Solar and heliospheric geoeffective disturbances

N.U. Crooker *

Center for Space Physics, Boston University, Boston, MA 02215, USA

Received 27 January 2000; accepted 3 April 2000

Abstract

The past decade has brought advances in several areas of solar-terrestrial physics which, when combined, provide nearly all of the pieces necessary for predicting geomagnetic storms. Advances in techniques for observing the Sun in X-rays and white light allow identification of solar disturbances headed toward Earth. Advances in our understanding of how the resulting heliospheric disturbances reflect aspects of the Sun’s magnetic field allow predictions of their magnetic topology and, hence, provide some measure of the geoeffective southward component which they carry. Advances in our understanding of the relationship between transient heliospheric disturbances and high-speed streams and how storm strength depends upon solar wind density and the magnetic polarity of streams allow substantial refinement for prediction schemes. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Space weather; Geomagnetic storms; Coronal mass ejections

1. Introduction

When on May 14, 1973, Skylab climbed above Earth’s atmosphere in order to view the Sun in X-rays and its corona in white light on a routine basis, it ushered in a new era in understanding the solar causes of geomagnetic storms. In the course of the next 15–20 years, with significant contributions from subsequent missions, it became clear that recurrent storms, those that recur with every 27-day rotation of the Sun, were governed by high-speed streams emanating from coronal holes, observed in X-rays, and that coronal mass ejections (CMEs), rising bubbles observed in the white-light coronagraph, were responsible for major geomagnetic storms (e.g., Hundhausen, 1977; Gosling et al., 1990, 1991).

A comparable period of progress in understanding the causes of storms began roughly a decade ago. The launch of Yohkoh’s soft X-ray telescope (SXT) in 1991 and SOHO’s large angle and spectrometric coronagraph (LASCO) and extreme ultraviolet imaging telescope (EIT) in 1995, instruments reflecting Skylab’s achievements, marked major milestones along the way. This paper reports on those aspects of progress in this decade which place us on the brink of being able to construct a viable empirical model for predicting geomagnetic storms. The next three sections focus on the three kinds of solar wind disturbances which either cause storms or affect storm strength. In order of importance and space allotted to them, these are the two traditional causes of storms, CMEs and high-speed streams, and the newly identified storm enhancers, high-density events. It is important to note at the outset, however, that these disturbances can be closely related to each other and that the largest storms are created by their interactions, as first suggested by Burlaga et al. (1987).

2. Coronal mass ejections

Coronal mass ejections can be geoeffective, in the sense that they can cause geomagnetic storms, primarily because they can bring to Earth strong southward magnetic fields of long duration. Southward fields merge with Earth’s northward fields at the dayside boundary of the magnetosphere and, as a consequence, allow solar wind energy, momentum, and mass access to the magnetosphere.

The speed of a CME also determines how geoeffective it will be, but not because speed itself is particularly
geoeffective (cf. Tsurutani et al., 1992). Speed is a factor in the solar wind electric field, which controls the merging rate at the boundary of the magnetosphere; but its overall contribution to storm strength as an electric field factor is not large because speed varies much less than the other controlling parameter, the strength of the southward magnetic field. CMEs which are faster than the ambient solar wind are more geoeffective primarily because they compress any southward fields in the vicinity of their leading edges.

Coronal mass ejections are not geoeffective because of their mass, in spite of their name, except in a secondary way. By the time CMEs reach Earth, they usually are no more massive than the surrounding solar wind, owing to expansion. Although high-density features often accompany CMEs, it is only recently that high density has been found to increase storm strength, and this is a modulating rather than primary effect (see Section 4).

Given that CMEs are responsible for the most geoeffective solar wind disturbances, an ultimate goal of solar-terrestrial physics is to be able to predict when a CME will arrive at Earth and whether or not it will be geoeffective. This section reports on present capabilities. Sections 2.1 and 2.2 describe solar and interplanetary signatures of CMEs, respectively, and illustrate the geomagnetic response. Section 2.3 reviews what we have learned about the magnetic topology of CMEs relative to the solar magnetic field and how that knowledge is beginning to be used to predict geoeffectivity. Further information on CMEs can be found in reviews by Webb (1995), Gosling (1997), Hundhausen (1997), and Forbes (2000).

2.1. Solar signatures

Although coronal mass ejections rising from the solar limb have been observed directly and routinely with white-light coronagraphs since the early 1970s, the ability to detect those directed at Earth is a recent development. Here we describe two solar signatures of Earth-directed CMEs which have gained wide attention, summarize several others, and illustrate how, in combination, these signatures are being tested as predictors of geomagnetic storms.

The first signature of Earth-directed CMEs came with Skylab’s ability to detect arcade events in X-ray images of the solar disk (e.g., Webb et al., 1976; Kahler, 1977; Sheeley et al., 1983); but it was the launch of Yohkoh’s SXT, with its high cadence appropriate for movie-making, which brought these images to the attention of the broader community (e.g., Tsuneta et al., 1992; McAllister et al., 1996). Arcade events are sudden brightenings of what are presumed to be magnetic arcades reforming through magnetic reconnection under CMEs as they leave the Sun (e.g., Pneuman and Kopp, 1971; Svestka and Cliver, 1992; Hiei et al., 1993; Forbes, 2000). Thus arcade formations are post-eruptive events. In traditional X-ray detectors at geostationary orbit, which measure whole-disk X-ray emission as a function of time, arcade events correspond to flares called “long-duration events” (LDEs), lasting over periods ranging from hours to days. To be detected there, however, they must be bright enough to rise above background emissions from the ubiquitous solar flare activity elsewhere on the Sun. The Yohkoh SXT provided a major advance in LDE detection, being able to resolve faint arcade events independent of routine flaring elsewhere.

