AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2017SW001609

Testing the current paradigm for space weather prediction with heliospheric imagers

Luke A. Barnard[1], Curt A. de Koning[2,3], Christopher J. Scott[1], Mathew J. Owens[1], Julia Wilkinson[4], and Jackie A. Davies[5]

[1]Department of Meteorology, University of Reading, Reading, UK, [2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA, [3]Space Weather Prediction Center, NOAA, Boulder, Colorado, USA, [4]Zooniverse, c/o Astrophysics Department, University of Oxford, Oxford, UK, [5]RAL Space, Didcot, UK

Key Points:
- J-map analysis methods may be inappropriate for analyzing CME dynamics with Heliospheric Imagers
- Different CME geometric models are indistinguishable under realistic observational errors
- Heliospheric Imager data may be better exploited with more sophisticated techniques, such as data assimilation

Correspondence to:
L. A. Barnard,
l.a.barnard@reading.ac.uk

Citation:
Barnard, L. A., C. A. de Koning, C. J. Scott, M. J. Owens, J. Wilkinson, and J. A. Davies (2017), Testing the current paradigm for space weather prediction with heliospheric imagers, *Space Weather*, *15*, 782–803, doi:10.1002/2017SW001609.

Received 9 FEB 2017
Accepted 17 APR 2017
Accepted article online 24 APR 2017
Published online 1 JUN 2017

Abstract Predictions of the arrival of four coronal mass ejections (CMEs) in geospace are produced through use of three CME geometric models combined with CME drag modeling, constraining these models with the available Coronagraph and Heliospheric Imager data. The efficacy of these predications is assessed by comparison with the Space Weather Prediction Center (SWPC) numerical MHD forecasts of these same events. It is found that such a prediction technique cannot outperform the standard SWPC forecast at a statistically meaningful level. We test the Harmonic Mean, Self-Similar Expansion, and Ellipse Evolution geometric models, and find that, for these events at least, the differences between the models are smaller than the observational errors. We present a new method of characterizing CME fronts in the Heliospheric Imager field of view, utilizing the analysis of citizen scientists working with the Solar Stormwatch project, and we demonstrate that this provides a more accurate representation of the CME front than is obtained by experts analyzing elongation time maps for the studied events. Comparison of the CME kinematics estimated independently from the STEREO-A and STEREO-B Heliospheric Imager data reveals inconsistencies that cannot be explained within the observational errors and model assumptions. We argue that these observations imply that the assumptions of the CME geometric models are routinely invalidated and question their utility in a space weather forecasting context. These results argue for the continuing development of more advanced techniques to better exploit the Heliospheric Imager observations for space weather forecasting.

Plain Language Summary Predicting the arrival of coronal mass ejections (CMEs) in near-Earth space is a challenging and important problem, as CMEs are the main driver of severe space weather. The Heliospheric Imager (HI) cameras provide observations of the plasma between the Sun and Earth and can be used to track the evolution of CMEs as they flow away from the Sun, toward Earth. Techniques have been developed to use the HI observations to predict the arrival of CMEs in near-Earth space, potentially allowing the HI observations to be used in a space weather forecasting context. We assess how well these methods work for four CMEs observed by the HI cameras. We found that for these four events the techniques that use HI observations do not perform as well as the standard CME forecasting techniques used by the Space Weather Prediction Center. We try to improve the HI-based predictions by using a citizen science approach to develop a better method of tracking the CMEs evolution. The improved CME tracking does not significantly improve the CME predictions and suggests that some of the assumptions in the HI based CME prediction techniques are poor assumptions.

1. Introduction

Coronal mass ejections (CMEs) are large-scale eruptions of plasma and magnetic field from the solar atmosphere into the solar wind. They are a major driver of space weather [see, e.g., *Gosling*, 1993; *Hapgood*, 2011], which can cause geomagnetic storms that adversely affect, for example, communications and radio navigation, the electric power grid, and satellite operations [*Cannon et al.*, 2013]. As humanity's evolving needs continue to be increasingly reliant on technology, it becomes even more important to accurately predict CME impact at Earth in a timely manner. The Space Weather Prediction Center (SWPC) of the National Oceanic and Atmospheric Administration (NOAA) issues CME arrival time forecasts [*Pizzo et al.*, 2011] using numerical models initiated by photospheric and coronagraph data only (i.e., observations out to approximately a tenth of

the Sun-Earth distance). Heliospheric imaging, as provided by, for example, the Solar Mass Ejection Imager [*Eyles et al.*, 2003; *Howard et al.*, 2006] and the Heliospheric Imagers (HIs) aboard the two STEREO spacecraft [*Eyles et al.*, 2008; *Howard et al.*, 2008], provides a wide-angle view of the outer corona and inner heliosphere. The HI instruments in particular currently offer the only routine direct observations of solar wind plasma between the solar corona and Earth orbit. Thus, HI observations potentially allow improvement of the current forecast schemes by allowing assimilation of new information.

As with coronagraph images, heliospheric images result from Thomson scattering of photospheric white light by free electrons, with the strongest scattering occurring on the Thomson Plateau [*Howard and Tappin*, 2009; *Howard and Deforest*, 2012]. Within the Heliospheric Imager field of view (FOV), this extended surface can no longer be approximated as a spherical shell, which complicates the interpretation of HI data. However, in an effort to contribute to the space weather enterprise, there is continuing interest in techniques that can quickly and easily retrieve gross CME characteristics from heliospheric images, particularly the CME speed and direction of propagation.

Since the launch of the two STEREO spacecraft, STEREO-A (STA) and STEREO-B (STB) [*Kaiser et al.*, 2008], various techniques based on simple CME geometries have been used in an effort to interpret CME tracks in the HI FOV in order to improve space weather forecasts. These techniques include point P (PP), fixed phi (FP), harmonic mean (HM), self-similar expansion (SSE), and Ellipse Evolution (ElCon) [*Howard et al.*, 2006; *Kahler and Webb*, 2007; *Sheeley et al.*, 1999; *Rouillard et al.*, 2008; *Lugaz et al.*, 2009; *Davies et al.*, 2012; *Möstl and Davies*, 2013; *Rollett et al.*, 2016]. Each makes different assumptions about the geometry of a CME, but all are effectively methods to estimate the radial distance of the CME apex from the observed elongation of the CME in the HI FOV. The FP model assumes that the feature being tracked is a point source, with no cross-sectional extent. The HM assumes a CME with circular cross section that expands with one point tied to Sun center. These two geometries were united by the introduction of the SSE model in which the half width of the CME was introduced as an additional free parameter. Setting this parameter to zero replicates the FP geometry while setting it to 90° replicates the HM geometry; intermediate values approximate the CME cross section as a self-similar expanding circle of constant half width (as discussed by *Davies et al.* [2012]). As a result, some studies have set the half width to a value that represents a compromise between the two extremes; *Möstl et al.* [2014] used 45°, *Rollett et al.* [2016] used 35°, while *Barnard et al.* [2015a] and *Mishra et al.* [2014] used 30°. Such techniques can be applied to CMEs observed from a single spacecraft or to simultaneous observations from two spacecraft with distinct viewpoints [*Lugaz*, 2010; *Liu et al.*, 2010, 2013]. Most recently, the SSE model has been extended further to consider the CME front as elliptical rather than circular, introducing the ellipse aspect ratio as an additional free parameter, resulting in the ElCon model [*Rollett et al.*, 2016].

In practice, all these techniques, to date, have relied on estimating the propagation of a CME based on elongation time (ϵ-t) profiles extracted, generally manually from "J-maps." A J-map (elongation time map) is a means of visualizing the time evolution of a solar wind structure along a given position angle. J-maps are created by extracting a lateral profile, usually centered on the ecliptic plane, from a sequence of heliospheric images. A temporal sequence of these lateral cuts are then displayed vertically to create a J-map [*Davies et al.*, 2009]. *Sheeley et al.* [1999] were the first to recognize that a compact solar wind transient propagating radially away from the Sun out to large elongations would have a nonlinear ϵ-t profile in such a J-map, even for transients propagating in a constant radial direction at constant speed. Therefore, in principle, this characteristic shape of the ϵ-t profile can be used to obtain the speed and direction of propagation of a solar wind transient, even when observed from only a single viewpoint. Unfortunately, when viewed from a single vantage point, one can only obtain the speed and propagation direction of a solar wind transient from analysis of its ϵ-t profile by assuming that the speed and direction remain constant throughout the observed propagation; an assumption that is not generally valid for a CME [*Gopalswamy et al.*, 2001; *Vršnak*, 2001]. Furthermore, although the process of constructing a J-map seems straightforward, the width of the cut in position angle is typically never 1 pixel wide; rather, a range of position angles is chosen. The average of this range is used in the generation of a J-map, resulting in fine structure that becomes progressively less pronounced as elongation increases.

A further refinement of such techniques is to prescribe the direction of propagation. In doing so, the geometric models can be used to instead return the radial distance of a CME from the Sun as a time series, the differential of which provides information on the CME speed profile. The merits of such a refinement were first investigated by *Harrison et al.* [2012] for a complex series of Earth-directed CMEs. Considering the FP and HM geometric models, they noted that by assuming a range of different propagation angles, a series of

differing CME distance-time profiles was produced, many members of which included unphysical accelerations at large distances. They interpreted such features as resulting from the use of unrepresentative propagation directions. *Wood et al.* [2010] also used geometric models to investigate the kinematics of a CME observed on 1 June 2008 by STEREO-A and demonstrated that, for this one CME, the PP and FP geometric models returned physically implausible kinematics estimates, while the HM geometry was physically reasonable; this highlights the importance of ensuring that a geometric model is representative of a studied event.

Subsequent attempts have been made to constrain the direction of CME propagation using additional information to HI data, in order to better estimate a CME's radial speed. These studies have had varied success, with a common feature being noisy speed profiles and unphysical accelerations at large elongations. *Mishra et al.* [2014] computed distance and speed profiles for PP, FP, HM, and SSE geometries from ϵ-t profiles manually extracted from J-maps derived from STEREO-HI data. By constraining the CME propagation direction using stereoscopic methods with STA and STB COR2 coronagraph data, they were able to generate distance-time and speed-time profiles for each technique. From their study of 10 different variants of these techniques, applied to three CMEs, they concluded that each CME was best modeled by a different technique.

Because of these difficulties, additional refinements have been introduced including approaches that combine standard analysis of ϵ-t profiles from J-map with other models, such as a drag force model [*Tucker-Hood et al.*, 2015; *Mishra et al.*, 2014, 2015; *Rollett et al.*, 2016]. Others have attempted to constrain the CME direction such that the distance-time and velocity-time profiles were most consistent with in situ measurements of the arrival time and velocity [*Rollett et al.*, 2012]. However, all such refinements are ill suited to space weather forecasting because they introduce additional parameters that cannot be measured directly during the early life of a CME and hence must be guessed at.

In this paper, we will first use the standard J-map analysis combined with CME geometric modeling to estimate the kinematics of a CME and will use this example to argue that there are inconsistencies with the results that imply either a breakdown in the applicability of the geometric models or a limited ability to employ such models due to errors in the CME tracking (or both). We will then introduce a citizen science-based approach to tracking CMEs in HI that provides a more comprehensive picture of the CME front and tracks the CME more stably than in a J-map analysis. We note that the tracking technique could also be employed by an expert observer or team. We then use the more detailed tracking of the CME to revisit the CME kinematics estimates, revealing that many of the same issues persist. Following this, we extrapolate the CME kinematics estimates farther into the heliosphere by combining them with the drag-based model (DBM) of CME propagation [*Vršnak et al.*, 2013], which allows comparison with both the observed CME arrival time at L1 and the SWPC forecast arrival. Section 2 introduces the data sources used throughout this study. Section 3 details the standard expert tracking of CMEs in J-maps and the computation of CME kinematics from the J-map ϵ-t profiles. The citizen science CME tracking and revised estimation of the CME kinematics are presented in sections 4–4.3. We conclude the article with a discussion of these results in section 5.

2. Data

Here we briefly review the data sources used throughout this study, including the SWPC CME forecast list and data generated by the STEREO-HI instruments. Later, we also introduce the three sources of ϵ-t profiles used to characterize the location of CME fronts in the HI data: the Reading expert identifications, the HELCATS catalog, and the Solar Stormwatch identifications.

