# Prediction of shock arrival times from CME and flare data

Marlon Núñez[1*], Teresa Nieves-Chinchilla[2], and Antti Pulkkinen[2]

[1]Universidad de Málaga, Málaga, Spain

[2] Heliophysics Science Division, GSFC, NASA, Greenbelt, MD USA

* Corresponding author: mnunez@uma.es

## Abstract

This paper presents the Shock ARrival Model (SARM) for predicting shock arrival times for distances from 0.72 AU to 8.7 AU by using coronal mass ejections (CME) and flare data. SARM is an aerodynamic drag model described by a differential equation that has been calibrated with a dataset of 120 shocks observed from 1997 to 2010 by minimizing the mean absolute error (MAE), normalized to 1 AU. SARM should be used with CME data (radial, earthward or plane-of-sky speeds), and flare data (peak flux, duration, and location). In the case of 1 AU, the MAE and the median of absolute errors were 7.0 h and 5.0 h respectively, using the available CME/flare data. The best results for 1 AU (an MAE of 5.8 h) were obtained using both CME data, either radial or cone-model-estimated speeds, and flare data. For the prediction of shock arrivals at distances from 0.72 AU to 8.7 AU, the normalized MAE and the median were 7.1 h and 5.1 h respectively, using the available CME/flare data. SARM was also calibrated to be used with CME data alone or flare data alone, obtaining normalized MAE errors of 8.9 h and 8.6 h respectively for all shock events. The model verification was carried out with an additional dataset of 20 shocks observed from 2010 to 2012 with radial CME speeds to compare SARM with the empirical ESA model [*Gopalswamy et al.*, 2005a] and the numerical MHD-based ENLIL model [*Odstrcil et al*., 2004]. The results show that the ENLIL's MAE was lower than the SARM's MAE, which was lower than the ESA's MAE. The SARM's best results were obtained when both flare and true CME speeds were used.

**Key words:** Space Weather / modeling / shock arrival time prediction

## 1. Introduction

One of the targets for space weather forecasters is to improve the interplanetary (IP) CME-driven shock time arrival predictions. CMEs often drive interplanetary (IP) shocks that impart the first pressure pulse on the magnetosphere resulting in sudden storm commencements [*Chao and Lepping*, 1974]. IP shocks are also drivers of high-energy solar energetic particle events, which can, for example, damage space-based equipment [*Miller et al*., 2003; *Wilson et al*., 2005], and interact with the Earth's atmosphere to produce penetrating neutrons that irradiate passengers and flight crews in commercial aircraft flying polar routes [*Beck et al*., 2005].

Several models have been proposed to predict shock arrival times, from empirical approaches, like the ESA (Empirical Shock Arrival) Model [*Gopalswamy et al*., 2005a, 2013], to numerical MHD-based models such as the WSA-ENLIL + Cone Model [*Odstrcil et al*., 2004], HAFv.3 model [*Fry et al*., 2001, 2003; *McKenna-Lawlor et al*., 2006; *Smith et al*., 2009], the STOA (Shock Time of Arrival) [*Dryer et al*., 2004; *Fry et al*., 2001; *McKenna-Lawlor et al*., 2006; Zhao and Dryer, 2014], and the Interplanetary Shock Propagation Model (ISPM) [*Smith and Dryer*, 1990]. The ESA and the ENLIL models will be used for comparison of the results later in the paper.

The use of flare data for predicting CME-related IP phenomena, such as shocks and related SEP events is a topic that has been debated and researched for decades [*Smith and Dryer*, 1990; *McKenna-Lawlor et al*., 2006; *Núñez*, 2011, 2015; *Liu and Qin*, 2012]. It is widely accepted that CMEs and flares are not causally related to each other; however, there is empirical evidence of a close relationship between flares and CMEs. *Yashiro and Gopalswamy* [2009] reported nearly a one-to-one correspondence between flares and CMEs, when energy fluence exceeds 0.1 J $m^{-2}$. *Jain et al.* [2010] showed that the speed of CMEs increases with the plasma temperature of X-ray flares, having a correlation coefficient r = 0.82. *Núñez* [2011] presented a SEP forecasting model, called UMASEP, that uses flare data to predict the occurrence of well-connected SEP events (obtaining a Probability of Detection (POD) of 90%), and the intensity of the prompt component of those well-connected events.

From a physical perspective, some studies have been carried out to obtain a broader view of the whole CME-flare eruption process. For example, *Chen et al*. [2010] analyzed the physical connection between the acceleration of CMEs and associated flare energy release. They use the erupting flux rope model [*Chen and Garren*, 1993; *Krall et al*., 2000], in which the driver is a poloidal flux injection. They concluded that injection of poloidal flux produces an electromotive force around the flux rope that can accelerate particles to X-ray energies. In summary, *Chen et al.* [2010] hypothesized that the poloidal flux injection, driver of flux rope eruptions, is also related to X-ray signatures. This can explain why empirical relationships exist between flare manifestations and CME travel times (e.g., [*Caroubalos*, 1964]; *Pick and Vilmer* [2008], and *Reeves and Moats* [2010] addressed this relationship quantitatively)

In this paper, we use the SARM model to predict IP shock arrival times from CME and/or flare data. We assume that the restraining IP force on CMEs is the aerodynamic drag caused by a lower-speed ambient solar wind [*Cargill*, 2004; *Vršnak et al*., 2010]. The main goal of this paper is to calibrate a formula that predicts the interplanetary shock speed from CME data (radial, cone-model-estimated or plane-of-sky speed) and/or flare data (duration, peak flux and location). We use the term IP shock speed to mean the physical speed of the shock discontinuity in the solar wind detected by spacecrafts. This study presents empirical evidence that IP shocks (directly driven by CMEs) are correlated with flares. There are important justifications for using a combination of CME and flare data in SARM: As noted above, flare and CME are physically related to the energy release process. Flare data are easy to obtain for real-time and historic data, radial CME-speed data are less readily available. For most of the multi-spacecraft observations we used to calibrate the SARM model, only plane-of-sky speeds were available. Since these speeds are rough earthward CME-speed approximations they are affected by projection effects.

Section 3.1 shows that the best shock arrival time prediction results were obtained with either radial or cone-model-estimated CME data, and flare data; however, satisfactory results were obtained using flare data alone. It is important to mention that due to the current low availability of radial CME speeds, the SARM model is being applied with flare data alone for the real-time prediction of shock arrival times as part of a larger model that predicts >10 MeV SEP event occurrence, peak flux and duration for the European Space Agency [*Garcia-Rigo et al*., 2016].

The SARM model has been calibrated with a dataset of 120 observations of shocks by an in-situ spacecraft from 1997 to 2010 from 0.72 to 8.7 AU. The shock arrival time prediction errors are presented in terms of mean absolute errors normalized to 1 AU, which are calculated as the mean absolute error divided by the distance (in AUs) at which the shock was detected. Finally, in this paper, we have compared SARM with the empirical ESA model and the MHD-based ENLIL model using a dataset of 20 shocks analyzed in other studies [*Taktakishvili et al*., 2009; *Gopalswamy*, 2013] by using radial CME speeds and/or flare data during the interval 2010-2012, and with/without the consideration of multi-CME complex shock events.

The paper is organized as follows: section 2 presents the empirical SARM model, and the approach used to calibrate its coefficients from a dataset of 120 shock events; section 3 presents the results, including the validation experiments with shock data that were not used in the calibration of the model; and, finally, section 4 presents the conclusions.

## 2. Shock arrival time prediction model

The Shock ARrival Model (SARM) that we present in this paper uses an equation of motion of a body under a drag force. A peculiarity of these problems is seen when the deceleration of the body is proportional to the square of its initial speed. For the case of the CME propagation, *Cargil* [2004] suggested that the drag force is $F_{drag} = -k\ A\,\rho(V_{CME} - V_{SW})^2$, where $k$ is the solar wind-induced drag coefficient for the CME, $A$ is the cross-sectional area of the CME, $\rho$ is the ambient solar wind density, $V_{CME}$ is the CME speed, and $V_{SW}$ is the ambient solar wind speed.

Several authors *[McKenna-Lawlor et al.,* 2002; *Zhao and Feng*, 2014] concluded that solar wind speed measurements $V_{SW}$ do not significantly improve the shock arrival predictions, and their use in shock models is less of an advantage than might at first appear. For the few benefits of using solar wind speed measurements $V_{SW}$, and for the sake of simplicity in the posterior calibration process, we decided to construct a drag-based model that does not take into account the solar wind speed. With the aim of building this model, we know that for these problems in which the drag force is $-kv^2$ [*Conrad*, 2002; *Herman*, 2013], the speed $v$ of the body influenced by the drag force may be expressed with an exponential decrease that is a function of the distance $x$ traveled from the initial location, and the drag coefficient $k$. Observational data [*Wang et al*., 2001, 2003] show that CMEs with very high initial speeds have corresponding IP shock speeds at distances > 1 AU that decrease very slowly out to several tens of AU. Thus we assume that the IP shock speed decreases gradually to an *asymptotic* shock speed, $V_a$, whose value depends on CME and flare data. This speed $V_a$ is a mathematical approximation that is necessary to simplify the model and its calibration. Several multi-spacecraft studies have reported on shocks detected at 1 AU and later, at several AUs [*Riley et al*., 2003; *Gonzalez-Esparza et al*., 1998]. From these studies, one can conclude that: 1) the shock speeds at distances of ~8.7 AU associated with most of the high-speed CME events are notably greater than the average solar wind speed (e.g., see Figure 1), and 2) the speeds of the aforementioned shocks decreased very slowly in most of their time traveling to those distances.

To fit the observational data, we also adopt a simple mathematical model in which the shock speed has an exponential decrease until an asymptotic speed. We use observational data to calibrate the initial conditions, the drag coefficient, and the asymptotic speed of this model. Based on these approximations, the instantaneous IP shock speed (*Vs*) may be estimated as a function of the radial distance *x* from the sun, as:

$$Vs(x) = \frac{dx}{dt} = \quad Vcme_x \, e^{-kx} + V_a \tag{1}$$

where *x* is the heliospheric distance from the sun to the IP shock, and $Vcme_x$ is the component of the radial CME initial speed (*Vcme*), projected on the axis from the sun to the spacecraft for which the arrival time will be predicted. $Vcme_x$ is calculated as *Vcme cos* ($\alpha$)**cos* ($\beta$), where $\alpha$ and $\beta$ are the longitude and latitude of the associated flare from the spacecraft's viewpoint. We assume that the shock's front propagates radially from the associated flare's location. The use of the location of the associated flare as the propagation direction of the IP shock is an approximation to simplify the model and its calibration.

In other words, *Vcme* is calculated along a vector that is normal to the solar flare site, and $Vcme_x$ is calculated along a vector in a different direction (i.e., the sun-spacecraft axis). Therefore, we need to project the initial CME speed on the sun-spacecraft axis by using the cosine function, which is an approach adopted by *Gopalswamy* [2013] for estimating earthward CME speeds from STEREO-observations-based radial CME speeds. This is appropriate for keeping scalar SARM and ESA equations; all of the parameters in the equations are parallel vectors.

In order to properly use SARM, radial CME speeds should be used to calculate its component $Vcme_x$ in the direction of the target goal (i.e., the location of the shock observation); however, other CME speed types may be the only available data. Section 2.1 presents the used approach to calculate $Vcme_x$ using plane-of-sky speeds without correction and with the cone-model-based correction [*Xie et al*., 2004].

Note that equation (1) does not simulate the propagation speed of the CME; in fact, the CME speed is not used for any calculation, nor included in the calibration process. The differential equation (1) is used to simulate the shock displacements $dx$ from the sun, to a distance $x_T$, where the spacecraft is located. SARM needs to solve the propagation formula described in equation (1) by using a numerical method (e.g., Runge Kutta 4th order) for iteratively simulating the shock location of every time step. The numerical solution calculates the instantaneous IP shock speed as a function of $x$, that is $dx/dt = f(x)$, where $f(x)$ is the differential equation (1), and where the initial conditions ($x = 0$ and $t = 0$) are also taken into account. This simulation may be summarized as follows: let us assume that at a time $t_1$ the shock is at a distance $x_1$ and propagates with a speed $v_1$, calculated according to equation (1) for $x = x_1$. With the purpose of performing a simulation until the time $t_1 + dt$, the integration method uses $v_1$ and $dt$ to calculate the shock-traveled distance $dx_1$. The simulation of the next time step $dt$ takes into account the new shock location (i.e., $x_1 + dx_1$), and the new IP shock speed, say $v_2$, calculated at the new location by using equation (1). The simulation process continues until $x = x_T$, that is, until the shock reaches the target distance. Therefore the simulated time $t$ at $x_T$ will be the SARM's predicted arrival time at the target location.

Figure 1 illustrates the design strategy of SARM in terms of an MHD-simulated shock speed profile (red curve) and the SARM's target shock speed profile (blue curve) for the case of the Bastille Day CME-driven shock on July 14th, 2000. The MHD-simulated profile was derived from the results obtained by *Von Steiger & Richardson* [2006] from 1 AU to 63 AU using a 2.5-D MHD numerical model. The speeds for intermediate distances show a decrease in the CME deceleration to an observed shock speed at 63 AU of 460 km $s^{-1}$. The deceleration is higher during the first hours of the CME expansion throughout the heliosphere, while the deceleration is very low at distances > 1 AU. With the aim of predicting a similar behavior, the SARM's target shock speed profile assumes an asymptotic shock speed, $V_a$, that is proportional to the released energy of the associated solar event. The initial CME speed and the peak flux and duration of the associated flare are well-documented manifestations of the released energy. For this reason $V_a$ is derived from CME and flare data. The rest of section 2 explains the calibration approach of SARM (including how the drag coefficient $k$, as well as the formula to calculate $V_a$ from CME and flare data, were empirically found).

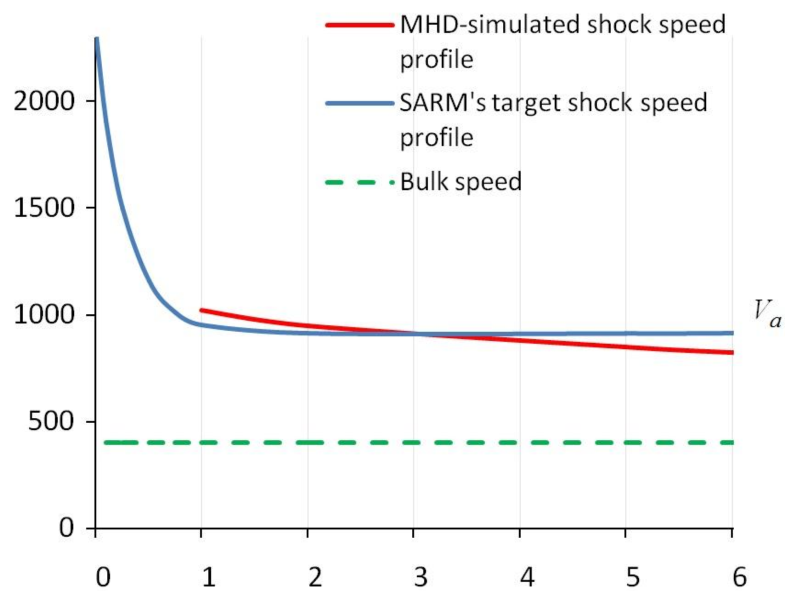



**Figure 1**. This figure illustrates the design strategy of the SARM model. The solid red curve shows an MHD-simulated shock speed profile as a function of distance (AU) from the sun to 6 AU for the case of the Bastille Day CME-driven shock on July 14th, 2000. This profile was derived from the results obtained by *Von Steiger & Richardson* [2006], using a 2.5-D MHD numerical model [*Wang et al.*, 2001; *Wang and Richardson*, 2003]. The green dashed line shows the mean solar wind speed, which is approximately 400 km $s^{-1}$ according to several authors [*Fleishman and Toptygin*, 2013]. The blue curve shows the SARM's target shock speed profile: the IP shock speed decreases due to the solar-wind induced drag on its driver (i.e. the CME) until it reaches an asymptotic speed $V_a$, which is calculated from the initial CME speed and the peak flux and duration of the associated flare.

### 2.1. Empirical approximation of $Vcme_x$ from radial and non-radial CME speeds

The procedure to estimate the CME speed component $Vcme_x$ depends on the target location of the spacecraft (e.g., Mars) for which the arrival time is going to be predicted, as well as the available CME data (radial, cone-model-estimated or plane-of-sky speeds). In general, we may say that if the CME speed in the direction of the spacecraft is not known, an estimate of the radial CME speed *Vcme* is needed to project it onto the direction from the Sun to the spacecraft, located at Earth or elsewhere in the solar system.

If we know the radial CME speed (i.e., *Vcme*), $Vcme_x$ is *Vcme* projected on the sun-spacecraft axis; that is, $Vcme_x = Vcme\ cos(\alpha) * cos\ (\beta)$, where $\alpha$ is the longitudinal distance between the flare location and the spacecraft location and $\beta$ is the latitudinal distance between the flare location and the spacecraft location. To predict shock arrival times at Earth, $\alpha$ and $\beta$ are the associated flare's longitude and latitude. To make predictions for another place in the solar system, $\alpha$ is calculated as $\alpha' - \Delta\alpha$, where $\alpha'$ is the flare's longitude from the earth's point of view, and $\Delta\alpha$ is the heliocentric longitudinal distance between the earth and the spacecraft. $\beta$ is calculated the same way in terms of latitudes. The longitudinal distance between the earth and the spacecraft is well-documented data for shock events observed at long distances.

In those cases for which the earthward speed $Vcme_E$ is known (e.g., by using the Cone model approach), we use it to calculate $Vcme_x$ depending on the case. To predict shock arrival times at Earth, $Vcme_x = Vcme_E$. To make predictions for another place in the solar system, $Vcme_E$ is deprojected to make an estimation of the radial speed; that is, $Vcme = Vcme_E\ /\ (\cos(\alpha').\cos(\beta'))$, where $\alpha'$ and $\beta'$ are the flare's longitude and latitude from the earth's point of view. Then, the radial speed is projected onto the direction from the Sun to the spacecraft, by using the approach mentioned in the previous paragraph.

In those shock events for which plane-of-sky CME speed ($V_{POS}$) is the only available information, we need to infer the earthward speed, $Vcme_E$, *Michałek et al.* [2003] reported that actual earthward speeds are 20% higher than plane-of-sky speeds. We used their finding as a first approximation of a statistical conversion factor: $Vcme_E = 1.2\ V_{POS}$. After completing SARM's calibration process (explained in section 2.3), we empirically confirmed *Michałek et al's* finding. Taking advantage of the use of a larger dataset, we refined the statistical conversion factor to $Vcme_E = 1.26\ V_{POS}$, by minimizing the arrival time prediction errors. To predict shock arrival times at Earth, $Vcme_x = Vcme_E$. To make predictions for another place in the solar system, $Vcme_E$ is deprojected to make an estimation of the radial speed, which is projected onto the direction from the Sun to the spacecraft, by using the approach mentioned in the previous paragraphs.

