Solar Phys (2011) 269: 111–127
DOI 10.1007/s11207-010-9689-y

An Analysis of the Sunspot Groups and Flares of Solar Cycle 23

Donald C. Norquist

Received: 7 December 2009 / Accepted: 25 November 2010 / Published online: 31 December 2010

Abstract Designing a statistical solar flare forecasting technique can benefit greatly from knowledge of the flare frequency of occurrence with respect to sunspot groups. This study analyzed sunspot groups and Hα and X-ray flares reported for the period 1997 – 2007. Annual catalogs were constructed, listing the days that numbered sunspot groups were observed (designated sunspot group-days, SSG-Ds) and for each day a record for each associated Hα flare of importance category one or greater and normal or bright brightness and for each X-ray flare of intensity C 5 or higher. The catalogs were then analyzed to produce frequency distributions of SSG-Ds by year, sunspot group class, likelihood of producing at least one flare overall and by sunspot group class, and frequency of occurrence of numbers of flares per day and flare intensity category. Only 3% of SSG-Ds produced a substantial Hα flare and 7% had a significant X-ray flare. We found that mature, complex sunspot groups were more likely than simple sunspot groups to produce a flare, but the latter were more prevalent than the former. More than half of the SSG-Ds with flares had a maximum intensity flare greater than the lowest category (C-class of intensity five and higher). The fact that certain sunspot group classes had flaring probabilities significantly higher than the combined probabilities of the intensity categories when all SSG-Ds were considered suggest that it might be best to first predict the flaring probability. For sunspot groups found likely to flare, a separate diagnosis of maximum flare intensity category appears feasible.

Keywords Flare probability · Solar flares · Sunspot groups · Sunspot group classes

1. Introduction

Emissions from solar flares pose a risk of serious damage to vulnerable ground, air and space assets. Flares originate in solar sunspot groups, also known as active regions. Flare prediction seeks to specify which sunspot groups will produce a flare with a lead time sufficient for adequate warning. Empirical/statistical techniques for flare forecasting have been developed in lieu of future physical flare models.

D.C. Norquist (✉)
AFRL/RVBXS, 29 Randolph Road, Hanscom AFB, MA 01731-3010, USA
e-mail: AFRL.RVB.PA@hanscom.af.mil

Current flare prediction schemes use sunspot group properties to estimate flare probability. There are two general types: categorical techniques and statistical diagnostic methods. Examples of the former include techniques reported by McIntosh (1990), Bornmann and Shaw (1994), Gallagher, Moon, and Wang (2002), Colak and Qawaji (2008), and Contarino *et al.* (2009). Categorical schemes employ past assessments of flaring rates *vs.* categories of sunspot group characteristics to estimate flare probability of current sunspot groups. Falconer (2001), Wheatland (2004, 2005), McAteer, Gallagher, and Ireland (2005), Cui *et al.* (2006), Georgoulis and Rust (2007), Barnes *et al.* (2007), Schrijver (2007), and Yu *et al.* (2009) describe statistical techniques. They compute a statistical analysis of the flare-related sunspot group properties and flare intensity. Resulting statistical relationships are applied to independent sunspot groups to estimate flare likelihood. To date, flare prediction methods have been at best competitive with the subjective methods used in space weather forecast operations.

This study attempts to document the basis for a statistical approach to flare prediction. Any reliable diagnostic scheme to predict the flare probability must be able to accurately reproduce the frequency distribution of the flares from previously observed sunspot groups. The flare distribution history should be known in order to properly design and prove a statistical prediction technique.

Several previous studies examined historical observations of sunspot groups and flares to ascertain their association. Kildahl (1980) analyzed the frequency of M- and X-class X-ray flares by sunspot group class over the period 1969 – 1976. Certain classes of sunspot groups were found to be more likely to flare than others. McIntosh (1990) reanalyzed the same data according to a 60-class classification system. He determined the frequency distribution of the sunspot group classes and flare occurrence frequency of the major class categories. Gallagher, Moon, and Wang (2002) calculated the average daily flare rate of C-, M-, and X-class flares for the period 1988 – 1996 in each of the McIntosh (1990) sunspot group classes.

In the current study we analyzed the sunspot groups and their associated flares of 1997 – 2007. This extends the studies of Kildahl (1980) and Gallagher, Moon, and Wang (2002) to the most recent solar cycle so as to determine if sunspot groups and flares have similar statistical characteristics among several cycles. If so, trends documented in solar cycle 23 will be generally representative of the typical solar cycle and would provide more justification to base the development of a statistical diagnostic technique for flare probability prediction on data from just that cycle.

2. Data

Observations of sunspot groups are archived in the National Geophysical Data Center (NGDC) web site in the "Sunspot region data (positions, area, *etc.*) from solar observatories" page: http://www.ngdc.noaa.gov/nndc/struts/form?t=102827\&s=40\&d=8,470,9. The telescopes from which data were drawn for this study were selected based on the availability of archived data for the years 1997 – 2007. They were: Kandilli, Turkey, Tashkent, Uzbekistan, United States Air Force (USAF) Solar Optical Observing Network sites (USAF_MWL), and Voroshilov, Russia. Data from all of these telescopes were available for 1998 – 2007, but only from Kandilli and the USAF sites in 1997. A description of the data is found in the "Data Documentation" hyperlink in the USAF_MWL subsection.

Ground-based solar telescopes also serve as the source for Hα flare observations. One of the conventional signatures of flare activity is a sudden localized brightening of the Hα (6563 Å) wavelength intensity as observed by telescope. The NGDC "H-Alpha Flares" web page http://www.ngdc.noaa.gov/nndc/struts/form?t=102827\&s=23\&d=8,210,9 served as the data base for this analysis. The data are taken from the "Events" subsection of the web page. In this subsection, the "Data Documentation" link provides a document describing the items included in each Hα flare report record. It also includes a document describing the Hα flare observations in more detail.

The other common flare activity indicator is a sudden spike in the X-ray flux as measured by sensors on geostationary satellites. Records of flare events measured by the X-ray sensor aboard the Geostationary Operational Environmental Satellite (GOES) were used as a source of X-ray flare observations. They were taken from the annual data files in the NGDC "Solar X-ray Flare" web page http://www.ngdc.noaa.gov/nndc/struts/results?t=102827&s=25&d=8,230,9. The document itemizing the contents of each X-ray flare report record can be accessed within each year's subsection under the GOES format hyperlink.

3. Analysis and Results

We constructed a sunspot group catalog for each calendar year using the NGDC archive data from the selected optical solar telescopes. The catalog has a record of the NOAA sunspot group number (subscripted sunspot group numbers were treated as separate sunspot groups) of each sunspot group observed by at least one telescope during its lifetime, along with the number of days the sunspot group was observed. Following each such record, all dates on which that sunspot group was observed are listed chronologically, with each one designated a "sunspot group-day" (SSG-D). Only days that a sunspot group was observed were included to ensure that its characteristics were known so that they might be associated with any observed flares.

A single observation represented the characteristics of the sunspot group on a given day in the catalog. The observation with the highest value of the quality flag was chosen, deemed the "best observation." If two or more observations had the highest quality flag, then the tie-breaker was the highest climatological flare probability of the reported sunspot group classification as determined by Gallagher, Moon, and Wang (2002). Sunspot classification (Zurich class, penumbra class, compactness class) as described by McIntosh (1990) was included in the catalog record for the best observation. A report of observation date, time, latitude, longitude and sunspot group number were required for the record to be included in the catalog.