Examples of three arcade events are shown in Fig. 1. They appear in a single image from Yohkoh, framed by superposed dashed rectangles. The array of faint arches in the lower right corner is the most spectacular type of arcade event owing to its size. Arcades like these arch over neutral lines that form channels for “polar crown” filaments, so-called because they border the dark (in X-rays) polar coronal holes, away from active regions. When one of these filaments erupts with CME liftoff, the associated arcade event spans large areas of the disk, as much as from limb to limb in a few rare cases (e.g., McAllister et al., 1996). The other two arcade events in Fig. 1 are smaller and brighter. Compact arcades like these are associated with active regions and, often, constitute major flares. Although in this single image the compact arcades look similar to the other bright regions, which are active regions, they can be clearly distinguished in movies by their temporal behavior. Also, in the figure, they are somewhat distinguished by their shape. The event in the top rectangle has cusped loops, presumably reflecting the x-line geometry of reconnection. Other observations suggest that the two kinds of arcade events in Fig. 1, a large-scale event associated with a polar crown...
filament and compact events associated with active regions, are connected by a continuum of events in which size depends inversely upon local magnetic field intensities, where intensity is highest at lower latitudes near active regions and weakest at higher latitudes near the polar crown (A. McAllister, private communication, 1999).

For obvious reasons, arcade events were hailed as a major improvement over traditional solar signatures of impending storms observed lower in the solar atmosphere in Hz light: flares energetic enough to be observed at that wavelength and disappearing filaments (e.g., Joselyn and McIntosh, 1981). Arcades are widely believed to be signatures of CMEs themselves rather than signatures of these lower-level associated phenomena, which have considerably smaller scale sizes and which may be located far from the centers of the CMEs (e.g., Gosling, 1993). Furthermore, filament eruptions are assumed to accompany most CMEs, many are difficult to detect (e.g., Webb et al., 1998), and Hz flares accompany only some CMEs. Most flares, even some major ones, occur independent of CMEs (e.g., Feynman and Hundhausen, 1994). While an arcade event is a clear sign that a CME has left the Sun, and its location is a good indicator of whether or not the CME will reach Earth, the ability to see CMEs themselves headed toward Earth, directly in white light on a routine basis, came with the launch of LASCO on SOHO. Called “halo” CMEs, owing to the Sun-centered circular patterns of light which form as the CMEs expand beyond the edge of the occulting disk (which blocks out the bright solar disk), their light is so diffuse compared to light from CMEs observed at the limb that earlier coronagraphs had difficulty resolving their signal (Howard et al., 1982).

Examples of halo CMEs from two well-studied events, on 6 January 1997 (e.g., Burlaga et al., 1998; Webb et al., 1998) and 7 April 1997 (e.g., Berdichevsky et al., 1998), are shown in Figs. 2a and b, respectively. These two events and their associated interplanetary and geomagnetic disturbances will serve as examples throughout the text to illustrate various points. The images in Fig. 2, from the C2 coronagraph, each show a solid circular area marking the occulting disk and a superposed white circle indicating the blocked solar disk. The brightest areas, emanating from the equatorial regions, mark the belt of dense streamers that encircle the Sun, possibly compressed as a result of CME lift off. The streamer belt is the source of most CMEs (e.g., Hundhausen, 1993; McAllister and Hundhausen, 1996), and CME outflows appear brightest when headed away from the east or west limb. In Fig. 2a, the CME material headed toward Earth forms a diffuse ring, or halo, outlined with a black circle. The signature is brightest in the lower right-hand quadrant. Since the halo is off-center, the CME was headed slightly southward as well as Earthward. The relatively uniform signature of the halo CME in Fig. 2a contrasts sharply with the signature in Fig. 2b. There the halo CME is marked by weakly enhanced, structured brightness extending outward from the edge of the occulting disk in all directions, nearly to the edge of the field of view in the southward direction. Unlike Fig. 2a, there is no ring of light marking the outer edge of the CME. Although halo CMEs are more difficult to discern in single printed images compared to movies and series of images, which can be viewed on the web at http://sohowww.nascom.nasa.gov/gallery/LASCO/index.html, Fig. 2 at least illustrates why earlier coronagraphs with resolution poorer than that of LASCO failed to detect them on a routine basis.

