



Separating the aa-index into Solar and Hale Cycle Related Components Using Principal Component Analysis

Jouni Takalo¹

Received: 1 February 2021 / Accepted: 20 April 2021 / Published online: 10 May 2021
© The Author(s) 2021

Abstract

We decompose the monthly aa-index of Cycles 10 to 23 using principal component analysis (PCA). We show that the first component (PC1) is related to the 11-year solar cycle, and accounts for 41.5% of the variance of the data. The second component (PC2) is related to 22-year Hale cycle, and explains 23.6% of the variance of the data. The PC1 time series of the aa-index for Cycles 10–23 has only one peak in its power spectrum at the period 10.95 years, which is the average solar cycle (SC) period for the interval SC10–SC23. The PC2 time series of the same cycles has a clear peak at period 21.90 (Hale cycle) and a smaller peak at $3/4$ of that period. We also study the principal components of the sunspot number (SSN) for Cycles 10–23, and compare the mutual behavior of the PC2 components of the aa-index and SSN PCA analyses. We note that they are in the same phase in all other cycles but Solar Cycles 15 and 20. The aa-index of Cycle 20 also differs from other even aa-index cycles in its shape, especially in anomalously high peaks during its descending phase. Even though there is a coherence in the PC2 time series phases of the aa-index and sunspot number, this effect is too small to be the origin of all the differences between the shape of even and odd aa cycles. We estimate that 30% of the shape of the PC2 component of the aa-index is due to the shape of the PC2 of the sunspot number and the rest to other recurrent events in the Sun and solar wind. The first maximum of the aa-index (typical to odd cycles), during sunspot maximum, has been shown to be related to coronal mass ejections (CME), while the second maximum (typical to even cycles) in the cycle descending phase, is probably related to high-speed streams (HSS). The last events increase the activity level such that the minimum between even and odd cycle pairs is always higher than the minimum between succeeding odd and even cycle pairs.

Keywords Sun: sunspot number index · Earth: geomagnetic disturbances · Earth: aa-index · Methods: principal component analysis

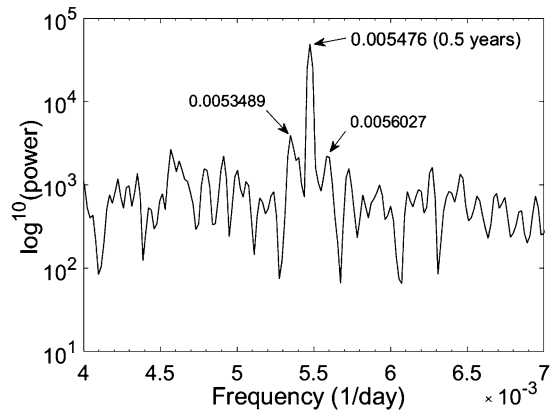
1. Introduction

The 22-year periodicity or double sunspot cycle in the geomagnetic activity has been studied since Chernosky (1966) noticed that the international geomagnetic index C_i revealed characteristically different patterns in even- and odd-numbered solar cycles. This 22-year cycle

✉ J. Takalo
jojuta@gmail.com

¹ Space Physics and Astronomy Research Unit, University of Oulu, POB 3000, 90014, Oulu, Finland

Figure 1 The amplitude modulation of the semiannual peak with the 22-year period in the daily aa-index in the interval 1868–2018.



is found in the recurrence of the 27-day geomagnetic activity index A_p and the ionospheric F2 layer variability (Apostolov, Altadill, and Todorova, 2004). Shnirman, Le Mouël, and Blanter (2009) applied the wave packet technique to the geomagnetic aa-index and found a 22-year variation in the interval 25–31 days centered at 27-day solar rotation period. More commonly, the 22-year modulation is observed in the semiannual variation in the geomagnetic indices. This modulation has been attributed to three various explanations: axial mechanism (Cortie, 1912; Bohlin, 1977), equinoctial hypothesis (Bartels, 1932; McIntosh, 1959; Svalgaard, 1977), and Russell–McPherron effect (Russell and McPherron, 1973) (see also the reviews by Cliver, Kamide, and Ling, 2000, 2002). The main reason for this 22-year variation is not within the scope of this article. In Figure 1 we, however, show the amplitude modulation of the semiannual peak with the 22-year period in the daily aa-index in the interval 1868–2018. The highest peak corresponds to the half-year frequency and the other two peaks at both sides of the main peak are side peaks due to the modulation. According to the amplitude modulation, the frequency difference of the side peaks is twice the modulation frequency (Takalo, Lohikoski, and Timonen, 1995). That is why we can calculate the modulation period as $2/(0.0056027 - 0.0053489) = 7880$ days = 21.6 years.

There have been some studies about decomposing the aa-index to separate components (Feynman, 1982; Echer et al., 2004; Hathaway and Wilson, 2006; Du, 2011). These studies use mostly annually averaged aa-index data. Feynman (1982) analyzed the relationship between the annual aa and sunspot-number R series from 1869 to 1975 and found that the aa-index values are all above a baseline (aaR) that is linearly related to R . Then she decomposed the aa-index into two equally strong periodic components: aaR and the remainder $aaI = aa - aaR$. The aaR component is associated with the transient phenomena and follows the sunspot cycle, while the aaI (interplanetary) component is associated with recurrent phenomena and is almost 180 degrees out of phase with the sunspot cycle. Hathaway and Wilson (2006) presented quite similar decomposition to that of Feynman (1982). One component is proportional to, and in phase with, the sunspot number and another, interplanetary, component aaI , which is out of phase with the sunspot cycle. This second component peaks some years before solar minimum, and is one of the most reliable indicators for the amplitude of the following solar maximum. Echer et al. (2004) stated that the aa-index has a clear double-peak structure and decomposed the aa-index data into the sunspot number R_z and fast solar-wind ($SW > 500 \text{ km s}^{-1}$) related terms. They found that 71% of the variability in the aa-index can be explained by a linear dependency on these terms. Du (2011) presented two different ways to decompose the aa-index into components. The first model

was similar to the aa model by Feynman (1982), but the components are (almost) perpendicular (90 degree phase shift) to each other. The second model is based on the logarithmic relationship between annual values of the aa-index and R_z . According to this model all aa-index values are between the lines $aa_t = e^{2.44} R_z^{2/7}$ (top-line) and $aa_b = e^{1.36} R_z^{2/7}$ (baseline), and the aa-index can be decomposed into independent terms using these lines (see a more thorough presentation in Du, 2011).

