Predicting the magnetic vectors within coronal mass ejections arriving at Earth: 2. Geomagnetic response: BZ VALIDATION

Article *in* Space Weather · October 2016

DOI: 10.1002/2016SW001458

9 authors, including:

Angelos Vourlidas
Johns Hopkins University
384 PUBLICATIONS **6,273** CITATIONS
SEE PROFILE

I. G. Richardson
University of Maryland, College Park
245 PUBLICATIONS **5,833** CITATIONS
SEE PROFILE

Antti Pulkkinen
NASA
146 PUBLICATIONS **1,561** CITATIONS
SEE PROFILE

Teresa Nieves-Chinchilla
IACS/CUA & GSFC/NASA
52 PUBLICATIONS **399** CITATIONS
SEE PROFILE

Predicting the magnetic vectors within coronal mass ejections arriving at Earth: 2. Geomagnetic response

N. P. Savani,[1,2] A. Vourlidas,[3] I. G. Richardson,[4,2] A. Szabo,[2] B. J. Thompson,[2] A. Pulkkinen,[2] M. L. Mays,[5,2] T. Nieves-Chinchilla,[5,2] V. Bothmer[6]

[1]Goddard Planetary Heliophysics Institute, University of Maryland, Baltimore County, 5523 Research Park Drive, Baltimore, MD 21228 USA

[2]NASA, Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

[3]Solar Section, Applied Physics Laboratory Johns Hopkins University, Laurel, MD USA

[4]University of Maryland, College Park, MD 20742, USA

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:10.1002/2016SW001458

Abstract. This is a companion to [*Savani et al.*, 2015] that discussed how a first-order prediction of the internal magnetic field of a coronal mass ejection (CME) may be made from observations of its initial state at the Sun for space weather forecasting purposes (BSS model). For eight CME events, we investigate how uncertainties in their predicted magnetic structure influence predictions of the geomagnetic activity. We use an empirical relationship between the solar wind plasma drivers and Kp index together with the inferred magnetic vectors, to make a prediction of the time variation of Kp ($Kp^{(BSS)}$). We find a 2σ uncertainty range on the magnetic field magnitude ($|B|$) provides a practical and convenient solution for predicting the uncertainty in geomagnetic storm strength. We also find the estimated CME velocity is a major source of error in the predicted maximum Kp. The time variation of $Kp^{(BSS)}$ is important for predicting periods of enhanced and maximum geomagnetic activity, driven by southerly-directed magnetic fields, and

[5]Catholic University of America, 620
Michigan Ave NE, Washington, DC 20064,
USA.

[6]Institute for Astrophysics,
Georg-August-University of Göttingen,
Friedrich-Hund-Platz 1, 37077 Göttingen,
Germany

periods of lower activity driven by northerly directed magnetic field. We compare the skill score of our model to a number of other forecasting models, including the NOAA/SWPC and CCMC/SWRC estimates. The BSS model was the most unbiased prediction model while the other models predominately tended to significantly over-forecast. The True skill score of the BSS prediction model ($TSS = 0.43 \pm 0.06$) exceeds the results of 2 baseline models and the NOAA/SWPC forecast. The BSS model prediction performed equally with CCMC/SWRC predictions while demonstrating a lower uncertainty.

Key Points

- Long lead-time magnetic vector predictions of 8 CME are converted into Kp predictions for comparisons to forecasts
- Model retained a more consistently strong skill for the more stringent 'Hit' criteria, while also displaying significantly lower uncertainty
- Uncertainty in magnetic field strength and velocity is the most practical solution for Kp forecast uncertainty

1. Introduction

In this second of two companion papers, we study how predictions of the magnetic vectors within coronal mass ejections (CMEs), the topic of the Paper 1 [*Savani et al.*, 2015], may be used to forecast the geomagnetic response from these CMEs. In particular, we investigate the effects on predicting the Kp geomagnetic index by varying the inputs to the CME magnetic field vector prediction method.

Extended periods of southward magnetic field, often found within and surrounding Earth-directed CMEs, result in periods of enhanced reconnection on Earth's dayside and energy input into the magnetosphere. In contrast, northward-directed fields inhibit reconnection, resulting in a weaker magnetospheric response [*Dungey*, 1961]. As CMEs usually generate the most geo-effective events [e.g., *Tsurutani and Gonzalez*, 1997; *Zhang et al.*, 2007; *Zhang and Moldwin*, 2014], inferring the magnetic field direction along the trajectory of Earth through the CME would be a major advance in geomagnetic activity prediction.

Theory and observations suggest that the magnetic fields inside CMEs are often twisted "flux ropes" [*Vourlidas*, 2014]. Such twisted fields are also often found in in-situ during the passage of CMEs [e.g., *Lepping et al.*, 1990; *Marubashi*, 1997; *Hidalgo et al.*, 2002; *Hu and Sonnerup*, 2002; *Owens et al.*, 2006; *Savani et al.*, 2013a]. However, the challenge has been to predict the structure of in-situ flux ropes from observations of the source of the CME at the Sun.

A link between solar active regions or eruptive filaments and flux rope structures has already been proposed by *Marubashi* [1997]. Further supporting evidence is provided

by the chirality of these structures [*Rust*, 1994; *Bothmer and Schwenn*, 1994; *Bothmer and Rust*, 1997] and their axis orientation in the ecliptic plane [*Yurchyshyn et al.*, 2007] and over a large range of heliographic latitudes [*Rees and Forsyth*, 2003]. Evidence has shown that the relationship of interplanetary flux ropes and their solar progenitors is more straightforward for filaments than active regions [*Leamon et al.*, 2002, 2004]. However, a large fraction of CMEs are associated with active regions.

Bothmer and Schwenn [1998] suggested that the magnetic topology of CMEs should be driven by conditions below the solar surface [*Hale*, 1925; *Wang*, 2013]. In particular they proposed a scheme that predicted the magnetic topology based on two parameters: whether the solar cycle number is even or odd, and whether the CME originated in the northern or southern hemisphere. However, it is now clear that this scheme can fail in some cases. For example, *Liu et al.* [2006] and *Harra et al.* [2007] reported an event whose overlying field arcade from which a CME erupted spanned two solar active regions. The resulting interplanetary magnetic flux rope structure was shown to be different from that of a prior CME that erupted from just one of these active regions. In response, Paper 1 introduced an additional parameter to the hemisphere-handedness scheme (adjusted Bothmer-Schwenn scheme, BSS) to allow for a single or double active region source.

The Earth trajectory through a model CME strongly influences the predicted magnetic field vectors at Earth. The vectors are also affected by evolution of the CME as it propagates to Earth. Radial propagation of CMEs have suggested a flattening of a circular cross section CME [*Riley and Crooker*, 2004; *Savani et al.*, 2011a, 2012]. However, in situ evidence often from model fittings, suggest a circular cross section may be appropriate on local scales [*Démoulin and Dasso*, 2009; *Savani et al.*, 2011b, 2013b]; the non-radial

extent is less well constrained [*Owens*, 2008]. CME deflections [*Cremades and Bothmer*, 2004; *Kay et al.*, 2013] and rotations [*Vourlidas et al.*, 2011; *Nieves-Chinchilla et al.*, 2013] may also influence the magnetic structures that encounter the Earth, as can deformations of a CME due to interactions with the ambient solar wind [*Savani et al.*, 2010] or other interplanetary CMEs [*Lugaz and Farrugia*, 2014; *Liu et al.*, 2014]. Studies have also shown evidence that the magnetic flux rope structure within a CME may also erode away during propagation [*Ruffenach et al.*, 2012; *Manchester et al.*, 2014; *Ruffenach et al.*, 2015].

In this paper, we discuss how the scheme for predicting the magnetic fields in CMEs described in Paper 1 may be used to forecast geomagnetic activity. We first describe the current geomagnetic storm forecasts provided by NOAA and NASA (Section 2). Second, we outline our empirical approach for predicting the Kp values (Section 3) and directly compare our predictions with current space weather forecasting products. Thirdly, we discuss how uncertainty in the trajectory of Earth through a model CME structure changes the predicted magnetic vectors and how this influences the predicted Kp values (Section 5). Additionally, we also investigate how other parameters affect the final Kp prediction and quantify the predictive capabilities with skill scores (Section 6). Finally, we discuss the results and best approach for developing the forecasting system in Section 7 and discuss future improvements in Section 8.

2. Kp Forecasting

The NOAA Space Weather Prediction Center (SWPC) is the official U.S. source for space weather forecasts. For geomagnetic storms, their current procedure is to: 1. Release a geomagnetic "watch" notice after a solar eruption is observed, i.e., 36-72 hrs prior to expected storm arrival; 2. Send out a "warning" $\sim$ 1hr prior to storm onset using data

from ACE/Wind spacecraft at the L1 point to formulate a robust Kp prediction based on established relations between the solar wind parameters and Kp; and 3. Provide "Alerts" in a now-casting/real time format during the geomagnetic storm using a real-time proxy of the official Kp index. SWPC also publishes routine 3-day Kp forecasts based on a heuristic approach that is heavily dependent on the skill and personal experience of the on-duty forecaster, for example, in interpreting ENLIL simulation results and incorporating knowledge of historical events.

NASA's Space Weather Research Center (SWRC) hosted by the CCMC (Community Coordinated Modeling Center) at the Goddard Space Flight Center (GSFC) employs more experimental forecasting procedures using the latest scientific techniques that have only recently been deployed in an operational-like setting. The CCMC is an inter-agency partnership to facilitate community research and accelerate implementation of progress in research into space-weather operations. CCMC/SWRC are tasked with addressing the space weather needs of NASA's robotic missions through experimental research forecasts, notification and analysis. As part of its activities, CCMC gathers predictions from researchers around the world and compares them with the actual observed space weather in order to understand the strengths and limitations of the different forecasting techniques.

Both NOAA/SWPC and CCMC/SWRC use ENLIL to model the propagation of CMEs between the Sun and Earth, and estimate arrival times [*Odstrcil*, 2003; *Odstrcil and Pizzo*, 2009]. However, the currently released ENLIL models do not include CME magnetic fields; only the magnetic structure of the ambient solar wind is included in the model. Thus, ENLIL cannot provide any forecasts of the magnetic field vectors inside CMEs, at the moment. CCMC are implementing an experimental method to forecast Kp with quan-

titative uncertainties using the maximum field strength from ENLIL model runs [*Mays et al.*, 2015a]. Recent developments have shown that real-time operations of numerical simulations that include CME magnetic fields are now possible [*Shiota and Kataoka*, 2016].

3. Kp Theory

Kp is the global magnetospheric index often used by forecasters to indicate the severity of a space weather event [*Wing et al.*, 2005, and references therein]. As discussed in Paper 1, a function is required that characterizes the coupling of the solar wind to the magnetosphere. Many functions that couple the solar wind to a wide variety of magnetospheric activity have been proposed in the past, often incorporating the magnetic field orientation [*Lockwood et al.*, 2013, and references therein]. A recent study suggests that one parameter correlates best with 9 out of 10 indices of terrestrial activity [*Newell et al.*, 2007]. This parameter, $d\Phi/dt$, represents the rate at which magnetic flux is opened at the magnetopause and is defined as

$$d\Phi/dt = v^{4/3} \mid \mathbf{B} \mid^{2/3} \sin^{8/3}(\theta_c/2) \tag{1}$$

where v is the velocity of the solar wind; $\mid \mathbf{B} \mid$ is the magnetic field magnitude; and the interplanetary magnetic field (IMF) clock angle is defined by $\theta_c \equiv \tan^{-1}(By/Bz)$. The correlation coefficient of $d\Phi/dt$ with the Kp index is r= 0.76 [see table 3, *Newell et al.*, 2007]. The predicted magnetic field time-series in GSM coordinate system is used to calculate a theoretical magnetic flux rate. A Kp prediction is then generated by the empirical correlation.

$$Kp^{(BSS)} = 9.5 - e^{A - B(d\Phi/dt)} \quad (2)$$

where $A = 2.18$, $B = 5.20 * 10^5$, and with the velocity and magnetic field measured in km/s and nT, respectively [*Emmons et al.*, 2013; *Mays et al.*, 2015a]. The Kp predictions are easily converted to the official NOAA geomagnetic storm scale (G1 to G5) by a linear mapping of the Kp values. (http://www.swpc.noaa.gov/noaa-scales-explanation).

Equation 1 displays how the predicted values of Kp are influenced by the solar wind speed. In this article we will use the average velocity during passage of the CME, as measured from the ACE and WIND data sets at L1. The reason for implementing the observed velocity from L1 is a simple one; we wish to understand and isolate the effects of the predicted magnetic vectors without prejudice from uncertainties in other model parameters that influence estimated arrival time or velocity [e.g., *Taktakishvili et al.*, 2010; *Colaninno et al.*, 2013; *Vršnak et al.*, 2014; *Tucker-Hood et al.*, 2015].

The derived Kp estimates using real time in situ data from L1 are currently implemented within the CCMC integrated Space Weather Analysis System (iSWA). The real L1 data has been used to derive the Kp for several CMEs and compared to a proxy L1 time-series created from an ensemble of ENLIL runs [*Mays et al.*, 2015a]. The level of geomagnetic activity predicted using the L1 data was shown to have a root mean square error of less than a single geomagnetic storm scale.

Currently CCMC/SWRC use a simple and arbitrary process for the long-lead time geomagnetic activity forecasts. The magnetic field strength from ENLIL in the compressed sheath material ahead of a CME is assumed to be constantly pointing West, Southwest or South during the passage of the CME, and the expected Kp is estimated for each

case. The average Kp estimate is taken from the Southwest direction. The Kp forecast for NASA robotic operations issued by CCMC/SWRC is then the range between these extreme cases, with the highest Kp prediction originating from a southward field [*Emmons et al.*, 2013; *Mays et al.*, 2015a].

In this paper, the observed arrival time of the CME from L1 data is used only as a guide to define the time-interval for the $Kp^{(BSS)}$ prediction. This is because we define our quantitative skill score metrics based upon a time interval where the predicted and observed Kp have a maximum correlation (see Section 6 for more details). This time shifting process enables a consistent analysis of Kp skill for all CMEs. However, implementation of this tool in an operational setting will require the arrival time to be fixed and dictated by a separate arrival time forecast process (which is likely to be determined through multiple ENLIL model runs).

4. Events

This article discusses eight Earth directed CME events between 2010 and 2014, whose criteria for selection are further detailed in Paper 1. Table 1 displays the solar initiation time of the CMEs. For these events, the estimated magnetic vectors overlaid onto solar wind magnetic fields, measured at L1 using ACE and WIND spacecraft [*Lepping et al.*, 1995; *Smith et al.*, 1998], are displayed in Figure 1. Both observed and predicted values are displayed in the GSE coordinate system (See Appendix A for the cartesian Bz component of the magnetic field). The technique of estimating the vectors is detailed in Paper 1.

After converting the predicted magnetic field vectors into the GSM coordinated system, equations 1 and 2 are used to obtain the predicted Kp values ($Kp^{(BSS)}$). In Figure 2, predicted Kp values are displayed alongside the observed Kp values for the January

2014 CME event highlighted in Paper 1. The predicted times-series have been time-averaged ($Kp^{(BSS)}$ prediction, red curve) to match the resolution of the observations. For this event a significant false alarm was noted by the operational forecasters. The maximum Kp forecast from both NOAA/SWPC and CCMC/SWRC are also displayed with uncertainties in their Kp magnitude estimation and with an uncertainty in their CME arrival time. The uncertainty of the maximum $Kp^{(BSS)}$ prediction is generated from varying the magnitude of the magnetic field strength, which is explained further in section 5.5. The predicted Kp values of all eight CMEs overlaid on the observed Kp measurements are displayed in Figure 3.

