

A STATISTICAL STUDY OF SOLAR ACTIVE REGIONS THAT PRODUCE EXTREMELY FAST CORONAL MASS EJECTIONS

YUMING WANG^{1,2} AND JIE ZHANG¹

Received 2007 December 16; accepted 2008 February 13

ABSTRACT

We present statistical results on the properties of the solar source regions that produced the 57 fastest ($\geq 1500 \text{ km s}^{-1}$) front-side coronal mass ejections (CMEs) from 1996 June to 2007 January. The properties of these fast-CME-producing regions, 35 in total, are compared with those of all 1143 active regions (ARs) in the period studied. An automated method, based on *SOHO* MDI magnetic synoptic charts, is used to select and characterize the ARs. For each, a set of parameters is derived that includes the areas (positive, negative, and total, denoted A_P , A_N , and A_T , respectively), the magnetic fluxes (positive, negative, and total, F_P , F_N , and F_T), the average magnetic field strength (B_{avg}), a quasi elongation (e) characterizing the overall shape of the AR, the number and length of polarity inversion lines (PILs, or neutral lines, N_{PIL} and L_{PIL} , respectively), and the average and maximum magnetic gradient on the PILs (GOP_{avg} and GOP_{max}). Our statistical analysis shows a general trend between the scales of an AR and the likelihood of its producing a fast CME; that is, the larger the geometric size (A_T), the larger the magnetic flux (F_T), the stronger the magnetic field (B_{avg}), and the more complex the magnetic configuration (N_{PIL} and L_{PIL}), the greater the possibility of producing a fast CME. When all the ARs are sorted into three evenly sized groups with low, intermediate, and high values of these parameters, we find that for all the parameters, more than 60% of extremely fast CMEs are from the high-value group. The two PIL parameters are the best indicators of fast-CME production, with more than 80% coming from the high-value group.

Subject headings: Sun: activity — Sun: coronal mass ejections (CMEs) — Sun: photosphere

1. INTRODUCTION

Coronal mass ejections (CMEs) are the most spectacular eruptive phenomenon in the solar corona, which ejects large amounts of mass and magnetic flux into interplanetary space and may cause adverse disturbances in geospace. From 1996 to 2007, more than 10,000 CMEs were observed by the LASCO instrument (the Large Angle and Spectrometric Coronagraph; Brueckner et al. 1995) on board the *SOHO* (*Solar and Heliospheric Observatory*) spacecraft. Based on the CDAW CME catalog,³ the measured CME speeds can be as low as $\sim 21 \text{ km s}^{-1}$ and as high as $\sim 3387 \text{ km s}^{-1}$, with the average being about 450 km s^{-1} . Among them, fast CMEs are of particular interest. These originate from solar active regions with a large amount of free magnetic energy, which gives rise to the kinetic energy of a CME. Further, fast CMEs are usually much more geoeffective than slow CMEs, for example, producing large solar energetic particle events and causing major geomagnetic storms. In this paper, we focus on extremely fast CMEs originating from the front side of the solar disk, that is, faster than 1500 km s^{-1} , or about 3 times higher than average. These fast CMEs make up a very small fraction ($\sim 0.5\%$) of all such events.

The speed of a CME may be closely tied to the magnetic properties of its surface source region. While CMEs can originate from quiet-Sun regions with a filament (or filament channel), it is believed that most, like most flares, originate from active regions (ARs), where a strong magnetic field occupies a relatively large area. In a statistical study of 32 CME events, Subramanian & Dere (2001) found that 84% were associated with ARs. Yashiro et al. (2005) suggested that almost all CMEs associated with sizable flares (stronger than class C3.0) occur in ARs, and the percentage

could reach up to 99% (S. Yashiro 2007, private communication). According to the flare catalogs compiled by the NOAA Space Weather Prediction Center (SWPC) from 2004 to 2007, for which the surface locations of almost all flares are listed, 79% of all recorded flares originated from ARs, and the rate is higher for stronger flares: 89% for class C and above, and 98% for M- and X-class flares. Thus, the properties of ARs are of major concern in studying the production of CMEs and flares.

The relationship between flare production and AR properties has been studied extensively by many researchers (e.g., Sammis et al. 2000; Leka & Barnes 2003, 2007; Maeshiro et al. 2005; Jing et al. 2006; Ternullo et al. 2006; Schrijver 2007; Georgoulis & Rust 2007). In general, larger flares tend to occur in ARs with more complicated morphology, larger magnetic flux, larger magnetic energy, larger helicity, and stronger and longer neutral lines.

Nevertheless, CMEs are different phenomena from flares even though they are related in many respects (see, e.g., Harrison 1995, 2003; Zhang et al. 2001). The relations between CME production/properties and AR properties have not been much pursued until recently. Examining six active regions with vector magnetogram observations, Falconer et al. (2006) demonstrated a good correlation between AR magnetic properties (including the total magnetic flux, three other parameters related to main neutral lines, and two more parameters related to magnetic field twist) and CME production. Guo et al. (2007) investigated 55 flare=CME-producing active regions by use of four magnetic parameters (tilt angle, total magnetic flux, length of main neutral lines, and effective distance) and found that fast CMEs tend to occur in ARs with high magnetic flux and long lengths of the main polarity inversion lines (PILs).

In this paper, we focus on the fastest 0.5% of front-side CMEs and characterize their source ARs. We use an automated method to select and characterize all ARs observed by the *SOHO* Michelson Doppler Imager (MDI) from 1996 June to 2007 January, in order to create a matrix for comparison. The events and ARs studied cover almost the whole of solar cycle 23. The data gathering and

¹ Department of Computational and Data Sciences, College of Science, George Mason University, Fairfax, VA 22030; ywangf@gmu.edu, jzhang7@gmu.edu.

² School of Earth and Space Science, University of Science and Technology of China, Hefei, 230026 Anhui, China; ymwang@ustc.edu.cn.

³ See http://cdaw.gsfc.nasa.gov/CME_list.

