



Can Solar Limb Flare Prediction Be Properly Made by Extreme-ultraviolet Intensities?

Jaewon Lee¹ , Yong-Jae Moon^{1,2} , Hyun-Jin Jeong² , Kangwoo Yi² , and Harim Lee² ¹School of Space Research, Kyung Hee University, Yongin, 17104, Republic of Korea; moonjy@khu.ac.kr²Department of Astronomy and Space Science, Kyung Hee University, Yongin, 17104, Republic of Korea

Received 2024 April 13; revised 2024 August 4; accepted 2024 August 5; published 2024 August 19

Abstract

We address the question of whether the solar limb flare prediction can be properly made by EUV intensity, which has less projection effects than solar white light and magnetogram data. We develop empirical and multilayer perceptron (MLP) models to forecast the probability of a major solar limb flare within a day. We use Solar Dynamics Observatory (SDO)/Atmospheric Imaging Assembly (AIA) 94 and 131 Å that have high correlations and large slopes with X-ray flare fluxes from 2010 to 2022. We select 240 flares stronger than or equal to the M1.0 class and located near the limb region (60° or more in heliographic longitude). For input data, we use the limb intensity as the sum of SDO/AIA intensities in the limb region and the total intensity of the whole image. We compare the model performances using metrics such as the receiver operating characteristic—area under the curve. Our major results are as follows. First, we can forecast major solar limb flare occurrences with only SDO/AIA 94 and/or 131 Å intensities. Second, our models show better probability prediction than the climatological model. Third, both empirical (AUC = 0.85) and MLP (AUC = 0.84) models have similar performances, which are much better than a random forecast (AUC = 0.50). Finally, it is interesting to note that our model can forecast the flaring probability of all 52 events during the test period, while the models in the NASA/CCMC flare scoreboard can forecast only 22 events. From the above results, we can answer that the solar limb flare prediction using EUV intensity can be properly made.

Unified Astronomy Thesaurus concepts: [The Sun \(1693\)](#); [Solar extreme ultraviolet emission \(1493\)](#); [Solar flares \(1496\)](#); [Neural networks \(1933\)](#)

1. Introduction

Solar flares are one of the most energetically significant eruptive phenomena in the solar atmosphere (Shibata & Magara 2011). Solar flares are probable causes of solar energetic particles (SEPs), along with coronal mass ejections (CMEs; Papaioannou et al. 2016). Solar flares can occur not only on the front side but also near the limbs and backside of the Sun. The source location of SEP-associated flares tends to be close to the western limb (Dierckx et al. 2015). On the other hand, flares occurring near the eastern limb are difficult to predict because the evolution of active regions (ARs) on the farside is unseen from the Earth (Pangestu et al. 2023). Hence, eastern limb flare may lead to sudden radio blackouts (Yasyukevich et al. 2018). In these senses, the effort to forecast the solar limb flare is crucial from a space weather perspective. Historically, one of the most commonly used data for solar flare prediction is McIntosh sunspot group classification (McIntosh 1990). This sunspot group classification is based on solar white light and magnetogram data. The National Oceanic and Atmospheric Administration (NOAA) currently provides the McIntosh classification through the Solar Region Summary (SRS). Most flare prediction models from the Community Coordinated Modeling Center (CCMC; Hesse et al. 2001) of the National Aeronautics and Space Administration (NASA) use McIntosh classification from SRS and magnetograms from the Helioseismic and Magnetic Imager (HMI; Schou et al. 2012) of Solar Dynamics Observatory (SDO; Pesnell et al. 2012; Leka et al. 2019, A-EFFORT;

Georgoulis & Rust 2007, AMOS; Lee et al. 2012, ASAP; Colak & Qahwaji 2009, ASSA; Hong et al. 2014, BOM; Steward et al. 2017, DAFFS, DAFFS-G; Leka et al. 2018b, MAG4; Falconer et al. 2011, MCSTAT, MCEVOL; Gallagher et al. 2002; Bloomfield et al. 2012; McCloskey et al. 2018, MOSWOC; Murray et al. 2017, NICT; Kubo et al. 2017, NJIT; Park et al. 2010, NOAA; Crown 2012, SIDC; Berghmans et al. 2005; Devos et al. 2014). However, near the solar limb, solar magnetograms and white light data are strongly affected by projection effects adding distortion. Despite attempts to correct these projection effects through coordinate transformations, the process introduces a significant amount of error (Wilkinson et al. 1989). The inaccuracy in predicting flare occurrences near the solar limb region is attributed to the limitations of existing models (Park et al. 2020) and much of the previous research is limited to ARs in the restricted longitudes: $\leq 30^\circ$ in Cui et al. (2006) and Huang et al. (2018); $\leq 45^\circ$ in Ahmed et al. (2013), Bobra et al. (2014), Zheng et al. (2019), and Deng et al. (2021); and $\leq 68^\circ$ in Wang et al. (2020).

Here we present solar limb flare models within the next 24 hr using solar ultraviolet (EUV) data to find answers to the following question: can solar limb flare prediction be properly made by EUV intensities? The difference in optical depth between traditional input data (magnetograms and white light data) of the flare prediction models and the EUV data may provide a breakthrough in predictions of solar limb flare. Magnetogram and white light data, which reflect solar photosphere (Boyer et al. 1985) and chromosphere (Hudson 1972), are optically thick, while solar EUV data, reflecting the solar corona (Battaglia & Kontar 2011), are optically thin. Thus, EUV data are expected to be less affected by projection effects near the solar limb compared to solar magnetograms and white light data, which results in fewer distortions in the information.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

In this Letter, we present empirical and multilayer perceptron (MLP; Gardner & Dorling 1998) models using solar EUV data to forecast the occurrence of major solar limb flares within the next 24 hr. This study is organized as follows. Specific data used in our study are covered in Section 2. The structures of the empirical and MLP models used in this study are presented in Section 3. Section 4 describes the prediction performance of each model and its comparison. Finally, in Section 5, we provide a brief summary and conclusion of our study.

