

# Time Lag Between Cosmic-Ray and Solar Variability: Sunspot Numbers and Open Solar Magnetic Flux

Sergey A. Koldobskiy<sup>1,2</sup> · Riikka Kähkönen<sup>1</sup> · Bernhard Hofer<sup>3</sup> · Natalie A. Krivova<sup>3</sup> · Gennady A. Kovaltsov<sup>2,4</sup> · Ilya G. Usoskin<sup>1</sup>

Received: 21 December 2021 / Accepted: 25 February 2022 © The Author(s) 2022

# Abstract

Solar magnetic activity drives the dominant 11-year cyclic variability of different space environmental indices, but they can be delayed with respect to the original variations due to the different physical processes involved. Here, we analyzed the pairwise time lags between three global solar and heliospheric indices: sunspot numbers (SSN), representing the solar surface magnetic activity, the open solar flux (OSF), representing the heliospheric magnetic variability, and the galactic cosmic-ray (GCR) intensity near Earth, using the standard crosscorrelation and the more detailed wavelet-coherence methods. All the three indices appear highly coherent at a timescale longer than a few years with persistent high coherence at the timescale of the 11-year solar cycle. The GCR variability is delayed with respect to the inverted SSN by about eight 27-day Bartels rotations on average, but the delay varies greatly with the 22-year cycle, being shorter or longer around positive A+ or negative Asolar polarity epochs, respectively. The 22-year cyclicity of the time lag is determined by the global heliospheric drift effects, in agreement with theoretical models. The OSF lags by about one year behind SSN, and is likely determined by a combination of the short lifetime of active regions and a longer ( $\approx$ 3 years) transport time of the surface magnetic field to the poles. GCRs covary nearly in antiphase with the OSF, also depicting a strong 22-year cycle in the delay, confirming that the OSF is a good index of the heliospheric modulation of GCRs. This provides an important observational constraint for solar and heliospheric physics.

Keywords Cosmic rays, galactic · Solar cycle · Magnetic fields, interplanetary

 S.A. Koldobskiy sergey.koldobskiy@oulu.fi
 I.G. Usoskin

ilya.usoskin@oulu.fi

- <sup>1</sup> University of Oulu, Oulu, Finland
- <sup>2</sup> National Research Nuclear University MEPhI, Moscow, Russia
- <sup>3</sup> Max Planck Institute for Solar System Research, Göttingen, Germany
- <sup>4</sup> Ioffe Physical-Technical Institute, St. Petersburg, Russia

## 1. Introduction

The flux of galactic cosmic rays (GCRs) outside the heliosphere is generally assumed to be constant at the time scales shorter than a hundred thousand years. At the same time, the flux measured near/at Earth varies at different time scales, reflecting the process of the heliospheric modulation of GCRs. Although the physics of the modulation process is well understood (e.g. Potgieter, 2013), full quantitative modeling is difficult. Therefore, it is important to use observational data to validate models and accurately constrain the range of possible parameters. One such constraint relates to the time lag (delay) of the GCR flux measured near Earth with respect to solar activity, typically represented by the sunspot number (SSN) (e.g. Clette and Lefèvre, 2016). The GCR flux variability is known to be delayed with respect to solar activity, leading to the so-called "hysteresis" effect of the phase shifts in the development of the 11-year solar cycle in both indices.

The long-term variability of GCRs is routinely monitored by a global network of groundbased neutron monitors (NMs) that are sensitive to the nucleonic component of the cosmicray-induced atmospheric cascades and located, as a worldwide network, around the globe since the 1950s (e.g. Belov, 2000; Simpson, 2000). Long-term records of NM count rates are typically used for the time-lag analysis (see more details in Section 3.1). Traditionally, the time lag between the GCR flux and SSN is defined via the cross-correlation analysis, but this approach cannot distinguish between different timescales and is thus not easy to interpret.

Soon after establishing the NM network, it was found that the GCR intensity lags behind the solar variability (Forbush, 1958; Dorman and Dorman, 1967; Mavromichalaki and Petropoulos, 1984) and that the lag varies from one cycle to another (Nagashima and Morishita, 1979). A detailed study of the lag between GCR and SSN for four solar cycles (1953–1995) was first performed by Usoskin et al. (1998) and updated for five cycles (1951-2000) by Usoskin et al. (2001), using the standard cross-correlation and a sophisticated 2D cycle-projection analysis. It was shown that the time lag varies significantly, depicting a 22-year pattern (Hale et al., 1919) so that it is longer (10-20 months) for negative (A-) and shorter (several months or even formally negative) for positive (A+) polarity cycles. The robustness of this main result has been confirmed by numerous subsequent studies (e.g. Inceoglu et al., 2014; Tomassetti et al., 2017; Mishra and Mishra, 2018; Iskra et al., 2019), mostly based on the cross-correlation analysis. Most of the previous works studied the lags for individual 11-year cycles (Schwabe, 1844) without providing a detailed temporal behavior of the lag. Thus, the fact that GCR intensity lags behind the solar activity by several months and the lag varies between subsequent 11-year cycles is well established; however, details are still unclear and depend on the exact method of analysis.

