Forecast evaluation of the coronal mass ejection (CME) geoeffectiveness using halo CMEs from 1997 to 2003

R.-S. Kim,[1,2] K.-S. Cho,[3] Y.-J. Moon,[3] Y.-H. Kim,[3] Y. Yi,[1] M. Dryer,[4,5] Su-Chan Bong,[3] and Y.-D. Park[3]

Received 4 May 2005; revised 4 August 2005; accepted 7 September 2005; published 19 November 2005.

[1] In this study we have made a forecast evaluation of geoeffective coronal mass ejections (CMEs) by using frontside halo CMEs and the magnetospheric ring current index, *Dst*. This is the first time, to our knowledge, that an attempt has been made to construct contingency tables depending on the geoeffectiveness criteria as well as to estimate the probability of CME geoeffectiveness depending on CME location and/or speed. For this, we consider 7742 CMEs observed by SOHO/LASCO and select 305 frontside halo CMEs with their locational information from 1997 to 2003 using SOHO/EIT images and GOES data. To select CME-geomagnetic storm ($Dst < -50$ nT) pairs, we adopt a CME propagation model for estimating the arrival time of each CME at the Earth and then choose the nearest *Dst* minimum value within the window of ±24 hours. For forecast evaluation, we present contingency tables to estimate statistical parameters such as probability of detection yes (PODy) and false alarm ratio (FAR). We examine the probabilities of CME geoeffectiveness according to their locations, speeds, and their combination. From these studies, we find that (1) the total probability of geoeffectiveness for frontside halo CMEs is 40% (121/305); (2) PODys for the location ($L < |50°|$) and the speed (>400 km s^{-1}) are estimated to be larger than 80% but their FARs are about 60%; (3) the most probable areas (or coverage combinations) whose geoeffectiveness fraction is larger than the mean probability (∼40%), are $0° < L < +30°$ for slower speed CMEs (≤800 km s^{-1}), and $-30° < L < +60°$ for faster CMEs (>800 km s^{-1}); (4) when the most probable area is adopted as the new criteria, the PODy becomes slightly lower, but all other statistical parameters such as FAR and bias are significantly improved. Our results can give us some criteria to select geoeffective CMEs with the probability of geoeffectiveness depending on the location, speed, and their combination.

Citation: Kim, R.-S., K.-S. Cho, Y.-J. Moon, Y.-H. Kim, Y. Yi, M. Dryer, S.-C. Bong, and Y.-D. Park (2005), Forecast evaluation of the coronal mass ejection (CME) geoeffectiveness using halo CMEs from 1997 to 2003, *J. Geophys. Res.*, *110*, A11104, doi:10.1029/2005JA011218.

1. Introduction

[2] Coronal mass ejections (CMEs) are huge bubbles of magnetized plasma ejected from the Sun. They propagate through the heliosphere and then interact with geomagnetic fields when they arrive at the Earth. The arrival of their interplanetary counterparts, interplanetary CMEs (ICMEs), implies the possible initiation of a geomagnetic storm if there exist sufficiently large-magnitude southward components of the interplanetary magnetic field [*Dryer*, 1994]. We may refer to the observed near-Sun CME, which results in the geomagnetic storm, as the eventual geoeffective ICME.

[3] The arrival of the CME at the Earth can be predicted by using empirical models. There are several CME propagation models [e.g., *Gopalswamy et al.*, 2001; *Zhang et al.*, 2003]. The model of *Gopalswamy et al.* [2001] uses initial speed (plane-of-sky) and first appearance time of CME to predict the arrival time at 1 AU by assuming the acceleration/deceleration of CMEs as they propagate through the solar wind [*Michalek et al.*, 2004]. The models usually assume that slow CMEs are accelerated and fast CMEs are decelerated toward the ambient solar wind speed (≃400 km s^{-1}). *Cho et al.* [2003] showed that the predictability of *Gopalswamy et al.*'s [2001] empirical model is comparable with those of shock propagation models [*Fry et al.*, 2003;

[1]Department of Astronomy and Space Science, Chungnam National University, Daejeon, South Korea.
[2]Also at Korea Astronomy and Space Science Institute, Division of Space Science, Daejeon, South Korea.
[3]Korea Astronomy and Space Science Institute, Division of Space Science, Daejeon, South Korea.
[4]NOAA Space Environment Center, Boulder, Colorado, USA.
[5]Also at Exploration Physics International, Inc., Huntsville, Alabama, USA.

0148-0227/05/2005JA011218$09.00

Smith et al., 2004, 2005; *Dryer and Smart*, 1984; *Smith and Dryer*, 1990; *Moon et al.*, 2002]. Then they suggested that the CME propagation model can be used for real-time forecasts of geomagnetic disturbances. However, they noted that there would be numerous false alarms on the CME geoeffectiveness if all CMEs are supposed to arrive at the Earth [*St. Cyr et al.*, 2000], even though only halo CMEs are considered. To avoid such false alarms, some criteria to select geoeffective CMEs are required. Therefore a predictability test based on contingency tables [*Smith et al.*, 2000] would be meaningful for the CME prediction.

[4] It is not well understood what kind of physical parameters control the CME geoeffectiveness and what kind of criteria are important for effective forecasting. A frontside halo CME, which appears to envelope the Sun, is known to be more geoeffective because it is directed toward the Earth [*Zhao and Webb*, 2003]. *Cane et al.* [2000] showed that only about half of frontside halo CMEs encountered the Earth, and their associated solar events typically occurred from east 50° to west 50° in heliolongitude. *Wang et al.* [2002] found that only 45% of the total 132 frontside halo CMEs from March 1997 to December 2000 caused geomagnetic storms with $Kp > 5$, and almost 83% of events took place within ±30° of the central meridian. They also found an asymmetry in longitudinal distribution of solar source regions for the CMEs responsible for major geomagnetic storms. *Zhang et al.* [2003] found that 26 of the 27 major geomagnetic storms ($Dst < -100$ nT) were identified with frontside halo or partial halo CMEs. Recently, *Srivastava and Venkatakrishnan* [2004] showed that the CME speeds in the LASCO field of view were roughly correlated with the strength of geomagnetic storms. In this respect, we feel that it is necessary to make a statistical study on the forecast of the geoeffective CMEs depending on CME physical parameters such as location and speed.

