RATES OF FLARING IN INDIVIDUAL ACTIVE REGIONS

M. S. WHEATLAND
Department of Physics, Macquarie University, NSW 2109, Australia
(e-mail: wheat@physics.mq.edu.au)

(Received 29 January 2001; accepted 2 July 2001)

Abstract. Rates of flaring in individual active regions on the Sun during the period 1981 – 1999 are examined using United States Air Force/Mount Wilson (USAF/MWL) active-region observations together with the Geostationary Operational Environmental Satellite (GOES) soft X-ray flare catalog. Of the flares in the catalog above C1 class, 61.5% are identified with an active region. Evidence is presented for obscuration, i.e. that the increase in soft X-ray flux during a large flare decreases the likelihood of detection of soft X-ray events immediately following the large flare. This effect means that many events are missing from the GOES catalog. It is estimated that in the absence of obscuration the number of flares above C1 class would be higher by (75 ± 23) %. A second observational selection effect – an increased tendency for larger flares to be identified with an active region – is also identified. The distributions of numbers of flares produced by individual active regions and of mean flaring rate among active regions are shown to be approximately exponential, although there are excess numbers of active regions with low flare numbers and low flaring rates. A Bayesian procedure is used to analyze the time history of the flaring rate in the individual active regions. A substantial number of active regions appear to exhibit variation in flaring rate during their transit of the solar disk. Examples are shown of regions with and without rate variation, illustrating the different distributions of times between events (waiting-time distributions) that are observed. A piecewise constant Poisson process is found to provide a good model for the observed waiting-time distributions. Finally, applications of analysis of the rate of flaring to understanding the flare mechanism and to flare prediction are discussed.

1. Introduction

The rate of occurrence of flares in individual active regions may provide clues to the flare mechanism, and to the process(es) of energy storage in active regions. The flaring rate presumably reflects the amount of free magnetic energy available in an active region and the rate of supply of energy to the region (e.g., Litvinenko and Wheatland, 2001), as well as the presence of magnetic field configurations suitable for flare occurrence.

A number of previous studies have focused on the relationship between flare rates and details of photospheric field measurements for statistical samples of active regions. Mayfield and Lawrence (1985) found that flare occurrence is correlated with the photospheric magnetic flux of an active region. Recently Sammis, Tang, and Zirin (2000) investigated the relationship between magnetic complexity of active regions and the production of large flares. They found that almost all

Solar Physics **203:** 87–106, 2001.

substantial flares occur in regions classified $\beta\gamma\delta$, which indicates bipolar spots with mixed polarities, including a penumbra enclosing umbrae of both polarities.

A different approach to studying rates of flaring is to consider the waiting-time distribution (WTD), or distribution of times between flares, for individual active regions. The reciprocal of the average waiting-time gives the mean rate of flaring for the active region. If flares occur with a constant probability per unit time, then flaring represents a Poisson process and the WTD is an exponential:

$$P(\Delta t) = \lambda e^{-\lambda \Delta t}, \tag{1}$$

where λ is the probability of flare occurrence per unit time (the flaring rate), and Δt denotes a waiting time. The assumption that flaring occurs with a constant probability per unit time implies that flares are independent of one another. If the rate of flaring varies with time (but flares are still independent), and if the rate variation can be well modelled by piecewise constant rates $\lambda_1, \lambda_2, \ldots$, then the WTD is approximately given by

$$P(\Delta t) = \sum_i \varphi_i \lambda_i e^{-\lambda_i \Delta t}, \tag{2}$$

where φ_i is the fraction of events corresponding to the rate λ_i.

Most studies of flare WTDs have involved times between flares from all flare-producing active regions present on the Sun during the observing period (Biesecker, 1994; Pearce, Rowe, and Yeung, 1993; Wheatland, Sturrock, and McTiernan, 1998; Boffeta *et al.*, 1999; Wheatland, 2000). This approach is appropriate if the Sun is considered as a flaring system, but it has the disadvantage that observed rate variations may include intrinsic variation within individual active regions (e.g., due to changing physical conditions in those regions) as well as variation due to the changing number of active regions on the disk.

Recently Boffeta *et al.* (1999) pointed out that, based on 20 years of soft X-ray flare observations, the flare WTD from all active regions on the Sun follows a power law for long waiting times (greater than a few hours). Based on this evidence, Boffeta *et al.* argued for a turbulence model to account for flare statistics. However, Wheatland (2000) showed that the observed WTD is qualitatively consistent with Equation (2), using rates λ_i estimated from the data with a Bayesian procedure. The observed time distribution of rates (fraction of time spent with a given flaring rate) was found to approximately follow an exponential. This suggests that there is no intrinsic significance to the appearance of a power law in the WTD, but that it arises because of the variation in the rate of flaring observed from all active regions on the Sun. The exact origin of this variation remains to be determined.

A few authors have examined flare WTDs in individual active regions. Crosby (1996) calculated times Δt between hard X-ray events in individual active regions, and then produced a distribution including values of Δt for a number of active regions. The resulting WTD was characterized as a power law with an exponential

rollover for large waiting times. Boffeta *et al.* (1999) also followed this procedure in their study of soft X-ray events and found that the resulting distribution was similar to the WTD they obtained for flares from all active regions present on the Sun. Recently Moon *et al.* (2001) constructed WTDs for individual active regions based on soft X-ray flare observations. They found that six very flare productive active regions observed in the years 1989 to 1991 had WTDs consistent with a simple exponential, implying that the rate of flaring was constant in these regions.

In this paper the rate of soft X-ray flare occurrence in individual active regions is examined for active regions observed on the Sun during the period 1981 to 1999. The goal of this paper is to characterize the observed flaring rates and rate variations, and to examine the waiting-time distributions in individual regions. The presentation is divided as follows. Section 2.1 describes the catalogs used in this study and the selection procedures involved. In Section 2.2 an observational effect that causes events to be missing from the flare catalog used is discussed. The distribution of average flaring rates among all of the active regions is determined in Section 2.3 (to our knowledge this is the first time this result has appeared in the literature). In Section 2.4 flaring rates in all of the moderately flare-producing active regions are examined in detail. A Bayesian procedure is used to determine whether the rate of flaring in each region is better represented by a constant rate Poisson or a multiple rate Poisson model. The observed waiting-time distributions for the active regions are then tested against the appropriate Poisson model. In Section 3 conclusions and directions for further work are presented.

2. Data and Analysis

2.1. Details of the data

The data used consists of the United States Air Force/Mount Wilson Observatory (USAF/MWL) catalog of active-region observations for the years 1981 to 1999, and the Geosynchronous Operational Environmental Satellite (GOES) soft X-ray flare catalog for the same years. The data are maintained by the National Geophysical Data Center of the National Oceanic and Atmospheric Administration (NOAA)[1]. Active regions are labelled with NOAA/USAF numbers, and a total of 4944 regions are listed in the USAF/MWL catalog. The flare catalog for the same years includes 40 240 flares.

It is important to consider the event selection procedure used at NGDC/NOAA to create the GOES catalog. The start of a flare is defined by four consecutive one-minute soft X-ray fluxes[2] in the GOES 1 to 8 Å band that meet the following criteria: (a) all four values are above the B1 threshold (flux greater than

[1]The data are available on the web at *ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA*.

[2]The one-minute flux is the average soft X-ray flux recorded by the GOES instrument during a minute.

