Quantitative Forecasting of Major Solar Flares

Manolis K. Georgoulis[1] & David M. Rust[1]

ABSTRACT

We define the effective connected magnetic field, B_{eff}, a single metric of the flaring potential in solar active regions. We calculated B_{eff} for 298 active regions (93 X- and M-flaring, 205 nonflaring) as recorded by SoHO/MDI during a 10-year period covering much of solar cycle 23. We find that B_{eff} is a robust criterion for distinguishing flaring from nonflaring regions. A well-defined twelve-hour conditional probability for major flares depends solely on B_{eff}. This probability exceeds 0.95 for M-class and X-class flares if $B_{eff} > 1600\ G$ and $B_{eff} > 2100\ G$, respectively, while the maximum calculated B_{eff}-values are near 4000 G. Active regions do not give M-class and X-class flares if $B_{eff} < 200\ G$ and $B_{eff} < 750\ G$, respectively. We conclude that B_{eff} is an efficient flare-forecasting criterion that can be computed on a timely basis from readily available data.

Subject headings: Sun: CMEs — Sun: flares — Sun: magnetic fields — Sun: photosphere

1. Introduction

From the thousands of active regions (ARs) that emerge, evolve, and decay in the solar atmosphere within a typical solar cycle, only a very small percentage ($\sim 10\%$, at most) will ever host one or more major flares. Most ARs are only involved in subflaring activity. The strongest flares, namely those classified as M- and X-class in the GOES $1-8$ Å bandwidth, are the ones mostly associated with coronal mass ejections (CMEs - Andrews 2003). Therefore, predicting major solar flares contributes to the prediction of fast active-region CMEs, which are the principal drivers of adverse space weather. Efficient flare prediction also provides an additional warning for high-energy particles that can reach 1 $AU \sim 20\ min$ after injection.

Solar flares are of magnetic origin, so active-region magnetic field measurements should hold some clues as to which ARs could be their sources. So far, little success has been

[1]The Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Rd. Laurel, MD 20723-6099, USA. manolis.georgoulis@jhuapl.edu, dave.rust@jhuapl.edu

reported in tackling this problem. Falconer, Moore, & Gary (2002; 2003; 2006) attempted to predict CMEs from various nonpotentiality indices of the source ARs, but have yet to show statistically significant results. Cui et al. (2006) provided some correlation of flare productivity with photospheric parameters. Leka & Barnes (2003a; b) examined a large number of photospheric parameters and found no single parameter capable of efficiently distinguishing flaring from nonflaring ARs. In a later study, Barnes & Leka (2006) concluded that the calculated coronal topology in ARs is more useful than the measured magnetic field on the (photospheric or chromospheric) boundary. Only very recently, Schrijver (2007) described a single photospheric metric, namely the unsigned magnetic flux within 15 Mm from a polarity inversion line, that shows some forecasting capability.

Here we describe a physically intuitive, quantitative, and readily calculated index of the flaring potential in solar active regions. In doing this, we use the conventional wisdom that most major flares are triggered near intense magnetic polarity inversion lines. Our analysis (i) abstracts the photospheric magnetic flux distribution in ARs, and (ii) finds the magnetic connectivity in the active-region photosphere by identifying footpoint pairs and the flux committed to the connections between these pairs. Vector magnetograms are not required, so we tested our derivations on a substantial part of the line-of-sight magnetogram archive of the Michelson-Doppler Imager (MDI - Scherrer et al. 1995), on board the Solar and Heliospheric Observatory (SoHO), thereby covering most of solar cycle 23.

2. The analysis method

For a measured magnetic flux distribution on a boundary, be it the photosphere or the chromosphere, knowledge of the *magnetic connectivity* means knowledge of the footpoint pair locations of magnetic flux tubes anchored in the boundary and extending above it. Finding the connectivity of each flux tube would be impractical; instead we start with a simplification of the boundary flux distribution and then we reproduce the magnetic connections by a new technique based on *simulated annealing*.

As Barnes, Longcope, & Leka (2005) suggest, a meaningful way of partitioning the boundary flux into an ensemble of magnetic monopoles, or "charges", is by utilizing the magnetic charge topology (MCT) model. We followed these authors' recipe to develop an MCT model that identifies magnetic flux concentrations above a lower flux limit and with area coverage above a lower area limit. For each identified flux concentration, we also calculated the corresponding flux-weighted centroid location, thus obtaining a sequence of N flux concentrations Φ_k, each with centroid position $\mathbf{r_k}$; $k \equiv \{1, N\}$.