Flare reports were then added to the catalog, assigned to a SSG-D according to their reported date and sunspot group number. First, selected characteristics of Hα flares were listed under the associated SSG-D in chronological order by flare starting time. This requires that the date and sunspot group number of the flare event be identified, which was virtually always the case in the NGDC Hα flare event data set. Because of our emphasis on consequential flare events, reports with an importance category of "subflare" or a brightness category of "faint" were excluded. Flare reports with an observational completeness designation of "S" (for only a fragmentary visual record) were skipped. Only Hα flares reported for SSG-Ds in the catalog were included, constituting 97.2% of all remaining reported flares.

In a similar manner, X-ray flare event records were also included in the sunspot group catalog. Here, only flare events of X-ray class C and intensity 5.0 or greater were included in the sunspot group catalog to exclude non-substantial flares from the analysis. Any X-ray flare records without a sunspot group number had it assigned from the Hα event with the closest starting time. Only X-ray flares for observed SSG-Ds were included in the catalog, comprising 95.8% of all reported substantial flares.

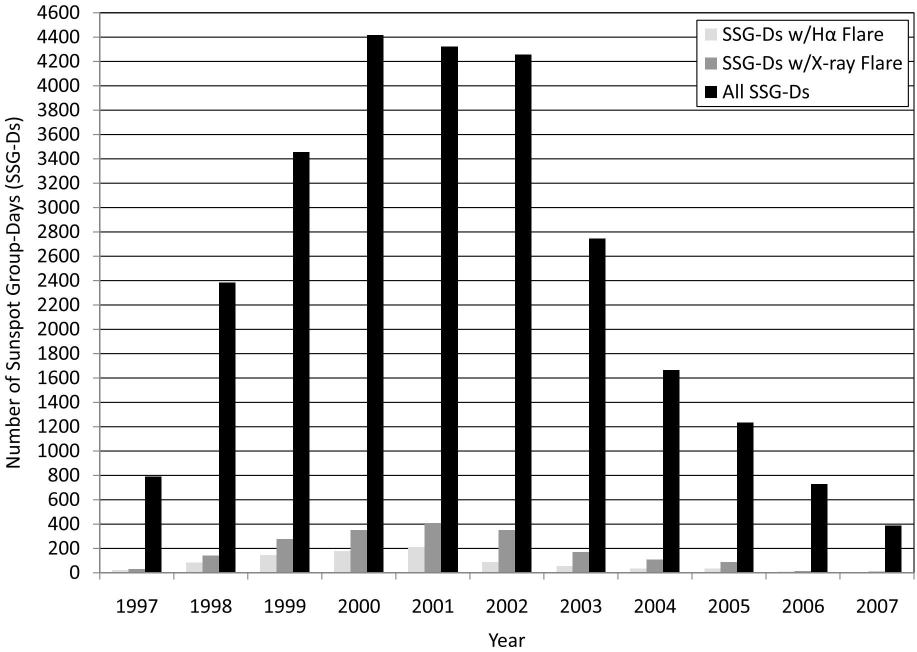


Figure 1 Total number of observed sunspot group-days (SSG-Ds) along with the number of SSG-Ds in which at least one Hα flare was observed, and the number of SSG-Ds having at least one observed X-ray flare. The year of Cycle 23 is indicated on the abscissa of the graph.

We then analyzed the sunspot group catalog with Hα and X-ray flare events for a given year to determine the frequency distribution of sunspot groups and flare events. We first tallied the total number of observed sunspot group-days (SSG-Ds) in each year. A bar graph of this analysis is shown in Figure 1. Also shown in the figure is the number of observed SSG-Ds having at least one Hα flare and having at least one X-ray flare that met the intensity threshold. There is a clear trend toward an increasing number of SSG-Ds with greater solar activity as depicted by sunspot number (see http://www.ngdc.noaa.gov/nndc/struts/results?t=102827\&s=5\&d=8,430,9, "Smoothed Plot Data"). The years 2000 – 2002 represent the maximum presence of sunspot groups during solar cycle 23. A small minority of the observed SSG-Ds produced an Hα and/or X-ray flare. The percentage of SSG-Ds with at least one Hα or X-ray flare is shown in Figure 2. Between 5% and 10% of the SSG-Ds had an X-ray flare in all but three of the years analyzed. The flare occurrence percentage does not follow the same trend with cycle year as does the SSG-D count. For example, the likelihood of a sunspot group producing a flare of either type was greater in 2005 than in 2003, even though 2003 had a much greater number of SSG-Ds. Nor are the trends for Hα and X-ray flare occurrence alike. In the Hα flare incidence there is a big drop from 2001 to 2002 and then a leveling off in the next two years, while X-ray flare occurrences show a less steep decline to 2003; then a rise. However, the three years with less than 5% X-ray flare production were the years with the lowest sunspot group counts.

As mentioned in Section 1, previous studies have shown an association between sunspot group class and probability of flaring. Next we look at the frequency of occurrence of SSG-Ds categorized by their sunspot group classification. We counted the SSG-Ds by sunspot

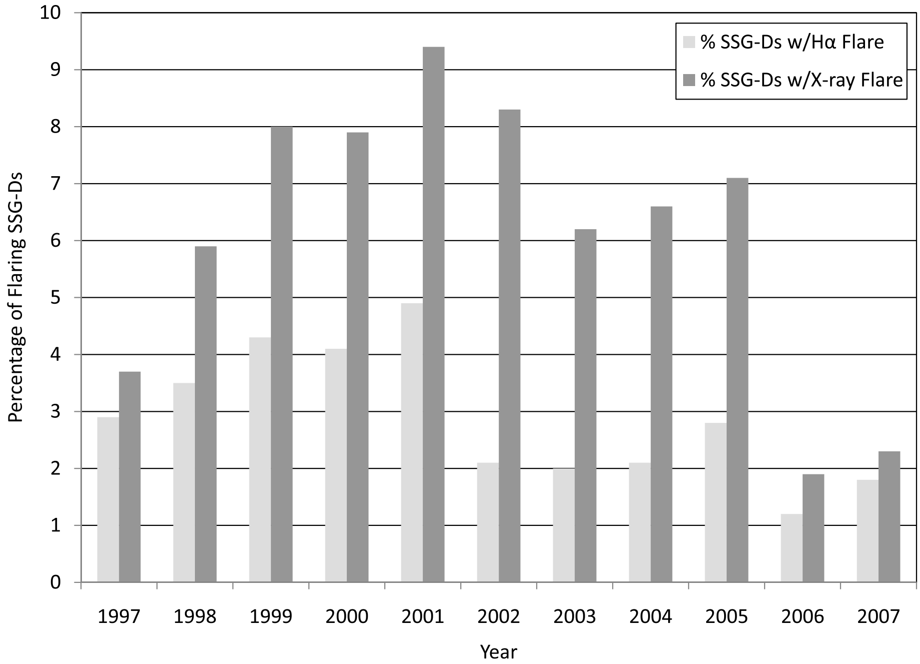


Figure 2 Percentages of SSG-Ds with at least one Hα flare and of SSG-Ds with at least one X-ray flare as computed from the flaring and total SSG-D counts in Figure 1.

group classes following the three-category, 60-class convention of McIntosh (1990). Three-category designation is *i*) Zurich, *ii*) penumbra, *iii*) compactness. Zurich categories are: A – unipolar, no penumbra; B – bipolar, no penumbra; C – bipolar, one end penumbra; D, E, F – bipolar of increasing size, penumbra both ends; H – unipolar with penumbra. Penumbra categories are: X – none; R – rudimentary; S – small symmetric; A – small asymmetric; H – large symmetric; K – large asymmetric. Compactness categories are: X – undefined for unipolar groups; O – open interior; I – intermediate interior; C – compact interior. Figure 3 shows counts by sunspot group class for each of the years of Cycle 23, ordered the number of SSG-Ds over all years. For all years and all classes (including those with unknown class), there were 26386 SSG-Ds.

