On the solar origin of in situ observed energetic protons

Rositsa Miteva
Space Research and Technology Institute - Bulgarian Academy of Sciences, BG-1113, Sofia rmiteva@space.bas.bg

(Submitted on 11.11.2018. Accepted on 29.11.2018)

Abstract. The aim of this study is to evaluate the differences in the reported solar origin identifications to the same in situ observed energetic proton event and to estimate the possible consequences. In order to assess such issue in a quantitative manner, both, the solar origin estimation (e.g., the observer selection of the solar flare and coronal mass ejection to each particle event) and the level of certainty on this procedure are used. Pearson correlation coefficients are calculated between the parameters of the protons and their solar origin. For the purpose of this study, the 20 MeV SOHO-ERNE proton catalog is finalized and used as a reference proton event list.

Key words: solar energetic protons, solar flares, coronal mass ejections

Introduction

The dilemma in the solar energetic particle (SEP) research gravitates around the still unresolved issue on their solar origin: is the solar flare (SF) or coronal mass ejection (CME) the sole/primary driver (Bazilevskaya et al. 2017). From one side, whether or not SFs and CMEs are able to accelerate particles is not scrutinized: both solar eruptive phenomena are known (directly, by their ability to generate electric fields or indirectly, through the generation of shock waves) to produce non-thermal particle distributions from the background thermal particle ensemble in the solar corona (Aschwanden 2002; Krauss-Varban 2010; Klein and Dalla 2017). Electromagnetic (EM) signatures, from gamma to radio emissions, support both the flare-driven magnetic reconnection and the CME-driven shock scenarios (Nindos et al. 2008; Pick and Vilmer 2008). Nevertheless, when the escaping protons and electrons are considered (or at least the portions that could be observed with satellites near Earth, at 1 AU, or, occasionally, out of the ecliptic place), the debate on the SFs vs. CMEs as the SEP-solar origin is far from being resolved.

There are several main reasons. The proton, electron or heavy ions intensities (termed SEP events) measured in situ, at energies ranging from keV to MeV, are modulated due to the following factors: variable efficiency and duration of the particle acceleration process; possibility for trapping and delayed injection of particles in the solar corona due to the event-specific magnetic field topology; scattering or/and re-acceleration during their subsequent interplanetary (IP) transport; loss of magnetic connection between the field line guiding the SEPs and the detector in order to be registered. Thus the SEP profiles observed at Earth are indicative only to a certain degree of the multitude of processes that take part. Numerous approximations are being imposed in order to tackle the problem, e.g., single, short in time period of particle injection; scatter-free propagation; simultaneous accelerations of different particle species and at different energies. This is mainly due to the fact that the time of injection, the duration of

transport, the amount of loss or re-acceleration of the particle flux are unknown parameters and thus a number of assumptions are implied. Such series of oversimplifications could well contribute to the ambiguous results on the solar origin identification of SEP events. A recent summary on the limitations in the SEP studies is given in (Miteva et al. 2018a).

Contributing aspects to the uncertainties are the limited set of parameters involved, lack of their accurate measurements, reduction of the dimensions of the problem or neglecting the evolution of the system. From one side, the particles are usually represented by the peak value of their intensity−time profile or their fluence profile (where fluence is the time integrated proton intensity). For the latter, the time of integration varies in different reports: onset-to-peak or onset-to-end. The end of the SEP event is in general difficult to evaluate, especially when there is a sequence of several particle events in close succession and the end of one SEP is masked by the onset of the next. In addition, the time markers (especially the particle onset) are also subject to adopted definition by the authors (see discussion in Miteva et al. 2017c; Miteva et al. 2018a). Usually in the literature, the peak value of the SEP intensity is adopted as the representative value for the strength of the particle flux.

Solar flare strength is commonly represented by the flux in GOES soft X-ray (SXR) 1−8 Å channel, known as flare class. The EM emission of solar flares, however, cover almost the entire range, between gamma to radio emissions (Fletcher et al. 2011). Although, radio signatures at various wavelengths are employed in the SEP studies (Miteva et al. 2017d), the lack of long uninterrupted series of solar hard X-ray (HXR), gamma, visible light data is the reason for these wavelengths to be dropped form large statistical studies. Interestingly, H-alpha and extreme ultraviolet (EUV) data are neither explored for large number of events. Thus, the multi-energy, multi-scale and time-dependent flare process is reduced to the peak value of the thermal plasma emission signature.

Similar is the issue with the CME properties (Webb and Howard 2012). As a representative value for the CME kinematics, almost exclusively, is taken the projected on the sky linear speed. The measured angular width is adopted to represent the spatial extent of the potentially driven shock wave. Both parameters are evaluated in a specific moment in time. Again, the temporal development, acceleration, mass are not considered in joint studies with SEP events. Recent efforts have been made to use the deprojected, 3D speed instead (Park et al. 2017). The use of different catalogs for the CME projected speed is considered permissible for statistical works (Richardson and Cane 2015).

In the present study, an alternative approach is followed. The main aim here is to compare the solar origin identifications made by independent teams and to quantify the reported differences. For this task, the quality of solar origin association is evaluated using a set of standard guidelines. The purpose of such analysis is to address the issue and effects of the observer subjectivity that cannot be avoided during the process of SEP identification and the selection of their solar origin.

1 Data analysis and event categories

For the present study, current science level data from SOHO/ERNE High Energy Detector (HED) is used as provided by the instrument team. The description of the instrument is presented in (Torsti et al. 1995) and the 55−80 MeV energy proton data is already been exploited and discussed in Vainio et al. (2013), Paassilta et al. (2017), respectively. The HED covers a broader energy range, 14−131 MeV, separated into 10 energy channels, whereas for this work the second channel is selected. At energies lower than about 20 MeV, the proton intensity profiles could be well contaminated from local IP sources (shocks, co-rotating/streaming interaction regions). At larger energies, this contribution is negligible if any (and also the observations show that the shock signatures are more prominent at lower energies), however the number of high energy protons is naturally fewer than the lower energy sample. Thus, as a compromise, the 17−22 MeV energy channel is selected to search and identify the SOHO/ERNE events for the present study. Henceforth, this energy range will be regarded as the ∼20 MeV reference channel.