5. Results

In this section, the Earth-directed CME detected in January 2014 is used as a case study to highlight how various adjustments effect the predicted magnetic vectors and hence the resulting Kp predictions. Adjustments to the hypothetical trajectory through the model flux rope structure are manually made to investigate the changes. Details of the implemented constant alpha flux rope model are given in *Savani et al.* [2013a], which closely follows the model of *Lepping et al.* [1990]. As this model has a circular cross-section cylindrical morphology, all possible trajectories can be generated through adjustments in two perpendicular directions: parallel and perpendicular to the flux rope axis.

Sub-sections below investigate "thought experiments" that manually adjust all the different degrees of freedom within the BSS model. Thus the uncertainty of our time-varying Kp predictions is studied by changing parameters such as the chirality and CME orientation.

5.1. Chirality of CME Axis

Bothmer and Schwenn [1998] suggested that the magnetic topology of CMEs above the solar surface should be driven by the subsurface behavior of the Sun [*Hale*, 1925; *Wang*, 2013]. The original scheme employed a prediction of the magnetic topology based on two parameters (odd/even solar cycle and north/south hemisphere). Paper 1 describes a new methodology (BSS scheme) to improve the reliability by introducing a third parameter; single/double active region. The introduction of an extra parameter does not change the chirality of the structure, but flips the field direction of the central axis. Cases where the overlying field arcade, from which a CME is released, emanate between two active regions are significantly less frequent than a single active region scenario. Panel G and F in Figure 1 and 3, show observations related to two events originating in double active regions.

Figure 4 displays the fictitious magnetic vector and $Kp^{(BSS)}$ predictions had the standard Bothmer-Schwenn scheme been used for the January 2014 event (i.e., only a single active region) in red, along with the adjusted Bothmer-Schwenn (BSS scheme) prediction in gray which takes into account that two active regions were associated with this CME. Whether the standard or adjusted scheme is used has a significant effect on the predictions made. In particular, the north-south component of the field is flipped such that with the adjusted scheme, the field rotates from southward to northward with time, whereas assuming the standard scheme, the rotation is in the opposite direction as suggested in Figure 4.

This delay in the southward field excursion, evident in the second panel of Figure 4, causes the predicted maximum $Kp^{(BSS)}$ to be delayed by approximately 6 hours but

the maximum $Kp^{(BSS)}$ remains the same (5-). Uncertainty in the arrival time of this maximum magnitude is of the same order as the uncertainty in the CME Earth arrival time predictions even though the scientific mechanisms are unrelated. Thus, the correct inference of the flux rope topology that emanates from double active regions is crucial for improving peak geomagnetic forecasts.

5.2. Variations of Earth Trajectory Parallel to CME Axis

The hypothetical trajectory through the model flux rope structure is manually adjusted to better understand the effect on the predicted magnetic vectors and $Kp^{(BSS)}$ prediction. This new trajectory through the CME is then the assumed path of Earth. The morphology of this simplified model is a circular-cross section cylinder with a straight central axis, which is a reasonable local approximation for a flux rope at 1 AU [*Owens et al.*, 2012]. The vector direction of the central axis is determined observationally and through an empirical relationship (see Paper 1 for further details).

All hypothetical trajectories can be generated by simple modifications along an orthogonal vector set (e.g., East/West and North/South). We choose to define our orthogonal vector set to be parallel and perpendicular to the central flux rope axis. Trajectory movements along the parallel direction increase the radial component of the predicted magnetic flux rope axis as the trajectory moves from the center to either legs of the CME. Figure 5 displays the effect of moving the Earth trajectory parallel to the flux rope axis towards the southern leg of the CME. A manual adjustment of $\approx 50\%$ of the axis half-length is added to the original prediction for the January 2014 event.

Figure 6 displays the resulting magnetic vector prediction and $Kp^{(BSS)}$ forecast for a manual change to the Earth-trajectory. In this scenario, the predicted magnetic vector

direction has a stronger and slightly more prolonged excursion in the southerly direction. However, the predicted peak in $Kp^{(BSS)}$ remains unchanged at 5-. For some other CMEs, the duration of the southerly field excursion is increased, extending the interval of enhanced Kp values.

5.3. Variations of Earth Trajectory Perpendicular to CME Axis

The perpendicular distance of a trajectory from the flux rope axis may be expressed as the impact parameter Y_0 defined as 0 if the trajectory intercepts the central axis of the flux rope, to 1 if it grazes the outer edge. Figure 7 schematically displays a manual change for the January 2014 event, by decreasing the impact parameter from 0.91 to 0.21.

Figure 8 displays the resulting magnetic vector prediction and $Kp^{(BSS)}$ prediction for this smaller impact parameter. In this scenario, the predicted magnetic vector smoothly rotates over a much larger angle. For this event, the larger field rotation is clearly shown in the B_θ component. This faster rotation of the field direction means the field rotates into a northerly direction earlier and for a longer duration, which in turn reduces the maximum $Kp^{(BSS)}$ prediction from 5- to 4-.

5.4. Tilt rotation of CME Axis

The predicted magnetic field vectors along a given trajectory can also be changed by varying the tilt of the flux rope axis with respect to the ecliptic. As discussed in paper 1, the tilt is determined near the Sun from multi-point coronagraph images using the graduated cylindrical shell (GCS) model [*Thernisien et al.*, 2009], though any other elliptical cone model could be used. While the morphology of a CME can be reasonably well constrained by triangulation using observations from three well-separated spacecraft,

uncertainties increase if these spacecraft are in a less than optimal configuration or if observations are not available from every spacecraft.

Using Figure 9, we test the robustness of the time-varying Kp prediction to a change in the tilt angle of the CME axis. The axis is manually increased from 40° to 60°. With this change in CME tilt, the predicted maximum Kp decreases slightly, from 5- to 4+. When the same tilt angle change was combined with the lower impact parameter discussed in Section 5.3, the predicted maximum Kp decreased further, to 3+ (figure not shown). For this case, the Earth's trajectory changes from being slightly East of the CME's central nose to slightly West. This has an important consequence on the predicted heliocentric radial component of the CME axis direction, which changes sign.

5.5. Variations of the Magnetic Field Strength

For the purposes of providing an early lead-time estimate of the magnetic vectors arriving at L1, we estimate the field strength along the central axis of the CME by incorporating results from ENLIL simulation runs provided by CCMC/SWRC. CCMC/SWRC make predictions and forecasts based on the maximum field strength and CME velocity at L1 from the WSA-ENLIL+Cone model. These maximum values usually pertain to the sheath plasma pile-up region ahead of a CME. Essentially, this method correlates the estimated CME field strength to the assumed CME momentum (from observations and ENLIL) and the different background solar wind conditions. Several other techniques to estimate the magnetic field strength in CMEs have been suggested, but do not provide a consistently reliable estimate for all observable CME events [e.g., *Longcope et al.*, 2007; *Kunkel and Chen*, 2010; *Poomvises et al.*, 2012; *Tun and Vourlidas*, 2013].

A statistical approach is taken to estimate the uncertainty in the magnetic field strength. *Lepping et al.* [2006] investigated 82 CME events using in situ observations from the WIND spacecraft. The authors fitted a constant alpha flux rope model to observed CME magnetic fields and determined the standard deviation of the magnetic field strength at the central flux rope axis to be $\sigma = 6.9nT$. For this study, we then choose to vary the field strength as $\mid \mathbf{B} \mid^{+2\sigma}_{-1\sigma}$. The reason for choosing a larger positive uncertainty is that predictions of maximum geo-effectiveness are of particular interest for storm forecasting, and with this choice, the uncertainty is biased towards larger fields that will lead to higher geomagnetic activity if the field is favorably orientated.

Figure 10 displays the resulting magnetic vector prediction and $Kp^{(BSS)}$ prediction for a manual increase in the maximum field strength along the central axis by 2σ ($\mid \mathbf{B} \mid = 31.8nT$). In this scenario, the predicted direction of the magnetic vector in the B_θ and B_ϕ panels remains unchanged. However, a dramatic and uniform increase in the field strength extending over the duration of the event has a relatively simple effect on Kp: Since $d\Phi/dt$ increases, $Kp^{(BSS)}$ also increases (cf equations 1 and 2). The double peak in the predicted $Kp^{(BSS)}$ remains but with a faster rise and reaching a larger maximum Kp (increasing from 5- to 6).

5.6. Variations to CME Velocity

In Paper 1, we used the average speed within the CME from the measured in situ data at L1. Here, to simulate a real-time forecasting scenario, we test the BSS magnetic field technique in a real-time forecasting setting. We use the maximum velocity (along the Earth trajectory) from a real time WSA-ENLIL+Cone model run by CCMC/SWRC.

Figure 11 displays the $Kp^{(BSS)}$ generated by the CCMC/SWRC 840 km/s velocity estimate. This is an increase of 470km/s from the observed speed at WIND (370 km/s), and above the 2σ variation in the predicted velocity. The effect of this velocity is to increase the rate at which the Kp values rise and to reach a higher maximum value (increasing from 5- to 8). For this case study, the change to the maximum estimated $Kp^{(BSS)}$ is very large. Usually the uncertainty of the CME velocity (relative to the measured speed) is smaller than the relative uncertainty of the field magnitude.

A simple view of equation 1 indicates changes to the CME velocity (via the effect of $v^{4/3}$) would have a greater impact on the $Kp^{(BSS)}$ than the magnetic field strength. However, the dynamic range of expected maximum magnetic field is larger than the velocity, and the uncertainty in the velocity decreases as the arrival time predictions become more accurate. Both arrival-time and CME speed at the Sun are separate estimates that are well investigated with many attempts to improve both of these prediction parameters in the current literature [e.g., *Gopalswamy et al.*, 2001; *Crosby et al.*, 2012; *Bothmer et al.*, 2013; *Vršnak et al.*, 2013; *Iju et al.*, 2014; *Tucker-Hood et al.*, 2015; *Shi et al.*, 2015]. In this paper, we wish to understand and isolate the effects of the predicted magnetic vectors without prejudice from these arrival-time estimates.

6. Skill scores

To test the capability of the new BSS technique to predict Kp, a variety of skill scores will be applied to the eight CMEs discussed above. As with all skill scoring systems, the definition of an 'event' must be carefully defined in such a way that the same definition can also be used in future tests of a larger number of events or other forecasting. The definitions must also consider time-discretization of an event as well as relevant thresholds

for a 'Hit'. We define an 'event' as a 3-hour period corresponding to a single Kp value and the interval investigated as the period of 'active solar wind' during the passage of a CME and its associated structures, e.g. shock and sheath.

6.1. Current Static skills

The Government Performance and Results Act metrics for NOAA/SWPC are based on the performance of geomagnetic storm forecasts, which monitor the Kp index. NOAA/SWPC geomagnetic Storm Forecast Accuracy is measured as the percentage of times that the 24 hour geomagnetic storm forecast is correct for the 60 most recent geomagnetic storms. The 24 hour geomagnetic storm forecast is considered accurate if a Kp=5 (G1 storm) or greater storm event was predicted. This measure is based on the next-day forecast of maximum Kp, and is verified against the NOAA Kp estimated from ground-based magnetometer observations [*NOAA*, 2014].

By monitoring the forecast skill in a 24 hour 'event' block, the duration of an entire solar storm passing over Earth, or 'active solar wind' period, is individually assessed as a single event within a contingency table (or skill score matrix). This time-static approach to measuring the skill score has two major implications in regards to the work presented here: 1) A comparison of the BSS model forecast to NOAA forecasts is less pertinent, as one focuses on how the Kp index varies during a solar storm and, the other on the overall maximum Kp impact of a solar storm. 2) The significantly lower time-resolution of the forecasting skill used by NOAA drastically reduces the total number of events to below where reliable statistical analysis can be performed. Indeed, NOAA/SWPC assess over the 60 most recent geomagnetic storms to maintain statistical significance for their important probability of detection (POD) metric [*NOAA*, 2014].

We investigated the 'Hit' criteria based on the NOAA/SWPC methodolgy. An event is considered to be the entire duration of a CME interval, and a 'Hit' is defined when both predicted and observed have $Kp \geq 5$. As the lead-time forecast for the BSS model is greater than 24 hours, the first available NOAA/SWPC 3-day forecast that includes an influence of the CME is investigated.

As NOAA/SWPC transitioned to forecasting the G scale (Kp Index) in 2012, we investigate the skill for the six CME events between 2012-2014. The NOAA/SWPC forecasting accuracy, the probability of detection, were POD = 0.8 for these CMEs. The official NOAA/SWPC Geomagnetic Storm Forecast Accuracy between 2012 and 2014 were POD = 47% and 40% respectively; while also achieving their target accuracy of 53% in 2015 [*NOAA*, 2013]. Therefore, our six CME events can be considered as between typical and relatively better forecasted than the annual averages; albeit using a smaller sample of the NOAA forecasts.

Table 2 displays further skill metrics using the same 'Hit' criteria for forecasts made by NOAA/SWPC, CCMC/SWRC and the equivalent new BSS model. Due to the limitation of statistics in our sample size, care is required prior to drawing strong conclusions. For this reason, the majority of skill tests shown in table 2, show comparable results between the different forecast techniques. Although, there appears significant improvement in the True skill score metric, that warrants further investigation with better statistics.

6.1.1. Model Static Skills

Here we discuss a more appropriate metric skill score designed to investigate how the Kp forecast varies during the passage of a CME (i.e. within the active solar wind period). This

new skill criterion also addresses the limitation of statistics in the prior NOAA/SWPC methodology.

In this article, the CCMC/SWRC Kp prediction is considered as the baseline to which our predicted $Kp^{(BSS)}$ is compared. On a practical sense, the long lead-time Kp predictions by CCMC/SWRC are generally a static maximum Kp value that will be reached at some point within the uncertainty of the arrival time. This methodology by CCMC/SWRC provides a comparable forecast to NOAA/SWPC alerts.

Table 3 displays the contingency table for each CME when a single event is a 3-hour Kp interval and a 'Hit' is defined when both observed and predicted (BSS) have $Kp \geq 5$. (I.e., $Kp^{(o)} \geq 5$ and $Kp^{(BSS)} \geq 5$). A 'Correct Null' is defined when neither the observed nor predicted Kp is found to be ≥ 5 i.e., meets or exceeds the NOAA G1 storm level.

Although Kp is defined for 3 hour periods, predicted variations in the IMF are made on shorter time scales, and can influence the level of geomagnetic activity (e.g. geomagnetically induced currents). Hence, averaging the IMF over intervals that match those of Kp may suppress features that are important drivers of geomagnetic activity. Nevertheless, for the purpose of this paper, we choose to average the Kp estimate from the field vectors predicted by the BSS model over the same 3-hour intervals as the Kp values; as this currently represent the service provided by NOAA/SWPC. NOAA/SWPC storm level provides a consistent threshold for future development, without which ambiguity may persist in defining a succesful forecast threshold magnitude for magnetic field components, e.g. By or Bz.

The comparison between observed and predicted Kp is performed after time shifting the $Kp^{(BSS)}$ to a region of maximum correlation. The start time of the predicted $Kp^{(BSS)}$

is placed at the estimated observed storm 'sudden commencement' and allowed to shift to maximize the correlation with the observed, with a maximum time shift restricted to 9 hours (3 Kp events). By enabling this time adjustment, each forecast method will provide their best possible skill metrics, without introducing skill score uncertainty that is driven by the uncertainty of arrival time predictions.

For the January 2014 event, the CME arrival time is poorly defined since this made a glancing encounter with the Earth and there are no clear signatures of CME arrival. In fact the delayed arrival suggested by *Mays et al.* [2015b] falls outside the time shifting window around the official initially forecasted arrival. For completeness, the skill scores for the later arrival time of this CME (i.e., a different assumed 'active solar wind' period) were estimated for all the different 'Hit' criteria discussed below. In every case, no changes to the skill metrics were observed.

