Beyond Traditional Flare Forecasting: A Data-driven Labeling Approach for High-fidelity Predictions

Jinsu Hong[0009-0002-4383-1376], Anli Ji[0000-0002-1551-2370], Chetraj Pandey[0000-0002-4699-4050], and Berkay Aydin[0000-0002-9799-9265]

Georgia State University, Atlanta GA 30303, USA
{jhong36, aji1, cpandey1, baydin2} @gsu.edu

Abstract. Solar flare prediction is a central problem in space weather forecasting. Existing solar flare prediction tools are mainly dependent on the GOES classification system, and models commonly use a proxy of maximum (peak) X-ray flux measurement over a particular prediction window to label instances. However, the background X-ray flux dramatically fluctuates over a solar cycle and often misleads both flare detection and flare prediction models during solar minimum, leading to an increase in false alarms. We aim to enhance the accuracy of flare prediction methods by introducing novel labeling regimes that integrate relative increases and cumulative measurements over prediction windows. Our results show that the data-driven labels can offer more precise prediction capabilities and complement the existing efforts.

Keywords: Solar Flares · Metadata · Space Weather Analytics

1 Introduction

Solar flares are one of the most pivotal space weather events that can have a significant influence on Earth and the near-Earth environment when accompanied by other eruptive phenomena such as coronal mass ejections. Some effects include but are not limited to power grid outages, disruption of navigation and positioning satellites, and increased radiation levels at high altitudes or space missions [5]. Thus, a robust prediction of solar flares and the accompanying events is essential to alert and prevent catastrophic impacts on Earth.

In flare forecasting literature, there are two main approaches: active region (AR) based models and full-disk models. The AR-based models are often formulated as image or time series classification [8] and each flare is associated with one active region. Full-disk solar flare prediction models [9] are usually framed as image classification problems and make use of all flares regardless of the AR associations. Solar flares, commonly detected using X-ray flux data from Geostationary Operational Environmental Satellites (GOES), are logarithmically categorized into five major classes (X, M, C, B, or A) by their peak X-ray flux measured in the 1-8Å passbands [4]. In traditional solar flare prediction tasks (be it active region-based or full-disk), time series data or solar full-disk images are typically labeled based on the maximum intensity of the flares. However, these labels have limited ability to reflect the solar cycle, despite some existing extensions and alternatives [11], [12].

The flare prediction models attempt to predict the occurrence of a pre-defined 'strong' flaring event within a prediction window (usually 12, 24, or 48h). The flare with maximum intensity (i.e., X-ray flux) in the prediction window is used for labeling the instances. However, there are three different limitations in the flare detection and labeling methods. First, quantifying the magnitude of X-ray flux for each AR is not feasible because X-ray flux measurements are global [4]. These global X-ray flux values can be misleading as they do not accurately represent the emitted radiation from individual ARs. Second, relying solely on the maximum intensity of a flare in the prediction window, however, disregards the background X-ray flux, fluence (integrated flux), or remaining flares' information. Lastly, the use of arbitrary thresholds for binary labels, such as >M1.0 or >C5.0, can further reduce the generalization capabilities of prediction models [10]. In this paper, we propose a more comprehensive approach that considers background conditions and cumulative indices to enhance solar flare prediction.

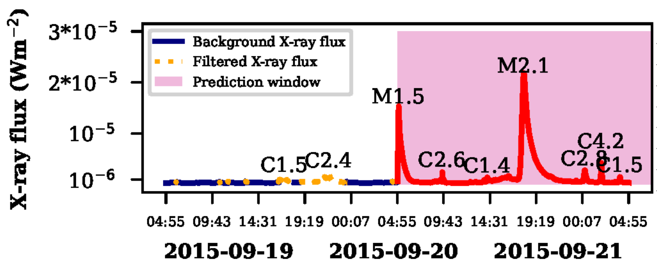
The rest of the paper is organized as follows: In Sec. 2, we describe the generation of new solar flare labels. In Sec. 3, we present a case study that reveals the feasibility of using these new labels in prediction tasks. Finally, in Sec. 4, we present our conclusion with future work.