The procedure for the proton identification is summarized in Miteva (2017b). For the present analysis a 5-min averaged proton data are utilized and 586 events in the period 1996−2016 are isolated. The onset, peak times and the peak of the background subtracted proton intensity have been identified. The distributions of proton number, intensity and rise time can be seen in Miteva (2017b). There, the yearly distribution of all identified proton events at SOHO/ERNE ∼20 MeV energy range is shown, using 6-month binning period. The solar cycle (SC) behavior is clearly evident also in terms of proton events, with a clear decrease (in peak value and number of events) during the ongoing SC24 with respect to SC23. Occasionally, short-time increases in the number of proton events are noticeable during the second half of 2002 and during 2005, related to the increased activity periods during the declining phase of the solar cycle.

The key part of any SEP study is the association of flare and CME to each proton events. The adopted procedure by different authors for solar origin identification is similar. As a rule, the strongest in SXR class flare and the fastest and widest CME are selected as the probable particle accelerator. Standard catalogs are used, both for the SFs (e.g., GOES listings[1]) and for the CMEs (e.g., SOHO LASCO CME catalog[2]). Occasionally, the (strength of the) radio emission signatures are used to time the escaping electrons and subsequently to select specific flare−CME pair over another. In spite of the care taken during the solar origin association, subjectivity issues are intrinsic to the process. Thus, the level of certainty of flare/CME identification can be evaluated at least in a qualitative manner. Here, the analysis follows the procedure proposed by Miteva (2017a) where three level of certainty for the solar origin association were specified, termed high, average and low. Here, the following criteria are imposed prior the SEP onset at 1 AU.

[1] https://hesperia.gsfc.nasa.gov/goes/goes_event_listings/
[2] https://cdaw.gsfc.nasa.gov/CME_list/

With a *high* certainty are denoted the cases when solar origin identification is straightforward: a single flare−CME pair exists prior the solar proton onset as measured at 1 AU (and any preceding pairs are about 2 hours earlier). An *average* certainty is used for those cases when multiple flare−CME pairs are observed before the proton onset, however the strongest one is adopted as solar origin. For the remaining cases, a *low* level of certainty has to be set due to the relatively large amount of subjectivity present and when a weaker candidate is selected. In this work we apply this classification separately to the flare and CME identification as the proton origin.

2 Results

There are 586 proton events identified in the period of interest (1996−2016) using 5-min averaged data. The list of the new catalog is given as online material. The identified SOHO/ERNE peak proton intensities range from 0.000185 to about 19.5 protons/(cm^2 s sr MeV), with mean/median value for the entire sample of 0.526/0.013, respectively. Then the solar events could be divided according to the median value for the proton sample, the strong proton events defined to have values larger or equal than the median value for the sample, wheres the weak proton events have intensities less than the median for the sample. SOHO/ERNE instrument is known to saturate at large proton intensities, which cannot be accounted for using the provided proton data. In order to minimize the possible instrumental effects and avoid such a bias from the results, the calculations between the proton intensity and the solar origin parameters are done using proton data from an alternative proton instrument, Wind/EPACT (Miteva et al. 2018b), with 429 proton events in the same time period. The results are summarized in the Appendix. Unless otherwise specified, the SOHO/ERNE proton sample is considered here as the (primary) proton event sample.

With respect to the solar origin identification, the proton events with solar origin at western helio-longitudes constitute 63% (370/586) of the entire sample, whereas those with eastern origin − 22% (129/586). The remaining number of events (87/586) are with uncertain solar origin association (15% of the entire sample, respectively). If we exclude the uncertain cases, the western-to-eastern ratio of solar origin of SEP events over the entire period of interest is 74-to-26%. All results based on this new proton catalog are given on the top row of the respective tables, denoted by 'any' or 'this study'.

2.1 Properties of the flare and CME events in the different certainty categories

For the study here, the certainty criteria is imposed separately on the identification of flares and CMEs. Based on the different levels of certainty evaluated, the yearly distribution of the proton-related solar origin (uncertain cases are dropped) is produced and shown in Fig. 1. The color-code there denotes the certainty of the solar origin association. Namely, black

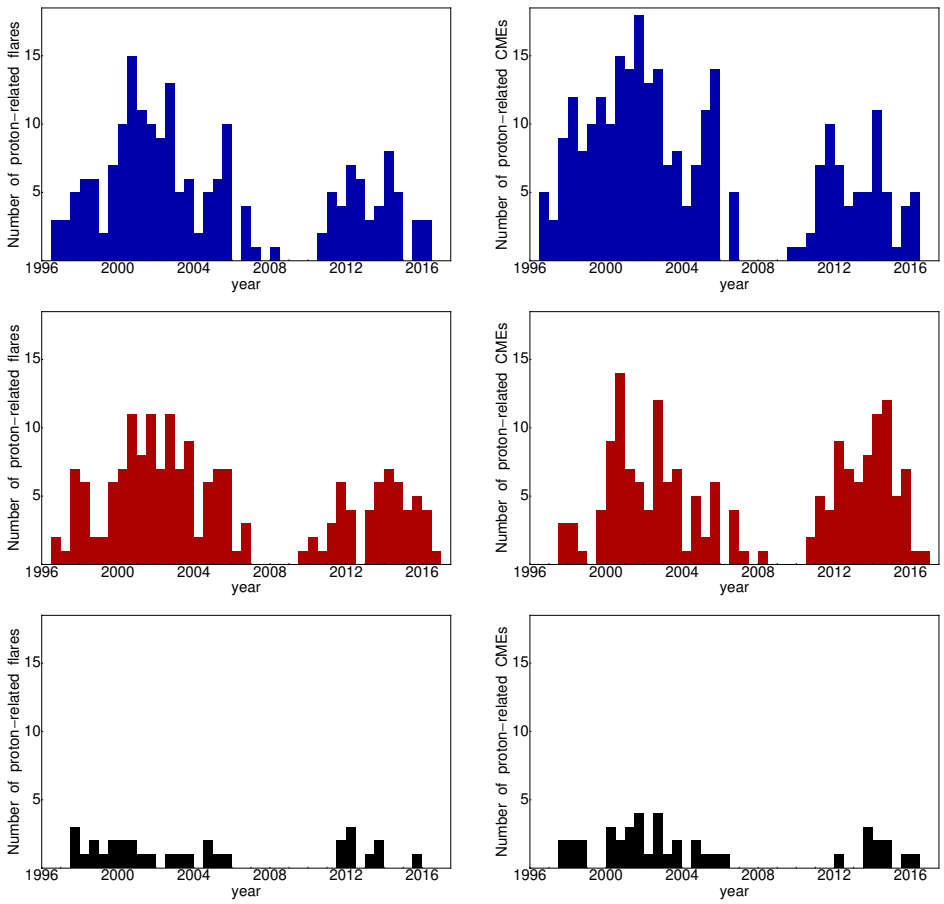

Fig. 1. Histograms of the proton-related flares (on the left) and CMEs (on the right) with identified solar origin vs. year, using 6-month binning. The length of the color bar denotes the number of events in each category. Color code used: blue – high certainty (upper row); red – average certainty (middle row); black – low certainty for solar origin association (lower row), respectively.

color is used for the events with the lowest level of certainty, red – for the average level and blue is used for the high level of certainty. Overall, the (sum of the) distributions are consistent with the shape of the distribution for all proton events (compare with Miteva 2017b). It is expected that during the peak of the SC, the larger number of solar eruptions will make the solar origin association more difficult for the observer and thus the amount of subjectivity will rise. It is found that the distribution of the low certainty events is relatively flat, although with some clustering over the SC maxima, whereas the distributions of the average and high certainty events has a clear SC trend.

