Kalman filter technique for defining solar regular geomagnetic variations: Comparison of analog and digital methods at Sodankylä Observatory

D. Martini,[1,2] M. Orispää,[3] T. Ulich,[3] M. Lehtinen,[3] K. Mursula,[2] and D.-H. Lee[1]

Received 2 December 2010; revised 11 February 2011; accepted 3 April 2011; published 24 June 2011.

[1] Motivated by recent attempts to derive geomagnetic activity from hourly mean data in long-term studies, we test the recursive Kalman filter method to obtain the regular solar variation curve of the geomagnetic field. Using a simple algorithm, we are able to assign a quiet day curve to every day separately, without the need for additional input parameter(s) to define the geomagnetically quiet days. We derive a digital counterpart A_{hK} of the analog range index *Ak* at the subauroral Sodankylä station and compare it to the earlier digital estimate *Ah* and the local *Ak* index. We find that the new method outperforms the former estimate in every aspect studied and provides a robust, straightforward manner of estimating and verifying the manually scaled *Ak* index, based on readily available hourly values. The model is independent of sampling; thus, for shorter-term studies where high-sampling data are available, more accurate estimates can also be obtained when needed. Therefore, in contrast to other recent approaches, we do not provide a method to quantify irregular activity directly but derive the actual quiet day curves in the traditional manner. In future applications the same algorithm may be used to define a wide variety of geomagnetic indices (such as *Ak*, *Dst*, or *AE*).

Citation: Martini, D., M. Orispää, T. Ulich, M. Lehtinen, K. Mursula, and D.-H. Lee (2011), Kalman filter technique for defining solar regular geomagnetic variations: Comparison of analog and digital methods at Sodankylä Observatory, *J. Geophys. Res.*, *116*, A06102, doi:10.1029/2010JA016343.

1. Introduction

[2] One can separate two distinctly different types of variations in the geomagnetic field; the regular (also often called Sq (solar quiet)) and irregular variations. While the former is mainly driven by the solar UV/EUV radiation and manifests itself as a smooth daily change in the magnetograms due to the Earth's rotation (hence the name regular), the latter is a result of the dynamic fluctuations of solar wind and HMF (heliospheric magnetic field). These fluctuations may lead to global magnetic storms and more local substorms. Until recently a number of indices have been implemented to characterize and quantify this geomagnetic activity. The key task of this procedure is to quantify and separate the fundamentally unknown regular quiet day curve (QDC) from the irregular activity carrying important information about the near-Earth space, as well as about the dynamics of ionospheric/magnetospheric current systems [*Bartels et al.*, 1939; *Menvielle and Berthelier*, 1991; *Nevanlinna*, 2004].

[3] The QDC has traditionally been defined by hand scaling of magnetograms, hence producing the so-called analog indices of geomagnetic activity. More recently, a new digital measure of geomagnetic activity, the *Ah* index, has been implemented, using hourly mean values of the magnetic field [*Mursula and Martini*, 2007a] (its derivation is shortly discussed in more detail). The explicit aim was to produce an *Ak*-type digital index that follows as closely as possible and appropriate the derivation method of *K*-type indices [*Bartels et al.*, 1939; *Mayaud*, 1980] but which, by using the available digitized hourly values and a definite technique, is more straightforward and verifiable than the hand-scaled indices, thus better suited for long-term studies. To define the regular variations they used monthly averaged QDCs, defined from the 5 quietest days of each month. The method is often referred to as iron curve method, due to its rigidity in taking into account day-to-day QDC variations that occasionally can be significant. This averaged QDC was then used for the given month to calculate the 3 h ranges in each day (very much similarly to the *K* method, fitting the QDC as upper and lower envelope to the actual data separately in each 3 h sectors), to be assigned as the *Ah* value. *Mursula and Martini* [2007a] used the local IHV index [*Le Sager and Svalgaard*, 2004; *Svalgaard et al.*, 2004] to define the local quiet days, therefore the *Ah* index was dependent on an additional input measure of geomagnetic activity. Using a new approach the Kalman filter algorithm, on the other hand, we are able to assign QDC for each and every day separately (to be called daily QDC), and

[1]School of Space Research, Kyung Hee University, Yongin, South Korea.
[2]Department of Physics, University of Oulu, Oulu, Finland.
[3]Sodankylä Geophysical Observatory, Sodankylä, Finland.

0148-0227/11/2010JA016343

the method is independent from any other geomagnetic index. The amplitude of the irregular activity is thereafter defined in each three-hourly section of a given day separately, as the difference of the upper and lower envelope fitted daily QDC to this 3 h; following closely the principle of the traditional *Ak* method (and that of *Ah*). In order to separate the two differently derived *Ah* indices, we call the new 3 h digital range index A_{hK}, where *h* refers to the use of hourly mean values of magnetic *H* component, while *K* refers to the Kalman method of QDC quantification.

