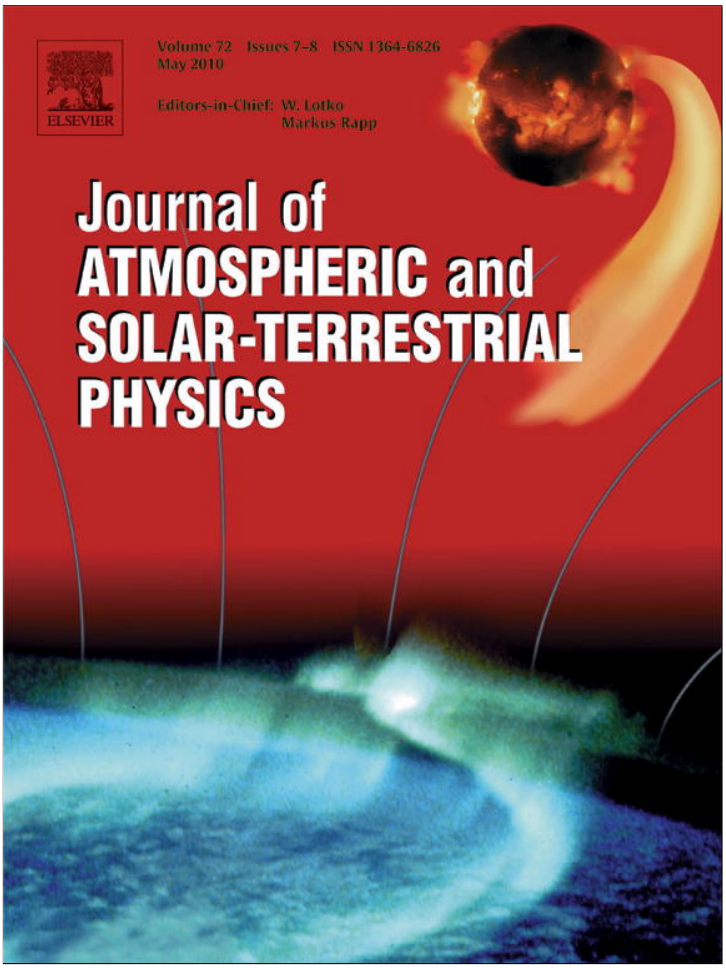

Journal of Atmospheric and Solar-Terrestrial Physics 72 (2010) 625–632

Contents lists available at ScienceDirect

Journal of Atmospheric and Solar-Terrestrial Physics

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

Estimation of Sq variation by means of multiresolution and principal component analyses

I. Maslova [a,*], P. Kokoszka [a], J. Sojka [b], L. Zhu [b]

[a] Department of Mathematics and Statistics, Utah State University, Old Main Hill 3900, Logan, UT 84322-3900, USA
[b] Department of Physics and Center for Atmospheric and Space Science, Utah State University, Old Main Hill 4405, Logan, UT 84322-4405, USA

ARTICLE INFO

Article history:
Received 21 May 2009
Received in revised form
25 January 2010
Accepted 6 February 2010
Available online 18 February 2010

Keywords:
Solar quiet variation estimate
Principal component analysis
Wavelets
Multiresolution analysis

ABSTRACT

A wavelet based method of deconvoluting the solar quiet variation from the low and mid-latitude H-component records is proposed. The resulting daily variation is non-constant, and its day-to-day variability is quantified by functional principal component scores. The procedure removes the signature of an enhanced ring current by comparing the scores at different stations. The method is fully algorithmic and is implemented in the statistical software R.

Published by Elsevier Ltd.

1. Introduction

We propose a new method of estimating the Sq component of low latitude magnetometer records. While our method builds on the method recently proposed by Chen et al. (2007), it is different in that it uses multiple stations and wavelet filtering.

An important part of the algorithm used to compute the Dst (Sugiura, 1964) is estimation of the quiet daily variation, which is essentially defined as the average of a few quiet days in a month. The Sq component calculated this way is the same for all days in a month, and it has been recognized that such an assumption is not accurate. Even on quiet days, the daily variation changes very visibly from day to day, both in its amplitude and its shape. This is attributable to multiple dynamic drivers which include not only tidal ionospheric winds, but also the effect of the Chapman–Ferraro current, the Sq current, and the magnetotail current, Xu and Kamide (2004), Chen et al. (2007), and references therein. On storm days, the interactions of these drivers are even more complex.

We propose a technique of isolating the low-latitude Sq variation which relies on wavelet and functional data analysis methods, and which is a refinement of an algorithmic step in the construction of an index of symmetric equatorial storm activity developed in Maslova et al. (2009). While the focus of the procedure of Maslova et al. (2009) was on extracting symmetric global features mainly attributable to the (enhanced) ring current, the present paper aims at isolating LT features attributable to ionospheric winds driven by solar heating. We do not estimate LT features due to the storm time enhancements of the partial ring current. A good technique for isolating Sq features should produce curves with a high degree of appropriately measured similarity (in LT) for neighboring stations. The LT H-component looks very different at different stations, but during the same UT day these different shapes are formed by approximately the same solar drivers, so the "correlation" of LT curves approximating the Sq for neighboring stations should be high. On the other hand, changes symmetric in UT, should not be reflected in the Sq estimates. These are our guiding principles for the construction of Sq estimates.

