Predicting CME Ejecta and Sheath Front Arrival at L1 with a Data-Constrained Physical Model

Phillip Hess and Jie Zhang

School of Physics, Astronomy and Computational Sciences, George Mason University 4400
University Dr. Fairfax, VA 22030

`phess4@gmu.edu`

Received ________________________; accepted ________________________

ABSTRACT

We present a method for predicting the arrival of both a CME flux rope driver in-situ, as well as the sheath of solar wind plasma accumulated ahead of the driver. For faster CMEs, the front of this sheath will be a shock. The method is based upon using measurements to superimpose geometries onto both the CME ejecta and sheath separately. These measurements are used to constrain a Drag-Based Model, which is improved by including both a height dependence and accurate de-projected velocities. We also attempt to constrain the geometry of the model to determine the error introduced as a function of the deviation of the CME from the Sun-Earth line. Combining all these factors allows us to create predictions for both fronts at the L1 point and compare against observations. We demonstrate an ability to predict the sheath arrival with an average error of under 4 hours, with an RMS error of about 1.5 hours. For the ejecta the error is less than two hours with an RMS error within an hour. In addition to testing the ability to recreate the arrival of the CME structures at L1, we discuss the physical implications of our model for CME expansion and density evolution. We show the power of our method when the best data possible is available and demonstrate the practical implications a permanent L5 observer on space weather forecasting capabilities, while also discussing the limitations of the method that will have to be addressed to create a realtime forecasting tool.

1. Introduction

Coronal mass ejections (CMEs) are among the most significant drivers of space weather at the Earth. Considering the wide array of negative consequences, including damage to satellites, power grids and communication interruptions (Schwenn 2006; Pulkkinen 2007), a full understanding of CMEs is an increasingly important area of research. The potential damage of a severe space weather caused by a CME interacting with the magnetosphere of the Earth can be mitigated by accurately predicting when the CME will arrive at the Earth ahead of time so that some of the affected systems can be protected. Current predictive efforts include numerical, analytical and empirical models. The current standard of prediction is on the order of 6-10 hours of average error with a standard deviation on similar scales (Gopalswamy et al. 2013; Colaninno et al. 2013; Vršnak et al. 2014; Möstl et al. 2014).

The inputs for all models are based on observations of the corona from white light coronagraphs aboard different spacecraft. The Solar Heliospheric Observatory (SOHO) was launched in 1996 and observes from the Lagrangian L1 point, on the Sun-Earth line at a distance of about 90% of 1 AU from the Sun. SOHO contains the Large Angle Solar Corona Observatory (LASCO) with two still operational coronagraphs that observe the corona between 2 and 32 $R_\odot$ (Brueckner et al. 1995).

This was the only white light viewpoint in space until the launch of the Solar Terrestrial Relations Observatory (STEREO) in 2006. STEREO consisted of two nearly identical spacecraft traveling in approximate 1 AU orbits in opposite directions relative to the Earth. Among the instruments onboard the STEREO spacecraft was the Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI), which combines two coronagraphs and two heliospheric imagers that allow for continuous white light observations in the heliosphere from the Sun to beyond 1 AU (Howard et al. 2008).

The advantages of STEREO are the unprecedented distance coverage in the heliosphere that can be viewed, and the multiple simultaneous viewpoints allowing for a three dimensional reconstruction of transient objects, rather than just the two dimensional plane of sky projections offered with just a single view, a significant improvement for CME tracking (Mishra & Srivastava 2013). These multiple observations have been used throughout the STEREO era to try and determine the structure of the CME in the heliosphere and the manner in which CMEs propagate and improve predictive efforts (Wood & Howard 2009; Lugaz et al. 2009; Rouillard et al. 2009; Tappin & Howard 2009; Möstl et al. 2009; Liu et al. 2010; Poomvises et al. 2010; Byrne et al. 2013; Colaninno et al. 2013).

As observations have improved, so has the understanding of the complete CME structure as it propagates. The eruptive material itself, which we will refer to as the ejecta, is widely considered to be a flux rope in the heliosphere (Zhang et al. 2013). However, it is known that as the ejecta propagates, it will accumulate ambient solar wind plasma ahead of the ejecta body. This region of ambient plasma driven by the ejecta is called the sheath region. If the difference in velocity between the ejecta and the ambient solar wind is faster than the local Alfvén speed, the boundary of this sheath will become a fast mode shock wave traveling in front of the flux rope driver.

Sheath evolution is important to understand, as the shocks can also generate Solar Energetic Particles (SEPs), another growing space weather concern (Gopalswamy et al. 2014). Both the sheath and ejecta can contribute to geomagnetic disturbances, though the majority of the energy contributing to these storms will come from the ejecta (Zhang et al. 2008), so understanding the independent evolution of both ejecta and sheath is crucial to space weather forecasting.

Observationally, both the flux rope material and the associated brightenings in the sheath can be viewed in white light (Vourlidas et al. 2003; Ontiveros & Vourlidas 2009; Bemporad & Mancuso 2010; Maloney & Gallagher 2011) and for many events it is possible to use different image processing techniques to allow either the sheath front or the flux rope portion of the structure to be emphasized (Hess & Zhang 2014). In Figure 1, the different features observed using different running and base differencing techniques are shown. Doing this for a number of events, both the ejecta and sheath front can be tracked and their evolutions determined, independently of one another. Also, by knowing the location of each front in the corona the standoff-distance, or the distance between the ejecta and sheath front along the nose of the CME can be determined. This is important for both understanding the physics that govern the sheath front generation and evolution and can be used as a proxy for probing the low corona, as the standoff-distance has been shown to allow for the calculation of coronal magnetic fields (Kim et al. 2012).

In addition to imaging data, signatures of CMEs can also be observed in in-situ measurements. These signatures have long been studied at the L1 point from measurements from spacecraft such as the Advanced Composition Explorer (ACE) (Stone et al. 1998). The signature of the flux rope is generally thought to be a magnetic cloud, or a structure defined by strong, rotating magnetic field and low density and temperature (Zurbuchen & Richardson 2006). The sheath front is given by sharp increases in velocity, density and magnetic field (Jackson 1986). An example of these in-situ signatures is shown in Figure 2.

The goal of this paper is the demonstration of a method capable of both capturing the propagation characteristics and dynamics of a CME in the heliosphere accurately enough to predict an arrival at L1. Additionally, rather than just predicting one part of the CME structure, we will show that it is possible to predict both the ejecta and sheath front independently. The rest of this paper will be organized as follows: Section 2 will detail the events that have been studied as well as describing the techniques for measuring and modeling ejecta and sheath front height in the heliosphere. We will also discuss the physics on which the predictive model is built. In section 3 we will analyze the ability of our model to predict these events as well as comparing the results of our model to others currently in use. Section 4 will discuss some of the important implications of this work as well as the limitations of our model and key steps needed to work towards making our method applicable for realtime space weather forecasting and section 5 will present the basic conclusions that we can derive from our method at this time.

2. Data and Methods

2.1. Events

The seven events utilized in this study are CMEs that occurred between April 2010 and March 2013, when the STEREO spacecraft were well positioned to observe CMEs throughout the heliosphere. The chosen events were well observed upon eruption and also had in-situ signatures visible at ACE. The in-situ arrivals were divided into the flux-rope and compression/shock front arrival. Some of the events at ACE, especially those propagating further from the Sun-Earth line, do not meet the classic magnetic cloud signature of a flux rope. In the past there was debate as to whether this was an indication of a non-flux rope CME, but we now believe that this is just the flux rope flank passing through the observer (Vourlidas et al. 2013; Zhang et al. 2013). This means that while the sheath front is usually unambiguous in its arrival, the flux rope arrival time is more subjective, especially for the more complex events.

Basic information of each event is summarized in Table 1. These 7 events include a wide range of CMEs in terms of propagation direction, speed and in-situ signatures, allowing us to limit the effect of selection bias tuning the model towards specific types of events. These include CMEs with initial speeds ranging from 400 to 1500 km/s, upstream solar wind speeds between 289 and 513 km/s and deviations of the CME propagation direction as high as 28° in longitude and 26° in latitude from the Sun-Earth line. Our sample also includes both filament eruptions without flare associations and CMEs coming from active regions correlating to flares of differing magnitudes. All of the measurements in the sample begin between 4 and 8 $R_\odot$, corresponding to the first appearance of the CME in the COR2 FOV, while the final measurements are as close as 28.1 $R_\odot$ and as deep into the heliosphere as 76.6 $R_\odot$. The general trend of the distance from the Sun for the final usable measurement is that the faster CMEs tend to be stronger and therefore are observable further from the Sun.

2.2. Drag Model

As discussed in Hess & Zhang (2014), the true heights of both the ejecta and sheath front can be obtained through the use of a forward modeling technique that superimposes geometries onto images from different spacecraft taken at approximately the same time. For the ejecta, a Graduated Cylindrical Shell (GCS) Model is used (Thernisien et al. 2006, 2009) and for the sheath front, we use a prolate spheroid bubble (Kwon et al. 2014). Using these geometries for a number of time steps for each event, we get a series of height-time measurements for each individual CME. The parameters of the GCS model are: the propagation longitude and latitude, the angular separation of the legs, the aspect ratio which defines the thickness, the tilt of the ejecta relative the line of latitude and the height of the ejecta. The spheroid shock is defined by the same propagation direction, and tilt. On the other hand, the major and minor axis size and the nose height of this spheroid are distinct from the GCS parameters. Both geometries are shown with real event data in Figure 1.

Using the aerodynamic-drag based model (Vršnak et al. 2013; Hess & Zhang 2014), the measurements can be fit and average propagation characteristics can be determined. The drag model is based on a quadratic acceleration, with full equations of motion

$$a(t) = -\Gamma(v(t) - v_{sw})|v(t) - v_{sw}| \tag{1}$$

$$v(t) = \frac{v_0 - v_{sw}}{1 + \Gamma(v_0 - v_{sw})t} + v_{sw} \tag{2}$$

$$R(t) = \frac{1}{\Gamma} ln[1 + \Gamma(v_0 - v_{sw})t] + v_{sw}t + R_0 \tag{3}$$

where v_0 is the initial CME speed, v_{sw} is the upstream solar wind speed, R_0 is the initial CME height and Γ is the drag parameter that controls the rate at which the CME reaches the solar wind speed.

The model is therefore a function of four parameters. The initial height is easily obtained from the measurements and, as long as the measurements are accurate, the initial speed can also be determined reliably. The solar wind speed, which would have to be predicted in an operational mode, can be measured in-situ for the purposes of developing the predictive model. This leaves the Γ term as the only true unknown.

We can fit all the data points to one static set of parameters, determining an average Γ throughout the heliosphere. However, it is known that Γ is not one constant value as the CME propagates, instead the CME undergoes more drag closer to the sun and Γ decreases as the distances from the Sun increases.

The height-dependent Γ is (Vršnak et al. 2013)

$$\Gamma = \frac{c_d A \rho_{sw}}{M + M_v} \tag{4}$$

where A is the cross-sectional area and M is the mass of the CME, ρ_{sw} is the ambient solar density and M_v is the virtual mass, approximated as $\rho_{sw}V/2$ where V is the CME volume (Cargill 2004). By taking CME mass to be given by ρV and V to be approximated as rA where r is the CME minor radius, the drag parameter can now be approximated as

$$\Gamma = \frac{c_d}{r(\frac{\rho}{\rho_{sw}} + 1/2)} \tag{5}$$

To determine the CME density, as it propagates, we assume it evolves as $\rho = \rho_0 r_0^3/r^3$, where r is the CME minor radius. (Vršnak et al. 2014) assumes the exponent in the CME density evolution is a two rather than a three. We believe two is too low and that the CME must undergo a more significant drop in density than the solar wind, which is known to follow an inverse square law and can be approximated as $\rho_{sw} = \rho_{sw0} R_0^2/R^2$. This would be consistent with observations, which indicate a CME of higher density than the ambient in coronagraphs but lower than the ambient in-situ.