[4] *Mursula and Martini* [2007b] made a thorough analysis of the characteristics of the *Ah* index, using data from the subauroral Sodankylä Geophysical Observatory, Finland (SOD, 67°22′ GGlat, 26°38′ GGlong, 63.9° GMlat). This station has a high-quality series of analog *Ak* indices, and its data have often been used for long-term comparisons of the geomagnetic activity [e.g., *Clilverd et al.*, 2002, 2005]. Therefore, the SOD station is an obvious choice for comparing the new Kalman-filtered index A_{hK} with the earlier digital *Ah* index and using the analog *Ak* index as a reference. The primary aim of this paper is to present a reliable proxy to the analog *Ak* index, based on hourly averaged magnetic measurements. However, the Kalman algorithm introduced is not limited to hourly sampling. By using higher sampling raters of raw data, one could aim to make even more accurate proxies of *Ak*, or use it as a derivation method of QDC for other indices, such as *Dst*. Besides, since the algorithm is mathematical and not physical, it can also be used to resolve technically similar problems, such as the periodic changes in the satellite orbits in space-borne measurements. These approaches, however, are subjects of a forthcoming paper.

2. The Kalman Filter

[5] The Kalman filter [*Kalman*, 1960] is a powerful recursive method to estimate the state of a process by minimizing the mean of the squared error. The filter can provide estimates of past, present and future states, even if the exact nature of the modeled system is unknown (G. Welch and G. Bishop, An introduction to the Kalman Filter, TR 95-041, Univ. of N. C. at Chapel Hill, Chapel Hill, 2006, http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf). We have used a very simple Kalman filter (*Kaipio and Somersalo* [2005]; for a more thorough treatise, see work by *Grewal and Andrews* [1993]) to estimate the QDC. The hourly mean magnetometer values were divided into daily 24 h (1 day) bins after which we ran the Kalman filter using these 24 data points one at a time. We assumed that the estimated QDC does not vary much from one day to the next, so we set the evolution model matrix of the Kalman filter equal to identity. Also, we set the observation model matrix to be identity since we are filtering the plain measurement data. In addition, we took the evolution model error covariance matrix to be diagonal with a given variance. The observation model covariance matrix was also taken to be diagonal, but the variances were calculated from the data (see below). This leads to a very simple linear Kalman filter, given as

$$X_{k+1} = X_k + W, k = 0, 1, 2, \dots \tag{1}$$

$$Y_k = X_k + V_k, k = 1, 2, \dots, \tag{2}$$

where (1) is the evolution model, (2) is the observation model, $X_k \in \mathbf{R}^{24}$ is the estimated QDC for day number k, $W \in \mathbf{R}^{24}$ is the evolution model error vector, $Y_k \in \mathbf{R}^{24}$ is the magnetometer data for day number k and $V_k \in \mathbf{R}^{24}$ is an observation model error vector for day number k. Here we assume that the error vectors have Gaussian probability distributions

$$W \sim N(0, \sigma I) \tag{3}$$

$$V_k \sim N(0, \Sigma_k), \tag{4}$$

where σ is the given predetermined evolution model error variance "evovar" and Σ_k is the calculated observation error covariance matrix for day number k. For the initial estimate X_0 we set

$$X_0 \sim N(E(X_0), C(X_0)) = N(Y_0, I),$$

where X_0 is a Gaussian random vector with expectation value $E(X_0)$ equal to the first measurement Y_0 and identity covariance matrix, $C(X_0) = I$. After these assumptions and initial settings, the Kalman filter is run as follows (k = 1, 2, …).

[6] 1. We calculate the a priori value $\overline{X}_k$ for X_k using the evolution model (1) and the previous estimate X_{k-1}, to get

$$\overline{X}_k \sim N(E(X_{k-1}), C(X_{k-1}) + \sigma I) =: N\left(E\left(\overline{X}_k\right), C\left(\overline{X}_k\right)\right).$$

[7] 2. We calculate the pointwise differences between the measurement data and the estimate of the previous quiet day curve

$$\Delta_k = |Y_k - E(X_{k-1})|$$

and discard any data points for which Δ_k exceeds an empirically predetermined threshold value "rangethres." Note that the "rangethres" is dependent on the characteristic day-to-day QDC amplitude variation of a given station, and is station specific (and above all changes with latitude). Missing data points are always discarded. Also, the discarded data points are not used in calculating the sample variance for the data points, as explained below.

[8] 3. The observation model (2) together with the a priori value $\overline{X}_k$ and the measurement data Y_k is used to calculate an estimate for X_k:

$$X_k \sim N(E(X_k), C(X_k)),$$

where

$$E(X_k) = E\left(\overline{X}_k\right) + K_k\left(Y_k - E\left(\overline{X}_k\right)\right), \tag{5}$$

$$= E(X_{k-1}) + K_k(Y_k - E(X_{k-1})), \tag{6}$$

$$C(X_k) = (I - K_k)\, C\left(\overline{X}_k\right), \tag{7}$$

$$= (I - K_k)(C(X_{k-1}) + \sigma I). \tag{8}$$

And K_k is the so-called Kalman gain matrix given by the formula

$$K_k = C\left(\overline{X}_k\right)\left(C\left(\overline{X}_k\right) + \Sigma_k\right)^{-1} \tag{9}$$

$$= (C(X_{k-1}) + \sigma I)(C(X_{k-1}) + \sigma I + \Sigma_k)^{-1}. \tag{10}$$

The observation model error covariance matrix Σ_k is constructed by calculating the hourly sample variance from the measurement data using a predetermined number of previous days "varint." In the beginning of the filter run, when enough previous days are not available, a predetermined constant value "measvar" is used as the variance.

