Geophysica (2008), 44(1–2), 3–13

Prediction of the Geomagnetic K Index Based on its Previous Value

Ari Viljanen[1], Antti Pulkkinen[2] and Risto Pirjola[1]

[1] Finnish Meteorological Institute, P.O.B. 503, FI-00101 Helsinki, Finland
[2] University of Maryland and NASA/Goddard Space Flight Center,Code 674, Greenbelt, MD 20771, USA

(Received: October 2007; Accepted: November 2007)

Abstract

The statistical behaviour of the geomagnetic K index is studied at two observatories in Finland: Nurmijärvi in the subauroral region and Sodankylä near the auroral region. Using data of 1953–2006, we show that the present K index can be used as a proxy for predicting the future indices. Especially, the statistical prediction of the nearest future index yields a clearly different value compared to the overall average of the given UT time. Long-term statistical predictions up to three months forward also reveal the 27 days periodicity related to the solar rotation. We tested the prediction method by deriving statistics for the years 1953–2002 and applied the results for 2003. About 72% of the predicted expectation values of the nearest future K are within one unit from the observed K and about 96% differ at most two units. Analysis of the 20 cases with the prediction error larger than 3 units shows that a sudden start of a large storm is impossible to forecast by using ground magnetic field alone, as expected.

Key words: geomagnetic forecasting, geomagnetic variations, K index

1. Introduction

The *K* index is a 3-hour quasi-logarithmic local indicator of the geomagnetic activity relative to an assumed quiet-day curve for the recording site. It measures the deviation of the more disturbed horizontal component on the scale from 0 to 9 (*Rangarayan*, 1989). Although the concept of the quiet-day variation is somewhat vague, the *K* index is a useful proxy for characterising long-term variations of the magnetic activity back to the earliest magnetic recordings in 1840s (*Nevanlinna et al.,* 1993). Moreover, digital data of field variations with a high time resolution are available only for the latest decades.

The goal of this study is to develop a simple procedure to predict the geomagnetic activity based purely on local magnetic recordings. Thanks to its long temporal coverage, the *K* index provides a handy way for this purpose. We start by describing the overall occurrence statistics of *K* and then assess its prediction based only on previous values.

One practical use of the *K* index is related to geomagnetically induced currents (GIC) in power systems and pipelines. There has been debate against and for the usefulness of traditional magnetic indices for GIC classification, especially concerning

K (*Oler*, 2004; *Simpson*, 2004; *Kappenman*, 2005). More quantitative activity indicators, directly related to field amplitudes or their time derivatives, are obviously better for such a specific purpose as suggested by *Trichtchenko and Boteler*, (2004), *Viljanen et al.* (2006) or *Menvielle and Marchaudon* (2007). Experiences of the simple prediction of the *K* index presented in this paper will be used for developing more accurate predictions of other more quantitative activity indicators.

2. *Results*

2.1 *Description of data*

We use data from two observatories in Finland: Nurmijärvi (geomagnetic latitude 57 N) and Sodankylä (geomagnetic latitude 64 N). Indices are available since 1953 from Nurmijärvi and since 1914 from Sodankylä (with a gap in September 1944 to December 1946). For a direct comparison, we use mostly years 1953–2006 for both observatories. The *K* index was previously determined by hand-scaling, but since about 1990 automatic computer methods have been preferred (e.g. *Menvielle et al.* 1995). This also reduces the ambiguity of the definition of the quiet-day field. The computer algorithm used in Finland is described by *Sucksdorff et al.* (1991).

2.2 *Statistics of the K index*

The overall occurrence of *K* is shown in Fig. 1. The mean value (weighted average) of all *K* indices at Nurmijärvi is 2.06 and at Sodankylä 2.36. Note that although the original *K* values are integers, the expectation values are decimal numbers. The diurnal variation is given in Table 1 and illustrated in Fig. 2. Results of Nurmijärvi agree with the earlier study by *Nevanlinna et al.* (1992). For Sodankylä, we also show

