Variation in Cosmic-Ray Intensity Lags Sunspot Number: Implications of Late Opening of Solar Magnetic Field

Yuming Wang[1,2,3] , Jingnan Guo[1,2] , Gang Li[4] , Elias Roussos[5] , and Junwei Zhao[6]

[1] CAS Key Laboratory of Geospace Environment, School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, People's Republic of China; ymwang@ustc.edu.cn, jnguo@ustc.edu.cn

[2] CAS Center for Excellence in Comparative Planetology, University of Science and Technology of China, Hefei 230026, People's Republic of China

[3] Mengcheng National Geophysical Observatory, University of Science and Technology of China, Mengcheng 233500, People's Republic of China

[4] Department of Space Science and CSPAR, The University of Alabama in Huntsville, AL, USA

[5] Max Planck Institute for Solar System Research, Goettingen, Germany

[6] W.W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA, USA

Abstract

Galactic cosmic rays (GCRs), the highly energetic particles that may raise critical health issues for astronauts in space, are modulated by solar activity, with their intensity lagging behind the variation in sunspot number (SSN) by about one year. Previously, this lag has been attributed to the combined effect of outward convecting solar wind and inward propagating GCRs. However, the lag's amplitude and its solar-cycle dependence are still not fully understood. By investigating the solar surface magnetic field, we find that the source of heliospheric magnetic field —the open magnetic flux on the Sun—already lags behind SSN before it convects into the heliosphere along with the solar wind. The delay during odd cycles is longer than that during sequential even cycles. Thus, we propose that the GCR lag is primarily due to the very late opening of the solar magnetic field with respect to SSN, though solar wind convection and particle transport in the heliosphere also matter. We further investigate the origin of the open flux from different latitudes of the Sun and find that the total open flux is significantly contributed by that from low latitudes, where coronal mass ejections frequently occur and also show an odd–even cyclic pattern. Our findings challenge existing theories, and may serve as the physical basis of long-term forecasts of radiation dose estimates for manned deep-space exploration missions.

Unified Astronomy Thesaurus concepts: Cosmic rays (329); Solar magnetic fields (1503); Solar cycle (1487); Heliosphere (711)

1. Introduction

Galactic cosmic rays (GCRs), which are omnipresent, charged, and energetic particles coming from outside the heliosphere, are affected by the heliospheric magnetic flux as they propagate inward from the heliospheric boundary at about 120 au (Krimigis et al. 2013). Over several decades we have learned that GCR fluxes are constantly affected by variations in the heliospheric magnetic fields, on both short and long timescales. In the short term of days or months, the GCR flux can be altered in the form of Forbush decreases (as first reported by Forbush 1937) due to transient heliospheric structures with more turbulent and intense magnetic fields, such as interplanetary coronal mass ejections (ICMEs, Cane 2000) and stream interaction regions (Richardson 2004). As GCRs can interact with Earth's atmosphere via ionization processes, such disturbed GCR variations have also been argued to be the link in Sun–climate correlations (Pittock 1978) by changing the global electric circuit and modifying cloud properties (Harrison et al. 2011; Laken et al. 2012; Laken & Čalogović 2013). In the long term of a few years, the GCR flux was first observed to anticorrelate with sunspot variations (Forbush 1958) since the transport of GCRs is modulated by heliospheric field strength and irregularities that evolve following the quasi-11 yr solar cycle (Parker 1965; Potgieter 1998). Specifically, enhanced magnetic flux is more efficient in

preventing charged GCR particles from penetrating deeply into the heliosphere, causing a decrease in GCR fluxes toward solar maxima. The variation in GCR fluxes at Earth has been correlated with various solar and heliospheric parameters, such as the sunspot number (SSN), the strength and turbulence level of heliospheric magnetic field (HMF), the tilt angle of the heliospheric current sheet (HCS), the open solar magnetic flux, the solar polarity, and so on (Usoskin et al. 1998; Cliver & Ling 2001; Rouillard & Lockwood 2004; Alanko-Huotari et al. 2007; Potgieter 2013, etc.), and empirical functions describing the GCR dependence on different solar cycle parameters have been proposed (e.g., Dorman 2001; Usoskin et al. 2011; Guo et al. 2015).

In particular, when correlating the temporal variations in GCRs and SSN, the strongest anticorrelation appears when the GCR profile is shifted backward in time, suggesting a delay in the GCR variation with respect to the evolution of the solar activity. The classic picture to explain this time lag involves the solar wind convection and the GCR transport in the heliosphere (Parker 1965; Van Allen 2000; Cliver & Ling 2001; Dorman 2001; Usoskin et al. 2001; Thomas et al. 2014). That is, GCRs propagate inward throughout the heliosphere and are affected by the magnetic field carried by the outward solar wind during their journey. Meanwhile, different paths of GCRs during different phases of the solar cycle could also influence the arrival times of GCRs.

It has also been observed that the modulation of GCRs has an odd–even cycle dependence (Webber & Lockwood 1988; Van Allen 2000; Cliver & Ling 2001; Thomas et al. 2014). This

implies that the GCR–SSN lag is much longer (one year or more) during odd solar cycles than that (no more than two months) during even solar cycles. This has been explained, following the above theory of the cause of the delay, by the different drift patterns and diffusion process of GCR propagation through the heliosphere when the solar magnetic poles have predominantly positive or negative spatial orientation following the 22 yr Hale cycle (Jokipii et al. 1995, 1977; Ferreira & Potgieter 2004). Specifically, the polarity of the solar field (often represented by the symbol A) is positive/negative when the dominant polar field is outward/inward in the northern hemisphere. The polar field reversal occurs around the time when SSN reaches a maximum and divides each solar cycle into A+ and A− half-cycles. Influenced by the curvature and gradient drift pattern, GCR protons arrive at the inner heliosphere after approaching the solar poles and moving out along the heliospheric current sheet during the A+ phase. In opposition, during the A− phase, GCR protons propagate inward along the HCS plane and leave via the poles. Around solar maximum, disturbances in the solar wind are much stronger, making it more difficult for particles to propagate inward along the HCS plane. Odd solar cycles start with A+ polarity and switch to A− so that GCRs experience a slower recovery during the declining phase of the cycle when they predominantly enter the heliosphere along the HCS. Even cycles (first A− and then A+) experience a faster GCR recovery.

However, the above GCR transport model alone can hardly predict the GCR–SSN lags at the precision of a few months that could be revealed from observations. A recent study (Ross & Chaplin 2019) has shown that during Cycle 24, GCR–SSN lag is about 4 months, which is slightly longer than the lags during preceding even-numbered cycles, which were 1–2 months, although not as long as those observed in previous odd-numbered cycles, which were longer than a year. To explain such exceptions, theoretical models would need to rely on various ad hoc and adjustable parameters whose values cannot be directly verified but are chosen to fit the modeling results with observations.

In this study, we focus on the solar open field (OF) that forms the large-scale HMF through which the GCRs propagate. Via an investigation of GCR–SSN correlation and the OF–SSN correlation throughout several solar cycles (Cycle 21–24) and at various heliospheric distances (from 1 to about 30 au), we find that the temporal evolution of GCRs closely mirrors the temporal variations in the open field, which is already delayed with respect to SSN. Thus, we propose that the GCR–SSN lag is primarily due to the late opening of the solar magnetic field at the Sun, while solar wind convection and GCR transport probably make a secondary contribution to cause a longer delay on top of the OF–SSN delay. Besides, we find that the OF–SSN delay during odd cycles is longer than that during sequential even cycles. Instead of the transport effect in the heliosphere, we suggest that the behavior of the open field on the Sun could be the main driver of the odd–even cycle dependence of the GCR–SSN delay. We further investigate the origin of the OF from different latitudes of the Sun and find that the total OF is largely contributed by that from low latitudes, where CMEs are frequently launched from active regions and also show an odd–even cycle dependence like the GCR–SSN delay. These findings advance our understanding of the solar cycle evolution and its impact on the heliosphere.

## 2. GCR Delays at Various Heliocentric Distances

To clearly show the GCR delay in the heliosphere, we investigate the energetic particle measurements at Earth, Mars, Rosetta swinging around 3 au, Cassini near Saturn, and New Horizons (NH) approaching Pluto during Cycles 23 and 24 (see panels 3–7 in Figure 1(a) for these five GCR data sets). The description of the GCR data is given briefly in the caption of Figure 1 and in more detail in Appendix A. We also use the daily SSNs (Clette et al. 2014), which are directly obtained from the Solar Influences Data and Analysis Centre (SIDC) of the Royal Observatory of Belgium (http://www.sidc.be/silso/datafiles; the latest version 2.0 is used), as shown in the first panel of Figure 1(a). The second panel shows the OF and will be explained later. All of these data are prepared in each Bartels rotation (BR, one BR is 27 days) by calculating the median value of the daily values of these data for each BR.

First, we use the peak-alignment method to estimate the GCR time delays with respect to SSN at various distances. In this method, we first locate a notable extremum (maximum or minimum) in SSN, and then set a reasonable time range, from −1 to 2.5 yr, to search for the corresponding extremum in the parameter of interest, e.g., the GCR flux at Earth. The time difference between the two peaks is the time delay. A positive value means a delay and a negative value a lead. Since short-term fluctuations due to temporary solar transients may influence the positions of the extrema, we smooth the data over a certain size and repeat the above procedures to obtain the time delay again. We change the smoothing window from 1 BR to 13 BRs, close to one year, and get 13 peak locations for each parameter as well as 13 time delays, of which the median value is adopted as the final time delay (the colored circles in Figure 1(a) or dots in Figure 1(b)) and the minimum and maximum values are used as the uncertainty (the associated bars in Figure 1).