One drawback to using halo CMEs as storm predictors is that the signatures look the same for CMEs headed away from Earth (backside events) as for those headed toward Earth (fronside events). This ambiguity can be resolved to some extent by searching for associated surface features expected for frontside cases. A number of these features, in addition to the arcade events described above, have been identified in X-rays and extreme ultraviolet light (e.g., Hudson and Webb, 1997; Thompson et al., 1998; Hudson et al., 1998; Webb et al., 2000). These include sigmoids, dimmings, and EIT waves. Sigmoids are pre-eruptive, bright, S- or inverse-S-shaped features in active regions. Canfield et al. (1999) showed that active regions with sigmoids are more likely to produce arcade events than those without sigmoids, although the time scale for the association was not addressed. Dimmings are dark regions that form near-arcade events, presumably signaling mass loss during CME eruption. They take on a variety of forms, which Hudson and Webb (1997) have attempted to categorize. EIT waves, named after the acronym for the telescope in which they have been observed, are global-scale disturbances that propagate from compact arcade sites in active regions. Probably related to the much rarer Moreton waves observed lower in the solar atmosphere, they provide dramatic images at coronal heights (Thompson et al., 1999). Thus if a halo CME is associated with an arcade event, a sigmoid, a dimming, and/or an EIT wave, and/or with a traditional surface feature, that is, a disappearing filament or a major flare, then it can be classified as Earth-directed, or frontside.

Returning to our example cases, both of the halo CMEs in Fig. 2 were classified as frontside events. In the 6 January 1997 case, the classification was based on marginal surface signatures (Webb et al., 1998, 2000). These were an obscure disappearing filament and a minor arcade event. In contrast, classification of the 7 April 1997 halo CME was based on a nearly complete set of surface signatures (Berdichevsky et al., 1998; Thompson et al., 1999; Webb et al., 2000). These were a compact arcade event, an EIT wave, sigmoid structure, dimming, and a disappearing filament. The differences between the two halo CMEs, both in the degree of complexity of their white light signatures in Fig. 2 and in the intensity of their surface signatures, support the recent claim that there are two types of CMEs (Sheeley et al., 1999). Whether there is any continuum between the two types, as there appears to be between arcade event types, discussed above, remains to be tested.
Fig. 2. White light images of halo CMEs on (a) 6 January 1997, 18:50 UT and (b) 7 April 1997, 15:52 UT, from the large angle spectrometric coronagraph (LASCO) on SOHO, courtesy of O.C. St. Cyr. The halo in (a) is a simple ring, outlined with a black circle, whereas the halo in (b) is highly structured.
Initial analyses of the relation between halo CMEs and geomagnetic storms indicate a high degree of correlation (e.g., Brueckner et al., 1998; Webb et al., 2000). Fig. 3, adapted from Webb et al. (2000), shows halo CME occurrences marked as vertical lines on a 27-day recurrence plot of the Dst index of geomagnetic activity. The vertical lines are solid or dashed according to whether the halo CME was classified as frontside or backside, respectively, and storms dipping to Dst values of $-50$ nT (moderate) or below are marked by triangles. A total of 14 halo CMEs and 12 storms occurred during the period. The most important result is that six of the seven frontside halo CMEs were followed within 5 days by the onset of storms. These include our example cases on 6 January and 7 April. The frontside halo CME that was not followed by a storm was skewed toward the limb and, thus, may have missed Earth. Curiously, three of the four halo CMEs presumed to have occurred on the backside (treating the triplet as one and ignoring the 7 April mixed doublet) also were followed within 5 days by storms. (The 16 April halo CME was too late to be linked with the 17 April storm and too early to be linked with the 22 April storm.) Whether these CMEs were actually frontside events but lacked visible surface signatures, were backside events with global consequences such as expulsions triggered by global waves (e.g., McComas et al., 1991), or were not related to the ensuing storms are open questions. In any case, this leaves three of the 12 storms without any halo-CME associations, or, more conservatively, half of the storms without confirmed Earth-directed halo CME associations. From these results one can conclude that halo CMEs with associated surface features nearly always presage geomagnetic storms but that at least half as many storms occur without halo-CME forewarning.

2.2. Interplanetary signatures

Passage of ejecta from CMEs through the interplanetary medium is usually marked by a number of distinctive signatures, but the degree of variation from event to event and within the events themselves make identification by a single parameter less reliable than identification by a group of parameters (e.g., Zwickl et al., 1983; Gosling, 1990;
Neugebauer and Goldstein, 1997). Nevertheless, single-parameter identifications are convenient, especially for statistical studies. The most reliable single-parameter signature of ejecta may be a cosmic ray depression (e.g., Richardson, 1997). Cosmic ray depressions occur because cosmic rays travel along magnetic field lines and, consequently, do not have access to the magnetically closed field lines which constitute most CMEs.

On the other hand, the most widely used single-parameter signature of ejecta seems to be the occurrence of counterstreaming (or bidirectional) suprathermal electrons (e.g., Gosling et al., 1987), illustrated in Fig. 4. Since suprathermal electrons carry electron heat flux away from the Sun along magnetic field lines, indicated by the broad arrows, when found streaming in both directions along the field, they are interpreted as signatures of closed magnetic field lines. Fig. 4 shows the most likely closed configuration, with both ends connected to the Sun.

The most widely used compound signature of ejecta is a magnetic cloud (e.g., Burlaga, 1991). Magnetic clouds are defined as large-scale rotations in the magnetic field accompanied by low ion temperatures and strong magnetic fields. Since all three criteria must be met to qualify as a cloud, clouds form perhaps the most pristine subset of ejecta encountered by spacecraft (Gosling et al., 1990; Burlaga, 1991). The magnetic field data from clouds usually provide good fits to flux rope models and, thus, are usually assumed to take that form (e.g., Burlaga, 1991). Flux ropes contain helical fields with increasing pitch from core to boundary, as sketched in Fig. 4.