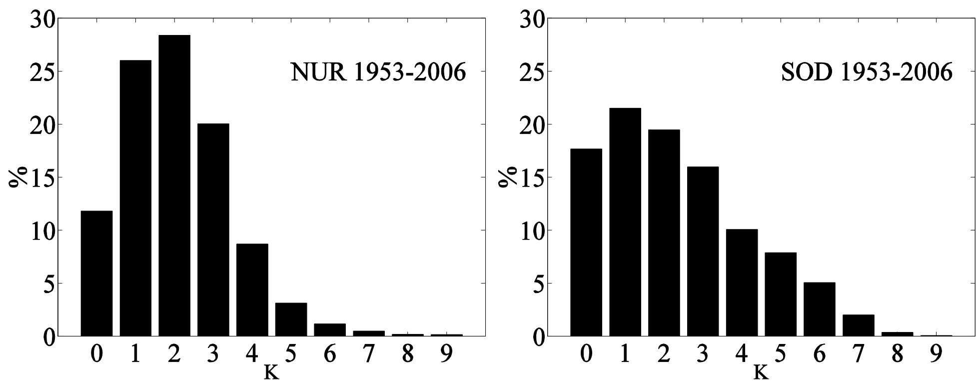

Fig. 1. Occurrence of the *K* index at the Nurmijärvi (NUR) and Sodankylä (SOD) observatories in 1953–2006.

the diurnal curve using data in 1914–1952. The difference to the data of 1953–2006 is systematic: on average, the *K* index during the former period is 0.36 units smaller. This manifests the higher geomagnetic activity during the latter period, as characterised by several other indicators such as the sunspot number, the number of auroral nights in

Finland or the planetary *aa* index (*Pulkkinen et al.*, 2001). The shape of the diurnal behaviour at Sodankylä is nearly identical in both periods. This is an evidence of a consistent determination of the *K* index throughout decades.

Both at Nurmijärvi and Sodankylä, the magnetic activity peaks around the local midnight and is lowest in the morning (the local magnetic time is about UT+2.5h). The peak-to-peak diurnal variation is somewhat larger in Sodankylä, which is due to its location closer to the auroral region. The largest magnetic storms (*K* = *8,9*) are rare during the local late morning and just after midday (UT 06–12). Adopting the terminology of *DelSole* (2004), we will call the mean diurnal variation as the "climatological average" hereafter.

2.3 Statistical prediction of the K index

The simplest prediction of *K* is obtained by using its climatological average as shown in Fig. 2, but this is not very illuminative. The conditional probability of the next *K* values given the present *K* is considerably better. As indicated by Table 1 and Fig. 2, it is also necessary to take into account the diurnal variation, so we studied the UT periods 00-03, ..., 21-24 separately. We considered each *K* and counted the number of the next *K* values up to 10 days. A compact way to summarise the result is to calculate the expectation values of the future indices and compare them to the average diurnal variation in Fig. 2.

Table 1. The diurnal occurrence (%) of *K* indices as a function of the UT period (00–03, 03–06, ...) in 1953–2006 at Nurmijärvi (NUR) and Sodankylä (SOD). The sum of each column is 100%. See Fig. 1. for the relative overall occurrence of *K* values.

NUR	0	1	2	3	4	5	6	7	8	9
00-03	16.3	12.4	11.4	12.3	11.7	11.7	12.9	13.7	13.0	16.2
03-06	16.8	15.6	12.9	9.5	7.0	5.6	5.6	8.0	5.8	4.2
06-09	12.3	16.8	14.7	9.5	5.2	2.8	4.0	3.5	2.4	1.7
09-12	8.4	13.2	15.1	13.6	8.8	5.9	6.2	5.9	2.0	3.3
12-15	10.6	11.2	12.4	14.1	14.9	14.3	12.7	12.9	14.0	12.0
15-18	11.8	10.5	11.2	13.4	17.3	20.3	18.8	16.8	23.5	17.8
18-21	11.7	9.9	10.8	13.9	18.6	22.6	21.1	18.4	18.8	20.3
21-24	12.1	10.4	11.4	13.9	16.5	16.7	18.7	20.7	20.5	24.5
SOD	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
00-03	13.3	10.4	9.4	11.0	14.6	17.1	19.9	20.8	23.8	27.8
03-06	16.5	14.8	12.8	11.7	10.0	7.4	5.4	6.5	12.1	18.6
06-09	17.8	17.8	14.9	10.2	6.3	2.9	1.5	1.4	1.6	4.1
09-12	12.6	15.5	16.0	14.2	9.6	5.2	2.7	1.2	1.1	2.1
12-15	10.3	12.2	14.8	15.5	12.8	11.6	7.8	4.0	2.7	6.2
15-18	10.5	12.0	12.8	13.9	15.0	13.7	10.7	10.1	10.7	7.2
18-21	9.2	9.1	10.2	12.6	15.8	19.3	22.7	26.7	26.9	14.4
21-24	9.6	8.2	9.1	10.9	15.9	22.7	29.1	29.3	21.2	19.6

Examples in Figs. 3–4 show that the larger the present value the larger are the nearest expected future values compared to the climatological ones. Vice versa, if the present value is small it is likely to remain such very systematically during the nearest days. Concerning especially the next *K*, the expected distribution can be clearly different from climatology as shown in Fig. 5, where all UT times are taken into account. When different 3-hour UT periods are considered separately, this result still varies slightly from time to time (not shown here).

