Space Weather

RESEARCH ARTICLE
10.1029/2024SW004165

Decent Estimate of CME Arrival Time From a Data-Assimilated Ensemble in the Alfvén Wave Solar Atmosphere Model (DECADE-AWSoM)

Hongfan Chen[1], Nishtha Sachdeva[2], Zhenguang Huang[2], Bart van der Holst[2], Ward Manchester IV[2], Aniket Jivani[1], Shasha Zou[2], Yang Chen[3], Xun Huan[1], and Gabor Toth[2]

[1]Dept. of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA, [2]Dept. of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA, [3]Dept. of Statistics, University of Michigan, Ann Arbor, MI, USA

Key Points:

- Coronal white light observations are used to constrain an ensemble of coronal mass ejection simulations
- Automated methods are developed to identify and compare important features in observed and simulated white light images
- We find a marked improvement of coronal mass ejection arrival time predictions with the new approach for multiple events

Correspondence to:
H. Chen,
chenhf@umich.edu

Citation:
Chen, H., Sachdeva, N., Huang, Z., van der Holst, B., Manchester, W., IV, Jivani, A., et al. (2025). Decent estimate of CME arrival time from a data-assimilated ensemble in the Alfvén wave solar atmosphere model (DECADE-AWSoM). *Space Weather*, *23*, e2024SW004165. https://doi.org/10.1029/2024SW004165

Received 13 SEP 2024
Accepted 18 DEC 2024

Author Contributions:

Conceptualization: Hongfan Chen, Aniket Jivani, Yang Chen, Xun Huan, Gabor Toth
Data curation: Hongfan Chen, Nishtha Sachdeva, Zhenguang Huang, Bart van der Holst, Ward Manchester IV, Aniket Jivani
Formal analysis: Hongfan Chen, Ward Manchester IV, Aniket Jivani, Shasha Zou, Yang Chen, Xun Huan
Funding acquisition: Bart van der Holst, Yang Chen, Xun Huan, Gabor Toth
Investigation: Hongfan Chen, Nishtha Sachdeva, Zhenguang Huang, Aniket Jivani
Methodology: Hongfan Chen, Nishtha Sachdeva, Bart van der Holst, Ward Manchester IV, Aniket Jivani, Shasha Zou, Yang Chen, Xun Huan, Gabor Toth

Abstract Forecasting the arrival time of Earth-directed coronal mass ejections (CMEs) via physics-based simulations is an essential but challenging task in space weather research due to the complexity of the underlying physics and limited remote and in situ observations of these events. Data assimilation techniques can assist in constraining free model parameters and reduce the uncertainty in subsequent model predictions. In this study, we show that CME simulations conducted with the Space Weather Modeling Framework (SWMF) can be assimilated with SOHO LASCO white-light (WL) observations and solar wind observations at L1 prior to the CME eruption to improve the prediction of CME arrival time. The L1 observations are used to constrain the model of the solar wind background into which the CME is launched. Average speed of CME shock front over propagation angles are extracted from both synthetic WL images from the Alfvén Wave Solar atmosphere Model (AWSoM) and the WL observations. We observe a strong rank correlation between the average WL speed and CME arrival time, with the Spearman's rank correlation coefficients larger than 0.90 for three events occurring during different phases of the solar cycle. This enables us to develop a Bayesian framework to filter ensemble simulations using WL observations, which is found to reduce the mean absolute error of CME arrival time prediction from about 13.4 to 5.1 hr. The results show the potential of assimilating readily available L1 and WL observations within hours of the CME eruption to construct optimal ensembles of Sun-to-Earth CME simulations.

Plain Language Summary Accurately predicting the arrival time of coronal mass ejections (CMEs) continues to be a challenge in space weather forecasting. To enhance the accuracy and reliability of predictions, data assimilation techniques can be employed. In this study, we investigate assimilating CME simulations with SOHO LASCO white-light solar corona observations. By establishing a correlation between CME speed extracted from remote white-light images and CME arrival time at Earth, we are able to perform data assimilation at an early stage of the CME simulation process. This enables us to effectively constrain the simulation and improve the overall quality of the prediction.

1. Introduction

Coronal mass ejections (CMEs) are violent solar events where a massive amount of plasma and magnetic fields are expelled from the solar atmosphere into the interplanetary space. Fast CMEs originating from active regions are particularly violent affairs, which are strong drivers of space weather. Timely and accurate predictions of Earth-oriented CMEs are crucial to mitigating the detrimental impacts of CMEs, such as satellite loss and damage (Baruah et al., 2024), risks in power systems and pipelines (Pirjola et al., 2000), and radiation hazards to aviation safety (Meier et al., 2020). Accurate CME arrival time prediction is one of several crucial and challenging goals of space weather forecasting, alongside predicting the temporal variation of plasma velocity, density and the B_z component of the magnetic field impacting Earth or other points of interest.

Many models have been proposed to predict CME arrival times, and they can be roughly categorized as first-principles (analytical and numerical) and data-driven (empirical or machine learning) models (Manchester et al., 2017). First-principles models involve solving governing equations derived from known physical principles. Drag-based model (DBMs) are the simple yet effective analytical models (Cargill et al., 1996; Owens

Project administration: Gabor Toth
Software: Hongfan Chen, Nishtha Sachdeva, Zhenguang Huang, Bart van der Holst, Ward Manchester IV, Gabor Toth
Supervision: Shasha Zou, Yang Chen, Xun Huan, Gabor Toth
Validation: Hongfan Chen, Nishtha Sachdeva, Zhenguang Huang
Visualization: Hongfan Chen, Ward Manchester IV, Aniket Jivani
Writing – original draft: Hongfan Chen, Aniket Jivani
Writing – review & editing: Hongfan Chen, Nishtha Sachdeva, Zhenguang Huang, Ward Manchester IV, Aniket Jivani, Shasha Zou, Yang Chen, Xun Huan, Gabor Toth

et al., 2005; Vršnak et al., 2004) based on the assumption that the aerodynamic dominates CME velocity after the eruption phase. One representative of numerical models is the WSA-ENLIL-Cone model (Emmons et al., 2013) that combines the empirical WSA and Cone models with the first-principles magnetohydrodynamic (MHD) ENLIL model, and is currently used by the Space Weather Prediction Center of National Oceanic and Atmospheric Administration (SWPC NOAA). This model was tested and compared alongside several other empirical models, all achieving a mean absolute error (MAE) of around 13 hr on arrival time prediction across 28 CME events (Gopalswamy et al., 2013; Riley et al., 2018). While solving for a fully detailed physical description including the solar corona would be highly computationally expensive, these models leverage empirical rules and simplified physics, making them less expensive but also lower in fidelity.

In contrast, data-driven models forego the governing physics, and instead use available data to learn an empirical correlation or a machine-learned mapping from near-sun observations to CME Sun-Earth transit times through statistical and machine learning (ML) algorithms. Such a relationship is presented by Vršnak and Žic (2007), which relates CME transit time to CME speed derived from white-light (WL) observations from Large Angle and Spectrometric Coronagraph (LASCO) instrument on-board SOlar and Heliospheric Observatory (SOHO). As an example of machine learning, Liu et al. (2018) employed support vector machine (SVM) to predict CME arrival time where they manually designed CME features such as propagation speed and angular width as model inputs. Sudar et al. (2016) applied the neural network predict CME transit times using only the coronagraphic CME speed and source-region location (central meridian distance) to train the model. Wang et al. (2019a, 2019b) proposed to use a convolutional neural network (CNN) that directly takes in the entire WL image and avoids the need for the manual design of CME features. Alobaid et al. (2022) further incorporated ensemble learning, which couples regression algorithms and CNN to extract CME features from different observed data. The main advantage of data-driven models is that they are computationally inexpensive to evaluate and suitable for real-time predictions. However, their accuracy is highly dependent on the quantity, quality, and generalizability of the available training data. For example, limited observations near the Sun and in interplanetary space, and few CME events available for training, all pose a challenge to these techniques. Existing data-driven models generally achieve an MAE of around 10 hr for arrival time predictions.

Recent studies have shown the potential of a hybrid approach combining newer-generation physics-based modeling with data assimilation (DA) to improve CME arrival time prediction. Data assimilation (DA) is the concept for integrating observational data into numerical models to improve predictive accuracy. DA commonly refers to the task of state estimation in geoscience (Carrassi et al., 2018), where the model state is sequentially updated with new observations while model parameters remain fixed. This is widely applied in numerical weather prediction and is particularly popular in meteorology and oceanography (e.g., Ghil & Malanotte-Rizzoli, 1991; Navon, 2009). However, DA can also be used for parameter estimation (e.g., Evensen, 2009; Evensen et al., 1998; Smith et al., 2009). In either case, challenges in applying existing DA frameworks to space weather prediction include, but are not limited to, the sparse and uneven observational coverage of the Sun-Earth space (Lang et al., 2017; Temmer et al., 2023), and the quality and inconsistency of observational data (Turner et al., 2023). These challenges make short-term and/or early-stage sequential updates difficult and introduce large uncertainty in the model parameters. As a result, DA in space weather is often coupled with uncertainty quantification (UQ) (e.g., Iwai et al., 2021; Singh et al., 2023). A common strategy involves first generating an ensemble of simulations that manifest from the uncertainty in the model parameters or inputs and then down-selecting a subset of simulations that best match the observational data. For instance, Iwai et al. (2021) demonstrated that by picking simulations of the SUSANOO-CME model (Shiota & Kataoka, 2016) that most closely agree with the interplanetary scintillation (IPS) observations, the MAE of the predicted CME arrival time can be reduced from 6.7 to 5 hr for 12 halo CME events. Singh et al. (2023) used ensemble simulations to train a ML surrogate model, and showed that the assimilation of Solar Terrestrial Relations Observatory (STEREO) Heliospheric Imager (HI) data into simulations of the constant-turn flux rope model (Singh et al., 2022) reduces the MAE of the MS-FLUKSS simulations from around 8 hr to 4–5 hr for six CME events. While these results were obtained from a small number of CME events and thus warrant further validation, they displayed great promise for first-principles modeling strategies with DA. One drawback, however, is that these studies primarily focused on the accuracy improvement from DA, while uncertainty quantification and reduction were often secondary considerations or largely overlooked in the final predictions.

The main contributions of this paper are the following: (a) proposing a probabilistic prediction of CME arrival time through ensemble simulations and UQ, and (b) proposing DA of L1 observations of the background solar

wind and SOHO LASCO C3 WL coronagraph observations, which are readily available at early times of a CME propagation. We begin by conducting full Sun-to-Earth ensemble CME simulations using first-principles models in the Space Weather Modeling Framework (SWMF; Gombosi et al., 2021; Tóth et al., 2012). The L1 observations are used to constrain the model parameters of the background solar wind simulations. We then assimilate WL observations into the CME simulations and illustrate its effectiveness in reducing the uncertainty of CME flux rope parameters. Specifically, the assimilation uses synthetic WL images obtained from the model through the line-of-sight integration of the density structure of the simulated CME. We demonstrate that the CME propagation speed estimated from the WL observations can be used to filter the initial ensemble simulations, such that only an optimal subset of the ensemble needs to be propagated to 1 au in order to construct an accurate ensemble forecast for the arrival time and other quantities of interest with quantified uncertainty. Our approach leverages observational data to reduce uncertainty and improves both predictive accuracy and numerical model efficiency, thereby aligns with the general concept of DA. While this approach is also closely related to the concept of *model calibration* that optimizes parameters to best match observations, it does not account for *model inadequacy* (Kennedy & O'Hagan, 2001).

This filtering procedure is established based on the result of correlation analysis between the apparent propagation speed inferred from CME WL images and the CME arrival time in historical CME events, which enables us to construct a down-select criterion from CME propagation speed overshoot under a Bayesian framework. The analysis calls for the estimation of CME propagation speed as well as the estimation of CME arrival time from both ensemble simulations and observations. Estimating CME propagation speed through manual identification of CME front edges in coronagraph images can be inefficient and challenging, especially when dealing with a large ensemble size. This underscores the necessity for an automated procedure. Over the years, many methods for automatic CME detection have been proposed, such as those by Robbrecht et al. (2009), Olmedo et al. (2008), Morgan et al. (2012), Wang et al. (2019a, 2019b), Shan et al. (2020), and Alshehhi and Marpu (2021). However, these methods have primarily been designed for LASCO WL observations, while in our study we need a method that can flexibly accommodate both observed and synthetic WL images. We propose a new CME edge detection algorithm based on K-means clustering, which is an unsupervised clustering technique. The algorithm is similar to that of Alshehhi and Marpu (2021), but we apply clustering algorithms directly to coronagraph images without using any dimension reduction techniques. Finally, we automate the estimation of CME arrival time through a shock detection algorithm with a background solar wind speed correction method.

We demonstrate the overall proposed method for predicting CME arrival time on the following three CME events:

1. CME on 2014-09-10 18:00 UT (referenced as CME1 hereafter),
2. CME on 2015-03-15 01:48 UT (hereafter CME2), and
3. CME on 2017-07-14 01:25 UT (hereafter CME3).

This method reduces the mean absolute error of CME arrival time prediction from 13.4 to 5.1 hr.

The remainder of this paper is organized as follows. Section 2 describes the simulation setup of the SWMF and the AWSoM and EEGGL models within it. Section 3 provides details of the processing pipeline for the WL images and their subsequent comparisons, and the UQ and DA methods that lead to the final CME arrival time prediction. Section 4 contains key results and their discussion. Section 5 presents concluding remarks and directions for future work.