The more significant drop in density also follows from a physical argument. If the CME underwent no internal expansion and was simply pulled out by the solar wind, the CME plasma would be frozen in to the solar wind plasma in a radial direction and the CME would undergo the same $1/r^2$ evolution in density as the solar wind. However, the CME, because of its internal magnetic field, undergoes significant expansion. The GCS model assumes that this expansion is self-similar and constant. If this were the internal radial expansion of the flux rope, which can be thought of as a cylinder, were to maintain this self-similarity the thickness of the flux rope would have to expand at the same rate as the lateral expansion of the flux rope in solar wind and the CME density would evolve on the order of $1/r^3$. This may be true in the inner heliosphere, but as the CME travels further from the Sun this internal expansion, rather than continuing to be on the same order as the solar wind expansion will decrease as the internal magnetic pressure weakens, and is probably closer to 20-40% of the lateral expansion. This would lead to something approximating the $1/r^{2.32}$ shown in Liu et al. (2005). For this work we will use $1/r^3$, but it is possible that the model would be improved from a physical standpoint by assuming less internal CME expansion.

Assuming $\rho = \rho_0 r_0^3/r^3$ and $\rho_{sw} = \rho_{sw0} R_0^2/R^2$, and relating the CME minor radius to the height with the aspect ratio, the two density equations can be combined and the density ratio can now be expressed as

$$\frac{\rho}{\rho_{sw}} = \frac{\rho_0 \kappa R_0}{\rho_{sw0} R} \tag{6}$$

which then leaves

$$\Gamma = \frac{c_d}{\frac{\rho_0 \kappa R_0}{\rho_{sw0}} + \frac{\kappa R}{2}} \tag{7}$$

c_d is an unknown, but can reasonably be used to be a number between 1 and 1.5 (Poomvises 2010; Subramanian et al. 2012) and we use 1.35. κ is, for every event a number known parameter of the model and is around .4. This leaves the initial density ratio as the only term in the drag model that cannot be observed or easily approximated, and even this ratio may able to to be obtained from white light images.

In this work, to determine Γ, a series of fittings of the measurements are carried out to approximate Γ at each point. This provides a series of values of Γ throughout the heliosphere that can be fit with Equation 7 with just the density ratio unknown. Once the drag parameter profile is determined, we can use Equation 6 to approximate the evolution of the CME density relative to the ambient, with an admittedly large error.

Once Γ is known throughout the heliosphere, it can be used to create a more accurate representation of the drag model. In works using the drag model with a constant Γ, it is essentially a two point model between the initial CME height and the L1 point. In our method, we instead use the drag model iteratively. At time t=1, we calculate R and v as functions of ($\Gamma(R_0)$, v_0, v_{sw}, R_0). At t=2, R and v are now functions of (Γ(R(1)), v(1), v_{sw}, R(1)) and so on until the CME reaches L1, to calculate the complete propagation of the

CME.

2.3. Propagation Direction Geometric Correction

One correction that should be considered is the difference between the GCS height measurements along the CME nose and the in-situ data points, which are on the Sun-Earth line. If the CME nose is near the Sun-Earth line, this difference may be negligible, but the further a CME is propagating from the Earth, a more noticeable difference can be noticed. To test this, we can apply the geometry of the GCS model.

However, this will not be a perfect comparison, as the GCS model is an idealized geometry assuming a constant self-similar expansion. This is a simplification of a complex, highly variable process. There are a number of factors that will distort the CME structure, including the frozen in effect of the CME and the solar wind (Riley & Crooker 2004), different part of the CME encountering different ambient environments, and possible errors in the GCS geometry itself. These potential errors must be considered, and comparing the arrival of the CME with its propagation direction and the curvature of the GCS geometry to see if the model geometry can accurately reproduce the propagation in the Sun-Earth line.

The GCS model is defined by two conic legs, of height, h, connected by a central axis, varying with the angular distance from the leg (β). At each β angle, there is a circular cross-section that also varies in size. At the leg, this circle is defined by the base of the cone and gets progressively wider as the cross section approaches the CME nose ($\beta = 90°$) the radius of this cross section gets larger (Thernisien et al. 2009). Considering for now just the z=0 plane of the flux rope central axis, the difference in the distance from the origin point of the legs (the center of the Sun) to the nose and to another point, P, on the outermost edge of the shell can be determined based on the free parameters of the model by determining the central axis location and cross-section size at each β.

Following the geometric formulas in Thernisien (2011), this distance from the center to P can be determined by four parameters, the CME height, the aspect ratio (κ), the half-angular width (α) and the angular difference between CME nose and the point in question (θ). κ is a key parameter for the construction of the front, as it is the relation between width of the CME and the length of the leg height, and provides a consistent, self-similar ratio between CME size and distance from the origin at each β.

To get the distance of P from the center of the Sun, we can form a triangle of the central point of the GCS geometry, the distance of that point to the edge of the shell and the distance to that same shell point from the center of the Sun. The central point, denoted B, is given by triangle with the CME nose-axis and conic height, where the latter is the hypotenuse and the angle between them is the width of the model. B then, is given by $B = h/cos\alpha$. The third side of this triangle, the distance from the center of the conic base to the central point on the nose axis is ρ.

The distance from B to P is given by a combination of the distance to the central axis point and the radius at each β. Again from Thernisien (2011), the central axis distance, denoted X_0 is given by known parameters as $X_0 = \frac{\rho + B\kappa^2 sin\beta}{1-\kappa^2}$. The radius of the cross section centered at X_0 is given by $R^2 = X_0^2 + \frac{B^2\kappa^2-\rho^2}{1-\kappa^2}$. It should be noted that we are using R in order to keep the same labels used in (Thernisien 2011), but this R is distinct from the previous section. For a more detailed derivation of all these terms, refer to (Thernisien et al. 2006, 2009; Thernisien 2011), but essentially the different quantities come from the self-similarity relation of the aspect ratio and a coordinate transform along β.

With all of these parameters, we have the full triangle shown in Figure 3 defined by sides, B, $X_0 + R$, P with angles $\beta + \pi/2$ and θ where P is the distance from the origin to the point on the outer shell θ from the y-axis. P is the unknown, but with the known sides and angles, the law of cosines can be used to calculate P.

$$P = b^2 + (X_0 + R)^2 - 2b(X_0 + R)cos(\pi/2 + \beta) \tag{8}$$

where the proper β is determined by calculating the entire outer shell of the front and determining which point interacts with the known θ.

As predictions were first attempted, it was obvious from linking the remote sensing measurements to the actual in-situ arrivals that this curvature was being greatly over-estimated for CMEs where θ was more than 10°, causing predicted arrivals as much as 25% after the observed arrival. However, ignoring the geometric correction entirely showed a lesser, but still consistent under-estimation of these events by about 10%. Therefore, we use a weighted average of the GCS corrected height and the uncorrected height, based on these relative errors. The height of the CME where the height is measured at the nose, h_N, combined with the height along θ, h_G can be combined

$$h_f = .645h_N + .355h_G \tag{9}$$

to give the final geometric correction used, h_f.

As the propagation direction of the CME as well as the other necessary parameters are calculated in the fitting, the difference in height between the nose edge and the edge in the Sun-Earth line. To use this correction, the measurements taken with the GCS model are considered to be along the CME nose, and are therefore left uncorrected. The geometric correction is applied to the in-situ arrival point. For example, according to this formula when a hypothetical CME with a width of 30°, and an aspect ratio of 0.4 propagating $20 - 30°$ from the Sun-Earth line reaches the L1 point, the CME nose will be about 5-10%

further into the heliosphere. The correction factor for this same CME for $0^\circ \leq \theta \leq 35^\circ$ is plotted in Figure 4.

This correction seems relatively minor for CMEs propagating within 35° of the Sun-Earth line, which will cover the large majority of CMEs that go on to impact the Earth (Zhang et al. 2007). Still, for a CME with an average speed of 500 km/s, an extra $15R_\odot$ (7.5% of the distance from the Sun to L1) from the distance correction will cause a difference of about 6 hours. As we have stated, the standard error of prediction is currently on the order of 6 hours, so ignoring this term will make it impossible to improve upon the current state of prediction.

2.4. Treating the Sheath Front

The propagation model and geometry in the previous sections were based solely on the flux rope. Since the sheath front precedes the arrival of the ejecta, it is arguably more important to predict the sheath front arrival, though most of the geomagnetic activity will be caused by the ejecta.

Attempting the same methodology with the sheath front as the ejecta consistently failed to properly fit the Γ observations of the sheath front, which lead to the sheath front arrivals being predicted well behind the observed arrivals. This is likely due to the sheath undergoing a different evolution than the flux rope. The drag model does not properly address the evolution of sheath front, which is often a non-linear shock wave.

To predict the arrival of the sheath front, a new model was used that combined both the ejecta measurements and fittings and the sheath front measurements. The basis for this model is that during the propagation, the sheath front will be a combination of two factors, the flux rope driver that pushes the sheath as well the independent motion of the sheath front itself.

It has been observed before (Corona-Romero et al. 2013; Hess & Zhang 2014) that there is an independent element to the propagation of the sheath front that provides additional momentum. To determine this term, the standoff-distance between the ejecta and sheath front noses was calculated at each point for which each front was measured. For nearly all the events, a linear trend in the standoff-distance was observed. An example of one of these events is shown in Figure 5. A linear fit was performed on these measurements which could be extrapolated from the Sun to the Earth.

Using this standoff-distance evolution as well as the propagation profiles determined for the ejecta, it is now possible to combine the two into a model for the sheath front

$$R_{CF}(t) = R_{FR}(t) + At + B \tag{10}$$

$$V_{CF}(t) = V_{FR}(t) + A \tag{11}$$

where A and B are the terms of the linear standoff-distance fitting. It is very likely that the standoff-distance does not actually increase linearly and the independent motion of the sheath front would be expected to, at some point, cease continuously increasing at the same rate. Therefore this model can likely be improved and refined, but it does work for an initial approximation.

Another factor in the handling of the sheath front is the geometric correction factor The geometry of the sheath front is not the same as that of the flux rope, and trying to use the same correction may introduce more error into the process. However, the curvature of the sheath and the ejecta should be similar enough that using the same geometric correction is still an improvement over ignoring geometric effects altogether, though refinement is possible. For this study, the correction is applied to the sheath based on the θ deviation the same as it was for the ejecta.

3. Results

3.1. Predictions

With our method, both fronts were fit for each of the seven events. An example of the measurements with modeled height profiles, velocities and the derived Γ profile for the July 2012 CME are plotted in Figures 6, 7, 8. As is typical for the events, the Γ fitting deviates noticeably from the measurements. For many events these individual Γ values scatter quite a bit, so the purpose of this fitting is not an attempt to capture the physical evolution for the model, but rather trying to constrain the value of Γ for each individual event, and these fittings can reach this purpose.

Figures 9 and 11 show the predicted arrival for each event plotted against the actual arrival time. Furthermore, we want to test not just that the arrival time can be predicted, but also that the method is accurately capturing the physics governing ejecta evolution profile. Figures 10 and 12 show the velocities of the model when the respective front is at L1 and the observed in-situ velocities. The in-situ velocity used is the average velocity over the entire sheath region for the sheath front, and likewise, ejecta velocity is also the average over the duration of the flux rope passing through ACE. For the ejecta there is more uncertainty in the velocity, as the sheath velocities are often fairly constant, but because of the expansion of the ejecta the average velocity of the in-situ signature corresponds to the bulk motion ejecta velocity and ignores the expansion component.

