Using recursive algorithm for defining solar regular geomagnetic variations

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

*[1]School of Space Research, Kyung Hee University, Republic of Korea; *e-mail: dmartini@khu.ac.kr*

[2]Sodankylä Geophysical Observatory, Sodankylä, Finland

[3]University of Oulu, Oulu, Finland

Motivation

One can separate two distinctly different types of variations in the geomagnetic field registered on the ground; the regular and irregular (classified as geomagnetic activity) variations. The former is mainly driven by the solar UV/EUV radiation and manifests itself as a smooth daily change in the magnetograms. The latter is a result of the dynamic fluctuations of solar wind and IMF (interplanetary magnetic field); hence it is of great interest for space climate studies. Motivated by recent attempts to characterize and quantify this geomagnetic activity from hourly mean data for long-term studies, we test the recursive Kalman filter method to obtain the regular solar variation curve of the geomagnetic field. In contrast to other recent approaches, we do not provide a method to quantify irregular activity directly but derive the actual quiet day curves first in the traditional manner. Therefore, in future applications the same algorithm may be used to define a wide variety of geomagnetic indices (such as Ak, Dst, or AE). Here we compare the Kalman method with former analog (Ak) and digital (Ah) activity estimates, based on Sodankylä station data.

The Kalman Filter

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 hourly mean magnetometer values (H-component) were divided into daily 24-hour (one day) bins after which we ran the Kalman filter using these 24 data points one at a time. Our goal is to find the **quiet day curve (QDC)** using Kalman filter. We made the following assumptions: *1) QDCs do not change very much from day to day; 2) QDC is quite smooth; 3) daily magnetic disturbances are considered as "noise".* From these assumptions we chose the evolution model to be simply identity matrix and the evolution noise is Gaussian with expectation value 0 and diagonal covariance matrix with diagonal equal to the variance of the evolution model noise.

$$X_{k+1} = X_k + W, \quad k = 0, 1, 2, \ldots \qquad (1)$$

$$Y_k = X_k + V_k, \quad k = 1, 2, \ldots, \qquad (2)$$

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

$$W \sim N(0, \sigma I) \qquad (3)$$

$$V_k \sim N(0, \Sigma_k), \qquad (4)$$

where σ is the given predetermined evolution-model error variance 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, ...):

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

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

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.

The observation model (2) together with the a priori value $\dot{X}_k$ 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

$$E(X_k) = E(\dot{X}_k) + K_k(Y_k - E(\dot{X}_k)), \qquad (5)$$

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

$$C(X_k) = (I - K_k)\, C(\dot{X}_k), \qquad (7)$$

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

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

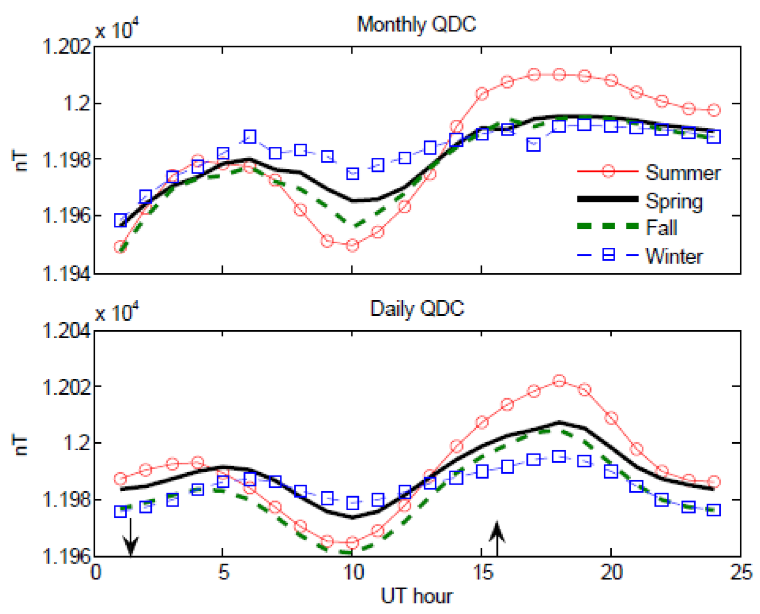
$$K_k = C(\dot{X}_k)\,(C(\dot{X}_k) + \Sigma_k)^{-1}, \qquad (9)$$

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

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. 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, $\dot{X}_k$, for X_{k+1} and so forth until the measurement data ends.

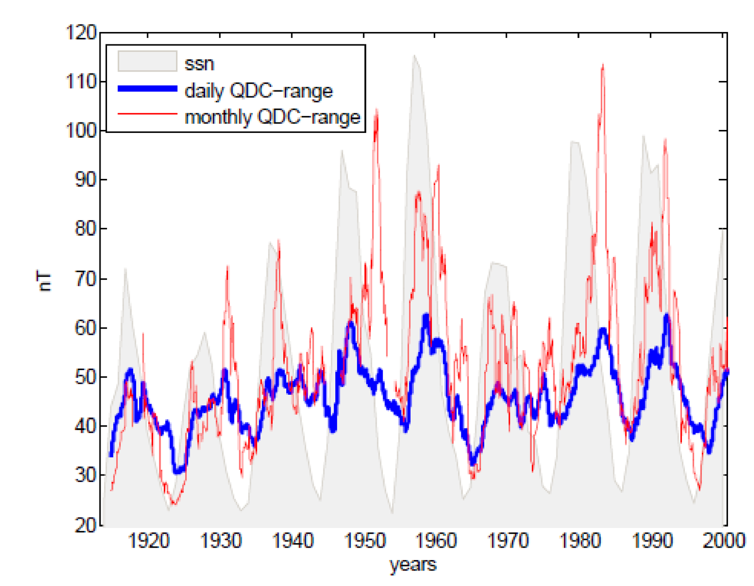
In addition, if the measurement data for two or more consecutive days are completely discarded due to data missing or exceeding the threshold, the Kalman filter is reset, i.e. the initial estimate is set back to X_0.