2. Numerical Models and Simulation Setup

2.1. Space Weather Modeling Framework (SWMF)

The modeling and simulation utilized in this paper are based on the University of Michigan's Space Weather Modeling Framework (SWMF; Gombosi et al., 2021; Tóth et al., 2012), a fully-functional, state-of-the-art, well-documented software for high-performance computing. SWMF covers different physical regimes of space through various model components and couples them together to model the space environment from the Sun to Earth and/or outer heliosphere in one framework. The SWMF suite is open-source (http://github.com/SWMFsoftware) and also available for runs-on-request through the Community Coordinated Modeling Center (CCMC) at the NASA Goddard Space Flight Center (GSFC) (https://ccmc.gsfc.nasa.gov).

2.2. Alfvén Wave Solar Atmosphere Model (AWSoM)

The solar corona (SC) and inner heliosphere (IH) components of the SWMF are represented by the Alfvén Wave Solar atmosphere Model (AWSoM; Sokolov et al., 2013, 2021; van der Holst et al., 2014, 2022). The SC extends from the upper chromosphere, through the transition region into the corona (up to 24 $R_\odot$ radial distance) and the IH component couples with the SC at 18–21 $R_\odot$ and extends up to 1 au and beyond. AWSoM is a self-consistent, global three-dimensional (3D) extended MHD model and includes the non-linear interaction of forward-propagating and reflected Alfvén waves (van der Holst et al., 2014). The solar wind is accelerated due to the pressure gradient of the Alfvén waves and heated by their dissipation. The energy is partitioned over the isotropic electron temperature and anisotropic parallel and perpendicular proton temperatures (van der Holst et al., 2022). AWSoM also includes stochastic heating and linear wave damping (Chandran et al., 2011), radiative losses based on the Chianti model (Dere et al., 1997) and electron heat conduction for both collisional and collisionless regimes. The MHD equations of AWSoM are solved using the Block-Adaptive-Tree Solar wind Roe-type Upwind Scheme (BATSRUS; Powell et al., 1999).

AWSoM is driven by observations of the solar photospheric magnetic field. The electron and proton temperatures (both parallel and perpendicular) are set equal to 50,000 K and the density at the inner boundary is set to 2×10^{17} m^{-3}. There are only a few free input parameters in AWSoM. These include the Poynting flux (S_A) of the outward traveling Alfvén wave (Fisk, 1996; Fisk & Schwadron, 2001; Sokolov et al., 2013), the wave correlation length ($L_\perp$) (Hollweg, 1986), and the stochastic heating amplitude and exponents (Chandran et al., 2011). Previous works by van der Holst et al. (2014), Huang et al. (2023, 2024), Sachdeva et al. (2019, 2021), and Jivani et al. (2023) describe in detail the typical values for these parameters and their variation during different phases of the solar cycle.

2.3. Eruptive Event Generator

The SWMF is equipped with an Eruptive Event Generator (EEG) module that facilitates the eruption of a magnetic flux-rope inserted at the inner boundary and superimposed on the modeled background solar corona. This flux-rope can be specified by an analytical flux-rope model, including the Gibson-Low (GL, Gibson and Low (1998)) and Titov-Démoulin (TD, Titov and Démoulin (1999)) flux-ropes. In this work, we use the GL model that has been employed in various previous studies (Jin et al., 2017a; Lugaz et al., 2005; Manchester et al., 2004, 2014) of CMEs using AWSoM. This eruptive event generator based on the Gibson-Low magnetic configuration (EEGGL) model is described in detail in Jin et al. (2017b), and is available as a user-friendly tool via CCMC. The analytical GL flux-rope is inserted along the polarity inversion line (PIL) of the CME-producing active region (AR) in a state of force imbalance leading to an immediate eruption. The flux-rope model is data-driven and parameterized with observational constraints. The location (longitude/latitude) of the erupting AR is obtained from flare observations, the radius of the flux-rope is constrained by the size of the active region obtained from the observations of the photospheric magnetic field (via magnetograms). The observed speed of the CME near the Sun is used to determine the magnetic strength of the flux-rope. The flux-rope also has an associated helicity and orientation. Two more geometric parameters associated with this flux-rope are the distance of the center of the torus to the center of the Sun and the stretching parameter used to obtain the analytical solution of a stretched axisymmetric spherical magnetic flux configuration (Jin et al., 2017a, 2017b).

The use of an out of equilibrium flux-rope inserted into the background solar wind is advantageous for both scientific and operational investigations. The GL flux-rope can mimic the three-part (front, cavity, core) CME structure commonly seen in coronagraph images (Hundhausen, 1993). Since the flux-rope is not in equilibrium it is also computationally less expensive to model the eruption in comparison to a time-consuming energy buildup process.

2.4. Numerical Simulation Setup

To model the solar wind and CMEs using AWSoM, the grid blocks in the SC and IH domains consist of $6 \times 8 \times 8$ and $8 \times 8 \times 8$ grid cells, respectively. The SC component uses the 3D spherical grid extending from the Sun to 24 $R_\odot$ while the IH, on a Cartesian grid, is coupled with SC through a spherical buffer grid between 18 and 21 $R_\odot$ and extending up to ± 250 $R_\odot$. Adaptive mesh refinement (AMR) is used to increase the resolution of the domain where needed, including the current sheet, around the CME eruption site and along the Sun-Earth line to better resolve the CME. The angular resolution in SC is 0.35° around the initial flux rope and the Sun-Earth

10.1029/2024SW004165

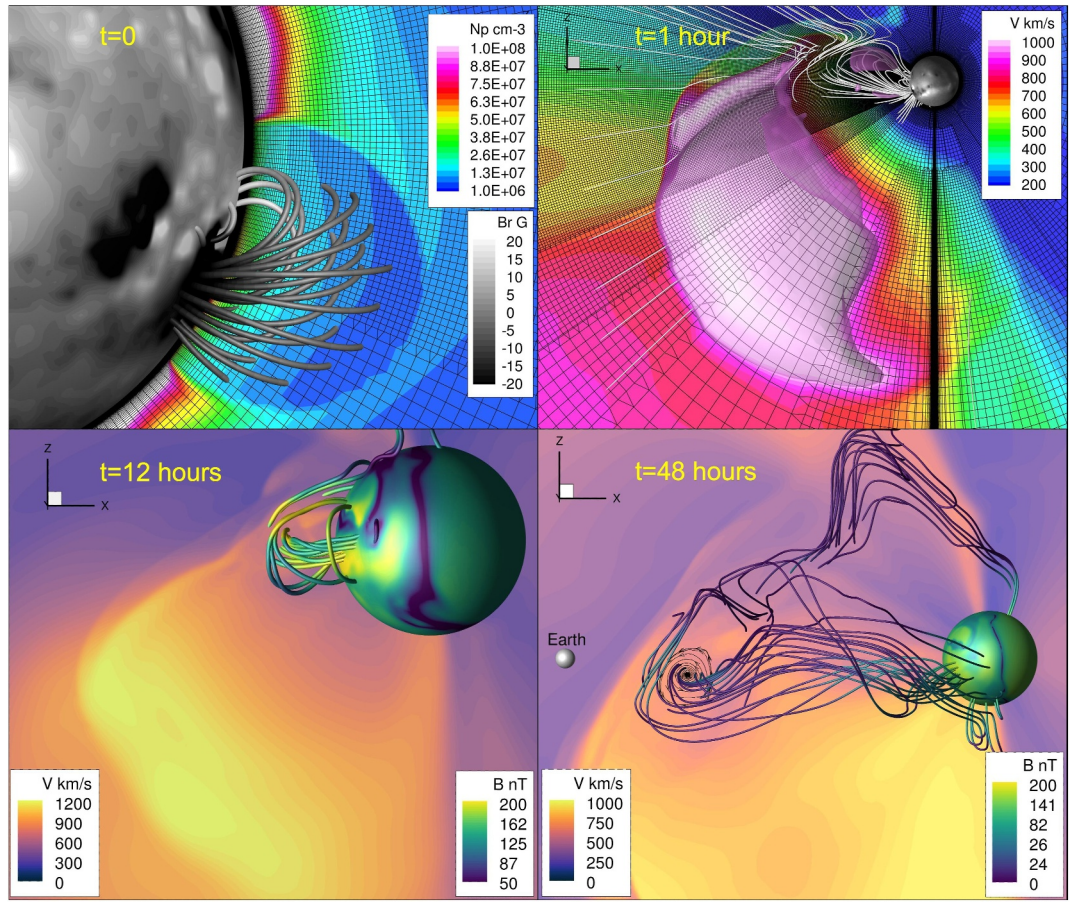

Figure 1. Simulation of the 2015 March 15 CME event (CME2). Four panels show the time evolution of flux rope and plasma on a central meridional plane at times $t = 0, 1, 12$ and 48 hr. *Top-left* $t = 0$: The GL flux-rope is shown in 3D with magnetic field lines above the solar surface showing B_r in gray scale. The central meridional plane shows the plasma mass density in color where a cavity is found embedded in the flux rope. The numerical mesh is shown in black. *Top-right* $t = 1$ hour: Magnetic field lines are shown in 3D. The plasma velocity on the central meridional plane is shown in color and an isosurface for velocity at 1,000 km/s captures the location of the CME driven shock. The numerical mesh is also shown. *Bottom-left* $t = 12$ hours: The erupting magnetic field are shown in 3D at with the plasma velocity shown in color on the central plane. The inner boundary of the IH domain is colored to show the magnetic field strength. *Bottom-right* $t = 48$ hours: The CME shown in the same format used at bottom-left.

line. In the IH component, the smallest cell size is 0.24 $R_\odot$ along the Sun-Earth line. To model the background solar wind, AWSoM uses local time-stepping for 80,000 iterations in SC to obtain a converged steady-state solution and couples it with the IH domain followed by 5,000 iterations in the IH component to get the 3D solar wind solution. The GL flux-rope is then superimposed onto the steady-state solar wind solution in the corona as a perturbation to the density, magnetic field, and pressure at the inner boundary, which leads to a force-imbalanced state and the flux-rope erupts as the simulation advances in time. The CME evolution is simulated in the SC domain only for 1 hr following which the solution is coupled with the IH domain and both domains are evolved for the next 11 hr. At the end of first 12 hr, it is assumed that the CME has completely traveled out of the SC domain into the IH domain. The SC component is then switched off and the solution is propagated in the IH component only for up to 4 days of physical time covering the entire Sun-to-Earth propagation of the CMEs. Figure 1 illustrates the evolution of simulated 2015 March 15 CME event (CME2) with the flux rope and the resulting disturbance passing through the SC and IH domains.

All numerical simulations are run on Frontera, a petascale computing system (Stanzione et al., 2020). We use 19 nodes for each background solar wind simulation and 59 nodes for each CME simulation, with each node equipped with 56 cores. The steady state phase takes 5 hr of wall time, resulting in about 5300 CPU hr. The time

dependent phase on average takes about 16 hr of wall time, using about 53,000 CPU hours. Note that the wall time is significantly shorter than the 4 days of simulated time.

Over the years, AWSoM has been extensively validated against remote and in situ observations during various phases of the solar cycle. AWSoM produces synthetic extreme ultra-violet (EUV) images that have been compared to EUV observations from STEREO/EUVI, SDO/AIA and SOHO/LASCO instruments (Jin et al., 2017a; Meng et al., 2015; Sachdeva et al., 2019, 2021; Van Der Holst et al., 2010). The AWSoM predicted structure of the solar corona also compares well with the tomographic reconstructions of the density and temperature of electrons near the Sun determined using the Differential Emission Measure Tomography (DEMT) during the quiescent phase (Lloveras et al., 2017, 2020, 2022). In addition, comparisons with Interplanetary Scintillation (IPS) data at various heliospheric distances as well as solar wind plasma observations at 1 au have successfully validated the capability of the AWSoM model to reproduce the solar wind structure near the Sun as well as in the inner heliosphere (Sachdeva et al., 2019). Multiple studies have also modeled observed CME events using AWSoM and the EEG module with SWMF with both the GL (Jin et al., 2017b; Manchester et al., 2014) and TD flux-rope descriptions (Lugaz et al., 2005; Manchester et al., 2008; Tóth et al., 2007). The modeled synthetic WL images have been compared with multi-viewpoint coronagraph observations of CMEs from LASCO C2, C3 and STEREO A/B COR1, COR2 instruments. CME driven shock arrival signatures at 1 au were also compared with the plasma observations from WIND and ACE satellites.

3. Methodology

In this section, first, we show the DA of L1 solar wind observation and present a systematic UQ study on the flux rope parameters that leads to the experimental design of ensemble simulations. Second, we present a six-step CME edge detection algorithm for calculating the CME propagation speed from WL images. Then, we demonstrate the automation of CME arrival time calculation from 1 au solar wind speed using our shock detection algorithm with background correction. Finally, we show the correlation between CME arrival time and WL speed, and use it to develop a UQ and DA procedure to improve the CME arrival time prediction by constraining the ensemble with WL observations.