In this study we present a new way to decompose the aa-index into components using principal component analysis (PCA). We extract the main three principal components from the seven-month trapezoidal-smoothed aa-data. The first component is due to the 10.95-year solar cycle (SC), and the second component is related to the 21.90-year Hale cycle. According to the definition of PCA, these components are mutually orthogonal. The third (and upper) PCs show shorter periods and are related to features of some specific cycles. This paper is organized as follows. Section 2 presents the data and methods used in this study. In Section 3 we present the results of the PC analysis for the aa-index and the sunspot-number index of Solar Cycles 10–23 and discuss their connection in Section 4. We give our conclusions in Section 5.

2. Data and Methods

2.1. aa-index

Mayaud (1972) presented the geomagnetic activity aa-index, which is based on the K-indices of two antipodal stations, one in Australia and another in southern England, which started measurements in 1868. For the northern hemisphere the sites have been Greenwich (for 1868–1925), Abinger (1926–1956), and Hartland (1957–present), and for the southern hemisphere they are Melbourne (1868–1918), Toolangi (1919–1979), and Canberra (1980–present). Later this index has been extended for two solar cycles between 1844–1868 using measurements made in Helsinki (Nevanlinna and Kataja, 1993; Lockwood et al., 2013). There exists a provisional aa-index until the end of the year 2020 (<http://isgi.unistra.fr>), but the main analysis is made for the aa Cycles 10–23 (from August, 1856 to December, 2009), because the aa Cycle 24 is still incomplete, and for a complete Hale cycle analysis we need also the next odd Cycle 25.

The aa-index is, however, quite noisy and we have made a seven-month trapezoidal smoothing for the index before analysis. This also removes the semiannual variation in the aa-index. Trapezoidal smoothing is a moving average smoothing such that the end points of the window have only half of the weight from the other points. Also the cycles for the aa-index lag somewhat from those of the sunspot number (SSN). This is shown in Figure 2, where the seven-month smoothed sunspot number (SSN2.0, <http://www.sidc.be/silso/>) and the aa-index are plotted together. Note that the difference between aa cycles and solar cycles is largest in the middle of the period SC10–SC23, i.e. for Cycles 16–19 corresponding to the years 1923–1955. The correlation analysis, using seven-month smoothed indices, shows that the mean lag between the aa-index and the SSN for C10–C23 is about 10–11 months. In order to determine the minima of the aa Cycles 10–23, we used 13-month trapezoidal smoothed aa time series. Note that 13-month smoothing is used only for determining the aa-index minima, otherwise we use seven-month smoothing for both aa-index and SSN. The lags of these minima compared to SSN minima are shown in the last column of Table 1. Note that the aa time series starts eight months later than the SSN time series as shown in the first row in Table 1. It is also clear that the aa cycle minima lag solar cycle minima most in the middle of 20th century, i.e. for Cycles 16–19 (Kane, 2007).

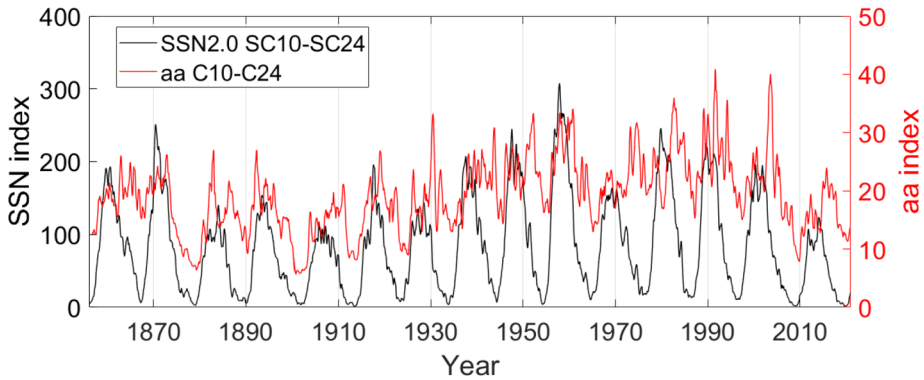


Figure 2 Comparison of aa-index and SSN2. The black line shows the SSN2 (y-axis left) and the red line shows the aa-index (y-axis right). Notice that aa-index maxima (minima) lag the maxima (minima) of SSN2.

Table 1 Sunspot cycle lengths and dates of starting minima used in the calculation of the analysis. The last column shows the lag of corresponding aa cycle minima.

Sunspot cycle number	Year and month of starting min.	Cycle length (years)	Lag of corresponding aa cycle minimum
10	1855 December	11.2	+8
11	1867 March	11.8	+8
12	1878 December	10.6	+5
13	1889 August	12.1	+14
14	1901 September	11.8	+3
15	1913 July	10.1	+4
16	1923 August	10.1	+17
17	1933 September	10.4	+11
18	1944 February	10.2	+16
19	1954 April	10.5	+15
20	1964 October	11.7	+9
21	1976 June	10.2	+9
22	1986 September	10.1	+6
23	1996 October	12.2	+12
24	2008 December		+12

2.2. Solar Sunspot Index

Rudolf Wolf collected the Zürich series of the sunspot number (SSN), which started in 1749. The first complete sunspot cycle included in SSN started in March 1755. Wolf started the numbering of sunspot cycles from this cycle and this numbering is still in use. The initial sunspot-number series (SSN1) was reconstructed at the Zürich Observatory until 1980 and at the Royal Observatory of Belgium since 1981. Following the change of the reconstruction method in 1981, the current version of the SSN series is called the international sunspot number (ISN). Recently the ISN series was modified to a version 2.0 (SSN2) that is supposed to present a preliminary correction of the known inhomogeneities in the SSN1 series (Clette et al., 2014). We have used the version SSN 2.0 in this study and call it later just SSN.

2.3. Principal Component Analysis Method

Principal component analysis is a useful tool in many fields of science, chemometrics (Bro and Smilde, 2014), data compression (Kumar, Rai, and Kumar, 2008), and information extraction (Hannachi, Jolliffe, and Stephenson, 2007). PCA finds combinations of variables that describe major trends in the data. PCA has earlier been applied, e.g., to studies of the geomagnetic field (Bhattacharyya and Okpala, 2015), geomagnetic activity (Holappa, Mursula, and Asikainen, 2014; Holappa et al., 2014), ionosphere (Lin, 2012), and the solar background magnetic field (Zharkova, Shepherd, and Zharkov, 2012; Zharkova et al., 2016). Before we can carry out the PCA, the vectors, in this case the cycles, must have equal lengths to form a matrix. In this article we estimate that the average length of the cycle is 131 months, and use it as a representative solar cycle length. We also use this same length (131 months) as a common length for aa-index cycles. To this end, we first resample the monthly sunspot and aa values such that all cycles have the same length of 131 time steps (months). This effectively elongates or abridges the cycles to the same length. Before applying the PCA method to the resampled sunspot cycles and aa cycles we standardize each individual cycle to have zero mean and unit standard deviation. This guarantees that all cycles will have the same weight in the study of their common shape. After applying the PCA method to these resampled and standardized cycles, we revert the cycle lengths and amplitudes to their original values (see also Takalo and Mursula, 2018).