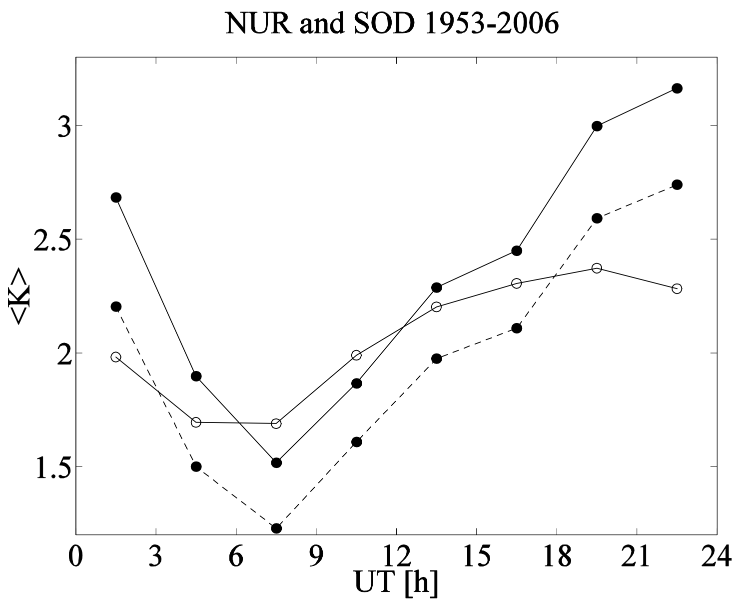

Fig. 2. Expectation value of the *K* index at the Nurmijärvi (circles) and Sodankylä (dots) observatories as a function of UT in 1953–2006. The dashed curve is for Sodankylä in 1914–1952. The local magnetic time is about UT+2.5h.

A closer inspection of the predicted *K* compared to the climatological average shows further features. Difference curves up to 90 days in Fig. 6 reveal the 27-day periodicity caused by the solar rotation. Although only the UT period 18-21 is shown in Fig. 6, a similar behaviour is found for all times. An unexpected feature, visible also in Figs. 3–4, is that the predicted *K* tends to stay below the climatogical average if the present *K = 1*. For *K* larger than 3, this is reversed. A possible explanation is as follows: *K* values 1–2 are the most probable ones and thus dominate climatology. If the present *K* happens to be (rather unlikely) large then the solar and magnetospheric activity is at a high level, which settles down slowly in days or weeks. Typically, the largest *K* indices are related to magnetospheric storms whose recovery phase is several days.

