



# ASR: Archival Solar Flares Catalog

M. Berretti<sup>1,2</sup> , S. Mestici<sup>3</sup> , L. Giovannelli<sup>1</sup> , D. Del Moro<sup>1</sup> , M. Stangalini<sup>4</sup> , F. Giannattasio<sup>5</sup> , and F. Berrilli<sup>1</sup>

<sup>1</sup>University of Rome Tor Vergata, Department of Physics, Via della Ricerca Scientifica 3, 00133 Rome, Italy; [michele.berretti@unitn.it](mailto:michele.berretti@unitn.it), [francesco.berrilli@roma2.infn.it](mailto:francesco.berrilli@roma2.infn.it)

<sup>2</sup>Univeristà di Trento, Via Calepina 14, 38122 Trento, Italy

<sup>3</sup>University of Rome "La Sapienza," Department of Physics, P.le A. Moro 5, 00185 Rome, Italy

<sup>4</sup>ASI Italian Space Agency, Via del Politecnico snc, 00133 Rome, Italy

<sup>5</sup>Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Roma, Italy

Received 2025 March 14; revised 2025 March 20; accepted 2025 March 27; published 2025 April 22

## Abstract

Solar flares result from the rapid conversion of stored magnetic energy within the Sun's corona. These energy releases are associated with coronal magnetic loops, which are rooted in dense photospheric plasma and are passively transported by surface advection. Their emissions cover a wide range of wavelengths, with soft X-rays being the primary diagnostic for the past 50 years. Despite the efforts of multiple authors, we are still far from a complete theory capable of explaining the observed statistical and individual properties of flares. Here, we exploit the availability of stable and long-term soft X-ray measurements from NASA's Geostationary Operational Environmental Satellites mission to build a new solar flare catalog, with a novel approach to linking sympathetic events. Furthermore, for the most energetic events since 2010, we have also provided a method to identify the origin of the observed flare and eventual link to the photospheric active region by exploiting the array of instruments on board NASA's Solar Dynamic Observatory. Our catalog provides a robust resource for studying space weather events and training machine learning models to develop a reliable early warning system for the onset of eruptive events in the solar atmosphere.

*Unified Astronomy Thesaurus concepts:* [Catalogs \(205\)](#); [Astronomical methods \(1043\)](#); [Solar physics \(1476\)](#); [Solar flares \(1496\)](#); [Solar x-ray flares \(1816\)](#)

## 1. Introduction

It is well known that space weather events such as solar flares, coronal mass ejections (CMEs), and solar energetic particle emissions, pose a significant threat to human infrastructure both on Earth and in space (e.g., F. Berrilli et al. 2014; L. Di Fino et al. 2014; M. H. MacAlester & W. Murtagh 2014; E. J. Oughton et al. 2017). Hence, understanding and predicting the hazards linked to the Sun's magnetic activity have grown in importance over the years. In the last decade, a plethora of authors have trained different machine learning models to forecast the likelihood that the Sun will produce a flare in a precise time frame, usually ranging from a few hours to 1–2 days (C. J. Schrijver 2007; R. Qahwaji & T. Colak 2007; D. Yu et al. 2010; J. Li 2011; M. G. Bobra & S. Couvidat 2015; G. Barnes et al. 2016; E. Jonas et al. 2016; X. Huang et al. 2018; E. Park et al. 2018; K. D. Leka et al. 2019; D. Cicogna et al. 2021; Z. Deng et al. 2021; N. Nishizuka et al. 2021; C. Pandey et al. 2023, to name a few). For an extensive analysis of the challenges of machine learning applied to space weather and the prediction of the occurrence of solar flares, we refer the reader to the review by E. Camporeale (2019).

A complete data set is crucial for enhancing the performance of machine learning models while also mitigating biases and optimizing computational efficiency, as discussed in N. Mehrabi et al. (2021). Traditionally, many works on flare forecasting are based on the flare catalog provided by the NOAA,<sup>6</sup> a live service

that has provided flare events since 2002 January with the launch of the GOES-08 satellite. However, while the NOAA catalog remains the primary reference for flaring events, standard flare identification methods can miss some events and fail to provide further insights into the more complex processes linking flares to the global-scale dynamics of the solar magnetic field. Alternative catalogs have been proposed, often relying on different detection algorithms or instruments.

In M. J. Aschwanden & S. L. Freeland (2012) the authors conducted a comprehensive analysis of solar flare activity spanning 37 yr (1975–2011). Using an automated flare detection algorithm applied to soft-X-ray data from Geostationary Operational Environmental Satellites (GOES), they identified and cataloged over 300,000 solar flare events. From now on, we will refer to this catalog as ASC.

In K. van der Sande et al. (2022), the authors present a solar flare catalog based on EUV images from the Atmospheric Imaging Assembly (AIA, J. R. Lemen et al. 2012) on board the Solar Dynamics Observatory (SDO, W. D. Pesnell et al. 2012). This approach increases the number of recorded events and improves the characterization of the total energy released. However, its applicability is limited by the relatively short operational window of the SDO mission compared to GOES, restricting statistical studies across multiple solar cycles.