[9] 4. The expectation value $E(X_k)$ is taken to be the estimate for the quiet day curve for day number k when calculating in Step 1 the a priori estimate, X_k, for X_{k+1} and so forth until the measurement data ends. Also, the previously discarded data points are not used in the variance calculation.

[10] In addition, if the measurement data for two or more consecutive days are completely discarded due to data missing or exceeding the threshold, the Kalman filter is reset; that is, the initial estimate is set back to X_0.

[11] The actual values of the input parameters for the Kalman filter used for the study are summarized in Table 1. The role of these parameters can be understood in terms of station specific "fine tuning," leaving the fundamental properties of the particular QDC estimate unaffected. We note that the filter gives robust results for wide range of tuning parameter values.

3. Geomagnetic Quiet Daily Curve

[12] Figure 1 shows the seasonally averaged QDC estimates of both models (monthly averaged and daily QDC) for the period of 1914–2000. Perhaps the most apparent difference is that monthly QDCs in all seasons depict their minima at about 01:00 UT. At high latitudes the QDC formation is much more complex than at low or midlatitudes. We get strong signal from the usual Sq current system around local noon (10:00 UT for SOD), but monthly QDC also shows the joint effect of field-aligned currents and the westward electrojet peaking around 03:25 magnetic local time LT (local time), i.e., 01:25 UT [*Allen and Kroehl*, 1975; *Finch et al.*, 2008]. The effect of westward electrojet should be a negative depression in ground-based *H* data. This can indeed be seen in Figure 1, marked with downward pointing arrow. Note that the effect is even more dominant in the monthly estimates than that of the Sq current system. As a result of the more profound daily QDC calculation of Kalman filter the effect of westward electrojet, although not completely absent, is much more moderate on the average.

[13] The other dominant difference occurs after about 18:00–19:00 UT until 24:00 UT, during the falloff of the sudden enhancement somewhat after 15:00 UT. The enhancement is the imprint of the eastward electrojet peaking around 15:30 UT in SOD [*Allen and Kroehl*, 1975; *Finch et al.*, 2008]. The monthly QDC estimate does not return to the approximate average level but remains excited during the night. This is either caused by the lasting effect of eastward electrojet, or, more likely, by the current wedge activity peaking in the local premidnight to postmidnight hours. As result we find that the monthly QDC does not follow the theoretical expectations for QDC in SOD (and for any station for that matter), as not having its minimum at local noon, but shows a clear increasing trend from a minimum at about 03:25 LT, being strongly affected by the depression due the combined effect of field-aligned currents and the westward electrojet at this time, and the enhancement due the eastward electrojet around 17:30 LT. This results in a clear discontinuity of the consecutive QDC estimates from 24:00 UT to 01:00 UT hours. We note that this discontinuity is clearly apparent not only in the overall seasonal averages depicted in Figure 1, but remains an important issue when comparing QDCs from month to month (not shown here). Such a trend is not present in the daily QDC estimates.

[14] We note that field-aligned effects are expected to be present even in the quiet data (see, e.g., the manually scaled QDCs in the annual Sodankylä Yearbooks), since, even during globally quiet days, at this latitude some level of activity is likely to occur. Therefore, any attempt to estimate QDC at the high latitudes faces considerable difficulty. This makes the SOD data a very good choice in testing the viability and comparability of various QDC methods.

[15] Figure 2 depicts the overall effect of the above discussed feature for the studied period of 1914–2000, showing the annual (365 days) running means of the daily ranges (amplitudes) of daily QDCs defined by the Kalman filter, and of the average monthly QDCs. The amplitudes are of the actual size, thus it is easy to see that while the QDC minimum levels of the two methods are roughly the same overall, the maxima differ radically. Since the magnetic QDC is largely formed by the UV/EUV radiation, its amplitude is expected to closely follow the evolution of the SSN (sunspot number) in the long term. The daily QDCs consistently show the SSN dependence better, thus fulfilling the theoretical expectations. On the other hand, the previously used monthly QDC systematically overestimates QDC most usually at the early to late declining, phases, depicting an "activity-like" long-term pattern, reaching its maxima typically shortly after the SSN maxima. This overshooting is the result of the inefficient QDC definition during the activity preferred LT sectors, as discussed in relation with Figure 1. Although the daily QDC follows the same pattern during some cycles, this problem is far less strident there.

[16] It is evident that any activity index based on the monthly QDC definition is unreliable and effectively assigns activity to the QDC variation dominantly during the early descending phase of the solar activity.

4. *Ah* Indices Based on Monthly and Daily QDCs

[17] Based on Figures 1 and 2, one would expect that the various QDC definitions, including daily versus monthly QDCs, as well as digital versus manual scaling, have a dominant effects in quantifying geomagnetic activity especially in the maximum/early descending phases of solar cycle, since this is the period when the different methods show the largest effects.