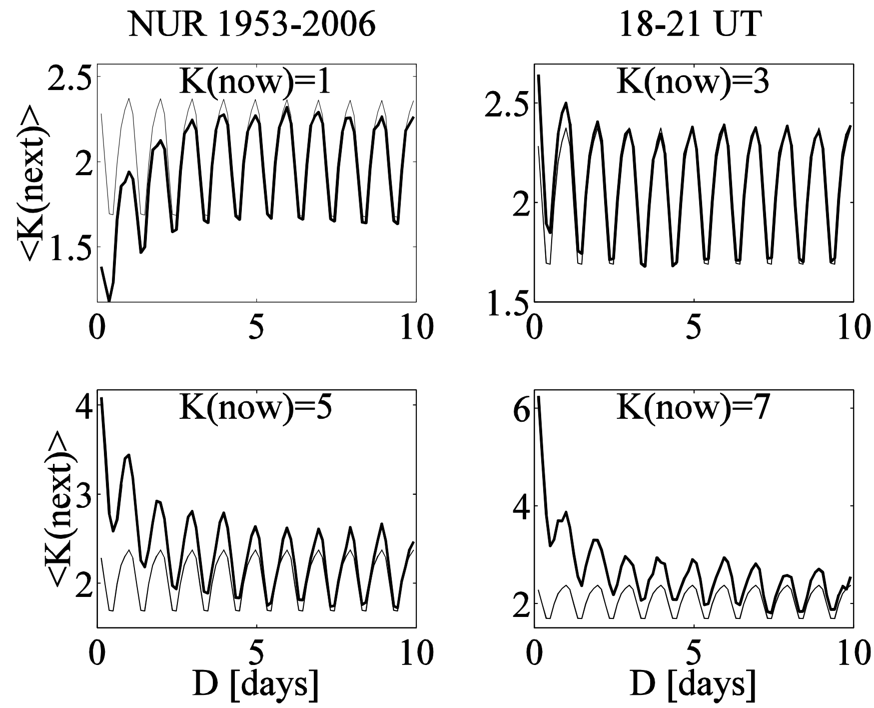

Fig. 3. Expectation value of the next *K* indices at Nurmijärvi in 1953-2006 given the present value ($K(now)$ = 1,3,5,7) for 18–21 UT. The time *D* in days after the present is given on the horizontal axis. The thin curve is the climatological average.

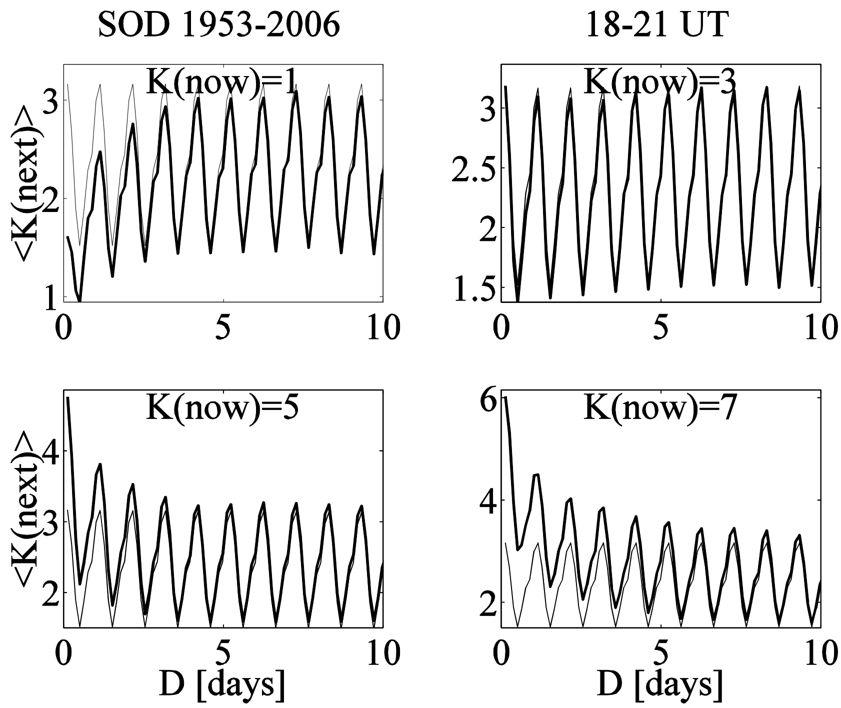

Fig. 4. As Fig. 3, but for Sodankylä.

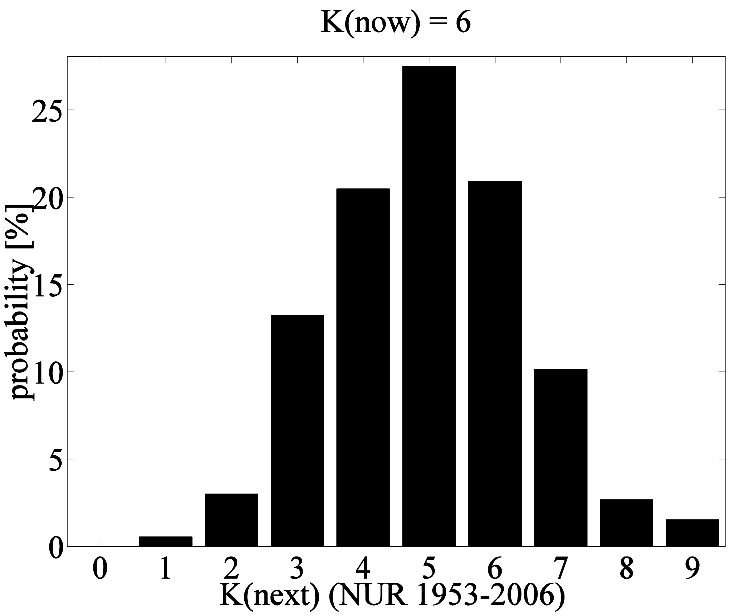

Fig. 5. The conditional probability of the next *K* at Nurmijärvi in 1953-2006 given that the present value is 6. All UT times are considered. Compare to the climatological distribution in Fig. 1.