We select four SSN extrema (two maxima and two minima, the color-coded circles in the top panel of Figure 1(a)) during Cycles 23 and 24 for the peak-alignment method. It should be noted that during the cycle maxima, SSN depicts a clear double-peak structure (see the solar maxima during Cycles 23 and 24), and the second peak of the SSN around a solar maximum is used in this study to make sure that we have as many extremum values as possible across different data sets that cover various durations (other panels of Figure 1(a)).

The estimated GCR delays are summarized in Figure 1(b), in which the dots with the error bars give the time delays for the extrema identified in Figure 1(a) with the same colors. It can be found in this panel that the GCR delay is generally longer during solar maxima (marked by purple and red dots) than minima (marked by green and blue dots), and could be as long as more than 22 solar BRs or 1.6 yr (see the purple dots beyond 1 au). According to existing particle transport theories (e.g., Potgieter 2013), such delays are thought to be mainly due to the outward convection of solar cycle conditions of the HMF with the solar wind and the inward diffusive propagation of GCRs. Here, we look at this issue from another point of view. We go to the source of the HMF, i.e., the open magnetic field on the Sun, to see whether the delay with respect to SSN already happens before the outward convection of the solar wind, which may provide an alternative explanation.

Since there is no technique to directly observe solar open field, we use the CSSS model (Zhao & Hoeksema 1995) to extrapolate and locate the open magnetic fields based on the line-of-sight (LOS) photospheric magnetic fields observed by
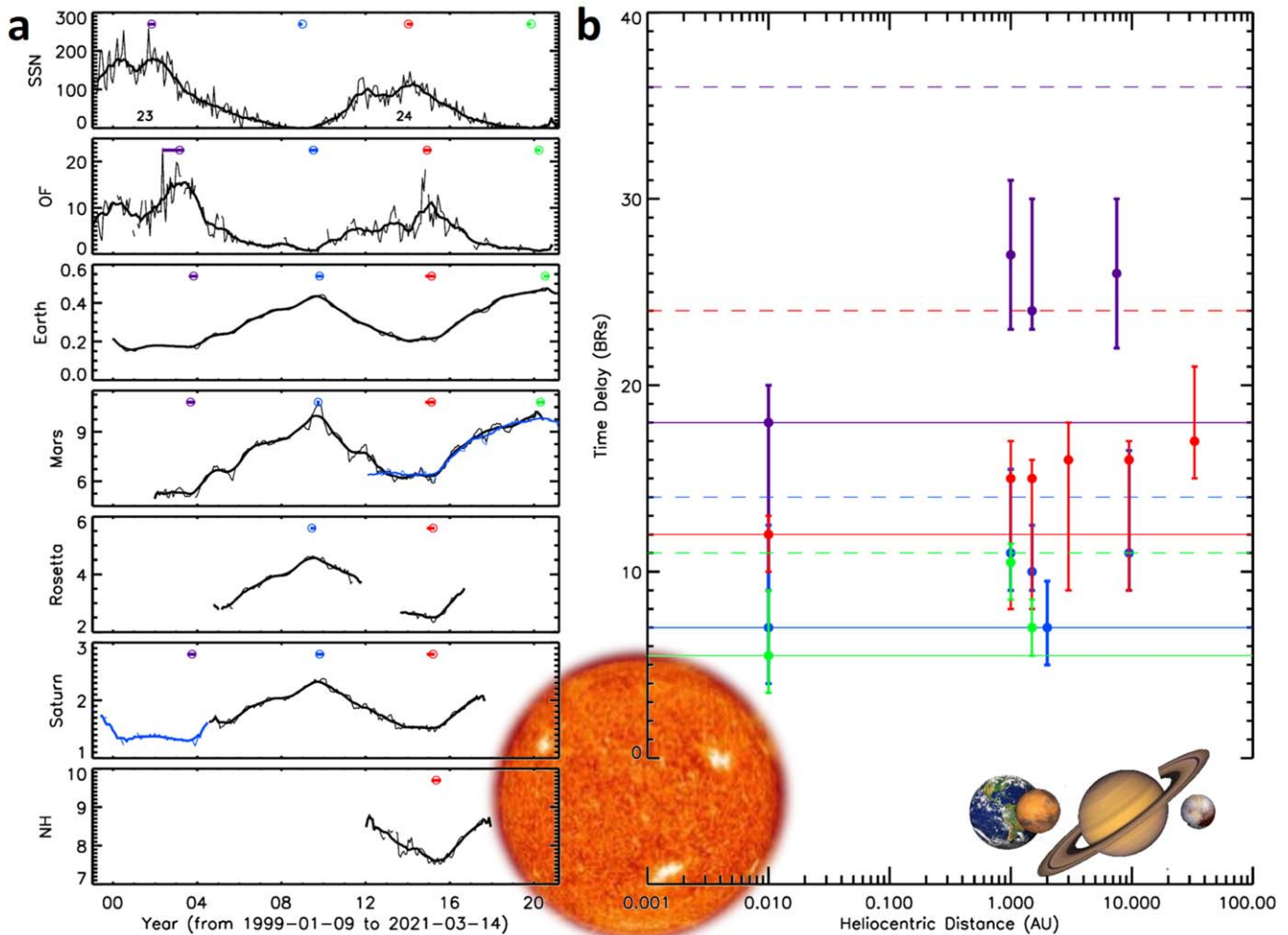
Figure 1. Time delays of solar open flux and GCR fluxes at five different heliocentric distances with respect to the SSN variation for solar Cycles 23 and 24. (a) From top to bottom, the panels show the solar cycle variations of the SSN (Clette et al. 2014, version 2.0), the open magnetic flux (OF) on the Sun (in units of 10^{13} Wb, derived from the current-sheet source–surface (CSSS) model (Zhao & Hoeksema 1995) based on the WSO magnetic synoptic charts (Duvall et al. 1977), see the main text for details), the GCR flux at Earth (measured by the EPHIN (Müller-Mellin et al. 1995) on board the Solar and Heliospheric Observatory (SOHO, Domingo et al. 1995) of protons above 53 MeV), the GCR flux at Mars (∼1.5 au, measured by HEND (Boynton et al. 2004) on board Odyssey orbiter of the Martian albedo neutrons with energies 10 eV—1 MeV in black and by RAD (Hassler et al. 2012, 2014) on board the Curiosity rover of Mars's surface dose rates induced by original protons above ∼160 MeV in blue (Guo et al. 2018)), the GCR flux within 4.5 au (measured by SREM (Evans et al. 2008) on board the Rosetta spacecraft of ions above 49 MeV), the GCR flux near Saturn at about 9.5 au (measured by Cassini spacecraft of ions above 120 MeV (Roussos et al. 2020); cruise data are shown in blue until Cassini arrived at Saturn in 2004 July), and the GCR flux from about 22 au to 40 au measured by New Horizons (NH, of 120 MeV–1.4 GeV ions (Hill et al. 2020)). The GCR flux recorded by New Horizons is in units of # keV^{-1} cm^{-2} sr^{-1} s^{-1}; the GCR data of RAD are scaled by 32 μGy/day, and the GCR fluxes from other instruments are all in units of counts per second. The temporal resolution of all the data is one Bartels rotation (BR) as shown by the thin curves. The thick curves are obtained after the running smooth using a 13 BR window. The purple, blue, red, and green circles in the first panel mark the four selected extrema (either maximum or minimum) of the SSN, and the circles in the other panels mark the positions of the corresponding extrema of the data shown. The horizontal bars across the circles give the uncertainties. (b) The dots give the time delays for the extrema identified in the left column (a) with the same colors, and the error bars are the uncertainties. The solar distance of the open flux is set at 0.01 au, which is close the cusp surface in the CSSS model where the solar magnetic fields open into the heliosphere. The horizontal solid lines mark the open flux delay times of the four extrema, and the dashed lines mark twice the corresponding delay times.

Wilcox Solar Observatory (WSO, Duvall et al. 1977, http://wso.stanford.edu/synopticl.html). The WSO synoptic charts are stored in Carrington rotations. To obtain the synoptic charts in Bartels rotations, we shift the starting time of the charts and reassemble them. Here, we use WSO data rather than the data from the Michelson Doppler Imager (MDI, Scherrer et al. 1995) on board the SOHO spacecraft and the Helioseismic and Magnetic Imager (HMI, Hoeksema et al. 2014) on board the Solar Dynamics Observatory (SDO, Pesnell et al. 2012) because WSO data cover the past four solar cycles, much longer than the MDI or HMI data. The difference in the extrapolation results between them and the influence on our conclusion can be found in Appendix B.

