



RESEARCH ARTICLE

10.1002/2015SW001170

Key Points:

- To show the connection between X-ray flare class and background level
- To show the usefulness of the X-ray background and ratio in flare prediction
- To build a framework for an X-ray flare forecast using the GOES satellite

Correspondence to:

L. M. Winter,
lwinter@aer.com

Citation:

Winter, L. M., and K. Balasubramaniam (2015), Using the maximum X-ray flux ratio and X-ray background to predict solar flare class, *Space Weather*, 13, 286–297, doi:10.1002/2015SW001170.

Received 16 JAN 2015

Accepted 1 APR 2015

Accepted article online 8 APR 2015

Published online 21 MAY 2015

Using the maximum X-ray flux ratio and X-ray background to predict solar flare class

L. M. Winter¹ and K. Balasubramaniam²

¹Atmospheric and Environmental Research, Superior, Colorado, USA, ²Air Force Research Laboratory, Albuquerque, New Mexico, USA

Abstract We present the discovery of a relationship between the maximum ratio of the flare flux (namely, 0.5–4 Å to the 1–8 Å flux) and nonflare background (namely, the 1–8 Å background flux), which clearly separates flares into classes by peak flux level. We established this relationship based on an analysis of the Geostationary Operational Environmental Satellites X-ray observations of ~ 50,000 X, M, C, and flares derived from the NOAA/Space Weather Prediction Center flares catalog. Employing a combination of machine learning techniques (K-nearest neighbors and nearest centroid algorithms) we show a separation of the observed parameters for the different peak flaring energies. This analysis is validated by successfully predicting the flare classes for 100% of the X-class flares, 76% of the M-class flares, 80% of the C-class flares, and 81% of the B-class flares for solar cycle 24, based on the training of the parametric extracts for solar flares in cycles 22–23.

1. Introduction

Solar flares release intense amounts of energy into the interplanetary medium. A statistical concept of how this energy is released is through an avalanche model [Lu and Hamilton, 1991; Lu et al., 1993; Lu, 1995]. According to common consensus, the creation of a solar flare begins with magnetohydrodynamic instabilities that release energy stored in the local magnetic field lines through an untwisting of the field lines. Conservation of magnetic energy leads to instabilities in nearby regions through an avalanche process that accelerates energetic particles along the large-scale magnetic field lines. Soft X-rays in the solar corona, which are the topic of the presented analysis, are emitted from the associated magnetic loops, which are, in turn, connected to the active-region surface magnetic fields, as seen in the photosphere. In this model, all flares are the result of the same physical processes and the ultimate strength of the flare is related to the cascaded number of reconnection events creating the flare.

While the basic mechanism of solar flares is believed to be understood, the detailed physical processes are still too complex to be modeled in a deterministic way for predictions of when a flare will occur. This is a challenge, since the space weather effects of the energy release can lead to diverse problems to human technology such as satellite damage or inoperability, ionospheric communication interference, and power grid failures. Therefore, understanding the solar conditions (e.g., magnetic activity and coronal temperature) that precede and lead to solar flares is of vital importance to enhance current predictions of space weather phenomena.

Current solar flare prediction models rely upon empirical observations, with many predictions based on tracking the properties of solar active regions [e.g., Gallagher et al., 2002; Barnes et al., 2007; Falconer et al., 2011; Ahmed et al., 2013; Balasubramaniam, 2013]. Since the magnetic active regions are the source of the magnetic energy released in the flare, this approach is well justified. However, other observable properties may also provide a diagnostic for the underlying physical conditions in the solar corona leading to flares. In particular, we present evidence for two easily observable soft X-ray measurements that diagnose the nonflare magnetic energy and coronal temperature (in section 2). The parameter space of these observables distinguishes properties of solar flares of different classes, as described in section 3. Machine learning techniques are used to build classification models applied to historical flares in section 4. Finally, the results from the statistical soft X-ray analysis of solar flares are discussed in section 5.

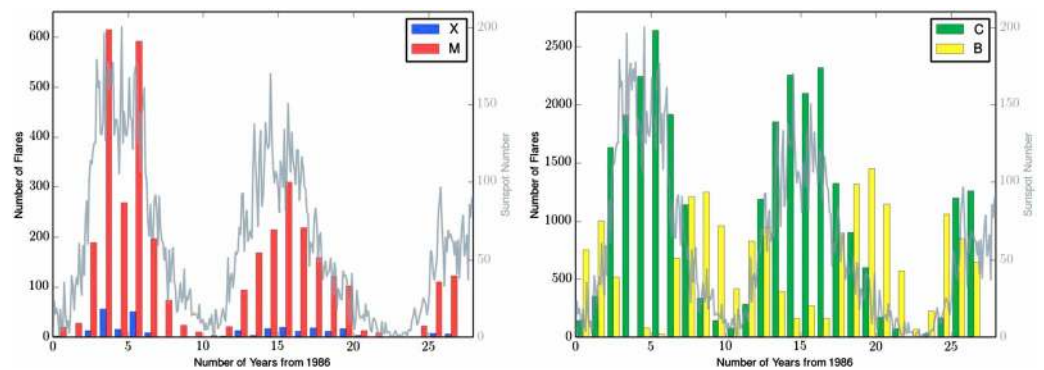


Figure 1. The distribution of flares of each type in the NOAA flare list, including observations from 1986 to July 2014. Gray lines trace the monthly sunspot number. The majority of X-, M-, and C-class flares occurs close to solar maximum, the maximum in sunspot number.

2. X-ray Flare Data

We analyzed the historical X-ray data from NOAA’s *GOES* X-Ray Sensor (XRS). The *GOES* observations include X-ray flux measurements averaged over every 1 min observed in both a short-wavelength and long-wavelength X-ray band (short: 0.5–4 Å and long: 1–8 Å) from 1986 to 2014. The NOAA flare lists include nearly 50,000 X-ray flares in this time span which covers solar cycles 22–24. The flare classifications fall into the following classes based on the peak flux level: X ($> 10^{-4} \text{ W m}^{-2}$), M ($> 10^{-5} \text{ W m}^{-2}$), C ($> 10^{-6} \text{ W m}^{-2}$), and B ($> 10^{-7} \text{ W m}^{-2}$). Figure 1 shows the distribution of the flares of each type by year. The majority of the stronger flares occur close to solar maximum, indicated in the plots by the maxima in sunspot number (obtained from the Solar Influences Data Analysis Center in Belgium).

The NOAA flare lists include the start, peak, and end time of the flares along with the flare location, if known, and X-ray class. For our analysis, we use the start and peak time along with the flare class. The start of a flare is defined as when four consecutive 1 min 1–8 Å flux measurements meet all of the following conditions: (1) All four values $> 10^{-7} \text{ W m}^{-2}$, (2) each consecutive measurement has a higher flux than the previous measurement, and (3) the last value is $> 1.4\times$ the measurement from 3 min earlier.