As a different approach, we determined the power spectrum of *K*. The data were divided into segments of 250 days and the power spectrum was computed individually for each window. The final spectrum shown in Fig. 7 is the mean value taken over all windows. The three dashed lines indicate the most apparent characteristic time scales present in *K*: 27 days, 160 hours and 24 hours. The 27-day and 24-hour periods seen also in the analysis above have strong harmonics that dominate the peaks in the spectrum. Ignoring the peaks due to them, the spectrum has a break point at about the period of 160 hours (about 7 days). The slope changes there gradually from a steeper power-law scaling behaviour at higher frequencies to a different scaling at lower frequencies. This indicates a change in the correlations of the signal which can be interpreted consistently with the results above: at periods of about 160 hours, the local temporal persistence becomes weaker and the signal converges to the average diurnal behaviour as in Figs. 3–4. This is also equivalent to the example shown in Fig. 6, where the difference curve gradually approaches zero with increasing time after the present.

2.4 Test of the prediction method

As a test for the year 2003 at Nurmijärvi, we derived occurrence statistics using data of years 1953–2002 and predicted the expectation value of the next *K* given the present value in 2003. The difference distribution is shown in Fig. 8. About 72% of predicted values are within one unit from the observed *K* and about 96% differ at most two units. The predicted values tend to be slightly smaller than the observed one, as the skewness of the distribution indicates. The skewness is probably due to the high magnetic activity of the year 2003.

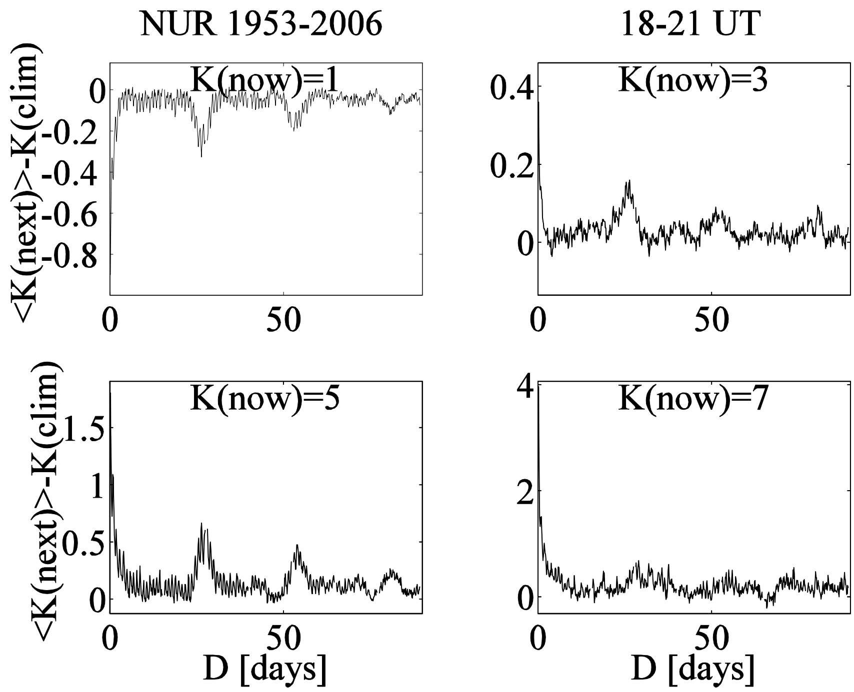

Fig. 6. Difference between the expectation value of the next *K* indices and the climatological average at Nurmijärvi in 1953–2006. The time *D* in days after the present is given on the horizontal axis. The expectation value is determined given the present value *(K(now)* = 1,3,5,7) for 18–21 UT.