3.1. Background Solar Wind Simulations

We perform background solar wind simulations using the AWSoM model in the SWMF. Our previous study (Jivani et al., 2023) performed global sensitivity analysis (GSA) on background solar wind simulations and identified the three most influential parameters that impact background speed and density to be: the multiplicative correction factor applied to the magnetogram ("FactorB0"), Poynting flux per magnetic field strength constant ("PoyntingFluxPerBSi") used at the inner boundary, and the coefficient of the perpendicular correlation length ("LperpTimesSqrtBSi") in the turbulent cascade model in AWSoM. Since our primary focus here is to investigate and incorporate the uncertainty of the flux rope parameters, we do not propagate uncertainty from these background parameters fully to the CME simulations in order to control computational cost. Instead, we identify the optimal background parameters that best reproduce the background solar wind observations at L1 under the curve distance (Sachdeva et al., 2019), and then perform CME simulations with varied flux rope parameters under this optimal background. Specifically, this first step of DA is described below and illustrated in Figure 2:

1. Following the same approach as in our previous study (Jivani et al., 2023), we vary the three most influential background parameters using maximum projection (MaxPro) designs (Joseph et al., 2015) to launch 10 background wind simulations for each CME event. The MaxPro design ensures that the space-filling properties on projections to all subsets of parameters are maximized (Joseph et al., 2015). The algorithm is implemented using R package MaxPro (Ba & Joseph, 2015). For CME3, which occurs during solar minimum conditions, we also vary the ion stochastic heating profile exponent ("StochasticExponent") parameter.
2. We calculate the curve distance between the simulation and the OMNI observations at 1 au for the plasma velocity (U) and proton number density (Np) over a 10 day time window prior to the map time. The map time refers to the time associated with the input magnetic field map of EEGGL, which is the closest map available to the CME eruption time. All observations in this window are equally weighted to calculate the distance. The background with the minimum averaged curve distance is then used to insert flux ropes and launch the ensemble of CME simulations.

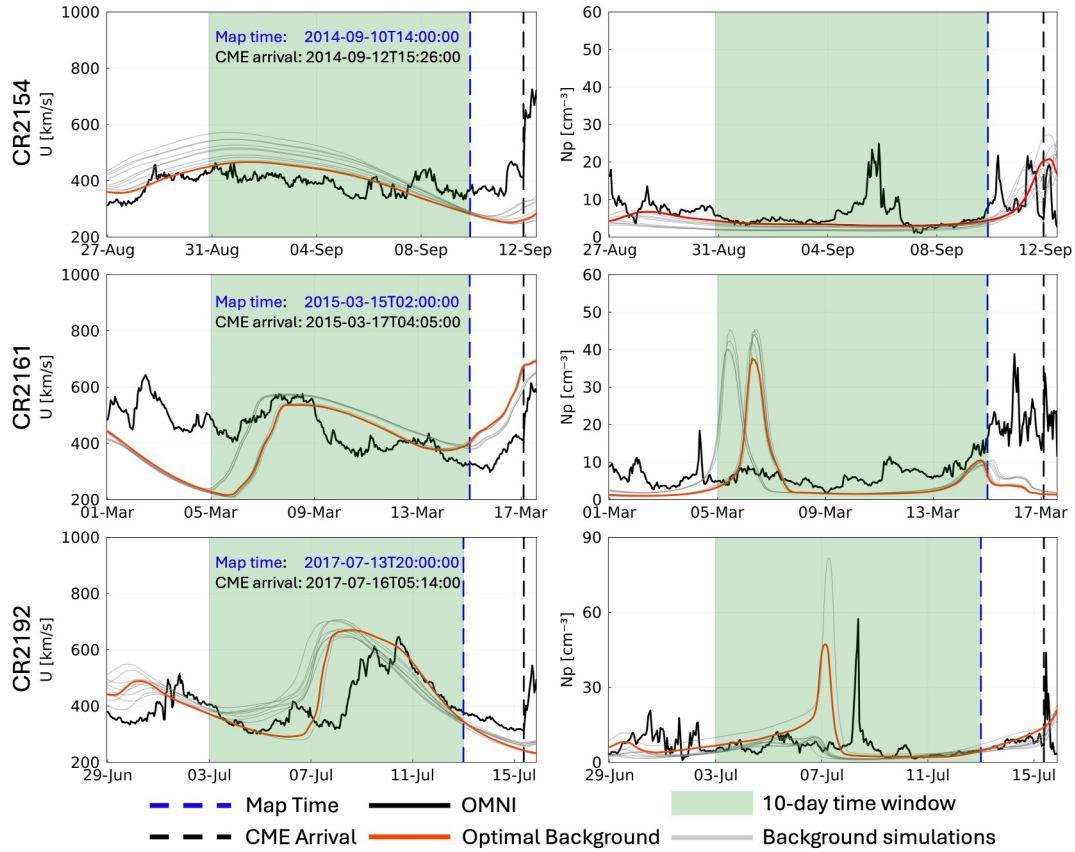

Figure 2. Background simulations of three Carrington rotations. We select the optimal background (in red) from the background ensemble (in gray) based on average curve distance calculated for plasma velocity (U) and proton number density (Np) over a 10-day time window prior to the map time (blue dashed line).

3.2. Ensemble CME Simulations

We parametrize GL flux ropes using the EEGGL module in the SWMF, and the flux rope parameters are listed in Table 1. These parameters correspond to the geometric representation of a GL flux rope in 2D, as depicted in Figure 3. While the baseline point estimates of these parameters can be obtained from EEGGL, a single CME simulation using these values would not capture the uncertainty effect of these parameters; therefore, we will create an ensemble of simulations while varying the values of certain parameters. Specifically, this is done by introducing a set of perturbation parameters: ΔOrientation, ApexCoeff, RelativeStrength and Helicity as shown in Table 2. These perturbation parameters are used either as-is or in combination with their corresponding baseline

Table 1
Flux Rope Parameters That Control the Flux Rope Geometry

EEGGL parameter	Description
Lat_{CME}, Lon_{CME}	Location of the flux-rope on the solar surface
Orientation	Orientation of the flux-rope on the solar surface
Radius	Radius of the GL torus that is stretched into the flux-rope
Stretch	Stretching parameter applied to the torus
ApexHeight	Height of the stretched torus above the solar surface
B	Magnetic field strength of the flux-rope
Helicity	Sign of the magnetic helicity of the flux-rope

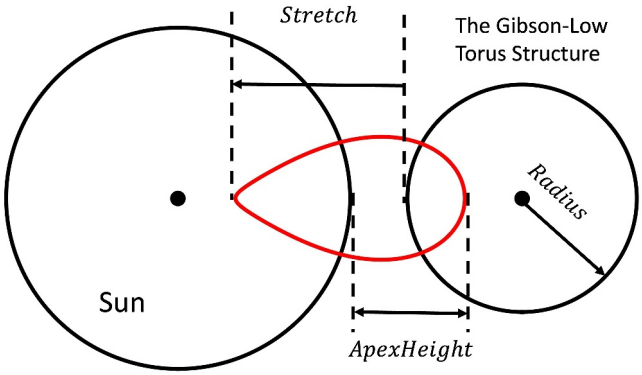

Figure 3. A schematic of Gibson-Low flux rope in 2D. Three EEGGL parameters are shown.

estimates (which are denoted by the subscript baseline in the following) to obtain the flux rope parameters that are inputs to the code (Jin et al., 2017a), while the other flux rope parameters are fixed to the baseline estimates:

1. $(\mathrm{Lat_{CME}}, \mathrm{Lon_{CME}}) = (\mathrm{Lat_{CME,baseline}}, \mathrm{Lon_{CME,baseline}})$
2. $\mathrm{Orientation} = \mathrm{Orientation_{baseline}} + \Delta\mathrm{Orientation}$
3. $\mathrm{Radius} = \mathrm{Radius_{baseline}}$
4. $\mathrm{Stretch} = \mathrm{Stretch_{baseline}}$
5. $\mathrm{ApexHeight} = \mathrm{Radius} \times \mathrm{ApexCoeff}$
6. $\mathrm{BStrength} = \mathrm{RelativeStrength} \times |B_{\mathrm{baseline}}|$
7. Helicity (-1 or $+1$)

We use MaxPro designs to generate the entire ensemble of perturbation parameters based on their feasible ranges. The limits of ΔOrientation mean that the angle between the line connecting the centers of the positive and negative polarities and the direction of the flux rope is at most 45°. The Helicity can only be ±1. The range of ApexCoeff limits the center of the stretched flux rope to be at $[-1/8, 1/4]\cdot$Radius relative to the surface of the Sun, which results in a reasonable shape and aspect ratio of the stretched flux rope above the surface. The RelativeStrength range is the least certain since the baseline value of Bstrength is based on empirical formulas in EEGGL and/or previous simulations results. Our range of ±50% is probably too narrow and may need to be expanded based on the comparison of simulated and observed WL images as we will discuss later. These values should not be considered hard limits and warrant further calibration in future work.

We conducted two batches of CME simulations in this work. The first batch has the largest sample size and includes an ensemble of 120 CME simulations for CME3, while the second batch consists of 72 simulations in total: 24 each for CME1, CME2, and CME3. The first batch was designed for sensitivity analysis, which will be discussed in detail in Section 4.1. Figure 4 illustrates the design of the resulting flux rope parameters (i.e., obtained by combining baseline estimates with perturbations) for the first batch. The first row of the figure shows the pairwise scatter plots of parameters under two levels of Helicty, with each point representing a single simulation. The ensemble produces synthetic WL images and CME arrival time predictions at 1 au for each member in chronological order as the CME propagates from the Sun to Earth. The design of the second batch of CME simulations, which were launched based on the results of sensitivity analysis and therefore varying only RelativeStrength and Helicity, will be introduced in Section 4.1.

3.3. White-Light Coronagraph Image Processing

WL images play a central role in our procedure to ultimately predict the CME arrival time. Both synthetic and observed WL images are involved. The former are generated from the SWMF described in Section 2; the latter are publicly available LASCO WL coronagraph data retrieved from https://sdac.virtualsolar.org. This section describes the image processing steps for extracting features that allow us to compare observed and synthetic images. In particular, WL observations usually show fine structures of CMEs but are subject to noise arising from irrelevant solar activities or measurement errors; synthetic WL images, on the other hand, do not have these features. Therefore, a few additional steps are included to process the observation images. We introduce our six-step algorithm below.

Step 1: Ratio Enhancement. Starting from a WL image (observation or synthetic), the total brightness is divided by the background brightness before CME eruption in order to reveal the relative enhancement. This step helps

Table 2
Flux Rope Perturbation Parameters

Parameter	Value range	Description
ΔOrientation	$[-45°, +45°]$	Change in orientation relative to baseline
ApexCoeff	$[0.875, 1.25]$	Multiplicative factor for ApexHeight from baseline
RelativeStrength	$[0.5, 1.5]$	Field strength relative to baseline from baseline
Helicity	-1 or $+1$	Sign of magnetic helicity

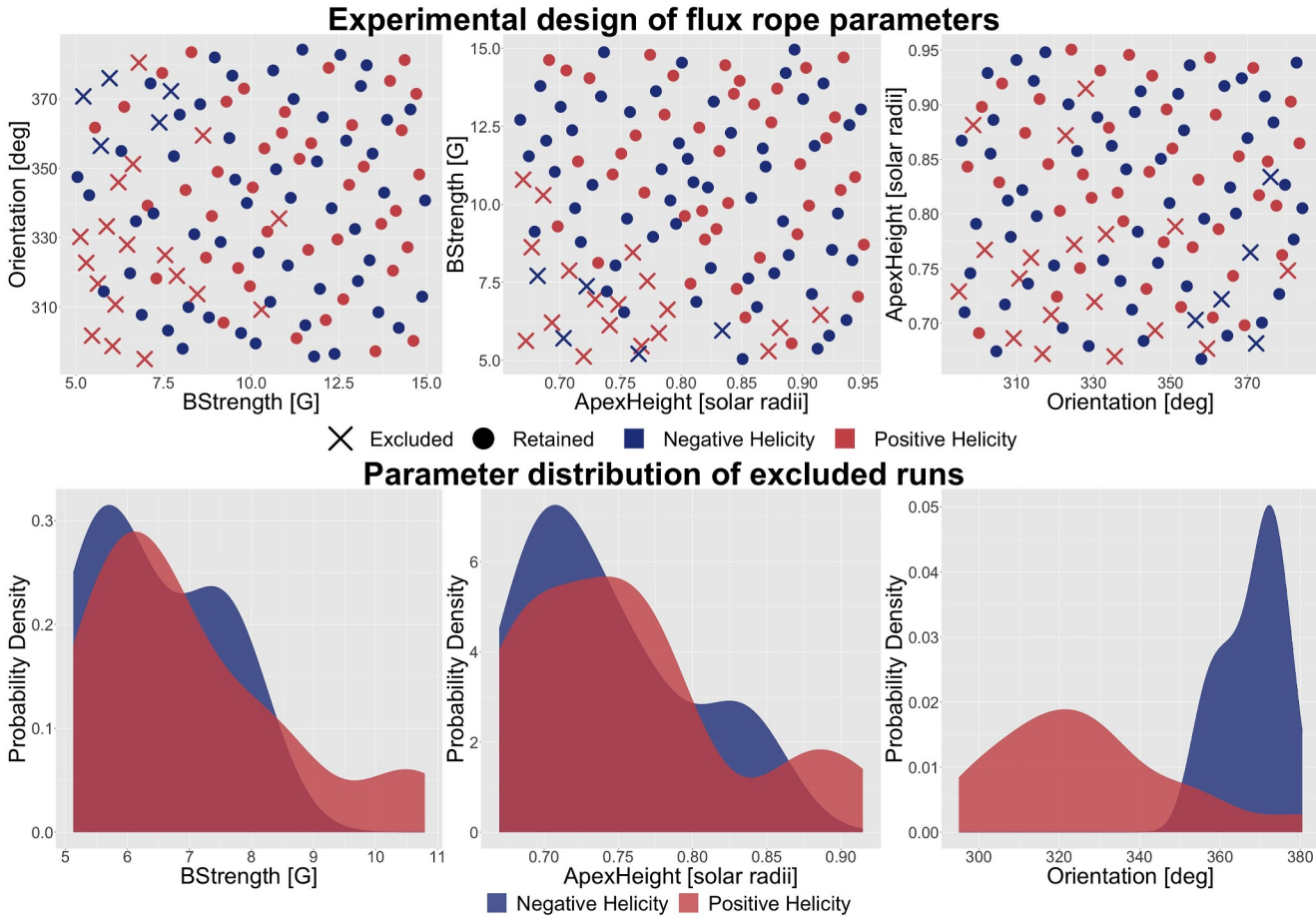

Figure 4. MaxPro design of flux rope parameters for the 120 ensemble CME simulations conducted for CME3. The first row shows the pairwise scatter plots of the parameters, with each point representing a single simulation. Excluded runs are annotated. The second row shows the parameter marginal distribution of excluded runs.