Using the downloaded XRS data in both the short and long bands, we measured the long X-ray background flux (B) and the ratio of the short to long bands:

$$R = \frac{F_{0.5-4\text{\AA}}}{F_{1-8\text{\AA}}}. \tag{1}$$

A full description of our method to determine B is included in *Winter and Balasubramaniam* [2014], where we show that the long X-ray background, but not the short X-ray background, varies along with the solar cycle for solar cycles 22–24. This soft X-ray background variation was also observed for solar cycle 21 by *Wagner* [1988] and *Aschwanden* [1994]. The background is measured as the minimum 1–8 Å flux in the preceding 24 h for each 1 min XRS measurement, following the procedure of *Hock et al.* [2013]. This background measurement is similar to the X_{b10} index, used in operational forecasting as an integrated irradiance proxy by *Tobiska and Bouwer* [2006]. The ratio R is computed from the start time of each flare until the peak, using the dates from the NOAA flare lists. In a small number of cases, $\sim 2\%$ of the total flares, inaccuracies in the flare list show a start time that occurs after the peak time. These cases were not considered in the final analysis. The maximum R value, R_{max} , was next computed for each flare. Figure 2 shows an example of the X-ray flux, X-ray background level, and R for a flare. In computing R_{max} , we required that more than one R measurement must exist and that both the short and long X-ray flux $> 10^{-9} \text{ W m}^{-2}$. These criteria excluded 22% of the total flares from further analyses with the R_{max} parameter, including near-instantaneous flares with short (~ 1 min) rise times.

Table 1 includes the average statistics for each flare class. These statistics include the total number of flares, time between R_{max} and 1–8 Å peak flux (where measurements were possible following the criteria from the previous paragraph), peak flux, background flux in the short and long band, and R_{max} . Since the ratio R is

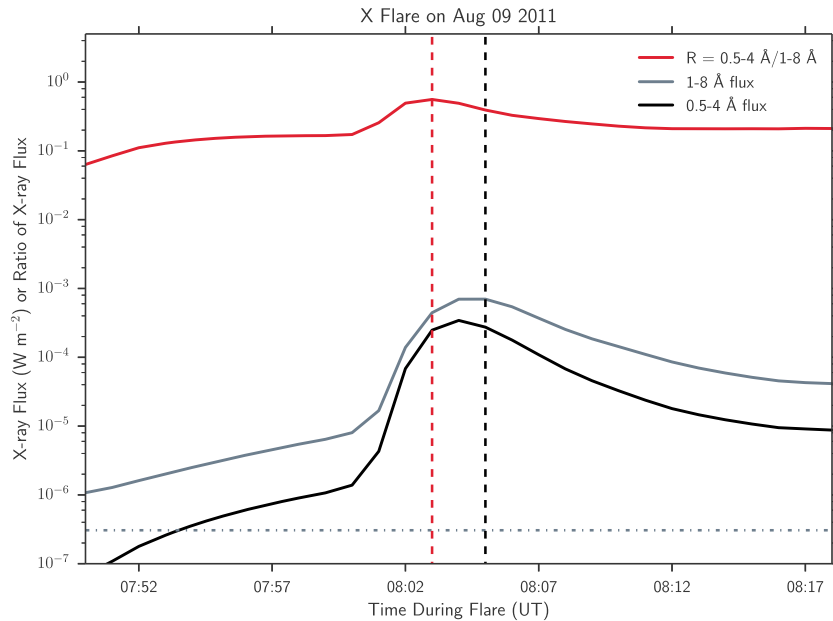


Figure 2. For the ~50,000 flares in the NOAA flare lists, occurring from 1986 to present, we calculated the maximum ratio of the short (0.5–4 Å) to long (1–8 Å) X-ray bands. An example is shown for an X-class flare from 2011. The ratio R (red), long-wavelength (gray), and short-wavelength (black) X-ray flare profiles are shown, with the maximum in R marked with a red dashed line and the 1–8 Å flare peak marked with a black dashed line. The dash-dotted line marks the 1–8 Å nonflare background, B .

related to coronal temperature [see, e.g., Thomas et al., 1985; Garcia, 1994; Feldman et al., 1996; White et al., 2005; Ryan et al., 2012], the R_{\max} value is related to the maximum temperature occurring during the flare, where

$$T_{\max} = A_0 + A_1 R_{\max}^2 + A_3 R_{\max}^3 \text{ MK.} \quad (2)$$

A_n are coefficients found in Table 2 of White et al. [2005]. We note that this direct assumption is a simplification based on an isothermal flare. As discussed in White et al. [2005], more extensive investigations of the multi-thermal flare properties are not possible with the two flux measurements provided with the GOES XRS but the ratio is still a useful tool in studying the overall energetics of the flares. We find that the maximum ratio is significantly higher for the strongest flares (e.g., 0.33 for X flares and 0.05 for B flares). As shown in the table, this maximum temperature, on average, occurs before the peak in the 1–8 Å flux, in part due to the fact that the 0.5–4 Å flux peaks ahead of the 1–8 Å flux by up to 20 min. The rise time from flare onset to R_{\max} is longer for X and M flares than C and B flares, consistent with statistical results presented by, e.g., Veronig et al. [2002].

We find that the average long-wavelength background flux is higher for the stronger flares (X and M) than the weaker flares (C and B). This is also shown in Figure 3, with contour plots of the multivariate density estimates for the peak flux and background flux. The density estimates are created with the kernel density method,

Table 1. X-ray Flare Statistics^a

Class	N	$\langle T(P - R_{\max}) \rangle$ (min)	$\langle P_{1-8\text{\AA}} \rangle$ (10^{-7} W m^{-2})	$\langle B_{0.5-4\text{\AA}} \rangle$ (10^{-8} W m^{-2})	$\langle B_{1-8\text{\AA}} \rangle$ (10^{-6} W m^{-2})	$\langle R_{\max} \rangle$
X	290	6.7 ± 13.2	2400 ± 1700	1.9 ± 2.1	1.2 ± 0.7	0.33 ± 0.08
M	3,742	4.8 ± 9.9	240 ± 180	1.6 ± 1.7	1.2 ± 0.7	0.18 ± 0.06
C	28,803	3.0 ± 7.0	31 ± 20	0.8 ± 1.0	0.8 ± 0.5	0.08 ± 0.04
B	15,751	2.7 ± 6.3	4.9 ± 2.6	0.2 ± 0.2	0.2 ± 0.1	0.05 ± 0.15

^aThe statistics include number of flares (N), average and standard deviation of the time between the 1–8 Å Peak and R_{\max} ($\langle T(P - R_{\max}) \rangle$), average and standard deviation of the peak flux ($\langle P_{1-8\text{\AA}} \rangle$), average and standard deviation of the background 0.5–4 Å flux ($\langle B_{0.5-4\text{\AA}} \rangle$), average and standard deviation of the background 1–8 Å flux ($\langle B_{1-8\text{\AA}} \rangle$), and average and standard deviation of the maximum ratio of 0.5–4 Å/1–8 Å ($\langle R_{\max} \rangle$), for each flare class.