The so-identified proton-related flares and CMEs are now investigated

Table 1. Table of the median values of the proton-related SF class and CME speed in the different certainty categories. The number of events used in each calculation is given in brackets.

Level of	Flares			CMEs		
certainty	All	Western	Eastern	All	Western	Eastern
	1996−2016					
any	M1.8 (396)	M1.6 (287)	M2.2 (98)	916 (481)	886 (360)	1119 (90)
high	M1.6 (190)	M1.7 (144)	M1.5 (42)	920 (266)	916 (214)	1016 (40)
high+average	M1.9 (365)	M1.8 (270)	M2.2 (85)	944 (440)	911 (333)	1120 (80)
average	M2.2 (175)	M2.0 (126)	M2.9 (43)	966 (174)	903 (119)	1187 (40)
low	C4.4 (31)	C2.6 (17)	M2.6 (13)	669 (41)	618 (27)	876 (10)
	SC23': 11/1996−10/2004					
any	M1.7 (235)	M1.7 (174)	M2.2 (51)	894 (272)	899 (213)	1056 (46)
high	M1.6 (116)	M1.6 (88)	M2.9 (24)	890 (163)	890 (135)	1133 (23)
high+average	M1.9 (215)	M1.9 (164)	M2.5 (42)	909 (245)	903 (197)	1095 (38)
average	M2.3 (99)	M2.6 (76)	M2.3 (18)	957 (82)	957 (62)	1071 (15)
low	C4.7 (20)	C3.5 (10)	C1.3 (9)	669 (27)	588 (16)	757 (8)
	SC24': 01/2009−12/2016					
any	M1.8 (113)	M1.3 (80)	M2.0 (32)	928 (156)	818 (109)	1090 (32)
high	M1.9 (50)	M1.9 (38)	M1.5 (12)	949 (68)	967 (54)	963 (10)
high+average	M1.8 (104)	M1.3 (74)	M1.9 (29)	949 (146)	885 (101)	1063 (31)
average	M1.7 (54)	M1.2 (36)	M1.9 (17)	940 (78)	772 (47)	1200 (21)
low	C4.4 (9)	C1.2 (6)	M2.6 (3)	732 (10)	732 (8)	1825 (1)

in details as a function of the level of certainty (in terms of median values of the SF class and CME projected speed). The results are summarized in Table 1 for three periods of interest, the entire period, and the first eight years from SCs 23 and 24. There is a slight tendency of the average certainty events to consists of the strongest flares in SXR class and fastest CMEs in projected linear speed, together with the high and high+average categories. Overall, the low certainty events consists of the weakest flares and slowest CMEs in any of the considered time periods. The mean values for the SF class and CME speed for the events with eastern origin are of larger values compared to the western cases, respectively. This is consistent with the selection effect the solar origin at eastern helio-longitudes to be of larger magnitudes in order for the generated energetic proton events to be finally observed at 1 AU.

2.2 Certainty issues on the solar origin identification

Following the procedure in Miteva (2017a), the scatter-plots between the peak proton intensity and the SF class (Fig. 2) or CME projected speed (Fig. 3) are produced, shown here for the entire proton sample. The following notation is used for the level of certainty when identifying the SEP-solar origin: black denote the events with low level of certainty, red − average certainty and blue color for the high level of certainty.

Upon visual assessment, there seems to be a large overlap of the three categories of events on the scatter-plot with the SF class (Fig. 2). In the plot with the CME speed (Fig. 3), the overlap is better noticed for the events with high and average certainty categories, whereas the category of events with low certainty are overrepresented towards slow CME speeds.

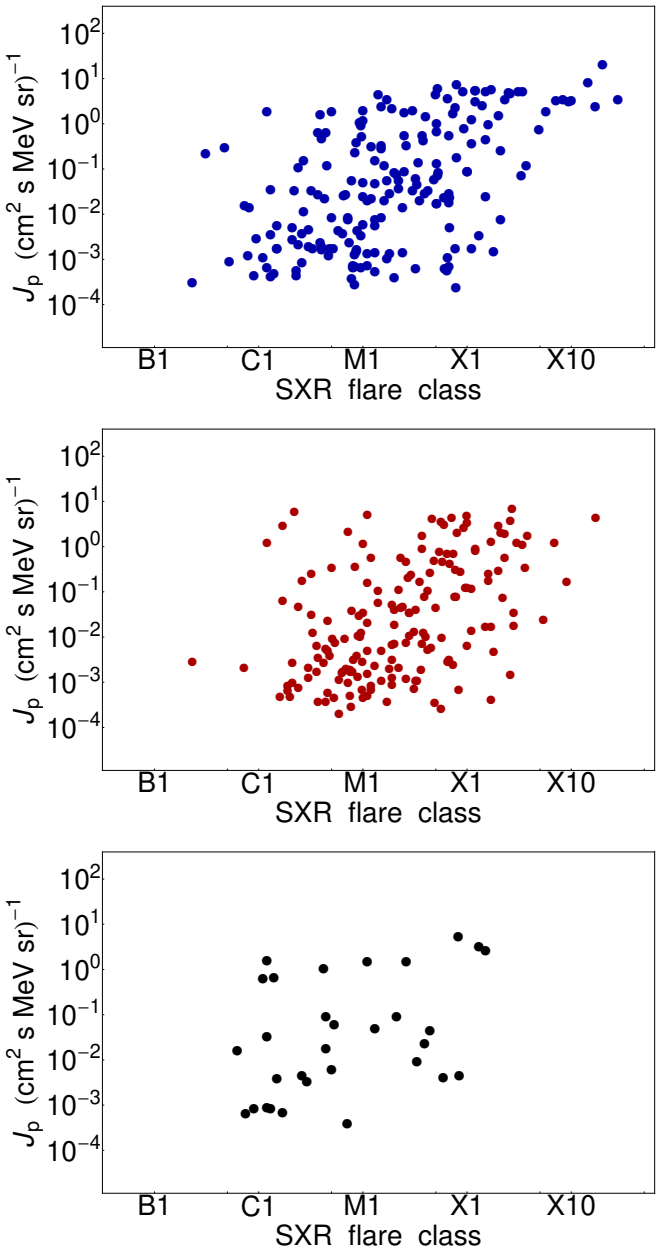

Fig. 2. Scatter plots of the SOHO/ERNE ∼20 MeV proton peak intensity and the proton-related flare class. Blue circles denote high level of certainty (upper row), red color − average (middle row) and black circles − low level of certainty (lower row).