TABLE 1
FRONT-SIDE FAST CMEs IN SOLAR CYCLE 23
AND THEIR SURFACE SOURCE REGIONS

No.	Date and Time ^a (UT)	V_{CME}^b (km s^{-1})	Location (deg)	NOAA AR	AR ^c
1.....	1997 Nov 6, 12:10:41	1556	S17, W62	8100	1
2.....	1998 Apr 20, 10:07:11	1863	S42, W63	...	No
3.....	1998 Dec 18, 18:09:47	1749	N26, E35	8415	2
4.....	1999 May 3, 06:06:05	1584	N20, E45	8525	3
5.....	1999 Jun 4, 07:26:54	2230	N19, W75	8552	4
6.....	2000 Jan 6, 07:31:24	1813	N22, W34	8816	No
7.....	2000 May 15, 08:50:05	1549	N22, E85	9002	5
8.....	2000 Jun 25, 07:54:13	1617	N15, W54	9046	6
9.....	2000 Jul 14, 10:54:07	1674	N17, W01	9077	7
10.....	2000 Sep 12, 11:54:05	1550	S19, W06	9163	8
11.....	2000 Nov 8, 23:06:05	1738	N18, W71	9213	9
12.....	2000 Nov 25, 01:31:58	2519	N07, E49	9240	10
13.....	2001 Jan 20, 21:30:08	1507	S09, E37	9313	11
14.....	2001 Apr 2, 22:06:07	2505	N16, W67	9393	12
15.....	2001 Apr 10, 05:30:00	2411	S24, W07	9415	13
16.....	2001 Jun 11, 04:54:05	1647	S13, E81	9501	14
17.....	2001 Jul 19, 10:30:22	1668	S05, W60	9537	15
18.....	2001 Sep 24, 10:30:59	2402	S17, E25	9632	16
19.....	2002 Apr 21, 01:27:20	2393	S15, W81	9906	17
20.....	2002 May 22, 03:50:05	1557	S14, W56	...	No
21.....	2002 Jul 23, 00:42:05	2285	S12, E69	10039	18
22.....	2002 Aug 16, 12:30:05	1585	S11, E19	10069	19
23.....	2002 Sep 5, 16:54:06	1748	N10, E28	10102	No
24.....	2002 Nov 9, 13:31:45	1838	S10, W27	10180	20
25.....	2002 Nov 10, 03:30:11	1670	S11, W37	10180	20
26.....	2003 Mar 18, 12:30:05	1601	S14, W46	10314	21
27.....	2003 Mar 23, 12:06:05	1505	S13, E71	10318	22
28.....	2003 Jun 2, 00:30:07	1656	S06, W85	10365	23
29.....	2003 Jun 15, 23:54:05	2053	S14, E82	10386	24
30.....	2003 Jun 17, 23:18:14	1813	S07, E58	10386	24
31.....	2003 Oct 26, 17:54:05	1537	N02, W37	10484	25
32.....	2003 Oct 28, 11:30:05	2459	S16, E06	10486	26
33.....	2003 Oct 29, 20:54:05	2029	S17, W04	10486	26
34.....	2003 Nov 2, 17:30:05	2598	S16, W57	10486	26
35.....	2003 Nov 4, 19:54:05	2657	S19, W81	10486	26
36.....	2003 Nov 18, 08:50:05	1660	N02, E18	10501	27
37.....	2004 Jan 7, 04:06:07	1581	N01, E75	10537	28
38.....	2004 Jan 7, 10:30:29	1822	N06, E73	10537	28
39.....	2004 Jan 8, 05:06:05	1713	N00, E65	10537	28
40.....	2004 Apr 11, 04:30:06	1645	S15, W45	10588	29
41.....	2004 Nov 7, 16:54:05	1759	N09, W16	10696	30
42.....	2004 Nov 9, 17:26:06	2000	N09, W49	10696	30
43.....	2004 Nov 10, 02:26:05	3387	N10, W49	10696	30
44.....	2005 Jan 15, 06:30:05	2049	N14, E03	10720	31
45.....	2005 Jan 15, 23:06:50	2861	N14, W00	10720	31
46.....	2005 Jan 17, 09:30:05	2094	N14, W23	10720	31
47.....	2005 Jan 17, 09:54:05	2547	N13, W23	10720	31
48.....	2005 Jan 19, 08:29:39	2020	N16, W45	10720	31
49.....	2005 May 13, 17:12:05	1689	N11, E10	10759	32
50.....	2005 Jul 30, 06:50:28	1968	N10, E52	10792	33
51.....	2005 Aug 22, 17:30:05	2378	S12, W60	10798	No
52.....	2005 Aug 23, 14:54:05	1929	S12, W77	10798	No
53.....	2005 Sep 9, 19:48:05	2257	S12, E67 ^d	10808 ^d	34
54.....	2005 Sep 10, 21:52:07	1893	S13, E47 ^d	10808 ^d	34
55.....	2005 Sep 11, 13:00:53	1922	...	10808 ^c	34
56.....	2005 Sep 13, 20:00:05	1866	...	10808 ^c	34
57.....	2006 Dec 13, 02:54:04	1774	S07, W22	10930	35

^a First appearance of the CME in the field of view of the LASCO C2 telescope.

^b Linear CME speed in the combined fields of view of LASCO C2 and C3.

^c Whether or not a CME originated from the MDI AR we identified. The sequential numbers indicate an AR, and "No" indicates negative association with an AR.

^d Information based on the NOAA/SWPC flare list.

^e Date, time, and V_{CME} are adopted from the CDAW LASCO CME catalog.