In N. Plutino et al. (2023), which is referred to as the PLU catalog from here on, the authors propose an alternative flare identification algorithm to detect events in the GOES X-ray signal. Instead of relying on the ratio between the background and the peak fluxes, their approach defines a true event based on an instrumental threshold, leading to a higher number of detected flares.

In this work, we present a new solar flare catalog based on soft-X-ray measurements from GOES and named the Archival

<sup>6</sup> <ftp://ftp.swpc.noaa.gov/pub/warehouse/>

Solar flares (ASR) catalog. We refine many of the events already recorded by both NOAA and PLU and employ a novel approach to identify and connect both sympathetic flares and homologous flares. This analysis allows us to investigate statistical relationships between distant flaring regions and explore the global properties of these energetic solar events. Thanks to the availability of stable and long-term X-ray measurements, we are able to cover two full solar cycles, and, for the most recent one, we also provide the origin point of the most intense events. Our catalog serves as a valuable resource for both space weather studies and machine learning applications. Indeed, thanks to a versatile and complete data set, we can support the development of more reliable flare prediction models and early-warning systems for potentially disruptive solar activity.

## 2. Data Set

The GOES, first launched in 1974 with GOES-1 and currently active with GOES-16/17/18, is a NASA/NOAA mission focused on monitoring both Earth and solar activity. Regarding the Sun observation side of the mission, the satellites are equipped with two X-ray sensors (XRS), one measuring radiation between 0.5 and 4 Å (short channel) and the other between 1 and 8 Å (long channel). See F. A. Hanser & F. B. Sellers (1996) for an introduction to GOES XRS. In this work, we considered the 1 minute averaged science-quality X-ray signal of the long channel of the primary satellite from 2002 January 1 to 2024 December 31. The science-quality data, provided directly by the NOAA science team, is thoroughly processed and includes several corrections with respect to the operational data, including those for Space Weather Prediction Center (SWPC) scaling factors, XRS-A bandpass, time stamps, temperature effects, better quality flags, and other issues. Specific information on the available data for each satellite and which to consider for a specific time range can be found in the extensive guide provided by the NOAA in the dedicated XRS web portal.<sup>7</sup> In general, when two or more satellites are actively collecting data at the same time, one of them is indicated as the primary source, while the other is indicated as the secondary one. Finally, the downloader function of the ASR catalog code provides yearly data sets containing the 1 minute averaged time series of the soft-X-ray flux along with the quality flag, the data source, and the time stamp in plain text format. With this approach, our method differs from the previous ASC and PLU catalogs, which use the GOES operational data, characterized by a temporal resolution ranging between 1 and 4 s later averaged to 12 s and with an additional smoothing window applied.

For the most energetic events, we are able to determine the origin point of the solar eruption thanks to the availability of stable, seeing-free, and continuous images of the solar corona captured by the AIA on board SDO. To determine the location of solar flares identified through our GOES X-ray analysis, we chose to utilize images from the Fe IX 17.1 nm spectral line, corresponding to images of the lower quiet corona to provide a reliable picture of the coronal magnetic fields, with little interference from other transient events such as shocks (K. van der Sande et al. 2022). Here, we estimated the position of 138 X-class flares, 1735 M-class flares, and 1958 C-class flares, for a total of 3831 events.

## 3. Methods

The ASR catalog code consists of a significant improvement over the previous PLU flare catalog, with several of the same authors contributing to both works. Although many of the novel features of the former catalog have been inherited, such as accounting for sympathetic/homologous events and computing each flare’s flux integral, the ASR catalog further offers many improvements. In this section, we will thoroughly review the various steps involved in the identification, definition, and characterization of flares in our catalog.

### 3.1. Flare Identification

The first step in building the flare catalog consists of identifying maxima and minima in the yearly GOES time series. To this end, we exploited the “argrelextrema” function of the SciPy library (P. Virtanen et al. 2020). For each sampled point, the function checks its nearest neighbors; if both values are larger or smaller than the considered point, then it is classified as a relative minimum or maximum, respectively. The index of the detected extreme is then recorded in a “pandas dataframe” data structure and saved as a separate comma-separated values file. A drawback of this approach is the loss of flat extremes, that is extremes characterized by a series of the same value. However, such situations are not commonly found in the soft-X-ray signal and are mainly associated with instrumental saturation. Hence, saturated events are handled by replacing the flat region in the measured GOES X-ray signal with a second-degree polynomial fitted by overestimating the peak by 2.5% in the center of the saturated region and adjusting the edges. It is worth noting that, in the 22 yr covered by the whole catalog, instrumental saturation occurred only once, during the famous storm of 2003 October.