2.1. SWPC Forecasts of CME Arrival in Geospace

Space weather forecasts for operational use are produced by SWPC, including forecasts of CME arrival in geospace. The first step in the forecasting process is to estimate initial CME parameters (i.e., propagation direction, CME speed, and width) from available coronagraph data using the CME Analysis Tool (CAT) [*Millward et al.*, 2013]. The near-Sun CME parameters are then used to initialize a WSA-ENLIL model [*Pizzo et al.*, 2011] run of this CME (or sequence of CMEs). A similar forecasting process is also implemented by the UK Met Office Space Weather Operations Centre (MOSWOC). As these forecasts are produced by world-leading space weather operations centers, we consider it fair to regard this process as the "industry standard" by which other CME forecasting techniques should be benchmarked. In this work, we use a record of the SWPC Earth-directed CME forecasts, which has since been validated by SWPC through comparison with the observed evolution of the event with spacecraft monitors at L1 and the subsequent effects on geospace. Specifically, we analyze four of the forecasts provided by SWPC, the details of which are presented in Table 1. These events

Table 1. SWPC Forecast CME Properties

Appearance Time[a] (UT)	HEE Lat. (deg)	HEE Lon. (deg)	Half-Width (deg)	Speed (km s^{-1})	SWPC Forecast Arrival (UT)	Observed Arrival (UT)
2012-08-31T20:36	0	−30	33	1010	2012-09-03T17:00	2012-09-03T11:23
2012-09-28T00:48	4	20	55	872	2012-09-30T15:00	2012-09-30T22:13
2012-10-05T07:42	−24	9	42	698	2012-10-08T15:00	2012-10-08T04:31
2012-11-20T12:36	20	22	47	664	2012-11-23T17:00	2012-11-23T21:12

[a]Dates are formatted as yyyy-mm-ddTHH:MM.

were selected as they spanned a range of initial CME speeds, are well defined in both the STEREO-A HI1 and STEREO-B HI1 FOV, and were identified in the L1 in situ observations. However, we note that they are closely clustered in time and come from a limited period of the STEREO mission in 2012, when the relative positioning of STEREO-A, STEREO-B, and Earth is less than optimal for viewing Earth-directed CMEs. However, the WSA-ENLIL CAT model only became operational at SWPC in October 2011, and so before this period we cannot compare against genuine SWPC operational forecasts, when the viewing geometry of geoeffective events may have been more favorable. Clearly, it would also be beneficial to study events from this earlier epoch of the STEREO mission, and we have started to do so. However, this will necessarily involve comparing against WSA-ENLIL CAT model hindcasts of these events, whereas here we choose to focus on comparing against the genuine SWPC forecasts. Figure 1 shows a schematic layout of the inner heliosphere for these four events. Each panel displays a view of the ecliptic plane from above the ecliptic north pole. The green circle and purple ellipse show the CMEs represented by the SSE and ElCon geometric models when the CME apex is at a distance of 30 R_{Sun}, with the required estimates of the propagation direction, and CME half width, close to, albeit not exactly in, the ecliptic, derived from the SWPC coronagraph analysis. The ellipse aspect ratio was set to 1.4, which *Möstl et al.* [2015] demonstrated was a reasonable value to assume in the absence of any other information to constrain this parameter.

2.2. STEREO-HI

The STEREO spacecraft, launched in late 2006, occupy near 1 AU ecliptic heliocentric orbits, one ahead (STEREO-A: STA) and the other behind (STEREO-B: STB) the Earth. Both spacecraft are equipped with the Sun Earth Connection Coronal Heliospheric Investigation (SECCHI) instrument package [*Howard et al.*, 2008]. The SECCHI package on each of STA and STB includes the HI instrument [*Eyles et al.*, 2008], which consists of two wide-field white-light cameras (HI1 and HI2) that can image solar wind structures such as CMEs propagating over a total elongation angle range from near 4° to around 90° from the Sun. In nominal science operations, the 20° FOV of HI1 is centered at 14° in the ecliptic plane and the 70° FOV of HI2 is centered at 53.8°, also in the ecliptic plane. Note that the presence of a trapezoidal occulter, which was intended to block the intense light from Earth at the start of the mission [*Eyles et al.*, 2008], limits the outer edge of the central portion of the HI2 FOV to an elongation of around 74°. The nominal cadence of HI1 and HI2 science images is 40 and 120 min, respectively, while their binned pixel size is 70 arc sec and 4 arc min, respectively. The CMEs analyzed in this work span the period from August until November 2012. Over this period, the separation between STA and STB was approximately 115° of longitude (in Heliocentric Earth Ecliptic (HEE) coordinates), while the separation of both STA and STB from Earth was approximately 120° of longitude. Throughout this work we will discuss the location of features in the HI FOV in terms of Helioprojective-Radial-Coordinates: position angle (PA), the anticlockwise angle from solar north, and elongation (ϵ), the angular distance from Sun center.

3. CME Kinematics Derived From J-Maps

In this section we will examine the kinematics of a CME using the standard techniques of identifying and extracting the CME ϵ-t profile from J-maps derived from HI observations and then applying several CME geometric models. Before we discuss the CME kinematics we will first introduce the ϵ-t profile data sets used, as well as the CME geometric models and relevant aspects of the data processing.

3.1. University of Reading Identifications

The authors manually tracked the fronts of the four selected CME in J-maps formed from both HI1 and HI2 differenced image data, in a 5° PA band, both along the PA corresponding to the ecliptic plane, and also along the same PA as tracked by the HELCATS project (see section 3.2 below). Each CME was tracked 6 times, and a

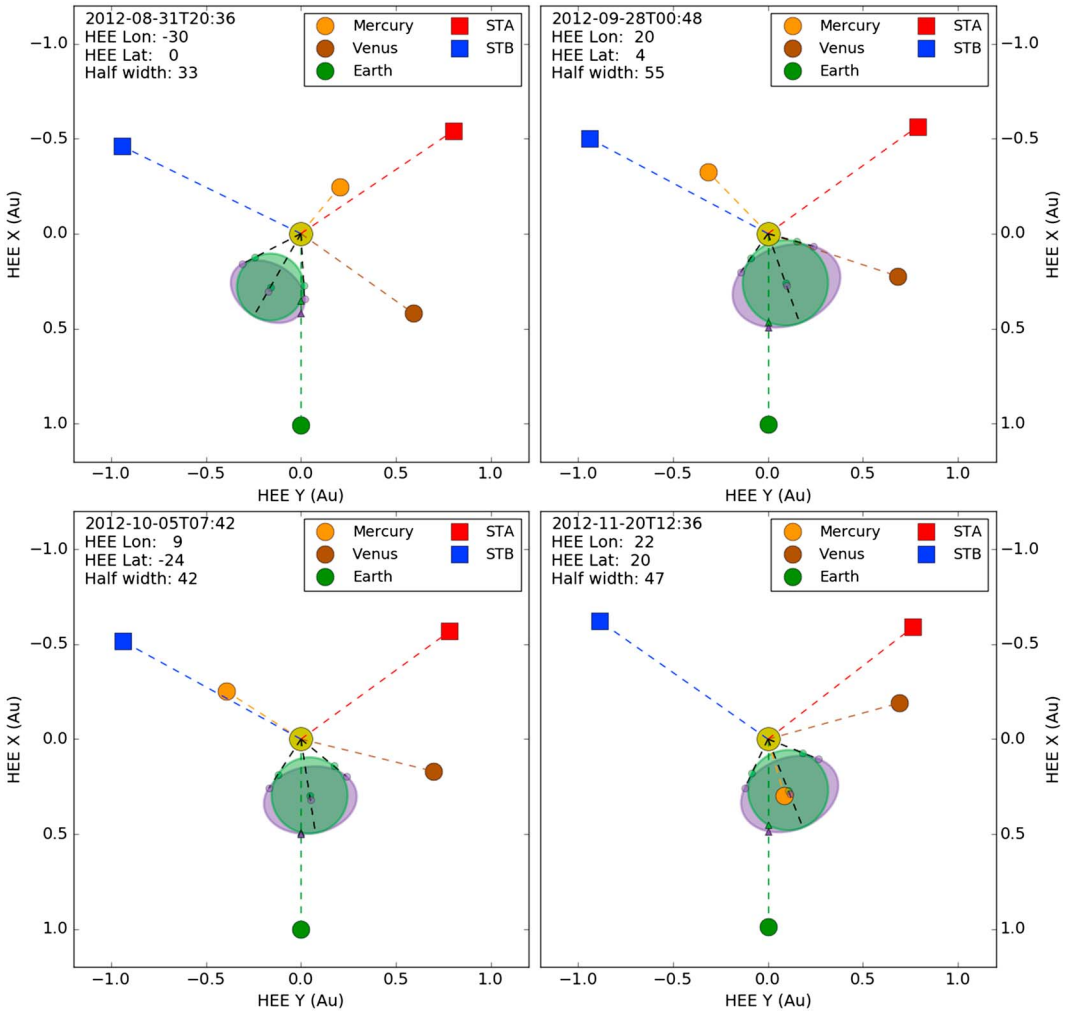

Figure 1. Schematic layout of the inner heliosphere for the four events under investigation, looking down upon the ecliptic plane (the *x*-*y* plane of the HEE coordinate system) from the ecliptic north pole. The large green circle and purple ellipse show the CMEs represented by the SSE and ElCon geometric models, respectively, with the CME half width derived from the SWPC coronagraph analysis and the ellipse aspect ratio set to 1.4.

consensus ϵ-*t* profile was calculated by averaging together the six profiles. This was calculated using a sliding window mean of the full set of time and elongation coordinates, using a window width of 120 min, centered on the central time of each of the HI1 frames. Cubic-spline interpolation was used to estimate the ϵ-*t* coordinates at frame times for which there were no track coordinates in the sliding window. We refer to the authors' profiles with the acronym RDG.

3.2. HELCATS

The HELCATS project (https://www.helcats-fp7.eu/) provides the ϵ-*t* profiles of CMEs identified in HI observations and manually extracted from J-maps formed from HI1 and HI2 differenced image data [*EU HELCATS et al.*, 2015]. HELCATS uses a bank of J-maps corresponding to distinct PA bands to track CMEs, with the central PA of each PA band being separated by 5°. The CMEs were tracked along the PA band that the expert deemed most appropriate, being as close to the CME apex as practicable, therefore being within ±5° of the CME apex PA (as estimated by the expert observer). Each event was tracked 6 times by the HELCATS experts. From these six profiles we calculate a consensus ϵ-*t* profile for each CME, by averaging together the multiple profiles using the same algorithm as described in section 3.1. We refer to the HELCATS profiles with the acronym HCT.

3.3. Single-Spacecraft CME Geometric Models

A range of CME geometric models were introduced in section 1, which can be used to estimate the radial distance of the CME apex from the elongation in the HI FOV. In this study we focus on applying the geometric

models to HI observations from a single observatory, rather than stereoscopic methods. With the continuing lack of STB observations, and the potential for HI-like observations from only a single future operational space weather monitor, it seems reasonable to assume we will most likely only have HI or HI-like observations from only one perspective in the future. Here we focus on the HM, SSE, and ElCon models. *Mishra et al.* [2014] presented evidence that the HM and SSE models outperformed the FP model for estimating CME kinematics, and we include the ElCon model as it is the latest published development in this area [*Möstl et al.*, 2015; *Rollett et al.*, 2016; *Amerstorfer et al.*, 2016]. We implement the SSE model as given by *Davies et al.* [2012], as well as the off-apex corrections to the CME front location given by *Möstl and Davies* [2013]. To compute the HM model, we use our implementation of SSE model but fix the CME half width to be 90°. The ElCon model is implemented as detailed by *Rollett et al.* [2016]. For the ElCon model we also fix the ellipse aspect ratio at 1.4, based on the results presented in *Möstl et al.* [2015]

In this work, we constrain the CME direction and half width using estimates provided by the SWPC analysis of available coronagraph data performed using the CME Analysis Tool [*Millward et al.*, 2013]. These are the same estimates of CME direction and half width that were used as input to the SWPC WSA-ENLIL forecast of these events. With estimates of these parameters we may then apply the geometric models to the ϵ-t profiles to compute estimates of the CME apex distance and speed. We emphasize that, in principle, these geometric models can recover the CME apex distance from the measured ϵ-t profile for any viewing direction. In other words, although our four events occur during a time period that is suboptimal for viewing Earth-directed CMEs with the STEREO-HIs, nevertheless, this should, in theory, not be an impediment to the geometric models, since the models directly relate the CME's lateral expansion to its radial expansion.

As pointed out by *Möstl et al.* [2015] and *Rollett et al.* [2016], we note that for each of these geometric models, it is possible, and perhaps potentially more realistic, to use time-dependent estimates of the CME direction, half width, and, for ElCon, also the ellipse aspect ratio. Other research has also concluded that including time dependence could be important [*Lugaz et al.*, 2010; *Conlon et al.*, 2014]. However, without observations or a physical model to realistically constrain the time dependencies of these parameters, we choose to leave them fixed, as is most often done when applying these techniques to HI observations from a single perspective [*Mishra et al.*, 2014; *Möstl et al.*, 2015; *Rollett et al.*, 2016]. In fact, without external estimates or constraints on the basic CME parameters, it is not possible to infer realistic CME kinematics based on these geometric models with HI-derived ϵ-t profiles from a single spacecraft; each of the geometric models requires that at least one free parameter be constrained before the CME kinematics profile can be estimated from a ϵ-t profile.

3.4. Error Calculations

In using the different geometric models to estimate the CME kinematics, it is important to account for the different sources of error in the calculations. For the RDG and HCT consensus ϵ-t profiles, error estimates are computed from the spread of the six individual ϵ-t profiles of each event. The gridding of the data in time and elongation in the generation of the J-maps from which they are extracted means that these profiles have errors in each coordinate. The long cadence of the HI images means that there is also uncertainty due to aliasing introduced by the CME propagation during the image integration. We do not consider this error in this analysis.

We also make use of the CME propagation directions and half widths derived from the SWPC coronagraph analysis. Error estimates are not provided for these quantities and so we assume a $\pm 5°$ error in each quantity throughout this work. We note that these are relatively small errors, and they could plausibly be larger. *Pizzo et al.* [2015] used an error of $\pm 10°$ in each quantity, which they considered to be representative of the errors, based on the experience of expert analysis of coronagraph observations. Therefore, the error estimates presented here are likely a "best-case scenario."