10^{-7} W m^{-2}), (b) all four values are strictly increasing, and (c) the last value is greater than 1.4 times the value three minutes earlier. The time of maximum indicates when the measured flux is a maximum, and the maximum flux value is the measured maximum in the 1–8 Å channel (no background subtraction is done). Note that all times are given to the nearest minute. The end time of an event is defined as the time when the flux reading returns to one half of the maximum value. Flares are classified according to peak flux $\mathcal{F}$ into classes with the letters B ($\mathcal{F} < 10^{-6}$ W m^{-2}), C ($10^{-6} < \mathcal{F} < 10^{-5}$ W m^{-2}), M ($10^{-5} < \mathcal{F} < 10^{-4}$ W m^{-2}) and X ($\mathcal{F} > 10^{-4}$ Wm^{-2}). In addition a number is added to the letter as a suffix to denote the multiplier of the power of ten indicated by the letter, so that, e.g., M3 indicates a flare with peak flux around 3×10^{-5} W m^{-2}.

In the analysis described below the set of flares is restricted to those with a soft X-ray classification greater than C1, which leaves 27 887 events. This restriction attempts to remove artifactual rate variations due to changes in the detection threshold for flares produced by a time-varying soft X-ray background. The C1 value is chosen because this flux is typical of background levels for the active Sun (Hudson, 1991).

Finally, not all flares in the catalog are identified with active regions. The GOES soft X-ray detectors have no spatial resolution, and the matching of events with active regions in the GOES catalog is based on observations of contemporaneous optical flares. These observations are made by ground based observatories around the world. Many events in the GOES catalog have no active region number because the flare did not produce an optical event, because the optical event was not observed, or because the flare occurred away from active regions. Also, it is possible that there are some misidentifications of flares and active regions in the catalog. In total 17 138 flares of greater than C1 class are identified with active regions, i.e., 61.5% of the total number. Larger flares are more likely to be identified with active regions. Of the C class flares 58% have active region numbers, whilst of the M and X class flares the fractions are 82% and 94%, respectively. This result means that there is a peak-flux dependent selection effect in the data, which should be kept in mind in assessing the results of this paper.

2.2. Evidence for obscuration

Before looking at the rate of flaring in individual active regions, it is important to consider how observational effects might influence the flaring rate determined from the dataset. As mentioned above, flares of C1 and lower peak flux are omitted in an attempt to account for secular variations in the background flux, which are expected to alter the threshold for flare detection. However, this step does not compensate for the increase in background brought about by individual large flares. Soft X-ray flare time histories typically involve a rapid rise to maximum followed by a slow, approximately exponential decline in soft X-ray flux (Feldman, 1996). There is clear evidence in the data that the increased flux during the declining phase of

large flares makes the identification of subsequent flares less likely. We will refer to this effect as obscuration.

The GOES event selection procedure imposes a constraint on whether flares that occur in the wake of large flares are detected. For an event to be detected there must be a monotonic increase in four one-minute flux values $\mathcal{F}_i$ ($i = 1, 2, 3, 4$) such that $\mathcal{F}_4 > 1.4\mathcal{F}_1$. Suppose that during the three-minute observing interval the background flux declines by an amount $\Delta\mathcal{F}_{\rm dec}$ due to a decrease in emission from a recent large flare. If the increase in flux associated with a second flare during the three minute interval is $\Delta\mathcal{F}_{\rm fl}$, then for the second flare to be detected requires $\mathcal{F}_4 = \mathcal{F}_1 + \Delta\mathcal{F}_{\rm fl} - \Delta\mathcal{F}_{\rm dec} > 1.4\mathcal{F}_1$. Hence detection requires $\Delta\mathcal{F}_{\rm fl} > \Delta\mathcal{F}_{\rm dec} + 0.4\mathcal{F}_1$. If the first flare is large then the decline in flux during three minutes can be neglected, and we have the condition $\Delta\mathcal{F}_{\rm fl} > 0.4\mathcal{F}_1$. If the second flare is sufficiently small it will not satisfy this condition and will not be detected. Assuming that the second flare occurs very soon after the first flare and that the second flare attains its peak flux during the three minute observing period, the condition for detection becomes very simple: the second flare must have a peak flux greater than about 40% of the peak flux of the first event. For example, if the first flare is X1 class, then a second flare that attains its maximum flux shortly after the first flare must be above M4 class to be detected. Although this argument is simplistic, it indicates that even quite large flares may escape detection.

To see the effect of obscuration, consider the sample of GOES flares with a peak flux greater than M5 (5×10^{-5} W m^{-2}). For these flares label the interval from the time of maximum flux to the time of maximum flux of the previous flare (the second flare can have any size) $\Delta t_{\rm bef}$. Similarly, label the interval to the subsequent flare (this flare can also have any size) $\Delta t_{\rm aft}$. Figure 1 shows a plot of $\Delta t_{\rm bef}$ versus $\Delta t_{\rm aft}$ for the 583 flares greater than M5 class in the GOES catalog. The dashed line follows $t_{\rm bef} = t_{\rm aft}$. It is clear from the figure that $t_{\rm aft}$ tends to be larger than $t_{\rm bef}$: 421 points fall below the dashed line and only 162 fall above the line. The average of $t_{\rm aft}$ is 5.3 hours, and the average of $t_{\rm bef}$ is 3 hours (since the averages are strongly influenced by the largest values it is perhaps better to calculate the median values, which are 3.6 hours and 1.7 hours, respectively).

Of course, Figure 1 could in principle illustrate a real physical effect, namely that the occurrence of a large flare reduces the likelihood of a subsequent flare. However, this is unlikely, for the following reasons. First, a catalog of eight years of hard X-ray flares does not reveal the same effect (Wheatland, Sturrock, and McTiernan, 1998). Hard X-ray emission in flares is much more impulsive than soft X-ray emission, and so the problem of obscuration is reduced with hard X-ray events. Second, the values of $t_{\rm aft}$ tend to be systematically larger than the time from the maximum of the large flare to the end of the large flare as recorded in the GOES catalog (we will refer to this as the decay time, $t_{\rm dec}$). Figure 2 illustrates this effect. For all of the GOES flares $t_{\rm aft}$ is plotted versus $t_{\rm dec}$. The solid line indicates $t_{\rm aft} = t_{\rm dec}$, and the dashed line is a linear fit to all of the points. There is a clear tendency for the time to the next flare to increase with the decay time of the first

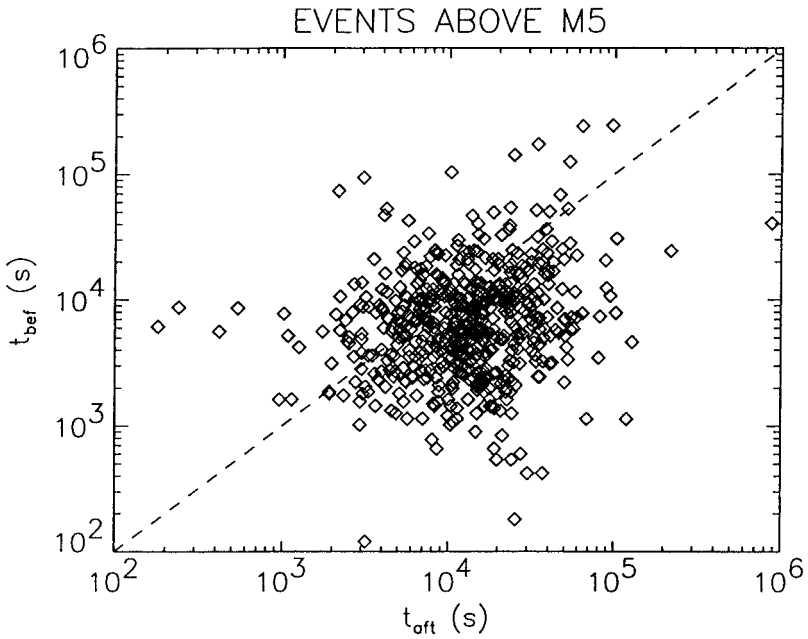

Figure 1. Time to the previous flare versus time to the next flare, for middle flares with peak flux above M5.

flare (the dashed line has a positive slope) and also $t_{\rm aft}$ tends to be larger than the decay time (the points tend to lie above the solid line).

