Solar Phys (2017) 292:180
DOI 10.1007/s11207-017-1212-2

Effective Acceleration Model for the Arrival Time of Interplanetary Shocks driven by Coronal Mass Ejections

Evangelos Paouris[1] · **Helen Mavromichalaki**[1]

Received: 22 August 2017 / Accepted: 16 November 2017 / Published online: 24 November 2017

Abstract In a previous work (Paouris and Mavromichalaki in *Solar Phys*. **292**, 30, 2017), we presented a total of 266 interplanetary coronal mass ejections (ICMEs) with as much information as possible. We developed a new empirical model for estimating the acceleration of these events in the interplanetary medium from this analysis. In this work, we present a new approach on the effective acceleration model (EAM) for predicting the arrival time of the shock that precedes a CME, using data of a total of 214 ICMEs. For the first time, the projection effects of the linear speed of CMEs are taken into account in this empirical model, which significantly improves the prediction of the arrival time of the shock. In particular, the mean value of the time difference between the observed time of the shock and the predicted time was equal to +3.03 hours with a mean absolute error (MAE) of 18.58 hours and a root mean squared error (RMSE) of 22.47 hours. After the improvement of this model, the mean value of the time difference is decreased to −0.28 hours with an MAE of 17.65 hours and an RMSE of 21.55 hours. This improved version was applied to a set of three recent Earth-directed CMEs reported in May, June, and July of 2017, and we compare our results with the values predicted by other related models.

Keywords Coronal mass ejections · Initiation and propagation · Coronal mass ejections · Interplanetary · Solar wind · Disturbances · Space weather

1. Introduction

Coronal mass ejections (CMEs) and their interplanetary counterparts inside the heliosphere known as interplanetary CMEs (ICMEs) are the main drivers of intense geomagnetic storms (Gosling, 1993; Zhang *et al.*, 2007; Paouris and Mavromichalaki, 2017). It is of great interest to predict the arrival of these events at Earth with the best possible accuracy. In recent years, an effort to develop models for estimating the arrival time of CMEs at Earth

✉ E. Paouris
evpaouris@phys.uoa.gr

H. Mavromichalaki
emavromi@phys.uoa.gr

[1] Faculty of Physics, National and Kapodistrian University of Athens, 15784 Athens, Greece

has been in full progress. These models, such as the ENLIL + cone model (Odstrcil, 2003; Odstrcil *et al.*, 2004), the drag-based model (Vrsnak *et al.*, 2013), the empirical shock-arrival model (Gopalswamy *et al.*, 2001, 2005), the enhanced drag-based model (Hess and Zhang, 2015), and others (see https://swrc.gsfc.nasa.gov/main/cmemodels), provide good results in general.

In this work we present a new model, which is an improvement of the effective acceleration model first introduced by Paouris and Mavromichalaki (2017; hereafter Paper I). In particular, in Paper I, a new database of interplanetary coronal mass ejections (ICMEs) with as much information as possible was presented in the form of a table of 266 events. In this table information was provided about the CME and the associated solar flare, the preconditions of the ambient solar wind (magnetic fields, speed, density, temperature, *etc*.) before the arrival of the CME at 1 AU, and the conditions of the solar wind in the sheath and the main part of the CME. Finally, the geomagnetic effects produced by the arrival of the CME at Earth, which provides the values of the Ap and Dst indices, were presented. Using *in situ* observations from the *Advanced Composition Explorer* (ACE) spacecraft and observations from the *Solar and Heliospheric Observatory/Large Angle and Spectrometric Coronagraph* (SOHO/LASCO), we were able to estimate the duration (τ) for each event from the beginning on the Sun up to the arrival at Earth. The initial speed (u) of each event available in the LASCO CME list, the distance from the Sun (d), the time duration (τ), and the *in situ* speed of the CME (during the main part of CME) at 1 AU (υ) based on ACE data were used to calculate the effective acceleration (α) of each event. The best relation between the acceleration (α) and the initial speed of the CME from the LASCO list (u) was a second-order equation of the form

$$\alpha = 1.45\left[\mathrm{m\,s^{-2}}\right] - 27.40 \times 10^{-4}\left[10^{-3}\ \mathrm{s^{-1}}\right]u - 1.28 \times 10^{-6}\left[10^{-6}\ \mathrm{m^{-1}}\right]u^2, \qquad (1)$$

where α is in $\mathrm{m\,s^{-2}}$ and u is in $\mathrm{m\,s^{-1}}$. The acceleration against the initial speed has a very good cross-correlation coefficient with $r = 0.98$.