The traditional method of computing the quiet daily variation consists in finding five most quiet UT days (as measured by an index like Kp), averaging the H-components over these five days and smoothing. This results in a daily curve which differs across stations, but is constant for every day in a month at a given station. This approach has recently been improved upon by Jach et al. (2006) and Janzhura and Troshichev (2008) who use very different approaches and motivation, but produce estimates of the daily variation with very similar properties. Jach et al. (2006) focus on equatorial stations, and define the daily component as the median of the sum of three levels in a wavelet multiresolution analysis (MRA). The appropriate levels of the MRA isolate the time scales characteristic of equatorial Sq, whereas the median

* Corresponding author. Current address: Department of Mathematics and Statistics, American University, 4400 Massachusetts Ave., NW, Washington, DC 20016-8050, USA.
E-mail address: Maslova@american.edu (I. Maslova).

1364-6826/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.jastp.2010.02.005

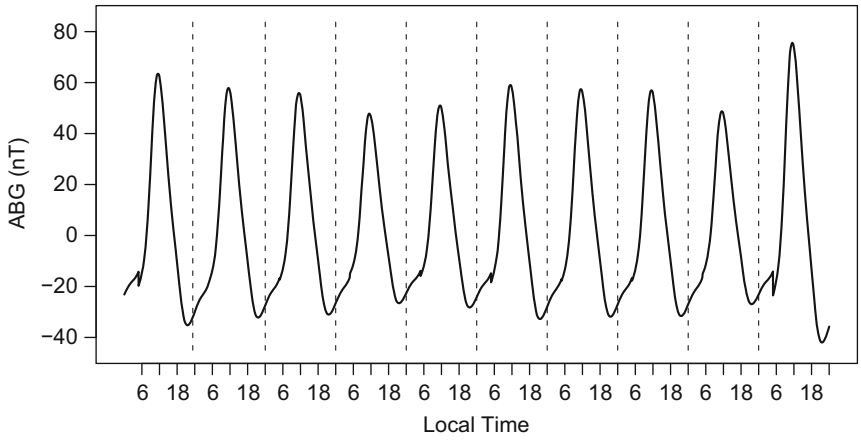

Fig. 1. Estimated Sq, Alibag (ABG) station during March 21–March 30, 2001.

produces a typical daily shape which is not affected by outliers (large disturbances). By using a moving window of flexible length (e.g. 30 days) an Sq which changes slowly from day to day can be calculated. Janzhura and Troshichev (2008) focus on auroral currents and use a more complex smoothing and outlier removal technique to determine a gradually changing quiet daily variation using a moving window of 30 days. By imposing certain conditions on the smoothness of the quiet daily curve, they can algorithmically identify quiet days. The time scales involved in their approach are 20 min and 3 h, and, appropriately to the task (auroral activity), are much smaller than the time scales used by Jach et al. (2006). Both techniques are fully algorithmic, and produce a slowly changing pattern (the curves on two consecutive days look almost identical). Bhardwaj and Rangarajan (1998) develop an algorithm for predicting such patterns, which incorporates the long term secular trend, the solar cycle and other long term variations. Their data are 24 time series (one for each UT hour) collected over a period of 70 years, producing 840 observations (one per month) for each of the 24 series.

Our goal is not the long term modeling of the evolution of a "typical" Sq pattern. We aim at good estimates of dynamic Sq over relatively short intra-annual periods capable of reflecting its often spectacular day-to-day variability (Butcher, 1989). By providing free software which implements our procedure, available at http://wami.usu.edu, we would like to offer a new tool for the community. An example of the variable Sq obtained using our software is shown in Fig. 1.

First, in Section 2 we describe measures of similarity of curves which we use to assess the quality of Sq estimates. We illustrate their applicability to synthetic Sq curves in Section 3. In Section 4, we describe the proposed new method in detail, and contrast it with the method of Chen et al. (2007). Section 5 focuses on the comparison of the two methods, while Section 6 concludes.

2. Measures of time-aligned similarity of curves

In this section we discuss measures of curve similarity that we use to compare methods of Sq estimation.

Correlations are often used as a measure of association. Let $Q^{(1)}$ and $Q^{(2)}$ be two samples of size N. The sample mean is $\overline{Q}=(1/N)\sum_{i=1}^{N} Q_i$, and sample variance is $s_Q^2=(1/(N-1))(Q-\overline{Q})^2$. The sample correlation between $Q^{(1)}$ and $Q^{(2)}$ is

$$r_{Q^{(1)},Q^{(2)}}=\frac{\sum_{i=1}^{N}(Q_i^{(1)}-\overline{Q}^{(1)})(Q_i^{(2)}-\overline{Q}^{(2)})}{(N-1)s_{Q^{(1)}}s_{Q^{(2)}}}.$$

Correlation is affected by extremely large values (outliers, see e.g. Freedman et al., 1991, Chapters 8 and 9). In the context of this paper, such extreme values can occur when the Sq estimate contains storm related features, which would result in a highly correlated data. We therefore propose a new measure of similarity of two samples of curves.

Let the first sample is denoted $Q_n^{(1)}(t)$, $t\in[0,T]$, $n=1,2,\ldots N_D$, the second $Q_n^{(2)}(t)$, $t\in[0,T]$, $n=1,2,\ldots N_D$. In the context considered in this paper, the interval $[0,T]$ represents an LT day, and N_D the number of days in each sample.

We define

$$D_n(Q_n^{(1)},Q_n^{(2)};a)=\frac{1}{T}\sum_{t=1}^{T}|Q_n^{(1)}(t)-aQ_n^{(2)}(t)|,\quad n=1,2,\ldots,N_D$$

and

$$\hat{D}(Q^{(1)},Q^{(2)})=\min_a\frac{1}{N_D}\sum_{n=1}^{N_D}D_n(Q_n^{(1)},Q_n^{(2)};a). \tag{1}$$

The measure $\hat{D}(Q^{(1)},Q^{(2)})$ is small if the two samples have similar time aligned features. Unlike correlation, $\hat{D}$ is not shift invariant, i.e. if a constant is added to all curves in one of the samples, the value of $\hat{D}$ will change. This is a desirable property because base line fields should be removed form estimates of Sq variations. The two samples do not have to be identical for $\hat{D}$ to be zero. If for each n, $Q_n^{(2)}(t)= cQ_n^{(1)}(t)$, for some constant c, then $\hat{D}(Q^{(1)},Q^{(2)})=0$.