Figure 2 is consistent with the hypothesis that the increase in soft X-ray background produced by large flares obscures flares that occur during the declining phase of the large flare. Many flares are expected to be missing from the GOES catalog as a result. It is possible to estimate the number of missing flares using Figure 2, as follows. Consider the points in Figure 2 corresponding to small values of $t_{\rm dec}$, say $t_{\rm dec} < t_{\rm small}$, where $t_{\rm small}$ is a suitable small value. These points should provide an unbiased sample of the values of $t_{\rm aft}$, because they are not strongly affected by obscuration. Denoting these values $t_{\rm aft}(t_{\rm dec} < t_{\rm small})$, we note that the average of these values, $\langle t_{\rm aft}(t_{\rm dec} < t_{\rm small})\rangle$ provides an estimate of the mean time between flares, in the absence of obscuration. If the total observing time for the GOES data is T, then the expected number of flares, based on this mean time between flares is

$$N_{\rm exp} = \frac{T}{\langle t_{\rm aft}(t_{\rm dec} < t_{\rm small})\rangle}. \tag{3}$$

To illustrate the use of Equation (3) we take $t_{\rm small} = 200$ s, which by reference to Figure 2 is a value large enough to include a reasonable number of points (3700), but sufficiently small that the effects of obscuration should be limited. For this value of $t_{\rm small}$ we evaluate $N_{\rm exp}$ using the bootstrap method (Press *et al.*, 1992) with 2000 resamplings to determine the mean and standard deviation of $t_{\rm aft}(t_{\rm dec} < t_{\rm small})$.

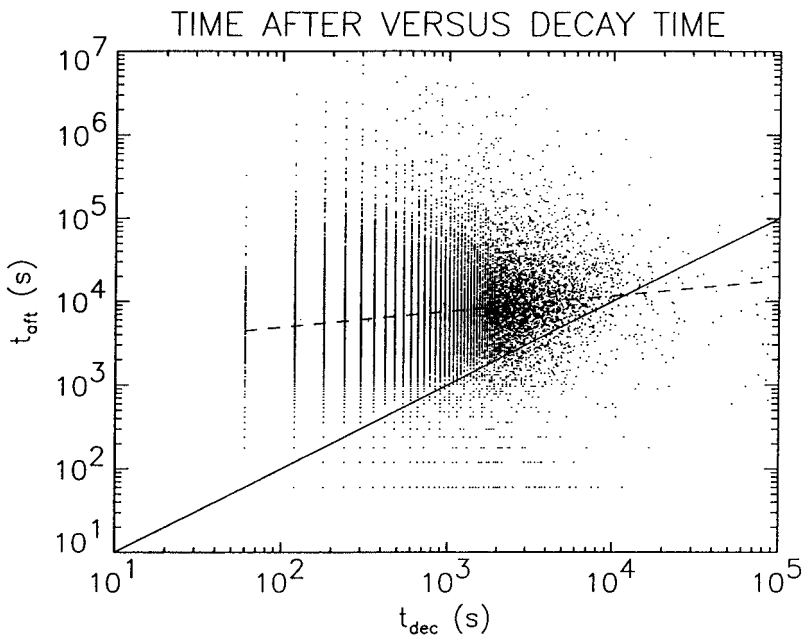

Figure 2. Time from the maximum of one flare to the maximum of the next flare versus the time from the maximum of the first flare to the end time of the first flare.

The total observed number of points in Figure 2 is $N_{obs} = 27608$ (this is less than the total number of GOES flares above C1 class because some of the flares are missing end times). The expected number of flares, based on the points with $t_{dec} < 200$ s is $N_{exp} = 42000 \pm 4000$. On this basis we would conclude that in the absence of obscuration the number of flares above C1 class would be higher by (52 ± 14) %.

Since the choice of t_{small} is arbitrary, we have used the following procedure to arrive at an improved estimate for N_{exp}. The steps outlined above were repeated for t_{small} equal to 120 s, 180 s, 240 s, 300 s and 360 sec. The resulting plot of N_{exp} versus t_{small} was extrapolated to $t_{small} = 0$ s. The result was $N_{exp} = 48000 \pm 6000$, i.e., the expected number of flares is higher than the observed number by (75 ± 23) %. Although we take this as our best estimate, it is worth remembering that it involves extrapolation from results using relatively small numbers of points in Figure 2.

One consequence of flares being missed due to obscuration is that the flares in the catalog are not strictly independent of one another, and hence the Poisson model for occurrence (e.g., Equations (1) and (2)) is not strictly valid. However, for most purposes the Poisson model appears to be a reasonable approximation, as demonstrated below.

2.3. RATE DISTRIBUTIONS

For each active region listed in the USAF/MWL catalog, the number n of GOES flares (of greater than C1 class) is counted. The upper panel of Figure 3 shows the cumulative distribution of flare numbers, $N_{\rm AR}(\geq n)$, i.e., the number of active regions producing at least n flares. The most flare producing active region observed during 1981–1999 (AR 6368) produced 110 flares larger than C1 class[3]. The majority of the active regions listed in the USAF/MWL catalog (3004) have no flares above C1 listed in the GOES catalog. The plot is log-linear so the distribution $N_{\rm AR}(\geq n)$ is approximately exponential, although there is a clear excess of active regions with low flare numbers. The diamond symbol on the vertical axis indicates the value $N_{\rm AR}(\geq 0) = 4944$. There are two biases in the flare selection process that contribute to the large numbers of active regions with few flares. First, many flares are missing from the catalog because of obscuration, as discussed in Section 2.2. Second, flares that are listed in the catalog but were not attributed to their active region (as discussed in Section 2.1) are not represented in Figure 3. Both of these effects act to increase the lower end of the distribution in the top panel of Figure 3.

An interesting question is whether, if the true number distribution of flares among active regions is exponential, the observational biases would act to skew the observed distribution away from an exponential, and produce the excess seen in Figure 3. It seems likely that the answer to this question is no, as follows. Assume that a given active region produces N_0 flares above peak flux $\mathcal{F}_0$. The flare frequency-peak flux distribution from all active regions on the Sun is observed to follow a power law with index $\gamma \approx 1.8$ (Hudson, 1991). Assuming that the same distribution holds in individual active regions, the expected frequency-peak flux distribution from the observed active region is

$$\mathcal{N}(\mathcal{F}) = \lambda_0(\gamma - 1)\mathcal{F}_0^{\gamma-1}\mathcal{F}^{-\gamma}, \tag{4}$$

where $\lambda_0 = N_0/T$ is the observed flaring rate and T is the (fixed) period of observation. The quantity $\mathcal{N}(\mathcal{F})d\mathcal{F}$ is the expected rate of occurrence of flares with peak flux in the range $\mathcal{F}$ to $\mathcal{F} + {\rm d}\mathcal{F}$. Observational effects causing flares to be missed may be represented by a probability $P_{\rm det}(\mathcal{F})$ for detection of flares of a given peak flux. The observed number of flares above peak flux $\mathcal{F}_0$ is then expected to be $N_{\rm obs} = T\int_{\mathcal{F}_0}^{\infty}\mathcal{N}(\mathcal{F})P_{\rm det}(\mathcal{F})\ {\rm d}\mathcal{F}$, and it follows that $N_{\rm obs} = aN_0$, where the constant of proportionality a is independent of the active region. Assuming N_0 is distributed exponentially over the observed active regions, $N_{\rm obs}$ will also be exponentially distributed, with an exponent steeper by the factor a. This argument relies on the probability for detection of flares depending only on the size of the event, and on all active regions following the same power-law frequency-peak flux