eliminate any bright background effects and highlights the subtle intensity changes caused by the outward propagation of the CME that may be otherwise hard to observe. Next, we add F and K corona contributions to the synthetic image as they are not captured by AWSoM. The left panel in Figure 5 shows a ratio enhancement image

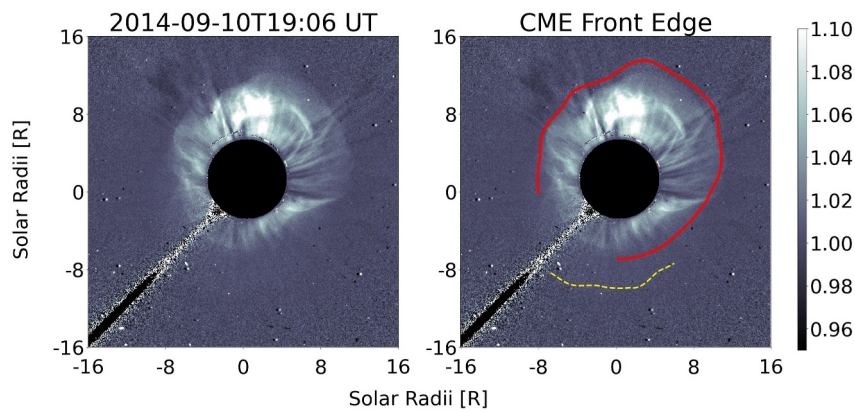

Figure 5. Edge detection demonstration on CME1. *Left*: One snapshot of the LASCO C3 WL ratio enhancement image of 2014-09-10 18:00 UT CME event. *Right*: Automatically detected CME front edges. The red solid line represents the edge used in this study. The yellow dashed line corresponds to the fast moving segment discussed in Section 4.2.

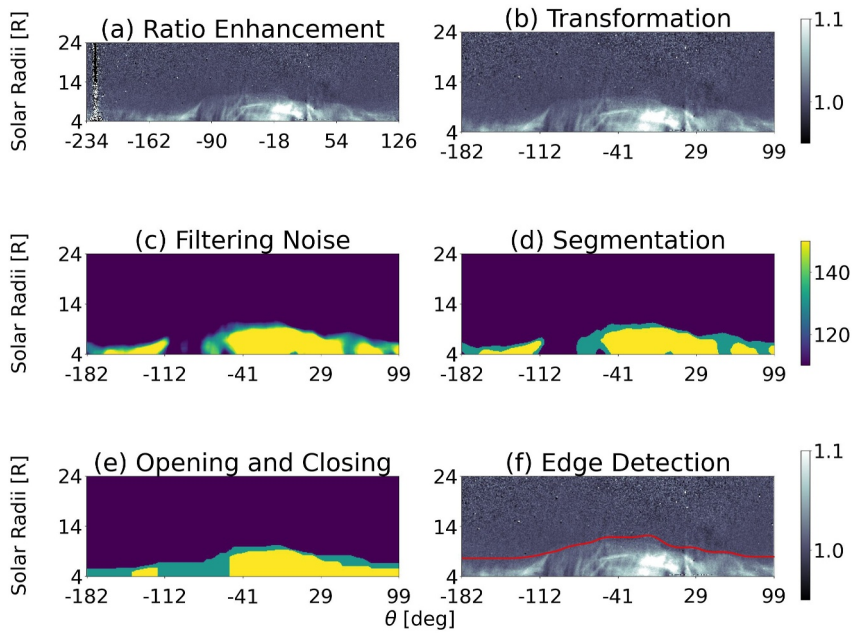

Figure 6. Edge detection workflow on CME1. *Panel* (a): after Step 1, LASCO C3 observation ratio image in polar coordinate. *Panel* (b): after Step 2, image clipping to retain angles at which CME front edge are traceable. *Panel* (c): after Step 3, application of the median filter which has edge-preserving property. *Panel* (d): after Step 4, image segmentation separates CME region from the background. *Panel* (e): after Step 5, opening operation and closing operations for morphological post-processing of segmented image. *Panel* (f): after Step 6, WL images in polar coordinate with detected edges in red.

for CME1 (2014-09-10 18:00 UT CME event). We will use this event as an example to demonstrate the subsequent algorithm steps.

Step 2: Transformation. In order to facilitate comparison and better estimate the CME propagation speed in the radial direction, one common practice (Kirnosov et al., 2015; Olmedo et al., 2008; Wang et al., 2019a, 2019b) is to transform the WL ratio enhancement image from the original ring-shaped image in Cartesian coordinates, $I_o(x, y)$, to a rectangular image in polar coordinates, $I_n(\theta, r)$. We transform the image into 512 $\times$ 128 pixels in the θ and r directions. Subsequently, the transformed image is constrained in the θ coordinate to the interval of interest where the CME has a traceable front. For halo CME observations, we only remove the angles blocked by the physical support of the instrument. This step is done manually due to the potential presence of irrelevant brightness features in the observational image, which may lead to artificial or incorrect boundaries. The coronagraph observation of CME2 shown in Figure 7 illustrates such cases, where a tail-like structure exists that may cause errors if not removed. See Figure 6 panels (a) and (b). The resulting image is an $N' \times M$ matrix with $M = 128$ and $N' \leq 512$ being the number of pixels in the angular direction covering the angles of interest.

Step 3: Noise Filtering. Observation WL images are contaminated with artifacts such as salt-and-pepper (SAP) noise and minor stray marks primarily induced by the LASCO instrument (Brueckner et al., 1995). A common approach to remove SAP noise is by applying a median filter, which additionally, has the edge-preserving property. Figure 6 panel (c) illustrates this idea, where the SAP noise is removed and the CME front edges are well-preserved after applying a median filter of size 7 $\times$ 7. As the median filter also smooths out local extreme pixel values, we apply it to both the observation and the synthetic WL images.

Step 4: Segmentation. This step aims at segmenting CME features from the filtered WL image. While various image segmentation techniques are available (Zaitoun & Aqel, 2015), we choose a clustering-based segmentation approach (Coleman & Andrews, 1979) since we anticipate CME features to be connected rather than isolated as the CME propagates outward away from the Sun. Specifically, we choose the K-means clustering (Lloyd, 1982; MacQueen, 1967), an unsupervised clustering method that assigns to the ith pixel a cluster label $Z_i \in \{1, 2, \ldots, K\}$ (from K candidate clusters) based on its pixel value X_i. For a given K, the optimal pixel cluster labels Z_i^* and cluster center locations $\mu_k^*, k = 1, 2, \ldots, K$ can be found by minimizing the within-cluster-sum-of-square (WCSS) value:

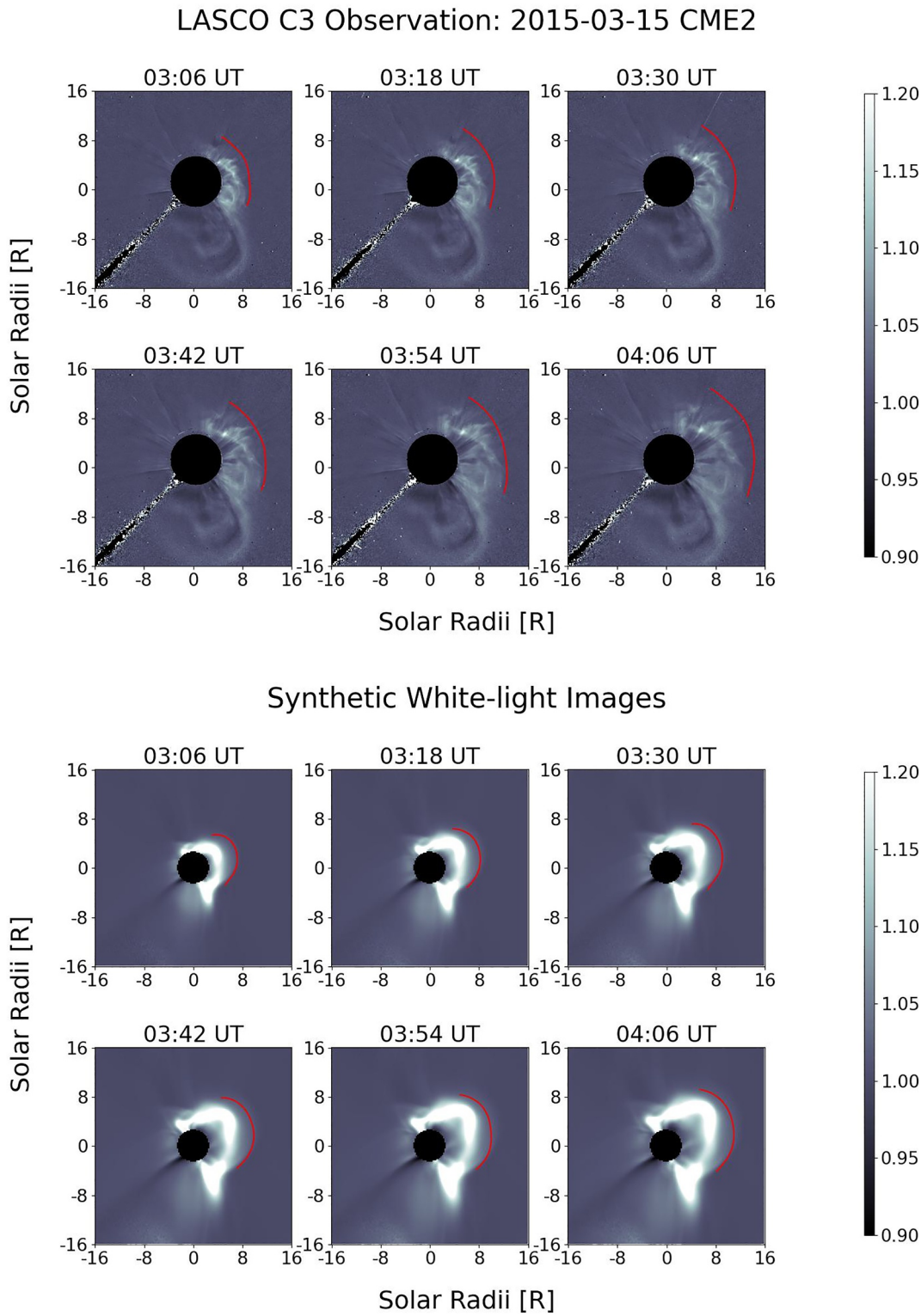

Figure 7. Snapshots of WL images from CME2 (2015-03-15 01:48 UT CME event). Red curves represent the detected edges. *Top*: LASCO C3 observation WL images. *Bottom*: Synthetic WL images from a simulation.

$$\{Z^*, \mu^*\} = \underset{Z,\mu}{\operatorname{argmin}} \left\{ W_K(Z,\mu) := \sum_{k=1}^{K} \sum_{i:Z_i=k} \|X_i - \mu_k\|^2 \right\}. \quad (1)$$

The number of clusters, K, is a tunable hyperparameter. In order to find the best K that avoids overfitting (e.g., choosing a large K may result in many small clusters, each with only a few points or even a single point, that are usually associated with background noise), we seek a K such that W_K explains 95% of the total variance V, which can be written as:

$$V = \sum_{i=1}^{MN'} \|X_i - \bar{X}\|^2 = W_K(Z^*, \mu^*) + \sum_{k=1}^{K} N_k \|\mu_k^* - \bar{X}\|^2, \quad (2)$$

where N_k is the number of points assigned to the k-th cluster. In other words, we choose the smallest K such that $W_K(Z^*, \mu^*)/V \geq 0.95$. Figure 6 panel (d) shows the image after segmentation.

Step 5: Opening and Closing. Occasionally the segmented image contains artifact spikes and blobs in the background. We include two morphological operations, *opening* and *closing*, to further clean the segmented image. An opening operation is a composite of two basic morphological operations: erosion followed by dilation (Soille, 2004), while using the same structuring element (SE). It removes noise and small objects from an image while preserving the overall shape and connectivity of larger objects. A closing operation is the reverse of opening, involving dilation followed by erosion. It is effective at closing small gaps in object boundaries and connecting objects that are almost touching. Empirically, we find that when applied to observation WL images, the combination of the two operations effectively removes noise while preserving the connectivity of CME features. When applied to synthetic WL images, the closing operation is often unnecessary. Figure 6 panel (e) shows the effect of this two-step procedure on LASCO C3 observation WL image.

