



ON THE STATE OF A SOLAR ACTIVE REGION BEFORE FLARES AND CMEs

M. B. KORSÓS^{1,2} AND R. ERDÉLYI^{1,2}

¹ Debrecen Heliophysical Observatory (DHO), Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, Debrecen, P.O. Box 30, H-4010, Hungary; korsos.marianna@csfk.mta.hu

² Solar Physics & Space Plasma Research Center (SP2RC), University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK; robertus@sheffield.ac.uk

Received 2015 August 6; accepted 2016 March 7; published 2016 June 1

ABSTRACT

Several attempts have been made to find reliable diagnostic tools to determine the state prior to flares and related coronal mass ejections (CMEs) in solar active regions (ARs). Characterization of the level of mixed states is carried out using the Debrecen sunspot Data for 116 flaring ARs. Conditional flare probabilities (CFPs) are calculated for different flaring classes. The association with slow/fast CMEs is examined. Two precursor parameters are introduced: (i) the sum of the (daily averaged) horizontal magnetic gradient G_S (G_{DS}) and (ii) the separation parameter S_{l-f} . We found that if $S_{l-f} \leq 1$ for a flaring AR then the CFP of the expected highest-intensity flare being X-class is more than 70%. If $1 \leq S_{l-f} \leq 3$ the CFP is more than 45% for the highest-intensity flare(s) to be M-class, and if $3 \leq S_{l-f} \leq 13$ there is larger than 60% CFP that C-class flare(s) may have the strongest intensity within 48 hr. Next, from analyzing G_S for determining CFP we found: if $5.5 \leq \log(G_S) \leq 6.5$, then it is very likely that C-class flare(s) may be the most intense; if $6.5 \leq \log(G_S) \leq 7.5$ then there is ~45% CFP that M-class could have the highest intensity; finally, if $7.5 \leq \log(G_S)$ then there is at least 70% chance that the strongest energy release will be X-class in the next 48 hr. ARs are unlikely to produce X-class flare(s) if $13 \leq S_{l-f}$ and $\log(G_S) \leq 5.5$. Finally, in terms of providing an estimate of an associated slow/fast CME, we found that, if $\log(S_{l-f}) \geq 0.4$ or $\log(G_{DS}) \leq 6.5$, there is no accompanying fast CME in the following 24 hr.

Key words: Sun: coronal mass ejections (CMEs) – Sun: flares – sunspots

1. INTRODUCTION

The production of flares and coronal mass ejections (CMEs) from solar active regions (ARs) is still not well understood. The size and morphological properties of ARs seem to be decisive. To the best of our knowledge, Waldmeier (1938) published the first classification scheme to examine the role of the size and morphology of sunspot groups in relation to determining the capacity of their flare-productivity. This scheme contained eight classes thought to be characteristic representatives of consecutive states in the development of a sunspot group. The system was somewhat modified by Waldmeier (1947) and is known today as the modified Zürich classification system (see also Kiepenheuer 1953). The classification was further developed by McIntosh in the 1960s and the version that is still in wide use today was released by McIntosh (1990). The McIntosh classification is a three-component form describing the (modified) Zürich class, the largest spot, and the sunspot distribution in an AR. The classification is based on scrutinizing white-light observations, but no magnetic information is considered. The scheme was used in several flare-prediction studies (see, e.g., McIntosh et al. 1986; Bornmann & Shaw 1994; Colak & Qahwaji 2008). Lee et al. (2012) attempted to develop this scheme further by including data on sunspot areas and their temporal variations. The classification of sunspot groups and tracking the evolution of ARs seem to be important tasks to unveil potentially existing connections between sunspot activity and various eruptive solar phenomena (including, e.g., solar flares and CMEs).

Probably the first magnetic classification scheme was introduced in the classic paper of Hale et al. (1919), known today as the Mount Wilson classification. It is simpler than the Zürich–McIntosh system, because it only distinguishes unipolar, bipolar, and mixed configurations, denoted by the letters α , β , and γ , respectively. These classes were later

complemented, by Künzel (1960), with the δ -class, denoting those cases where two spots of opposite polarity are very close to each other within a common penumbral feature. Künzel also found that the δ -class configurations are the most productive sources of energetic flares. This simple but suggestive early finding is a worthwhile hint that magnetic fields may play a role in the development of localized solar energetic events. Later on, more focused investigations confirmed this conjecture (see, e.g., Warwick 1966; Sammis et al. 2000).

Solar flares and CMEs are likely part of a single, magnetically driven occurrence put forward by the standard 2D CSHKP model (Carmichael 1964; Sturrock 1966; Hirayama 1974; Kopp & Pneuman 1976). A flare and CME could occur independently of each other. The vast majority of CMEs have associated flares, with only a much smaller population of non-flaring filament lift-offs leading to CMEs (Gosling et al. 1976; Harrison 1995). Further, there are even fewer cases of “stealth” CMEs that have no apparent solar activity at all (Howard & Harrison 2013). On the other hand, the association (or connection) rate of flares and CMEs increases sharply with the strength of the flare event (Yashiro et al. 2006). Therefore, in summary, flares and CMEs may need to have a different classification scheme. The most generally known flare classification is currently based on data from the *Geostationary Operational Environmental Satellite (GOES)*. Measurements of the maximum flux of X-rays at wavelengths from 0.1 to 0.8 nm are classed as arising from flares of A, B, C, M, or X type; however, there are only two separate types of CMEs and they depend on speed. If the linear speed of a CME is smaller/larger than the velocity of the solar wind, we call it slow (500–800 km s^{−1})/fast (over 800 km s^{−1}; Webb & Howard 2012). Evans et al. (2013) introduced a more detailed classification of CMEs that is based on the linear speed of the CMEs from the *SOHO/LASCO*

(*Solar and Heliospheric Observatory/Large Angle Spectrometric Coronagraph*³) catalog. In this work, we use the *SOHO/LASCO* catalog because it contains all identified CMEs from 1996. This catalog contains two speed values of a CME, the linear speed and the quadratic speed. These speeds are usually determined by the height–time measurements for the fastest moving part of the CME front. This part is projected on the plane of the sky. In the *SOHO/LASCO* CME catalog the values for velocity in the plane of the sky are usually lower than the real radial velocity of the CME front, which causes some uncertainties in our investigation.

The classification schemes introduced above were useful in revealing potential connections between the size and structure of sunspot groups and their flare- and CME-productivity. However, it is somewhat ambitious that these classification procedures contain subjective factors to be determined by visual inspection besides objective measures, such as the size of a group and that of the largest spot. With the current schemes only specific classes of sunspots can be examined. We suggest that it would be more advantageous to develop a scheme based on properly defined proxies of the non-potentiality of magnetic field associated with ARs, with parameters characterizing the flaring capability, more objectively. The aim of the present paper is to find and test such possible parameters.

2. CLASSIFICATIONS OF MORPHOLOGICAL COMPLEXITY

The current study is based on SDD⁴ (*SOHO/MDI-Debrecen Data*, Győri et al. 2011; Győri 2015), the detailed sunspot catalog in the *SOHO/Michelson Doppler Imager* (MDI) era. This catalog contains valuable data on position, area, and mean magnetic field for each sunspot and sunspot group.

Korsós et al. (2014), henceforth Paper I, investigated the pre-flare behavior of the horizontal magnetic gradient (G_M) parameter. They found that G_M exhibits characteristic pre-flare patterns: a steep rise until a maximum, followed by a gradual decrease for several hours and then the outburst. This behavior is promising for flare-forecast perspectives because, when G_M starts to decrease after a pronounced maximum, this pattern may be considered to be a warning signal of eruption(s). Later, the method of Paper I was further developed and the weighted horizontal magnetic gradient, or WG_M (Korsós et al. 2015a), was introduced. This new proxy enables the potential to forecast flares stronger than M5. The improved capability includes (i) the prediction of flare onset time and (ii) an assessment of whether a flare will be followed by another event within about 18 hr. In the future, we would like to automate the method outlined in Korsós et al. (2015a). Driven by this desire, therefore, in this paper we seek to find even more reliable precursors for identifying with high probability ARs with the potential for flare/CME activities. It is our aim that only ARs passing a certain set of evaluation criteria would be further investigated, analyzed, and followed by such automated methods (e.g., WG_M or other) in the not too distant future. In what follows, we discuss only flaring ARs, because flaring and non-flaring ARs were tested separately on a large sample by Korsós et al. (2015a). We found differences in the behavior of the separation, between the area-weighted centers of opposite polarities, of flaring and non-flaring ARs.