6.1.2. Stricter criteria

Here we consider defining a 'Hit' criteria under a stricter regime that not only predicts the occurance of a geomagnetic storm, but also correctly predictes the storm size under the NOAA G-scale. As the predicted size of an observed storm is relevant to stakeholders, testing this 'Hit' criteria is insightful for future forecasting development.

We define a 'Hit' both when either the observed or predicted goes above the threshold (defined as $Kp \geq 5$), and when the observed and predicted Kp values are of similar values to each other. In this case, the requirement is such that $|Kp^{(o)} - Kp^{(BSS)}| \leq 1.5$. A more stringent criterion of $|Kp^{(o)} - Kp^{(BSS)}| \leq 1.0$ is also investigated. The Kp difference of 1.5 is tested as there is evidence that a limitation in accuracy is present in the underlying empirical Kp formulation [*Mays et al.*, 2015a]. However, there will naturally be a long-

term desire to correctly forecast the precise G-scale storm classification, which drives the most stringent Hit criteria in the article. The contingency table and skill metrics for these results using the BSS model and CCMC/SWRC predictions are shown in table 4.

By inspecting the metrics of both the CCMC/SWRC and BSS models we note a number of interesting insights [e.g., *Barnes and Leka*, 2008; *Wilks*, 2011]. Under a time-static forecast regime, CCMC/SWRC has a perfect hit rate. With closer inspection, the results are extremely biased and significantly over-forecast, which also leads to a large false alarm ratio. As a result, the CCMC/SWRC displays no skill under a Hanssen-Kuiper discriminant (True skill score test). Alternatively, although the BSS model shows a lower hit rate, the results are significantly closer to being unbias; with the True skill score varying between 0.4 and 0.6 for the two stringent criteria described above.

6.2. Time-varying Skills

Here we correct the time-static forecast model that led to zero skill level under the TSS test. Thus the single value for an entire geomagnetic storm is improved by incorporating the predicted time-varying Kp values for both NOAA/SWPC and CCMC/SWRC forecasts.

Both NOAA/SWPC and CCMC/SWRC have a methodology that provides time-varying Kp values with more than 24-hours in advance. NOAA/SWRC based their technique with more emphasis on a heuristic approach of the on-duty forecasters skill, while CCMC/SWRC drive their estimates using the ENLIL simulation and equations 1 and 2. To test the Kp formulae, a separate baseline approach can be directly taken from the real-time in-situ data(T0) and from solar wind conditions 27 days earlier (T27). Therefore it is possible to compare these four time-varying predictions with the BSS method.

Table 5 displays the overall metric skills for T0, T27, BSS, NOAA/SWPC and CCMC/SWRC. The compilation of the scores are for the three 'Hit' thresholds used earlier. For both NOAA/SWPC and CCMC/SWRC, the first possible release of the time-varying Kp values with CME information were used to compare against the observed. For both NOAA/SWPC and CCMC/SWRC the six CME events between 2012 and 2014 were used in estimating the skill scores, because data prior to 2012 were unavailable. The uncertainty and averages were estimated using the jackknife method from all 28 combinations of choosing six CMEs from the total of eight CMEs available for investigation. For NOAA/SWPC and CCMC/SWRC, the uncertainty was determined from 20 combinations of choosing three CMEs from the total of six CME events.

The skill score for the BSS display a general improvement on the T0, T27, NOAA/SWPC and CCMC/SWRC methodologies when we focus on the more stringent requirements that the predicted and observed magnitudes must be of comparable size. The skill of the T0 baseline also significantly and smoothly decreases as the 'Hit' criteria becomes more stringent, which is approximately consistent with recent results by *Mays et al.* [2015a], the Kp empirical formulation is accurate to about Kp=1.5.

7. Discussion

7.1. BSS model

The analysis of varying the predicted magnetic field within a CME for all eight events (figures not all shown) showed that the majority of the effects on $Kp^{(BSS)}$ were caused by the changes to the CME velocity and field strength. In our case study January 2014 event, where the velocity was significantly overestimated during the real-time forecasting

process, the uncertainty in velocity had a larger effect on the $Kp^{(BSS)}$ than the field strength.

For CME D, in June 2012, the predictive limitations of the very strong field strength reduced the accuracy of the Kp predictions. Prior evidence has suggested the reason be due to southerly IMF strength being the most dominant solar wind parameter for driving geomagnetic activity at Earth. The strong correlation of solar wind speed to geomagnetic activity is evident during CME conditions, but is only considered the most dominant parameter during high speed streams [*Holappa et al.*, 2014].

Thus we have chosen to focus our metrics towards estimating the $Kp^{(BSS)}$ variability with the magnetic field strength. This simplified method for uncertainty analytics is chosen as a practical solution to quickly implement within forecasting procedures. With the hope of more sophisticated techniques being designed as a larger number of events are investigated and a suitable methodology for direct comparison between the predicted and observed magnetic vectors can be developed.

As a simple test of the magnetic Bz component, a mean square error (MSE) was calculated between the real time in-situ data (T0) and a baseline predictor created by using the solar wind conditions 27 days earlier (T-27). A similar MSE was estimated with the BSS prediction model values and both MSE's were compared while using 10 minute resolution. Using Laplace's law of succession for small sample sizes [*Wilson*, 1927], the BSS technique is likely to provide a more accurate prediction for 60% of CMEs when estimating the magnetic field Bz component and clock angle. However, these results would strongly benefit from a larger sample size.

Figures 5-10 illustrate the effects of changing the hypothetical Earth-trajectory through the magnetic flux rope model. For cases when reliable triangulation of the CME direction is not possible and reliance of CME tilt is determined by solar disc imagery, then large differences in the hypothetical trajectory are possible. This does have a significant impact on the estimated magnetic vectors at L1.

Assuming multi-spacecraft analysis of a CME is possible, a smaller range of uncertainty in the hypothetical trajectory is possible. Small increases in the impact parameter (Y_0, perpendicular to the FR axis) causes the angular change of the vectors during passage of the CME to decrease. For an event where the spacecraft trajectory is close to the central axis of the flux rope, the field vectors are relatively unaffected by small uncertainties in the trajectory, whereas larger changes occur for trajectories well outside the central axis. However, the largest geomagnetic impacts result from near-central axis encounters for which the field vectors are evidently less influenced by the spacecraft trajectory.

By using statistical analysis *Riley and Richardson* [2013] suggest that a distinct magnetic flux rope object may be contained within an overall propagating CME structure and that the plasma-β is a good predictor variable. *Savani et al.* [2013b] used simulations to confirm that such a scenario of a distinct coherent core obstacle can occur within the CME structure and that a transition in plasma-β can be observed. Therefore the BSS model contains a region outside the central axis region but within the 'volume of influence' which is considered to be also composed of the sheath solar wind that has been disturbed by interactions with the CME and contains magnetic field lines that are draped around the CME, or CME plasma which is outside the flux rope structure.

The constant alpha flux rope model [*Lepping et al.*, 1990; *Savani et al.*, 2013a] is often used as a benchmark structure from which more sophisticated models are tested. Events that deviate away from this ideal topology are frequently observed. Allowing future models to change the alpha parameter is a simple adjustment that can replicate a non-symmetric time-series of the CME profile. Reducing the alpha parameter may also provide a simple solution to replicate possible flux erosion that has been observed during the CME evolution [*Ruffenach et al.*, 2012, 2015]. Panels E and F in Figure 1 display two clear examples of where the CME field magnitude is highly non-symmetric.

Figures 1 and 3, Panel A, displays a CME-driven storm in April 2010 with maximum observed Kp=8-. However, maximum Kp occurred just prior to the arrival of the magnetic flux rope, and was due to strong southward fields in the sheath upstream of the flux rope. This event highlights the need to forecast the field vectors not only in the flux rope but also in the sheath [e.g., *Huttunen and Koskinen*, 2004].

Figures 1 and 3, Panel D, displays a geomagnetic storm during June 2012. This event displays unusually strong magnetic field magnitude that illustrates another problem with the field vector prediction scheme discussed here. This field strength was due to a complex CME-CME interaction. The BSS model is not ideal under such cases, but can indirectly adapt to such a situation if a real-time WSA-ENLIL+cone forecast is provided with an enhanced field strength from interacting CMEs. A CCMC/SWRC forecasting of a CME-CME interaction with the WSA-ENLIL+cone run was not performed, but is likely to increase the magnetic field strength inferred from the model. The uncertainty analysis for the BSS model also accounts for situations of large magnetic field by way of the 2σ field variation.

7.2. Kp Model Skills

In this article we have focused on whether the BSS model can predict the Kp index during the passage of CMEs, and have estimated skill scores for these predictions. It is also important to be able to predict Kp during the intervening quiet periods. This suggests that when estimating the skill of a complete forecasting system, it would be useful to measure a sets of skills in two-steps: 1. 24-hour probability of detection skill for the periods of quiet and active (or disturbed) solar wind; and 2. a skill estimated using higher time resolution within the active period. Under a proposed 2-step skill scoring system for forecasting, the skill of any CME prediction model can be tested at higher time resolution during important storm epochs where the Kp variability persists within the G1-G5 storm range.

We first investigated how our 8 CME case study events measured against the official U.S. government metrics used by NOAA/SWPC. The results and geomagnetic forecast accuracy from our small sample size were similar to the national annual trends measured by NOAA/SWPC, thereby suggesting our event sample is typical.

Then we gathered skill scores metrics using more complex 'Hit' criteria which centered around defining each event to be the same 3-hour duration as a Kp value. We defined three 'Hit' criteria: 1. When both the observed or predicted goes above the threshold (defined as $Kp \geq 5$); 2. When either the observed and predicted Kp values are above 5 and $|Kp^{(o)} - Kp^{(BSS)}| \leq 1.5$; 3. When either the observed and predicted Kp values are above 5 and $|Kp^{(o)} - Kp^{(BSS)}| \leq 1.0$.

We compared 5 different prediction model techniques. Two were baseline methodolgies defined as: 1. T0, the real-time data from L1 spacecraft; 2. T27, a L1 data time-series

taken 27 days prior to the CME arrival. The other three methods were the NOAA/SWPC 3-day forecasts, CCMC/SWRC forecast tests and the BSS prediction model. The full collection of metrics are displayed in table 5.

Unsuprisingly, the baseline T0 model (i.e. real-time prediction) performed with the best skill for a significant number of the metrics tested, while the T27 was more frequently the worst performing and often showing little improvement over having no skill. The T0 model Hit rate was above 70% for all other than the most stringent Hit criteria where the success dropped to above 50%. The T0 model does show a dramatic increase in the false alarm ratio (FAR) as the Hit criteria became more stringent, and this can be partially explained by the significant increase in the bias to over-forecast (FBI). The decrease in skill can be partly interpreted as a limitation in accuracy of the underlying empirical relationship between the magnetic vectors to the kp Index.

The Hit rate (POD) for the 3-day lead-time NOAA/SWPC forecast, where we investigate the individual Kp values during a geomagnetic storm, shows similar results to their official metrics (see Section 6.1). Moreover, these results appear robust to the more stricter Hit criteria. However, the large false alarm ratio and false alarm rate with very bias over-forecasting has led to an extremely low True skill score (TSS) value.

The CCMC/SWRC demonstrated a Hit rate as consistently good (above 70%) as the real-time T0 forecast model. However, this model also showed the worst over-forecasts bias against all models under all Hit criteria. The threat score (which does not account for correct nulls) also remained strong and out performed the T0 model at the more stringent Hit criteria. The TSS value remained consistently high and again out performed the T0

model; but the uncertainty in this value remains high and improvement to the statistics is warranted.

The new BSS model predictions have performed consistently amongst the best metrics demonstrated by the other models. Across all Hit crtieria, the BSS predictions, have consistently demonstrated the most unbiased results. Under the stricter Hit criteria, the Hit rate of the BSS predictions are comparable to the T0 predictions but not as good as the CCMC/SWRC predictions. The True skill score under the least stringent Hit criteria, were not as good as the T0 or CCMS/SWRC predictions. However, under the stricter Hit criteria, the BSS model predictions beat the T0 predictions and performed equally with CCMC/SWRC while demonstrating a lower uncertainty.

8. Conclusions and Future Work

In this article, we test the robustness of the time-varying magnetic vector predictions (BSS model) in relation to an empirically-driven model of the geomagnetic Kp index ($Kp^{(BSS)}$ forecast). We conclude that quantifying the uncertainty in this experimental Kp forecast is best achieved by varying the maximum magnetic field strength and the associated velocity within the predicted CME. Uncertainty in the field strength represents the largest obstacle to replicating the in situ field vectors and providing a reliable Kp forecast. However, the velocity estimate also has a significant impact of the predicted Kp values using equations 1 and 2.

The implementation of BSS scheme, described in Paper 1, as a component within a heuristic approach is also expected to qualitatively aid space weather forecasters, and provide scientific insight by statistically quantifying the terrestrial response for future events. Further statistical analysis of more events both from a historical data set [e.g.

Marubashi et al., 2015] and ongoing real-time CMEs without discrimination will build up the necessary statistics for skill score estimates using the BSS scheme.

Time variations in the Kp should also be considered when assessing the reliability of forecast. We find that the time of peak predicted Kp value can vary by more than 6 hours, depending on the time variation of the magnetic field direction and strength. Considering the uncertainty in the best estimates for CME arrival times are also approximately 6 hours, the time-varying nature of the BSS prediction can provide significant assistance towards geomagnetic forecasting.

Other sources of error lie in the estimated CME speed, direction and size. These parameters can be estimated most reliably through triangulation using CME observations from multiple spacecraft, as discussed in Paper 1. In the future, observations from a spacecraft at L5 could be utilized in addition to observations at L1 [*Vourlidas*, 2015].

By treating forecasting skill as two separate measures of 'all solar wind' and 'within active solar wind' periods, then the BSS tool for predicting the magnetic vectors shows great promise towards better forecasting within the 'active solar wind' period. Using various skill metrics (tabulated in table 5), the BSS model predictions were the most unbias results. Out of the 5 models tested, two of which were baseline tests, four models were consistently over-forecasting. The BSS model retained a more consistently strong skill score for the more stringent 'Hit' criteria, while also displaying significantly lower uncertainty. The True skill score for the BSS model prediction under the stricter Hit criteria ($(Kp^{(o)} \cup Kp^{(BSS)}) \geq 5$ and $|Kp^{(o)} - Kp^{(BSS)}| \leq 1$) beat the T0 real-time in situ data model prediction and performed equally with CCMC/SWRC while demonstrating a lower uncertainty.

Further statistical analysis of both historical data sets alongside tests within a real-world forecasting setting will go a long way to supporting the skill scores and the uncertainty estimated here.

Appendix A: Bz Component of the magnetic field

Figure 12 displays the Bz component of the solar wind magnetic field at L1 in cartesian GSE coordinates. The estimates from the BSS model are overplotted in red. The field displayed in Figure 1 represents the full coordinate system, but in spherical coordinates and not cartesian. As space weather concerns are often described by the magnitude of a cartesian based Bz component driver, this figure below is provided to assist the readership.

Acknowledgments. This work benefited from the NASA/LWS Coordinated Data Analysis Workshops on CME flux ropes in 2011. We acknowledge the workshop support provided by NASA/LWS, Predictive Sciences Inc. (San Diego, CA), University of Alcalá (Alcalá de Henares, Spain), and Ministerio de Ciencia e Innovacin (Reference number AYA2010-12439-E), Spain.

N.P.S. and A.V. were supported by NASA SECCHI Contract to NRL. N.P.S and T.N.C. were supported by NSF SHINE award 1433202 and 1433086, respectively. The authors wish to acknowledge Lutz Rastaetter and the rest of the CCMC staff for their generous support throughout the work discussed in the paper.