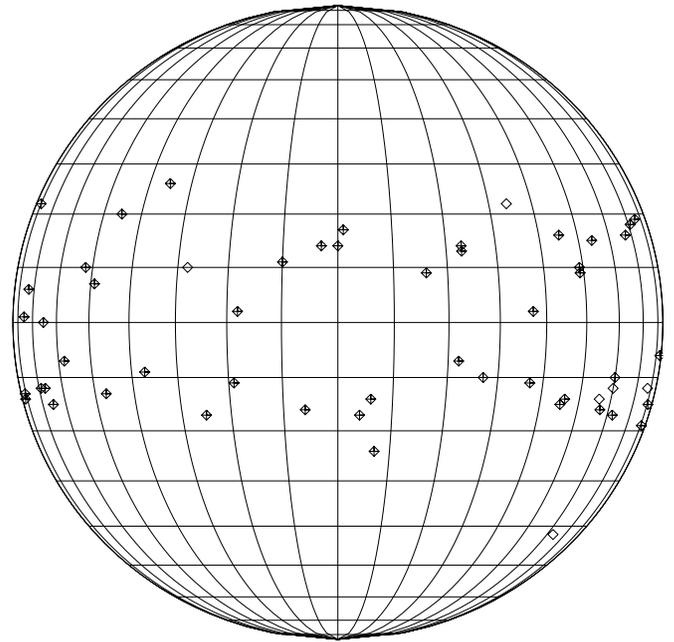


FIG. 1.— Distribution of heliographic coordinates of the surface source regions of the extremely fast CMEs. ARs are indicated by diamonds with plus signs, and quiet-Sun regions are indicated by open diamonds.

processing methods are introduced in the following section. The statistical results are presented in § 3. Section 4 summarizes the work.

2. DATA AND METHODS

2.1. CME Selection and Source-Region Identification

Based on the CDAW CME catalog, the numbers of CMEs with speeds over 1000, 1500, 2000, and 2500 km s^{-1} from 1996 June to 2007 January are 496, 122, 37, and 9, respectively. For the sake of avoiding a time-consuming manual examination of all the fast CMEs, we chose the category with speeds $\geq 1500 \text{ km s}^{-1}$, which has a limited population of 122 events and is more than 3 times faster than the CME average. Through a visual examination of *SOHO* EIT (EUV Imaging Telescope; Delaboudinière et al. 1995) and LASCO movies, 57 CMEs were found to originate from the front side of the solar disk and thus be usable to study the source regions. Table 1 lists these front-side fast CMEs, including their speeds and source-region locations. These events (except Nos. 55 and 56, which were determined according to the flare information in the absence of EIT data) have definite eruptive signatures, including large-scale dimming, compact brightening, or posteruption loop arcades as seen in the EIT coronal images. Figure 1 shows the heliographic coordinate distribution of the surface source regions of these CMEs. Apparently, the sample is uniform in longitude, along which direction they are almost equally distributed; there is no bias toward the limbs, from which a CME may appear faster because of the smaller effect of projection on speed measurements.

2.2. Determining the Properties of Active Regions

A catalog of solar ARs is compiled on a daily basis by the SWPC.⁴ It is based on a near-real time active region report from approximately half a dozen ground-based observatories around the globe, including observations in white light and $\text{H}\alpha$ and magnetograms. The compiled Solar Region Summary, for each identified AR (or sunspot group) on the front disk, lists the assigned NOAA

⁴ See <http://www.swpc.noaa.gov/ftpmenu/forecasts/SRS.html>.

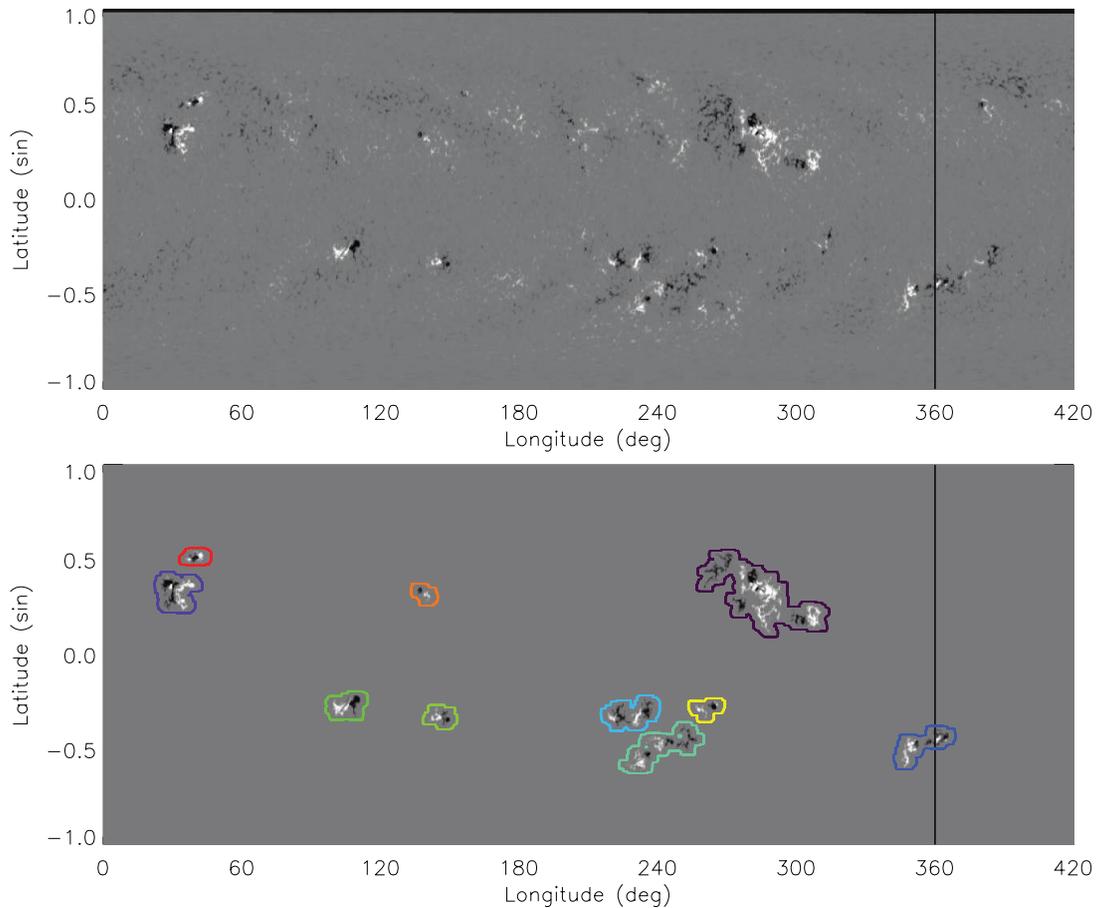


FIG. 2.—Example (CR 1948, 00:54 UT 1999 April 4 to 07:11 UT May 1) illustrating the original MDI synoptic chart (*top*) and the map showing only ARs extracted with our automated method (*bottom*).