It is important to say that, while not very accurate, radial and cone-model estimated CME speeds are estimations that are closer to true values; for this reason, in this paper, we say that they are *true* speeds. In section 3, we study the effect on the accuracy of shock arrival time predictions by using true and plane-of-sky speeds (see Figure 5).

### 2.2. SARM calibration steps

The SARM calibration process uses equation (1) with interplanetary shock arrival times and solar associations in Table 1 to carry out an iterative data-driven three-step analysis to determine the coefficients that minimizes the mean absolute error (MAE), normalized to 1 AU. The calibration process finally leads to equation (2). Table 1 presents the shock events that were used to calibrate SARM. It shows the observed IP shocks and the reported solar associations (i.e., CME and/or flare data) of 120 shocks observed from 1997 to 2010 at distances from 0.72 to 8.7 AU.

**Table 1**. List of shock events observed from 0.72 to 8.7 AU. The last column includes the references of the studies that investigated the shocks' solar associations and IP observations from spacecraft data, including ACE, IMP, Stereo, Mars Global Surveyor, Ulysses, and Cassini. [a]

| | Shock | | | | Shock's associated CME | | | | | | Shock's Associated Flare | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Relative direction[e] | | | | | | | | | |
| Event | Dist (AU) | Shock Time (UTC) | Transit Time (h) | Long Distance (deg) | Vcme[b] (km/s) | $Vcme_E$[c] Axis (km/s) | $V_{POS}$[d] (km/s) | Lat (deg) | Long (deg) | CME time (UTC) | Location | Class | Start | Peak | end | s/c[f] | Ref[g] |
| Event1 | 0.72 | 8/2/10 11:30 | 27.7 | -53.2 | 1284 | | | 20 | 19.18 | 8/1/10 7:50 | N19E34 | C3.2 | 7:55 | 8:26 | 9:35 | 5 | 23, 15 |
| Event2 | 0.72 | 8/1/10 14:41 | 54.2 | -53.2 | 619 | | | 20 | -4.8 | 7/30/10 8:30 | | | | | | 5 | 23, 15 |
| Event3 | 1 | 1/10/97 0:52 | 81.7 | 0 | | | 136 | -18 | -6 | 1/6/97 15:10 | | | | | | 6 | 9, 10 |
| Event4 | 1 | 2/9/97 12:40 | 60.2 | 0 | | | 490 | -20 | 4 | 2/7/97 0:30 | | | | | | 6 | 9, 10 |
| Event5 | 1 | 4/10/97 12:55 | 70.5 | 0 | | 790 | | -30 | -19 | 4/7/97 14:27 | S30E19 | C6.8 | 13:50 | 14:07 | 14:19 | 6 | 24 |
| Event6 | 1 | 4/10/97 12:58 | 70.5 | 0 | | | 830 | -30 | -19 | 4/7/97 14:27 | S30E19 | C6.8 | 13:50 | 14:07 | 14:19 | 6 | 24, 17, 18 |
| Event7 | 1 | 5/15/97 1:15 | 66.8 | 0 | | | 464 | 21 | 19 | 5/12/97 6:30 | N21W19 | C1.3 | 4:42 | 4:55 | 5:26 | 6 | 9, 10 |
| Event8 | 1 | 11/6/97 22:18 | 64.1 | 0 | | | 785 | -14 | 33 | 11/4/97 6:10 | S14W33 | X2.1 | 5:52 | 5:58 | 6:02 | 6 | 9, 10, 11,17 |
| Event9 | 1 | 11/22/97 9:10 | 68.7 | 0 | | | 150 | -17 | 12 | 11/19/97 12:27 | | | | | | 6 | 9, 10 |
| Event10 | 1 | 11/30/97 7:14 | 65.3 | 0 | | | 441 | 20 | 60 | 11/27/97 13:57 | N20W60 | X2.6 | 12:59 | 13:17 | 13:20 | 6 | 17, 18 |
| Event11 | 1 | 12/30/97 1:15 | 94.7 | 0 | | | 197 | 24 | -14 | 12/26/97 2:31 | | | | | | 6 | 9, 10 |
| Event12 | 1 | 1/6/98 13:30 | 86.0 | 0 | | | 438 | 47 | 3 | 1/2/98 23:28 | | | | | | 6 | 9, 10 |
| Event13 | 1 | 3/4/98 11:05 | 94.3 | 0 | | | 176 | -24 | 1 | 2/28/98 12:48 | | | | | | 6 | 9, 10, 16 |
| Event14 | 1 | 5/1/98 21:00 | 52.0 | 0 | | 1448 | | -18 | -20 | 4/29/98 16:58 | S18E20 | M6.8 | 16:06 | 16:37 | 16:59 | 6 | 24 |
| Event15 | 1 | 5/1/98 21:20 | 52.4 | 0 | | | 1374 | -18 | -20 | 4/29/98 16:58 | S18E20 | M6.8 | 16:06 | 16:37 | 16:59 | 6 | 9, 10 |
| Event16 | 1 | 5/4/98 6:00 | 39.9 | 0 | 1418 | | | 11 | 18 | 5/2/98 14:06 | | | | | | 7 | 25 |
| Event17 | 1 | 10/18/98 19:28 | 81.4 | 0 | | | 262 | 22 | 1 | 10/15/98 10:04 | | | | | | 6 | 9, 10 |
| Event18 | 1 | 11/8/98 4:20 | 55.3 | 0 | | | 1124 | 18 | 21 | 11/5/98 20:59 | N18W21 | M8.4 | 19:00 | 19:55 | 20:12 | 6 | 24, 17, 18 |
| Event19 | 1 | 11/8/98 4:42 | 55.7 | 0 | | 1230 | | 18 | 21 | 11/5/98 20:58 | N18W21 | M8.4 | 19:00 | 19:55 | 20:12 | 6 | 9, 10, 24 |
| Event20 | 1 | 2/11/99 8:58 | 51.4 | 0 | | | 808 | -25 | 24 | 2/9/99 5:33 | S25W24 | C2.3 | 4:54 | 5:08 | 5:31 | 6 | 17, 18 |
| Event21 | 1 | 3/10/99 0:38 | 41.7 | 0 | | | 664 | -22 | -76 | 3/8/99 6:54 | | | | | | 7 | 17, 18 |
| Event22 | 1 | 4/16/99 11:10 | 68.5 | 0 | | | 291 | 16 | 0 | 4/13/99 3:30 | | | | | | 6 | 9, 10 |
| Event23 | 1 | 7/2/99 0:25 | 66.5 | 0 | | | 589 | 18 | 7 | 6/29/99 5:54 | N18W07 | M1.4 | 5:01 | 5:10 | 5:17 | 7 | 17, 18 |
| Event24 | 1 | 7/2/99 0:25 | 64.9 | 0 | | 771 | | 18 | 7 | 6/29/99 7:31 | N18W07 | M1.4 | 5:01 | 5:10 | 5:17 | 6 | 24 |
| Event25 | 1 | 7/13/99 8:45 | 55.2 | 0 | | | 318 | 13 | -32 | 7/11/99 1:31 | | | | | | 7 | 17, 18 |
| Event26 | 1 | 7/22/99 9:50 | 78.7 | 0 | | | 430 | 12 | 16 | 7/19/99 3:06 | N12W16 | C4.2 | 1:49 | 2:53 | 2:53 | 7 | 17, 18 |
| Event27 | 1 | 8/8/99 17:50 | 107.4 | 0 | | | 405 | 16 | 64 | 8/4/99 6:26 | N16W64 | M6 | 5:45 | 5:57 | 6:14 | 7 | 18 |
| Event28 | 1 | 8/23/99 11:30 | 60.1 | 0 | | | 812 | -25 | -64 | 8/20/99 23:26 | S25E64 | M9.8 | 23:01 | 23:10 | 23:17 | 7 | 17, 18 |
| Event29 | 1 | 8/31/99 1:31 | 55.1 | 0 | | | 462 | -26 | 14 | 8/28/99 18:26 | S26W14 | X1.1 | 17:52 | 18:05 | 18:18 | 7 | 17, 18 |
| Event30 | 1 | 9/2/99 9:35 | 72.8 | 0 | | | 404 | 18 | 16 | 8/30/99 8:50 | | | | | | 7 | 17, 18 |
| Event31 | 1 | 9/15/99 20:05 | 50.6 | 0 | | | 444 | 15 | 17 | 9/13/99 17:31 | | | | | | 6 | 17, 18 |
| Event32 | 1 | 2/11/00 2:33 | 65.0 | 0 | | 1089 | | 25 | -26 | 2/8/00 9:30 | N25E26 | M1.3 | 8:42 | 9:00 | 9:18 | 6 | 24 |
| Event33 | 1 | 2/11/00 21:28 | 49.6 | 0 | | 954 | | 31 | -4 | 2/9/00 19:54 | | | | | | 6 | 24 |

Table 1 (Cont.)

| | Shock | | | | Shock's associated CME | | | | | | Shock's Associated Flare | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Relative direction[e] | | | | | | | | | |
| Event | Dist (AU) | Shock Time (UTC) | Transit Time (h) | Long Distance (deg) | Vcme[b] (km/s) | $Vcme_E$[c] Axis (km/s) | $V_{POS}$[d] (km/s) | Lat (deg) | Long (deg) | CME time (UTC) | Location | Class | Start | Peak | end | s/c[f] | Ref[g] |
| Event34 | 1 | 2/11/00 23:18 | 44.8 | 0 | | | 944 | 31 | -4 | 2/10/00 2:30 | | | | | | 7 | 17, 18 |
| Event35 | 1 | 2/14/00 6:56 | 50.4 | 0 | | | 1107 | 26 | 23 | 2/12/00 4:31 | N26W23 | M1.7 | 3:51 | 4:10 | 4:31 | 7 | 17, 18 |
| Event36 | 1 | 2/20/00 20:50 | 71.3 | 0 | | | 550 | -29 | -7 | 2/17/00 21:30 | S29E07 | M1.3 | 20:17 | 20:35 | 21:07 | 6 | 9, 10, 11 |
| Event37 | 1 | 2/20/00 21:00 | 71.5 | 0 | | 719 | | n/a | n/a | 2/17/00 21:30 | n/a | M1.3 | 20:17 | | 21:07 | 6 | 24 |
| Event38 | 1 | 4/7/00 1:00 | 56.5 | 0 | 2038 | | | 16 | 66 | 4/4/00 16:32 | N16W66 | C9.7 | 15:12 | 15:41 | 16:05 | 7 | 25 |
| Event39 | 1 | 5/2/00 10:44 | 49.8 | 0 | | | 540 | -11 | 18 | 4/30/00 8:54 | | | | | | 7 | 17, 18 |
| Event40 | 1 | 6/8/00 8:40 | 40.8 | 0 | | | 1119 | 33 | -25 | 6/6/00 15:54 | N33E25 | X2.3 | 14:58 | 15:25 | 15:40 | 7 | 11, 17, 18 |
| Event41 | 1 | 6/8/00 9:04 | 41.2 | 0 | | 1282 | | 33 | -25 | 6/6/00 15:54 | N33E25 | X2.3 | 14:58 | 15:25 | 15:40 | 6 | 24 |
| Event42 | 1 | 6/18/00 17:02 | 68.9 | 0 | | | 1081 | 20 | 65 | 6/15/00 20:06 | | | | | | 7 | 17, 18 |
| Event43 | 1 | 7/13/00 9:18 | 59.5 | 0 | | | 1352 | 18 | -49 | 7/10/00 21:50 | N18E49 | M5.7 | 21:05 | 21:42 | 22:27 | 7 | 17, 18 |
| Event44 | 1 | 7/14/00 15:32 | 43.0 | 0 | | | 820 | 17 | 65 | 7/12/00 20:30 | | | | | | 7 | 17 |
| Event45 | 1 | 7/14/00 15:39 | 74.2 | 0 | | | 1078 | 18 | -27 | 7/11/00 13:27 | | | | | | 6 | 9, 10, |
| Event46 | 1 | 7/15/00 14:18 | 27.4 | 0 | | | 1674 | 18 | 27 | 7/14/00 10:54 | N18W27 | X5.7 | 10:03 | 10:24 | 10:43 | 6 | 9, 10, 11, 12 |
| Event47 | 1 | 7/16/00 1:00 | 38.1 | 0 | 1686 | | | 22 | 25 | 7/14/00 10:54 | | | | | | 7 | 25 |
| Event48 | 1 | 7/28/00 5:41 | 74.2 | 0 | | | 528 | 6 | 3 | 7/25/00 3:30 | N06W03 | M8 | 2:43 | 2:49 | 2:54 | 7 | 17, 18 |
| Event49 | 1 | 8/11/00 18:51 | 50.3 | 0 | | | 702 | 11 | 11 | 8/9/00 16:30 | N11W11 | C2.3 | 15:19 | 16:22 | 17:00 | 6 | 9, 10 |
| Event50 | 1 | 8/12/00 10:00 | 65.5 | 0 | 960 | | | 11 | 11 | 8/9/00 16:30 | N11W11 | C2.3 | 15:19 | 16:22 | 17:00 | 7 | 25 |
| Event51 | 1 | 9/6/00 16:12 | 117.3 | 0 | | | 411 | 10 | 60 | 9/1/00 18:54 | N10W60 | C9.1 | 18:05 | 18:20 | 18:30 | 7 | 17, 18 |
| Event52 | 1 | 9/17/00 16:57 | 35.6 | 0 | | | 1215 | 14 | 7 | 9/16/00 5:18 | N14W07 | M5.9 | 4:06 | 4:26 | 4:48 | 7 | 18 |
| Event53 | 1 | 9/17/00 17:00 | 35.7 | 0 | | 1327 | | 14 | 7 | 9/16/00 5:18 | N14W07 | M5.9 | 4:06 | 4:26 | 4:48 | 6 | 24 |
| Event54 | 1 | 9/18/00 0:00 | 42.7 | 0 | 1493 | | | 14 | 7 | 9/16/00 5:18 | N14W07 | M5.9 | 4:06 | 4:26 | 4:48 | 7 | 25 |
| Event55 | 1 | 10/3/00 1:05 | 55.0 | 0 | | | 703 | -20 | -42 | 9/30/00 18:06 | S20E42 | M1 | 17:38 | 18:27 | 19:05 | 6 | 9, 10 |
| Event56 | 1 | 10/12/00 21:44 | 69.9 | 0 | | | 798 | 1 | 14 | 10/9/00 23:50 | N01W14 | C6.7 | 23:19 | 23:43 | 0:21 | 6 | 9, 10, 16, 17 |
| Event57 | 1 | 11/6/00 9:20 | 62.9 | 0 | | | 291 | 2 | 2 | 11/3/00 18:26 | N02W02 | C3.2 | 18:35 | 19:02 | 20:06 | 6 | 9, 10 |
| Event58 | 1 | 3/19/01 11:37 | 79.8 | 0 | | | 271 | 11 | 9 | 3/16/01 3:50 | | | | | | 6 | 9, 10 |
| Event59 | 1 | 3/31/01 9:00 | 46.6 | 0 | 913 | | | 20 | 19 | 3/29/01 10:26 | N20W19 | X1.7 | 9:57 | 10:15 | 10:32 | 7 | 25 |
| Event60 | 1 | 4/11/01 13:40 | 45.8 | 0 | | | 1192 | -21 | 4 | 4/9/01 15:54 | S21W04 | M7.9 | 15:20 | 15:34 | 16:00 | 8 | 1 |
| Event61 | 1 | 4/11/01 14:06 | 46.2 | 0 | | 1210 | | -21 | 4 | 4/9/01 15:54 | S21W04 | M7.9 | 15:20 | 15:34 | 16:00 | 6 | 24 |
| Event62 | 1 | 4/12/01 0:00 | 42.5 | 0 | 1260 | | | 9 | -52 | 4/10/01 5:30 | N09E52 | X2.3 | 5:06 | 5:26 | 5:42 | 7 | 25 |
| Event63 | 1 | 4/21/01 15:31 | 51.0 | 0 | | | 392 | 20 | 20 | 4/19/01 12:30 | | | | | | 6 | 9, 10 |
| Event64 | 1 | 4/28/01 5:02 | 40.5 | 0 | | 1003 | | 17 | 31 | 4/26/01 12:30 | N17W31 | M7.8 | 11:26 | 13:12 | 13:19 | 6 | 24 |
| Event65 | 1 | 4/28/01 5:06 | 40.6 | 0 | | | 1006 | 17 | 31 | 4/26/01 12:30 | N17W31 | M7.8 | 11:26 | 13:12 | 13:19 | 6 | 9, 10 |
| Event66 | 1 | 5/27/01 14:52 | 58.8 | 0 | | | 569 | -9 | 4 | 5/25/01 4:06 | | | | | | 6 | 9, 10 |
| Event67 | 1 | 9/25/01 20:18 | 33.8 | 0 | | 1773 | | -16 | -23 | 9/24/01 10:30 | S16E23 | X2.6 | 9:32 | 10:38 | 11:09 | 6 | 24 |
| Event68 | 1 | 9/25/01 20:45 | 34.3 | 0 | | | 2402 | -16 | -23 | 9/24/01 10:30 | S16E23 | X2.6 | 9:32 | 10:38 | 11:09 | 8 | 1 |
| Event69 | 1 | 10/21/01 22:00 | 53.2 | 0 | 1452 | | | 15 | 29 | 10/19/01 16:50 | N15W29 | X1.6 | 16:13 | 16:30 | 16:43 | 7 | 25 |
| Event70 | 1 | 10/31/01 13:47 | 50.0 | 0 | | | 598 | 12 | -25 | 10/29/01 11:50 | N12E25 | M3.6 | 10:56 | 11:13 | 11:19 | 6 | 9, 10 |

Table 1 (Cont.)