2. Data

2.1. Data Selection

Our study focuses on major limb flare prediction. Thus, we need to define "major" and "limb" for the data selection. Solar flares are classified as A, B, C, M, and X based on the soft X-ray (0.1–0.8 nm) flux and X-class flares having the highest flux. In this study, we define major flares as those with a class of M1.0 or higher. A limb flare is defined when its source location is at a heliographic longitude of 60° or more. We have a couple of reasons for this longitudinal criterion. First, there are serious projection effects near the limb: misidentification of sunspot classification and magnetic field vectors (Venkatakrishnan & Gary 1989). Second, many flare studies and prediction models are based on the data within a heliographic longitude of 60° to minimize the projection effect near the solar limb region (Park et al. 2010; Falconer et al. 2011; Liu et al. 2012; Lim et al. 2019). Following these criteria, we collect flare data from the Heliophysics Event Knowledgebase (HEK; Hurlburt et al. 2012) provided by Lockheed Martin Solar and Astrophysics Laboratory (LMSAL). We use solar EUV data from the Atmospheric Imaging Assembly (AIA; Lemen et al. 2012) instrument on SDO. Among the seven EUV wavelengths (94, 131, 171, 193, 211, 304, and 335 Å) of SDO/AIA, we select 94 and 131 Å since they represent high-temperature coronal regions, and their peak intensities have stronger correlations having larger slopes with Geostationary Operational Environmental Satellite (GOES; Menzel & Purdom 1994) X-ray fluxes than other wavelengths (Van der Sande et al. 2022).

2.2. Data Collection and Preprocessing

The EUV data are collected from 2010 May 13 to 2022 December 31 at UT 00:00 each day, and if a major limb flare occurs within the subsequent 24 hr from the observation time, they are labeled as flaring samples. We remove the data whose data are partially saturated due to solar activity such as strong flares using the information given in their FITS header. Our models consider two types of solar EUV data: (1) total intensity, which is the sum of intensity for all pixels, not only in the solar disk, and (2) limb intensity, which is the sum of intensity for pixels in regions with heliographic longitude of 60° or more. The limb intensity also includes the outer region of the solar disk. We anticipate that our models can predict flare occurrences in the limb region by considering variations in limb intensity relative to total intensity. As a result, we can get a total of 4222 nonflaring data and 240 flaring data. We divide the data into train and test data sets to construct the empirical and MLP models, considering both flare occurrences and the solar cycle. For each year, April, August, and December are used for testing. February and June serve as the validation set for MLP models. The remaining months are used for training.

To accommodate computing resources, we downsample the original FITS data from 4096×4096 pixels to 1024×1024 pixels while fixing the solar radius to 392 pixels (Park et al. 2019). Then we divide all SDO/AIA pixel intensities by the exposure time to ensure consistent exposure conditions. We also perform the north-up and recentering processes during the resizing. To mitigate bias toward large values, we apply a logarithm transformation to both summed intensities.

Optical instruments like SDO/AIA suffer degradation in their sensitivity over time (Boerner et al. 2012). To consider long-term data, we should correct the degradation effect. Conventionally, degradation can be corrected by cross calibration between SDO/AIA and the Extreme Ultraviolet Variability Experiment (EVE; Woods et al. 2012) of SDO (Boerner et al. 2014). SDO/EVE is also calibrated with observations of NASA's sounding rockets (Wieman et al. 2016). This approach is not problematic when using data within the solar disk. However, utilizing data from outside the solar disk for multiple channels, there are considerable discrepancies.

Figure 1 shows the distribution of summed intensities at 94 (y-axis) and 131 Å (x-axis). After applying traditional degradation correction using the AIA response table (Boerner et al. 2014) version 10 from the Joint Science Operations Center (JSOC), the ratios of summed intensities inside the solar disk show a relatively constant at the two wavelengths across years (see panel (a) of Figure 1). However, when adding the intensities outside the disk, the ratios of the limb intensities at the two wavelengths for each year are divided into two groups, as shown in panel (b) of Figure 1. This implies that the ratio varies over time, which can potentially impact the performance of models using multichannel inputs. To address this issue, we cross-calibrate all data using the nonflaring data of 2011, which is close to the starting time of SDO operation. Fitting the data to the 2011 nonflaring data set shows a quite good coherent relationship, as shown in the right panel (c) of Figure 1, demonstrating that the combined use of multichannel data has been successfully made.

3. Method

3.1. Empirical and MLP Models

We develop two kinds of solar limb flare models: an empirical model and an MLP one. To construct the empirical model, we examine the total intensity and limb intensity of 94 and 131 Å for the training data set and divide them into several bins. Then we investigate the probability of flare occurrence in each bin, which is carried out by dividing the number of flaring data by the number of all samples for each bin. To provide meaningful probabilistic predictions, each bin should have the counts of sufficient samples. In this study, each bin contains a minimum of 24 samples, corresponding to 10% of the total flaring data. To satisfy this, the size of all bins is fixed at 0.5 [log(DN/s)]. We use the same bins of the training data set for test data. Out of a total of 4222 nonflaring data and 240 flaring data, 1077 and 52 of each were used for testing our empirical model performance.