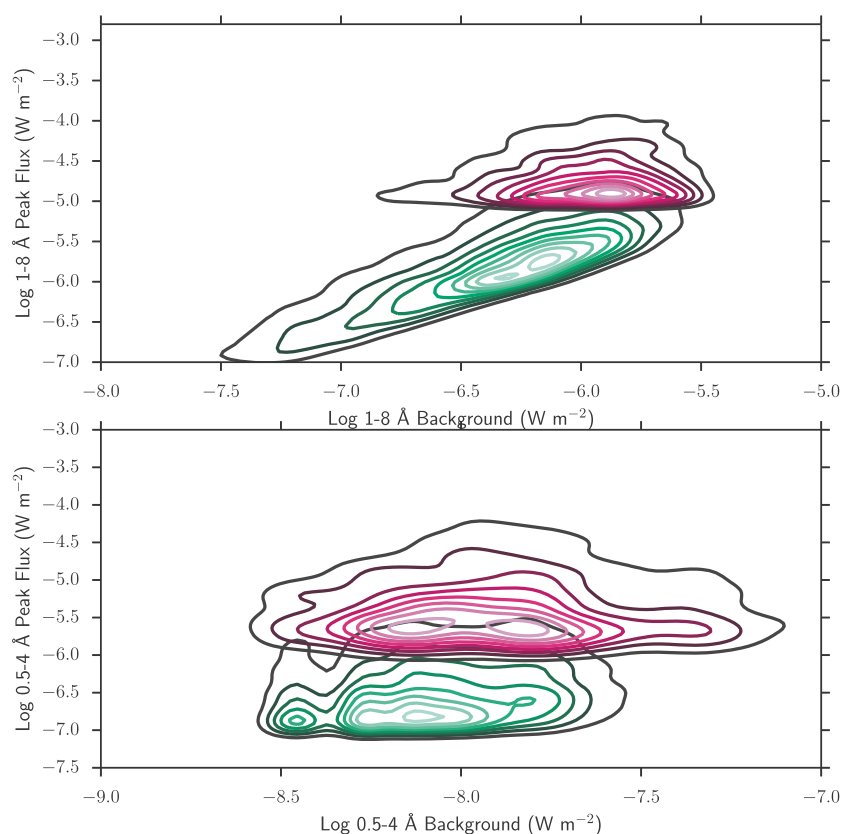


Figure 3. Kernel density estimates showing the two-dimensional probability distribution of peak flux and background flux in both the (top) long-wavelength and (bottom) short-wavelength X-ray bands. The green contours show the density distributions for all flares, which are dominated by the more numerous C- and B-class flares, while the red contours show the density distributions for X- and M-class flares. There is no correlation between short-wavelength peak flux and background, while there is a positive correlation in the relationship between long-wavelength peak flux and background ($R^2 = 0.45$, see section 2 for details).

using Gaussian filters, through the scientific Python (scipy) `gaussian_kde` function [Scott, 2009]. A linear correlation is found between the peak flux and background in the long-wavelength band with

$$\log B = (1.03 \pm 0.01) \times \log F_{\text{peak},1-8\text{\AA}} + (0.86 \pm 0.04) \quad (3)$$

and a correlation coefficient $R^2 = 0.45$. This is particularly evident in the full sample of flares, whose statistics are dominated by the more numerous B- and C-class flares (shown in green). However, the trend does not exist when examining the distribution of M- and X-class flares alone (shown in red, $R^2 = 0.07$). At high peak flux levels, there is no difference between the background levels (this is discussed again in the following section). Similarly, we computed density plots for the short-wavelength peak and background. Given the issue of the 0.5–4 Å background being close to or below the instrumental limit, we chose to examine flares occurring from 1999 to 2006, a time period with higher measured backgrounds that includes the rise through fall phases of solar cycle 23 solar maximum. No correlation exists between peak flux and background in the short-wavelength band ($R^2 < 0.01$). This further illustrates that the background in the long-wavelength band and not the short-wavelength band is an appropriate observational parameter that is tied to the peak flare flux.

3. Separation Into Solar Flare Classes

With the extensive database of X-ray flare properties, we investigated whether measurements based on the 1 min GOES XRS observations showed properties useful for predicting the X-ray class. Specifically, in order to build a classification model based on the properties of past flare events, we identified properties that lend themselves to the use of classification techniques by showing a separation between different flare classes in

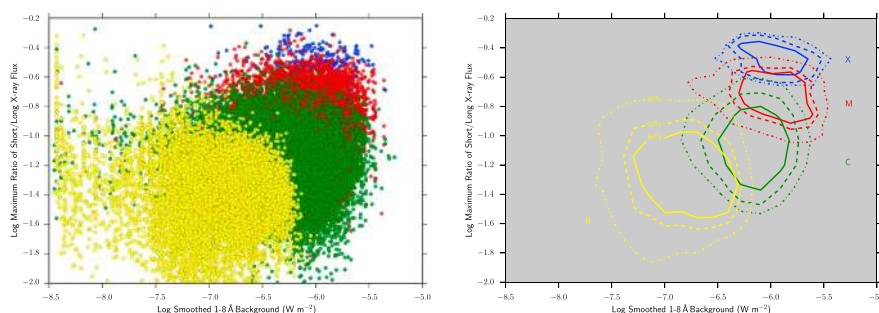


Figure 4. The observational parameters of 1–8 Å nonflare X-ray background flux and the maximum ratio of the 0.5–4 Å/1–8 Å flux (R_{\max}) separate the NOAA flares effectively into different parameter space based on the peak flux. (left) A scatterplot of the measured parameters of the 50,000 NOAA flares (with color coding corresponding to flare class as blue = X, red = M, green = C, and yellow = B). (right) Contour levels display levels enclosing 50% (solid line), 68% (dashed line), and 85% (dash-dotted line) of the X- (blue), M- (red), C- (green), and B- (yellow) class flares. Lower peak flux occurs when the background is also low. High peak flux occurs when the background flux and R_{\max} parameters are high.

their parameter space. The 1–8 Å nonflare background and the R_{\max} parameters yielded such a separation, shown in Figure 4. For each of the flare classes, the contours represent the parameter space of the diagram including 50%, 68% (the 1 sigma contour level), and 85% of the flares in the given class. The X-class flares occupy the right corner of the diagram, indicating high background flux and high R_{\max} . The M-class flares share a similar range of background flux, but with lower R_{\max} than the X-class flares. This is also evident in Table 1, where the average and standard deviation in B are consistent between X and M flares while the average R_{\max} is significantly higher for the X class. Similarly, C-class flares share the range of backgrounds with X- and M-class flares but have lower R_{\max} values. However, the B-class flares have significantly lower measurements of the background flux, with a similar range of R_{\max} to the C-class flares. Due to the upper flux limit of B-class flares ($< 10^{-6} \text{ W m}^{-2}$), they cannot be observed when the background is high.

To further test the apparent difference in the B and R_{\max} parameters of different flare classes, we computed Kolmogorov-Smirnov statistics [Kolmogorov, 1933; Smirnov, 1948]. The Kolmogorov-Smirnov two sample test determines the probability of two samples being drawn from the same distribution. To do this, the cumulative distribution of each sample is computed and the maximum distance between the two chosen distributions is determined as the Kolmogorov-Smirnov (KS) statistic. When the KS statistic is small and the two-tailed p value approaches 1, the null hypothesis of both samples being drawn from the same distribution cannot be rejected. The two sample KS test was run using the `scipy ks_2samp` function [Oliphant, 2007]. The test was run on all combinations of two flare classes (e.g., sample 1 as X-class flares and sample 2 as M-class flares) for each of the parameters B and R_{\max} . Results are given in Table 2, including the KS statistic and the two-tailed p value. These statistics show that the long X-ray background is consistent

with being drawn from the same population for X- and M-class flares. To a lesser degree, the KS statistic is low (~ 0.3) for comparisons between the distribution of background flux of X- and M-class flares with C-class flares, but the low p values (< 0.001) indicate that these distributions are distinct. Additional comparisons between the distributions of B and R_{\max} for the flare classes result in high KS statistics (from 0.395 to 0.985) and low p values (< 0.001). This suggests that the distributions of values in the B - R_{\max} parameter space are distinct. This is also evident in Figure 4, where we show that the majority of flares of each class (e.g., the 1 sigma or 68% contour level) occupy a distinct parameter space.