The CSSS model is a widely-used model for extrapolation of the coronal magnetic field (e.g., Wang & Zhang 2007; Gui et al. 2011), and compared to the potential field source surface (PFSS) model, it includes the effects of the large-scale horizontal current sheet in the inner corona, the warped heliospheric current sheet in the upper corona, and the volume currents in the outer corona, and can predict the strength and polarity of HMF better (Zhao & Hoeksema 1995). Using this model, we calculate the spherical harmonic coefficients up to nine orders, and trace magnetic field lines from the solar surface. Field lines returning to the surface are closed field lines, while those crossing over the cusp surface at 2.5 solar radii are open field lines. Since spatial resolution of the input
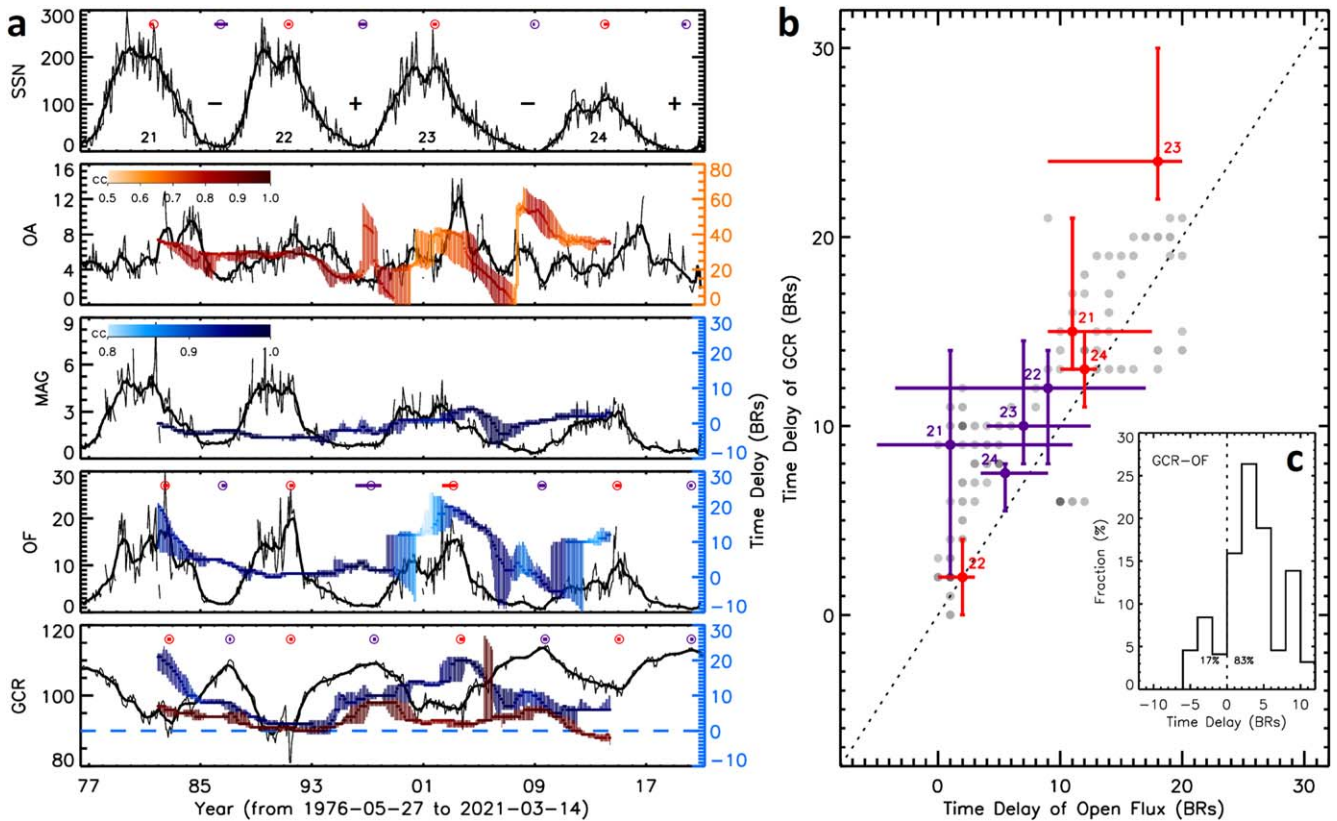
Figure 2. Time delays of the open magnetic field area, open field strength, open flux on the Sun and GCR fluxes at Earth with respect to the SSN evolution for the previous four solar cycles. (a) From top to bottom, the panels show the solar cycle variations of the SSN, open area (OA, as a percentage of the total solar surface), open field strength (MAG, in units of gauss), and open flux (OF) on the Sun and the Oulu monitor count rate at Earth (GCR, counts s^{-1}), respectively, from 1976 May to 2021 March. As in Figure 1(a), the circles and crossing bars are the selected extrema and their uncertainties. The red ones are for the second major maximum of each cycle and the purple ones for the solar minima. The color-coded ribbon-like curves in the second through the fifth panels show the solar cycle variations (from 1981 December 12 to 2014 May 26) of the time delays of the shown parameter relative to the SSN by using cross-correlation analysis on the 13 BR smoothed data. The last panel has two ribbons, with the blue one showing the GCR–SSN delay and the red one for the GCR–OF delay. These ribbons are composed of dots (the median values of the time delays) and error bars (the 20th/80th percentiles of the time delays), scaled by the right y-axis of each panel. The colors of the ribbons represent the correlation coefficients, scaled by the color bars in the second and third panels. (b) The red and purple dots with error bars mark the time delays of the open flux and GCRs for the eight extrema. The solar cycle numbers are also indicated. The gray dots show all the time delays obtained from the cross-correlation analysis. Multiple data points with the same time delays result in a darker dot. The dotted line is the diagonal line. (c) The histogram shows the distribution of the GCR–OF delays based on the cross-correlation analysis, i.e., the distribution of the red ribbon dots in the last panel in the left column (a). For about 83% of the data points, the GCR delay is longer than the open flux delay.

WSO chart is 72 (in longitude) by 30 (in latitude), we trace one field line from each grid point, resulting in a total of $72 \times 30 = 2160$ field lines. Meanwhile, we calculate the surface area of each grid point. The total area of all the grid points is the total solar surface area. With this information, we can easily derive the open flux, open area, and the averaged open magnetic field strength. It should be noted that there are sometimes missing data in WSO LOS synoptic charts. If the number of missing data points exceeds 60, i.e., about 3% of the total number of data points, in a BR chart, we simply omit the open field data corresponding to this BR chart.

We find that the total flux of the open fields shows clear solar cycle variations (the second panel of Figure 1(a)—OF) and looks well anticorrelated with the GCR fluxes. Using the previous peak-alignment method, we determine the time delay of the open flux after the SSN to be about 18, 8.5, 12, and 5.5 BRs for the four extrema, respectively, as summarized in Figure 1(b) (see the dots with error bars located at 0.01 au; the four horizontal solid lines crossing the dots are plotted for comparison with the time delay of GCRs in the heliosphere). The time delay of each GCR data set is longer than the delay of the corresponding open flux, except for the distance-varying Rosetta data point in blue. Considering that

the GCR delay may consist of two sources—one is the process in the heliosphere including heliospheric magnetic field convection by the solar wind and particle transport, and the other is at the very beginning on the Sun, i.e., the delay of the solar open flux—we can find in the figure that the latter may play an important role because none of the estimated GCR delays is longer than twice the corresponding open flux delay within the orbit of Pluto (indicated by the dashed lines in Figure 1(b), which correspond to twice the time of the open flux delay marked by the solid lines).

3. GCR Delays at Earth over the Past 45 Years

We further extend the analysis to the past 45 years from 1976 May to 2021 March (Cycles 21–24) as shown in Figure 2(a), during which the WSO observations of the solar photospheric magnetic fields are available but continuous GCR observations are only available at Earth from ground-based neutron monitors. We use the Oulu data (http://cosmicrays.oulu.fi) for this study. Previous studies show that the delay time can be energy-dependent with shorter delays for higher-energy GCRs (Shen et al. 2020), an effect that may be attributed to the energy-dependent transport of GCRs in the heliosphere (Moloto & Engelbrecht 2020). Compared to the SOHO/EPHIN data in
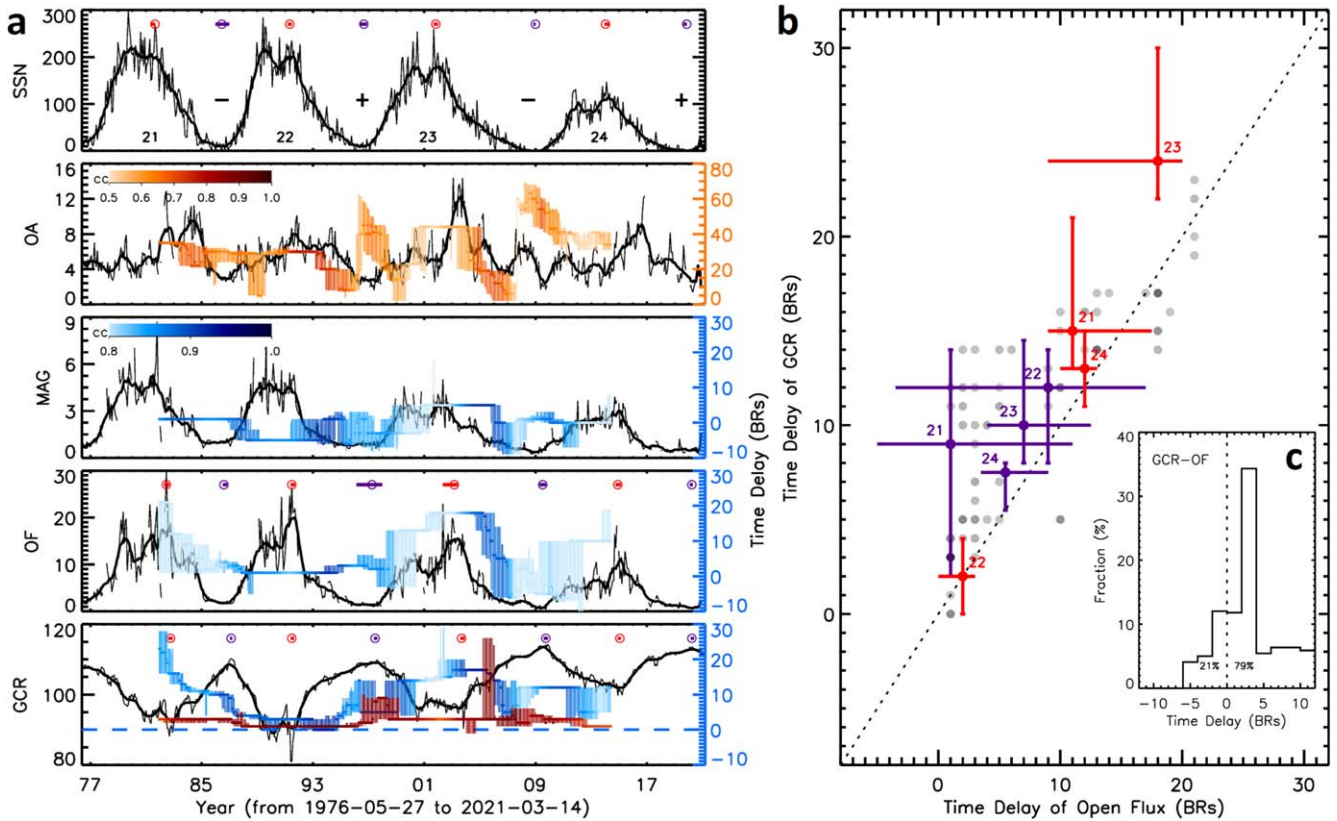
Figure 3. The same as Figure 2 except the data used for cross-correlation analysis are not smoothed.