AR number, its heliographic coordinates (rotated to 24:00 UT), Carrington longitude, total corrected area of the sunspot group, Modified Zurich classification (McIntosh 1990), longitudinal extent of the group in heliographic degrees, total number of visible sunspots in the group, and magnetic classification of the group. For the 11 yr period from 1996 to 2006 inclusive, there are in total 2999 active regions listed by NOAA. Evidently, the NOAA AR catalog does not provide the necessary quantitative information, such as magnetic field strength and total magnetic flux, for the purpose of studying the source regions of fast CMEs. For this paper, we generated a version of our own AR catalog from the long-term *SOHO* MDI observations from 1996 to 2006.

To simplify the process of identifying and quantifying ARs, we used the Carrington rotation (CR) synoptic MDI charts instead of individual magnetograms. When an AR crosses the front disk of the Sun, over about 2 weeks, its appearance may vary as a result of both intrinsic evolution and the projection effect in the time-lapse magnetograms. On the other hand, a synoptic chart is assembled from a stripe of data in the central meridian from individual magnetograms observed over the course of one Carrington rotation. With the use of synoptic charts, the projection effect is greatly reduced. However, the possible evolution of ARs is thus simply ignored. Since the launch of *SOHO* in 1995 December, high-resolution synoptic charts of photospheric magnetic fields from MDI are routinely made available by the MDI team at Stanford University.⁵ There are 143 synoptic charts from 1996 May to 2007 January, covering almost the whole of solar cycle 23.

⁵ See <http://soi.stanford.edu/magnetic/index6.html>.

We have developed an automated method to uniformly identify ARs in MDI synoptic charts. Attempts to detect ARs automatically have been made before, for example, based on simple thresholdings (e.g., Worden et al. 1998; Brandt & Steinegger 1998; Preminger et al. 2001). Turmon et al. (2002) utilized pattern recognition techniques to distinguish different features on the solar surface. Benkhalil et al. (2006) used a more sophisticated approach that applied the region-growing segmentation technique in addition to thresholding. In this study, we use the region-growing method to automate the AR identification and characterization processes, as follows:

1. Find the kernel or seed pixels in the synoptic chart. We choose a threshold of 15 times of the standard deviation, σ , of the quiet-Sun region (typically $|B_{\text{los}}| \leq 50$ G).
2. Apply the morphological closure procedure. Isolated small features, including anomalously noisy pixels, are effectively removed by applying a morphological erosion operation followed by a dilation operation with a size of 15 pixels on the 3600×1080 pixel map.
3. Apply the region-growing procedure. Starting from the kernels defined above, all connected pixels with $|B_{\text{los}}| \geq 5\sigma$ are included in the region that is grown.
4. Apply the morphological opening procedure. Regions that are very close in space are merged, by first applying the dilation operation followed by the erosion operation with a size of 60 pixels.

Figure 2 illustrates the result of this morphological analysis process. The top panel shows the MDI synoptic chart for CR 1948, corresponding to the period from 00:54 UT on 1999 April 4 to

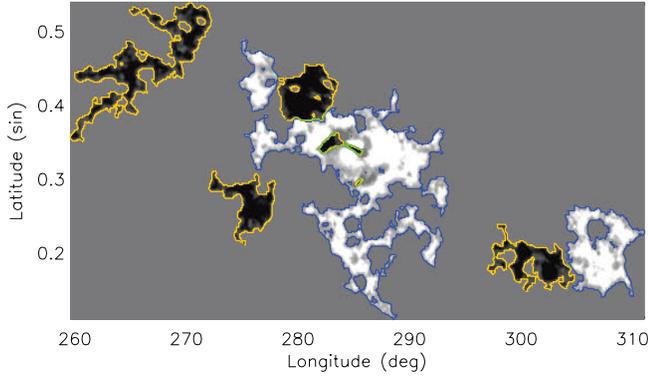


FIG. 3.— An AR from Carrington rotation 1948, showing the areas of positive (white patches with blue boundaries) and negative (black with yellow boundaries) polarity, and the inversion lines (green curves) between them.

07:11 UT on 1999 May 1, and the bottom panel highlights all the extracted active regions. Ten individual ARs are indicated by the colored lines. The synoptic chart actually processed is an enlarged chart that is patched on the right with the leftmost portion of the previous synoptic chart (separated by the vertical black line); this treatment is to prevent those ARs crossing the Carrington rotation boundary from being cropped. The details of the image processing method, including how sensitively the results depend on various intensity thresholds and area boxes chosen, will be pursued in another paper. Changing the AR selection thresholds does not have a significant impact on the results of this paper, since this study mainly concerns the relative properties of ARs.

Once an AR is identified and its boundary determined, it is straightforward to extract many useful parameters. We use one AR in CR 1948 (Fig. 3) to illustrate the characterization. The white and black patches, enclosed by blue and yellow lines, correspond to the regions of positive and negative polarity in the region. We calculate the areal size of the positive polarities (A_P), the size of the negative polarities (A_N), and the total size ($A_T = A_P + A_N$). We also calculate the total positive magnetic flux (F_P), the total negative magnetic flux (F_N), and the total unsigned flux (F_T). We derive the average magnetic field strength ($B_{\text{avg}} = F_T/A_T$), and we also define a quasi-elongation parameter for ARs, $e = 1 - D_s/D_{\text{max}}$, where D_s is the radius of an assumed circle of total size A_T and D_{max} is the greatest distance between two points in the AR. For the region shown in Figure 3, we find that $A_P, A_N, A_T, F_P, F_N, F_T, B_{\text{avg}}$, and e are $1.9 \times 10^4 \text{ Mm}^2$, $1.2 \times 10^4 \text{ Mm}^2$, $3.1 \times 10^4 \text{ Mm}^2$, $505.7 \times 10^{12} \text{ Wb}$, $280.7 \times 10^{12} \text{ Wb}$, $786.4 \times 10^{12} \text{ Wb}$, 72.8 G, and 0.32, respectively. These parameters quantify the intensity and morphology of the AR, which may be useful in relating ARs to eruptive phenomena.