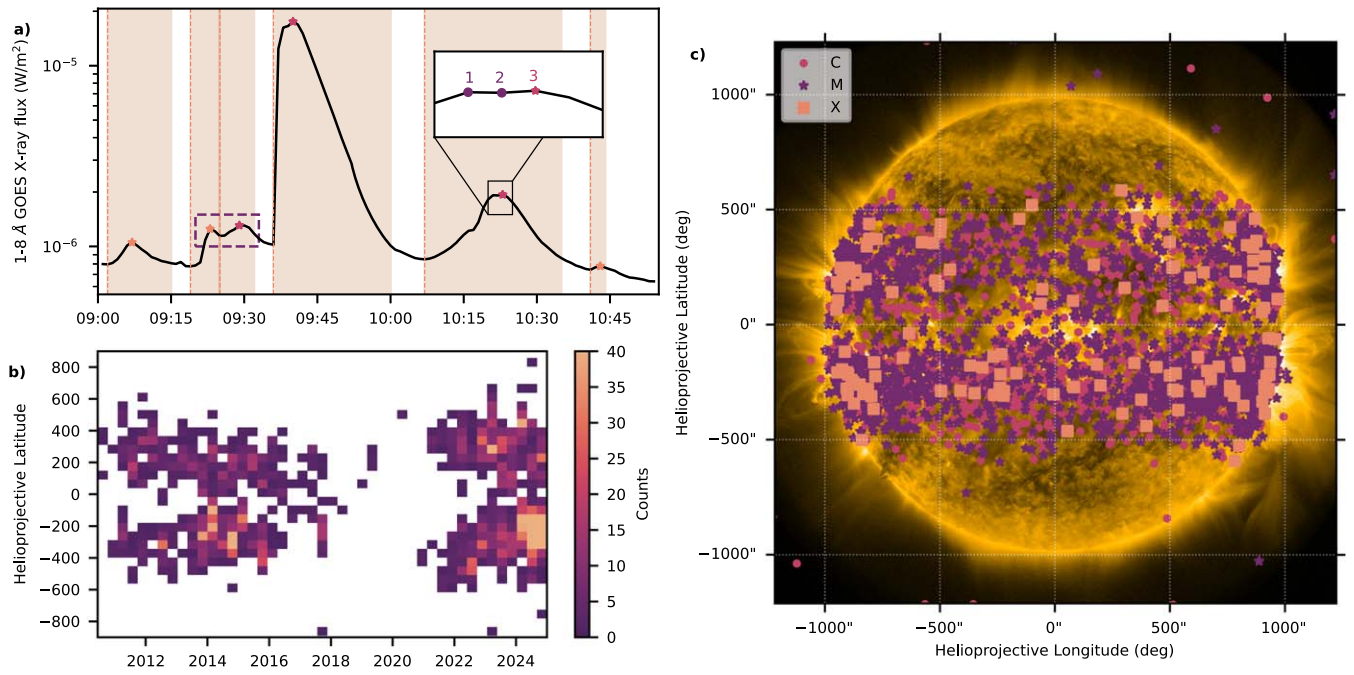
Due to the inherently noisy nature of the GOES signal, even in the softer long-channel and 1 minute averaged science-quality data, we need to account for the presence of consecutive extrema. These neighboring extrema are rarely associated with true flare events, which are generally characterized by a sharp increase in flux followed by a slower energy decay. Thus, considering these extrema as events would lead to the presence of false flares, compromising the statistic and goodness of the catalog. In our code, this is taken into account by checking the indexes of the extrema: if three consecutive ones are detected, the first two are dropped. We point out that this might lead to an offset in either the flare start, end, or peak time of no more than 2 minutes.

### 3.2. Flare Definition

Once the relative extrema of the entire GOES time series are identified, we check how many of these events are true flares and not noise-driven fluctuations. Over the years, various definitions have been proposed to reliably identify solar flares in the GOES signal. The NOAA catalog defines a flare using three criteria: (1) the event starts where four consecutive points increase monotonically and exceed  $10^{-7} \text{ W m}^{-2}$ ; (2) the flux ratio between the first (start) and fourth (peak) points is greater than 1.4; and (3) the end time of the flare is defined as the point in time successive to the flare peak where the flux is half that observed at the peak.

In our catalog, we adopt the same flare definition as in PLU. Each minimum represents a potential flare start, while the following maximum represents the possible flare peak. A

<sup>7</sup> <https://www.ngdc.noaa.gov/stp/satellite/goes-r.html>



**Figure 1.** (a) Application of the flare detection algorithm to GOES X-ray data from 2015 March 9. The violet dashed rectangle highlights two homologous events correctly detected by the algorithm, while the magnified inset shows the difficulty of flare detection due to noise-induced fluctuations. Orange stars mark the flare peaks detected by our algorithm, while pink ones highlight those identified in both the NOAA catalog and our own. Orange-shaded areas denote the duration of the corresponding flare. (b) Reconstructed butterfly diagram using solar flares, showing the latitudinal spread of flare occurrences over different solar cycles. (c) Spatial distribution of flares on the solar disk, with flare types represented by color. Stray values are due to oversaturated AIA images.

provisional flare event is classified as a true flare if the flux at the peak exceeds the background flux, i.e., the averaged flux 2 minutes prior to the minimum, by a threshold equal to  $2 \times 10^{-8} \text{ W m}^{-2}$ . Properly estimating the flux background can greatly improve the detection of weaker events and the estimation of their energy. The threshold value, related to the intrinsic sensibility of the XRS instrument, defines the noise floor for typical GOES flux measurements, established during quiescent periods of solar activity (e.g., M. J. Aschwanden & S. L. Freeland 2012). Accurately estimating the background flux is crucial. Indeed, this flux can be considered as the superposition of the contributions coming from the active region where the flare originates, as well as the integrated flux coming from the entire solar disk (including beyond the visible limb).