Given a flux distribution with m positive-polarity and l negative-polarity concentrations ($N = m + l$), one may define a $m \times l$ magnetic connectivity matrix $\mathcal{M}$ containing the net flux $|\Phi_{ij}|$ committed to each connection (i, j) ($i \equiv \{1, m\}$ and $j \equiv \{1, l\}$). For non-existing connections, $\Phi_{ij} = 0$. In analogy to $\mathcal{M}$, a matrix $\mathcal{L}$ may contain the distances $L_{ij} = |\mathbf{r_i} - \mathbf{r_j}|$ between the two centroid locations of the i- and j-concentrations. If $\Phi_{ij} \neq 0$, the pair $(\mathbf{r_i}, \mathbf{r_j})$ provides the footpoint location of the connection, while L_{ij} provides the footpoint separation length of the connection. Next step is to calculate $\mathcal{M}$ and $\mathcal{L}$.

Calculation of each field line would be both impractical and unnecessary, because our analysis does not require a detailed coronal topology. Instead, we calculate the connectivity matrices on the boundary as follows: each magnetic flux concentration is normalized with respect to the smallest concentration present, say $|\Phi_{min}|$, and is rounded to the nearest integer. We thus obtain a sequence of flux concentrations Φ'_k, ranging from ± 1 to several flux "units". Each flux unit is assigned to a surface element (for magnetograms, this would be a single pixel) adjacent to the centroid $\mathbf{r_k}$ of its corresponding flux concentration Φ_k. In this manner, the boundary magnetogram reduces to a distribution of point sources, each with flux ± 1. To find out how these sources are connected we use a simulated annealing method (see, e.g., Press et al. (1992)) that identifies connections by simultaneously minimizing the overall magnetic flux imbalance and the overall separation length of these connections. As a result, we minimize the function

$$F = \sum_{i=1}^{m} \sum_{j=1}^{l} \left(\frac{|\mathbf{r_i} - \mathbf{r_j}|}{|\mathbf{r_i}| + |\mathbf{r_j}|} + \frac{|\Phi'_i + \Phi'_j|}{|\Phi'_i| + |\Phi'_j|} \right) \ . \tag{1}$$

Simulated annealing is an extremely robust technique. Starting from an arbitrary initial guess, it proceeds iteratively and converges to a solution where no more connections are changed between iterations. It can be shown that the method absolutely minimizes the chosen function without being misled by local minima. Each connection incurs a "penalty" on the basis of equation (1) and a Boltzmann-type probability function with a gradually decreasing "temperature" (Metropolis et al. 1953). The minimum penalty is achieved when the function is minimized. For perfectly flux-balanced configurations, annealing will restrict like-polarity connections to only those between flux units belonging to the same flux concentration. These connections are, of course, ignored. In case of a reasonable (i.e., up to 20%) flux imbalance, annealing will converge by assigning the weakest possible fluxes to like-polarity connections. These are also ignored and considered open connections, i.e. closing beyond our field of view.

Minimizing the flux imbalance implies that ARs are, to a good approximation, closed magnetic structures. Minimizing the separation length incorporates the empirical rule that

if bipolar flux is concentrated around a polarity inversion line (e.g., a δ-sunspot or an arcade), then, most probably, the flux concentrations involved are connected to each other. A minimum flux imbalance and overall separation length naturally suggest a minimum-energy solution for the studied magnetic structure. The typical time required to converge for a single MDI magnetogram is $\lesssim 20$ *min* and depends on the number of the identified flux concentrations.

During annealing, each paired flux unit is connected to two other flux units, so the flux committed to each connection is $(1/2)|\Phi_{min}|$. Upon convergence, the point sources are clustered back to their corresponding flux concentrations Φ'_k at $\mathbf{r_k}$ and they are multiplied by $|\Phi_{min}|$ to yield Φ_k in physical units. Then, the matrices $\mathcal{M} = \Phi_{ij}$ and $\mathcal{L} = L_{ij}$ are populated with nonzero values where connections $(\mathbf{r_i}, \mathbf{r_j})$ exist. For each existing connection between the flux concentrations i and j, a value of $(1/2)|\Phi_{min}|$ is added to M_{ij}. Figure 1 shows an example of the convergence solution for NOAA AR 10486.

