The Astrophysical Journal, 885:35 (9pp), 2019 November 1 https://doi.org/10.3847/1538-4357/ab45e7
© 2019. The American Astronomical Society. All rights reserved.

Ensemble Forecasting of Major Solar Flares with Short-, Mid-, and Long-term Active Region Properties

Daye Lim[1], Yong-Jae Moon[1], Eunsu Park[1], Jongyeob Park[2], Kangjin Lee[1,3], Jin-Yi Lee[4], and Soojeong Jang[2]
[1] School of Space Research, Kyung Hee University 1732, Deogyeong-daero, Giheung-gu, Yongin-si Gyeonggi-do 17104, Republic of Korea; moonyj@khu.ac.kr
[2] Korea Astronomy and Space Science Institute 776, Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea
[3] Electronics and Telecommunications Research Institute 218, Gajeong-ro, Yuseong-gu, Daejeon 34129, Republic of Korea
[4] Department of Astronomy & Space Science, Kyung Hee University 1732, Deogyeong-daero, Giheung-gu, Yongin-si Gyeonggi-do 17104, Republic of Korea
Received 2019 July 30; revised 2019 September 5; accepted 2019 September 17; published 2019 October 29

Abstract

We apply an ensemble technique for major flare prediction by considering short-, mid-, and long-term active region (AR) properties and their relative contributions. For this, we consider magnetic parameters from *Solar Dynamics Observatory*/Helioseismic and Magnetic Imager and flare lists from *Geostationary Operational Environmental Satellites*. In this study, we simultaneously consider flaring rates during short- (1 day), mid- (several days), and long-term (several years) timeframes. In our model, the predicted rate is given by a weighted combination of the three rates such that the sum of their weights is 1. We calculate the Brier skill scores (BSSs) for investigating prediction performance and weights of these three terms to provide optimal results. The BSS (0.22) of the model with only long-term properties is higher than that with only short-term (0.07) or mid-term (0.08) properties. When short-/mid-term properties are additionally considered, the BSS is improved to 0.28/0.24. Our model has the best performance (BSS = 0.29) when all terms are considered, and their relative contributions to the short-, mid-, and long-term rates are 20%, 20%, and 60%, respectively. In addition, the model with three terms is more effective at predicting major flares in strong ARs. In view of the energy storage and release process, long-term magnetic properties may indicate the storage of magnetic free energy, while short- and mid-term flare history may reflect a recent trend of energy release process. Our results suggest that the performances of other existing flare models based on long-term properties should be improved by considering short- and/or mid-term flare history.

Unified Astronomy Thesaurus concepts: The Sun (1693); Space weather (2037); Solar flares (1496); Solar magnetic fields (1503)

1. Introduction

Solar flares release an enormous amount of energy stored in magnetic fields in the form of electromagnetic radiation with various wavelengths. Such energetic events affect space-weather hazards, which can hinder human activities and satellite operations (Tsurutani et al. 2005; Schwenn 2006; Bocchialini et al. 2018). Thus, predicting solar flare occurrence plays an increasingly important role. Most solar flares erupt from active regions (ARs), which are areas of complex and intense photospheric magnetic fields. Thus, the various characteristics (e.g., morphological information and magnetic parameters from line-of-sight and vector magnetic fields) of ARs have been investigated and used for predicting flares (Hale et al. 1919; McIntosh 1990; Bornmann & Shaw 1994; Sammis et al. 2000; McAteer et al. 2005; Cui et al. 2006; Leka & Barnes 2007; Qahwaji & Colak 2007; Schrijver 2007; Li et al. 2008; Colak & Qahwaji 2009; Yu et al. 2009; Yuan et al. 2010; Ahmed et al. 2013; Huang & Wang 2013; Li & Zhu 2013; Falconer et al. 2014; Bobra & Couvidat 2015; Barnes et al. 2016; McCloskey et al. 2016; Liu et al. 2017; Nishizuka et al. 2017, 2018; Raboonik et al. 2017; Huang et al. 2018; Leka et al. 2018; Park et al. 2018; Lim et al. 2019).

Flare probabilities depending on AR magnetic properties have been studied by several authors. Following Giovanelli (1939), Gallagher et al. (2002) developed a flare prediction system that provides daily flare probabilities based on Poisson statistics using the McIntosh classification from 1988 to 1996. Barnes et al. (2007) provided flare probabilities based on Bayes's theorem using magnetic parameters such as the vertical current, current helicity, twist parameter α, and magnetic shear angles for discriminating flaring and flare-quiet ARs from 2001 to 2004. Falconer et al. (2011) predicted flare probabilities from an empirical relationship between the flare event rates and a proxy of magnetic free energy based on line-of-sight magnetograms from the *Solar and Heliospheric Observatory* (*SOHO*; Domingo et al. 1995)/Michelson Doppler Imager (MDI; Scherrer et al. 1995) from 1996 to 2004. The Poisson probabilities using the McIntosh classification from 1996 to 2010 were also reported in Lee et al. (2012). Barnes et al. (2016) compared a number of prediction algorithms using common data sets consisting of line-of-sight magnetic field and continuum intensity images from 2000 to 2005 using *SOHO*/MDI data. They showed that no one method clearly outperformed all others, which may be due to the strong correlations among the AR parameters. Lim et al. (2019) studied forecasting methods of the daily probabilities of major flares based on the relationships between 10 photospheric magnetic parameters and flaring rates. The magnetic parameters are taken from Space-weather HMI Active Region Patch (SHARP; Bobra et al. 2014) data made by the *Solar Dynamics Observatory* (*SDO*; Pesnell et al. 2012)/Helioseismic and Magnetic Imager (HMI; Scherrer et al. 2012). Among the parameters, the total unsigned parameters achieved better performance for predicting flares over the test period (about 2.5 yr). However, the models using total unsigned parameters show relatively low probabilities for strong ARs. These magnetic properties are mostly based on long-term data (typically several years or even more). The long-term-based model can predict the average activity with high statistical

Table 1
Descriptions and Formulas of the Six SHARP Magnetic Parameters Used in This Paper

Keyword	Description	Formula
TOTUSJH	Total unsigned current helicity	$H_{C,\mathrm{total}} = \sum\lvert B_z \cdot J_z\rvert$
TOTUSJZ	Total unsigned vertical current	$J_{z,\mathrm{total}} = \sum\lvert J_z\rvert\, dA$
TOTPOT	Total photospheric magnetic free energy density	$\rho_{\mathrm{tot}} = \sum(\boldsymbol{B}_{\mathrm{obs}} - \boldsymbol{B}_{\mathrm{pot}})^2\, dA$
USFLUX	Total unsigned magnetic flux	$\Phi = \sum\lvert B_z\rvert\, dA$
ABSNJZH	Absolute value of the net current helicity	$H_{C,\mathrm{abs}} = \lvert\sum B_z \cdot J_z\rvert$
SAVNCPP	Sum of the net current emanating from each polarity	$J_{z,\mathrm{sum}} = \lvert\sum^{B_z^+} J_z dA\rvert + \lvert\sum^{B_z^-} J_z dA\rvert$

Note. Constant terms are omitted.

stability, and this predicted flaring activity can be underestimated for strong ARs with many major flares.