This work was supported by NASA grant NNH14AX40I. N.P.S. was also supported by NASA LWS Institute 2015 for geomagnetically induced currents, and the SCOSTEP Varsiti program.

We thank S. Patsourakos (University of Ioannina) for constructive comments about active region helicity. We thank C. Balch (NOAA/SWPC) for providing historical NOAA forecast alerts. K.D. Leka (NorthWest Research Associates) for her advise and expertise with Skill score metrics. And, last but not least, W. Thompson (Adnet Systems inc.) for assistance with SECCHI products within SolarSoft.

We thank the Space Physics Data Facility of NASAs Goddard Space Flight Center for use of OMNIWeb service and OMNI data.

References

Barnes, G., and K. D. Leka (2008), Evaluating the Performance of Solar Flare Forecasting Methods, *Astrophys. J.*, , *688*, L107–L110, doi:10.1086/595550.

Bothmer, V., and D. M. Rust (1997), The field configuration of magnetic clouds and the solar cycle, *Washington DC American Geophysical Union Geophysical Monograph Series*, *99*, 139–146, doi:10.1029/GM099p0139.

Bothmer, V., and R. Schwenn (1994), Eruptive prominences as sources of magnetic clouds in the solar wind, *Space Science Reviews*, *70*, 215–220, doi:10.1007/BF00777872.

Bothmer, V., and R. Schwenn (1998), The structure and origin of magnetic clouds in the solar wind, *Annales Geophysicae*, *16*, 1–24, doi:10.1007/s00585-997-0001-x.

Bothmer, V., J. Hesemann, M. Venzmer, and E. Bosman (2013), CME arrival, Kp, aurora and GNSS-error forecast.

Colaninno, R. C., A. Vourlidas, and C. C. Wu (2013), Quantitative comparison of methods for predicting the arrival of coronal mass ejections at Earth based on multiview imaging, *Journal of Geophysical Research (Space Physics)*, *118*, 6866–6879, doi:

10.1002/2013JA019205.

Cremades, H., and V. Bothmer (2004), On the three-dimensional configuration of coronal mass ejections, *Astron. Astrophys.*, , *422*, 307–322, doi:10.1051/0004-6361:20035776.

Crosby, N. B., A. Veronig, E. Robbrecht, B. Vrsnak, S. Vennerstrom, O. Malandraki, S. Dalla, L. Rodriguez, N. Srivastava, M. Hesse, D. Odstrcil, and Comesep Consortium (2012), Forecasting the space weather impact: The COMESEP project, in *American Institute of Physics Conference Series*, *American Institute of Physics Conference Series*, vol. 1500, edited by Q. Hu, G. Li, G. P. Zank, X. Ao, O. Verkhoglyadova, and J. H. Adams, pp. 159–164, doi:10.1063/1.4768760.

Démoulin, P., and S. Dasso (2009), Magnetic cloud models with bent and oblate cross-section boundaries, *Astron. Astrophys.*, , *507*, 969–980, doi:10.1051/0004-6361/200912645.

Dungey, J. W. (1961), Interplanetary Magnetic Field and the Auroral Zones, *Physical Review Letters*, *6*, 47–48, doi:10.1103/PhysRevLett.6.47.

Emmons, D., A. Acebal, A. Pulkkinen, A. Taktakishvili, P. MacNeice, and D. Odstrcil (2013), Ensemble forecasting of coronal mass ejections using the WSA-ENLIL with CONED Model, *Space Weather*, *11*, 95–106, doi:10.1002/swe.20019.

Gopalswamy, N., A. Lara, S. Yashiro, M. L. Kaiser, and R. A. Howard (2001), Predicting the 1-AU arrival times of coronal mass ejections, *J. Geophys. Res.*, , *106*, 29,207–29,218, doi:10.1029/2001JA000177.

Hale, G. E. (1925), A Test of the Electromagnetic Theory of the Hydrogen Vortices Surrounding Sun-Spots, *Proceedings of the National Academy of Science*, *11*, 691–696, doi:10.1073/pnas.11.11.691.

Harra, L. K., N. U. Crooker, C. H. Mandrini, L. van Driel-Gesztelyi, S. Dasso, J. Wang, H. Elliott, G. Attrill, B. V. Jackson, and M. M. Bisi (2007), How Does Large Flaring Activity from the Same Active Region Produce Oppositely Directed Magnetic Clouds?, *Solar Physics*, *244*, 95–114, doi:10.1007/s11207-007-9002-x.

Hidalgo, M. A., T. Nieves-Chinchilla, and C. Cid (2002), Elliptical cross-section model for the magnetic topology of magnetic clouds, *Geophys. Res. Lett.*, , *29*(13), 1637, doi:10.1029/2001GL013875.

Holappa, L., K. Mursula, and T. Asikainen (2014), A new method to estimate annual solar wind parameters and contributions of different solar wind structures to geomagnetic activity, *Journal of Geophysical Research (Space Physics)*, *119*, 9407–9418, doi:10.1002/2014JA020599.

Hu, Q., and B. U. Ö. Sonnerup (2002), Reconstruction of magnetic clouds in the solar wind: Orientations and configurations, *Journal of Geophysical Research (Space Physics)*, *107*, 1142, doi:10.1029/2001JA000293.

Huttunen, K., and H. Koskinen (2004), Importance of post-shock streams and sheath region as drivers of intense magnetospheric storms and high-latitude activity, *Annales Geophysicae*, *22*, 1729–1738, doi:10.5194/angeo-22-1729-2004.

Iju, T., M. Tokumaru, and K. Fujiki (2014), Kinematic Properties of Slow ICMEs and an Interpretation of a Modified Drag Equation for Fast and Moderate ICMEs, *Solar Physics*, *289*, 2157–2175, doi:10.1007/s11207-014-0472-3.

Kay, C., M. Opher, and R. M. Evans (2013), Forecasting a Coronal Mass Ejection's Altered Trajectory: ForeCAT, *Astrophys. J.*, , *775*, 5, doi:10.1088/0004-637X/775/1/5.

Kunkel, V., and J. Chen (2010), Evolution of a Coronal Mass Ejection and its Mag-

netic Field in Interplanetary Space, *Astrophys. J.,* , *715*, L80–L83, doi:10.1088/2041-8205/715/2/L80.

Leamon, R. J., R. C. Canfield, and A. A. Pevtsov (2002), Properties of magnetic clouds and geomagnetic storms associated with eruption of coronal sigmoids, *Journal of Geophysical Research (Space Physics)*, *107*, 1234, doi:10.1029/2001JA000313.

Leamon, R. J., R. C. Canfield, S. L. Jones, K. Lambkin, B. J. Lundberg, and A. A. Pevtsov (2004), Helicity of magnetic clouds and their associated active regions, *Journal of Geophysical Research (Space Physics)*, *109*, A05106, doi:10.1029/2003JA010324.

Lepping, R. P., L. F. Burlaga, and J. A. Jones (1990), Magnetic field structure of interplanetary magnetic clouds at 1 AU, *J. Geophys. Res.,* , *95*, 11,957–11,965, doi:10.1029/JA095iA08p11957.

Lepping, R. P., M. H. Acũna, L. F. Burlaga, W. M. Farrell, J. A. Slavin, K. H. Schatten, F. Mariani, N. F. Ness, F. M. Neubauer, Y. C. Whang, J. B. Byrnes, R. S. Kennon, P. V. Panetta, J. Scheifele, and E. M. Worley (1995), The Wind Magnetic Field Investigation, *Space Science Reviews*, *71*, 207–229, doi:10.1007/BF00751330.

Lepping, R. P., D. B. Berdichevsky, C.-C. Wu, A. Szabo, T. Narock, F. Mariani, A. J. Lazarus, and A. J. Quivers (2006), A summary of WIND magnetic clouds for years 1995-2003: model-fitted parameters, associated errors and classifications, *Annales Geophysicae*, *24*, 215–245, doi:10.5194/angeo-24-215-2006.

Liu, Y., D. F. Webb, and X. P. Zhao (2006), Magnetic Structures of Solar Active Regions, Full-Halo Coronal Mass Ejections, and Geomagnetic Storms, *Astrophys. J.,* , *646*, 1335–1348, doi:10.1086/505036.

Liu, Y. D., Z. Yang, R. Wang, J. G. Luhmann, J. D. Richardson, and N. Lugaz (2014),

Sun-to-Earth Characteristics of Two Coronal Mass Ejections Interacting Near 1 AU: Formation of a Complex Ejecta and Generation of a Two-step Geomagnetic Storm, *Astrophys. J.,* , *793*, L41, doi:10.1088/2041-8205/793/2/L41.

Lockwood, M., L. Barnard, H. Nevanlinna, M. J. Owens, R. G. Harrison, A. P. Rouillard, and C. J. Davis (2013), Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yr - Part 1: A new geomagnetic data composite, *Annales Geophysicae*, *31*, 1957–1977, doi:10.5194/angeo-31-1957-2013.

Longcope, D., C. Beveridge, J. Qiu, B. Ravindra, G. Barnes, and S. Dasso (2007), Modeling and Measuring the Flux Reconnected and Ejected by the Two-Ribbon Flare/CME Event on 7 November 2004, *Solar Physics*, *244*, 45–73, doi:10.1007/s11207-007-0330-7.

Lugaz, N., and C. J. Farrugia (2014), A new class of complex ejecta resulting from the interaction of two CMEs and its expected geoeffectiveness, *Geophys. Res. Lett.,* , *41*, 769–776, doi:10.1002/2013GL058789.

Manchester, W. B., IV, B. van der Holst, and B. Lavraud (2014), Flux rope evolution in interplanetary coronal mass ejections: the 13 May 2005 event, *Plasma Physics and Controlled Fusion*, *56*(6), 064006, doi:10.1088/0741-3335/56/6/064006.

Marubashi, K. (1997), Interplanetary magnetic flux ropes and solar filaments, in *Coronal Mass Ejections*, *Geophys. Monogr. Ser.*, vol. 99, edited by N. Crooker, J. A. Joselyn, and J. Feynman, pp. 147–156, AGU, Washington, D. C., doi:10.1029/GM099p0147.

Marubashi, K., S. Akiyama, S. Yashiro, N. Gopalswamy, K.-S. Cho, and Y.-D. Park (2015), Geometrical Relationship Between Interplanetary Flux Ropes and Their Solar Sources, *Solar Physics*, *290*, 1371–1397, doi:10.1007/s11207-015-0681-4.

Mays, M. L., A. Taktakishvili, A. Pulkkinen, P. J. MacNeice, L. Rastätter, D. Odstrcil, L. K. Jian, I. G. Richardson, J. A. LaSota, Y. Zheng, and M. M. Kuznetsova (2015a), Ensemble Modeling of CMEs Using the WSA-ENLIL+Cone Model, *Solar Physics*, doi: 10.1007/s11207-015-0692-1.

Mays, M. L., B. Thompson, L. K. Jian, R. Colaninno, D. Odstrcil, C. Möstl, M. Temmer, N. Savani, G. Collinson, A. Taktakishvili, P. J. MacNeice, and Y. Zheng (2015b), Propagation of the 7 January 2014 CME and Resulting Geomagnetic Non-Event, *Astrophys. J.,* , *812*, 2, doi:10.1088/0004-637X/812/2/145.

Newell, P. T., T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich (2007), A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, *Journal of Geophysical Research (Space Physics)*, *112*, A01206, doi: 10.1029/2006JA012015.

Nieves-Chinchilla, T., A. Vourlidas, G. Stenborg, N. P. Savani, A. Koval, A. Szabo, and L. K. Jian (2013), Inner Heliospheric Evolution of a "Stealth" CME Derived from Multiview Imaging and Multipoint in Situ observations. I. Propagation to 1 AU, *Astrophys. J.,* , *779*, 55, doi:10.1088/0004-637X/779/1/55.

NOAA, Space Weather Prediction Center, (2013), SpaceWx GPRA January 2013, Available from: http://www.swpc.noaa.gov/sites/default/files/images/u30/swpc%20gpra%20metric%20description.pdf. (Accessed: 08/09/2016)

NOAA, Space Weather Prediction Center, (2014), Max Kp and GPRA, Available from: http://www.swpc.noaa.gov/sites/default/files/images/u30/max%20kp%20and%20gpra.pdf (Accessed: 08/09/2016)

Odstrcil, D. (2003), Modeling 3-D solar wind structure, *Advances in Space Research*, *32*, 497–506, doi:10.1016/S0273-1177(03)00332-6.

Odstrcil, D., and V. J. Pizzo (2009), Numerical Heliospheric Simulations as Assisting Tool for Interpretation of Observations by STEREO Heliospheric Imagers, *Solar Physics*, *259*, 297–309, doi:10.1007/s11207-009-9449-z.

Owens, M. J. (2008), Combining remote and in situ observations of coronal mass ejections to better constrain magnetic cloud reconstruction, *Journal of Geophysical Research (Space Physics)*, *113*, A12102, doi:10.1029/2008JA013589.

Owens, M. J., V. G. Merkin, and P. Riley (2006), A kinematically distorted flux rope model for magnetic clouds, *Journal of Geophysical Research (Space Physics)*, *111*, A03104, doi:10.1029/2005JA011460.

Owens, M. J., P. Démoulin, N. P. Savani, B. Lavraud, and A. Ruffenach (2012), Implications of Non-cylindrical Flux Ropes for Magnetic Cloud Reconstruction Techniques and the Interpretation of Double Flux Rope Events, *Solar Physics*, *278*, 435–446, doi:10.1007/s11207-012-9939-2.

Owens, M. J., T. S. Horbury, R. T. Wicks, S. L. McGregor, N. P. Savani, and M. Xiong (2014), Ensemble downscaling in coupled solar wind-magnetosphere modeling for space weather forecasting, *Space Weather*, *12*, 395–405, doi:10.1002/2014SW001064.

Poomvises, W., N. Gopalswamy, S. Yashiro, R.-Y. Kwon, and O. Olmedo (2012), Determination of the Heliospheric Radial Magnetic Field from the Standoff Distance of a CME-driven Shock Observed by the STEREO Spacecraft, *Astrophys. J.,* , *758*, 118, doi:10.1088/0004-637X/758/2/118.

Rees, A., and R. J. Forsyth (2003), Magnetic clouds with east/west orientated axes observed by Ulysses during solar cycle 23, *Geophys. Res. Lett.*, , *30*, 8030, doi:10.1029/2003GL017296.

Riley, P., and N. U. Crooker (2004), Kinematic Treatment of Coronal Mass Ejection Evolution in the Solar Wind, *Astrophys. J.*, , *600*, 1035–1042, doi:10.1086/379974.

Riley, P., and I. G. Richardson (2013), Using Statistical Multivariable Models to Understand the Relationship Between Interplanetary Coronal Mass Ejecta and Magnetic Flux Ropes, *Solar Physics*, *284*, 217–233, doi:10.1007/s11207-012-0006-9.

Ruffenach, A., B. Lavraud, M. J. Owens, J.-A. Sauvaud, N. P. Savani, A. P. Rouillard, P. Démoulin, C. Foullon, A. Opitz, A. Fedorov, C. J. Jacquey, V. Génot, P. Louarn, J. G. Luhmann, C. T. Russell, C. J. Farrugia, and A. B. Galvin (2012), Multispacecraft observation of magnetic cloud erosion by magnetic reconnection during propagation, *Journal of Geophysical Research (Space Physics)*, *117*, A09101, doi:10.1029/2012JA017624.

Ruffenach, A., B. Lavraud, C. J. Farrugia, P. Démoulin, S. Dasso, M. J. Owens, J.-A. Sauvaud, A. P. Rouillard, A. Lynnyk, C. Foullon, N. P. Savani, J. G. Luhmann, and A. B. Galvin (2015), Statistical study of magnetic cloud erosion by magnetic reconnection, *Journal of Geophysical Research (Space Physics)*, *120*, 43–60, doi:10.1002/2014JA020628.