The present work is a generalization and further development of Paper I. It demonstrates the usefulness and tests the introduction of different predefined parameters characterizing the level of mixed states of sunspots by estimating their conditional probabilities before the onsets of flares (also presented as a single-case pilot study in Korsós et al. 2015b). The probabilities presented here are conditional upon already knowing that the AR being considered will definitely flare. This is a result of data selection criteria being applied so that only flaring ARs have been included in all of our analysis, with similar selection quotas being used for ARs that produced at most B-/C-class flares, at most M-class flares, and at most X-class flares. The first suggested complexity parameter and its derivative, denoted by G_S and its daily average G_{DS} , respectively, is the generalization of the G_M proxy. G_S is the sum of G_M taken at every 1.5 hr for all spot-pairs of opposite polarities within a group, while G_{DS} is actually the daily average of the sum of G_M .

The sum of the horizontal magnetic gradient is

$$G_S = \left| \sum_{i,j} \frac{B_{p,i}A_{p,i} - B_{n,j}A_{n,j}}{d_{i,j}} \right|, \quad (1)$$

and the daily average of the sum of G_M is

$$G_{DS} = \left| \sum_{i,j} \frac{B_{p,i}A_{p,i} - B_{n,j}A_{n,j}}{d_{i,j}} \right|, \quad (2)$$

where B is the mean magnetic field determined by $f(A)$ in Paper I. A is the area of the umbra. The indices p and n denote positive and negative polarities, i and j are their running indices in the entire sunspot group, and d is the distance between two spots of opposite polarity.

The second proposed complexity parameter of this study characterizes the separation of opposite-polarity subgroups defined by

$$S_{l-f} = \frac{d_{lc-fc}}{2\sqrt{\sum A_g/\pi}}. \quad (3)$$

Here, l and f refer to the leading and following polarities. A_g represents the total area of the sunspot group. The numerator denotes the distance between the area-weighted centres (thus the index c) of the spots of leading and following polarities. The denominator is the diameter of a hypothetical circle whose area is equal to the total area of all umbrae constituting the sunspot group.

To test and demonstrate the informativeness of formulae (1)–(3), we selected three typical ARs with different levels of complexity. Of course, the method is consecutively applied to a much larger data set of SDD, containing 116 samples of ARs that produced flare(s) of different intensity with or without an associated CME covering the years of MDI operations, from 1996 to 2010. Since we prefer to use a homogeneous catalog we opted for this time interval for the investigation presented because SDD covers only this period. In the future we would expand the examination to include other catalogs (e.g., the *SDO/HMI—Debrecen Data*, known as HMIDD⁵).

First of all, we introduce three typical ARs that cannot be distinguished very well using the Mount Wilson classification. AR 10693 is a beta–gamma magnetic region, which means that

³ http://cdaw.gsfc.nasa.gov/CME_list/

⁴ <http://fenyi.solarobs.unideb.hu/SDD/SDD.html>

⁵ <http://fenyi.solarobs.unideb.hu/ESA/HMIDD.html>

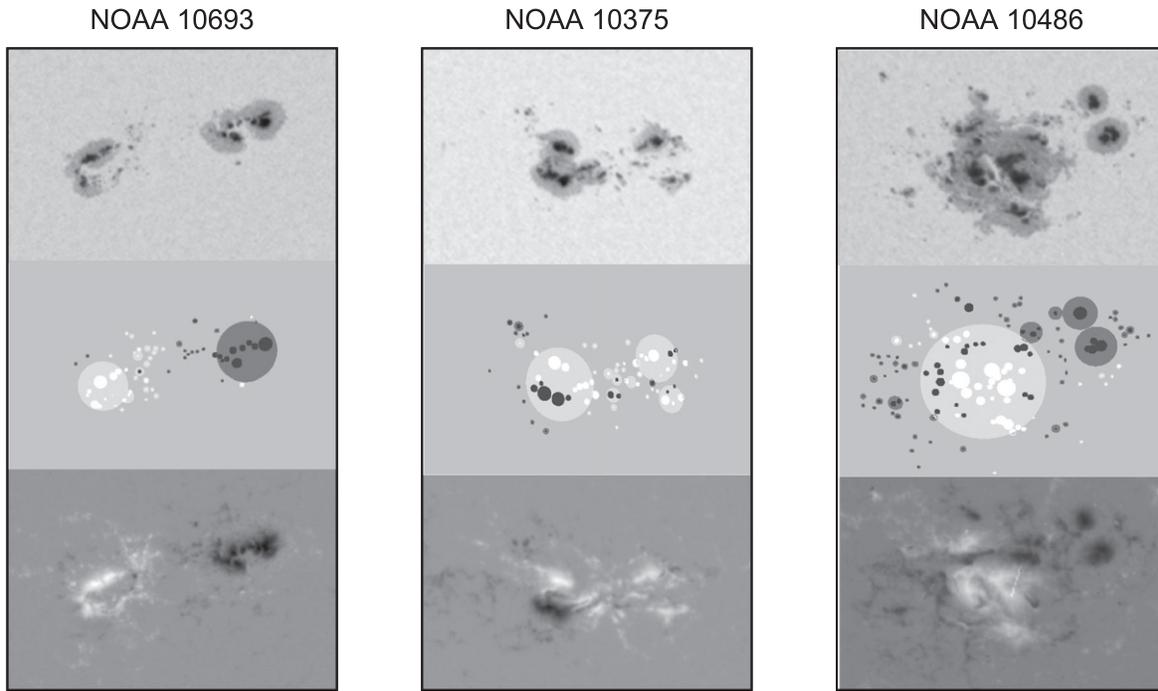


Figure 1. Snapshots of NOAA AR 10693, AR 10375, and AR 10486. Top panels: white-light images; middle panels: reconstructions of the ARs from the SDD data; bottom panels: *SOHO*/MDI magnetograms.

it is bipolar, and a continuous line can be drawn between the positive and negative polarities. ARs 10375 and 10486 are beta–gamma–delta magnetic configurations. This latter classification means that the sunspot group is sufficiently complex, having umbrae of opposite polarity within the penumbra. Let us now turn to our approach: Figure 1 depicts AR 10693 (left panel), a clearly bipolar structure, where the opposite polarities are well separated; next, AR 10375 is an example of a medium level of polarity mixture (middle panel); finally, AR 10486 (right panel) has numerous very large flares and is highly mixed. All panels cover the same spatial extent, and in each panel the top image is the white-light snapshot of the corresponding AR and the bottom image is its magnetogram; all these observations have been made by the *SOHO*/MDI instrument. The middle panels of Figure 1 show a synthetic polarity drawing of the sunspot group that is reconstructed from the *SOHO*/MDI Debrecen sunspot Data catalog. The middle images are cartoons reconstructing the relevant ARs by using the data on position, area, and mean magnetic polarity (separately for umbra and penumbra) of the individual spots. These cartoons may visualize well the complexity level of the sunspot polarity distributions. Generally, black and white (and their shades) distinguish the polarity. Black and white circles are for umbrae. Often, much larger, shaded circles represent penumbrae and occasionally contain even a mix of positive and negative (i.e., black and white) polarities (for more details see the Debrecen Heliophysical Observatory website⁶).

Let us now analyze these typical ARs by applying to them the complexity parameters introduced above. Figure 2 shows the variation of G_S , G_{DS} , and S_{l-f} . The pre-flare behaviors of G_S (top row) or G_{DS} (middle row), and S_{l-f} (bottom row) are rather different for these typical ARs. They demonstrate a

correlation with the level of complexity of polarities. At first sight G_S (and G_{DS}) may seem to become larger in time in all cases; therefore, one may expect stronger flare intensity. However, on closer inspection, for the bipolar AR 10693, the actual values of G_S (G_{DS}) are rather low. The parameter(s) G_S (G_{DS}) for the medium level of polarity mixture, i.e., for AR 10375, is (/are) higher than their counterpart(s) in the bipolar case, but still lower than their counterpart(s) for AR 10486 with highly mixed polarity. Let us keep these trends in mind. It is worth noting, however, that the fluctuation of G_S is too strong. Thus, it may be advisable to use G_{DS} instead for forecasting flares and CMEs. A posteriori, actually this is why we have introduced the daily average of the horizontal magnetic gradient (e.g., Section 4).