Step 6: Edge Detection. The coordinates of edge points are found by scanning each column from top to bottom and identifying the first value that differs from the background. The resulting set of edge points is a vector of f_i values describing the discrete radius values of edge occurrence at each polar angle θ_i. The edge detection is then finalized by applying the Savitzky–Golay filter (Savitzky & Golay, 1964) to the set of edge points, which fits successive subsets of adjacent edge points with a cubic polynomial by the method of linear least squares to ensure smoothness. Figure 6 panel (f) shows the edge detection in polar coordinates while the right panel of Figure 5 depicts the corresponding edge curve in the original Cartesian coordinates.

We further demonstrate the overall edge detection algorithm applied to the CME2 (2015-03-15 01:48 UT CME event) WL images in Figure 7, showing six snapshots of both LASCO C3 observations and synthetic WL images from a simulation, together with their detected edges. The angles are selected according to the observation such that the CME part is covered.

3.4. Observation-Simulation Comparison

The resulting CME front edges allow quantitative comparison between LASCO C3 observations and synthetic WL images from simulations. In this study, we target comparisons in terms of height-time measurement and angle-averaged speed overshoot.

CME catalogs typically contain the observed CME speed based on height-time plots, where the height is defined as the maximum edge value over the angles θ:

$$h(t) = \max_{\theta} f(\theta, t), \quad (3)$$

with $f(\theta, t)$ denoting the extracted edge radial value at angle θ and time t. Subsequently, the CME propagation speed v can be estimated directly by performing a linear fit of these height-time measurements $h(t)$ and then taking the slope. Although this quantity is a popular choice to characterize CME propagation, it also has several limitations described below.

1. The maximum CME edge values (i.e., the height values defined in Equation 3) generally occur at different angles for different times. Consequently, the CME propagation speed based on such mixture of different angles lacks a straightforward interpretation. It does not necessarily correspond to the maximal speed over all propagation angles, nor does it represent an average speed of a specific angle.
2. The height defined in Equation 3 loses angular information, since the maximum operator is a summary statistic that only reports the single extreme value across all angles. In fact, angular width is already recognized as an important feature of CMEs in several studies, for example, by Liu et al. (2018).
3. While a general agreement in morphology can be achieved between synthetic WL images and LASCO observations, an exact match is always challenging, maybe even impossible to achieve. Using this quantity to compare simulation with observation emphasizes the discrepancy in the outermost segment of CMEs.

We propose instead to characterize the CME propagation speed by the average speed over all angles of interest:

$$\bar{v} = \frac{1}{N'}\sum_i v(\theta_i), \tag{4}$$

where $v(\theta_i)$ is the speed estimate at angle θ_i obtained by performing a linear least squares fit to $f_i(t) = f(\theta_i, t)$ for each i and then taking the slope. The comparison between simulation and observation is then made through the error metric:

$$\eta = \frac{\bar{v}_{\text{sim}} - \bar{v}_{\text{obs}}}{\bar{v}_{\text{obs}}} \times 100\%, \tag{5}$$

where $\bar{v}_{\text{sim}}$ and $\bar{v}_{\text{obs}}$ are the calculated average speeds for the simulation and observation, respectively. This quantity reflects the relative overshoot error of angle-averaged propagation speed of simulation compared to observation. We refer to η as the WL speed overshoot hereafter. Note that this metric primarily measures the radial propagation speed of the CME front, while lateral velocities only contribute indirectly.

3.5. CME Arrival Time Initial Estimation at 1 au

The CME arrival time at Earth's position is initially estimated (i.e., prior to any background corrections and DA) for each individual simulation or observational time-series based on the solar wind speed simulated/measured at 1 au. For simulation, the CME arrival time is defined to be the time that corresponds to the first significant change in simulated solar wind speed. For observation, in general such property holds as well. Therefore, the arrival time can be viewed as a *change point* (Basseville & Nikiforov, 1993) in the time series of solar wind speed. It is important to note that even without a CME there can be gradual or even sharp changes in speed, for example, due to corotating interaction regions (CIRs). Consequently, identifying the CME shock arrival is not always straightforward.

We propose a shock detection algorithm to automatically determine the CME transit time for both observational and simulated data. The CME arrival time can then be obtained by adding CME transit time to the event start time. The algorithm is based on statistical hypothesis testing. We track the hourly-averaged solar wind speed over a 10-hr-wide sliding window. The CME arrival time is identified by a sudden increase in the speed of the 10-hr window that has a probability less than 2.5% to be random change. We formalize the algorithm in Appendix A. Figure 8 demonstrates this algorithm on CME2 (2015-03-15 01:48 UT CME event) using 1 hr averaged OMNI data. The algorithm detects the *change point* when the F statistics of the sliding window ($F = 25.09$) exceeds the threshold value ($F_{0.975}(9,9) = 4.03$) under statistical significance level $\alpha = 0.025$. We validated that the detected change point matches true (i.e., expert identified) arrival time.

3.6. Background Solar Wind Correction

The CME arrival time estimated for a simulation can be further improved by taking into account the difference in the background solar wind speed between the simulation and the observation. This difference is typically not very large since we constrain the parameters of the background model to match L1 observations, but often it is not negligible either, as shown in Figure 2.

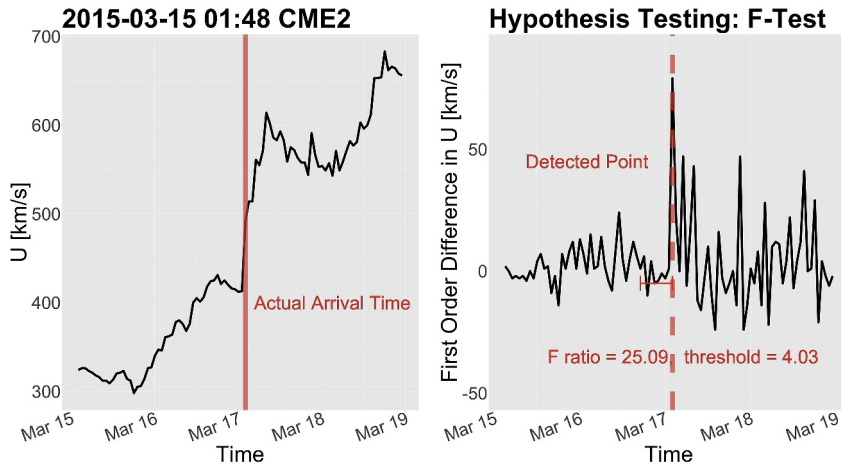

Figure 8. Applying the shock detection algorithm to 1 au plasma speed of CME2 (2015-03-15 01:48 UT CME event) to determine the CME arrival time. *Left*: plasma velocity of 4 days of CME propagation. The solid line represents the actual arrival time. *Right*: The first order difference in plasma speed. The dotted line represents the arrival time determined by the algorithm, which is detected when the F-statistic calculated from a 10-hr sliding window exceeds the threshold value under a statistical significance level of $\alpha = 0.025$. See Appendix A for details.

The CME arrival time is the start time plus the transit time T. The former is defined to be the start time of the flare associated with the CME t_{flare}, which is always known, and the latter can be written as

$$T = \frac{L}{u_{\text{CME}} + u_b}, \tag{6}$$

where u_{CME} is the average CME speed relative to the background over the propagation period, u_b is the background solar wind speed that is assumed to be constant, and L is the distance between Sun and Earth (1 au). Changing the background speed u_b from u_{sim} to u'_{sim} results in a change of T from T_{sim} to

$$T'_{\text{sim}} = \frac{L}{u_{\text{CME}} + u'_{\text{sim}}} = \frac{L}{\frac{L}{T_{\text{sim}}} - u_{\text{sim}} + u'_{\text{sim}}} = \frac{1}{1 + \frac{T_{\text{sim}}}{L}\left(u'_{\text{sim}} - u_{\text{sim}}\right)} T_{\text{sim}}. \tag{7}$$

As expected, the correction makes $T'_{\text{sim}} > T_{\text{sim}}$ if $u'_{\text{sim}} < u_{\text{sim}}$ and decreases it if $u'_{\text{sim}} > u_{\text{sim}}$. Also, the relative correction $|T'_{\text{sim}} - T_{\text{sim}}|/T_{\text{sim}}$ is smaller for fast CMEs, that is, when T_{sim}/L is smaller than for slow CMEs when T_{sim}/L is larger. The error $\left(u'_{\text{sim}} - u_{\text{sim}}\right)$ in the simulated background speed can be estimated by $u_{\text{obs}}\left(t_{\text{flare}}, L_1\right) - u_{\text{sim}}\left(t_{\text{flare}}, L_1\right)$, the difference between the observed and simulated solar wind speeds at L1 at the time of the CME eruption. We use the modified transit time T'_{sim} to better approximate T_{obs}.

3.7. Correlation Analysis and Data Assimilation

The prediction error of CME arrival time from a single run is the same as its prediction error in transit time, which we denote by y:

$$y = T_{\text{obs}} - T'_{\text{sim}}. \tag{8}$$

Using an ensemble of simulations, we conduct a linear correlation analysis between y and WL speed overshoot η. As will be shown later in Section 4.3, the data indicates a strong linear relationship with only a small amount of dispersion. This finding motivates us to build a linear regression model:

$$y = k\eta + b, \tag{9}$$

where k is the slope, and b is the y-intercept which can be interpreted as the bias of the physics-based model.

When making a prediction for a new, previously unseen event, ideally one would like to select a simulation whose η is close to $-\frac{b}{k}$, which leads to a near-zero error for the arrival time prediction (i.e., $y = 0$) per Equation 9. However, as will be illustrated in Figure 12, both k and b differ from event to event. Therefore, we take a probabilistic approach to capture a *distribution* of possible η values that emerges from the event-to-event uncertainty, and constrain the range of η using observational data from historical events through a Bayesian framework. Subsequently, we retain only a subset of simulations from the initial ensemble that fall within this range of η (i.e., the most promising simulations for achieving $y = 0$) and propagate only these simulations all the way to 1 au to finally construct a probabilistic prediction for the CME arrival time.

Under this approach, we treat η as a random variable, and η now has a distribution. Consider a particular historical CME event E, the posterior distribution of η conditioned on $y = 0$ (i.e., the probability density for overshoot of η to lead to zero-error arrival time prediction) can be expressed following Bayes' rule:

$$p(\eta|\, y = 0, E) = \frac{p(\eta|E)p(y = 0|\eta, E)}{\int p(\eta, y = 0|E)\, d\eta} \propto p(\eta|E)p(y = 0|\eta, E) \tag{10}$$

where $p(\eta|E)$ is the prior distribution and $p(y = 0|\eta, E)$ is the likelihood. The prior represents our belief in the overshoot distribution before seeing any data for this historical CME. Leveraging knowledge from previous works on model validation (Jin et al., 2017a; Manchester et al., 2008), we place a Gaussian prior with mean 0 and standard deviation $\sigma = 0.1$. According to the empirical rule of Gaussian distribution, this means that we expect the overshoot to be within $3\sigma = 30\%$. We further assume the prior does not depend on the specific event since one would expect the same model performance if given a randomly-selected CME event. Hence, the prior is $p(\eta|E) \sim \mathcal{N}(0, \sigma^2)$. The likelihood $p(y = 0|\eta, E)$, in principle, can be obtained by running many different simulations and using them to estimate the conditional probability density; however, this would be a highly expensive endeavor. Instead, we use the fitted linear model in Equation 9 as a density estimator to obtain an approximation of the true likelihood, which we denote as $\hat{p}(y = 0|\eta, E)$. Forming this linear model needs training data of η and y, which in turn requires simulations to 1 au, WL observations, and the true arrival time, of this historical event. With the prior and likelihood, we can draw samples from the posterior distribution via Markov chain Monte Carlo (MCMC), specifically the Metropolis–Hastings algorithm (Hastings, 1970; Metropolis et al., 1953).

The above posterior can be computed for multiple different historical events. We can combine them together, by taking the expectation over multiple events, to obtain the final posterior:

$$p(\eta|\, y = 0) = \sum_E p(\eta|\, y = 0, E)\mathbb{P}(E). \tag{11}$$

We assume each historical CME event contributes equally to our knowledge of η, so we take $\mathbb{P}(E)$ to be a constant for all events E. Samples from this final posterior can be obtained by a direct mixing of the (same number of) posterior samples from individual events.

For the prediction on a new CME event, we constrain the WL speed overshoot η to the 95% credible interval of the posterior, which is constructed by taking the $\frac{\alpha}{2}$-th and $\left(1 - \frac{\alpha}{2}\right)$-th quantiles of the MCMC posterior samples with $\alpha = 0.05$. We use 10^5 posterior samples per event to estimate the credible interval at high resolution. Lastly, we use this credible interval to filter the simulation ensemble, and only retain ensemble members whose η values fall within this acceptance range.

3.8. Probabilistic CME Arrival Time Prediction

We now form the probabilistic prediction for CME arrival time at the ensemble level. Given the final ensemble of transit time predictions $T'_{\text{sim};i}$ for ensemble members $i = 1, \ldots, M$, we use the ensemble mean as the final estimate:

$$\overline{T} = \frac{1}{M}\sum_{i=1}^{M} T'_{\text{sim};i}. \tag{12}$$

Under a normal distribution assumption, a confidence interval (CI) for the mean can be established:

$$\mathrm{CI} = \left[\overline{T} - t_{\frac{\alpha}{2}}(M-1)\frac{S_T}{\sqrt{M}},\ \overline{T} + t_{\frac{\alpha}{2}}(M-1)\frac{S_T}{\sqrt{M}}\right], \tag{13}$$

where $S_T = \sqrt{\sum_{i=1}^{M}\left(T'_{\mathrm{sim};i} - \overline{T}\right)^2/(M-1)}$ is the square root of sample variance and $t_\alpha(N)$ represents the α-th quantile of Student's t-distribution with N degrees of freedom. We construct CI under a statistical significance level of $\alpha = 0.05$. The CME arrival time prediction is then obtained by adding the estimated CME transit time to the event start time.