The observed time-lag pattern is well understood in the sense of the physical processes driving the modulation (Potgieter, 2013). Sunspots do not affect cosmic-ray intensity directly, but the surface magnetic activity, especially large and long-lived active regions, can produce disturbances in the properties of the solar wind (Stansby et al., 2021) that modulates the GCR transport in the heliosphere. The delay of cosmic-ray variability, measured near Earth, with respect to solar activity is generally related to the outward propagation of solar wind with a frozen-in heliospheric magnetic field (HMF) and to inward diffusive transport of GCR particles, the latter being affected by large-scale drifts. Different time lags for different cycles are understood as the effect of large-scale drifts of GCR in the heliosphere (e.g. Jokipii and Levy, 1977) and can be generally reproduced by modern numerical models (e.g. Gervasi et al., 1999; Alanko-Huotari et al., 2007; Strauss et al., 2011; Boschini et al., 2018). However, many details are still missing. This particularly concerns the parameters of



**Figure 1** Datasets used here: (a) International sunspot number, v.2.0 (SSN). (b) Open solar flux (OSF), including its fast (red shading) and slow (blue) components. (c) Standardized NM records for the three cutoff rigidity  $P_c$  groups: low (green), moderate (blue) and high (red) (see details in Figure 2). All datasets are shown with the Bartels rotation (27 days) averaging. Periods of the reversals of the heliospheric magnetic field polarity are shaded with gray (Thomas, Owens, and Lockwood, 2014; Krainev et al., 2016).

the models (e.g. the diffusion tensor) that cannot be measured or computed directly but need to be inferred through largely empirical relations. Accordingly, it is important to quantify the observed time lags and provide constraints to the models.

One of the global physical parameters of the heliosphere is the open (solar) magnetic flux (OSF), which is the magnetic flux that leaves the top of the solar atmosphere (the so-called source surface), is dragged by the radially expanding solar wind and propagates through the heliosphere (Lockwood, 2013). It was proposed (e.g. Usoskin et al., 2004) that the OSF can be considered as a primary solar/heliospheric index of cosmic-ray variability when studying the long-term behavior of GCRs using cosmogenic proxy data (e.g. Owens, Usoskin, and Lockwood, 2012; Wu et al., 2018; Usoskin et al., 2021).

Here, we update an empirical study of the time relation between GCRs and two global solar/heliospheric indices, SSN, and the OSF, using a modern analysis based on wavelet coherence with emphasis on the 11-year solar cycle time-scale and the odd–even-cycle effect. We also discuss the obtained results in the context of the modulation theory.

# 2. Datasets and Their Sources

#### 2.1. Sunspot Numbers (SSN)

The first dataset is the international sunspot number (ISN, Clette et al., 2014) version 2.0 as available at Sunspot Index and Long-term Solar Observations (SILSO, https://wwwbis.sidc. be/silso/). The relative sunspot number (SSN) is a synthetic (not physical) index composed as follows. The number of sunspot groups g and individual spots s are counted daily by



**Figure 2** Data-availability plot for the 28 NMs considered in this article (NM names, sorted by cutoff rigidity  $P_c$  value, are given on the left). Full information about these NMs can be found elsewhere (Väisänen, Usoskin, and Mursula, 2021). NMs are divided into three groups depending on their average cutoff rigidity: low (< 2 GV), moderate (2–6 GV), and high (>6 GV).

many observers around the world and reported to the SILSO data center, where the relative sunspot number, R, is computed as a weighted average of the counts  $R = k \cdot (10 \cdot g + s)$  for individual observers with the personal "quality" factor k (see Clette et al., 2007, for details). This dataset is the commonly used index of solar activity whose high quality since the mid-20th century is undoubted (Clette and Lefèvre, 2016). We used the daily SSN that has no gaps over the studied period. The SSN dataset is shown in Figure 1a.

## 2.2. Open Solar Magnetic Flux (OSF)

Here, we use the daily OSF as computed using the newest version of the SATIRE-T (Spectral and Total Irradiance Reconstructions) model based on direct solar observations for the last century (Krivova et al., 2021). This reconstruction is consistent with the OSF record from in situ space-borne measurements since the 1980s (Owens et al., 2017). The OSF involves two components: the slow one, mostly related to the flux accumulation at the solar poles, and the fast one, mostly related to the magnetic flux taken away by coronal mass ejections (CMEs) propagation (Owens and Crooker, 2006). The OSF dataset is shown in Figure 1b, including both components.

## 2.3. Neutron Monitors

NMs have been in operation since 1951 with a worldwide network of standard instruments set in 1964 (e.g. Simpson, 2000). The count rate of the NM depends on two factors: the GCR flux on top of the atmosphere and the local thickness of the atmosphere above the NM. The latter is directly related to the barometric pressure measured at the NM location. In order to represent the GCR intensity, the NM count rate must be first corrected for the barometric effect (Dorman, 2004). The correction reduces the NM count rate to the nominal pressure

value and is straightforward to apply, using a nearly perfect exponential relation between the barometric pressure and the NM count rate. These pressure-corrected count rates are used in the forthcoming analysis. The GCR flux on top of the atmosphere depends also on the geomagnetic shielding, as quantified through the geomagnetic cutoff rigidity,  $P_c$ , (Cooke et al., 1991). NMs located at different latitudes are sensitive to different energies/rigidities of GCRs (Asvestari et al., 2017) making the global NM network a rough spectrometer.

Data from NMs are collected in several sources: the World Data Centre for Cosmic Rays (WDC CR, https://cidas.isee.nagoya-u.ac.jp/WDCCR), Neutron-Monitor Data Base (NMDB, https://www.nmdb.eu), IZMIRAN Database (http://cr0.izmiran.ru/common/links. htm), and individual NM websites (e.g. http://cosmicrays.oulu.fi). The reliability of NM data from different datasets was recently analyzed by Väisänen, Usoskin, and Mursula (2021), who provided a list of the most stable NMs (called "primary") and their data sources. Here, we use data from long-operational NMs from this dataset, averaged over 27-day Bartels rotations (BR). NM data were grouped into three groups according to their geomagnetic cutoff rigidity  $P_c$ : low ( $P_c < 2$  GV), moderate ( $2 < P_c < 6$  GV), and high ( $P_c > 6$  GV) rigidity, as shown in Figure 2. The full dataset covers the period from 1951–2019, namely 6.5 solar cycles, but there is no single NM that was continuously operational throughout the entire period.