In addition to the measure of the curve similarity introduced above we use the wavelet power spectrum to analyze the estimated Sq.

Let Q_t, t = 0, 1, …, N be the one minute estimate of the Sq of length N.

The discrete wavelet power spectrum associated with a scale $\tau_j=2^{j-1}$, where j = 1, 2, …, J is

$$P_{\tilde{W}}(\tau_j)=\frac{1}{N}\|\tilde{\boldsymbol{W}}_j\|^2,$$

where $\tilde{\boldsymbol{W}}_j$ is a vector the wavelet coefficients obtained filtering Q_t (see Percival and Walden, 2000, Chapter 5 for more details). The empirical power spectrum is the variance of the wavelet coefficients. We use it to see which frequency of the estimated Sq is most pronounced. The values of $P_{\tilde{W}}(\tau_j)$ should be higher for the levels j that capture daily variations.

We test measures of curve similarity using synthetic Sq estimates introduced in the following section.

3. Application to synthetic Sq curves

In this section we construct synthetic examples of "good" and "bad" Sq estimates, and apply to them the curve similarity measures described in Section 2. We simulate several pairs of synthetic Sq estimate curves (see Figs. 2 and 3). Assume that these pairs of curves are the estimates found using data from neighboring stations aligned in local time (LT). The "bad" Sq estimate includes some extreme values or even patterns that are present at all stations and aligned in UT. It means that the storm features were not completely removed. The "good" Sq estimate consists of the daily pattern that stays roughly the same each day. The amplitude of such an estimate varies slightly from day to day. Therefore, it is an estimate of a nonconstant SQ. Finally, there are no global features present.

Define the synthetic daily pattern as

$$sQ(t) = 0.6\sin\left(\frac{\pi t}{1440}\right) + 0.2\sin\left(\frac{2\pi t}{1440}\right), \quad t = 1,\ldots,1440.$$

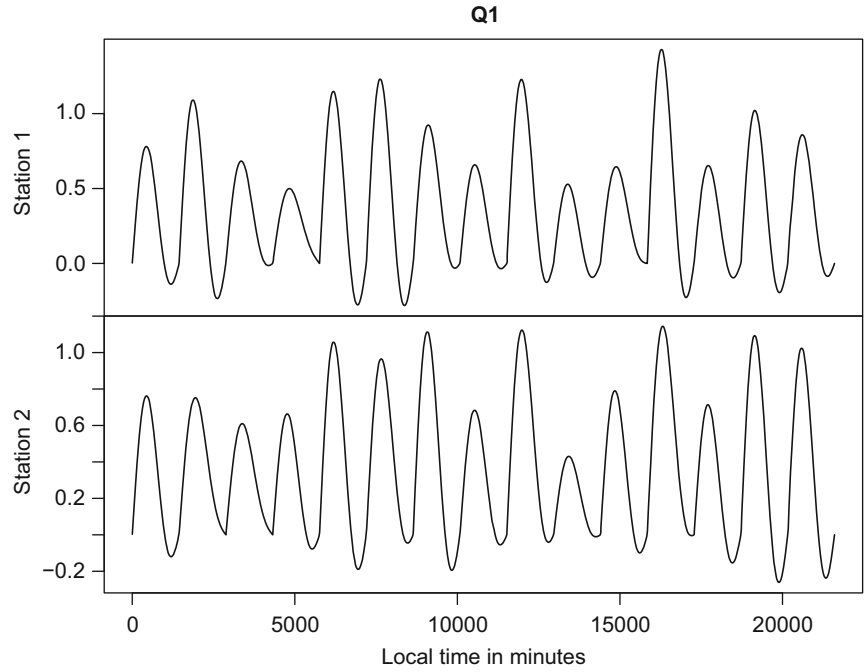

Fig. 2. Synthetic "good" Sq example. The most pronounced features are LT aligned.

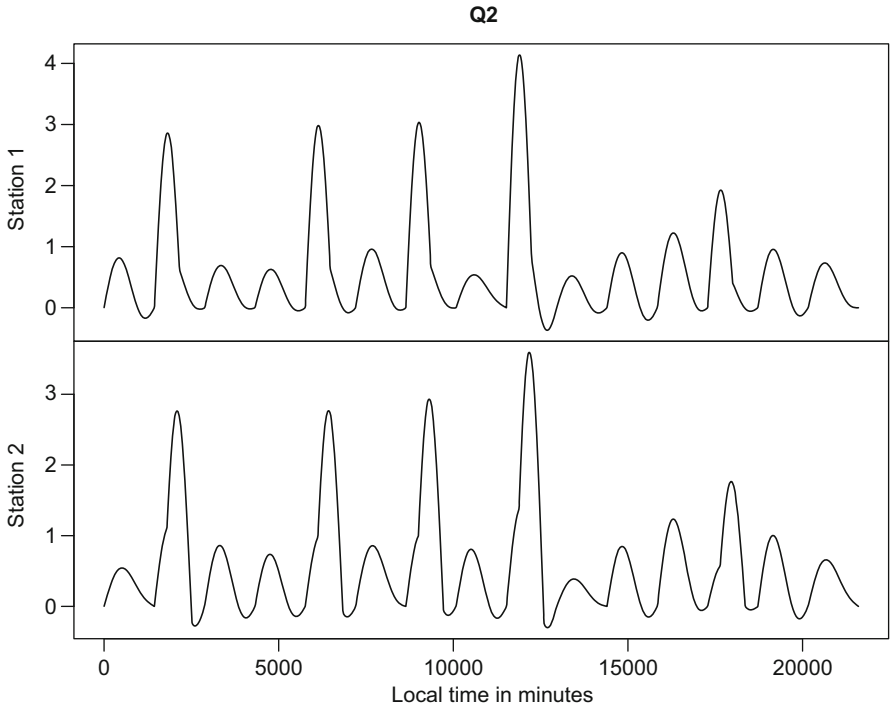

Fig. 3. Synthetic "bad" Sq example. Storm features are aligned in UT but shifted in LT.