The overall DA procedure can be summarized as follows:

1. Assimilate L1 background solar wind observation into an ensemble of background solar wind simulations, and select the best matching background model.
2. Launch an ensemble of CME simulations based on perturbed EEGGL estimate.
3. Calculate WL speed overshoot for each ensemble member using WL observation.
4. Retain ensemble members whose WL speed overshoot fall within the acceptance range derived from historical events.
5. Simulate the filtered ensemble to 1 au and construct probabilistic prediction.

4. Results

In this section, we demonstrate the proposed method to assimilate WL observational data for constraining the ensemble of CME simulations and finally predicting the CME arrival times.

4.1. Sensitivity Analysis

We begin by performing sensitivity analysis to identify the most important flux rope parameters associated with CME arrival time prediction. The result can allow us to focus on varying a smaller set of perturbation parameters, thereby improving the statistical power of ensembles. In order to avoid contamination of the sensitivity results from event-to-event variations, we carry out the sensitivity analysis using the first batch of 120 CME simulations only from CME3. The experimental design of these runs are shown in Figure 4 and discussed in Section 3.1. While all simulations completed successfully, we further exclude those that provide physically implausible results. Specifically, we extract remote LASCO C3 images and the plasma state along Earth's orbit. A simulation is discarded if it meets any of the following exclusion criteria:

1. the plasma velocity (U) at 1 au exceeds 2,000 km/s;
2. the number density (Np) of the background solar wind at 1 au exceeds 250 cm^{-3};
3. the CME propagation speed overshoot (η) exceeds 200% or falls below −50%;
4. no CME arrival detected based on the solar wind speed at 1 au.

As a result, 23 runs were excluded and 97 runs are retained. Most of the excluded runs meet exclusion criteria 3 and 4 due to their slow solar wind speeds. Figure 4 highlights the excluded runs with cross signs and shows the probability density of the excluded flux rope parameters. These cases are primarily associated with small BStrength and ApexHeight values. Interestingly, specific combinations of Orientation and Helicity also play a significant role in determining exclusion. However, the generality of these exclusion patterns requires additional investigations.

One approach to sensitivity analysis is by calculating variable importance through regression analysis, where a regression model is created to map from the flux rope parameters to the CME arrival time. A specific class of regression models, known as decision tree-based models, can approximate the mapping while simultaneously providing variable importance scores for predictors. In this study, we employ the Bayesian additive regression tree (BART) (Chipman et al., 2010), a Bayesian non-parametric model that uses a sum of decision trees to learn flexible representations of data. Each decision tree consists of a set of splitting rules that partition the parameter space into the subgroups, and a regularization prior is placed to avoid overfitting. The tree predicts the same value for points within the same subgroup. Since every split depends on the value of only a single variable, the variable

importance score for any variable can be estimated by its inclusion proportion—that is, the proportion of times that a variable is chosen as a splitting rule out of all splitting rules among the posterior draws of the BART model (Chipman et al., 2010; Kapelner & Bleich, 2016).

We adopted the BART implementation from the R package `bartMachine` (Kapelner & Bleich, 2016). The inclusion proportion results, reflecting the importance of each flux rope parameter, are shown in Table 3. Among all continuous parameters, BStrength, which is the magnetic field strength of the flux-rope, has the highest inclusion proportion in the trained models thus it is identified as the most important variable. Physically, BStrength is also the most obvious parameter impacting the arrival time, and we vary it in future ensemble runs. Although the inclusion proportions of ApexHeight and Orientation suggest that varying these two variables in future ensemble runs could benefit UQ, we have decided not to vary them in order to keep the ensemble small and fit within our computational budget. With a realistic ensemble size of around 20 members, including more parameters would lead to poor coverage of the sample space, thus increasing the difficulty of parameter inference. Note that, we should not compare the inclusion proportion of Helicity with other continuous variables. This is because a categorical variable with two levels can at most show up once in the splitting rules of a decision tree, while there is no such constraint for continuous variables. However, we choose to vary Helicity for future runs as it can have a significant effect on Bz prediction, which is crucial for quantifying the geo-impact of a CME event in addition to predicting CME arrival time. It is also relatively computationally inexpensive to vary Helicity, as it only increases the ensemble size by a factor of two.

For all three events, CME1–3, we now conduct the second batch of CME simulations with smaller ensembles that only varies BStrength and Helicity, which corresponds from varying the RelativeStrength and Helicity perturbation parameters, while fixing the remaining continuous parameters to the baseline estimate. We extend the range of RelativeStrength from $[0.5, 1.5]$ in Table 2 as needed to make sure we cover both positive and negative values of the velocity overshoot parameter η as shown in Figure 12. In total, each ensemble uses 12 values of BStrength, each with Helicity set to -1 or $+1$, resulting in 24 runs for each event. All runs completed successful. Note that the first batch of simulations for CME3 is used solely for sensitivity analysis and is not included in the subsequent analysis. For the prediction of future events, we cannot afford such a large ensemble size due to the computational cost. Excluding the first batch ensures a fair comparison between events with affordable ensemble sizes.

The CME arrival time error versus BStrength for these simulations are shown in Figure 9. Note that per Equation 8, a positive error means that the model predicts an arrival time earlier than the actual arrival time, and vice versa. The predicted arrival time becomes early when BStrength increases, while for CME1 we observe that the predicted arrival time stops changing when BStrength exceeds around 55 G. Different signs of magnetic helicity also introduce difference to the prediction of the arrival time. Figure 10 shows the difference in the quantities of interest between the run with the smallest arrival time prediction error and its counterpart with flipped sign of helicity. These quantities include the plasma velocity (U), the proton number density (Np), and the magnetic field component in the North/South direction (Bz). We see that both U and Bz can change drastically when the helicity is flipped, while having the helicity set to EEGGL estimate does not always yield better predictions. This suggests that investigating CME with both values of helicity is important.

4.2. Height-Time Measurement Comparison

In this section, we validate our edge detection algorithm against the CDAW catalog. In particular, we will compare the CME height as defined in Equation 3 and the resulting CME (scalar) propagation speed, as they are widely used and available in many CME catalogs. However, discrepancies often exist for the same event across multiple catalogs. For the sake of consistency, we make our validation comparisons with the CDAW CME catalog (Gopalswamy et al., 2009).

We compare our estimates of CME maximal height and CME propagation speed to those in the CDAW catalog where these quantities are estimated from visually identified CME attributes. The hyperparameters of the edge detection algorithm are shown in Table 5. Figure 11 shows the height-time profile comparisons, with the linear fit slopes being the propagation speed estimates. The Spearman's rank correlation coefficients between CME maximal height estimated by our method and that from the CDAW catalog is greater than 0.99 for all three CME events considered in this paper. We see that our estimates are close to that of CDAW for most of CME events we studied, except for CME1. For CME1, processing the available LASCO images becomes more complicated as

Table 3
Inclusion Proportion of Flux Rope Perturbation Parameters in the BART Model for Sensitivity Analysis Using CME3 Simulations

BStrength	ApexHeight	Orientation	Helicity
0.314	0.251	0.244	0.191

several snapshots contain features of another CME prior to CME1. The event time of CME1 is 2014-09-10 18:00, while the prior CME has event time is 2014-09-10 17:24. The data on CDAW show that heights of CME1 were measured at the fastest segment of the leading edge and the measurement position angle of the maximal height is at around 175° counter-clockwise to solar north (see yellow dashed line in Figure 5). However, it is unclear if the outermost segment around this angle, which is faint and blends with the background very quickly after a few snapshots, is truly a part of CME1. In contrast, our method is tracking the bright front of CME1, which has a clear edge and whose maximal height is obtained at around 335° (see solid red line in Figure 5). Maharana et al. (2023) also studied this event and they reached a similar conclusion that CME1 propagated mainly north of the ecliptic. Therefore, the discrepancy mainly comes from the complexity of this event. It is also worth noting that using propagation speed derived from maximal height to describe a CME can be problematic in such cases, since it does not reflect the propagation globally. In fact, for CME1, the estimate of the speed is much larger than that averaged from the major part (brightest region in coronagraph images) of the CME so the result can be misleading if we include this estimate for comparisons of simulation output.

Considering the above comparisons and explanation for CME1, the results overall support that our algorithm is able to provide reliable estimate of the CME front edge in an automated manner.

4.3. Correlation Results and Arrival Time Prediction Without Data Assimilation

In this section, we show the correlation results between arrival time error and WL speed overshoot, and we analyze CME arrival time prediction at both the individual and the ensemble level.

We first apply the edge detection algorithm and shock detection algorithm with background correction introduced in Section 3 to simulation and observational data of each event. The parameters of the two algorithms used for simulation data are listed in Table 4, and the parameters used for observations are listed in Table 5. For the WL edge detection algorithm, these include the range of position angle at which the speed is measured, the size of the median filter, the size of Savitzky–Golay filter, the size of structuring element for opening operation, the size of transformed image in polar coordinate, and the proportion of variance explained by WCSS in K-means clustering. The position angles are measured from solar north in degrees counter-clockwise. For the L1 shock detection algorithm, these include the size of the sliding window, and the background correction factor.

We calculate the WL speed overshoot and the arrival time error for the second batch of simulations of three events. All successful runs are then subjected to the exclusion criteria in Section 4.1. For CME1, out of 24 runs, 4 runs are excluded and 20 runs are retained. For CME2, out of 24 runs, 5 runs are excluded and 19 runs are retained. For CME3, out of 24 runs, 1 run is excluded and 23 runs are retained.

The CME arrival time predictions at the individual ensemble member level are shown in the left column of Figure 12, where we plot scatter plots of speed overshoot η against the arrival time error y for each event. Each point represents a single ensemble member. We find a strong positive correlation between the two quantities with

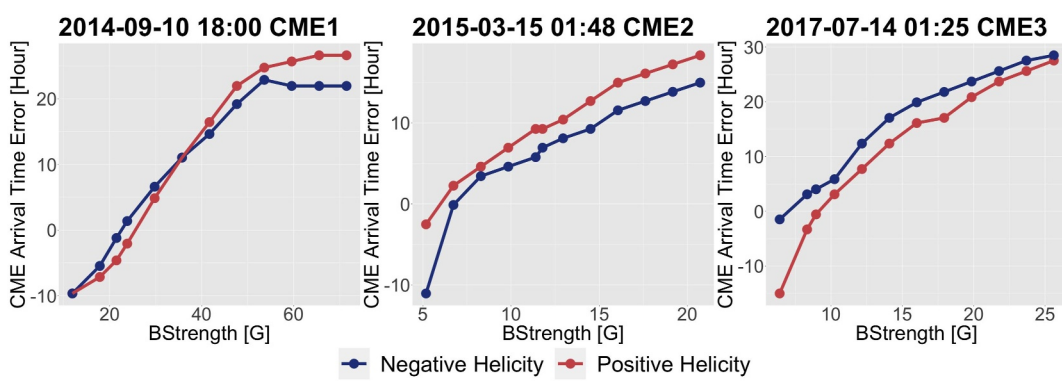

Figure 9. CME arrival time error versus BStrength with positive and negative helicity choices for CME1–3. A positive error means that the model predicts an arrival time earlier than the actual arrival time.