Our MLP model takes the limb intensity and total intensity of 94 and 131 Å as input. It is developed using supervised learning, where the flare occurrence status serves as the label. We train MLP models using 2472 nonflare samples and 161 flare samples and validate them using 673 nonflare samples and 27 flare samples. We use the same test data as the empirical

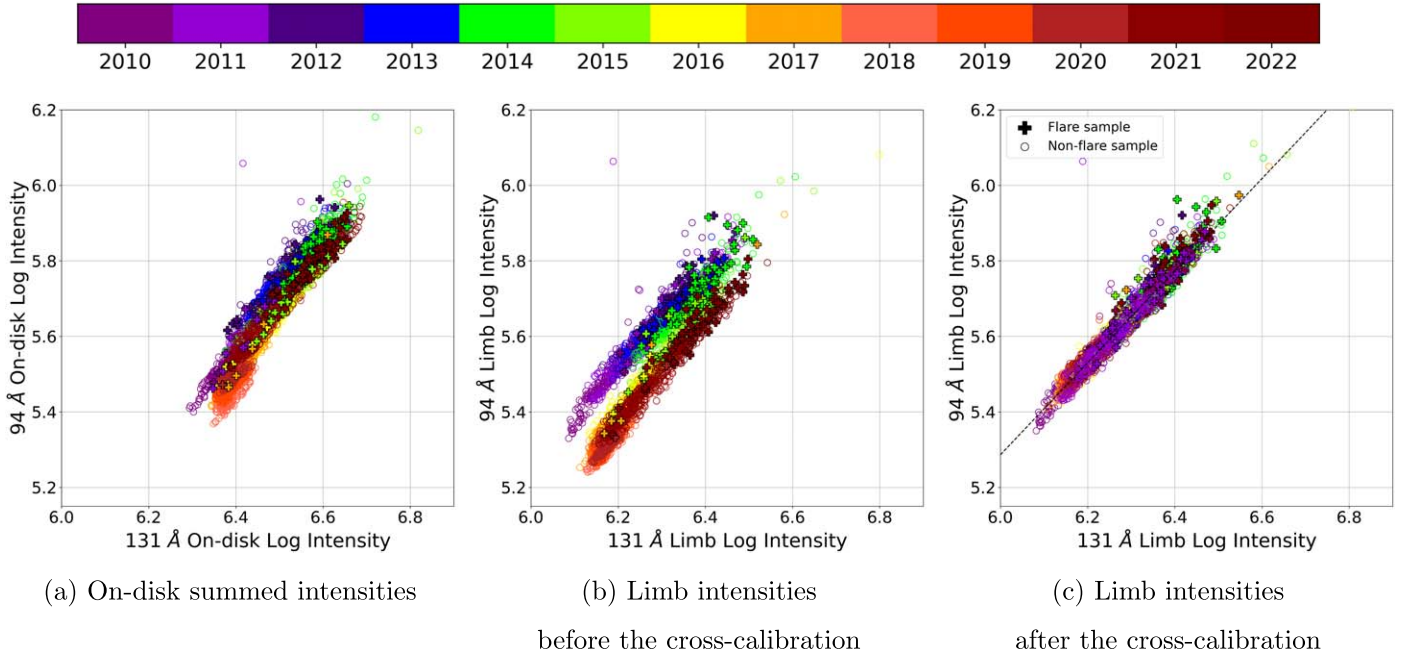


Figure 1. The distribution of summed intensities at 94 (y-axis) and 131 Å (x-axis). The data of each year are represented by a different color. Cross and circle symbols represent flaring and nonflaring samples, respectively. Panel (a) is the summation of pixels inside the solar disk. Panel (b) is the defined limb intensity and panel (c) is the cross-calibrated one by the 2011 nonflaring data set. All data are corrected by AIA response table version 10 and presented on a logarithmic scale.

model. Our MLP model consists of two layers with 128 nodes, batch normalization (Ioffe & Szegedy 2015), activation function, and dropout. Batch normalization and dropout prevent the model from biasing specific values or nodes. Nonlinear rectified linear unit (Nair & Hinton 2010) activation functions are used to deal with nonlinearity. The input data processed through these layers undergo a sigmoid function to generate a value between 0 and 1, representing the probability of a flare occurrence. If the model’s output probability is equal to or greater than the threshold, it is classified as a flare occurrence; otherwise, it is classified as nonflaring in a binary classification. The loss is calculated using the binary cross entropy (BCE; Cox 1958) loss function, which considers the model’s predicted results against the actual flare occurrence status. Through multiple epochs, the model updates the weights associated with connections between each node to reduce the BCE loss.

When the data set is imbalanced, the BCE loss tends to be biased toward the dominant set of data. In such cases, the probabilities predicted by the model are limited to a narrower range (e.g., $[0, 0.3]$) instead of $[0, 1]$. Additionally, since the metric results used to evaluate the model’s performance vary depending on the threshold, there is also a need to find an appropriate threshold. Meanwhile, the Score-Oriented Loss (SOL; Guastavino et al. 2022a; Marchetti et al. 2022) function, which simultaneously performs score maximization and minimization of loss during the training process, has no need to find a threshold where the metric is maximized. In the process of optimizing a metric that is effective for imbalanced data sets, the model’s probabilities would be distributed between 0 and 1. However, the probabilistic predictions made by a model trained with SOL are mainly clustered around 0 or 1, which would be like deterministic classification rather than probabilistic prediction. The BCE loss allows for a relatively continuous distribution of predicted probabilities. Since each approach has its own advantages and disadvantages, and the probabilistic prediction is

just as important as deterministic prediction for risk assessment in operations, we present both results: BCE loss and SOL. Here the optimal thresholds are obtained from the train data set for each metric, which will be described in the next section (3.2). We also have tried four types of models: (1) simple convolutional neural networks (Li et al. 2021) models using EUV images, (2) convolutional block attention module (Woo et al. 2018) architectures that incorporate attention mechanisms (Yohanandan et al. 2018), (3) Deep Flare Net (DeFN; Nishizuka et al. 2018), which is an MLP-based residual network (He et al. 2016), and (4) DeFN-R (Nishizuka et al. 2020) architectures. Since our simple MLP model is better than the other model architectures from the perspective of the Brier skill score (BSS; Brier 1950; Nishizuka et al. 2020, Section 3.2), we will mainly present the MLP model.