[5] In this study, we consider 7742 CMEs observed by SOHO/LASCO (Large Angle Spectroscopic Coronagraph) [*Brueckner et al.*, 1995] from 1997 to 2003 and select 305 frontside halo CMEs. We apply the empirical CME propagation model (mentioned above) to these CMEs in order to estimate their arrivals at the Earth and then select *Dst* minimum values within the window of ±24 hours from the arrival time predicted by the model. As a result, we find 121 CME-geomagnetic storm pairs ($Dst < -50$ nT) from 305 CME-*Dst* minimum pairs. By using the 305 CME-*Dst* pairs, we make contingency tables and evaluate statistical parameters such as probability of detection yes (PODy), false alarm ratio (FAR), and critical success index (CSI).

[6] The paper is organized as follows. Data and methodology of our study are given in section 2. We present results of statistical evaluation of CME geoeffectiveness forecasts in section 3. A brief summary and discussion are derived in section 4.

2. Data and Methodology

2.1. CME Selection

[7] The SOHO/LASCO has much improved instrument capability compared with previous coronagraphs such as low stray-light, low noise levels, and large dynamic range. In this study, we considered 7742 CMEs with their speed information from 1997 and 2003. Measured properties of the CMEs such as their speeds, angular widths, and position angles are well compiled in the CME online catalog (http://cdaw.gsfc.nasa.gov/CME_list/index.html) [*Yashiro et al.*, 2004]. Then we selected 883 halo or partial halo CMEs (11.4%) whose angular width is greater than 120° as noted in the table of the CME catalog.

[8] The identification of frontside CMEs is an important task in the space weather context because the earthward direction of a CME is a necessary condition for the CME to be geoeffective. To select frontside halo CMEs, we carefully compared the SOHO/LASCO images with SOHO/EIT running difference images. EIT provides spectro-heliograms of the corona and transition region on the solar disk and up to 1.5 solar radii above the solar limb. It allows us to diagnose solar plasma at certain temperatures in the range of 6×10^4 to 3×10^6 K [*Delaboudiniere et al.*, 1995]. By determining whether the CME position and EUV features (brightening or ejecting loops) have spatial and temporal closeness, we selected 521 frontside halo CMEs.

[9] To identify the location of a CME, we first looked for the GOES X-ray flare whose starting time is approximately coincident with the extrapolated CME onset time. Then we used the locational information of the flare list observed by the Solar Flare Telescope (SOFT)[*Park et al.*, 1997; *Moon et al.*, 2000] in Korea Astronomy and Space Science Institute (KASI) (http://www.boao.re.kr/yjmoon/flarelist.htm) as well as the flare list compiled by National Geophysical Data Center (NGDC) (ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/FLARES). Otherwise, we identified the location of the active region by examining whether EIT brightening/dimming coincided spatially and temporally with the eruption of a LASCO CME. We finally selected 305 frontside halo CMEs with their locational information.

2.2. Geomagnetic Storm Selection

[10] In this study, we assumed that CMEs are the main driver of geomagnetic storms; it was also assumed that the start time of the storm would be near the arrival time of the CME at the Earth. According to *Gonzalez et al.* [1994], the geomagnetic storms can be defined in terms of their intensity by *Dst* minimum value as follows: (1) weak or minor storm, minimum *Dst* falls between −30 and −50 nT; (2) moderate storms, minimum *Dst* falls between −50 and −100 nT; and (3) strong storms, minimum *Dst* of −100 nT or less. In this study we defined an event having $Dst < -50$ nT as a geomagnetic storm.

[11] We applied the empirical CME propagation model suggested by *Gopalswamy et al.* [2001] to 305 frontside halo CMEs using linear speeds and first appearance times in LASCO C2 or C3 field of view. Although the CME propagation model predicts the arrivals of interplanetary CMEs (ICMEs), our investigation showed that most of the geomagnetic storms occurred near the arrival time of ICMEs [*Cane and Richardson*, 2003]. Thus we adopted the model to look for CME and geomagnetic storm pairs by using the time window of ±24 hours from the predicted CME arrival time at the Earth. Practically, we searched for the minimum value of *Dst* index provided by National Space Science Data Center (NSSDC) (http://nssdc.gsfc.nasa.gov/omniweb/ow.html) within the time window. Figure 1 illustrates how to determine the pairs. The detailed

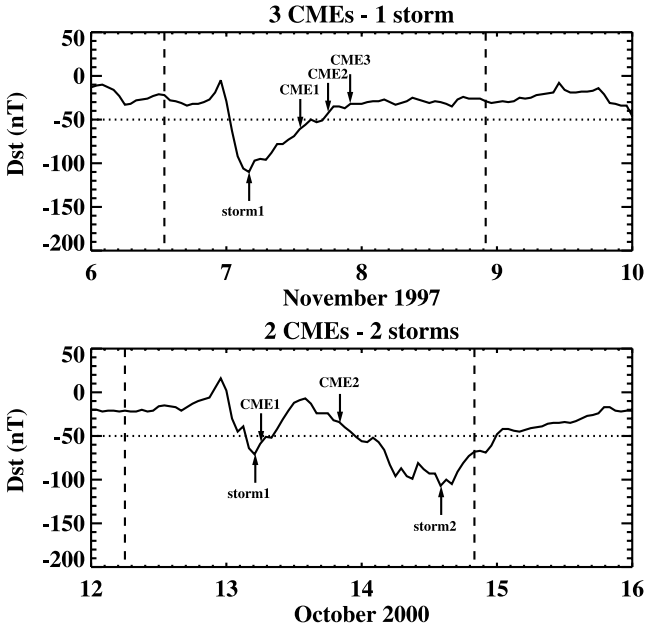

Figure 1. Examples are given to show how to select coronal mass ejection (CME)-storm pairs. Upper panel shows the cases that three CMEs are predicted to arrive at the Earth, but a single geomagnetic storm was detected. Lower panel shows the arrivals (CME1 and CME2) of two CMEs and when two geomagnetic storms (storm1 and storm2) took place. Left vertical dashed lines indicates the negative window (−24 hours) from the first CME arrival and the right vertical dashed line indicates the positive window (+24 hours) from the last CME arrival. Dotted line across the panel indicates the Dst value of −50 nT.