The first major feature of Figure 3 is the dominance of the simplest sunspot categories. Just eight of 60 sunspot group classes, those that are unipolar or that have no penumbra or both (Zurich class A, B and H), contribute 47.6% of the sunspot groups of known classification. Of the 52.4% of the SSG-Ds of Zurich class C, D, E, and F, occurrences of the asymmetric penumbra categories A and K (32.8%) exceed the bipolar symmetric S and H (16.3%) and rudimentary R penumbra (3.3%) groups. In the compactness category, open interior O (36.4%) surpasses intermediate interior I (12.4%) and compact interior C (3.6%) Zurich class C – F groups. Unknown sunspot group classes in 1999 – 2001 represent 5 – 8% of the total observed SSG-Ds for their respective years. Knowledge of their classification would make only a minor change in the distribution.

Figure 4 compares this study's distribution of sunspot group classes (computed from the SSG-Ds of known classification) from 1997 – 2007 with the frequency of occurrence of SSG-Ds found by Kildahl (1980) for the period 1969 – 1976. We found that 70% of the

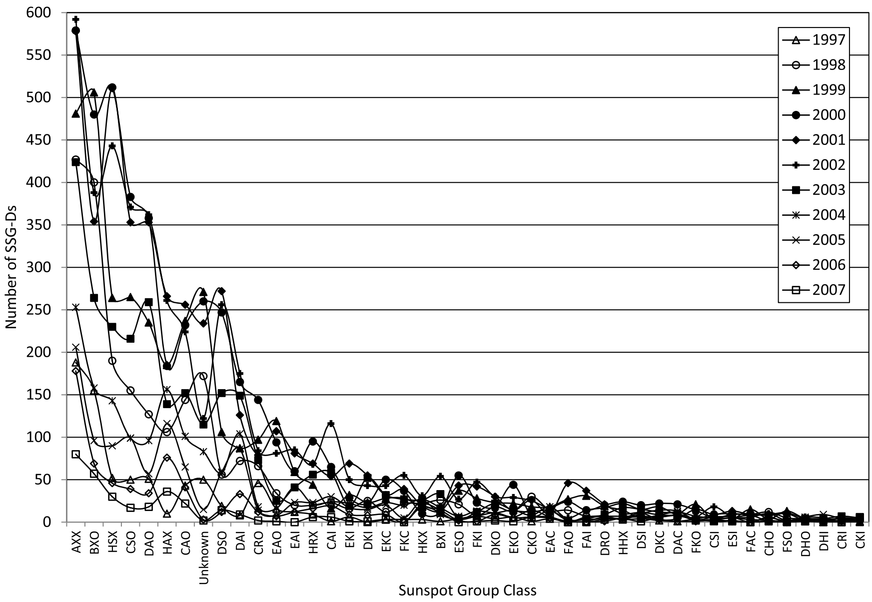


Figure 3 Number of SSG-Ds by 60-class sunspot group classification. The order of the classes left to right is by total number of SSG-Ds over all years. Only the classes with at least 0.1% of the total number of SSG-Ds with known classification are shown, along with the total for each year of unknown sunspot group class.

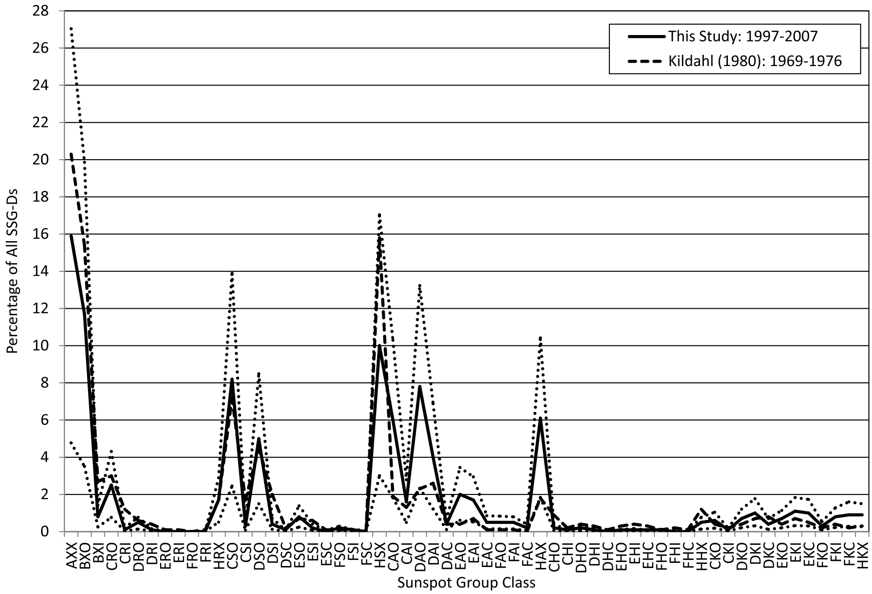


Figure 4 Frequency distribution (percent) of SSG-Ds by known McIntosh (1990) classification from the current study for the years 1997 – 2007 compared to the same found by Kildahl (1980) for the years 1969 – 1976. The dotted curves represent the upper and lower bounds of the values from the current study when accounting for the uncertainty in assigning the sunspot group class to the observed sunspot groups.

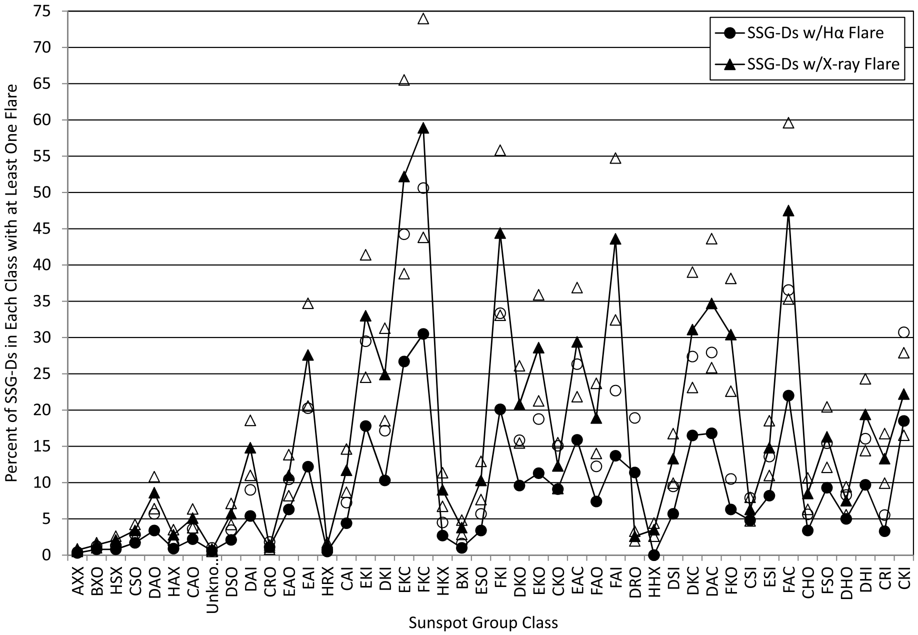


Figure 5 Percentage of SSG-Ds in each sunspot class having at least one Hα flare or at least one X-ray flare. The order of the classes left to right is by total number of SSG-Ds over all years as in Figure 3. Only the classes with at least 0.1% of the total number of SSG-Ds with known classification are shown, along with the percentage of flaring SSG-Ds of unknown class. The open symbols represent the range of uncertainty when accounting for the Hα flares (upper bounds only) and X-ray flares (both bounds) unassigned to sunspot groups in the flare reports.