In Section 2 we present the method we used to estimate the arrival time of an ICME shock. In Section 3 we establish the improvement of the model by taking for the first time the projection effects into account, as the linear speed of the CME is a projection on the plane of the sky (Gopalswamy *et al.*, 2000), and we also provide the statistical analysis for 214 events from Paper I when possible. Finally, in Section 4 we provide the results of the model for a case study of three Earth-directed CMEs that occurred in May, June, and July of 2017, as well as comparison with other related models. In the conclusions section we discuss the results of the current work, similar results by other researchers, and potential improvements of our model.

2. Effective Acceleration of the ICME Shock

The previous model based on Equation 1 is a useful tool for space weather applications and gives the possibility to predict the arrival of a CME (the main part of a CME, not the preceding shock) at 1 AU, although it is equally important to predict the arrival time of the shock that is driven by the CME, as many geomagnetic storms occur with the arrival of the shock. In Paper I we showed that almost 47.7% of the examined events had a maximum three-hour value of the geomagnetic Ap-index during the sheaths (the space between the shock and the main part of the CME). In order to recalculate the acceleration or deceleration of each event, we used variables such as the duration of the event from the onset on the Sun

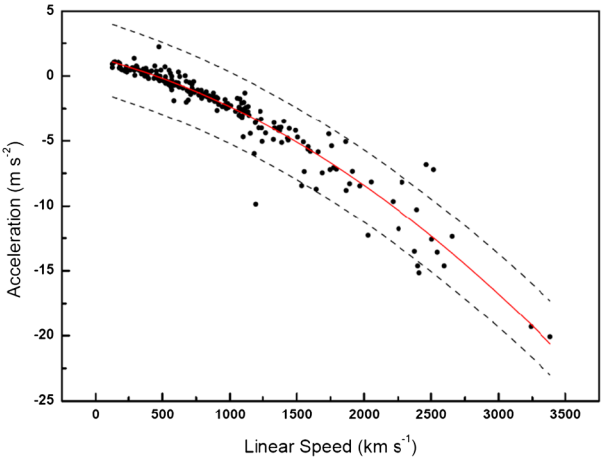

Figure 1 Scatter plot of the effective acceleration of the preceding shock driven by a CME against the CME linear speed (plane-of-the-sky expansion). The *solid line* is a second-order polynomial fit with $r = 0.981$ from Equation 3 when the seven points outside of the 99% prediction levels (*black dashed lines*) are excluded.

up to the arrival at 1 AU using ACE data (τ), and the initial speed of the CME (u_0) replacing only the final speed of the CME with the speed of solar wind inside the sheath. This first approach in order to develop a model of the effective acceleration (hereafter EAMv1) of the shock leads to a relation slightly different from Equation 1:

$$\alpha = 1.43138\left[\mathrm{m\,s^{-2}}\right] - 26.10 \times 10^{-4}\left[10^{-3}\ \mathrm{s^{-1}}\right]u_0 - 1.15834 \times 10^{-6}\left[10^{-6}\ \mathrm{m^{-1}}\right]u_0^2. \quad (2)$$

Furthermore, we continued to study this model by excluding from our sample all the CMEs that were not associated with a preceding shock. In a total of 266 ICMEs, 222 were associated with a shock, and of these, 214 events had no data gaps in linear speed or in the speed of the solar wind inside the sheath. From this sample, the previous relation took the final form:

$$\alpha = 1.41392\left[\mathrm{m\,s^{-2}}\right] - 26.30 \times 10^{-4}\left[10^{-3}\ \mathrm{s^{-1}}\right]u_0 - 1.14717 \times 10^{-6}\left[10^{-6}\ \mathrm{m^{-1}}\right]u_0^2. \quad (3)$$

The initial speed of each CME (u_0) on Sun and the onset date and time are the only inputs for estimating the acceleration or deceleration of the CME shock in interplanetary space. The scatter plot of the acceleration (α) against the initial speed of the CME (u_0) is presented in Figure 1. The black dashed lines are the 99% upper and lower prediction levels. Only seven points are outside these levels, and when we exclude them from our sample, the correlation coefficient presents an improvement from $r = 0.960$ up to $r = 0.981$. The red solid line is the acceleration as a function of the initial speed according to Equation 3. The point where $\alpha = 0$ occurs is for a critical velocity of $u_c = 449.5\ \mathrm{km\,s^{-1}}$. This velocity is a typical solar wind velocity and is the boundary for the acceleration of slower events where $u < u_c$, or for the deceleration of faster events where $u > u_c$.