We define the end time of an event either as the time successive to the peak where the flux value returns to the background value plus the threshold value, or the point when a new true event starts. The choice between the two falls to whichever occurs first. In the former case, a unique label is assigned to the flare. In the latter case, that is, in case of consecutive true flare events, a shared label is assigned to the group of events, allowing direct identification of homologous or sympathetic events.

### 3.3. Flare Characterization

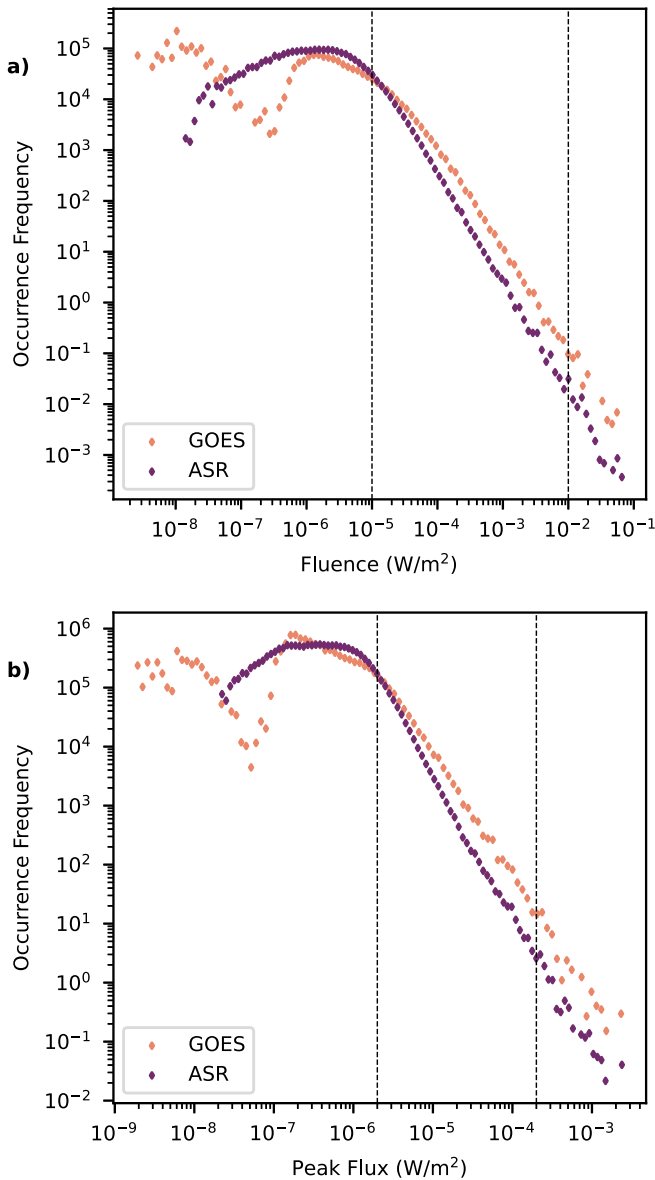
In addition to the start, peak, and end times and their associated flux values, our catalog offers other properties related to the flare that can be used to characterize each event. The integrated flux of each flare is computed as the integral, using the trapezoidal rule, of the flux values between the start and end times. This information is critical to understand the

**Table 1**  
Summary of the Flare Statistics for the NOAA, Plutino, and ASR Catalogs  
Divided by Class up to the Year 2022

Class	GOES	PLU	ASR	ASR <sub>REL</sub>
A	2690	2800	5077	147,696
B	15,980	88,005	118,384	93,963
C4-	8543	68,980	125,051	17,248
C5+	1564	4807	10,352	2698
M	1581	2803	5189	2490
X	159	187	274	232

energy released during the entire flare event. We also provide a corrected integrated flux value that corresponds to the difference between the integrated and background fluxes. This quantity takes into account the background solar activity and offers insight into the “rescaled” energy released.

To further compare our catalog with the one released by the NOAA, we also provide the classification of flares according to the NOAA standard procedure. Indeed, flares are traditionally divided into five classes depending on the intensity of the peak emission in the soft-X-ray signal, starting from the weakest A-class flares, with peak flux emissions between  $10^{-8}$  and  $10^{-7} \text{ W m}^{-2}$ , up to X-class flares with peak emissions over  $10^{-4} \text{ W m}^{-2}$ . Similarly to the flux integral, we provide two different class definitions for any flare event. The traditional class is determined similarly to the NOAA catalog by considering the peak flux of the flare event. We then introduce a relative flare class that is estimated by rescaling the event peak flux to the background flux (referred to as delta flux from here on), so that it is not biased by the solar activity cycle. A detailed description of the estimated background is provided in



**Figure 2.** Probability density functions of the estimated flux integral, also known as fluence, (top) and peak flux (bottom) for the GOES catalog and the ASR catalog, in orange and violet, respectively. The dashed black lines highlight the range used for the linear fit to estimate the slope of the distributions. In the top panel, this range spans from  $3 \times 10^{-5}$  to  $10^{-2} \text{ W m}^{-2}$ , and the fitted slopes are  $-2.1 \pm 0.1$  and  $-1.9 \pm 0.01$ , respectively, for the ASR and GOES ones. In the bottom panel, it extends from  $3 \times 10^{-6}$  to  $2 \times 10^{-4} \text{ W m}^{-2}$ , and the estimated slopes are  $-2.3 \pm 0.01$  for the ASR catalog and  $-2.1 \pm 0.01$  for the GOES one.