3.2. Metrics for Evaluation

The evaluation of our models uses various performance metrics commonly employed in many flare prediction studies. These include deterministic prediction metrics such as true skill statistics (TSS; Allouche et al. 2006), the Heidke skill score (HSS; Heidke 1926), and the receiver operating characteristic—area under the curve (ROC-AUC; Fawcett 2006). The probabilistic prediction metric such as BSS is also involved. TSS is given by

$$\text{TSS} = \frac{\text{TP}}{\text{TP} + \text{FN}} - \frac{\text{FP}}{\text{FP} + \text{TN}} = \text{POD} - \text{POFD}, \quad (1)$$

where TP, FN, FP, and TN are true positive (hit), false negative (miss), false positive (false alarm), and true negative (correct rejection). Here, POD and POFD represent the probability of detection, and the probability of false detection, respectively. TSS ranges from -1 to 1 , where 1 represents a perfect prediction. TSS is a comprehensive metric that evaluates prediction performance by considering both the POD and POFD. It is useful in binary

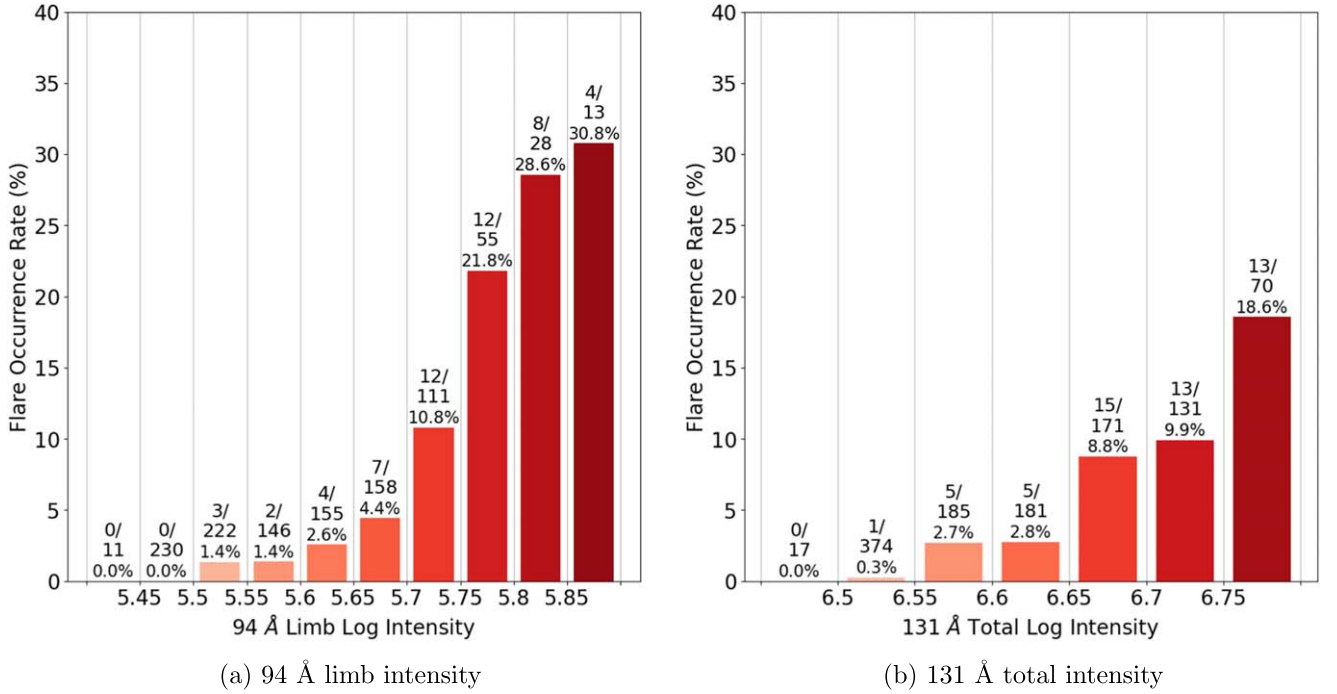


Figure 2. The flare occurrence rates of the test data set as a function of 94 Å limb intensity (a) and 131 Å total intensity (b). The ratios above each bin indicate the number of flaring events divided by that of the total ones.

classification tasks and exhibits robustness against imbalanced data sets (i.e., flare class) compared to other metrics (Guastavino et al. 2022b). On the other hand, HSS is defined by

$$\text{HSS} = \frac{2[(\text{TP} \times \text{TN}) - (\text{FP} \times \text{FN})]}{(\text{TP} + \text{FN})(\text{FN} + \text{TN}) + (\text{TP} + \text{FP})(\text{FP} + \text{TN})}, \quad (2)$$

HSS ranges from $-\infty$ to 1, and it represents a perfect prediction when it equals 1. HSS is useful as it measures the relative improvement of predictions compared to random guessing across various data sets (Hyvärinen 2014). ROC-AUC represents the area under the curve when plotted as the ROC curve with POFD on the x -axis and POD on the y -axis. It ranges from 0 to 1, the closer the value is to 1, the better the model is at classification. For a random classifier, ROC-AUC is 0.5. BSS is given by

$$\text{BSS} = \frac{\text{BS} - \text{BS}_C}{0 - \text{BS}_C} = 1 - \frac{\text{BS}}{\text{BS}_C}, \quad (3)$$

where BS is the Brier score of our model and BS_C is the BS of the climatological model. In this study, we use the climatological model as a reference model of BSS calculation, where the climatological model is a model that predicts the probability of event occurrence for all samples as equal to the event rate of the test data set. The BS and BS_C are expressed by

$$\text{BS} = \sum_{n=1}^N (p(y_n) - y_n^*)^2 \quad (4)$$

$$\text{BS}_C = \sum_{n=1}^N (S - y_n^*)^2, \quad (5)$$

where N , $p(y_n)$, and y_n^* correspond to the total number of samples, model's prediction, and observed value, respectively.