The acceleration (α) is used in the set of equations:

$$\upsilon = u_0 + \alpha \Delta t, \quad (4)$$

and

$$d = u_0 \Delta t + \frac{1}{2}\alpha \Delta t^2, \quad (5)$$

where d is the distance of Earth from the Sun in km, υ is the final speed of the CME shock at 1 AU in $\mathrm{km\,s^{-1}}$, and α is the effective acceleration calculated by Equation 3 in $\mathrm{m\,s^{-2}}$.

The time Δt is the difference between the liftoff of the CME from Sun (t_0) and the arrival time of the shock at 1 AU ($t_{\rm pred}$) in seconds, based on the relation $\Delta t = t_{\rm pred} - t_0$. From Equations 3 and 5, we calculate Δt, and as a result, the predicted arrival time of the shock ($t_{\rm pred}$) at the distance of 1 AU. The outputs of this simple model are the acceleration or deceleration, the velocity of the shock at 1 AU, and the arrival date and time of the shock.

The method described using Equations 3 and 5 was applied for each of the 214 ICMEs from Paper I in order to calculate the predicted arrival time of the shock. Then, the duration between the predicted arrival time $t_{\rm pred}$ and the actual arrival time of the shock $t_{\rm shock}$ was calculated using the relation $T = t_{\rm pred} - t_{\rm shock}$. The time difference T ranged from -64.23 to $+55.98$ hours, where the negative value means that the actual shock time was earlier than the predicted arrival time. The mean value of the time duration between the predicted and the actual arrival time was found to be equal to $\langle T \rangle = +3.03 \pm 1.53$ hours. The standard error of 1.53 hours was calculated as the mean of all events. In order to estimate the standard errors of the other calculated statistical variables, a simple bootstrap method was applied with replacement for 10^6 runs. The standard deviation was found to be equal to 22.32 ± 0.98 hours and the median equal to 4.17 ± 1.84 hours. The mean absolute error (MAE) for these events was 18.58 ± 0.87 hours, and RMSE was 22.47 ± 0.95 hours.

3. Projection Effects

As described above, the initial speed of the CME u_0 was obtained from the LASCO CME list. The speed of a CME that is calculated by white-light coronagraph data may be slightly different from the actual radial speed of the CME. This is due to the projection effects as this speed is measured on the plane of the sky (Gopalswamy *et al.*, 2001). It is generally assumed that this speed is the same as the radial speed of the CME (Gopalswamy *et al.*, 2000). An effort to calculate the radial velocity using the coordinates of the active region that produced the CME was presented by Leblanc *et al.* (2001) based on an earlier work by Sheeley *et al.* (1999). According to this work, the radial speed (u_r) and the speed that is a projection on the plane of the sky (u_0) are associated through the relation

$$u_r = u_0 \frac{1 + \sin a}{\sin \varphi + \sin a}, \tag{6}$$

where the angle φ is calculated from the coordinates of the active region on the surface of the Sun using the relation

$$\cos \varphi = \cos \psi \cos \lambda, \tag{7}$$

where ψ is the longitude and λ is the latitude of the active region.

Finally, the angle a is the cone angle (half angular width) of the CME. The angular width is not the actual 3D angular width of the CME, but is the projection of the actual angular width on the plane of the sky. At this point, a new problem occurs as the cone angle for halo CMEs with an angular width equal to 360° cannot be measured (Gopalswamy *et al.*, 2001). The majority (55.6%) of the events from Paper I were halo events, and as a result, at a first approach, all the halo CMEs were excluded from our sample in order to apply Equation 6 to calculate the radial speed of these events.

Moreover, the association between CMEs and solar flares was also investigated for Paper I. In particular, when the CME was accompanied by a solar flare, we also presented the characteristics of the flare such as the magnitude, the peak time, and the coordinates (heliographic latitude and longitude). As described above, in order to use Equation 6 and 7, it is

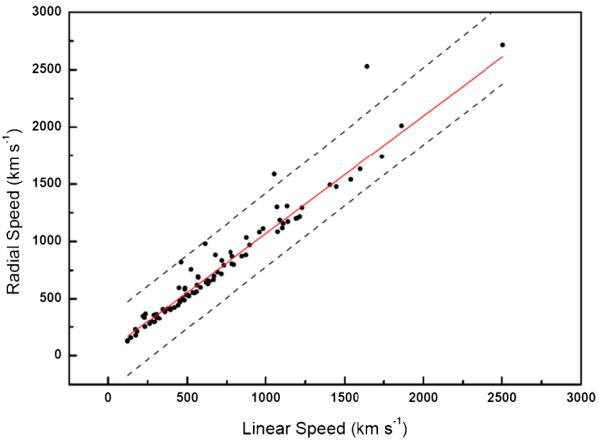

Figure 2 Scatter plot of the radial speed as calculated from Equation 6 against the CME linear speed. The *solid line* is a linear fit with $r = 0.987$ described by Equation 8.