The relationship between these two widely used ejecta signatures, counterstreaming electrons and magnetic clouds, is usually more complicated than pictured in Fig. 4. The two signatures should be coincident if both ends of all field lines in the flux rope loop constituting the magnetic cloud are connected to the Sun, as implied. A number of studies, however, show that often this is not the case. Counterstreaming electrons extend well beyond the boundaries of some magnetic clouds, implying that those clouds are parts of larger ejecta (e.g., Bothmer et al., 1996; Crooker et al., 1998). Also, counterstreaming can be intermittent within clouds, implying that while flux rope structure can be intact locally, some field lines can open remotely through reconnection (Gosling et al., 1995; Larson et al., 1997; Crooker et al., 1998b; Shodhan et al., 2000).

Interplanetary signatures of the ejecta from our example halo CME cases are shown in Fig. 5. The 6 January 1997 CME produced a clear magnetic cloud signature on 10 January near Earth, which provided a good fit to a flux rope model (see Burlaga et al., 1998, for detailed discussion). Evidence of the characteristic field rotation is shown in the fifth panel of Fig. 5a, where the north-south component of the interplanetary magnetic field (IMF) changes gradually from strongly southward to strongly northward in the shaded interval identifying the cloud. The other two defining characteristics, low temperature (proton thermal speed) and high magnetic field strength, are evident in third and fourth panels. Counterstreaming electrons are indicated by the hatched bars at the top of each panel (data available on the web, courtesy of R. Lin and D. Larson). In this case they are confined to the cloud, as expected from the sketch in Fig. 4; but significant gaps are present, signifying a mix of open and closed field lines. A more detailed analysis of the electron data suggests that some of these field lines may actually be detached from the Sun at both ends (Larson et al., 2000). The top panel of Fig. 5a shows that the cloud’s speed, about 450 km/s, was moderate by absolute standards but faster than the ambient medium ahead of it. As a result, a shock preceded the cloud at 0100 UT on 10 January. The shock caused the increases in speed, density, thermal speed, and field magnitude in the top four panels, respectively.

The bottom panel of Fig. 5a shows the magnetosphere’s response to the cloud in the form of the Dst index. Dst decreases with increasing ring current strength, a measure of storm strength, and increases with increasing magnetopause shielding currents, a measure of magnetospheric compression produced by an increase in solar wind dynamic pressure (proportional to the product of density and the square of the speed). Dst rises in response to the shock compression, then decreases in response to the southward IMF. The storm recovery, usually regarded as primarily a magnetospheric process, is obscured in this case by solar wind compression created by an extraordinarily high-density feature near the end of the cloud, thought to be composed of prominence material from the disappearing filament (Burlaga et al., 1998). More about the effects of this density feature and the following high-speed flow is discussed in Sections 3 and 4.
Fig. 5a shows the interplanetary signatures of ejecta from
the 7 April halo CME. No magnetic cloud has been
documented for this case, owing to the lack of any gradual field
rotation; but the shaded interval labeled “cloud-like ejecta”
has the other two cloud-defining characteristics, low tem-
perature (thermal speed, third panel) and high field strength
(fourth panel), as well as the commonly found lack of field
fluctuations in clouds (cf. Berdichevsky et al., 1998). It is
thus possible that the cloud-like interval marks some skimming
passage through a flux rope structure. As in Fig. 5a,
counterstreaming electrons are indicated by hatched bars. In
contrast to Fig. 5a, they are continuous (with some minor
exceptions which are not shown), and they are not confined
to the cloud-like structure. Counterstreaming began about 9
h prior to the start of the cloud-like structure and ended more
than 4 days later (at 0330 UT on 14 April), well beyond the
end of the plot. The fifth panel shows that the southward IMF responsible for the decrease in Dst in the bottom panel occurred not in the cloud-like interval, where the field was steady and strongly northward, but in the ejecta preceding it. These ejecta may be the result of a skimming encounter with another CME observed earlier on 7 April, off the southwest limb of the Sun (cf. Berdichevsky et al., 1998). The rise in Dst preceding the storm decrease reflects compression created by the higher speed of the ejecta compared to the ambient wind ahead of it, as seen in the top panel of Fig. 5a. In this case, unlike the January case, the speed rise occurred gradually rather than at a shock, although a weak shock does mark the start of the rise (at 1300 UT on 10 April) (Berdichevsky et al., 1998).

In summary, Fig. 5 shows that the interplanetary ejecta signatures for our two example cases differ in much the same way as do their halo and surface signatures. The pristine magnetic cloud in Fig. 5a seems consistent with the simple halo in Fig. 2a and the lack of energetic surface activity, whereas the complex ejecta in Fig. 5b seem consistent with the structured halo in Fig. 2b and the full complement of energetic surface features that accompanied it. These events were not chosen for consistency, so that its existence suggests that the degree of complexity of interplanetary signatures may to some extent be predicted from solar signatures.