procedure is as follows: (1) if a CME is predicted to arrive at the Earth and there is a geomagnetic storm within the time window, we selected the storm as the counterpart of the CME without any ambiguity; (2) if a CME is predicted and there are multiple storms within the time window, we selected the event having the smallest *Dst* minimum value; (3) if two or more CMEs are predicted and there is a storm within the time window, we selected a CME that has the nearest prediction time from the *Dst* minimum time as shown in upper panel of Figure 1. In this figure, three CMEs are predicted to arrive at the Earth sequentially, but only a single storm occurred. In this case we selected the first CME (CME1) as a driver of the geomagnetic storm (storm1) and the other two CMEs are regarded to be nongeoeffective; (4) if two or more CMEs are predicted

Table 1. Forecast Contingency Table and Verification Statistics[a]

	Observation			
Prediction	Yes	No	Total	Remarks
Yes	a	b	a + b	a = hits, b = false alarm
No	c	d	c + d	c = misses, d = correct nulls
Total	a + c	b + d	a + b + c + d	

[a]PODy (a/(a + c)) is the proportion of yes observations that were correctly forecast. PODn (d/(b + d)) is the proportion of no observations that were correctly forecast. FAR (b/(a + b)) is the proportion of yes predictions that were incorrect. Bias ((a + b)/(a + c)) is the ratio of number of yes forecasts to yes observations. CSI (a/(a + b + c)) is the proportion of hits that were either predicted or observed.

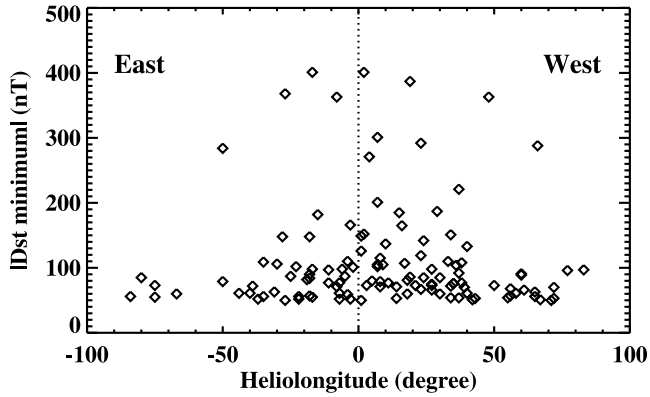

Figure 2. Longitudinal distribution of 121 geoeffective CMEs.

and there are two or more *Dst* minimum value, we selected the CME-storm pairs according to their temporal sequence as follows. As shown in the lower panel of Figure 1, while the first CME arrival (CME1) can be easily paired with the first storm (storm1), the second CME arrival(CME2) is paired with the *Dst* minimum time (storm2). As a result, we found 121 CME-storm pairs. About one-third (41/121) of the events are strong geomagnetic storms characterized by $Dst < -100$ nT.

2.3. Contingency Table

[12] For the statistical evaluation of CME geoeffectiveness, we adopted a contingency table that has been widely used in the meteorological forecasting literature. The table can provide us with the information of the success or failure (or degree thereof) of the forecasting experience in real time [*Smith et al.*, 2000]. Table 1 is a general form of the contingency table. In this table, "a" is the number of hits that, in the table's entry position, means "yes" predicted and "yes" observed. The letter "b" is the number of false alarms that means "yes" predicted but not observed. Similarly, "c" is the number of misses that means not predicted but "yes" observed, and the letter "d" is the number of correct nulls that means not predicted or observed.

[13] The statistics may then be computed from the contingency table. Thus "the probability of detection yes: PODy" is the proportion of "yes" observations that were

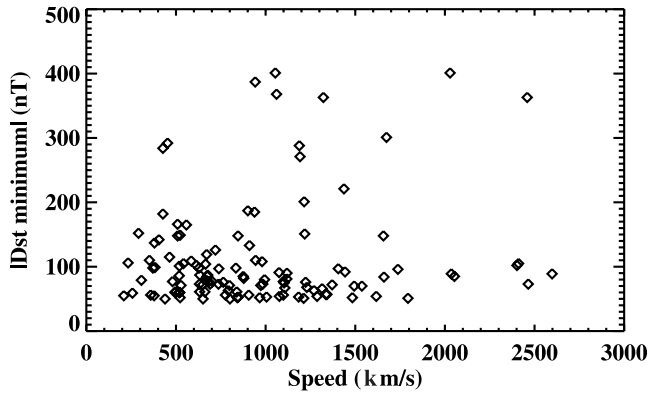

Figure 3. Speed distribution of 121 geoeffective CMEs.