2 Methodology

2.1 Relative Increase of Background X-ray Flux

To address the limitations of the existing label based on GOES classification, we propose a new methodology for generating labels based on the relative increase of background X-ray flux and complementary cumulative indices. We utilize 1-minute averaged GOES X-ray flux data to generate the labels for solar flares based on relative X-ray flux increase (referred to as *rxfi*). To ensure accuracy in determining *rxfi*, we first establish the background X-ray flux for a specific flare. We consider the 24-hour period prior to the start of the flare and exclude certain intervals from the background X-ray flux calculation. These exclusions include: (1) periods between the start and end times of other flares, (2) X-ray flux measurements that exceed the initial X-ray flux value at the start time of the flare of interest, and (3) measurements identified as low-quality by the instrument. After applying these filtering steps, we calculate the background X-ray flux by averaging the valid X-ray flux values within the specified period. Finally, the *rxfi* value is obtained by dividing the peak X-ray flux by the background X-ray flux.

We provide a practical example of how we calculate the *rxfi* for the M1.5 flare on 2015-09-20 in Fig. 1, which shows the X-ray flux from 2015-09-19T04:55 to 2015-09-21T04:55. Measurements from 2015-09-19T04:55 to 2015-09-20T04:55 are used for calculating the background X-ray flux for the M1.5 flare. The orange dotted line displays the filtered X-ray flux obtained using Cases 1 and 2. The blue line shows the cleaned/calibrated background X-ray flux. In this example, we obtain a background X-ray flux of $3.74 \times 10^{-7} Wm^{-2}$ for the M1.5 flare. Since the peak X-ray flux of the M1.5 flare is $1.5 \times 10^{-5} Wm^{-2}$, we find $rxfi = \frac{1.5 \times 10^{-5}}{3.74 \times 10^{-7}} = 40.11$.



GOES class	rxfi
M1.5	40.11
C2.6	6.29
C1.4	3.34
M2.1	40.26
C2.8	6.5
C4.2	9.67
C1.5	3.88

Fig. 1: An illustration of GOES 1-8 Å Solar X-ray flux observation.

2.2 Data-driven Labeling for Solar Flares

We utilize aggregated indices for our data-driven labels, which are as follows: (1) GC^{Max}: the GOES class of the maximum intensity flare in a given prediction window (24h in this work), (2) $rxfi^{Max}$: the flare with the maximum *rxfi* value in the prediction window, (3) $GC^{\sum}$: the weighted sum of the GOES subclass values in the prediction window, calculated as $\sum C_i + 10 \times \sum M_j + 100 \times \sum X_k$, and (4) $rxfi^{\sum}$: the sum of the *rxfi* values in the prediction window.

The new indices are obtained for both individual ARs and full-disk. In the AR-based approach, for a given time point (t_i), the procedure checks whether there exists a set of flares in the prediction window (in $[t_i, t_i + 24h]$). Based on the given set of flares, each time point (which is associated with an instance) is labeled as maximum or cumulative indices. Note that for time series-based flare classifiers we trained for our case study (presented in Sec. 3, we use an observation window of 12h (meaning. for a time point t_i, multivariate time series instances are obtained from $[t_i - 12h,\, t_i]$.

To better describe our new labels, we present an example of a prediction window in Figure 1, which shows seven flares from 2015-09-20T04:55 to 2015-09-21T04:55. We create three new labels in this example: In the case of $rxfi^{Max}$, the value of 40.26 (from M2.1 flare) will be the label, as it is the highest. For the $GC^{\sum}$, there are two M class flares and five C-class flares; therefore, the index for the $GC^{\sum}$ is 48.5 ((1.5 + 2.1)×10 + (2.6 + 1.4 + 2.8 + 4.2 + 1.5) = 48.5). Lastly, the $rxfi^{\sum}$ over this time period is calculated as 110.05 (40.11 + 6.29 + 3.34 + 40.26 + 6.5 + 9.67 + 3.88). These three different labels are used for a time series of the 12-hour observations or a solar full-disk image at a specific time point. The sliced time series dataset and full-disk labels can be found in the data repository [6].

3 Case Study: Flare Prediction with Data-Drive Labels

3.1 Data collection

Our baseline is active region-based and implemented by using the SWAN-SF [2], which is a multivariate time series dataset comprising 24 magnetic field parameters, covering from 2010 to 2018. For the purpose of solar flare prediction in our case study, we utilized six magnetic field parameters studied in [7]. We partitioned our data using the tri-monthly partitioning introduced in [10] and further split the SWAN-SF into four partitions where each one covers three months of data over the entire given dataset; i.e., Partition 1 covers instances from January to March, Partition 2 from April to June, Partition 3 from July to September, and Partition 4 from October to December. In this work, we used Partition 4 for testing and the rest of the three partitions for training.