The different formulae used to calculate the CME geometric models contain up to 13 terms, and it would be cumbersome to calculate a full error equation for each model, accounting for the errors in each term. Alternatively, we estimate the errors by computing the geometric model calculations with every combination of the upper and lower error bounds of each observed or derived quantity in the equations. The final error estimate is given by the maximum and minimum limits of this set of error calculations.

3.5. Kinematics of 28 September 2012 CME

The SWPC forecast registered a CME that first appeared in coronagraph images at 2012-09-28T00:48 UT. The SWPC analysis of available real-time coronagraph data yielded estimates of the coronal propagation direction

as radially away from 4°N, 20°W in HEE coordinates, as well as a CME speed and half width of 872 km s^{-1} and 55°, respectively.

Using the commonplace J-map analysis, estimates of the kinematics of this CME were calculated for each of the SSE, HM, and ElCon geometric models, based on the RDG consensus ϵ-*t* profile of the CME in the ecliptic plane. The time profile of the CME apex distance (*R*) was computed for each model, using the SWPC estimate of the CME half width and propagation direction to constrain the free parameters of each model. For the ElCon model a CME aspect ratio of 1.4 was also assumed. Profiles of CME speed (*V*) were calculated by the pointwise numerical differentiation of the *R*-*t* profile. As stated, we track the CME along the ecliptic, but the SWPC propagation direction estimate was directed 4° above the ecliptic; we assume that this small latitudinal offset of the CME apex from the ecliptic has a negligible effect on this analysis—it is within the assumed error of the coronal propagation direction.

The results are presented in Figure 2; the left and right columns show the results for STA and STB, respectively. Figure 2 (row a) shows the ϵ-*t* profiles for the event. Figure 2 (row b) shows the *R*-*t* profile for the SSE (green circles), ElCon (purple squares), and HM (blue triangles). The colored stars at $\approx$70 h show the estimated heliocentric distance of the CME apex for each geometric model, based on the SWPC determination of the CME arrival time at L1 from in situ measurements, and correcting for the separation of the estimated CME propagation direction from the Sun-Earth line (which has previously been identified as a significant source of error in CME arrival time forecasts [*Owens and Cargill*, 2004]). Estimates of the CME *V*-*t* profile are presented in row C. The horizontal solid black line is the CME apex speed used in the SWPC forecast, while the horizontal dashed lines mark estimates of the average CME apex speed for each model, derived using the estimate of the apex distance at the observed L1 arrival time, and assuming that the CME was first recorded as appearing halfway through the COR2 FOV at 7.5 R_{Sun}.

There are several sanity checks we can use to assess whether this analysis gives plausible results. First, if the CME was tracked precisely, and if the geometric models are applied appropriately, and if the model assumptions hold, we would expect the *R*-*t* profile to pass through, within the errors, the *R*-*t* coordinate of the CME apex calculated from the observed arrival at L1. The limited extent of the ϵ-*t* profile means that this event could not be tracked out to this distance but, at least for STA, it seems reasonable to assume this is likely. Second, we would expect to see the CME speed vary from somewhere close to the SWPC speed estimate, to somewhere close to but less than the average speed estimate based on the observed L1 arrival time. For STA this appears to work reasonably well until $\approx$35 h, but after this point, each geometric model returns a significant acceleration that would appear to be unphysical. The issue is worse for STB, where the speed profile appears to be nonsense.

As discussed in section 3.3, in our implementation of the geometrical models the CME geometry and propagation direction are fixed. These approximations are frequently used in other research using these methods. If these assumptions are valid, then we would expect the kinematics derived from the STA and STB profiles to be similar. This is not the case, with STA and STB yielding very different CME speeds early in the event and generally opposite speed trends throughout the event. Finally, we would like to draw attention to the lack of distinction between the results for each geometric model. Within the observational errors it is generally not possible to distinguish these models from each other, and at later times when they can be, they appear to show speed variations that are probably unphysical. Although we present only this example here, we have found through looking at a range of events that these types of issues routinely occur.

All of the highlighted issues could result from two factors; poor tracking of the CME front and a breakdown of the applicability of the CME geometric models. These factors are not mutually exclusive and may both interfere with the analysis of a specific event. For example, the *R*-*t* and *V*-*t* profiles depend sensitively on the ϵ-*t* profile, and minimizing the error in the ϵ-*t* profile and improving the stability with which it tracks a feature could possibly resolve the peculiar speed variations at later event times. Furthermore, reducing the errors would make it easier to resolve the different geometric models from each other. The subjectivity and error in manual expert identification of a ϵ-*t* profile is highlighted in Figure 3, which presents a series of STA-HI1 differenced images for the event studied in Figure 2. Overlaid on these frames are the PA-ϵ coordinates of the consensus CME profile tracked in J-maps, along the same PA, by the authors (pink) and the HELCATS project (green). Provided that uncertainty is taken into account, the consensus tracks overlap each other; however, it can be seen that both experts have tracked different regions in the J-map, with neither accurately identifying the front as well as can be done from the full image itself. We are not suggesting that either one or the other

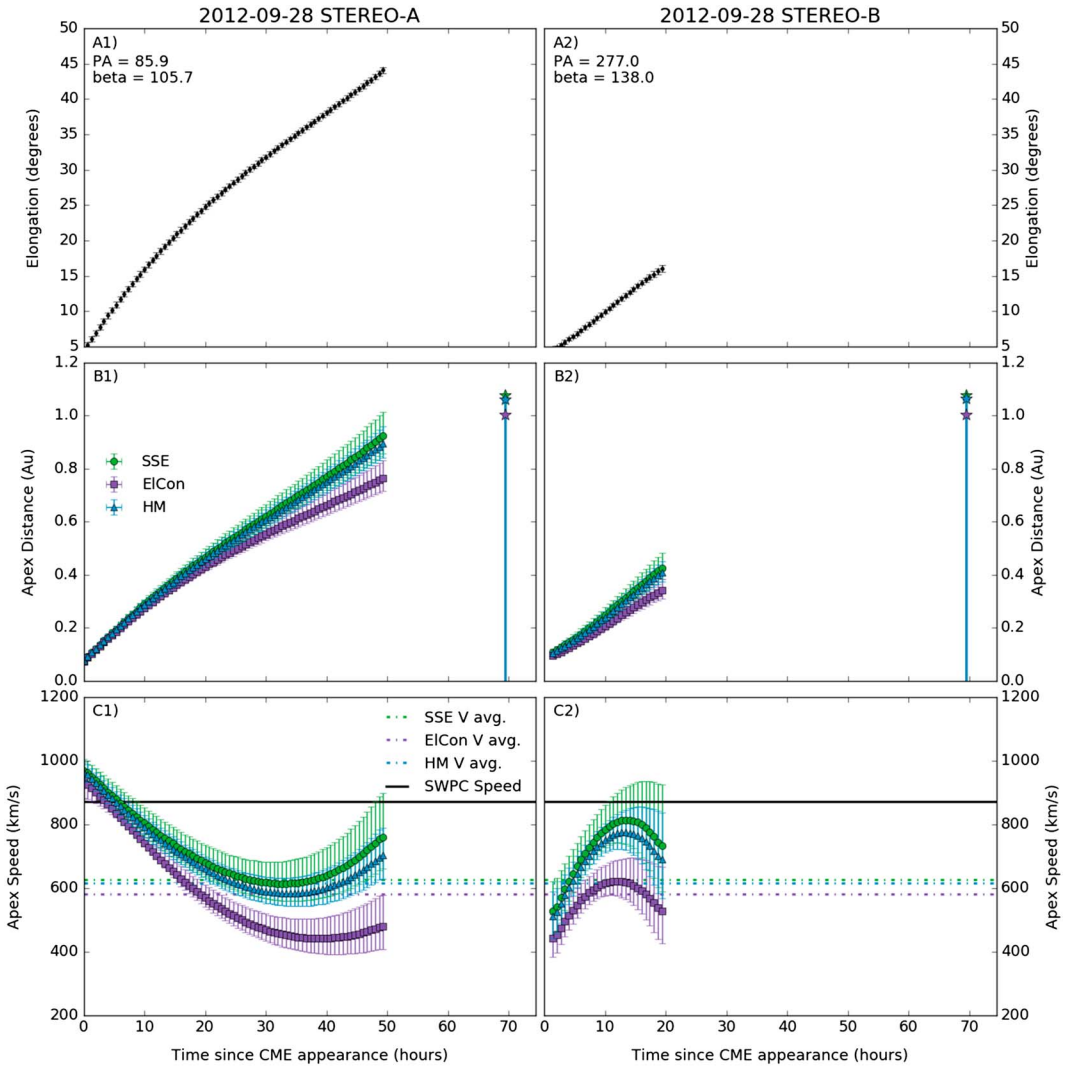

Figure 2. Kinematics of the CME first observed on 28 September 2012, derived from RDG ϵ-t profiles extracted from J-maps of (left column) combined STA HI1 and HI2 images and (right column) combined STB HI1 and HI2 images. (row a) The ϵ-t profiles are the average of six independent identifications along a fixed PA corresponding to the ecliptic plane (see section 3.1). (row b) The CME apex distances are calculated for the SSE (green circles), ElCon (purple squares), and HM (blue triangles) geometric models, using a CME half width and propagation direction fixed by the SWPC coronagraph observations and assuming radial propagation. The colored stars on the vertical lines mark the predicted heliocentric distance of the CME apex for each geometric model at the observed L1 arrival time, assuming the CME geometry persists unchanged. Errors in the distance and speed estimates are based on assumed errors in the propagation direction and CME half width as well as the calculated errors in the average elongation profile. (row c) The CME speed profile is estimated by pointwise numerical differentiation. The black horizontal line marks the CME speed estimate from the SWPC coronagraph analysis, while the green, purple, and blue horizontal dashed lines show an estimate of the average CME apex speed, for the SSE, ElCon, and HM geometries, given the observed CME arrival time at L1 and assuming radial propagation from the coronal signature.

profile is (in)correct, but wish to highlight that this is a difficult and subjective classification, which has significant impact on the estimates of the CME kinematics—a situation that is not ideal should such methods be used for operational space weather forecasting purposes. A method for improved tracking of the CME leading edge, and its impact on CME kinematics, is presented in the next section.

4. Investigating the Effect of Improved CME Tracking on Kinematics Estimates

In section 3, we concluded that the unphysical features in the estimates of the CME kinematics could be due to either or both of poor tracking of the CME in the HI observations, or a failure in the application of the geometric

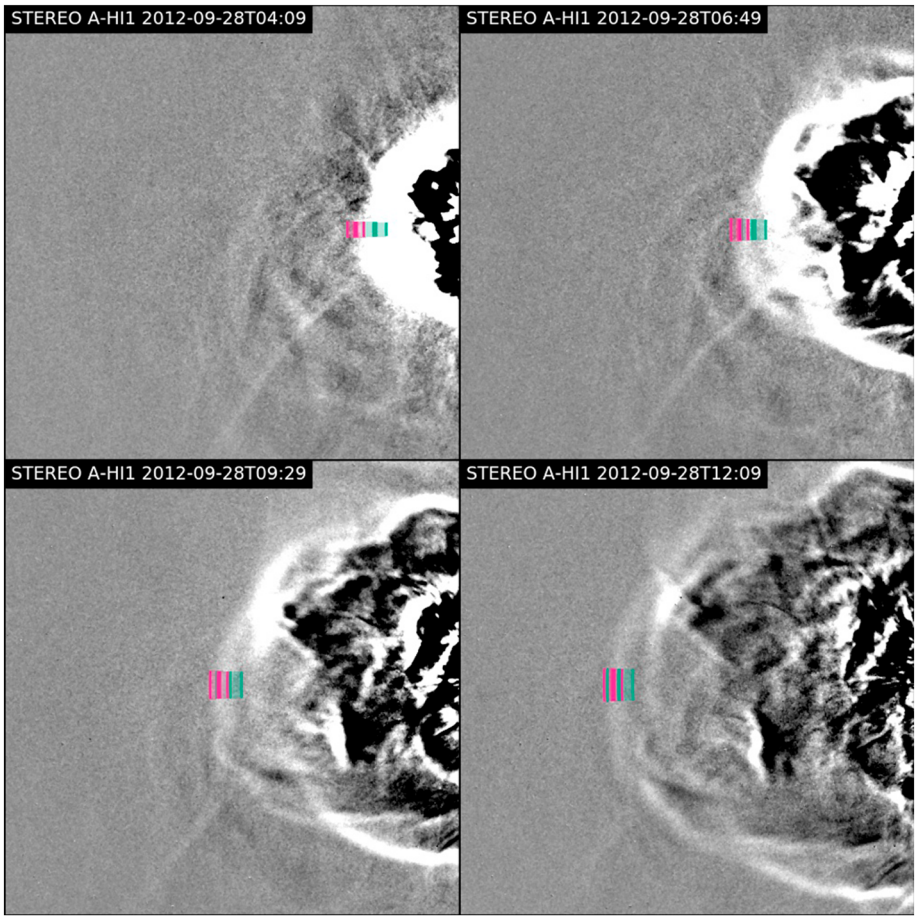

Figure 3. A sequence of STA HI1 differenced images corresponding to the 28 September 2012 CME shown in Figure 2 at different observation times. Overlaid on each image are PA elongation contours marking the elongation of the CME front at that time, derived from its tracking in J-maps, along the same PA, by the authors (pink), and the HELCATS project (green). The thick solid lines show the CME front elongation estimated from averaging the RDG and HCT profiles, while the shaded region shows the uncertainty in the average elongation. The agreement between the RDG and HCT tracking is fair, with the error regions typically overlapping; however, it is clear there is a varying difference between CME front identifications.

models. Here we investigate the effect of improved tracking on the CME kinematics estimates, by developing a new citizen science activity, "Storm Front," which is an extension of the Solar Stormwatch project [*Barnard et al.*, 2015b]. The aim of this activity is to provide a more accurate representation of the CME front in the HI FOV, and in doing so, we aim to provide the geometric models with a better opportunity to reconstruct the kinematics of a CME. We then present estimates of the CME kinematics computed from ϵ-t profiles derived from Storm Front, as is done in section 3.