As stated earlier, before applying the PCA to the aa-index data we standardize the time series of each individual cycle of the aa-index to have zero mean and unit standard deviation. Standardized data are then collected into the columns of the matrix X , which can be decomposed as (Hannachi, Jolliffe, and Stephenson, 2007; Holappa et al., 2014; Takalo and Mursula, 2018)

$$X = U D V^T, \quad (1)$$

where U and V are orthogonal matrices, V^T a transpose of matrix V , and D a diagonal matrix $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ with λ_i the i th singular value of matrix X . The principal components are obtained as the column vectors of

$$P = UD. \quad (2)$$

The column vectors of the matrix V are called empirical orthogonal functions (EOF) and they represent the weights of each principal component in the decomposition of the original normalized data of each cycle X_i , which can be approximated by

$$X_i = \sum_{j=1}^N P_{ij} V_{ij}, \quad (3)$$

where j denotes the j th principal component (PC). The explained variance of each PC is proportional to the square of the corresponding singular value λ_i . Hence the i th PC explains a percentage

$$\frac{\lambda_i^2}{\sum_{k=1}^n \lambda_k^2} \cdot 100\% \quad (4)$$

of the variance in the data.

Figure 3 The three first PCs of the aa-index for Cycles 10–23.

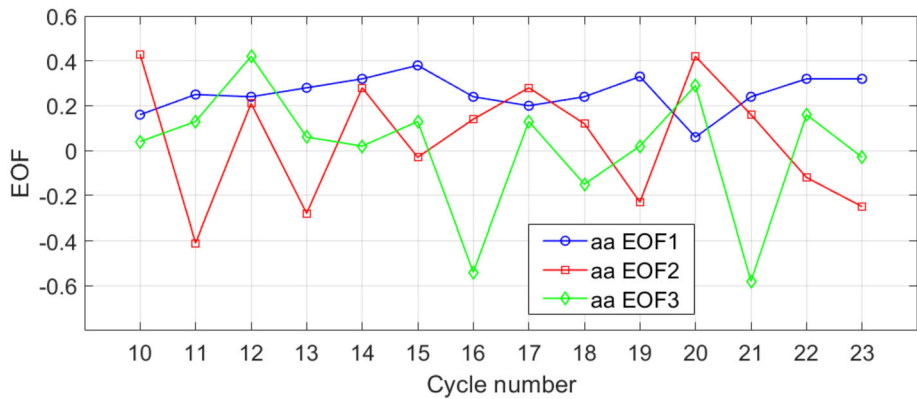
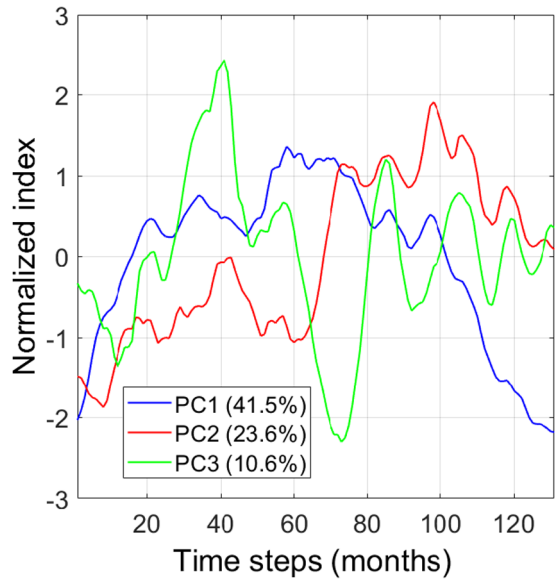


Figure 4 The three first EOFs of the aa-index for Cycles 10–23.

3. Results

3.1. PCA of the aa-index

Using modified (solar related) cycles for the aa-index C10–C23, we made a PCA by equalizing the aa cycles to 131 time steps (months) to get the three main principal components shown in Figure 3. The first, second, and third PCs explain 41.5%, 23.6%, and 10.6% of the total variation of the data, respectively. Hence the three main PCs account together for 75.7% of the smoothed aa-index. Although PC1 accounts only for 41.5% of the variance of the aa-index, the PC1 has correlation coefficient 0.998 ($p < 10^{-100}$) with the mean aa cycle (C10–C23). Note, however, that PC2 explains almost a quarter of the total variance in the aa-index.

Figure 4 shows the three EOFs of the aa cycles. Note especially the almost sawtooth like shape of the EOF2. All even numbered EOF2s are positive, except for Cycle 22, and all the

odd-numbered EOF2s are negative, except for Cycle 17 and Cycle 21. Looking to the shape of the aa PC2, we note that a positive EOF2 means positive phase in the second half of the cycle, i.e. in the descending part of the cycle, and a negative EOF2 means a positive phase in the ascending part of the cycle.

Reverting the PC1, PC2, and PC3 cycles of Figure 3 back to their original cycle amplitudes and lengths and concatenating them we get first, second, and third principal component time series for the aa C10–C23 shown in Figure 5a, b, and c, respectively. It is evident that the PC1 is dominated by the sunspot cycle related period. Interestingly, the size of PC1 is almost equal up to Cycle 19, although the accumulation of the activity level raises the background and the total height of the index. Cycle 20 seems to have smallest PC1 component (although its height is similar to the heights of the PC1s of Cycle 16 and 17 due to the background baseline) and after that the PC1 gradually increases for Cycles 21–23. Note that at the end of the time series of Figure 5 we show the Cycle 24 components as calculated from the PCA of C10–C24. It is evident that the shape of the PCs of Cycle 24 do not differ much from the other PCs, although with quite low amplitude.

In contrast with PC1, the aa PC2 shows clear Hale cycle period (see Figure 5b). Note that the consecutive even–odd cycle pairs show similar structure for three Hale cycles (the Hale cycle is traditionally referred to as even–odd cycle pair (Gnevyshev and Ohl, 1948; Wilson, 1988; Makarov, 1994; Cliver, Boriakoff, and Bounar, 1996)), i.e. for pairs 10–11, 12–13, and 14–15, although Cycle 15 has quite a small PC2 component. Cycles 16–17 are mutually in phase, which could be seen also from the EOF2 of the odd Cycle 17 (the first odd cycle with positive EOF2). The cycle pair 18–19 is again similar to earlier Hale cycles, but Cycles 20–21 are again in the same phase (C21 is another odd cycle with positive EOF2). The PC2 of Cycle 22 is the only even cycle with positive phase in the first half and negative phase in the second half of the cycle., i.e. Cycle 22 is the only even cycle with negative EOF2. Cycle 23 is a common odd cycle with negative EOF2. It seems that Cycle 24 is again a common even aa cycle (its PC2 is similar to the common even cycle PC2, although with quite low amplitude), but as stated earlier we need the next odd Cycle 25 to have a complete even–odd Hale cycle for our analysis.