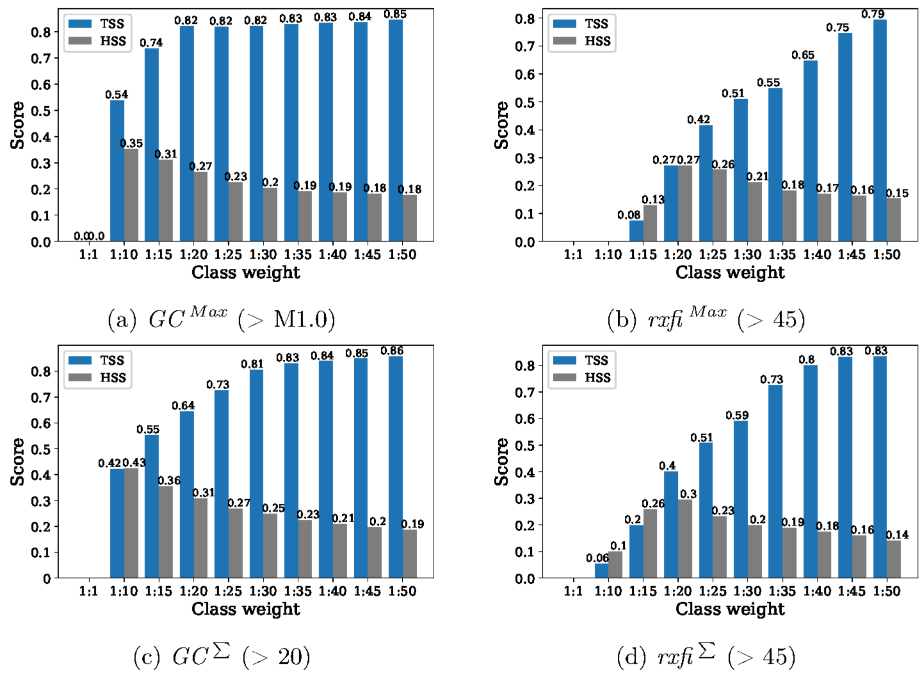
3.2 Classification Method and Evaluation

In our case of analyzing multivariate time series from active region patches, we utilized a time series forest (TSF) with column ensemble technique to work with multiple parameters given [3]. The outputs of classifiers from each separate parameter are then aggregated to form a final prediction using equal voting. The other notable hyperparameters are as follows: number of estimators is set to 50 and maximum tree depth is set to 3. More details on our implementation can be found in the project repository [1].

Utilizing data-driven labels: We utilized four different AR-based indices to label our instances: GC^{Max}, $rxfi^{Max}$, GC^{Σ}, and $rxfi^{\Sigma}$. For instances labeled with a GC^{Max} of M- or X-class flares, we designated them as flaring instances, while instances labeled with B- or C-class flares or flare quiet regions were designated as non-flaring instances. To label instances with the $rxfi^{Max}$, we assigned each instance a tentative numeric label and discretized them using predefined thresholds, where instances with a $rxfi$ greater than the threshold were considered flaring, and those below the threshold were considered non-flaring. We also set thresholds for the GC^{Σ} and $rxfi^{\Sigma}$. This resulted in the creation of 19 thresholds (from 10 to 100 with a step size of 5) for binary classification and four different prediction tasks.

Model Evaluation: To evaluate the models, we utilized a binary confusion matrix and used popular forecast skill scores in solar flare prediction: these are True Skill Statistics (TSS) and Heidke Skill Score (HSS) [8]. We set nine different class weights (i.e., 1:10, 1:15, ..., and 1:50) for each label type to counter the class imbalance issue and report our results.

As depicted in Figure 2, we present a comparison of models with four distinct labels: GC^{Max} with a threshold of M1.0, $rxfi^{Max}$ with a threshold of 45, GC^{Σ} with a threshold of 20, and $rxfi^{\Sigma}$ with a threshold of 45. Each of the three new labels has 19 models trained with 19 different thresholds ranging from 10 to 100 with a step size of 5. However, for brevity, we report only one result for each. Additional results from the models, along with saved candidate models, can be found in our project repository [1]. We also note that the TSF models are treated as black-box ensembles in the scope of this paper, and individual results from univariate time series classifiers are not independently reported.