4.1. Development of the Solar Stormwatch "Storm Front" Elongation Time Profiles

Although section 3 demonstrated that it is worthwhile determining whether a CME front can be more accurately tracked from the HI images, rather than from J-maps at a single position angle, the time required to track the whole CME front in each image would quickly become prohibitive if many events were to be considered. In order to enable a timely study, a citizen science experiment was set up using open-source web-based tools made available via the Zooniverse project (www.zooniverse.org). This task, named Storm Front, presented volunteers with sets of three consecutive HI1 difference images containing the four CMEs analyzed here, drawn from the catalog of Earth-directed events for which SWPC had made forecasts (section 2.1).

In Storm Front, participants are required to draw a polygon around the CME front in the HI images, with the freedom to place up to 50 points anywhere in the HI1 FOV to characterize the shape of the CME front.

This provides, for each HI1 frame analyzed, a distribution of points in PA and ϵ, corresponding to the participants best estimates of where the CME front is located. Each frame was analyzed up to a maximum of 30 times and a minimum of 6 times, while the mean number of characterizations was 20. A consensus estimate of the CME front is then derived through analyzing this distribution of points. Kernel density estimation is used to provide a nonparametric estimate of this distribution [*Silverman*, 1986], from which the elongation of the CME front is estimated along lines of fixed PA, spaced every 1°. For each PA, a transect of the full PA-ϵ distribution is obtained, and the CME front location along this PA is taken as the maximum of the distribution. The uncertainty in the CME front location is calculated as the full width at half maximum of the distribution. The results of this process are demonstrated for the 28 September 2012 CME in Figure 4, while an animation is provided to illustrate the data processing algorithm (https://doi.org/10.6084/m9.figshare.4295141.v1). In Figure 4 (top left), the points show the PA-ϵ coordinates of the CME front as identified by each contributing citizen scientist. The points are color coded according to the frame time, revealing the temporal evolution of the CME front across the STA-HI1 FOV. The correspondingly colored lines mark the consensus estimate of the location of the CME front. The three remaining panels show the consensus estimate of the CME front overlaid onto STA-HI1 differenced images (solid orange), with the uncertainty estimates provided by the orange dashed lines. Generally, the consensus estimate of the CME front location tracks the differenced image brightness enhancement closely. From the presented frames of this event, it is clear that the CME front identification is less reliable at smaller position angles, with the consensus front location becoming more variable and the errors increasing significantly. It is possible that the characterization of this region could be improved by including more classifications. However, this region of increased uncertainty does not affect our subsequent analysis, which focuses on the CME evolution near the ecliptic plane. The Storm Front activity has provided estimates of the CME front for each of the four events being investigated, using HI1 differenced images from STA and STB. As this is only a beta version of the activity, no data yet exists for CMEs other than the four events under investigation. Animations of the CME front estimates for each event, from both the STA and STB perspectives, are provided at https://doi.org/10.6084/m9.figshare.4296239.v1.

As discussed below, this method of tracking CMEs has several advantages over the standard J-map analysis. First, it provides a more detailed picture of the CME front than is offered by a J-map analysis, and the error estimates are more reliable than those obtained from averaging the multiple profiles of one expert. Multiple-observer estimates are less subject to systematic error than repeated estimates from a single observer and thus provide a better quantification of the random error. Second, the Storm Front method removes uncertainty in the time coordinate of the resulting ϵ-t profile, which is present in J-map analysis due to the gridding/interpolation of the data. We refer to the Solar Stormwatch profiles with the acronym SSW. However, one disadvantage is that Storm Front only analyzed data from HI1, whereas the J-map analysis included HI2 observations, which enables CMEs to be tracked out to greater elongations. CMEs cay appear very faint in HI2 images (a low signal-to-noise ratio), and using a J-map can aid the CME identification in the HI2 FOV. We note that although here, for the sake of accuracy and expediency, we have used a citizen science approach, in principle it could be practical for an expert observer/team to characterize CME fronts in HI in a similar manner.

The four events studied here are all quite clear bright CMEs in the HI1 FOV and were also quite simple sequences of activity. It remains to be seen how the citizen science approach we advocate for here will perform on more complex or fainter ejecta and more complex sequences of activity. It is possible that experts may outperform the citizen science approach for complex events. However, we note that in an earlier version of Solar Stormwatch, which characterized CMEs in J-maps, there was no evidence that the citizen science approach was inferior to expert identification at characterizing complex or faint events [*Barnard et al.*, 2014, 2015a].

We are currently preparing to launch the Storm Front activity for a larger number of CMEs, taken from throughout the duration of the STEREO mission. This will provide the opportunity to answer questions concerning the efficacy of the citizen science approach on more complex sequences of activity. Furthermore, this will allow us to assess how the evolving relative positions of STA, STB, and Earth affect the predictive skill of the CME geometric modeling techniques, which are investigated here for only a limited viewing geometry. Based on our experience with earlier Solar Stormwatch projects [*Barnard et al.*, 2014, 2015b], we expect that it will take 12–18 months to collect the classifications required to analyze these events.

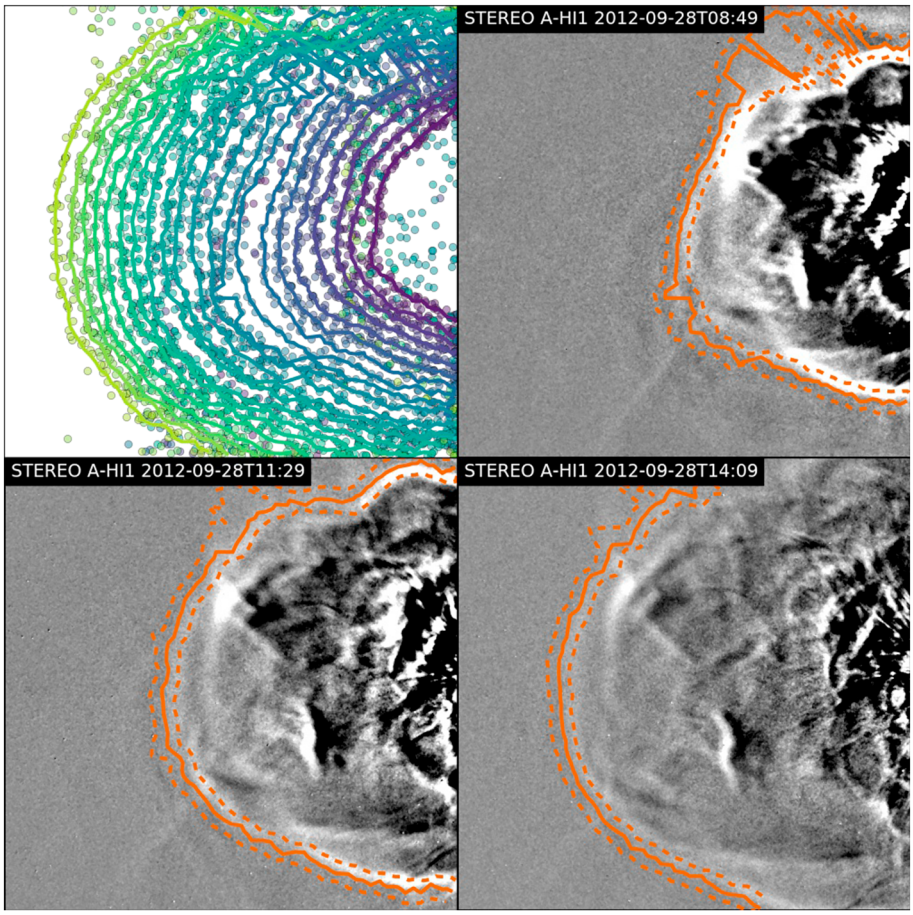

Figure 4. An illustration of how a CME front is characterized by the new Solar Stormwatch activity, Storm Front, for the 28 September 2012 CME, tracked in STA-HI1 differenced images (the same event as in Figure 2). (top left) The points show the PA-ϵ coordinates of the CME front as identified by all contributing citizen scientists. The points are color coded according to the frame time, revealing the temporal evolution of the CME front across the STA-HI1 FOV. The correspondingly colored lines mark the consensus estimate of the CME front location, from combining the identifications of each citizen scientist. The three remaining panels, show the consensus estimate of the CME front location overlaid onto the STA-HI1 differenced images (solid orange) at fixed times, with the uncertainty estimates marked by the orange dashed lines.

4.2. Validation of the Solar Stormwatch Elongation-Time Profiles

Having derived results from the Storm Front activity, it is prudent to compare them with the profiles obtained by the standard J-map analysis (section 3), to assess whether or not they provide a better representation of the CME front. We do this here by comparison with the RDG and HCT profiles.

To this end, animations of HI1 differenced images of each event, with the SSW, RDG, and HCT profiles overlaid, are provided at https://doi.org/10.6084/m9.figshare.4309670.v1. Although these animations are informative, it is desirable to also look more specifically along the PA of interest. Therefore, we have also constructed animations of the evolution of the HI1 difference image intensity as a function of elongation, along the PA corresponding to the HCT ϵ-t profiles. The HI1 image intensity decreases with approximately the square of the elongation, so in these animations we rescale the intensity by the elongation squared, providing a more uniform intensity variation across the HI1 FOV. When averaging the rescaled HI1 differenced image intensity, we use a PA band that is the same width as is used to form the J-maps (5°), and elongation bins 0.1° wide. These animations are available here https://doi.org/10.6084/m9.figshare.4309655.v1, where the solid vertical lines mark the elongation of the CME front determined by SSW, RDG, and HCT. The intensity enhancement associated with the propagating CME front is clearly visible in each movie. These movies indicate that SSW tends to track the same feature of the CME front more stably than either RDG or HCT, which tend to jump around more.

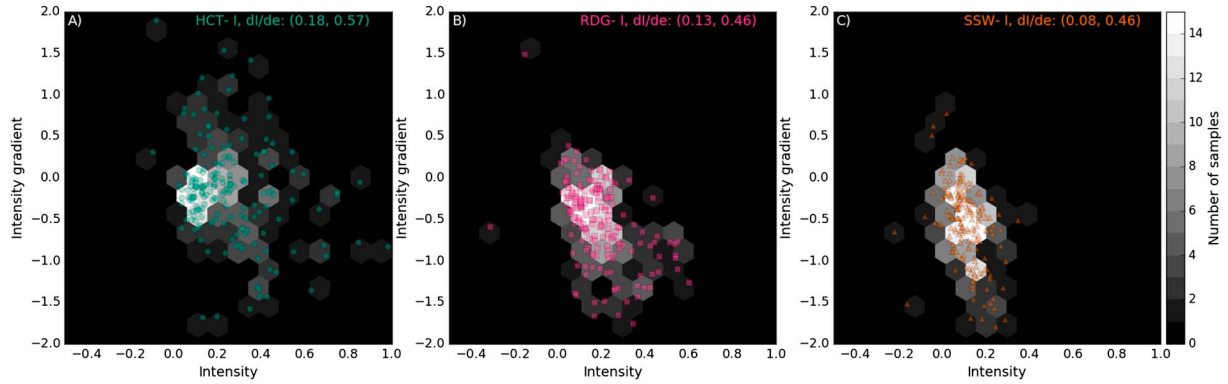

Figure 5. Scatterplots of the differenced image intensity versus the intensity gradient at the time, elongation, and position angle coordinates of the ϵ-t profiles (along the PA tracked by HCT). Shown are the values corresponding to (a) HCT, (b) RDG, and (c) SSW. Beneath the scatterplots are 2-D histograms of the distribution of the samples in each scatterplot. Each histogram uses the same color scaling, as displayed in the color bar. For each panel the HCT, RDG, and SSW samples have been matched, so that only HI frames where each source has identified the CME front are included, thus ensuring a fair comparison between sources. In the top right corner of each plot, the standard deviation of the intensity and intensity gradient coordinates of the scattered points is printed, demonstrating that, overall, SSW displays less scatter than either HCT or RDG.