[3]It should be noted that long-lived active regions may last several solar rotations, and at each reappearance on the disk they are assigned a new NOAA/USAF number. We consider each active region with a different number as being distinct, so that the time histories describe flaring during one transit of the disk.

distribution. The first assumption appears reasonable, but is difficult to test based on the data. Regarding the second assumption, Kucera *et al.* (1997) have presented evidence that small active regions depart from the global power-law distribution. This effect is attributed to an upper limit to the available free energy for flaring in small regions. However, the study by Kucera *et al.* found that this was a small effect. Hence it seems unlikely that the large departure from an exponential apparent in the upper panel of Figure 3 is due solely to this effect. The conclusion is then that the true distribution of flare number does not follow an exponential, but exhibits an excess of active regions producing small numbers of flares. This provides suggestive evidence for two kinds of active regions: flare producing, and non-flare producing active regions. However, it is difficult to make definitive statements about the true distribution of flare number among active regions, given the observational selection effects influencing the observed distribution.

The flare numbers for each active region can be turned into mean flaring rates $\lambda = n/T$ if the period T of observation of the active region is known. For this to be a physically meaningful rate, the flare production rate must be stationary (see Section 2.4). We take $T = t_f - t_i$ where t_i is the time of the first USAF/MWL observation of the active region, or (if it is earlier) is the time of the first GOES flare from the region. Similarly t_f is taken to be the time of the last USAF/MWL observation, or (if it is later) is the time of the last flare. There are often GOES events listed before (after) the first (last) USAF/MWL observation, presumably because the active region observations are ground-based, and so are more likely to suffer from data gaps than the GOES observations.

The lower panel in Figure 3 shows the cumulative distribution of flare rates, $N_{AR}(\geq \lambda)$, i.e., the number of active regions with a mean rate of at least λ. This distribution is also approximately exponential, although once again there is an excess of active regions with low flaring rates. The remarks made above concerning the contribution of missing flares to the observed distribution of flare number apply also to the observed distribution of flaring rates.

2.4. RATES IN INDIVIDUAL ACTIVE REGIONS

Next we consider the question of whether the rate of flaring in individual active regions is stationary, or whether it varies with time. The data used are the start times for GOES events for active regions that produced at least 10 flares, together with the earliest and latest times of observation of the active regions, based on both the GOES and USAF/MWL catalogs. The rates of flare occurrence in these active regions are examined using the Bayesian Blocks procedure (Scargle, 1998). This approach was previously applied to flare data in Wheatland, Sturrock, and McTiernan (1998) and Wheatland (2000). The Bayesian procedure determines whether the observed time history of flaring is more likely to be produced by a constant rate Poisson process, or by a piecewise constant Poisson process with two rates, where the change in rate occurs at the time of one of the events. If the dual rate model is

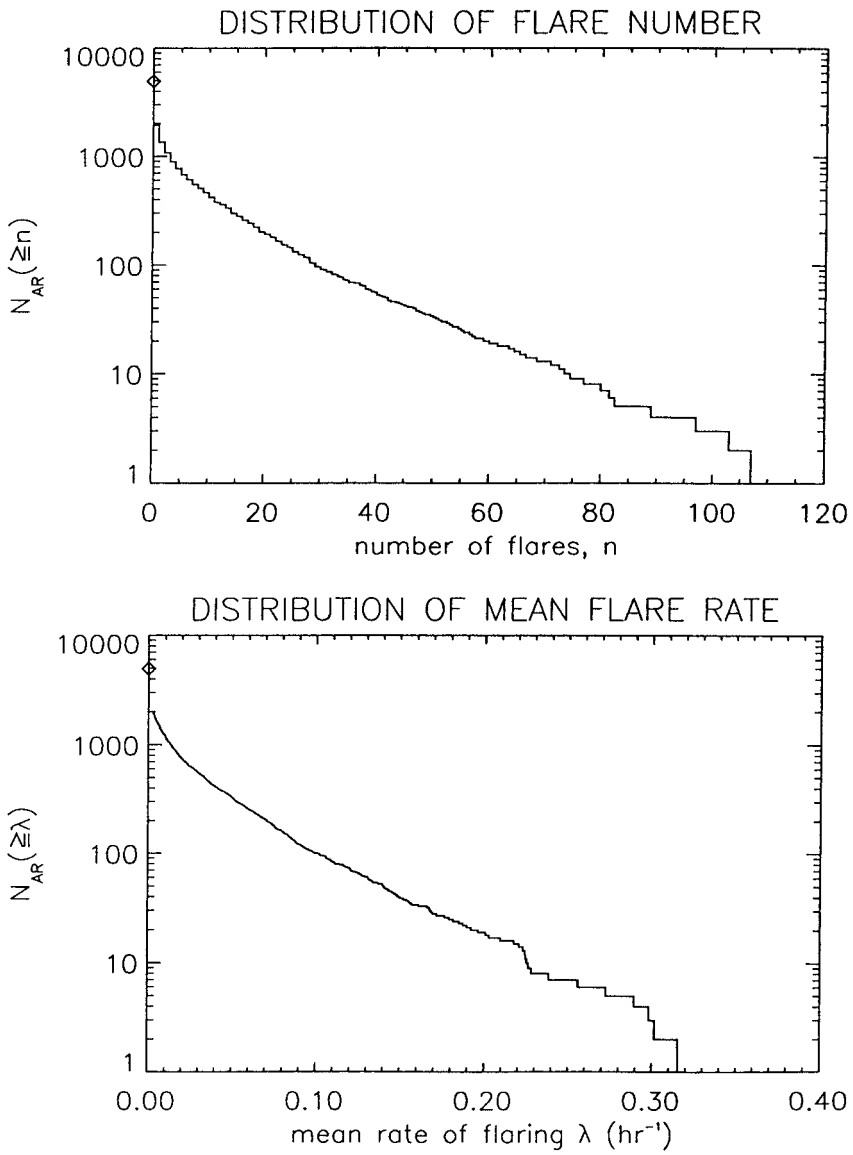

Figure 3. Cumulative distributions of flare number (*top panel*) and of mean flaring rate (*bottom panel*) among active regions.

more likely, then the two sets of data during the periods of time with different rates are each subject to the same decision making process: are they better modelled by a constant rate Poisson process, or by a dual rate process? This procedure is continued iteratively until a final decomposition into piecewise intervals of constant rate ('Bayesian Blocks') is achieved. The only free parameter in the method is the ratio of probability for the dual rate model over the single rate model that must be

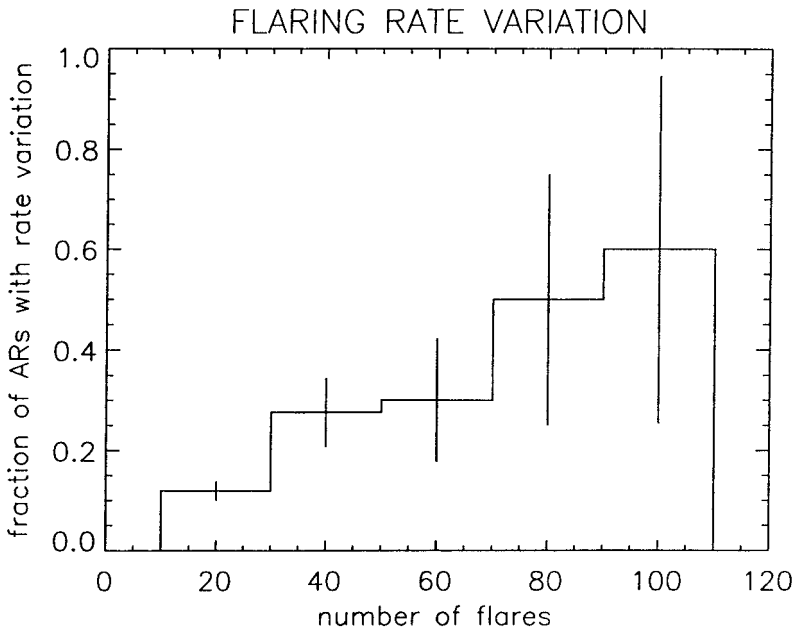

Figure 4. Fraction of GOES active regions exhibiting variation in the rate of flare occurrence, according to the Bayesian Blocks procedure.

met for the dual rate model to be accepted. This ratio, called the prior odds, is taken to be two.