The PC3 time series of Figure 5c shows higher order periods. Note that there are four cycles (C10, C14, C19, and C23) with very small PC3 component. The higher PC-components usually describe some special features of only a few cycles. Figure 6 shows the original aa-index time series, PC1 + PC2 + PC3 proxy time series, and the residual time series. Notice that some high peaks are cut out from the PC1 + PC2 + PC3 proxy, because they exist only in some individual cycles.

Figure 7a and b show the power spectra of PC1 and PC2 time series of Figure 5a and b, respectively. It is clear that the PCA effectively decomposes the solar cycle and Hale cycle (even–odd cycle) related periods from the aa-index. The solar cycle period in the aa-index is 10.95 years and the Hale cycle 21.90 years. In Figure 7b, there is another smaller peak at a period of three quarters of the Hale cycle. We believe that this peak is due to the cycles where the phase of the common even–odd sequence changes (the cycles with “wrong” phase are marked in light gray in Figure 5b). If we remove the incomplete Hale Cycles 16–17 and 21–22 such that succeeding cycle pairs are all in opposite phases (note that now the last pair is 20–23) we get only a higher peak at period 21.90 (shown with a dashed black line in Figure 7b).

3.2. Detrended aa-index

Figure 8 shows the original smoothed aa-index (red) and two background envelopes: solar cycle background, which goes through all minima (dash black broken-line), and Hale

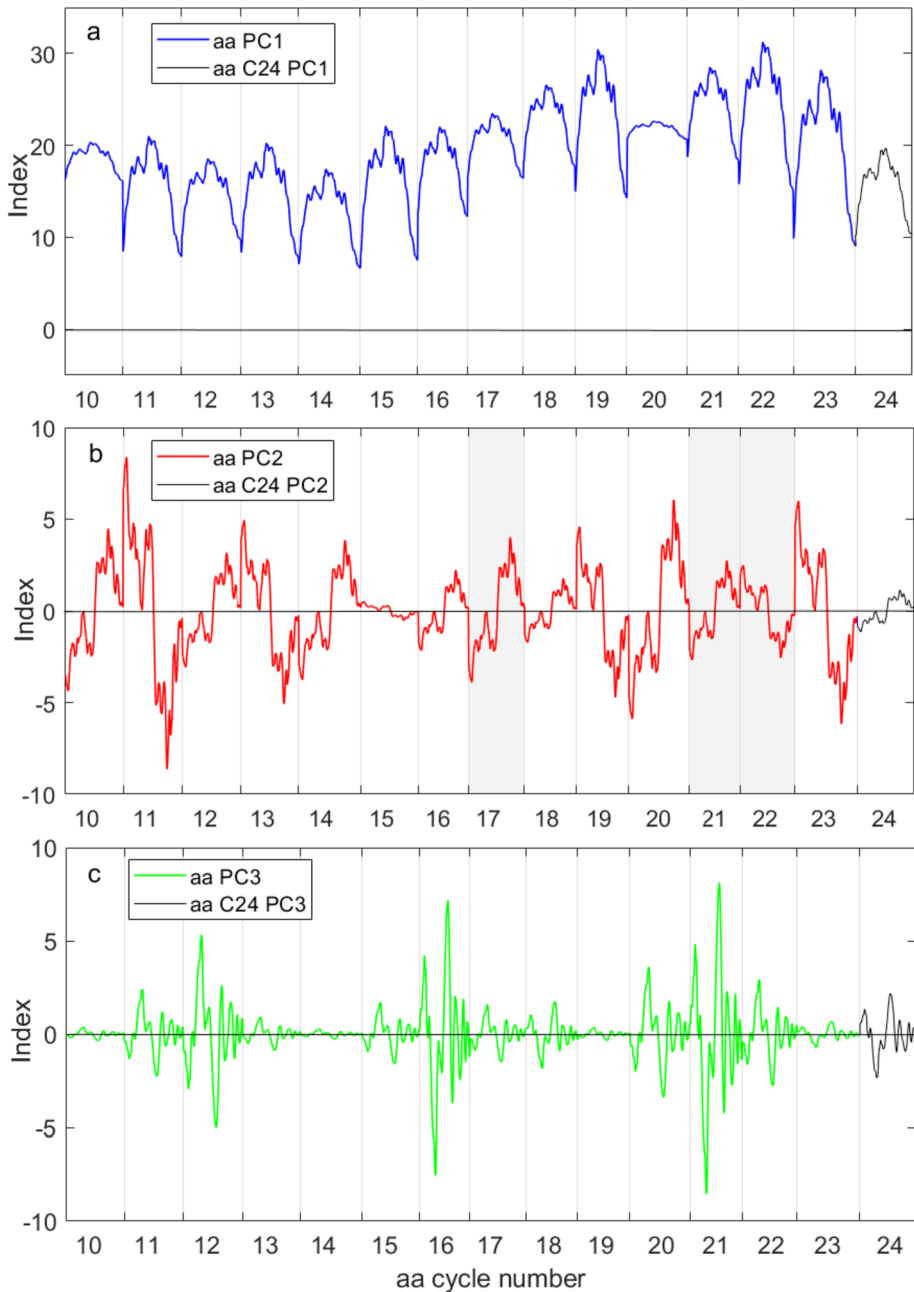


Figure 5 a) PC1, b) PC2, and c) PC3 time series of the PCA of the aa-index for Cycles 10–23. Note that Cycle 24 PCs are calculated from the PCA of Cycles 10–24.

cycle background, which goes through the minima between odd and even aa cycles (dash magenta broken-line). When subtracting the Hale cycle background envelope from the aa-index (blue) we get a detrended aa-index (red line). It is evident that there exist succeeding