Appendix A. For the sake of clarity, from now on we will refer to the catalog which uses the traditional class definition as the ASR catalog and to the one using the relative class definition as the ASR<sub>REL</sub> catalog.

Finally, to align with the NOAA procedure and enable further refinement of observed flaring events, we also include the ratio between peak and background flux. Adjusting the filters in our catalog makes it easy to reconstruct the NOAA one.

### 3.4. Flare Location

For the most energetic events, that is, those with a delta flux greater than  $5 \times 10^{-6} \text{ W m}^{-2}$ , we estimate the point of origin of the flare. To determine the position of a given event, we

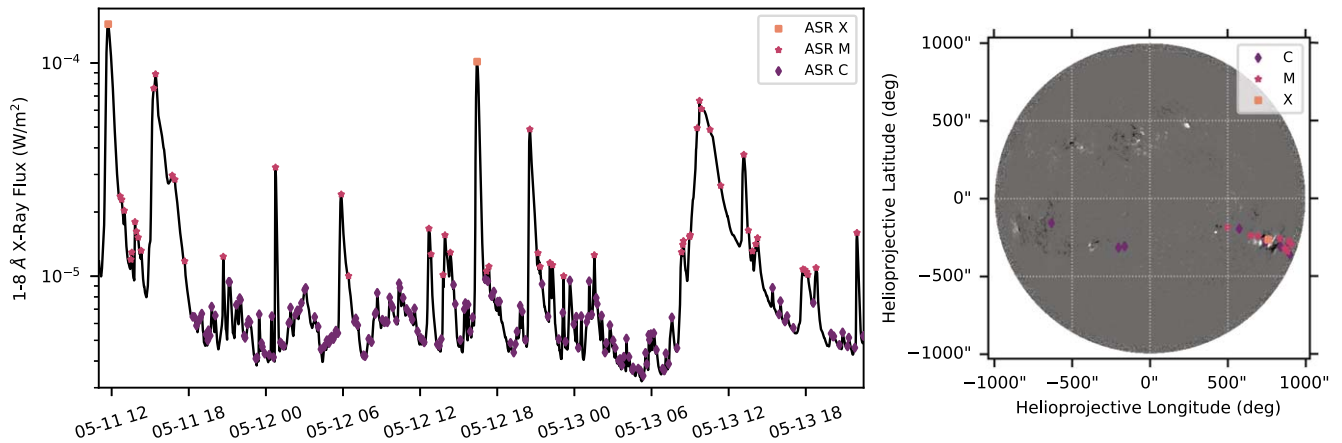
consider the barycenter of the brightest cluster of pixels in the difference AIA Fe IX 17.1 nm channel image between the peak and start of the flare. Moreover, as we consider only the largest and brightest cluster of pixels, we can safely assume that other energetic events in the solar atmosphere, such as weaker flares and shocks, can be neglected and do not risk affecting the flare localization process. The flare location is then provided in pixel, helioprojective, and heliographic coordinates to allow for a straightforward connection to the corresponding active region.

## 4. Results and Discussion

The total number of flares detected by our algorithm as a function of the flare class is shown in Table 1. In particular, the second, third, and fourth columns refer to the number of flares identified by the NOAA, PLU, and ASR catalogs, respectively. Instead, the fifth column reports the number of flares per class after accounting for the background flux, based on a “relative” flux value obtained by subtracting the background flux from the peak one. We wish to stress that the background we intended here is related to the magnetic activity of the Sun and is not related to the background given by other Galactic sources, which is already removed in the science-quality data provided by NOAA. The number of flares in our catalog is consistently higher with respect to the NOAA one in each class. These differences, especially in the lower-energy classes, stem from the radically different detection method used in our work, which is based on an instrumental threshold and does not require a specific ratio between the flux measured at the start and peak points. The increase in flares in the M and X classes can also be attributed to our definition of a flare. In fact, secondary peaks in the tail of M- or X-class events may also be classified with a similar class. Table 1 also shows that the number of flares identified in our catalog is also greater than that presented in PLU. Although the identification and definition processes are the same, the main difference lies in the data set used. In fact, the former uses the operational data provided by GOES, which requires further processing. We also observe that the relative class attribution, which takes into account the background solar activity, drastically changes the distribution of flares, especially at lower energies. In fact, most of the standard flares populating the B and C classes are shifted to a lower-energy class in the adjusted flux-integral scenario. The X and M classes are instead less affected by the flux rescaling, meaning that most of the large events, which are generally observed during high solar activity, largely overshoot the background-level flux.

It is worth pointing out that the number of flare events for a catalog is not a significant proxy of its quality. Indeed, this number is mainly related to the flare-event definition, which can be quite arbitrary. Here, we aimed to provide the most general and versatile definition of a flare, leaving the users the possibility and freedom to decide which events to consider and which not. Furthermore, the grouping of homologous and/or sympathetic events under the same ID offers the opportunity to consider these events as a single longer event or multiple separated ones. We wish to stress yet again that the novelty and strength of this work lies in its versatility, making it easily adaptable to many different tasks.