We combined the records of NMs in each rigidity group. First, the overlap periods were defined when all NMs within each group were taking data during most of the time. These periods are 1970 - 1988, 1982 - 1998, and 1976 - 1992 for low-, moderate-, and high-cutoff NM groups, respectively. For these periods, the mean count rate,  $\langle N \rangle_i$ , and its standard deviation,  $\sigma_i$ , were calculated for each *i*th NM. Next, the (pressure-corrected) count rate of the *i*th NM,  $N_i$ , was standardized using these values,  $\langle N \rangle_i$  and  $\sigma_i$  defined for the overlap period, for the entire NM time series. The standardized count rate of the *i*th NM,  $N_i^*(t)$ , was calculated from its recorded pressure-corrected count rate,  $N_i(t)$ , as

$$N_i^*(t) = \frac{N_i(t) - \langle N \rangle_i}{\sigma_i}.$$
 (1)

Finally, all standardized NM count rates were averaged within each  $P_c$  group to produce a homogeneous combined time series. These combined datasets are shown in Figure 1c. It is interesting that they are very close to each other, implying that the same modulation process operated at different rigidities. The only small but systematic discrepancy can be observed between the mid-rigidity and the two other datasets over the period 1957–1965. This is probably due to the small number of NMs.

#### 2.4. Direct Cosmic-Ray Measurements

We also made use of the direct GCR proton-flux measurements performed by space-borne spectrometers at low orbits: PAMELA (Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics: Adriani et al., 2013; Martucci et al., 2018), which performed measurements from June 2006 through January 2016, and AMS-02 (Alpha Magnetic Spectrometer v.2: Aguilar et al., 2018), which has been in operation since May 2011. Data were obtained from the Cosmic-Ray Database https://tools.ssdc.asi.it/CosmicRays/. Data from PAMELA and AMS-02 experiments are slightly different in their time cadences, which are Carrington rotations (CaR, 27.2753 days, corresponding to the mean synodic solar-rotation period as used in solar astronomy) and Bartels rotations (BR, 27 days, often used in helio-spheric physics), respectively, as well as in energy/rigidity channels. Accordingly, the data need calibration to be merged together.



Figure 3 Normalized GCR proton fluxes as measured by PAMELA (red curves) and AMS-02 (blue) experiments for the three rigidity channels studied here (panels a through c). The period used for normalization is indicated with the green shading.

For further analysis, we produced a homogeneous series of cosmic-ray proton variability at three rigidity channels: low-rigidity (around 1.1 GV), mid-rigidity ( $\approx$ 5.1 GV) and highrigidity ( $\approx$ 10.7 GV). However, the exact rigidity channels are different in the considered experiments: 1.08  $\pm$ 0.08 GV, 5.125  $\pm$ 0.245 GV, 10.55  $\pm$ 0.55 GV in AMS-02 and 1.06  $\pm$ 0.03 GV, 5.21  $\pm$ 0.24 GV, 10.88  $\pm$ 0.5 GV in PAMELA. Accordingly, we renormalized the PAMELA data to AMS-02 channels by a simple scaling as follows. First, we considered the period of their data overlap from May 2011 (CaR 2110, BR 2426) through June 2014 (CaR 2146, BR 2462) and calculated the mean value of the measured cosmic-ray flux in each rigidity bin mentioned above for the overlap period. Then, all data for each experiment were divided by the mean value during the overlap period, as shown in Figure 3. For the date of transition from PAMELA to AMS-02 data, we took 5 November 2012 (beginning of BR 2446 and CaR 2130), so that (normalized) PAMELA data are used before that date and AMS-02 data from that day onward.

#### 2.5. Data Processing

All datasets were averaged over BRs to avoid short-term transient effects. Data gaps were not filled (marked as NaN) for the cross-correlation analysis (Section 3.1) and filled with linear interpolations for the wavelet-coherence analysis (Section 3.2). Days with solar energetic-particle events registered by NMs (ground level enhancements, GLEs, Usoskin et al., 2020) were removed from both NM data and PAMELA+AMS-02 datasets.



**Figure 4** Pearson correlation coefficient, *r*, between GCR and SSN datasets, as a function of the delay for the period 2006-2017 as obtained for the combined moderate-cutoff NM record (blue) and for the PAMELA+AMS-02 flux measurements at  $R \approx 5.1$  GV (red). Maximum absolute values of |r| are denoted with the stars and their  $1\sigma$  uncertainties are denoted with horizontal dotted lines. The vertical arrows bound the corresponding uncertainties of the delay.

## 3. Analysis of Time Lags Between Different Datasets

#### 3.1. Correlation Analysis

A standard way of determining the time lag between two time series is to use crosscorrelation analysis with the following methodology. Let x(t) and y(t) be two time series with the same cadence, to be analyzed. The standard Pearson's cross-correlation coefficient, r, between the x series and y series, the latter lagged by time  $\tau$ , within a time window of Ntime points is:

$$r(\tau) = \frac{\sum_{j=1}^{N} (x(t_j) - \bar{x}) \times (y(t_j + \tau) - \bar{y})}{\sqrt{\sum_{j=1}^{N} (x(t_j) - \bar{x})^2 \times \sum_{j=1}^{N} (y(t_j + \tau) - \bar{y})^2}}$$

$$\bar{x} = \frac{1}{N} \sum_{j=1}^{N} x(t_j), \quad \bar{y} = \frac{1}{N} \sum_{j=1}^{N} y(t_j + \tau).$$
(2)

The time lag  $\tau^*$  between the two series is defined as the value of  $\tau$  that maximizes (the absolute value of) the lagged cross-correlation  $|r(\tau^*)|$ . For the analysis, we used the correlation window length N corresponding to 11 years, that is N = 149 BRs.