| | Shock | | | | Shock's associated CME | | | | | | Shock's Associated Flare | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Shock | Transit | Long | | | | Relative direction[e] | | | | | | | | | |
| | Dist | Time | Time | Distance | Vcme[b] | $Vcme_E$[c] | $V_{POS}$[d] | Lat | Long | CME time | | | | | | | |
| Event | (AU) | (UTC) | (h) | (deg) | (km/s) | Axis (km/s) | (km/s) | (deg) | (deg) | (UTC) | Location | Class | Start | Peak | end | s/c[f] | Ref[g] |
| Event71 | 1 | 11/6/01 7:00 | 38.4 | 0 | 1999 | | | 38 | 6 | 11/4/01 16:35 | N38W06 | X1 | 16:03 | 16:03 | 16:57 | 7 | 25 |
| Event72 | 1 | 11/19/01 18:15 | 60.8 | 0 | | | 1379 | -13 | -42 | 11/17/01 5:30 | S13E42 | M2.8 | 4:49 | 5:25 | 6:11 | 8 | 1 |
| Event73 | 1 | 11/24/01 4:52 | 29.4 | 0 | | 1551 | | n/a | n/a | 11/22/01 23:30 | n/a | M9.9 | 22:32 | | 0:06 | 6 | 24 |
| Event74 | 1 | 11/24/01 5:50 | 30.3 | 0 | | | 1437 | -17 | 36 | 11/22/01 23:30 | S17W36 | M9.9 | 22:32 | 23:30 | 0:06 | 6 | 9, 10 |
| Event75 | 1 | 11/24/01 17:00 | 41.5 | 0 | 2809 | | | -17 | 36 | 11/22/01 23:30 | S17W36 | M9.9 | 22:32 | 23:30 | 0:06 | 7 | 25 |
| Event76 | 1 | 3/18/02 12:33 | 61.5 | 0 | | | 907 | -8 | 3 | 3/15/02 23:06 | | | | | | 6 | 9, 10 |
| Event77 | 1 | 3/23/02 10:47 | 58.9 | 0 | | | 1049 | -19 | 60 | 3/20/02 23:54 | S19W60 | C5.7 | 23:46 | 0:23 | 0:46 | 6 | 9, 10 |
| Event78 | 1 | 4/19/02 8:00 | 47.6 | 0 | | | 1231 | -14 | 1 | 4/17/02 8:26 | | | | | | 6 | 9, 10 |
| Event79 | 1 | 5/18/02 19:50 | 67.0 | 0 | | | 600 | -22 | -14 | 5/16/02 0:50 | | | | | | 6 | 9, 10 |
| Event80 | 1 | 5/18/02 19:51 | 67.0 | 0 | | 870 | | -22 | -14 | 5/16/02 0:50 | | | | | | 6 | 24 |
| Event81 | 1 | 5/23/02 18:00 | 38.2 | 0 | 1957 | | | -30 | 34 | 5/22/02 3:50 | | | | | | 7 | 25 |
| Event82 | 1 | 5/29/03 19:10 | 45.1 | 0 | | | 1122 | -7 | 17 | 5/27/03 22:06 | S07W17 | X1.3 | 22:56 | 23:07 | 23:13 | 8 | 1 |
| Event83 | 1 | 5/29/03 19:10 | 42.3 | 0 | | | 964 | -7 | 20 | 5/28/03 0:50 | S07W20 | X3.6 | 0:17 | 0:27 | 0:39 | 8 | 1 |
| Event84 | 1 | 5/30/03 16:20 | 38.9 | 0 | | | 1366 | -6 | 37 | 5/29/03 1:27 | S06W37 | X1.2 | 0:51 | 1:05 | 1:12 | 8 | 1 |
| Event85 | 1 | 10/29/03 6:00 | 18.5 | 0 | | 2752 | | -16 | -8 | 10/28/03 11:30 | S16E08 | X17.2 | 9:51 | 11:10 | 11:24 | 6 | 24 |
| Event86 | 1 | 10/29/03 6:10 | 18.7 | 0 | | | 2459 | -16 | -8 | 10/28/03 11:30 | S16E08 | X17.2 | 9:51 | 11:10 | 11:24 | 8 | 1 |
| Event87 | 1 | 10/30/03 1:00 | 37.5 | 0 | 2868 | | | -16 | -8 | 10/28/03 11:30 | S16E08 | X17.2 | 9:51 | 11:10 | 11:24 | 7 | 25 |
| Event88 | 1 | 10/30/03 20:00 | 23.1 | 0 | | | 2029 | -15 | 2 | 10/29/03 20:54 | S15W02 | X10 | 20:37 | 20:49 | 21:01 | 8 | 1 |
| Event89 | 1 | 10/30/03 23:00 | 26.1 | 0 | 1872 | | | 19 | 30 | 10/29/03 20:54 | N19W30 | X10 | 20:37 | 20:49 | 21:01 | 7 | 25 |
| Event90 | 1 | 11/20/03 8:25 | 47.6 | 0 | | | 1660 | 0 | -18 | 11/18/03 8:50 | N00E18 | M3.9 | 8:12 | 8:31 | 8:59 | 8 | 1 |
| Event91 | 1 | 11/20/03 21:00 | 60.2 | 0 | 1215 | | | 0 | 11.8 | 11/18/03 8:50 | N00W12 | M3.9 | 8:12 | 8:31 | 8:59 | 7 | 25 |
| Event92 | 1 | 7/27/04 14:00 | 47.1 | 0 | 1289 | | | 8 | 33 | 7/25/04 14:54 | N08W33 | M1.1 | 14:19 | 15:14 | 16:43 | 7 | 25 |
| Event93 | 1 | 11/8/04 7:00 | 52.9 | 0 | 1525 | | | 7 | 0 | 11/6/04 2:06 | N07W00 | M3.6 | 1:40 | 1:57 | 2:08 | 7 | 25 |
| Event94 | 1 | 9/11/05 11:00 | 39.2 | 0 | 1903 | | | -10 | -58 | 9/9/05 19:48 | S10E58 | X6.2 | 19:13 | 20:04 | 20:36 | 7 | 25 |
| Event95 | 1 | 4/5/10 7:58 | 46.9 | 0 | 1011 | | | n/a | n/a | 4/3/10 9:05 | | - | | | | 6 | 13 |
| Event96 | 1 | 4/5/10 9:00 | 39.7 | 0 | 1071 | | | n/a | n/a | 4/3/10 17:16 | | | | | | 9 | 26 |
| Event97 | 1 | 5/29/10 22:00 | 56.3 | 0 | 600 | | | n/a | n/a | 5/27/10 13:39 | | | | | | 9 | 26 |
| Event98 | 1 | 6/3/10 9:00 | 55.3 | 0 | 760 | | | n/a | n/a | 6/1/10 1:40 | | | | | | 9 | 26 |
| Event99 | 1 | 8/3/10 5:00 | 45.2 | -71.3 | 1284 | | | 19 | -34 | 8/1/10 7:50 | N19E34 | C3.2 | 7:55 | 8:26 | 9:35 | 5 | 15 |
| Event100 | 1 | 8/17/10 18:00 | 79.8 | 78.57 | 600 | | | 17 | 52 | 8/14/10 10:12 | N17W52 | C4.4 | 9:38 | 9:20 | 10:31 | 5 | 16 |
| Event101 | 1.38 | 11/20/01 3:35 | 70.1 | -55.9 | | | 1379 | -13 | 13.94 | 11/17/01 5:30 | N13E42 | M2.8 | 4:49 | 5:25 | 6:11 | 4 | 1 |
| Event102 | 1.4 | 10/30/03 5:30 | 42.0 | -21.5 | | | 2459 | -16 | 13.47 | 10/28/03 11:30 | N16E08 | X17.2 | 9:51 | 11:10 | 11:24 | 4 | 1 |
| Event103 | 1.4 | 10/31/03 11:30 | 38.6 | -21.9 | | | 2029 | 19 | 30 | 10/29/03 20:54 | S05W02 | X10 | 20:37 | 20:49 | 21:01 | 4 | 1 |
| Event104 | 1.43 | 12/31/01 18:00 | 69.5 | -72.3 | 2216 | | | -24 | -17.7 | 12/28/01 20:30 | N50E90 | X3.4 | 20:02 | 20:45 | 21:32 | 4 | 1 |
| Event105 | 1.43 | 11/21/03 4:50 | 68.0 | -29.8 | | | 1660 | 0 | 11.83 | 11/18/03 8:50 | N00E18 | M3.9 | 8:12 | 8:31 | 8:59 | 4 | 1 |
| Event106 | 1.43 | 5/31/03 3:45 | 50.3 | 31.35 | | | 1237 | -6 | 5.65 | 5/29/03 1:27 | S06W37 | X1.2 | 0:51 | 1:05 | 1:12 | 4 | 1 |
| Event107 | 1.44 | 9/25/01 19:55 | 33.4 | -36.5 | | | 2402 | -16 | 13.45 | 9/24/01 10:30 | S16E23 | X2.6 | 9:32 | 10:38 | 11:09 | 4 | 1 |

Table 1. (Cont)

| Event | Shock | | | | Shock's associated CME | | | | | | Shock's Associated Flare | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Shock | Transit | Long | | | | Relative direction[e] | | | | | | | | | |
| | Dist | Time | Time | Distance | Vcme[a] | $Vcme_E$[c] | $V_{POS}$[d] | Lat | Long | CME time | | | | | | | |
| | (AU) | (UTC) | (h) | (deg) | (km/s) | Axis (km/s) | (km/s) | (deg) | (deg) | (UTC) | Location | Class | Start | Peak | end | s/c[f] | Ref[g] |
| Event108 | 1.44 | 5/29/03 23:20 | 47.5 | 32.1 | | | 964 | -7 | -15.1 | 5/27/03 23:50 | S07W17 | X1.3 | 22:56 | 23:07 | 23:13 | 4 | 1 |
| Event109 | 1.44 | 5/29/03 23:20 | 46.5 | 31.72 | | | 1366 | -7 | -11.7 | 5/28/03 0:50 | S07W20 | X3.6 | 0:17 | 0:27 | 0:39 | 4 | 1 |
| Event110 | 1.53 | 4/19/01 3:35 | 85.5 | 26.15 | 1199 | | | 20 | 58.85 | 4/15/01 14:06 | S20W85 | X14.4 | 13:19 | 13:50 | 13:55 | 4 | 1 |
| Event111 | 1.57 | 4/12/01 11:00 | 53.5 | 29.02 | | | 2411 | 9 | -52 | 4/10/01 5:30 | N09E23 | X2.3 | 5:06 | 5:26 | 5:42 | 4 | 1 |
| Event112 | 1.57 | 4/12/01 11:00 | 67.1 | 29.02 | | | 1192 | -21 | -25 | 4/9/01 15:54 | S21W04 | M7.9 | 15:20 | 15:54 | 16:00 | 4 | 1 |
| Event113 | 1.58 | 3/9/89 20:15 | 78.0 | -72 | 1260 | | | 35 | 3 | 3/6/89 14:15 | N35E69 | X15 | 13:50 | 14:05 | 14:40 | 3 | 2 |
| Event114 | 5 | 3/2/99 0:00 | 332.8 | 0 | | | n/a | 22 | 14 | n/a | S23W14 | M3.2 | 2:49 | 3:12 | 3:45 | 2 | 3, 20 |
| Event115 | 5.24 | 11/13/03 16:19 | 212.4 | 100 | 2657 | | | -34 | -17 | 11/4/03 19:54 | S19W83 | X17.4 | 19:29 | 19:53 | 20:06 | 1 | 4 |
| Event116 | 5.24 | 11/15/03 0:03 | 176.2 | 100 | 2237 | | | 0 | -10 | 11/7/03 15:54 | | | | | | 1 | 4 |
| Event117 | 5.3 | 1/26/05 17:00 | 154.1 | 29 | | 3675 | | -17 | 32 | 1/20/05 6:54 | N14W61 | X7.1 | 6:36 | 7:01 | 7:26 | 2 | 21, 5 |
| Event118 | 5.4 | 3/23/98 21:30 | 560.7 | 5.5 | | | 176 | -3 | -3.5 | 2/28/98 12:48 | | | | | | 2 | 7 |
| Event119 | 6.6 | 12/7/01 12:00 | 348.5 | 21 | | | 1437 | 15 | 13 | 11/22/01 23:30 | S15W34 | M9.9 | 22:32 | 0:06 | 0:06 | 1 | 19 |
| Event120 | 8.7 | 11/10/03 0:00 | 174.5 | 60 | | | 2981 | -14.4 | -1 | 11/2/03 17:30 | S18W59 | X8.3 | 17:03 | | | 1 | 4 |

a This table presents the 120 shocks observed from 1997 to 2010, at distances from 0.72 to 8.7 AU, that were used to calibrate SARM. The following columns also present the reported solar associations. Column 1 lists the identifiers of the shocks. Columns 2-4 contain the shock data: column 2 presents the distances at which shocks were detected, column 3, the shock arrival time, and column 4, the IP shock transit time. Column 5 lists the heliocentric longitudinal distance in degrees between Earth and the spacecraft. Columns 6-11 present data from the CME associated with the corresponding shock: the radial CME speed *Vcme* (if available) is listed in column 6; the cone-model-based earthward CME speed $Vcme_E$ (if available), using the *Xie et al*. [2004]'s approach, is listed in column 7; if neither *Vcme* or $Vcme_E$ is available, column 8 lists plane-of-sky speed from in the Catalog of CME events observed by the Large Angle and Spectrometric Coronagraph Experiment (LASCO) instrument in the Solar and Heliospheric Observatory (SOHO) Catalog (linear speed); columns 9 and 10 present the shock propagation direction in terms of the longitude and latitude of the associated solar event from the spacecraft's point of view; the first time of CME detection by LASCO is listed in column 11. Columns 12-16 present X-ray-related data from the flare associated with the corresponding shock; the location in column 12, the flare peak flux in column 13, the peak start, time and end time are listed in columns 14, 15, and 16, respectively, according to the NOAA/SWPC's edited event list (i.e., http://legacy-www.swpc.noaa.gov/ftpdir/indices/events/events.txt). Finally, column 17 contains the bibliographic references from where each shock event, and its corresponding CME and flare associations, was taken.

*b* *Vcme* is the radial CME speed.

*c* $Vcme_E$ is the earthward CME speed obtained by using the cone-model approach explained in *Xie et al.* [2004].

*d* $V_{POS}$ is the linear speed in the SOHO/LASCO catalog (http://cdaw.gsfc.nasa.gov/CME_list).

e Columns 9 and 10 list the shock propagation direction from the target spacecraft's point of view. See Section 2.1 about the approach used to calculate the values of this column.

*f* Spacecraft or interplanetary shock database: 1: Cassini; 2: Ulysses; 3: Phobos-2; 4: Mars Global Surveyor; 5: Stereo A/B; 6: WIND; 7:ACE; 8: NASA's OMNI-DB; 9: NASA's DONKI-DB.

*g* Bibliographic references: 1: [*Falkenberg, et al*., 2011]; 2: [*Aran et al*., 2007]; 3: [*Riley et al.*, 2003]; 4: [*Jian*, 2008]; 5: [*Lepri et al.*, 2012]; 6: [*Richardson et al*., 2006]; 7: [*Skoug et al.*, 2000]; 9: [*Gopalswamy et al*., 2005a]; 10: http://lepmfi.gsfc.nasa.gov/mfi/mag_cloud_pub1p.html; 11: [*McKenna-L. et al*., 2002]; 12: [*Burlaga et al*., 2001]; 13: [*Xie et al*., 2013]; 15: [*Möstl et al*., 2012]; 16: [*Steed et al*., 2010]; 17: [*Fry et al*., 2003]; 18: [*Cho et al*., 2003]; 19: [*Lario et al*., 2004]; 20: [*Lario et al*., 2001]; 21: [*Gopalswamy et al*., 2005b]; 23: [*Webb et al*., 2013]; 24: [*Xie et al*., 2006]; 25:[*Taktakishvili et al*., 2011]; 26: http://kauai.ccmc.gsfc.nasa.gov/DONKI/.

The strategy of the SARM's calibration approach is to get first estimates of the speed $V_a$ in Eq. 1 by using shocks observed at distances > 1 AU, and then, starting from these, refine the estimate of all parameters in Eq. 1 by using all shocks in Table 1.

The overall idea is that: at distances > 1 AU, $V_a$ can be approximated by the distance Sun-observer divided by the observed travel time (i.e., the IP shock transit speed); therefore, the purpose of the first two calibration steps is to obtain an approximate predictor of the speed $V_a$ from CME data alone and from flare data alone at distances > 1 AU. To this end, this approach uses coronagraphic measurements of CME speeds (explained in detail in Section S.1) and a proxy inferred from soft X-ray observations (explained in Section S.2). In the last calibration step (explained in detail in Section S.3), we found that the mean value of the two CME speeds (i.e., the observed CME speed and the X-ray-inferred proxy) is most appropriate, with the further advantage that a CME speed can be guessed even when the data coverage is incomplete. Starting from these first guesses of $V_a$, all 120 shocks in Table 1 (from 0.72 to 8.9 AU) are used to encounter the drag coefficient $k$ and refine $V_a$. During the last calibration step, full numerical simulations are undertaken, from the sun to the observed IP shock's distance, to calibrate the coefficients of equation (1) by minimizing the mean absolute error of arrival time predictions, normalized to 1 AU.

In summary, section S describes the details of the coarse-to-fine optimization process approach used to transform equation (1) into equation (2), which presents the resulting formula for calculating the instantaneous IP shock speed:

$$\frac{dx}{dt} = Vdriver_x\, e^{-7x} + 0.42\, Vdriver_x + 330 \quad \text{km s}^{-1} \tag{2}$$

where:

- $x$ is the heliospheric distance (in AUs) from the sun to the IP shock
- $Vdriver_x$ is the *Average* ($Vcme_x$, $saVcme_x$) in km s$^{-1}$. That is, if no flare data are available, $Vdriver_x = Vcme_x$; if no CME data are available, $Vdriver_x = saVcme_x$
- $Vcme_x$ is the radial CME speed $Vcme$ projected on the sun-spacecraft axis; that is, $Vcme_x = Vcme \cos(\alpha) * \cos(\beta)$, where $\alpha$ and $\beta$ are the longitude and latitude of the associated flare from the

spacecraft's viewpoint. If radial CME speed is not available, the cone-model speed $Vcme_E$ may be used; that is, $Vcme_x = Vcme_E$. For more information about how to obtain $Vcme_x$, see section 2.1.

– *saVcme* is calculated as 1015 $\log_{10}$ (*PFswpc x FDswpc*) +5500, where *PFswpc* and *FDswpc* are the flare peak flux and the flare duration, using the NOAA/SWPC data. This empirical formula is introduced in this paper (see section S.2 for details about how this log-linear equation was obtained).

### 2.3. SARM's triggering conditions for issuing shock arrival time predictions

If SARM issues a forecast for each CME or flare occurrence, it will generate a high number of false alarms; therefore, SARM has to filter out some forecasts to maximize the number of issued forecasts and minimize the mean absolute error. We realized that the angular distance between the locations of the solar event and the spacecraft is an important CME geometry-oriented triggering condition. We also realized that a minimum CME speed is an important triggering condition of forecasts that are only based on CME data; and a minimum flare peak flux is another important triggering condition of forecasts that are only based on flare data.

We empirically found the following three forecast triggering rules, which were used in section 3 as necessary conditions to trigger forecasts: the minimum speed of $Vcme_x$ to issue a CME-related shock forecast is 330 km $s^{-1}$; the minimum flare peak flux to issue an X-ray-related shock forecast is C4; and, the maximum Euclidean angular distance ω between the flare and the spacecraft locations is 60 degrees, where $\omega$ is calculated as the square root of the sum of the squares of the longitudinal distance and the latitudinal distance between the aforementioned locations.

As we mentioned in section 2, the asymptotic shock speed $V_a$ is calculated in terms of CME and flare data. Thus, equation (2) shows that the term $V_a$ is calculated as 0.42 $Vdriver_x$ + 330 (i.e., the non-exponential summand in equation (2)), where $Vdriver_x$ is a function of CME and flare data. That is, $V_a$ is estimated as the sum of two terms: the fixed term 330 km $s^{-1}$, and the *Vdriver*-dependant term. If the minimum speed of $Vcme_x$ to trigger a SARM prediction is 330 km $s^{-1}$ (as we mentioned in the previous paragraph), and there are no flare data, $Vdriver_x$ is also 330 km $s^{-1}$; therefore, the asymptotic shock speed $V_a$ for the minimum-speed CME is 468.6 km $s^{-1}$ (i.e., 0.42 x 330 + 330), which is larger than mean solar wind speed, as it should be.