The prediction numbers are shown in Table 2. The ejecta has an average error of 1.46 hours, with an RMS error of 0.76 hours. As can be seen in Figures 9 and 11, there is more error in the sheath front predictions. For the sheath front, the average error is

3.47 hours with an RMS of 1.52 hours. The most significant deviation in the sheath front numbers come from the slower events being predicted to arrive a few hours too early. This is likely due to the linear standoff-distance fitting. Comparison between standoff-distance measurements and in-situ sheath sizes indicate that while the increase is roughly linear closer to the Sun the increase tapers off before the sheath hits the Earth (Hess & Zhang 2014). For faster CMEs that spend less time in the heliosphere before reaching the Earth, this effect is minimal enough that a linear fitting is good enough to approximate the arrival. The results of the slower events indicate that the model for the sheath front could still be improved with some sort of damping term, perhaps also based on aerodynamic drag, for diminishing the rate of sheath growth.

3.2. Sensitivity of the Predictions

The results provided by feeding the GCS and shock measurements into a distance dependent model for Γ and then creating an iteratively calculated drag model were able to successfully model the arrival of both fronts for a number of events. The next test was the sensitivity of the model predictions to the input data. Specifically, we wish to determine if we can avoid the measurement phase and determine Γ empirically or, if the measurements are needed, what the smallest subset of the data necessary to get accurate predictions would be.

The first step, to determine the necessity of the measurements, requires looking at the degree to which the fitted model parameters vary. In general, if we see a slight variation, or a pattern to the variation in model inputs it would be possible to predict the input parameter and thus the Γ value. As shown in Equation 7, the only input for the ejecta is the density ratio of the ejecta to the ambient at the initial height of the model. The initial height of each event is different, but they should be close enough that a comparison can be made.

The range of $\rho_{ratio}(R_0)$ values vary from 3.24 to 32.17 with 6 of the 7 events in between 3.24 and 18.61, as seen in Table 2. The values are scattered seemingly at random with no noticeable pattern that could be used for predictive purposes.

We also want to test the sensitivity of the model to this parameter. To this end, we create a series of hypothetical CMEs with initial speed of 1000 km/s and encountering an ambient solar wind with a speed of 350 km/s. The difference between all the different runs is a varied initial density ratio, over the range of observed initial density values. We can then study the effect the derived initial density ratio has on arrival time, plotted in Figure 13. The change in arrival time is begins to diminish as $\rho_{ratio}(R_0)$ approaches the highest value in our sample, but between 3 and 18, the difference in arrival time is 16 hours. Obviously this is a crucial effect on the arrival time, and the initial density ratio and by extension Γ must be constrained to improve prediction.

To achieve an accuracy within 5 hours, and using the median $\rho_{ratio}(R_0)$ value of the event set, 10.31, the density ratio needs be accurate to be within a range of approximately 6-15, or about 50%. While predictions would likely be capable without any of the measurements, in order to achieve consistent results Γ must be constrained.

To determine $\rho_{ratio}(R_0)$ accurately enough to reach this level of prediction, the number of data points differs from event to event. For most of the events, there a great deal of scatter in the Γ measurements in the low corona to indicate that at least 5 or 6 observations must be taken to see a trend emerge and possibly be able to fit through the noise, indicating that a number of measurements be made to create a usable prediction. However, the main advantage to this type of model is the ease and speed with which it can be calculated, meaning predictions could be made almost immediately and then refined and improved as the CME continues to propagate and more measurements are possible, with every observation improving the accuracy.

For the sheath front, at least as much data is necessary. While the ejecta is fit with only one free parameter, the fact that there are three unknowns for the sheath front leaves more uncertainty in the fittings and therefore stable fittings will require, based on current data a minimum of 6 measurements, with the results being consistently improved as more data is fed into the model.

For a forecasting service, it is also desirable to provide the maximum amount lead time possible, so the ability to create a prediction as the CME is closer to the Sun is of the most value. All of these predictions are based on data where the CME is within $80R_{\odot}$, and for many events the fronts become so faint beyond $50R_{\odot}$ that measurements beyond these heights are not taken into account. Slower events tend to stabilize closer to the Sun, as there be more distinct observational points that can be fit. Faster events need to be measured further out into the Heliosphere, just to get a suitable number of measurements. For all events presented here, we would have a lead time of at least 36 hours based upon the last SECCHI image used for each event.

4. Discussion

4.1. CME Density in the Heliosphere

Apart from the predictive capabilities of this model, there are more basic scientific implications to CME research that can be derived, the most fundamental of these is the evolution of CME density as it propagates. As shown in Equation 7, the one free parameter of the ejecta fittings will scale to the ratio of the density of the ejecta to the ambient solar wind.

To verify that the scientific assumptions of the model are valid and providing meaningful results, we can derive the density ratio in the heliosphere from Equation 6 and compare these profiles with theoretical models of CME density. The profiles of the 7 events are shown in Figure 14.

As was explained in the Methods section, the assumptions that lead to these density ratios having a dependence of approximately $1/r$ is probably an overestimation, especially deeper in the heliosphere and the correct value is probably between 0 and 1, and possibly varies with distance.

The in-situ density ratios from the model are seen in Table 1, with the possible exception of the extremely low value of the September 2011 event, are all between .15 and .9, reasonable values given the average solar wind in-situ density of approximately $6cm^{-3}$ (Lepping et al. 2015) and the measured ejecta densities in-situ, which were all less than $4cm^{-3}$. Still, the majority of the events show density ratios under .5 at L1, which is lower than was expected. This despite the initial density ratios being largely higher than we expected. This is likely indicative that the density evolution we are using is too extreme and the CME density evolution does not trend with R^{-3}.

To get a better idea of how accurate these ratios may be, we also determined the average ejecta and ambient density and compared that ratio to the model, the values of which are also in Table 1. These values are all generally in similar ranges, though there does not seem to be much correlation between the modeled values. This could be indicating the assumptions of the model, very likely in c_d are leading to a distortion of the true density profiles, but again is likely a sign that either the density should be improved or may just be too complex of an evolution to simply model.

Interesting to note is that the biggest outlier of the densities is that of the April 2010 CME. In the modeled data, this was the only event in the sample in which a high speed stream of solar wind was observed in-situ ahead of the ICME, which has long been known to correspond to a lower solar wind density (Hundhausen 1972). This would likely give rise to this large ratio between the CME density and the ambient and show why this CME propagates more linearly than the other events. However, in the in-situ data, the solar wind density is indeed below average, but the ratio is not noticeably high, and is even well below average. With the lower solar wind density, it is possible the CME underwent more significant expansion by the time the CME reaches L1 causing a lower CME internal density.

It may be possible to use the white light coronagraph data to calculate the density inside the the flux rope and use a solar wind model to determine both the initial density ratio and the ratio between the two as the CME propagates observationally, as well as gaining a better understanding of how the CME internal density evolves. This goes beyond the current scope of this paper, but may be attempted in the future to try and further constrain the model.

4.2. Comparison with Other Models

In general, our ability to predict the two distinct fronts within 3 hours with a variance of less than 2 hours is a significant improvement over other models. Colaninno et al. (2013), also using GCS fittings and considering a number of model fits including a similar drag model, showed an average time error of 8.1 hours with a 6.3 hour variance. Gopalswamy et al. (2013) used the Empirical Shock Arrival (ESA) Model to accurately predict CME arrival within 7.3 hours with a 3.2 hour deviation. Möstl et al. (2014) used time-elongation plots, or j-maps (Sheeley et al. 1999), with an empirical prediction to predict CMEs to an average absolute error 6.1 hours with a deviation of 5.1 hours. Lastly, using the drag based model, (Vršnak et al. 2014) predicted arrival times with an average absolute error of 14.8 hours, with deviations on a similar order. Clearly our method improves upon all of these results, though it should be noted that given the sample size of events and quality of data used, all of the results are not an ideal comparison. It still shows the effectiveness of our method as a proof of concept.

Using our method, as well as the published formulas of the ESA and drag model, we can compare our results to those, acknowledging that every forecasting technique will be most effective in the hands of those who developed the technique using their own tracking methods, and simply plugging our initial velocities into the formulas of the other models will probably do worse than those modelers themselves.

The numbers for each event are presented in Table 3 and can be viewed graphically in Figure 15. As can be seen, our method significantly outperforms the others, with perhaps the most positive thing about it our prediction being the lack of an extreme outlier event where the prediction is extremely inaccurate. If, moving forward, these results can be repeated our method will be a massive step forward for space weather forecasting.

Of course the model we have is not currently operational and uses data not available in realtime, making a comparison to those works that attempt to predict arrivals in real time is not really fair. In the next section we will detail the hurdles that will have to be addressed to make real time predictions. It does however, help demonstrate what the ceiling of effective predictions could be with an ideal set of observational data available in realtime.

4.3. Creating an Operational Model

As we attempt to transfer this model into an operational mode, there will be a few difficulties that will have to be overcome. The most obvious of these is the difficulty in getting accurate heights of the fronts in real-time. All observations have currently been performed with science quality data. It will be more difficult to accurately get these measurements in real time from the lower quality data available.

The observational part of this technique is also made more difficult by the lack of a permanent observer somewhere other than L1. From 2006-2014 when STEREO was consistently providing multiple viewpoints, especially during the 2010- 2013 period of these events when the STEREO spacecraft was more optimally positioned for observing Earth directed CMEs, it was much easier to create an accurate reconstruction. For now, we may be restricted to a cone model using only LASCO data (Na et al. 2013) and it is unknown if this will be able to provide enough accuracy for our model. This concern is one facing all space weather forecasting operations, and the accuracy of our method using stereoscopic techinques is one more piece of evidence indicating why an L5 mission with a coronagraph and heliospheric imager would be a great benefit to the space weather community.

Another issue will be our ability to predict the solar wind. For this method, we have used ACE measurements ahead of the transients to get an average value for the solar wind speed. In real-time, this would obviously be impossible and we would have to estimate the solar wind speed. The widely-used Wang Sheeley Arge (WSA) model (Arge & Pizzo 2000) can be used to predict solar wind speed in the heliosphere, and is being continually refined to provide more accurate results (Hickmann et al. 2015). Using a model, provided it is accurate, would also give us the added benefit of, rather than assuming a constant solar wind speed in front of the CME, having a height-dependent solar wind speed that could account for the CME interacting with varied solar wind regimes.

The last issue is an acknowledgement that we are focusing only on simple events propagating freely in the heliosphere. For multiple CME events such as the well studied August 2010 events (Webb et al. 2013), this method as currently constructed will not be able to properly capture the propagation. However, before these complex events can be understood, we must make sure our ability to reconstruct the most basic CMEs is complete.

We also plan to perform further comparison with numerical models, such as ENLIL (Odstrčil & Pizzo 1999) for purposes of cross-validation. Not only can we test the ability of our model to capture the propagation of observed CMEs, we can also test the ability to capture the propagation of the modeled CMEs. which may also serve to improve the inputs used in those models. A collaborative effort between observers and modelers can greatly benefit those in each field.

Despite these potential operational problems, what we have shown here is useful as a proof of concept for the method, and could be the first step in providing valuable insight to the the heliophysics community. If our model is indeed able to consistently provide accurate predictions for simple events, it means that even if our exact method can't be put into operation, it is at least accurately capturing the physics of a CME in the heliosphere making it a valuable addition to the space weather community.

5. Conclusions

For seven CMEs observed by STEREO and ACE, we have developed a method based upon a height-dependent Γ drag model that accurately predicts the arrival of the presumed flux rope in-situ. We also combine measurements of the sheath front and ejecta to calculate the standoff-distance in the heliosphere, which we fit linearly. This provides an accurate arrival prediction within 6 hours for all events for the compression front, and does better for faster events that arrive at L1 within 70 hours. All the predictions for the compression front on the slowest events are consistently early, so it is possible we can use a more accurate function that will slow the sheath front down properly. Still, this error is as good as or better than all currently used models, with the added distinction of differentiating the two different fronts.

Not only does this distinction of the ejecta and sheath front improve the forecasting capabilities of the model, it also allows us to infer more of the physics governing CME evolution. In addition to the ability to infer the ratio of CME density to the ambient, the more we study the standoff distance evolution it will tell us more about the exact mechanism through which the sheath wave is generated and propagated. From the geometric correction we apply, we can also say that, assuming the GCS model is indeed an accurate representation of the ejecta, that the ejecta loses about 65% of its curvature by the time it reaches L1.