2.3. Solar imprint on magnetic topology

Within the last decade or so, our understanding has advanced significantly with regard to the degree to which the magnetic topology of the Sun, CMEs, their ejecta in interplanetary space, and the heliosphere are interrelated. Contrary to a common tendency for issues to become more complicated with further observations, in this case, issues have become simpler. It has become increasingly clear that CMEs tend to carry with them the imprint of the lower harmonics of the Sun’s magnetic field and that, as a result, their magnetic structure tends to blend into the magnetic structure of the heliosphere. As will be illustrated below, these findings have important implications for predicting the geoeffectiveness of CMEs. This section reports on preliminary findings, which are still in the process of being verified.

Fig. 6 provides a simplified schematic view of solar topology for the purpose of placing the findings in a larger context as they are reviewed. The figure shows dipolar field lines which form an arcade from whose apex stems the heliospheric current sheet (HCS). Together the arcade and HCS constitute the skeleton of the helmet streamer belt which enircles the Sun and extends out into the heliosphere. North and south of the streamer belt lie the open field line regions of the polar coronal holes, marked ‘+’ and ‘−’. The heavy curve is the projection of the HCS onto the solar surface, indicating the heliomagnetic equator.

The two short, heavy arrows in Fig. 6 mark the locations of CMEs forming under the umbrella of the helmet streamer belt. Each is assumed to be taking the form of a flux rope in which arcade field lines will become the outermost coils through reconnection behind the CME as it lifts off the Sun (e.g., Gosling, 1990; Crooker et al., 1998a). (It is this reconnection which presumably creates the arcade events seen in X-rays.) The direction and tilt of each arrow represents the orientation of the projected axis of the flux rope, which is the magnetic field line threading its core. For example, the axis of the flux rope in Fig. 4 is the heaviest curve in the center of the loop. Projected back to the Sun, from the perspective of Fig. 6 it would become a right-pointing arrow along the heliomagnetic equator in the center of the figure, with its head below the equator, consistent with the sense of the solar dipolar field. The directions of the two arrows in Fig. 6 determine the chirality (sense of helicity) of the resulting flux rope, given the overall polarity of the dipolar field, away from the Sun in the north and toward the Sun in the south, in this case. The axis fields are drawn to produce a left-handed rope in the northern hemisphere and a right-handed rope in the southern hemisphere in order to match a hemispherically dependent magnetic field pattern associated with filaments lower in the solar atmosphere. This key point is discussed further below. In addition, as mentioned for the projection of the Fig. 4 axis, each axis is tilted slightly with respect to the heliomagnetic equator in order to match the Sun’s dipolar field. Thus, Fig. 6 indicates how a CME flux rope might blend into the Sun’s topological configuration, with its axis aligned with the heliomagnetic equator and its leading fields in the direction of the Sun’s dipolar fields. The following describes results supporting this view.

Marubashi (1986, 1997) made one of the first connections between interplanetary and solar aspects of CMEs. He fit a cylindrical flux rope model to magnetic clouds and showed that the deduced tilts of the rope axes with respect to the ecliptic plane were correlated with the tilt angles of the associated erupting filaments. This pattern was quantified
by Zhao and Hoeksema (1997, 1998), who extended the Marubashi studies and showed that the tilt angles of cloud axes are proportional to those of filaments by a factor of 0.7 (±18°). In other words, filaments have somewhat higher inclinations. Although erupting filaments are usually assumed to occupy only a small volume on the trailing edge of magnetic clouds (cf. Fig. 5a) (e.g., Illing and Hundhausen, 1986; Gopalswamy et al., 1998), and the topological relationship between filaments and clouds is not well understood (see below), the above results can be understood in terms of the fact that the heliomagnetic equator maps down to the neutral line in the chromosphere along which filaments form (cf. Section 2.1). Under the assumption that cloud axes align with the heliomagnetic equator, as pictured in Fig. 6, the higher inclinations of filaments would arise from the higher harmonics of the solar magnetic fields in the chromosphere. Zhao and Hoeksema (1997, 1998) also recognized that knowledge of the tilt angle of a cloud axis implies knowledge of southward IMF in that cloud. For example, if a cloud is a perfectly cylindrical flux rope with its axis pointing directly northward, it will contain no southward IMF; if it lies in the ecliptic plane, half of its fields will have southward IMF; and if it points directly southward, the IMF will be southward throughout. Zhao and Hoeksema (1997, 1998) confirmed this pattern by showing a correlation between tilt angles for clouds fit with flux ropes and the duration and maximum intensity of southward IMF in those clouds. The derived relationships between these parameters combined with the above relationship between the tilt angles of cloud axes and erupting filaments provides a predictive formula for southward IMF in a magnetic cloud based on the orientation of the associated filament. For our example case on 6 January 1997, even though the erupting filament was obscure, Zhao and Hoeksema (1997) measured its tilt angle and obtained a maximum southward IMF strength of −13 ∓ 5 nT and a duration of 14 ± 5 h from the predictive formula. These values compare remarkably well with the in situ observations of −15 nT and 13 h.