very important to know the coordinates of the active region on the Sun. These coordinates were known for 87 events in our ICME list, as these CMEs were associated with a solar flare with known coordinates. From this sample of 87 events, a new linear relation between the radial speed and the speed on the plane of the sky was obtained, with the form

$$u_r = 41.516 + 1.027u_0, \tag{8}$$

where both speeds are in km s^{-1}. In Figure 2 we show the radial speed, using Equation 6, against the CME linear speed on the plane of the sky from LASCO. This relation is the result obtained by applying a linear fit method with 99% upper and lower prediction levels. Only two points are outside of these levels (black dashed lines), and when we exclude them from our sample, the correlation coefficient is improved from $r = 0.971$ to $r = 0.987$. The solid red line is the radial speed as a function of the CME linear speed based on Equation 8.

Moreover, with Equation 8 applied for the 214 events and using the new radial speed as an input in our EAM model in Equation 3 and not the initial speed (u_0) from LASCO directly, a new form of the model is obtained (EAMv2):

$$\alpha = 1.41392\left[\mathrm{m\,s^{-2}}\right] - 26.30 \times 10^{-4}\left[10^{-3}\ \mathrm{s^{-1}}\right]u_r - 1.14717 \times 10^{-6}\left[10^{-6}\ \mathrm{m^{-1}}\right]u_r^2. \tag{9}$$

This method improves the results, thus proving that the projection effects were more important than first assumed. The time difference ranged from -65.17 to $+52.02$ hours, where the negative value means that the actual shock time was earlier than the predicted arrival time. The new mean value between the arrival time of the shock and the predicted arrival time by the model based on Equation 9 was found to be equal to $\langle T \rangle = -0.28 \pm 1.48$ hours. The standard deviation is 21.60 ± 0.96 hours and the median is 0.35 ± 1.88 hours. The MAE is 17.65 ± 0.85 hours, and the RMSE is 21.55 ± 0.96 hours. The histogram with the results using Equation 3 and after taking into account the projection effects using also Equation 9 is given in Figure 3.

4. Case Study of Three Events

A slightly different version of the effective acceleration empirical model, first introduced in Paper I to predict the arrival time of the CME, was developed in order to predict the arrival time of a shock driven by a CME based on Equation 2.

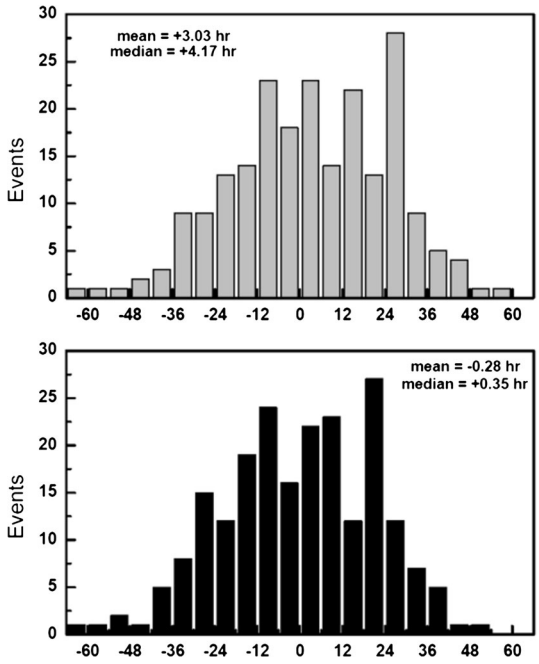

Figure 3 Histograms of the time duration of 214 ICMEs from Paper I between the observed shock arrival time and the predicted times from Equation 3 (*top panel*) and EAM v2 from Equation 9 (*bottom panel*). The bin size is 6.0 hours. The statistical variables of mean and median are also presented.

This version (EAMv1) is one of the registered models available in the Community Coordinated Modeling Center (CCMC). This multi-agency partnership was developed for the support of next-generation space science and space weather models (https://ccmc.gsfc.nasa.gov/). In particular, the EAM model contributes to the CME Scoreboard with predictions for the arrival time of CME shocks, among other models. Up to now, this version is not available for online runs, but it will be very soon. The CME Scoreboard (developed at the CCMC – https://kauai.ccmc.gsfc.nasa.gov/CMEscoreboard/) is a research-based forecasting activity. Through this, it is possible for every registered forecaster to submit a forecast, and all the forecasts are instantly available to everyone. In addition, it is very helpful to at once compare in real time the predictions when the event has arrived.