We not only compare arrival times, but also relative densities and arrival velocities to try and prove that, in addition to being able to accurately predict the arrivals our model also accurately captures the propagation of the CME in the heliosphere. This is especially true for the CME ejecta or flux rope, for which we have a physical basis for the model we use to interpret our measurements. For the sheath front, the model is more ad hoc and may have more room for improvement and physical refinement.

This work is supported by NSF AGS-1156120 and NSF AGS-1249270 We acknowledge the use of data from the STEREO mission, the SOHO mission and also the use of ACE data.

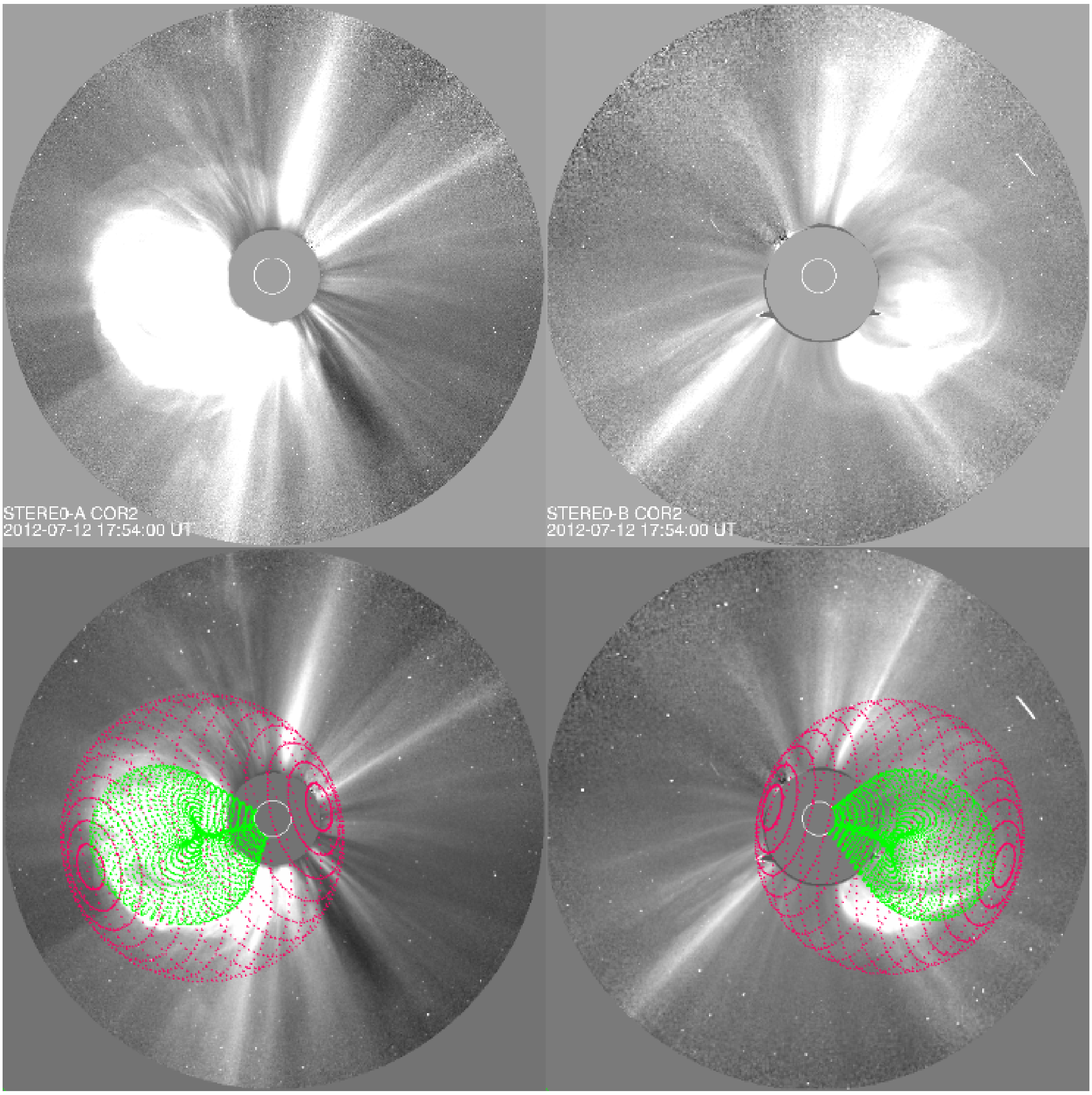

Fig. 1.— Model fitting of CME ejecta and shock front. Images at 17:54 UT on July 12, 2012 from STEREO A COR2 (Left) and STEREO B COR2 (Right) are shown along without (top) and with (bottom) the raytrace mesh. The green mesh shows the GCS fitting to the CME ejecta, while the red mesh shows the spheroid fitting to the CME shock front.

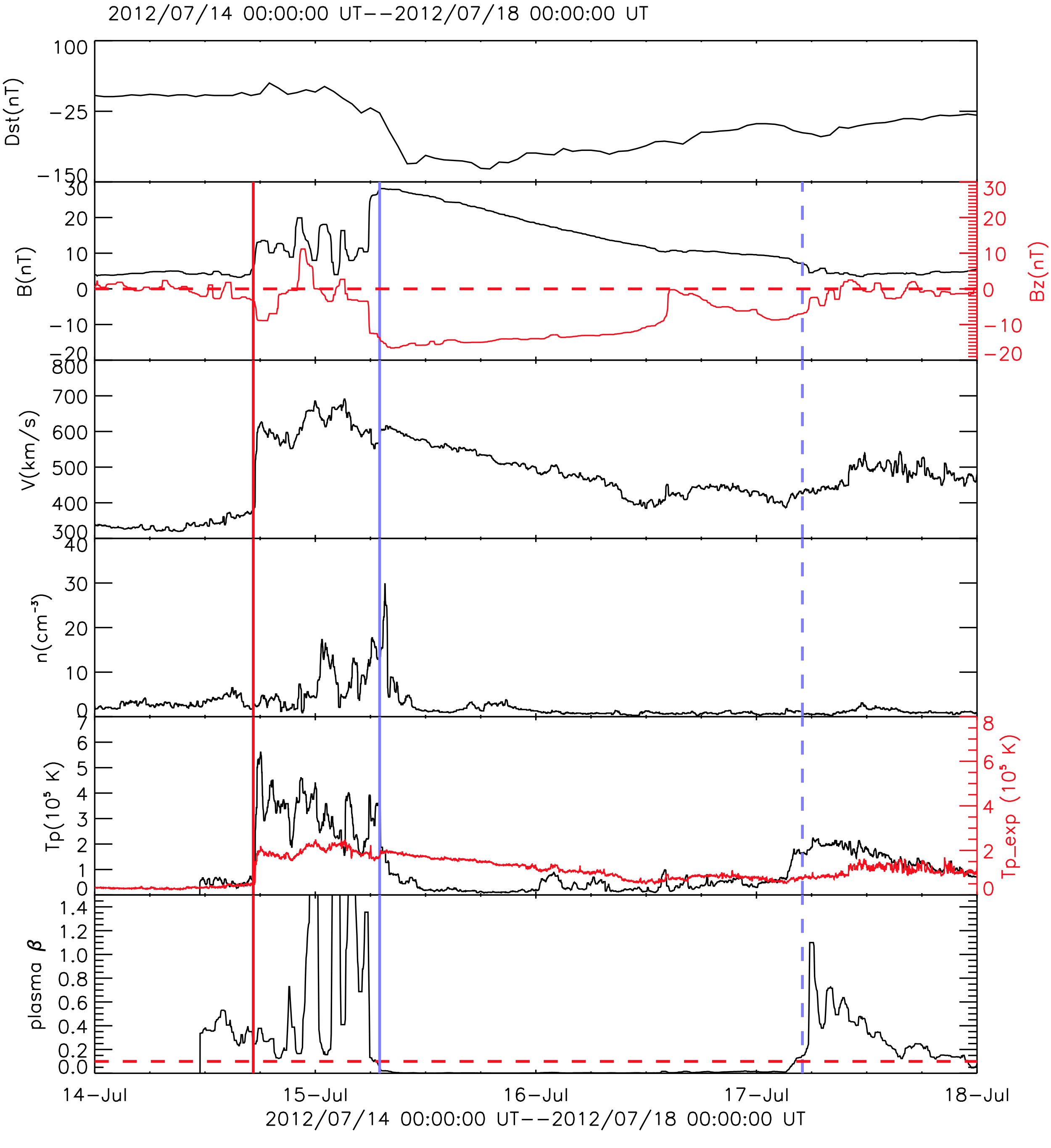

Fig. 2.— Solar wind data from ACE for the July 14 ICME. The sheath onset is denoted with the red line, the ejecta with the solid blue line and the passing of the ejecta with the dotted blue line. Plots from top to bottom show the Dst index, total magnetic field and the z component of the magnetic field (red), total velocity, density, temperature and the expected temperature (red) based on the velocity, and the plasma β. The sheath arrival was determined by the sudden increases in B, V, and T, while the ejecta boundaries were constrained mostly by the period of low plasma β.

First Measurement^a	SH Arrival^b	EJ Arrival^b	Direction^c	Width^c	κ^c	V_0^d	V_{sw}^e	R_0^f	R_f^f	AR^g	Flare^h	Flare Peak^h
04/03/2010 10:24	04/05 08:00	04/05 11:30	E06S26	25.0	.379	854.7	512.4	5.5	62.8	11059	B7	04/03 09:54
05/24/2010 14:54	05/28 02:00	05/28 07:00	E28N03	22.6	.468	605.7	362.3	4.6	45.0	-	-	-
09/14/2011 00:24	09/17 02:00	09/17 19:00	W20S16	24.0	.489	519.5	396.9	5.3	28.1	11289	-	-
07/12/2012 16:54	07/14 17:00	07/15 07:15	W00S09	25.4	.370	1492.0	353.7	4.3	76.6	11520	X1	07/12 15:36
09/28/2012 00:24	09/30 23:00	10/01 06:00	E28N17	33.3	.327	1230.5	310.4	6.3	74.1	11577	C3	09/28 00:00
10/27/2012 17:24	10/31 15:00	11/01 00:00	E12N12	31.9	.407	400.1	289.8	6.2	49.0	-	-	-
03/15/2013 07:24	03/17 15:30	03/18 00:00	W24S07	27.1	.389	1220.2	429.3	7.4	37.0	11692	M1	03/15 05:46

Table 1: a- The time step of the first SECCHI and LASCO images used for GCS fitting. Given the time offsets between the different satellites, the time given refers to the SECCHI observations

b- Sheath and Ejecta arrivals at ACE as manually determined

c- Parameters given by the GCS/Ellipsoid model

d- Initial Velocity (km/s) is obtained by performing a fit of the data using the drag based model over all observations

e- Solar wind speed (km/s) is determined by taking an average value of the ACE data preceding the arrival of the sheath signature

f- The measured ejecta height ($R_\odot$) of the first and last point used for fitting

g- Associated Active Region given by tracking CME back to the surface using EUV data. Not all CMEs can be linked with an active region

h- Flare Strength and Peak determined by comparing the EUV observation to X-Ray flux from GOES