The past flare occurrence history of ARs, together with the magnetic properties, has also been considered for predicting flares (e.g., Wheatland 2004). Falconer et al. (2012) examined the relationships between the free magnetic energy proxy and prior flaring rates using the data from 1996 to 2004 collected by *SOHO*/MDI. They showed that ARs with major flares on the previous day have much greater chances of producing major flares the next day. Shin et al. (2016) developed a flare forecast model using the McIntosh sunspot group, Mount Wilson magnetic classifications, sunspot areas, daily maximum flare flux, and weighted total flux of the previous day from 1996 to 2004. Nishizuka et al. (2017) presented a forecast model based on machine-learning algorithms using several features, including magnetic properties and flaring history from 2010 to 2014 collected by *SDO*/HMI. They showed that previous flare activity (such as the flare number in the AR and the maximum X-ray intensity of the previous day) is the most important feature. Jonas et al. (2018) considered the magnetic parameters from *SDO*/HMI, a timeline of the flare history, and EUV images of 80% of data from 2010 to 2014. They found that photospheric vector magnetic field data, combined with flaring history, yield the best performance. Leka et al. (2018) used magnetic parameters and the flare peak soft X-ray fluxes of 6, 12, and 24 hr prior from 2010 to 2017 collected by *SDO*/HMI. They showed that these prior flare peak fluxes are among the top-performing parameters for predicting flares.

An interesting approach is considering a combination of different forecasts, which is known as *ensemble forecasting*. Guerra et al. (2015) presented the first ensemble prediction model for major solar flares using four forecast models based on AR magnetic properties. They investigated combinations of weights for each model to optimize the Heidke Skill Score and found that a linear combination of various methods can improve the probabilistic prediction. Knipp (2016) summarized results in which ensemble studies exhibited the potential to improve space-weather forecasting. Murray (2018) addressed the importance of ensemble techniques for operational space-weather forecasting.

In this study, we apply the ensemble technique for flare forecasting by considering both recent flaring information and long-term magnetic properties. We simultaneously consider flare occurrence rates during short- (1 day), mid- (several days), and long-term (several years) properties. Through this study we address the three main following questions. First, how is the performance of our prediction model improved when all three terms are considered, compared to when only a single term is used? Second, what are the weights of all terms that provide the best prediction performance? Third, how well does the model with three terms predict major flares in strong solar ARs? To our knowledge, the present study is the first to consider these three rates and their relative contributions for predicting major (M- and X-class) flares. This paper is organized as follows. Section 2 describes the data. The method is explained in detail in Section 3. Results and a discussion are presented in Section 4. A summary and the conclusions are presented in Section 5.

2. Data

2.1. GOES *X-Ray Flares*

Geostationary Operational Environmental Satellites (*GOES*) have measured solar X-rays in the passband ranges of 1–8 Å and 0.5–4 Å. Solar X-ray flares are classified according to the peak flux of X-rays with wavelength bands ranging from 1 to 8 Å as measured by *GOES*. We use *GOES* major X-ray flare (M1.0 or greater) data from 2010 May to 2018 April, and their locations are identified via the Lockheed Martin Solar and Astrophysics Laboratory (LMSAL).[5] Our data include 448 M-class flares and 27 X-class flares. Our data have information on 262 event days (with major flares) and 2658 non-event days (without major flares). There were 1440 M-class flares and 94 X-class flares in solar cycle (SC) 23 (from 1996 August to 2008 December). This cycle has 922 event days and 3580 non-event days.

2.2. SDO/HMI

SDO, which is the first space-weather mission in NASA's Living With a Star (LWS) Program, was launched in 2010. Among the three scientific instruments of *SDO*, HMI (Scherrer et al. 2012; Schou et al. 2012; Hoeksema et al. 2014) provides full-disk Dopplergrams and vector magnetograms with high cadence.

The HMI team developed a set of derivative data called Space-weather HMI Active Region Patch (SHARP) data (Bobra et al. 2014). These data contain automatically identified AR patches and magnetic parameters that summarize the size, distribution, and non-potentiality of vector magnetic fields in each AR with a 12 minute cadence. These parameters may be useful for forecasting space-weather events.

Various SHARP magnetic parameters have been used for flare forecasting (Bobra & Couvidat 2015; Liu et al. 2017) based on machine-learning algorithms. Among the SHARP parameters, we consider only six parameters that have linear Pearson correlation coefficients (CCs) relating them with

[5] https://www.lmsal.com/solarsoft/latest_events_archive.html

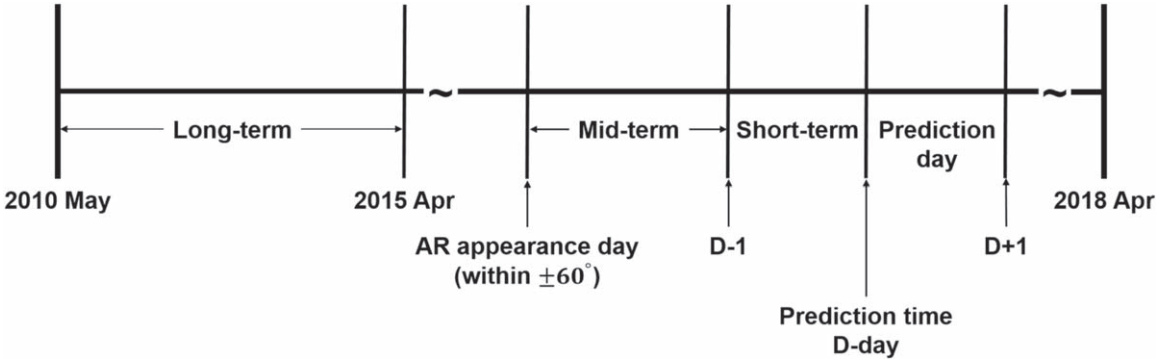

Figure 1. A schematic description of the short-, mid-, and long-term timeframes.

flaring rates higher than 0.86, as described in Section 3. Descriptions and formulas of these parameters are listed in Table 1.