Further, we extract the polarity inversion lines (PILs, or neutral lines) of the host AR. It is generally believed that the properties of PILs are related to the stored free energy of the AR. For each AR, we determine the number of PILs (N_{PIL}) and their length (L_{PIL}). The PILs are indicated by green curves in Figure 3. A PIL is a collection of linearly linked pixels that lie between pixels of opposite polarity. In this AR, there are six separate PILs with a total length (L_{PIL}) of 185.2 Mm. We also determine the average magnetic gradient across the PILs (GOP_{avg}) and the maximum gradient (GOP_{max}); the numbers are 152.7 and 496.4 G Mm^{-1} , respectively.

In summary, a total of 12 magnetic parameters are derived; A_T indicates the size of an AR, F_T and B_{avg} suggest how strong the region is, e indicates how much the overall shape deviates from a perfect circle, N_{PIL} and L_{PIL} indicate the morphological complexity

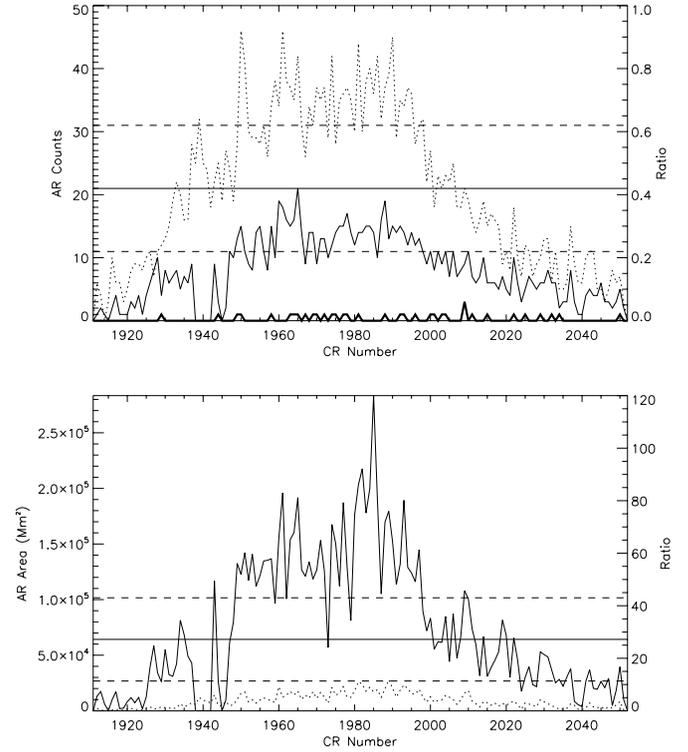


FIG. 4.— Comparison of our identified ARs with NOAA ARs from CR 1911 to CR 2052: *top*, AR counts; *bottom*, total AR area. The solid lines present our results and the dotted lines show the NOAA results. The horizontal solid and dashed lines indicate the average ratio of our results to the NOAA results and the standard deviations. The heavy solid line in the top panel denotes the number of ARs hosting extremely fast CMEs.

of an AR, and GOP_{avg} and GOP_{max} measure the free energy stored between opposite polarities. These parameters are expected to have certain relationships with the properties of eruptive events that result. For instance, F_T was ranked first in effectiveness at predicting flare occurrence by a comprehensive discriminant data analysis (Leka & Barnes 2007). A similar positive correlation between F_T and large flare productivity was obtained by Schrijver (2007). In addition, ARs with larger areas, more complicated configurations, or both tend to produce larger flares (see, e.g., Sammis et al. 2000; Ternullo et al. 2006).

2.3. Comparison with NOAA Active Regions

We compare our MDI AR catalog with the NOAA AR catalog in Figure 4. The top panel displays the AR counts in each CR from our catalog (*solid line*) and the NOAA catalog (*dotted line*). The variation in counts along the solar cycle seems to be the same for the two lines: small near solar minimum and reaching a peak near solar maximum. However, the number of our ARs is smaller than that of NOAA ARs. The ratio of the two counts is 0.42 averaged over the entire period studied, with a standard deviation of 0.20. The smaller number of ARs identified by our method can be attributed to several selection effects: (1) We use magnetogram data. Regions that are sufficiently close to each other in the distribution of magnetic field may be grouped into a single large AR. The selection of NOAA ARs is largely based on the visual appearance of sunspots, which are usually compact in size and discrete in distribution. On the other hand, the surface magnetic field, which we use to define the size, may extend to a much larger area than that of the sunspots. (2) ARs with weak magnetic fields are not selected, because of the adoption of a high kernel threshold. (3) Some short-lived ARs (emerging from the western hemisphere