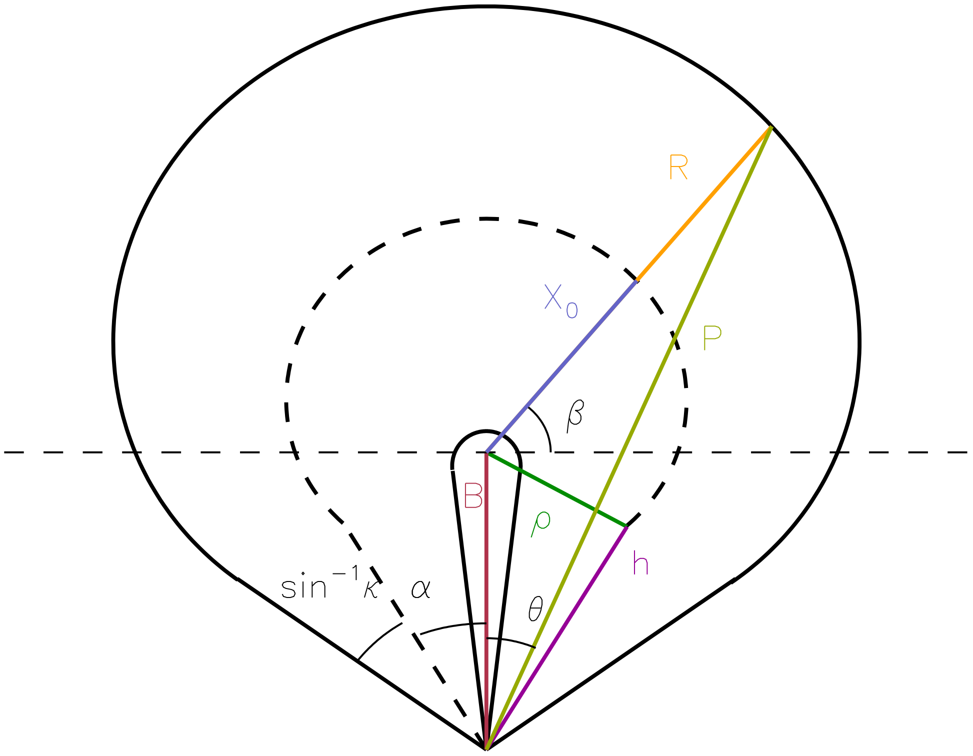

Fig. 3.— The calculated quantities of the GCS model used to get the height of the CME along the propagation direction of the eruption. The dashed curve represents the ejecta central axis.

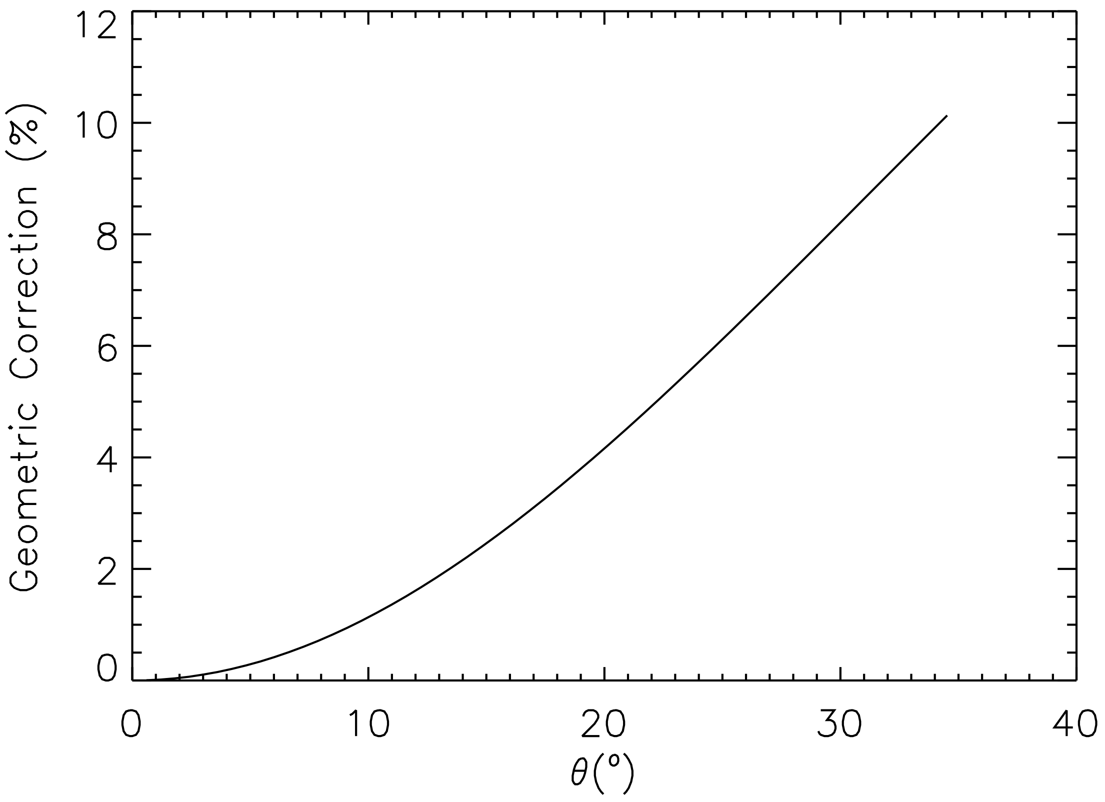

Fig. 4.— The effect of the geometric correction as a function of the θ angle between the CME nose and the Sun- Earth line.

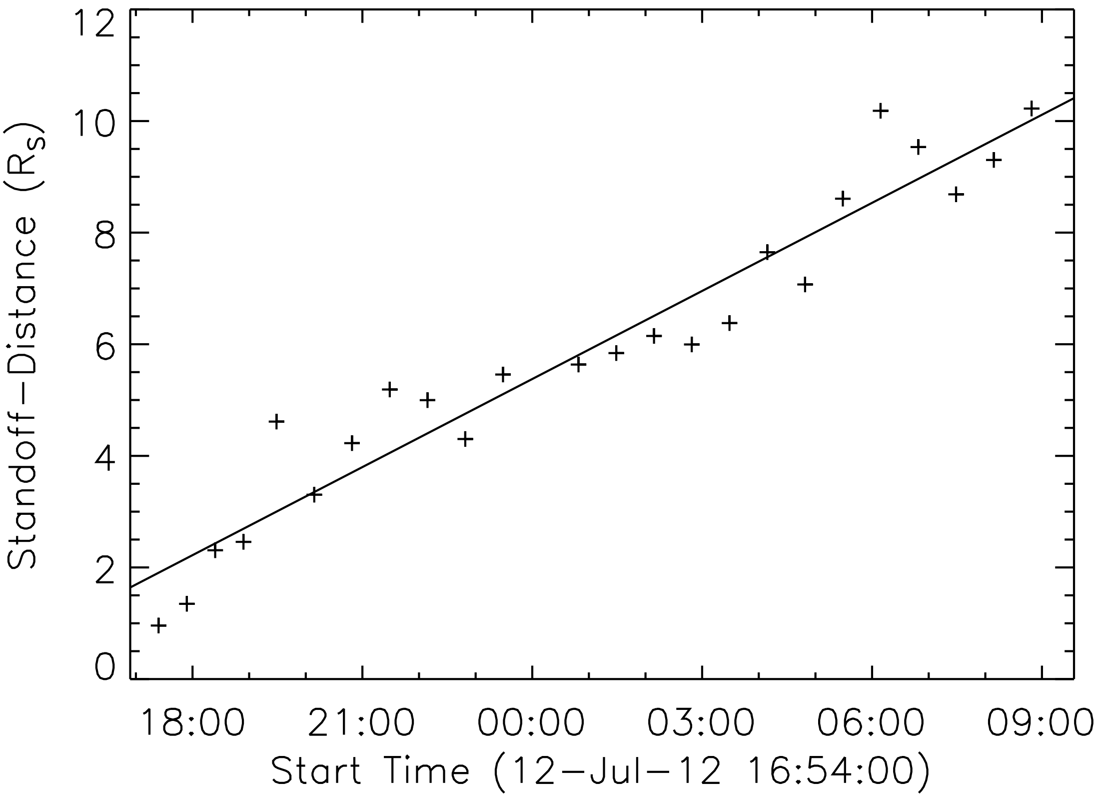

Fig. 5.— The standoff-distance of the July 2012 CME as determined by subtracting the ejecta measurements from the shock measurements. The solid line is the linear fit to the data.

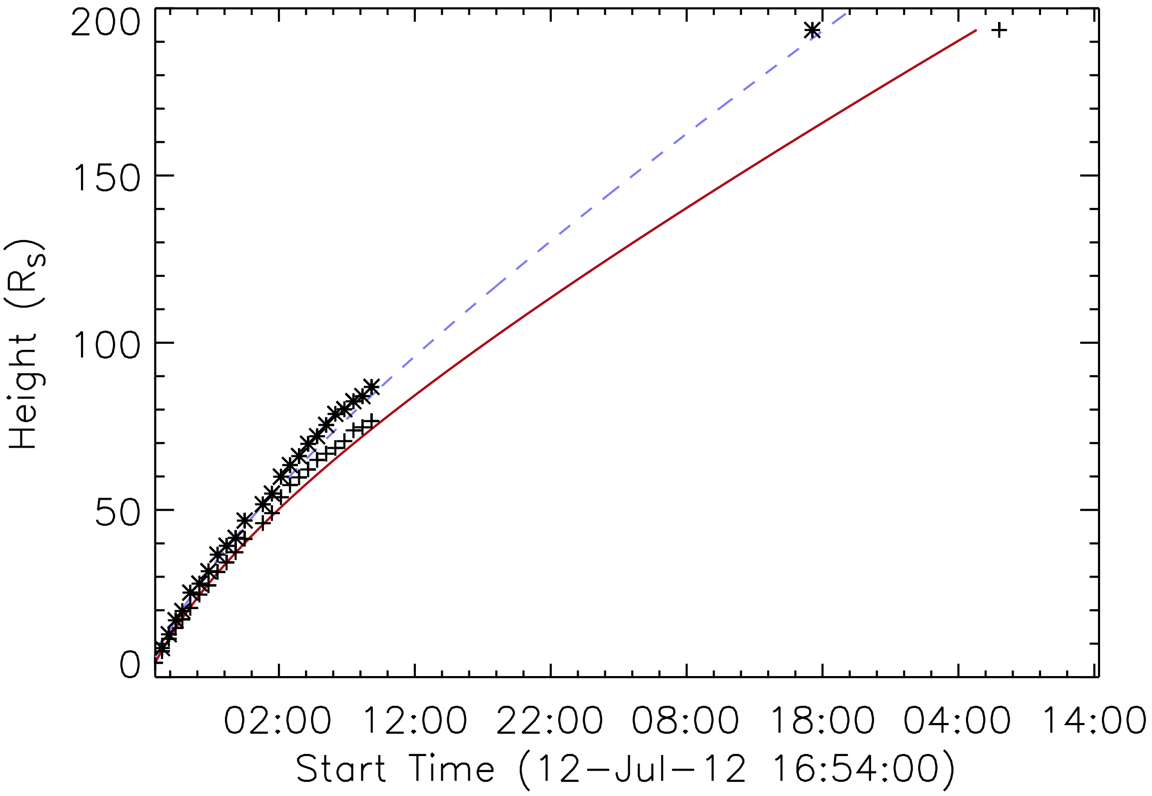

Fig. 6.— The measurements for the ejecta (crosses) and sheath front (stars) for the July 2012 CME. The derived profiles for each front (red solid line, blue dashed line respectively) are also plotted. The maximum height in each plot is the arrival point of the front at ACE.