Next, the diagrams of S_{l-f} (bottom row) show variations that have a much smoother character than that of G_S (G_{DS}). In the quiet, bipolar case the magnitude of the separation parameter S_{l-f} is notably higher than (a threshold value of; we come back to this later) three. In the moderately complex case S_{l-f} is lower than three. Finally, in the most mixed case the values of S_{l-f} would be around unity or less. We may, therefore, conclude that the higher the value of the separation parameter (i.e., S_{l-f}), the more bipolar the character of the sunspot group. If S_{l-f} is low it may mean that the corresponding subgroup is mixed. In summary, so far, we suggest that the two proposed complexity parameters (G_S [G_{DS}] and S_{l-f}) may also be tracked as functions of time during the development of sunspot groups.

The featured examples indicate that if we track these two parameters (G_S [G_{DS}] and S_{l-f}) during the evolution of sunspot groups then we may be able to conclude with some conditional likelihood (that still needs to be determined) whether the two parameters are capable of predicting flares. The next section examines exactly these capabilities.

⁶ http://fenyi.solarobs.unideb.hu/ESA/explanation/20130104-005854.20/20130104-005854.20_11641_w.html

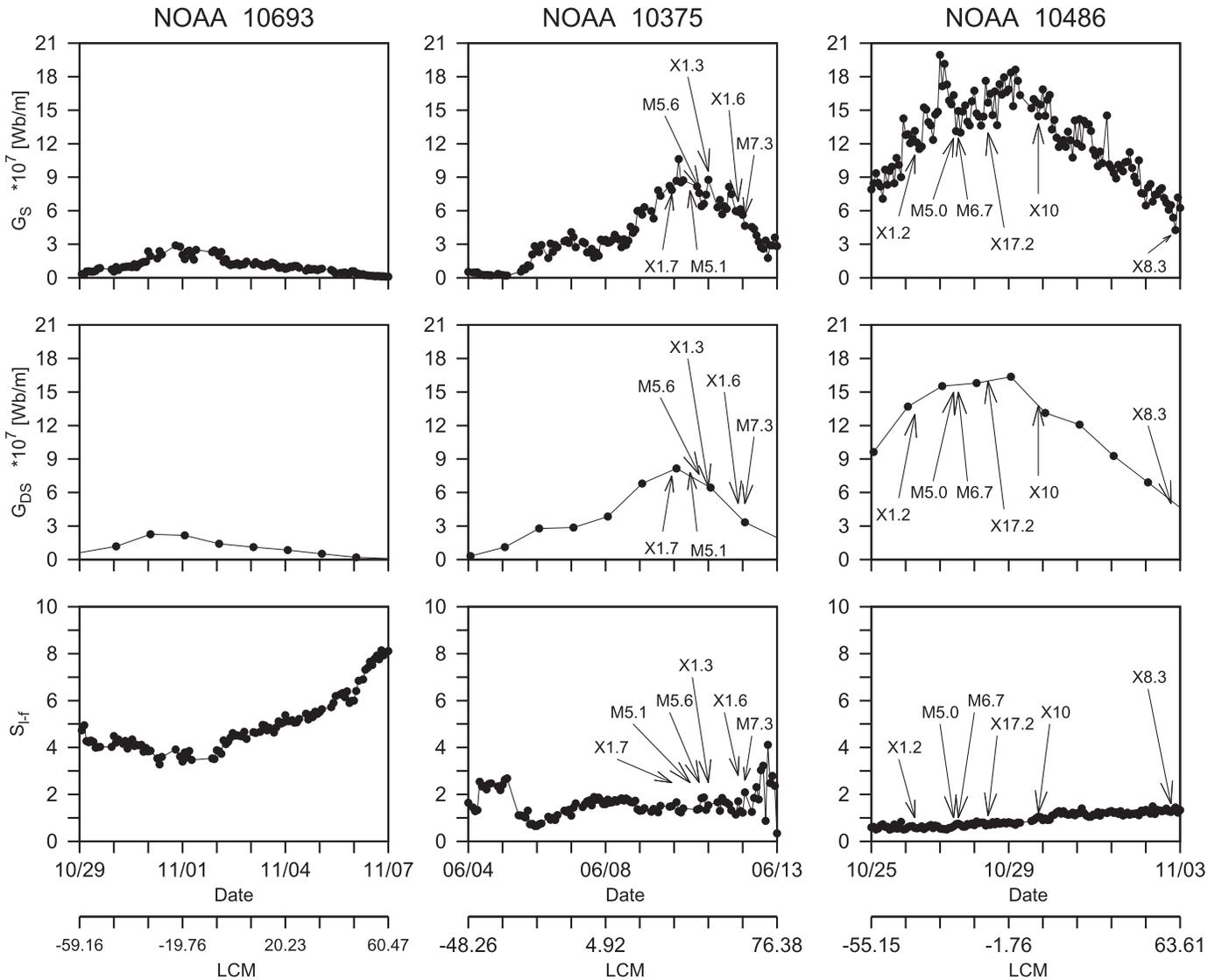


Figure 2. Evolution of parameters G_S , G_{DS} , and S_{l-f} of NOAA AR 10693 (left), AR 10375 (middle), and AR 10486 (right). The pre-flare evolution of G_S (upper row), G_{DS} (middle row), and S_{l-f} (lower row) is shown.

3. FLARE PRECEDED BY CONSECUTIVE DAILY VARIATIONS OF G_{DS} AND S_{l-f}

To test the conditional flare probability (CFP) of the variation of the daily average of the sum of G_M and the separation parameter, we have carried out statistical studies on a selected sample of 116 ARs. The selection criterion was that about a third of the selected ARs produced only B- and C-class flares, another third showed M-class flares, and the remaining third showed X-class flares. The considered quantity S_{l-f} is from 24, 48, and 72 hr before flare onset (often with just a few minutes deviation). We calculate the daily average of the sum of G_M in the intervals 0–24, 24–48, and 48–72 hr before the flare event.

The left panels of Figure 3 show the dependence of the highest-intensity flare on the separation parameter in logarithmic scaling. S_{l-f} has been determined from the moment of first available observation of sunspot groups by using the *SOHO*/MDI-Debrecen Data catalog (recall that SDD provides data on sunspot groups with a cadence of 1.5 hr). The power-law correlation suggests that if $\log(S_{l-f})$ is smaller than about 0.47,

i.e., there is a higher concentration of magnetic polarities, then we may expect flare intensity above the level of M-class flares as opposed to lower-class flares when the value of $\log(S_{l-f})$ is larger than ~ 0.47 (see the upper left panel) a day before flare onset. We can draw a similar conclusion even two days before a flare occurs but in less reliable ways than within 24 hr. Furthermore, we cannot make a reliable conclusion earlier in time because the data become very scattered (e.g., three days before flare onset, see bottom left panel). We have not fitted a linear regression to log–log data taken three days before flare occurrence, because the scatter of these data is too large.

The right panels of Figure 3 show the relationship between the highest-intensity flare of ARs and the daily average of the sum of G_M in the intervals 0–24, 24–48, and 48–72 hr before flare onset. Unfortunately, the fluctuation of G_S itself turned out to be often too large even during an interval as short as one day (see, for example, AR 10486) prior to flare onset; therefore one may not be able to draw a firm conclusion statistically even within the preceding day. Instead, we suggest here, as a better proxy, to apply the daily average of the sum of the horizontal magnetic gradient (G_{DS}) as concluded from following the