Figure 1(a), the GCR delay time at Earth is about one to two BRs shorter if the Oulu data are used (more details in Appendix A and Figure A1).

Using the same peak-alignment method, we select the second major maximum (to maintain consistency with what we did for Cycles 23 and 24) of each cycle and every solar minimum, resulting in a total of eight extrema (the top panel of Figure 2(a)), to find the delay of each corresponding extremum in the solar open flux and the GCR flux (the last two panels of Figure 2(a)). The results displayed by colored dots with error bars and numbers in Figure 2(b) show that, for all the extrema, the GCR delays were roughly equal to (when the dots lie close to the diagonal line) or slightly longer than (when the dots lie above the diagonal line) the corresponding open flux delays and were much shorter than twice the open flux delays, suggesting that the primary role of the open flux in causing the GCR delay holds for the previous four solar cycles. We note that the purple point numbered 21 (the solar minimum between Cycles 21 and 22) is an exception, with the GCR delay much longer. However, the large uncertainties in both time delays make it difficult to discuss this outlier.

Benefiting from the continuous collection of GCR data at Earth throughout this period, we employ the cross-correlation method to reveal the cycle variation of the delay time. For any given time, we first select an N-year wide segment centered on this time in SSN, and then set the same wide time window on the parameter of interest, e.g., the open flux or the GCR flux, to calculate the correlation coefficient (cc) between the selected segments of the parameter and the SSN. The N-year window rolls through the parameter over a reasonable time range in order to locate the best correlation within the range. The time delay is the shifted time of the window away from its original position. Meanwhile the associated cc value is also obtained. Since the correlation depends

on the size of the time window, we change the value of N from 5 to 11 in steps of 0.074 (i.e., one BR), gathering the data of half of a solar cycle to a complete solar cycle, and repeat the above steps to achieve a total of 82 time delays and cc values, among which the time delays associated with a cc less than 0.5 are discarded. The median values of these delays and these cc values are chosen as the final time delay and the final cc value, respectively. The 20th/80th percentile is chosen as the uncertainty range of the time delay. Applying the above technique to the data throughout solar Cycles 21–24, we then get the cycle variations of the time delays as shown by the color-coded ribbons in Figure 2(a).

The smoothness of the data has little influence on the result of the cross-correlation analysis. The color-coded ribbons in Figure 2(a) are obtained based on the 13 BR smoothed data, and those in Figure 3(a) are based on the unsmoothed data. They look similar at the scale of a solar cycle. The percentage of the data points with the GCR delay longer than or equal to the open flux delay is 79% (Figure 3(c)), similar to that based on the smoothed data, which is 83% (Figure 2(c)). But we also notice that the smoothness of the data is probably an error source for a few gray dots in Figures 2(b) and 3(b), of which the GCR delays are shorter than the open flux delay. For example, comparing the red color-coded ribbons in the last panels of Figures 2(a) and 3(a), we find that during 2013–2014 the GCR flux seems to evolve ahead of the open flux in Figure 2(a), but behind it in Figure 3(a). This illustrates the uncertainty caused by the smoothness of the data sets.

We obtain the cycle evolution of the OF and GCR delays with respect to SSN as shown by the color-coded ribbons, which are scaled by the y-axes on the right, in the last two panels of Figure 2(a). In particular, the red color-coded ribbon in the last panel (GCR–OF delay) lying below the blue
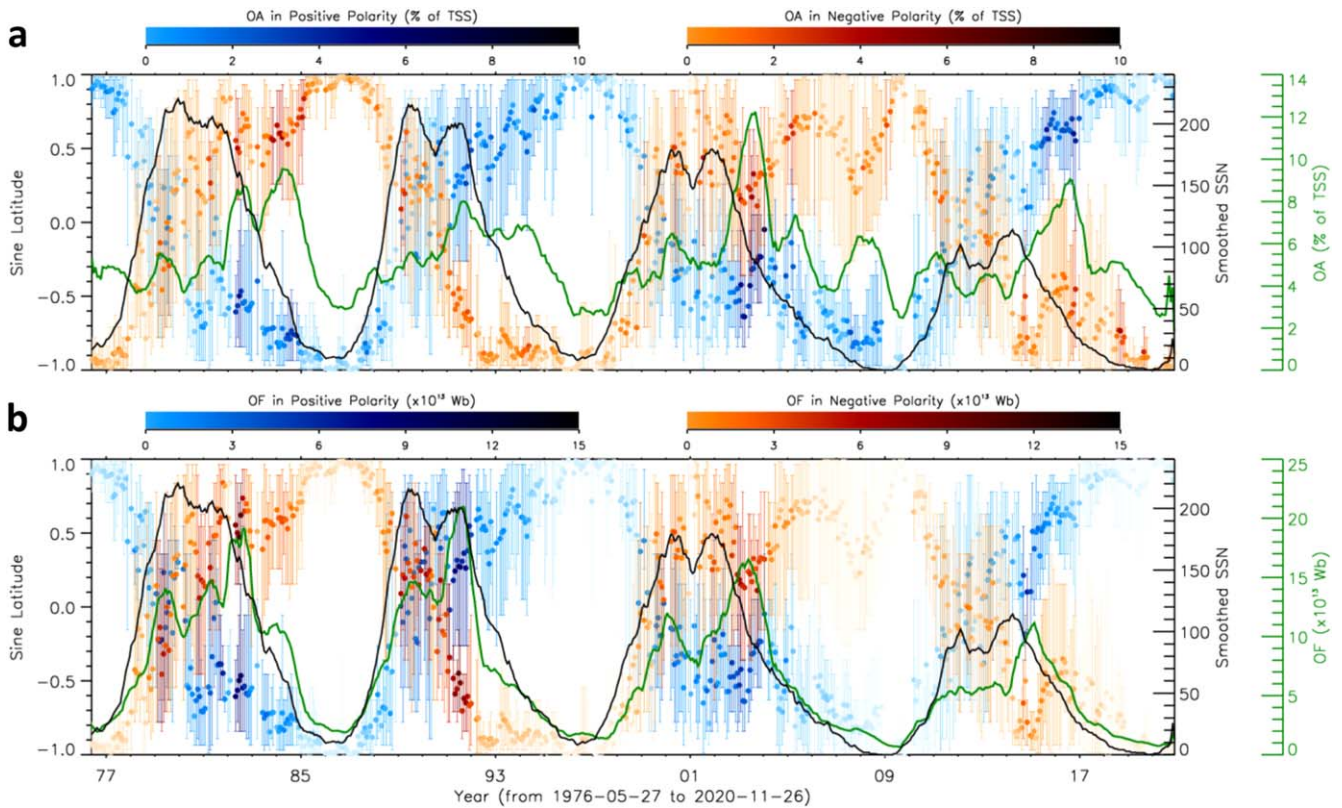
Figure 4. Cyclic diagrams of the solar open area and open flux. (a) The latitudinal and size evolution of the open area with the solar cycles. The blue/red dots indicate the latitudinal centroid of the positive/negative open field lines. The darker the color, the bigger the total area, which is scaled by the color bars on top of the panel. The error bars give the 20th/80th percentile of the open field lines in latitude. The smoothed SSN (black curve) and total open area (green curve) are overplotted. (b) Similar to (a), but for the open flux.

color-coded ribbon (GCR–SSN delay) clearly shows that the GCR delays with respect to SSN are longer than the GCR delays with respect to the open flux. The gray dots in Figure 2(b) represent the GCR–SSN delays (the blue ribbon in the last panel of Figure 2(a)) versus the OF–SSN delays (the blue ribbon in the second last panel of Figure 2(a)). Similar to the eight colored dots obtained using the peak-alignment method, the delay time based on the cross-correlation analysis is generally longer for GCR–SSN than for OF–SSN. The histogram in Figure 2(c) further shows the distribution of GCR–OF delay, i.e., dots forming the red ribbon in the last panel of Figure 2(a). It shows that about 83% of the GCR delays are longer than or equal to the open flux delays, and if we choose only those delays with the correlation coefficient $\geqslant 0.9$, the fraction increases to 93% (not shown in the figure), again supporting the idea that open flux delay is the main driver for the corresponding GCR lags.