Next, we generate a "good" Sq estimate $Q^{(1)}_{s,n}(t)$, for which the main features are adjusted in local time. Let

$$Q^{(1)}_{s,n}(t) = \{sQ(t) + R_s(t)\}U_n,$$

where U_n is a uniformly distributed on [0.5, 1.5] daily noise that is the same for all stations s. Here,

$$R_s(t) = 0.5V_{s,n}\sin\left(\frac{2\pi t}{1440}\right),$$

where $V_{s,n}$ is uniformly distributed on [0, 1] noise. This is an additive daily noise function that is different for all stations s.

Thus, $Q^{(1)}_{s,n}(t)$ simulates a "good" non-constant Sq estimate that captures daily pattern aligned in local time. Day-to-day variability is introduced by the noise variables U_n and $R_s(t)$ (see Fig. 2).

Next, we present an example of a "bad" non-constant Sq estimate. Define

$$G(t) = 2\sin\left(\frac{2\pi t}{1440}\right)I_{[0;720]}.$$

Let

$$G_s(t) = G(t + \Delta_s),$$

where $\Delta_1 = 0$, $\Delta_2 = 360$, $\Delta_3 = 520$, $\Delta_4 = 720$. Here, Δ represents the shift of the pattern that occurs when some UT features are not removed properly.

Set

$$Q^{(2)}_{s,n} = \{sQ(t) + R_s(t) + G_s(t)X_n\}U_n,$$

where X_n is a Binomial random variable with $p = 0.3$.

Fig. 3 presents an example of a "bad" Sq estimate where the extreme spikes are not aligned in time. The same happens when the global storm features are not removed from the quiet day component estimate.

Next, we simulate both "good", $Q^{(1)}_{s,n}$, and "bad", $Q^{(2)}_{s,n}$, Sq estimate curves for four stations during 15 days, i.e. s = 1, …, 4, n = 1, …, 15. These curves are used to evaluate the curve similarity measures introduced in previous section. The associations between simulated data sets for different stations are roughly linear, therefore the correlation analysis is appropriate. Panel (a) of Fig. 4 provides the correlations for two simulated Sq estimates. Since we generate data for four stations we get six different combinations, hence, six correlation values for each Sq method. We conclude that correlations separate these two synthetic Sq estimates. Correlations capture the shifts in time Δ, but are not sensitive towards changes in the amplitude.

However, in case of real data simple correlation does not perform as good as for simulated data. First, the association of Sq extracted from real magnetometer data from different stations is not linear. Second, if the method of Sq estimation fails to remove most of the storm features the correlation values become very high. In that case the high correlation is due to the extremely large values in Sq estimate which are attributable to the remaining storm signature. Therefore, a clearly "bad" Sq estimate gives high correlation values. Panel (b) of Fig. 4 shows that minimized average distance clearly distinguishes between the "good" and "bad" Sq estimates.

4. Estimation of a non-constant solar quiet daily variation

A method of natural orthogonal components, which we call here the principal components, to identify the solar quiet variation is introduced by Golovkov et al. (1978). Like for Chen et al. (2007), this observation is the starting point of our method. We, however, introduce a number of important refinements.

First, a wavelet-based representation of the data is introduced. Let, $X_s = \{X_{s,t}, t = 1, \ldots, N\}$ be the magnetogram recorded at station s = 1, …, m, where N is the length of the record in minutes, e.g. two months. We can write it as

$$X_s(t) = \sum_{j=1}^{J} D_{s,j}(t) + S_{s,j}(t),$$

where $D_{s,j} = \{D_{s,j}(1), \ldots, D_{s,j}(N)\}$ are the details, and $S_{s,j} = \{S_{s,j}(1), \ldots, S_{s,j}(N)\}$ is the smooth. Here, j = 1, …, J is the multiresolution analysis (MRA) level. The details capture the part of the records that corresponds to the frequencies in the range from 2^{-j-1} to 2^{-j} cycles per minute. For further details see Percival and Walden (2000, Chapter 5).

The Sq component is most clearly pronounced in the MRA details $D_{s,j}$ for levels j = 8, 9, 10. These levels capture different parts of the Sq spectrum; level j = 8 captures approximately the 6 h periodic component, j = 9–12 h component, and j = 10–24 h component. However, these details are enhanced during a storm. Comparing the records from several stations we can clearly see that these disturbances are aligned in UT. Therefore, they are not the part of Sq and should be removed. To remove storm associated features we use the storm index introduced in Jach et al. (2006) and improved by Maslova et al. (2009). Four Dst stations are used to construct the storm activity index (see Table 1). However, one can use any roughly equispaced equatorial stations. The stations