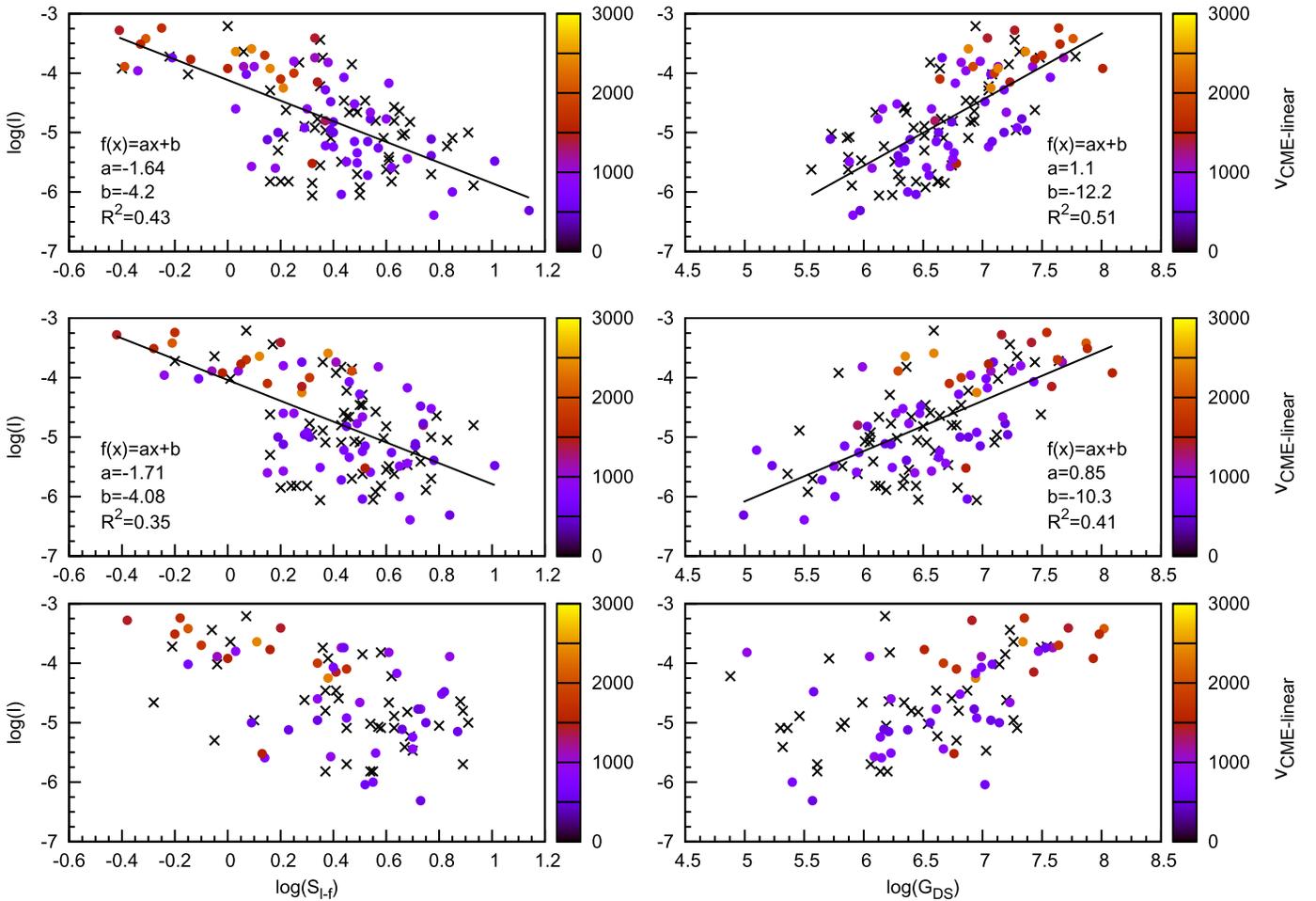


Figure 3. Dependence of the highest-intensity flare of ARs on S_{l-f} (left column) and G_{DS} (right column) in logarithmic scaling. The upper panels refer to data taken a day before the flares occurred, and the middle/lower panels to two/three days before. The black crosses refer to flares that occurred without CME and the colored circles indicate flares associated with CME. The color bars denote the value of the linear speed of the associated CME.

evolution of G_{DS} (see the upper and middle panels of Figure 2). In a first approximation, it is easier to track the daily average of the sum of the horizontal magnetic gradient because of the smoothing of the data. It appears that G_{DS} may also be able to yield a quick (and more practical) estimate of the intensity of the expected flare on the days preceding flare onset, since G_S itself may rise or fall in time faster than S_{l-f} (see Figure 2). Similarly to the case of S_{l-f} , we found a power-law relationship between G_{DS} and the *GOES* intensity of the largest flare produced in the considered time interval (Figure 3 right column, top/middle panels). The linear regression is a better fit to log-log data a day (or two) before flare onset. We may conclude that one may provide an estimate of the flare intensity one or even two days before the expected onset time, but it is almost impossible to do so three days before (see lower right panel of Figure 3).

Next, Figure 4 depicts the relationship between the intensity of the first flare within the highest-intensity flare-class produced by the ARs and S_{l-f}/G_{DS} . If the most intense flare was, e.g., an X-class flare of the AR then we selected the first X-class flare that occurred for further study. If the most intense flare was, e.g., M-strength in the AR then we chose the first M-intensity flare that happened in the AR. We applied the same process for C- and B-class cases as well. Similarly to the case of the most intense flares, we found from the linear correlation

that when $\log(S_{l-f})$ is smaller than approximately 0.47 then $\sim 80\%$ of these flare cases are above M-class in the 48 hr interval. In this latter case, G_{DS} also shows a power-law relationship with the first highest-intensity flare produced in ARs in the interval(s) 0–24 (and 24–48) hr before the flare onset. The correlation coefficient a day before ($R^2 = 0.52$) is much stronger than that two days in advance ($R^2 = 0.38$).

Based on the above analysis, it is worth mentioning that the similarity between Figures 3 and 4 is not surprising, because we studied the same level of the *GOES* flare intensity class in the same AR with two different/complementary approaches.

4. FLARE PRECEDED BY THREE-HOUR TIME STEPS OF G_S AND S_{l-f}

In order to test the newly introduced precursor parameters in relation to analysing the pre-flare state of ARs, we now study the variation of G_S and S_{l-f} over a period of three days prior to a flare. Here, we will also determine the conditional percentage probability of flare activity for each observed *GOES* class. After a number of trials and errors we arrived at determining three suitable bands of the parameters G_S and S_{l-f} applicable a day before flare onset. The bands are statistically predicted by the power-law correlations found for the case of the highest-intensity flare of ARs (see Figure 3, top panels). Note that earlier studies have estimated the probability of flare

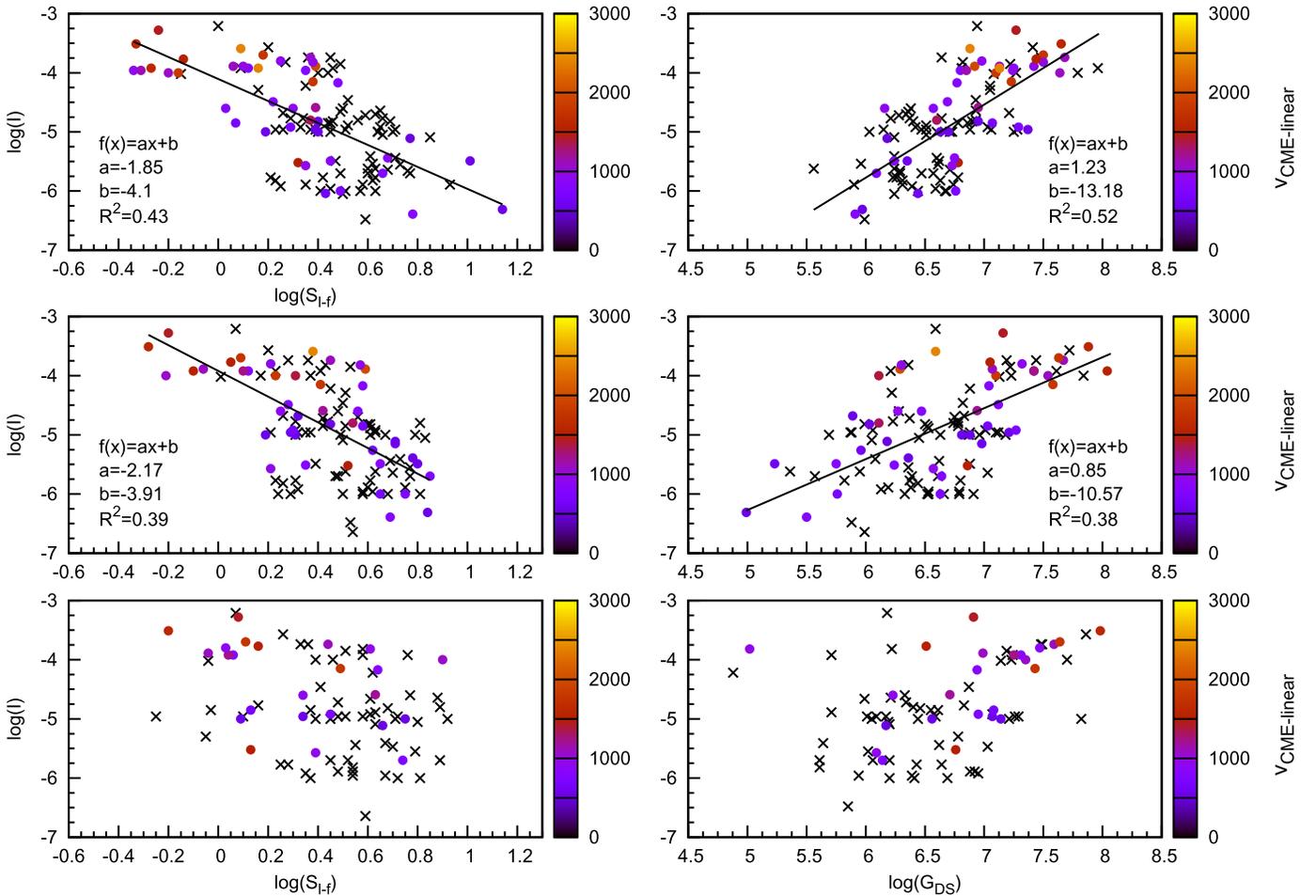


Figure 4. Dependence of the intensity of the first flare within the highest-intensity flare-class produced by the ARs on S_{l-f} (left column) and G_{DS} (right column) in logarithmic scaling. The upper/middle/lower panels are 1/2/3 days before the flare occurs. The black crosses refer to flares that occurred without CME and the colored circles indicate flares associated with CME. The color bars denote the value of the linear speed of the associated CME.