4.1. Annual Averages

[18] Figure 3 shows the annual averages of the analog *Ak* indices, and the digital *Ah* and A_{hK} indices. Both *Ah* and A_{hK} were normalized to the *Ak* index. The best linear fits lay very close to each other, $Ak = 0.44*Ah\text{-}1.4$; and $Ak = 0.44*A_{hK}\text{-}0.73$. The average standard deviation between the *Ak* and the fitted *Ah* (A_{hK}) indices is 1.19 (1.10) nT. This is only about 6.9% (6.4%, respectively) of the *Ak* mean level (17.3 nT). The three indices follow each other so closely that it is hard to distinguish among them at most parts. Probably the only major deviation, which can be seen at this resolution, is that *Ak* tends to depict smaller values before about

Table 1. Input Parameters for the Kalman Algorithm Used

Rangethres	Evovar	Measvar	Varint
46	10	100	30

1940s during activity maxima. This is probably due to a conservative approach in the manual scaling of the original *K* values during the early part of last century, when the usual activity level was relatively low (D. Martini et al., Comparing indices of geomagnetic activity at a high-latitude station, submitted to *Journal of Geophysical Research*, 2010, hereinafter Martini2010). Apart of this, all the three indices show qualitatively the same centennial behavior: superposed on the solar cycle variation there is a systematic increase of the background level until about 1960 that is followed by a significant dropout, and a slower increase thereafter. At this time scale the two digital indices correlate roughly equally well with *Ak*, with correlation coefficients of r = 0.982 for *Ah* and r = 0.985 for A_{hK}.

[19] In Figure 4 the annual residuals of linear regression fits are shown for *Ak* and *Ah*, or *Ak* and A_{hK}. As discussed above, the differences are very small but systematic. The new Kalman method observably reduces the earlier reported (Martini2010) excess in *Ah* with respect to *Ak* around some solar cycle maxima. Based on Figure 2 this improvement is expected, since this is the solar activity phase when the difference between the monthly and daily QDCs was found to be most dominant. The regular relative deviation in the late declining phase of solar cycle usually remains unaffected, or only moderately affected by the new method. Occasionally nonsystematic changes are observed; the most prominent is an increase in the residuals at a later descending phase around 1963 and a decrease after 1994. As Martini2010 discuss in more detail, the systematic negative deviation between *Ak* and *Ah* is due to the digital nature of the *Ah* index. Since *Ah* uses hourly mean values, it is less sensitive to high-frequency phenomena most likely to occur during the declining phase driven by the high-speed solar wind streams. This is a limitation that is inherent in indices using hourly mean values and cannot be significantly improved by a more elaborate QDC definition.

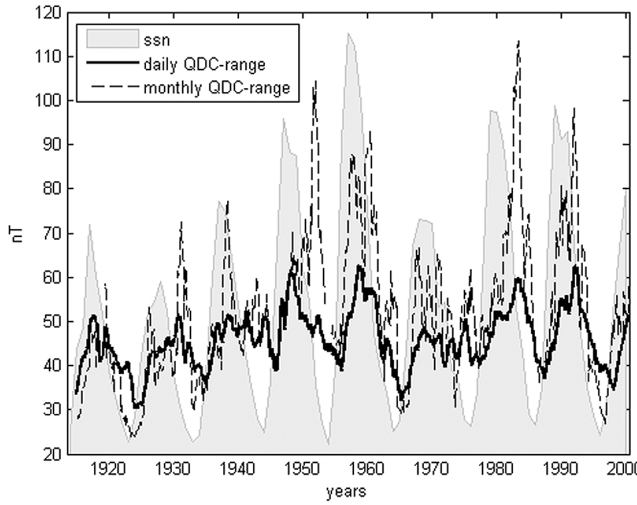

Figure 2. Annual running averages of the monthly and daily QDC ranges at SOD in 1914–2000. For comparison the qualitative annual sunspot numbers are also indicated with the shaded area.

4.2. Daily Averages

[20] Figure 5 shows the daily averaged indices in an arbitrarily selected period of late 1993. As it is expected the

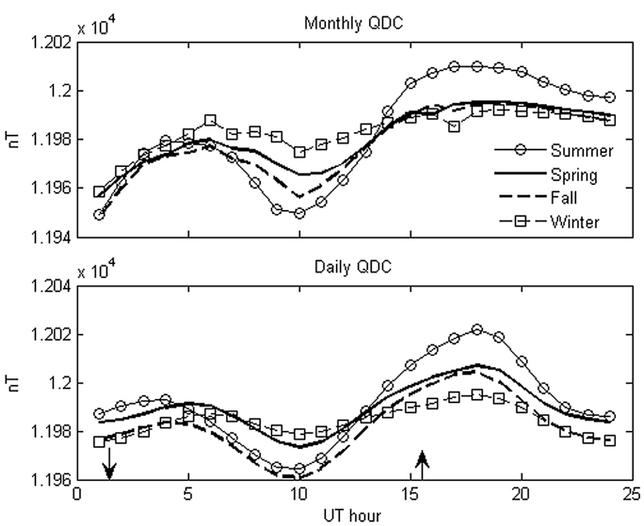

Figure 1. Average seasonal daily curves obtained from the magnetic *H* component in 1914–2000 at SOD, as defined by the Kalman filter (daily QDC), and the formerly used monthly QDC methods. The approximate peak time and the expected effect of westward (eastward) electrojet is marked by a downward (upward) arrow at 01:25 (15:30, respectively) UT. Note that SOD LT is 2 h ahead of UT.