A web-based version of SARM that uses equation (2) to make shock arrival time forecasts is available in http://spaceweather.uma.es/sarm/index.html

## 3. Results and arguments

The calibration data includes the shock data from Table 1 (i.e., from 1997 to2010), describing solar situations that are very diverse; therefore, section 3.1 includes the analysis of the prediction results with calibration data in order to give a better idea of the model’s expected strengths and weaknesses; and section 3.2 includes the validation results with shock data that were not used in the model’s calibration.

### 3.1. Analysis of prediction results with calibration data

Table 2 presents the SARM's predicted shock transit times for the events of Table 1. The last three columns show the normalized prediction errors (i.e. observed transit time - predicted transit time) for the 120 shocks, by using CME and/or flare data. Note that some of the predictions are not issued, because the properties of the associated solar event do not fulfill the SARM's triggering conditions explained in section 2.3. The last three columns are the main reference for presenting the analysis of the results in Figures 2 to 5. As shown in this table, the use of CME and flare data obtained a normalized mean absolute error of 7.1 h for all the events in Table 1 that meet SARM's triggering conditions, which is better than the mean absolute errors of the individual models, by using data from the CME alone (8.9 h), or by using flare data alone (8.6 h).

Figure 2 shows the distribution of the errors listed in the last column of Table 2 (i.e., normalized errors of SARM's predictions from CME and flare data) for all shocks of Table 1 (top chart) and for shocks at 1 AU (bottom chart). For both figures, the most frequent interval is [-2.5 h, 2.5 h]. Regarding normalized absolute errors, Figure 2a shows that the mean, median and standard deviation are 7.1 h, 5.1 h and 6.0 h, respectively. Figure 2b shows that the mean, median and standard deviation are 7.0 h, 5.0 h and 6.3 h, respectively. Although the performance of SARM's predictions for 1 AU is similar to the performance for all the analyzed distances, a closer analysis of the error performance as a function of the distance (see Figure 3) shows that there are important differences in the performance of SARM's shock arrival predictions for different distances, taking into account shock events in Table 1.

**Table 2**. Model's input data for each shock and its forecast error with flare and/or CME data (normalized to 1 AU) [a]

| | | SARM's outputs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Predicted Shock Travel Times | | | Normalized prediction error | | |
| Event | Shock Distance AU | *Vcme$_x$* km/s | *saVcme$_x$* kms$^{-1}$ | Average *Vcme$_x$* & *saVcme$_x$* kms$^{-1}$ | With CME Data Only hours | With Flare Data Only hours | With CME & Flare Data hours | With CME Data Only hours | With Flare Data Only hours | With CME & Flare Data hours |
| Event1 | 0.7 | 1140 | | 1140 | 32.1 | | 32.1 | -4.42 | c | -4.42 |
| Event2 | 0.7 | 580 | | 580 | 41.9 | | 41.9 | 12.30 | c | 12.30 |
| Event3 | 1.0 | 171 | | 171 | | | | B | c | b, c |
| Event4 | 1.0 | 617 | | 617 | 63.5 | | 63.5 | -3.31 | c | -3.31 |
| Event5 | 1.0 | 790 | 350 | 570 | 55.9 | 86.1 | 67.7 | 14.57 | -15.60 | 2.77 |
| Event6 | 1.0 | 1046 | 350 | 698 | 47.6 | 86.1 | 61.1 | 22.97 | -15.55 | 9.39 |
| Event7 | 1.0 | 585 | | 585 | 65.2 | | 65.2 | 1.57 | c | 1.57 |
| Event8 | 1.0 | 989 | 977 | 983 | 49.2 | 55.7 | 52.2 | 14.98 | 8.46 | 11.91 |
| Event9 | 1.0 | 189 | | 189 | | | | b | c | b, c |
| Event10 | 1.0 | 556 | 1398 | 977 | | | | | d | d |
| Event11 | 1.0 | 248 | | 248 | | | | b | c | b, c |
| Event12 | 1.0 | 552 | | 552 | 67.0 | | 67.0 | 19.06 | c | 19.06 |
| Event13 | 1.0 | 222 | | 222 | | | | b | c | b, c |
| Event14 | 1.0 | 1448 | 1215 | 1332 | 38.5 | 46.4 | 42.1 | 13.55 | 5.60 | 9.95 |
| Event15 | 1.0 | 1731 | 1215 | 1473 | 34.0 | 46.4 | 39.2 | 18.42 | 5.94 | 13.14 |
| Event16 | 1.0 | 1325 | | 1325 | 40.9 | | 40.9 | -0.97 | c | -0.97 |
| Event17 | 1.0 | 330 | | 330 | 82.2 | | 82.2 | -0.75 | c | -0.75 |
| Event18 | 1.0 | 1416 | 1444 | 1430 | 39.1 | 41.8 | 40.4 | 16.27 | 13.58 | 14.97 |
| Event19 | 1.0 | 1230 | 1444 | 1337 | 42.9 | 41.8 | 42.3 | 12.81 | 13.96 | 13.40 |
| Event20 | 1.0 | 1018 | | 1018 | 48.3 | | 48.3 | 3.10 | c | 3.10 |
| Event21 | 1.0 | 837 | | 837 | | | | | c | d |
| Event22 | 1.0 | 367 | | 367 | 79.2 | | 79.2 | 0.47 | c | 0.47 |
| Event23 | 1.0 | 742 | 350 | 546 | 57.8 | 82.2 | 67.8 | 8.70 | -15.63 | -1.30 |
| Event24 | 1.0 | 771 | 350 | 561 | 56.6 | 82.2 | 67.0 | 8.27 | -17.25 | -2.12 |
| Event25 | 1.0 | 401 | | 401 | 76.6 | | 76.6 | -21.39 | c | -21.39 |
| Event26 | 1.0 | 542 | 350 | 446 | 67.5 | 82.2 | 74.1 | 11.20 | -3.50 | 4.60 |
| Event27 | 1.0 | 510 | 894 | 702 | | | | | d | d |
| Event28 | 1.0 | 1023 | 848 | 936 | | | | | d | d |
| Event29 | 1.0 | 582 | 1113 | 848 | 65.3 | 49.7 | 56.4 | -10.20 | 5.38 | -1.32 |
| Event30 | 1.0 | 509 | | 509 | 69.4 | | 69.4 | 3.35 | c | 3.35 |
| Event31 | 1.0 | 559 | | 559 | 66.5 | | 66.5 | -15.96 | c | -15.96 |
| Event32 | 1.0 | 1089 | 350 | 720 | 46.4 | 86.3 | 60.2 | 18.70 | -21.20 | 4.83 |
| Event33 | 1.0 | 954 | | 954 | 50.2 | | 50.2 | -0.65 | c | -0.65 |
| Event34 | 1.0 | 1189 | | 1189 | 43.9 | | 43.9 | 0.95 | c | 0.95 |
| Event35 | 1.0 | 1395 | 480 | 937 | 39.5 | 76.9 | 52.1 | 10.92 | -26.50 | -1.70 |
| Event36 | 1.0 | 693 | 460 | 577 | 59.9 | 76.6 | 67.3 | 11.43 | -5.29 | 4.08 |
| Event37 | 1.0 | 719 | 460 | 590 | 58.8 | 76.6 | 66.5 | 12.72 | -5.12 | 4.97 |
| Event38 | 1.0 | 791 | 357 | 574 | | | | | d | d |
| Event39 | 1.0 | 680 | | 680 | 60.5 | | 60.5 | -10.65 | c | -10.65 |
| Event40 | 1.0 | 1410 | 1650 | 1530 | 39.2 | 42.6 | 40.8 | 1.55 | -1.78 | -0.03 |
| Event41 | 1.0 | 1282 | 1650 | 1466 | 41.8 | 42.6 | 42.2 | -0.60 | -1.38 | -0.98 |
| Event42 | 1.0 | 1362 | | 1362 | | | | | c | d |
| Event43 | 1.0 | 1704 | 1330 | 1517 | 34.4 | 54.6 | 42.2 | 25.12 | 4.89 | 17.32 |
| Event44 | 1.0 | 1033 | | 1033 | | | | | c | d |
| Event45 | 1.0 | 1358 | | 1358 | 40.2 | | 40.2 | 34.00 | c | 34.00 |
| Event46 | 1.0 | 2109 | 2028 | 2069 | 29.4 | 34.2 | 31.6 | -1.95 | -6.75 | -4.17 |
| Event47 | 1.0 | 1415 | | 1415 | 39.1 | | 39.1 | -1.00 | c | -1.00 |
| Event48 | 1.0 | 665 | 594 | 630 | 61.2 | 64.9 | 63.0 | 13.00 | 9.26 | 11.20 |

**Table 2 (cont.)**

| | | SARM's outputs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Predicted Shock Travel Times | | | Normalized prediction error | | |
| Event | Shock Distance AU | $Vcme_x$ km/s | $saVcme_x$ $kms^{-1}$ | Average $Vcme_x$ & $saVcme_x$ $kms^{-1}$ | With CME Data Only hours | With Flare Data Only Hours | With CME & Flare Data hours | With CME Data Only hours | With Flare Data Only hours | With CME & Flare Data Hours |
| Event49 | 1.0 | 885 | | 885 | 52.5 | | 52.5 | -2.15 | c | -2.15 |
| Event50 | 1.0 | 926 | | 926 | 51.1 | | 51.1 | 14.37 | c | 14.37 |
| Event51 | 1.0 | 518 | 350 | 434 | | | | | d | d |
| Event52 | 1.0 | 1531 | 1050 | 1291 | 37.1 | 48.5 | 42.0 | -1.40 | -12.83 | -6.35 |
| Event53 | 1.0 | 1327 | 1050 | 1189 | 40.8 | 48.5 | 44.3 | -5.12 | -12.78 | -8.63 |
| Event54 | 1.0 | 1439 | 1050 | 1245 | 38.7 | 48.5 | 43.0 | 4.03 | -5.78 | -0.32 |
| Event55 | 1.0 | 886 | 589 | 737 | 52.5 | 76.0 | 62.0 | 2.51 | -20.99 | -7.04 |
| Event56 | 1.0 | 1005 | 350 | 678 | 48.7 | 81.4 | 60.9 | 21.22 | -11.48 | 9.05 |
| Event57 | 1.0 | 367 | | 367 | 79.2 | | 79.2 | -16.30 | c | -16.30 |
| Event58 | 1.0 | 341 | | 341 | 81.2 | | 81.2 | -1.44 | c | -1.44 |
| Event59 | 1.0 | 811 | 1436 | 1124 | 55.1 | 41.9 | 47.6 | -8.53 | 4.67 | -1.00 |
| Event60 | 1.0 | 1502 | 1157 | 1330 | 37.6 | 46.6 | 41.6 | 8.22 | -0.85 | 4.17 |
| Event61 | 1.0 | 1210 | 1157 | 1184 | 43.4 | 46.6 | 45.0 | 2.83 | -0.42 | 1.25 |
| Event62 | 1.0 | 773 | 2597 | 1685 | 56.6 | 36.1 | 44.0 | -14.10 | 6.45 | -1.52 |
| Event63 | 1.0 | 494 | | 494 | 70.4 | | 70.4 | -19.35 | c | -19.35 |
| Event64 | 1.0 | 1003 | 1610 | 1306 | 48.8 | 41.0 | 44.5 | -8.22 | -0.45 | -4.00 |
| Event65 | 1.0 | 1268 | 1610 | 1439 | 42.1 | 41.0 | 41.5 | -1.48 | -0.38 | -0.93 |
| Event66 | 1.0 | 717 | | 717 | 58.9 | | 58.9 | -0.13 | c | -0.13 |
| Event67 | 1.0 | 1773 | 2073 | 1923 | 33.4 | 32.6 | 33.0 | 0.43 | 1.25 | 0.85 |
| Event68 | 1.0 | 3027 | 2073 | 2550 | 22.1 | 32.6 | 26.3 | 12.17 | 1.70 | 7.93 |
| Event69 | 1.0 | 1228 | 1342 | 1285 | 43.0 | 45.2 | 44.1 | 10.19 | 7.97 | 9.10 |
| Event70 | 1.0 | 753 | 567 | 660 | 57.4 | 69.8 | 63.0 | -7.42 | -19.80 | -13.00 |
| Event71 | 1.0 | 1574 | 1394 | 1484 | 36.3 | 46.2 | 40.7 | 2.09 | -7.73 | -2.25 |
| Event72 | 1.0 | 1738 | 1017 | 1377 | 33.9 | 58.0 | 42.8 | 26.88 | 2.73 | 18.00 |
| Event73 | 1.0 | 1551 | 1633 | 1592 | 36.7 | 42.2 | 39.3 | -7.33 | -12.83 | -9.90 |
| Event74 | 1.0 | 1811 | 1633 | 1722 | 32.9 | 42.2 | 37.0 | -2.55 | -11.87 | -6.62 |
| Event75 | 1.0 | 2172 | 1633 | 1903 | 28.7 | 42.2 | 34.2 | 12.80 | -0.70 | 7.35 |
| Event76 | 1.0 | 1143 | | 1143 | 45.0 | | 45.0 | 16.45 | c | 16.45 |
| Event77 | 1.0 | 1322 | 350 | 836 | | | | | d | d |
| Event78 | 1.0 | 1551 | | 1551 | 36.7 | | 36.7 | 10.87 | c | 10.87 |
| Event79 | 1.0 | 756 | | 756 | 57.2 | | 57.2 | 9.77 | c | 9.77 |
| Event80 | 1.0 | 870 | | 870 | 53.0 | | 53.0 | 14.05 | c | 14.05 |
| Event81 | 1.0 | 1388 | | 1388 | 39.6 | | 39.6 | -1.46 | c | -1.46 |
| Event82 | 1.0 | 1414 | 1000 | 1207 | 39.1 | 50.4 | 44.1 | 5.94 | -5.31 | 1.02 |
| Event83 | 1.0 | 1215 | 1562 | 1389 | 43.3 | 38.3 | 40.6 | -0.94 | 4.03 | 1.70 |
| Event84 | 1.0 | 1721 | 1058 | 1389 | 34.1 | 53.9 | 41.8 | 4.78 | -15.04 | -2.89 |
| Event85 | 1.0 | 2752 | 2883 | 2817 | 23.9 | 23.9 | 23.9 | -5.35 | -5.38 | -5.37 |
| Event86 | 1.0 | 3098 | 2883 | 2990 | 21.7 | 23.9 | 22.7 | -2.98 | -5.21 | -4.05 |
| Event87 | 1.0 | 2732 | 2883 | 2808 | 24.0 | 23.9 | 23.9 | 13.52 | 13.62 | 13.57 |
| Event88 | 1.0 | 2557 | 2051 | 2304 | 25.3 | 30.8 | 27.8 | -2.18 | -7.67 | -4.65 |
| Event89 | 1.0 | 1531 | 2051 | 1791 | 37.0 | 34.7 | 35.9 | -10.93 | -8.63 | -9.75 |
| Event90 | 1.0 | 2092 | 917 | 1504 | 29.5 | 52.9 | 37.9 | 18.05 | -5.29 | 9.70 |
| Event91 | 1.0 | 1190 | 917 | 1054 | 43.9 | 52.0 | 47.6 | 16.32 | 8.15 | 12.60 |
| Event92 | 1.0 | 1074 | 853 | 963 | 46.8 | 59.2 | 52.2 | 0.32 | -12.07 | -5.12 |
| Event93 | 1.0 | 1514 | 654 | 1084 | 37.4 | 62.0 | 46.6 | 15.55 | -9.08 | 6.33 |
| Event94 | 1.0 | 1003 | 2387 | 1695 | 48.8 | 42.3 | 45.3 | -9.57 | -3.08 | -6.10 |
| Event95 | 1.0 | 917 | | 917 | 51.4 | | 51.4 | -4.52 | c | -4.52 |
| Event96 | 1.0 | 892 | | 892 | 52.2 | | 52.2 | -12.49 | c | -12.49 |
| Event97 | 1.0 | 588 | | 588 | 65.0 | | 65.0 | -8.67 | c | -8.67 |