occurrence in a 24 hr window after obtaining AR observations (e.g., Bloomfield et al. 2012), which differs from the work considered here by including time windows with no flares occurring. Furthermore, the choice of the highest-intensity flare gives a good approximation of which flare-class may be expected. We estimated, from the power-law correlation (see the upper left panel of Figure 3), that if S_{l-f} is smaller than ~ 1 then one can expect X-class flare(s) with a high probability. Also, if the separation parameter is between ~ 1 and ~ 3 then energetic flare(s) of M-class could happen. If S_{l-f} is between ~ 3 and ~ 13 then flare event(s) with C-class intensity may take place (see Figure 3) with a reasonable probability. We carried out another test employing G_S as well. It was found that flare intensities of C, M, or X class are expected if $5.5 \leq \log(G_S) \leq 6.5$, $6.5 \leq \log(G_S) \leq 7.5$, or $7.5 \leq \log(G_S)$, respectively (see the right upper panel of Figure 3).

In what follows, for the two approaches outlined in Section 3, i.e., for the case of the highest-intensity flare and for the case of the first flare within the highest-intensity flare-class produced by the ARs, we followed what percentage of all investigated flares have X-, M-, and C-class intensity within the above specified band of S_{l-f} and G_S in every three-hour interval. In this respect, we examine the behavior of CFP by using the parameters S_{l-f} and G_S . Figures 5 and 6 show how CFP varies in time prior to flare onset. In these figures the red/green/blue

symbols represent the probability of X-/M-/C-class flares in specified bands of S_{l-f} and G_S over three consecutive hours in a 72 hr interval before the investigated flare onset. In general, one can say that the CFP values do not change radically within a given interval for each parameter. So, we could estimate the lower limits of CFP using the two approaches (i.e., the highest-intensity flare and the first flare within the highest-intensity flare-class) if we summarize the average percentage from Figures 5 and 6 in the defined intervals by the two parameters in the intervals 0–24, 24–48, and 48–72 hr prior to flares. These actual numerical evaluations of Figures 5 and 6 are given in a concise summary format in Tables 1 and 2.

From Table 1 and Figure 5 one may conclude as follows for the highest-intensity flare/the first flare within the highest-intensity flare-class, respectively. If the S_{l-f} is smaller than one the conditional probability of the highest flare being X-class, assuming it occurs, is about $\sim 85\%/70\%$ in the period 0–48 hr before flare(s) onset. The CFP of M-class flare(s), on average, is $\sim 15\%/30\%$. Finally, the strongest flare-class has practically nil probability of being C-class. Next, if S_{l-f} is between 1 and 3 then the average CFP is about $45\%/50\%$ for the strongest flare intensity to be of M-class. For X-class it is $\sim 25\%/25\%$ and for C-class it is $\sim 30\%/25\%$ in the preceding 0–48 hr interval. If the separation parameter is between 3 and 13, the likelihood of the most energetic class being C-class is $\sim 60\%/65\%$ on

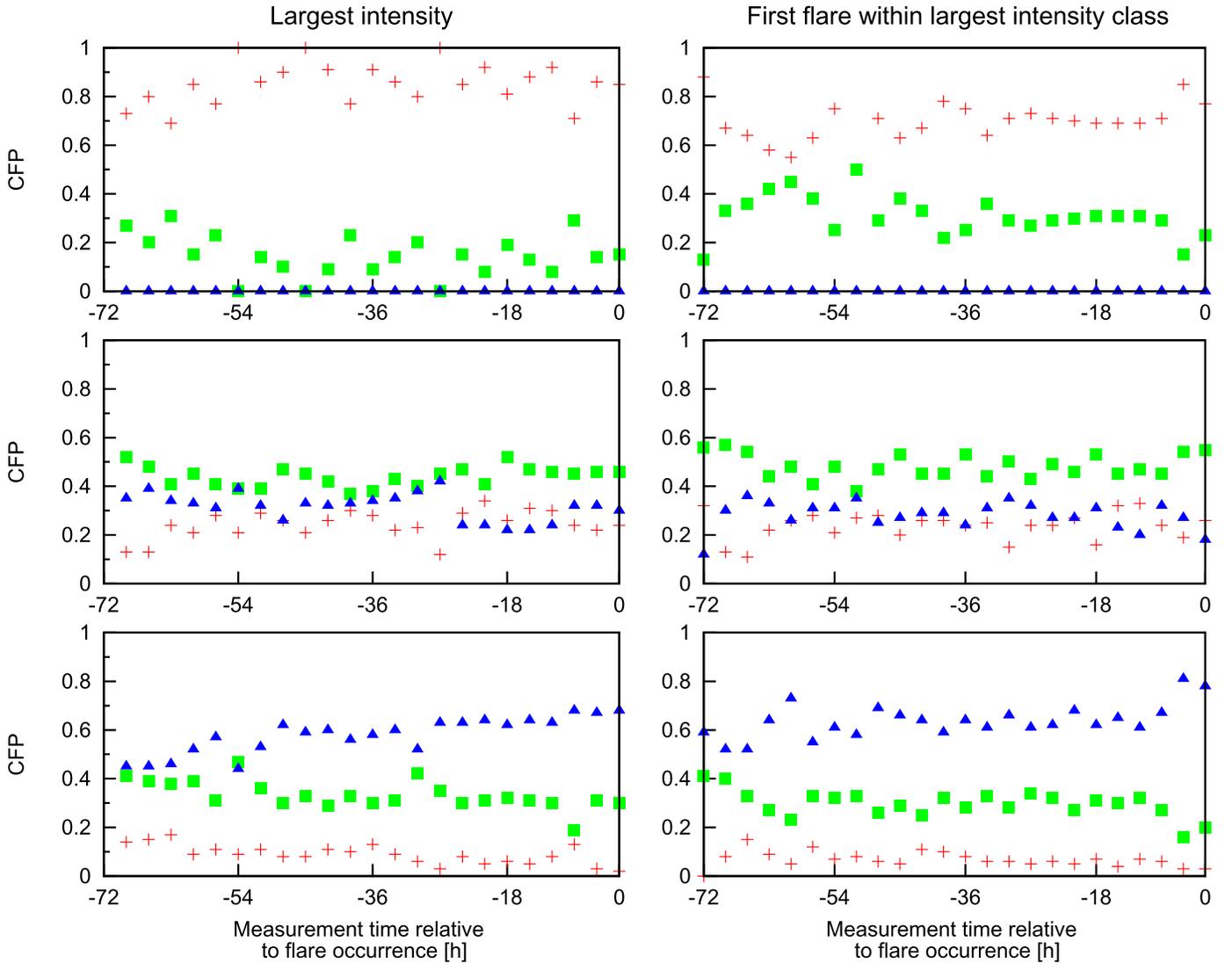


Figure 5. Conditional flare probability, CFP, of S_{l-f} : the figures show the probability of flaring as a function of time. The reference point of the x-axis is the 72 hr prior to flare onset. The left column is for cases with the highest-intensity flare of ARs and the right column is for the case of the first flare within the highest-intensity flare-class. The red/green/blue symbols are the probabilities of X-/M-/C-class flares. The upper, middle, and lower panels are for the intervals $S_{l-f} \leq 1$; $1 \leq S_{l-f} \leq 3$, and $3 \leq S_{l-f} \leq 13$.

average. For M-class it is only $\sim 30\%/30\%$, and for the X-class it is $\sim 5\%/5\%$.