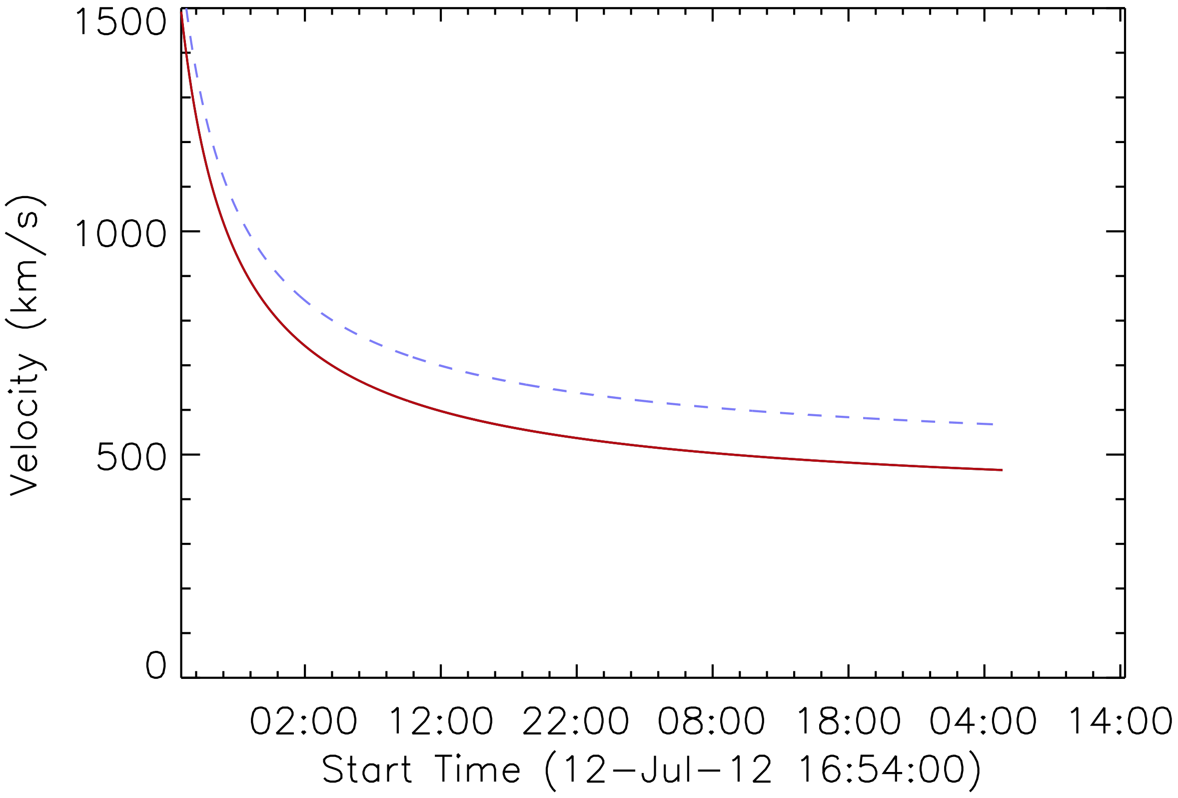

Fig. 7.— The velocity of the two fronts for the July 2012 CME. The sheath front is given by the dashed blue line, the ejecta by the solid blue line.

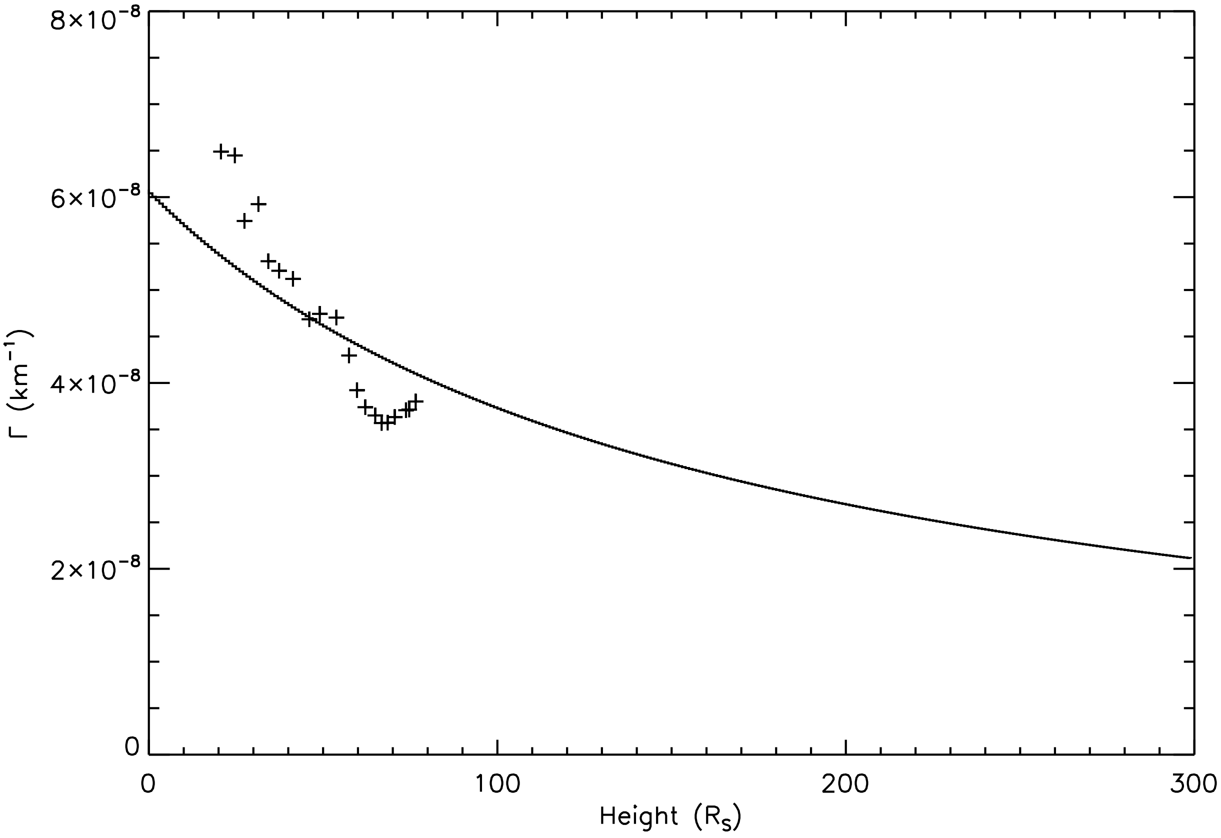

Fig. 8.— The derived Γ profile for the July 2012 CME is shown with the solid line, as well as the ejecta Γ values determined at each point by fitting the measurements up to each individual point. It should be noted that the derived Γ profile is not being fit to the measured values, but rather the measured values are used to constrain the theoretical Γ profile.

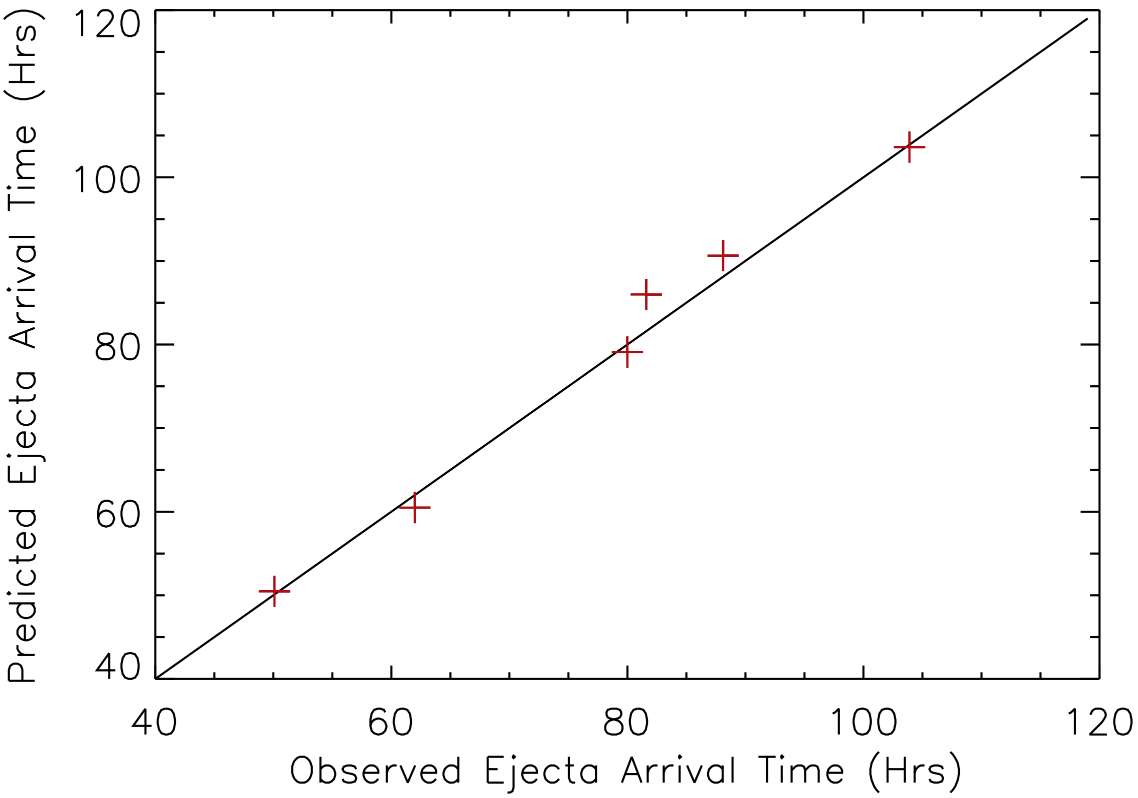

Fig. 9.— The predicted vs. observed arrival for the ejecta for each event. The solid line represents perfect accuracy in prediction.

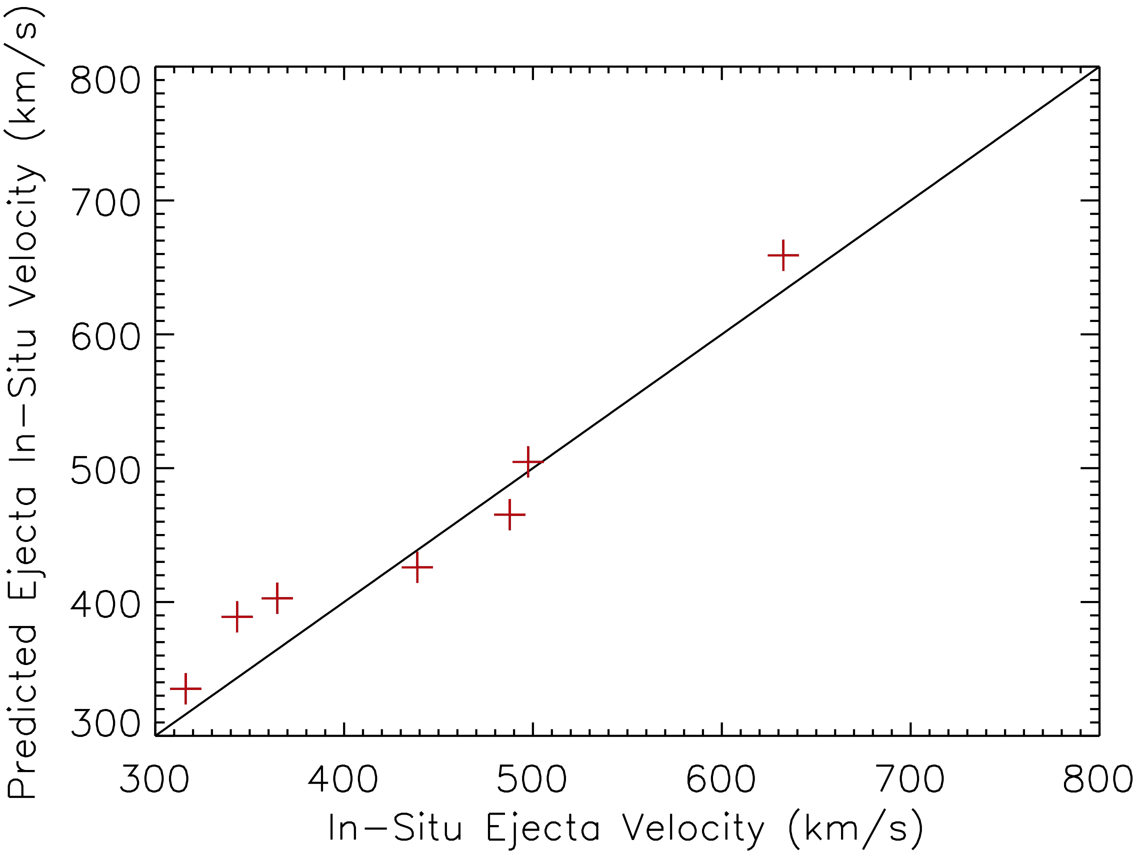

Fig. 10.— The predicted vs. observed velocities at L1 for the ejecta for each event. The solid line represents perfect accuracy in prediction.