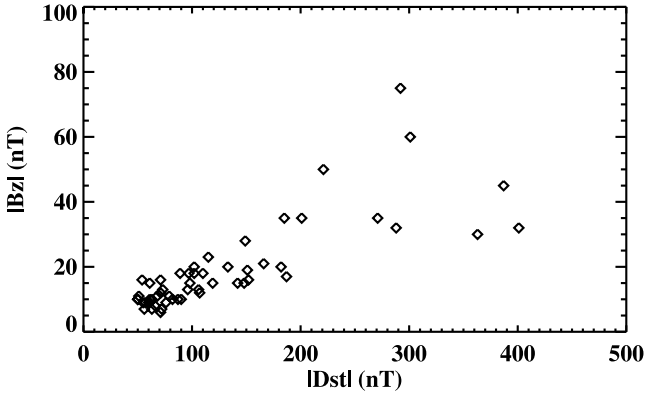

Figure 4. The relationship between *Bz* and *Dst* index using 55 CME-*Dst* pairs which have *Bz* information.

correctly forecast, estimated by a/(a + c). "The probability of detection no: PODn" is the proportion of "no" observations that were correctly forecast, estimated by d/(b + d). "The false alarm ratio: FAR" is the proportion of "yes" predictions that were incorrect, estimated by b/(a + b); and the "Bias" is the ratio of "yes" predictions to "yes" observations, and it is estimated by (a + b)/(a + c). Finally, the "critical success index: CSI" is the proportion of hits that were either predicted or observed, estimated by a/(a + b + c). All of these definitions are given in Table 1. It should be noted that statistical parameters such as PODy and FAR are used for the statistical evaluation of past data but not used for real forecast.

3. Forecast Evaluation

3.1. Dependency on Location or Speed

[14] Among 305 frontside halo CMEs, we found 121 geoeffective CMEs that have corresponding geomagnetic storms characterized by $Dst < -50$ nT. The total probability of CME geoeffectiveness for frontside halo CMEs is about 40%. To improve the forecasting capability, we may need to set up some criteria to select geoeffective CMEs. First, we present the longitudinal distribution of the 121 geoeffective CMEs in Figure 2. It is found that about 80% (96/121) of the CMEs for all geomagnetic storms ($Dst < -50$ nT) and all the CMEs except for one for strong storms ($Dst < -100$ nT) were located within ±50°. Such a distribution is quite similar to that of *Wang et al.* [2002]. The first geoeffective criterion that we set up is that geoeffective CMEs should originate near the central meridian (±50° in heliolongitude).

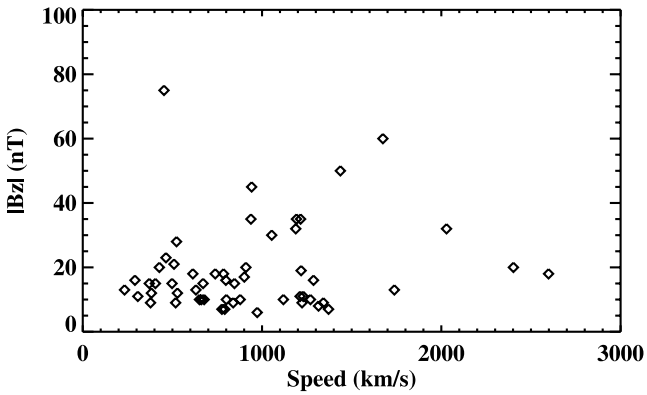

Figure 5. The relationship between *Bz* and CMEs speed using 55 CME-*Dst* pairs which have *Bz* information.

Table 2. Contingency Table Based on the Criteria of the CME Location (L < |50°|) for the 305 Events

	Observation		
Prediction	Yes	No	Total
Yes	96	117	213
No	25	67	92
Total	121	184	305

[15] Second, we consider the linear CME speed (i.e., in the plane-of-sky) determined in LASCO C2 or C3 field of view. Figure 3 shows the CME speed distribution for the 121 events. Even though it is hard to determine a speed criterion for geoeffective CMEs, our second geoeffective criterion is that the CME speed is faster than the mean solar wind speed (400 km s^{-1}).

[16] As shown in the figure, it is noted that most (11/13) of the very strong geomagnetic storms ($Dst < -200$ nT) were caused by very fast CMEs whose speeds are higher than 900 km s^{-1}. We found a weak correlation between CME speed and *Dst* index, which is quite similar to *Srivastava and Venkatakrishnan* [2004], who used 54 events. These results are contrasted with *Yurchyshyn et al.* [2004], who found a good relationship between the CMEs speed and the IMF southward component (*Bz*), as well as an even better relationship between *Bz* and *Dst* index. They used only 14 full halo CMEs, associated with an ejecta whose solar source region was located within 55 degrees from the central meridian. On the other hand, we considered all frontside partial and full halo CMEs (121) that are likely to be associated with geomagnetic storms during the period from 1997 to 2003. Figures 4 and 5 show the relationship between *Bz* and *Dst* index and the relationship between *Bz* and CME's speed using 55 events whose relationship between *Bz* and *Dst* storm is clearly identified. It is shown that *Bz* and *Dst* index has a good relationship but *Bz* and CME's speed shows a worse relationship than that of *Yurchyshyn et al.* [2004]. The major difference may come from the number of events under consideration and limb events.

[17] By using the above two criteria, we made two contingency tables for the 305 frontside halo CMEs. Table 2 shows the 2 × 2 contingency table based on the criteria of the CME location. In this case, we defined the disk CMEs (L < |50°|) to be the "yes predicted" and the geomagnetic storms (Dst < −50 nT) to be the "yes observed." In the table, hits are 96, false alarms are 117, misses are 25, and correct nulls are 67, respectively, and total data number (N = a + b + c + d) equals 305. Table 3 shows the contingency table depending on the criteria of the CME

Table 3. Contingency Table Based on the Criteria of the CME Speed (>400 km s^{-1}) for the 305 Events

	Observation		
Prediction	Yes	No	Total
Yes	110	163	273
No	11	21	32
Total	121	184	305

Table 4. Statistical Comparison of Disk CME and Fast CME for Geomagnetic Storms (Dst < −50 nT)

	Disk CME	Fast CME
Probability of detection, yes (PODy)	0.79	0.91
Probability of detection, no (PODn)	0.36	0.11
False alarm ratio (FAR)	0.55	0.60
Bias	1.76	2.26
Critical success index (CSI)	0.40	0.39

speed. The CME, whose speed is greater than 400 km s^{-1}, is defined to be the "yes predicted" and the geomagnetic storms ($Dst < -50$ nT) to be the "yes observed." In this case, hits are 110, false alarms are 163, misses are 11, and correct nulls are 21, respectively, and total data number equals 305.