Figure 4 summarizes the results, showing a histogram of the fraction of active regions that were determined to exhibit rate variation as a function of the number of flares produced by the active regions. The error bars correspond to the square root of the number of active regions in each bin. The figure shows that significant fractions of active regions appear to exhibit rate variation. The incidence of detection of rate variation increases with the number of events. Of the 13 active regions with more than 70 events, seven were determined to have a time-varying rate.

To understand the results in Figure 4 it is helpful to know how reliable the Bayesian Blocks procedure is at detecting rate variations for the relatively small numbers of events in the individual active regions. Figure 5 summarizes a series of simulations designed to test the Bayesian procedure. A series of N events was generated over 14 days from a two rate Poisson process with rates λ and $r\lambda$, where the change in rate was taken to be equally likely to occur at any time during the time series. The Bayesian Blocks procedure was applied to the resulting time history of simulated flaring, to determine whether the rate change could be detected. This procedure was repeated 1000 times for choices of N in the range 10–100, and for the values $r = 1, 2, 3, 5$. Figure 5 shows the fraction of rate-change detections among each set of 1000 simulations, as a function of N. The curves show the results for $r = 1$ (solid), $r = 2$ (dotted), $r = 3$ (dashed) and $r = 5$ (dot-

dashed). A number of points may be noted from the simulations. First, the solid curve shows that the Bayesian procedure is fallible, and makes a small number of false detections of rate variation. The fraction of false detections increases with N, and for $N = 100$ is around 10%. Hence, we conclude that the analysis of the GOES active regions summarized in Figure 4 is likely to contain some false rate-change detections. However, as shown in Figure 4 the fraction of active regions exhibiting rate variation is significantly larger than 10% for all active regions with more than 30 flares. This result indicates that the Bayesian procedure is detecting real rate variation among the GOES active regions. Another point to note from the simulations in Figure 5 for the cases $r = 2, 3, 5$ is that the Bayesian procedure often fails to detect genuine rate variation, especially for low numbers of events and for low values of r. As expected, the procedure is more successful for larger values of r. Still, for $r = 5$ the procedure only detects the rate variation about 60% of the time for $N = 100$ events. (Part of the reason this fraction is low is that if the rate change occurs near the beginning or end of the time series, then there are few events at a different rate, and it is intrinsically hard to detect the variation.) Comparing these results with Figure 4 we conclude that there are probably many instances of rate variation among the GOES active regions that are not being detected.

The simulations described above demonstrate that the observed incidence of rate variation greatly exceeds that expected if there is no intrinsic variation, and also provides an indication of how often real variation is missed by the Bayesian procedure in specific cases. It does not provide a detailed model of the response of the Bayesian procedure to the observed data, because the precise nature of the rate variation in the observed data is unknown. For example, for many active regions the Bayesian procedure chose three Bayesian blocks to represent the observed flaring rate, suggesting that active region flaring rates may vary repeatedly during a transit of the disk. The simulations involve a single rate change with a fixed ratio of rates.

Next we examine the statistics of flare occurrence in a number of individual active regions. As an example of a very flare-productive active region with an apparently stationary rate, consider AR 4474. Figure 6 shows an analysis of the rate of flare occurrence in this active region. The top panel in Figure 6 shows a schematic of the 99 flares above C1 class produced by this active region: each flare is represented by a spike at its start time, with height equal to the peak flux of the event. The vertical dashed lines in the top panel show the earliest and latest times of observation of the active region in the USAF/MWL catalog. The middle panel shows the Bayesian Blocks analysis of the observed flares. The flaring rate is determined by the procedure to have the constant value $\lambda = 0.28$ hr^{-1} (above C1) over the 14.7 days of flare observations. In the lower panel of the figure the waiting-time distribution for the active region is shown as a histogram in log-log space. The error bars on the WTD correspond to the square root of the number of waiting times in each bin. The solid curve in the bottom panel represents the exponential Poisson model for the WTD, Equation (1). The Poisson WTD is seen to well represent the observed WTD for AR 4474. A Kolmogorov–Smirnov

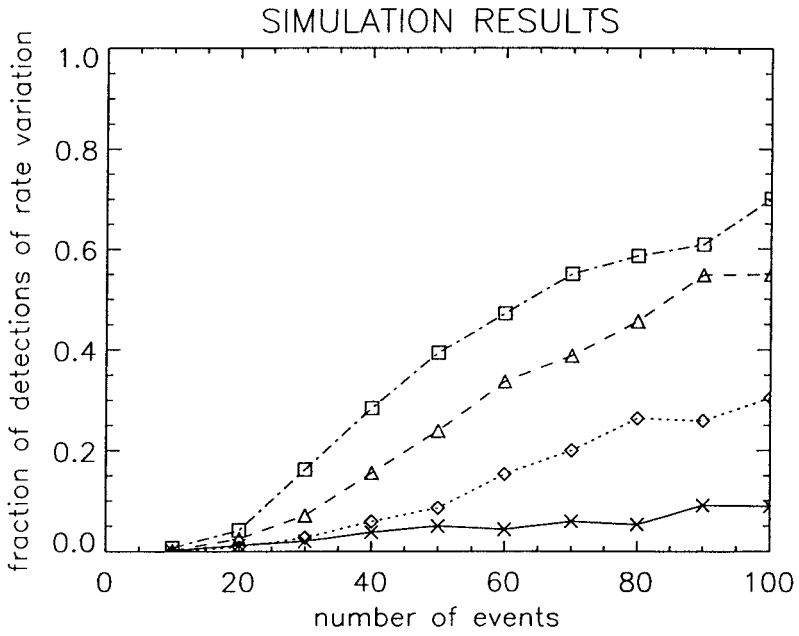

Figure 5. Result of application of Bayesian Blocks procedure to simulated flares produced by a dual rate Poisson process. The fraction of times the rate variation is detected is shown as a function of the number of events for ratios of rates of unity (*solid curve*), two (*dotted*), three (*dashed*) and five (*dot-dashed*).

(K−S) test confirms that the model and observed distributions are compatible (the significance of the difference between the distributions is 0.92). Hence, we see that the Poisson model is adequate to describe the observed WTD, despite the problem of flare inter-dependence due to obscuration discussed in Section 2.2. There is some evidence in the WTD for a lack of short waiting times by comparison with the model. For example, there are four waiting times less than 0.3 hours, whereas the model predicts around eight. This effect is not particularly significant for just this active region, as indicated by the K–S test. However, this effect is common among the active regions with large numbers of flares, and is probably a consequence of obscuration, which results in some short waiting times between events in the same active region being missed.