Meanwhile, we find that the open flux delays during the odd cycles were notably longer than the delays during the sequential even cycles (see the blue color-coded ribbon in the second last panel of Figure 2(a) or Figure 3(a)). A similar odd–even pattern was observed in the GCR delays with respect to SSN (Van Allen 2000; Cliver & Ling 2001; Thomas et al. 2014), but does not appear in the GCR delays with respect to the open flux (see the ribbons in the last panel of Figure 2(a) or Figure 3(a)). Thus, the open flux delays could also be a driver of the observed odd–even cycle behavior in the GCR lags, which has been solely attributed to the different drift patterns of GCR propagation through the heliosphere following the 22 yr Hale cycles, i.e., GCRs drift in through the poles during an A+ cycle or along the

HCS close to the ecliptic plane during an A− cycle as previously reported (Jokipii et al. 1995; Ferreira & Potgieter 2004; Potgieter 1998, 2013). According to the analysis here, the drift effect due to the changes in solar polarity is perhaps only a secondary effect in driving the global variation of GCRs.

4. Origin of the Open Flux Delay on the Sun

To further find out the origin of the delay of the open flux variation with respect to SSN variation, we investigate the area (OA) and the averaged magnetic field strength (MAG) of the solar open fields (the second and third panels of Figure 2(a) or Figure 3(a)), whose product is the open magnetic flux (OF, shown in the fourth panel). Using the same cross-correlation analysis, we find that the MAG changes with the SSN almost synchronously, with a small delay varying between ±3 BRs (the blue curve), but the OA lags behind the SSN substantially, by about 31 BRs or 2.3 yr on average, but sometimes by more than 40 BRs corresponding to 3 yr (see the red curve). Such great delays of the open area cause the open flux to significantly lag behind the SSN. Note that the delay of open area around the solar minimum between Cycles 23 and 24 changed abruptly. This segment is not reliable, because during that period there were multiple local maxima and minima even in the smoothed data from the open area, making it difficult to determine a unique association between the SSN minimum and the open area minima.

The detailed evolution of the open area and flux over solar cycles is shown in Figures 4(a) and (b), respectively. The open field of positive or negative polarities appears near one polar region at solar minima, gradually migrates toward low latitudes with increasing solar activity until solar maxima, and then
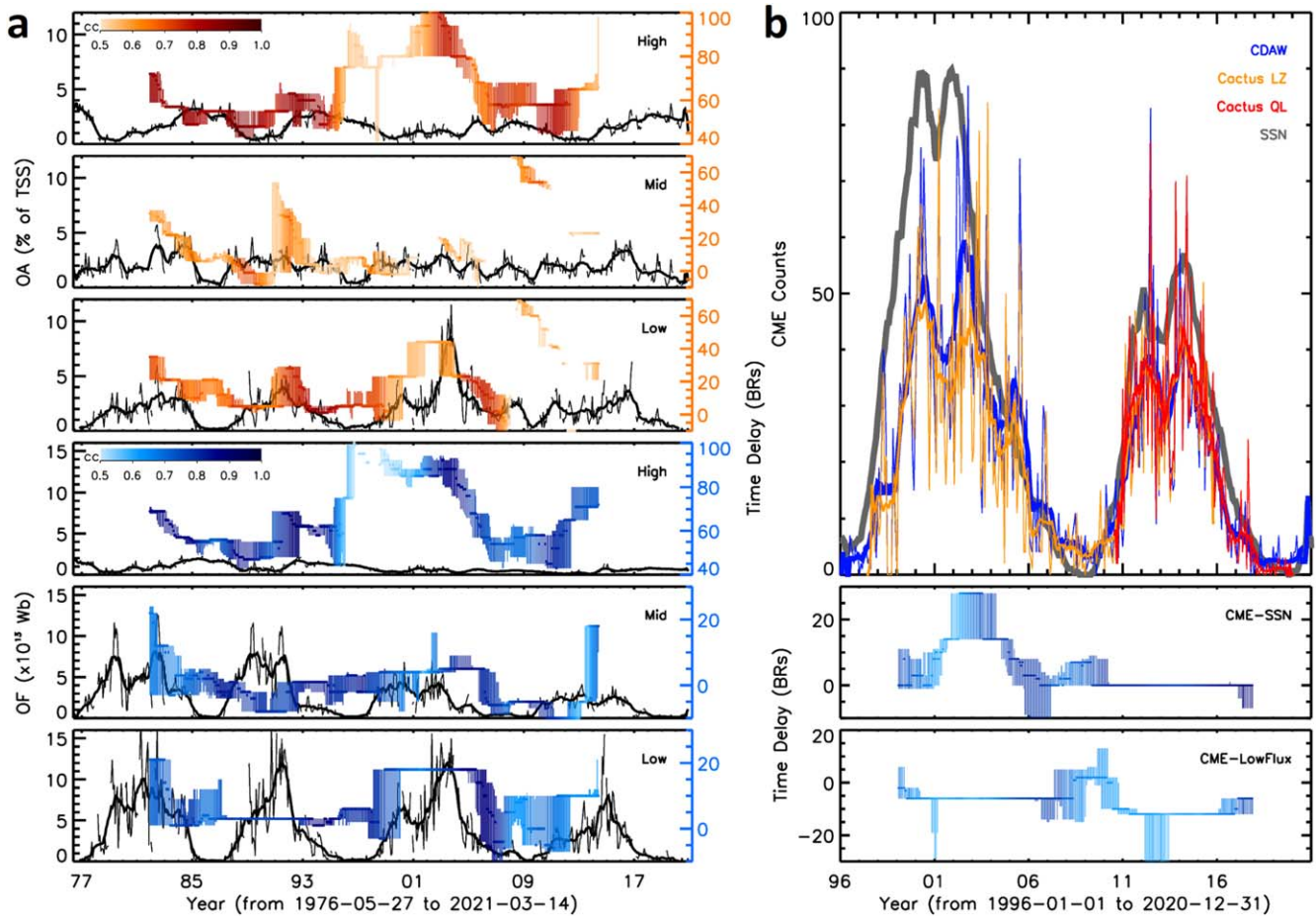
Figure 5. Solar cycle variations of the high, mid, and low latitudinal open fields and the CME occurrence. (a) From top to bottom are the open areas at the high (beyond ±60°), mid (between ±30° and ±60°), and low (within ±30°) latitudes and the open fluxes at the high, mid, and low latitudes. The color-coded ribbons have the same meaning as those in Figure 2(a), showing the delays relative to the SSN. (b) The top panel shows the number of CMEs per BR with a speed higher than 500 km s^{-1}. Thin curves are the original counts and thick ones are the 13 BR smoothed values. CME counts derived from the CDAW CME catalog (Yashiro et al. 2004) are indicated in blue and those from the Cactus catalog (Robbrecht & Berghmans 2004) are in orange and red. The smoothed SSN is overplotted as the gray curve (its value is scaled by 0.4). The bottom two panels show the time delays of the CDAW CME counts relative to the SSN and low latitudinal open flux, respectively. The color bar in the fourth panel of (a) is used.

continuously moves to the other polar region. During this course, the latitudinal variations of the total open area and the open flux follow the SSN, but their amplitudes evolve with a lag behind the SSN and reach the maximum generally during the declining phase of each cycle.

Solar open field originates mainly from coronal holes (normally located at high latitudes, Wang et al. 1996) and sometimes from active regions (distributed at low latitudes, Schrijver & Derosa 2003). We then investigate the OA and OF rooted in the high (beyond ±60°), mid (between ±30° and ±60°), and low (within ±30°) latitudinal zones, as displayed in Figure 5(a). It is found that the open areas at different latitudes are of the same order of magnitude. The variations of the open areas at low and mid latitudes follow the SSN with a delay, and those in polar regions are more or less reversed in phase. Consequently, the phases of the open flux at high and low latitudes are reversed too. This can be understood by noticing that the solar open field is mainly contributed by the lowest order of the multipolar magnetic fields (Wang et al. 2005; Jiang et al. 2011), i.e., the axial dipole field and the equatorial dipole field that provide most of the open fluxes at high and low latitudes, respectively, with reversed phases.

The open flux at low latitudes should be paid attention because it is overall about twice the high-latitude open flux. Considering

that low latitudes are the main locations of active regions, which can also contribute to the heliospheric magnetic flux through CMEs (Luhmann et al. 1998; Owens & Crooker 2006), we check the solar cycle variation of the CME occurrence number and its relationship with the low-latitude open flux.