[18] Table 4 summarizes statistical parameters from two contingency tables (Tables 2 and 3). In the case of location criterion (or Disk CME), PODy, PODn, FAR, bias, and CSI are estimated to be 0.79, 0.36, 0.55, 1.76, and 0.40, respectively. The PODy, PODn, FAR, bias, and CSI for the CME speed criterion (or Fast CME) are estimated to be 0.91, 0.11, 0.60, 2.26, and 0.39, respectively. We found that the PODys are relatively high for both cases, but their FARs and CSIs are rather poor. The obtained high false alarms are consistent with previous qualitative arguments [*St. Cyr et al.*, 2000; *Cho et al.*, 2003]. Also, their large biases (1.76, 2.26) mean that we did too many "predict yes." In fact, the bias should be equal to 1 for good forecasting. We tried several attempts to improve these statistical parameters using different speed criteria, but we failed.

[19] In addition to the forecast evaluation based on the contingency tables, we examined the probability of CME geoeffectiveness according to the different ranges of CME location and plane-of-sky speed. Figure 6 shows the probability of geoeffectiveness depending on the heliolongitude of the CME source region for 305 frontside halo CMEs. The probability of geoeffective CMEs' coming from the west hemisphere is found to be a little higher that of CMEs' arriving from the east hemisphere; that is, 63% of the geoeffective CMEs originate in western hemisphere. Such an asymmetry has been noted by several authors [e.g., *Wang et al.*, 2002]. The most probable region is 0° < L < +30°, where the probability is about three times higher than that of east limb CMEs and about two times higher than that of west limb CMEs. Figure 7 shows the probability of the geoeffectiveness depending on CME speed for 305 frontside halo CMEs. As seen in the figure, the probability slightly increases with CME speed for the CMEs whose speeds are lower than 1200 km s^{-1}. It is noted that the probability of very fast CMEs (>1200 km s^{-1}) is the highest, which is about 50%. The speed of a CME may be one of the criteria required to select a geoeffective CME, but it does not seem to be critical.

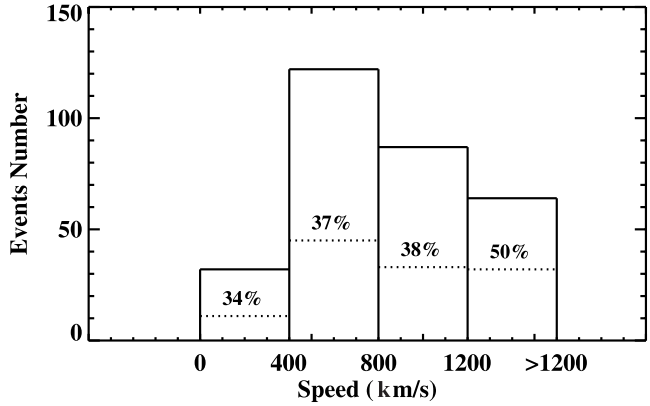

Figure 7. Probability of CME geoeffectiveness depending on the speed for 305 frontside halo CMEs. Dotted line means the number of geoeffective CMEs for a given speed range.

3.2. Dependency on the Combination of Location and Speed

[20] Figure 8 and Table 5 show the probability of the CME geoeffectiveness depending on the combination of CME location and speed. As shown in Table 5, some probabilities with small event numbers are regarded to be meaningless. The most probable "areas" (or coverage combinations), whose geoeffectiveness fraction is larger than the mean probability (about 40%), are 0° < L < +30° for slower CMEs (≤800 km s^{-1}) and −30° < L < +60° for faster CMEs (>800 km s^{-1}). That is, the area becomes

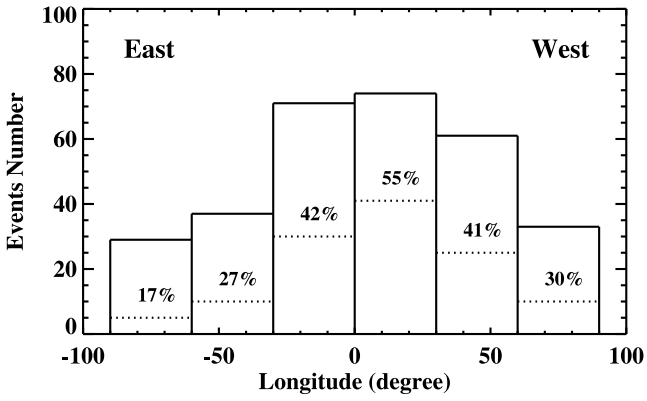

Figure 6. Probability of CME geoeffectiveness depending on the heliolongitude for 305 frontside halo CMEs. Dotted line means the number of geoeffective CMEs for a given longitudinal range.

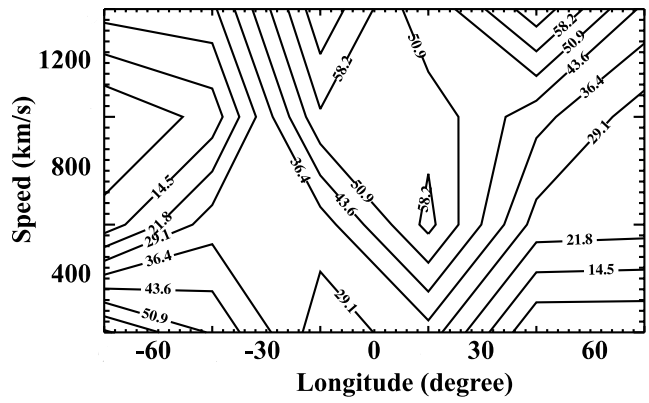

Figure 8. Probability of CME geoeffectiveness depending on the combination of CME location and speed for moderate geomagnetic storm ($Dst < -50$ nT).