As mentioned above, active region 6368 was the most flare productive active region during the observing period, producing a total of 110 flares. Figure 7 shows the rate analysis for this active region. The Bayesian procedure breaks up the history of flaring into three Blocks, as shown in the middle panel. The observed waiting-time distribution (histogram in lower panel) appears to be reasonably well represented by the piecewise constant Poisson model of Equation (2) (solid curve in lower panel). A Kolmogorov–Smirnov test indicates there is some discrepancy, but only at the 10% significance level. Once again the discrepancy is greatest for short waiting times: the model predicts that there should be around 14 waiting times less

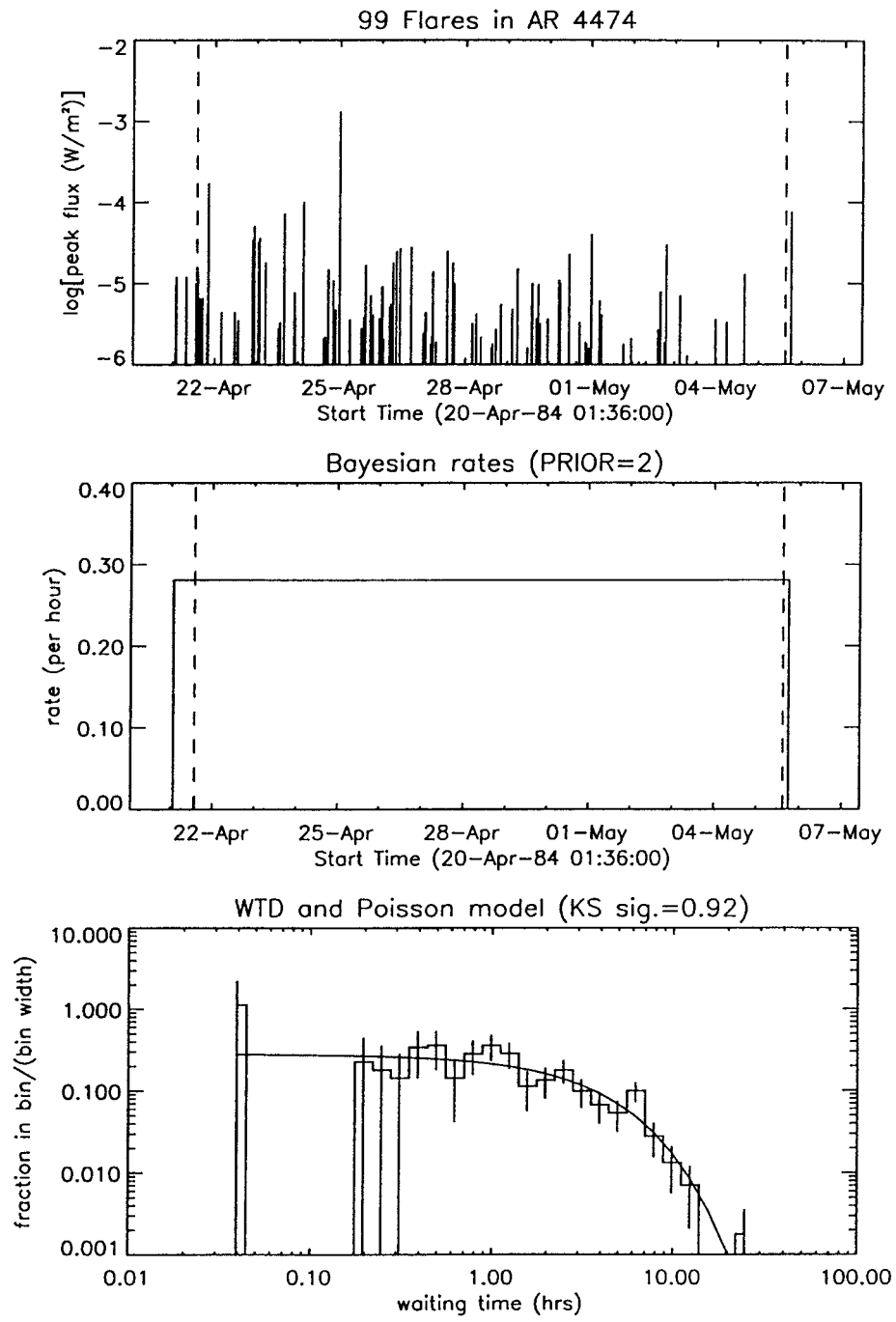

Figure 6. Analysis of flares in AR 4474. *Top panel*: schematic of time history of flaring. *Middle panel*: Bayesian analysis. *Bottom panel*: waiting-time distribution.

than 0.3 hours, whilst only three are observed. This effect may be a consequence of obscuration.

As an example of an active region with a dramatic change in flaring frequency, consider active region 7260. Figure 8 shows the analysis of flares from this region, which was observed from 12 August until 24 August 1992. The flaring rate appears to increase suddenly after 19 August, and the Bayesian analysis confirms this impression, indicating that the rate changes from 0.03 flares per hour to around 0.51 flares per hour, an increase by more than a factor of 15. In the latter part of the observing period the rate is determined to drop again. The piecewise constant Poisson model is found to provide a good model for the observed WTD (K-S significance 0.84). Six waiting times less than 0.3 hours are expected, and four are observed.

Moon *et al.* (2001) looked at the waiting-time distributions for GOES flares in six highly flare productive active regions: 5395 (with 107 flares above C1 class), 5747 (53 flares), 6233 (48 flares), 6545 (57 flares), 6659 (72 flares), and 6891 (81 flares). They reported that the WTDs of these active regions were consistent with exponentials, implying a constant rate of flare production. Application of the Bayesian procedure confirms a constant rate of flaring in these active regions with the exception of AR 6545, which is decomposed into three Bayesian Blocks. The Bayesian procedure indicates that the rate of flaring in AR 6545 is lower near the beginning and end of the observing period, although the periods with lower rate do not greatly influence the WTD. Given the frequency with which active regions exhibit rate variation (Figure 4), Moon *et al.* (2001) appear to have been lucky to have chosen a sample of active regions with little or no detectable rate variation.

The rate analysis was applied to all active regions with greater than 10 flares, as summarized in Figure 4. In general the piecewise constant Poisson process was found to provide a good model for the observed WTDs. There is evidence for some departure from the Poisson model in the statistics of the significances returned by the Kolmogorov–Smirnov test. Out of the 462 active regions with greater than 10 events, 117 regions ($\approx$ 25%) had a K–S significance less than 5% (if the model is correct, then by definition 5% of regions should have a K–S significance below 5%). However, for active regions with larger numbers of events, the agreement is better. For example, out of the 35 active regions with more than 50 events, only 4 ($\approx$ 11%) have a K–S significance less than 5%. The smallest significance out of those 35 is about 1%. For small numbers of events, the Bayesian procedure is less reliable at detecting rate variations (e.g., see Figure 5). With this in mind, these results may indicate that the Bayesian procedure is missing some rate variation (which is influencing the WTD), in particular for small numbers of events. It is also possible that the piecewise constant Poisson model is inadequate. For example, the Bayesian Blocks may not well represent smooth variations in the flaring rate. Also, as argued in Section 2.2, obscuration strictly invalidates the Poisson assumption of independence. However, in general the piecewise constant Poisson model provides a good representation of the observed WTDs, and among the regions with large