Here S indicates the actual event rate of the data. S is given by

$$S = \frac{\text{TP} + \text{FN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \quad (6)$$

BSS ranges from -1 to 1 , and the unity is only attainable in a perfect deterministic prediction. To compare with the probabilistic prediction of flare occurrences from models in the NASA/CCMC flare scoreboard, we used the BSS as the criterion for selecting the best model.

4. Result and Discussion

4.1. Comparison between Empirical and MLP Models

We combine two wavelengths (94 and 131 Å) and two types of intensities (total intensity and limb intensity) to construct various empirical and MLP models, comparing them using BSS to find the best model. The best empirical model utilizes only the limb intensity of 94 Å, yielding a BSS = 0.11.

Figure 2 shows the flare occurrence rates of the test data set as a function of 94 Å limb intensity and 131 Å total intensity. As shown in Figure 2, the probability distribution of 94 Å limb intensity is more contrasted than that of 131 Å total intensity. In the case of 94 Å limb intensity, the Pearson correlation coefficient (CC; Pearson 1895) with the flare occurrence ratio is 0.91 for the training data set and 0.94 for the test data set, respectively. Additionally, the CC of the flare occurrence probabilities between the train data set and test data set is approximately 0.99, demonstrating that the test data sets have a similar trend with the training ones.

The best MLP model (described in Section 3 and using BCE loss) is utilizing the limb intensity and total intensity of 94 and 131 Å, with a BSS = 0.10. It is interesting to note that the simplest MLP model shows the best performance. We think of a possible reason as follows. As observed in the empirical

Table 1

The Contingency Tables of Our Empirical and MLP Models with BCE or SOL, along with the Thresholds (@th) Optimized to TSS and HSS

	Empirical (@th = 5.65)	MLP (BCE) (@th = 0.07)	MLP (SOL) (@th = 0.42)
TSS optimized			
TSS	0.53	0.53	0.54
HSS	0.14	0.15	0.18
TP	43	42	39
FP	322	296	231
FN	9	10	13
TN	755	781	846
HSS optimized			
TSS	0.32	0.36	0.45
HSS	0.27	0.27	0.26
TP	19	22	28
FP	54	68	99
FN	33	30	24
TN	1023	1009	978

model, intensities at 94 and 131 Å of EUV wavelengths exhibit high correlations with flare occurrence probability. Hence, they may not require complex nonlinearity, and a simpler model structure can perform well. On the other hand, unlike the empirical model, the MLP model achieves the best BSS when both intensities of the two EUV wavelengths are used. It may be attributed to the ability of the MLP model's architecture to effectively extract features from each wavelength and combine them better, rather than simply combining the data into 2D data as in the empirical model. In Equation (3), the BSS is ultimately related to the ratio of mean square error (MSE) between our model and the reference climatological model, which forecasts all the instances as the flaring ratio of the test data set (0.048). Our models show better probabilistic predictions than the climatological model. It is also confirmed that there is not a significant difference in the performances of the empirical and MLP models from a probabilistic forecasting perspective.

We also compare the deterministic prediction performances of our empirical and MLP models. When examining the model's performance by changing the threshold of deterministic prediction, we find that HSS and TSS have slightly different thresholds at which they reach their maximum values.

Table 1 shows the best HSS and TSS of our empirical and MLP models with BCE or SOL, along with their corresponding thresholds. For TSS, the empirical model achieves TSS = 0.53 with a threshold of 5.65, while the MLP model with SOL reaches a maximum TSS = 0.54 with a probability threshold of 0.42. As for HSS, the empirical model obtains HSS = 0.27 with a threshold of 5.77, and the MLP model with BCE shows HSS = 0.27 with a probability threshold of 0.18. Even when optimized for TSS and HSS, MLP models using the SOL function show better performance in terms of HSS and TSS compared to the MLPs using BCE or empirical models. This is because the SOL function leads to less variation in the contingency table components (TP, FP, FN, TN) depending on the threshold, making it more robust. Both TSS and HSS show that there is no significant difference in the performance among the models.

Figure 3 shows the ROC-AUC of our models and random forecast. The AUC values of our empirical and MLP models

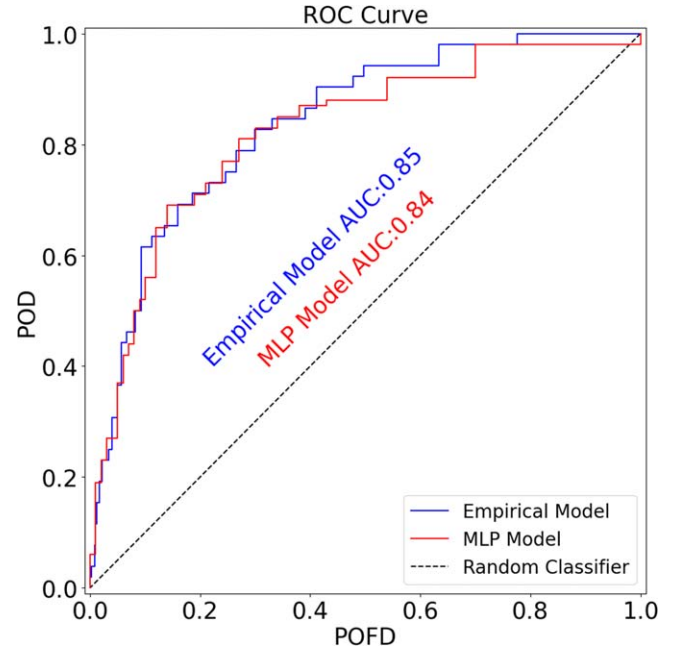


Figure 3. ROC curves of our empirical and MLP models. Both the empirical and the MLP models show better performance than a random forecast, and the difference in performance between the two models is not significant.

are 0.85 and 0.84, respectively, while the random forecast has AUC = 0.50. Similar to the previous results, there is no significant difference in performance between the two models.