10.1029/2024SW004165

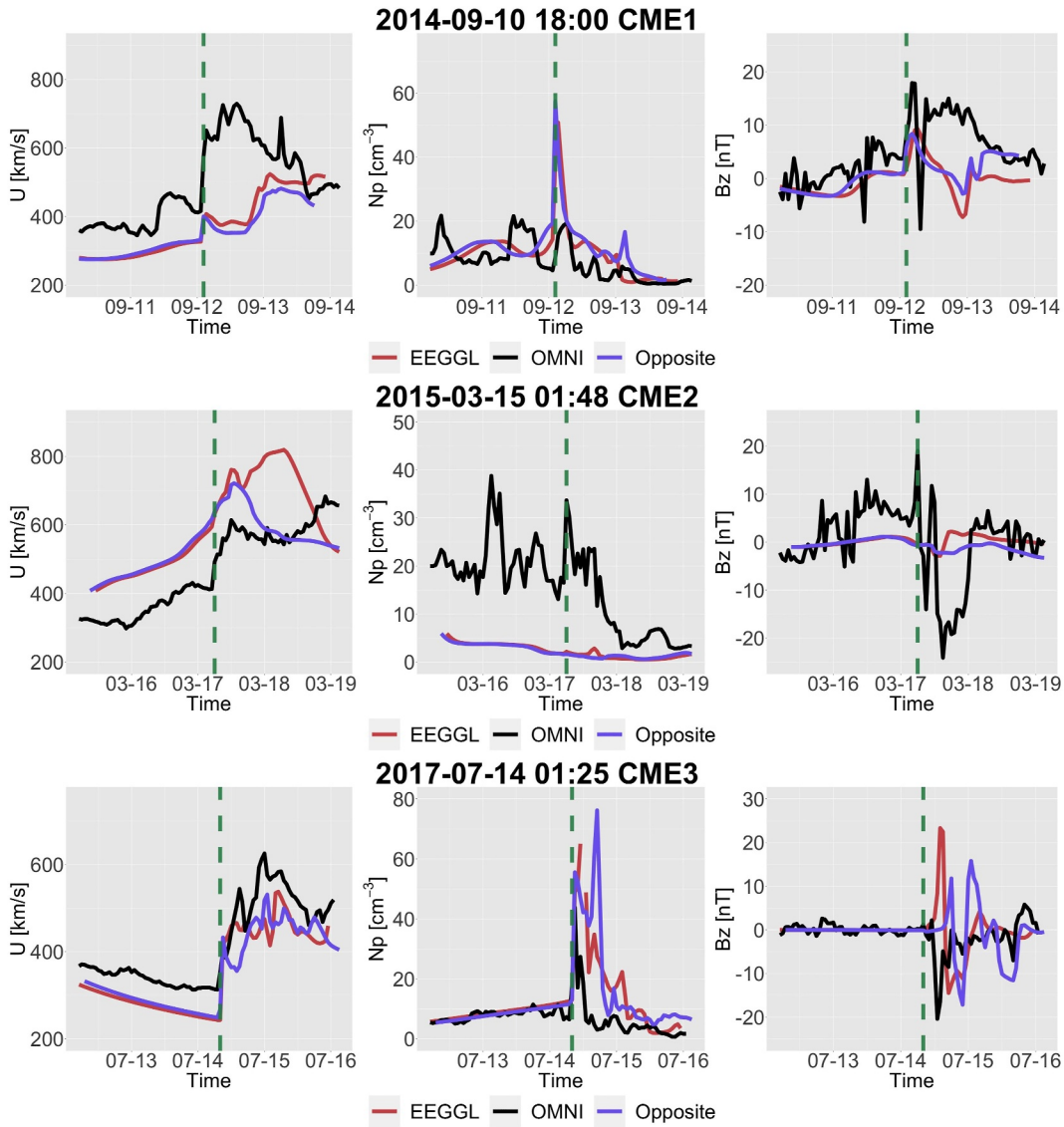

Figure 10. Quantities of interest for CME1–3 in the 1 hr averaged OMNI data (black), simulations with magnetic helicity set to the EEGGL estimate (red), and simulations with sign of helicity opposite from the EEGL estimate (purple). Each row represents a different CME event and each column represents a different quantity. Simulations are shifted to the actual CME arrival time (green dashed line) for better comparisons.

Spearman's rank correlation coefficient exceeding 0.90 for all three events. This indicates that a CME simulation that propagates rapidly at the early stage of simulation would also arrive early at 1 au in the simulation. However, the vertical wide spread of scatter points also indicates that there is uncertainty in the predicted arrival time among simulations even with similar initial propagation speeds. The uncertainty may be due to different reasons:

- WL coronagraph images represent only the integral of brightness along the line of sight, which means they are projections and cannot fully represent the actual 3D objects. In fact, coronagraph images show the plasma velocity in close proximity to the image plane, while the velocity in the direction of Earth is approximately orthogonal to this plane and located behind the occulting disk. Consequently, we have no direct measurement of the speed of the Earth-directed plasma.
- The WL propagation speed is averaged over a range of propagation angles, so two simulations with the same averaged propagation speed do not necessarily match in every propagation angle.

Table 4
Parameters of WL Edge and L1 Shock Detection Algorithms for Simulation Data

Algorithm parameters	CME1	CME2	CME3
White-light Edge Detection			
Position angle	[−182°, 99°]	[244°, 329°]	[209°, 300°]
Size of median filter		13 × 13	
Size of Savitzky–Golay filter		51	
Size of structuring element in opening		11 × 11	
Size of transformed image in polar coordinates		128 × 512	
Proportion of variance explained by WCSS		0.99	
L1 Shock Detection			
Size of sliding window [h]		10	

The least-squares linear regression captures this uncertainty. Prediction uncertainty intervals are given alongside the fitted line, as shown in Figure 12. The results of the least squares fit are summarized in Table 6. The Y-intercept represents the expected arrival time error conditioned on a zero WL speed overshoot. These errors are 7.7, 11, and 10 hr for CME1, CME2, and CME3, respectively. These indicate that the model has a tendency to predict early arrival for all three events. This tendency is further illustrated in the right column of Figure 12, where we compare the WL averaged propagation speed across position angles between the simulations and the observation. The dashed line represents the observation, while each solid line represents a single ensemble member, with color indicating the error in arrival time prediction.

Table 7 contains the summary statistics of the CME transit time prediction errors $T_{\text{obs}} - T'_{\text{sim}}$ based on the initial ensemble without DA. The prediction errors of CME1, CME2, and CME3 are 14.6, 10.9, and 14.7 hr, respectively. The CIs fail to cover the actual arrival time for all three events. We will next discuss and compare the prediction results after DA is performed.

4.4. Arrival Time Prediction With Data Assimilation

The accuracy of CME arrival time prediction can be improved by assimilating WL observations into the simulations, which allows us to down-select the initial simulation ensemble based on the WL speed overshoot. Recall that, the ideal overshoot, which is expected to give no error in arrival time prediction, is the X-intercept of the fitted linear equation, as shown in Figure 12. The variation in X-intercepts across different events indicates additional uncertainties induced by the nature of events, aside from the uncertainty mentioned in Section 4.3. Our proposed Bayesian procedure can systematically incorporate uncertainty from these sources.

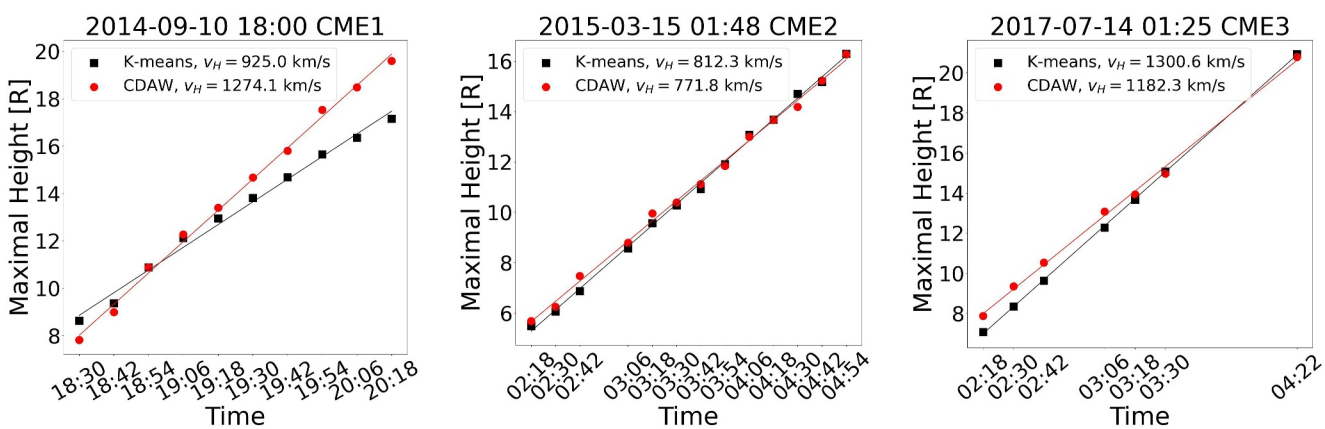

Figure 11. Height-time profiles of three CME events produced by our method and by CDAW catalog. The straight lines are fitted from scatter points by the method of linear least squares regression. The average speeds are derived based on regression coefficients.

Table 5
Parameters of WL Edge and L1 Shock Detection Algorithms for Observational Data

Algorithm parameters	CME1	CME2	CME3
White-light edge detection			
Position angle	[−182°, 99°]	[244°, 329°]	[209°, 300°]
Size of median filter	15 × 15	7 × 7	11 × 11
Size of Savitzky-Golay filter	51	51	51
Size of structuring element in opening	11 × 11	7 × 7	11 × 11
Size of transformed image in polar coordinates		128 × 512	
Proportion of variance explained by WCSS		Optimized by snapshots	
L1 Shock Detection			
Size of sliding window [h]		10	

To evaluate the DA method, we perform leave-one-event-out cross-validation across the three events. Specifically, we use data of two events to construct the acceptance range of the WL speed overshoot, then we filter the ensemble of the third event to construct the final ensemble solution. The results are presented in Table 7. The DA procedure reduces the prediction error from 14.6, 10.9, and 14.7 hr to 0.8, 9, and 5.4 hr for CME1, CME2 and CME3, respectively. The mean absolute error is reduced from 13.4 to 5.1 hr. In addition, the CI for arrival time is improved in each event. However, we note that the CIs still fail to cover the actual arrival time for CME2 and CME3 even after DA, suggesting that the model has additional errors that are not captured by the analysis. We also include the results without background solar wind correction in Table 7. The mean absolute error of three events increases to 6.1 hr (from 5.1) when the correction is not applied. There is a significant decrease in prediction performance for CME2 that has the slowest initial WL speed among the three events as shown in Figure 11, while there is a large discrepancy between the simulated (about 404 km/s) and observed (about 327 km/s) background solar wind speeds. This highlights the usefulness of the background solar wind correction.

Although the results are obtained from only three CME events, we observe that the acceptance ranges of WL speed overshoot have similar values, indicating their transferability across events to a certain extent. This shows the prospect of learning the acceptance range from past events and leveraging early-stage data assimilation to constrain CME flux rope parameters for predicting new events, thereby improving CME arrival time predictions at 1 au. As data from additional CME events become available, we can continually update the acceptance range following this Bayesian framework. Importantly, the DA procedure offers computational savings, as the filtering only requires synthetic WL simulation of 3 hr of physical time, currently taking approximately 13 hr of wall-clock time on 3,304 CPUs on Frontera. Completing the simulation to obtain 1 au results would take approximately 16 hr on the same number of cores. Therefore, down-selecting a smaller ensemble of simulations for completion to 1 au reduces the overall computational cost for accurate CME arrival time prediction.

Lastly, we present the ensemble prediction for three quantities of interest (QoIs), U, Np, and Bz, with and without DA in Figures 13–15 for CME1–3. The point estimate, given by the ensemble average and shown as the solid red line in the figures, is calculated after aligning all simulated quantities. The prediction interval is constructed based on 0.025-th and 0.975-th quantiles of ensemble members. These QoIs predictions are then aligned with the observations to facilitate comparison, and the CME arrival predictions are shown separately as dashed red lines. We observe that while there is significant improvement in CME arrival time prediction with DA, the QoIs prediction with the filtered ensemble does not necessarily outperform the prediction without filtering in terms of accuracy to the OMNI observational data. This shows the difficulty of predicting these QoIs and the misalignment of calibrating the model to predict the QoIs versus the arrival time.

Table 6
Linear Regression Results and Correlation Coefficients for the Three CME Events

	CME1	CME2	CME3
Slope (k)	15	15	22
Y intercept (b) in hours	7.7	11	10
Spearman's rank correlation	0.94	0.98	0.99

5. Conclusions

In this study, we demonstrated the benefit of combining MHD CME simulations with DA to improve CME arrival time prediction. We performed a correlation analysis between the LASCO WL edge speed overshoot

Table 7
Arrival Time Prediction Errors (in Hours) for the Three CMEs, With and Without DA, and With and Without Background Correction

With background correction	CME1	CME2	CME3	MAE
Prediction error with DA	0.8 ± 3.7	9 ± 1.4	5.4 ± 3.9	5.1
Prediction error without DA	14.6 ± 4.9	10.9 ± 2.0	14.7 ± 4.5	13.4
Acceptance range of η	[−52%, −1%]	[−30%, 10%]	[−52%, 10%]	[−45%, 6%]
Ensemble size without DA	20	19	23	–
Ensemble size with DA	6	10	11	–
Without background correction	**CME1**	**CME2**	**CME3**	**MAE**
Prediction error with DA	−2.7 ± 4.2	12.1 ± 1.2	3.5 ± 4.2	6.1
Prediction error without DA	12.8 ± 5.5	13.5 ± 1.8	13.4 ± 4.8	13.2
Acceptance range of η	[−43%, 3%]	[−27%, 12%]	[−43%, 12%]	[−38%, 9%]
Ensemble size without DA	20	19	23	–
Ensemble size with DA	6	9	11	–

(simulation vs. observation) and the arrival time error at L1, and a very high degree of correlation was observed. The Spearman's rank correlation coefficients are greater than 0.90 for the three CME events considered in this paper. We developed a Bayesian DA procedure to quantify the uncertainty in the WL speed overshoot and filter the initial ensemble based on LASCO WL observations. Fusing DA into a physics-based model framework substantially reduced the MAE of CME arrival time prediction from 13.4 to 5.1 hr over the three CME events. Another benefit of this DA procedure is to reduce the computational cost by only propagating a smaller and more reliable ensemble to 1 au. While CME arrival time predictions with similar accuracy can be achieved using simpler and computationally less expensive models (e.g., Chierichini et al. (2024); Zhao & Dryer (2014)), the Sun-to-Earth 3D MHD simulation provides comprehensive predictions for plasma parameters and magnetic field in the full 3D domain and is also suitable for physical interpretation of the CME evolution and propagation through the heliosphere.