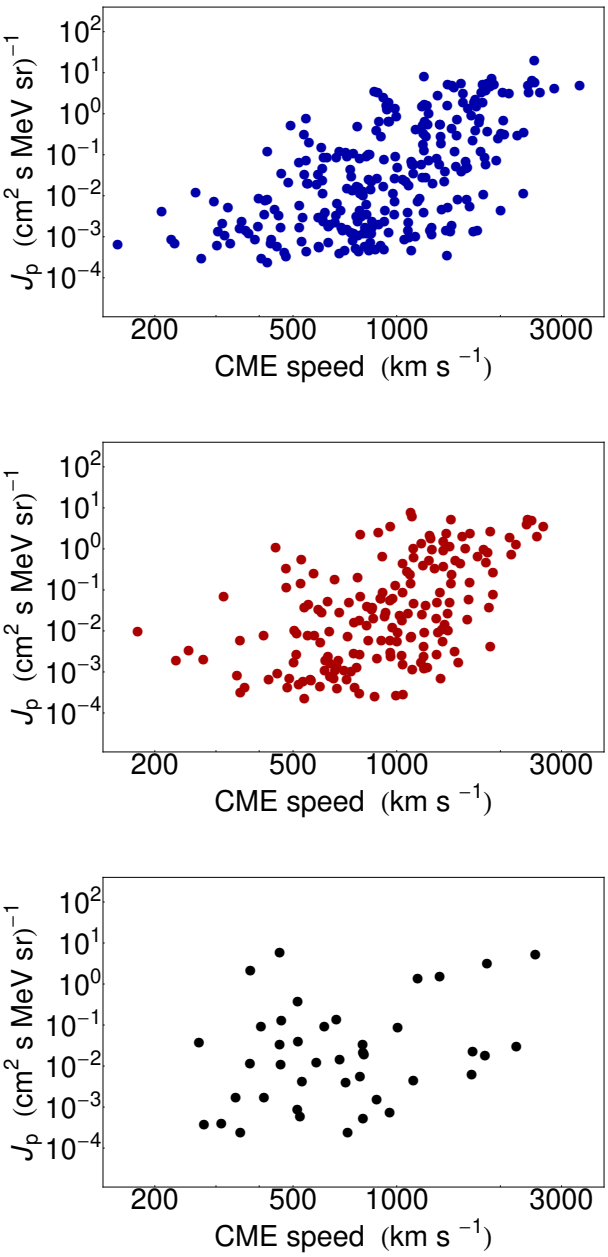

Fig. 3. Scatter plots of the SOHO/ERNE ∼20 MeV proton peak intensity and the proton-related CME speed. Symbols as in Fig. 2.

Quantitatively, the scatter can be evaluated in terms of linear (Pearson) correlation coefficients. The coefficients are calculated between $\log_{10}$ of the peak proton intensity and $\log_{10}$ of the reported by GOES SXR flare class or $\log_{10}$ of the reported by LASCO CME linear speed. The results are summarized in Table 2 shown for three different periods of interest: the entire sample 1996−2016 and for the first eight years from SC23 and SC24, respectively. In addition the longitudinal dependence is evaluated, by calculating Pearson coefficients for all events, and for those with origin in the West or East. Finally, each evaluation is done for a sub-sample with specific level of certainty: All events (no level of certainty considered); only events with high, average or low certainty; and combined sample of events with high and average certainty level.

content...

The following trends are noticed. Overall, for the period 1996−2016, the correlations decrease when the level of certainty is the lowest (see Table 2). The decrease is however within the uncertainties. When only the sample of low certainty is selected, the correlation coefficients, both with SFs and with CMEs, are lower compared to the coefficients for all events (around 0.5). This difference is either not or just marginally statistically significant, due to the large uncertainty of the low certainty event group (note the low sampling as given in brackets). The same trend is obtained for the samples of Western origin events. For the Eastern events these trends are either not present (in the case with CME speed) or the opposite behavior is noticed (for the flare class). However due to the large uncertainty (mostly due to the fewer events in the specific sample) the differences are not always significant. For statistical significance we use the bootstrapping method (Wall and Jenkins 2003), as implemented by Miteva et al. (2013).

The highest correlation coefficient (both with SFs and CMEs) is often obtained when only the sample of high certainty is considered, however with some variation between the categories of high, average and high+average. Overall, the decrease of correlations with the decrease of the certainty to average is small, if any, since the differences are often within the statistical uncertainties. The exact values with their error margins can be inspected from Table 2. When the correlations between protons and SFs vs. CMEs are compared, the correlations with the CME speed are in general higher, however either not statistically significant or only marginally larger, depending on the compared pair of samples.

The SC differences are also evaluated for different levels of certainty, based on time sample of eight years from the start of SC23 and SC24, and the details can be inspected from the respective columns in Table 2. The results in the first eight years of either SC tend to be consistent with the overall behavior.

2.3 Subjectivity issues on the solar origin identification

In order to compare the solar origin identifications from various sources, one needs to answer the question on the identity of the same proton event phenomena observed by various spacecraft. This is a relatively straightforward task when the proton event is an isolated case. When several particle increases are present, however, the one-to-one correspondence becomes

Table 2. Table with $\log_{10}-\log_{10}$ correlation coefficients and their uncertainties calculated between the SOHO/ERNE peak proton intensity with SF class or CME speed, respectively, in the period 1996−2006, and the first eight years from SC23 and SC24, respectively. In brackets are given the number of events over which the correlation is performed (u: uncertain).

