A Catalog of Solar Flare Events Observed by the SOHO/EIT

Sumanth A. Rotti1, Petrus C. H. Martens1, and Berkay Aydin2

1 Department of Physics and Astronomy, Georgia State University, Atlanta, GA 30303, USA; rotti@astro.gsu.edu
2 Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA

Received 2020 March 29; revised 2020 June 4; accepted 2020 June 5; published 2020 July 24

Abstract

We have compiled a catalog of solar flares (SFs) as observed by the Extreme ultraviolet Imaging Telescope (EIT) on board the Solar and Heliospheric Observatory (SOHO) spacecraft and the Geostationary Operational Environmental Satellites (GOES) spacecraft over a span from 1997 to 2010. During mid-1998, the cadence of EIT images was revised from two images per day to 12 minutes. However, the low temporal resolution causes significant data gaps in capturing much of the flaring phenomenon. Therefore, we monitor possible errors in flare detection by flare parameters such as temporal overlap, observational wavelength, and considering full field-of-view images. We consider the GOES flare catalog as the primary source. We describe the technique used to enhance the GOES detected flares using the extreme ultraviolet image captured by the EIT telescope. In order to detect brightenings, we subtract the images with a maximum cadence of 25 minutes. We have downloaded and analyzed the EIT data via the Virtual Solar Observatory. This flare data set from the SOHO/EIT period proves indispensable to the process of the SF predictions as the instrument has covered most of the Solar Cycle 23.

Unified Astronomy Thesaurus concepts: Solar flares (1496); Catalogs (205); Solar physics (1476)

1. Introduction

Solar flares (SFs) are sudden brightening features across the electromagnetic spectrum observed due to the release of stored magnetic energy in the solar atmosphere (Benz 2016). Such an instant explosion of energy is followed by a release of intense radiation into space. SFs are one of the most prominent activities that are often accompanied by coronal mass ejections (CMEs). However, not all flares have accompanying CMEs. Larger CMEs lead to solar energetic particle (SEP) emissions. These high-energy particles consist of electrons, protons, and higher ions accelerated to near-relativistic speeds. Often, SEP events are classified as gradual and impulsive where the latter focus on SFs while the former are relevant to those SEPs originating from CMEs (Kouloumvakos et al. 2015; Desai & Giacalone 2016; Kahler et al. 2017). Even though SEPs are associated with large flares, not every large flare has SEPs associated with it. In this regard, the X-ray peak contends as a strong indicator for a large SF that is capable of accelerating high-energy particles toward the Earth.

In space weather (SWx) research, the solar events are studied for (i) their influences on the nature of geomagnetic fields and (ii) how they can interact with the electromagnetic systems of the Earth (Pevtsov 2017). Some extreme events affect the radiation environment traveled by many spacecraft and astronauts outside the Earth’s magnetosphere, posing an extreme threat to humans on space missions (Dartnell 2011; Melott & Thomas 2011). High-energy rays from SFs can cause disturbances in the upper atmosphere of the Earth at ~80 km, which can cause radio blackouts. The degradation of high-frequency radio signals occurs due to ionization in the D region of a more dense and lower layers of the ionosphere, primarily impacting the 3–30 MHz band (Benson 1964; McNamara 1979; Frissell et al. 2019) This includes affecting certain types of global military communications and air traffic control (Knipp & Gannon 2019).

SFs are monitored continuously in white light, radio waves, soft and hard X-rays, and extreme ultraviolet (EUV) wavelengths by several instruments on the ground and in space (Benz 2016). The Geostationary Operational Environmental Satellites3 (GOES; Bornmann et al. 1996) have collected a database of SFs providing temporal and spatial data along with intensity classifications. Most of the solar missions, such as Yohkoh4 (Bentley 1994), RHESSI5 (Hurford 2003), Hinode6 (Kosugi et al. 2007), and the Solar Dynamics Observatory7 (SDO; Lemen et al. 2012) have their flare catalogs, whose events are curated using onboard instruments. These catalogs use GOES X-rays mostly to detect flares but imaging from other ones to localize the flare. The metadata obtained from various observational instruments has been processed and produced in the form of a catalog by many researchers for scientific study essential to solar physicists and SWx researchers (e.g., Yohkoh Flare Catalog: Sato et al. 2006; Hinode Flare Catalog: Watanabe et al. 2012; SDO/EUV Variability Experiment Catalog: Hock 2012). There are also efforts by Sadykov et al. (2017) in combining metadata from various instruments into one database useful for predictive analytics. The integration of data well supports researchers alongside the technology of Big Data.8 To analyze flares, the data from a specific instrument and observation information are essential for statistical as well as analytical studies. For example, an active region (AR) producing a GOES >M1.0 flare has a larger probability of producing flares in the next 24 hr (Park et al. 2018). The location of the parent SF is an essential parameter for the efficient prediction of a possible