To quantify the rule that major flares occur near polarity inversion lines, one may define a matrix $\mathbf{T} = (\Phi_{ij}/L_{ij})$, whose elements are larger for stronger connection fluxes and shorter separation lengths. To weight potentially flaring structures even more, we define the matrix $\mathbf{B} = (\Phi_{ij}/L_{ij}^2)$, each element of which is expressed in magnetic field units. We further define the *effective connected magnetic field* B_{eff} as

$$B_{eff} = \sum_{i=1}^{m} \sum_{j=1}^{l} \frac{\Phi_{ij}}{L_{ij}^2} \quad , \quad for \quad \Phi_{ij} \neq 0 \quad , \tag{2}$$

that is, by summing all nonzero, finite values of the matrix $\mathbf{B}$. In this sense, the entire AR and its magnetic connections at a given time are described by a single magnetic field strength. We sum over all nonzero values of $\mathbf{B}$ to implement the familiar principle that flux-massive ARs are statistically the most flare-productive ones.

The definition of B_{eff}, equation (2), effectively handles the problem of limited spatial resolution in magnetogram measurements. Obviously, if the spatial resolution of a given magnetogram changes, so does the result of the MCT flux partitioning. The total magnetic flux of the AR, however, does not change significantly, at least for strongly magnetized areas where the filling factor approaches unity. Therefore, the robust simulated annealing should continue to capture the essential magnetic connections in the AR, thus making B_{eff} fairly insensitive to the spatial resolution.

3. Data reduction and selection criteria

We acquired the full-disk $96-min$ MDI magnetogram data from mid 1996 to mid 2005. To minimize projection effects, we restricted calculations to a zone of 41^o EW from disk center. We multiplied the magnetic field measurements by 1.56, to implement the correction of Berger & Lites (2003). To enhance the signal-to-noise ratio, we took $24-hr$ averages of the magnetograms, centered alternatively at UT noon and midnight, thus obtaining an averaged magnetogram every nearly 12 hr. To detect the ARs present in each magnetogram, we used an automatic AR identification technique described in LaBonte, Georgoulis, & Rust (2007). Briefly, the technique looks for flux-balanced magnetic structures and compares their flux-weighted centroid locations with NOAA's Space Environment Center (SEC) archives, to assign a NOAA AR number to each of them.

Each selected AR was then extracted from the full-disk magnetogram. Its magnetic field values were divided by the cosine of the angular distance from disk center, to estimate the radial magnetic field component. Finally, the resulting field distribution was rotated to the heliographic plane, following the formulation of Gary & Hagyard (1990).

Setting the maximum tolerated magnetic flux imbalance to 20% for a normal field strength > 50 G, we identified 298 ARs between 07/05/96 and 07/12/05. All cycle-23 ARs linked to major flares were within the selected sample. We obtained up to thirteen $12-hr$ averaged magnetograms for each AR, since an AR needs 6-7 days to traverse the 82-degree zone of the calculations. To document the major flare history for each AR, we browsed (i) NOAA's GOES archives, and (ii) the *Yohkoh*/HXT flare catalogue[2]. Of the selected ARs, 47 were unambiguously X-flare-productive, with a total of 113 X-class flares, and 46 were unambiguously M-flare-productive, with a total of 455 M-class flares. The remaining (assumed nonflaring) 205 ARs were not unambiguously linked to major flares. We performed an additional sorting of our flaring ARs with respect to *when* a given AR flared. We defined four categories, namely, **CI-regions**, i.e., ARs with one or more major flares during the observing interval (57 ARs), **CIIa-regions**, i.e., ARs with flares prior to the observations (21 ARs), **CIIb-regions**, i.e., ARs with flares after the observations (11 ARs), and **CIIab-regions**, i.e., ARs with flares before *and* after, but not during, the observations (4 ARs).

We calculated B_{eff} for each of the 2140 twelve-hour averaged magnetograms of the 298 selected ARs. As a flux threshold for the MCT model, we took the flux of a 10×10

[2]The catalogue is available at *http://gedas22.stelab.nagoya-u.ac.jp/HXT/catalogue/index.html* and has been compiled by J. Sato, K. Yoshimura, T. Watanabe, M. Sawa, M. Yoshimori, Y. Matsumoto, T. Watanabe, S. Masuda, & T. Kosugi.

pixel ($19''.8 \times 19''.8$) MDI flux concentration with each pixel having a normal field of 50 G ($\sim 1 \times 10^{20}$ Mx). This limits the minimum flux committed to a single connection to $\sim 5 \times 10^{19}$ Mx. Moreover, we did not consider flux concentrations with areas smaller than 40 pixels, even when their absolute flux exceeded 10^{20} Mx.