Level of certainty	Pearson correlations ($\log_{10}-\log_{10}$) between peak proton intensity and Flare class			CME speed		
	All	Western	Eastern	All	Western	Eastern
	1996−2016					
any	0.48±0.04 (396)	0.51±0.05 (287)	0.43±0.09 (98)	0.54±0.03 (481)	0.58±0.03 (360)	0.49±0.06 (120)
high	0.53±0.06 (190)	0.59±0.06 (144)	0.34±0.15 (42)	0.59±0.04 (266)	0.70±0.04 (141)	0.46±0.11 (52)
high+average	0.50±0.04 (365)	0.53±0.05 (270)	0.43±0.09 (85)	0.57±0.03 (440)	0.66±0.03 (264)	0.51±0.07 (106)
average	0.46±0.07 (175)	0.46±0.08 (126)	0.51±0.11 (43)	0.52±0.05 (174)	0.59±0.06 (123)	0.55±0.08 (54)
low	0.41±0.17 (31)	0.12±0.23 (17)	0.58±0.23 (13)	0.32±0.16 (41)	0.21±0.23 (17)	0.48±0.27 (14)
	SC23': 11/1996−10/2004					
any	0.49±0.06 (235)	0.50±0.06 (174)	0.45±0.12 (51)	0.53±0.04 (272)	0.55±0.05 (213)	0.53±0.11 (58)
high	0.52±0.08 (116)	0.57±0.08 (88)	0.35±0.17 (24)	0.59±0.05 (163)	0.61±0.05 (135)	0.37±0.24 (21)
high+average	0.51±0.06 (215)	0.52±0.06 (164)	0.50±0.12 (42)	0.56±0.04 (245)	0.57±0.04 (197)	0.52±0.15 (37)
average	0.49±0.09 (99)	0.45±0.11 (76)	0.66±0.15 (18)	0.47±0.09 (73)	0.47±0.10 (62)	0.69±0.17 (16)
low	0.22±0.23 (20)	0.11±0.37 (10)	0.37±0.38 (9)	0.32±0.19 (27)	0.28±0.23 (16)	0.65±0.25 (9)
	SC24': 01/2009−12/2016					
any	0.40±0.08 (113)	0.51±0.10 (50)	0.45±0.14 (32)	0.55±0.05 (156)	0.61±0.06 (109)	0.47±0.11 (47)
high	0.51±0.09 (50)	0.56±0.10 (38)	0.52±0.18 (16)	0.60±0.07 (68)	0.74±0.07 (38)	−0.02±0.27 (14)
high+average	0.43±0.08 (104)	0.47±0.09 (74)	0.52±0.11 (37)	0.56±0.05 (146)	0.68±0.06 (74)	0.46±0.10 (45)
average	0.35±0.15 (54)	0.36±0.18 (36)	0.57±0.10 (21)	0.53±0.08 (78)	0.60±0.12 (36)	0.57±0.09 (31)
low	0.50±0.32 (9)	u (6)	u (4)	0.38±0.39 (10)	u (6)	u (2)

more difficult. Overall, the first increase is easier to identify, whereas the subsequent ones could be specified with respect to the level of their proton intensity. Irrespectively on the details of the analysis, a set of general criteria is followed here, as defined by Miteva et al. (2018a), namely the two proton onsets to be within one day and both peak intensities to be either smaller or larger compared to the median value for the specific sample. Then, it is accepted that these two reports relate to the same proton phenomena.

For the analysis in the present study, five different proton catalogs are utilized: SOHO/ERNE 20 MeV reference list (termed 'this study'), NOAA GOES proton event list[3] ('NOAA'), GOES proton list reported by Papaioannou et al. 2016 ('GOES'), the SOHO/ERNE 55−80 MeV from Paassilta et al. 2017 ('SOHO') and the Wind/EPACT event list from Miteva et al. 2018b ('Wind').

The reference proton list is compared to the remaining four databases in order to quantify the differences in terms of number of alternatively reported solar origin, SFs and CMEs. The results are reported in percentage, normalized to the number of events shared by each pair of catalogs. The cases with uncertain solar origin (flares and CMEs) in either of the compared pair of catalogs is also counted as a different association in order to estimate the upper limit of the differences. The analysis starts with a list of proton events identified to be the same phenomena but observed by both satellites. The outcome for the alternative solar origin as reported by different sources is as follows:

- This study vs. NOAA: When the solar origin is considered, 100 proton cases have been identified with flares and 94 are with CMEs. In the majority of the cases the solar origin association coincides, apart from 13 cases for the flares and 14 cases (15%) for the CMEs.
- This study vs. GOES: In this case the GOES catalog does not extend after March-2013, thus the reported values can be considered as lower limits. There are 137 flare and 142 CME associations for the common proton events. In 15/21 of the cases (11%/15%), the identified flares/CMEs are different with respect to the current 20 MeV proton event study, respectively.
- This study vs. SOHO: Out of all protons present in both listings (note the larger proton energy channel for the SOHO event list), 1/118 flares (less than 1%) and 6/157 CMEs (about 4%) have alternative entires.
- This study vs. Wind: With respect to these catalogs, no concurrence is reached for 8/238 cases for the flare (about 3%) and 11/283 for the CME associations (about 4%).

The differences range from just a few to about 15% depending on compared pair of catalogs and solar origin type. For all of the different cases, however, namely when alternative solar origin is identified, the proton intensity is above 0.01 in SOHO/ERNE units, i.e. mostly large proton events are involved. Since the number of alternative flares/CMEs depends on the compared pair of catalogs, the results for their helio-location range are calculated for a representative sample as described below.

[3] https://umbra.nascom.nasa.gov/SEP/

Table 3. Table of the Pearson correlation coefficients ($\log_{10}-\log_{10}$) between the SOHO/ERNE proton peak intensity and the flare and CME parameters, respectively, using different solar origin association.