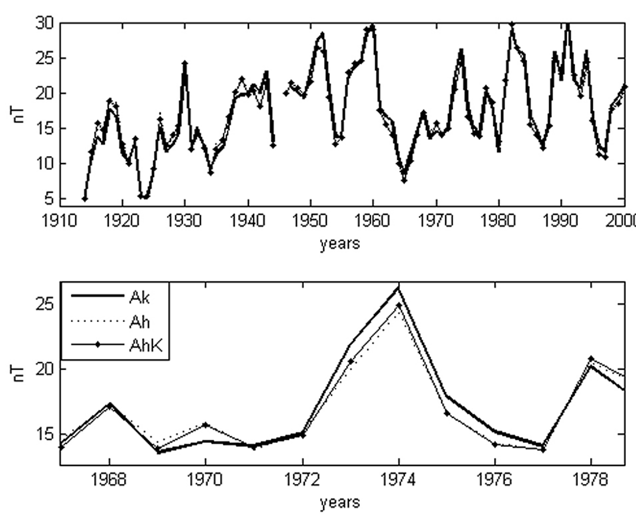

Figure 3. (top) The annual averages of the *Ak* (thick black line), *Ah* (dotted line), and A_{hK} (thin solid line with dots) indices at Sodankylä in 1914–2000. *Ah* and A_{hK} indices were normalized to *Ak*. (bottom) An enlarged period in the 1970s that includes both maximum (around 1970) and declining (around 1974) phases of solar activity, where the systematic deviations between the analog *Ak* and the digital *Ah* and A_{hK} indices can be observed.

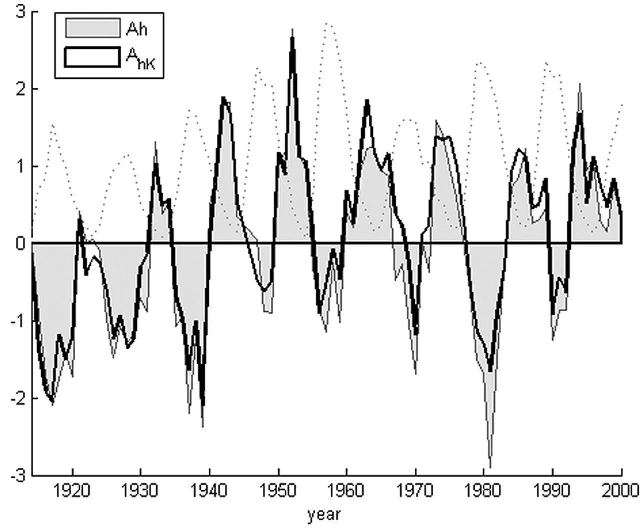

Figure 4. Residuals of linear regression fits to *Ak* of *Ah* (thick curve) and of A_{hK} (shaded) in nT. The zero base level is obtained when the best fitting value of the compared index is equal to *Ak*; negative (positive) value is found if the index value is larger (smaller, respectively) than *Ak*. For comparison the qualitative annual sunspot numbers are also indicated with a dotted line.

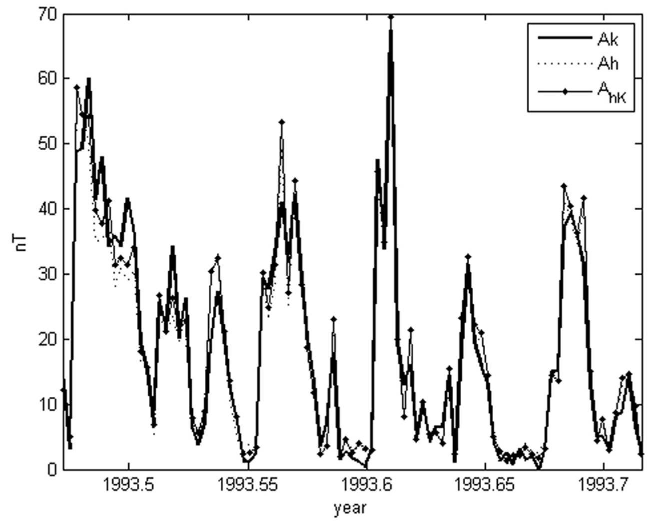

Figure 5. The daily averages of the *Ak* (thick solid line), *Ak*-normalized *Ah* (dotted line), and A_{hK} (thin solid line with dots) indices at Sodankylä during an arbitrarily selected late 1993 period.