We use 00:00 TAI definitive SHARPs in cylindrical equal area (CEA) coordinates (hmi.sharp_720s_cea data series) when their longitudes are within ±60° of the central meridian and the corresponding six magnetic parameters from the Joint Science Operations Center (JSOC).[6] According to Hoeksema et al. (2014), the number of high-confidence pixels in SHARP data decreases significantly beyond ±60° of the central meridian.

3. Method

For this study, we consider the prediction of daily major flare occurrences of 4724 ARs from 2015 April 21 to 2018 April 30. The predicted flare occurrence rates are determined considering the short-, mid-, and long-term history. Figure 1 is a schematic description of what constitutes the short-, mid-, and long-term timeframes. Descriptions of the predicted rate terms are as follows.

First, the short-term flare occurrence rate (R_S) is assumed as the number of major flares that occurred in a given AR within one day before the prediction time. We expect that this reflects the temporary evolution of a given AR. Kubo et al. (2017) showed that such recent information is useful for forecasting flares, which agrees with the daily persistence model in which the flare occurrence of the next will be the same as that of the present day (Zirin & Marquette 1991). Second, the mid-term rate (R_M) is assumed as the mean flaring rate of the day for which a given AR appearance occurs (within ±60° of the central meridian) to one day before the prediction time. We expect that this reflects the mean evolution of a given AR. The mid-term period ranges from 1 to 9 days for a given AR.

Third, the flare history for a given AR is not applicable for the case of the long-term period (i.e., longer than the mid-term period) because in such cases the AR goes behind the Sun and disappears within several weeks or months (van Driel-Gesztelyi & Green 2015). Thus, we instead consider past ARs with similar magnetic properties to the given ARs we want to forecast. We use the long-term flaring rate based on empirical models of six magnetic parameters as described in Lim et al. (2019). These models are based on 11040 ARs from 2010 May 1 to 2015 April 20. For each parameter, the AR data are divided into 50 subgroups with equal numbers of ARs. Then, the average parameter values and numbers of major flares within a day from each daily AR are determined in each subgroup. Each mean major flare occurrence rate (R_i) of the ith group (G_i) is given by

$$R_i = \frac{\#\text{ of major flares of } G_i}{\#\text{ of ARs of } G_i}. \quad (1)$$

To find relationships between the magnetic parameters and flaring rates, these rates are plotted as a function of each parameter in log–log scales. Our study considers power-law functions as shown in Figures 1(a)–(f) in Lim et al. (2019). The fitting function is given by

$$\log(R_L) = a\log(x) + b, \quad (2)$$

where R_L is assumed to be a long-term flare occurrence rate in this study, x is a parameter value of the AR, a is a power-law slope, and b is a fitting constant. The fitting coefficients and uncertainties of six SHARP parameters are shown in Table 2 in Lim et al. (2019). They investigated the dependence of the binning size of the subgroups by considering four cases of binning size: 10, 20, 50, and 100. For all four cases, power-law functions are well fitted with the data, and the differences between these cases are very small. Thus, they use power-law fitting coefficients and uncertainties with a binning size of 50, and we follow their approach in the present study.

In our model, we consider linear combinations of the three predicted flare occurrence rates as an ensemble technique, given by

$$R = w_S R_S + w_M R_M + w_L R_L \quad \text{for } w_S + w_M + w_L = 1, \quad (3)$$

where R is a rate from linear combinations, and w_S, w_M, and w_L are the weights of the short-, mid-, and long-term rates, respectively. The National Aeronautics and Space Administration (NASA) Flare Scoreboard currently implements an ensemble average across multiple forecast models. In this study, we investigate the combination of weights that provides the best performance, as did Guerra et al. (2015).

From these rates, we calculate the flare probabilities using the Poisson distribution (Wheatland 2000; Moon et al. 2001; Bloomfield et al. 2012; Lee et al. 2012). The probability (P_R) that at least one flare occurs within a day is given by

$$P_R = 1 - \exp(-R), \quad (4)$$

where R is the rate as mentioned above.

[6] http://jsoc.stanford.edu/

Table 2
Descriptions of Seven Cases

Case	Description
Case 1	$w_S = 1, w_M = 0, w_L = 0$
Case 2	$w_S = 0, w_M = 1, w_L = 0$
Case 3	$w_S = 0, w_M = 0, w_L = 1$
Case 4	$w_S + w_M = 1, w_L = 0$
Case 5	$w_M + w_L = 1, w_S = 0$
Case 6	$w_S + w_L = 1, w_M = 0$
Case 7	$w_S + w_M + w_L = 1$

4. Results and Discussion

To investigate the prediction performance and find the weights of each term that provide the best performance, we calculate the Brier skill score (BSS) for seven cases, as described in Table 2. This measure has been used to verify the performance of probabilistic forecast models (Wheatland 2005; Barnes et al. 2007, 2016; McCloskey et al. 2018). The BSS, which represents the relative skill compared to the model using the climatological event rate during the testing interval, is given by

$$\mathrm{BSS} = \frac{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2 - \frac{1}{N}\sum_{i=1}^{N}(\overline{O_i} - O_i)^2}{0 - \frac{1}{N}\sum_{i=1}^{N}(\overline{O_i} - O_i)^2}, \quad (5)$$

where Pi is the predicted probability and O_i is the observation that an event occurred ($O_i = 1$) or did not occur ($O_i = 0$). When a perfect prediction occurs, the BSS is 1, and exhibiting "no-skill" compared to the climatological forecast results in 0. Several forecast studies have mentioned the necessity of considering uncertainties for verification measures (Barnes et al. 2016; Kubo et al. 2017; Leka et al. 2018) because these measures are calculated from finite numbers of samples. To estimate the uncertainties of BSS, we use a bootstrap method that accounts for random errors using resampling. We consider the size of the resampled data, which is the difference between the number of non-events (4686) and events (38) during a prediction period. We create a resampled data set that was randomly selected with a size of 4448. Then, we calculate the BSS from the resampled data for all seven cases. This process is iterated 1000 times and the uncertainties are estimated using the standard deviation of the resampled values of the BSS.