This separation into flare classes hints at differences in the physical conditions of the solar corona. The long X-ray background, B , is the nonflare flux level associated with active regions. It can be construed as a proxy

Table 2. Kolmogorov-Smirnov Statistics for Two Sample Comparisons^a

Classes	B		R_{\max}	
	KS	p	KS	p
X, M	0.056	0.600	0.776	0.000
X, C	0.298	0.000	0.976	0.000
X, B	0.857	0.000	0.985	0.000
M, C	0.283	0.000	0.706	0.000
M, B	0.806	0.000	0.888	0.000
C, B	0.695	0.000	0.395	0.000

^aThe K-S statistic (KS) and probability value (p) are listed for tests on the distributions of the long X-ray background and R_{\max} for all combinations of comparisons of two X-ray flare classes.

for the magnetic energy of the corona. The R_{\max} measurement is associated with coronal temperature, as well as radiative losses and emission measure [see, e.g., Thomas et al., 1985; Garcia, 1994; Feldman et al., 1996; White et al., 2005; Ryan et al., 2012]. A possible explanation for the separation of flare classes in the B - R_{\max} is that the built-up energy of the regions that produce the flares (measured by B) is directly related to the amount of energy released in the flare (measured by R_{\max}). Since the background measurement is an average over the entire Sun and not just the flare site, we expect that a more careful analysis where B is replaced by a measurement of the energy/flux of the flare site alone would reveal a tighter correlation with R_{\max} . This, however, is a more difficult measurement to make for a real-time forecast situation.

4. Machine Learning Classification

To quantify the separation into flare classes shown in the B - R_{\max} parameter space, we used machine learning classification techniques. Specifically, we used the K-nearest neighbors and nearest centroid algorithms from the Python machine learning library, `scikit-learn` [Pedregosa et al., 2011]. These classifiers build predictions using input data from a training set of data with known classes. Our training set included X-ray flares from solar cycles 22 to 24. The input parameters were the X-ray background and the maximum ratio of short to hard X-ray flux (R_{\max}), with the classes labeled as X, M, C, and B. The classifier algorithms then use the training set to predict what the class of a new flare event will be. In section 4.1, the machine learning classification techniques are described. The machine learning algorithms applied to the input data create what is termed a model, which is a statistical model based on training data that is used to make predictions for new data sets. Results of our analysis applying our statistical models are included in section 4.2.

4.1. Statistical Model Descriptions

For the K-nearest neighbor classifier, the parameter space of the logarithm of X-ray background versus the logarithm of R_{\max} is broken up into a grid. Along each point in the grid, the number of data points of each possible class in the k nearest neighboring points are counted, where k is a user-defined integer. Whichever class corresponds to the most neighboring data points is adopted in the model as the likely class of any new flares with the same values of the X-ray background and R_{\max} at that grid point. For our analysis, the grid size was 0.01 and the k neighboring points used was 5. As an example of computing a classification for one grid point, we consider the point where $\log R_{\max} = -0.8$ and \log of the X-ray background = -8 . Figure 4 shows that the five nearest data points to the selected point include four B-class flares and one C-class flare. Therefore, an unknown flare at the selected point would be classified as a B-class flare.

Alternatively, the nearest centroid algorithm uses the distance from the centroid of the distribution of points in each class of the training set for predictions. For example, we consider the classification of a point based on the distance from the centroid of the X-class flares ($\log R_{\max} = -0.48$ and \log X-ray background = -5.92) and M-class flares ($\log R_{\max} = -0.72$ and \log X-ray background = -5.92). To classify an unknown flare with $\log R_{\max} = -0.5$ and \log X-ray background = -5.9 , we calculate the Euclidean distance of the point to the centroid of each of the classes. The selected point is a distance of $\sqrt{(-0.5 - -0.48)^2 + (-5.9 - -5.92)^2} = 0.03$ from the X-class centroid and, similarly computed, 0.22 from the M-class centroid. Since it is closer to the X-class centroid, the selected point is classified as an X-class flare.

Since the K-nearest neighbor approach weights according to the number of points along the grid, it is more accurate for classifying C- and B-class flares, which include 4–100× more flares than the M- and X-class flares. Meanwhile, since it is based solely on the distance from the centroid of the parameters for a given class, the nearest centroid method does a better job at classifying the X-class flares. Using these machine learning methods, classification models were built with the B and R_{\max} measurements from flares from solar cycles 22 to 24. Since there are relatively few, less than 300, X-class flares in the entire sample spanning nearly three decades, the advantage of using the cycle 22–24 data as a training set is that the resultant model will include as many X-class flares as possible. These statistical models built with the full flare data set are shown in Figure 5.

To create statistical models that do not include the training set data but still include a larger number of X-class flares, we also built models using only the flare parameters from solar cycles 22 and 23. These models were then used to predict the solar cycle 24 flare classifications (as shown in Figure 6). A concern with using this approach is whether the flare behavior during the much weaker solar cycle 24 is different from the flares in cycles 22 and 23 that were used to create the model. To test whether there are differences in flare

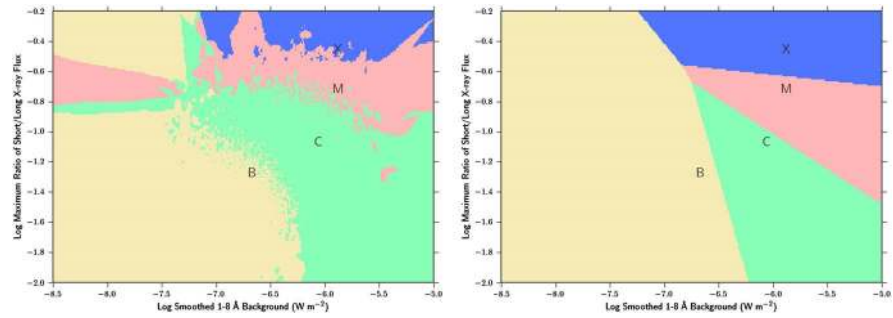


Figure 5. Classification models built from the solar cycle 22–24 flare parameters for the (left) K-nearest neighbor and (right) nearest centroid algorithms. The K-nearest neighbor model classifications use the classes (X, M, C, and B) of the five nearest points in the training set to each grid point (0.01×0.01) to predict the class of a flare with the B and R_{\max} values at that grid point. The nearest centroid model predicts classes based on the Euclidean distance; each grid point is from the centroid of B and R_{\max} for each class (X, M, C, and B). See the text for more details on the classification algorithms.