The alignment of cloud axes with the heliomagnetic equator has been indirectly confirmed with observations by Mulligan et al. (1998). They demonstrated that there is a solar cycle variation in the orientation of cloud axes, with high inclinations dominating during solar maximum and the declining phase and low inclinations dominating during solar minimum and the ascending phase. This pattern approximately matches the well-known solar cycle variation of HCS inclination (e.g., Hoeksema, 1991; Shodhan et al., 1994), consistent with Fig. 6.

As mentioned above, from Fig. 6 one can see that the direction of the leading field in a magnetic cloud, made up of the large-scale arcade fields, should point opposite the direction of the solar dipole. Since the dipolar field changes polarity at the maximum of every solar cycle, it follows that the direction of the leading field in clouds should change at solar maximum, as well. Bothmer and Rust (1997) and Bothmer and Schwenn (1998) have documented roughly that pattern. About 80% of magnetic clouds during 1974–1981 had southward leading fields and during 1982–1991 had northward leading fields. The pattern was also confirmed by Mulligan et al. (1998).

The cloud chirality dependence on hemisphere illustrated in Fig. 6 has been demonstrated statistically by Rust (1994) and Bothmer and Schwenn (1994, 1998). By identifying the cloud flux rope as the filament itself, Rust (1994) attributed the dependence to the fact that filament skew patterns, fine structure resembling the threads on screws, are statistically sorted by hemisphere (Martin et al., 1994) and that the skew patterns imply that filaments are flux ropes with left chirality in the northern hemisphere and right in the southern hemisphere. Later Martin and McAllister (1997) argued that if filaments are flux ropes at all, their chirality is opposite to that assumed by Rust (1994), that magnetic clouds are composed of the much larger-scale coronal arcade fields, as illustrated in Fig. 6, and that the skew pattern of the coronal fields, for which they found a one-to-one correlation with the skew pattern of the underlying filaments, unambiguously determines cloud chirality. Irrespective of these opposing interpretations of filament skew patterns, which hinge on whether the fine structure extending from filaments is concave upward or downward, all studies agree that cloud chirality is well correlated with the hemisphere of origin of the associated erupting filament, as illustrated in Fig. 6, and can be predicted precisely from observations of the filament or arcade skew pattern, even for the occasional cases where the hemispheric pattern does not apply (McAllister et al., 1998).

Further support for the Fig. 6 view has been found in interplanetary observations of the relationship between sector structure and ejecta from CMEs. Contrary to what might be expected, most ejecta do not disrupt sector structure (Zhao and Hoeksema, 1996; Kahler et al., 1999). That is, they do not impose large intervals of random polarity on what otherwise would be a relatively stable two- or four-sector polarity pattern. Instead they blend into the sector structure, consistent with a solar imprint on their topology. Fig. 4 illustrates this point. The flux rope loop constituting the magnetic cloud in essence carries the sector boundary (cf. Crooker et al., 1998a). The leading leg has fields pointing toward the Sun (toward polarity), and the trailing leg has fields pointing away from the Sun (away polarity). Under the assumption that the rope axis is tilted with respect to the ecliptic plane, such that its axis would project back to the heliomagnetic equator at the center of Fig. 6, as discussed above, then a spacecraft being overtaken by the leading edge of the flux rope loop would first sense toward polarity, then the rotating fields of the cloud, and then away polarity. The cloud itself thus becomes the sector boundary and can be considered as a bulge in the heliospheric current sheet (Crooker and Intriligator, 1996).

In summary, recent studies with limited data sets have demonstrated that CMEs carry the imprint of the solar magnetic field out into the heliosphere. These findings imply that many characteristics of ejecta from CMEs can be
predicted, at least statistically, from characteristics of the solar field and of associated solar source features. Potentially predictable characteristics include the orientation of magnetic cloud axes and their leading fields, the chirality of their helical fields, and the maximum strength and duration of their southward IMF.

3. High-speed streams

This section focuses on the second of the two traditional causes of geomagnetic storms, high-speed streams. Three aspects of high-speed streams are emphasized here: their ability to compress, their impact on CMEs, and the dependence of their geoeffectiveness on the polarity of their magnetic fields.

A common misunderstanding about high-speed streams is that the high-speed flow itself causes geomagnetic storms. This misunderstanding was exacerbated during the Skylab era when coronal holes were hailed as the mysterious M-regions which cause recurrent storms (e.g., Bohlin, 1977). Earlier studies had shown that peak storm strength coincides with passage of the leading edge of a high-speed stream, not with the subsequent high-speed flows, and ascribed this pattern to compression of pre-existing southward IMF at the leading edge (e.g., Hirshberg and Colburn, 1973). More recent reviews of the earlier studies have attempted to rectify the misunderstanding in terms of this explanation (Crooker and Cliver, 1994; Tsurutani et al., 1995).

More important for space weather, however, is the recent recognition that CMEs often make the biggest contribution to peak storm strength at the leading edges of high-speed streams (Crooker and Cliver, 1994; Crooker and McAllister, 1997). The reason for this is primarily geometrical. Since CMEs arise from the closed field line regions of the Sun, as pictured in Fig. 6, they travel through the slow flows associated with the streamer belt, which border fast flows from coronal holes. For a tilted streamer belt which gives high-speed flows access to the ecliptic plane, a spacecraft located there will observe any CMEs in the slow flow immediately preceding the leading edge of a high-speed stream.