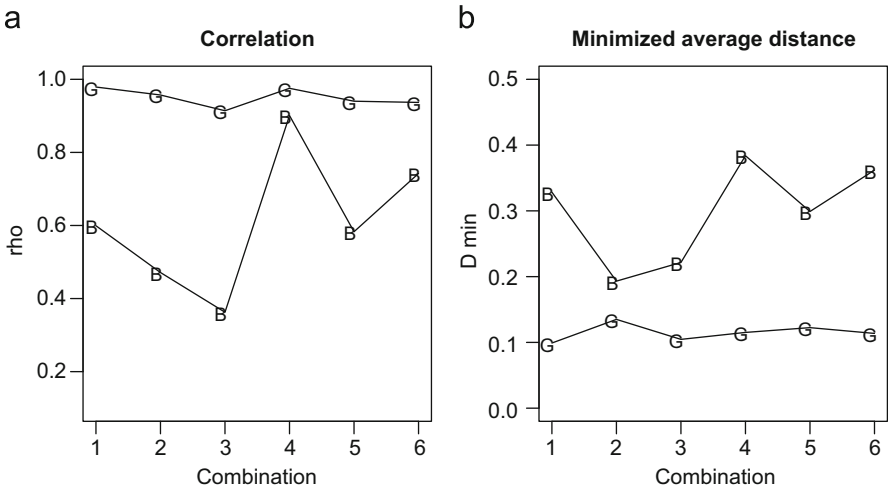

Fig. 4. Measures of curve similarity for "good" (G) and "bad" (B) Sq estimates: (a) correlation, (b) minimal average distance (N_D=15).

Table 1
Geomagnetic observatories used in this study.

s	Name	Colatitude	Longitude
1	Hermanus* (HER)	124.43	19.23
2	Alibag (ABG)	71.38	72.87
3	Phuthuy (PHU)	68.97	105.95
4	Kakioka* (KAK)	53.77	140.18
5	Honolulu* (HON)	68.68	202.00
6	Tucson (TUC)	57.82	249.27
7	Fredericksburg (FRD)	51.80	282.63
8	San Juan* (SJG)	71.89	293.85

Stations used to estimate the ring current activity are labeled with *.

used for storm index estimation must not include the stations where Sq is to be estimated.

Let, $I(t)$, $t = 1, \ldots, N$ be the storm index of Maslova et al. (2009) reflecting the strength of the ring current. We remove $I(t)$ from the H-component of all the stations used in the study.

After removing the storm index from the data, we perform the multiresolution analysis (MRA), using maximum overlap discrete wavelet transform MODWT and LA(8) filter. We can write it as

$$X_s(t) - I(t) = \sum_{j=1}^{J} D_{s,j}(t) + S_{s,j}(t), \quad t = 1, \ldots, N.$$

Let

$$D_{s,Q}(t) = D_{s,8}(t) + D_{s,9}(t) + D_{s,10}(t), \quad t = 1, \ldots, N$$

be the part of the signal that includes practically all frequencies of the daily component spectrum. The subscript "Q" stands for "quiet" daily component.

We therefore postulate that

$$D_{s,Q}(t) = Q_s(t) + R_s(t), \quad t = 1, \ldots, N,$$

where Q_s is identified with the solar quiet daily periodic component and R_s is the residual effect of a storm. Next, we apply principal component analysis techniques to estimate the daily variation Q_s. We convert $D_{s,Q}$ into functional objects, i.e. daily functions that start at UT midnight. Using principal component analysis we can write (t' is the time in minutes within one day)

$$D_{s,Q}(t') = \mu_s(t') + \sum_{j=1}^{\infty} \gamma_{s,j} u_{s,j}(t'), \quad t' = 1, \ldots, 1440,$$

where $\mu_s(t')$ is the daily mean, $\gamma_{s,j}$ is a score vector for j th PC, and $u_{s,j}$ is the j th PC for station s.

Denote the number of days by $N_D = N/1440$. We assume that the periodic component for day $i = 1, \ldots, N_D$ is

$$Q_{s,i}(t') = \mu_s(t') + \gamma^*_{s,1,i} u_{s,1}(t'), \quad t' = 1, \ldots, 1440, \tag{2}$$

where $\mu_s(t')$ is the daily mean, $\gamma^*_{s,1,i}$ is a filtered score for the i th day described below. The function $u_{s,1}(t')$ is the first PC for station s. In (2) $\mu_s(t')$ and $u_{s,1}(t')$ are deterministic functions defined over the 24-h interval, and $\gamma^*_{s,1}$ are random weights that change from day to day. Hence, the extracted Sq, $Q_s(t)$, is non-constant. Note that $Q_{s,i}(t')$ where $t' = 1, \ldots, 1440$ and $i = 1, \ldots, N_D$ is the same daily periodic component as $Q_s(t)$ where $t = 1, \ldots, N$ split into daily functions.

Even after removing the storm signature from the magnetometer records the three selected MRA levels may still contain residual storm features. Daily scores of the first PC, $\gamma_{s,1}$, show extreme values during the days when a storm occurred. Therefore, they contain the residual signature of the storm which should be removed.

Let $M_{s,1}$ = median $(\gamma_{s,1,i}, i \le N_D)$ be the median first principal component score for station s. Further, let $p_{0.90,s}$ denote the 90th percentile of the absolute value of the daily median adjusted scores for station s, i.e. $|\gamma_{s,1} - M_{s,1}|$. We define

$$\gamma^*_{s,1} = \begin{cases} M_{s,1} & \text{if } |\gamma_{s,1} - M_{s,1}| > p_{0.90,s} \text{ for all } s, \\ \gamma_{s,1} & \text{otherwise,} \end{cases} \tag{3}$$

where $M_{\gamma_{s,1}}$ is the median score of station s. This means that to eliminate the residual storm effect from the daily scores we find the largest 10% of the scores $|\gamma_{s,1} - M_{s,1}|$ for each station s individually. If the extreme value is captured by all stations we replace it by the median score, $M_{s,1}$, of the corresponding station. The scores $\gamma^*_{s,1}$ defined in (3) are used to compute the daily periodic component P_s defined in (2). In order to get better Sq estimate it is necessary to use records from at least two stations. Otherwise there is no way to eliminate the residual global features, as described above.