CMEs have been continuously observed by the Large Angle and Spectrometric Coronagraph (LASCO, Brueckner et al. 1995) on board the SOHO spacecraft since 1996. There are several CME catalogs covering Cycles 23 and 24. Here we choose one manually-maintained CME catalog and one machine-automated catalog. The CDAW catalog (https://cdaw.gsfc.nasa.gov/CME_list/) is manually maintained (Yashiro et al. 2004), and the Cactus catalog (http://sidc.be/cactus/) is based on a machine-automated algorithm (Robbrecht & Berghmans 2004) using either the level zero (LZ) data (until 2015-10-31) or the quick look (QL) data (from 2010-07-09). The CME linear speed in the LASCO field of view is provided in the CDAW catalog and the median velocity is provided in the Cactus catalog. They are mostly comparable for CMEs without significant acceleration. CME counts below a certain speed threshold are calculated for each BR as shown in Figure 6. The comparison of the two different catalogs shows better consistency with a larger threshold. As a result, we choose the CMEs with a speed higher than 500 km s^{-1} for analysis
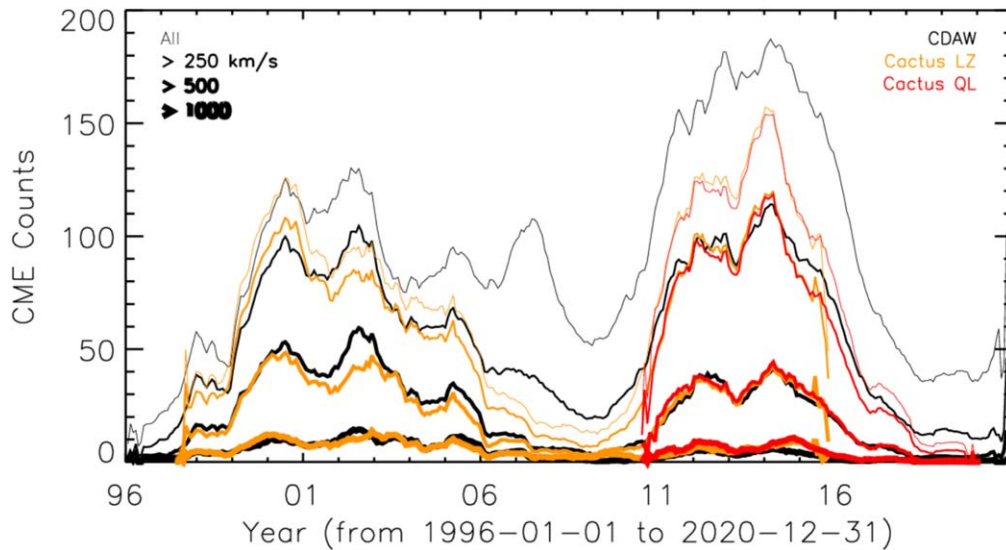
Figure 6. The solar cycle variations of the CME numbers according to the CDAW CME catalog (Yashiro et al. 2004) (black) and the Cactus catalog (Robbrecht & Berghmans 2004) based on the LASCO level zero data (orange) and on the LASCO quick look data (red). From thin to thick, the lines show the numbers of CMEs with the speed higher than 0, 250, 500, and 1000 km s^{-1}.

(Figure 5(b)) because the two catalogs are consistent and the CME number is large enough.

It is remarkable that the CME occurrence rate also shows the odd–even solar cycle pattern as revealed by the cross-correlation analysis (the top and middle panels of Figure 5(b)), i.e., it notably lagged behind the SSN in Cycle 23, but changed synchronously with the SSN in Cycle 24. Previous studies presented a model for dynamical energy balance in the flaring solar corona, which predicted a time lag between flare occurrence and the supply of energy to the corona (Wheatland & Litvinenko 2001). Studies of solar flare rates and SSN over five solar cycles confirmed this time delay and also found it to be much longer for odd solar cycles than for even cycles (Temmer et al. 2003). Both flares and CMEs represent the process of the gradual accumulation and impulsive release of the magnetic energy in the solar corona. This process may explain the flare/CME lagging behind the SSN, but not their odd–even behavior, which deserves further studies.

We further compare the CME numbers with open flux at low latitudes (the last panel of Figure 5(b)), and find that the open flux lagged behind the CME rate almost constantly, by about 6 BRs in Cycle 23 and 10 BRs in Cycle 24. This implies that part of the extrapolated solar open flux should result from the reconfigured magnetic fields after CMEs, consistent with previous studies that found that up to 30% of the heliospheric field could come from active regions (Schrijver & Derosa 2003), and CMEs could contribute to the open solar magnetic fields (Luhmann et al. 1998). According to the time delay between CME counts and open flux in Figure 5(b), we conjecture that the timescale of the eruption-driven field opening that can be reflected in the photospheric magnetogram and the extrapolation is about half a year. This finding sheds light on the origin and opening of the solar magnetic flux at low latitudes.

5. Conclusion

In summary, by investigating the GCR fluxes at various heliocentric distances and the solar open flux over the previous four solar cycles, we found that the GCR–SSN delay is longer than the OF–SSN delay, and both of them show the same odd–even cyclic pattern. We further found that the open flux delay is mainly due to the significant delay of the solar magnetic field opening, and CMEs make important contributions to the open flux at low latitudes. We also showed that the frequency of CME occurrence has an odd–even cycle dependence. Considering that the solar open flux is the source of HMF, we conclude that the GCR–SSN delay has its major origin on the Sun and is dominated by the solar cycle evolution at and below the solar surface. In other words, the delay of the open flux on the Sun is the primary contributor to the GCR delay in the heliosphere, and consequently is also the major cause of the odd–even cycle pattern of the GCR delay. This finding adds a new dimension (open flux delay) to the existing theories that address the GCR–SSN delay mainly based on the solar wind convection and transport of GCRs in the heliosphere, and should be folded into existing theoretical models. The correlations established among SSN, open flux, and GCR intensity in this study provide a basis for the long-term forecast of the GCR radiation levels in the heliosphere, which are important concerns for future human space missions (Cucinotta et al. 2017).

Appendix A
GCR Data Sets

The GCR data at Earth in Figure 1 are measured by the Electron Proton Helium Instrument (EPHIN), which is part of the Comprehensive Suprathermal and Energetic Particle Analyzer (COSTEP, Müller-Mellin et al. 1995) on the SOHO spacecraft. As a proxy for the measurement of GCRs, we utilize the channel for particles penetrating through the last detector with minimum ion energies of 53 MeV/nucleon. The GCR data at Earth in Figures 2 and 3 are measured by the neutron monitor at Oulu and the averaged daily values are downloaded from http://cosmicrays.oulu.fi. As GCRs enter Earth's atmosphere, they collide with atmospheric particles, producing secondary particles such as neutrons, which are then observed at detectors situated around the globe. Ground-level enhancement induced by solar energetic particles (SEPs) is removed and the daily data are binned into each BR with the median values used in the following analysis. The Oulu neutron monitor has been recording data since 1964. Given the Earth's geomagnetic field and atmospheric shielding, the cutoff rigidity for Oulu is 1 GV, corresponding to about 430 MeV, much

higher than that for SOHO/EPHIN. Thus, the delay time of the Oulu GCRs will be shorter than that of SOHO particles as illustrated in Figure A1(a), in which the upper panel shows the comparison between the BR data of the Oulu GCR flux and the data of the SOHO/EPHIN energetic particle flux, and the lower panel the time delay of the SOHO/EPHIN data with respect to the Oulu GCR data based on the cross-correlation analysis, with the same meaning as the ribbons in Figure 2(a). It is found that the amplitude of the time shift between the two data sets is steadily one or two BRs. When a shift of 2 BR is applied, the Oulu data and the SOHO/EPHIN data are well correlated, with a correlation coefficient as high as 0.95 (see Figure A1(b)).

The GCR data at Mars include the data from the High Energy Neutron Detector (HEND, Boynton et al. 2004) on board the Mars Odyssey spacecraft (https://pds-geosciences.wustl.edu/missions/odyssey/grs_edr.html) and the data from the Radiation Assessment Detector (RAD, Hassler et al. 2012) on board the Curiosity Mars rover (https://pds.nasa.gov/ds-view/pds/viewDataset.jsp?dsid=MSL-M-RAD-3-RDR-V1.0). HEND measures the albedo neutrons that are generated by primary GCRs in the Martian environment and scattered upwards to the orbit. The detector with the thickest moderator
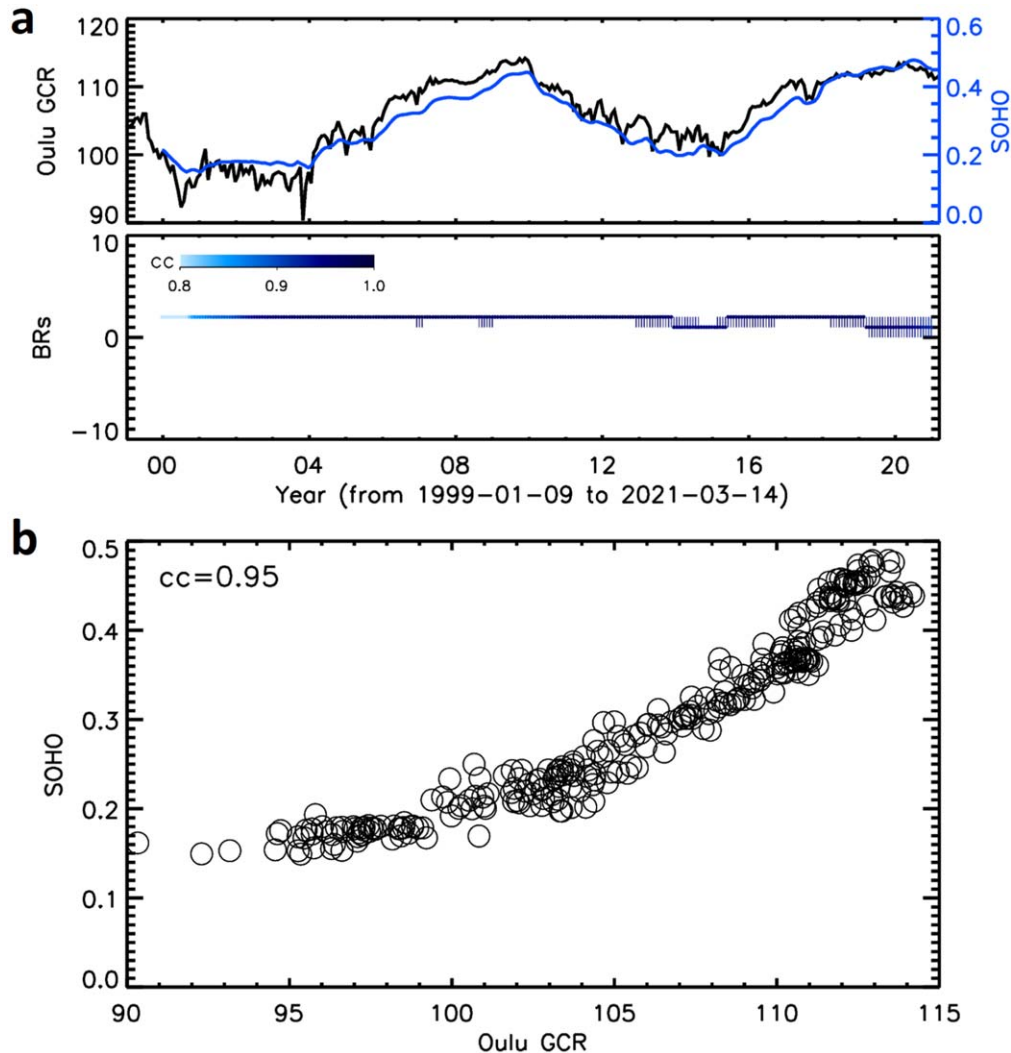
Figure A1. The upper panel of (a) shows the comparison between the BR data of the Oulu GCR flux (black) and the data of the SOHO/EPHIN energetic particle flux (blue). The lower panel of (a) shows the time delay of the SOHO/EPHIN data with respect to the Oulu GCR data based on the cross-correlation analysis, with the same meaning as in Figure 2(a). Panel (b) shows how well the Oulu data and the SOHO/EPHIN data are correlated when a shift of 2 BR is applied.