In Figure 1(a) we show the detection algorithm process for a particularly challenging case study. We selected the GOES X-ray signal from 2015 March 9, during the peak of solar



**Figure 3.** In the left panel, we show the measured GOES X-ray signal from 2024 May 11 to 2024 May 14. The detected flares are marked in violet, pink, and orange diamonds for C-, M-, and X-class events, respectively. The right panel shows the positions of the most energetic flares in the solar disk over a magnetogram captured by the Helioseismic and Magnetic Imager on board SDO on 2024 May 14. As expected, the majority of flares originated in a well-defined active region, while weaker flares are scattered across the solar disk.

activity of the 24th cycle, between 09:00 and 11:00 UTC, due to its complex features, which our algorithm successfully handled. The continuous black line is the 1 minute average of the soft-X-ray flux in the long channel, and the star points represent the flare peaks. In particular, the pink stars highlight events that are identified both in the NOAA and ASR catalogs, whereas the orange ones are unique to our catalog. The shaded regions mark the time of the identified flare events, and the dashed vertical line is the starting point. In the violet dashed box, we show that the algorithm successfully links two homologous events, as the second one starts before the first can return to preflare flux levels. The enlarged inset provides an example of three consecutive extrema, which were identified as provisional maxima and minima. As explained previously, depending on the order of the sequence, this case may lead to a shift up to 2 minutes of the flare peak. In this case, the flare peak is attributed to the second maximum of the sequence.

Figures 1(b) and (c) present two statistical properties of our catalog. In panel (b), we show the distribution of solar flares across the Sun’s surface throughout the period covered by our catalog. Similarly to sunspot behavior, flares initially appear at midlatitudes and migrate toward the equator as the solar cycle progresses. This pattern resembles the well-known butterfly diagram for sunspots (e.g., S. Cloutier et al. 2024), but our analysis includes the positions of thousands of flares. The occurrence rate and the spatial distribution of flare events on the disk are correlated to solar activity, with higher latitudes seeing more flares when solar activity is higher. The solar activity minimum of the 24th solar cycle between 2018 and 2020 is also evident, as almost no intense flare events are observed.

We display the flare spatial distribution across the entire solar disk in panel (c). Flares that belong to the M, C, and X standard classes are indicated by orange squares, pink crosses, and violet dots, respectively. Few stray values, which differ from the majority of the data set constrained to midlatitudes, are the result of the oversaturated AIA images.

The probability density distribution of the flux integral and the peak flux for the ASR and NOAA catalogs are shown in Figure 2. The shapes of the distributions associated with the two data sets clearly differ at low energies. The flux-integral distribution of the NOAA catalog has a first peak around

$10^{-8} \text{ W m}^{-2}$ , followed by a minimum at  $5 \times 10^{-7} \text{ W m}^{-2}$  and another maximum at  $\sim 10^{-6} \text{ W m}^{-2}$ . As XRS detection is effective for fluxes greater than  $10^{-8} \text{ W m}^{-2}$ , it could be argued that the first peak is driven purely by noise. Instead, the ASR probability distribution increases exponentially from lower energies to a peak around  $\sim 10^{-6} \text{ W m}^{-2}$ . This more “consistent” behavior can be attributed to the effectiveness of the algorithm in identifying true flare events while discarding noise-driven fluctuations. At higher energies, both distributions are characterized by a similar exponential decay, covering more than two decades in energy. The dashed lines in Figure 2 show the range in which we performed a linear fit of the distributions in the log–log space. The slopes associated with the two distributions are  $-2.1 \pm 0.1$  and  $-1.9 \pm 0.1$  for the ASR and NOAA catalogs, respectively. Our results are consistent with not only NOAA, but also with previous observations (ASC, PLU, A. Veronig et al. 2002a, 2002b; S. Yashiro et al. 2006) and the scaling index of the fractal diffusive self-organized criticality model described in ASC and M. J. Aschwanden (2011).

Most of the points discussed for the flux-integral distribution can also be applied to the peak flux distribution shown in Figure 2(b). Once again, the dashed lines are the linear fit of the distributions. The associated slopes are  $-2.3 \pm 0.1$  and  $-2.1 \pm 0.1$ , which are consistent.

#### 4.1. 2024 May Case Study

In this section, we present an in-depth focus on the most intense space weather event of the last two decades that occurred on 2024 May 11–14. The solar active regions (ARs) classified as AR 13664 and AR 13668 by the NOAA, which eventually merged into a single one, produced a series of geoeffective space weather events comprising flares and fast CMEs. In the leftmost panel of Figure 3, we show the GOES X-ray flux from 2024 May 11 to 2024 May 14, marking the peaks of the detected flare in violet, pink, and orange, respectively, for C, M, and X classes. According to the absolute class, i.e., the standard NOAA definition for flare class, our algorithm detected a total of 171 C-class flares, 53 M-class flares, and 2 X-class flares. In the right panel of Figure 3 we show the position of the detected events on the solar disk. As we can see, the majority of flares originated from

AR 13664 and AR 13668 with a handful of less intense events coming from the other side of the solar disk.