Table 5. Probability Forecast of Frontside Halo CMEs Based on Their Locations and Speeds for the 305 Events

	Heliolongitude, degrees						
Speed, km s^{-1}	$-90 < L \leq -60$	$-60 < L \leq -30$	$-30 < L \leq 0$	$0 < L \leq 30$	$30 < L \leq 60$	$60 < L \leq 90$	Total
$v \leq 400$	67% ($\frac{2}{3}$)	50% ($\frac{2}{4}$)	25% ($\frac{3}{12}$)	44% ($\frac{4}{9}$)	0% ($\frac{0}{3}$)	0% ($\frac{0}{1}$)	34% ($\frac{11}{32}$)
$400 < v \leq 800$	10% ($\frac{1}{10}$)	33% ($\frac{4}{12}$)	32% ($\frac{10}{31}$)	61% ($\frac{23}{38}$)	22% ($\frac{5}{23}$)	25% ($\frac{2}{8}$)	37% ($\frac{45}{122}$)
$800 < v \leq 1200$	0% ($\frac{0}{8}$)	10% ($\frac{1}{10}$)	57% ($\frac{12}{21}$)	56% ($\frac{10}{18}$)	39% ($\frac{7}{18}$)	25% ($\frac{3}{12}$)	38% ($\frac{33}{87}$)
$v > 1200$	25% ($\frac{2}{8}$)	27% ($\frac{3}{11}$)	71% ($\frac{5}{7}$)	44% ($\frac{4}{9}$)	76% ($\frac{13}{17}$)	42% ($\frac{5}{12}$)	50% ($\frac{32}{64}$)
Total	17% ($\frac{5}{29}$)	27% ($\frac{10}{37}$)	42% ($\frac{30}{71}$)	55% ($\frac{41}{74}$)	41% ($\frac{25}{61}$)	30% ($\frac{10}{33}$)	40% ($\frac{121}{305}$)

wider for fast CMEs. Figure 9 shows the probability for strong storms ($Dst < -100$ nT). The most probable area for strong storm is $-30° < L < +30°$ (top of figure) for fast CMEs (>1200 km s^{-1}) with the geoeffectiveness probability of about 40% (7/16). The probability of this area is about three times higher than the mean probability (about 13%).

[21] On the basis of the above results of the most probable area, we made a contingency table depending on the combination of location and speed. That is, we used two location criteria depending on different speeds: $0° < L < +30°$ for slower speed ($\leq$800 km s^{-1}) and $-30° < L < +60°$ for faster speed (>800 km s^{-1}). Table 6 shows the contingency table depending on the combination of location and speed. Hits are 78, false alarms are 59, misses are 43, and correct nulls are 125, respectively, and total data number equals 305. Table 7 summarizes statistical parameters: PODy, PODn, FAR, bias, and CSI are estimated to be 0.64, 0.68, 0.43, 1.13, and 0.43, respectively. When comparing these values with those from Table 4, we found that the PODy becomes slightly lower, but FAR and bias are significantly improved.

4. Summary and Discussion

[22] In this paper we have presented the forecast evaluation of the CME geoeffectiveness using contingency tables and the probability of CME geoeffectiveness. For this, we have selected 305 frontside halo CMEs from 1997 to 2003 and then applied a CME propagation model to these CMEs.

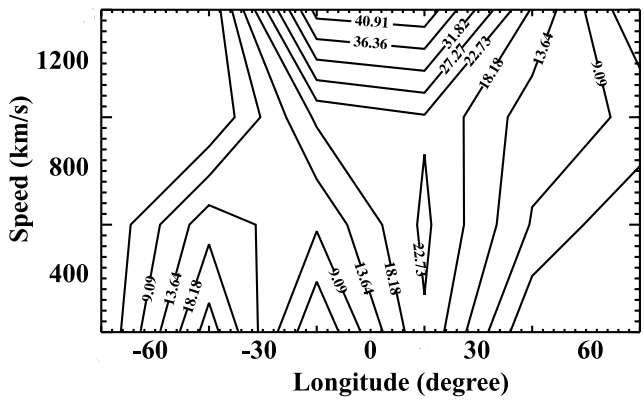

Figure 9. Probability of CME geoeffectiveness depending on the combination of CME location and speed for strong geomagnetic storm ($Dst < -100$ nT).

We finally selected 121 CME-geomagnetic storm pairs by comparing the model predicted arrival time of the CME with the minimum time of *Dst* index. Our main results can be summarized as follows.

[23] 1. The total probability of CME geoeffectiveness for frontside halo CMEs is about 40%.

[24] 2. When we defined the criterion of geoeffective CMEs as either disk CMEs or faster CMEs, the PODys from the contingency tables are relatively high. However, all other statistical parameters such as false alarm ratio and CSI are not so good.

[25] 3. The disk CMEs are more geoeffective than limb CMEs, and the CMEs that originated from the west hemisphere tend to be more geoeffective than those from eastern hemisphere.

[26] 4. The probability of CME geoeffectiveness tends to increase with CME speed with the maximum value of 50% for very fast CMEs (>1200 km s^{-1}).

[27] 5. The most probable areas whose geoeffectiveness fraction is larger than 0.4 are $0° < L < +30°$ for slower speed CMEs ($\leq$800 km s^{-1}) and $-30° < L < +60°$ for faster speed CMEs (>800 km s^{-1}). The most probable area for strong storms ($Dst < -100$ nT), is $-30° < L < +30°$ for faster CMEs (>1200 km s^{-1}) with the geoeffectiveness of about 40%.

[28] 6. According to the above results, we defined new criteria of geoeffective CMEs as the combination of location and speed. The PODy becomes slightly lower, but all other statistical parameters such as false alarm ratio and bias are significantly improved.