layer (about 30 mm) is most sensitive to neutrons with energies of 10 eV–1 MeV and its count rate is used in this study. RAD data are the absorbed dose rate, which is the rate of deposition of energy by all surface GCRs (including both primary and secondary particles generated in the atmosphere) in the plastic detector. SEPs are removed from both data sets before the daily values are averaged into those for each BR (median values are used).

The data of the Standard Radiation Environment Monitor (SREM, Evans et al. 2008) aboard Rosetta (Glassmeier et al. 2007) are available at https://spitfire.estec.esa.int/ODI/dplot_SREM.html. The TC2 channel with an energy threshold of about 49 MeV for protons is used in this study. SEPs are removed and daily values are binned into BR-averaged values. It should be noted that the distance of the Rosetta spacecraft varied greatly between 1 and 4.5 au after 2009 (Honig et al. 2019), and therefore its GCR profile contains the effects of spatial gradient due to the varying heliospheric distance. The same effect applies for the Cassini spacecraft in its cruise phase (blue line in the sixth panel of Figure 1(a), from 1 au to 9.5 au before 2004 July) and the New Horizons data (between 22 and 40 au). The gradient of the GCR flux is only about 2%–4% per au (at distances <10 au) according to observations from multiple spacecraft (Honig et al. 2019; Roussos et al. 2020). The heliospheric distance for each dot shown in Figure 1(b) is marked using the distance of the spacecraft when the extremum is identified, and extra caution should be taken for those obtained from these distance-changing data sets.

The data from the Cassini spacecraft were obtained through the Low Energy Magnetospheric Measurement System (LEMMS), one of the three energetic particle detectors of the Magnetosphere Imaging Instrument (MIMI)

suite (Krimigis et al. 2004, https://pds-ppi.igpp.ucla.edu/search/view/?f=yes&id=pds://PPI/CO-S-MIMI-4-LEMMS-CALIB-V1.0). LEMMS is a double-sided energetic particle telescope, with 57 counters designed to measure electrons and ions above several tens of keV and up to about 10 MeV (electrons) and several hundred MeV/nucleon (ions). GCR protons penetrate the shielding of LEMMS and get recorded as a low-intensity noise signal, which can be isolated from other magnetospheric or solar wind particle sources, as described in detail in earlier studies (Roussos et al. 2020). The measurements shown here are from LEMMS channel E6, the GCR response of which is dominated by >120 MeV protons.

The Pluto Energetic Particle Spectrometer Science Investigation (PEPSSI, McNutt et al. 2008) on board New Horizons (Stern 2009) was designed to measure keV–MeV pick-up ions from Pluto's outgassing atmosphere. Recently, the channels that were originally used to detect Jovian electrons below 1 MeV have been successfully used to approximate the deep-space GCR fluxes (Hill et al. 2020), and the data are described and accessible at http://sd-www.jhuapl.edu/pepssi/analysis/reducedData/. The channel corresponding to ions between 120 MeV and 1.4 GeV is used in this analysis. A few SEP events are removed and daily values are binned into BR-averaged values.
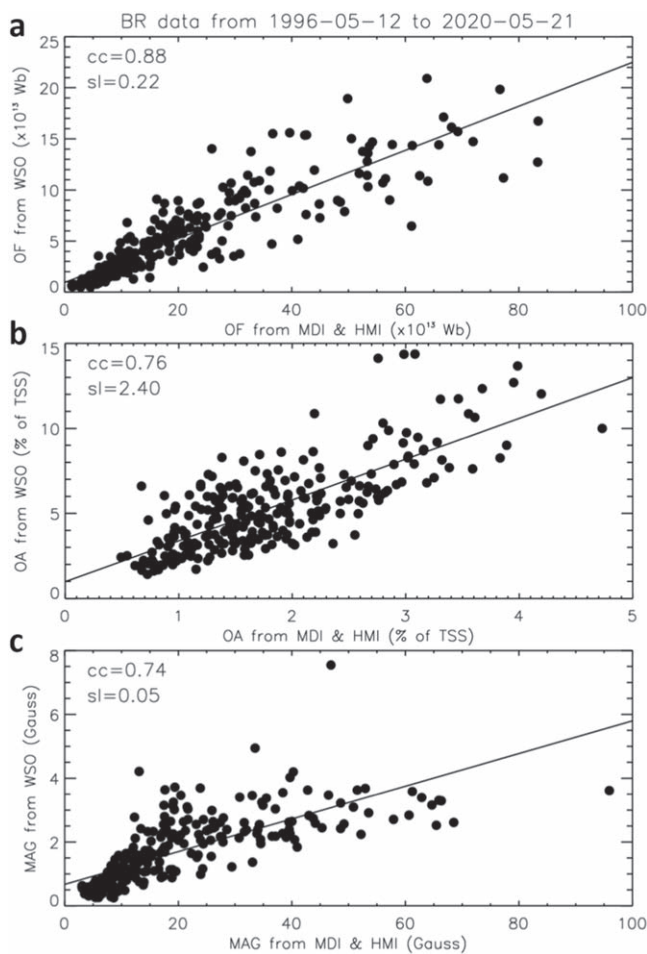
Appendix B
Comparison between WSO and HMI/MDI Synoptic Charts

Compared to the magnetograms from spaceborne instruments, e.g., the SOHO/MDI and the SDO/HMI, which provide the synoptic charts with resolutions of 3600 by 1080 and 3600 by 1440 in Cycles 23 and 24, respectively, the spatial resolution of
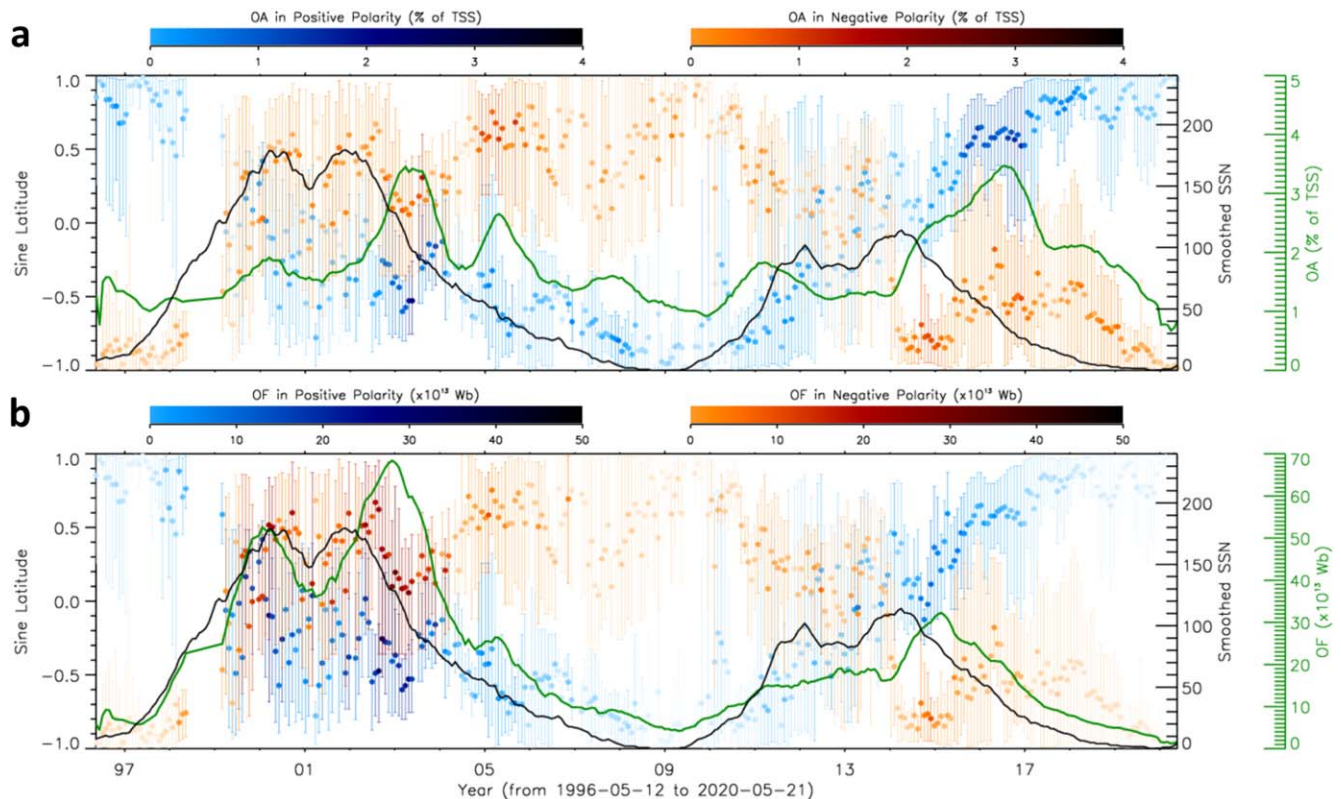
Figure B1. Similar to Figure 4. The data points before 2011 June 25, which is the beginning of BR 2414, are derived based on the MDI synoptic charts, and those after 2011 June 25 are obtained based on the HMI synoptic charts.