As an example, our new method was applied here to three Earth-directed CMEs reported on 23 May, 28 June, and 14 July 2017.

The CME on 23 May 2017 was reported at 06:00 UT, and it was a slow CME with an angular width smaller than 60°. The associated shock arrived on 27 May 2017 at 14:47 UT, and this CME produced a strong G3 geomagnetic storm (according to the National Oceanic and Atmospheric Administration (NOAA) scales: http://www.swpc.noaa.gov/noaa-scales-explanation), with Kp and Dst indexes reaching the values of 7 and −125 nT, respectively. The predictions provided from the different models from the CME Scoreboard of the CCMC that are available for this event are presented in Table 1. A positive time difference means that the prediction time was later than the actual arrival time of the shock, and *vice versa* for negative values.

The CME on 28 June 2017 was reported at 16:24 UT, and it was also a slow CME with an angular width smaller than 50°. The associated shock arrived on 1 July 2017 at 16:26 UT, and this CME produced a weak impact on Earth, with Kp and Dst indexes reaching the values of 4 and −14 nT, respectively. In Table 2 the predictions by three models that were available for this event are presented. For this event, it is noteworthy that all the models

Table 1 Predicted arrival time of the shock driven by the CME on 23 May 2017 06:00 UT by three models. The difference time between this time and the actual arrival time of the shock (in hours) is also given.

Model	Prediction	Difference (hours)
WSA-ENLIL + Cone (GSFC SWRC)	2017-05-26 18:00 UT	−20.78
WSA-ENLIL + Cone (Met Office)	2017-05-26 15:00 UT	−23.78
EAMv1 (ASWFC)	2017-05-27 21:07 UT	+6.33

Table 2 Predicted arrival time of the shock driven by the CME of 28 June 2017 16:24 UT by three models. The difference between this time and the actual arrival time of the shock (in hours) is also presented.

Model	Prediction	Difference (hours)
WSA-ENLIL + Cone (Met Office)	2017-07-02 21:00	+28.60
WSA-ENLIL + Cone (GSFC SWRC)	2017-07-02 19:30	+27.10
EAMv1 (ASWFC)	2017-07-03 04:50	+36.40

Table 3 Predicted arrival time of the shock driven by the CME on 14 July 2017 01:36 UT by different models. The difference between this time and the actual arrival time of the shock (in hours) is also presented.

Model	Prediction	Difference (hours)
WSA-ENLIL + Cone (NOAA/SWPC)	2017-07-16 18:00	+12.77
DBM (SIDC)	2017-07-16 12:00	+6.77
WSA-ENLIL + Cone (Met Office)	2017-07-16 15:00	+9.77
WSA-ENLIL + Cone (GSFC SWRC)	2017-07-16 21:42	+16.47
SARM	2017-07-16 08:53	+3.65
Ensemble WSA-ENLIL + Cone (GSFC SWRC)	2017-07-16 16:51	+11.62
EAMv1 (ASWFC)	2017-07-16 08:48	+3.57

that provide a predicted arrival time gave the arrival of the shock as more than one day later than the actual arrival time of the shock. At this point, we would like to mention that among the multiple sources of error in the prediction of the CME arrival time, the most probable from our point of view is the CME input speed. This CME reached Earth after almost 72 hours, which indicates a transit velocity (a constant speed from the Sun to Earth) of about 580 km s^{-1}. When we reverse the procedure, using the times of the CME onset and the real arrival time, we find an initial speed of about 710 km s^{-1}. These speeds are both very different from the speeds that all the previous models (Table 2) have used.

Finally, the CME on 14 July 2017 was reported at 01:36 UT, and it was a fast partial-halo CME with an angular width of almost 240°. The associated shock arrived on 16 July 2017 at 05:14 UT, and this CME produced a moderate G2 geomagnetic storm, with Kp and Dst indexes reaching values of 6 and −69 nT, respectively. In Table 3 the predictions that are available from different models for this event are presented.

As mentioned before, for the first time, we took into account in the EAMv1 model the projection effects on the initial linear speed of the CME. This approach significantly improved the statistics of this model, as is shown in the test case of the above three Earth-directed CMEs. In Table 4 we list the prediction of our model (EAMv1) before the improvement – using Equation 2 – and after the improvement taking into account the projection

Table 4 Predicted and actual arrival times of the three CMEs shocks with the EAMv1 and EAMv2 models. The difference between these times (in hours) is also presented.