Rust, D. M. (1994), Spawning and shedding helical magnetic fields in the solar atmosphere, *Geophys. Res. Lett.*, , *21*, 241–244, doi:10.1029/94GL00003.

Savani, N. P., M. J. Owens, A. P. Rouillard, R. J. Forsyth, and J. A. Davies (2010), Observational Evidence of a Coronal Mass Ejection Distortion Directly Attributable to a Structured Solar Wind, *Astrophys. J.*, , *714*, L128–L132, doi:10.1088/2041-

8205/714/1/L128.

Savani, N. P., M. J. Owens, A. P. Rouillard, R. J. Forsyth, K. Kusano, D. Shiota, and R. Kataoka (2011a), Evolution of Coronal Mass Ejection Morphology with Increasing Heliocentric Distance. I. Geometrical Analysis, *Astrophys. J.*, , *731*, 109, doi:10.1088/0004-637X/731/2/109.

Savani, N. P., M. J. Owens, A. P. Rouillard, R. J. Forsyth, K. Kusano, D. Shiota, R. Kataoka, L. Jian, and V. Bothmer (2011b), Evolution of Coronal Mass Ejection Morphology with Increasing Heliocentric Distance. II. In Situ Observations, *Astrophys. J.*, , *732*, 117, doi:10.1088/0004-637X/732/2/117.

Savani, N. P., J. A. Davies, C. J. Davis, D. Shiota, A. P. Rouillard, M. J. Owens, K. Kusano, V. Bothmer, S. P. Bamford, C. J. Lintott, and A. Smith (2012), Observational Tracking of the 2D Structure of Coronal Mass Ejections Between the Sun and 1 AU, *Solar Physics*, *279*, 517–535, doi:10.1007/s11207-012-0041-6.

Savani, N. P., A. Vourlidas, A. Pulkkinen, T. Nieves-Chinchilla, B. Lavraud, and M. J. Owens (2013a), Tracking the momentum flux of a CME and quantifying its influence on geomagnetically induced currents at Earth, *Space Weather*, *11*, 245–261, doi:10.1002/swe.20038.

Savani, N. P., A. Vourlidas, D. Shiota, M. G. Linton, K. Kusano, N. Lugaz, and A. P. Rouillard (2013b), A Plasma β Transition within a Propagating Flux Rope, *Astrophys. J.*, , *779*, 142, doi:10.1088/0004-637X/779/2/142.

Savani, N. P., A. Vourlidas, A. Szabo, M. L. Mays, B. J. Thompson, I. G. Richardson, R. Evans, A. Pulkkinen, and T. Nieves-Chinchilla (2015), Predicting the magnetic vectors within coronal mass ejections arriving at Earth: 1. Initial Architecture, *ArXiv*

e-prints, doi:10.1088/0004-637X/779/2/142.

Shi, T., Y. Wang, L. Wan, X. Cheng, M. Ding, and J. Zhang (2015), Predicting the Arrival Time of Coronal Mass Ejections with the Graduated Cylindrical Shell and Drag Force Model, *Astrophys. J.,* , *806*, 271, doi:10.1088/0004-637X/806/2/271.

Shiota, D., and R. Kataoka (2016), Magnetohydrodynamic simulation of interplanetary propagation of multiple coronal mass ejections with internal magnetic flux rope (SUSANOO-CME), *Space Weather*, *14*, 56–75, doi:10.1002/2015SW001308.

Smith, C. W., J. L'Heureux, N. F. Ness, M. H. Acuña, L. F. Burlaga, and J. Scheifele (1998), The ACE Magnetic Fields Experiment, *Space Science Reviews*, *86*, 613–632, doi:10.1023/A:1005092216668.

Taktakishvili, A., P. MacNeice, and D. Odstrcil (2010), Model uncertainties in predictions of arrival of coronal mass ejections at Earth orbit, *Space Weather*, *8*, S06007, doi:10.1029/2009SW000543.

Thernisien, A., A. Vourlidas, and R. A. Howard (2009), Forward Modeling of Coronal Mass Ejections Using STEREO/SECCHI Data, *Solar Physics*, *256*, 111–130, doi:10.1007/s11207-009-9346-5.

Tsurutani, B. T., and W. D. Gonzalez (1997), The Interplanetary causes of magnetic storms: A review, in *Washington DC American Geophysical Union Geophysical Monograph Series*, *Washington DC American Geophysical Union Geophysical Monograph Series*, vol. 98, edited by B. T. Tsurutani, W. D. Gonzalez, Y. Kamide, and J. K. Arballo, pp. 77–89, doi:10.1029/GM098p0077.

Tucker-Hood, K., C. Scott, M. Owens, D. Jackson, L. Barnard, J. A. Davies, S. Crothers, C. Lintott, R. Simpson, N. P. Savani, J. Wilkinson, B. Harder, G. M. Eriksson, E. M.

L Baeten, and L. L. Wan Wah (2015), Validation of a priori CME arrival predictions made using real-time heliospheric imager observations, *Space Weather*, *13*, 35–48, doi: 10.1002/2014SW001106.

Tun, S. D., and A. Vourlidas (2013), Derivation of the Magnetic Field in a Coronal Mass Ejection Core via Multi-frequency Radio Imaging, *Astrophys. J.,* , *766*, 130, doi: 10.1088/0004-637X/766/2/130.

Vourlidas, A. (2014), The flux rope nature of coronal mass ejections, *Plasma Physics and Controlled Fusion*, *56*(6), 064001, doi:10.1088/0741-3335/56/6/064001.

Vourlidas, A. (2015), Mission to the Sun-Earth L5 Lagrangian Point: An Optimal Platform for Space Weather Research, *Space Weather*, *13*(0), doi:10.1002/2015SW001173.

Vourlidas, A., R. Colaninno, T. Nieves-Chinchilla, and G. Stenborg (2011), The First Observation of a Rapidly Rotating Coronal Mass Ejection in the Middle Corona, *Astrophys. J.,* , *733*, L23, doi:10.1088/2041-8205/733/2/L23.

Vršnak, B., T. Žic, D. Vrbanec, M. Temmer, T. Rollett, C. Möstl, A. Veronig, J. Čalogović, M. Dumbović, S. Lulić, Y.-J. Moon, and A. Shanmugaraju (2013), Propagation of Interplanetary Coronal Mass Ejections: The Drag-Based Model, *Solar Physics*, *285*, 295–315, doi:10.1007/s11207-012-0035-4.

Vršnak, B., M. Temmer, T. Žic, A. Taktakishvili, M. Dumbović, C. Möstl, A. M. Veronig, M. L. Mays, and D. Odstrčil (2014), Heliospheric Propagation of Coronal Mass Ejections: Comparison of Numerical WSA-ENLIL+Cone Model and Analytical Drag-based Model, *Astrophys. J. (Supp.),* , *213*, 21, doi:10.1088/0067-0049/213/2/21.

Wang, Y.-M. (2013), On the Strength of the Hemispheric Rule and the Origin of Active-region Helicity, *Astrophys. J.,* , *775*, L46, doi:10.1088/2041-8205/775/2/L46.

Wilks, D. S. (2011), *Statistical methods in the atmospheric sciences*, vol. 100, Academic press.

Wilson, E.B. (1927), Probable inference, the law of succession, and statistical inference, *JASA*, *22*, 209212, doi:10.1080/01621459.1927.10502953.

Wing, S., J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman, K. Costello, M. Balikhin, and K. Takahashi (2005), Kp forecast models, *Journal of Geophysical Research (Space Physics)*, *110*, A04203, doi:10.1029/2004JA010500.

Yurchyshyn, V., Q. Hu, R. P. Lepping, B. J. Lynch, and J. Krall (2007), Orientations of LASCO Halo CMEs and their connection to the flux rope structure of interplanetary CMEs, *Advances in Space Research*, *40*, 1821–1826, doi:10.1016/j.asr.2007.01.059.

Zhang, J., I. G. Richardson, D. F. Webb, N. Gopalswamy, E. Huttunen, J. C. Kasper, N. V. Nitta, W. Poomvises, B. J. Thompson, C.-C. Wu, S. Yashiro, and A. N. Zhukov (2007), Solar and interplanetary sources of major geomagnetic storms (Dst = -100 nT) during 1996-2005, *Journal of Geophysical Research (Space Physics)*, *112*, A10102, doi: 10.1029/2007JA012321.

Zhang, X.-Y., and M. B. Moldwin (2014), The source, statistical properties, and geoeffectiveness of long-duration southward interplanetary magnetic field intervals, *Journal of Geophysical Research (Space Physics)*, *119*, 658–669, doi:10.1002/2013JA018937.

Table 1. Geomagnetic Storm Parameters Driven by CMEs

CME	Flare Detection		Max Kp Index			
#	Date	Time	Observed	NOAA prediction	NASA estimate	BSS prediction
a	3 Apr 2010	2:33	8−	−	−	6
b	25 Mar 2011	16:47	3+	−	−	4+
c	10 Mar 2012	2:33	6+	7	6−	7
d	14 Jun 2012	2:49	6+	5	6−	5−
e	12 Jul 2012	7:23	7	4	7+	7+
f	27 Sep 2012	23:36	7−	6	6	4−
g	7 Jan 2014	18:04	3	7	7	4+
h	10 Sep 2014	13:57	6+	7	6	7−

Table 2. Skill Metrics for 6 Events (c-h), Using NOAA/SWPC Criteria

Skill Metrics	**Metric Range**	**Perfect Score**	**BSS**	**CCMC**	**SWPC**
Proportion Correct	$[0 < \mathrm{PC} < 1]$	1	0.67	0.83	0.67
Hit Rate, H	$[0 < \mathrm{POD} < 1]$	1	0.6	1	0.8
False Alarm Ratio	$[0 < \mathrm{FAR} < 1]$	0	0	0.17	0.2
Freq Bias Index, B	$[0 < \mathrm{FBI} < 8]$	1	0.6	1.2	1
Threat Score, TS	$[0 < \mathrm{CSI} < 1]$	1	0.6	0.83	0.67
True Skill Score	$[-1 < \mathrm{TSS} < 1]$	1	0.6	0	-0.2

Table 3. Contingency Table and Metrics for CME Events. An event is each Kp value during the predicted interval. A 'Hit' is defined as an event where the Kp = 5.

(TOTAL) BSS		**(TOTAL) CCMC**		**Skill Metrics**	**BSS**	**CCMC**
12	**7**	**20**	**45**	Proportion Correct	0.76	0.31
15	**58**	**0**	**0**	Hit Rate, H	0.44	1
				False Alarm Ratio	0.37	0.7
	Hit		False Alarm	Freq Bias Index, B	0.7	3.25
	Misses		Correct Nulls	Threat Score, TS	0.35	0.31
				True Skill Score	0.34	0
	(A) BSS		**CCMC**	**(B) BSS**	**CCMC**	
1	2	–	–	0	0	–	–
6	3	–	–	0	15	–	–
	(C) BSS		**CCMC**	**(D) BSS**	**CCMC**	
2	2	2	7	0	0	4	4
0	5	0	0	4	4	0	0
	(E) BSS		**CCMC**	**(F) BSS**	**CCMC**	
7	0	10	3	0	0	2	10
3	3	0	0	2	10	0	0
	(G) BSS		**CCMC**	**(H) BSS**	**CCMC**	
0	0	0	11	2	3	2	10
0	11	0	0	0	7	0	0

Table 4. Contingency Table and Metrics for Eight CMEs. Assuming time-static Kp prediction for CCMC.

An event is each Kp value within the predicted interval. A hit for $BSS^{(1.5)}$ is defined when either Observed or Predicted is Kp $\geq$ 5 AND $|Kp^{(o)} - KpBSS| = 1.5$. Similarly, $BSS^{(1.0)}$ is for $|Kp^{(o)} - KpBSS| = 1.0$.

$BSS^{(1.5)}$		**CCMC**†		**$BSS^{(1.0)}$**		**CCMC**†	
19	6	9	56	14	8	8	57
9	58	0	0	12	58	0	0

† The combined CCMC skill score are a combined total from 6 events between (C) and (H)

Skill Metrics	**$BSS^{(1.5)}$**	**CCMC**	**Skill Metrics**	**$BSS^{(1.0)}$**	**CCMC**
Proportion Correct	0.84	0.14	Proportion Correct	0.78	0.12
Hit Rate, H	0.68	1	Hit Rate, H	0.54	1
False Alarm Ratio	0.24	0.86	False Alarm Ratio	0.36	0.88
Freq Bias Index, B	0.89	7.2	Freq Bias Index, B	0.85	8.1
Threat Score, TS	0.56	0.14	Threat Score, TS	0.41	0.12
True Skill Score	0.58	0	True Skill Score	0.42	0

Table 5. Metric Skill Score for the CMEs. Each Kp value within the predicted CME interval is considered an event. T0 is the baseline skill score generated by implemeting the real-time solar wind conditions for each of the eight events. BSS, T0 and T27 combined all eight events; while the NOAA/SWPC and CCMC/SWRC are a result of six events after 2012. The 2-3 day lead-time time-varying Kp values are used for both NOAA and CCMC forecasts. PC = Proportion correct; POD = Probability of detection (Hit rate); FAR = False alarm ratio; FBI = Frequency bias index; CSI = Critical success index (Threat score); TSS = True skill score.

PC	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$	FBI	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$
T0	0.8 ± 0.04	0.75 ± 0.07	0.70 ± 0.07	T0	1.24 ± 0.15	1.77 ± 0.45	1.95 ± 0.49
T27	0.68 ± 0.06	0.69 ± 0.06	0.69 ± 0.06	T27	0.25 ± 0.14	0.29 ± 0.13	0.29 ± 0.15
BSS	0.76 ± 0.05	0.84 ± 0.03	0.78 ± 0.04	BSS	0.71 ± 0.14	0.90 ± 0.13	0.85 ± 0.12
NOAA	0.42 ± 0.10	0.53 ± 0.12	0.44 ± 0.11	NOAA	2.20 ± 1.78	1.57 ± 0.90	1.78 ± 1.30
CCMC	0.75 ± 0.12	0.75 ± 0.11	0.72 ± 0.10	CCMC	1.59 ± 0.80	1.84 ± 10.5	2.36 ± 1.32
POD	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$	CSI	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$
T0	0.74 ± 0.09	0.71 ± 0.05	0.53 ± 0.09	T0	0.49 ± 0.08	0.37 ± 0.10	0.23 ± 0.05
T27	0.04 ± 0.03	0.08 ± 0.02	0.08 ± 0.05	T27	0.03 ± 0.02	0.07 ± 0.02	0.07 ± 0.04
BSS	0.44 ± 0.11	0.67 ± 0.08	0.54 ± 0.09	BSS	0.35 ± 0.10	0.55 ± 0.05	0.41 ± 0.06
NOAA	0.47 ± 0.32	0.62 ± 0.27	0.45 ± 0.32	NOAA	0.17 ± 0.07	0.33 ± 0.122	0.19 ± 0.09
CCMC	0.71 ± 0.18	0.76 ± 0.133	0.78 ± 0.14	CCMC	0.43 ± 0.18	0.43 ± 0.17	0.35 ± 0.13
FAR	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$	TSS	$Kp \geq 5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.5$	$\lvert Kp^{(O)} - Kp^{(P)} \rvert \leq 1.0$
T0	0.40 ± 0.08	0.56 ± 0.13	0.71 ± 0.08	T0	0.55 ± 0.09	0.47 ± 0.09	0.25 ± 0.1
T27	0.69 ± 0.38	0.59 ± 0.35	0.62 ± 0.36	T27	-0.04 ± 0.05	0.00 ± 0.05	0.00 ± 0.07
BSS	0.38 ± 0.13	0.25 ± 0.08	0.37 ± 0.06	BSS	0.33 ± 0.12	0.58 ± 0.08	0.41 ± 0.07
NOAA	0.75 ± 0.11	0.55 ± 0.18	0.72 ± 0.13	NOAA	-0.06 ± 0.3	0.16 ± 0.28	-0.05 ± 0.30
CCMC	0.46 ± 0.23	0.49 ± 0.22	0.60 ± 0.16	CCMC	0.49 ± 0.25	0.52 ± 0.21	0.49 ± 0.20