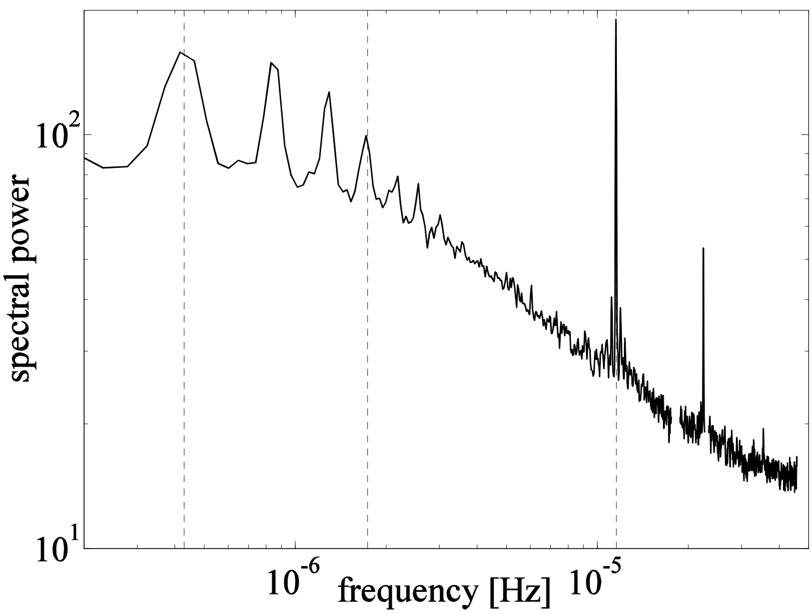

Fig. 7. Power spectrum of the *K* index at Nurmijärvi. The three vertical dashed lines indicate the time scales of 27 days, 160 hours and 24 hours.

A closer inspection to failed predictions is given in Table 2. It lists cases where the predicted expectation value of *K*(*next*) differs more than 3 units from the observed one. In all cases, the observed *K*(*next*) is larger than the predicted one. Nine cases are related to a beginning of a magnetic storm ($K \geq 7$). The rest 11 cases are related to a moderate activation (*K* = 4–6), which in most cases did not evolve into a storm. The largest error occurred in the beginning of the Halloween storm on 29 October 2003

(*Rosenqvist et al.* 2005): *K* index at 03–06 was 4 giving an expectation value of 4.10 for 06–09 UT. However, the solar wind shock caused variations larger than 1000 nT just after 06 UT leading to K = 9. It is clear that such a huge impulse directly due to the solar wind cannot be predicted only by using the previous magnetic field recordings on the ground.

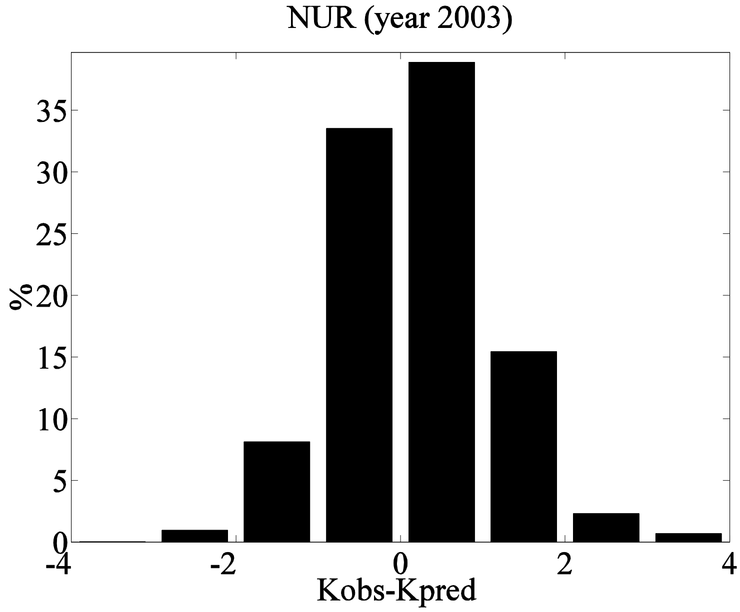

Fig. 8. Difference distribution between the observed and predicted expectation value of the next *K* index at Nurmijärvi in 2003. All UT times are included. The prediction is based on statistics of the *K* index in 1953–2002.

We also compared the pure climatological prediction to the conditional one. The previous method means that the expectation value of *K* for each UT period is considered as the "prediction". The average absolute error of the prediction is then 1.29 units, whereas for the conditional method the error is 0.75. The climatological guess also leads to a much larger underestimation of the next *K* value than the conditional statistics.

2.5 Possible further developments

A possible improvement could be to consider more past values than just the present one. We studied this by taking into account both the present value ($K(now)$) and the previous one ($K(now - 1)$). Then we calculated the expectation value of the next *K* index. As Fig. 9 shows, the previous value does not practically yield much more information. If the present *K* is smaller than 6 then there is a slight increase of the expected next *K* with an increasing value of the previous *K*. For the largest value (6–9) this does not hold, but the expectation value of the next index seems quite independent of $K(now - 1)$.