Solar origin	Pearson correlations			
association	Flare class	Flare fluence	CME speed	CME kinetic energy
1996−2016				
this study	0.48±0.04 (395)	0.50±0.05 (373)	0.54±0.03 (480)	0.53±0.03 (411)
NOAA	0.29±0.09 (102)	0.16±0.10 (90)	0.39±0.07 (92)	0.45±0.10 (76)
GOES	0.32±0.07 (137)	0.19±0.09 (121)	0.42±0.07 (143)	0.38±0.09 (114)
SOHO	0.26±0.08 (121)	0.39±0.11 (113)	0.52±0.06 (158)	0.42±0.09 (131)
Wind	0.38±0.06 (237)	0.39±0.07 (220)	0.51±0.05 (282)	0.48±0.05 (236)
SC23': 11/1996−10/2004				
this study	0.49±0.06 (234)	0.48±0.07 (225)	0.53±0.04 (271)	0.53±0.05 (232)
NOAA	0.31±0.11 (62)	0.14±0.14 (53)	0.48±0.09 (56)	0.57±0.13 (41)
GOES	0.26±0.10 (96)	0.15±0.12 (84)	0.36±0.08 (94)	0.45±0.08 (75)
SOHO	0.26±0.11 (70)	0.30±0.14 (68)	0.58±0.07 (88)	0.51±0.10 (72)
Wind	0.34±0.08 (144)	0.34±0.10 (134)	0.53±0.06 (162)	0.56±0.06 (134)
SC24': 01/2009−12/2016				
this study	0.40±0.07 (113)	0.52±0.07 (106)	0.55±0.05 (156)	0.56±0.05 (137)
NOAA	0.30±0.10 (31)	0.23±0.13 (30)	0.26±0.14 (33)	0.31±0.14 (29)
GOES	0.36±0.15 (21)	0.20±0.19 (21)	0.43±0.17 (31)	0.50±0.13 (26)
SOHO	0.17±0.14 (39)	0.56±0.12 (36)	0.43±0.11 (54)	0.51±0.10 (46)
Wind	0.35±0.11 (69)	0.41±0.10 (65)	0.44±0.07 (96)	0.45±0.07 (83)

If one considers all proton events (from the table in the online material) for which alternative origin (either flare or/and CME) is proposed, a total of 60 cases are selected. For this sample (based on the proton/flare/CME values identified for the present catalog), the mean/median proton intensity is 0.273/0.44, the mean/median flare class is X1.4/M5.4 and the median CME speed is 1050/920. When origin location could be identified, the result is 30% (15/50) for eastern and 70% (35/50) for western longitude. Compared to the 75% western location for the entire sample, the eastern origin events are not significantly overrepresented.

In order to quantify the influence of alternative solar origin associations to the correlation coefficients, we use the SOHO/ERNE proton intensity and vary only the solar origin associations (SFs and CMEs) by using the same five difference sources as above: this study, NOAA, GOES, SOHO and Wind. Additionally to the SF class and CME speed, the correlation analysis is extended by adding the SF fluence and, for the first time, the CME kinetic energy as an alternative CME parameter. In Tables 3 and 4 are listed the correlation coefficients between SOHO/ERNE 20 MeV proton peak intensity and the parameters of the solar origin adopted from several different sources (catalogs), for the entire sample and for events from western helio-longitude, respectively. The values for SF class, CME speed and CME kinetic energy are adopted from their respective catalogs. For the SF fluence the following analysis is completed. Using the goes-routine in IDL/SolarSoft, the flare data (GOES 1−8 Å SXR channel with 2−3 sec resolution) are collected and a background subtraction is performed for each case. Finally, a temporal integration from the reported in the catalog

Table 4. Table of the Pearson correlation coefficients ($\log_{10}-\log_{10}$) between the SOHO/ERNE proton peak intensity and the flare and CME parameters, respectively, using different solar origin association at western helio-longitudes.

Solar origin	Pearson correlations			
association	Flare class	Flare fluence	CME speed	CME kinetic energy
1996−2016				
this study	0.51±0.05 (288)	0.52±0.05 (270)	0.58±0.04 (361)	0.59±0.04 (313)
NOAA	0.31±0.11 (74)	0.23±0.14 (68)	0.41±0.10 (71)	0.44±0.09 (60)
GOES	0.32±0.08 (106)	0.22±0.11 (92)	0.38±0.08 (112)	0.40±0.10 (94)
SOHO	0.24±0.09 (108)	0.35±0.11 (100)	0.51±0.06 (137)	0.45±0.10 (118)
Wind	0.38±0.07 (185)	0.39±0.08 (170)	0.50±0.05 (224)	0.50±0.06 (191)
SC23': 11/1996−10/2004				
this study	0.50±0.06 (175)	0.46±0.08 (168)	0.54±0.05 (214)	0.58±0.05 (185)
NOAA	0.27±0.13 (49)	0.11±0.16 (45)	0.47±0.12 (44)	0.44±0.11 (35)
GOES	0.29±0.10 (76)	0.13±0.13 (66)	0.40±0.08 (76)	0.49±0.08 (62)
SOHO	0.22±0.13 (63)	0.22±0.15 (61)	0.52±0.08 (78)	0.58±0.08 (66)
Wind	0.34±0.09 (116)	0.31±0.11 (108)	0.52±0.07 (134)	0.58±0.06 (113)
SC24': 01/2009−12/2016				
this study	0.43±0.09 (80)	0.58±0.07 (74)	0.61±0.06 (109)	0.60±0.06 (95)
NOAA	0.37±0.18 (20)	0.60±0.12 (20)	0.22±0.24 (22)	0.40±0.22 (20)
GOES	0.29±0.17 (16)	0.43±0.19 (16)	0.06±0.25 (21)	0.48±0.19 (20)
SOHO	0.11±0.15 (35)	0.55±0.14 (32)	0.48±0.12 (45)	0.53±0.11 (39)
Wind	0.36±0.12 (52)	0.51±0.11 (48)	0.46±0.08 (70)	0.49±0.08 (61)

flare onset time to peak time is performed and further referred to as 'flare fluence'.

Moreover, we investigated the influence of the time period (SC dependence) for the entire sample, and for the first eight years of SCs 23 and 24 (the exact duration is given in the Tables). Finally, we explore separately the events with a solar origin from western helio-longitudes (Table 4).

When comparing the correlations between the SOHO/ERNE proton intensity with SF fluence, we notice increase of the correlations with the fluence for this study, SOHO and Wind and a decrease when using association according to NOAA and GOES, Table 3. The same results are obtained when considering different time periods (1996−2016, SC23' and SC24') and also for the case of western events only, Table 4. When using SF class, the results based on using SOHO association are closer to those of NOAA and GOES. Thus, opposite conclusions can be reached depending on how the solar origin association is made. When comparing CME speed and kinetic energy, the results in terms of correlations coefficients are closer in values, some increasing or decreasing trends are noticed that are sample-specific and within the uncertainty.

Overall, weaker correlations are obtained for the correlations with the flare parameters compared with the CME parameters in any of the time periods. This trend is kept, especially for the lists with low number of data points (NOAA, GOES and SOHO), also when one considers the SF fluence for western events in SC23'. When the event sample size is larger, however (see 'this study', and also 'Wind'), there are no statistical differences in the correlations between the SEPs with SF fluence, CME class and CME kinetic energy.

When considering western events (Table 4), the main finding obtained for the entire sample (Table 3) are also valid here. Overall, slightly larger correlations are obtained for the western events compared to the respective values for the entire sample (compare the two tables), however not statistically significant. The above findings are valid when the SEP sample and peak intensities are adopted from another instrument (i.e., Wind/EPACT, see results in the Appendix), thus the conclusions reached here should not vary significantly after accounting for possible bias of the used particle instrument.