rate between the solar cycles, we determined the occurrence frequency rate as a function of 1–8 Å peak flux during the rise to solar maximum and solar maximum phases for each of the solar cycles. The occurrence frequency distributions were fit with a power law of the form $N(F_{1-8\text{Å}}) = N_1 \times (F_{1-8\text{Å}})^{-\alpha+1}$, where N

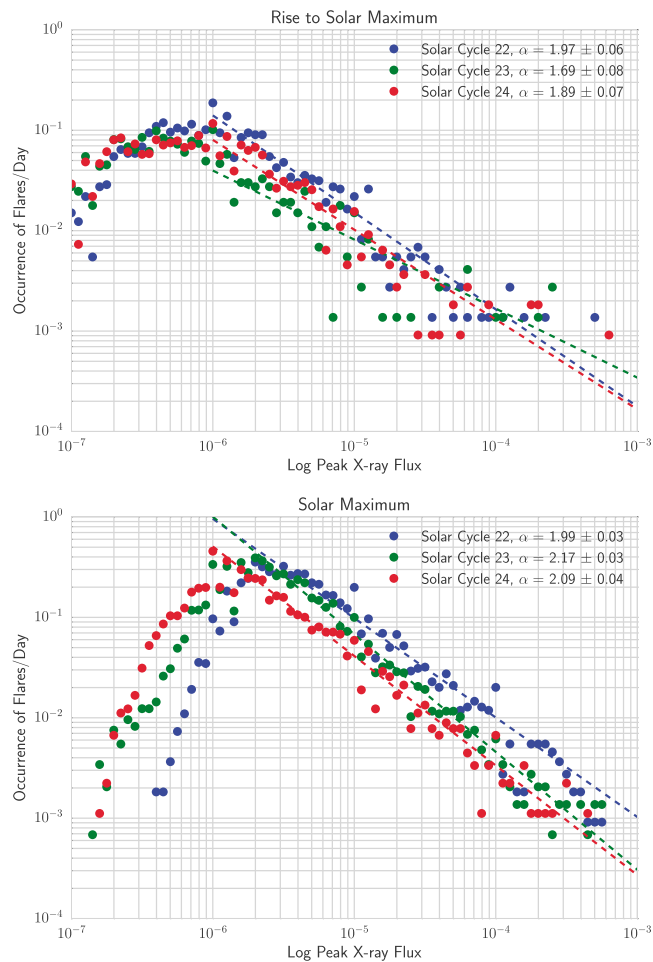


Figure 6. Occurrence frequency distributions of the peak 1–8 Å flux are shown for solar cycles 22–24 both during the rise to solar maximum and during solar maximum. Flare rates are similar between solar cycles for the rise to solar maximum and during solar maximum.

Table 3. Best Fit Parameters for Power Law Fits to the Occurrence Flare Rates^a

Solar Cycle Phase	N_1	α	χ^2/dof
22 Rise	-6.65 ± 0.29	1.97 ± 0.06	33.2/35
23 Rise	-5.53 ± 0.40	1.69 ± 0.08	40.8/23
24 Rise	-6.46 ± 0.34	1.89 ± 0.07	34.7/29
22 Maximum	-5.96 ± 0.15	1.99 ± 0.03	74.7/47
23 Maximum	-7.03 ± 0.16	2.17 ± 0.03	28.7/46
24 Maximum	-6.84 ± 0.20	2.09 ± 0.01	28.6/41

^aOccurrence frequency distribution for the rise phase (Rise, from the beginning of the solar cycle toward maximum) and during solar maximum (Maximum) were fit with a power law model (see section 4.1 for details). The best fit value and errors are shown for the normalization factor (N_1 , the normalization factor for the logarithm of the 1–8 Å peak flux in W m^{-2}) and the power law index (α). The goodness of fit is assessed with the reduced χ^2 statistic (χ^2 divided by the degrees of freedom, dof).

is the occurrence frequency (flares rate/d), $F_{1-8\text{Å}}$ is the logarithm of flare peak flux (W m^{-2}), and the fit parameters include the normalization factor, N_1 , and the power law index, α . We utilize the Levenberg-Marquardt least squares minimization technique [Levenberg, 1944; Marquardt, 1963] to determine the best fit function parameters, fitting the occurrence rate where peak flux $> 10^{-6} \text{W m}^{-2}$. We omit the B-class flares from these fits since the low end of the power law distribution is near the GOES detection threshold. Goodness of fit is assessed with the χ^2 statistic, defined as $\chi^2 = \Sigma(\text{data} - \text{model})^2 / \text{std}^2$, where the data are the frequency distribution values (N), the model is the power law fit, and std is the standard deviation for each of the measurements of N . Good fits are those where χ^2/dof are close to unity, where dof are the degrees of freedom or number of data points – number of free parameters that are fit.

Using the solar X-ray background analysis from Winter and Balasubramaniam [2014], the rise phases are defined as occurring from August 1986 to August 1988 (solar cycle 22), May 1996 to May 1998 (solar cycle 23), and December 2008 to December 2011 (solar cycle 24), and the solar maximum phases are defined as August 1988 to August 1991 (solar cycle 22), May 1999 to May 2003 (solar cycle 23), and December 2011 to December 2014 (solar cycle 24, noting that this is an incomplete solar maximum phase including flares up until the end of the period examined in the paper). Plots of the frequency occurrence rates are included in Figure 6, and the best fit parameters are in Table 3. Aschwanden [2011] includes power law estimates from past analyses of frequency distributions of X-ray flares, which find α to range from 1.58 to 2.0. They find that the slopes change throughout the solar cycle, with flatter slopes during solar maximum. Our rates are consistent with these values. Additionally, as in Aschwanden [2011], we find that power law slopes are similar during the same phase (e.g., solar maximum) but have different normalizations. Therefore, we conclude that the power law slopes of the frequency rates for flares are consistent between solar cycles, and as a result, we can effectively utilize the solar cycle 24 flares as an appropriate test set for classification models built with the solar cycle 22–23 flare parameters.

4.2. Results

Table 4 presents statistics on the percent of correct identifications (PC), the number of true classifications (TC; number of flares where the correct flare class is predicted) divided by the total number of flares (N) examined. The PC computations show the ability of the models, built with solar cycle 22–24 and solar cycle 22–23 flare parameters, to correctly classify flares from the test sets of flare parameters, solar cycle 22–24, and 24 flares. The models built with the solar cycle 22–24 data correctly classify $\sim 90\%$ of the flares with the K-nearest neighbor model and $\sim 75\%$ with the nearest centroid model. The nearest centroid model correctly classifies 95.9% of X-class flares with the solar cycle 22–24 tested model and correctly classifies all of the solar cycle 24 X-class flares. The K-nearest neighbor model better predicts the M-, C-, and B-class flares, correctly classifying 66.6% of M-class flares, 91.8% of C-class flares, and 89.1% of B-class flares from solar cycles 22 to 24. The model

Table 4. Classification Model Statistics^a

Class	<i>N</i>	KNN PC	NC PC	<i>N</i>	KNN PC	NC PC	<i>N</i>	KNN PC	NC PC
Model	22–24	22–24	22–24	22–24	22–24	22–24	22–23	22–23	22–23
Test set	22–24	22–24	22–24	24	24	24	24	24	24
All	39,391	88.9	75.0	7032	88.9	73.7	7032	83.4	73.0
X	196	59.2	95.9	23	73.9	100	23	73.9	100
M	2,964	66.6	70.5	349	80.5	53.6	349	75.6	51.0
C	23,425	91.8	72.9	3987	90.6	73.7	3987	85.5	72.8
B	12,806	89.1	79.7	2673	87.4	76.2	2673	81.4	76.0