An example of the thus determined time lag, along with its 68% confidence interval, of GCR variability with respect to SSN is shown in Figure 4 for the period 2006–2017, as obtained for the moderate-cutoff NM record and for the PAMELA+AMS-02 proton-flux measurements in the rigidity channel around 5.1 GV. It is noteworthy that the dependence of r on  $\tau$  is not smooth. Although it has a roughly inverted bell-like shape allowing for a general definition of the lag, the exact lag value is not precisely determined. The standard way to define the lag is to take the  $\tau^*$ -value corresponding to the formally maximum value of  $|r^*|$  (shown by the stars in Figure 4). However, this formal  $r^*$ -value does not guarantee, for



Figure 5 Time lag between cosmic-ray variability and SSN. Color curves represent count rates from NMs in different rigidity groups (Figure 2) as specified in the legend. Colored symbols with error bars correspond to the results based on different rigidity bins of the combined proton-flux data from PAMELA and AMS-02 experiments.  $R_2$  and  $R_3$  points are slightly offset on the time axis for better visibility. Error bars and shaded areas denote the 68% confidence intervals for the time lags. Each point corresponds to the center of the considered correlation 11-year interval. Vertical gray bars denote the approximate epochs of the HMF polarity change, as in Figure 1.

a real limited-size and noisy dataset, the true best correlation and is not statistically distinguishable from the neighboring values. Thus, it cannot guarantee the true value of the time lag. Accordingly, we have considered, as the uncertainties of the time-lag determination, the range of  $\tau$ -values corresponding to the range of |r|-values between  $|r^*|$  to  $|r^* - \sigma_r|$ , where  $\sigma_r$  is the 68% confidence level standardly defined for the best-fit Pearson's correlation coefficient,  $r^*$ , as denoted by the dashed lines in the figure. While the best-fit  $\tau^*$ -values (stars in the figure) are close to each other for NMs and the combined PAMELA+AMS-02 datasets, being 4–6 BRs, their 68% confidence intervals, marked with the vertical arrows in the figure, cover the ranges of [2-3] and [-1-14] BRs, for the two datasets, respectively. These values of the time lags were ascribed to the center of the considered 11-year correlation time window. Following this methodology, we computed the time lags, along with their uncertainties, as a function of time, between different GCR datasets and SSN for all individual datasets. All deduced time-lag values for both NMs vs. SSN and PAMELA+AMS-02 vs. SSN were found to be significant (*p*-value < 0.01), with the Pearson correlation coefficient |r| varying in the range from  $\approx 0.5$  to almost 1.

The results of the time-lag correlation analysis applied to GCR and SSN data are shown in Figure 5. Being on average about 8 BRs, the time lag of the combined NM records with respect to SSN depicts a clear bimodal distribution with alternating short (about 3 BRs) and long (12-15 BRs) delays between NMs and SSN. The switch between the two modes is fast and roughly corresponds to the change of the large-scale HMF polarity, leading to a pronounced 22-year pattern in the time lag. Approximate periods of the polarity change are marked with gray shading in Figure 5. The results are not accurate before 1965, because of the smaller number of NMs involved in the analysis. The pattern is consistent for all



**Figure 6** Time lag between the moderate-rigidity combined NM record (blue curve in Figure 1c) and SSN (Figure 1a). Black and blue dashed curves represent the time lag determined using the cross-correlation analysis (Section 3.1) for the raw and bandpass [5-16 years] filtered data, respectively. The magenta curve represents the time lag for the same dataset determined using the wavelet-coherence method (Section 3.2). Shaded areas denote the uncertainties as described in the text. Some earlier results are shown by colored horizontal bars as denoted in the legend. The used references are: I2014 – Inceoglu et al. (2014), T2017 – Tomassetti et al. (2017), M2018 – Mishra and Mishra (2018), I2019 – Iskra et al. (2019).

the studied rigidity intervals, implying no dominant rigidity/energy dependence of the delay (different colored curves in Figure 5 agree within the error bars). However, there is a tendency of a longer delay for higher-rigidity (>6 GV) GCR during the last decade, as represented by the high-cutoff NMs (red curve) and the high-rigidity ( $\approx$ 10 GV) proton channel of PAMELA+AMS-02 dataset, but the difference is not statistically significant.

The correlation method does not distinguish between different timescales since it combines slow large-scale modulation and fast transients, whose impacts during the periods of high solar activity can be significant. The effect of recurrent (27-day) variability is suppressed by the BR averaging of data but nonrecurrent transients such as CMEs, (merged) interaction regions, fast solar-wind streams, etc., can still play a significant role in variability and hence in the correlations. Accordingly, it is unclear from the above analysis whether, for example, the short formal delay around a solar-cycle maximum reflects real fast access of GCRs to the Earth's vicinity, or is rather caused by the local modulation by fast transients. To investigate this we also performed a similar correlation analysis but with smoothing of the analyzed time series with a bandpass fast Fourier transform (FFT) filter with the period ranging from 5 to 16 years. Results of this analysis are shown by the dashed blue line in Figure 6. In general, the results based on the bandpass filtered data are consistent with the results based on raw data, but show a smoother variability of the time lag. This is important in the light of the comparison with the wavelet-coherence analysis that will be discussed in the next section.

Figure 6 also shows the cycle-averaged time lags between NMs and SSN data after 1964 as published previously by different teams (Inceoglu et al., 2014; Tomassetti et al., 2017; Mishra and Mishra, 2018; Iskra et al., 2019). Our results, obtained by the same cross-correlation method, agree well with the previous results but provide more detailed time variability of the lag.