Decomposition (2) is akin to the ideas of Xu and Kamide (2004) and Chen et al. (2007), who argued that the first principal component follows the pattern of the daily Sq-variation. However, while these authors work with the raw magnetometer records, we first remove the global storm signature to eliminate the storm effect and then apply a wavelet filter to the data and use just the levels that contain the periodic component. So, in our paper, to estimate daily periodic component Q_s, we eliminate the global storm features that are not the part of Sq. We compute the first PC of $D_{s,Q}$ rather than the first PC of the raw magnetometer data with some seasonal adjustments which do not remove the storm activity from the Sq. Our method combines data from multiple low latitude stations.

5. Comparison of the Sq estimates

The goal of this section is to provide a detailed comparison of the Sq estimate introduced in Section 4 and the method proposed by Chen et al. (2007). We refer to the approach proposed here as the *new method* and the approach of Chen et al. (2007) as the *alternative method*.

First, we introduce the data used for this comparison. We use the H-component of the ground-based magnetometer records. Table 1 provides the list of the geomagnetic stations used in our study. As mentioned in Section 4, first, we remove the storm index from the raw data. The storm index is computed from four Dst stations: Hermanus, Kakioka, Honolulu, San Juan.

We compare the estimates of the daily variation for the following pairs of stations: (1) Alibag (ABG) and Phuthuy (PHU), (2) Tucson (TUC) and Fredericksburg (FRD). Note that the stations in each pair are relatively close to each other, so that the data used for Sq extraction is generated by the same ionospheric configuration (see Fig. 5).

We present the results for two periods of time: a one month period, February 2001, and two month—March–April 2001. The main difficulty of the Sq estimation occurs during geophysically disturbed time, hence the choice of the time intervals. February 2001 is a relatively quiet period of time, the second period, March–April, contains several extremely strong storms, one of them took place on March 31, 2001 (see Fig. 6).

Fig. 7 provides a visual comparison of the two Sq estimates to the raw magnetometer records. One can see that the new Sq (dashed line) follows the quiet daily pattern (solid line) for both quiet and disturbed times. However, the alternative Sq (dotted line) is very affected by the extreme values of the storm signature. During the storm it drops up to -400 nT, which definitely contradicts the Sq definition. The new procedure eliminates the global storm effects, therefore, the Sq remains stable during very strong storms with only a mild enhancement during the sudden commencement phase (see bottom panel of Fig. 7). The poor

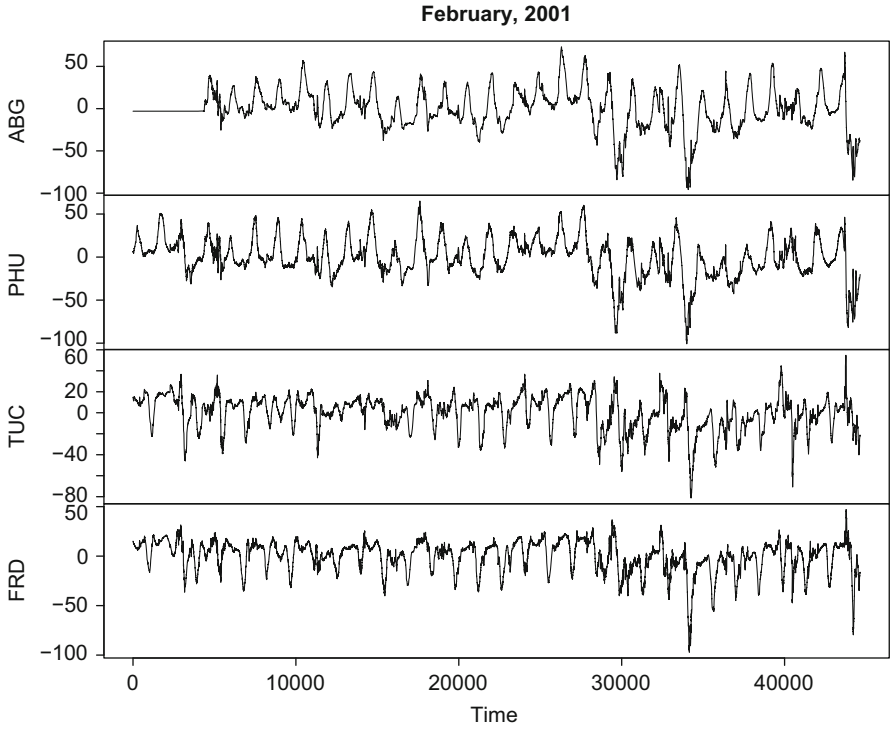

Fig. 5. Magnetic field H-component recorded at ABG, PHU, TUC, and FRD stations during February 2001.