(a) GC^{Max} ($>$ M1.0) (b) $rxfi^{Max}$ (> 45)

(c) GC^{Σ} (> 20) (d) $rxfi^{\Sigma}$ (> 45)

Fig. 2: The forecast skill scores for AR-based flare prediction models trained using instances labeled with data-driven labels and four distinct thresholds

Remarks: As the class weights increase, the trends of TSS and HSS scores vary. This is because the number of true positives and false positives increases. The models trained with new data-driven labels have similar results to the model trained with the GC^{Max}. Importantly, direct comparison between the labels may not be appropriate due to their inherent differences, such as varying imbalance ratios, background X-ray flux, and thresholds. However, it is anticipated that the *rxfi* would reflect the solar cycle, which influences the background X-ray flux. Models utilizing $rxfi^{Max}$ are intrinsically less prone to false alarms during solar minima. Therefore, *Max rxfi* can be served as an alternative label due to the high degree of flexibility in selecting thresholds and superior performance during solar minima. The evaluation metrics used in this study indicate that the proposed labels can enhance the performance of flare forecasting models beyond random guessing. Moreover, with the optimal threshold and class weights, the models trained with the new labels have the potential to outperform the existing labeling techniques.

4 Conclusion and Future Work

In this work, we have generated a new collection of flare labels (i.e., $rxfi^{Max}$, GC^{Σ}, $rxfi^{\Sigma}$) that can be used to complement the existing labeling techniques in space weather forecasting, specifically in solar flare prediction, for both active region-based and full-disk classification. We have presented a preliminary study evaluating the feasibility of using these new labels for active region-based prediction. Our results indicate that the proposed labels can serve as valuable additions to the existing labeling techniques, and their combination can improve the capability of solar flare prediction. We plan to extend this work to test different discretization thresholds and apply them to full-disk flare prediction models.
Acknowledgements: This work is supported in part under two grants from NSF (Award #2104004) and NASA (SWR2O2R Grant #80NSSC22K0272).

References

1. Source code, https://bitbucket.org/gsudmlab/data_driven_labels/src/main/
2. Angryk, R.A., Martens, P.C., Aydin, B., Kempton, D., Mahajan, S.S., Basodi, S., Ahmadzadeh, A., Cai, X., Boubrahimi, S.F., Hamdi, S.M., Schuh, M.A., Georgoulis, M.K.: Multivariate time series dataset for space weather data analytics. Scientific Data **7**(1) (Jul 2020)
3. Deng, H., Runger, G., Tuv, E., Vladimir, M.: A time series forest for classification and feature extraction. Information Sciences **239**, 142–153 (Aug 2013)
4. Fletcher, L., Dennis, B.R., Hudson, H.S., Krucker, S., Phillips, K., Veronig, A., Battaglia, M., Bone, L., Caspi, A., Chen, Q., Gallagher, P., Grigis, P.T., Ji, H., Liu, W., Milligan, R.O., Temmer, M.: An observational overview of solar flares. Space Science Reviews **159**(1-4), 19–106 (Aug 2011)
5. Fry, E.K.: The risks and impacts of space weather: Policy recommendations and initiatives. Space Policy **28**(3), 180–184 (Aug 2012)
6. Hong, J., Ji, A., Pandey, C., Aydin, B.: A data-driven Labels for solar flare predictions (2023). https://doi.org/10.7910/DVN/1U2Q3C
7. Ji, A., Arya, A., Kempton, D., Angryk, R., Georgoulis, M.K., Aydin, B.: A modular approach to building solar energetic particle event forecasting systems. In: 2021 IEEE Third International Conference on Cognitive Machine Intelligence (CogMI). pp. 106–115. IEEE (2021)
8. Ji, A., Aydin, B., Georgoulis, M.K., Angryk, R.: All-clear flare prediction using interval-based time series classifiers. In: 2020 IEEE International Conference on Big Data (Big Data). pp. 4218–4225. IEEE (2020)
9. Pandey, C., Angryk, R.A., Aydin, B.: Solar flare forecasting with deep neural networks using compressed full-disk hmi magnetograms. In: 2021 IEEE International Conference on Big Data (Big Data). pp. 1725–1730. IEEE (2021)
10. Pandey, C., Angryk, R.A., Aydin, B.: Deep neural networks based solar flare prediction using compressed full-disk line-of-sight magnetograms. In: Information Management and Big Data, pp. 380–396. Springer International Publishing (2022)
11. Ursi, A., Parmiggiani, N., Messerotti, M., Pellizzoni, A., Pittori, C., Longo, F., Verrecchia, F., Argan, A., Bulgarelli, A., Tavani, M., Tempesta, P., D'Amico, F.: The first agile solar flare catalog (2023)
12. Zhang, H., Li, Q., Yang, Y., Jing, J., Wang, J.T.L., Wang, H., Shang, Z.: Solar flare index prediction using SDO/HMI vector magnetic data products with statistical and machine-learning methods. The ApJ Supplement Series **263**(2), 28 (Dec 2022)