Among the three rate terms, only the long-term rate depends on the magnetic parameters. We calculate the BSS from the long-term flaring rate (Case 3) for 4724 ARs and the results are as follows: 0.22 ± 0.005 for $H_{C,\mathrm{total}}$, 0.17 ± 0.004 for $J_{z,\mathrm{total}}$, 0.12 ± 0.007 for ρ_{tot}, 0.12 ± 0.004 for Φ, 0.14 ± 0.007 for $H_{C,\mathrm{abs}}$, and 0.15 ± 0.007 for $J_{z,\mathrm{sum}}$. From now on, we only consider $H_{C,\mathrm{total}}$ as it gives the best prediction results among the six magnetic parameters. This parameter also achieved the highest F-score in the classification between flaring events and non-flaring events (Bobra & Couvidat 2015).

We calculate the BSS values of seven cases for the weights at 0.05 intervals. Figure 2 shows the maximum BSSs and their uncertainties for all cases. In the case of only a single term, the BSS of Case 3 (0.22 ± 0.005) is higher than those of Cases 1 (0.07 ± 0.005) and 2 (0.08 ± 0.005). Kubo et al. (2017) showed that major flare forecasts in the Regional Warning Center (RWC) of Japan, which are operated under the influence of human judgment based on the X-ray flux, sunspot magnetic

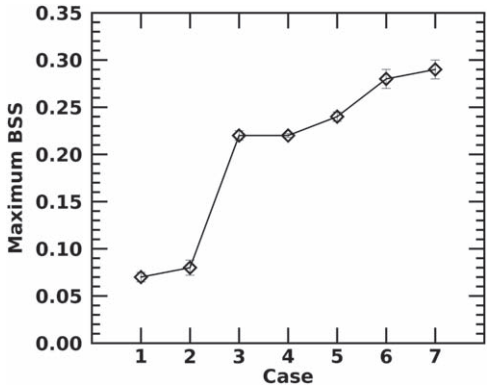

Figure 2. Comparison of the maximum BSSs for seven cases and their uncertainties.

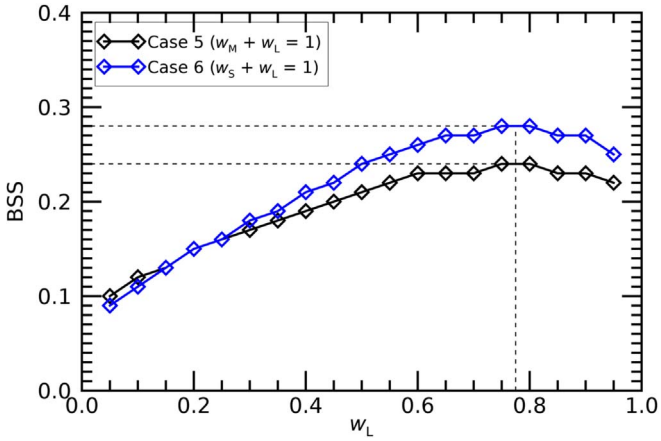

Figure 3. BSSs as a function of long-term weight w_L for Cases 5 and 6.

field configurations, and chromospheric brightening of the AR, are slightly better than those of the daily persistence model (Case 1 in this study). This BSS variation between the three cases demonstrates the statistical stability for the long-term based model. This result could be the reason why many forecast models consider the properties based on long-term periods rather than only daily persistence. When short- or mid-term periods are additionally considered, the BSSs of Cases 5 (0.24 ± 0.005) and 6 (0.28 ± 0.01) are higher than that of Case 3. The improvement of Case 5, compared to the result of Case 3, implies that AR evolution during the period that the AR located on the disk (a maximum of 14 days) plays an important role. The horizontal magnetic field variations between opposite polarity spots (such as converging–diverging motions in 2–3 days before strong flares) have been shown to have a relationship with strong flare occurrences (Korsós et al. 2015; Tlatov et al. 2018). Unlike the cases for only short- or mid-term periods, short-term periods are more effective at predicting flaring tendency than mid-term periods when they are supported by the stable long-term periods. Kilcik et al. (2018) reported that flaring ARs with the three modified Zurich classes (A, B, and C) rapidly evolve over one day to higher classes, unlike non-flaring ARs. The combination of short- and long-term histories results in a higher performance than the combination of mid- and long-term histories. When all terms are considered together (Case 7: 0.29 ± 0.01), the best prediction results are obtained. This result is consistent with Guerra et al. (2015), who showed that linear combination

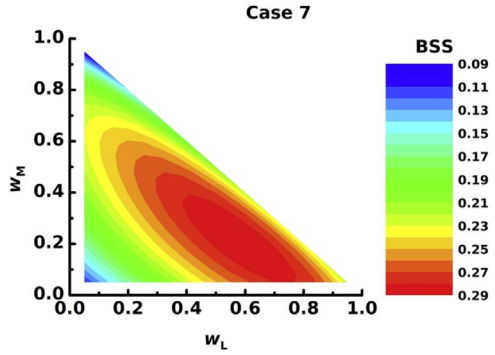

Figure 4. BSS contour map as a function of two weights, w_M and w_L, for Case 7.

methods can improve the probabilistic prediction. This implies that the recent (short- and mid-term) flaring tendency with the long-term flaring rate is effective at predicting flares.

Figure 3 shows variations of the BSSs as a function of w_L for the two cases when short- or mid-term histories are additionally considered alongside the long-term history. The long-term history contributes about 78% to the best performance. When all terms are considered together, the distribution of the BSSs as a function of the two weights is shown in Figure 4. In this case, the relative contributions from the short-, mid-, and long-term rates to the best prediction performance are 20%, 20%, and 60% on average, respectively. The longer-term results contribute more in each case, which could be interpreted as the statistical stability tending to increase as the size of the data sample becomes larger.

We also investigate how the results differ between Cases 3 and 7 as described by a reliability plot and its sharpness (McCloskey et al. 2018). The reliability plot shows the observed occurrence rates against the predicted occurrence rates. When a perfect forecast occurs, all points in the reliability plot lie on the diagonal line. The sharpness diagram represents the relative frequency with which events have been predicted with different levels of probability. Figure 5 shows that the points tend to be close to the diagonal line for both cases. The maximum predicted probability is 54% when only the long-term rate is considered. This increases up to about 84% when short- and mid-term rates are additionally considered. These results imply that the long-term-based model can predict only the average activity with statistical stability, and this predicted flaring activity can be underestimated for strong ARs with many major flares.