SSG-Ds with two or more observations of quality fair or better had differing assigned classes of at least the quality of the observation cataloged. Figure 4 depicts the corresponding uncertainty in the frequency of occurrence for the current study. Even though the current study covers all 11 years of a solar cycle, while Kildahl (1980) used just eight years, the major features of the distributions are the same. Both found the major contributions to be from the simplest classes (Zurich A, B, and H). Kildahl (1980) found a higher percentage of classes AXX and HSX while this study had a greater proportion of Zurich D and E sunspot groups with small asymmetric penumbra and open or intermediate compactness, and of the HAX class.

Figure 5 shows the percentage of SSG-Ds in each class having at least one substantial flare following the ordering of sunspot group classes in Figure 3. About 26% of the reported X-ray flares were not assigned to a sunspot group, and the range plotted in Figure 5 for the SSG-Ds with at least one X-ray flare reflects this level of uncertainty in making the assignment. Similarly, there were 66% fewer Hα flares than X-ray flares reported overall, so Figure 5 shows the maximum that could be realized in percentage of SSG-Ds with at least one Hα flare when these are taken into account. The overall pattern of the distribution is generally the same for Hα and X-ray flare reports. The first seven sunspot classes shown are dominant in SSG-Ds (Figure 3) but have less than a 10% chance of producing a flare on a given day (Figure 5). Every sunspot group class with a greater than 20% likelihood of producing an X-ray flare is Zurich C through F with an asymmetric penumbra. Bipolar asymmetric classes shown in Figure 5 have an average of a 27% chance of producing an X-ray flare compared to a 11% probability for the symmetric classes shown. This along with

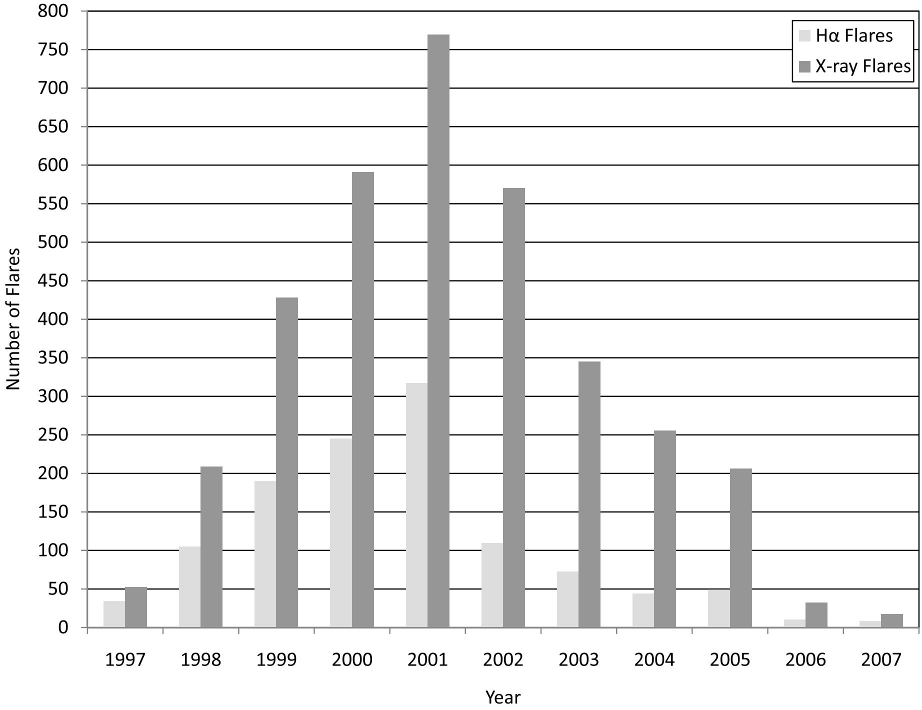


Figure 6 Number of flares reported for each year of solar cycle 23.

the more than double occurrence of asymmetric over symmetric penumbra bipolar SSG-Ds makes the asymmetric penumbra sunspot groups the dominant flare producers. Zurich class D through F with intermediate interior compactness have an average 25% chance of an X-ray flare *vs.* 42% for compact interior. The classes AXX, BXO, BXI, CRO, HSX, HAX, HRX, and HHX have a negligibly small probability of producing an Hα flare. Sunspot groups without penumbra (Zurich A, B) and unipolar groups with penumbra (Zurich H), regardless of the penumbra category, are the least likely to flare. These flare probability trends among the sunspot group classes agree with those of Gallagher, Moon, and Wang (2002) in every detail.

In Figure 6 we tally the individual flares occurring in each year of the analysis for both Hα and X-ray flares. Here again we see the inordinate drop in Hα flares from 2001 to 2002. The prevalence of X-ray flares follows the trend in sunspot number through the solar cycle for the most part. The major exception is that in 2000 and 2002 the number of X-ray flares are appreciably less than in 2001, while the number of SSG-Ds stay nearly the same. Figure 6 shows that substantial X-ray flares occurred at a rate of over two per day on average in 2001. Approximately a quarter of all substantial solar cycle 23 flares occurred in just that year. On the other extreme, only a total of 17 X-ray flares of intensity $\geq$ C 5 were reported in all of 2007.

Figure 7 shows percentages of all flares contributed by each sunspot group class for both Hα and X-ray flares. The uncertainty of the flare assignment to sunspot group is applied in the same manner as in Figure 5. Consistent with the discussion above about the sunspot groups most likely to flare, the asymmetric penumbra classes dominate over the symmetric penumbra classes in the percentage of total flares produced. The same peaks are apparent in Figure 7 as Figure 5, but the maxima on the left side of the chart are amplified while those