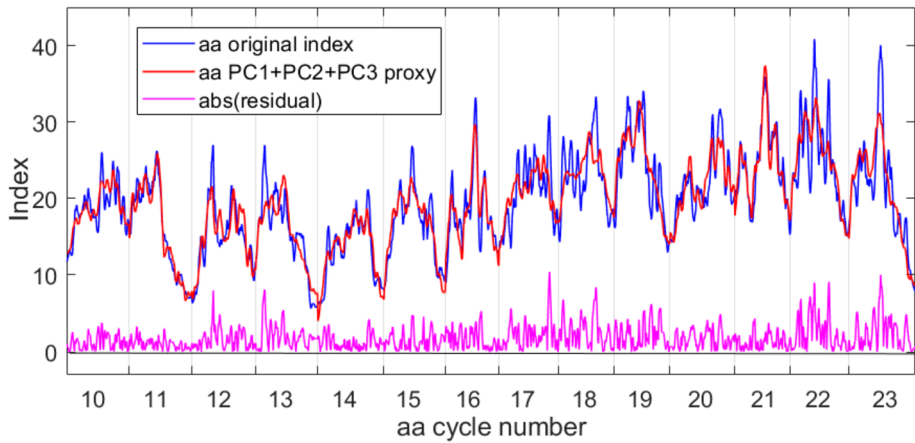


Figure 6 Original monthly aa time series (blue), its PC1 + PC2 + PC3 proxy time series (red), and the absolute value of their difference (residual) time series (magenta) for Cycles 10–23.

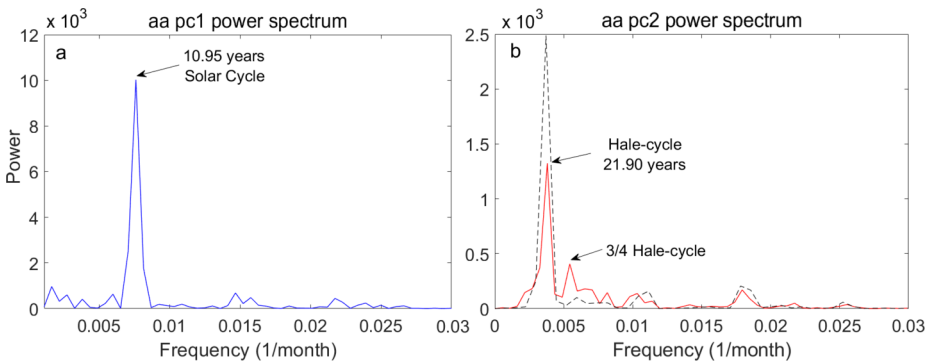


Figure 7 The power spectra of the **a)** PC1 and **b)** PC2 time series of Figure 5.

even–odd cycle pairs. The minima between the even–odd pairs are always higher than the minima between succeeding odd–even pairs. Furthermore, when the long-term trend is removed from the aa-index, the autocorrelation function (ACF) shows better the mutual intensity of the succeeding cycles. The inset in Figure 8 shows the ACF of the modified aa-index (magenta line, when the Hale cycle background is subtracted). Note that every second peak is higher in the ACF showing the Hale cycle in the aa-index.

3.3. PCA of the SSN for Cycles 10–23

In this section we study the Solar Cycles 10–23 because we compare the principal components of sunspot numbers with the principal components of the aa-index, which exists only since 1844. In Figure 9 we show the first two main principal components (PC1 and PC2) of the SSN for SC10–SC23 and the corresponding time series PC1 and PC2 proxy in Figure 10. In this study we are not interested in the SSN PC1 time series proxy. It is clear that the PC1s of the aa-index and SSN are related with each other. More interesting is the relation of the PC2 time series of the aa-index and the SSN. The relevance of the PC2

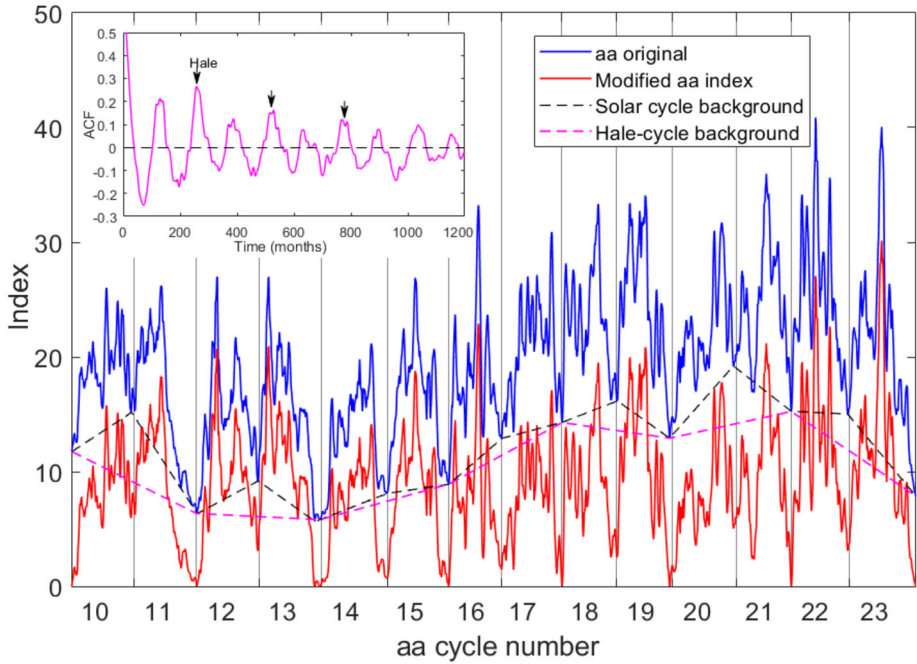
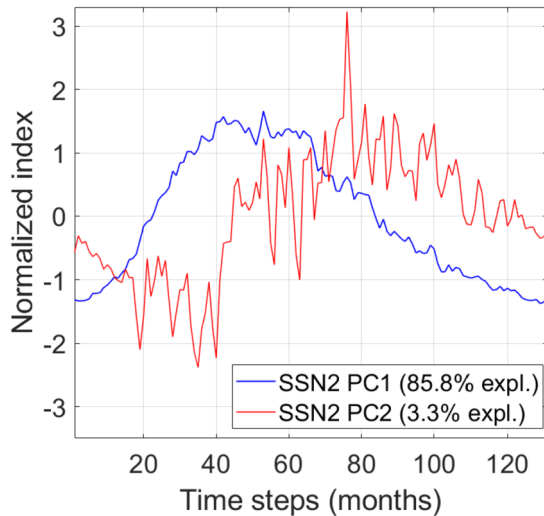


Figure 8 The original aa-index (blue) and the detrended aa-index (red) of Cycles 10–23. The inset shows the ACF of the detrended aa-index.