^aModels were built using the K-nearest neighbor (KNN) and nearest centroid (NC) methods, using the solar cycle 22–24 flare parameters. The number of flares in each category (*N*), along with the percent of correct classifications are shown for the model built and applied to the flare data (e.g., KNN PC is the percent correct for the KNN model). The solar cycles used to build each of the models are listed in the row labeled *Model*, and the listed statistics are for testing the model on the solar cycles listed in the row labeled *Test Set*.

does a better job of classifying the solar cycle 24 flares, with correct classifications of 80–90% of M through B flares. From the tests of the solar cycle 22–24 and solar cycle 22–23 built models, the performance is similar in correctly identifying solar cycle 24 flares, but the classifications are slightly better for the solar cycle 22–24 built models (by ~ 5% overall) since the training set includes the solar cycle 24 flares.

Additional skill scores were computed to better quantify the results, shown in Table 5. These skill scores include the probability of detection (POD), false alarm rate (FAR), Heidke skill score (HSS; see *Heidke* [1926]), and true skill score (TSS; defined in *Hanssen and Kuipers* [1965]). For each flare class, the following values were computed: the true classifications (TC), false null classifications (FN; number of flares in the class incorrectly predicted not to be in the flare class), false classifications (FC; number of flares not in the class incorrectly predicted to be in the flare class), and the true null classifications (TN; number of flares not in the flare class and correctly predicted not to be in the flare class). These definitions are similar to those defined in forecasting solar energetic particle events and solar flares (recent examples include *Laurenza et al.* [2009] and *Bloomfield et al.* [2012]). Using these definitions,

$$\begin{aligned}
 \text{POD} &= \text{TC} / (\text{TC} + \text{FN}), \\
 \text{FAR} &= \text{FC} / (\text{TC} + \text{FC}), \\
 \text{HSS} &= \frac{2 \times [(\text{TC} \times \text{TN}) - (\text{FN} \times \text{FC})]}{(\text{TC} + \text{FN})(\text{FN} + \text{TN}) + (\text{TC} + \text{FC})(\text{FC} + \text{TN})}, \text{ and} \\
 \text{TSS} &= \frac{\text{TC}}{(\text{TC} + \text{FN})} - \frac{\text{FC}}{(\text{FC} + \text{TN})}.
 \end{aligned}$$

The POD values indicate that, as shown with the PC statistic, the nearest centroid model correctly predicts the X- and M-class flares better than the K-nearest neighbor model. However, the FAR shows that the K-nearest neighbor model makes fewer false predictions of a flare incorrectly being classified in the X or M class. For

Table 5. Skill Scores^a

Class	<i>N</i>	KNN Model				NC Model			
		POD	FAR	HSS	TSS	POD	FAR	HSS	TSS
X	196	0.59	0.28	0.65	0.59	0.96	0.75	0.39	0.94
M	2,964	0.67	0.22	0.70	0.65	0.71	0.66	0.40	0.60
C	23,425	0.92	0.10	0.78	0.77	0.73	0.14	0.53	0.55
B	12,806	0.89	0.11	0.84	0.84	0.80	0.19	0.71	0.71

^aThe number of flares (*N*), probability of detection (POD), false alarm rate (FAR), Heidke skill score (HSS), and true skill score (TSS) presented for the K-nearest neighbor (KNN) and nearest centroid (NC) models built with the solar cycle 22–24 flare parameters and used to classify the same sample of flares.

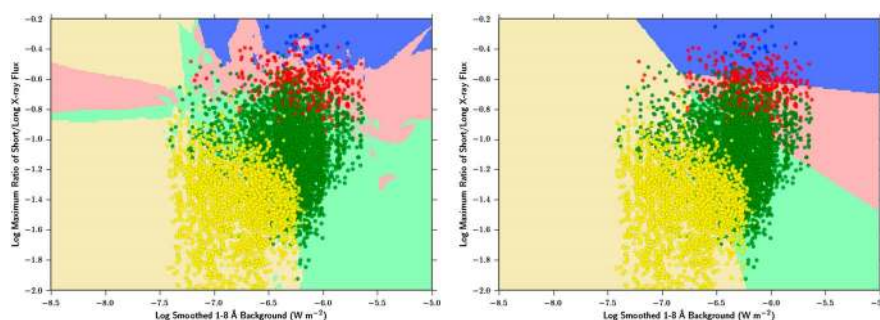


Figure 7. Results from classification models built with the solar cycle 22 and 23 flare parameters for the (left) K-nearest neighbor and (right) nearest centroid algorithms, applied to the solar cycle 24 data (points, color coded as in Figure 4). The background shading indicates the model-predicted class (color coded as in Figure 5). The K-nearest neighbor algorithm correctly classifies 83% of the flares. However, the nearest centroid algorithm classifies the highest peak flux flares (X) with 100% accuracy compared to the 74% accuracy from the K-nearest neighbor model.

the X-class flares, for instance, even though all of the X-class flares are correctly classified, the FAR is high with the nearest centroid model since there are 567 non-X-class flares incorrectly classified as being in the X class. From a forecast stand point, this argues that both a combination of the K-nearest neighbor predictions, which have low FAR, and the nearest neighbor predictions, which have high POD, would be necessary for making solar flare predictions. For the C- and B-class flares, the combination of low FAR and high POD for the K-nearest neighbor model proves it to be the superior model for predicting lower flux flares.

The final statistics, HSS and TSS, are commonly used skill scores with an advantage over the PC, POD, and FAR statistic in that they incorporate all of the parameters TC, FN, FC, and TN. By using all combinations, HSS takes into account the expected number of correct identifications due to chance. An advantage of TSS over HSS, as pointed out in *Bloomfield et al.* [2012], is that the TSS does not change depending on the number of flares in the sample size. The results of the TSS show the nearest centroid model as the best performer for the X-class flares, while the K-nearest neighbor is better for the weaker flares. The HSS roughly agree with TSS, with the exception of the X- and M-class flares in the nearest centroid model. The lower values of HSS are likely due to the smaller number of flares in these categories compared to the number of false predictions (for instance, there are 196 X-class flares, but 567 non-X-class flares were incorrectly predicted as X class) (Figure 7).

5. Discussion

Many of the previous studies of the properties of solar flares rely upon detailed analyses of a single or small group of flares. However, statistical analyses of large samples of flares offer new insights into the physical properties associated with these events. For instance, studies of the GOES X-ray light curves by *Aschwanden and Freeland* [2012] tested a theoretical model (fractal-diffusive self-organized criticality) for flare generation, finding that nanoflares are not likely to play a major role in flare heating, and *Ryan et al.* [2012] built upon previous studies to refine peak temperature and emission measure statistics. In this paper, we present a new way to separate solar flares into the NOAA flare classes based on a statistical analysis of the GOES X-ray observations of the ~50,000 flares occurring from 1986 to mid-2014.