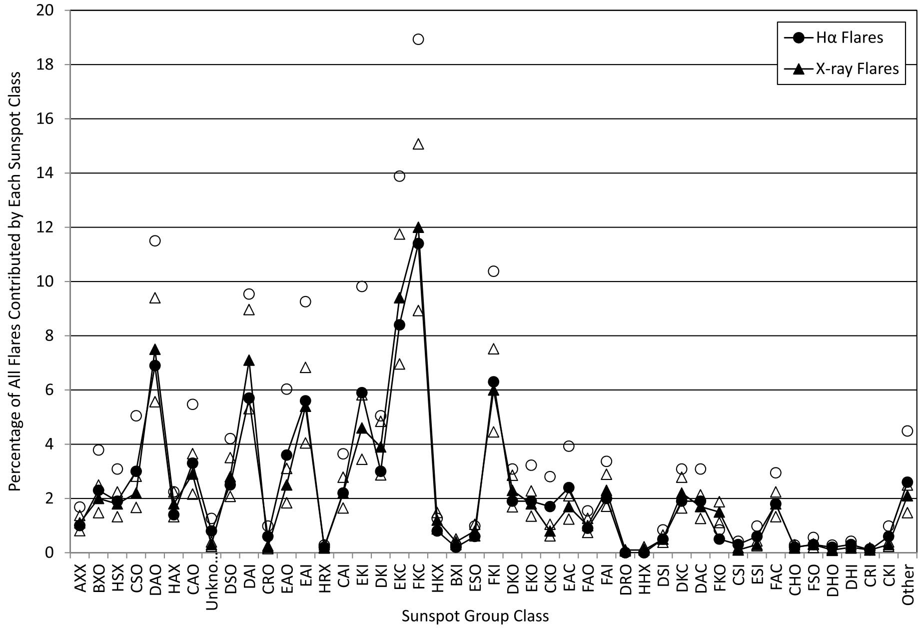


Figure 7 Percentage of all flares that originated from each sunspot group class. The order of the classes left to right is by total number of SSG-Ds over all years as in Figure 3. Only the classes with at least 0.1% of the total number of SSG-Ds with known classification are shown, along with the percentage flare contribution of SSG-Ds of unknown class. The class designated "Other" includes the percentage flare contributions from all sunspot group classes with less than 0.1% of SSG-Ds with known classification. Open symbols indicate uncertainty of the flare assignment to sunspot groups as in Figure 5.

on the right are reduced. This is because the classes on both charts are ordered in decreasing number of SSG-Ds, so that there are more sunspot groups to make flares in the classes on the left than for the classes on the right.

To investigate the flaring rate for those SSG-Ds that produce at least one flare, we divided the total number of flares each sunspot group class produced over the study period by the number of SSG-Ds with one or more flares. The results are plotted in Figure 8, where the sunspot group classes are ordered in decreasing X-ray flare rate. The uncertainty estimate for each class based on association of unassigned flares with sunspot groups is also shown. This flaring rate is different from the flaring rate of Kildahl (1980) and McIntosh (1990) because those studies included all SSG-Ds as the divisor. Figure 8 depicts the number of flares that were produced by a sunspot group of a given class averaged over all days on which it produced at least one flare. These average rates for flaring SSG-Ds would typically be higher than rates for all SSG-Ds, especially for sunspot classes with lower probability of flaring (and thus more days without a flare). An advantage of this type of analysis is that it shows which sunspot classes are most likely to produce one or more additional flares after having flared on a given day. While Figures 5 and 7 indicated that by percentage flaring likelihood and percentage flare contribution the sunspot classes had similar trends for both Hα and X-ray flares, this is clearly not the case when flaring rates are computed as shown in Figure 8. Note that in the ten classes with the highest X-ray flaring rates, nine have a large, asymmetric (K) penumbra. Contrarily, of the 12 classes with the lowest X-ray flaring rate ten have a rudimentary or small symmetric (R or S) penumbra.

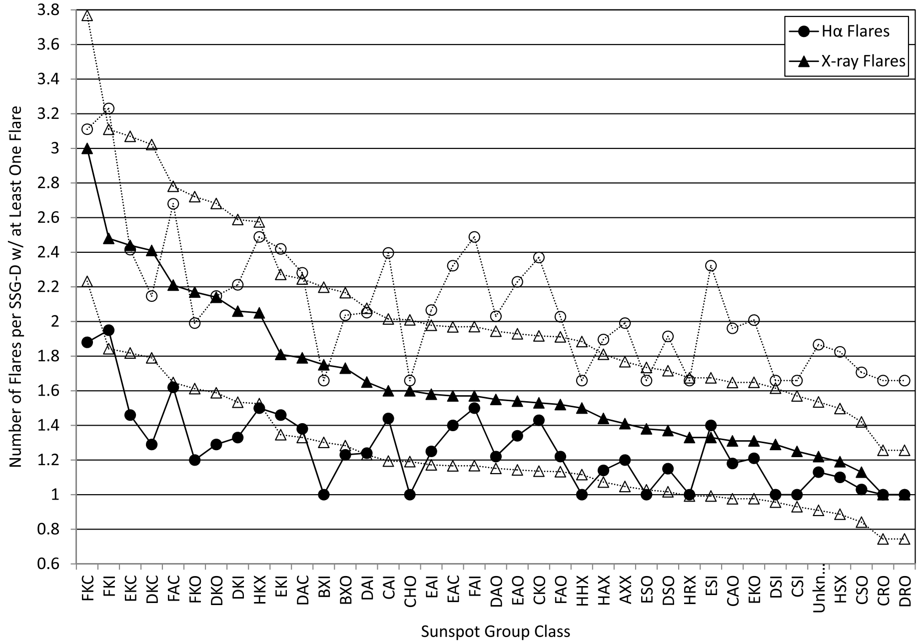


Figure 8 Number of flares divided by the number of SSG-Ds having at least one flare, displayed by sunspot group class for all classes with at least 0.1% of the total number of SSG-Ds with known classification, along with the ratio for the SSG-Ds with unknown class. The classes are ordered left to right in decreasing order of the ratio for X-ray flares. This ratio quantifies the flaring rate (flares per SSG-D of flaring SSG-Ds) by sunspot group class found in this study for the years 1997 – 2007. Open symbols indicate uncertainty of the flare assignment to sunspot groups as in Figure 5.

To allow for a comparison of the X-ray flaring rate of the sunspot classes with those reported by Kildahl (1980) for 1969 – 1976, the M- and X-class flares and the SSG-Ds were counted for the period 2000 – 2007. We divided the number of M- and X-class flares by the number of SSG-Ds for each sunspot group class to get a directly comparable flaring rate. The top five classes found in this study were: FKC, EKC, FAC, FKI, and FKO. The top five found by Kildahl (1980) were: EKC, FKI, FKC, EKI, and DKC. The lowest flare rates (for classes with at least 0.1% of the total number of SSG-Ds) belonged to DRO, HHX, CHO, CRI, and CRO in this study and to FAO, FAC, AXX, BXO, and HRX in Kildahl (1980). While there is good agreement between the two studies in the most active classes (large, asymmetric penumbra, compact interior), Kildahl (1980) identified two classes as inactive (FAO, FAC) that were either moderately or highly active in this study.

We also investigated the distribution of the number of observed sunspots per day, and how this affected the flaring rate. In Figure 9 we show the number of days in each year that fell within specified ranges of numbers of observed sunspots. There is a clear shift to a more common occurrence of higher numbers of visible sunspots with solar activity level. From 1997 to 2001, the peak number of days increases by one sunspot count category per year. From 2002 to 2005 it decreases by one count category per year. In the solar maximum years of 2000 – 2002, it was most common to see 10 – 12 sunspot groups per day. In the quiet years of 1997 and 2005 – 2007, only one – three sunspot groups were most often present.