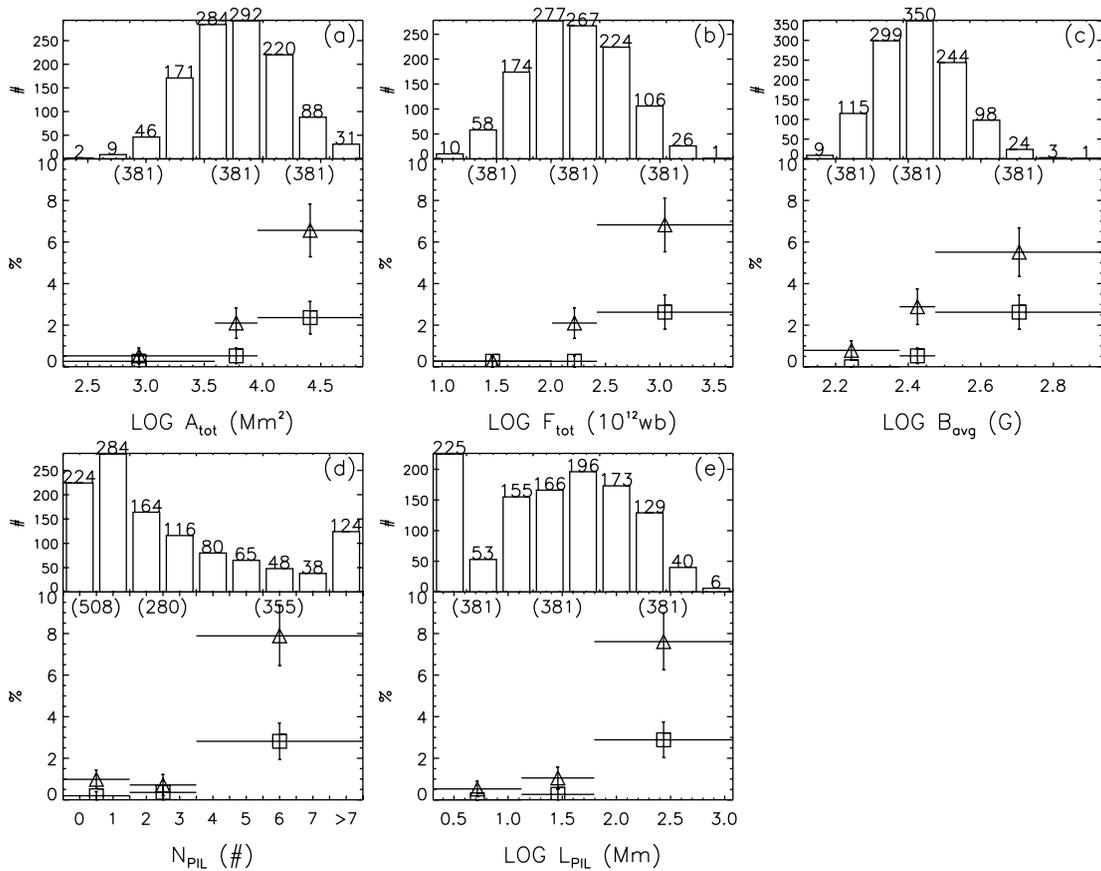


FIG. 5.—Distributions of AR numbers for various AR parameters: (a) total area; (b) total magnetic flux; (c) average magnetic field; (d) number of PILs; (e) total length of PILs. The first three parameters and the last one are plotted in logarithmic scale along the x-axis. In each panel, the histogram for all ARs is shown at the top in nine bins, while the ratios of the ARs hosting extremely fast CMEs to all ARs in three equally numbered general AR groups are shown below (see text for details). The numbers in parentheses give the number of ARs in the three groups. The triangles are for the extremely fast CMEs, and the squares are for the fast CMEs that originated at longitudes of $\pm 30^\circ$.

and lasting less than one Carrington rotation) are not caught in synoptic charts. A full comparison between our catalog and the NOAA catalog will be presented in a separate paper.

The bottom panel of Figure 4 presents the total AR areas during each CR for both the MDI ARs we identified and the NOAA ARs. In contrast to the AR counts, the areas from our method are much larger than those from NOAA. The ratio suggests that the total AR area from our method is on average 27.2 times that of NOAA ARs. The AR area estimated by NOAA concerns only sunspots, which usually correspond to regions with extremely strong magnetic fields. On the other hand, we include all regions with $|B_{\text{los}}| \geq 5 \sigma$ that connect to the AR kernels.

3. RESULTS

Using our method, we have identified and characterized 1143 MDI ARs for Carrington rotations 1911–2051, covering the period from 1996 June 28 to 2007 January 8. It seems that the five parameters that characterize the overall scales of an AR (A_T , F_T , B_{avg} , N_{PIL} , and L_{PIL}) are all fairly related to the production of fast CMEs. The distributions of numbers of ARs for these five parameters are shown in Figure 5. All parameters except N_{PIL} are plotted on a logarithmic scale and divided into nine bins. One can see that all the distributions, except for L_{PIL} , are unimodal, with a central peak. Figures 5a and 5b, for A_T and F_T , indicate the distributions of geometric size and magnetic flux value, respectively. Most ARs have a size of $\sim 10^{3.8} \text{ Mm}^2$, with the total magnetic flux peaking at $\sim 10^{14} \text{ Wb}$. This kind of centrally peaked distribution

is much different from that of the size distribution of sunspot groups studied before, which is found to be exponential (e.g., Schrijver 1988; Howard 1996) or to follow a power law (e.g., Harvey & Cornelis 1993). A detailed analysis of the causes of these differences is beyond the scope of this paper and will be pursued separately. Here we focus on the ARs that produce extremely fast CMEs.

Table 1 lists all the fast CMEs studied, the heliographic coordinates of the surface source regions, and the NOAA numbers of the host ARs. Fifty-five of the 57 events studied originated from NOAA ARs. The remaining two, Nos. 2 and 20, originated neither in a NOAA AR nor in one of our MDI ARs. Careful inspection shows that these two fast CMEs were associated with the eruption of giant filaments. Four more CMEs (Nos. 6, 23, 51, and 52) did not originate from one of our MDI ARs. The surface source regions of these CMEs are too small or weak in the MDI synoptic charts to qualify as MDI ARs using our method. A summary is given in Table 2. In total, 89.5% (51 out of 57) of the fast CMEs originated from our MDI ARs (their locations are indicated by diamonds with plus signs in Fig. 1). The percentage for AR origination is as high as 96% when NOAA ARs are used.

These fast CMEs originated from 35 individual MDI ARs. Compared with the total of 1143 MDI ARs during the period under study, the percentage of ARs hosting these extremely fast CMEs is quite small, only about 3.1%. The distribution over the solar cycle of these fast-CME–hosting ARs is plotted by the thick solid

TABLE 2

STATISTICS OF EXTREMELY FAST CMEs AND THE HOST ACTIVE REGIONS

Quantity	No.
Total ARs.....	1143
All front-side fast CMEs:	
Front-side CMEs.....	57
Front-side CMEs originating from ARs.....	51 (89.5%)
ARs hosting these CMEs.....	35 (3.1%)
Fast CMEs originating from E30° to W30°:	
Front-side CMEs.....	17
Front-side CMEs originating from ARs.....	16
ARs hosting these CMEs.....	12

line in Figure 4 (*top*). Apparently, the majority of these ARs occurred during solar maximum.