Next, let us follow the CFP of G_5 to draw some conclusions, with the help of Figure 6 and Table 2, on the occurrence probability of the highest-intensity flare-class in the highest-intensity flare/the first flare within the highest-intensity flare-class, respectively. If $5.5 \leq \log(G_5) \leq 6.5$, the highest-intensity flare-class may be the C-class with a CFP, on average, $\sim 70\%/60\%$ within the 0–48 hr interval. The probability for the M-class, on average, is $\sim 25\%/35\%$ and for the X-class there is a minimal likelihood. If $6.5 \leq \log(G_5) \leq 7.5$ then the M-class could be the strongest intensity with a higher probability (on average $\sim 45\%/45\%$) than for the C-class ($\sim 30\%/35\%$). For the X-class, however, the CFP is $\sim 25\%/20\%$ on average. If $\log(G_5)$ is larger than 7.5 within 48 hr prior to the flare, then there is a CFP of around $85\%/70\%$ for X-class flare(s) to occur. Note that from the two approaches ARs do not produce X-class flare(s) if $13 \leq S_{l-f}$ and $\log(G_5) \leq 5.5$. If $S_{l-f} \leq 1$ and $7.5 \leq \log(G_5)$ then there is above 70% CFP of the major event(s) being X-class 48 hr before flare(s) occurrence.

5. ON LINEAR VELOCITY ESTIMATION OF A CME USING G_{DS} AND S_{l-f}

In the next step, let us now study the connection between the linear velocity of the CME associated with ARs and the two complexity parameters introduced earlier (G_{DS} and S_{l-f}). We are aware of which flares occurred with (or without) CME from the combination of *GOES*, *LASCO*, and *SDD* catalogs. We identified which flares are associated with CMEs from the *GOES* and *LASCO* catalogs. Also, for a better identification, we cross-check the position angles of CMEs and source ARs from the *SDD* catalog. Next, we establish the linear velocity of the CME. If a flare does not have a CME, the black crosses indicate zero speed on Figures 3 and 4. Based on the generally accepted CME classification (Webb & Howard 2012) and on the CME SCORE Classification System (Evans et al. 2013) we draw a boundary between slow ($500 \text{ km s}^{-1} \leq v_{\text{lin}} \leq 1000 \text{ km s}^{-1}$) and fast ($1000 \text{ km s}^{-1} \leq v_{\text{lin}}$) CMEs. We found that 22 flares out of 116 occurred with a fast CME and 94 events are associated with slow or no CME in the

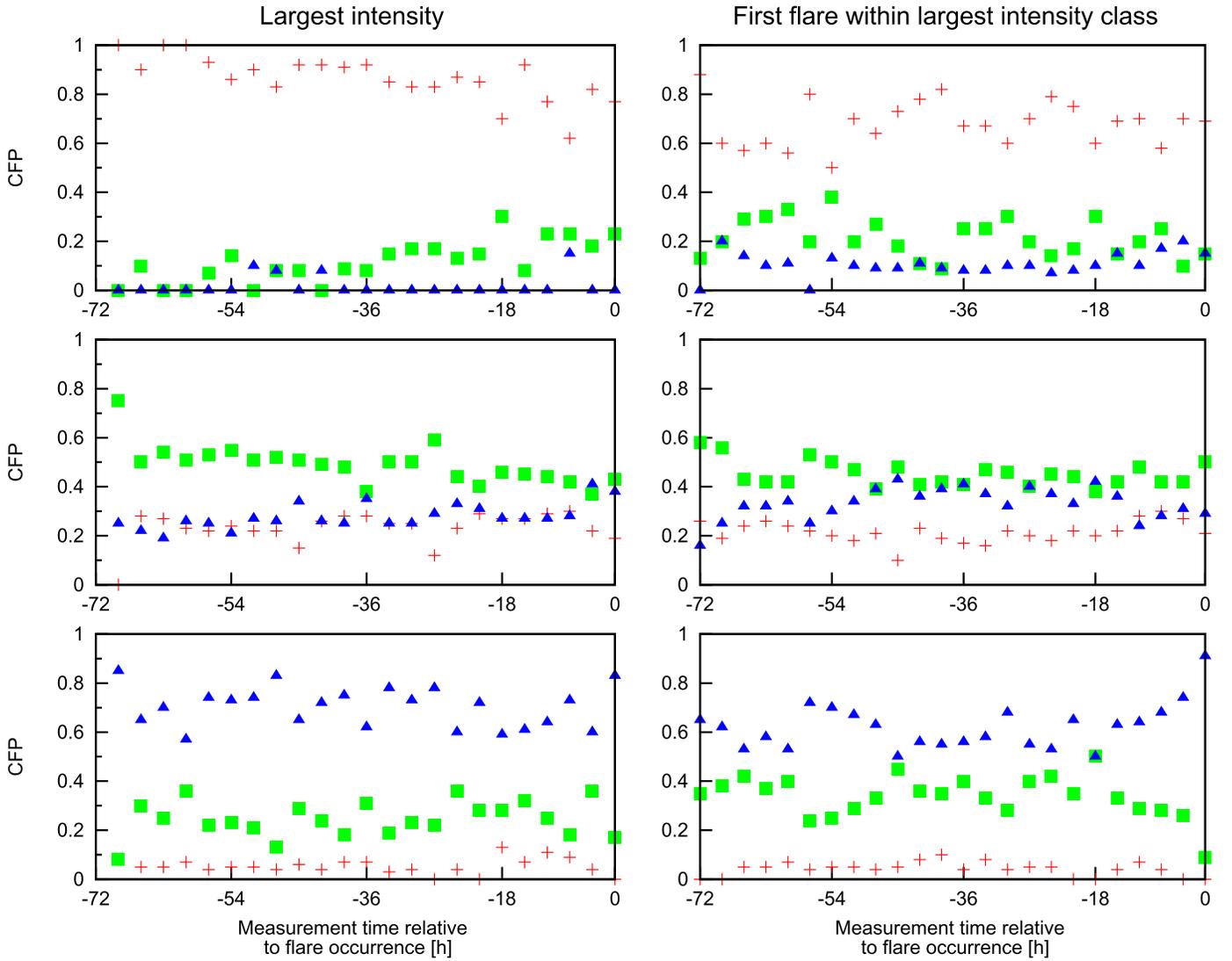


Figure 6. Conditional flare probability, CFP, of G_S : the figures show the temporal variation of the conditional probability of flare activity. The left and right columns are for the highest-intensity flare and for the first flare within the highest-intensity flare-class of ARs. The red/green/blue symbols are conditional probabilities for X/M/C flares in terms of *GOES* classification. The upper, middle, and lower panels depict data for the bands $7.5 \leq \log(G_S)$; $6.5 \leq \log(G_S) \leq 7.5$, and $5.5 \leq \log(G_S) \leq 6.5$. The reference point of the x -axis is the same as that of Figure 5.

Table 1

CFP of S_{l-f} for the Case of the Highest-intensity Flare (Highest) and for the Case of the First Flare within the Highest-intensity Flare-class (First) Produced by the ARs

Intensity Class	Time Interval (hr)	$S_{l-f} \leq 1$		$1 \leq S_{l-f} \leq 3$		$3 \leq S_{l-f} \leq 13$	
		Highest	First	Highest	First	Highest	First
X	0–24	0.85	0.73	0.28	0.25	0.06	0.05
	24–48	0.89	0.7	0.24	0.24	0.08	0.07
	48–72	0.82	0.65	0.22	0.23	0.12	0.08
M	0–24	0.15	0.27	0.46	0.49	0.29	0.27
	24–48	0.11	0.3	0.43	0.48	0.33	0.3
	48–72	0.18	0.35	0.43	0.48	0.4	0.33
C	0–24	0	0	0.26	0.26	0.65	0.68
	24–48	0	0	0.33	0.29	0.59	0.64
	48–72	0	0	0.35	0.29	0.48	0.59

highest-intensity flare cases. In the case of the first flare within the highest-intensity flare-class produced by the ARs 18 flares out of 116 happened with a fast CME and 98 events show slow or no CME. For the same reasons as in Section 5, let us analyze

the two upper panels of Figures 3 and 4. These panels clearly support the idea that a fast CME may occur when $\log(S_{l-f})$ is smaller than 0.4 and the value of $\log(G_{DS})$ is larger than about 6.5 a day before the estimated flare onset time.