We then totaled all flares on the days in these sunspot group count categories, and divided each category total by the number of sunspot groups in each category. The resulting flare

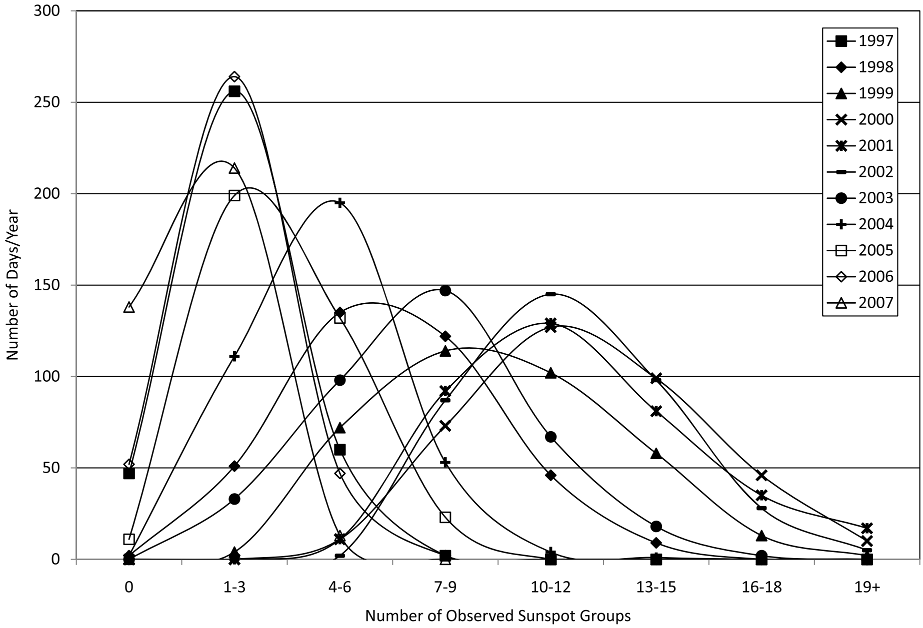
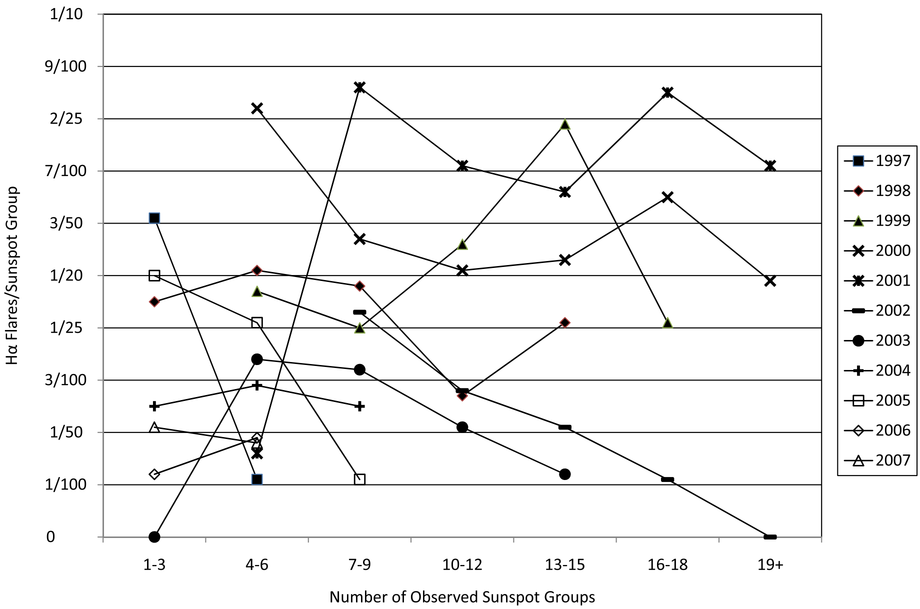


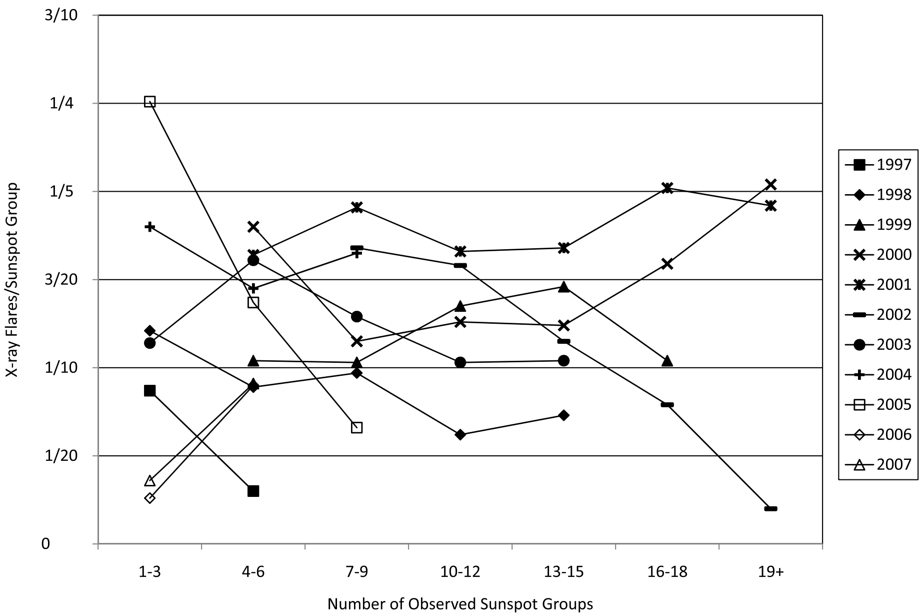
Figure 9 Number of days in each year that had the indicated number of observed sunspot groups. The smoothed curves are included to better highlight the frequency distribution for each year.

rates are shown in Figure 10. The values are shown only when a sunspot count category had at least 50 sunspot groups. There does not appear to be a clear association between flaring rate and number of visible sunspot groups per day. While some years show a persistent declining flaring rate with a greater density of visible sunspots (*e.g.*, 2002), others are mixed between increasing and decreasing rates. What does appear a significant trend is the flaring rate by year. There is a tendency for the years during or near solar maximum to have higher flaring rates and the years closer to solar minimum to have lower flaring rates for most of the count categories. In the Hα flaring rates (Figure 10a) we see the big drop from 2001 to 2002 that was evident in comparing the same years in Figures 1 and 2 (similar number of SSG-Ds but many fewer with at least one Hα flare). The rallying of X-ray flaring rate in 2004 and 2005 seen in Figure 10b is similarly apparent in comparing the two years with the three preceding years in Figures 1 and 2.

Are the increased flaring rates at solar maximum due to a greater proportion of the more flare-productive sunspot groups or to an across-the-board increase in flare productivity? It is difficult to compute the flaring rate for the less common sunspot group classes by year, since there are so few of them at solar minimum. However, we determined the percent change in the contribution of each class to the total number of SSG-Ds per year between the years 2001 and 1997 and the years 2006 and 2001. During the solar activity rise from 1997 and 2001, just the three simplest classes AXX, BXO and CRO accounted for a nearly 28% decrease in total sunspot contribution. Figure 5 indicates that these categories are among the least likely to produce even one flare on a given day. However, the biggest gaining classes from 1997 to 2001 were HAX, DSO, HSX and CSO, which according to Figure 5 are not much more flare-productive. What seemed to be boosting the flare rate the most from 1997 to 2001 is the modest increase in the number of the more flare-productive classes such as FKC, DKI,



(a)



(b)

Figure 10 Total number of (a) Hα flares and (b) X-ray flares divided by the total number of sunspot groups in each daily total sunspot group category. Values are shown only when a sunspot count category had at least 50 sunspot groups.

Table 1 Percentage of SSG-Ds in the period 1997 – 2007 with a specific number of flares.