4.2. Comparison between Our Model and NASA/CCMC Flare Scoreboard

Our model can provide not only deterministic prediction expressed as “yes” or “no” but also probabilistic prediction as previously shown with BSS. For the more objective evaluations, as well as the performance comparisons between the empirical and MLP models, we compare the results of our model with those of the existing models in the NASA/CCMC flare scoreboard. Most existing models in the NASA/CCMC flare scoreboard forecast the probability of flares with M-class or higher within 24 hr like our model. It is noted that few models show the probability of limb flares. Hence, we match the source locations of limb flares during the test period according to the NASA/CCMC flare scoreboard. For our comparison, we use our MLP model with BCE loss since its results are similar to those of the empirical one and give us relatively continuous probabilities. If at least one model in the NASA/CCMC flare scoreboard provides a prediction for ARs near the solar limb regions at UT 00:00 before the flare onset, we conduct a comparison. If multiple predictions are available for the NASA/CCMC flare scoreboard, we compare the average probability of those predictions with our model's probability. It is very impressive to say that out of a total of 52 events, 30 events (about 60%) are not provided by the NASA/CCMC flare scoreboard. Among these, 24 events occur near the solar eastern limb and 6 near the western limb. Only 22 events are available for comparison, with 5 occurring near the eastern limb and 17 near the western limb.

Figure 4 shows the scatterplot for these 22 events. For 9 events, the NASA/CCMC flare scoreboard has a higher average flare occurrence probability, while our MLP model's probability is higher for the remaining 13 events. For about

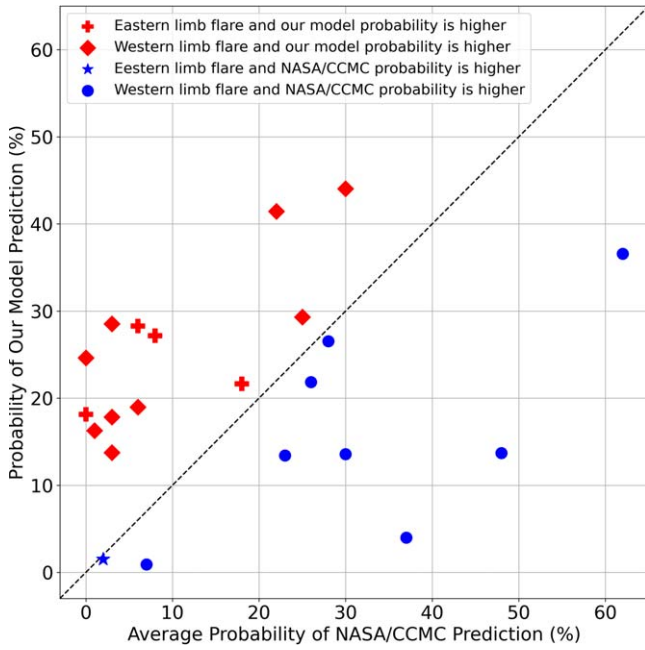


Figure 4. Scatterplot of 22 events for the available models in NASA/CCMC flare scoreboard. The x-axis represents the average flare occurrence probability of the NASA/CCMC flare scoreboard, and the y-axis represents the flare occurrence probability predicted by our MLP model.

53% of the western limb flares, our model predicts higher probabilities than those of the NASA/CCMC flare scoreboard. For the eastern limb flares, this rate is 80%. It is noted that flares occurring near the eastern limb are relatively more challenging to predict using existing flare models such as ones in NASA/CCMC flare scoreboard, as the ARs associated with them may not be visible on magnetogram and white light data. In contrast, solar EUV data can forecast flare occurrence near the eastern limb, even before they have significant information on possible flaring activity since they are optically thin. For flares occurring near the western limb of the Sun, the corresponding ARs may be better observed using white light and magnetogram data that are conventionally used for solar flare prediction. When calculating BSS values using the climatological model as a reference, we find that our MLP model and the NASA/CCMC flare scoreboard achieve BSS values of 0.30 and 0.22, respectively. This indicates that our model can provide relatively accurate probabilistic predictions of limb flares compared to the NASA/CCMC flare scoreboard for the 22 events.

5. Summary and Conclusion

In this study, we present empirical and MLP models using solar EUV data to forecast the occurrence of major solar limb flares within the next 24 hr. Our major results are summarized as follows. First, without magnetogram or white light data, we can forecast major solar limb flares using EUV data. There are strong positive linear correlations between the EUV intensities and the probabilities of flare occurrence. This can be attributed to the use of 94 and 131 Å whose peak intensities have strong correlations having large slopes with GOES X-ray fluxes. Second, when compared to the reference climatological model, which has an event ratio of 0.048, both empirical and MLP models show better probabilistic predictions. Third, the empirical model (AUC = 0.85) and the MLP model (AUC = 0.84) outperform

the random forecast (AUC = 0.50) for ROC-AUC. Additionally, both models show very similar performance from the perspective of TSS, HSS, BSS, and ROC-AUC. Finally, when comparing the flare probability predictions between the NASA/CCMC flare scoreboard and our models, we find that 30 events are not provided by the NASA/CCMC flare scoreboard, while our model can provide probability predictions for all 52 events. It is notable that for the flares occurring near the eastern limb of the Sun, our model tends to predict a higher probability.