To investigate this quantitatively, we first posit that a feature is tracked more stably if the resulting profile tends to lie in a similar region of the intensity enhancement associated with the CME front, meaning that similar values of elongation-scaled intensity and intensity gradient are tracked with increasing elongation. For example, if the tracked feature corresponded to the middle of the intensity enhancement, we may expect the profile to result in high intensity values, but low intensity gradient values, as it is near the intensity maximum. Whereas if the tracked feature was on the leading edge of the intensity enhancement, we may expect generally lower intensity values, but higher intensity gradient values, as the intensity decreases from the intensity peak to background levels. When tracking a feature, ideally a system should try to consistently identify one aspect of that feature. Otherwise the profile will have elongation variability resulting from tracking different parts of the CME, which could be an important source of error when estimating the CME kinematics.

For all four events, and for both STA and STB, we have calculated the elongation-scaled intensity and intensity gradient at each ϵ coordinate for the HCT, RDG, and SSW ϵ-t profiles, all tracked along the same PA as the HCT profile. Figures 5a–5c presents these data showing the results for HCT, RDG, and SSW, respectively. The panels display scatterplots of intensity versus intensity gradient, and also 2-D histograms of these data (each panel uses the same scaling for the histogram colors). For this analysis, the data from the HCT, RDG, and SSW sources have been matched, so that only frames where the CME front was tracked by every source are included. Consequently, each distribution is composed of the same number of 192 samples, from frames where each source tracked a CME, ensuring a fair comparison between sources. For all three sources, a small fraction of samples fall outside of the axis limits, which have been chosen to provide a sensible presentation of the bulk of distribution. From the HCT, RDG, and SSW profiles, 10, 1, and 2 samples fall outside the axis limits, which is in each case a small proportion. Visually comparing these panels, there appears to be less scatter in the distribution of the SSW values than in the HCT and RDG distributions. This is confirmed by calculating the standard deviation in intensity and intensity gradient for each panel (values listed in the panels), which were only calculated for samples within the axis limits. Each of SSW, RDG, and HCT has a similar spread in intensity gradient values, but SSW has less spread in intensity values, resulting in a generally more compact distribution. Therefore, on the basis of the analysis presented in this section, we conclude that SSW system yields a more stable tracking of the intensity enhancement associated with a CME in HI1 differenced images than is provided by the consensus profiles derived from a J-map analysis.

4.3. CME Kinematics Derived From Storm Front Tracking

The CME kinematics estimates for the four events in Table 1, derived from the SSW ϵ-t profiles, are presented in Figures 6–9. Each figure has the same format. Row a displays the R-t profile of the CME apex for the SSE (green circles), ElCon (purple squares), and HM (blue triangles) models. Row b presents the R-t profile of the

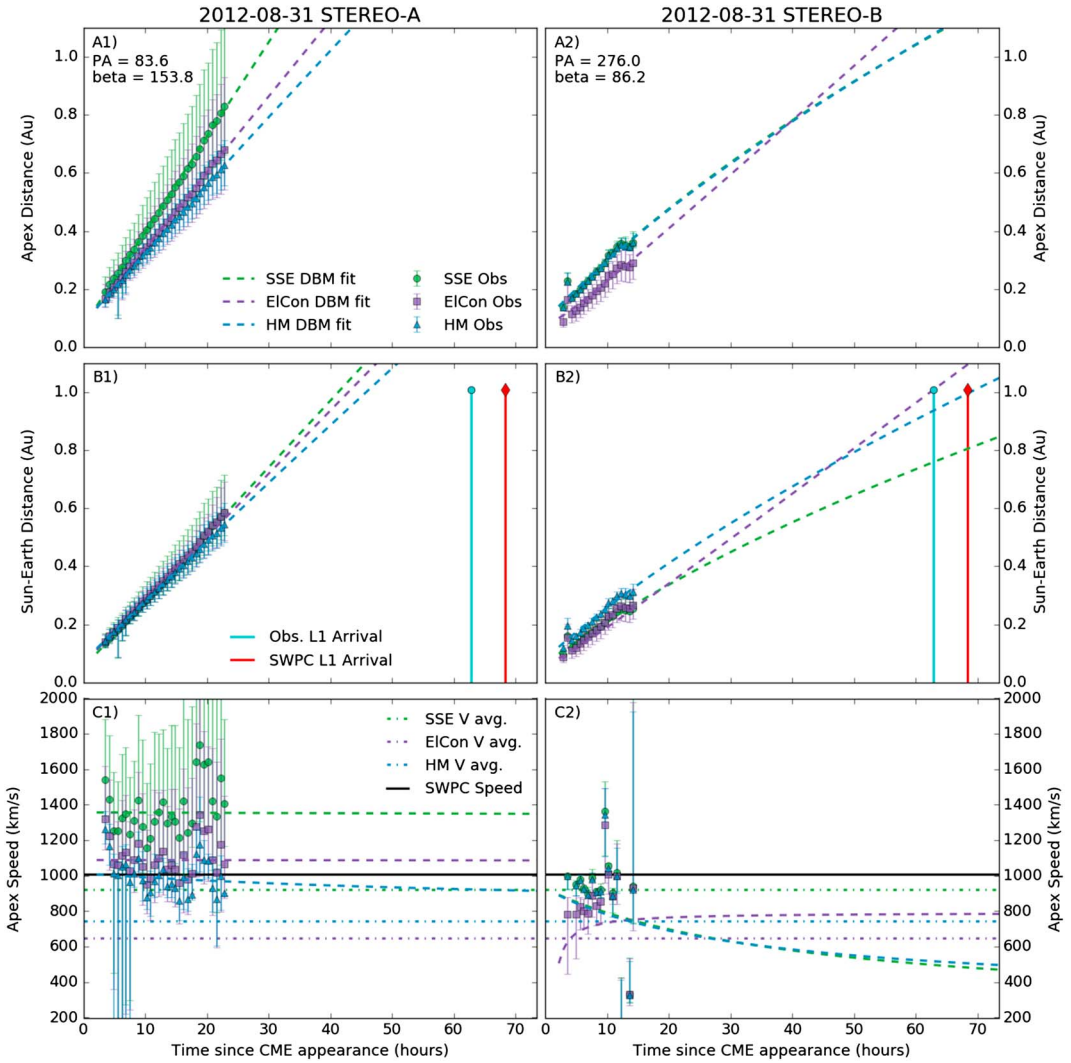

Figure 6. The kinematics estimates for the 31 August 2012 CME, derived from the SSW profiles. (left column) The results for STA, and (right column) the STB results. In each panel, the observational estimates of the CME parameters for the SSE, ElCon, and HM models are shown by the green circles, purple squares, and blue triangles, respectively. The DBM fits to the observed profiles are shown by the correspondingly colored dashed lines. (row a) The CME apex distance profile. (row b) The distance profile of the CME front along the Sun-Earth line and additionally the observed L1 arrival (cyan line) and SWPC forecast arrival (red line). (row c) The CME apex speed profile, while the horizontal black line shows the SWPC estimate of the CME speed in the coronagraph field of view. The dash-dotted lines mark the estimated average speed of the CME apex for each geometric model, computed from the estimated CME apex distance at the observed L1 arrival time.

CME front along the Sun-Earth line for each model, as well as the observed L1 arrival (circle-topped cyan line) and the SWPC forecast L1 arrival (diamond-topped red line). The CME apex *V-t* profile is shown in row c, as is the SWPC estimate of the CME speed established from the coronagraph analysis (horizontal black line), and the average apex speeds estimated from the observed L1 arrival and assuming the CME appeared halfway through the COR2 FOV, at 7.5 R_{Sun} (dash-dotted lines). The results for STA and STB are presented in the left and right columns, respectively.

In addition to presenting the kinematics derived from the SSW ϵ-*t* profiles, we have also computed a best fit drag-based model (DBM) profile [*Vršnak et al.*, 2013] of the CME apex distance. A description of the DBM, and our implementation of it, is given in Appendix A. Fitting the DBM to these events provides a prediction of the CME kinematics outside of the tracked range. In particular, it allows us to compute an estimate of the CME

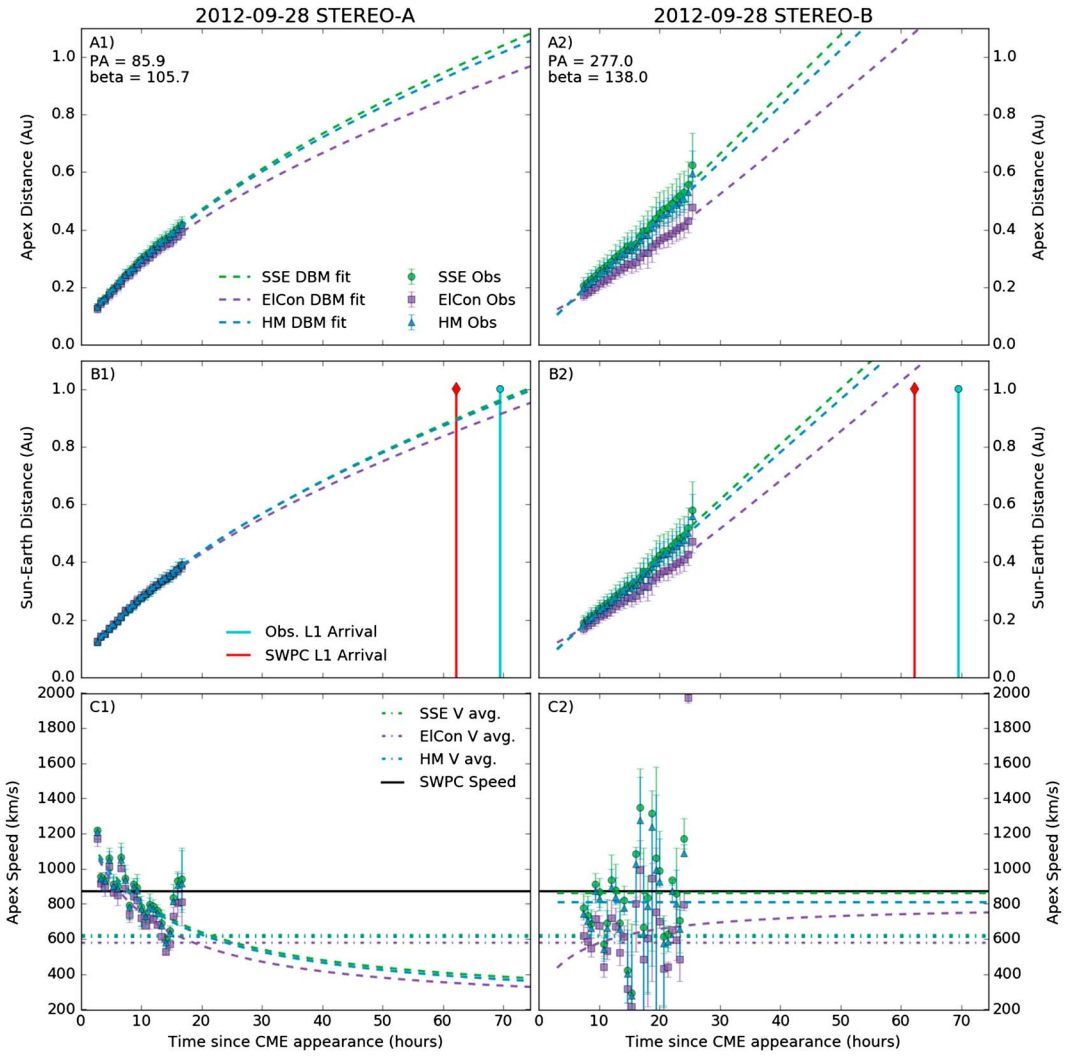

Figure 7. The same as Figure 6, except for the 28 September 2012 CME. This is the same event as analyzed in Figure 2.

arrival in geospace, facilitating a comparison with the SWPC forecast and observed arrival times. The DBM fits to each of these modeled profiles are marked by the correspondingly colored dashed lines in Figures 6–9.

Similar issues persist in the CME kinematics estimates derived from Storm Front profiles as were found in section 3. Visually, there is no clear advantage to either of the HM, SSW, or ElCon models; each shows much variability and none of them appears to systematically provide better predictions. It is typically the case that the differences between the CME kinematics derived from the SSE, ElCon, and HM models are less than the observational uncertainties. This suggests that, at least for these events, given the typical observational errors there is little advantage to using the more complex CME geometric models with additional free parameters.

The 31 August 2012 and 28 September 2012 CMEs (Figures 6 and 7) show large differences between the kinematics estimates derived from STA and STB, but there is reasonable agreement between the STA and STB kinematics estimates for the 5 October 2012 and 20 November 2012 CMEs (Figures 8 and 9).

In all but two cases (STEREO-B on 31 August 2012 and STEREO-A on 28 September 2012), using the geometric models in combination with the best fit DBM results in predicted L1 arrival times that are earlier than observed, implying that such methods are overestimating the CME speed. This is supported by the apex speed profiles, which in six of eight examples show little variation from the SWPC coronagraph speed, and in some instances significant increases above the SPWC coronagraph speed.