CME (date and time UT)	Shock arrival time (date and time UT)	Predicted arrival date and time			
		EAM-v1	Difference (hours)	EAM-v2	Difference (hours)
23-05-2017 06:00	27-05-2017 14:47	27/05/2017 21:07	+6.33	27/05/2017 17:55	**+3.15**
28-06-2017 16:24	01-07-2017 16:26	03/07/2017 04:50	+36.40	03/07/2017 01:27	**+33.03**
14-07-2017 01:36	16-07-2017 05:14	16/07/2017 08:48	+3.57	16/07/2017 05:33	**+0.33**

effects (EAMv2) – using Equation 9 – as well as a comparison with the times where the shocks actual arrived.

5. Discussion and Conclusions

We presented a new simple and accurate model for predicting the arrival time of a shock that precedes a CME at 1 AU. This empirical model is based on a previous version (developed in Paper I) for predicting the arrival of the main part of a CME at 1 AU, which uses as inputs the onset date and time and the linear speed of the CME. The first version (EAMv1) was tested for three Earth-directed CMEs, and the results are available on the CME Scoreboard.

Many researchers discussed and took into account the projection effects in their models, as the linear speed of a CME is a projection on the plane of the sky of the coronagraph (see Gopalswamy *et al.*, 2001; Leblanc *et al.*, 2001; Xie *et al.*, 2006; Vrsnak *et al.*, 2007). A significant improvement was noted in our results when the projection effects of the linear speed of the CME were taken into account. The linear speed of the CME from the LASCO list may be slower than the radial speed due to projection effects as the linear speed of the CME is a speed measured on the plane of the sky (Gopalswamy *et al.*, 2001). A new relation between the radial speed and the linear speed of the CME is presented in our work, and the predictions improved when this new relation was combined with the EAM model. This version of our model (EAMv2) also has better statistics than the previous version using Equation 2. In particular, the mean value of the time difference between the predicted and the real arrival times for 214 events from Paper I is -0.28 ± 1.48 hours with an MAE of 17.65 ± 0.85 hours and an RMSE of 21.55 ± 0.96 hours, while before, the improvement the mean value was $+3.03 \pm 1.53$ hours with an MAE of 18.58 ± 0.87 hours and an RMSE of 22.47 ± 0.95 hours. The standard error is on the mean of all events. In Table 5 we list the statistical results of the two models using a simple bootstrap method with replacement.

This model was also applied in a test case of three Earth-directed CMEs. In all three cases, the new predicted arrival times were improved in comparison with the times of the previous version of the EAMv1 model. In particular, the time difference between the observed arrival time and the predicted times for the CME on 23 May decreased from 6.33 hours to 3.13 hours, for the CME on 28 June, they decreased from 36.40 hours to 33.02 hours, and for the CME on 14 July, they decreased from 3.57 hours to 0.32 hours.

In a recent study, Liu and Qin (2015) also studied very many eruptions associated with solar flares during Solar Cycle 23 and used the shock time of arrival (STOA) model (Dryer and Smart, 1984) to estimate the shock arrival times. In a sample of 225 events, they found an RMSE using the STOA model of 18.96 hours and from the modified STOA model, of 17.89 hours. The majority of the published works used in their studies a much smaller sample. Makela, Gopalswamy, and Yashiro (2016) selected 19 CMEs with as much information

Table 5 Summary of the statistical parameters for both models using a sample of 214 ICMEs of the duration (T) between the predicted arrival time and the actual arrival time of the shock $T = t_{\rm pred} - t_{\rm shock}$. The error of the average is on the mean of all events. The error bars for the other parameters are estimated using a simple bootstrap method with replacement for 10^6 times. All results are in hours.

Statistical parameter	EAMv1	EAMv2
Average	$+3.03 \pm 1.53$	-0.28 ± 1.48
Median	$+4.17 \pm 1.84$	$+0.35 \pm 1.88$
Standard deviation	$+22.32 \pm 0.98$	$+21.60 \pm 0.96$
MAE	$+18.58 \pm 0.87$	$+17.65 \pm 0.85$
RMSE	$+22.47 \pm 0.95$	$+21.55 \pm 0.96$