1 https://www.ngdc.noaa.gov/stp/solar/solarflares.html
2 https://helioperiod.nasnas.nasa.gov/}
3 https://hesperia.gsfc.nasa.gov/legacy/HXT_catalog/index.html
4 https://hesperia.gsfc.nasa.gov/hessidata/hbase/hessi_flare_list.txt
5 http://xt.cfa.harvard.edu/flare_catalog/
6 https://hesperia.gsfc.nasa.gov/hessidata/dbase/hessi_flare_list.txt
7 https://hesperia.gsfc.nasa.gov/hessidata/hbase/hessi_flare_list.txt
8 Interactive Multi-Instrument Database of Solar Flares https://helioperiod.nas.nasa.gov/
SEP event (Gopalswamy et al. 2014). Since 1974, two X-ray sensors on each GOES satellites measure solar X-ray fluxes for the wavelength bands of 0.5–4 Å (short channel) and 1–8 Å (long channel; Garcia 1994). The flares from GOES are detected through hard X-ray detectors and provides the flare location on the Sun using the Ha images and the Solar X-ray Imager instrument uninterrupted from 2003 (Bornmann et al. 1996). GOES has traced flaring activity since the beginning of 1974 and has a catalog available with spatial and temporal specifications along with an associated AR number. However, some flare locations and/or AR numbers are missing from the GOES catalog or probably have errors (Milligan & Ireland 2018; Angryk et al. 2020a, 2020b).

Among several operational instruments, the Extreme ultra-violet Imaging Telescope (EIT) on board the Solar and Heliospheric Observatory (SOHO) observed the Sun in EUV wavelengths (Delaboudinière et al. 1995). Although EIT was not designed to observe flares, SOHO did capture a large number of flares during Solar Cycle 23. Nonetheless, no catalog of flares has been available based on the EIT data until now. Although our work is not exclusive to EIT, we are enhancing the GOES detected flares using EIT data. The enhancement is rather the important localization aspect as it increase the quality of the flare database. Such a catalog would be very useful to researchers to investigate for statistical and other SF studies further.

In this paper, we present a catalog of SFs, as seen by the SOHO/EIT instrument. This part of our work is introduced in Section 2. We choose the duration of the EIT’s primary operations, i.e., from 1997 to 2010, to get the flare locations. The main purpose behind this effort is to improve the quality of the flare locations in the existing database curated for machine learning (ML) and to make the flare metadata during the SOHO-era available for interested researchers. We have examined all the spatiotemporal attributes and/or characteristics of the EIT flares alongside the GOES SF catalog that is explained with the results in Section 3. We summarize our results in Section 4. Our catalog will be made publicly available on Harvard dataverse at [do:10.7910/DVN/C9H34R].

2. Data

In addition to the traditional statistical survey, for ML-based SWx forecasts, it is crucial to verify and establish the input data in the GOES catalog with other space-borne instruments (Angryk et al. 2020b). Flare prediction is one of the most prominent SWx forecasting applications for which flare localizations are of utmost importance. In this regard, Angryk et al. (2020b) developed a benchmark data set of ARs, together with flares consisting of a clear, verified, and easily accessible database put together from several sources using SDO data. In order to extend the quality of this benchmark data set, we choose to analyze the pre-SDO period utilizing the SOHO data.

2.1. EIT Imaging

Level Zero science data from the EIT instrument was downloaded by querying the Virtual Solar Observatory (VSO) using the Solarsoft10 IDL routines. The EIT images are already corrected for pointings in the header. The EIT captures the Sun in four different EUV wavelengths, which are 171, 195, 284, and 304 Å. The EIT instrument has a Nyquist limited resolution of 5º/2 (276 pixels; Delaboudinière et al. 1995). The initial cadence for observing was six hours but was improved to 12 minutes from 1998 July on (ESA 2010). Nonetheless, the available 195 Å synoptic images have a varying cadence from seven minutes to 13 minutes.

In Figure 1, the number of available EIT images per month from 1996 to 2010 is shown. At the specified cadence of 12 minutes, there should be approximately 3600 images captured by EIT every month. However, that is not the case. Noticeably, there is a significant irregularity in the available EIT images, which is a primary challenge for finding flare locations. There is an intrinsic difference between X-ray observations and those in the EUV images and they do not necessarily correspond one-to-one with each other. In the 195 Å band, Fe XII, Ca XVII, and Fe XXIV lines are dominated, which have their maximum fractional abundances at 1.4 × 10⁻⁶ K, 5 × 10⁻⁹ K, and (1–2) × 10⁻⁹ K, respectively (Feldman et al. 1999). The 195 Å wavelength shows significantly hot flaring regions on the Sun, giving a better choice to detect flares. However, 304 Å also shows the transition region of the solar atmosphere, effectively reducing the contrast between flares and the background Sun compared to the other three wavelengths, and therefore was not considered in this work.