**Table 2 (cont.)**

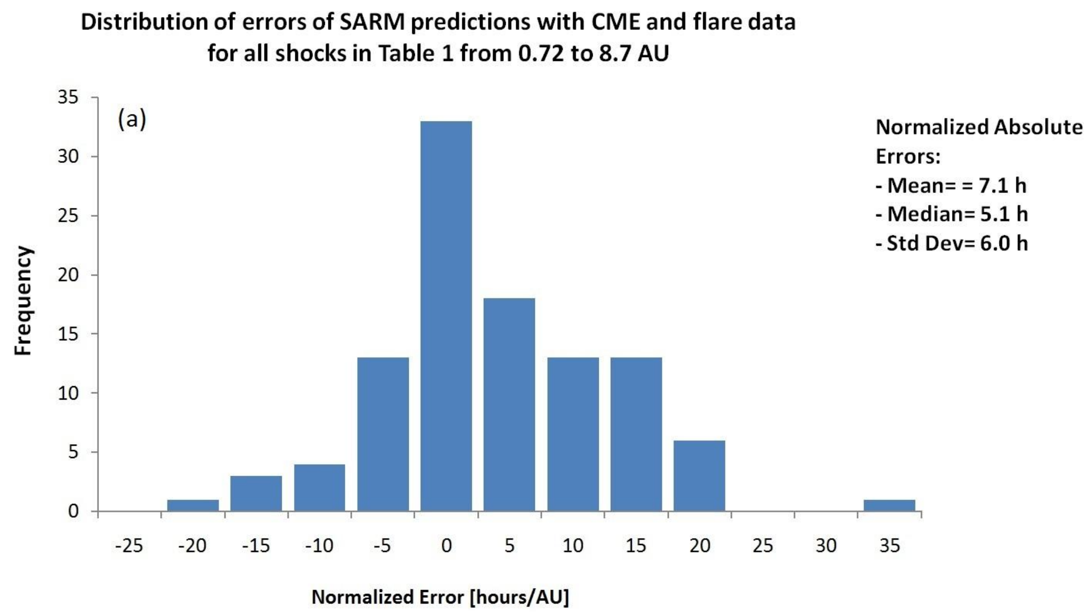
| | | SARM's outputs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Predicted Shock Travel Times | | | Normalized prediction error | | |
| Event | Shock Distance AU | $Vcme_x$ km/s | $saVcme_x$ kms[-1] | Average $Vcme_x$ & $saVcme_x$ kms[-1] | With CME Data Only hours | With Flare Data Only Hours | With CME & Flare Data hours | With CME Data Only hours | With Flare Data Only hours | With CME & Flare Data hours |
| Event98 | 1.0 | 732 | | 732 | 52.4 | | 52.4 | -2.94 | c | -2.94 |
| Event99 | 1.0 | 957 | | 957 | 40.2 | | 40.2 | -5.00 | c | -5.00 |
| Event100 | 1.0 | 528 | 350 | 439 | 90.7 | 74.8 | 83.6 | 10.89 | -5.03 | 3.82 |
| Event101 | 1.4 | 2271 | 1017 | 1644 | 48.4 | 71.1 | 56.6 | 21.68 | -1.05 | 13.52 |
| Event102 | 1.4 | 3045 | 2883 | 2964 | 35.0 | 37.3 | 36.1 | 7.04 | 4.69 | 5.93 |
| Event103 | 1.4 | 2101 | 2051 | 2076 | 41.8 | 47.2 | 44.3 | -3.18 | -8.58 | -5.66 |
| Event104 | 1.4 | 1929 | 2158 | 2044 | 53.6 | 54.1 | 53.8 | 15.94 | 15.40 | 15.68 |
| Event105 | 1.4 | 2154 | 917 | 1535 | 50.6 | 74.2 | 59.1 | 17.42 | -6.23 | 8.94 |
| Event106 | 1.4 | 1943 | 1058 | 1500 | 47.8 | 64.5 | 54.4 | 2.49 | -14.19 | -4.14 |
| Event107 | 1.4 | 3200 | 2073 | 2637 | 32.3 | 42.8 | 36.5 | 0.72 | -9.35 | -3.12 |
| Event108 | 1.4 | 1227 | 1000 | 1113 | 59.1 | 66.3 | 62.4 | -11.60 | -18.80 | -14.91 |
| Event109 | 1.4 | 1794 | 1562 | 1678 | 48.6 | 52.9 | 50.6 | -2.14 | -6.38 | -4.14 |
| Event110 | 1.5 | 573 | 2391 | 1482 | | | | | d | d |
| Event111 | 1.6 | 2048 | 2597 | 2322 | 51.5 | 57.3 | 54.1 | 2.03 | -3.84 | -0.62 |
| Event112 | 1.6 | 1363 | 1157 | 1260 | 66.7 | 76.2 | 71.0 | 0.40 | -9.06 | -3.86 |
| Event113 | 1.6 | 1035 | 2553 | 1794 | 78.9 | 59.7 | 67.0 | -0.86 | 18.30 | 10.99 |
| Event114 | 5.0 | n/a | 907 | 907 | | 327.2 | 327.2 | | 5.59 | 5.59 |
| Event115 | 5.2 | 2105 | 2486 | 2296 | 205.3 | 207.3 | 206.1 | 7.16 | 5.10 | 6.32 |
| Event116 | 5.2 | 2204 | | 2204 | 174.9 | | 174.9 | 1.28 | c | 1.28 |
| Event117 | 5.3 | 4532 | 2224 | 3378 | 146.0 | 162.4 | 149.6 | 8.07 | -8.25 | 4.46 |
| Event118 | 5.4 | 242 | | 242 | | | | B | c | b, c |
| Event119 | 6.6 | 2130 | 1633 | 1882 | 329.0 | 337.5 | 332.8 | 19.51 | 10.98 | 15.74 |
| Event120 | 8.7 | 3753 | 2148 | 2950 | 175.9 | 188.5 | 180.8 | -1.41 | -13.98 | -6.27 |

[a] This table lists the normalized mean absolute errors of the SARM's predictions by using the CME and flare data. Column 1 shows the identifiers of the events. Column 2 presents the distance at which the shock was detected. Columns 3, 4 and 5 list the used $Vcme_x$, $saVcme_x$ and the average ($Vcme_x$, $saVcme_x$), respectively. Columns 6, 7 and 8 present the prediction by using the speeds listed in columns 3-5, and columns 7-9 show the normalized absolute errors (i.e., absolute error / distance in AU) of the predictions using CME data only (column 9), flare data only (column 10) and both CME and flare data (column 11).

[b] Prediction using CME and flare data was filtered out because the Euclidean angular distance between the flare and the spacecraft locations is greater than 60º.

[c] Prediction using CME data only was filtered out because CME initial speed is lower than 330 km $s^{-1}$.

[d] Prediction using flare data only was filtered out because the associated flare peak flux is < C4 or no flare data is available.



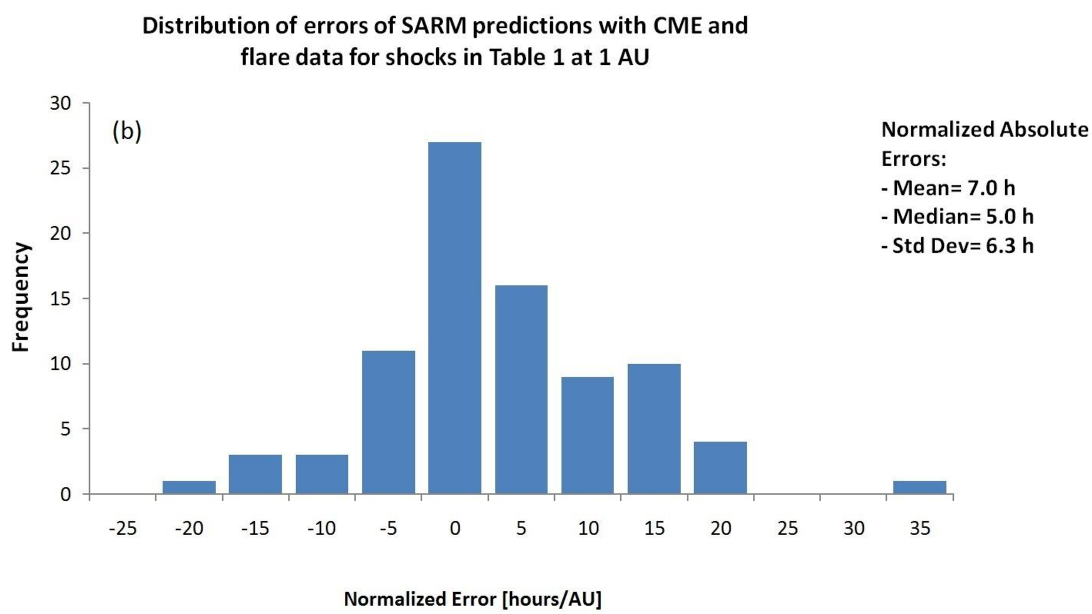




Figure 2 shows the distribution of errors of SARM predictions using CME and flare data. These errors were extracted from the results presented in Table 2. The top chart shows the distribution of errors for all shocks in Table 1 from 0.72 AU to 8.7 AU. The top chart also shows that the mean, median and standard deviation of normalized absolute errors are 7.1 h, 5.1 h, and 6.0 h, respectively. The bottom chart shows the distribution of errors for shocks at 1 AU. The bottom chart also shows that the mean, median and standard deviation of the normalized absolute errors for 1 AU are 7.0 h, 5.0 h, and 6.3 h, respectively. For both figures, the most frequent interval of errors is [-2.5 h, 2.5 h].

Figure 3 shows the normalized mean absolute errors for five groups of distances (i.e., 072 AU, 1 AU, 1.4-1.6 AU and 5.0-5.4 AU and 6.6-8.7 AU) from predictions using CME and/or flare data. It can be seen that the error averages using CME and flare data are less than 8.5 h, except for large distances 6.6-8.7 AU. The best normalized mean absolute error was obtained from predictions of shock arrival times for distances in the interval 5.0-5.4 AU, which was 4.5 h. In the case of 1 AU, the normalized mean absolute error is 7.0 h using both flare and CME data; if only CME speeds are used, the MAE was 9.2 h; and if only flare data are used, the MAE was 8.4 h.

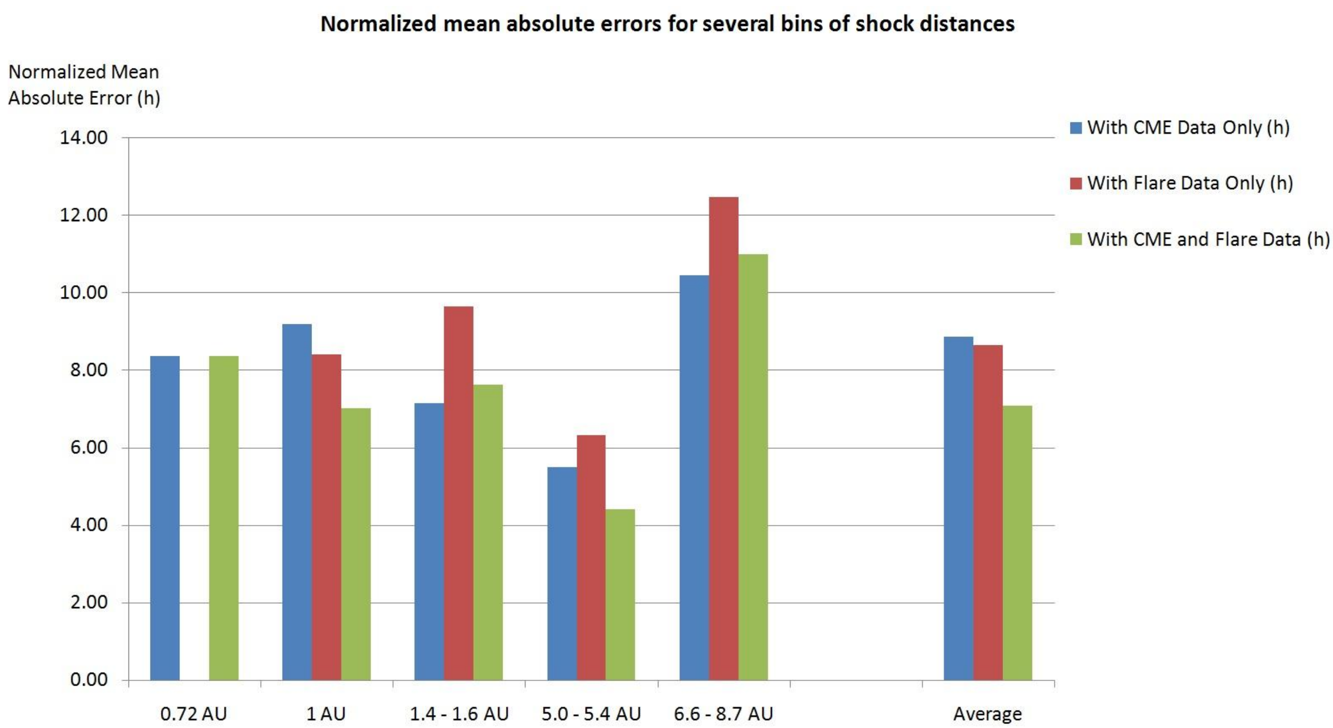



**Figure 3.** Distribution of normalized mean absolute errors for several groups of shock distances from CME and/or flare data. Blue, red and green bars show the normalized mean absolute errors using CME data only, flare data only, and both CME and flare data, respectively. Regarding the use of both CME and flare data, the least normalized mean absolute error was obtained from predictions of shock arrival times for distances in the interval 5 AU-5.4 AU, which was 4.5 h; the highest normalized mean absolute error was obtained from predictions for distances in the interval 6.6-8.7 AU, which was 11.0 h.

Figure 4 shows the forecasting error for those cases where flares were used. It presents the normalized mean absolute errors for several ranges of flare peak flux. It can be seen that SARM's forecasts using CME & flare data (green bars) are better when shocks are associated with >M4 flares. It is interesting to see that the SARM propagation model is also good for predicting the arrival time from CME data only when the shocks are associated with >M4 flares. Section 3.2, obtained a similar conclusion with validation experiments that use the earthward speed calculated from radial speeds.

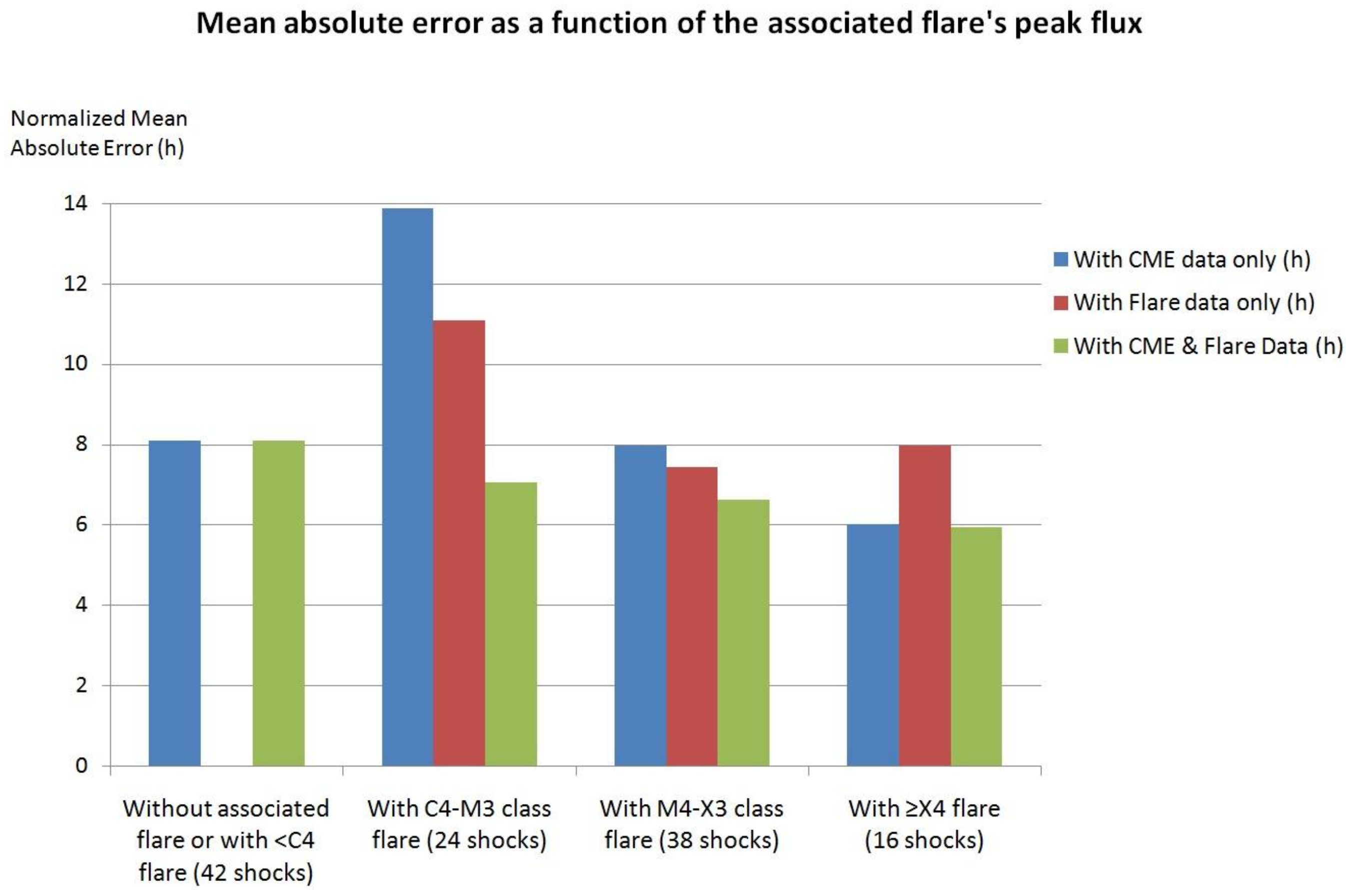



**Figure 4**. This chart shows the normalized mean absolute error depending on the associated flare's peak flux. The left group of bars shows the errors for the case in which the shock is not associated with a flare or it is associated with a <C4 flare. If the shock is associated with a ≥C4 flare, the figure shows the normalized mean absolute error depending on whether the flare's peak flux is between an interval (i.e. C4 - M3, M4 - X3, and ≥X3). It can be seen that SARM's forecasts are better as shocks are associated with >M4 flares.

The ESA model [*Gopalswamy et al*., 2005a] was introduced with 29 events that are included in Table 1. Gopalswamy *et al*. [2005a] reported a mean absolute error of 12 hours. For these data, SARM's combined approach (i.e., using CME and flare data) obtained a lower mean absolute error, which was 9.1 hours. Since ESA and SARM models used the same shock data during their calibration, this comparison is valid.

In *Falkenberg, et al*. [2011], a study was carried out using ENLILv2.6 for predicting shock arrivals at Earth and Mars. *Falkenberg et al*. [2011] reported a mean absolute error of 19.16 hours (by using manual parameter settings) while SARM obtained a mean absolute error of 12.86 hours; however, these results are not conclusive because SARM's triggering conditions filtered out several shocks, and these shocks were part of SARM's calibration data set. In section 3.2, a comparison is made with 1-AU shock data not included in the SARM calibration dataset which will allow us to provide a valid conclusion with regard to the comparison of SARM and the ENLIL models.

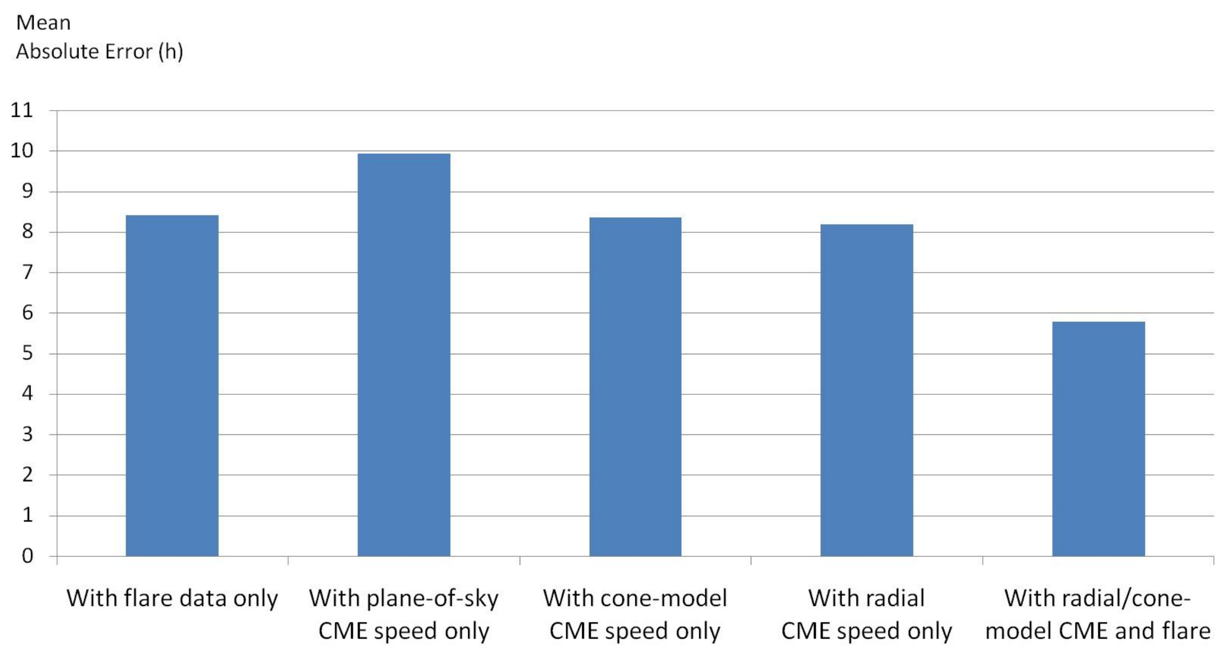



**Figure 5.** Mean absolute error of arrival predictions to 1 AU as a function of the data availability. The first bar shows the mean absolute errors if only flare data are used. The second, third and fourth bars show the mean absolute errors if only plane-of-sky, cone-model-estimated, or radial CME speeds are used. The fifth column presents the mean absolute errors if both flare and true CME speed (radial or cone-model speeds) are used. This figure shows that the worst results were obtained if only plane-of-sky speed data are used (9.9 h). The best results are obtained when both flare and true CME speeds are used (5.8 h).