Other probability forecast verification measures include the relative operating characteristic (ROC) curve, in which the probability of detection (POD) is plotted as a function of probability of false detection (POFD), and the area under the ROC curve (AUC). When the curve passes through the upper left corner, i.e., when the AUC is 1, the forecast is perfect. The ROC curve plots are conditioned on observations, while the reliability plots are conditioned on the forecast. ROC curves are therefore a good companion to reliability plots. The AUCs of Cases 3 and 7 as in Figure 6 are about 0.97, implying that the ability of the forecast to discriminate between events and non-events for Case 7 is similar to that for Case 3.

As an example, Figure 7 shows a comparison between the major flare occurrence probabilities and the *GOES* X-ray flux in the two ARs. For this, we consider the two ARs with the most flaring during our test period (total 81 major flares): HARP 5983 (NOAA AR 12422) with 12 total major flares from 2015 September 24 to 30 and HARP 7115 (NOAA AR 12673) with 25 total major flares from 2017 August 31 to September 8. For Case 7, we calculate their flare probabilities using the weights from Figure 4. In HARP 5983, the probabilities of two cases are lower than 13% until 2015 September 26. Note that there was no major flare during this period. On September 27, the highest probability is found in Case 3 (only long-term considered), which is about 16%, because there was no major flare on previous days. From September 28 to 30, the probabilities of Case 7 are about two times higher than those of Case 3. There were several major flares for each day. In HARP 7115, the probabilities of the two cases are almost 0% until September 3 and there was no major flare. On September 4, the highest probability (18%) is shown in Case 3, as similarly shown in HARP 5983. The probabilities of Case 7 from September 5 to 8 are higher than those of Case 3. There were several major flares for each day. Case 3 provides the highest probability in the first flaring day for both of the ARs. This indicates that flare prediction based on magnetic properties from long-term data is useful for such a sudden flare occurrences. When flares occurred consecutively, Case 7 gives the best prediction, as we expected. These results imply that the consideration of short-, mid-, and long-term rates together is more useful, especially when there are consecutive active flaring activities such as homologous flares.

The magnetic properties related to AR size and non-potential magnetic fields represent the potential to generate flares in terms of energy storage. In the general view of a process of energy storage and release processes, the stored energy may be released all at once or be released in a fragmented way (see Figure 1 in Aschwanden et al. 1998). This means that an AR with a large amount of magnetic free energy may have a major flare and/or several small flares. In this respect, the release process is likely to be very complicated. There have been many models proposed for the triggering processes of solar eruptions, such as kink instability (Fan & Gibson 2003), break-out model (Antiochos et al. 1999), emerging magnetic flux (Chen & Shibata 2000), tether-cutting reconnection (Moore et al. 2001), and sunspot rotating (Su et al. 2007). Recent observations have shown that major solar flares were triggered by single or multiple triggering processes (Chen 2011; Benz 2017; Hernandez-Perez et al. 2019). Moon et al. (2001) found that there are no systematic relationships between peak fluxes of flares and their waiting-time distributions, which is not consistent with the simple scenario of storage and release (Rosner & Vaiana 1978). Their results support the idea that the solar corona is in a self-organized critical (SOC) state. In this respect, flare history (especially short-term and mid-term) may represent the tendency (or characteristics) of recent energy release processes for a given AR. Summing up, long-term magnetic properties may indicate magnetic free energy is stored by some storage processes such as shearing motions and flux emergences, while short- and mid-term flare history may reflect the recent trends of energy release processes. In view of this fact, we can understand that the model considering short-, mid-, and long-term flare history is much better than the model considering only long-term magnetic properties.

Even though we used long-term data from 2010 to 2015, this does not cover the whole SC 24. The period used for the long-term rates includes the ascending and maximum phase of SC 24. Our prediction period almost coincides with the descending phase of SC 24. Lee et al. (2016) investigated the dependence of flare occurrence on the SC phase for the McIntosh classes.

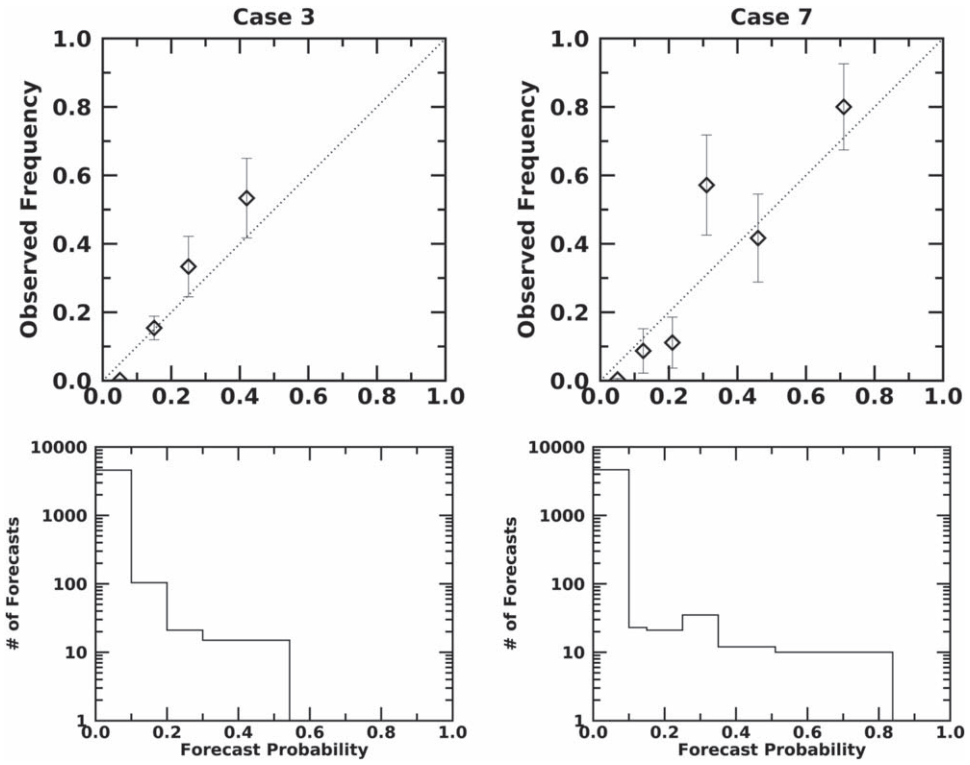

Figure 5. Comparison of reliability plots (top) and their sharpness (bottom) between Cases 3 (left) and 7 (right). The reliability plots are the observed occurrence rates against the predicted occurrence rates. The uncertainties of the observed rates for each probability bin are represented by error bars. When a perfect prediction occurs, all the points in the reliability plot lie on the diagonal line. The sharpness diagram represents the relative frequency with which an event has been predicted with different levels of probability.