#### 3.2. Wavelet-Coherence Analysis

The correlation analysis presented above compares the analyzed datasets in the time domain and provides a single number, the Pearson's correlation coefficient, without distinguishing



**Figure 7** Wavelet-coherence analysis of the SSN and the standardized moderate-cutoff NM dataseries time series. (**a**): Analyzed time series identical to those in Figures 1a and c. (**b**): Wavelet coherence between these series. Color represents the coherence as depicted in the color bar on the right, arrows depict the phase of the coherence, and the shaded area denotes the cone of influence. Black contours bound the regions with highly significant (p < 0.01) coherence. White horizontal dashed lines denote the 11-year and 22-year periods. (**c**): Time lag in Bartels rotations (as quantified by the color scale on the right) between the two time series, for the areas with highly significant (p-value < 0.01) coherence as shown in panel b. The horizontal dashed line denotes the 11-year period, while the gray-shaded region corresponds to the interval of periods considered for uncertainty calculation (see text). (**d**): Temporal evolution of the time lag within the range of periods defined in panel c (see text).

between different timescales. Accordingly, the result can depend on the length of the series, time resolution, sampling/smoothing (as was shown above for the bandpassed data), etc. More detailed information on the covariance of the analyzed datasets can be provided by the coherence analysis, which projects the concept of correlation into the frequency domain. The next level of detail is provided by the wavelet-coherence analysis, which unfolds the coherence into a 2D pattern in both frequency and time domains. The wavelet coherence yields the power and phase of the normalized crosswavelet spectrum of the two data series

so that it varies between zero (no coherence) and unity (full coherence) at any given time and frequency, accounting also for the relative phase. For example, the lag of one quarter of the cycle would lead to the formal zero cross-correlation between the series but would be correctly identified as a full coherence with a phase shift of 90° by the coherence analysis.

We used the wavelet-coherence Matlab package based on the algorithm described by Grinsted, Moore, and Jevrejeva (2004), applying the Morlet wavelet basis with the scale parameter k = 3. The statistical significance of the coherence C(f, t) between series x(t) and y(t) is estimated at each time t and frequency f location, using a standard procedure based on the Monte Carlo method that can be briefly described as follows. First, one of the analyzed series (let it be x) is replaced by a synthetic one  $x^*$  composed of the red noise with the same AR1 (autoregression with unity time shift) coefficient as the original x series. Then, the wavelet coherence between the synthetic  $x^*$  and the original y series is computed as  $C^*$ . This procedure is repeated N = 1000 times, randomly choosing the series to be randomized, and the significance of the coherence C(f, t) at each time–frequency location is estimated as  $N^*/N$  where  $N^*$  is the number of simulated cases when  $|C(f, t)^*| > |C(f, t)|$ . A full description of the method is available elsewhere (Grinsted, Moore, and Jevrejeva, 2004).

An example of the wavelet coherence is shown in Figure 7b for the SSN and the standardized moderate-cutoff NM count rate (see Figures 1a and c). Color represents the coherence (see the color scale on the right), and arrows depict the phase of the coherence, so that the arrows pointing to the right/left indicate inphase/antiphase coherence, respectively. The shaded area denotes the cone of influence where the results are unreliable because of the proximity to the edge of the series. Black contours bound the regions of highly significant (p < 0.01) coherence defined as described above. One can see that the coherence is highly significant for the range of periods between a few and 30 years, but it is systematically high throughout the entire studied time interval only in a narrow range of periods around 11 years, being nearly in antiphase.

From the coherence phase (clockwise angle  $\phi$  of arrows in Figure 7b) at a given period T, we determined the time lag between the two analyzed series as  $\tau = \min(T \cdot (\phi \pm \pi)/2\pi)$ . Accordingly, we calculated the time lag  $\tau$  between the SSN and NM series as a function of time and period, as shown in Figure 7c. Thus, the defined time lag is meaningful only if the coherence is highly significant (*p*-value < 0.01), otherwise the exact phase, and thus the lag value, can be spurious. One can see that the time lag can be determined only for a range of periods around 11-years for the considered data (SSN vs. NMs). This period is of particular interest, and we calculated the time lag  $\tau$  within the 11-year cut (shown with a dotted black line). As the time-lag uncertainty, we considered a full range of  $\tau$ -values computed at each time point in the range of periods from 9.4 to 12.6 years, as denoted by the gray-shaded region in the figure. The resulting time-lag series is shown in Figure 7d. A clear  $\approx$ 22-year cyclic variation of the time lag from 3 to 12 BRs can be observed, in general agreement with the results of the correlation analysis for the bandpass filtered data (Figure 6). The time lag becomes very short ( $2 \pm 3$  BRs) around 1972, formally implying almost no (or even slightly negative within the confidence interval) delay, indicating that the GCR level was recovering faster than normal during the declining phase of Solar Cycle 20 (see discussion in Usoskin et al., 1997). That period is known for the unusual heliospheric current-sheet structure leading to the distortion of the regular GCR modulation, which is known as the *minicycle* (Webber and Lockwood, 1988; Usoskin et al., 1998). Therefore, this short delay was caused by the unusual heliospheric structure.

Next, we performed a similar analysis for the OSF vs. SSN datasets. The OSF time series is almost perfectly coherent (Figure 8b) with SSN in the entire range of periods longer than a few years, but with a systematic delay of about 12 BRs (Figure 8d). Since the OSF is



Figure 8 The same as Figure 7 but for SSN vs. the OSF time series.

composed of two components, fast and slow (Figure 1b), we also analyzed them separately. The corresponding delays with respect to SSN are shown in Figure 9. One can see that while the fast component is delayed with respect to SSN by a few BRs, the smooth slow component is delayed by about 30 BRs. The fast-component lag is consistent with the opening time scale of  $\approx$ 50 days for the CME contribution to the OSF (Owens and Crooker, 2006). The slow-component lag of  $\approx$ 30 BRs is consistent with the migration time of the surface flux from active regions to the solar poles, considering the mean meridional transport velocity at the surface of  $\approx$ 10 m/s (Gizon et al., 2020), where it builds up the poloidal magnetic field. The combined effect from the delays for the two OSF components produces a nearly constant delay of  $\approx$ 12 BRs of the total OSF.

Wavelet coherence between the moderate-rigidity combined NM record and the OSF time series is shown in Figure 10. One can see that it resembles the pattern observed for the NMs vs. SSN relation (Figure 7) but with the different absolute values of the time lag so that cosmic-ray variability formally leads that of the OSF by several months.



Figure 9 Wavelet-coherence-based time delays of the OSF (black solid curve, identical to that shown in Figure 8d) and its fast (red dash-dotted curve) and slow (blue dotted curve) with respect to SSN. Shaded areas denote the uncertainty range (see text).