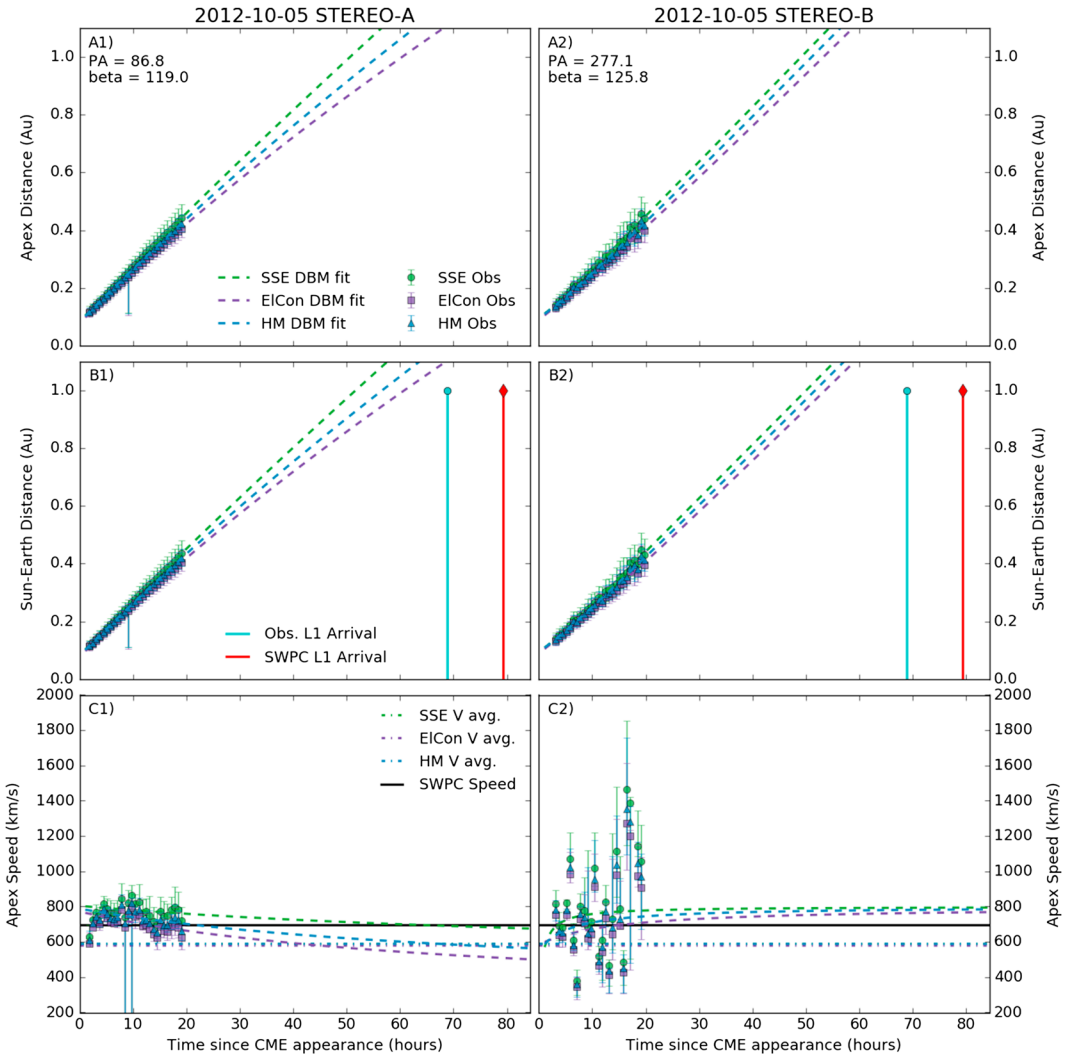

Figure 8. The same as Figure 6, except for the 5 October 2012 CME.

4.4. Comparing the + DBM Predictions With the SWPC Forecasts

Recently, methods such as ElEvoHI [*Rollett et al.*, 2016; *Amerstorfer et al.*, 2016] have been presented as a way to forecast CME arrival, by using the CME kinematics estimated from geometric modeling of HI observations, in conjunction with a best fit DBM. Although we have demonstrated that there are inconsistencies with the CME kinematics derived from the geometrical modeling of the studied events, we still considered it appropriate to test whether using these kinematics estimates in conjunction with a DBM, in a manner similar to ElEvoHi, could provide useful CME arrival time predictions. Here this is achieved by computing the skill of such predictions S_{model} relative to the SWPC forecast, according to

$$S_{\text{model}} = 1 - \frac{e_{\text{model}}}{e_{\text{swpc}}}, \tag{1}$$

where e_{swpc} is the absolute error in the SWPC forecast arrival time and e_{model} is the absolute error in the geometric model + DBM prediction. When $S_{\text{model}} < 0$, the combined geometric modeling + DBM prediction has a larger arrival time error than the SWPC forecast, and when $S_{\text{model}} > 0$, the prediction has a smaller error than the SWPC forecast. The geometric model + DBM arrival time predictions, errors, and skill are presented in Table 2 for STA and Table 3 for STB. Instances where S_{model} is positive have been highlighted in bold. Only 3 of these 24 predictions show positive skill. Perhaps counterintuitively, given that STB-HI ϵ-t profiles are typically noisier, two of these three positive skills scores correspond to STB-HI data. Furthermore, although visually it was unclear if any of the geometric models offered improved performance relative to the others, two of

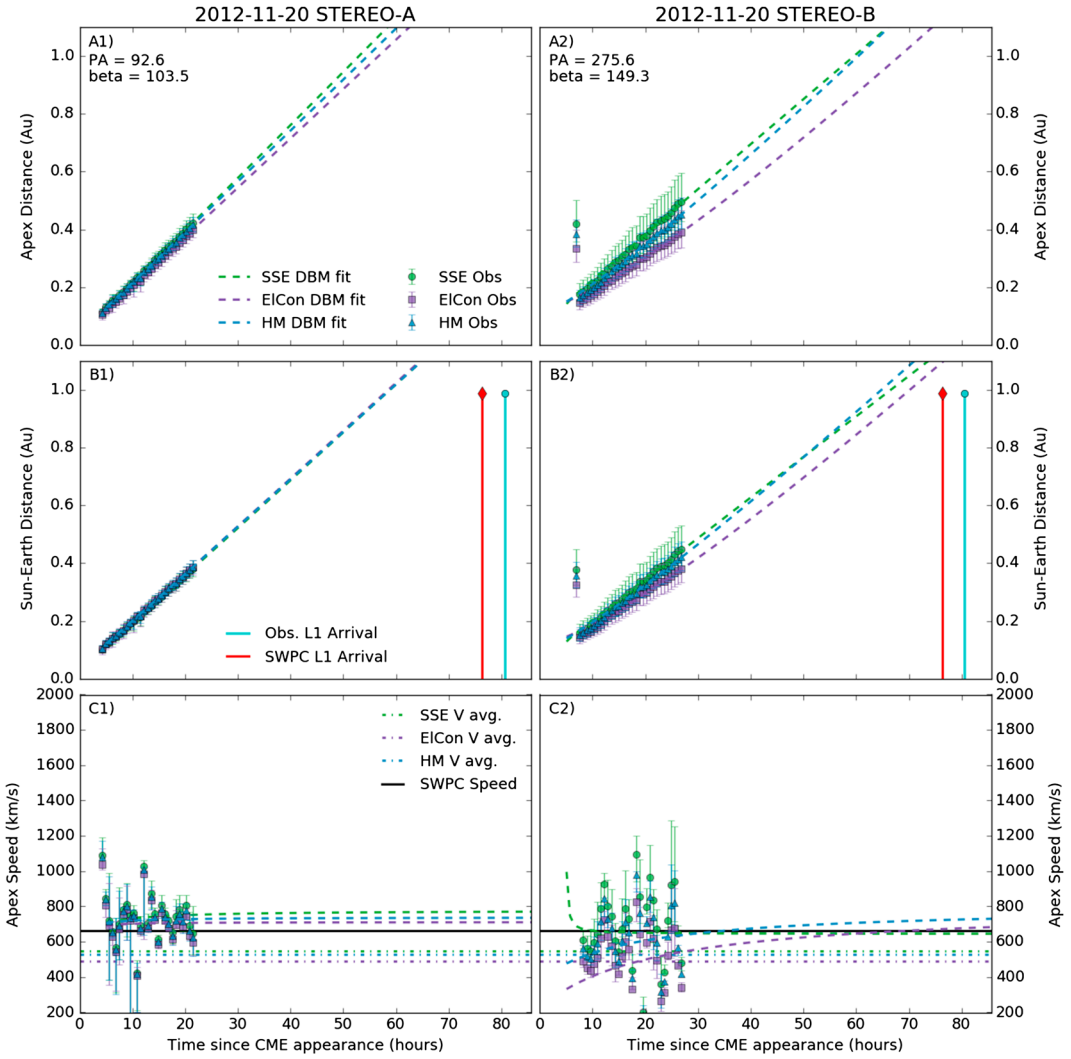

Figure 9. The same as Figure 6, except for the 20 November 2012 CME.

the three positive skill values were obtained with the ElCon model. This is perhaps tentative evidence that ElCon does offer a performance improvement over the SSE and HM models. Table 4 presents the mean skill for the SSE, ElCon, and HM models, averaged across the four events, for STA, STB, and both spacecraft pooled. The error values are 1 standard error of the mean. All values are less than zero, demonstrating that on average the combined geometric modeling + DBM models provide worse predictions than the SWPC forecasts. Considering the average skills from the pooled results, the ElCon model is technically the least negative, but the uncertainties of the average skill for each geometric model are sufficiently large that it is not possible to identify one geometric model as providing better predictions than the others.

Note that with this analysis we are not arguing that this directly identifies issues with the geometric modeling, as there are also significant uncertainties in the application of the DBM. Furthermore, the few events considered means these statistics are likely not robust, and care must be taken not to overinterpret them. However, we interpret this as tentative evidence that combining CME geometric modeling with a simple DBM is unlikely to provide better predictions than the standard SWPC forecast. A future study will extend this work to analyze a larger number of events.

4.5. Variation of the Geometric Model + DBM Predictions With Maximum Tracked Elongation

In the previous section we assessed the performance of the combined geometric model + DBM predictions computed from the full HI1 ϵ-t profile for each event. This was based on the assumption that using all of the

Table 2. Prediction Results for STEREO-A[a]

Obs. Arrival	SWPC Arrival	SWPC Error (min)	Model	Arrival	Error (min)	Skill
			HM	2012-09-02T19:02	−980	−1.91
2012-09-03T11:23	2012-09-03T17:00	−337	SSE	2012-09-02T14:26	−1256	−2.73
			ElCon	2012-09-02T16:02	−1160	−2.44
			HM	2012-10-01T06:13	480	−0.11
2012-09-30T22:13	2012-09-30T15:00	433	SSE	2012-10-01T05:35	442	−0.02
			ElCon	2012-10-01T11:13	780	−0.80
			HM	2012-10-07T16:14	−736	−0.17
2012-10-08T04:31	2012-10-08T15:00	629	SSE	2012-10-07T11:17	−1033	−0.64
			ElCon	2012-10-07T20:02	−508	**0.19**
			HM	2012-11-22T22:22	−1369	−4.43
2012-11-23T21:12	2012-11-23T17:00	−252	SSE	2012-11-22T22:18	−1373	−4.45
			ElCon	2012-11-22T22:06	−1385	−4.50

[a]Dates are formatted as yyyy-mm-ddTHH:MM. Instances where S_{model} is positive have been highlighted in bold.

available ϵ-t profile to constrain the geometric models and DBM would provide the best representation of the CME evolution. However, it is also informative to analyze how the skill of the geometric model + DBM predictions vary as a function of the maximum tracked elongation. For example, with this analysis we can begin to answer interesting questions such as: Is there a minimum elongation over which the CME kinematics must be characterized to obtain a useful DBM fit? When used with the DBM, do any of the geometric models offer a performance improvement in different elongation ranges?

To address this, we considered it prudent to synthesize such a process with the four events studied here. For each event, profiled from both the STA and STB perspectives, we computed the predicted CME L1 arrival time iteratively, from a minimum of the first four HI frames in which the CME was tracked (the minimum required to provide a least squares DBM fit), adding subsequent frames sequentially. For each iteration we calculated the prediction skill relative to the SWPC forecast. Figure 10 presents these skill scores, where each panel shows the skill of the SSE (green circles), HM (blue triangles), and ElCon (purple squares) models as a function of the maximum tracked elongation. The red horizontal line marks the zero skill threshold. The top row shows results for STA, while the bottom row presents results for STB; the four columns correspond to the four events.

Table 3. Prediction Results for STEREO-B[a]

Obs. Arrival	SWPC Arrival	SWPC Error (min)	Model	Arrival	Error (min)	Skill
			HM	2012-09-03T03:34	−468	−0.39
2012-09-03T11:23	2012-09-03T17:00	−337	SSE	2012-09-03T16:42	319	**0.05**
			ElCon	2012-09-03T10:51	−31	**0.91**
			HM	2012-09-30T05:47	−985	−1.28
2012-09-30T22:13	2012-09-30T15:00	433	SSE	2012-09-30T04:12	−1080	−1.50
			ElCon	2012-09-30T12:59	−553	−0.28
			HM	2012-10-07T11:07	−1043	−0.66
2012-10-08T04:31	2012-10-08T15:00	629	SSE	2012-10-07T09:38	−1132	−0.80
			ElCon	2012-10-07T12:39	−951	−0.51
			HM	2012-11-23T14:35	−396	−0.57
2012-11-23T21:12	2012-11-23T17:00	−252	SSE	2012-11-23T07:55	−796	−2.16
			ElCon	2012-11-23T10:01	−670	−1.66

[a]Dates are formatted as yyyy-mm-ddTHH:MM. Instances where S_{model} is positive have been highlighted in bold.