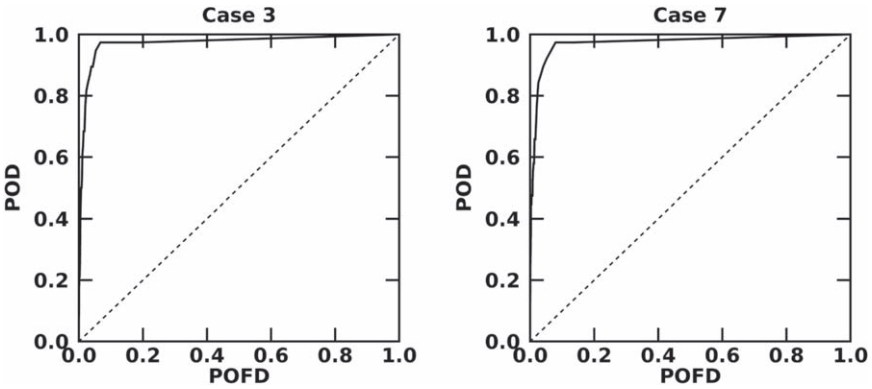

Figure 6. Comparison of ROC plots in which the POD is plotted as a function of POFD between Cases 3 (left) and 7 (right). When a perfect forecast occurs, the curve passes through the upper left corner.

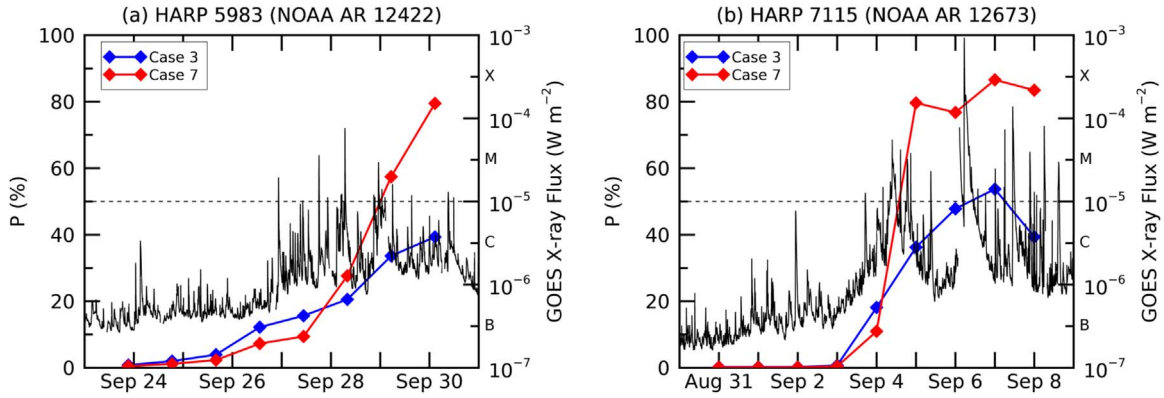

Figure 7. Comparison of predicted flare probabilities (P) and *GOES* X-ray fluxes for the two ARs with the most flaring during our test period: (a) HARP 5983 (NOAA AR 12422) from 2015 September 24 to 30 and (b) HARP 7115 (NOAA AR 12673) from 2017 August 31 to September 8. The predicted flaring probabilities (left axis) are displayed as filled-in diamonds for Cases 3 (blue) and 7 (red). The *GOES* X-ray fluxes (right axis) during the period are shown as black solid lines.

They found that the flare occurrence rates of the descending phase are noticeably higher than those during the other phases. In the future, we expect SC phase effects will be taken into account as another long-term effect.

5. Summary and Conclusion

We have investigated the prediction of major flares based on short-, mid-, and long-term AR properties and their relative contributions from 2015 April to 2018 April. For this, we considered the magnetic parameters, which are characterized by the field distribution and non-potentiality, from the *SDO*/HMI and the flare list from *GOES*. In this study, the predicted rate is given as combinations of the three weighted rates such that the sum of weights is 1, considered as an ensemble technique. For determining the weights of the three terms that provide the best prediction performance, we estimated the BSSs as a function of the weights for seven cases. Reliability plots and sharpness diagrams were also considered to compare the prediction results.

The major results of this study are as follows. The total unsigned current helicity provides a better BSS than the other magnetic parameters (Table 1). This parameter has also been considered a useful parameter in previous forecasting studies (Bobra & Couvidat 2015; Liu et al. 2017). The BSS (0.22) of the model with only the long-term rate is higher than that (0.07) with only the short-term or (0.08) with only the mid-term rate. When short-/mid-term rates are additionally considered alongside the long-term rate, the BSSs are improved to 0.28/0.24. Our model has the best performance (BSS = 0.29) when all three terms are considered, and the relative contributions from short-, mid-, and long-term rates are 20%, 20%, and 60%, respectively. Sharpness diagrams show that the model considering all three terms can predict a wider range of flaring probabilities than that with only the long-term rate. ROC curves show that the ability of the forecast to discriminate between events and non-events for Case 7 is similar to that for Case 3. Comparisons of our model considering the *GOES* X-ray fluxes for given ARs show that the model with short-, mid-, and long-term rates seems to be more effective when predicting major flares in strong solar ARs.

Most previous flare forecasting models have used AR magnetic properties and/or flaring history based on only long-term data (Falconer et al. 2011; Bloomfield et al. 2012; Barnes et al. 2016). In this study, we considered a model with short-term and mid-term in addition to long-term flare history data. This provides improved prediction performance over the case considering only long-term rate. We also found the relative contributions of the three rates that provide the best prediction performance among our models. Our results suggest the possibility that the performances of other existing flare models based on long-term properties can be improved by also considering short- and/or mid-term properties. In this respect, we expect that other existing models can be accomplished through their optimizing the relative contributions. The relative contributions in our study may change depending on the forecasting period and/or data sets (Barnes et al. 2016; Nishizuka et al. 2017).