Finally, we present in Figure 11 wavelet-based time lags for the three datasets, that is GCRs (combined NM records), the OSF, and SSN. One can see that, while the temporal variability of the NMs vs. SSN and NMs vs. the OSF lags have a similar 22-year cyclic shape, the absolute values of the time lag are different so that NM variability lags by about 8 BRs on average (varying between 3-13 BR) behind SSN, but formally leads the OSF by about 5 BRs on average (-2 through 9 BR). While the low- and moderate-rigidity NM records depict nearly identical time lags after 1970, the high-rigidity NM records tend to yield slightly longer delays, which is counterintuitive as the propagation time of more energetic cosmic rays is expected to be shorter (Alanko-Huotari et al., 2007). However, the difference is insignificant most of the time between 1970 and 2005 and may be related to the smaller magnitude of variability for high-rigidity NM data and, thus, to larger uncertainties of the time-lag determination. Thus, the data does not allow us to claim any significant rigidity dependence of the time lag, which is consistent with earlier estimates by other methods (Ross and Chaplin, 2019).

Meanwhile, the time lag between the OSF and SSN is almost constant at about 11-12 BRs (SSN leads) with a small increase up to 14 BRs for the last solar cycle. It does not depict any significant cyclic variations.

## 4. Discussion and Conclusion

In this work, we have calculated the time lag between SSN, the OSF, and GCR intensity using the method of wavelet coherence, newly applied to this type of analysis. It was found that GCR variability lags behind SSN by about 8 BR on average and exhibits a clear 22-year cycle, so that the time lag is short (about 3 BRs) at around early A+ epochs and long (about 12 BRs) at around early A- epochs. This strong 22-year cycle is in agreement with earlier findings and theoretical models that predict faster/slower inward heliospheric transport of GCRs during A+/A- polarity epochs because of the large-scale drifts of positively charged particles in the heliosphere. On the one hand, the standard theory predicts that the inward diffusion propagation time depends on the GCR energy so that the transport time, and thus the time lag, is shorter for higher-energy particles (e.g. Gervasi et al., 1999). However, on the other hand, we did not find such a dependence in the analyzed data (cf. Ross and Chaplin,



Figure 10 The same as Figure 7 but for the OSF vs. moderate-rigidity NM time series.

2019), suggesting that the lag is largely driven by the drifts rather than by the straightforward diffusion transport. This provides an important observational constraint for the heliospheric modulation theory of cosmic rays.

It is also found that the OSF lags the SSN by about one year (about 12 BRs) without any observable variability, except for a small increase of the delay to 14 BRs during Solar Cycle 24. This could be a feature of the weak solar cycle or just a spurious coincidence – the data does not allow us to distinguish this. The 1-year time lag of the OSF is formed by a combination of the nearly instantly changing fast OSF component (time lag of a few BRs), probably defined by the opening time scale for the CME-driven component, and a smooth and delayed slow component with the time lag of about 30 BRs, likely related to the meridional transport of the surface magnetic field to the poles.

The time difference between the GCR variability and the OSF depicts a pattern very similar to that for GCRs vs. SSN, namely a clear 22-year cycle, but GCRs are slightly ahead of the OSF variations – by 3 BRs on average and varying between no delay and 10 BRs.



**Figure 11** Time lag between SSN, the OSF, and the NM data series, computed using the wavelet-coherence method. (a): Three NM datasets (low-, moderate- and high-rigidity groups indicated by green, blue and red colors, respectively) vs. SSN. (b): The same three NM datasets as in panel a vs. the OSF. (c): The OSF vs. SSN.

Although this excludes that OSF directly drives GCRs, GCR variability is well coherent also with the OSF with little time difference, confirming that the OSF is a good proxy of cosmic-ray variability at the 11-year time scale, as proposed earlier (e.g. Usoskin et al., 2004; Wu et al., 2018; Usoskin et al., 2021).

Summarizing, we have analyzed, using the method of the wavelet coherence, the pairwise time lags between three global solar and heliospheric indices: sunspot numbers, representing the solar surface magnetic activity, the open solar flux, representing the heliospheric magnetic field, and the cosmic-ray intensity near Earth. All the three indices appear highly coherent at timescales longer than a few years with the persistent high coherence at the timescale of the 11-year solar cycle. The GCR variability is delayed with respect to the inverted SSN by about 8 BRs on average, but the delay varies greatly with the 22-year cycle, being shorter/longer around positive A+/negative A- solar polarity epochs, respectively. GCRs covary nearly in antiphase with the OSF, also depicting a strong 22-year cycle in the delay, confirming that the OSF is a good proxy for the heliospheric modulation of GCRs.

Acknowledgments This work was partly supported by the Academy of Finland (projects ESPERA No. 321882 and QUASARE No. 330064), University of Oulu (Project SARPEDON), and the Russian Science Foundation (RSF Project No. 20-72-10170).

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

Data Availability All data used here are publicly available elsewhere, as described in Section 2.

#### Declarations

Disclosure of Potential Conflicts of Interest The authors declare that they have 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