The overall distributions of these fast-CME–hosting ARs are shown along with that for all ARs in Figure 5. While the distributions of all the ARs (*histograms*) are shown in nine bins thanks to their large number, the distributions of fast-CME–hosting ARs are displayed in only three groups with low, intermediate, and high values of the derived parameters, because of their limited number. The boundary values of the parameters for each group are chosen such that there are equal numbers of ARs in each. The values used to make these equal divisions are listed in the second column of Table 3. Note that for N_{PIL} the numbers cannot be exactly equalized, but are kept close to each other, because N_{PIL} is highly discrete. Then we determine, for each of the three groups, the number of ARs that produced fast CMEs, which indicates the effectiveness of the parameters used in discriminating fast-CME–hosting ARs from the general population. The triangles show the ratio of ARs hosting fast CMEs to all ARs, that is, the probability of an AR’s producing a fast CME. The horizontal bars indicate the range of values for the groups, and the vertical bars indicate the uncertainty of the calculated percentage probability, defined as $[p(1-p)/N]^{1/2}$, where p is the probability and N is the total sample number; this uncertainty is based on the assumption that the occurrence of fast CMEs obeys a binomial distribution.

It is obvious that the probability of producing fast CMEs increases as the value of the AR parameters increases. This trend is true for all five parameters plotted in the figure. The differences between the low/middle groups and the high-value groups all exceed the estimated uncertainty. In particular, N_{PIL} and L_{PIL} manifest the largest increase from low/middle to high-value groups; the probability increases by a factor as large as 8. This kind of trend indicates that the larger the AR parameters, that is, the larger the geometric size, the larger the magnetic flux, the stronger the magnetic field, and the more complex the magnetic configuration, the higher the likelihood of producing a fast CME.

Table 3 lists the occurrence numbers and percentages of fast-CME–hosting ARs in the three groups discussed above (the third through fifth columns). For all the parameters studied, more than 60% of the fast-CME–hosting ARs lie in the high-value group. The PIL length (L_{PIL}) has the largest percentage (82.9%). These results indicate that the PIL-related parameters are the best in determining whether an AR will produce extremely fast CMEs.

As mentioned above, the evolution of ARs was not taken into account when the AR parameters were derived, since the calculations are based on synoptic charts in order to minimize the projection effect. An AR may have evolved to some extent before the time of CME occurrence and the time of central meridian crossing; the difference may be as much as 7 days for those CMEs originating from the limbs. To find out how significantly temporal evolution affects our statistical results, we made a similar study but considered only those CMEs that were close to the central meridian. The squares in Figure 5 and the last three columns in Table 3 show the results for these near–central meridian CMEs (longitudes within $\pm 30^\circ$). It is found that there is no significant difference from those for all front-side CMEs.

Some studies have shown that flux variations (e.g., emergence or cancellation) are related to CME initiation (e.g., Feynman & Martin 1995; Lara et al. 2000; Green et al. 2003; Sterling et al. 2007). We checked the Hale classes of all the corresponding NOAA ARs that produced a fast CME in the period studied. During the interval between CME launch and the central meridian crossing of the corresponding source AR, 15 out of 35 ARs changed from one Hale class to another. However, of these 15 ARs nine were located very close to the limb when the CME erupted, indicating a maximal projection effect, which may affect the AR classification. Therefore, most of the ARs concerned probably did not change their Hale class. While it may be related to CME initiation, the small-scale flux variation may not alter the overall structure (and the stored energy) of the coronal magnetic field where CMEs originate. Considering that this study is a rather coarse quantitative statistical work, we believe that the effect of AR evolution, when fully addressed, will not change our results.

4. SUMMARY

By applying the region-growing segmentation method, a total of 1143 ARs were extracted from 141 MDI synoptic charts covering the period from 1996 June 28 to 2007 January 8. Twelve quantities (A_p , A_N , A_T , F_p , F_N , F_T , B_{avg} , e , N_{PIL} , L_{PIL} , GOP_{avg} , and GOP_{max}) are derived from the photospheric magnetic field distribution, characterizing the size, strength, morphology, complexity, and free energy of ARs. These active regions produced more than 10,000 CMEs, among which 122 were as fast as or faster than 1500 km s^{-1} . We studied 57 of these extremely fast CMEs that originated from front-side ARs. Through a comparison

TABLE 3
PERCENTAGES OF FAST-CME–HOSTING ARs IN THREE DIFFERENT VALUE GROUPS

Quantity	Boundaries	n^{low}	n^{mid}	n^{high}	$n_{\text{cen}}^{\text{low}}$	$n_{\text{cen}}^{\text{mid}}$	$n_{\text{cen}}^{\text{high}}$
A_T	3.90, 9.05 ^a	2 (5.7%)	8 (22.9%)	25 (71.4%)	1 (8.3%)	2 (16.7%)	9 (75.0%)
F_T	102.1, 263.1 ^b	1 (2.9%)	8 (22.9%)	26 (74.3%)	1 (8.3%)	1 (8.3%)	10 (83.3%)
B_{avg}	238, 298 ^c	3 (8.6%)	11 (31.4%)	21 (60.0%)	0 (0.0%)	2 (16.7%)	10 (83.3%)
N_{PIL}	1, 3	5 (14.3%)	2 (5.7%)	28 (80.0%)	1 (8.3%)	1 (8.3%)	10 (83.3%)
L_{PIL}	13.3, 62.4 ^d	2 (5.7%)	4 (11.4%)	29 (82.9%)	0 (0.0%)	1 (8.3%)	11 (91.7%)

^a In units of 10^3 Mm^2 .

^b In units of 10^{12} Wb .

^c In units of gauss.