[29] The probability used in this study has the same concept with the flare probability given in the Active Region Monitor by *Gallagher et al.* [2002]. It is conceptually similar to the probability of rainfall in weather forecasting procedures. For example, if a CME has a speed faster than 1200 km s^{-1} and its source location is between 30 and 60 degrees, the probability of CME geoeffectiveness ($Dst < -50$ nT) is 76% (see Table 5), while that of CME

Table 6. Contingency Table Based on the Combination of Location and Speed for the 305 Events

	Observation		
Prediction	Yes	No	Total
Yes	78	59	137
No	43	125	168
Total	121	184	305

Table 7. Statistical Parameters of the Contingency Table Depending on the Combination of Location and Speed

Statistical Parameter	Value
Probability of detection, yes (PODy)	0.64
Probability of detection, no (PODn)	0.68
False alarm ratio (FAR)	0.43
Bias	1.13
Critical success index (CSI)	0.43

whose speed is slower than 400 km s^{-1} and whose location is located between −30 and 0 degree is 25%. In this case the former has three times stronger potential to produce geomagnetic storm than the latter has.

[30] This is the first time, to our knowledge, that an attempt has been made to present contingency tables depending on the geoeffectiveness criteria as well as to estimate the probability of CME geoeffectiveness depending on CME location and/or speed. We hope that this work can be a good starting point for forecasting geoeffective CMEs and its evaluation. A recent work by *Schwenn et al.* [2005] has used a radial speed, V rad (directed to Earth) that is empirically based on a similar extensive study of LASCO CMEs and their expansion speed, V exp, in the plane-of-sky at right angles to the largest plane-of-sky speed, V *ps*, that we have used here. Their results, although not presented in the statistical terms used here, are generally similar to ours. Our results can give us some criteria to select geoeffective CMEs with the probability of CME geoeffectiveness depending on the CME location and/or speed.

[31] There are several physical parameters which have to considered to improve the forecast capability of the geoeffective CME. It has been suggested that the geoeffectiveness of frontside halo CMEs depends on (1) whether or not a CME is Earth-directed or able to hit the Earth and (2) whether or not the Earth-directed CME contains and/or generates southward interplanetary magnetic field. Regarding the first issue, we note that the geoeffectiveness (75%) of frontside full halo CMEs [*Zhao and Webb*, 2003] is significantly higher than that (40%) of frontside partial halo CMEs [*Wang et al.*, 2002]. This fact implies that many, if not most, frontside partial halo CMEs are not Earth-directed. Using frontside halo CMEs originating near the central meridian can exclude those halo CMEs that cannot hit the Earth, thus increasing the geoeffectiveness. Very recently, *Moon et al.* [2005] suggested a new geoeffective parameter representing the direction of a CME and then showed that it is well correlated with *Dst* index for very fast halo CMEs. As for the second issue, *Pevtsov and Canfield* [2001] found that if a coronal flux rope model is used to interpret magnetic structures, eruptions with a southward leading magnetic field are associated with stronger geomagnetic storms, but those with a northward leading field are associated with weak storms. *Song et al.* [2005] found a close relationship between the orientation of magnetic field in the source region and the hourly averaged ACE measurements of the *Bz* component of the interplanetary magnetic fields that is believed to be the indicator of geomagnetic storms. That means there is a significant correlation between the orientation of magnetic fields in source region and *Dst* index. Finally, it should be noted that we have not considered the effect of the ICME's shock nor the IMF's orientation in the sheath region [*Wu et al.*, 1996]. More extensive investigations are in preparation for the forecast evaluation of geoeffective CMEs using these CME physical parameters such as direction and field orientation.

[32] **Acknowledgments.** We really appreciate the referees' constructive comments. This work has been supported by the MOST grants (M1-0104-00-0059 and M1-0407-00-0001) of the Korean government. YJM and YDP were supported by the Korea Research Foundation (KRF-2005-070-C00059) of the Korean Government. CNU group was supported by the grant (F01-2003-000-00186-0) from the Korea Science and Engineering Foundation (KOSEF). MD was supported by NASA's Living With a Star Program via grant NAG5-12527 to Exploration Physics International, Inc. MD also thanks NOAA's Space Environment Center for their hospitality. The CME catalog used here is generated and maintained by NASA and the Catholic University of America in cooperation with the Naval Research Laboratory. SOHO is a project of international cooperation between ESA and NASA. The *Dst* index is provided by the World Data Center for Geomagnetism at Kyoto University.

[33] Lou-Chuang Lee thanks Haimin Wang and Xue Pu Zhao for their assistance in evaluating this paper.

References

Brueckner, G. E., et al. (1995), The Large Angle Spectroscopic Coronagraph (LASCO) Visible light coronal imaging and spectroscopy, *Solar Phys.*, *162*, 357–402.

Cane, H. V., and I. G. Richardson (2003), Interplanetary coronal mass ejections in the near-Earth solar wind during 1996–2002, *J. Geophys. Res.*, *108*(A4), 1156, doi:10.1029/2002JA009817.

Cane, H. V., I. G. Richardson, and O. C. St. Cyr (2000), Coronal mass ejections, interplanetary ejecta and geomagnetic storms, *Geophys. Res. Lett.*, *27*, 3591–3594.

Cho, K.-S., Y.-J. Moon, M. Dryer, C. D. Fry, Y.-D. Park, and K.-S. Kim (2003), A statistical comparison of interplanetary shock and CME propagation models, *J. Geophys. Res.*, *108*(A12), 1445, doi:10.1029/2003JA010029.

Delaboudiniere, J. P., et al. (1995), EIT: Extreme-Ultraviolet Imaging Telescope for the SOHO mission, *Solar Phys.*, *162*, 291–312.

Dryer, M. (1994), Interplanetary studies: Propagation of disturbances between the Sun and the magnetosphere, *Space Sci. Rev.*, *67*(3/4), 363–419.

Dryer, M., and D. F. Smart (1984), Dynamical models of coronal transients and interplanetary disturbances, *Adv. Space Res.*, *4*, 291–301.

Fry, C. D., M. Dryer, Z. Smith, W. Sun, C. S. Deehr, and S.-I. Akasofu (2003), Forecasting solar wind structures and shock arrival times using an ensemble of models, *J. Geophys. Res.*, *108*(A2), 1070, doi:10.1029/2002JA009474.