An ecliptic cross-section of this stream pattern is illustrated in Fig. 7. Ejecta from a CME, labeled “ICME”, for interplanetary CME, is expanding and traveling out into the heliosphere in the corridor of slow flow surrounding the HCS. Behind the slow flow is fast flow from a coronal hole. The field line with away polarity is meant to lie in the fast flow, near its leading edge. Its Parker spiral is more loosely wound, consistent with higher speed. Although all of the magnetic structures conform to the Parker spiral, the velocity of each parcel of solar wind is directed radially outward. Thus the fast flow pushes into the slow flow radially ahead of it and compresses it, creating a corotating interaction region (CIR) (see Gosling and Pizzo, 1999 for a tutorial on CIR structure). The CIR immediately follows the ICME so that southward IMF both in the ICME and in the CIR contribute to storm strength.

In some cases, the ICME itself can be caught up in the CIR so that any southward fields in its trailing section intensify from compression (Crooker and Cliver, 1994; Odstrcil and Pizzo, 1999a). Fenrich and Luhmann (1998) demonstrated this effect for selected cases and pointed out that the effect will have a solar cycle variation: As discussed in Section 2.3, magnetic clouds with leading northward fields and trailing southward fields, favorable for the effect, occur between those solar maxima when the Sun’s dipolar field is opposite to that shown in Fig. 6.

The storm on 10–11 April 1997 from our example case and the following storm on 16–17 April both provide profiles characteristic of high-speed streams. Fig. 8 shows the Dst index for the two storms, replotted from the fifth panel of Fig. 3, along with a plot of flow speeds. As described above, peak storm strength in both cases coincides with the leading edge of a corresponding high-speed stream. For these cases, however, peak speeds reach only moderate levels (~ 550 km s⁻¹), and the streams are not recurrent, as is apparent from Fig. 3. These features are consistent with their occurrence in the ascending phase of the solar cycle, when the solar field varies considerably from rotation to rotation, compared to the declining phase, when recurrent storms reflect the stability of the solar field.

Although the Dst and speed profiles for the two storms in Fig. 8 are similar, there are substantial differences in their causes. From the analysis in Section 2, we know that the peak...
of the first storm was caused by ejecta from a CME. In this case, the southward IMF responsible for the storm occurred near the leading edge of the CME and, consequently, was not compressed by the subsequent high-speed flow. This is evident in the higher time resolution plot in Fig. 5b, where the declining field magnitude in the fourth panel indicates no compression. Moreover, the high-speed flow in this case may have been part of the ejecta rather than flow from a coronal hole, as evidenced by the counterstreaming electrons observed until early on 14 April. In contrast, no evidence of ejecta can be found in the solar wind parameters for the second storm (not shown), nor were any signatures of CMEs found in the corresponding solar data (Webb et al., 2000). The southward IMF responsible for the storm had a spiky profile and occurred within a CIR created by the oncoming higher-speed flow. Thus the pattern for this storm fits the traditional, steady-state explanation for storms created by high-speed streams (cf. Lindsay et al., 1995).

A secondary feature of a storm created by a high-speed stream is the sustained, low-level activity which follows peak activity. It is only this portion of the storm which is caused by the high-speed flow itself, and not all high-speed flows generate sustained activity. The southward IMF responsible for sustained activity has been attributed to the ubiquitous Alfvénic fluctuations in flow from coronal holes (Tsurutani and Gonzalez, 1987), but these can be ineffective owing to the Russell–McPherson effect (Crooker and Siscoe, 1986; Crooker and Cliver, 1994). The Russell–McPherson effect is a projection effect which depends upon season and IMF polarity (Russell and McPherson, 1973). The IMF, ordered in the Sun’s heliographic coordinate system, projects a southward component in Earth’s dipole-ordered coordinate system when it points toward the Sun during spring and away from the Sun during fall (northern hemisphere). High-speed streams cause significant sustained activity only when this polarity/season-dependent southward IMF adds to the effects of the Alfvénic fluctuations.

The Russell–McPherson effect on sustained activity is evident in Fig. 8, even though the first stream may be transient rather than coronal hole flow. The two dashed lines running through the Dst trace indicate that sustained activity for the second storm was more than twice as strong as for the first storm. Since the polarities of the two corresponding streams were away in the first and toward in the second, and the storms occurred in spring, the difference in sustained activity levels is as predicted by the Russell–McPherson effect.

4. High-density events

For many years, high density in the solar wind was thought to affect geomagnetic storms only insofar as it increases dynamic pressure. The most pronounced effect of increased dynamic pressure is a positive excursion of Dst caused by magnetopause currents, counter to the negative excursion caused by the ring current which measures storm strength, as discussed in Section 2.2. In phase with the ring current excursion, however, Murayama (1982) found a correlation between negative Dst and dynamic pressure, which was recently confirmed by Fenrich and Luhmann (1998); but this dependence of storm strength on dynamic pressure is weak. Nevertheless, Fenrich and Luhmann (1998) point out that had the IMF been southward during passage of our January 1997 example cloud, the increased dynamic pressure owing to the dramatic increase in density on the trailing edge of the cloud, combined with compression from the oncoming high-speed stream, would have doubled the strength of the ensuing storm.

