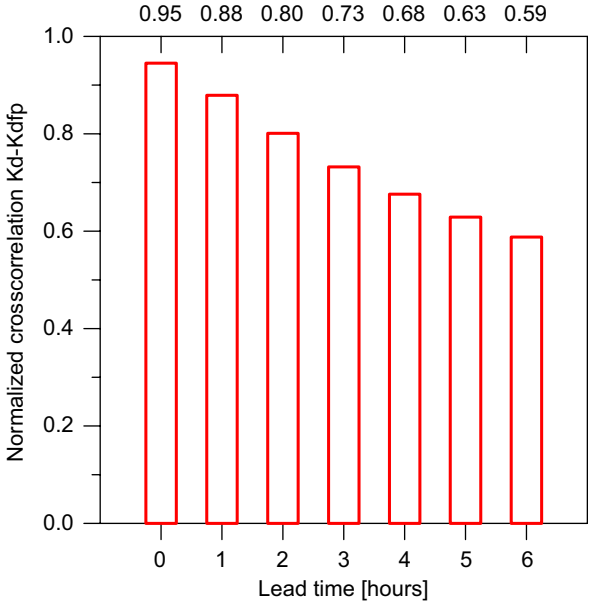

Fig. 10. Vertical bars represent correlation coefficient between Kd and predicted Kdfp for the lead time 1–6 h. The lead time = 0 shows the correlation coefficient between Kd and Kdf. The correlation coefficient values are shown at the upper axis.

Table 2
Correlation coefficient ***r*** estimated by different models.

Model	Type	*r*	Input parameters
Castello	NN	0.75	*V*, Bz, |**B**|
NARMAX	NN	0.77	*V*, Bz, |**B**|, Kp($t-3$)
Lund	NN	0.77	*V*, *n*, Bz,
APL	NN	0.92	*V*, Bz, |**B**|, nowcast Kp
HDK	Empirical	0.95	*V*, *P*, Bz, nowcast K

and the Kd part as data adjustment. Therefore, we consider the HDK model as both prediction and forecasting software.

How far ahead the forecasting can be made? Theoretically there is no restriction. Every user can decide how accurate forecast likes to make. Fig. 7 shows the expected error up to 6 h lead time. Theoretically, when the lead time approaches the time constant (in our case $T = 13$ h), the predicted value approaches the average (Kdfmean) and the prediction error becomes close to the standard deviation of data around the average. This means that for predictions, exceeding 10–13 h, it is better to use directly the average value of Kd (for our database it is 2.42), instead of making such a complex calculations.

Another question that should be addressed when we obtain nowcast values is to which time the nowcast refers. The nowcast value of K should be available at the time when it is going to be used. For the near-real-time operations, solar wind parameters from ACE are available shortly after being measured. The time of measurement precedes by 30–60 min the moment when the solar wind strikes the Earth's magnetosphere and eventually produce

geomagnetic disturbances. From the other side, having 1-h sampling rate, Ksw may represent solar wind conditions 1 h in the past, at worst. Consider this time schedule, it is reasonable to assume that the advance time of measurement compensates the delay due to the hourly sampling, and Ksw and consecutively the nowcast Kdf represent the current magnetospheric conditions.

7. Conclusion

A new empirical model for nowcasting and predicting the geomagnetic K index is developed, based on the combined use of solar wind parameters and ground-based magnetic data. The present approach implements the previously developed solar wind-based MAK model, calibrating its values with magnetogram-derived K index. The new model, applied to the K index issued at the Dourbes Center of Geophysics of Belgian Royal Meteorological Institute, is named Hybrid Dourbes K model. The HDK model combines the advantages of predicting the sudden changes of geomagnetic activity induced by solar wind with the longer term predictability of the K index.

MAK model coefficients were re-calculated by fitting the model expression (2) to IMF Bz, solar wind velocity and dynamic pressure and Dourbes K index data. The database used for modeling consists of the hourly values of solar wind parameters and the hourly interpolated 3-h Dourbes K index (Kd), collected in the period 1998–2004. The HDK nowcast model, the quantity Kdf, is obtained by MAK model Ksw, corrected with a combination of differences between several past values of Kd and Ksw. The model error of the nowcast Kdf is found to be 0.38 KU, or nearly twice less than that of the MAK model. Kdf has a good predictability. Prediction made by weighted extrapolation 6 h ahead carries an error of 1.0 KU, while for the first 1 h the error is 0.58 KU only.

Acknowledgements

This work has been performed in the frame of contract GJU/06/2423/CTR/GALOCAD with Galileo Joint Undertaking and in the frame of a bilateral collaboration agreement between Belgium and Bulgaria financed by Belgian Science Policy

References

Andonov, B., Muhtarov, P., Kutiev, I., 2004. Analogue model, relating Kp index to solar wind parameters. J. Atmos. Solar Terr. Phys. 66 (11), 927–932.

Balikhin, M.A., Boaghe, O.M., Billings, S.A., Alleyne, H., 2001. Terrestrial magnetosphere as a non-linear resonator. Geophys. Res. Lett. 28, 1123–1126.

Boaghe, O.M., Balikhin, M.A., Billings, S.A., Alleyne, H., 2001. Identification of nonlinear processes in the magnetospheric dynamics and forecasting of Dst index. J. Geophys. Res. 106, 30,047–30,066.

Boberg, F., Wintoft, P., Lundstedt, H., 2000. Real Kp predictions from solar wind data using neural networks. Phys. Chem. Earth 25 (4), 275–280 〈http://www.lund.irf.se/rwc/kp〉.

Costello, K.A., 1997. Moving the rice MSFM into a real-time forecast mode using solar wind driven forecast models. Ph.D. Dissertation, Rice University, Houston, TX, June 1997.

Korn, G.A., Korn, T.M., 1968. Mathematical Handbook, second ed. McGraw-Hill, New York.

Kutiev, I., Muhtarov, P., Cander, Lj., Levy, M., 1999. Short-term prediction of ionospheric parameters based on autocorrelation analysis. Ann. Geofis. 42 (1), 121–124.

Mayaud, P.N. (Ed.), 1980. Derivation, Meaning, and Use of Geomagnetic Indices, Geophys. Monogr. Ser, vol. 22. AGU, Washington, DC.

Muhtarov, P., Kutiev, I., Cander, Lj., Zolesi, B., de Franceschi, G., Levy, M., Dick, M., 2001. European ionospheric forecast and mapping. Phys. Chem. Earth 25 (5), 347–351.

Muhtarov, P., Kutiev, I., 1999. Autocorrelation method for temporal interpolation and short-term prediction of ionospheric data. Radio Sci. 34 (2), 459–463.

Muhtarov, P., Andonov, B., 2000. Improved relationship between the IMF component B_z and Kp index. Bulg. Geophys. J. 26 (1–4), 165–172.

Takahashi, K., Toth, B.A., Olson, J.V., 2001. An automated procedure for near-real-time Kp estimates. J. Geophys. Res. 106, 21,017–21,032.

Thomsen, M., 2004. Why Kp is a good measure of magnetospheric convection. Space Weather 2, S11004.

Wing, S., Johnson, J.R., Jen, J., Meng, C.-I., et al., 2005. Kp forecast models. J. Geophys. Res. 110, A04203.