These flare classification predictions are based upon observed X-ray properties—the 24 h nonflare X-ray background in the 1–8 Å band and the maximum ratio of the short- to long-band flux during the flare. These parameters reveal a separation between the X-, M-, C-, and B-class flares. The separation was quantified and verified through machine-learning algorithms and skill score statistics, applied to the solar flare parameters from solar cycles 22 to 24.

The R_{\max} parameter is related to the maximum temperature of the flare. Using the relations and constants from *White et al.* [2005], we find that the maximum temperatures of flares range from ~16 to 49 MK (X class), 6 to 15 MK (M class), 4 to 6 MK (C class), and ~ 4 to 11 MK (B class). These results are consistent with recent analyses, e.g., peak temperature from GOES presented in *Ryan et al.* [2012]. The maximum temperature is reached from a few minutes before to up to 25 minutes after the start of the flare. The stronger the flare, the more potential warning time we may have to predict the peak. For example, average R_{\max} for X flares occurs 6.7 min before the

peak while the time for C- and B-class flares is much shorter at ~ 3 min. One potential challenge to applying this technique for real-time flare forecasting is that the predictions are made after the maximum is reached. With XRS observations of 1 min resolution, as used in the current analysis, the short warning time is significantly decreased. In this case, however, the predictions are still useful in determining that the flare will be entering the declining phase. Also, NOAA SWPC currently provides finer time resolution XRS archival observations of a few seconds cadence. Accessing these data in real time would greatly enhance the warning time available from our technique.

While the maximum temperature is an indicator, like the peak flux of the flare, of the energy release, the nonflare background is less straightforward to interpret. This parameter is the integrated X-ray flux of the Earth-facing Sun, during the lowest flux period in 24 h preceding the flare. Integrated X-ray flux is dominated by active regions (e.g., *Acton* [1996] showed that more than 50% of the coronal luminosity is associated with 2% of the solar surface). The nonflare background is therefore a measure of the active regions, or areas of enhanced coronal heating. In past studies of X-ray imaging from SOHO and Yohkoh and full-disk magnetograms, *Fisher et al.* [1998] showed that X-ray luminosity is highly correlated with the active region's unsigned magnetic flux, with $L_x \sim \Phi_{\text{tot}}^{1.19}$. *Tan et al.* [2007] confirmed this in a study of 160 active regions and also found a strong correlation with the magnetic energy dissipation (also found earlier by *Abramenko et al.* [2006]), which they claim could be showing the importance of photospheric turbulent motions to heating of the corona above active regions. Therefore, the X-ray nonflare background flux is an indicator of the average magnetic energy of the active regions as well as the turbulent energy of the photosphere below these active regions. The phase separation between strong flares and weak flares is guided by the magnetic energy available to produce flares. Higher turbulence or stored magnetic energy leads to more energy release in the flare, measured by the maximum flare temperature.

Since the nonflare background is measured in the 24 h preceding the flare, it can be used for flare predictions in advance of the flare. One way the background can be used in real-time forecasting is as a threshold for predicting when strong flares may or may not occur. This is possible due to the large separation in the range of background values of strong flares (M and X class) versus the weakest flares (B class). From analysis of the distributions of the background for the different flare classes, we find that at the -2σ level (the 2.28th percentile) for M- and X-class flares, the background is $1.6 \times 10^{-7} \text{ W m}^{-2}$. Therefore, there is a low probability of a background flux below this level being associated with an X- or M-class flare. For C-, M-, and X-class flares, the -2σ level is $1.07 \times 10^{-7} \text{ W m}^{-2}$, meaning anything lower than this background flux is unlikely to be associated with a strong flare.

Based on these X-ray background results, during solar minimum when the background flux is low we expect no X- or M-class flares. For instance, the results from Figure 1 of *Winter and Balasubramaniam* [2014] show that solar cycle 24 had average 2 week X-ray background measurements below $1.07 \times 10^{-7} \text{ W m}^{-2}$ for the first 2 years of the cycle. This means that only B-class flares would have been expected in these years (2009–2011). From Figure 1, it is clear that relatively few strong flares had occurred during this time. During these 2 years, the NOAA list records 816 flares, including no X-class flares, 13 M-class flares (1.6% of flares), 103 C-class flares (12.6% of flares), and 700 B-class flares (85.8% of flares). Since this is a rough estimate of the background based on 2 week averages, we expect that the background occasionally rose above this low threshold, accounting for the small number of M-class flares observed during the last solar minimum. These flare rates during the beginning of solar cycle 24 are similar to those in the first 2 years of solar cycles 22 and 23 for X- and M-class flares, but, likely due to the higher backgrounds in cycles 22 and 23, there are more C-class and fewer B-class flares in cycles 22 and 23. For comparison, the beginning of solar cycle 22 had 1460 flares (2 X-class, 41 M-class, 410 C-class, and 1007 B-class flares from 1986 to 1988) and solar cycle 23 had 914 flares (4 X-class, 18 M-class, 262 C-class, and 630 B-class flares from 1996 to 1998).

Additional investigation into the relationship between temperature and nonflare X-ray background will lead to refinement of our flare predictions. Future work will investigate how the separation is affected by choice of the nonflare background. For instance, for forecasting purposes the goal is to have a longer lead time before the flare occurs. Different binning periods for nonflare background can be tested to determine how much lead time is possible while still preserving the phase separation between weak and strong flares. We will also look into increasing the warning time for flare class predictions, by determining whether other observables

help us determine what the R_{\max} /peak value will be further in advance. One such possible path is through determining characteristic flare shapes for the rise time, which can be used at the start of the flare to predict when the maximum will occur.

Acknowledgments

The GOES XRS data used in this paper are available at <http://www.ngdc.noaa.gov/stp/satellite/goes/dataaccess.html>. The NOAA X-ray Flare Lists were downloaded from NOAA NGDC through the FTP site linked from here: http://www.ngdc.noaa.gov/nndc/struts/results?t=102827*s=25*d=8,230,9. The monthly sunspot number was obtained from NASA Marshall Space Flight Centers compilation available here: http://solarscience.msfc.nasa.gov/greenwch/spot_num.txt. K.S.B. expresses gratitude to AFOSR for supporting the AFRL Task on Flares and CMEs. L.M.W. is supported by AFRL contract FA8718-05-C-0036. The authors thank Doug Biesecker (NOAA SWPC) for useful discussion during the 2014 SHINE conference that led to investigating the R_{\max} ratio, which became a focal point of this paper.