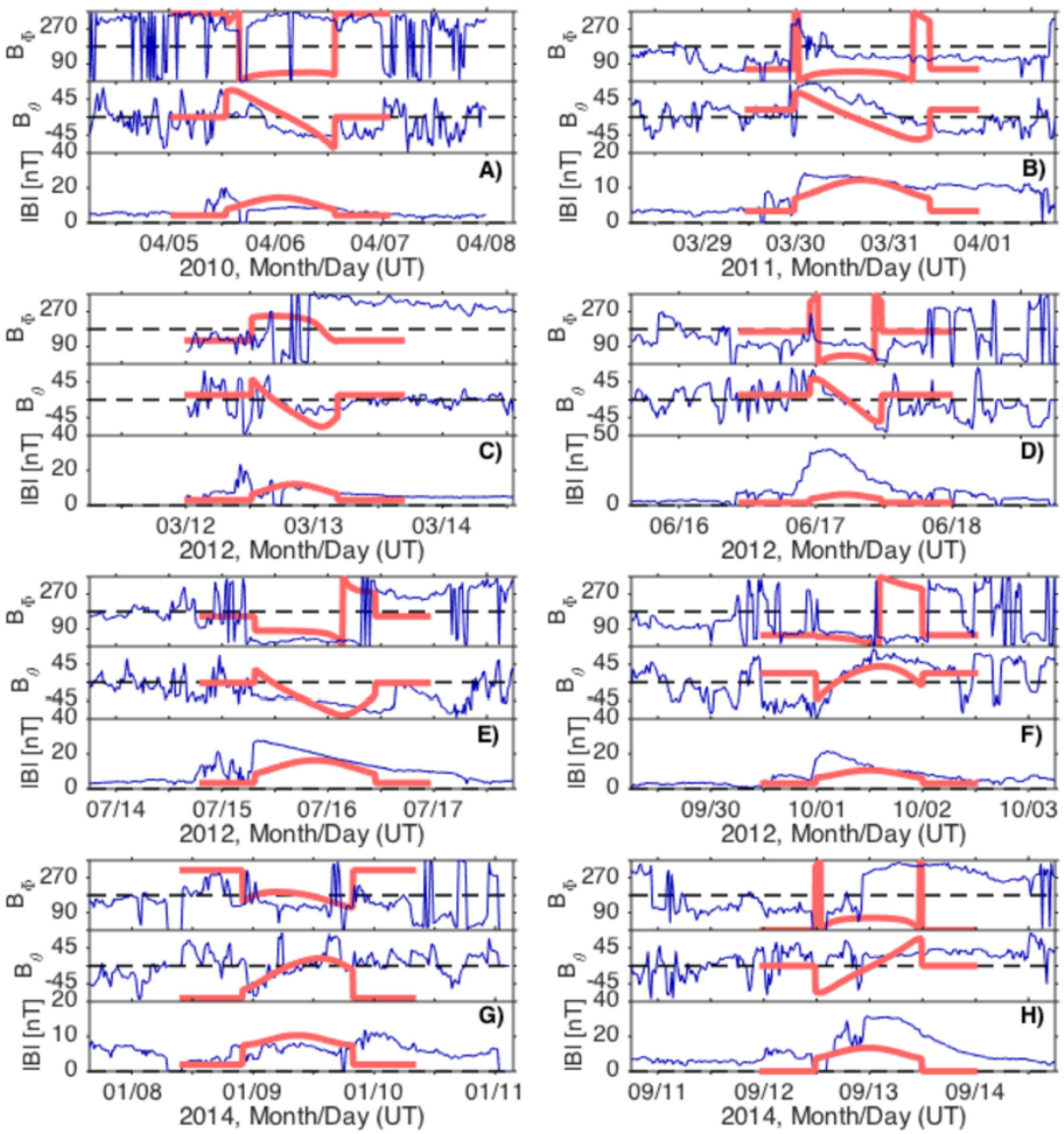

Figure 1. Predicted (red) and observed (L1; blue) magnetic vectors in spherical GSE coordinates system for 8 CME events between 2010 and 2014. $\mathbf{B}_\theta$ is the angular magnetic field direction out of the Sun-Earth plane. Further details of the BSS technique used to predict the magnetic vectors are provided in Paper 1. [From *Savani et al.*, 2015].

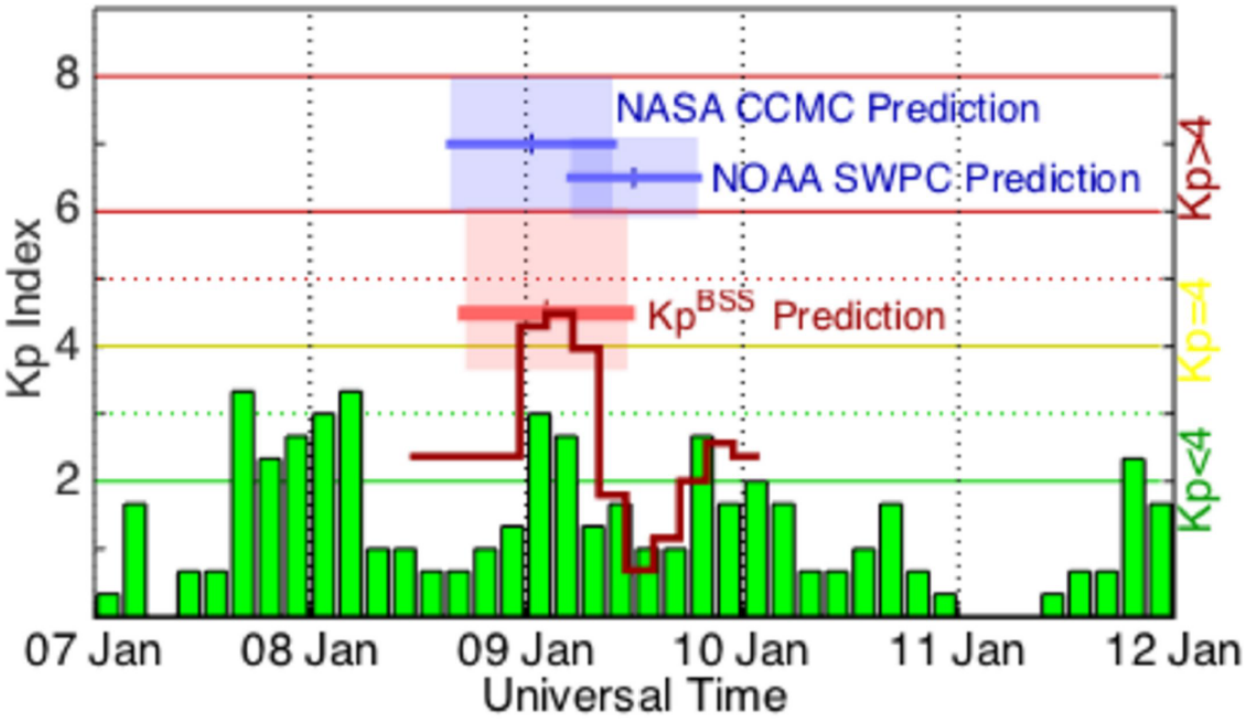

Figure 2. Geomagnetic Kp index for the January 2014 CME event (green). The forecast predictions by NOAA/SWPC and NASA/SWRC leading to a major false alarm forecast are displayed in blue horizontal bars. The red bar represents the maximum value in the $Kp^{(BSS)}$ forecast, which is derived using the empirical relationship of *Newell et al.* [2007] with the predicted magnetic field vectors (Figure 1G) as input. The shaded regions surrounding the forecast bars represent the uncertainty in the maximum Kp value and storm arrival time.

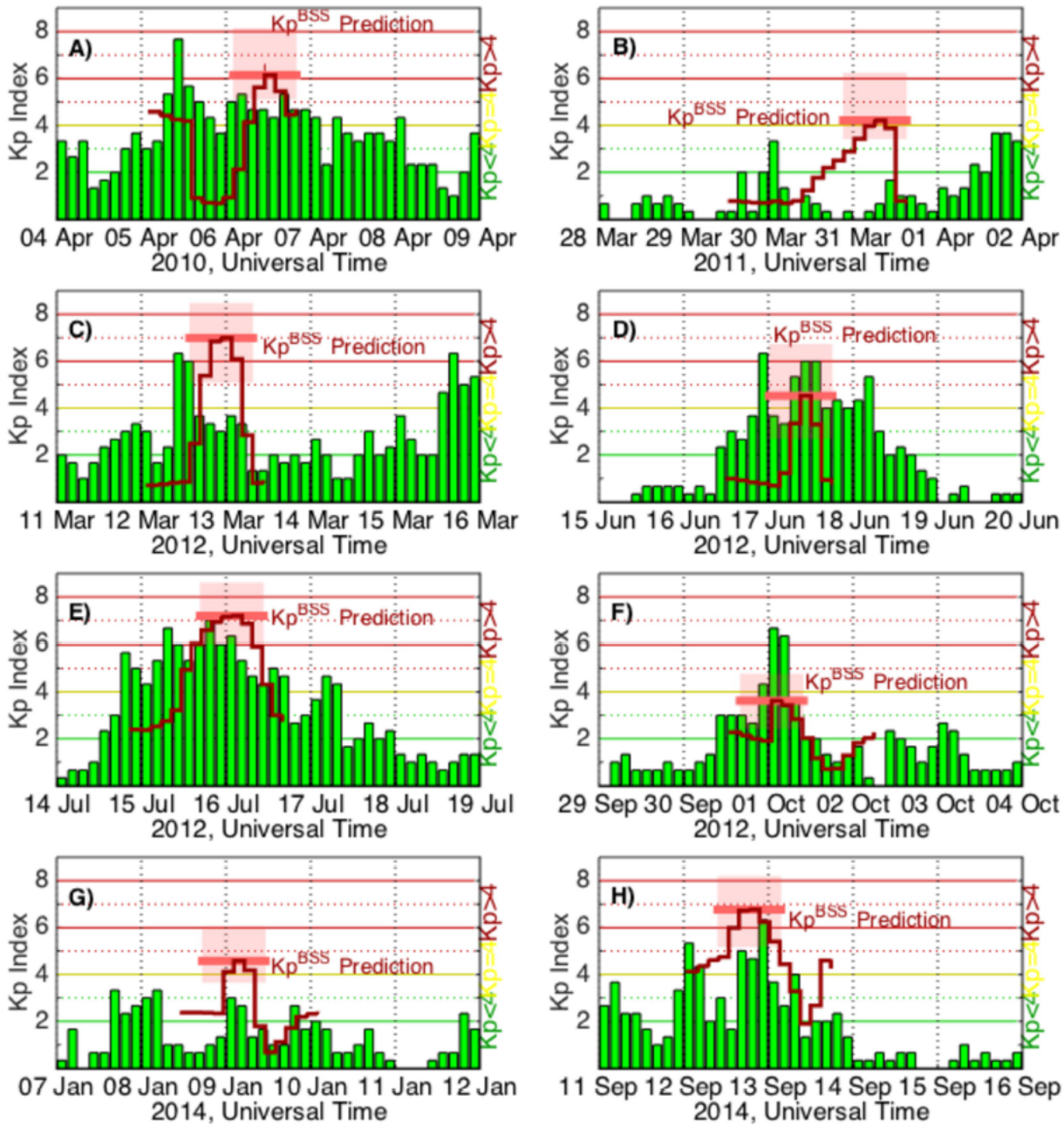

Figure 3. Kp index measuring the Earth's geomagnetic response to the passage of eight CMEs between 2010 and 2014 (green). The red curves display the $Kp^{(BSS)}$ forecast using the estimated magnetic vectors within the CME displayed in Figure 1. The red shaded regions indicate the range in the predicted Kp that results from varying the magnetic field magnitude as discussed in the text.

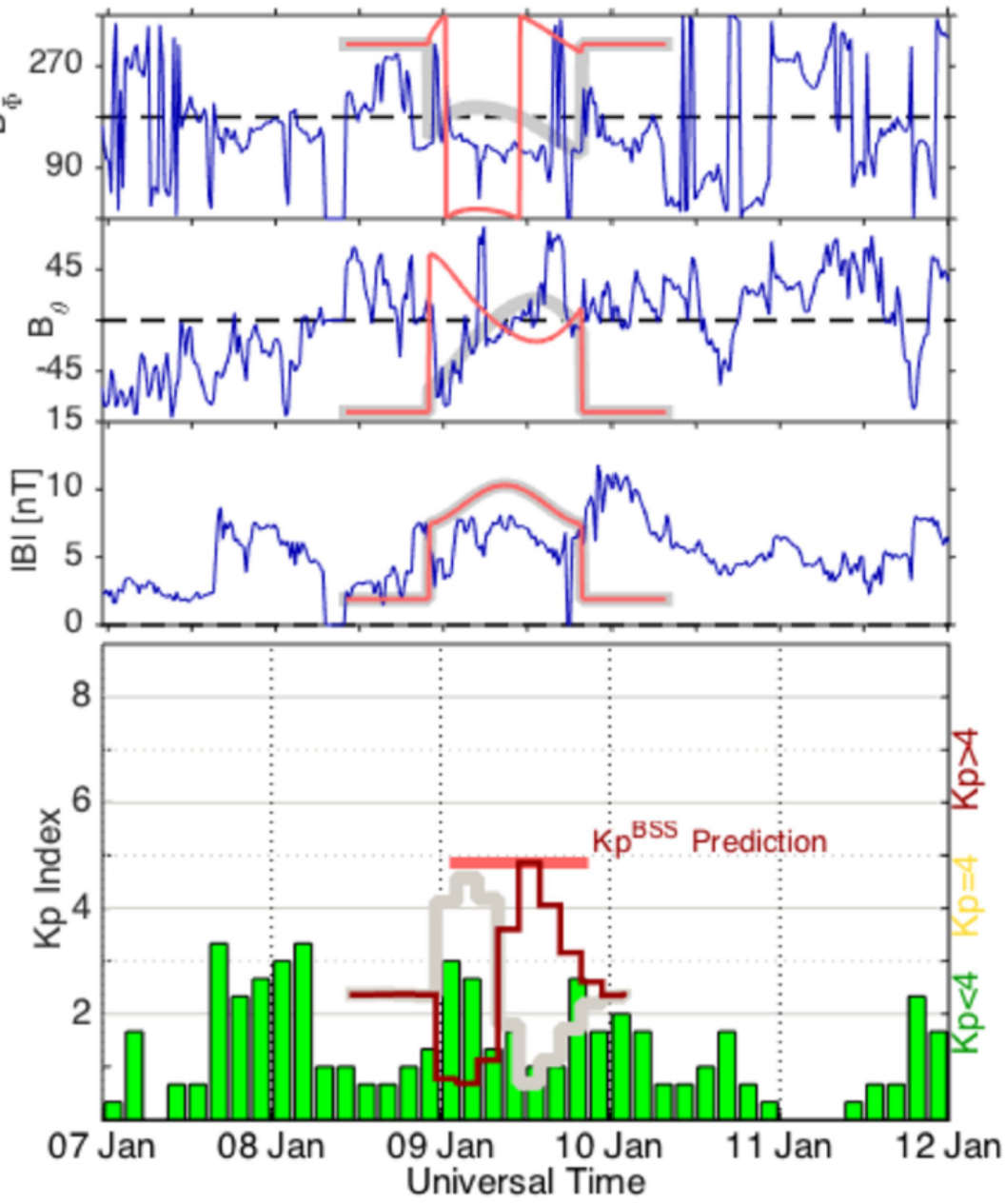

Figure 4. The observed magnetic vectors (blue, in top three panels) and Kp index (green bar chart in bottom panel) for the January 2014 CME event. The original predicted magnetic vectors using the BSS scheme is overlaid in gray (cf. Figure 1G). The chirality of the predicted magnetic vectors (red) is reversed back to the standard Bothmer-Schwenn system. In the bottom panel, the observed Kp index is in green with the original predicted values using the BSS scheme (gray). The $Kp^{(BSS)}$ forecast (red) for these different predicted magnetic vectors is displayed on the observed Kp values. A red horizontal red bar indicates the time and magnitude of the maximum predicted Kp.