	Zero	One	Two	Three	Four	Five+
% SSG-Ds w/specified number of Hα flares	96.69	2.49	0.59	0.15	0.05	0.03
% SSG-Ds w/specified number of X-ray flares	92.64	4.54	1.43	0.71	0.31	0.37

Table 2 Percentage of flaring SSG-Ds having a maximum intensity Hα flare of the indicated importance and brightness category.

	One	Two	Three	Four
% SSG-Ds w/max. Hα flare of 'normal' brightness	52.06	13.50	0.57	0
% SSG-Ds w/max. Hα flare of 'bright' brightness	16.71	13.50	3.32	0.34

Table 3 Percentage of flaring SSG-Ds with a maximum intensity X-ray flare of the indicated intensity category.

	C 5+	M 1 – 5	M > 5	X 1 – 5	X > 5	Unknown
% SSG-Ds w/max. X-ray flare of indicated intensity category	47.81	41.23	5.76	4.27	0.93	0.05

EKI and FAI from nearly 0 to $\frac{1}{2}$ to 1% of the SSG-Ds. There was a decrease of about the same size in these and in the more moderate flare producers FAI, DAO, and EAO from 2001 to 2006. So this relatively small change in the proportion of the more flare-active sunspot group classes and a large change in the simplest, least flare-active classes seems to be the biggest factor in the difference in flaring rates between solar maximum and minimum.

We next considered the frequency distribution of the number of flares in each of the SSG-Ds. Counting all SSG-Ds with zero, one, two, three, four, and five or more flares and then dividing each count by the total number of SSG-Ds gives a percent frequency of SSG-Ds with a specific number of flares. Results for 1997 – 2007 are shown in Table 1. Overall, just over 3% of SSG-Ds produced an Hα flare and slightly more than 7% had at least one X-ray flare. As can be seen in Figure 2, these percentages varied somewhat with year in the solar cycle. But overall, flares are an infrequent event relative to the prevalence of sunspot groups.

We next present the frequency distribution of flares by intensity category. Intensity categories of Hα flares are a combination of "importance" class, an indicator of the flare area, and "brightness" class, also called "brilliance": Normal brightness with importance = 1, 2, 3 or 4, then bright brightness with importance = 1, 2, 3, or 4. More information is given in the NOAA/NGDC web page for Hα flare observations mentioned in Section 2, in the "Data Documentation" link within the "flares. txt" document under the heading "Imp Opt." X-ray flares are categorized by X-ray class and intensity: C 5+, M 1 − 5, M > 5, X 1 − 5, X > 5. A description of these designations is also found in the same document under the heading "Imp X-ray." We counted the SSG-Ds with at least one flare by the above intensity categories according to their maximum intensity flare and computed the percentages for the intensity categories. The results are shown in Table 2 for Hα flares and Table 3 for X-ray flares.

Table 4 Percent occurrence frequency of all substantial Hα flares by importance class and brightness category.

	One	Two	Three	Four
% of Hα flares of 'normal' brightness	55.79	14.29	1.01	0
% of Hα flares of 'bright' brightness	15.22	10.90	2.45	0.34

Table 5 Percent occurrence frequency of all substantial X-ray flares by intensity category.

	C 5+	M 1 – 5	M > 5	X 1 – 5	X > 5
% of X-ray flares of indicated intensity category	57.34	35.40	3.89	2.85	0.52

About two-thirds of the 874 SSG-Ds with Hα flares displayed normal brightness in their most intense flare. In the 1943 X-ray flaring SSG-Ds, over half of them had their highest intensity in the categories greater than C class. The very intense flares are infrequent events (Hudson, 2007) even among all flares, which are in turn few in number compared to sunspot group appearances.

The frequency distribution of all flares by intensity category is shown for the entirety of cataloged flares: 1183 Hα flares in Table 4 and in Table 5 for 3474 X-ray flares. The frequency distribution of all Hα flares (Table 4) is very similar to the frequency distribution of the maximum flare intensity category of the SSG-Ds (Table 2). This is because SSG-Ds with at least one Hα flare were most likely to have just one (see Figure 8 and Table 1). The proportion of all flares of intensity M 1 – 5 or less (92.7%) is more than for maximum intensity flares (89%). Only 3.4% of all flares were in the X-class intensity category, but they occurred in all years of the study except 2007.

The value of Tables 2 and 3 is that they detail the frequency distribution of the strongest flare among intensity categories for those SSG-Ds that produced at least one flare. The discrete category frequency distribution allows a basis for evaluation of the frequency of the flare intensities predicted by a categorical statistical algorithm. In this way they provide some information about maximum flare intensity likelihood from sunspot groups about to produce a flare. Tables 4 and 5 give the full range of occurrence frequencies from all flares in the same intensity categories. For the purpose of comparison with past flare intensity distribution studies, we counted the 3,474 X-ray flares by bins of one unit of flare intensity (bin widths of 10^{-6}, 10^{-5} and 10^{-4} for C-, M- and X-class flares, respectively) and averaged the intensity in each bin. The best-fit power law of N flares for each bin of I average intensity was $N = 0.00109 I^{-1.09529}$. The power-law index of approximately 1.1, reflecting the slope of the decrease of log N with log I, compares with values from other studies in the range 1.5 – 2 (Wheatland, 2010). The more shallow slope in this study may be the result of omitting C-class flares of intensity < 5, which are generally included in other studies.

4. Discussion

The purpose of this study was to document the statistical characteristics of the sunspot groups and flares of solar cycle 23 to establish a basis for the design of a statistical flare prediction scheme. We conducted this study on the 11 years of solar cycle 23, 1997 – 2007. To the extent that the statistics from this solar cycle are similar to those of previous ones, a statistical diagnostic flare algorithm based on the most recent cycle should produce a flare frequency distribution representative of a typical solar cycle.

Recorded sunspot groups and Hα flares of importance/brightness of one/normal or greater and X-ray flares of intensity C 5 or higher were compiled in annual catalogs. Each sunspot group was marked by the date of its recorded appearance. This was designated a "sunspot group-day" (SSG-D). Each observed flare was assigned to its numbered sunspot group of origin according to its occurrence date. This allowed for a count of the number of both types of flares, each with recorded intensity. We analyzed the catalogs for each of the 11 years to examine the statistical characteristics of the sunspot groups and flares of solar cycle 23.

A total of 26386 SSG-Ds were observed. At least one Hα flare was observed on 874 SSG-Ds (3.3%), while at least one X-ray flare was determined to have occurred on 1943 SSG-Ds (7.4%). This latter figure is somewhat less than the approximately 20% found by McAteer, Gallagher, and Ireland (2005) when all C-, M- and X-class X-ray flares in the period 1996 – 2004. However, much of this difference can be explained by their use of all C-class flares *vs.* C 5 and higher in this study. Based on this study, an observed sunspot group on average demonstrated less than a one in ten chance of producing a significant flare on a particular day. In comparison to the occurrence of sunspot groups, the solar flare is an uncommon event. The frequency of appearance of sunspot groups on the visible solar disk had a maximum in 2000 and declined toward a minimum in 1997 and 2007. However, the likelihood of a flare occurring on a SSG-D does not similarly rise and fall with sunspot number. In the most active year for SSG-Ds, the likelihood of flaring was only about 4% greater than the seven other non-solar minimum years. This suggests that, at least for solar cycle 23, the cycle year is not expected to be a good predictor of the flare probability.