2.2. Data Selection

To localize the flares and associate them with ARs, three data resources are needed: (1) the GOES flare catalog,11 (2) the National Oceanic and Atmospheric Administration (NOAA) AR summary, and (3) the full field-of-view EIT images. For our study, we used 195 Å images of the EIT instrument with a resolution of 1024 × 1024 pixels. The EIT web catalog interface and VSO hold the data products from SOHO/EIT, which were downloaded to our server. More details about the instrument and calibration can be obtained from the official website.12 We utilized all the 195 Å EIT images from the beginning of 1997 until the end of 2010 July, covering the entire Solar Cycle 23.

3. Methodology

Flare detection schemes are often built upon (1) time series data and/or (2) image data, for automatic detection (Fernandez Borda et al. 2002; Qu et al. 2003; Kano et al. 2008; Grigis et al. 2009; Aranda & Caballero 2010; Bonte et al. 2013; Caballero & Aranda 2014; Kraaijik 2015). For this work, we utilized Solarsoft’s Flare Locator13 algorithm (Freehand & Handy 1998) that works on the second principle—that is, image-based data. The flares are detected by measuring a threshold of the peak light intensity per macro pixel of an EIT 195 Å image closest to the GOES start time and then taking the relative difference with an image closest to the GOES peak time. Milligan & Ireland (2018) report that this process of image subtraction does not actually guarantee a flaring emission is detected but only that

11 https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-flares/x-rays
12 https://umbra.nascom.nasa.gov/eit/eit_guide/
13 https://hesperia.gsfc.nasa.gov/sss/gen/idl/ssw_util/ssw_flare_locator.pro
the timing and pointing given in GOES data set are consistent with the timing and location of the flare. The image subtraction using the Flare Locator algorithm also extracted information about dynamic flaring events in the EIT data that are not seen in X-ray images and, hence, cannot be related to a GOES flare. These are not false reports but are multiple events or, in some cases, post-flare loops occurring over a different active region or part of the Sun. Timing and position information from each event can be cross-checked with the metadata from other instruments to determine the observation of these events at the same time in the same location. The procedure to obtain the input images to the Flare Locator is discussed below.

3.1. Flare Detection and Verification

The steps followed in preparing the EIT Flare Catalog include:

1. All the flares reported in the GOES catalog are downloaded from the NOAA site during SOHO/EIT’s primary operation from 1997 to 2010.
2. For each flaring event of GOES class >C1.0, a full-disk EIT image at 195 Å is identified, such that the imaging time corresponds to the GOES start time of the flare within the range of ±5 minutes.
3. Every available image is then input as a primary image to the Flare Locator with a reference image taken after the primary within the range of ±25 minutes. If multiple images are available within the considered time range, then the closest image observed after the GOES flare peak time is selected as the reference image.

The verification process between the EIT flare coordinates and GOES flare coordinates is two-fold based on the following convention:

1. Exclusive flare positions in the GOES catalog are compared with the EIT flare coordinates such that the flares overlap in their temporal occurrence.
2. The implicit location derived from the centroid of NOAA AR is used if the explicit point coordinates are not available.

In Table 1, a sample data set of verified X-class flares from 1997 to 2000 is shown. The data set provides the index, start and end times of the event detection, location of the flaring region, the GOES X-ray classification, and the NOAA AR number. The following convention is used to determine whether a flare is flagged as verified or not verified; if the flare position (X, Y) in the EIT image is within 25″ compared to the GOES location, then it is considered verified. The selection of the distance is based on a similar convention implemented by the Hinode flare catalog to identify the overlapping flaring regions (Watanabe et al. 2012). If the EIT flare coordinates within 25″ do not overlap with the GOES flare catalog, then they are considered not verified. There are some events detected in the EIT images where the corresponding locations vary largely from GOES coordinates.
This is due to co-occurring events around the same time on different parts of the Sun. Such events are also flagged as not verified. The not verified class would include the flares occurring near-limb or behind the limb that would be detected by GOES X-rays as a result and if there is no way here to localize those events. If there are no existing coordinates in the GOES catalog to cross-check the EIT coordinates, then such flares are also flagged as not verified. The structure of the database schema—that is, the headers in Table 1—are as follows:

1. start_time_detection corresponds to the observation time of the initial EIT image used for detection. This is the primary image input to the flare detection software.
2. end_time_detection corresponds to the observation time of the preceding EIT image used for detection where the second image is input as a reference to obtain the difference image.
3. eit_fl_location is the value of flare locations in Helio- graphic Stonyhurst coordinates where north and west are positive while south and east negative. The number of decimal places here is limited to two values.
4. goes_class represents the classification of the flare magnitude according to GOES standards.
5. noaa_active_region is a number assigned to the flaring region by the NOAA group available in the GOES catalog. Here, N/A refers to not available AR numbers in the GOES catalog.