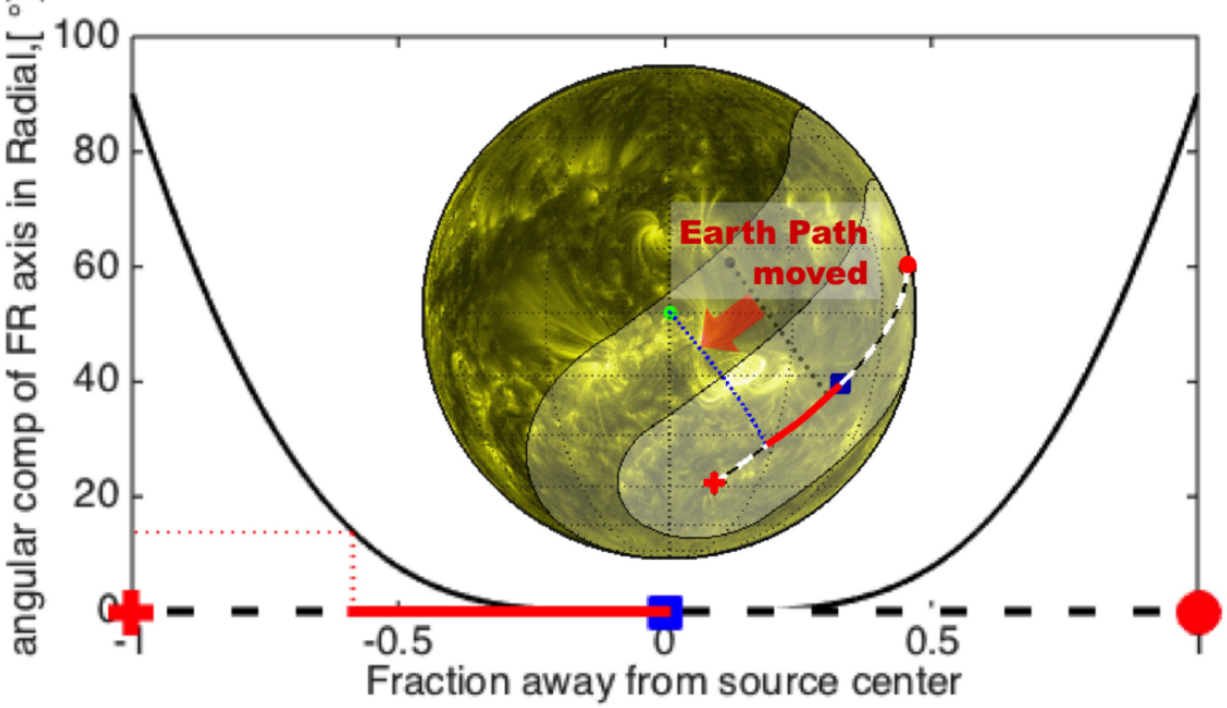

Figure 5. Empirical relationship for estimating the radial component of the CME axis for any given Earth trajectory from movements parallel to the central CME axis. This increase away from the CME nose (blue square), changes the orientation of the modeled CME central axis by increasing the radial component of the axis. Solar inset displays the position of a hypothetical CME moving relative to the Sun and Earth position. The Earth trajectory (green circle) moves towards the eastern leg parallel to the central CME axis (black/white curve). Therefore the original position of the Earth (gray circle) is no longer accurately represent in solar latitude and longitude.

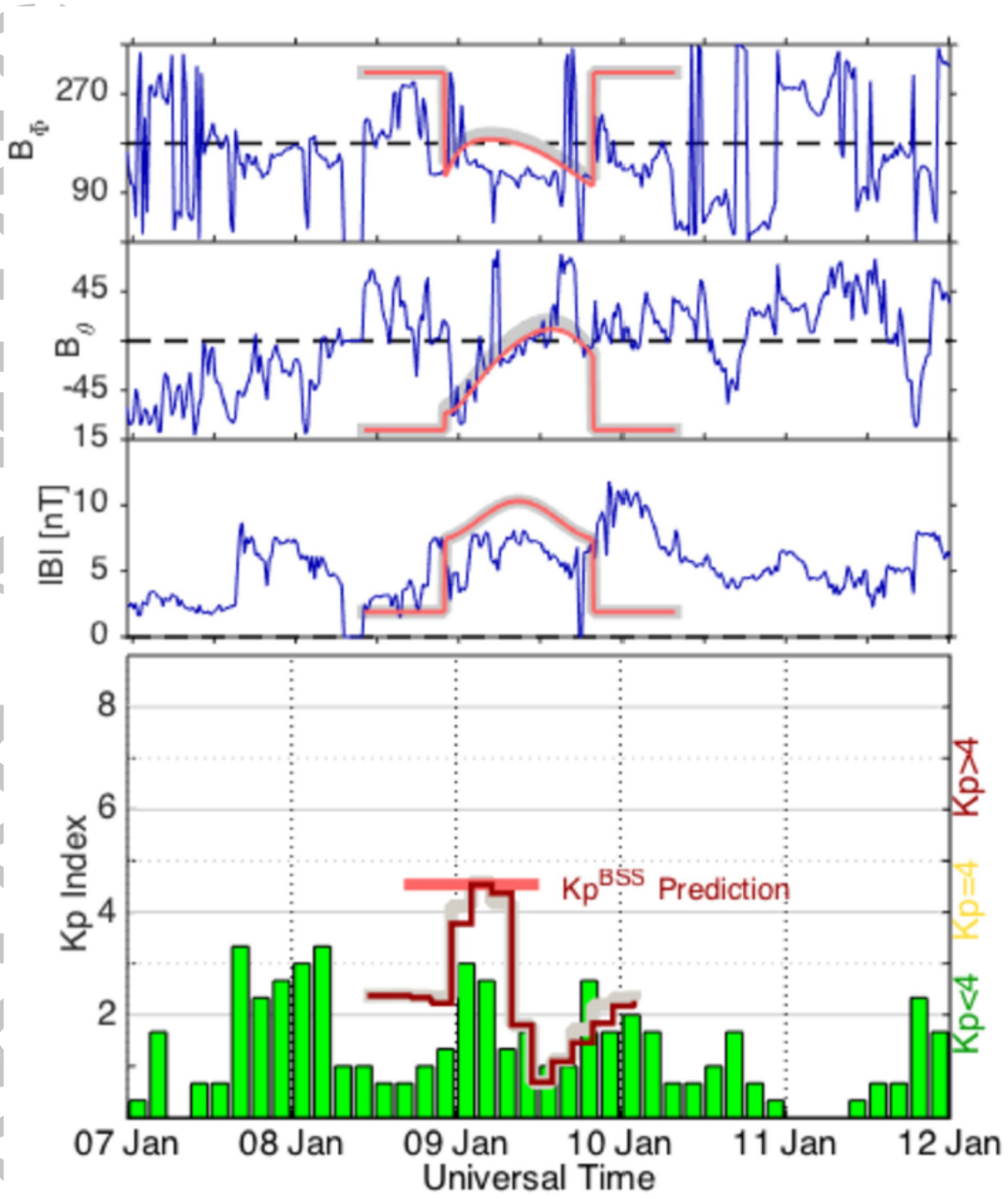

Figure 6. The figure is in the same format as for figure 4, again using the January 2014 CME event. Here we display the predicted magnetic vectors and $Kp^{(BSS)}$ if the trajectory of Earth moved parallel to the central CME axis as shown in Figure 5. The $Kp^{(BSS)}$ forecast (red) is displayed on the observed Kp values.

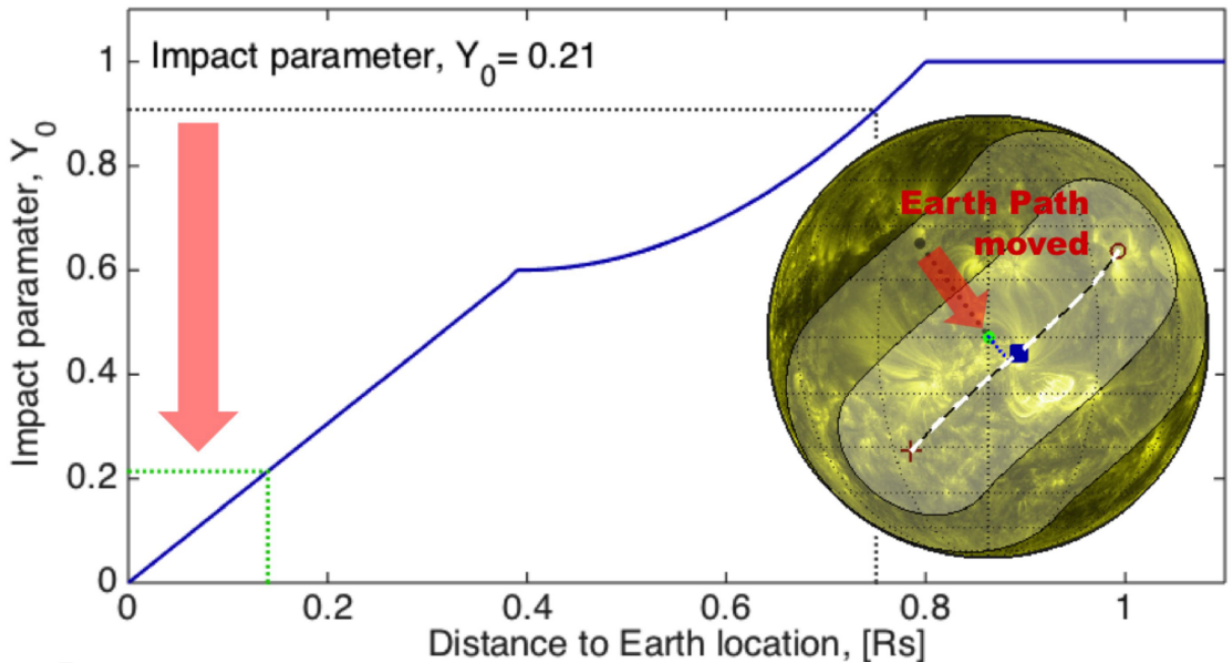

Figure 7. Empirical relationship for estimating the impact parameter to the CME axis for any given Earth trajectory from movements perpendicular to the central CME axis. Solar inset displays the position of a hypothetical CME moving relative to the Sun and Earth position. The Earth trajectory (green circle) moves towards central CME axis (black/white curve). Therefore the original position of the Earth (gray circle) is no longer accurately represent in solar latitude and longitude. The blue square is the location of the CME nose.

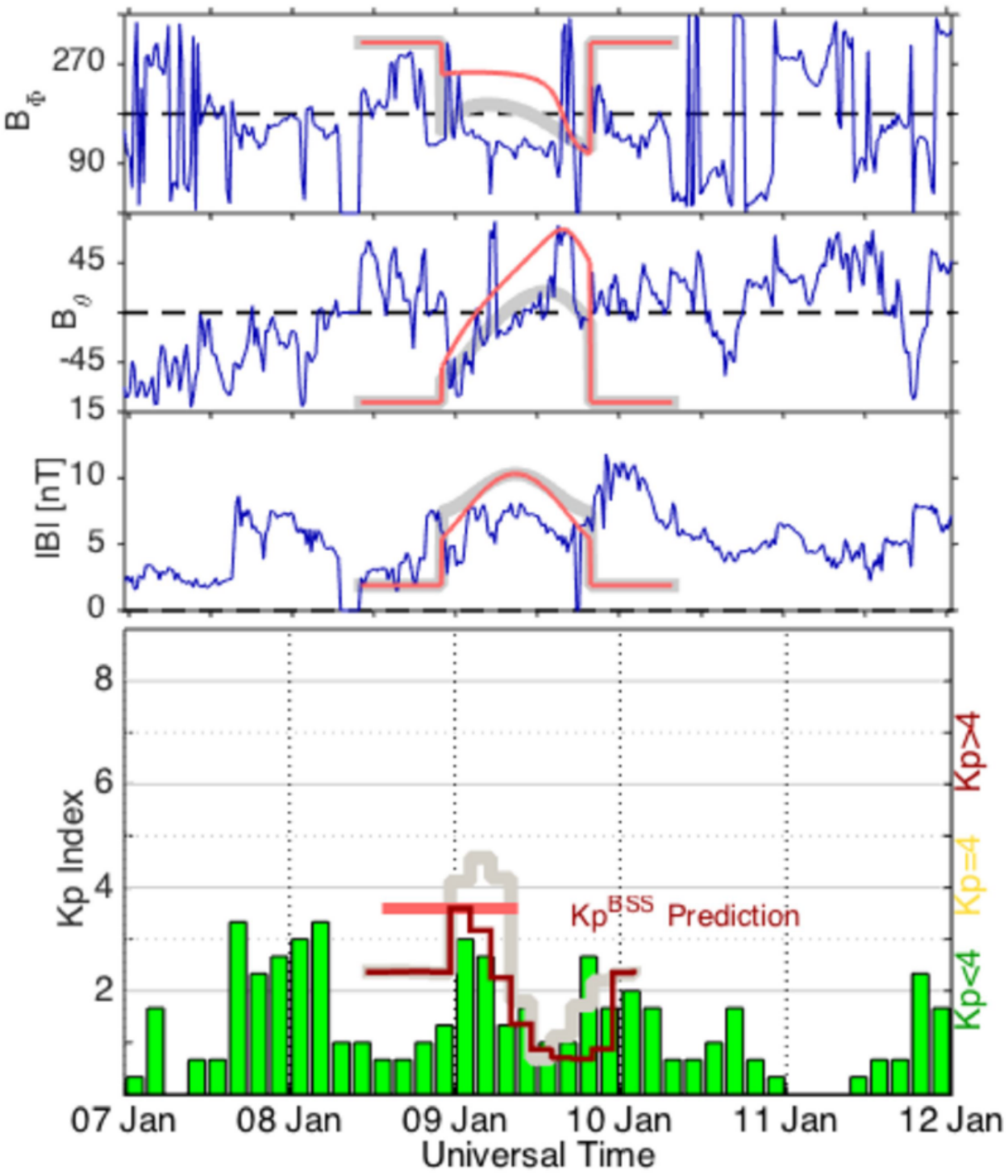

Figure 8. The figure is in the same format as for figure 4, again using the January 2014 CME event. Here we display the predicted magnetic vectors and $Kp^{(BSS)}$ if the trajectory of Earth moved perpendicular to the CME axis as shown in Figure 7.