Figure 9 The PC1 and PC2 of SSN for the Solar Cycles 10–23. SSN2 in this figure and in Figures 10 and 11 refer to version 2 of the SSN.



is to correct the shape of the cycle when the corresponding cycle differs from the average cycle shape. The main effect of the PC2 is to reduce (negative phase for PC2 in the second half of the cycle) or enhance (positive phase for PC2 in the second half) the activity level of the descending phase with respect to the ascending phase of the cycle (Takalo and Mursula, 2018). Although the PC2 of SSN explains only 3.3% of the variance of the data

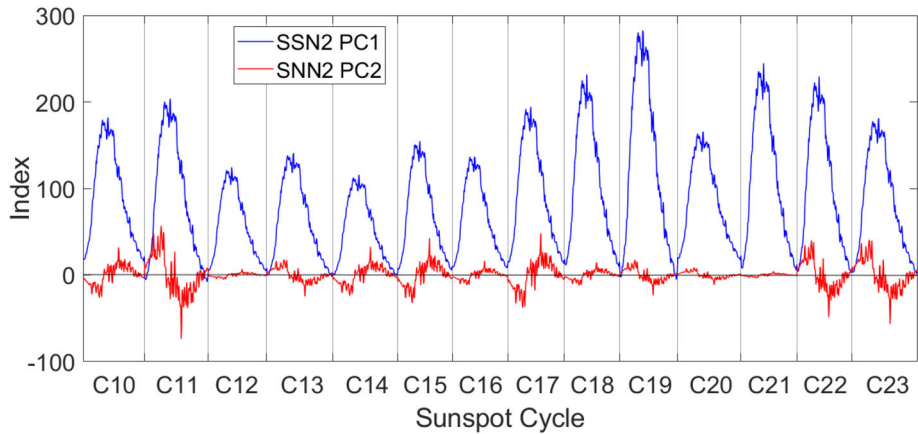


Figure 10 The PC1 (blue) and PC2 (red) time series of SSN for the Solar Cycles 10–23.

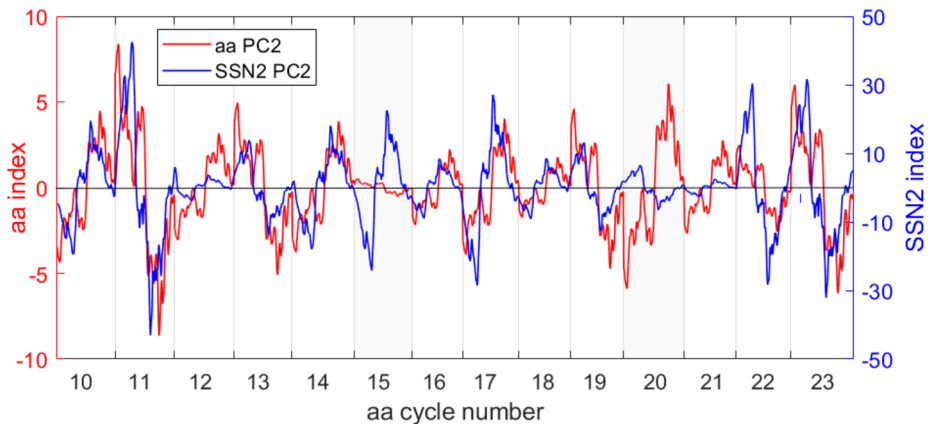


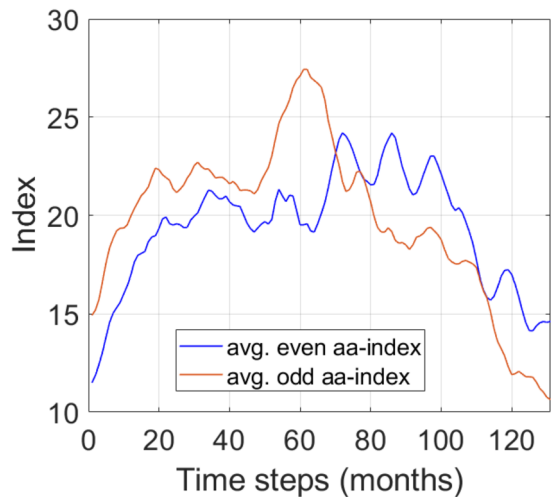
Figure 11 The PC2 time series of the aa-index (red) and SSN (blue) for the Solar Cycles 10–23.

its significance is more important for some cycles. Figure 11 shows the aa and SSN PC2 time series (seven-month smoothed) together. Note that the vertical lines are minima of the aa-index not of SSN. The essential feature here is the mutual phase of the aa and SSN PC2 time series. We notice that the phases are the same for all cycles other than SC15 and SC20. From these cycles, the aa PC2 is rather small for SC15 and the SSN PC2 is quite small for SC20. Furthermore, Cycle 20 is known to be anomalous compared to other aa cycles. The correlation coefficient (CC) between the aa and SSN PC2 time series is best when SSN PC2 lags by about two years ($CC = 0.587$ with $p < 10^{-100}$).

4. Discussion

The coherence in phase in Figure 11 shows that some part of the shape of the aa cycle PC2 is due to the sunspot-number PC2 component. However, most of the aa PC2 component must be caused by other recurrent phenomena in the Sun, and consequently in the solar wind.

Figure 12 The average shapes of the even (blue) and odd (red) aa cycles C10–C23.



We calculate that the average of the second half (descending phase) of the PC2 component of the even cycles is 45.1% compared to the total average of second half for the aa-index, but only 13.8% for the SSN. We thus consider that approximately about 30% of the shape of aa PC2 is due to the sunspot number and the rest for other recurrent processes. The first maximum of aa, during sunspot maximum, has been shown to be related to coronal mass ejections (CMEs), while the second maximum, in the descending phase of the SC, is probably related to high-speed streams (HSSs) (Gosling, Asbridge, and Bame, 1977; Simon and Legrand, 1986, 1989; Cliver, Boriakoff, and Bounar, 1996; Echer et al., 2004). For example, according to Richardson, Cliver, and Cane (2000) the CME level was unusually low during 1972, i.e. near Solar Cycle 20 maximum. On the other hand, Gosling, Asbridge, and Bame (1977) found that HSSs and streams with velocity in excess of 700 km/sec were common in 1973–1975, i.e. during the descending phase of SC20. These facts are probably the cause of the anomalous shape the Cycle 20 aa-index during the maximum and descending phase of the corresponding solar cycle. A counterexample is the strong CME on May 1921 during the descending phase of Solar Cycle 15, which caused very high values in the magnetometers on Earth (Kappenmann, 2006; Hapgood, 2019). This event was, however, so short that it does not appear in the PC2 of the odd Cycle 15. Note, however, that aa Cycle 15 has the largest weight of the PC1 component (see Figure 4).