- Adriani, O., Barbarino, G.C., Bazilevskaya, G.A., Bellotti, R., Boezio, M., Bogomolov, E.A., Bongi, M., Bonvicini, V., Borisov, S., Bottai, S., Bruno, A., Cafagna, F., Campana, D., Carbone, R., Carlson, P., Casolino, M., Castellini, G., De Pascale, M.P., De Santis, C., De Simone, N., Di Felice, V., Formato, V., Galper, A.M., Grishantseva, L., Karelin, A.V., Koldashov, S.V., Koldobskiy, S., Krutkov, S.Y., Kvashnin, A.N., Leonov, A., Malakhov, V., Marcelli, L., Mayorov, A.G., Menn, W., Mikhailov, V.V., Mocchiutti, E., Monaco, A., Mori, N., Nikonov, N., Osteria, G., Palma, F., Papini, P., Pearce, M., Picozza, P., Pizzolotto, C., Ricci, M., Ricciarini, S.B., Rossetto, L., Sarkar, R., Simon, M., Sparvoli, R., Spillantini, P., Stozhkov, Y.I., Vacchi, A., Vannuccini, E., Vasilyev, G., Voronov, S.A., Yurkin, Y.T., Wu, J., Zampa, G., Zampa, N., Zverev, V.G., Potgieter, M.S., Vos, E.E.: 2013, Time dependence of the proton flux measured by PAMELA during the 2006 July-2009 December solar minimum. *Astrophys. J.* **765**, 91. DOI.
- Aguilar, M., Ali Cavasonza, L., Alpat, B., Ambrosi, G., Arruda, L., Attig, N., et al.: 2018, Observation of fine time structures in the cosmic proton and helium fluxes with the alpha magnetic spectrometer on the international space station. *Phys. Rev. Lett.* **121**, 051101. DOI.
- Alanko-Huotari, K., Usoskin, I.G., Mursula, K., Kovaltsov, G.A.: 2007, Stochastic simulation of cosmic ray modulation including a wavy heliospheric current sheet. J. Geophys. Res. 112, A08101. DOI.
- Asvestari, E., Gil, A., Kovaltsov, G.A., Usoskin, I.G.: 2017, Neutron monitors and cosmogenic isotopes as cosmic ray energy-integration detectors: effective yield functions, effective energy, and its dependence on the local interstellar spectrum. J. Geophys. Res. 122, 9790. DOI.
- Belov, A.: 2000, Large scale modulation: view from the Earth. Space Sci. Rev. 93, 79. DOI.
- Boschini, M.J., Della Torre, S., Gervasi, M., La Vacca, G., Rancoita, P.G.: 2018, Propagation of cosmic rays in heliosphere: the HELMOD model. Adv. Space Res. 62, 2859. DOI. arXiv.
- Clette, F., Berghmans, D., Vanlommel, P., Van der Linden, R.A.M., Koeckelenbergh, A., Wauters, L.: 2007, From the wolf number to the international sunspot index: 25 years of SIDC. *Adv. Space Res.* 40, 919. DOI.
- Clette, F., Lefèvre, L.: 2016, The new sunspot number: assembling all corrections. Solar Phys. 291, 2629. DOI.
- Clette, F., Svalgaard, L., Vaquero, J., Cliver, E.: 2014, Revisiting the sunspot number: a 400-year perspective on the solar cycle. *Space Sci. Rev.* 186, 35. DOI.
- Cooke, D., Humble, J., Shea, M., Smart, D., Lund, N., Rasmussen, I., Byrnak, B., Goret, P., Petrou, N.: 1991, On cosmic-ray cut-off terminology. *Nuovo Cimento C* 14, 213. DOI.
- Dorman, I.V., Dorman, L.I.: 1967, Solar wind properties obtained from the study of the 11-year cosmic-ray cycle, 1. J. Geophys. Res. 72, 1513. DOI.
- Dorman, L.: 2004, Cosmic Rays in the Earth's Atmosphere and Underground, Kluwer Academic, Dordrecht.
- Forbush, S.E.: 1958, Cosmic-ray intensity variations during two solar cycles. J. Geophys. Res. 63, 651. DOI.
- Gervasi, M., Rancoita, P.G., Usoskin, I.G., Kovaltsov, G.A.: 1999, Monte-Carlo approach to galactic cosmic ray propagation in the heliosphere. Nucl. Phys. B, Proc. Suppl. 78, 26. DOI.
- Gizon, L., Cameron, R., Pourabdian, M., Liang, Z., Fournier, D., Birch, A., Hanson, C.: 2020, Meridional flow in the Sun's convection zone is a single cell in each hemisphere. *Science* 368, 1469. DOI.
- Grinsted, A., Moore, J.C., Jevrejeva, S.: 2004, Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* 11, 561. DOI.
- Hale, G.E., Ellerman, F., Nicholson, S.B., Joy, A.H.: 1919, The magnetic polarity of Sun-Spots. Astrophys. J. 49, 153. DOI.
- Inceoglu, F., Knudsen, M.F., Karoff, C., Olsen, J.: 2014, Modeling the relationship between neutron counting rates and sunspot numbers using the hysteresis effect. *Solar Phys.* 289, 1387. DOI.
- Iskra, K., Siluszyk, M., Alania, M., Wozniak, W.: 2019, Experimental investigation of the delay time in galactic cosmic ray flux in different epochs of solar magnetic cycles: 1959–2014. Solar Phys. 294, 115. DOI.