References

- Abramenko, V. I., A. A. Pevtsov, and P. Romano (2006), Coronal heating and photospheric turbulence parameters: Observational aspects, *Astrophys. J.*, *646*, L81–L84, doi:10.1086/506592.
- Acton, L. W. (1996), Coronal structures: Local and global, in *IAU Colloq. 153: Magnetodynamic Phenomena in the Solar Atmosphere—Prototypes of Stellar Magnetic Activity*, edited by Y. Uchida, T. Kosugi, and H. S. Hudson, pp. 3–11, Kluwer Acad., Dordrecht, Netherlands.
- Ahmed, O. W., R. Qahwaji, T. Colak, P. A. Higgins, P. T. Gallagher, and D. S. Bloomfield (2013), Solar flare prediction using advanced feature extraction, machine learning, and feature selection, *Sol. Phys.*, *283*, 157–175, doi:10.1007/s11207-011-9896-1.
- Aschwanden, M. J. (1994), Irradiance observations of the 1–8 Å solar soft X-ray flux from GOES, *Sol. Phys.*, *152*, 53–59, doi:10.1007/BF01473183.
- Aschwanden, M. J. (2011), The state of self-organized criticality of the Sun during the last three solar cycles: I. Observations, *Sol. Phys.*, *274*, 99–117, doi:10.1007/s11207-011-9755-0.
- Aschwanden, M. J., and S. L. Freeland (2012), Automated solar flare statistics in soft X-rays over 37 years of GOES observations: The invariance of self-organized criticality during three solar cycles, *Astrophys. J.*, *754*, 112, doi:10.1088/0004-637X/754/2/112.
- Balasubramaniam, K. (2013), Physics of solar flares and development of statistical and data driven models, *DTIC Tech. Rep. AFRL-RV-PS-TR-3013-0150*, Air Force Research Laboratory.
- Barnes, G., K. D. Leka, E. A. Schumer, and D. J. Della-Rose (2007), Probabilistic forecasting of solar flares from vector magnetogram data, *Space Weather*, *5*, S09002, doi:10.1029/2007SW000317.
- Bloomfield, D. S., P. A. Higgins, R. T. J. McAteer, and P. T. Gallagher (2012), Toward reliable benchmarking of solar flare forecasting methods, *Astrophys. J.*, *747*, L41, doi:10.1088/2041-8205/747/2/L41.
- Falconer, D., A. F. Barghouty, I. Khazanov, and R. Moore (2011), A tool for empirical forecasting of major flares, coronal mass ejections, and solar particle events from a proxy of active-region free magnetic energy, *Space Weather*, *9*, S04003, doi:10.1029/2009SW000537.
- Feldman, U., G. A. Doschek, W. E. Behring, and K. J. H. Phillips (1996), Electron temperature, emission measure, and X-ray flux in A2 to X2 X-ray class solar flares, *Astrophys. J.*, *460*, 1034–1041, doi:10.1086/177030.
- Fisher, G. H., D. W. Longcope, T. R. Metcalf, and A. A. Pevtsov (1998), Coronal heating in active regions as a function of global magnetic variables, *Astrophys. J.*, *508*, 885–898, doi:10.1086/306435.
- Gallagher, P. T., Y.-J. Moon, and H. Wang (2002), Active-region monitoring and flare forecasting—I. Data processing and first results, *Sol. Phys.*, *209*, 171–183, doi:10.1023/A:1020950221179.
- Garcia, H. A. (1994), Temperature and emission measure from GOES soft X-ray measurements, *Sol. Phys.*, *154*, 275–308, doi:10.1007/BF00681100.
- Hanssen, A., and W. Kuipers (1965), On the relationship between the frequency of rain and various meteorological parameters, *Meded. Verh.*, *81*, 2–15.
- Heidke, P. (1926), Berechnung der erfolges und der gute der windstarkevorhersagen im sturmwarnungsdienst, *Geogr. Ann.*, *8*, 301–349.
- Hock, R. A., D. Woodraska, and T. N. Woods (2013), Using SDO EVE data as a proxy for GOES XRS B 1–8 Å, *Space Weather*, *11*, 262–271, doi:10.1002/swe.20042.
- Kolmogorov, A. (1933), Sulla determinazione empirica di una legge di distribuzione, *G. Ist. Ital. Attuari*, *4*, 83–91.
- Laurenza, M., E. W. Cliver, J. Hewitt, M. Storini, A. G. Ling, C. C. Balch, and M. L. Kaiser (2009), A technique for short-term warning of solar energetic particle events based on flare location, flare size, and evidence of particle escape, *Space Weather*, *7*, S04008, doi:10.1029/2007SW000379.
- Levenberg, K. (1944), A method for the solution of certain non-linear problems in least squares, *Q. Appl. Math.*, *2*, 164–168.
- Lu, E. T. (1995), The statistical physics of solar active regions and the fundamental nature of solar flares, *Astrophys. J.*, *446*, L109, doi:10.1086/187942.
- Lu, E. T., and R. J. Hamilton (1991), Avalanches and the distribution of solar flares, *Astrophys. J.*, *380*, L89–L92, doi:10.1086/186180.
- Lu, E. T., R. J. Hamilton, J. M. McTiernan, and K. R. Bromund (1993), Solar flares and avalanches in driven dissipative systems, *Astrophys. J.*, *412*, 841–852, doi:10.1086/172966.
- Marquardt, D. (1963), An algorithm for least-squares estimation of nonlinear parameters, *SIAM J. Appl. Math.*, *11*(2), 431–441.
- Oliphant, T. (2007), Python for scientific computing, *Comput. Sci. Eng.*, *9*, 10–20.
- Pedregosa, F., et al. (2011), Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.*, *12*, 2825–2830.
- Ryan, D. F., R. O. Milligan, P. T. Gallagher, B. R. Dennis, A. K. Tolbert, R. A. Schwartz, and C. A. Young (2012), The thermal properties of solar flares over three solar cycles using GOES X-ray observations, *Astrophys. J. (Supp.)*, *202*, 11, doi:10.1088/0067-0049/202/2/11.
- Scott, D. W. (2009), *Multivariate Density Estimation: Theory, Practice, and Visualization*, vol. 383, John Wiley, New York.
- Smirnov, N. (1948), Table for estimating the goodness of fit of empirical distributions, *Annals Math. Stat.*, *19*(2), 279–281, doi:10.1214/aoms/1177730256.
- Tan, C., J. Jing, V. I. Abramenko, A. A. Pevtsov, H. Song, S.-H. Park, and H. Wang (2007), Statistical correlations between parameters of photospheric magnetic fields and coronal soft X-ray brightness, *Astrophys. J.*, *665*, 1460–1468, doi:10.1086/519304.
- Thomas, R. J., C. J. Crannell, and R. Starr (1985), Expressions to determine temperatures and emission measures for solar X-ray events from GOES measurements, *Sol. Phys.*, *95*, 323–329, doi:10.1007/BF00152409.
- Tobiska, W. K., and S. D. Bouwer (2006), New developments in {SOLAR2000} for space research and operations, *Adv. Space Res.*, *37*(2), 347–358, doi:10.1016/j.asr.2005.08.015.
- Veronig, A., M. Temmer, A. Hanslmeier, W. Otruba, and M. Messerotti (2002), Temporal aspects and frequency distributions of solar soft X-ray flares, *Astron. Astrophys.*, *382*, 1070–1080, doi:10.1051/0004-6361/20011694.
- Wagner, W. J. (1988), Observation of 1–8 Å solar X-ray variability during solar cycle 21, *Adv. Space Res.*, *8*, 67–76, doi:10.1016/0273-1177(88)90173-1.
- White, S. M., R. J. Thomas, and R. A. Schwartz (2005), Updated expressions for determining temperatures and emission measures from GOES soft X-ray measurements, *Sol. Phys.*, *227*, 231–248, doi:10.1007/s11207-005-2445-z.
- Winter, L. M., and K. S. Balasubramaniam (2014), Estimate of solar maximum using the 1–8 Å Geostationary Operational Environmental Satellites X-ray measurements, *Astrophys. J. Lett.*, *793*(2), L45.