daily resolution depicts much larger variability than that of annual averages, due to the short-term disturbances (such as storm, substorm) of the magnetosphere. The average standard deviation between *Ak* and the fitted *Ah* index is 6.96 nT; that is, it has increased considerably to about 40% of the average *Ak* level. Nevertheless, the agreement among the three indices is outstanding. *Mursula and Martini* [2007b] have already demonstrated that the correlation between the *Ak* and *Ah* indices remains remarkably high for daily or higher sampling. Using 31777 data points, the correlation is as good as r = 0.936. However, A_{hk} performs slightly, but significantly even better; its correlation with *Ak* is r = 0.944, while the standard deviation of the *Ak*-A_{hk} difference is 6.43 nT. The better correlation is visually demonstrated in Figure 6, where the low end of the scatterplots is shown for *Ak* and *Ah* (top), and for *Ak* and A_{hk} (bottom). Only every 20th points are depicted, but the best fitting lines were calculated by using all data points. Note that the correlation of the *Ak* and A_{hk} indices is moderately but observably improved compared to that of *Ah*, although the correlations of both digital indices with *Ak* are very good over the whole dynamic range (due to the minor difference only part of the dynamic range is shown). The best linear fits are $Ak = 0.42*Ah-0.42$; and $Ak = 0.43*A_{hk}-0.11$. The fitting parameters are still very close to each other, although A_{hk} consistently (including annual averages) depicts a somewhat smaller offset than *Ah*.

[21] We investigate the overall agreement of the indices, by showing the correlations with *Ak* at zero lag as a function of averaging time scale in Figure 7. The A_{hK} index has a significantly better correlation with *Ak* than *Ah*, at all time scales from daily to yearly. The difference between the old and the new method is smallest for annual averages, but, as we have seen in Figure 4 even on annual scale the daily QDC is a far more appropriate choice. The largest difference occurs at 27 days, where also local minima appear in the correlations with *Ak*. This is due to the fact that the analog and digital indices respond differently to disturbances driven by recurrent activity (dependent on high-speed solar wind streams).

4.3. Three-Hour Averages

[22] Figure 8 shows the average diurnal variation of the three mean-normalized indices, *ak*, A_{hK}, and *Ah* in 1914–2000 in the eight 3 h UT sector. (Note that *ak* notation stands for the highest sampling 3 h values, while *Ak* represents the daily

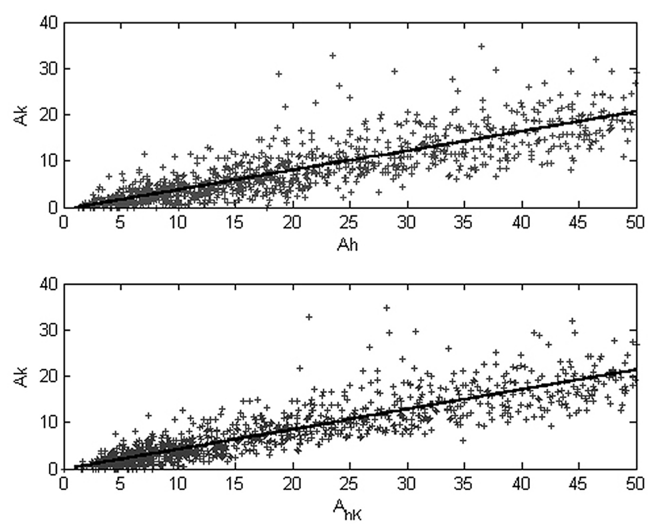

Figure 6. Scatterplot of the daily averages of (top) the *Ak* and *Ah* and (bottom) the *Ak* and A_{hK} indices in 1914–2000. Only every 20th point is depicted, but best fitting lines were calculated using all data points. The plots show only the low end of the total dynamic range (0–370 nT).

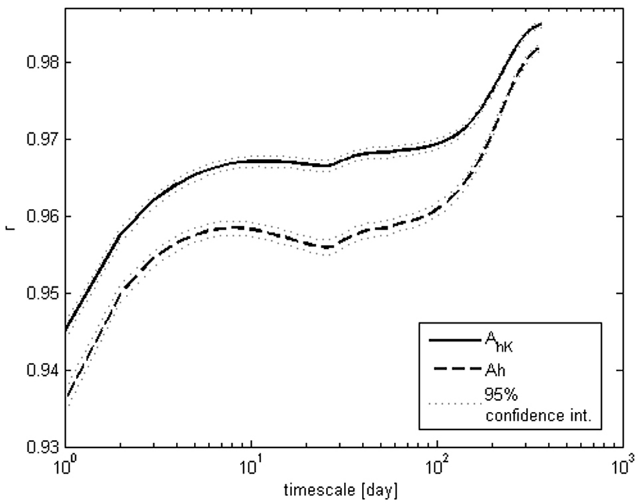

Figure 7. Correlation coefficients (r) between *Ak*, *Ah*, and A_{hK} as a function of averaging length. Because of the rapid change at time scales shorter than about a week, the time scale is plotted logarithmically. The significances of the coefficients are better than 99.9%, for both indices over the whole time scale.

or longer averages). This extremely good correlation yields compelling evidence for the detailed success of both the *Ah* and A_{hK} indices in general, in comprising the same magnetic phenomena as the *ak* index, allowing long-term studies up to 3 h resolution. We note, however, that the small differences occurring at the first UT sector exactly correspond to the westward electrojet peaking time, while deviations of the last two UT sectors coincide with the fallout following the eastward electrojet peaking time, discussed before.