Table 2. Largest differences in predicted and observed *K* indices at Nurmijärvi in 2003. Each row contains the date and the eight observed daily *K* values. Two successive values related to an inaccurate prediction are marked by boldface: *K*(*now*) and *K*(*next*). The last column gives the predicted expectation value of *K*(*next*) based on statistics of 1953–2002.

20030103	3	1	2	1	**3**	**6**	5	4	2.93
20030202	5	5	4	3	**4**	**8**	6	6	3.70
20030218	**2**	**5**	3	3	3	2	3	4	1.70
20030425	3	3	3	4	3	3	**2**	**6**	2.00
20030509	4	4	4	5	5	4	2	**4**	3.10
20030510	**7**	6	5	3	2	4	5	3	
20030529	4	3	2	**4**	**8**	8	9	9	4.10
20030719	4	3	3	4	5	4	**3**	**6**	2.63
20030817	1	0	1	**2**	**6**	4	3	6	2.15
20030817	1	0	1	2	6	4	**3**	**6**	2.63
20031014	3	3	4	4	4	**4**	**9**	8	3.51
20031024	0	2	2	2	**4**	**8**	6	4	3.70
20031029	4	**4**	**9**	7	9	9	9	9	3.12
20031030	9	8	6	5	6	**7**	**9**	9	5.96
20031104	3	**3**	**6**	6	3	2	3	2	2.52
20031120	2	**2**	**5**	6	9	9	9	8	1.87
20031120	2	2	5	**6**	**9**	9	9	8	5.79
20031122	4	2	2	2	**3**	**8**	7	6	2.93
20031220	0	1	1	3	**3**	**6**	3	3	2.93
20031227	1	2	1	3	1	1	**2**	**5**	2.00
20031231	3	2	1	2	**3**	**6**	5	4	2.93

An easily understandable forecast could be issued as follows: determine the latest *K* index and predict the next one based on statistics derived for each 3-hour UT period. Illustrate the result by stating the probability of the next value in three categories: smaller, equal, or larger than the latest one. For example, assume that at 15–18 UT, the *K* index was 6 at Nurmijärvi. The conditional occurrence probabilities of the *K* values from 0 to 9 for 18–21 UT are then 0.0, 0.6, 2.9, 11.6, 22.6, 31.0, 18.3, 10.1, 1.7 and 1.2%. The forecast issued at 18 UT would give the following probabilities of the three categories: 69%, 18%, 13%, respectively. The climatological probability of $K \geq 6$ at 18–21 UT is only 3%.

<K(next)> at NUR (1953-2006)

K(now-1) \ K(now)	0	1	2	3	4	5	6	7	8	9
9	NaN	NaN	2.0	3.8	2.7	**4.9**	**5.9**	**6.0**	**6.6**	**7.5**
8	NaN	NaN	2.0	3.5	**4.1**	**4.4**	**5.1**	**6.0**	**7.1**	**7.6**
7	NaN	NaN	**2.3**	**3.0**	**3.4**	**4.5**	**5.0**	**5.7**	**6.1**	**7.5**
6	NaN	**1.7**	**2.3**	**3.0**	**3.6**	**4.2**	**4.8**	**5.7**	**7.0**	**8.2**
5	0.8	**1.7**	**2.4**	**2.9**	**3.5**	**4.0**	**4.9**	**6.1**	**6.8**	**6.9**
4	**0.9**	**1.5**	**2.2**	**2.8**	**3.4**	**4.2**	**5.0**	**6.3**	**6.8**	8.5
3	**0.8**	**1.4**	**2.1**	**2.8**	**3.5**	**4.2**	**5.1**	**6.4**	7.3	7.1
2	**0.7**	**1.4**	**2.0**	**2.7**	**3.2**	**4.1**	**5.1**	6.0	NaN	7.0
1	**0.7**	**1.3**	**1.9**	**2.4**	**3.2**	**4.1**	**5.1**	6.0	NaN	9.0
0	**0.6**	**1.1**	**1.7**	**2.2**	**3.0**	4.2	NaN	NaN	NaN	NaN

Fig. 9. Expectation value of the next *K* index at Nurmijärvi given the present ($K(now)$) and the previous values ($K(now - 1)$) in 1953–2006. The light font denotes pairs ($K(now)$, $K(now - 1)$) which have occurred less than 10 times. Missing pairs are marked by NaN. All UT times are included.