## 5. Conclusion

In this work, we presented the ASR catalog, a novel flare catalog based on GOES soft-X-ray measurements that covers more than 20 yr of observations. Based on an already established detection algorithm, we implemented significant changes to the identification algorithm and offered additional properties related to the flare, such as the point of origin of the most energetic events. By exploiting the user friendliness of Python, we offer a robust and accessible alternative to already existing catalogs. The ASR catalog offers reliable support for training machine learning models for space weather forecasting and solar physics studies. Finally, it is a live service that provides a daily bulletin of flares in the past 24 hr.

## Acknowledgments

We acknowledge the use of data from GOES satellites, concerning solar signal in the soft-X-ray range (<https://www.ngdc.noaa.gov/stp/satellite/goes/dataaccess.html>) and the use of NOAA SWPC flare lists (<https://www.ngdc.noaa.gov/stp/solar/solarflares.html>). M.B. acknowledges that this article was produced while attending the PhD program in Space Science and Technology at the University of Trento, Cycle XXXIX, with the support of a scholarship financed by the Ministerial

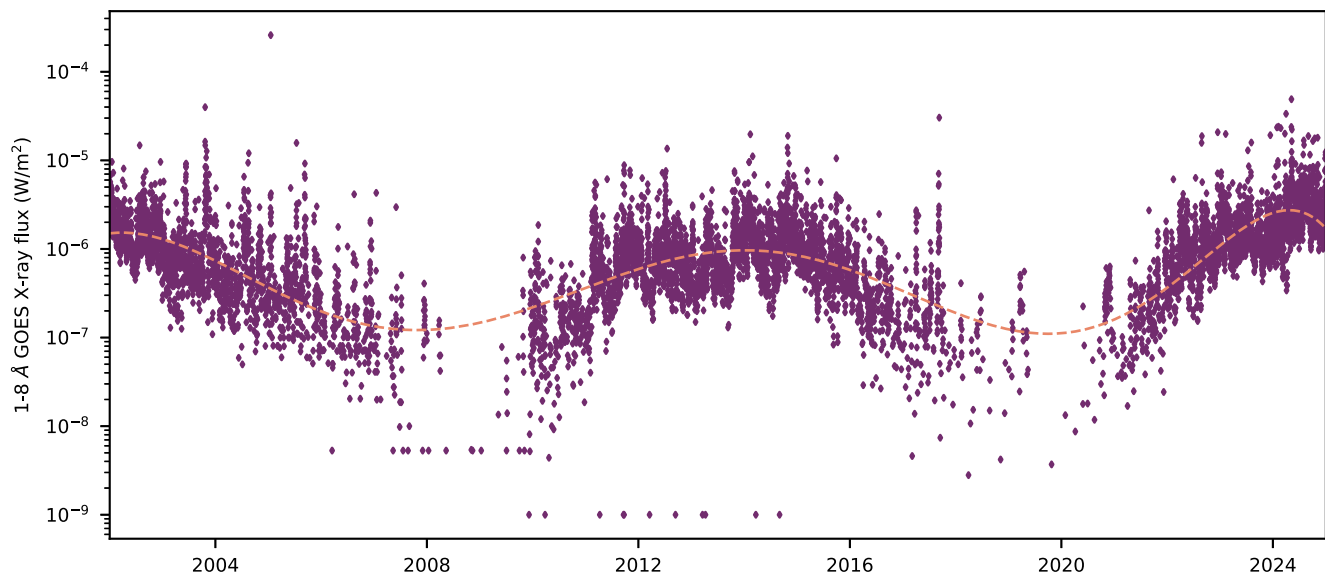
Decree no. 118 of 2023 March 2, based on the NRRP—funded by the European Union—NextGenerationEU—Mission 4 “Education and Research,” Component 1 “Enhancement of the offer of educational services: from nurseries to universities”—Investment 4.1 “Extension of the number of research doctorates and innovative doctorates for public administration and cultural heritage”—CUP E66E23000110001. S.M. acknowledges the PhD course in Astronomy, Astrophysics and Space Science of the University of Rome “Sapienza,” University of Rome “Tor Vergata,” and Istituto Nazionale di Geofisica e Vulcanologia, Italy.

## Data Availability

The ASR catalog and its code can be freely accessed at [https://github.com/helio-unitov/ASR\\_cat](https://github.com/helio-unitov/ASR_cat).

## Appendix A Background Characterization

In Figure 4 we show the observed flux background over the full temporal length of our data set. The clear reproduction of the solar activity cycle emphasizes the importance of accounting for background variations when estimating the event integral, ensuring an unbiased measure of flare energetics. The concentration of points at lower energies represents the full scale of the instrument.



**Figure 4.** Observed background flux over the entire duration of the data set. The orange dashed line shows a seventh-degree polynomial fit. For clarity, we included only one event out of every 30 to avoid overcrowding with overlapping flares.

## Appendix B

### Data Product Description

The head of the ASR catalog is organized in 13 columns. A full explanation for each tag is given in the paper, and in Table 2 we provide a brief description of each column.

Furthermore, the three catalogs that also report the position of the flare events (one for each class) have six additional columns (Table 3) that accommodate the flares coordinates in different reference frames.