[23] It is very interesting to further study the characteristics of the eight UT sectors. Figure 9 shows the correlation coefficients separately between the eight UT sectors of *ak* and *Ah*, and *ak* and A_{hK}. The 95% confidence interval is also depicted for each coefficient. Even at this resolution the correlations between *ak* and the digital measures are outstanding, ranging from r = 0.79 to r = 0.89. A_{hK} significantly outperforms *Ah* in the postmidnight hours, while *Ah* have significantly better correlation with *ak* around about local noon (which is also the sector of the overall minimum correlation). Since geomagnetic activity minimizes around the local noon hours at SOD (see Figure 8), the better correlation between *ak* and A_{hK} during the first two three-hourly UT sectors results in the overall significantly better performance of the A_{hK} index. It is interesting to note that the activity indices derived from the monthly QDC and the Kalman QDC methods practically do not differ during the afternoon to midnight sectors. While the monthly QDC is very inaccurate also in the premidnight hours (see Figure 1), this does not seem to lead to a significant degradation in the correlation between *ak* and *Ah*.

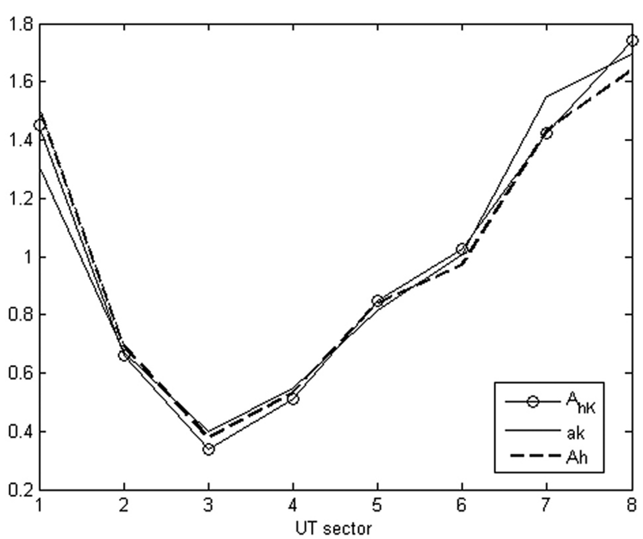

Figure 8. Average values in 1914–2000 of the self-normalized *ak*, *Ah*, and A_{hK} indices in the eight 3 h UT sectors separately. The 3 h sectors are centered to the middle hour.

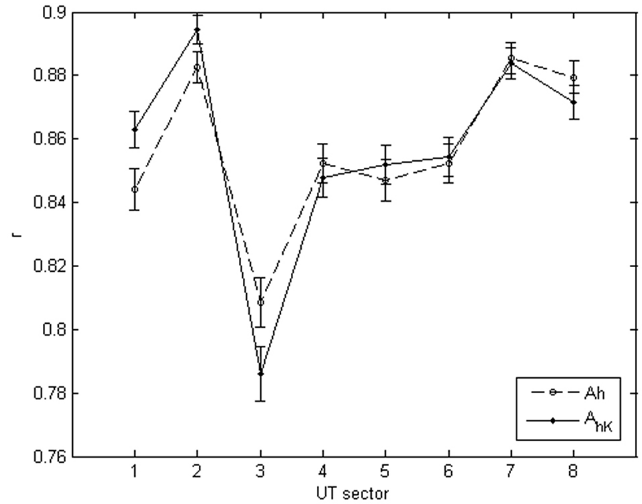

Figure 9. The correlation coefficients between the eight three-hourly UT sectors of *ak* and *Ah* (dashed line with circles) and *ak* and A_{hK} (solid line with dots) at Sodankylä in 1914–2000. The 95% confidence interval is indicated for each coefficient.

5. Discussion and Conclusions

[24] Unlike most of other recent attempts to derive digital geomagnetic measures [*Le Sager and Svalgaard*, 2004; *Svalgaard et al.*, 2004; *Finch*, 2008] our method gives the means to define the solar regular variation QDC much like in the traditional approach. This gives a unique flexibility to accommodate different preferences. Since the model is not restricted to hourly sampling, for shorter-term studies where good-quality high-sampling data are available, more accurate estimates can be obtained and deviations due to sampling differences can be minimized. Therefore, the algorithm can be used to derive a number of traditional indices of geomagnetic activity, such as the *Ak*, *Dst* or *AE* indices. This would provide a homogeneous derivation method over a wide variety of measures.

[25] Our study shows that the Kalman filter is an adequate method to define the regular variation from hourly data of

the geomagnetic field, even at high latitudes where such variation is strongly affected by the electrojet activity at all but the quietest days. Using the Kalman algorithm, the method implemented earlier to produce a digital *Ak*-type index *Ah* becomes self-consistent, free of the need for any additional input parameter to define geomagnetically quiet days. The new method of calculating a daily QDC outperforms in every aspect studied the previous *Ah* method of using monthly averaged QDCs. The Kalman filter is able to identify effects of electrojets that often mix with quiet time variation at high latitude, resulting in a QDC that more closely follows the sunspot number evolution in the long term and depicts the typical pattern of geomagnetic activity to a much lesser extent than monthly QDC. Therefore, it produces a more reasonable basis for calculating the 3 h range deviation, called the A_{hK} index.