as possible, using data from the LASCO/C3 and STEREO/COR2 coronagraphs from a list of Earth-directed CMEs in 2010 – 2012 published by Gopalswamy *et al.* (2013). They showed that the full ice-cream cone model of the CME should be used in order to estimate the radial speed of the CME from the expansion speed. Furthermore, they applied the ESA model and found an MAE and an RMSE between the observed arrival times and the predicted times of 8.4 and 5.8 hours, respectively, using measurements from STEREO/COR2. When they used inputs from LASCO, the MAE rose to 10.4 hours and the RMSE to 11.6 hours. Falkenberg *et al.* (2011) studied 16 shocks spotted at Mars by the *Mars Global Surveyor* and at Earth using data from the Operating Missions as a Node on the Internet (OMNI) database in 2001 – 2003. For this complex method, they used ENLILv2.6, for which the magnetohydrodynamics around a sphere (MAS) or Wang-Sheely-Arge (WSA) model was used in order to define the solution of the coronal solar wind. Of four of the total six input parameters defined by two methods, the first was the manual method by Xie, Ofman, and Lawrence (2004), and the second was the automated method by Pulkkinen, Oates, and Taktakishvili (2010). Standard values were set for the CME density and temperature equal to 1.2×10^3 cm^3 and 0.8×10^6 K, respectively. The MAE using the ENLIL model was 13 hours for the manual method and 15 hours for the automated method. Taktakishvili *et al.* (2011) studied 36 geomagnetic storms, and in 20 of them, they were able to define the input parameters of the associated CMEs to estimate the shock arrival time using the WSA-ENLIL model. They used the two methods of Xie, Ofman, and Lawrence (2004) and Pulkkinen, Oates, and Taktakishvili (2010) and found an MAE of 6.9 and 11.2 hours, respectively. In this work, they also used the ESA model for comparison, and they found an MAE of 8.0 hours. Millward *et al.* (2013) studied 25 CMEs during the period of October 2011 – October 2012 using the WSA-ENLIL model, and they found an MAE of 7.5 hours. Mays *et al.* (2015) used the WSA-ENLIL+Cone model to study a set of 30 CMEs observed in the period January 2013 – July 2014. The results of MAE and RMSE were 12.3 and 13.9 hours, respectively. Hess and Zhang (2015) studied a very small sample of 7 CMEs, although with very good results. They studied the propagation of the shock using a drag-based model and STEREO/HI data, extending the CME measurements from the Sun as far out as possible. They calculated the arrival time of the associated shocks and found an MAE of 3.5 hours.

In our case, using either the EAMv1 or EAMv2 model, the high values of MAE and RMSE in comparison with other similar works are a function mainly of the event selection, and second a function of the sample size. In the case of a small sample size with very good predictions, these values should be much lower.

According to Zhang *et al.* (2007), the prolonged and enhanced southward-directed magnetic field B_z inside the solar wind is the driver for geomagnetic storms, regardless of the

solar origin (CME or coronal hole). In a future work, the southward component of the interplanetary magnetic field B_z based on the data from the ICME list presented in paper I will be investigated in order to be develop an empirical model for the prediction of the magnetic field value. This attempt will be very useful for estimating the levels of the potential geomagnetic storm in the case of an Earth-directed CME.

Temmer *et al.* (2017) showed that the preconditioning of the interplanetary space due to transient CME disturbances is very important for CME propagation and for space weather forecasting. In Paper I, the background conditions (magnetic field and solar wind plasma data) before the arrival of 266 Earth directed CMEs have been made available for the first time. In a future work, these conditions will be investigated using also the data during the main part of the ICME in order to estimate the possible association between the initial/background and the final conditions of the solar wind, deriving thus the possible increase in speed or magnetic field.

The different models and/or the methods described above show that it is possible to obtain a good accuracy between the observed time of the arrival of the shock and the predicted time with a period of some hours. In some cases, the models are very simple in use as they have as inputs the date and time and the speed of the CME (EAM, ESA), and in other cases, they may be very complicated as further calculations are needed (ENLIL, DBM). In particular, the EAM and ESA models are similar in their calculations, while DBM and ENLIL are not similar. DBM involves solving two equations of motion in 1D, and ENLIL is a 3D magnetohydrodynamics (MHD) model. Summarizing, the main problem of predicting the arrival time of the shock preceding a CME probably is determining the CME characteristics with the greatest possible accuracy, such as the initial speed or the angular width. It is necessary to have 3D surveillance of the solar corona using satellites, not only at the Lagrange L1 libration point between the Sun and Earth, but also in other sites (*e.g.* STEREO), in order to have a full image of the CME expansion in interplanetary space.

This version of our model (EAMv2) is a very useful tool for space weather applications, and the Athens Space Weather Forecasting Center (ASWFC – http://cosray.phys.uoa.gr/index.php/space-weather-report) already uses it for the preparation of the daily space weather forecasting report.