**Table 2**  
Description of the Columns Included in the Flare Catalog

Column	Description
flare_id	Identifying number of a flare; not unique as consecutive flares share the same ID.
tstart	Flare starting time in ISO8601 date–time format.
tpeak	Flare peak time in ISO8601 date–time format.
tend	Flare end time in ISO8601 date–time format.
BG_flux	Background flux computed over the 2 minute interval before the event.
flux_integral	Flux integral of the flare event from tstart to tend.
ratio	Ratio between background and peak fluxes.
flux_int_corrected	Flux integral from tstart to tend, accounting for background flux.
delta_flux	Difference between peak flux and background flux.
fclass_simple	Flare class estimated based on delta_flux, accounting for background flux.
fclass_full	Full class of the flare estimated as in fclass_simple.
abs_class_simple	Flare class estimated according to NOAA.
abs_class_full	Full class of the flare estimated as in abs_class_simple.

**Table 3**  
Description of the Additional Columns Included in the Flare Catalogs with the Positions

Column	Description
flare_x	Position of the flare in pixel coordinates
flare_y	Position of the flare in pixel coordinates
Lon [Helioprojective]	Longitude of the flare in the helioprojective reference frame
Lat [Helioprojective]	Longitude of the flare in the helioprojective reference frame
Lon [Heliographic]	Longitude of the flare in the heliographic Stonyhurst reference frame
Lat [Heliographic]	Longitude of the flare in the heliographic Stonyhurst reference frame

### ORCID iDs

M. Berretti  <https://orcid.org/0009-0007-2465-1931>  
 S. Mestici  <https://orcid.org/0009-0000-2442-8878>  
 L. Giovannelli  <https://orcid.org/0000-0001-7369-8516>  
 D. Del Moro  <https://orcid.org/0000-0003-2500-5054>  
 M. Stangalini  <https://orcid.org/0000-0002-5365-7546>  
 F. Giannattasio  <https://orcid.org/0000-0002-9691-8910>  
 F. Berrilli  <https://orcid.org/0000-0002-2276-3733>

### References

- Aschwanden, M. J. 2011, *Self-organized Criticality in Astrophysics* (Berlin: Springer)
- Aschwanden, M. J., & Freeland, S. L. 2012, *ApJ*, **754**, 112
- Barnes, G., Leka, K. D., Schrijver, C. J., et al. 2016, *ApJ*, **829**, 89
- Berrilli, F., Casolino, M., Del Moro, D., et al. 2014, *JWSWC*, **4**, A16
- Bobra, M. G., & Couvidat, S. 2015, *ApJ*, **798**, 135
- Camporeale, E. 2019, *SpWea*, **17**, 1166
- Cicogna, D., Berrilli, F., Calchetti, D., et al. 2021, *ApJ*, **915**, 38
- Cloutier, S., Cameron, R. H., & Gizon, L. 2024, *A&A*, **691**, A9
- Deng, Z., Wang, F., Deng, H., et al. 2021, *ApJ*, **922**, 232
- Di Fino, L., Zaconté, V., Stangalini, M., et al. 2014, *JWSWC*, **4**, A19
- Hanser, F. A., & Sellers, F. B. 1996, *Proc. SPIE*, **2812**, 344
- Huang, X., Wang, H., Xu, L., et al. 2018, *ApJ*, **856**, 7
- Jonas, E., Shankar, V., Bobra, M., & Recht, B. 2016, *AGUFM*, **2016**, SH34A–02
- Leka, K. D., 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, J. 2011, *ApJ*, **735**, 130
- MacAlester, M. H., & Murtagh, W. 2014, *SpWea*, **12**, 530
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. 2021, *ACM Comput. Surv.*, **54**, 115
- Nishizuka, N., Kubo, Y., Sugiura, K., Den, M., & Ishii, M. 2021, *EP&S*, **73**, 64
- Oughton, E. J., Skelton, A., Home, R. B., Thomson, A. W. P., & Gaunt, C. T. 2017, *SpWea*, **15**, 65
- Pandey, C., Angryk, R. A., & Aydin, B. 2023, in *Machine Learning and Knowledge Discovery in Databases: Applied Data Science and Demo Track. ECML PKDD 2023*, ed. G. De Francisci Morales et al. (Berlin: Springer), 72
- 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
- Plutino, N., Berrilli, F., Del Moro, D., & Giovannelli, L. 2023, *AdSpR*, **71**, 2048
- Qahwaji, R., & Colak, T. 2007, *SoPh*, **241**, 195
- Schrijver, C. J. 2007, *ApJL*, **655**, L117
- van der Sande, K., Flyer, N., Berger, T. E., & Gagnon, R. 2022, *FrASS*, **9**, 354
- Veronig, A., Temmer, M., & Hanslmeier, A. 2002a, *HvaOB*, **26**, 7
- Veronig, A., Temmer, M., Hanslmeier, A., Otruba, W., & Messerotti, M. 2002b, *A&A*, **382**, 1070
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Yashiro, S., Akiyama, S., Gopalswamy, N., & Howard, R. A. 2006, *ApJL*, **650**, L143
- Yu, D., Huang, X., Hu, Q., et al. 2010, *ApJ*, **709**, 321