Table 4. Summary of Prediction Skill for HM, SSE, and ElCon Models With DBM

	Mean Skill		
Model	STA	STB	Both
HM	-1.66 ± 0.88	-0.72 ± 0.17	-1.19 ± 0.48
SSE	-1.96 ± 0.88	-1.10 ± 0.41	-1.53 ± 0.51
ElCon	-1.89 ± 0.89	-0.39 ± 0.46	-1.14 ± 0.57

The results are surprising; in general, they do not suggest that profiling a CME further into the heliosphere provides a more reliable prediction of the CME arrival. In fact, in three of the eight cases there is a clear negative trend in the prediction skill, for all models, with increasing maximum elongation. This plot also highlights the modest differences between the SSE, ElCon, and HM models. In most cases the skill of the models is similar and demonstrates the same trend. The 28 September 2012 CME tracked in STB is the only clear exception to this, where the ElCon model appears to perform significantly better than the others over the full elongation range; however, we note that this is not the case for the STA tracking of this event, where the models perform similarly.

It would be interesting to extend this analysis to include HI2 observations, tracking the CME out to larger elongations/heliocentric distances. Such an analysis might reveal what aspect of the combined CME geometric modeling and drag modeling is limiting the utility of this approach for these events. Furthermore, a better understanding of the usefulness of this approach could be obtained by repeating this analysis for events with different relative positions of STA, STB, and Earth, which are generally better for viewing Earth-directed CMEs.

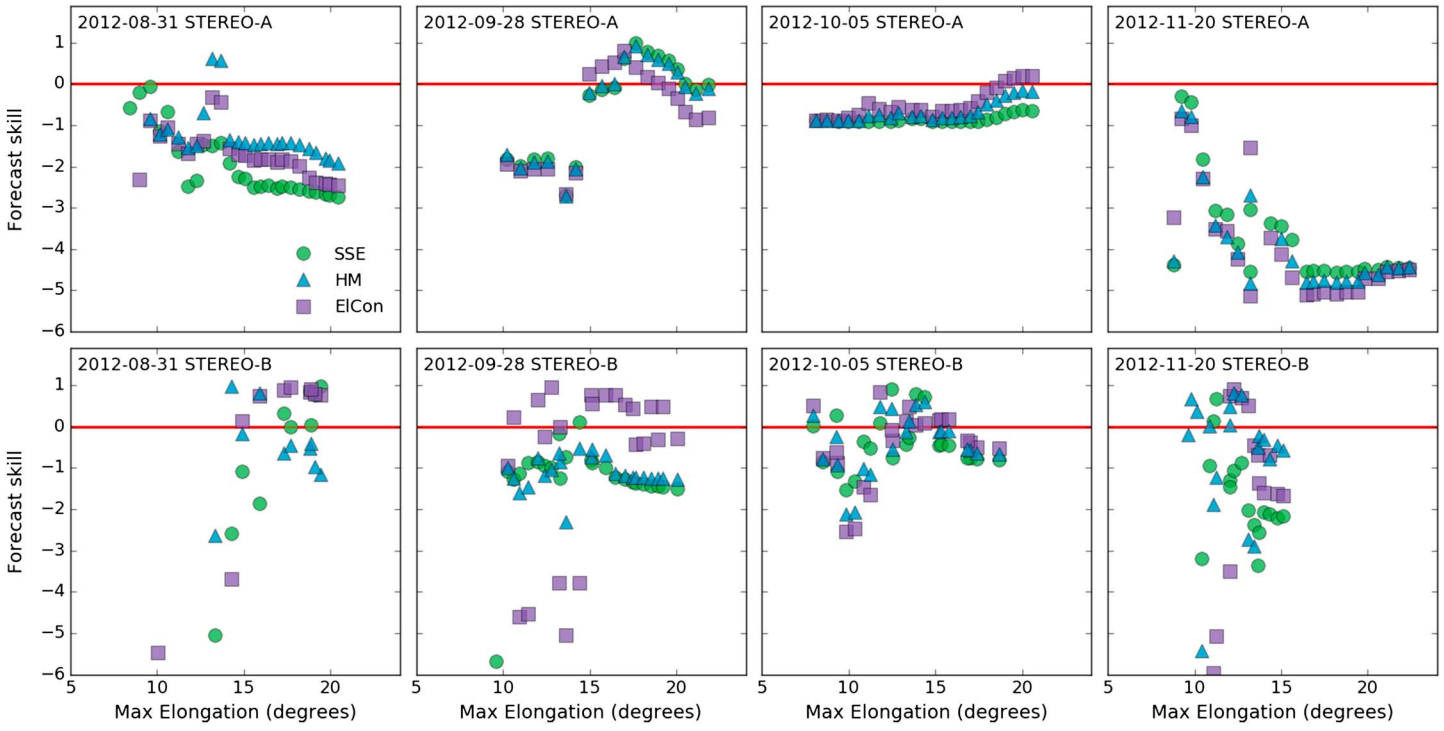

Figure 10. Skill of the combined geometric model + DBM predictions (relative to the SWPC forecast), as a function of the maximum elongation as the CME was tracked through the HI FOV. The four columns correspond to the four events studied, while the top and bottom rows provide the results for STA and STB, respectively. The horizontal red line marks the zero skill threshold. Points with positive skill (above the red line) offer an improvement in forecast accuracy over the SWPC forecast, while points below the line are indicative of poorer forecast accuracy.

5. Discussion

Throughout this report we have demonstrated that inconsistent results are obtained for the kinematics of CMEs estimated through the application of CME geometric models to ϵ-t profiles extracted from HI data, at least for the four events that form the basis of this study. Considering the kinematics estimates obtained from STA and STB in isolation, we have seen examples of late stage accelerations in the CME apex that appear unphysical, and CME apex speeds that are inconsistent with both the speeds obtained from the analysis of coronagraph data, and the mean transit speed estimated from the observed L1 arrival. This is the case for each of the geometric models considered, and often, the results are very similar for each model, with the differences between them being smaller than the observational errors. The situation is not improved when comparing the kinematics estimates obtained from STA and STB. If the assumptions of the geometric models are met, similar kinematics estimates should be derived from the STA and STB perspectives, which is rarely the case.

As discussed in section 3, it therefore seems that either the CMEs are being poorly tracked through the HI FOV or the geometric models are being inappropriately applied. However, through our development of the SSW CME characterizations, we are confident that we cannot obtain further significant improvements upon the tracking of the CMEs analyzed in this work. Despite the improved tracking developed in this paper, we still observe inconsistencies in the derived CME kinematics. This leads us to consider that the geometric models provide inconsistent results because assumptions inherent to each model are being invalidated. The one assumption common to our implementation of each geometric model is that the various CME geometries are fixed and persist unchanged throughout the CME propagation. We believe it is likely that this assumption is sufficiently invalid that it renders the geometric modeling results generally unreliable. Furthermore, we think it would be beneficial if future development of the single spacecraft fitting techniques focused on improving upon this time-independent geometry assumption. This time independence of CME parameters is not an inherent feature of the geometric models. It is theoretically possible to make parameters such as the CME direction, half width, and aspect ratio time dependent. Numerical studies suggest that the ambient solar wind can have a large effect on CME transit times [*Case et al.*, 2008]. Thus, solar wind structure may well be responsible for the limited skill of the combined CME geometric modeling and DBM predictions. However, in practice, when using only coronagraph and HI observations, it is common to assume that these parameters are stationary (such as CME half width and propagation direction), as there are not the observations or physical understanding to plausibly evaluate their time dependence. However, we are interested to see how the recently published FRiED model may contribute to this area [*Isavnin*, 2016]. Furthermore, the remote sensing package on Solar Orbiter, in particular the METIS and SoloHI instruments, may well provide the out-of-ecliptic CME observations needed to improve our understanding of the time dependence of these parameters.

We used the best fit DBM to extrapolate the CME kinematics derived from the geometric models to compare the predicted L1 arrival time with the SWPC forecast arrival. It was found that the predictions obtained from extrapolating the geometric models results with the DBM rarely offered an improvement over the SWPC forecast. Of the geometric models tested, the ElCon model provided the most positive results and, on average, provided slightly more skilful predictions than the SSE and HM models—all still worse than the SWPC forecast. These results are more difficult to interpret, due to the coupling of the geometric models with the DBM, which has its own limitations and uncertainties.

This analysis was then extended by calculating the combined geometric model + DBM predictions iteratively throughout the evolution of the event. This revealed that more reliable predictions are not necessarily obtained by tracking the CME further out through the HI FOV. Across the four events studied from both STA and STB, there was no clear positive trend in the predictive skill with the maximum elongation of the CME tracked in HI. Indeed, in several instances the skill of the predictions decreased with maximum tracked elongation. Differences in skill between the geometric models were generally small and skills for each model followed similar trends throughout events. Our interpretation of this result is similar to the arguments above; that this is evidence that the assumptions of the geometric models are invalid and hence using the models to estimate the CME kinematics through the heliosphere provides unreliable results. We should ask how useful these CME geometric modeling techniques are in an operational space weather forecasting context?

A potential limitation in this analysis arises from the heliospheric location of STA and STB relative to these Earth-directed CMEs. For the FP, HM, and SSE fitting techniques, where a least squares fit is obtained to the CME direction and speed, assuming these are both constant, it is known that the results become increasingly unreliable as the CME direction increases toward and beyond the plane of the sky [*Williams et al.*, 2009;

Lugaz et al., 2009; *Davies et al.*, 2012]. For each of the four CMEs studied, the SWPC estimate of the CME propagation directions suggests that the CME will be propagating significantly into the plane of the sky from either the STA or STB perspective. If plans for a space weather monitor situated at the L5 Lagrange point (such as the proposed Carrington mission) are realized, the HI viewing geometry for Earth-directed CMEs would typically be better than was the case in this study. However, for longitudinally broad ejecta, it is quite plausible that geoeffective events would be observed to propagate well into the plane of sky for a HI-like instrument from the L5 perspective. This motivates the development of techniques that continue to perform well under such viewing geometries.

In the near future we will launch the Storm Front activity with a larger number of CMEs, taken from throughout the STEREO mission. This will enable us to extend this study to examine the effect of viewing geometry on the predictive skill of CME geometric modeling combined with DBM to forecast L1 arrival of CMEs. As the events studied here have a suboptimal viewing geometry, it is possible that the extended analysis will reveal an improved performance of the CME geometric modeling techniques. However, we will generally be unable to compare CMEs observed with more favorable STEREO geometries against genuine SWPC operational forecasts and will instead have to compare against WSA-ENLIL CAT model hindcasts of these events.

Our study has focused on techniques applicable to single-spacecraft observations, rather than any stereoscopic techniques [e.g., *Liu et al.*, 2010], as in the future it is possible that there will be an HI-like instrument included on an operational space weather monitoring spacecraft, but it is less likely there will be two operational HI-like instruments for space weather forecasting. We do not intend that our results should be interpreted as arguing against the utility of HI-like observations for space weather forecasting; in fact, we are optimistic about their potential utility. The techniques investigated here are a first step in investigating what may be possible. For example, we would be excited to see how HI observations may be incorporated into a data assimilation scheme coupled to a heliospheric model, or to see what improvements in CME localization could be obtained with a polarized heliospheric imager [*DeForest et al.*, 2016].

Appendix A: Drag Model Fitting

In section 4 we employ the drag-based model (DBM) of *Vršnak et al.* [2013], using the implementation described by *Rollett et al.* [2016]. The DBM approximates the dynamics of a CME by assuming that a drag force acts upon it, the magnitude of which depends on the background solar wind mass flux and the CME size, speed, and density. Under the assumptions that the CME drag parameter γ is constant and that the background solar wind speed is constant and homogeneous, the DBM provides an analytical solution for the heliocentric distance of the CME apex as a function of time,

$$R(t) = \pm\frac{1}{\gamma} ln[1 \pm \gamma(v_{\text{init}} - w)t] + wt + r_{\text{init}} \tag{A1}$$

where t is time since CME initiation (strictly speaking time since the DBM becomes applicable, which is different from the CME initiation time in the corona), r_{init} and v_{init} are the initial CME starting distance and speed, and w is the background solar wind speed. Therefore, with an observed R-t profile of the CME apex, it is possible to calculate a best fit DBM, which can then be extrapolated to provide a forecast of the CME arrival (and speed) elsewhere in the heliosphere. In sections 3 and 4.3 we calculate a least squares fit to the DBM using the R-t series derived from the application of the CME geometric models to the ϵ-t profiles observed by HI.

We compute a bounded fit for w, γ, r_{init}, and v_{init}, using the bounds listed in Table A1, where r_{swpc} and v_{swpc} are the initial distance and speed used in the SWPC forecasts to insert the CME into the WSA-ENLIL model. There is no supporting evidence for the chosen bounds, but from our review of the literature, they appear to be reasonable without overconstraining the fit. We note that this is a relatively simple formulation of the DBM and are aware of recent works to develop this approach, for example, by including more realistic background speed profiles that vary both temporally and spatially. This is something we are interested to incorporate in our future studies.