3. *Conclusions*

Prediction of the magnetic *K* index can be based on the previous value. This method yields predictions which clearly differ from the simple climatological expectation value especially when the magnetic activity is high. This method also reproduces the previously known result that it takes several days for the magnetic field to recover from a disturbed state. Vice versa, a very quiet field does not typically turn into a storm rapidly.

The largest failures of the statistical prediction of the nearest *K* index are related to a rapid activation of the magnetic field. Such events cannot be predicted only by using the previous magnetic field recordings, but usage of solar wind observations would be necessary. At times, solar wind data are not available, in which case a prediction based on ground magnetic field data may be the best alternative.

This study has dealt only with two observatories in Finland, one being close to the auroral region and another located in the subauroral region. A straightforward extension could be to consider a few stations at higher and lower latitudes.

Acknowledgements

We appreciate very much the long-term careful work performed at the Nurmijärvi and Sodankylä geomagnetic observatories.

References

DelSole, T., 2004. Predictability and Information Theory. Part I: Measures of Predictability, *J. Atmosph. Sci*., **61**, 2425–2440.

Kappenman, J.G., 2005. An overview of the impulsive geomagnetic field disturbances and power grid impacts associated with the violent Sun-Earth connection events of 29-31 October 2003 and a comparative evaluation with other contemporary storms, *Space Weather*, **3**, S08C01, doi:10.1029/2004SW000128.

Menvielle, M., N. Papitashvili, L. Häkkinen and C. Sucksdorff, 1995. Computer production of K indices: review and comparison of methods, *Geophys. J. Int*., **123** (3), 866–886. doi:10.1111/j.1365-246X.1995.tb06895.x

Menvielle, M. and A. Marchaudon, 2007. Geomagnetic indices in solar-terrestrial physics and space weather, in: Space Weather, Research Towards Applications in Europe (ed. J. Lilensten), Springer, Astrophysics and Space Science Library, 344, 277–288.

Nevanlinna, H., A. Ketola, A. Viljanen, L. Häkkinen and K. Ivory, 1992. Magnetic Results from Helsinki Magnetic-Meteorological Observatory. Part II. Geomagnetic Activity 1844-1856, *Finnish Meteorological Institute Geophysical Publications*, No. **30**, 65 pp.

Nevanlinna, H., A. Ketola, A. Viljanen, L. Häkkinen and K. Ivory, 1993. Geomagnetic activity during solar cycle 9 (1844-1856), *Geophys. Res. Lett*., **20**, 743–746.

Oler, C., 2004. Imperfect K Index Satisfactory For Electric Power Industry, *Space Weather*, **2**, S06001, doi:10.1029/2004SW000083.

Pulkkinen, T.I., H. Nevanlinna, P.J. Pulkkinen and M. Lockwood, 2001. The Sun-Earth connection in time scales from years to decades and centuries, *Space Sci. Rev*., **95**, 625–637.

Rangarayan, G.K., 1989. Indices of Geomagnetic Activity, in: Geomagnetism, Volume 3 (ed. J.A. Jacobs), Academic Press, 323–384.

Rosenqvist, L., H. Opgenoorth, S. Buchert, I. McCrea, O. Amm and C. Lathuillere, 2005. Extreme solar-terrestrial events of October 2003: High-latitude and Cluster observations of the large geomagnetic disturbances on 30 October, *J. Geophys. Res*., **110**, doi: 10.1029/2004JA010927.

Simpson, S., 2004. Does the Electrical Power Industry Need Better Storm Warnings? *Space Weather*, **2**, S03001, doi:10.1029/2004SW000063.

Sucksdorff C., R. Pirjola and L. Häkkinen, 1991. Computer production of K-values based on linear elimination, *Geophys. Transact*., **36**, 333–345.

Trichtchenko, L. and D.H. Boteler, 2004. Modeling Geomagnetically Induced Currents Using Geomagnetic Indices and Data, *IEEE T. Plasma Sci*., **32**, 1459–1467, doi:10.1109/TPS.2004.830993.

Viljanen, A., A. Pulkkinen, R. Pirjola, K. Pajunpää, P. Posio and A. Koistinen, 2006. Recordings of geomagnetically induced currents and a nowcasting service of the Finnish natural gas pipeline system, *Space Weather*, **4**, S10004, doi:10.1029/2006SW000234.