Figure 5 summarizes the SARM's absolute errors of arrival predictions to 1 AU as function of the data availability. These MAE errors were extracted from the results presented in Table 2. The first bar shows that SARM obtained a MAE of 8.4.h if only flare data are available. The second bar shows a MAE of 9.9.h if only plane-of-sky CME speed data are available. Since plane-of-sky speed $V_{POS}$ are lower than the actual CME speeds, they were statistically adjusted to actual speeds (see section 2.1) by minimizing their normalized MAE error. The third bar shows a MAE of 8.3 h if only cone-model-corrected CME speeds are used. Cone-model speeds are considered *true* speeds, for this reason if only cone-model speed $V_E$ are available, $Vcme_E = V_E$. The fourth bar presents a MAE of 8.2 h if only radial CME speeds are used. And finally, the fifth bar presents a MAE of 5.8 h if true CME speeds (i.e., cone-model or projected-radial speeds) and flare data are used. The improvement of MAE using flare data and true CME data is nearly 30% compared to the MAE obtained with any true CME speed alone (i.e., projected radial speed or cone-mode-based CME speed).

Regarding Figures 2 to 5, the following conclusions may be drawn: in the case of 1 AU, the normalized mean and the median of absolute errors were 7.1 h and 5.1 h respectively; the mean absolute error using both true CME speeds and flare data was 5.8 h; for the prediction of shock arrivals at distances from 0.72 to 8.7 AU, the normalized mean and the median of absolute errors were 7.0 h and 5.0 h, respectively. If only CME data are available or if only flare data are available, the normalized MAE errors were 8.9 h and 8.6 h respectively, using all shock events.

### 3.2. Validation of the model

This section presents validation experiments with shock data that were not included in the calibration of the model. In this validation test, a list of 20 shock data selected in a recent study by *Gopalswamy et al*. [2013] is presented, in which radial CME speeds are obtained by using STEREO A/B data.

*Gopalswamy et al*. [2013] reported the forecasting error with CME speeds calculated for the ecliptic plane using STEREO data. An additional advantage of this shock data is that they reported a comparison between the ESA and ENLIL model, which was useful to compare SARM with these two state-of-the-art models. Of the 20 shocks, *Gopalswamy et al*. [2013] reported 2 shocks associated with complex CME-CME and CME-coronal hole interaction which could lead to large deviations from model predictions [*Nieves-Chinchilla et al*., 2012, 2013].

Table 3 presents the subset of 20 full halo events selected by *Gopalswamy et al*. [2013], and the forecast results after applying SARM and ENLIL with the same CME data.

Table 3. Shock data from *Taktakishvili et al*. [2009] and *Gopalswamy et al*. [2013] which includes radial and earthward CME speeds (using two measuring approaches). This table also presents the associated flare data. [a]

| | Shock data and time | Transit Time | CME Date and time | Radial Speed (*Vcme*) | Earthward speed ($Vcme_E$) | Earthward speed_1 ($Vcme1_E$) | Flare location | Flare class | Flare Duration (h) |
|---|---|---|---|---|---|---|---|---|---|
| eventT1 | 2/15/10 17:28 | 75.9 | 2/12/10 13:31 | 867 | 765 | 756 | N26E11 | M8.3 | 0.15 |
| eventT2 | 4/11/10 12:18 | 79.8 | 4/8/10 4:30 | 771 | 677 | 630 | N24E16 | | |
| eventT3 | 8/3/10 16:51 | 56.4 | 8/1/10 8:24 | 1031 | 784 | 1257 | N20E36 | C3.2 | 1.67 |
| eventT4 | 2/18/11 0:40 | 70.1 | 2/15/11 2:36 | 945 | 879 | 864 | N12W18 | X2.2 | 0.37 |
| eventT5 | 3/10/11 5:45 | 63.0 | 3/7/11 14:48 | 691 | 633 | 738 | N11E21 | M1.9 | 1.18 |
| eventT6 | 6/23/11 2:18 | 47.0 | 6/21/11 3:16 | 986 | 939 | 812 | N16W08 | | |
| eventT7 | 8/4/11 21:10 | 62.6 | 8/2/11 6:36 | 1015 | 951 | 883 | N14W15 | M1.4 | 1.48 |
| eventT8 | 8/5/11 17:23 | 52.1 | 8/3/11 13:17 | 1322 | 1062 | 1161 | N22W30 | M6.0 | 0.88 |
| eventT9 | 8/5/11 18:32 | 38.9 | 8/4/11 3:40 | 1709 | 1307 | 1945 | N19W36 | M9.3 | 0.38 |
| eventT10 | 9/9/11 11:49 | 81.4 | 9/6/11 2:24 | 513 | 494 | 521 | N14W07 | M5.3 | 0.50 |
| eventT11 | 9/17/11 3:05 | 75.1 | 9/14/11 0:00 | 577 | 534 | 467 | N22W03 | | |
| eventT12 | 11/12/11 5:10 | 63.6 | 11/9/11 13:36 | 1366 | 911 | 1210 | N22E44 | M1.1 | 1.13 |
| eventT13 | 1/22/12 5:18 | 62.9 | 1/19/12 14:25 | 1153 | 907 | 674 | N32E22 | | |
| eventT14 | 1/24/12 14:33 | 34.9 | 1/23/12 3:38 | 2002 | 1645 | 1245 | N29W20 | M8.7 | 0.93 |
| eventT15 | 2/26/12 21:07 | 65.3 | 2/24/12 3:46 | 779 | 623 | 678 | N25E28 | | |
| eventT16 | 3/8/12 10:53 | 33.3 | 3/7/12 1:36 | 2190 | 1866 | 1402 | N17E27 | X5.4 | 0.63 |
| eventT17 | 3/11/12 12:52 | 56.6 | 3/9/12 4:14 | 861 | 822 | 1176 | N17W03 | | |
| eventT18 | 3/12/12 8:45 | 39.1 | 3/10/12 17:40 | 1558 | 1361 | 1081 | N17W24 | M8.4 | 1.25 |
| eventT19 | 6/16/12 8:52 | 42.3 | 6/14/12 14:36 | 1207 | 1148 | 1317 | S17E06 | M1.9 | 3.07 |
| eventT20 | 7/14/12 17:27 | 48.6 | 7/12/12 16:49 | 1548 | 1502 | 1210 | S14W01 | X1.4 | 1.88 |

[a] This table presents the subset of 20 full halo events selected by *Gopalswamy et al*. [2013]. Column 1 presents the identifier of the event. Column 2 presents the date and time of shocks. The transit times are listed in column 3. The associated CMEs at the sun (date and time) are listed in column 4. The CME time refers to the first appearance of the CME in the STEREO/COR2 FOV. The solar source of the CME is identified as the heliographic coordinates of the eruption location observed in EUV images either from the Solar Dynamics Observatory (SDO) or STEREO. The speed measured in that STEREO/COR2 FOV in which the CME was closest to the limb is the radial speed (*Vrad*) of the CME is presented in column 5 as *Vcme*. The speed is the average speed within the COR2 FOV obtained by fitting a straight line to the height-time measurements. The earthward speed $Vcme_E$ (column 6) was obtained by applying a simple projection correction to the COR2 radial speed i.e., $Vcme_E = Vcme * cos\ (\alpha) * cos\ (\beta)$ km s$^{-1}$, where *a* and *b* are the heliolongitude and heliolatitude of the solar event location (SARM also uses this projection formula to calculate the earthward CME speed (see equation (2)); for this reason, we kept the same speed name, $Vcme_E$). Column 7 lists the earthward CME speeds in the ecliptic plane from STEREO A/B data ($Vcme1_E$) calculated by *Gopalswamy et al*. [2013] by making the CME height-time measurements at position angles 90º (STA) and 270º (STB), neglecting the solar B0 angle (the heliographic latitude of the ecliptic). The associated flare's location, peak and duration are listed in columns 8, 9 and 10, respectively.

Table 4 presents the absolute errors in two validation tests: Test A) by using the 20 events selected by *Gopalswamy et al.* [2013], and Test B) by using the 15 shock cases that are common between *Taktakishvili et al*. [2009] and *Gopalswamy et al*. [2013]. According to the summary of Table 4 (see last three rows), SARM's absolute errors are lower than ESA's absolute errors in all tests and statistical measures. The last three columns of Table 4 also show that ENLIL's absolute errors are the lowest.

**Table 4.** Validation test of SARM and comparison with ESA and ENLIL [a]

| | Test A [*Gopalswamy et al.*, 2013] | | | | Test B [*Taktakishvili et al.*, 2009] | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Using $Vcme_E$ | | Using $Vcme1_E$ | | Using $Vcme_E$ | | Using $Vcme1_E$ | | |
| | ESA | SARM | ESA | SARM | ESA | SARM | ESA | SARM | ENLIL |
| eventT1 | 11.7 | 12.3 | 11.0 | 12.02 | | | | | |
| eventT2 | 8.7 | 19.2 | 4.7 | 16.95 | 8.7 | 19.2 | 4.7 | 16.95 | 3.3 |
| eventT3 | 6.4 | 0.2 | 18.1 | 14.15 | 6.4 | 0.2 | 18.1 | 14.15 | 6.9 |
| eventT4 | 13.4 | 23.1 | 12.5 | 22.89 | 13.4 | 23.1 | 12.5 | 22.89 | 9.8 |
| eventT5 | 11.8 | 2.2 | 3.3 | 4.52 | 11.8 | 2.2 | 3.3 | 4.52 | 1.8 |
| eventT6 | 6.1 | 3.6 | 14.0 | 8.04 | 6.1 | 3.6 | 14.0 | 8.04 | 9.6 |
| eventT7 | 10.2 | 8.0 | 6.2 | 6.72 | | | | | |
| eventT8 | 5.5 | 3.1 | 9.9 | 4.55 | 5.5 | 3.1 | 9.9 | 4.55 | 0.4 |
| eventT9 | 2.3 | 9.2 | 16.9 | 1.53 | | | | | |
| eventT10 | 6.7 | 20.0 | 3.9 | 20.57 | 6.7 | 20.0 | 3.9 | 20.57 | 5.2 |
| eventT11 | 8.9 | 7.2 | 15.9 | 3.08 | 8.9 | 7.2 | 15.9 | 3.08 | 5.9 |
| eventT12 | 8.9 | 0.4 | 23.4 | 7.17 | 8.9 | 0.4 | 23.4 | 7.17 | 3.5 |
| eventT13 | 8.0 | 11.0 | 8.4 | 2.11 | | | | | |
| eventT14 | 7.6 | 5.0 | 3.9 | 9.26 | 7.6 | 5.0 | 3.9 | 9.26 | 0.5 |
| eventT15 | 10.4 | 2.1 | 5.7 | 4.77 | 10.4 | 2.1 | 5.7 | 4.77 | 1.0 |
| eventT16 | 10.3 | 0.0 | 14.5 | 3.52 | 10.3 | 0.0 | 14.5 | 3.52 | 0.8 |
| eventT17 | 3.7 | 2.0 | 15.1 | 12.46 | | | | | |
| eventT18 | 4.3 | 1.9 | 6.6 | 5.02 | 4.3 | 1.9 | 6.6 | 5.02 | 14.3 |
| eventT19 | 0.4 | 2.6 | 6.0 | 0.65 | 0.4 | 2.6 | 6.0 | 0.65 | 10.0 |
| eventT20 | 17.9 | 13.6 | 8.4 | 11.25 | 17.9 | 13.6 | 8.4 | 11.25 | 5.5 |
| Summary of results: | | | | | | | | | |
| - Mean: | 8.2 | 7.3 | 10.4 | 8.6 | 8.5 | 6.9 | 10.1 | 9.1 | 5.2 |
| - Mean w/o outliers: | 7.3 | 6.1 | 10.4 | 7.6 | 7.4 | 5.2 | 10.0 | 7.9 | 4.9 |
| - Median: | 8.4 | 4.3 | 9.2 | 6.9 | 8.7 | 3.1 | 8.4 | 7.2 | 5.2 |

a Table 4 presents the absolute errors in two validation tests: Test A) by using the 20 events selected by *Gopalswamy et al*. [2013], and Test B) by using the 15 shock cases that are common between *Taktakishvili et al*. [2009] and *Gopalswamy et al*. [2013]. Column 1 presents the shock identifier. Columns 2 to 5 present the mean absolute errors for Validation Test A using the ESA and SARM models with two CME speeds (i.e., $Vcme_E$ and $Vcme1_E$), and Columns 6 to 10 present the mean absolute errors of ESA, SARM and ENLIL (reported by *Taktakishvili et al*. [2009] for the Test B) using the mentioned speeds. The last three columns of Table 4 present the summary of the results. The first row of the summary presents the mean of the absolute errors of the respective columns. The second row of the summary presents the mean of absolute error without considering two outliers identified by *Gopalswamy et al.* [2013] because of CME-CME and CME-Coronal hole interactions (i.e., eventT4 and eventT20); the last row of the summary presents the median of the respective columns.

Figure 6 shows the distribution of errors of SARM predictions with CME data (using the speed estimation approach $Vcme_E$) and flare data with the dataset in *Gopalswamy et al.* [2013]. This figure also shows that the most frequent error interval is [2.5 h, 7.5 h], the MAE is 7.3 h, and the standard deviation of errors is 7.1 h. Note that the MAE for the calibration shock events at 1 AU (i.e., 7.0 h − see Figure 2b) is similar to the MAE with the validation shock events at 1 AU (i.e., 7.3 h − see Figure 6). These results also show that the overall performances of SARM with the calibration and validation shock cases were similar.

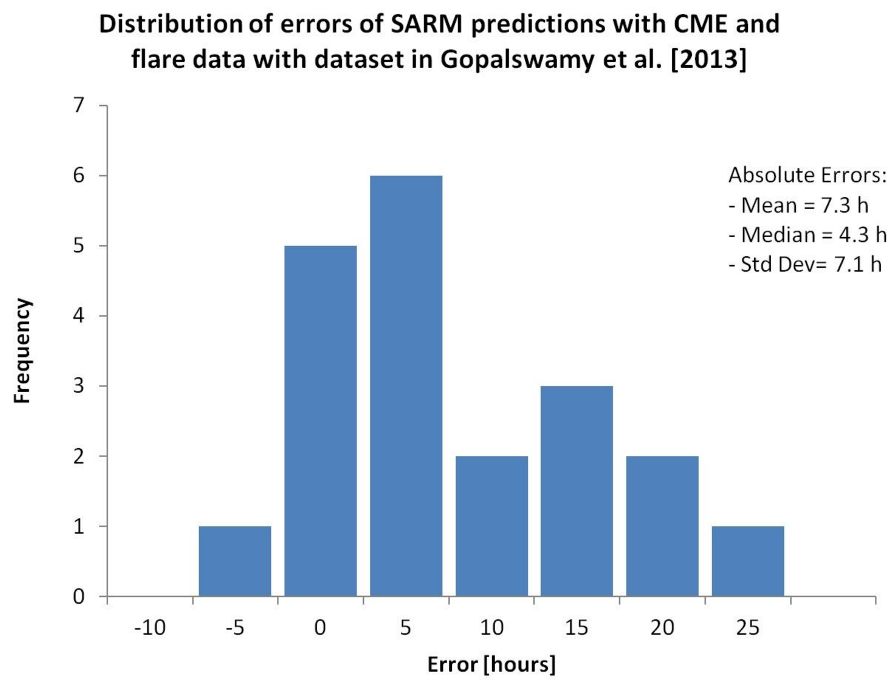



Figure 6. Distribution of errors of SARM predictions with CME data (using the speed estimation approach $Vcme_E$) and flare data with the test set using $Vcme_E$ in *Gopalswamy et al.* [2013]. These errors were extracted from the results presented in Table 4. This figure shows that the most frequent error interval is [2.5 h, 7.5 h], the MAE is 7.3 h, and the standard deviation of errors is 7.1 h.

From Table 4 we may say that ENLIL's median of absolute errors and MAE are lower than those errors yielded by the SARM and the ESA models, and that SARM's median of absolute errors and MAE are lower than those errors of the ESA models in tests A and B, and in every particular condition (i.e., with and without outliers). With the aim of studying the statistical support of the differences between the median of absolute errors of the three models shown in the last three columns in Table 4 (i.e., 8.4 h, 7.2 h, and 5.2 h for the ESA, SARM, and ENLIL models, respectively), we applied the Wilcoxon signed-rank test [*Wilcoxon*, 1945] with a statistical significance of 0.05. We conclude that there is statistical evidence of the difference between the aforementioned errors of ENLIL and ESA ($0.01 < p < 0.02$), i.e., the probability of obtaining similar results by chance is very low, so the conclusion that ENLIL yields a lower median of absolute errors is statistically supported; on the other hand, we also concluded that there is no statistical evidence of the difference between the medians of absolute errors of ENLIL and SARM ($0.05 < p < 0.1$), and between the aforementioned errors of SARM and ESA ($p > 0.2$); therefore, no statistically supported conclusions may be drawn with the later comparisons.

The mean absolute errors presented in section 3 may be verified by using data in Table 1 and Table 3 as input to the SARM model that is available in http://spaceweather.uma.es/sarm/index.html.

## 4. Conclusions

This paper has presented the SARM (Shock Arrival Model) tool's principles for the prediction of shock arrival times for distances from 0.72 AU to 8.7 AU. This drag-based model is the result of a comprehensive analysis of data, catalogs and observations of CMEs and flares from heliospheric observatories.

The SARM model is an empirical drag model that calculates the shock speed as a function of its location, and whose motion is subjected to a drag force, until a constant speed is reached. A dataset of 120 shocks observed from 1997 to 2010 was used to find the best coefficients that allow the absolute errors to be minimized. The coefficients were obtained by minimizing the normalized mean absolute errors, that is, those where the absolute error is divided by the shock distance in AUs.

The SARM model calculates the shock speed by using the differential equation (2) $dx/dt = Vdriver_x\, e^{-7x} + 0.42\, Vdriver_x + 330$ km s$^{-1}$, where $Vdriver_x$ is a function of the CME data (radial, earthward or plane-of-sky speeds) and the flare data (peak flux, duration, and location). This model may also be used with the CME data only or with flare data only.

For the prediction of shock arrivals at distances from 0.72 AU to 8.7 AU, the MAE error was 7.1 hours. This average was lower than the errors using the individual approaches: 8.9 h using the CME data only (radial, cone-model, plane-of-sky speeds) and 8.6 h using flare data only.