We thank the referee for their helpful comments. This work was supported by the BK21 plus program through the National Research Foundation (NRF) funded by the Ministry of Education of Korea, the Basic Science Research Program through the NRF funded by the Ministry of Education (NRF-2016R1A2B4013131, NRF-2016R1A6A3A11932534, NRF-2019R1A2C1002634), the NRF of Korea Grant funded by the Korean Government (NRF-2013M1A3A3A02042232), the Korea Astronomy and Space Science Institute under the R&D program supervised by the Ministry of Science, ICT and Future Planning, the Korea Astronomy and Space Science Institute under the R&D program Development of a Solar Coronagraph on International Space Station (Project No. 2019-1-850-02) supervised by the Ministry of Science, ICT and Future Planning, and the Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (2018-0-01422, Study on analysis and prediction technique of solar flares). The data used here are provided courtesy of NASA/*SDO* and the HMI science team, as well as the *GOES* team.

Appendix

So far we have presented the results of seven cases for the total unsigned current helicity $H_{C,\mathrm{total}}$. Figure 8 shows the maximum BSSs and their uncertainties of the seven cases for the other five parameters in Table 1.

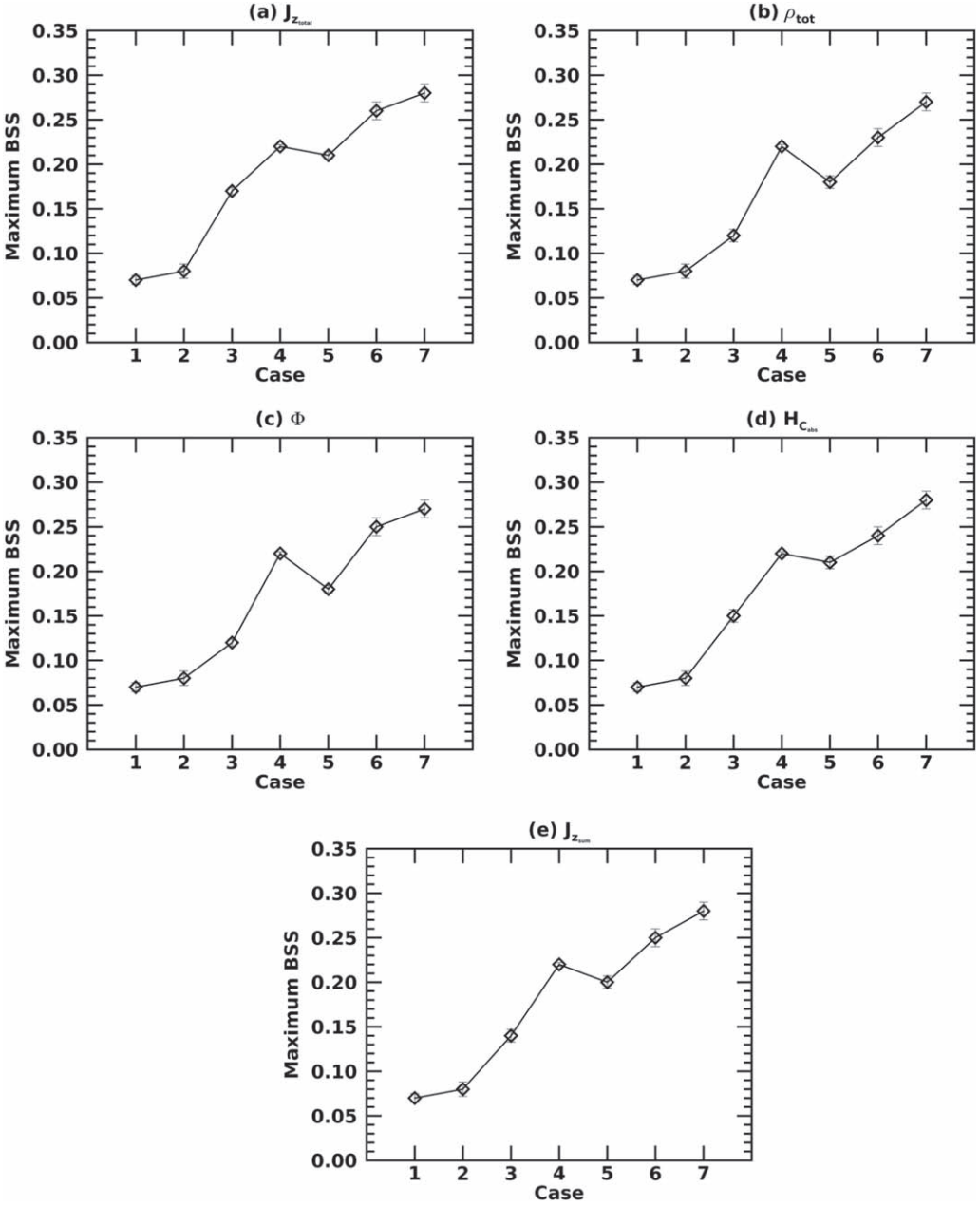

Figure 8. Same as Figure 2, but for the other magnetic parameters: (a) the total unsigned vertical current $J_{z,\mathrm{total}}$, (b) the total photospheric magnetic free energy density ρ_{tot}, (c) the total unsigned magnetic flux Φ, (d) the absolute value of the net current helicity $H_{C,\mathrm{abs}}$, and (e) the net current emanating from each polarity $J_{z,\mathrm{sum}}$.