Table A1. Bounds of Parameter Fitting to Drag-Based Model

	w	γ	r_{init}	v_{init}
Lower bound	200 km s^{-1}	10^{-10} m^{-1}	$0.5 \times r_{\text{swpc}}$	$0.5 \times r_{\text{swpc}}$
Upper bound	800 km s^{-1}	10^{-5} m^{-1}	$1.5 \times r_{\text{swpc}}$	$1.5 \times r_{\text{swpc}}$

Acknowledgments
This publication has been made possible by the participation of volunteers in the Solar Stormwatch project (http://www.solarstormwatch.com/authors). This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation. L.B. and M.J.O. thank the Science and Technology Facilities Council for support under grant ST/M000885/1. C.A.D.K thanks AFOSR and NASA LWS/TR, and T for support under grants FA9550-14-1-0401, NNX09AJ84G, and NNX15AF39G, respectively. We thank the STEREO/HI instrument team at the Rutherford Appleton Laboratory for providing the HI data. This research has made use of SunPy, an open-source and free community-developed solar data analysis package written in Python [*The SunPy Community et al.*, 2015]. The HELCATS data are available at https://www.helcats-fp7.eu/. The University of Reading identifications and the Solar StormWatch Storm Front CME front data are available on request from l.a.barnard@reading.ac.uk.

References

Amerstorfer, T., C. Möstl, A. Isavnin, J. A. Davies, M. Kubicka, U. V. Amerstorfer, and R. A. Harrison (2016), Erratum: "ElEvoHI: A novel CME prediction tool for heliospheric imaging combining an elliptical front with drag-based model fitting", *Astrophys. J.*, *831*(2), 210, doi:10.3847/0004-637X/831/2/210.

Barnard, L., et al. (2014), The solar stormwatch CME catalogue: Results from the first space weather citizen science project, *Space Weather*, *12*, 657–674, doi:10.1002/2014SW001119.

Barnard, L., C. J. Scott, M. Owens, M. Lockwood, S. R. Crothers, J. A. Davies, and R. A. Harrison (2015a), Differences between the CME fronts tracked by an expert, an automated algorithm, and the solar stormwatch project, *Space Weather*, *13*, 709–725, doi:10.1002/2015SW001280.

Barnard, L., C. J. Scott, M. Owens, M. Lockwood, K. Tucker-Hood, J. Wilkinson, B. Harder, and E. Baeten (2015b), Solar stormwatch: Tracking solar eruptions, *Astron. Geophys.*, *56*(4), 4.20–4.24, doi:10.1093/astrogeo/atv131.

Cannon, P., et al. (2013), Extreme space weather: Impacts on engineered systems and infrastructure, *Tech. Rep.*, Royal Acad. of Eng., London.

Case, A. W., H. E. Spence, M. J. Owens, P. Riley, and D. Odstrcil (2008), The ambient solar wind's effect on ICME transit times, *Geophys. Res. Lett.*, *35*, L15105, doi:10.1029/2008GL034493.

Conlon, T. M., S. E. Milan, and J. A. Davies (2014), Assessing the effect of spacecraft motion on single-spacecraft solar wind tracking techniques, *Sol. Phys.*, *289*, 3935–3947, doi:10.1007/s11207-014-0549-z.

Davies, J. A., R. A. Harrison, A. P. Rouillard, N. R. Sheeley, C. H. Perry, D. Bewsher, C. J. Davis, C. J. Eyles, S. R. Crothers, and D. S. Brown (2009), A synoptic view of solar transient evolution in the inner heliosphere using the Heliospheric Imagers on STEREO, *Geophys. Res. Lett.*, *36*, L02102, doi:10.1029/2008GL036182.

Davies, J. A., et al. (2012), A self-similar expansion model for use in solar wind transient propagation studies, *Astrophys. J.*, *750*, 23, doi:10.1088/0004-637X/750/1/23.

DeForest, C. E., T. A. Howard, D. F. Webb, and J. A. Davies (2016), The utility of polarized heliospheric imaging for space weather monitoring, *Space Weather*, *14*, 32–49, doi:10.1002/2015SW001286.

EU HELCATS, D. Barnes, J. Byrne, J. Davies, and R. Harrison (2015), HELCATS HCME WP2 V02, doi:10.6084/m9.figshare.1492351.v1.

Eyles, C. J., G. M. Simnett, M. P. Cooke, B. V. Jackson, A. Buffington, P. P. Hick, N. R. Waltham, J. M. King, P. A. Anderson, and P. E. Holladay (2003), The Solar Mass Ejection Imager (SMEI), *Sol. Phys.*, *217*(2), 319–347, doi:10.1023/B:SOLA.0000006903.75671.49.

Eyles, C. J., et al. (2008), The heliospheric imagers onboard the STEREO mission, *Sol. Phys.*, *254*(2), 387–445, doi:10.1007/s11207-008-9299-0.

Gosling, J. T. (1993), The solar flare myth, *J. Geophys. Res.*, *98*, 18,93–18,949, doi:10.1029/93JA01896.

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*(A12), 29,207–29,217, doi:10.1029/2001JA000177.

Hapgood, M. A. (2011), Towards a scientific understanding of the risk from extreme space weather, *Adv. Space Res.*, *47*(12), 2059–2072, doi:10.1016/j.asr.2010.02.007.

Harrison, R. A., et al. (2012), An analysis of the origin and propagation of the multiple coronal mass ejections of 2010 August 1, *Astrophys. J.*, *750*(1), 45, doi:10.1088/0004-637X/750/1/45.

Howard, R. A., et al. (2008), Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI), *Space Sci. Rev.*, *136*, 67–115, doi:10.1007/s11214-008-9341-4.

Howard, T. A., and S. J. Tappin (2009), Interplanetary coronal mass ejections observed in the heliosphere: 1. Review of theory, *Space Sci. Rev.*, *147*, 31–54, doi:10.1007/s11214-009-9542-5.

Howard, T. A., and C. E. DeForest (2012), The Thomson surface. I. Reality and myth, *Astrophys. J.*, *752*, 130, doi:10.1088/0004-637X/752/2/130.

Howard, T. A., D. F. Webb, S. J. Tappin, D. R. Mizuno, and J. C. Johnston (2006), Tracking halo coronal mass ejections from 0–1 AU and Space Weather Forecasting Using the Solar Mass Ejection Imager (SMEI), *J. Geophys. Res.*, *111*, A04105, doi:10.1029/2005JA011349.

Isavnin, A. (2016), FRiED: A novel three-dimensional model of coronal mass ejections, *Astrophys. J.*, *833*, 267, doi:10.3847/1538-4357/833/2/267.

Kahler, S. W., and D. F. Webb (2007), V Arc interplanetary coronal mass ejections observed with the solar mass ejection imager, *J. Geophys. Res.*, *112*, A09103, doi:10.1029/2007JA012358.

Kaiser, M. L., T. A. Kucera, J. M. Davila, O. C. St. Cyr, M. Guhathakurta, and E. Christian (2008), The STEREO mission: An introduction, *Space Sci. Revi.*, *136*, 5–16, doi:10.1007/s11214-007-9277-0.

Liu, Y., J. A. Davies, J. G. Luhmann, A. Vourlidas, S. D. Bale, and R. P. Lin (2010), Geometric triangulation of imaging observations to track coronal mass ejections continuously out to 1 AU, *Astrophys. J. Lett.*, *710*, L82–L87, doi:10.1088/2041-8205/710/1/L82.

Liu, Y., J. G. Luhmann, N. Lugaz, C. Möstl, J. A. Davies, S. D. Bale, and R. P. Lin (2013), On Sun-to-Earth propagation of coronal mass ejections, *Astrophys. J.*, *769*, 45, doi:10.1088/0004-637X/769/1/45.

Lugaz, N. (2010), Accuracy and limitations of fitting and stereoscopic methods to determine the direction of coronal mass ejections from heliospheric imagers observation, *Sol. Phys.*, *267*, 411–429, doi:10.1007/s11207-010-9654-9.

Lugaz, N., A. Vourlidas, and I. I. Roussev (2009), Deriving the radial distances of wide coronal mass ejections from elongation measurements in the heliosphere—Application to CME-CME interaction, *Ann. Geophys.*, *27*, 3479–3488, doi:10.5194/angeo-27-3479-2009.

Lugaz, N., J. N. Hernandez-Charpak, I. I. Roussev, C. J. Davis, A. Vourlidas, and J. A. Davies (2010), Determining the azimuthal properties of coronal mass ejections from multi-spacecraft remote sensing observations with STEREO SECCHI, *Astrophys. J.*, *715*, 493–499, doi:10.1088/0004-637X/715/1/493.

Millward, G., D. A. Biesecker, V. J. Pizzo, and C. A. de Koning (2013), Development of an operational software tool for the analysis of coronagraph images: Determining CME parameters for input into heliospheric models, *Space Weather*, *11*, 57–68, doi:10.1002/swe.20024.

Mishra, W., N. Srivastava, and J. A. Davies (2014), A comparison of reconstruction methods for the estimation of coronal mass ejections kinematics based on SECCHI/HI observations, *Astrophys. J.*, *784*, 135, doi:10.1088/0004-637X/784/2/135.

Mishra, W., and N. Srivastava (2015), Heliospheric tracking of enhanced density structures of the 6 October 2010 CME, *J. Space Weather Space Clim.*, *5*, A20, doi:10.1051/swsc/2015021.

Möstl, C., and J. A. Davies (2013), Speeds and arrival times of solar transients approximated by self-similar expanding circular fronts, *Sol. Phys.*, *285*(1), 411–423, doi:10.1007/s11207-012-9978-8.

Möstl, C., et al. (2014), Connecting speeds, directions and arrival times of 22 coronal mass ejections from the Sun to 1 AU, *Astrophys. J.*, *787*, 119, doi:10.1088/0004-637X/787/2/119.

Möstl, C., et al. (2015), Strong coronal channelling and interplanetary evolution of a solar storm up to Earth and Mars, *Nat. Commun.*, *6*, 7135, doi:10.1038/ncomms8135.

Owens, M. J., and P. Cargill (2004), Predictions of the arrival time of coronal mass ejections at 1 AU: An analysis of the causes of errors, *Ann. Geophys.*, *22*(2), 661–671, doi:10.5194/angeo-22-661-2004.

Pizzo, V. J., G. Millward, A. Parsons, D. A. Biesecker, S. Hill, and D. Odstrcil (2011), Wang-Sheeley-Arge-Enlil cone model transitions to operations, *Space Weather*, *9*, S03004, doi:10.1029/2011SW000663.

Pizzo, V. J., C. A. de Koning, M. Cash, G. Millward, D. A. Biesecker, L. Puga, M. Codrescu, and D. Odstrcil (2015), Theoretical basis for operational ensemble forecasting of coronal mass ejections, *Space Weather*, *13*(10), 676–697, doi:10.1002/2015SW001221.

Rollett, T., C. Möstl, M. Temmer, A. M. Veronig, C. J. Farrugia, and H. K. Biernat (2012), Constraining the kinematics of coronal mass ejections in the inner heliosphere with in-situ signatures, *Sol. Phys.*, *276*, 293–314, doi:10.1007/s11207-011-9897-0.

Rollett, T., C. Möstl, A. Isavnin, J. A. Davies, M. Kubicka, U. V. Amerstorfer, and R. A. Harrison (2016), ElEvoHI: A novel CME prediction tool for heliospheric imaging combining an elliptical front with drag-based model fitting, *Astrophys. J.*, *824*, 131, doi:10.3847/0004-637X/824/2/131.

Rouillard, A. P., et al. (2008), First imaging of corotating interaction regions using the STEREO spacecraft, *Geophys. Res. Lett.*, *35*, L10110, doi:10.1029/2008GL033767.

Sheeley, N. R., Jr., J. H. Walters, Y.-M. Wang, and R. A. Howard (1999), Continuous tracking of coronal outflows: Two kinds of coronal mass ejections, *J. Geophys. Res.*, *104*, 24,739–24,767, doi:10.1029/1999JA900308.

Silverman, B. W. (1986), *Density Estimation For Statistics and Data Analysis*, Chapman and Hall, London.

The SunPy Community et al. (2015), SunPy—Python for solar physics, *Comput. Sci. Discovery*, *8*, 014009, doi:10.1088/1749-4699/8/1/014009.

Tucker-Hood, K., et al. (2015), Validation of a priori CME arrival predictions made using real-time heliospheric imager observations, *Space Weather*, *13*, 35–48, doi:10.1002/2014SW001106.

Vršnak, B. (2001), Deceleration of coronal mass ejections, *Sol. Phys.*, *202*, 173–189, doi:10.1023/A:1011833114104.

Vršnak, B., et al. (2013), Propagation of interplanetary coronal mass ejections: The drag-based model, *Sol. Phys.*, *285*, 295–315, doi:10.1007/s11207-012-0035-4.

Williams, A. O., J. A. Davies, S. E. Milan, A. P. Rouillard, C. J. Davis, C. H. Perry, and R. A. Harrison (2009), Deriving solar transient characteristics from single spacecraft STEREO/HI elongation variations: A theoretical assessment of the technique, *Ann. Geophys.*, *27*(12), 4359–4368, doi:10.5194/angeo-27-4359-2009.

Wood, B. E., R. A. Howard, and D. G. Socker (2010), Reconstructing the morphology of an evolving coronal mass ejection, *Astrophys. J.*, *715*(2), 1524, doi:10.1088/0004-637X/715/2/1524.