4. Results

The peak B_{eff} values for each identified AR are shown in Figure 2. For CI-regions, with major flares during the observations, we plotted the closest *preflare* B_{eff}-value (i.e., at most 12 hr before the flare) instead of the overall peak. For multiple flares, the peak preflare B_{eff} was used. We notice that B_{eff} is a robust metric, with a dynamical range spanning ~ 2.5 orders of magnitude. The segregation between flaring and nonflaring ARs is evident, especially for X-flaring ARs (red symbols). X-class flares do not occur in ARs with $B_{eff} < 750$ G (red dotted line). Only 16 out of the 205 ($\sim 8\%$) nonflaring ARs satisfy $B_{eff} > 750$ G. M-class flares, on the other hand, do not occur in ARs with $B_{eff} < 200$ G (blue dotted line). For all, but one, M-flaring ARs, $B_{eff} > 350$ G, with $\sim 56\%$ of the nonflaring ARs showing $B_{eff} < 350$ G.

Using the preflare B_{eff}, we can calculate the likelihood of a major eruption in the host AR within 12 hr of the magnetic field observations. For a certain $B_{eff} = B_0$, with B_0 exceeding the lower B_{eff}-limit for the examined flare class, we simply calculate the ratio of the number of CI-ARs with $B_{eff} > B_0$ over the total number of ARs with $B_{eff} > B_0$. Here, all CIIa-, CIIb-, and CIIab-ARs are considered nonflaring because they did not flare within the next 12 hr after the MDI observations. The X-flaring ARs are naturally included into the calculation of the M-class flare likelihood. The twelve-hour X-class and M-class flare likelihoods can also be viewed as the twelve-hour conditional flare probabilities, P_M and P_X, respectively, and are shown in Figure 3. The probabilities can be readily fitted with respect to B_{eff}. The "no-flare" zones of B_{eff} are shaded green, while the "imminent-flare" ($P > 0.95$) zones are shaded red. The upper B_{eff}-limit in Figure 3 is 4000 G. This is nearly the maximum B_{eff}-value calculated from the MDI data.

For M-flaring ARs, the probability P_M is best fitted by a sigmoidal curve, similar to Cui et al. (2006), but shifted in higher probabilities. The curve is defined as $P_M = P_{max} - (P_{max} - P_{min})[1 + \exp[(\log B_{eff} - B_c)/w]]^{-1}$, where P_{min} and P_{max} are the lower and upper, respectively, asymptotic extremes of P_M, B_c is the logarithm of the central fitted value of B_{eff}, and w is the width of the distribution. For X-flaring ARs, a linear fitting is sufficient. We find, in particular,

$$P_M \simeq 1 - 0.75[1 + \exp[(\log B_{eff} - 2.87)/0.14]]^{-1} \ ; \qquad B_{eff} \in (200, 1600)\ G \qquad \text{(3a)}$$

for the M-flaring ARs, and

$$P_X \simeq -1.78 + 0.82 \log B_{eff} \ ; \qquad B_{eff} \in (750, 2100)\ G \tag{3b}$$

for the X-flaring ARs. Notice that the above P_M and P_X are conservative estimates of the actual flare probabilities, because they include only ARs **unambiguously** linked to major flares. In the last solar cycle, however, numerous, if not most, M-class flares (and even a few X-class ones) were not attributed to an AR. Given that B_{eff} gives rise to significant flare probabilities that tend to unity fairly quickly, especially for X-flaring ARs, it can be useful in forecasting major eruptions and in minimizing false alarms. For M- (X-) class flares, P_M (P_X) exceeds 0.95 for $B_{eff} \gtrsim 1600\ G$ ($B_{eff} \gtrsim 2100\ G$). About 32% (23%) of the CI-regions are already located in these imminent-flare zones.