[26] We find that A_{hK} is able to include disturbances of the directly driven system practically in the same way as the analog *Ak* index. The improvement is compelling, as A_{hK} shows a significantly better correlation with the *Ak* index than the former *Ah* index at all time scales.

[27] The only systematic deviation, due to the QDC definition, with respect to the analog *Ak* index is a moderate excess of A_{hK} values at or right after some of the solar maximum years. However, this excess is significantly reduced by the new daily QDC method.

[28] We find that even at high latitudes there seems to be a limit how well a digital index based on hourly data can match analog indices, due to the fact that mean values are less sensitive for high-frequency fluctuations of Alfvén waves of the high-speed solar wind streams, and somewhat more sensitive to HMF.

[29] **Acknowledgments.** This work has been supported by the WCU grant (R31-10016) funded by the Korean Ministry of Education, Science and Technology. We also acknowledge the support of the European Community Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme, LAPBIAT (RITA-CT-2006-025969). The work of Mikko Orispää was funded by the Academy of Finland (application 213476, Finnish Programme for Centres of Excellence in Research 2006–2011).

[30] Philippa Browning thanks the reviewers for their assistance in evaluating this paper.

References

Allen, J. H., and H. W. Kroehl (1975), Spatial and temporal distributions of magnetic effects of auroral electrojets as derived from *AE* indices, *J. Geophys. Res.*, *80*, 3667–3677, doi:10.1029/JA080i025p03667.

Bartels, J., N. H. Heck, and H. F. Johnston (1939), The three-hour-range index measuring geomagnetic activity, *Terr. Magn. Atmos. Electr.*, *44*, 411–454, doi:10.1029/TE044i004p00411.

Clilverd, M. A., T. D. G. Clark, E. Clarke, H. Rishbeth, and T. Ulich (2002), The causes of long-term change in the *aa* index, *J. Geophys. Res.*, *107*(A12), 1441, doi:10.1029/2001JA000501.

Clilverd, M. A., E. Clarke, T. Ulich, J. Linthe, and H. Rishbeth (2005), Reconstructing the long-term *aa* index, *J. Geophys. Res.*, *110*, A07205, doi:10.1029/2004JA010762.

Finch, I. D. (2008), The use of geomagnetic activity observations in studies of solar wind-magnetosphere coupling and centennial solar change, Ph.D. thesis, Univ. of Southampton, Southampton, U. K.

Finch, I. D., M. L. Lockwood, and A. P. Rouillard (2008), Effects of solar wind magnetosphere coupling recorded at different geomagnetic latitudes: Separation of directly driven and storage/release systems, *Geophys. Res. Lett.*, *35*, L21105, doi:10.1029/2008GL035399.

Grewal, M. S., and A. P. Andrews (1993), *Kalman Filtering: Theory and Practice*, Prentice Hall, Englewood Cliffs, N. J.

Kaipio, J., and E. Somersalo (2005), *Statistical and Computational Inverse Problems*, Springer, New York.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82*, 35–45.

Le Sager, P., and L. Svalgaard (2004), No increase of the interplanetary electric field since 1926, *J. Geophys. Res.*, *109*, A07106, doi:10.1029/2004JA010411.

Mayaud, P.-N. (1980), *Derivation, Meaning, and Use of Geomagnetic Indices*, *Geophys. Monogr. Ser.*, vol. 22, AGU, Washington, D. C.

Menvielle, M., and A. Berthelier (1991), The *K*-derived planetary indices: Description and availability, *Rev. Geophys.*, *29*, 415–432, doi:10.1029/91RG00994.

Mursula, K., and D. Martini (2007a), A new verifiable method of centennial geomagnetic activity: Modifying the K index method for hourly data, *Geophys. Res. Lett.*, *34*, L22107, doi:10.1029/2007GL031123.

Mursula, K., and D. Martini (2007b), New indices of geomagnetic activity at test: Comparing the correlation of the analogue *ak* index with the digital A_h and IHV indices at the Sodankylä station, *Adv. Space Res.*, *40*, 1105–1111.

Nevanlinna, H. (2004), Results of the Helsinki magnetic observatory 1844–1912, *Ann. Geophys.*, *22*, 1691–1704, doi:10.5194/angeo-22-1691-2004.

Svalgaard, L., E. W. Cliver, and P. Le Sager (2004), IHV: A new long-term geomagnetic index, *Adv. Space Res.*, *34*, 436–439, doi:10.1016/j.asr.2003.01.029.

D.-H. Lee and D. Martini, School of Space Research, Kyung Hee University, Gyeonggi-Do, 446-701 Yongin, South Korea. (dmartini@khu.ac.kr)

M. Lehtinen, M. Orispää, and T. Ulich, Sodankylä Geophysical Observatory, FI-99600 Sodankylä, Finland.

K. Mursula, Department of Physics, University of Oulu, FI-90014 Oulu, Finland.