Through the above results, we can conclude the following facts. First, for the first time, we provide the major solar limb flare prediction models using solar EUV data. They can offer the probability of solar flare occurrences near the limb regions, which are not provided by the traditional flare prediction models. Furthermore, since there was no study specifically targeting solar limb flare prediction before, we can anticipate that the results of this study can serve as a baseline model for future solar limb flare studies. Second, we compare our empirical and MLP models, and their performances do not show significant differences. In the case of solar limb flare occurrence probabilities, as observed in the results of the empirical model, there is already a strong linear correlation with the EUV intensities. Hence, MLP, which excels in solving nonlinear relationships, does not show a noticeably higher performance. The imbalance of the flare data, which is dominated by the nonflaring samples is a problem that we must solve to improve MLP. To solve this problem, we test the following methods: weighted cross-entropy loss, SOL function, and oversampling, but a trade-off occurs in which the forecast performance for the nonflaring data deteriorates. Finally, our model can provide predictions for events that the NASA/CCMC flare scoreboard models could not forecast. Existing models, due to distortions near the solar limb in magnetogram and white light data, provide predictions for a limited longitudinal range of ARs. In other words, predictions are often not provided for ARs close to the solar limb. However, our model can provide flare predictions for all cases since it can predict the probability of major solar limb flares occurrences based on the defined intensity from EUV data, regardless of the location of the ARs. Moreover, from the perspective of BSS for comparable events, our models show higher performance than the average probability of the models in the NASA/CCMC flare scoreboard, indicating that the use of EUV data on limb flare prediction may be more effective than traditional input data such as magnetogram and white light, especially for eastern limb events.

From an operational perspective, our models cannot only forecast solar limb flares but also construct a part of an ensemble model together with existing flare models. It would potentially provide better forecasting outcomes for flare occurrence near the solar limb. When future deep space missions like L4 (Cho et al. 2023) and L5 (Palomba & Luntama 2022) combine observational data, a comprehensive study on the front, limb, and backside flares of the Sun can be conducted. We can expect that the methods proposed in this study, along with solar EUV data, have the potential to guide diverse directions for that study.

Acknowledgments


We thank those who contributed to the SDO mission and successful AIA, EVE, and HMI operations. We give a grant to the community's devotion to developing the open-source

packages that were used in this work such as Python, Numpy, Matplotlib, Pandas, SunPy, and Pytorch. We also acknowledge the CCMC at Goddard Space Flight Center (GSFC) of NASA for the use of the Flare Scoreboard (<https://ccmc.gsfc.nasa.gov/scoreboards/flare/>), and the LMSAL for the use of flare data from HEK. This work was supported by the BK21 FOUR program through National Research Foundation of Korea (NRF) under Ministry of Education (MoE) (Kyung Hee University, Human Education Team for the Next Generation of Space Exploration), the Korea Astronomy and Space Science Institute under the R&D program (Project No. 2024-1-850-12) supervised by the Ministry of Science and ICT (MSIT), Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government MSIT (No. RS-2023-00234488, Development of solar synoptic magnetograms using deep learning, 15%), and Basic Science Research Program through the NRF funded by the MoE (NRF-2021R1I1A1A01049615, NRF-2021R1A6A3A01088835, and RS-2023-00248916).

Facility: SDO (AIA)

Software: Python 3 (Van Rossum & Drake 2009), Numpy (Harris et al. 2020), Matplotlib (Hunter 2007), Pandas (McKinney 2011), SunPy (Mumford et al. 2020), Pytorch (Paszke et al. 2019).

ORCID iDs

Jaewon Lee  <https://orcid.org/0000-0002-8287-956X>
 Yong-Jae Moon  <https://orcid.org/0000-0001-6216-6944>
 Hyun-Jin Jeong  <https://orcid.org/0000-0003-4616-947X>
 Kangwoo Yi  <https://orcid.org/0000-0003-4342-9483>
 Harim Lee  <https://orcid.org/0000-0002-9300-8073>