Analysis of the frequency distribution of the SSG-Ds by sunspot group class showed that the simplest stages made up nearly half of all SSG-Ds. These are the classes that have either unipolar sunspot groups or no penumbra or both. Of the remaining SSG-Ds that are bipolar with penumbra, those with asymmetric penumbra were twice as prevalent as symmetric penumbra groups, and 10 times more common than those with rudimentary (partial) penumbra. The interior of these sunspot groups were three times more likely to be open than intermediate and 10 times more likely open than compact. The major features of the distribution agreed with those determined from a prior solar cycle.

The likelihood of a flare from a SSG-D varied greatly by sunspot group, but the trends were similar for Hα and X-ray flares. This study found that the Zurich A, B, and H class sunspot groups were least likely to produce a flare. Of the Zurich C – F sunspot groups, those with asymmetric penumbra were $2\,\frac{1}{2}$ times more likely to produce a flare than groups with symmetric penumbra. The combination of more common and more likely to flare makes the asymmetric penumbra classes more flare-productive than symmetric classes. Zurich D – F sunspot groups with compact interiors were almost twice as likely to flare as those with intermediate interiors. These trends agree with at least one study of a earlier solar cycle and confirm that the structural properties of the sunspot group can act as a useful predictor for determining the flaring likelihood of a sunspot group.

The total frequency of occurrence of all recorded flares was also analyzed. This information is useful in gauging how common flaring is from any source on the solar disk. When broken down by cycle year, we found that flare count followed the sunspot number trend, with a maximum number of over two substantial X-ray flares/day in 2001 and less than one X-ray flare in every seven days in 1997, 2006 and 2007. The prevalence of all flares should not be confused with the likelihood of at least one flare on a SSG-D, which, as Figure 2 shows, does not reveal a similar decline with solar activity.

When the number of flares from each sunspot class was divided by the number of SSG-Ds with at least one flare, a measure of the likelihood of additional flares per day is derived. While probability of one or more Hα and X-ray flares varies similarly among the sunspot group classes, this is not true for the multiple flare rate measure. For X-ray flares, the sunspot groups with large, asymmetric penumbra had the highest probabilities of multiple flares per day, while those with small symmetric or rudimentary penumbra were least likely to produce more than one flare. These results may aid in attempting to predict the probability of multiple flaring in a statistical flare-forecast algorithm.

Our analysis showed that the number of sunspots visible per day on the solar disk increases with solar activity. And to a lesser degree, flaring rate is somewhat more elevated in the more active periods of the solar cycle. However, no definite relationship between the number of observed sunspots and flaring rate was detected. This suggests that the number of sunspot groups on the Sun at any one time may not in itself be a reliable indicator that any given sunspot group is more or less likely to produce a flare.

The frequency of occurrence of the 60 sunspot group classes found in this study of solar cycle 23 agree very well with those of Kildahl (1980) based on solar observations from 1969 – 1976. The primary difference was unipolar and small bipolar sunspots with penumbra: Kildahl found more with symmetric penumbra while this study had more with asymmetric penumbra. Trends in flaring probability by sunspot group class found here are in good agreement with those published by Gallagher, Moon, and Wang (2002) from 1989 – 1996 data. Basic flaring rates for X-ray flares (total flares/total SSG-Ds) for M- and X-class X-ray flares for the analogous period of solar cycle 23 indicated highest and lowest rate sunspot group classes similar to those of Kildahl (1980). The fact the major characteristics of sunspot groups and flares of solar cycle 23 are comparable to those of the earlier solar cycles suggest that they persist from cycle to cycle. This fact supports the use of data from a single cycle to design a generally representative flare-forecast scheme.

This study found that, during solar cycle 23, substantial Hα and X-ray flares occurred on just over 3% and 7% of SSG-Ds, respectively. Considering all SSG-Ds, this fact suggests an extremely low likelihood of a sunspot group producing a flare on a given day. This makes the prediction of high probability of flaring by a sunspot group without regard to its properties difficult to achieve. A prediction of "no flare" would almost always be correct. A similar conclusion was reached by Barnes and Leka (2008). But it is the times when the flares do occur that are of most importance. We saw that flaring incidence of certain sunspot group classes was much higher than these overall occurrence rates. A successful flaring probability tool would have to take into account these associations between sunspot group properties and flaring likelihood. The fact that certain classes have flare occurrence probabilities considerably greater than the likelihood of a particular flare intensity category (since the combined probability of all categories amounts to just 7% when all SSG-Ds are considered) suggests the potential benefit of initially predicting probability of the flare occurrence for each SSG-D. Our statistics showed that, once just the SSG-Ds that did flare are considered, more than 50% of X-ray flaring SSG-Ds produced their strongest flare with an intensity greater than C class. This gives a subsequent prediction of the maximum flare intensity category a chance to succeed, because the occurrence likelihood is more evenly distributed among the categories. Thus, this study's results suggest that a two-stage statistical flare probability scheme applied to each sunspot group-day is worthy of investigation.

Acknowledgements The National Geophysical Data Center web site, http://www.ngdc.noaa.gov/nndc/struts/results?t=102827&s=1&d=8,2,9, was the source for all of the data used in this study. I appreciate the helpful comments of my colleagues in the AFRL Solar and Solar Wind Disturbance Prediction Section. This work was funded by the 6.2 applied research program of the Air Force Research Laboratory.

References

Barnes, G., Leka, K.D., Schumer, E.A., Della-Rose, D.J.: 2007, *Space Weather* **5**, S09002.
Barnes, G., Leka, K.D.: 2008, *Astrophys. J. Lett.* **668**, 107.
Bornmann, P.L., Shaw, D.: 1994, *Solar Phys.* **150**, 127.
Colak, T., Qawaji, R.: 2008, *Solar Phys.* **248**, 277.
Contarino, L., Zuccarello, F., Romano, P., Spadaro, D., Guglielmeno, S., Battiato, V.: 2009, *Acta Geophys.* **57**, 52.
Cui, Y., Li, R., Zhang, L., He, Y., Wang, H.: 2006, *Solar Phys.* **237**, 45.
Falconer, D.A.: 2001, *J. Geophys. Res.* **106**, 25185.
Gallagher, P.T., Moon, Y.-J., Wang, H.: 2002, *Solar Phys.* **209**, 171.
Georgoulis, M.K., Rust, D.M.: 2007, *Astrophys. J. Lett.* **661**, 109.
Hudson, H.S.: 2007, *Astrophys. J. Lett.* **663**, 45.
Kildahl, K.: 1980, In: Donnelly, R.F. (ed.) *Solar-Terrestrial Predictions Proceedings* **3**, U.S. Dept. Commerce, Boulder, C-166.
McAteer, R.T.J., Gallagher, P.T., Ireland, J.: 2005, *Astrophys. J. Lett.* **631**, 628.
McIntosh, P.S.: 1990, *Solar Phys.* **125**, 251.
Schrijver, C.J.: 2007, *Astrophys. J. Lett.* **655**, 117.
Wheatland, M.S.: 2004, *Astrophys. J. Lett.* **609**, 1134.
Wheatland, M.S.: 2005, *Space Weather* **3**, S07003.
Wheatland, M.S.: 2010, *Astrophys. J.* **710**, 1324.
Yu, D., Huang, X., Wang, H., Cui, Y.: 2009, *Solar Phys.* **255**, 91.