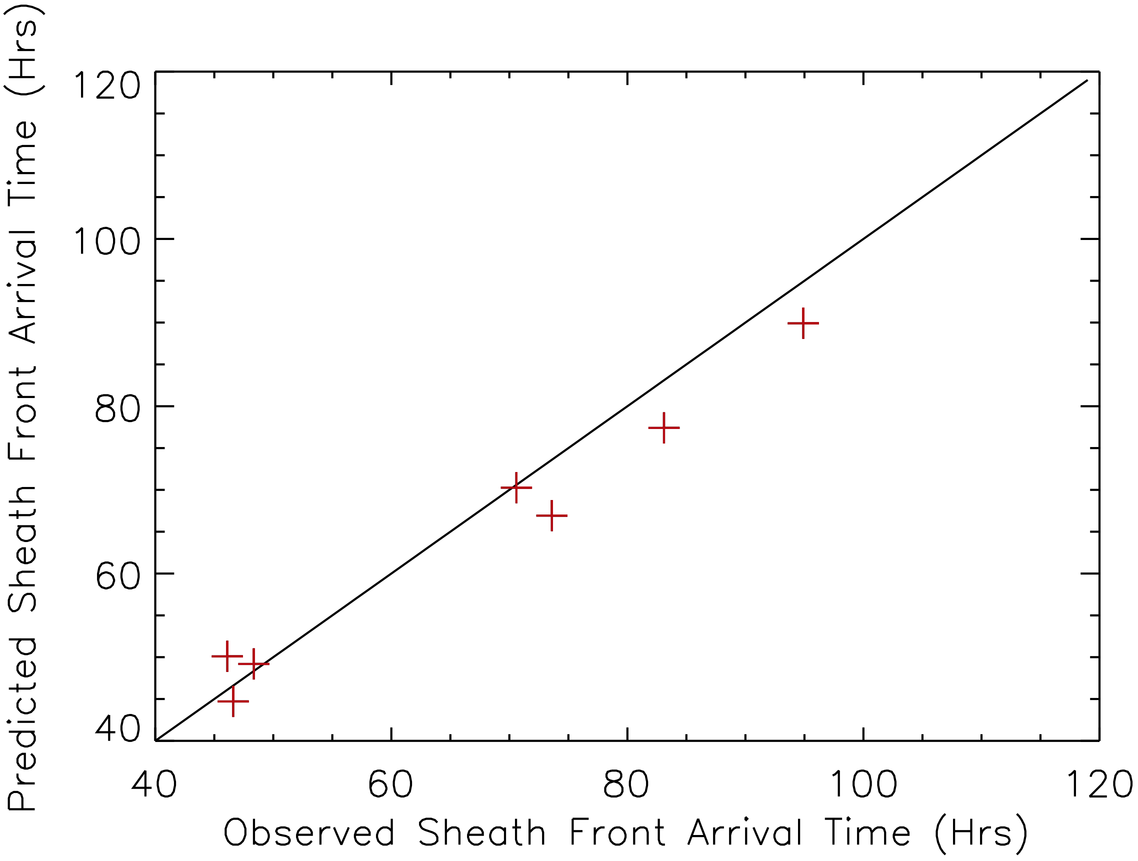

Fig. 11.— The predicted vs. observed arrival for the sheath front for each event. The solid line represents perfect accuracy in prediction.

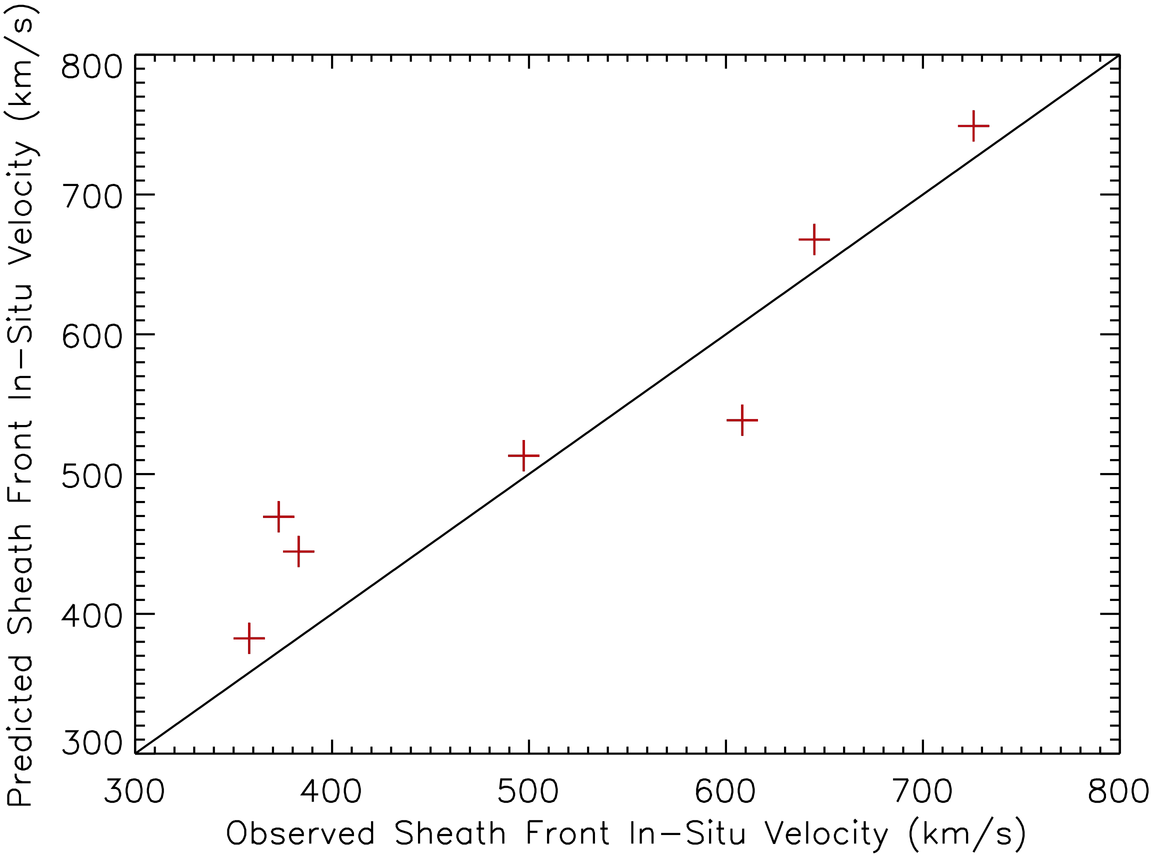

Fig. 12.— The predicted vs. observed velocities at L1 for the sheath front for each event. The solid line represents perfect accuracy in prediction.

ICME Date[a]	$\Delta T_{SF}{}^b$	$\Delta T_{EJ}{}^b$	$\Delta V_{SF}{}^c$	$\Delta V_{EJ}{}^c$	$\rho_{ratio}(R(0))^d$	$\rho_{ratio}(L1)^d$	$\rho_{ratio}(ACE)^e$
04/05/2010	1.89	0.38	23.3	26.4	32.17	0.91	0.41
05/24/2010	5.69	2.52	96.3	38.1	6.70	0.15	1.21
09/14/2011	6.68	4.39	15.8	13.0	3.24	0.09	0.71
07/12/2012	0.84	1.51	24.8	22.4	18.61	0.41	0.61
09/28/2012	0.34	0.9	61.6	45.6	10.31	0.31	0.97
10/27/2012	4.99	0.28	24.5	19.0	14.78	0.47	0.67
03/15/2013	3.91	0.26	22.9	7.2	5.98	0.21	0.38
Average	3.47	1.46	38.5	24.5	13.11	0.36	0.80
RMS	1.58	0.76	17.9	12.9	-	-	-

Table 2: a- The date of the ICME arrival at ACE

b- The absolute value of the difference in hours between the predicted and observed arrival of the sheath (SF) and Ejecta(EJ)

c- The difference in velocity in km/s between the speed of each feature as predicted by the model and as compared to the average speed observed for each feature in-situ

d- The derived density ratio from the model at the initial height of observation and at the point where the ejecta reaches L1 e- The ratio of the densities of the ejecta and solar wind, as determined from the average values of each from ACE.

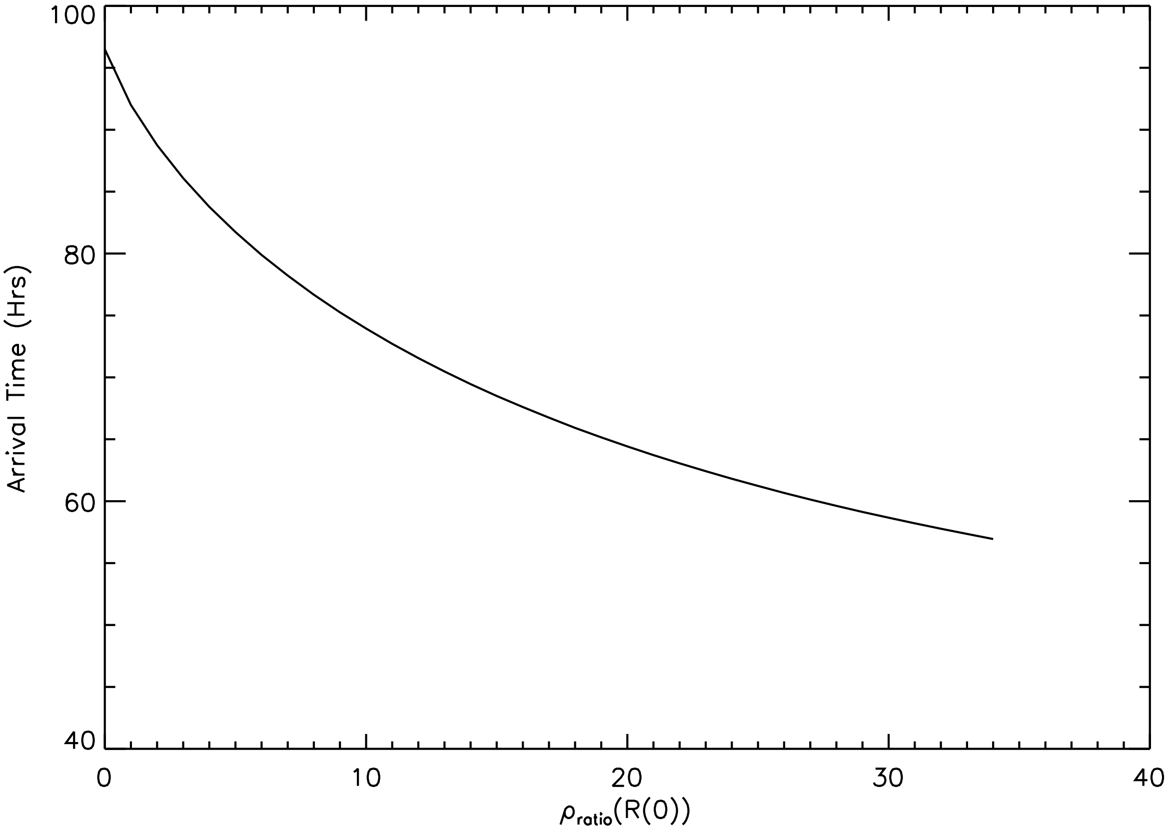

Fig. 13.— The difference in the arrival of the ejecta at L1 by varying the initial density ratio ranges among the set of modeled values obtained from fits to measurement.