Table 2CFP of G_S for the Case of the Highest-intensity Flare (Highest) and for the Case of the First Flare within the Highest-intensity Flare-class (First) Produced by the ARs

Intensity Class	Time Interval (hr)	$7.5 \leq \log(G_S)$		$6.5 \leq \log(G_S) \leq 7.5$		$5.5 \leq \log(G_S) \leq 6.5$	
		Highest	First	Highest	First	Highest	First
X	0–24	0.79	0.69	0.26	0.23	0.06	0.03
	24–48	0.88	0.71	0.23	0.19	0.04	0.06
	48–72	0.95	0.65	0.21	0.22	0.06	0.04
M	0–24	0.19	0.18	0.43	0.44	0.28	0.32
	24–48	0.11	0.2	0.49	0.43	0.24	0.37
	48–72	0.04	0.25	0.55	0.49	0.23	0.34
C	0–24	0.02	0.13	0.31	0.33	0.67	0.66
	24–48	0.02	0.09	0.28	0.38	0.72	0.57
	48–72	0.01	0.1	0.24	0.29	0.71	0.63

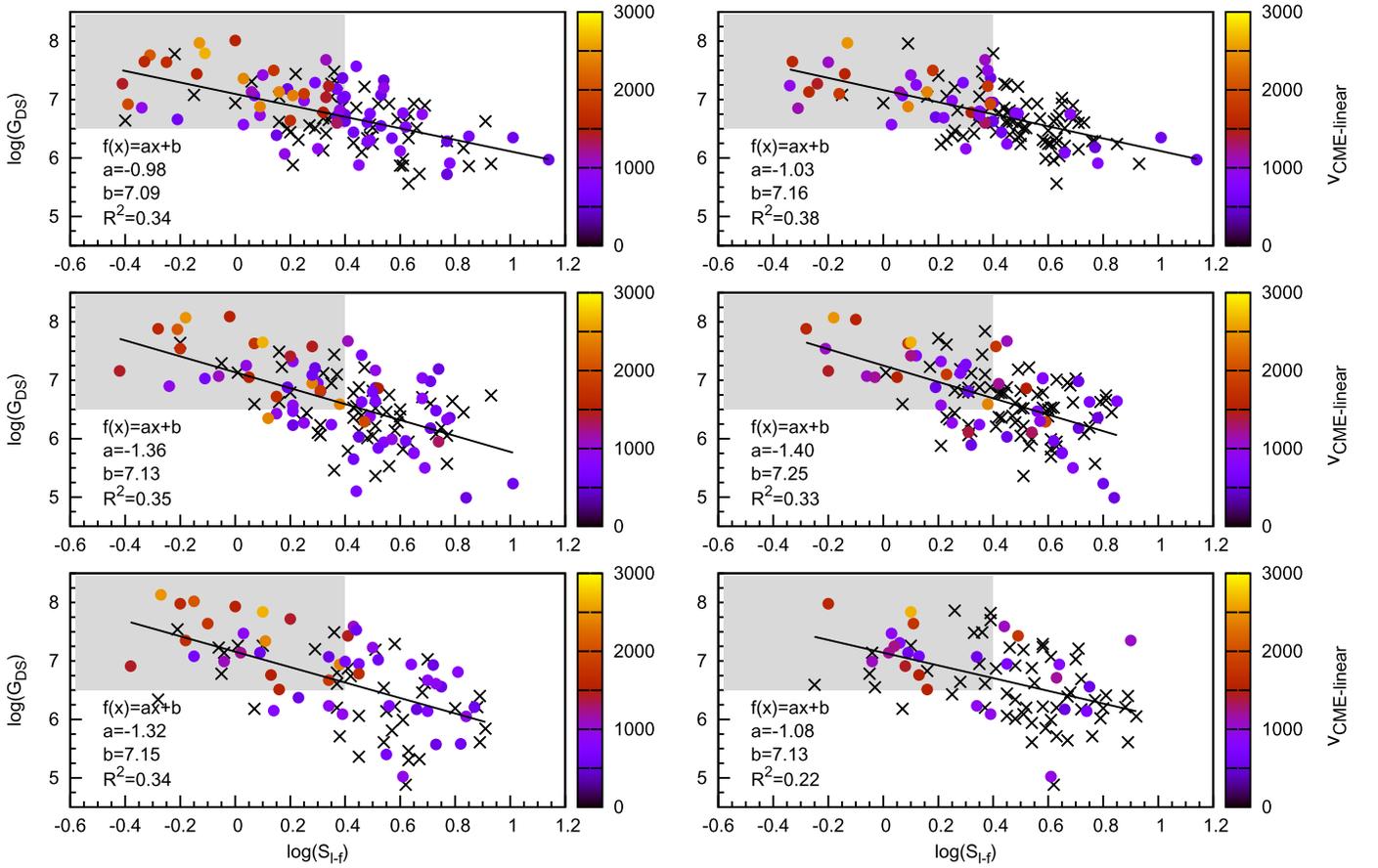


Figure 7. Relationship between S_{l-f} and G_{DS} in logarithmic scaling in the case of the highest-intensity flare (left column) and the first flare within the highest-intensity flare-class produced in the ARs (right column). The upper panels are distributions 1 day before the flare occurs, the middle panels 2 days before, and the lower panels 3 days before. The gray rectangle represents the interval $\log(S_{l-f}) \leq 0.4$ and $\log(G_{DS}) \geq 6.5$. Note that v_{lin} of the CME is the same CME speed representation as in Figures 3 and 4. Crosses/colored circles refer to flares associated without/with CME.

Next, Figure 7 demonstrates that one may establish a connection between S_{l-f} and G_{DS} with a linear fit in logarithmic scaling. We determined for how many ARs $\log(S_{l-f})$ is smaller than 0.4 and concurrently $\log(G_{DS})$ is larger than about 6.5 in cases of (i) the highest-intensity flare of ARs and (ii) the first flare within the highest-intensity flare-class produced in the ARs. Interestingly, it was found that all flares with fast CME events belong to the domain bounded by $\log(S_{l-f}) \leq 0.4$ and $\log(G_{DS}) \geq 6.5$ (this region is highlighted by a gray rectangle in Figure 7) a day before flare onset. It is

also fair to say that if, a day before, $\log(S_{l-f})$ is larger than a threshold value, 0.4, and $\log(G_{DS})$ is smaller than another threshold, i.e., about 6.5, there is no accompanying fast CME. However, about 25% of flares with fast CME were not in the gray range two/three days before the events. Thus, we suggest that these newly introduced parameters may be able to provide more accurate further information on the flare activity (and productivity) of ARs only a day before flare onset. In addition, we also demonstrate that under certain conditions a flare could be accompanied by a fast CME that may be predicted a day or

even 2–3 days before flare onset. In the two cases, i.e., the highest-intensity flare/the first flare within the highest-intensity flare-class produced by the ARs, we also examined what portion of all flares with fast CME fall into the range of $\log(S_{l-f}) \leq 0.4$ and $\log(G_{DS}) \geq 6.5$. We found that $\sim 30\%$ of flare events occurred with fast CME(s) within this range of $[S_{l-f}, G_{DS}]$.

6. DISCUSSION

The two suggested complexity parameters, i.e., the sum of the horizontal magnetic gradient (G_S) (and its daily averaged derivative, i.e., G_{DS}) and the separation parameter (S_{l-f}), are potential indicators of the measures of non-potentiality of the magnetic fields of sunspot groups. These parameters may be suitable new prediction indicators besides the traditional (e.g., Zürich, McIntosh, Mount Wilson) classification schemes that are characterized by alpha-numerical parameters based on morphological data of sunspots and their magnetic fields. The newly defined parameters seem to be viable, as demonstrated by testing them with the most detailed sunspot database currently available, the SDD sunspot catalog (Győri et al. 2011; Győri 2015).

Qahwaji & Colak (2007) developed a similar-purpose automated tool based on the McIntosh classification for flare forecasting; however, an advantage of the methods presented above is that the risks of flare and CME can also be assessed directly from the values of S_{l-f} and G_S (G_{DS}) without the need for classification.