5. Conclusions and future work

By defining the effective connected magnetic field strength B_{eff}, we properly quantified the magnetic complexity in the lower boundary of solar active regions. By calculating B_{eff} for thousands of active-region magnetograms from SoHO/MDI, we robustly related it to twelve-hour conditional probabilities for major flares. By extension, the respective CME probability is $\sim max\{0.6P_M, 0.95P_X\}$, because 60% (95%) of M-(X-) class flares are eruptive (Andrews 2003). Due to the fast, inexpensive, calculation of B_{eff}, our analysis can lead to a quantitative and real-time solar flare prediction scheme. Full-disk vector magnetography carried out by instruments of the Synoptic Optical Long-Term Investigation of the Sun (SOLIS - Henney, Keller, & Harvey 2007) and on board the Solar Dynamics Observatory (SDO - Scherrer 2002) will improve our forecasting capability, because they will allow calculation of the normal magnetic field component up to within $\sim 65^o$ EW from disk center, as opposed to the $\sim 40^o$ EW-zone that can be used with only line-of-sight magnetograms.

Besides proposing B_{eff} as a readily calculated and physically meaningful metric for solar flare prediction, we aim to fully explain (Georgoulis & Rust, 2007, in preparation) *why* B_{eff} works well in distinguishing flaring from nonflaring ARs, thus establishing the physics behind its usage as a space-weather forecasting tool.

We are grateful to Jeneen Sommers and the SoHO/MDI Team for providing us with a substantial part of the MDI $96-min$ magnetogram archive. We are also thank NOAA's SEC and GOES services and the *Yohkoh*/HXT Team for making their flare catalogue available online. This work is partially supported by NASA Grant NNG05-GM47G.

REFERENCES

Andrews, M. D. 2003, SoPh, 218, 261

Barnes. G., & Leka, K. D. 2006, ApJ, 646, 1303

Barnes, G., Longcope, D. W., & Leka, K. D. 2005, ApJ, 629, 561

Berger, Th. E., & Lites, B. W. 2003, SoPh, 213, 213

Cui, Y., Li. R., Zhang, L., He, Y., & Wang, H., 2006, SoPh, 237, 45

Falconer, D. A., Moore, R. L., & Gary, G. A. 2002, ApJ, 569, 1016

Falconer, D. A., Moore, R. L., & Gary, G. A. 2003, JGR, 108(A10), SSH 11-1

Falconer, D. A., Moore, R. L., & Gary, G. A. 2006, ApJ, 644, 1258

Gary, G. A., & Hagyard, M. J. 1990, SoPh, 126, 21

Henney, C. J., Keller, C. U., & Harvey, J. W., 2007, ASP Conf. Series, 358, in press

LaBonte, B. J., Georgoulis M. K., & Rust, D. M., 2007, ApJ, submitted

Leka, K. D., & Barnes, G. 2003a, ApJ, 595, 1277

Leka, K. D., & Barnes, G. 2003b, ApJ, 595, 1296

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E., 1953, J. Chem. Phys., 21(6), 1087

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterlink, W. T. 1992, Numerical Recipes: The Art of Scientific Computing (New York: Cambridge Univ. Press) 652