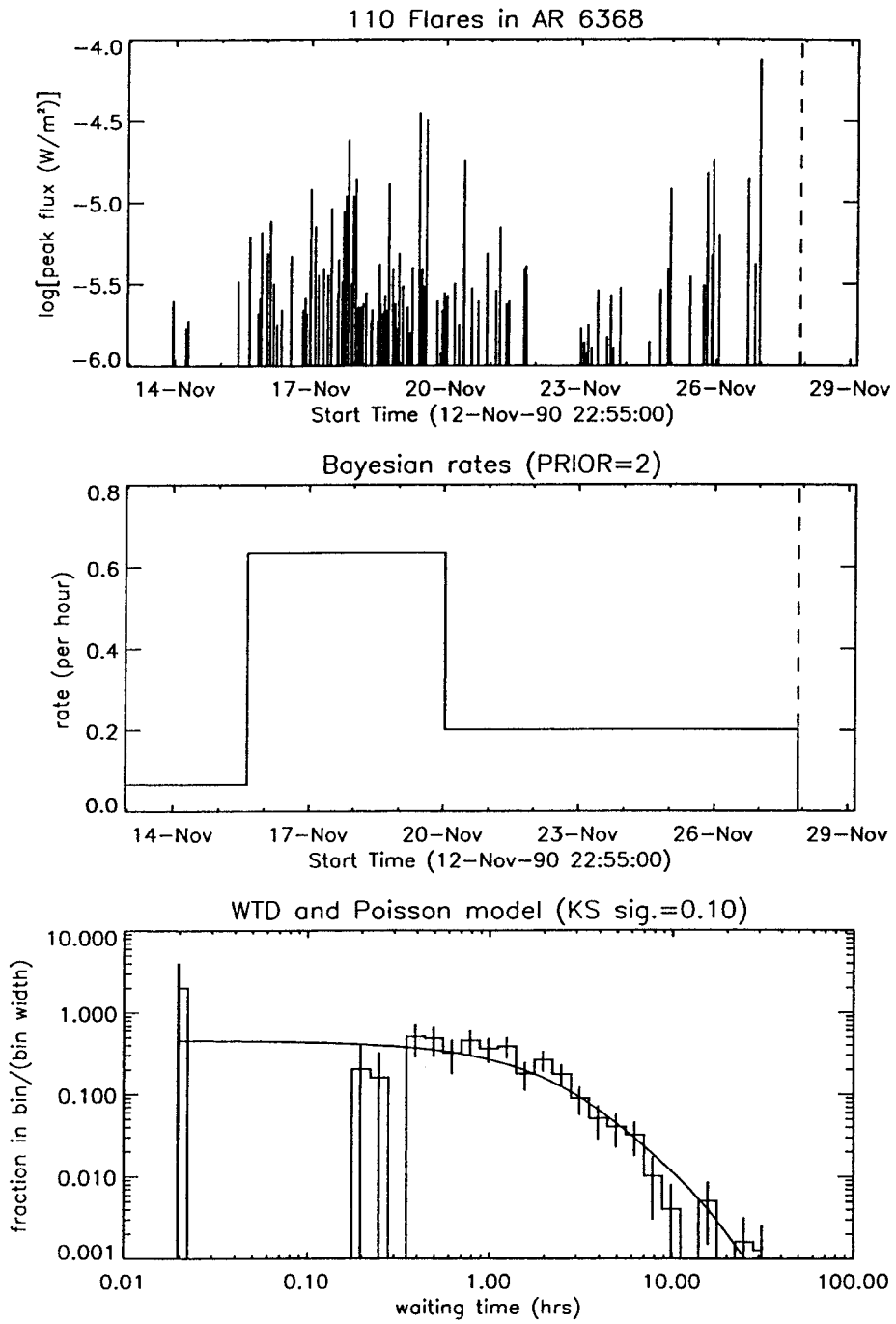

Figure 7. Analysis of flares in AR 6368, following the layout of Figure 6.

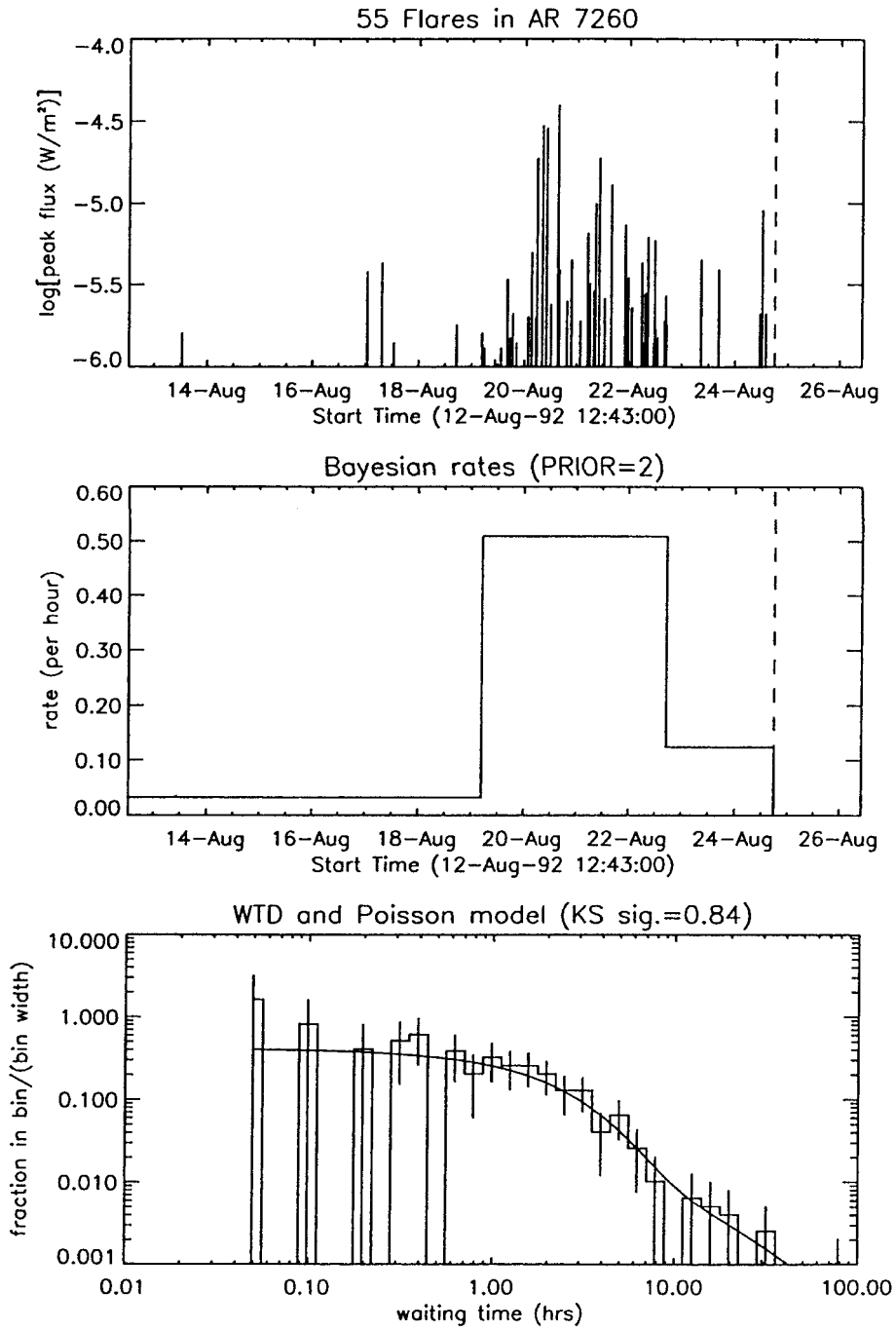

Figure 8. Analysis of flares in AR 7260, following the layout of Figure 6.

numbers of events there is no example of a very significant departure from the model.