References

- Ahmed, O. W., Qahwaji, R., Colak, T., et al. 2013, *SoPh*, **283**, 157
 Allouche, O., Tsoar, A., & Kadmon, R. 2006, *JApEc*, **43**, 1223
 Battaglia, M., & Kontar, E. P. 2011, *A&A*, **533**, L2
 Berghmans, D., van der Linden, R. A. M., Vanlommel, P., et al. 2005, *AnGeo*, **23**, 3115
 Bloomfield, D. S., Higgins, P. A., McAteer, R. J., & Gallagher, P. T. 2012, *ApJL*, **747**, L41
 Boerner, P., Testa, P., Warren, H., Weber, M., & Schrijver, C. 2014, *SoPh*, **289**, 2377
 Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, *SoPh*, **289**, 3549
 Boerner, P., Edwards, C., Lemen, J., et al. 2012, *SoPh*, **275**, 41
 Boyer, R., Machado, M., Rust, D., & Sotirovski, P. 1985, *SoPh*, **98**, 255
 Brier, G. W. 1950, *MWRv*, **78**, 1
 Cho, K.-S., Hwang, J., Han, J.-Y., et al. 2023, *JKAS*, **56**, 263
 Colak, T., & Qahwaji, R. 2009, *SpWea*, **7**, S06001
 Cox, D. R. 1958, *J. R. Stat. Soc. Ser. B Stat. Method*, **20**, 215
 Crown, M. D. 2012, *SpWea*, **10**, S06006
 Cui, Y., Li, R., Zhang, L., He, Y., & Wang, H. 2006, *SoPh*, **237**, 45
 Deng, Z., Wang, F., Deng, H., et al. 2021, *ApJ*, **922**, 232
 Devos, A., Verbeeck, C., & Robbrecht, E. 2014, *JSWSC*, **4**, A29
 Dierckxsens, M., Tziotziou, K., Dalla, S., et al. 2015, *SoPh*, **290**, 841
 Falconer, D., Barchouty, A. F., Khazanov, I., & Moore, R. 2011, *SpWea*, **9**, S04003
 Fawcett, T. 2006, *PaRel*, **27**, 861
 Gallagher, P. T., Moon, Y.-J., & Wang, H. 2002, *SoPh*, **209**, 171
 Gardner, M. W., & Dorling, S. 1998, *AtmEn*, **32**, 2627
 Georgoulis, M. K., & Rust, D. M. 2007, *ApJ*, **661**, L109
 Guastavino, S., Marchetti, F., Benvenuto, F., Campi, C., & Piana, M. 2022a, *A&A*, **662**, A105
 Guastavino, S., Piana, M., & Benvenuto, F. 2022b, *ITNN*, **35**, 1993
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
 He, K., Zhang, X., Ren, S., & Sun, J. 2016, in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (Piscataway, NJ: IEEE), **770**
 Heidke, P. 1926, *GeAnA*, **8**, 301
 Hesse, M., Bellaire, P., & Robinson, R. 2001, in Proc. of the Space Weather Workshop: Looking Toward a European Space Weather Programme, ed. R. Gendrin et al. (Paris: ESA), 17
 Hong, S., Kim, J., Han, J., & Kim, Y. 2014, *AGUFM*, **2014**, SH21A
 Huang, X., Wang, H., Xu, L., et al. 2018, *ApJ*, **856**, 7
 Hudson, H. S. 1972, *SoPh*, **24**, 414
 Hunter, J. D. 2007, *CSE*, **9**, 90
 Hurlburt, N., Cheung, M., Schrijver, C., et al. 2012, *SoPh*, **275**, 67
 Hyvärinen, O. 2014, *WtFor*, **29**, 177
 Ioffe, S., & Szegedy, C. 2015, arXiv:1502.03167
 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
 Leka, K., Barnes, G., & Wagner, E. 2018b, *JSWSC*, **8**, A25
 Leka, K., Park, S.-H., Kusano, K., et al. 2019, *ApJ*, **881**, 101
 Lemen, J. R., Title, A. M., Akin, D. J., et al. 2012, *SoPh*, **275**, 17
 Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. 2021, *ITNN*, **33**, 6999
 Lim, D., Moon, Y.-J., Park, J., et al. 2019, *JKAS*, **52**, 133
 Liu, Y., Hoeksema, J., Scherrer, P., et al. 2012, *SoPh*, **279**, 295
 Marchetti, F., Guastavino, S., Piana, M., & Campi, C. 2022, *PatRe*, **132**, 108913
 McCloskey, A. E., Gallagher, P. T., & Bloomfield, D. S. 2018, *JSWSC*, **8**, A34
 McIntosh, P. S. 1990, *SoPh*, **125**, 251
 McKinney, W. 2011, Python for High Performance and Scientific Computing, **14**, 1
 Menzel, W. P., & Purdom, J. F. 1994, *BAMS*, **75**, 757
 Mumford, S., Freij, N., Christe, S., et al. 2020, *JOSS*, **5**, 1832
 Murray, S. A., Bingham, S., Sharpe, M., & Jackson, D. R. 2017, *SpWea*, **15**, 577
 Nair, V., & Hinton, G. E. 2010, in Proc. of the 27th Int. Conf. on Machine Learning (ICML-10), ed. J. Furnkranz & T. Joachims (Madison, WI: Omnipress), 807
 Nishizuka, N., Kubo, Y., Sugiura, K., Den, M., & Ishii, M. 2020, *ApJ*, **899**, 150
 Nishizuka, N., Sugiura, K., Kubo, Y., Den, M., & Ishii, M. 2018, *ApJ*, **858**, 113
 Palomba, M., & Luntama, J.-P. 2022, in COSPAR Scientific Assembly, **44**, 3544, <https://www.cosparathens2022.org/>
 Pangestu, A. D., Muhamad, J., Nurzaman, M. Z., et al. 2023, *JApA*, **44**, 25
 Papaioannou, A., Sandberg, I., Anastasiadis, A., et al. 2016, *JSWSC*, **6**, A42
 Park, E., Moon, Y.-J., Lee, J.-Y., et al. 2019, *ApJL*, **884**, L23
 Park, S.-H., Chae, J., & Wang, H. 2010, *ApJ*, **718**, 43
 Park, S.-H., Leka, K., Kusano, K., et al. 2020, *ApJ*, **890**, 124
 Paszke, A., Gross, S., Massa, F., et al. 2019, arXiv:1912.01703
 Pearson, K. 1895, *RSPS*, **58**, 240
 Pesnell, W. D., Thompson, B. J., & Chamberlin, P. 2012, The Solar Dynamics Observatory (SDO) (Berlin: Springer)
 Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, *SoPh*, **275**, 229
 Shibata, K., & Magara, T. 2011, *LRSP*, **8**, 6
 Steward, G., Lobzin, V., Cairns, I. H., Li, B., & Neudegg, D. 2017, *SpWea*, **15**, 1151
 Van der Sande, K., Flyer, N., Berger, T. E., & Gagnon, R. 2022, *FrASS*, **9**, 1031211
 Van Rossum, G., & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace)
 Venkatakrishnan, P., & Gary, G. A. 1989, *SoPh*, **120**, 235
 Wang, X., Chen, Y., Toth, G., et al. 2020, *ApJ*, **895**, 3
 Wieman, S., Didkovsky, L., Woods, T., Jones, A., & Moore, C. 2016, *SoPh*, **291**, 3567
 Wilkinson, L. K., Emslie, A. G., & Gary, G. A. 1989, *SoPh*, **119**, 77
 Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. 2018, in Proc. of the European Conf. on Computer Vision (ECCV), ed. V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Springer), **3**
 Woods, T. N., Eparvier, F., Hock, R., et al. 2012, *SoPh*, **275**, 115
 Yasyukevich, Y., Astafyeva, E., Padokhin, A., et al. 2018, *SpWea*, **16**, 1013
 Yohanandan, S., Song, A., Dyer, A. G., & Tao, D. 2018, in Proc. of the European Conf. on Computer Vision (ECCV), ed. V. Ferrari et al. (Cham: Springer), 235
 Zheng, Y., Li, X., & Wang, X. 2019, *ApJ*, **885**, 73