The sum of the horizontal magnetic gradient (G_S) is a quantity similar to the so-called effective connected magnetic field (B_{eff}) analyzed by Georgoulis & Rust (2007). The B_{eff} method developed by Georgoulis and Rust is based on magnetograms. Our method, however, employs sunspots, white light, and magnetic data. Sunspots are discrete units instead of the continuous magnetic field distributions of magnetograms, and they are locations of high flux densities. G_S and B_{eff} are determined for all spot-pairs of opposite polarities within ARs and they both characterize the magnetic complexity of ARs. Lower limits of these parameters identify when it is unlikely for major flare(s) to occur in subsequent time intervals. If B_{eff} is less than 250/750 G then M/X-class flare(s) cannot occur; if B_{eff} is larger than 1600/2100 G then there is a high probability for M/X-class flare(s) to occur 12 hr before onset. Here, we also estimate lower and upper boundaries of G_S 48 hr before onset. When $\log(G_S)$ is smaller than 5.5 then ARs do not seem to produce X-class flare(s), and if $7.5 \leq \log(G_S)$ then the CFP of X-class flare(s) is $\sim 75\%$ in the subsequent 48 hr.

Guo et al. (2006) examined the so-called effective distance (d_E), whose meaning is comparable to that of our separation parameter introduced in this work. However, instead of using sunspots they have derived the effective distance from magnetograms. In their study, the centers of the leading and following parts are computed as the flux-weighted centers of two opposite-polarity regions from the magnetograms, while we derive this from photospheric intensity data. Guo et al. (2006) found that d_E is a quantitative measure of the McIntosh classification; therefore, it gives a good correlation between the magnetic complexity of ARs and flare class/CME speed in 24 AR cases. We also found a good relationship between S_{l-f} and flare class/CME speed in 116 AR cases. Nevertheless, the present separation parameter has the further advantage that S_{l-f}

provides limiting values and CFP for flare and CME forecasting (see below).

We propose that the application of the two parameters (S_{l-f} and G_S) may be complementary to other forecast methods. Here, we calculated the CFP of G_S and S_{l-f} for a period of three days prior to the event, sampled every three hours, to determine the probability of flare activity using the two approaches (i.e., the highest-intensity flare and that of the first flare within the highest-intensity flare-class produced by the ARs). We can draw the following conclusions from the two approaches: if $S_{l-f} \leq 1$, there is more than 70% CFP of the flare being X-class in the following two days. If $1 \leq S_{l-f} \leq 3$ then M-class flare(s) could be the highest-intensity increase (more than 45% likely). If $3 \leq S_{l-f} \leq 13$ the C-class flare(s) may be the main flare intensity class, with more than 60% likelihood in the next 48 hr.

We also estimated the CFP of G_S , and can draw the following conclusions: if $5.5 \leq \log(G_S) \leq 6.5$ then there is more than $\sim 60\%$ CFP that C energetic class flaring may be the largest flare intensity in the following two days. If $6.5 \leq \log(G_S) \leq 7.5$ then the chances of M-class being the expected highest intensity are about $\sim 45\%$ for the next two days. Finally, if $\log(G_S)$ is larger than 7.5 then the CFP for X-class flare(s) to develop is more than 70% within 48 hr. Importantly, we emphasize that: (i) X-class flare(s) do not occur if $13 \leq S_{l-f}$ and $\log(G_S) \leq 5.5$; (ii) the above-mentioned CFPs are indeed lower limits.

The two parameters introduced here can also be used in parallel for CME forecasting but, unfortunately, the quick fluctuations of G_S itself do not seem to allow this proxy to be chosen as a single parametric value for forecasting in a practical and simple way. Therefore, we suggest to determine its daily average of the sum of horizontal magnetic gradient, i.e., G_{DS} . Moreover, S_{l-f} and G_{DS} should actually be tracked simultaneously *together* in order to estimate the linear speed of a CME. Flares with accompanying fast CMEs (i.e., $1000 \text{ km s}^{-1} \leq v_{\text{lin}}$) are only found within 24 hr when $\log(S_{l-f}) \leq 0.4$ and $\log(G_{DS}) \geq 6.5$. It is also worth mentioning that in only $\sim 30\%$ of all events is the flare associated with a fast CME in the range of $\log(S_{l-f}) \leq 0.4$ and $\log(G_{DS}) \geq 6.5$; therefore, we need to search for additional precursor(s) of the slow/fast CMEs. So, if $\log(S_{l-f}) \geq 0.4$ or $\log(G_{DS}) \leq 6.5$, there is no accompanying fast CME in the following 24 hr.

A reliable forecast procedure is unlikely to be based on a single physical quantity. An accurate variation of probability should include simultaneous parallel methods (and parameters) for which the two proxies suggested above may be helpful, among (or complementary to) other methods. In the future, we would like to automate the identification of the weighted horizontal magnetic gradient applied by Korsós et al. (2015a). Therefore, here we have tried to find more reliable forecast parameters that could narrow down the identification of larger intensity flaring ARs with an associated fast CME.

M.B.K. and R.E. are grateful to Science and Technology Facilities Council (STFC) UK and to the Hungarian Academy of Science (HAS) for the support received. M.B.K. is grateful to the University of Sheffield for the support received. The authors also acknowledge the support received from the Royal Society, UK (ref nr: IE141493) for part of this work. The authors also express their gratitude toward their colleagues at DHO and Christopher J. Nelson and Michel S. Ruderman (both

at University of Sheffield, UK) for a number of useful discussions and improving the manuscript. Finally, we thank the unknown referee for the constructive comments, suggestions, and careful reading.

REFERENCES

- Bloomfield, D. S., Higgins, P. A., McAteer, R. T. J., & Gallagher, P. T. 2012, *ApJL*, **747**, L41
- Bornmann, P. L., & Shaw, D. 1994, *SoPh*, **150**, 127
- Carmichael, H. 1964, in Proc. AAS-NASA Symp. on the Physics of Solar Flares, NASA-SP 50, ed. W. N. Hess (Washington, DC: NASA, Science and Technical Information Division), 451
- Colak, T., & Qahwaji, R. 2008, *SoPh*, **248**, 277
- Evans, R. M., Pulkkinen, A. A., Zheng, Y., et al. 2013, *SpWea*, **11**, 333
- Georgoulis, M. K., & Rust, D. M. 2007, *ApJL*, **661**, L109
- Gosling, J. T., Hildner, E., MacQueen, R. M., et al. 1976, *SoPh*, **48**, 389
- Guo, J., Zhang, H., Chumak, O. V., & Liu, Y. 2006, *SoPh*, **237**, 25
- Győri, L. 2015, *SoPh*, **290**, 1627
- Győri, L., Baranyi, T., & Ludmány, A. 2011, *IAUS*, **273**, 403
- Hale, G. E., Ferdinand, E., Nicholson, S. B., & Joy, A. H. 1919, *ApJ*, **49**, 153
- Harrison, R. A. 1995, *A&A*, **304**, 585
- Hirayama, T. 1974, *SoPh*, **34**, 323
- Howard, T. A., & Harrison, R. A. 2013, *SoPh*, **285**, 269
- Kiepenheuer, K. O. 1953, (The Sun. Chicago, IL: Univ. Chicago Press), 322
- Kopp, R. A., & Pneuman, G. W. 1976, *SoPh*, **50**, 85
- Korsós, M. B., Baranyi, T., & Ludmány, A. 2014, *ApJ*, **789**, 107 (Paper I)
- Korsós, M. B., Gyenge, N., Baranyi, T., & Ludmány, A. 2015b, *JApA*, **36**, 111
- Korsós, M. B., Ludmány, A., Erdélyi, R., & Baranyi, T. 2015a, *ApJL*, **802**, L21
- Künzel, H. 1960, *AN*, **285**, 271
- Lee, K., Moon, Y.-J., Lee, J.-Y., Lee, K.-S., & Na, H. 2012, *SoPh*, **281**, 639
- McIntosh, P., Simon, P. A., Heckman, G., & Shea, M. A. 1986, in Solar Terrestrial Predictions, ed. P. A. Simon, G. Heckman, & M. A. Shea (Boulder, CO: NOAA), 357
- McIntosh, P. S. 1990, *SoPh*, **125**, 251
- Qahwaji, R., & Colak, T. 2007, *SoPh*, **241**, 195
- Sammis, I., Tang, F., & Zirin, H. 2000, *ApJ*, **540**, 583
- Sturrock, P. A. 1966, *Natur*, **211**, 695
- Waldmeier, M. 1938, *ZA*, **16**, 276
- Waldmeier, M. 1947, *Publ. Zürich*, Band IX, Heft 1
- Warwick, C. S. 1966, *ApJ*, **145**, 215
- Webb, D. F., & Howard, T. A. 2012, *SoPh*, **9**, 3
- Yashiro, S., Akiyama, S., Gopalswamy, N., & Howard, R. A. 2006, *ApJL*, **650**, L143