Scherrer, P. H. 2002, Fall AGU Meeting, abstract #SH52A-0494

Scherrer, P. H., et al., 1995, SoPh, 162, 129

Schrijver, C. J. 2007, ApJ, 655, L117

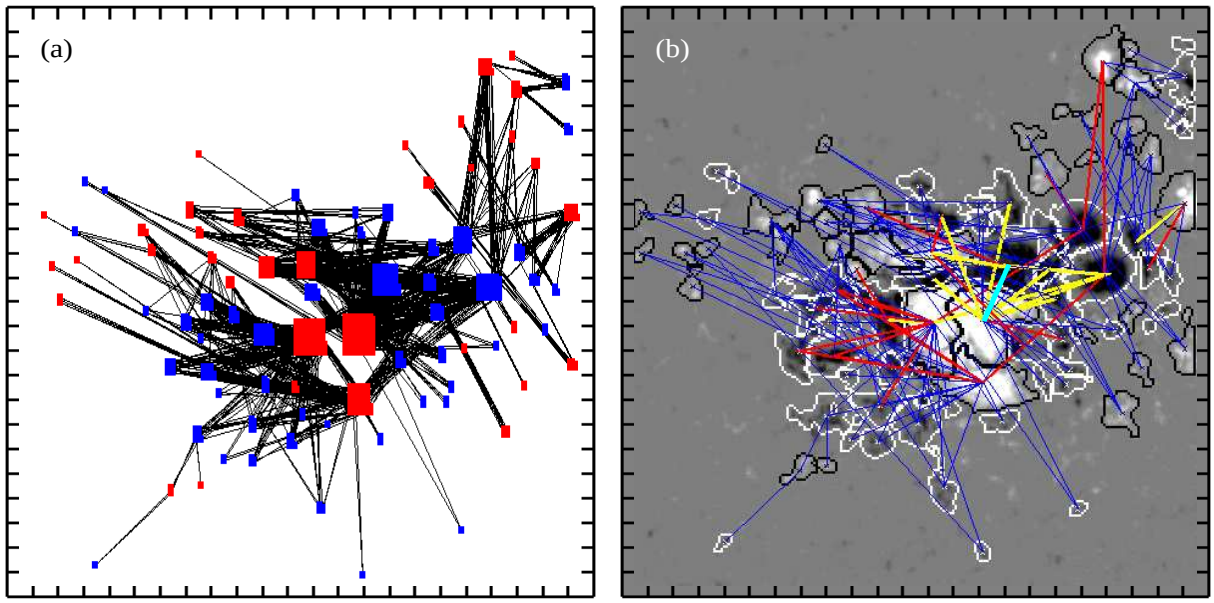

Fig. 1.— Reproducing the photospheric magnetic connectivity in a flux-partitioned MDI magnetogram of NOAA AR 10486 obtained on 2003 October 29 at around 14:27 UT. Tic mark separation is $20''$. (a) The reduced flux-partitioned magnetogram with positive (negative) flux concentrations shown by red (blue), together with the final connectivity returned by the simulated annealing. (b) Translation of the results to the original magnetogram. Centroids of flux concentrations are represented by crosses. Blue lines indicate connections with flux $\Phi \in (0.5 - 5) \times 10^{20}\ Mx$. Red lines indicate connections with $\Phi \in (5 - 10) \times 10^{20}\ Mx$. Yellow lines indicate connections with $\Phi \in (10 - 50) \times 10^{20}\ Mx$. For the cyan connection, $\Phi > 5 \times 10^{21}\ Mx$. Open (unmatched) connections are omitted for clarity. The black and white contours outline the various flux concentrations. The grayscale background shows the estimated normal magnetic field of the AR saturated at $\pm 1000\ G$.