3 Discussions and conclusions

In the present study, the validity of the solar association is estimated using qualitative factors, based primary on a observer judgment for the level of confidence and, quantitatively, on the time offset between pairs of SFs−CMEs. Here it is confirmed, similarly to the preliminary results Miteva (2017a), that the inclusion of events with uncertain SEP-origin association decreases the correlation coefficients between protons and SFs/CMEs despite the low percentage of the differences compared to the total event number. Thus, erroneously identified solar origin tends to randomize additionally the data and adds to the already largely scattered samples. Note that the statistical correlations are regarded as a proxy for the physical relationship between particles and their solar origin. Thus a drop in the correlations (with some variations depending on the sample), both when calculated with SF class and CME speed, may be misinterpreted, even though the difference between the correlations range from mostly within the statistical uncertainties to marginally different. In case of large statistics, it is advisable to drop events with doubtful origin association in order to reduce the scatter and to improve the quantitative comparison.

The analysis in this work confirms the relatively small number of alternative SF and CMEs as SEP solar origin as reported by the different authors of the NOAA, GOES, SOHO and Wind catalogs, respectively: the discrepancies range from 1−13% for SFs and 4−15% for CMEs. The sample of alternative events contains large protons (with peak intensity larger than the median value of the SOHO/ERNE sample) with associated flares of M4.2 (that are also larger compared to the median values for the entire sample) and CMEs of 890 $\mathrm{km\,s^{-1}}$ (both in median values). The longitudinal distribution cannot explain why different origin has been proposed by different authors. Since the events have well identified proton intensity (so-called large events), the probable explanation could be sought in the multitude of eruptions that take place in close succession prior the proton onset at 1 AU and the implicit subjectivity of the association procedure.

When different SFs and CMEs are identified for the same proton event, the respective Pearson correlation coefficients between properties of alternative flare/CME with the proton intensities are also different. Using the proton list based on 20 MeV SOHO/ERNE data, a quantitative assessment on the difference was performed. In summary, the variation in flare/CME associations leads to a decrease in the correlation coefficients, more pronounced for the flares and for the samples with low event number. Thus,

this effect could explain the lower correlations reported with flares, compared than CMEs in the past.

In summary, alternative solar origin associations do lead to different values for the correlation coefficients, and the effect is more pronounced when the sample is small, so the weight of the difference becomes larger. The difference is stronger when the correlations with SF class but also SF fluence are considered. Thus, alternative solar origin associations and small event sample could be reasons for the usually lower correlation coefficients reported with flares, compared to CMEs, especially when using NOAA and GOES based solar origin associations or catalogs. CME kinetic energy gives consistent results to those when using CME speed, since either can be regarded as a representative parameter of the CME proton acceleration and the two parameters are interdependent. The use of SF fluence over class usually improves the correlations, however this trend is specific to the event sample − both in origin and size − and to the time-period under consideration. Thus the observer subjectivity still plays an decisive role in the SEP analysis and their interpretation, since a chain of specific choices are made that are instrument, time and observer-specific, namely selecting which type of flare or CME parameter to be used for the analysis, the time coverage under investigation, the way of association between SEPs and SFs/CMEs, and the final sample size.

Acknowledgement

This work is supported by the National Science Fund of Bulgaria with contract No. DNTS/Russia 01/6 (23-Jun-2017). We acknowledge the use of SOHO/ERNE data in our study. SOHO is a project of international collaboration between ESA and NASA.

References

Aschwanden, M. J., 2002, *Space Sci. Rev.*, **101**, pp. 1–227
Bazilevskaya, G. A., 2017, *Journal of Physics Conference Series*, **798**, p. 012034
Fletcher, L., Dennis, B. R., Hudson, H. S., Krucker, S., Phillips, K. , Veronig, A. Battaglia, M., Bone, L., Caspi, A., Chen, Q., Gallagher, P., Grigis, P. T., Ji, H., Liu, W., Milligan, R. O., Temmer, M., 2011, *Space Sci. Rev.*, **159**, pp. 19–106
Klein, K.-L., Dalla, S., 2017, *Space Sci. Rev.*, **212**, pp. 1107–1136
Krauss-Varban, D., 2010, *Cambridge University Press*, 209pp.
Miteva, R., 2017a, *Proceedings of the 120 years Astronomical Observatory, Sofia University 'St. Kliment Ohridski'*, pp. 5–10
Miteva, R., 2017b, *Space, Ecology, Safety - SES 2017, Thirteenth International Scientific conference 'Space, Ecology, Safety-SES2017', held 2–4 November 2017 in Sofia, Bulgaria, G. Mardirossian, Ts. Srebrova and G. Jelev (eds.), ISSN: 1313-3888*, pp. 52–56
Miteva, R., Klein, K.-L., Malandraki, O., Dorrian, G., 2013, *Solar Physics*, **282**, pp. 579–613
Miteva, R., Samwel, S. W., Costa-Duarte, M. V., Malandraki, O. E., 2017c, *Sun and Geosphere*, **12(2)**, pp. 11–19
Miteva, R., Samwel, S. W., Krupar V., 2017d, *Journal of Space Weather and Space Climate*, **7**, id.A37, 15pp.
Miteva, R, Samwel, S. W., Costa-Duarte, M. V., 2018a, *JASTP*, **180**, pp. 26–34
Miteva, R., Samwel, S. W., Costa-Duarte, M. V., 2018b, *Solar Phys.*, **293**, id. 27
Nindos, A., Aurass, H., Klein, K.-L., Trottet, G., 2008, *Solar Physics*, **253**, pp. 3–41