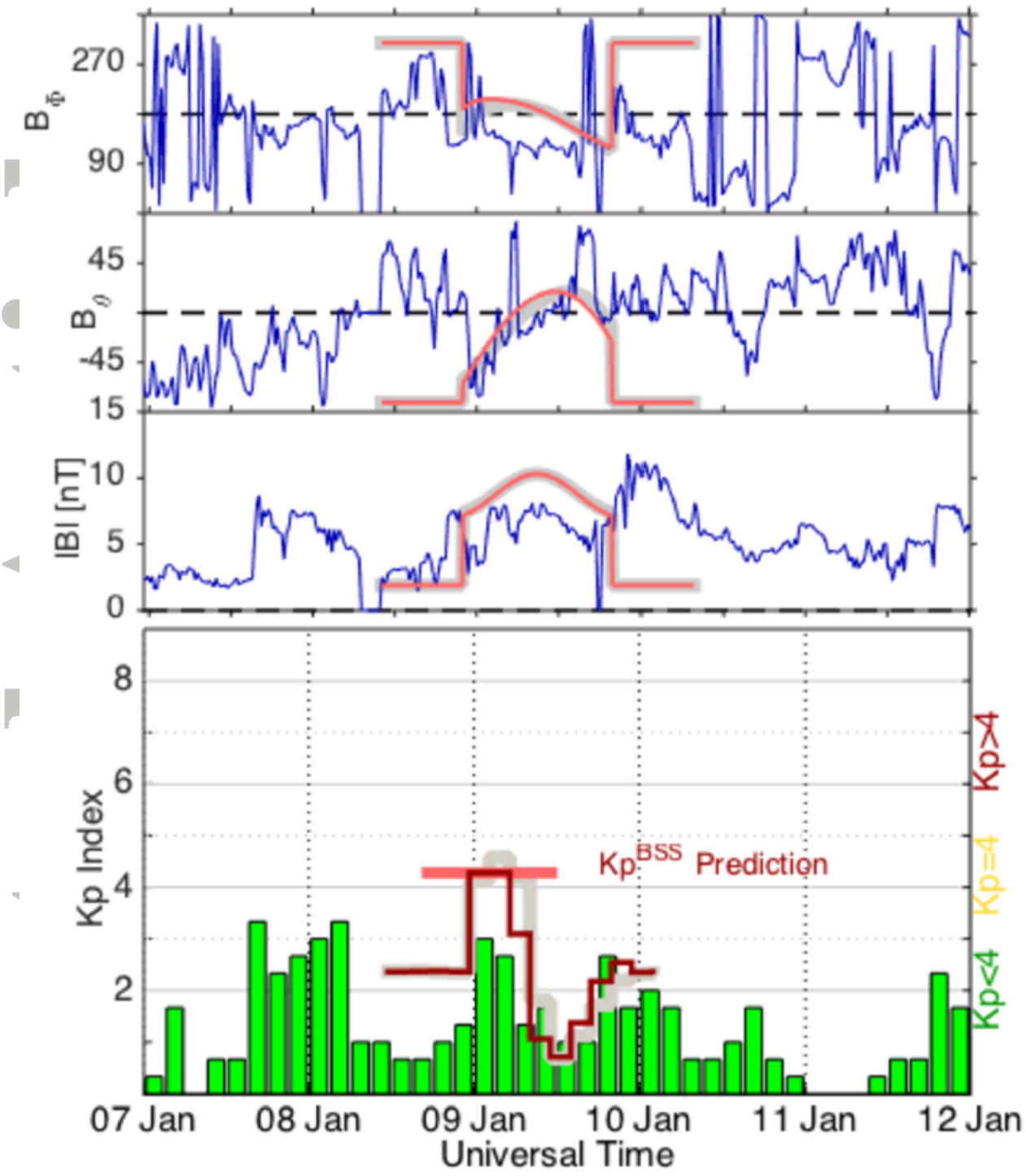

Figure 9. The figure is in the same format as for figure 4, again using the January 2014 CME event. Here we display the predicted magnetic vectors and $Kp^{(BSS)}$ if the theoretical tilt angle of the CME axis is rotated from 40 deg to 60 deg.

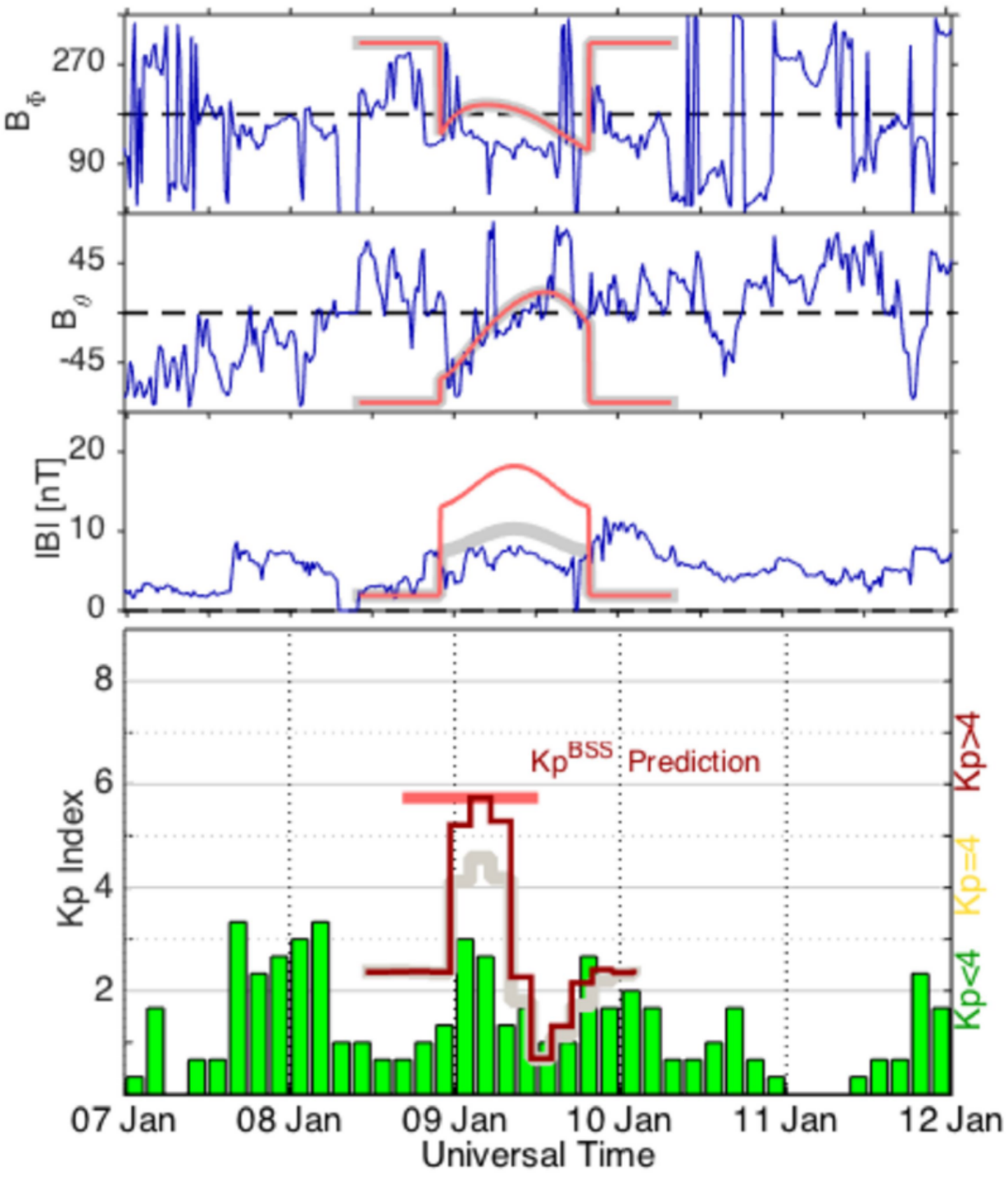

Figure 10. The figure is in the same format as for figure 4, again using the January 2014 CME event. Here we display the predicted magnetic vectors and $Kp^{(BSS)}$ if a 2σ increase in the magnetic field strength had occurred with no change in the field vectors directions.

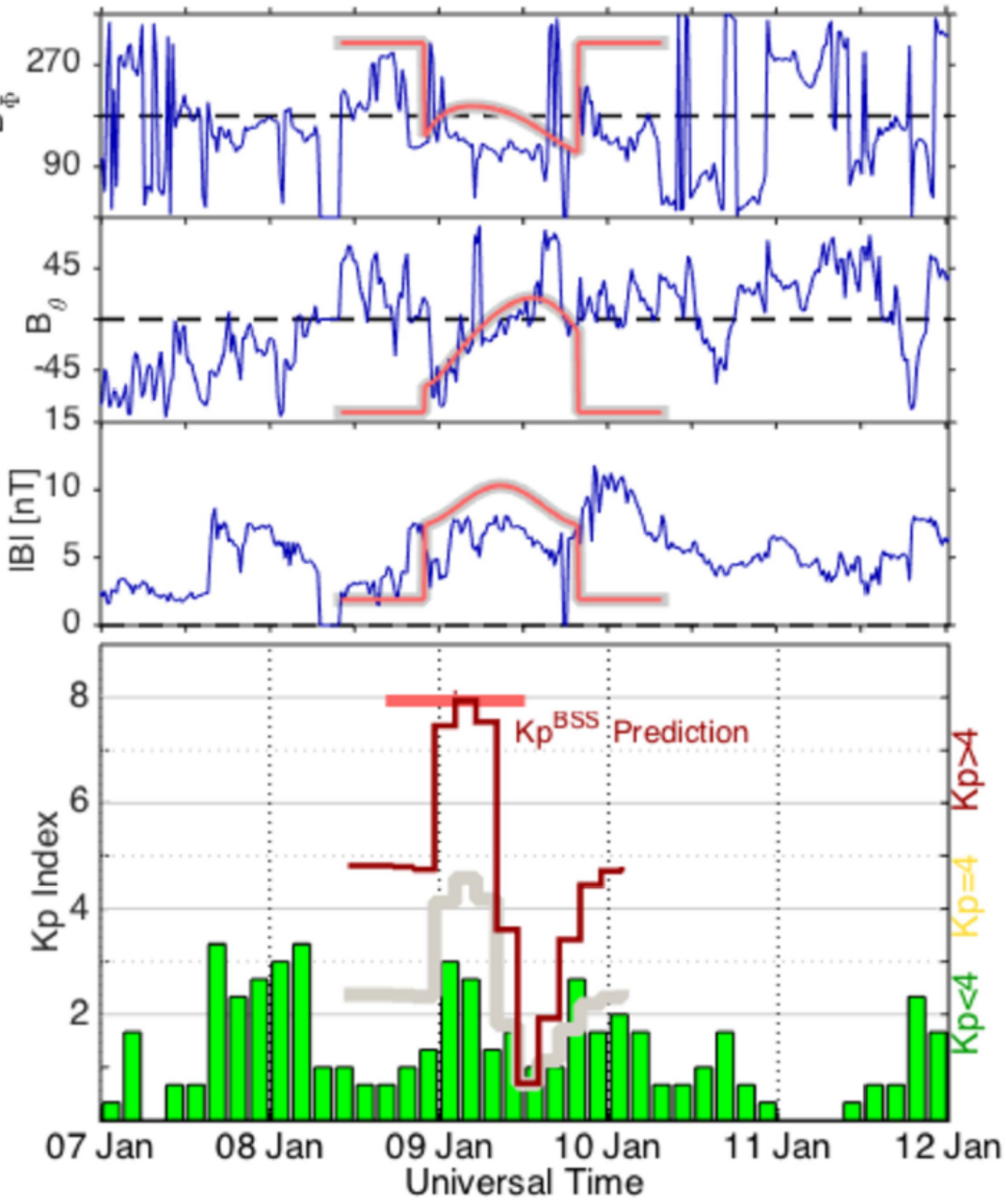

Figure 11. The figure is in the same format as for figure 4, again using the January 2014 CME event. Here we display the predicted magnetic vectors and $Kp^{(BSS)}$ if the maximum CME speed from the WSA-ENLIL+cone model of 840 km/s (assumed by a real-time NASA/SWRC forecast) rather than the observed speed of 370 km/s.

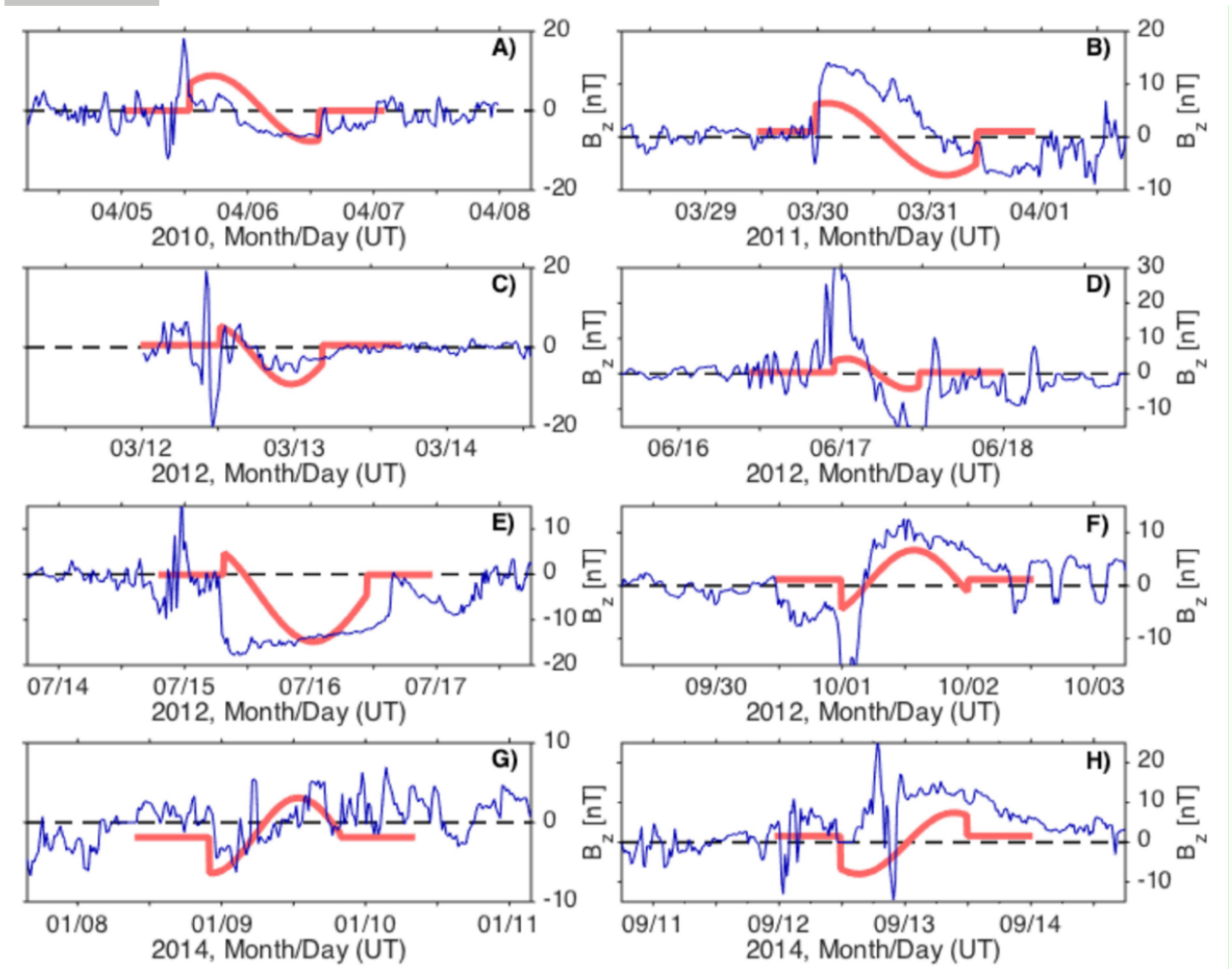

Figure 12. Predicted (red) and observed (L1; blue) magnetic vectors in the cartesian GSE coordinates system for the Bz component. (A) - (H) display the 8 CME events between 2010 and 2014 as shown in figure 1.