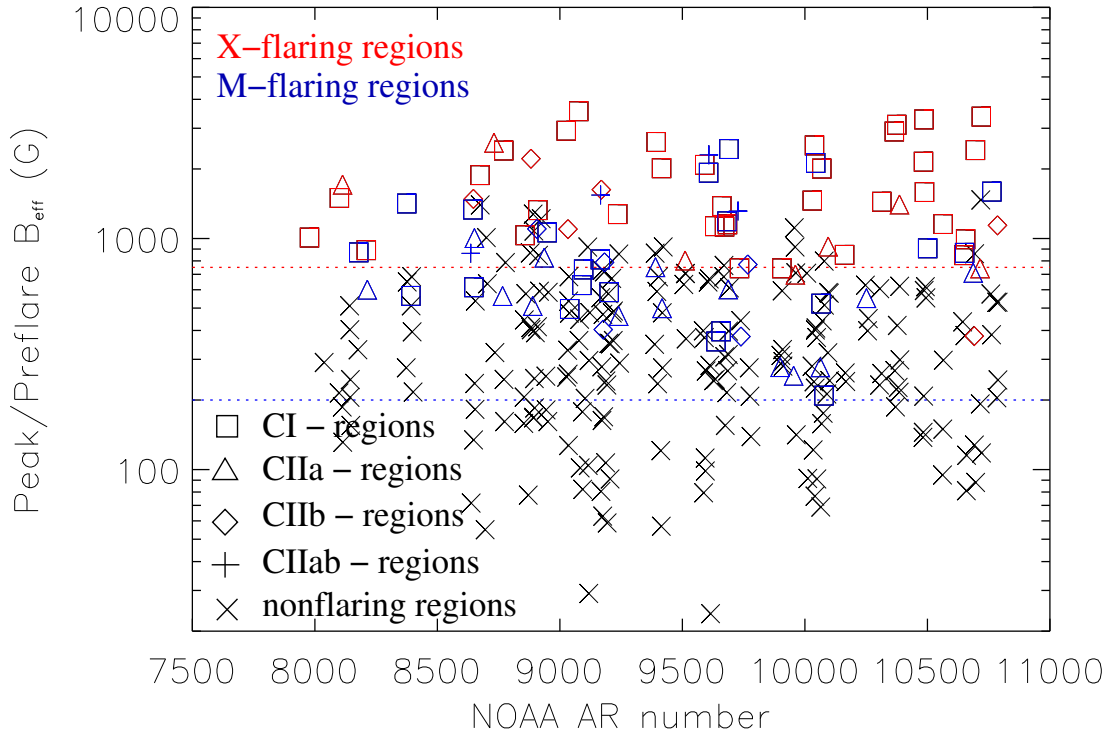

Fig. 2.— Peak B_{eff} values for nonflaring, CIIa-, CIIb-, and CIIab-ARs. For CI-ARs we plot the peak preflare B_{eff} values. X-flaring (M-flaring) ARs are denoted with red (blue) symbols. Their respective lower B_{eff}-limits are shown with the red ($B_{eff} = 750\ G$) and blue ($B_{eff} = 200\ G$) dotted lines.

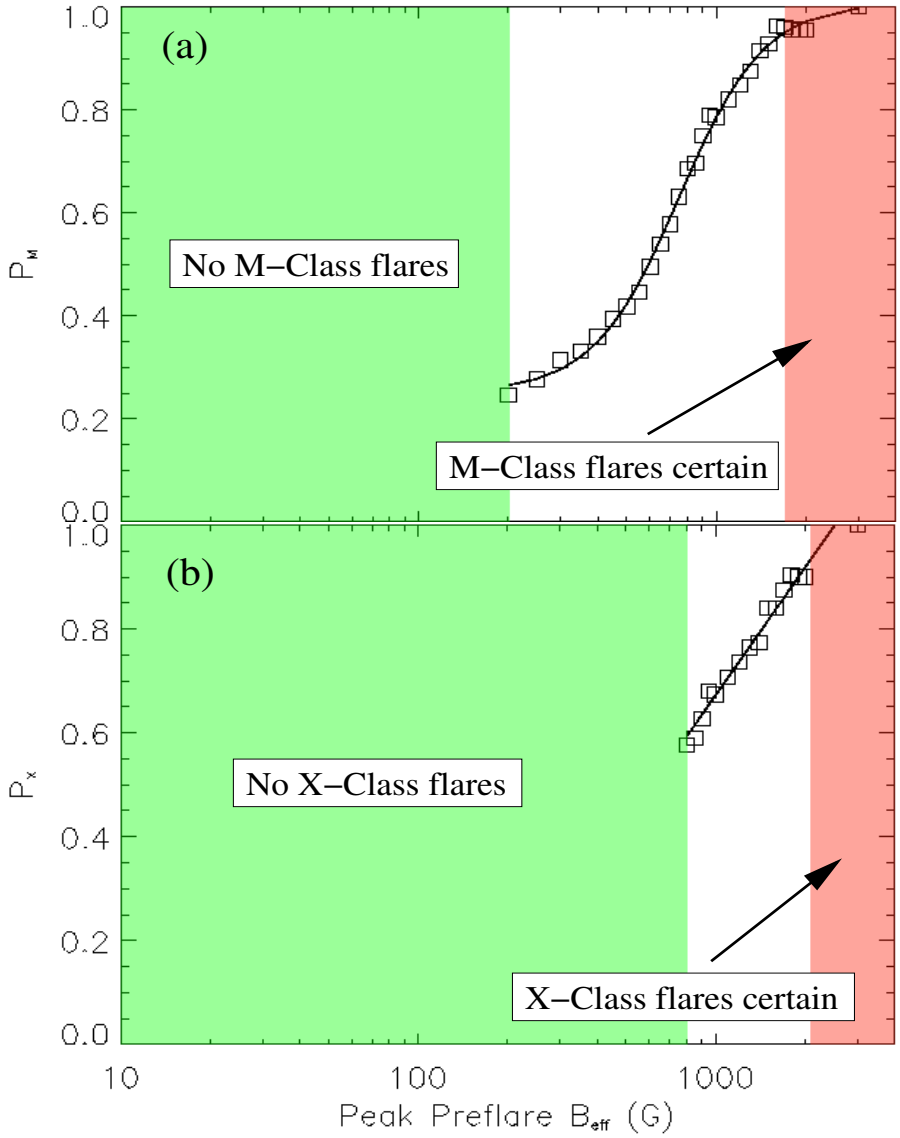

Fig. 3.— Twelve-hour conditional probabilities of major flares as a function of the peak preflare B_{eff}-value in the host ARs (equations (3)). The no-flare and imminent-flare zones are shaded green and red, respectively.