Acknowledgements We are grateful to the providers of the solar, interplanetary, and geomagnetic data used in this work. The coronal mass ejection data for the CMEs on 23 May, 28 June, and 14 July 2017 are taken from the CACTus CME list (http://sidc.oma.be/cactus/catalog.php). The prediction of the arrival time of the shocks for the Earth-directed CMEs used in Section 4 are taken from the CME Scoreboard (https://kauai.ccmc.gsfc.nasa.gov/CMEscoreboard/). Thanks are also due to the anonymous referee for the comments that improved this paper.

Conflict of interest The authors declare that they have no conflicts of interest.

References

Dryer, M., Smart, D.F.: 1984, *Adv. Space Res.* **4**, 291. DOI.

Falkenberg, T.V., Taktakishvili, A., Pulkkinen, A., Vennerstrom, S., Odstrcil, D., Brain, D., Delory, G., Mitchell, D.: 2011, *Space Weather* **9**, S00E12. DOI.

Gopalswamy, N., Lara, A., Lepping, R.P., Kaiser, M.L., Berdichevsky, D., St. Cyr, O.C.: 2000, *Geophys. Res. Lett.* **27**, 145. DOI.

Gopalswamy, N., Lara, A., Yashiro, S., Kaiser, M.L., Howard, R.A.: 2001, *J. Geophys. Res.* **106**(29), 207. DOI.

Gopalswamy, N., Lara, A., Manoharan, P.K., Howard, R.A.: 2005, *Adv. Space Res.* **36**, 2289. DOI.

Gopalswamy, N., Makela, P., Xie, H., Yashiro, S.: 2013, *Space Weather* **11**, 661. DOI.

Gosling, J.T.: 1993, *J. Geophys. Res.* **98**, 18937. DOI.
Hess, P., Zhang, J.: 2015, *Astrophys. J.* **812**, 2. DOI.
Leblanc, Y., Dulk, G.A., Vourlidas, A., Bougeret, J.-L.: 2001, *J. Geophys. Res.* **106**, 25301. DOI.
Liu, H.L., Qin, G.: 2015, *J. Geophys. Res.* **120**(7), 5290. DOI.
Makela, P., Gopalswamy, N., Yashiro, S.: 2016, *Space Weather* **14**, 5. DOI.
Mays, M.L., Taktakishvili, A., Pulkkinen, A., MacNeice, P.J., Rastatter, L., Odstrcil, D., *et al.*: 2015, *Solar Phys.* **290**, 1775. DOI.
Millward, G., Biesecker, D., Pizzo, V., de Konig, C.A.: 2013, *Space Weather* **11**, 57. DOI.
Odstrcil, D.: 2003, *Adv. Space Res.* **32**, 497. DOI.
Odstrcil, D., Pizzo, V.J., Linker, J.A., Riley, P., Lionello, R., Mikic, Z., Luhmann, J.G.: 2004, *J. Atmos. Solar-Terr. Phys.* **66**, 1311. DOI.
Paouris, E., Mavromichalaki, H.: 2017, *Solar Phys.* **292**, 30. DOI.
Pulkkinen, A., Oates, T., Taktakishvili, A.: 2010, *Solar Phys.* **261**, 115. DOI.
Sheeley, N.R., Walters, J.H., Wang, Y.M., Howard, R.A.: 1999, *J. Geophys. Res.* **104**, A11. DOI.
Taktakishvili, A., Pulkkinen, A., MacNeice, P., Kuznetsova, M., Hesse, M., Odstrcil, D.: 2011, *Space Weather* **9**, S06002. DOI.
Temmer, M., Reiss, M.A., Nikolic, L., Hofmeister, S.J., Veronig, A.M.: 2017, *Astrophys. J.* **835**, 2. DOI.
Vrsnak, B., Sudar, D., Ruzdjak, D., Zic, T.: 2007, *Astronomy and Astrophysics* **469**, 339. DOI.
Vrsnak, B., Zic, T., Vrbanec, D., Temmer, M., Rollett, T., Mostl, C., Veronig, A., Calogovic, J., Dumbovic, M., Lulic, S., Moon, Y.-J., Shanmugaraju, A.: 2013, *Solar Phys.* **285**, 295. DOI.
Xie, H., Ofman, L., Lawrence, G.: 2004, *J. Geophys. Res.* **109**, A03109. DOI.
Xie, H., Gopalswamy, N., Ofman, L., St. Cyr, O.C., Michalek, G., Lara, A., Yashiro, S.: 2006, *Space Weather* **4**, 10. DOI.
Zhang, J., Richardson, I.G., Webb, D.F., Gopalswamy, N., Huttunen, E., *et al.*: 2007, *J. Geophys. Res.* **112**, A10102. DOI.