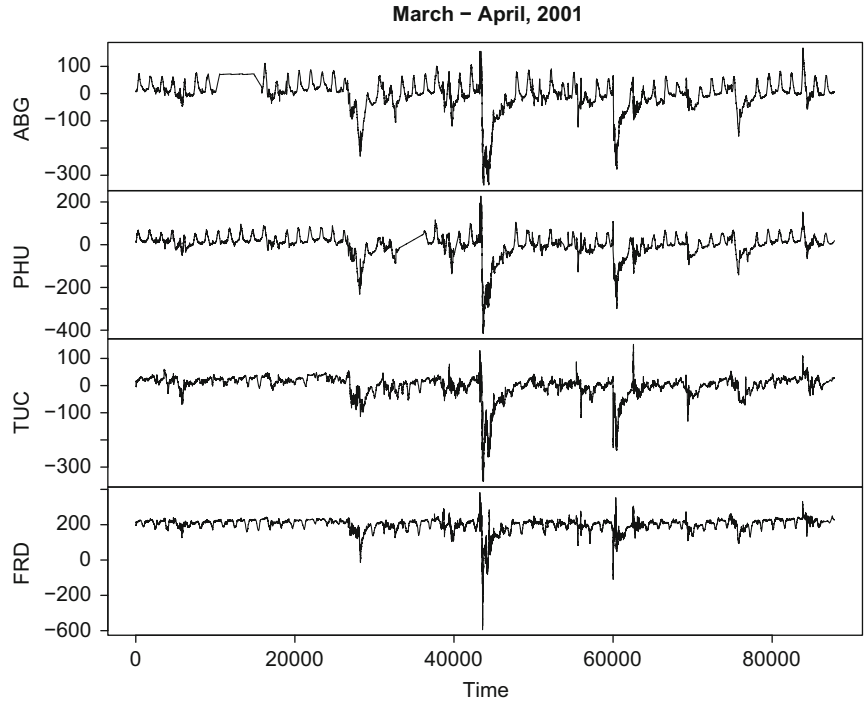

Fig. 6. Magnetic field H-component recorded at ABG, PHU, TUC, and FRD stations during March–April 2001.

performance of the alternative method during the quiet period shown in the top panel of Fig. 7 is due to the presence of storms during the two month period used to construct these Sq estimates. If a shorter period without strong storms is used for the estimation then the alternative method gives results comparable to the new procedure. Long quiet periods are, however, rare.

As part of the analysis, we compare the wavelet based power spectrums of the Sq obtained using our technique, the alternative methodology, and raw data. Fig. 8 presents an example of the empirical wavelet power spectrum based on MODWT for $j = 1, \ldots, 13$. Note that the daily variations are captured by levels $j = 8, 9, 10$. We can see that the raw data (circles) have a significant daily component present, as well as larger scale variations. A good Sq

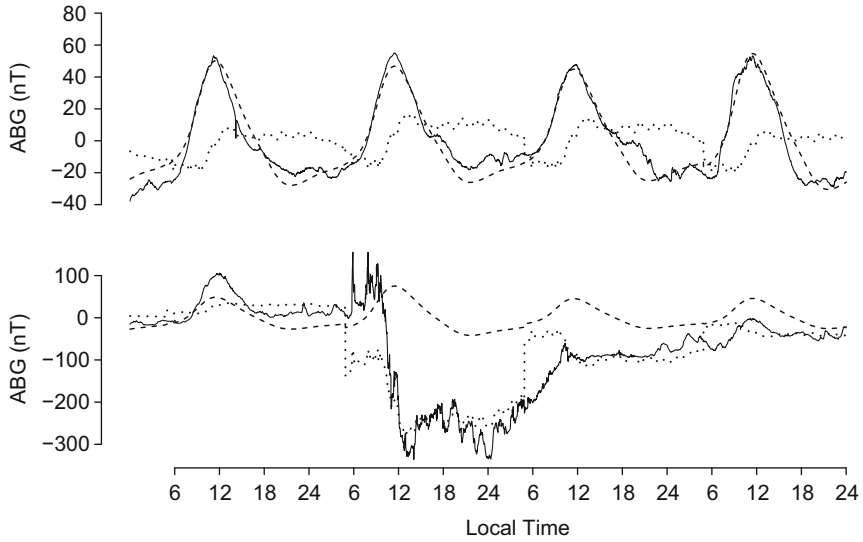

Fig. 7. Estimated Sq component using new methodology (dashed line), alternative approach (dotted line), and raw magnetometer data (solid line) at ABG station during quiet period of time: March 14–March 17, 2001 (top panel) and disturbed period of time: March 29–April 1, 2001 (bottom panel). The poor performance of the alternative method in the top panel is due to the presence of a storm in the two month period used to construct the estimates. Notice a moderate Sq enhancement of the new Sq estimate that follows the sudden storm commencement (bottom panel).