The data set consists of a “Remarks” column specifying whether the flare is verified or not. If not verified, then the co-occurring events or those with unavailable GOES location are specified correspondingly to maintain distinction. For such events, the GOES classification column in the EIT catalog retains the original values of that flare from the GOES catalog.

3.1.1. Statistics

Figure 2 shows a bar plot of the EIT flares from 1997 January to 2010 August. Overall, we have obtained 4602 flares
from the EIT data. In terms of numbers, 1881 flares are matching spatially and temporally with GOES flares, which is considered verified, while 1083 EIT flares match temporally. However, 1638 flares do not have a corresponding GOES location. Figure 3 shows a comparison between the number of flares in the GOES catalog for GOES class >C1.0 with and without coordinate values, respectively. The number of flares with no locations in the GOES catalog dominates for all the years except 2003. This gap greatly influences the verification process as a result, and the ratio of verified to non-verified EIT flares is sharply reduced. However, the overall ratio is similar between the verified EIT flares and the GOES flares with a location. It remains between ≈20% to ≈40% from 1997 to 2003, while it is >40% from 2004 to 2006 and 2010. The ratio falls below 20% for 2008 and 2009. Figure 4 compares the total number of GOES and EIT flares for our analysis. Figure 5 shows scatter plot of all the verified flares in accordance to the GOES classes. The two sub-types of non-verified flares are shown in Figure 6 but are not represented based on flare intensities as they cannot be confirmed. Also seen are some of the flares occurring near-limb.

Statistical analysis shows that 41% of the EIT flares are verified with the GOES catalog while 35.5% of the EIT flares do not have a corresponding GOES location. The remaining 23.5% of the EIT flares do not match spatially with GOES. Our statistics show satisfactory results despite many challenges, as stated earlier. We present a summary of our verified EIT flare findings in Table 2 along with the SDO data from Angryk et al. (2020b).

**4. Conclusions**

The SOHO/EIT flare catalog for 195 Å is developed for predictive and exploratory analytics using ML in flare prediction and to the scientific community for investigation of flares during the Solar Cycle 23. This catalog primarily serves to verify the locations of flares reported in the GOES catalog as a first step to generate a benchmark data set used for flare prediction. Here, verification implies the spatiotemporal matching of the flaring events between the GOES and EIT data. However, some events are not verified because (i) the GOES catalog has no report of exclusive coordinate locations for many flares to cross-check and (ii) multiple events are detected in the EIT images. Altogether, we identified 4602 events from the EIT images between 1997 January and 2010 July, out of which 1881 flares overlap with the GOES catalog in temporal and spatial coordinates. The SOHO/EIT Flare Catalog will be publicly available on Harvard dataverse for the solar physics community in [doi:10.7910/DVN/C9H34R]. We appreciate any suggestions concerning the catalog.

As far as the combining of EIT and GOES data is concerned, our foremost goal is to link the GOES flares with the correct magnetogram locations. So, the preflare magnetogram data can be used for testing flare prediction algorithms. Furthermore, it has become clear that there may be signs in preflare EUV emission predicting an upcoming flare. We can now verify those signatures from a large database of the EIT and Atmospheric Imaging Assembly (AIA) observed flares.

**Table 2**

The Number of Flares in the GOES Catalog Verified between 1997 and 2017 Using Secondary Sources

| Source          | Period             | C Class | M Class | X Class |
|-----------------|--------------------|---------|---------|---------|
| SOHO/EIT        | 1997 Jan 1 to 2010 Jun 30 | 1182    | 647     | 52      |
| SDO/AIA (Angryk et al. 2020b) | 2010 May 1 to 2018 Sep 1 | 6319    | 694     | 50      |
Figure 5. Distribution of all the verified EIT flares on heliographic coordinates.