For 1 AU, the MAE error was 7.0 h. This average was lower than the errors using the individual approaches: 9.2 h using the CME data only (radial, cone-model, plane-of-sky speeds) and 8.4 h using flare data only. The best combination for 1 AU was found using both flare and true CME data (radial or cone-model-estimated speed), which obtained a MAE of 5.8 h. It is important to note the very satisfactory results of SARM in terms of the median: for 1 AU, for example, the median of absolute errors was 5.0 h during the calibration phase (see Figure 2). For all shocks the median of normalized absolute errors was also low (5.1 h).

SARM model was compared with the empirical ESA model [*Gopalswamy et al*., 2005a] and the numerical MHD-based model ENLIL [*Odstrcil et al*., 2004] with a dataset of 20 shocks that were not used during the calibration phase (see Table 4). These shocks were observed at 1 AU and associated with true CME data from 2010 to 2012 [*Taktakishvili et al*., 2009; *Gopalswamy et al*. 2013]. The ESA model obtained a MAE error of 8.16 h taking into account all shock events and using true CME speeds. SARM obtained a MAE of 7.3 h with these shock events and using both true CME data (with the same simple projection approach) and flare data. *Gopalswamy et al*. [2013] also found that the predictions of two of these events obtained large deviations due to complex CME-CME and CME-coronal hole interactions, considered as outliers. For this reason, the main result that they reported was obtained without considering the aforementioned complex interplanetary events. They reported a MAE of 7.3 h for the 18 shocks without the outliers. SARM obtained a MAE of 6.1 h for the same 18 shock events (using CME and flare data).

Table 4 also shows that the median of absolute errors obtained by the ENLIL, SARM and ESA models were 5.2 h, 7.2 h, and 8.4 h, respectively. In order to test the statistical support of these results, we used the Wilcoxon signed-rank test with a statistical significance of 0.05. We concluded that the ENLIL's median of absolute error is significantly lower than the same error of the ESA model, and that there is no statistical support of the differences between the median of absolute errors of ENLIL and SARM, nor between the median of absolute errors of SARM and ESA.

Although promising, the SARM model needs to be tested with real-time data for a large period of continuous operations. In an operational mode, predictors yield higher errors than those using historical data (e.g., *Zhao and Dryer* [2014] concluded that operational CME/shock arrival time prediction models for 1 AU generally yield mean absolute errors of 10 h for a large number of data events).

Although there is no physical relationship between flares and CME-driven shocks, this study shows that it is possible to predict shock arrival times using flare data alone, and that the best results are obtained when true CME speeds and flare data are used.

**Acknowledgements.**

The SARM model was funded by the Plan Propio de Investigación of Universidad de Málaga / Campus de Excelencia Internacional Andalucía Tech, and by the European Space Agency's Technology Research Programme (project SEPsFLAREs; Contract No. 4000109626/13/NL/AK). As mentioned in the text, the shock arrival predictions presented in this paper are available by processing the SARM model at http://spaceweather.uma.es/sarm/index.html with data in Table 1 and Table 3.

## References

Aran, A., D. Lario, B. Sanahuja, R. G. Marsden, M. Dryer, C. D. Fry, and S. McKenna-Lawlor (2007), Modeling and forecasting solar energetic particle events at Mars: The event on 6 March 1989, *Astronomy & Astrophysics*, 469 , 3, 1123-1134, doi:10.1051/0004-6361:20077233.

Beck, P., M. Latocha, S. Rollet, and G. Stehno (2005), TEPC reference measurements at aircraft altitudes during a solar storm, *Adv. Space Res.*, 16(9), 1627 – 1633, doi:10.1016/j.asr.2005.05.035.

Cargill, P.J. (2004), On the aerodynamic drag force acting on interplanetary coronal mass ejections, *Solar Physics*, 221, 135–149, doi:10.1023/B:SOLA.0000033366.10725.a2.

Caroubalos, C. (1964), Contribution a l'etude de l'activité Solaire en relation avec ses effets geophysiques, *Annales d'Astrophysique*, 5.

Chao, J. K., and R. P. Lepping (1974), A correlative study of SSC's, interplanetary shocks, and solar activity, *J. Geophys. Res*., 79, 1799–1807.

Chen, J., and Garren, D. A. (1993), Interplanetary magnetic clouds: Topology and driving mechanism, Geophys. Res. Lett., 20, 2319, doi: dx.doi.org/10.1029/93GL02426.

Chen, J., and V. Kunkel (2010), Temporal and physical connection between coronal mass ejections and flares, Astrophys. J., 717, 1105–1122, doi:10.1088/0004-637X/717/2/1105.

Cho, K.-S., Moon, Y.-J., Dryer, M., Fry, C. D., Park, Y.-D., and Kim, K.-S. (2003), A statistical comparison of interplanetary shock and CME propagation models, *Journal of Geophysical Research*, 108.

Conrad, B. (2002), *Differential Equations A Systems Approach*, Prentice Hall, 2002, ISBN 978-0130460264.

Dryer, M., Z. Smith, C. D. Fry, W. Sun, C. S. Deehr, and S.-I. Akasofu (2004), Real-time shock arrival predictions during the ‘‘Halloween 2003 epoch’’, *Space Weather*, 2, S09001, doi:10.1029/2004SW000087.

Falkenberg, T. V., A. Taktakishvili, A. Pulkkinen, S. Vennerstrom, D. Odstrcil, D. Brain, G. Delory, and D. Mitchell (2011), Evaluating predictions of ICME arrival at Earth and Mars, *Space Weather*, 9, S00E12, doi:10.1029/2011SW000682.

Fleishman G. D., Toptygin, I. N. (2013), *Cosmic Electrodynamics, Electrodynamics and Magnetic Hydrodynamics of Cosmic Plasmas, Springer*, ISBN 978-1-4614-5782-4.

Fry, C. D., W. Sun, C. S. Deehr, M. Dryer, Z. Smith, S.-I. Akasofu, M. Tokumaru, and M. Kojima (2001), Improvements to the HAF solar wind model for space weather predictions, *Journal Geophysical Research*, 106, 20,985 – 21,001, doi:10.1029/2000JA000220.

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

García-Rigo, A., M. Núñez, R. Qahwaji, O. Ashamari, P. Jiggens, G. Pérez, M. Hernández-Pajares, and A. Hilgers (2016), Prediction and Warning system of SEP events and Solar Flares for risk estimation in space launch operations, *J. Space Weather and Space Climate*, 6, (A28), DOI: 10.1051/swsc/2016021.

Gonzalez-Esparza, J. A., and Bravo (1998), Two spacecraft observations of transient shocks and ejecta in the interplanetary medium, *Journal of Geophysical Research*, 103, A12, 29643—29650.

Gopalswamy, N., A. Lara, P. K. Manoharan, R. A. Howard (2005a), An empirical model to predict the 1-AU arrival of interplanetary shocks, *Advances in Space Research*, 36, 2289-2294.

Gopalswamy, N., H. Xie, S. Yashiro, I. Usoskin (2005b), Coronal Mass Ejections and Ground Level Enhancements, *Proceedings of the 29th International Cosmic Ray Conference - Pune*, 00, pp101-104.

Gopalswamy, N., P. Mäkelä, H. Xie, and S. Yashiro (2013), Testing the empirical shock arrival model using quadrature observations, *Space Weather*, 11, 661–669, doi:10.1002/2013SW000945.

Herman, R. L. (2013), *A Course in Mathematical Methods for Physicists*, Taylor & Francis, ISBN 978-1466584679.

Jain, R., M. Aggarwal, and P. Kulkarni (2010), Relationship between CME dynamics and solar flare plasma, *Res. Astron. Astrophys*., 10, 473.

Jian, L. (2008), Radial evolution of large-scale solar wind structures, PhD thesis, UCLA, USA.

Lario, D., D. K. Haggerty, E. C. Roelof, S. J. Tappin, R. J. Forstth, J. T. Gosling (2001), Joint Ulysses and ACE Observations of a Solar Energetic Particle Event Related to the Same CME, *Space Science Reviews*, 97, pg 277-280.

Lario, D., S. Livi, E. C. Roelof, R. B. Decker, and S. M. Krimigis (2004), Heliospheric energetic particle observations by the Cassini spacecraft: Correlation with 1 AU observations, *Journal of Geophysical Research*, 109, A09S02, doi:10.1029/2003JA010107.

Lepri, S. T, J. Martin Laming, Cara E. Rakowski, and Rudolf von Steiger (2012), Spatially dependent heating and ionization in an ICME, *The Astrophysical Journal*, 760, 105.

Liu, H. and G. Qin (2012), Using soft X-ray observations to help the prediction of flare related interplanetary shocks arrival times at the Earth, *Journal of Geophysical Research*, 117, A04108, doi:10.1029/2011JA017220.

Krall, J., Chen, J., Duffin, R. T., Howard, R. A., and Thompson, B. J. (2001), Erupting Solar Magnetic Flux Ropes: Theory and Observation, *The Astrophysical Journal*, 562:1045-1057, doi: dx.doi.org/10.1086/323844.

McKenna-Lawlor, S. M. P., M. Dryer, , Z. Smith, K Kecskemety, C. D. Fry, W. Sun, C. S. Deehr, D. Berdichevsky, K. Kudela, and G. Zastenker, (2002), Arrival times of Flare/Halo CME associated shocks at the Earth: comparison of the predictions of three numerical models with these observations, *Annales Geophysicae*, 20, 7, 917-935, doi: 10.5194/angeo-20-917-2002.

McKenna-Lawlor, S., M. Dryer, M. D. Kartalev, Z. Smith, C. D. Fry, W. Sun, C. S. Deehr, K. Kecskemety, and K. Kudela (2006), Near Real-time Predictions of the Arrival at the Earth of Flare-generated Shocks during Solar Cycle 23, *Journal of Geophysical Research*, 111, A11103,doi:10.1029/2005JA011162.

Miller, J., C. Zeitlin, F. A. Cucinotta, L. Heilbronn, D. Stephens, and J. W. Wilson (2003), Benchmark studies of the effectiveness of structural and internal materials as radiation shielding for the International Space Station, *Radiation Research*, 159, 381–390, doi:10.1667/0033-7587(2003)159[0381:BSOTEO]2.0.CO;2.

Michałek, G., N. Gopalswamy, A. Lara, and S. Yashiro (2003), A new method for estimating widths, velocities, and source location of halo coronal mass ejections, *Astrophys. J.*, 584, 472–478, doi:10.1086/345526.

Möstl, C. et al. (2012), Multi-point Shock and Flux Rope Analysis of Multiple Interplanetary Coronal Mass Ejections around 2010 August 1 in the Inner Heliosphere, *The Astrophysical Journal*, 758, 10, doi:10.1088/0004-637X/758/1/10.

Nieves-Chinchilla, T., Colaninno, R., Vourlidas, A., et al. (2012), Remote and in situ observations of an unusual Earth-directed coronal mass ejection from multiple viewpoints, *Journal of Geophysical Research*, 117, A06106.

Nieves-Chinchilla, T., Vourlidas, A., Stenborg, G., et al. (2013), Inner Heliospheric Evolution of a "Stealth" CME derived from Multi-view imaging and Multipoint in-situ Observation. I. Propagation to 1 AU, *The Astrophysical Journal*, 779, 55.

Núñez, M. (2011), Predicting solar energetic proton events (E > 10 MeV), *Space Weather*, 9, S07003, doi:10.1029/2010SW000640.

Núñez, M. (2015), Real-time prediction of the occurrence and intensity of the first hours of > 100 MeV solar energetic proton events, *Space Weather*, 13, doi:10.1002/2015SW001256.

Odstrcil, D., V. J. Pizzo, J. A. Linker, P. Riley, R. Lionello, Z. Mikic, and J. G. Luhmann (2004), Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes, *Journal of Atmospheric and Solar-Terrestrial Physics*, 66, 1311-1326, doi:10.1016/j.jastp.2004.04.007.

Pick M. and N. Vilmer (2008), Sixty-five years of solar radioastronomy: flares, coronal mass ejections and Sun–Earth connection, *Astron Astrophys Rev*, 16, 1, pp 1-153.

Reeves, K. K., & Moats, S. J. 2010, Relating coronal mass ejection kinematics and thermal energy release to flare emissions using a model of solar eruptions, *ApJ*, 712, 429.

Riley, P., J. A. Linker, Z. Mikić, D. Odstrcil, T. H. Zurbuchen, D. Lario, and R. P. Lepping (2003), Using an MHD simulation to interpret the global context of a coronal mass ejection observed by two spacecraft, *Journal of Geophysical Research*, 108, 1272, doi:10.1029/2002JA009760, A7.

Richardson, J. D., Y. Liu, C. Wang, D. J. McComas, E. C. Stone, A. C. Cummings, L. F. Burlaga, M. H. Acuna N. F. Ness (2006), Source and consequences of a large shock near 79 AU, *Geophysical Research Letters*, 33, L23107, doi:10.1029/2006GL027983.

Skoug, R. M., W. C. Feldman, J. T. Gosling, D. J. McComas, D. B. Reisenfeld, C. W. Smith, R. P. Lepping, A. Balogh (2000), *Journal of Geophysical Research*, 105, doi:10.1029/2000JA000095.

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

Smith, Z. K., M. Dryer, S. M. P. McKenna-Lawlor, C. D. Fry, C. S. Deehr, and W. Sun (2009), Operational validation of HAF's predictions of interplanetary shock arrivals at Earth: Declining phase of Solar Cycle 23, *Journal of Geophysical Research*, 114, 5, A05106, doi:10.1029/2008JA013836.

Steed, K., Long, D. Davies, J. Walsh, A. P. Lapenta, G. (2012), The origins and heliospheric evolution of CMEs on 7 and 14 August 2010 originating from the same solar source region, European Space Weather Week (Session 4a), Brussels, Belgium, 5 - 9 November 2012.

Taktakishvili, A., Kuznetsova, M., MacNeice, P., Hesse, M., Rastätter, L., Pulkkinen, A., Chulaki, A., and Odstrcil, D. (2009), Validation of the coronal mass ejection predictions at the Earth orbit estimated by ENLIL heliosphere cone model, *Space Weather*, 7, S03004, doi: 10.1029/2008SW000448.

Taktakishvili, A., A. Pulkkinen, P. MacNeice, M. Kuznetsova, M. Hesse, and D. Odstrcil (2011), Modeling of coronal mass ejections that caused particularly large geomagnetic storms using ENLIL heliosphere cone model, *Space Weather*, 9, S06002, doi:10.1029/2010SW000642.

Von Steiger, R., J. D. Richardson (2006), ICMEs in the Outer Heliosphere and at High Latitudes: An Introduction, Space Science Reviews, 123, Issue 1, pp 111-126, DOI. 10.1007/s11214-006-9015-z

Vršnak, B. and Gopalswamy, N. (2002), Influence of the aerodynamic drag on the motion of interplanetary ejecta. *J. Geophys. Res*., 107, SSH 2-1, pp. 1019.

Vršnak, B., T. Žic, T. V. Falkenberg, C. Möstl, S. Vennerstrom, and D. Vrbanec (2010), The role of aerodynamic drag in propagation of interplanetary coronal mass ejections, *Astron. Astrophys*., 512, A43, doi:10.1051/0004-6361/200913482.

Wang, C. and J. D. Richardson (2003), The interaction and evolution of interplanetary shocks from 1 to beyond 60 AU. Proceedings of the Tenth International Solar Wind Conference, Pisa, Italy, M. Velli, R. Bruno, F. Malara eds., *AIP Conference Proceedings*, 679, 725-778.

Wang, C., J. D. Richardson, and L. Burlaga (2001), Propagation of the Bastille Day 2000 CME shock in the outer heliosphere, *Solar Physics*, 204, 413-423.

Webb, D. F., C. Möstl, B. V. Jackson, M. M. Bisi, T. A. Howard, T. Mulligan, E. A. Jensen, L. K. Jian, J. A. Davies, C. A. de Koning, Y. Liu, M. Temmer, J. M. Clover, C. J. Farrugia, R. A. Harrison, N. Nitta, D. Odstrcil, S. J. Tappin, and H.-S. Yu (2013), Heliospheric Imaging of 3D Density Structures During the Multiple Coronal Mass Ejections of Late July to Early August 2010, *Solar Physics*, Volume 285, Issue 1-2, pp 317-348, http://dx.doi.org/10.1007/s11207-013-0260-5.

Wilcoxon, F. (1945), Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83. doi: 10.2307/3001968.

Wilson, J. W., C. J. Mertens, P. Goldhagan, W. Friedberg, G. De Angelis, J. M. Clem, K. Copeland, H. B. Bidasaria (2005), Atmospheric ionizing radiation and human exposure, Tech. Rep. NASA/TP-2005-213935, NASA, Washington, DC.

Xie, H., L. Ofman, and G. Lawrence (2004), Cone model for halo CMEs: Application to space weather forecasting, *J. Geophys. Res*., 109, A03109, doi:10.1029/2003JA010226.

Xie H., N. Gopalswamy, L. Ofman, O. C. St. Cyr, G. Michalek, A. Lara, and S. Yashiro (2006), Improved input to the empirical coronal mass ejection (CME) driven shock arrival model from CME cone models, *Space Weather*, 4, S10002, doi:10.1029/2006SW000227.

Xie, H., O. C. St. Cyr, N. Gopalswamy, D. Odstrcil, and H. Cremade (2013), Understanding shock dynamics in the inner heliosphere with modeling and type II radio data: A statistical study, *Journal of Geophysical Research - Space Physics*, 118, 4711–4723, doi:10.1002/jgra.50444.

Yashiro, S., and N. Gopalswamy (2009), Statistical relationship between solar flares and coronal mass ejections, in IAU Symposium, vol. 257, edited by N. Gopalswamy and D. F. Webb, pp. 233–243, Universal Heliophysical Processes, Cambridge Univ. Press, London.

Zhao, X. H and X. S. Feng (2014), Shock Propagation Model version 2 and its application in predicting the arrivals at Earth of interplanetary shocks during Solar Cycle 23, *Journal of Geophysical Research: Space Physics*, 119, 1–10, doi:10.1002/2012JA018503.

Zhao, X., and M. Dryer (2014), Current status of CME/shock arrival time prediction, *Space Weather*, 12, 448–469, doi:10.1002/2014SW001060.

Supporting Information S1

AGU PUBLICATIONS

**Space Weather**
**Supporting information for**

**Prediction of shock arrival times from CME and flare data**

Marlon Núñez[1*], Teresa Nieves-Chinchilla[2], and Antti Pulkkinen[2]

[1]Universidad de Málaga, Málaga, Spain
[2] Heliophysics Science Division. GSFC/NASA, Greenbelt, MD USA

The SARM calibration process, described in this document, is an iterative data-driven three-step analysis. This document explains how to get the first estimates of the speed $V_a$ in Eq. 1 by using shocks observed at distances > 1 AU, and how, starting from these, to refine the estimate of all parameters in Eq. 1 by using all shocks in Table 1.