Figure B2. Scatter plots showing the correlation between the parameters derived based on the WSO data and those based on the MDI and HMI data during 1996–2020.

4–5 times smaller than that based on the MDI and HMI data, the open area is about 2–3 times larger, and the averaged open field strength is about 20 times smaller. But the correlation coefficients of the open flux, open area, and open field strength are as high as 0.88, 0.76, and 0.74, respectively, suggesting the parameters derived from WSO data are well correlated with those from the MDI and HMI data. Thus, the low spatial resolution of the WSO data should not distort the statistical results presented in the main text.

ORCID iDs

Yuming Wang ● https://orcid.org/0000-0002-8887-3919
Jingnan Guo ● https://orcid.org/0000-0002-6957-3355
Gang Li ● https://orcid.org/0000-0003-4695-8866
Elias Roussos ● https://orcid.org/0000-0002-5699-0678
Junwei Zhao ● https://orcid.org/0000-0002-6308-872X

References

Alanko-Huotari, K., Usoskin, I., Mursula, K., & Kovaltsov, G. 2007, AdSpR, 7, 1064
Boynton, W. V., Feldman, W., Mitrofanov, I., et al. 2004, SSRv, 110, 37
Brueckner, G. E., Howard, R. A., Koomen, M. J., et al. 1995, SoPh, 162, 357
Cane, H. V. 2000, SSRv, 93, 55
Clette, F., Svalgaard, L., Vaquero, J. M., & Cliver, E. W. 2014, SSRv, 186, 35
Cliver, E. W., & Ling, A. G. 2001, ApJL, 551, L189
Cucinotta, F. A., To, K., & Cacao, E. 2017, LSSR, 13, 1
Domingo, V., Fleck, B., & Poland, A. I. 1995, SSRv, 72, 81
Dorman, L. I. 2001, AdSpR, 27, 601
Duvall, T. L. J., Wilcox, J. M., Svalgaard, L., Scherrer, P. H., & McIntosh, P. S. 1977, SoPh, 55, 63
Evans, H., Bühler, P., Hajdas, W., et al. 2008, AdSpR, 42, 1527
Ferreira, S., & Potgieter, M. 2004, ApJ, 603, 744
Forbush, S. E. 1937, PhRv, 51, 1108
Forbush, S. E. 1958, JGRA, 63, 651
Glassmeier, K.-H., Boehnhardt, H., Koschny, D., Kührt, E., & Richter, I. 2007, SSRv, 128, 1
Gui, B., Shen, C., Wang, Y., et al. 2011, SoPh, 271, 111
Guo, J., Zeitlin, C., Wimmer-Schweingruber, R. F., et al. 2015, A&A, 577, A58
Guo, J., Zeitlin, C., Wimmer-Schweingruber, R. F., et al. 2018, Astron. J., 155, 49
Harrison, R. G., Ambaum, M. H. P., & Lockwood, M. 2011, RSPSA, 467, 2777
Hassler, D. M., Zeitlin, C., Wimmer-Schweingruber, R. F., et al. 2012, SSRv, 170, 503
Hassler, D. M., Zeitlin, C., Wimmer-Schweingruber, R. F., et al. 2014, Sci, 343, 1244797
Hill, M. E., Allen, R. C., Kollmann, P., et al. 2020, ApJ, 905, 69
Hoeksema, J. T., Liu, Y., Hayashi, K., et al. 2014, SoPh, 289, 3483
Honig, T., Witasse, O. G., Evans, H., et al. 2019, AnGeo, 37, 903
Jiang, J., Cameron, R. H., Schmitt, D., & Schüssler, M. 2011, A&A, 528, A83
Jokipii, J., Levy, E., & Hubbard, W. 1977, ApJ, 213, 861
Jokipii, J. R., Kota, J., Giacalone, J., Horbury, T. S., & Smith, E. J. 1995, GeoRL, 22, 3385
Krimigis, S., Decker, R., Roelof, E., et al. 2013, Sci, 341, 144
Krimigis, S. M., Mitchell, D. G., Hamilton, D. C., et al. 2004, SSRv, 114, 233
Laken, B., & Čalogović, J. 2013, JSWSC, 3, A29
Laken, B. A., Pallé, E., Čalogović, J., & Dunne, E. M. 2012, JSWSC, 2, A18
Luhmann, J. G., Gosling, J. T., Hoeksema, J. T., & Zhao, X. 1998, JGRA, 103, 6585
McNutt, R. L., Livi, S. A., Gurnee, R. S., et al. 2008, SSRv, 140, 315
Moloto, K., & Engelbrecht, N. E. 2020, ApJ, 894, 121
Müller-Mellin, R., Kunow, H., Fleißner, V., et al. 1995, SoPh, 162, 483
Owens, M. J., & Crooker, N. U. 2006, JGRA, 111, A10104
Parker, E. N. 1965, P&SS, 13, 9
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Pittock, A. B. 1978, RvGeo, 16, 400
Potgieter, M. 1998, SSRv, 83, 147
Potgieter, M. S. 2013, LRSP, 10, 1
Richardson, I. G. 2004, SSRv, 111, 267
Robbrecht, E., & Berghmans, D. 2004, A&A, 425, 1097

WSO data is too low. However, MDI data just covered the period from 1996 to 2010, and HMI data from 2010 to date. Thus, we use WSO data for the long-term analysis here. But we also check whether the low resolution of WSO data has any influence on the extrapolation of coronal magnetic field and consequently affects the statistical results. To do so, we apply the same procedures on the MDI and HMI synoptic charts (http://soi.stanford.edu/magnetic/index6.html and http://jsoc.stanford.edu/HMI/LOS_Synoptic_charts.html, respectively) to derive the open flux, open area, and open field strength.

Here, we reduce the spatial resolution of the MDI and HMI data to 360 by 180, calculate the spherical harmonic coefficients up to 180 orders, and trace field lines on a 180 by 90 mesh to save computing time. These reduced synoptic charts are 30 times higher in resolution than WSO charts. Figure B1 shows the solar cycle variations of the open area and open flux based on the MDI and HMI data. Their variation patterns look almost the same as those displayed in Figure 4, except that the absolute values are different. The open area derived based on the WSO data is larger than that based on the MDI and HMI data, whereas the open flux based on the WSO data is smaller. The correlation between the WSO and MDI/HMI data is displayed in Figure B2. It can be found that the slopes of the fitting lines obviously deviate from unity, suggesting that the open flux based on the WSO data is about

Ross, E., & Chaplin, W. J. 2019, SoPh, 294, 8
Rouillard, A., & Lockwood, M. 2004, AnGeo, 22, 4381
Roussos, E., Dialynas, K., Krupp, N., et al. 2020, ApJ, 904, 165
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129
Schrijver, C. J., & Derosa, M. L. 2003, SoPh, 212, 165
Shen, Z., Qin, G., Zuo, P., Wei, F., & Xu, X. 2020, ApJ, 900, 143
Stern, S. A. 2009, SSRv, 140, 3
Temmer, M., Veronig, A., & Hanslmeier, A. 2003, SoPh, 215, 111
Thomas, S. R., Owens, M. J., & Lockwood, M. 2014, SoPh, 289, 407
Usoskin, I., Bazilevskaya, G., & Kovaltsov, G. 2011, JGRA, 116, A02104

Usoskin, I., Kananen, H., Mursula, K., Tanskanen, P., & Kovaltsov, G. 1998, JGRA, 103, 9567
Usoskin, I., Mursula, K., Kananen, H., & Kovaltsov, G. 2001, AdSpR, 27, 571
Van Allen, J. A. 2000, GeoRL, 27, 2453
Wang, Y., & Zhang, J. 2007, ApJ, 665, 1428
Wang, Y.-M., Hawley, S. H., & Sheeley, N. R. 1996, Sci, 271, 464
Wang, Y.-M., Lean, J. L., & Sheeley, N. R., Jr. 2005, ApJ, 625, 522
Webber, W. R., & Lockwood, J. A. 1988, JGRA, 93, 8735
Wheatland, M., & Litvinenko, Y. E. 2001, ApJ, 557, 332
Yashiro, S., Gopalswamy, N., Michalek, G., et al. 2004, JGRA, 109, A07105
Zhao, X., & Hoeksema, J. T. 1995, JGRA, 100, 19