Paassilta, M., Raukunen, O., Vainio, R., Valtonen, E., Papaioannou, A., Siipola, R., Riihonen, E., Dierckxsens, M., Crosby, N., Malandraki, O., Heber, B., Klein, K.-L., 2017, *Journal of Space Weather and Space Climate*, **7**, id.A14
Papaioannou, A., Sandberg, I., Anastasiadis, A., Kouloumvakos, A., Georgoulis, M. K., Tziotziou, K., Tsiropoula, G., Jiggens, P., Hilgers, A., 2016, *Journal of Space Weather and Space Climate*, **6**, id.A42
Park, J., Moon, Y.-J., Lee, H., 2017, *Astrophys. J.*, **844**, 17
Pick, M., Vilmer, N., 2008, *Astron. Astrophys. Rev.*, **16**, pp. 1–153
Richardson, I. G., von Rosenvinge, T. T., Cane, H. V., 2015, *Solar Physics*, **290**, pp. 1741–1759
Torsti, J., Valtonen, E., Lumme, M., Peltonen, P., Eronen, T., Louhola, M., Riihonen, E., Schultz, G., Teittinen, M., Ahola, K., Holmlund, C., Kelhä, V., Leppälä, K., Ruuska, P., Strömmer, E., 1995, *Solar Physics*, **162**, pp. 505–531
Vainio, R., Valtonen, E., Heber, B., Malandraki, O. E., Papaioannou, A., Klein, K.-L., Afanasiev, A., Agueda, N., Aurass, H., Battarbee, M., Braune, S., Dröge, W., Ganse, U., Hamadache, C., Heynderickx, D., Huttunen-Heikinmaa, K., Kiener, J., Kilian, P., Kopp, A., Kouloumvakos, A., Maisala, S., Mishev, A., Miteva, R., Nindos, A., Oittinen, T., Raukunen, O., Riihonen, E., Rodríguez-Gasén, R., Saloniemi, O., Sanahuja, B., Scherer, R., Spanier, F., Tatischeff, V., Tziotziou, K., Usoskin, I. G., Vilmer, N., 2013, *Journal of Space Weather and Space Climate*, **3**, id.A12
Wall, J. V., Jenkins, C. R., 2003, *Cambridge observing handbooks for research astronomers, Vol. 3, Cambridge University Press*
Webb, D. F., Howard, T. A., 2012, *Living Reviews in Solar Physics*, **9**, 3

Appendix

The following Tables 5 and 6 summarize the Pearson correlation coefficients calculated between the Wind/EPACT peak proton intensity and the solar origin as identified by various sources, separately for the entire sample and for events from western helio-longitude, respectively.

Table 5. Table of the Pearson correlation coefficients ($\log_{10}-\log_{10}$) between the Wind/EPACT proton peak intensity and the flare and CME parameters, respectively, using different solar origin association.

Solar origin	Pearson correlations			
association	Flare class	Flare fluence	CME speed	CME kinetic energy
1996−2016				
this study	0.42±0.06 (245)	0.44±0.07 (230)	0.50±0.04 (292)	0.50±0.04 (246)
NOAA	0.35±0.08 (101)	0.27±0.09 (89)	0.46±0.06 (95)	0.48±0.08 (77)
GOES	0.43±0.07 (137)	0.34±0.09 (122)	0.49±0.05 (142)	0.51±0.05 (113)
SOHO	0.32±0.08 (122)	0.46±0.09 (114)	0.55±0.05 (161)	0.55±0.05 (133)
Wind	0.42±0.05 (283)	0.43±0.06 (260)	0.51±0.04 (334)	0.47±0.04 (282)
SC23': 11/1996−10/2004				
this study	0.37±0.08 (147)	0.39±0.09 (139)	0.51±0.06 (166)	0.52±0.06 (139)
NOAA	0.32±0.11 (60)	0.25±0.12 (52)	0.47±0.08 (54)	0.55±0.10 (41)
GOES	0.34±0.09 (92)	0.27±0.11 (81)	0.45±0.07 (87)	0.49±0.07 (69)
SOHO	0.33±0.11 (70)	0.41±0.13 (68)	0.60±0.05 (90)	0.55±0.07 (74)
Wind	0.37±0.07 (170)	0.38±0.09 (156)	0.56±0.05 (186)	0.53±0.05 (157)
SC24': 01/2009−12/2016				
this study	0.37±0.10 (72)	0.45±0.10 (68)	0.51±0.07 (100)	0.46±0.07 (86)
NOAA	0.35±0.13 (32)	0.29±0.11 (30)	0.44±0.11 (34)	0.49±0.13 (30)
GOES	0.51±0.13 (25)	0.39±0.15 (24)	0.53±0.12 (36)	0.64±0.10 (31)
SOHO	0.26±0.12 (41)	0.57±0.10 (37)	0.52±0.10 (56)	0.63±0.08 (47)
Wind	0.37±0.09 (86)	0.40±0.09 (80)	0.43±0.07 (119)	0.44±0.07 (103)

Table 6. Table of the Pearson correlation coefficients ($\log_{10}-\log_{10}$) between the Wind/EPACT proton peak intensity and the flare and CME parameters, respectively, using different solar origin association at western helio-longitudes.

Solar origin	Pearson correlations			
association	Flare class	Flare fluence	CME speed	CME kinetic energy
1996−2016				
this study	0.40±0.06 (189)	0.45±0.08 (175)	0.51±0.05 (232)	0.53±0.05 (199)
NOAA	0.29±0.10 (75)	0.28±0.10 (69)	0.48±0.06 (72)	0.51±0.08 (61)
GOES	0.38±0.08 (106)	0.33±0.10 (93)	0.48±0.06 (112)	0.57±0.05 (94)
SOHO	0.28±0.08 (109)	0.42±0.10 (102)	0.54±0.06 (140)	0.58±0.05 (120)
Wind	0.42±0.06 (211)	0.44±0.07 (193)	0.49±0.05 (262)	0.49±0.05 (228)
SC23': 11/1996−10/2004				
this study	0.35±0.08 (115)	0.36±0.11 (108)	0.52±0.06 (135)	0.56±0.05 (115)
NOAA	0.24±0.12 (49)	0.22±0.13 (45)	0.50±0.07 (44)	0.50±0.09 (35)
GOES	0.31±0.10 (73)	0.20±0.12 (63)	0.48±0.07 (72)	0.54±0.07 (58)
SOHO	0.27±0.10 (63)	0.35±0.14 (61)	0.56±0.06 (80)	0.58±0.06 (68)
Wind	0.38±0.09 (128)	0.36±0.10 (117)	0.53±0.05 (153)	0.54±0.05 (132)
SC24': 01/2009−12/2016				
this study	0.37±0.11 (55)	0.54±0.10 (51)	0.52±0.08 (75)	0.57±0.08 (65)
NOAA	0.33±0.16 (21)	0.46±0.11 (21)	0.39±0.16 (23)	0.58±0.16 (21)
GOES	0.46±0.15 (19)	0.57±0.14 (19)	0.44±0.17 (25)	0.71±0.09 (24)
SOHO	0.19±0.13 (37)	0.58±0.10 (34)	0.57±0.10 (47)	0.66±0.09 (40)
Wind	0.38±0.10 (64)	0.50±0.09 (59)	0.43±0.09 (86)	0.47±0.08 (76)