A breakthrough in understanding the effects of solar wind density itself on storm strength came with the realization that the magnetosphere’s response to a change in density is a change in plasma sheet density, for which the time scale is much longer (~5 h) than the response to southward IMF (<1 h) (Smith et al., 1999, and references therein). Further, high density is geoeffective only when the IMF points southward. Thomsen et al. (1998), like Fenrich and Luhmann (1998), point out that our January 1997 example cloud would have been much more geoeffective had the extremely high density on its trailing edge occurred when the IMF was southward, but on a different time scale than the nearly instantaneous response to dynamic pressure and for another reason — an increase in plasma sheet density and, consequently, an increase in ring current density. These findings establish high density as a parameter of more than marginal impact on storm strength.

Fig. 7 illustrates where high-density regions can be found in the solar wind. As mentioned earlier, an ICME itself is a low- or ambient-density feature, but it is associated
with high-density features of both interplanetary and solar origin. The interplanetary features are the two compression regions. The first is the preceding sheath, which forms if the ICME is faster than ambient. If the speed difference exceeds the magnetosonic speed, a shock wave will form, indicated by the hatched line. The second compression region is the trailing CIR. CIRs, of course, form independent of ICMEs, but the high densities of those associated with ICMEs are more likely to be geoeffectve owing to the higher probability of the required concurrent southward IMF, especially during alternate solar cycles, as described above.

The high-density features of solar origin are the slow flow, in general, and the smaller-scale pressure balance structures embedded in the slow flow. These are indicated by shaded, irregular, elongated forms in Fig. 7, sketched to resemble bright, field-aligned features in coronagraphs. Pressure balance structures, as their name indicates, are in pressure balance with their immediate surroundings so that they neither expand nor contract. An increase in gas pressure from an increase in density is balanced by a decrease in either temperature or magnetic field pressure. Shodhan et al. (1999) showed that high-density pressure balance structures tend to become the highest-density features in the solar wind when they become compressed in sheaths or CIRs. Some of the most pronounced pressure balance structures lie within the trailing portions of ICMEs, as in our January 1997 example case, and carry composition signatures which tag them as remnants of the associated erupting filament (e.g., Burlaga et al., 1998; Gopalswamy et al., 1998). High-density pressure balance structures also encase the heliospheric current sheet, in which case they are called “heliospheric plasma sheets” (Winterhalter et al., 1994); but these tend not to be associated with strong southward IMF and, hence, are not particularly geoeffectve.

High-density pressure balance structures can be seen in the second panel of Fig. 5b, during passage of our April 1997 example ICME. Consistent with Shodhan et al. (1999), the two highest-density peaks occurred in the leading and trailing compression regions, indicated by the rising speeds there. It is possible that the trailing peak, which coincided with a brief southward excursion of the IMF, was responsible for the depth of the second dip in Dst. High density also may have been the ultimate cause of why the diffuse aurora over Boston were spectacularly visible during this event. Normally diffuse aurora are subvisual there, but in this case they were visible owing to high proton flux (M. Mendillo, private communication, 1999). The high proton flux, in turn, may have been the result of high plasma sheet density created by the high solar wind density.

5. Conclusions

In describing what makes solar and heliospheric disturbances geoeffectve, in the sense that they cause geomagnetic storms, the key words are “southward IMF” and “compression”. Southward IMF is important because it breaches Earth’s solar wind shield, and compression is important primarily because it strengthens any southward IMF and secondarily because it increases density. Coronal mass ejections, by far the most geoeffectve structures, usually have intrinsic southward IMF of long duration and, if faster than ambient, compress any southward IMF at their leading edges and in the preceding sheaths created by the speed difference. (In addition, they create southward IMF by distorting the ambient medium (e.g., Odstrcil and Pizzo, 1999b, and references therein), a topic not covered by this review.) High-speed streams are geoeffectve primarily because they compress any southward IMF in CIRs. This compression can be especially geoeffectve when CIRs encroach on CMEs. The secondary, sustained geomagnetic activity caused by southward IMF excursions in Alfvenic fluctuations in the high-speed flow itself reaches significant levels only when supplemented by southward IMF from the Russell–McPherron polarity effect. High solar wind density structures, often found preceding and trailing ICMEs, are geoeffectve only when coincident with southward IMF.

The prerequisite tools and basic scientific understanding for predicting geoeffectve heliospheric disturbances from solar signatures have improved dramatically over the past decade. CMEs headed toward Earth can be identified, and topological parameters relevant to the intensity and duration of their intrinsic southward IMF can be predicted. The major missing elements in prediction schemes are knowledge of CME speed, to determine if compression of southward IMF will play a role, and of transit time from Sun to Earth, which depends not only upon speed but upon acceleration or deceleration. These should be provided in the near future with the launch of the STEREO mission.

Acknowledgements

This paper was presented as a tutorial lecture at the Joint CEDAR-GEM-SHINE Workshop on Solar-Terrestrial Coupling Processes in Boulder, Colorado, June 1999. The author thanks D. Webb and A. McAllister for helpful discussions. Research was supported by the National Science Foundation under grants ATM-9805064 and ATM-9819483.

References