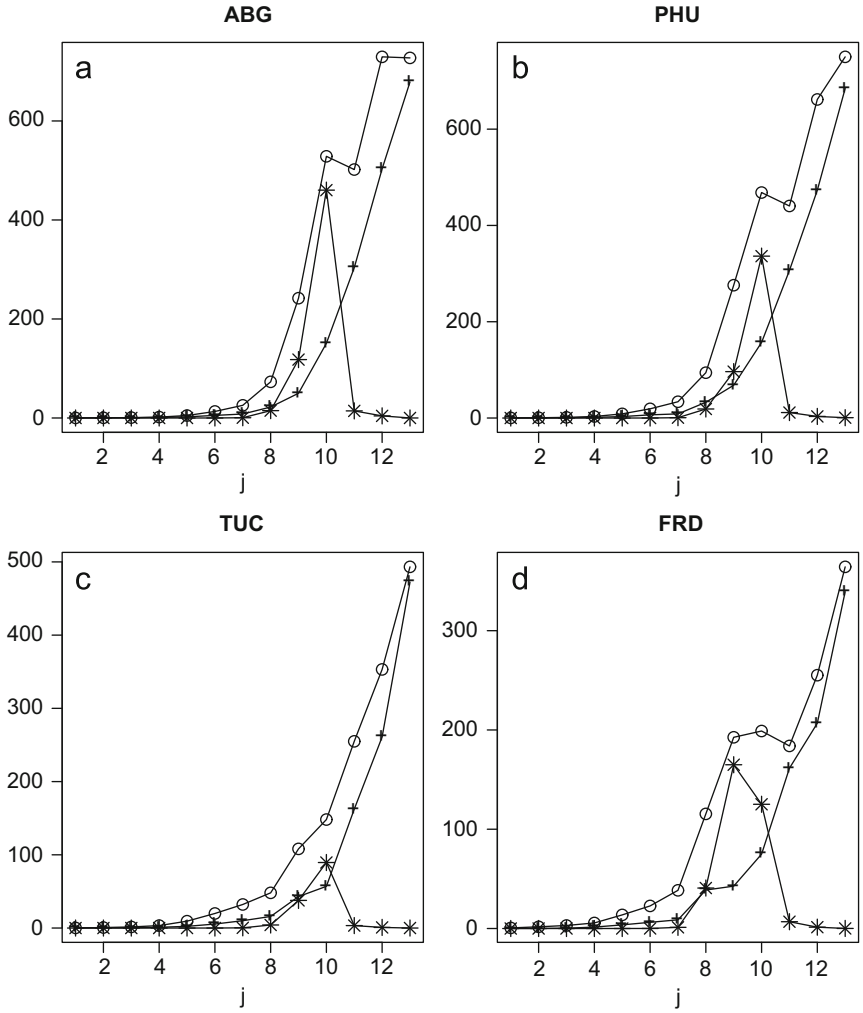

Fig. 8. Empirical wavelet power spectra based on MODWT levels $j = 1, \ldots, 13$ for estimated Sq component using new methodology (star), alternative approach (cross), and raw magnetometer data (circle) at (a) ABG, (b) PHU, (c) TUC, and (d) FRD stations during March–April 2001.

Table 2
Distance $\hat{D}$ (1) between the estimated Sq curves.

Method	February		March–April	
	ABG and PHU	TUC and FRD	ABG and PHU	TUC and FRD
New Sq	3.84	4.29	5.13	6.81
Alternative Sq	6.06	5.92	10.42	9.40

estimate should mostly capture the daily periodic component. The higher and lower frequency variations should be insignificant. One can see that our Sq estimate captures the daily variability, and the largest spectrum values are at j = 8, 9, 10, which is a desirable property (see the line with stars in Fig. 8). The spectra of the alternative Sq estimate roughly follows the raw data spectra. The daily component is not significant (see the line with crosses in Fig. 8). The highest values are observed for levels $j > 10$, which are associated with variations on scales larger than the daily scale.

Finally, we find the measure of the curve similarity $\tilde{D}$ for two pairs of the stations used in our study. Table 2 provides the quantitative results. The minimized average distance between the Sq estimates at neighboring stations is smaller for the technique introduced in this paper. This means that our proposed procedure extracts the local time aligned features better.

6. Conclusions

We introduce an automated procedure of extracting the Sq signature from the H-component of low-latitude magnetometer records. Wavelet and functional analysis techniques are applied. The methodology proposed here uses the index of storm activity to remove the main storm features as an initial step. In order to extract the daily variation we use the data from multiple stations. Multiresolution analysis is used to better isolate the part of the signal that captures the Sq component. We use the functional principal component approach to estimate the nonconstant daily variation. Our methodology gives the Sq estimate that extracts the local time features in a more accurate way and is more stable than other methods.

Acknowledgements

Research supported by NSF Grants DMS-0413653 and DMS-0804165.

Data were provided by USGS via the global network of observatories INTERMAGNET.

References

Bhardwaj, S.K., Rangarajan, G.K., 1998. A model for solar quiet variation at low latitude from past observations using singular spectrum analysis. Proceedings of the Indian Academy of Sciences (Earth and Planetary Science) 107, 217–224.

Butcher, E.C., 1989. Abnormal Sq behavior. Pure and Applied Geophysics 131, 463–483.

Chen, G.-X., Xu, W.-Y., Du, A.-M., Wu, Y.-Y., Chen, B., Liu, X.-C., 2007. Statistical characteristics of the day-to day variability in the geomagnetic sq field. Journal of Geophysical Research 112 doi:10.1029/2006JA012059.

Freedman, D., Pisani, R., Purves, R., 1991. Statistics, second ed. W.W. Norton and Company, New York, NY.

Golovkov, V.P., Papitashvili, N.Y., Tyupkin, Y.S., Kharin, Y.P., 1978. Separation of geomagnetic field variations into quiet and disturbed components by the method of natural orthogonal components. Geomagnetism and Aeronomy 18, 342–344.

Jach, A., Kokoszka, P., Sojka, J., Zhu, L., 2006. Wavelet-based index of magnetic storm activity. Journal of Geophysical Research 111, A09215.

Janzhura, A.S., Troshichev, O.A., 2008. Determination of the running quiet geomagnetic variation. Journal of Atmospheric and Solar-Terrestrial Physics 70, 962–972.

Maslova, I., Kokoszka, P., Sojka, J., Zhu, L., 2009. Removal of nonconstant daily variation by means of wavelet and functional data analysis. Journal of Geophysical Research 114, A03202 doi:10.1029/2008JA013685.

Percival, D.B., Walden, A.T., 2000. Wavelet Methods for Time Series Analysis. Cambridge University Press, Cambridge.

Sugiura, M., 1964. Hourly values of equatorial Dst for the IGY. Annals of the International Geophysical Year, vol. 35, no. 9, Pergamon Press, Oxford.

Xu, W.-Y., Kamide, Y., 2004. Decomposition of daily geomagnetic variations by using method of natural orthogonal component. Journal of Geophysical Research 109, A05218 doi:10.1029/203JA010216.