Figure 6. Distribution of all the non-verified EIT flares on heliographic coordinates.
S.R. thanks the members of DM Lab, Dr. Georgoulis and Ms. Aparna Venkataramanasastry, for the many discussions during the project. This work was supported in part by funding from NASA Space Radiation Analysis Group (under GSU’s Sponsor Award No. TXS0156755). All scientific images used in this work are courtesy of the SOHO/EIT Consortium; SOHO is a joint ESA-NASA program. We are very grateful to the anonymous reviewer who provided constructive feedback that greatly improved the quality and scope of this paper.

ORCID iDs
Sumanth A. Rotti @ https://orcid.org/0000-0003-1080-3424
Berkay Aydin @ https://orcid.org/0000-0002-9799-9265

References
Angryk, R., Aydin, B., Kempton, D., et al. 2020a, Multivariate Time Series Dataset for Space Weather Data Analytics, Harvard Dataverse, doi:10.7910/DVN/EBCFKM
Angryk, R. A., Martens, P. C., Aydin, B., et al. 2020b, Multivariate Time Series Dataset for Space Weather Data Analytics (Atlanta, GA; Data Mining Lab), http://dmlab.cs.gsu.edu/wp-content/uploads/mvts4swa.pdf
Aranda, M. C., & Caballero, C. 2010, in Computational Intelligence for Knowledge-Based Systems Design, ed. E. Hüllermeier, R. Kruse, & F. Hoffmann (Berlin: Springer), 69
Benson, R. 1964, JRNBS, 68D, 1123
Bentley, R. D. 1994, LNP, 432, 311
Benz, A. O. 2016, LRSP, 14, 2
Bonte, K., Berghmans, D., De Groof, A., Steed, K., & Poedts, S. 2013, SoPh, 286, 185
Borrmann, P. L., Speich, D., Hirman, J., et al. 1996, Proc. SPIE, 2812, 309
Caballero, C., & Aranda, M. C. 2014, SoPh, 289, 1643
Dartnell, L. R. 2011, AsBio, 11, 551
Delaboudinière, J. P., Artzner, G. E., Brunaud, J., et al. 1995, SoPh, 162, 291
Desai, M., & Giacalone, J. 2016, LRSP, 13, 3
ESA 2010, A Change of Pace for EIT, the Ground-breaking Sun-watching Camera (Paris: ESA), https://sci.esa.int/s/8DR1npw
Feldman, U., Laming, J. M., Doschek, G. A., Warren, H. P., & Golub, L. 1999, ApJL, 511, L61
Fernandez Bordas, R. A., Mininni, P. D., Mard rini, C. H., et al. 2002, SoPh, 206, 347
Freeland, S. L., & Handy, B. N. 1998, SoPh, 182, 497
Frisell, N. A., Vega, J. S., Markowitz, E., et al. 2019, SpWea, 17, 118
Garcia, H. A. 1994, SoPh, 154, 275
Gopalswamy, N., Xie, H., Akiyama, S., Mäkelä, P. A., & Yashiro, S. 2014, EP&S, 66, 104
Grigis, P. C., Davey, A. R., Martens, P. C., et al. 2009, AGUFM, 2009, HS51B-1277
Hock, R. A. 2012, PhD thesis, Univ. Colorado at Boulder
Hurford, G. J. 2003, Proc. SPIE, 4833, 31
Kahler, S. W., White, S. M., & Ling, A. G. 2017, JSWSC, 7, A27
Kano, R., Sakao, T., Hara, H., et al. 2008, SoPh, 249, 263
Knipp, D. J., & Gannon, J. L. 2019, SpWea, 17, 794
Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, SoPh, 243, 3
Kouroumvakos, A., Nindos, A., Valthon, E., et al. 2015, A&A, 580, A80
Kraaijkamp, Verbeeck 2015, JSWSC, 5, A18
Lemen, J. R., Title, A. M., Akin, D. J., et al. 2012, SoPh, 275, 17
McNamara, L. F. 1979, RaSc, 14, 1165
Melott, A. L., & Thomas, B. C. 2011, AsBio, 11, 343
Milligan, R. O., & Ireland, J. 2018, SoPh, 293, 18
Park, S.-H., Guerra, J. A., Gallagher, P. T., Georgoulis, M. K., & Bloomfield, D. S. 2018, SoPh, 293, 114
Pevtsov, A. A. 2017, Ge&Ae, 57, 769
Qu, M., Shih, F. Y., Jing, J., & Wang, H. 2003, SoPh, 217, 157
Sadykov, V. M., Kosovichev, A. G., Ori a, V., & Nita, G. M. 2017, ApJS, 231, 6
Sato, J., Matsumoto, Y., Yoshimura, K., et al. 2006, SoPh, 236, 351
Watanabe, K., Masuda, S., & Segawa, T. 2012, SoPh, 279, 317