Accuracy of QDC estimates; in comparison with 'iron-curve' method at SOD station



Average seasonal daily curves in 1914-2000 at SOD, as defined by Kalman filter (daily QDC), and the formerly used monthly QDC methods. In latter case QDC is approximated by the averaged QDCs from the 5 quietest days in each months. The approximate peak time (and the expected effect) of westward (eastward) electrojet is marked by a downward (upward) arrow at 1.25 (15.30, respectively) UT. Note that the monthly 'iron-curve' method remains excited during night and is strongly affected by the westward electrojet.

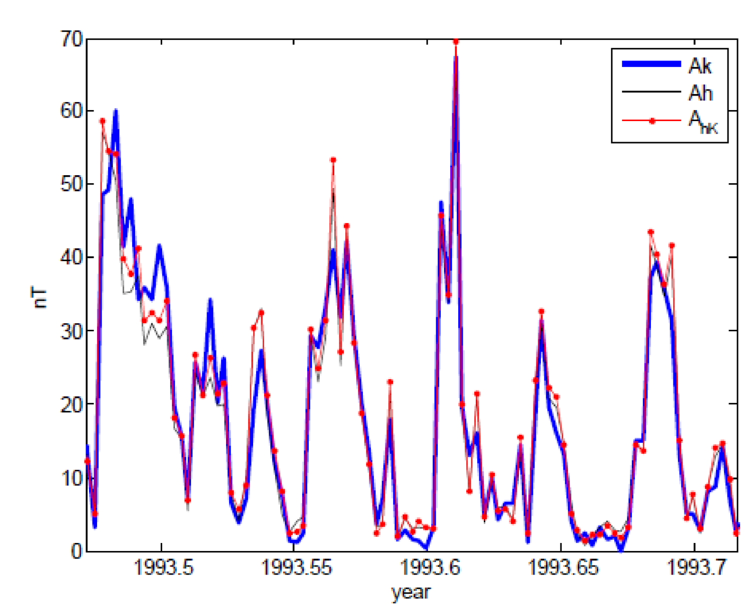
Kalman QDCs more closely depict the ideally expected quiet curve pattern



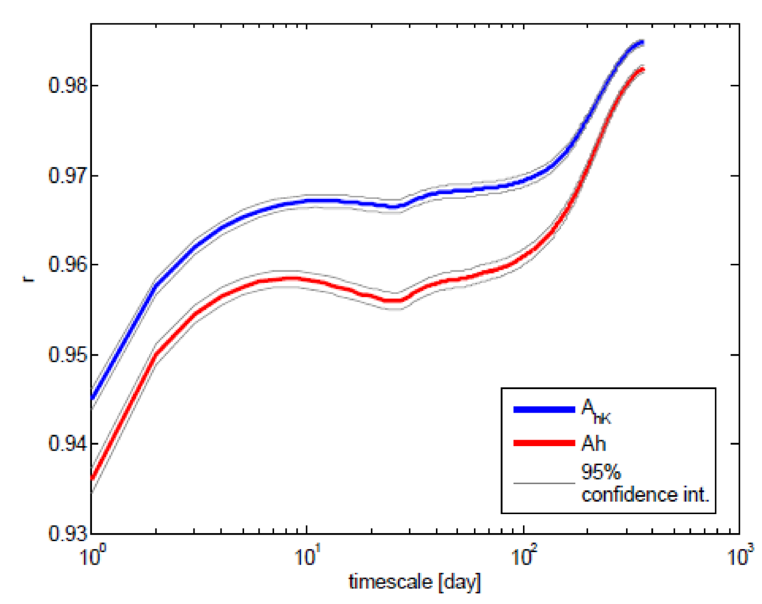
The annual (365 days) running means of the daily ranges (amplitudes) of daily QDCs defined by the Kalman filter, and of the average monthly QDCs.

While the QDC minimum levels of the two methods are roughly the same overall, the maxima differ radically. The iron-curves depict a more 'activity-like' pattern, indicating that above inaccuracies cause classifying actual irregular fluctuations as quiet pattern.

Derived magnetic indices from the iron-curve (Ah) and Kalman (A_{hK}) QDCs; Comparison with the local analog Ak index as reference measure



Daily averages of the Ak, Ak-normalized Ah, and A_{hK} indices at Sodankylä during an arbitrarily selected period. For the daily averages using 31777 data points, the correlation is as good as r=0.936 between Ak and Ah. However, A_{hk} performs slightly, but significantly even better; its correlation with Ak is r=0.944. (Note that even for daily averages inaccuracies in the QDC definition have a relatively minor effect in activity estimates).



The association, expressed by the linear correlation coefficients (r), between Ak and A_{hK} and that of Ah as a function of averaging timescales. The largest difference occurs at 27 days; Analog and digital indices respond differently to disturbances driven by recurrent activity (dependent on high-speed solar wind streams).

Using the daily quiet curve estimates of the Kalman algorithm the derived magnetic activity index, A_{hK}, has a significantly better correlation with Ak at all timescales from daily to yearly, than that of the former digital method, where monthly QDCs are averaged from the five quietest days of each month.

Summary and Conclusion

1. **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.**

2. **The new method of calculating a daily QDC outperforms in every aspect studied the previous Ah method of using monthly averaged QDCs (so called iron-curve method). Therefore, it produces a more reasonable basis for calculating the 3-hour range deviation, called the A_{hK} index.**

3. **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).**

For more details please find our paper on the subject:

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.**