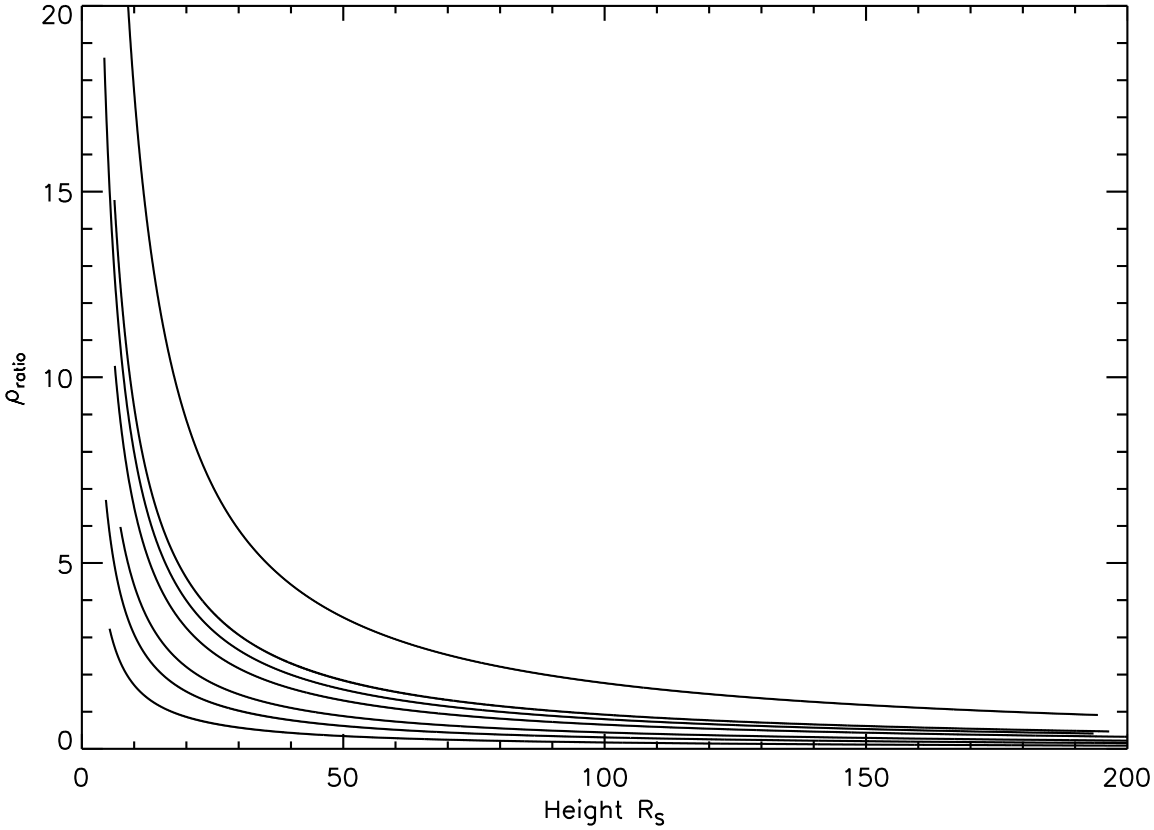

Fig. 14.— The profiles for the ratio of ρ_{CME} to ρ_{sw} for each of the seven events, as given by the Γ fittings and Equation 7.

ICME Date	Constrained Drag Model	ESA	Static DBM
04/05/2010	-1.89	-11.6	-14.0
05/24/2010	-5.69	7.91	10.6
09/14/2011	-6.68	-11.5	-6.00
07/12/2012	0.84	17.4	2.88
09/28/2012	-0.34	32.9	22.5
10/27/2012	-4.99	-3.70	2.11
03/15/2013	3.91	8.00	-1.45
Average	3.47	13.27	8.5
RMS	1.58	6.04	4.20

Table 3: A comparison of the error in hours between our method and the ESA and DBM models for each event. The average values for each model is the average of the absolute value of the error for each event.

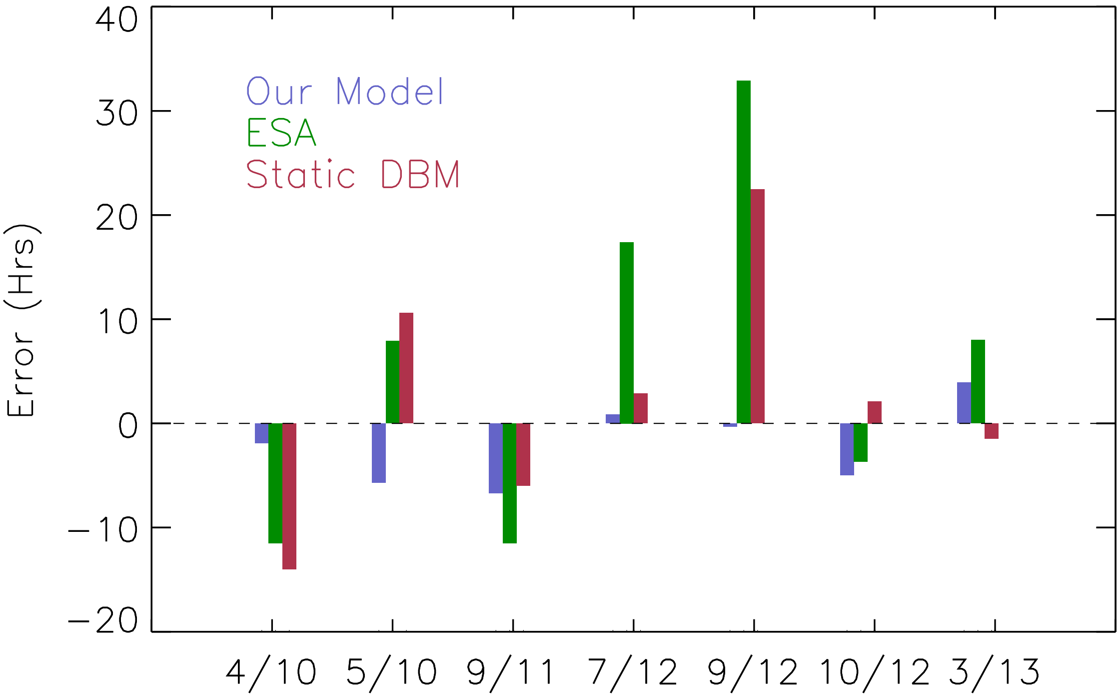

Fig. 15.— A comparison of the arrival times for the sheath front with our method (blue), the ESA method (yellow) and DBM model (red). A negative error time corresponds to the predicted arrival being after the observed arrival.

REFERENCES

Arge, C. N., & Pizzo, V. J. 2000, J. Geophys. Res., 105, 10465

Bemporad, A., & Mancuso, S. 2010, ApJ, 720, 130

Brueckner, G. E., Howard, R. A., Koomen, M. J., et al. 1995, Sol. Phys., 162, 357

Byrne, J. P., Long, D. M., Gallagher, P. T., et al. 2013, A&A, 557, A96

Cargill, P. J. 2004, Sol. Phys., 221, 135

Colaninno, R. C., Vourlidas, A., & Wu, C. C. 2013, Journal of Geophysical Research (Space Physics), 118, 6866

Corona-Romero, P., Gonzalez-Esparza, J. A., & Aguilar-Rodriguez, E. 2013, Sol. Phys., 285, 391

Gopalswamy, N., Mäkelä, P., Xie, H., & Yashiro, S. 2013, Space Weather, 11, 661

Gopalswamy, N., Xie, H., Akiyama, S., Mäkelä, P. A., & Yashiro, S. 2014, Earth, Planets, and Space, 66, 104

Hess, P., & Zhang, J. 2014, ApJ, 792, 49

Hickmann, K. S., Godinez, H. C., Henney, C. J., & Arge, C. N. 2015, Sol. Phys., arXiv:1410.6185

Howard, R. A., Moses, J. D., Vourlidas, A., et al. 2008, Space Sci. Rev., 136, 67

Hundhausen, A. J. 1972, Coronal Expansion and Solar Wind, Vol. 238 (New York: Springer-Verlag), 101

Jackson, B. V. 1986, Advances in Space Research, 6, 307

Kim, R.-S., Gopalswamy, N., Moon, Y.-J., Cho, K.-S., & Yashiro, S. 2012, ApJ, 746, 118

Kwon, R.-Y., Zhang, J., & Olmedo, O. 2014, ApJ, 794, 148

Lepping, R. P., Wu, C.-C., & Berdichevsky, D. B. 2015, Sol. Phys., 290, 553

Liu, Y., Davies, J. A., Luhmann, J. G., et al. 2010, ApJ, 710, L82

Liu, Y., Richardson, J. D., & Belcher, J. W. 2005, Planet. Space Sci., 53, 3

Lugaz, N., Vourlidas, A., & Roussev, I. I. 2009, Annales Geophysicae, 27, 3479

Maloney, S. A., & Gallagher, P. T. 2011, ApJ, 736, L5

Mishra, W., & Srivastava, N. 2013, ApJ, 772, 70

Möstl, C., Farrugia, C. J., Temmer, M., et al. 2009, ApJ, 705, L180

Möstl, C., Amla, K., Hall, J. R., et al. 2014, ArXiv e-prints, arXiv:1404.3579

Na, H., Moon, Y.-J., Jang, S., Lee, K.-S., & Kim, H.-Y. 2013, Sol. Phys., 288, 313

Odstrčil, D., & Pizzo, V. J. 1999, J. Geophys. Res., 104, 483

Ontiveros, V., & Vourlidas, A. 2009, ApJ, 693, 267

Poomvises, W. 2010, PhD thesis, George Mason University

Poomvises, W., Zhang, J., & Olmedo, O. 2010, ApJ, 717, L159

Pulkkinen, T. 2007, Living Reviews in Solar Physics, 4, 1

Riley, P., & Crooker, N. U. 2004, ApJ, 600, 1035

Rouillard, A. P., Davies, J. A., Forsyth, R. J., et al. 2009, Journal of Geophysical Research (Space Physics), 114, 7106

Schwenn, R. 2006, Living Reviews in Solar Physics, 3, 2

Sheeley, N. R., Walters, J. H., Wang, Y.-M., & Howard, R. A. 1999, J. Geophys. Res., 104, 24739

Stone, R. G., Frandsen, A. M., Mewaldt, R. A., et al. 1998, Space Sci. Rev., 86, 1

Subramanian, P., Lara, A., & Borgazzi, A. 2012, Geophys. Res. Lett., 39, 19107

Tappin, S. J., & Howard, T. A. 2009, Space Sci. Rev., 147, 55

Thernisien, A. 2011, ApJS, 194, 33

Thernisien, A., Vourlidas, A., & Howard, R. A. 2009, Sol. Phys., 256, 111

Thernisien, A. F. R., Howard, R. A., & Vourlidas, A. 2006, ApJ, 652, 763

Vourlidas, A., Lynch, B. J., Howard, R. A., & Li, Y. 2013, Sol. Phys., 284, 179

Vourlidas, A., Wu, S. T., Wang, A. H., Subramanian, P., & Howard, R. A. 2003, ApJ, 598, 1392

Vršnak, B., Žic, T., Vrbanec, D., et al. 2013, Sol. Phys., 285, 295

Vršnak, B., Temmer, M., Žic, T., et al. 2014, ApJS, 213, 21

Webb, D. F., Möstl, C., Jackson, B. V., et al. 2013, Sol. Phys., 285, 317

Wood, B. E., & Howard, R. A. 2009, ApJ, 702, 901

Zhang, J., Hess, P., & Poomvises, W. 2013, Sol. Phys., 284, 89

Zhang, J., Poomvises, W., & Richardson, I. G. 2008, Geophys. Res. Lett., 35, 2109

Zhang, J., Richardson, I. G., Webb, D. F., et al. 2007, Journal of Geophysical Research (Space Physics), 112, 10102

Zurbuchen, T. H., & Richardson, I. G. 2006, Space Science Reviews, 123, 31

This manuscript was prepared with the AAS LaTeX macros v5.2.