ORCID iDs

Daye Lim https://orcid.org/0000-0001-9914-9080
Yong-Jae Moon https://orcid.org/0000-0001-6216-6944
Eunsu Park https://orcid.org/0000-0003-0969-286X
Jongyeob Park https://orcid.org/0000-0002-1063-9129
Jin-Yi Lee https://orcid.org/0000-0001-6412-5556

References

Ahmed, O. W., Qahwaji, R., Colak, T., et al. 2013, SoPh, 283, 157
Antiochos, S. K., Devore, C. R., & Klimchuk, J. A. 1999, ApJ, 510, 485
Aschwanden, M. J., Dennis, B. R., & Benz, A. O. 1998, ApJ, 497, 972
Barnes, G., Leka, K. D., Schrijver, C. J., et al. 2016, ApJ, 829, 89
Barnes, G., Leka, K. D., Schumer, E. A., & Della-Rose, D. J. 2007, SpWea, 5, S09002
Benz, A. O. 2017, LRSP, 14, 2
Bloomfield, D. S., Higgins, P. A., McAteer, R. T. J., & Gallagher, P. T. 2012, ApJL, 747, L41
Bobra, M. G., & Couvidat, S. 2015, ApJ, 798, 135
Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549
Bocchialini, K., Grison, B., Menvielle, M., et al. 2018, SoPh, 293, 75
Bornmann, P. L., & Shaw, D. 1994, SoPh, 150, 127
Chen, P. F. 2011, LRSP, 8, 1
Chen, P. F., & Shibata, K. 2000, ApJ, 545, 524
Colak, T., & Qahwaji, R. 2009, SpWea, 7, S06001
Cui, Y., Li, R., Zhang, L., He, Y., & Wang, H. 2006, SoPh, 237, 45
Domingo, V., Fleck, B., & Poland, A. I. 1995, SoPh, 162, 1
Falconer, D. A., Barghouty, A. F., Khazanov, I., & Moore, R. L. 2011, SpWea, 9, S04003
Falconer, D. A., Moore, R. L., Barghouty, A. F., & Khazanov, I. 2012, ApJ, 757, 32
Falconer, D. A., Moore, R. L., Barghouty, A. F., & Khazanov, I. 2014, SpWea, 12, 306
Fan, Y., & Gibson, S. E. 2003, ApJL, 589, 105
Gallagher, P. T., Moon, Y.-J., & Wang, H. 2002, SoPh, 209, 171

Giovanelli, R. G. 1939, ApJ, 89, 555
Guerra, J. A., Pulkkinen, A., & Uritsky, V. M. 2015, SpWea, 13, 626
Hale, G. E., Ellerman, F., Nicholson, S. B., & Joy, A. H. 1919, ApJ, 49, 153
Hernandez-Perez, A., Su, Y., Veronig, A. M., et al. 2019, ApJ, 874, 122
Hoeksema, J. T., Liu, Y., Hayashi, K., et al. 2014, SoPh, 289, 3483
Huang, X., & Wang, H.-N. 2013, RAA, 13, 351
Huang, X., Wang, H., Xu, L., Liu, R., & Dai, X. 2018, ApJ, 856, 7
Jonas, E., Bobra, M., Shankar, V., Hoeksema, J. T., & Recht, B. 2018, SoPh, 293, 48
Kilcik, A., Yurchyshyn, V., Sahin, S., et al. 2018, MNRAS, 477, 293
Knipp, D. J. 2016, SpWea, 14, 52
Korsós, M. B., Ludmány, A., Erdélyi, R., & Baranyi, T. 2015, ApJL, 802, L21
Kubo, Y., Den, M., & Ishii, M. 2017, JSWSC, 7, A20
Lee, K., Moon, Y.-J., Lee, J.-Y., Lee, K.-S., & Na, H. 2012, SoPh, 281, 639
Lee, K., Moon, Y.-J., & Nakariakov, V. M. 2016, ApJ, 831, 131
Leka, K. D., & Barnes, G. 2007, ApJ, 656, 1173
Leka, K. D., Barnes, G., & Wagner, E. 2018, JSWSC, 8, A25
Li, R., Cui, Y., He, H., & Wang, H. 2008, AdSpR, 42, 1469
Li, R., & Zhu, J. 2013, RAA, 13, 1118
Liu, C., Deng, N., Wang, J. T. L., & Wang, H. 2017, ApJ, 843, 104
Lim, D., Moon, Y.-J., Park, J., et al. 2019, JKAS, 52, 133
McAteer, R. T. J., Gallagher, P. T., & Ireland, J. 2005, ApJ, 631, 628
McCloskey, A. E., Gallagher, P. T., & Bloomfield, D. S. 2016, SoPh, 291, 1711
McCloskey, A. E., Gallagher, P. T., & Bloomfield, D. S. 2018, JSWSC, 8, A34
McIntosh, P. S. 1990, SoPh, 125, 251
Moon, Y.-J., Choe, G. S., Yun, H. S., & Park, Y. D. 2001, JGR, 106, 29951
Moore, R. L., Sterling, A. C., Hudson, H. S., & Lemen, J. R. 2001, ApJ, 552, 833
Murray, S. A. 2018, SpWea, 16, 777
Nishizuka, N., Sugiura, K., Kubo, Y., et al. 2017, ApJ, 835, 156
Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., & Ishii, M. 2018, ApJ, 858, 113
Park, E., Moon, Y.-J., Shin, S., et al. 2018, ApJ, 869, 91
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Qahwaji, R., & Colak, T. 2007, SoPh, 241, 195
Raboonik, A., Safari, H., Alipour, N., & Wheatland, M. S. 2017, ApJ, 834, 11
Rosner, R., & Vaiana, G. S. 1978, ApJ, 222, 1104
Sammis, I., Tang, F., & Zirin, H. 2000, ApJ, 540, 583
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, SoPh, 162, 129
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, SoPh, 275, 207
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, SoPh, 275, 229
Schrijver, C. J. 2007, ApJL, 655, L117
Schwenn, R. 2006, LRSP, 3, 2
Shin, S., Lee, J.-Y., Moon, Y.-J., Chu, H., & Park, J. 2016, SoPh, 291, 897
Su, Y., Golub, L., van Ballegooijen, A., et al. 2007, PASJ, 59, S785
Tlatov, A. G., Abramov-Maximov, V. E., Borovik, V. N., & Opeikina, L. V. 2018, Ge&Ae, 58, 1087
Tsurutani, B. T., Judge, D. L., Guarnieri, F. L., et al. 2005, GeoRL, 32, L03S09
van Driel-Gesztelyi, L., & Green, L. M. 2015, LRSP, 12, 1
Wheatland, M. S. 2000, ApJL, 536, L109
Wheatland, M. S. 2004, ApJ, 609, 1134
Wheatland, M. S. 2005, SpWea, 3, S07003
Yu, D., Huang, X., Wang, H., & Cui, Y. 2009, SoPh, 255, 91
Yuan, Y., Shih, F. Y., Jing, J., & Wang, H.-M. 2010, RAA, 10, 785
Zirin, H., & Marquette, W. 1991, SoPh, 131, 149