In section S.1, we assume that at distances > 1 AU, $V_a$ can be approximated by the distance Sun-observer divided by the observed travel time. To this end Section S1 uses coronagraphic measurements of CME speeds. Section S2 refines $V_a$ by using a proxy inferred from soft X-ray observations. In section S.3 we present that the mean value of the two is most appropriate, with the further advantage that a CME speed can be guessed even when the data coverage is incomplete. Starting from these first guesses, section S.3. uses all shocks in Table 1 (i.e.,120 shocks observed from 0.72 to 8.9 AU) to encounter the drag coefficient $k$ and refine $V_a$ (eq. S.3).

## S.1. Correlation between IP shock transit speed and CME data alone

In the first analysis step, explained in detail in this section, we explore the correlation between CME data alone and the IP shock transit speeds for distances between 1.3 AU (by Mars) and 8.7 AU (by Saturn). We found that the IP shock transit speed is easily obtained with historical data (i.e., the radial distance of the spacecraft from the sun divided by the observed transit time); therefore, the purpose of this step has been to obtain a predictor of the IP shock transit speed from CME data alone for the aforementioned distances. In this section we obtain a linear regression formula that allow us to make rough predictions of the IP asymptotic speed $V_a$ in equation (1) from CME data alone.

In order to avoid the necessity of determining the drag coefficient in this first iteration of the model design, we use shock data associated with shocks observed at *long* distances (i.e., > 1.3 AU). Figure S.1 presents the events in Table 1 observed at distances > 1.3 AU in terms of the observed IP shock transit speed $\bar{v}$ and $V_{\mathrm{CME}x}$ (which is calculated by using the approaches presented in section 2.1). Note that we carry out this correlation with the purpose of finding a first approximation of the IP asymptotic speed from CME data alone that will be refined in sections S2 and S3. The IP shock transit speed $\bar{v}$ is calculated as $D/tt$, where $D$ is the distance (i.e., column 2 of Table 1) at which the shock was observed, and $tt$ is the transit time of the shock (i.e., column 4 of Table 1).

The linear dependence between the IP shock transit speed and the CME initial speed alone for shocks in Table 1 whose distances are greater than 1.3 AU is shown in Figure S.1. The solid line represents a linear formula whose coefficients were found by minimizing the mean of absolute errors. The found regression equation is to estimate the observed IP shock transit speeds:

$$\bar{v} = 0.39\, V_{\mathrm{CME}x} + 322 \text{ (km s}^{-1}\text{)} \quad \text{for distances between 1.3 and 8.7 AU} \qquad \text{(S.1)}$$

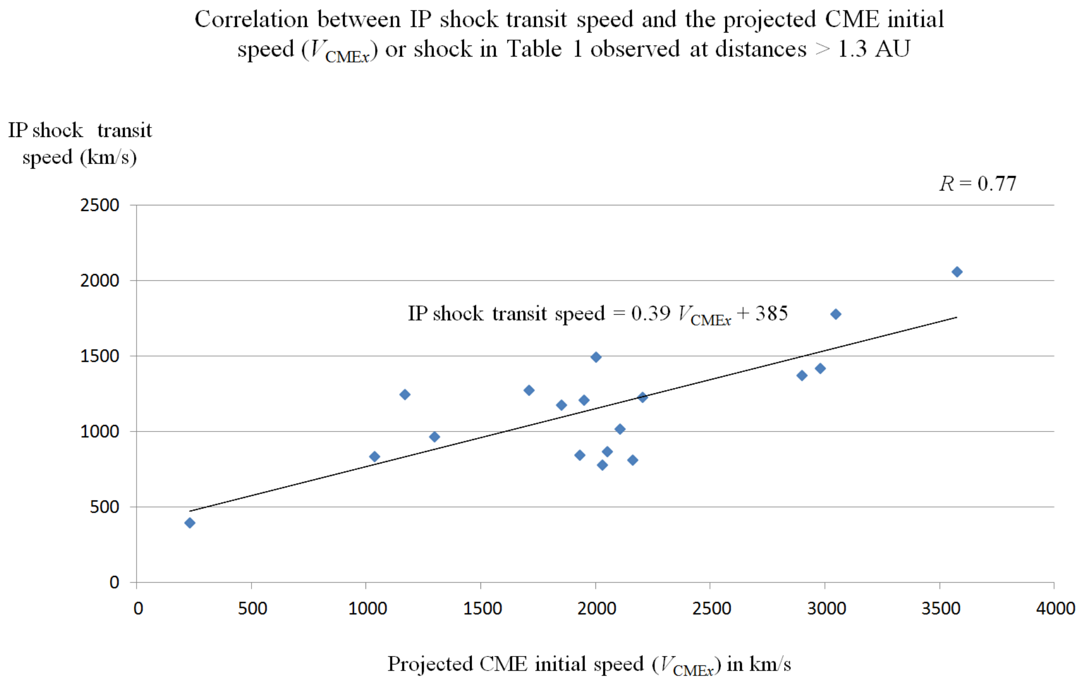



**Figure S.1**. Linear dependence between the observed IP shock transit speed and the projected CME initial speed ($V_{\mathrm{CME}x}$) alone for shocks in Table 1 whose distances are greater than 1.3 AU. The solid line shows the regression line calculated by minimizing the mean of absolute distances of each point to the solid line.

Although the correlation coefficient of Figure S.1 is not high ($R$ = 0.77), the associated regression line was useful to find $\bar{v}$, which is a first rough approximation of the asymptotical shock speed $V_a$ at large distances, from $V_{\mathrm{CME}}$ speeds only. Note that the regression line in Figure A.1 cannot be directly validated, because $V_a$ does not correspond to any observed measure at the spacecraft; therefore, what will be important at the end of the whole calibration process is the validation in terms of the absolute value of errors of shock arrival time predictions with all events in Table 1 and with new shock events (see section 3).

## S.2. Correlation between IP shock transit speed and flare data alone

The purpose of the second analysis step, explained in this section, is to obtain a radial CME speed $saV_{\mathrm{CME}}$ from flare data alone (duration, peak, and location) for distances between 1.3 AU and 8.7 AU. We finalize this section by refining the linear regression equation (S.1) that allows us to predict the IP asymptotic speed $V_a$ in equation (1) from CME and flare data.

Firstly, we want to predict the radial CME initial speed from the soft X-ray flux associated to the same solar-activity process. Therefore, in this section we want to predict the radial CME speed (i.e., $V_{\mathrm{CME}}$) from flare data. To differentiate the observed $V_{\mathrm{CME}}$ and the predicted CME initial speed, we refer to the latter as $saV_{\mathrm{CME}}$ to emphasize that it is predicted from the solar-activity data. We have to study the correlation between the $saV_{\mathrm{CME}}$ term and a flare-related term in the context of the prediction of the observed IP shock transit speed. In this context, we want to correlate the observed IP shock transit speed with $saV_{\mathrm{CME}}$. With the purpose of making a linear regression correlation, each shock event in Table 1 (for distances > 1.3 AU) should be used to construct an analysis instance in a 2D graph, in which an X-ray-related input is used to predict a $saV_{\mathrm{CME}}$-related output, as follows:

- There have been studies in which soft X-ray fluxes have been used to predict shock-related events: On the other hand, *Smith et al*. [1994] showed that the product of the soft X-ray flare peak and duration is proportional to the energy released by flares. Later, *Liu and Qin* [2012] used this product as a proxy to predict whether or not a IP shock will hit the Earth; therefore, the product of flare duration with flare peak seems to be a proxy that needs to be studied. Besides, these data are easy to consult (e.g., they are found in the NOAA's edited event list). The time-integration of the X-ray flux has also been used by the UMASEP system to forecast the >10 MeV and >100 MeV integral proton flux intensity of the prompt component of well-connected SEP events [*Núñez,* 2011, 2015]. Observationally, we know that X-class flares are correlated with high CME speeds and C-class flares with low CME speeds. These observations led us to study the linear correlation between CME speeds with the $log_{10}$ of the aforementioned product (i.e., flare duration *FD* times the flare peak *FP*). We used the NOAA/SWPC's edited event list as a reference to obtain the flare duration (abbreviated as *FDswpc)* and the flare intensity peak (abbreviated as *PFswpc*).
- Regarding the $saV_{\mathrm{CME}}$-related output, we could use $V_{CME}$ in Table 1 as the target prediction output; however, it is not advisable given the fact that we have a more accurate goal to predict: the *observed* IP transit speed $\bar{v}$, which is easy to obtain (see column 3 in Table 1) and is accurate. Since we know that equation (S.1) presents a formula that predicts $\bar{v}$ from $V_{\mathrm{CME}}$, $saV_{\mathrm{CME}}$ may be used in equation (S.1) to more accurately predict $\bar{v}$; that is, $\bar{v}$ = 0.39 $saV_{\mathrm{CME}}$ $cos(\alpha)$ cos $(\beta)$ + 322, for distances > 1.3 AU (where $\alpha$ and $\beta$ are the longitude and latitude of the associated flare from the spacecraft's viewpoint) by isolating $saV_{\mathrm{CME}}$ in this formula, the term $saV_{\mathrm{CME}} = (\bar{v} - 322) / (0.39 \; cos\,(\alpha)\, cos\,(\beta))$. Note that $V_{\mathrm{CME}x}$ in the original equation (S.1) depends on the radial CME speed; therefore, we want to predict the radial $saV_{\mathrm{CME}}$ that, used alone (without CME data) and with the corresponding projection i.e., $saV_{\mathrm{CME}x}$, may provoke that shocks arrive at the observe transit time.

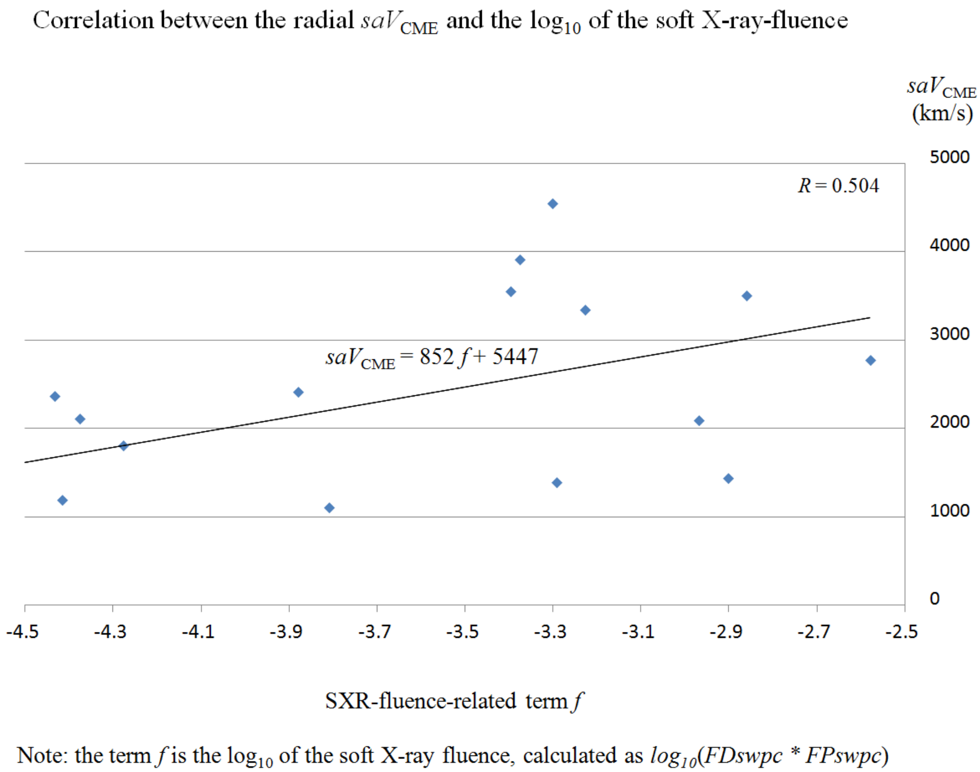



**Figure S.2.** Regression analysis for predicting the solar-activity-calculated CME initial speed ($saV_{\mathrm{CME}}$) from X-ray data for predicting the observed IP shock transit speed at distances between 1.3 to 8.7 AU. The coefficients of the regression line are discovered by minimizing the normalized mean absolute error.

Figure S.2 presents shocks at distances from 1.3 to 8.7 AU for each shock in terms of the aforementioned input and output terms, and the corresponding regression line. The regression line $saV_{\mathrm{CME}} = 852{*}f + 5447$ km s$^{-1}$ (where $f = \log_{10}(PFswpc * FDswpc)$) represents a linear formula whose linear coefficients were obtained by minimizing the normalized mean absolute error (see Figure S.2). Although the correlation coefficient of Figure S.2 is not high ($R = 0.504$), the associated regression line was useful to find a first rough approximation of the asymptotical shock speed $V_a$ at large distances in terms of flare data only, that will be refined in Section S.3. If we use the regression line shown in Figure S.2 in equation (S.1) to predict the IP shock transit speed (i.e., $\bar{v} = 0.39\ saV_{\mathrm{CME}x} + 322$), and calculate the observed IP shock arrival time (i.e., the spacecraft distance divided by the transit speed $\bar{v}$), the normalized mean absolute error of time arrival predictions is 10.5 h for the shocks in Table 1 detected at distances of between 1.3 and 8.7 AU, which means this second approximation (i.e. prediction of $\bar{v}$ from $saV_{\mathrm{CME}x}$) yields acceptable results.

## S.3. Estimating the IP shock speed from CME and/or flare data

The third analysis step uses all shock events in Table 1 to refine all the model coefficients (e.g., the drag coefficient $k$ and all the linear regression coefficients obtained in section S.1 and section S.2) by minimizing the mean absolute error of arrival time predictions.

In section S.1 and S.2 we found an estimation of $V_a$ as a function of $V_{\mathrm{CME}}$ (i.e. from coronographs), and from $saV_{\mathrm{CME}x}$ (i.e., a proxy inferred from soft X-ray observations). In this section we present that the mean value of the two is most appropriate, because, if no CME observations are available, the proxy $saV_{\mathrm{CME}x}$ may be used; that is $V_a$ may be approximated as average between ((0.39 $V_{\mathrm{CME}x}$ +322) and ((0.39 $saV_{\mathrm{CME}x}$ +322). Another way to express this equation is the equation (S.2):

$$\bar{v} = 0.39\, V_{DRIVERx} + 322 \quad \text{for shock distances between 1.3 and 8.7 AU} \qquad \text{(S.2)}$$

where $V_{DRIVERx}$ is the average between $V_{\mathrm{CME}x}$ and $saV_{\mathrm{CME}x}$, and $\bar{v}$ is the IP shock transit speed to be predicted. The term $saV_{\mathrm{CME}x}$ is calculated as the cosine projection of the radial term ($852 * \log_{10}(PFswpc * FDswpc) + 5447$) in the sun-spacecraft axis. If we replace $\bar{v}$ (i.e., equation (S.2)) in equation (1), we obtain $V_a$ in equation (S.3) to calculate the instantaneous IP shock speed. Instead of presenting a refined equation with the coefficients found from flare data alone and CME data alone, we present it in equation (1) in terms of its five coefficients (i.e., *a*, *b*, *c*, *d* and *k*), because we want to emphasize that the coefficients found are approximations that need to be tuned by using both CME and flare data with all shocks in Table 1.

$$V_{shock}(x) = \frac{dx}{dt} = V_{DRIVERx}\; e^{-k\,x} + \overbrace{a\; V_{DRIVERx} + b}^{V_a \;\; \text{(see equation (S.2))}} \qquad \text{(S.3)}$$

where *a* and *b* are the linear coefficients to calculate the IP shock transit speed, which is a first approximation of $V_a$, the asymptotic speed in equation (1).

In order to estimate the drag coefficient *k* of equation (S.3) we ran the Runge Kutta 4th order numerical method with different values of *k* for the 120 shock events in Table 1 with all CME and flare data, and by using the coefficients *a* and *b* in equation (S.2). We found a first approximation of the drag coefficient *k* that yielded the lowest mean absolute error with equation (S.3) taking into account the above analysis was -6.18.

The final optimization process also requires us to optimize the statistical finding reported by *Michałek et al*. [2003] that found that the actual earthward speeds $V_{\mathrm{CME}e}$ are 20% higher than plane-of-sky speeds, $V_{POS}$. They also reported that halo CMEs originating close to the Sun center, subjected to the largest projection effects, were not included in their results. Since we have 48 shock events (including CMEs close to the Sun center), we decided to refine the aforementioned percentage. We empirically found by using the refinement process presented in this section, that using equation (S.3) a ratio $V_{\mathrm{CME}e}/V_{POS}$ of 1.26 (instead of 1.2) minimizes the normalized mean absolute errors of SARM's arrival predictions using the shock events in Table 1 that use plane-of-sky speeds.

Now we need to do the final refinement of all coefficients in equation (S.3) by using both CME and flare data. This refinement is done by minimizing the normalized mean absolute error of the predictions on all shock data in Table 1 by using an optimization via a Coarse-to-Fine approach; that is, for each iterative refinement step, each shock in Table 1 needs to be simulated with the differential equation (S.3) configured with slightly different coefficients than those obtained in the previous refinement step. For each $^{simulated}$ shock propagation, single absolute errors (from observed shock arrivals in Table 1) are normalized by dividing them by the corresponding distance (in AU), and averaged. The best sets of coefficients are taken as input of the next refinement step. This iterative optimization approach finalizes when no more improvement is achieved. The resulting refined coefficients were: $a$ = 0.42, $b$ = 330, $c$ = 1015, $d$ = 5500, and $k$ = -7. We may take the common factor $V_{DRIVERx}$ out and transform equation (S.3) into equation (S.4), which corresponds to the SARM equation (see equation (2)) for calculating the instantaneous IP shock speed:

$$\frac{dx}{dt} = V_{DRIVERx}\, e^{-7x} + 0.42\, V_{DRIVERx} + 330 \qquad km\ s^{-1} \tag{S.4}$$

where

- $x$ is the heliospheric distance from the sun to the IP shock
- $V_{DRIVERx}$ is the *Average* ($V_{\mathrm{CME}x}$, $saV_{\mathrm{CME}x}$). That is, if no flare data are available, $V_{DRIVERx} = V_{\mathrm{CME}x}$; if no CME data are available, $V_{\mathrm{DRIVER}x} = saV_{\mathrm{CME}x}$
- $V_{\mathrm{CME}x}$ is the radial CME speed $V_{\mathrm{CME}}$ projected in sun-spacecraft axis; that is, $V_{\mathrm{CME}x} = V_{\mathrm{CME}}$ $cos(\alpha) * cos\ (\beta)$, where $\alpha$ and $\beta$ the longitude and latitude of the associated flare from the spacecraft's viewpoint. If radial CME speed is not available, the cone-model speed $V_E$ may be used; that is, $V_{\mathrm{CME}x} = V_E$. For more information about how to obtain $V_{\mathrm{CME}x}$, see section 2.1.
- $saV_{\mathrm{CME}}$ is calculated as 1015 $\log_{10}$ (*PFswpc x FDswpc*)+5500, where *PFswpc* and *FDswpc* are the flare peak and the flare duration, using the NOAA/SWPC data.