^d In units of Mm.

of these fast-CME–hosting ARs with all other ARs, the following conclusions are reached:

1. Of the fast CMEs, 89.5% (51 out of 57) originated from one of our MDI ARs. The number of ARs hosting these fast CMEs is 35, making up 3.1% of all ARs. A majority of these ARs occurred during solar maximum.

2. The distributions of the parameters, except for L_{PIL} , are unimodal, with a peak at a certain middle value. A typical MDI AR has a geometric size of $\sim 10^{3.8}$ Mm² and a magnetic flux of $\sim 10^{14}$ Wb.

3. The derived AR parameters A_T , F_T , B_{avg} , N_{PIL} , and L_{PIL} all show a positive correlation with the probability of an AR's producing a fast CME. When all ARs are equally divided into three groups with low, intermediate, and high values of the study parameters, we find that the majority of fast-CME–hosting ARs reside in the high-value groups; for example, the PIL-related parameters (N_{PIL} and L_{PIL}) are particularly effective at indicating the occurrence of fast CMEs, with a percentage as high as 82.9%. Compared with the low- to middle-value AR groups, the probability of producing a fast CME from a high-value AR can increase by a factor as large as 8. These results suggest that the larger, the stronger, and the more complex an AR, the more likely it is to produce an extremely fast CME.

The ARs hosting fast CMEs make up a small fraction of all ARs ($\sim 3.1\%$). The fraction is about 8% even for just the high-value AR groups. This indicates that there are many ARs that have large parameter values but never produce an extremely fast CME.

The current method is based on photospheric magnetic field observations only. However, we believe that coronal magnetic fields play a direct role in determining the dynamic and kinematic properties of CMEs. In a statistical study of 99 halo CMEs, Liu (2007) suggested that CMEs under unidirectional open field structures are significantly faster than those under the heliospheric current sheet. To develop a better capability of predicting the occurrence of fast CMEs, one needs to study the AR parameters from both the photosphere and the corona as well.

We acknowledge the use of solar data from the MDI, LASCO, and EIT instruments on board the *SOHO* spacecraft. The *SOHO*/LASCO data used here are produced by a consortium of the Naval Research Laboratory (US), the Max-Planck-Institut für Aeronomie (Germany), the Laboratoire d'Astronomie (France), and the University of Birmingham (UK). *SOHO* is a project of international cooperation between ESA and NASA. We also acknowledge the use of the CME catalog generated and maintained at the CDAW Data Center by NASA and the Catholic University of America in cooperation with the Naval Research Laboratory, and the flare list compiled by the Space Weather Prediction Center of NOAA. This work was supported by NSF SHINE grant ATM 04-54612 and NASA grant NNG05GG19G. Y. W. is also supported by grants from the National Natural Science Foundation of China (40525014) and the Ministry of Science and Technology of China (2006CB806304), and J. Z. is also supported by NASA grant NNG04GN36G.

REFERENCES

- Benkhalil, A., Zharkova, V. V., Zharkov, S., & Ipson, S. 2006, *Sol. Phys.*, 235, 87
- Brandt, P. N., & Steinegger, M. 1998, *Sol. Phys.*, 177, 287
- Brueckner, G. E., et al. 1995, *Sol. Phys.*, 162, 357
- Delaboudinière, J.-P., et al. 1995, *Sol. Phys.*, 162, 291
- Falconer, D. A., Moore, R. L., & Gary, G. A. 2006, *ApJ*, 644, 1258
- Feynman, J., & Martin, S. F. 1995, *J. Geophys. Res.*, 100, 3355
- Georgoulis, M. K., & Rust, D. M. 2007, *ApJ*, 661, L109
- Green, L. M., Démoulin, P., Mandrini, C. H., & van Driel-Gesztelyi, L. 2003, *Sol. Phys.*, 215, 307
- Guo, J., Zhang, H.-Q., & Chumak, O. V. 2007, *A&A*, 462, 1121
- Harrison, R. A. 1995, *A&A*, 304, 585
- . 2003, *Adv. Space Res.*, 32, 2425
- Harvey, K. L., & Cornelis, Z. 1993, *Sol. Phys.*, 148, 85
- Howard, R. F. 1996, *ARA&A*, 34, 75
- Jing, J., Song, H., Abramenko, V., Tan, C., & Wang, H. 2006, *ApJ*, 644, 1273 (erratum 652, 1796)
- Lara, A., Gopalswamy, N., & DeForest, C. 2000, *Geophys. Res. Lett.*, 27, 1435 (erratum 27, 1863)
- Leka, K. D., & Barnes, G. 2003, *ApJ*, 595, 1296
- . 2007, *ApJ*, 656, 1173
- Liu, Y. 2007, *ApJ*, 654, L171
- Maeshiro, T., Kusano, K., Yokoyama, T., & Sakurai, T. 2005, *ApJ*, 620, 1069
- McIntosh, P. S. 1990, *Sol. Phys.*, 125, 251
- Preminger, D. G., Walton, S. R., & Chapman, G. A. 2001, *Sol. Phys.*, 202, 53
- Sammis, I., Tang, F., & Zirin, H. 2000, *ApJ*, 540, 583
- Schrijver, C. J. 1988, *A&A*, 189, 163
- . 2007, *ApJ*, 655, L117
- Sterling, A. C., Harra, L. K., & Moore, R. L. 2007, *ApJ*, 669, 1359
- Subramanian, P., & Dere, K. P. 2001, *ApJ*, 561, 372
- Ternullo, M., Contarino, L., Romano, P., & Zuccarello, F. 2006, *Astron. Nachr.*, 327, 36
- Turmon, M., Pap, J. M., & Mukhtar, S. 2002, *ApJ*, 568, 396
- Worden, J. R., White, O. R., & Woods, T. N. 1998, *Sol. Phys.*, 177, 255
- Yashiro, S., Gopalswamy, N., Akiyama, S., Michalek, G., & Howard, R. A. 2005, *J. Geophys. Res.*, 110, No. A12S05
- Zhang, J., Dere, K. P., Howard, R. A., Kundu, M. R., & White, S. M. 2001, *ApJ*, 559, 452