Figure 12 shows the average shapes of the even and odd aa cycles during C10–C23. The even cycles have maxima between 70–95 months after the start of the aa cycle (related to HSS), while the odd cycles have a single maximum at about 60 months after the start of the cycle (related to CMEs). Note also that the first half of an average even cycle is lower than the first half of an average odd cycle. Takalo (2021) has shown that similar average profiles exist in the even and odd cycles of the Ap-index.

5. Conclusion

We have decomposed monthly aa-index of Cycles 10–23 using principal component analysis (PCA). Because of the noise of the data, we use seven-month trapezoidal-smoothed aa-index. We show that the first component (PC1) is related to the solar sunspot cycle, and

accounts for 41.5% of the variance of the data. The second component (PC2) is related to the 22-year Hale cycle, and explains 23.6% of the variance of the data. The PC3 and higher PCs show shorter periods due to only some individual cycles. PC3 explains, however, still 10.6%, but higher components under 6% of the variance. The PC1 time series of the Cycles 10–23 aa-index has only one peak in its power spectrum at the period 10.95 years, which is the average solar cycle period for the interval SC10–SC23. The PC2 time series of the same cycles has a clear peak at period 21.90 (Hale cycle) and a smaller peak at 3/4 of that period. If we remove the cycle pairs 16–17 and 21–22 in order to get time series with all succeeding cycles having opposite phases, the 3/4 Hale period disappears, and we get clearer peak at the period 21.90 (note that in this case the last cycle pair is 20–23). We also show that the first aa Cycle 24 of the ongoing Hale cycle has at least started with a common even cycle phase, i.e. negative phase in the first half of the cycle. We have also studied the principal components of solar sunspot numbers (SSN) for Solar Cycles 10–23, and we have compared the mutual behavior of the PC2 components of aa-index and SSN PCA analyses. We note that they are in the same phase in all other cycles than Solar Cycle 15 and 20. This shows that at least some part of the different shapes of the even and odd aa cycles is due to the solar sunspot cycle. The aa Cycle 20 also differs from the other even aa cycles in its shape, especially in anomalously high peaks during its descending phase. It has also, by far, the smallest weight on the aa PC1. On the other hand aa Cycle 15 has largest weight on the aa PC1 component. These reasons may be the cause of the phase difference of aa Cycles 15 and 20 compared to Solar Cycles 15 and 20, respectively. Even though there is a clear coherence in the phases of the PC2 time series of the aa-index and the sunspot number, we estimate that the sunspot-number variation explains only 30% the shape of PC2, and the rest is caused by other recurrent events in the Sun/solar wind. Note that the PC2 of sunspot numbers accounts only for 3.3% of the variance of the corresponding data. The first maximum of aa (typical to odd cycles), during sunspot maximum, has been shown to be related to CMEs, while the second maximum (typical to even cycles) in the descending phase of the SC, is probably related to HSSs (Gosling, Asbridge, and Bame, 1977; Simon and Legrand, 1986, 1989; Cliver, Boriakoff, and Bounar, 1996; Echer et al., 2004). The HSSs raise the activity level such that the minimum between even–odd cycle pairs is always higher than the minimum between succeeding odd–even cycle pairs. This can be seen also in the ACF of the detrended aa-index such that the Hale cycle related maximum is higher than the solar cycle related maximum in its ACF.

Acknowledgements The aa-index 1968–2010 is downloaded from ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/AASTAR/ and the aa-index for Cycle C24 from ISGI (isgi.unistra.fr). We are grateful to the Finnish Meteorological Institute/Nevalinna for the extended aa-data from 1844 onwards. The dates of cycle minima were obtained from the National Geophysical Data Center, Boulder, Colorado, USA (<ftp.ngdc.noaa.gov>). The daily sunspot-number data set were downloaded from the SILSO, World Data Center-Sunspot Number and Long-term Solar Observations, Royal Observatory of Belgium, online Sunspot catalog (<http://www.sidc.be/silso/datafiles>).

Funding Note Open access funding provided by University of Oulu including Oulu University Hospital.

Declarations

Disclosure of Potential Conflicts of Interest The author declares that there are no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons

licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Apostolov, E., Altadill, D., Todorova, M.: 2004, The 22-year cycle in the geomagnetic 27-day recurrences reflecting on the F2-layer ionization. *Ann. Geophys.* **22**, 1171. [DOI](#). [ADS](#).
- Bartels, J.: 1932, Terrestrial-magnetic activity and its relations to solar phenomena. *J. Geophys. Res.* **37**, 1. [DOI](#). [ADS](#).
- Bhattacharyya, A., Okpala, K.C.: 2015, Principal components of quiet time temporal variability of equatorial and low-latitude geomagnetic fields. *J. Geophys. Res.* **120**, 8799. [DOI](#).
- Bohlin, J.D.: 1977, Extreme-ultraviolet observations of coronal holes. I: locations, sizes and evolution of coronal holes, June 1973 – January 1974. *Solar Phys.* **51**, 377. [DOI](#). [ADS](#).
- Bro, R., Smilde, A.K.: 2014, Principal component analysis. *Anal. Methods* **6**, 2812.
- Chernosky, E.J.: 1966, Double sunspot-cycle variation in terrestrial magnetic activity 1884–1963. *J. Geophys. Res.* **71**, 965. [DOI](#). [ADS](#).
- Clette, F., Svalgaard, L., Vaquero, J.M., Cliver, E.W.: 2014, Revisiting the sunspot number. A 400-year perspective on the solar cycle. *Space Sci. Rev.* **186**, 35. [DOI](#). [ADS](#).
- Cliver, E.W., Boriakoff, V., Bounar, K.H.: 1996, The 22-year cycle of geomagnetic and solar wind activity. *J. Geophys. Res.* **101**, 27091. [DOI](#). [ADS](#).
- Cliver, E.W., Kamide, Y., Ling, A.G.: 2000, Mountains versus valleys: semiannual variation of geomagnetic activity. *J. Geophys. Res.* **105**(A2), 2413. [DOI](#).
- Cliver, E.W., Kamide, Y., Ling, A.G.: 2002, The semiannual variation of geomagnetic activity: phases and profiles for 130 years of aa data. *J. Atmos. Solar-Terr. Phys.* **64**(1), 47. [DOI](#).
- Cortie, A.L.: 1912, Sunspots and terrestrial magnetic phenomena, 1898–1911: the cause of the annual variation in magnetic disturbances. *Mon. Not. Roy. Astron. Soc.* **73**, 52.
- Du, Z.L.: 2011, The correlation between solar and geomagnetic activity – Part 1: two-term decomposition of geomagnetic activity. *Ann. Geophys.* **29**, 1331. [DOI](#). [ADS](#).
- Echer, E., Gonzalez, W.D., Gonzalez, A.L.C., Prestes, A., Vieira, L.E.A., dal Lago, A., Guarneri, F.L., Schuch, N.J.: 2004, Long-term correlation between solar and geomagnetic activity. *J. Atmos. Solar-Terr. Phys.* **66**, 1019. [DOI](#). [ADS](#).
- Feynman, J.: 1982, Geomagnetic and solar wind cycles, 1900–1975. *J. Geophys. Res.* **87**, 6153. [DOI](#). [ADS](#).
- Gnevyshev, M.N., Ohl, A.I.: 1948, On the 22-year cycle of solar activity. *Astron. Zh.* **25**, 18.
- Gosling, J.T., Asbridge, J.R., Bame, S.J.: 1977, An unusual aspect of solar wind speed variations during solar cycle 20. *J. Geophys. Res.* **82**, 3311. [DOI](#). [ADS](#).
- Hannachi, A., Jolliffe, I.T., Stephenson, D.B.: 2007, Empirical orthogonal functions and related techniques in atmospheric science: a review. *Int. J. Climatol.* **27**, 1119. [DOI](#).
- Hapgood, M.A.: 2019, The great storm of May 1921: an exemplar of a dangerous space weather event. *Space Weather* **17**, 950.
- Hathaway, D.H., Wilson, R.M.: 2006, Geomagnetic activity indicates large amplitude for sunspot cycle 24. *Geophys. Res. Lett.* **33**, L18101. [DOI](#). [ADS](#).
- Holappa, L., Mursula, K., Asikainen, T.: 2014, A new method to estimate annual solar wind parameters and contributions of different solar wind structures to geomagnetic activity. *J. Geophys. Res.* **119**, 9407. [DOI](#). [ADS](#).
- Holappa, L., Mursula, K., Asikainen, T., Richardson, I.G.: 2014, Annual fractions of high-speed streams from principal component analysis of local geomagnetic activity. *J. Geophys. Res.* **119**, 4544. [DOI](#). [ADS](#).
- Kane, R.P.: 2007, A preliminary estimate of the size of the coming Solar Cycle 24, based on Ohl's precursor method. *Solar Phys.* **243**, 205. [DOI](#).
- Kappenmann, J.G.: 2006, Great geomagnetic storms and extreme impulsive geomagnetic field disturbance events – an analysis of observational evidence including the great storm of May 1921. *Adv. Space Res.* **38**, 188. [DOI](#).
- Kumar, D., Rai, C.S., Kumar, S.: 2008, Principal component analysis for data compression and face recognition. *J. Comput. Sci.* **7**, 48.
- Lin, J.-W.: 2012, Ionospheric total electron content seismo-perturbation after Japan's March 11, 2011, $M = 9.0$ Tohoku earthquake under a geomagnetic storm; a nonlinear principal component analysis. *Astrophys. Space Sci.* **341**, 251. [DOI](#).

- Lockwood, M., Nevanlinna, H., Owens, M.J., Harrison, R.G., Rouillard, A.P., Davis, C.J.: 2013, Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr – Part 1: a new geomagnetic data composite. *Ann. Geophys.* **31**, 1957. [DOI](#).
- Makarov, V.I.: 1994, Global magnetic activity in 22-year solar cycles. *Solar Phys.* **150**, 359.
- Mayaud, P.-N.: 1972, The aa indices: a 100-year series characterizing the magnetic activity. *J. Geophys. Res.* **77**, 6870. [DOI](#). [ADS](#).
- McIntosh, D.H.: 1959, On the annual variation of magnetic disturbance. *Phil. Trans. Roy. Soc. London Ser. A* **251**, 525. [DOI](#). [ADS](#).
- Nevanlinna, H., Kataja, E.: 1993, An extension of the geomagnetic activity index series aa for two solar cycles (1844–1868). *Geophys. Res. Lett.* **20**, 2703. [DOI](#). [ADS](#).
- Richardson, I.G., Cliver, E., Cane, H.V.: 2000, Sources of geomagnetic activity over the solar cycle: relative importance of coronal mass ejections, high-speed streams, and slow solar wind. *J. Geophys. Res.* **105**, 18203. [DOI](#).
- Russell, C.T., McPherron, R.L.: 1973, Semiannual variation of geomagnetic activity. *J. Geophys. Res.* **78**, 92. [DOI](#). [ADS](#).
- Shnirman, M.G., Le Mouél, J.-L., Blanter, E.M.: 2009, The 27-day and 22-year cycles in solar and geomagnetic activity. *Solar Phys.* **258**, 167. [DOI](#). [ADS](#).
- Simon, P.A., Legrand, J.-P.: 1986, Some solar cycle phenomena related to the geomagnetic activity from 1868 to 1980. II – high velocity wind streams and cyclical behavior of poloidal field. *Astron. Astrophys.* **155**, 227. [ADS](#).
- Simon, P.A., Legrand, J.P.: 1989, Solar cycle and geomagnetic activity: a review for geophysicists. Part 2. The solar sources of geomagnetic activity and their links with sunspot cycle activity. *Ann. Geophys.* **7**, 579. [ADS](#).
- Svalgaard, L.: 1977, Geomagnetic activity: dependence on solar wind parameters. In: Zirker, J.B. (ed.) *Coronal Holes and High Speed Wind Streams*, 371. [ADS](#).
- Takalo, J.: 2021, Comparison of geomagnetic indices during even and odd solar cycles SC17–SC24: signatures of Gnevyshev gap in geomagnetic activity. *Solar Phys.* **296**, 19. [DOI](#).
- Takalo, J., Lohikoski, R., Timonen, J.: 1995, Structure function as a tool in AE and Dst time series analysis. *Geophys. Res. Lett.* **22**, 635. [DOI](#). [ADS](#).
- Takalo, J., Mursula, K.: 2018, Principal component analysis of sunspot cycle shape. *Astron. Astrophys.* [DOI](#).
- Wilson, R.M.: 1988, Bimodality and the Hale cycle. *Solar Phys.* **117**, 269.
- Zharkova, V.V., Shepherd, S.J., Zharkov, S.I.: 2012, Principal component analysis of background and sunspot magnetic field variations during solar cycles 21–23. *Mon. Not. Roy. Astron. Soc.* **424**, 2943. [DOI](#).
- Zharkova, V., Popova, H., Zharkov, S., Shepherd, S.: 2016, Two principal components of solar magnetic field variations and prediction of solar activity on multi-millennium timescale. In: *41st COSPAR Scientific Assembly*, abs. id. E2.1-18-16.