Gallagher, P., Y.-J. Moon, and H. Wang (2002), Active-region monitoring and flare forecasting I. Data processing and first results, *Solar Phys.*, *209*, 171–183.

Gonzalez, W. D., J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani, and V. M. Vasyliunas (1994), What is a geomagnetic storm?, *J. Geophys. Res.*, *99*, 5771–5792.

Gopalswamy, N., A. Lara, S. Yashiro, M. Kaiser, and R. A. Howard (2001), Predictig the 1-AU arrival times of coronal mass ejections, *J. Geophys. Res.*, *106*, 29,207–29,218.

Michalek, G., N. Gopalswamy, A. Lara, and P. K. Manoharan (2004), Arrival time of halo coronal mass ejections in the vicinity of the Earth, *Astron. Astrophys.*, *423*, 729–736.

Moon, Y.-J., Y. D. Park, H. C. Seong, C. W. Lee, K. J. Sim, and H. S. Yun (2000), A near real-time flare alerting system based on GOES soft X-ray flux, *J. Korean Astron. Soc.*, *33*(2), 123–126.

Moon, Y.-J., M. Dryer, Z. Smith, Y. D. Park, and K.-S. Cho (2002), A revised shock time of arrival (STOA) model for interplanetary shock propagation: STOA-2, *Geophys. Res. Lett.*, *29*(10), 1390, doi:10.1029/2002GL014865.

Moon, Y.-J., K.-S. Cho, M. Dryer, Y.-H. Kim, S. Bong, J. Chae, and Y. D. Park (2005), New geoeffective parameters of very fast halo coronal mass ejections, *Astrophys. J.*, *624*, 414–419.

Park, Y. D., Y.-J. Moon, B. H. Jang, and K. J. Sim (1997), Observation System of Solar Flare Telescope (in Korean), *Publ. Korean Astron. Soc.*, *12*(1), 35–45.

Pevtsov, A. A., and R. C. Canfield (2001), Solar magnetic fields and geomagnetic events, *J. Geophys. Res.*, *106*, 25,191–25,198.

Schwenn, R., A. Dal Lago, E. Huttunen, and W. D. Gonzalez (2005), The association of coronal mass ejections with their effects near the Earth, *Ann. Geophys.*, *23*, 1033–1059.

Smith, Z., and M. Dryer (1990), MHD study of temporal and spatial evolution of simulated interplanetary shocks in the ecliptic plane within 1 AU, *Solar Phys.*, *129*, 387–405.

Smith, Z., M. Dryer, E. Ort, and W. Murtagh (2000), Performance of interplanetary shock prediction model: STOA and ISPM, *J. Atmos. Sol. Terr. Phys.*, *62*, 1265–1274.

Smith, Z., T. Detman, M. Dryer, C. D. Fry, C.-C. Wu, W. Sun, and C. S. Deehr (2004), A verification method for space weather models using solar data to predict arrivals of interplanetary shocks at Earth, *IEEE Trans. Plasma Phys.*, *32*(4), 1498–1505.

Smith, Z., M. Dryer, and C. D. Fry (2005), Determining shock velocities for inputs to Sun-to-Earth models from radio and coronagraph data, *Space Weather*, *3*, S07002, doi:10.1029/2004SW000136.

Song, H., V. B. Yurchyshyn, and H. Wang (2005), Towards real-time automated prediction of geo-magnetic storms based on observations of source regions of halo CMEs, *Eos Trans. AGU*, *86*(18), Jt. Assem. Suppl., Abstract SP 23A-01.

Srivastava, N., and P. Venkatakrishnan (2004), Solar and interplanetary sources of major geomagnetic storms during 1996–2002, *J. Geophys. Res.*, *109*, A10103, doi:10.1029/2003JA010175.

St. Cyr, O. C., et al. (2000), Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998, *J. Geophys. Res.*, *105*, 18,169–18,186.

Wang, Y. M., P. Z. Ye, S. Wang, G. P. Zhau, and J. X. Wang (2002), A statistical study on the geoeffectiveness of Earth-directed coronal mass ejections from March 1997 to December 2000, *J. Geophys. Res.*, *107*(A11), 1340, doi:10.1029/2002JA009244.

Wu, C.-C., M. Dryer, and S. T. Wu (1996), Three-dimensional MHD simulation of interplanetary magnetic field changes a 1 AU as a consequence of simulated solar flares, *Ann. Geophys.*, *14*, 383–399.

Yashiro, S., N. Gopalswamy, G. Michalek, O. C. Cyr, S. P. Plunkett, N. B. Rich, and R. A. Howard (2004), A catalog of white light coronal mass ejections observed by the SOHO spacecraft, *J. Geophys. Res.*, *109*, A07105, doi:10.1029/2003JA010282.

Yurchyshyn, V. B., H. Wang, and V. Abremenko (2004), Correlation between speeds of CMEs and the intensity of geomagnetic storms, *Space Weather*, *2*, S02001, doi:10.1029/2003SW000020.

Zhang, J., K. P. Dere, R. A. Howard, and V. Bothmer (2003), Identification of solar sources of major geomadnetic storms between 1996 and 2000, *Astrophys. J.*, *582*, 520–533.

Zhao, X. P., and D. F. Webb (2003), Source regions and storm effectiveness of frontside full halo coronal mass ejections, *J. Geophys. Res.*, *108*(A6), 1234, doi:10.1029/2002JA009606.

S.-C. Bong, K.-S. Cho, Y.-H. Kim, Y.-J. Moon, and Y.-D. Park, Korea Astronomy and Space Science Institute 61-1, Whaam-dong, Yuseong-gu, Daejeon, 305-348, South Korea.

M. Dryer, NOAA Space Environment Center, 325 Broadway, Boulder, CO 80303, USA.

R.-S. Kim and Y. Yi, Department of Astronomy and Space Science, Chungnam National University, 220 Gung-dong, Yuseong-gu, Daejeon, 305-754, South Korea. (rskim@cnu.ac.kr)