The influence of obscuration on the results for individual active regions has been discussed, but the other selection bias – the flux-dependent identification of flares with active regions – has not been addressed. It is difficult to say precisely how this effect influences the results obtained. If the Poisson hypothesis is correct, i.e., active regions produce flares with a locally constant mean rate per unit time, independent of the size of the events, then this effect should act only to reduce the overall rates that are observed. The forms of the waiting time distributions should be unaffected.

3. Conclusions

The rate of occurrence of flares in active regions on the Sun during the period 1981–1999 is here analyzed using the GOES soft X-ray flare catalog together with the USAF/MWL active region catalog. The main results are as follows. In Section 2.2 evidence is presented for obscuration of flares during the slow decline phase of large flares, an effect which is likely to mean that many events are missing from the GOES catalog. It is estimated that in the absence of obscuration the number of flares above C1 class would be higher by $(75 \pm 23)\,\%$. A second observational selection effect – the flux-dependent identification of flares with active regions – is also identified. In Section 2.3 the distributions of flare number and of mean flaring rate are shown to be approximately exponential, although there are excess numbers of active regions with low flare numbers and flaring rates. The influence of flare selection effects on these distributions is discussed in detail. It is concluded that the departure of the observed distributions from exponentials is not due solely to the selection effects. In Section 2.3 evidence is presented that the flaring rate in individual active regions often changes with time. Regions with and without rate variation are examined in detail. The waiting-time distributions are generally found to be well represented by a piecewise constant Poisson process, with a history of rates estimated from the data using a Bayesian procedure. The influence of selection effects on the results for individual active regions is discussed.

The problem of obscuration illustrates the difficulties met in statistical studies of flares. It is important to be wary of observational effects and their influence on flare samples, and throughout this paper the results are discussed in light of the possible influence of these effects. For many statistical studies hard X-ray flare observations are more suitable than soft X-ray observations, because they do not suffer from the same problems with background emission. However, the GOES flare catalog provides a unique resource for statistical studies, because of its long observation period, and because of the relative absence of data gaps.

One outcome of the present study is a confirmation of the Poisson model of flare occurrence in individual active regions. This result indicates that flares in individual active regions occur as independent events, i.e., the occurrence of a flare does not make another flare more or less likely. The observed WTDs do not provide evidence for flare sympathy, or for long term correlations in flare occurrence (cf., Boffeta *et al.*, 1999). These findings are consistent with the avalanche model for flare statistics (e.g., see Wheatland, Sturrock, and McTiernan, 1998).

There are many possible applications of the rate analysis outlined above. To conclude this paper two examples are considered: the application to quantitative flare prediction, and a possible use of the method to gain insight into the flare mechanism.

Solar flare prediction is a problem of great practical significance. The Space Environment Center of the NOAA currently provides sophisticated three-day probability forecasts for flares with soft X-ray classes C, M, and X, based on a variety of observations. Here we consider a simpler approach using only observed flare statistics, which is an elaboration of a suggestion by Moon *et al.* (2000). Assume that an active region has been observed for a certain time since its appearance on the solar disk. If the Bayesian Blocks procedure is applied to the time history of GOES flares observed from the active region during that time, then a current rate of flaring λ_0 (above a threshold peak flux $\mathcal{F}_0$) is obtained. Suppose that the probability $P_1(T)$ of observing at least one flare above a peak flux $\mathcal{F}_1$ during a subsequent time interval T is required. Assuming the Poisson model for flare occurrence, we have

$$P_1(T) = 1 - e^{-\lambda_1 T}, \tag{5}$$

where λ_1 is the rate of flaring above $\mathcal{F}_1$. Based on the expected frequency-peak flux distribution for flares, Equation (4), the expected rate of flaring above $\mathcal{F}_1$ is

$$\lambda_1 = \lambda_0 \left(\frac{\mathcal{F}_1}{\mathcal{F}_0} \right)^{-\gamma+1}. \tag{6}$$

Equations (5) and (6) provide the required probability estimate. As new observations become available the Bayesian Blocks procedure can be reapplied to detect (if present) any variation in the flaring rate above λ_0, and this can be used to update the probability forecast.

This method relies on all active regions following the same power-law peak flux distribution. As discussed in Section 2.3, Kucera *et al.* (1997) have presented evidence for a departure from the global power-law distribution in small active regions, which is attributed to an upper limit to the available free energy in these regions. This effect may lead to the method overestimating the probability of large flare occurrence in small active regions. However, small regions typically have low rates of flaring, and so the probabilities assigned to large events will be small, and the error may not be significant. In future work the forecasting method will be tested on the GOES catalog, including assessment of its success with small active

regions. It is not expected that this simple method will improve on current forecasting techniques, but it may prove useful because it requires only trivial computing and is completely objective.

Finally, a second possible application of the flaring rate analysis presented in this paper concerns the flare mechanism. So far we have considered the GOES (and USAF/MWL) data in isolation. Other datasets may be used in conjunction with the GOES rate analysis to identify physical changes in an active region that are contemporaneous with a change in flaring rate. For example, magnetogram data could be used to test whether changes in flaring rate are associated with the appearance of new magnetic flux, or with an increase in the magnetic complexity of an active region. The Bayesian Blocks procedure provides an objective determination of when the flaring rate changes, and coincidence with changes in the physical parameters of active regions can be tested statistically. It is hoped that this approach will provide insight into long standing questions concerning the physical mechanisms of flare energy release and flare triggering.

Acknowledgements

Part of this work was performed whilst the author was the recipient of a U2000 Post-doctoral Fellowship at the University of Sydney. Thanks are due to an anonymous referee for suggestions which led to improvements in this paper, in particular in Section 2.2.

References

Biesecker, D. A.: 1994, 'On the Occurrence of Solar Flares Observed with the Burst and Transient Source Experiment', Ph.D. Thesis, University of New Hampshire.

Boffeta, G., Carbone, V., Giuliani, P., Veltri, P., and Vulpiani, A.: 1999, *Phys. Rev. Lett.* **83**, 4662.

Crosby, N.: 1996, 'Contribution à l'étude des Phénomènes Éruptifs du Soleil en Rayons X à partir des Observations de l'Expérience WATCH sur le Satellite GRANAT', Ph.D. Thesis, University Paris VII.

Feldman, U.: 1996, *Phys. Plasmas* **3**, 3203.

Hudson, H. S.: 1991, *Solar Phys.* **133**, 357.

Kucera, T. A., Dennis, B. R., Schwartz, R. A., and Shaw, D.: 1997, *Astrophys. J.* **475**, 338.

Litvinenko, Y. E. and Wheatland, M. S.: 2001, *Astrophys. J.* **550**, L109.

Mayfield, E. B. and Lawrence, J. K.: 1985, *Sol. Phys.* **96**, 293.

Moon, Y.-J., Choe, G. S., Yun, H. S., and Park, Y. D.: 2001, *J. Geophys. Res. Space Phys.* (submitted).

Pearce, G., Rowe, A., and Yeung, J.: 1993, *Astrophys. Space Sci.* **208**, 99.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: 1992, *Numerical Recipes*, 2nd ed., Cambridge University Press, Cambridge.

Sammis, I., Tang, F., and Zirin, H.: 2000, *Astrophys. J.* **540**, 583.

Scargle, J.: 1998, *Astrophys. J.* **504**, 405.

Wheatland, M. S.: 2000, *Astrophys. J.* **536**, L109.

Wheatland, M. S., Sturrock, P. A., and McTiernan, J. M.: 1998, *Astrophys. J.* **509**, 448.