- Jokipii, J.R., Levy, E.H.: 1977, Effects of particle drifts on the solar modulation of galactic cosmic rays. Astrophys. J. Lett. 213, L85. DOI.
- Krainev, M., Bazilevskaya, G., Kalinin, M., Svirzhevskaya, A., Svirzhevsky, N.: 2016, GCR intensity during the sunspot maximum phase and the inversion of the heliospheric magnetic field. In: *Proceedings of the* 34th International Cosmic Ray Conference – PoS(ICRC2015) 236, 081. DOI.
- Krivova, N., Solanki, S., Hofer, B., Wu, C., Usoskin, I., Cameron, R.: 2021, Modelling the evolution of the Sun's open and total magnetic flux. *Astron. Astrophys.* 650, A70. DOI.
- Lockwood, M.: 2013, Reconstruction and prediction of variations in the open solar magnetic flux and interplanetary conditions. *Living Rev. Solar Phys.* 10, 4. DOI.
- Martucci, M., Munini, R., Boezio, M., Di Felice, V., Adriani, O., Barbarino, G.C., Bazilevskaya, G.A., Bellotti, R., Bongi, M., Bonvicini, V., Bottai, S., Bruno, A., Cafagna, F., Campana, D., Carlson, P., Casolino, M., Castellini, G., De Santis, C., Galper, A.M., Karelin, A.V., Koldashov, S.V., Koldobskiy, S., Krutkov, S.Y., Kvashnin, A.N., Leonov, A., Malakhov, V., Marcelli, L., Marcelli, N., Mayorov, A.G., Menn, W., Mergè, M., Mikhailov, V.V., Mocchiutti, E., Monaco, A., Mori, N., Osteria, G., Panico, B., Papini, P., Pearce, M., Picozza, P., Ricci, M., Ricciarini, S.B., Simon, M., Sparvoli, R., Spillantini, P., Stozhkov, Y.I., Vacchi, A., Vannuccini, E., Vasilyev, G., Voronov, S.A., Yurkin, Y.T., Zampa, G., Zampa, N., Potgieter, M.S., Raath, J.L.: 2018, Proton fluxes measured by the PAMELA experiment from the minimum to the maximum solar activity for solar cycle 24. Astrophys. J. Lett. 854, L2. DOI. arXiv.
- Mavromichalaki, H., Petropoulos, B.: 1984, Time-lag of cosmic-ray intensity. Astrophys. Space Sci. 106, 61. DOI.
- Mishra, V., Mishra, A.: 2018, Long-term modulation of cosmic-ray intensity and correlation analysis using solar and heliospheric parameters. *Solar Phys.* 293, 141. DOI.
- Nagashima, K., Morishita, I.: 1979, Twenty-two year modulation of cosmic rays associated with polarity reversal of polar magnetic field of the Sun. *Internat. Cosmic Ray Conf.*, 325.
- Owens, M.J., Crooker, N.U.: 2006, Coronal mass ejections and magnetic flux buildup in the heliosphere. J. Geophys. Res. 111, A10104. DOI.
- Owens, M.J., Lockwood, M., Riley, P., Linker, J.: 2017, Sunward strahl: a method to unambiguously determine open solar flux from in situ spacecraft measurements using suprathermal electron data. J. Geophys. Res. 122, 10980. DOI.
- Owens, M.J., Usoskin, I., Lockwood, M.: 2012, Heliospheric modulation of galactic cosmic rays during grand solar minima: past and future variations. *Geophys. Res. Lett.* 39, L19102. DOI.
- Potgieter, M.: 2013, Solar modulation of cosmic rays. *Living Rev. Solar Phys.* 10, 3. DOI. arXiv.
- Ross, E., Chaplin, W.J.: 2019, The behaviour of galactic cosmic-ray intensity during solar activity cycle 24. Solar Phys. 294, 8. DOI.
- Schwabe, H.: 1844, Sonnenbeobachtungen im Jahre 1843. Von Herrn Hofrath Schwabe in Dessau. Astron. Nachr. 21, 233. DOI.
- Simpson, J.A.: 2000, The cosmic ray nucleonic component: the invention and scientific uses of the neutron monitor. Space Sci. Rev. 93, 11. DOI.
- Stansby, D., Green, L., van Driel-Gesztelyi, L., Horbury, T.: 2021, Active region contributions to the solar wind over multiple solar cycles. *Solar Phys.* 296, 116. DOI. arXiv.
- Strauss, R.D., Potgieter, M.S., Kopp, A., Büsching, I.: 2011, On the propagation times and energy losses of cosmic rays in the heliosphere. J. Geophys. Res. 116, A12105. DOI.
- Thomas, S.R., Owens, M.J., Lockwood, M.: 2014, The 22-year hale cycle in cosmic ray flux evidence for direct heliospheric modulation. *Solar Phys.* 289, 407. DOI. arXiv.
- Tomassetti, N., Orcinha, M., Barão, F., Bertucci, B.: 2017, Evidence for a time lag in solar modulation of galactic cosmic rays. Astrophys. J. Lett. 849, L32. DOI. arXiv.
- Usoskin, I., Mursula, K., Solanki, S., Schüssler, M., Alanko, K.: 2004, Reconstruction of solar activity for the last millennium using <sup>10</sup>Be data. Astron. Astrophys. 413, 745. DOI. arXiv.
- Usoskin, I.G., Kananen, H., Mursula, K., Tanskanen, P., Kovaltsov, G.A.: 1998, Correlative study of solar activity and cosmic ray intensity. J. Geophys. Res. 103, 9567. DOI.
- Usoskin, I.G., Koldobskiy, S., Kovaltsov, G.A., Gil, A., Usoskina, I., Willamo, T., Ibragimov, A.: 2020, Revised GLE database: fluencies of solar energetic particles as measured by the neutron-monitor network since 1956. Astron. Astrophys. 640, A17. DOI.
- Usoskin, I.G., Kovaltsov, G.A., Kananen, H., Mursula, K., Tanskanen, P.J.: 1997, Phase evolution of solar activity and cosmic-ray variation cycles. *Solar Phys.* 170, 447. DOI.
- Usoskin, I.G., Mursula, K., Kananen, H., Kovaltsov, G.A.: 2001, Dependence of cosmic rays on solar activity for odd and even solar cycles. Adv. Space Res. 27, 571. DOI.
- Usoskin, I.G., Solanki, S.K., Krivova, N., Hofer, B., Kovaltsov, G.A., Wacker, L., Brehm, N., Kromer, B.: 2021, Solar cyclic activity over the last millennium reconstructed from annual <sup>14</sup>C data. Astron. Astrophys. 649, A141. DOI.

Väisänen, P., Usoskin, I., Mursula, K.: 2021, Seven decades of neutron monitors (1951–2019): overview and evaluation of data sources. J. Geophys. Res. 126, e2020JA028941. DOI.

- Webber, W.R., Lockwood, J.A.: 1988, Characteristics of the 22-year modulation of cosmic rays as seen by neutron monitors. J. Geophys. Res. 93, 8735. DOI.
- Wu, C.J., Usoskin, I.G., Krivova, N., Kovaltsov, G.A., Baroni, M., Bard, E., Solanki, S.K.: 2018, Solar activity over nine millennia: a consistent multi-proxy reconstruction. Astron. Astrophys. 615, A93. DOI. arXiv.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.