The following are the key contributions of this work:

1. *Real-time availability*: We proposed DA of L1 observations of the background solar wind and SOHO LASCO WL coronagraph observations. Both data are readily available, which makes this approach highly viable for forecasts in an operational setting with further refinement and scaling.
2. *Reduction of uncertainty*: We have established the high correlation between the simulated CME's WL propagation speed and the CME arrival time, based on which we developed a data assimilation procedure to filter the initial ensemble by the WL speed overshoot and reduce uncertainties in flux rope parameters.
3. *Predictive uncertainty estimation*: We proposed a probabilistic CME arrival time prediction based on the data-assimilated ensemble to cover the uncertainty and achieve more accurate results at 1 au.
4. *Automation*: We developed an automated processing pipeline for both observational and synthetic LASCO WL images that is robust to complex CME morphology and also efficient when applied to ensemble simulations. We designed an automated processing pipeline for estimating CME arrival time in ensemble simulations with the background solar wind correction.
5. *Flexibility*: The down-selecting criterion in the DA is conducted within a Bayesian framework, which can be continually updated as more events are studied.

There are several avenues of future work that can further improve the current method. First, robustness of the DA procedure can be strengthened by incorporating a more extensive range of events to cover diverse types of CMEs. We note that the simulations for the three CMEs studied in this work exhibit similar characteristics, consistently predicting shorter transit times compared to the observed data, even when synthetic white light images align well with coronagraph observations. While the results may suggest a systematic model bias, the small sample size of CMEs could also skew the analysis, as the model may exhibit different transit time behaviors when tested on a broader variety of CMEs. Second, the spatial and temporal information contained in the sequence of WL images is lost when the image is summarized into a scalar quantity of WL speed overshoot, and other WL features can

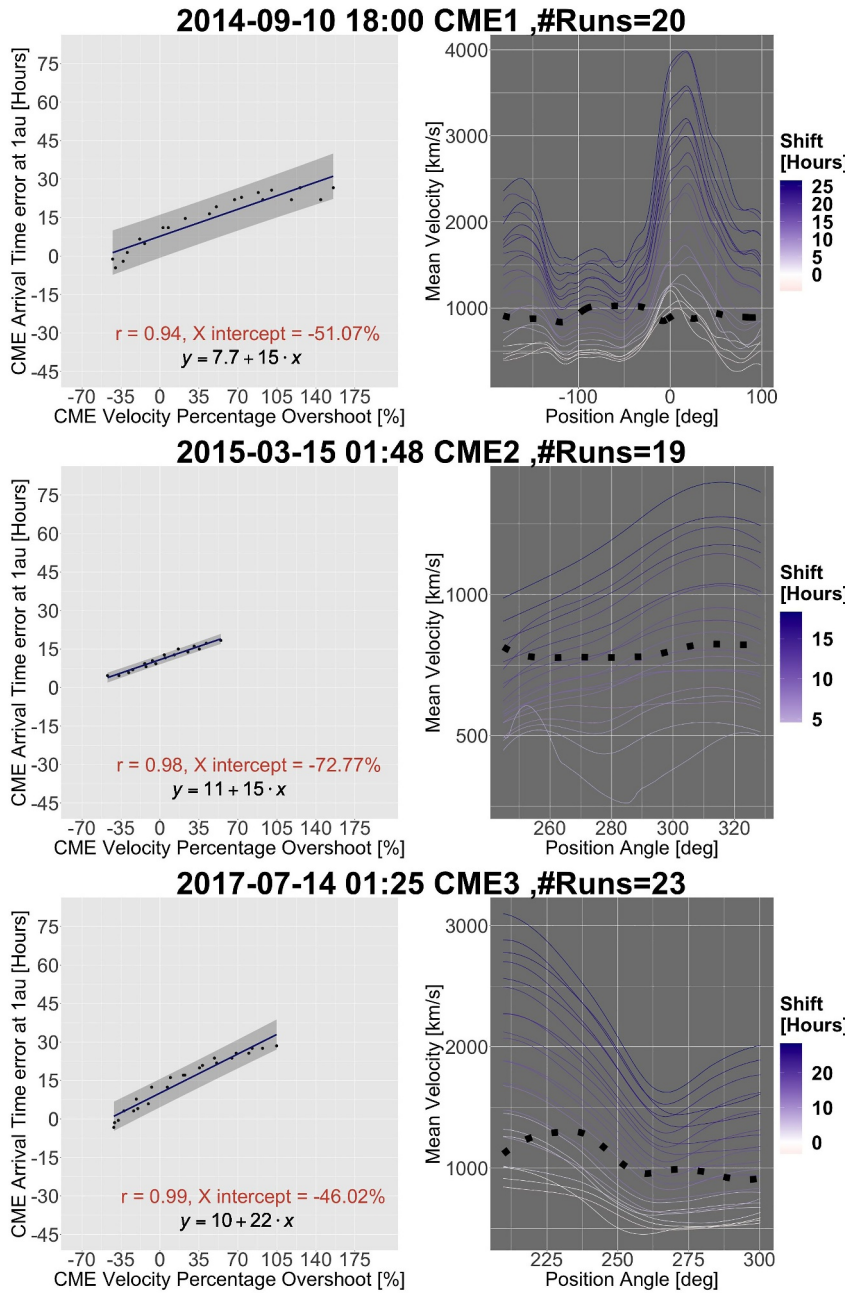

Figure 12. *Left column*: Linear regression analysis of the CME arrival time error versus the CME propagation speed overshoot. A positive error means that the model predicts an arrival time earlier than the actual arrival time. Shaded area represents the 95% prediction interval. *r* is the Spearman's rank correlation coefficient. *Right column*: The averaged CME propagation speed as a function of the position angle. Simulations are color-coded based on the prediction error in CME arrival time. The dashed line represents the observation.

potentially even better inform the CME arrival time as well as other quantities of interest such as U, Np, and Bz. A key challenge to finding informative features lies in identifying a relationship that is transferable across different events. Third, the relatively high spread in arrival time, such as in CME2, suggests the need of calibration of feasible range of flux rope parameters.

10.1029/2024SW004165

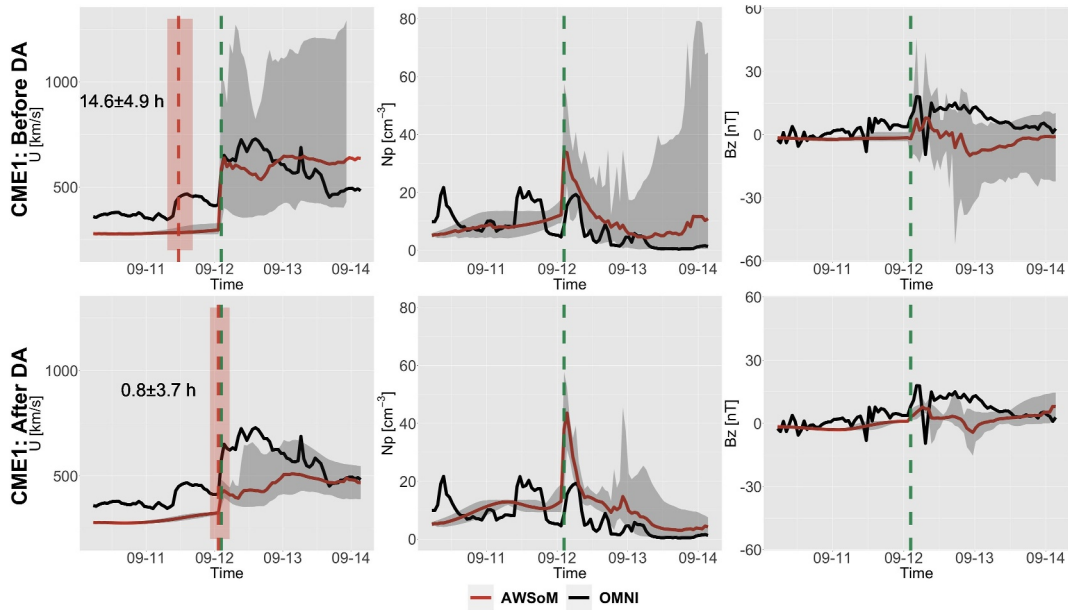

Figure 13. Ensemble predictions of CME arrival time and quantities of interest (QoIs) in the 1 hr averaged OMNI data (black) for CME1 before (the first row) and after (the second row) DA. QoIs predictions are aligned with the actual CME arrival time (green and dashed) for better comparison, with CME arrival time predictions shown separately as a dashed red line. Both the solid and dashed red lines indicate the ensemble average. The shaded area represents the 95% prediction interval.

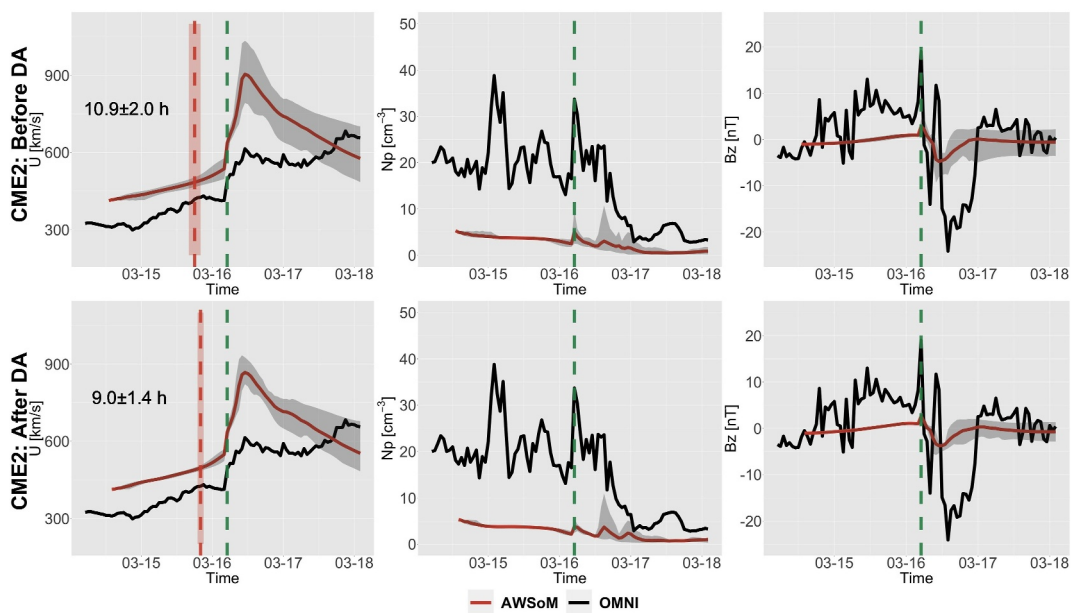

Figure 14. Ensemble prediction of CME arrival time and quantities of interest in the 1 hr averaged OMNI data for CME2. Notations are the same as in Figure 13.

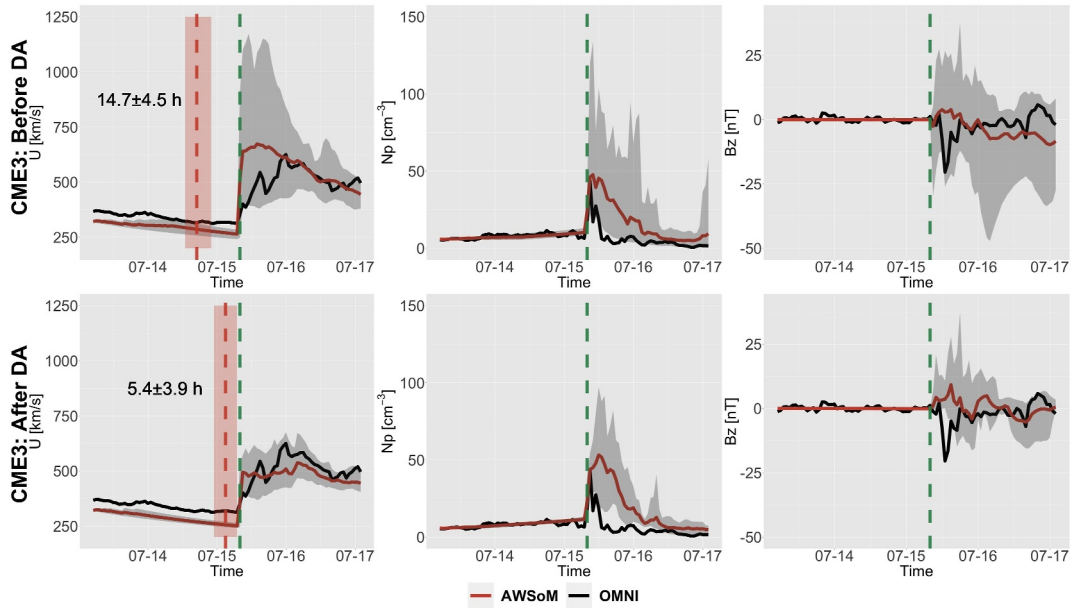

Figure 15. Ensemble prediction of CME arrival time and quantities of interest in the 1 hr averaged OMNI data for CME3. Notations are the same as in Figure 13.

Appendix A: Shock Detection Algorithm

Let t denote the time and $u(t)$ denote the solar wind speed at time t. We assume that the difference of solar wind speed:

$$x_t = u(t+1) - u(t), \tag{A1}$$

follows a Gaussian distribution $x_t \sim \mathcal{N}(\mu, \sigma^2)$, with mean μ and variance σ^2. Then if we apply a sliding window with bandwidth h to the time series x_t, and if we denote the sample variance calculated on the i-th window as

$$S_i^2 = \frac{1}{h-1}\sum_{m=1}^{h}(x_{i+m} - \bar{x})^2, \tag{A2}$$

it is easy to show that

$$\frac{(h-1)S_i^2}{\sigma^2} \sim \chi^2(h-1), \tag{A3}$$

and so

$$F_i = \frac{S_{i+1}^2}{S_i^2} \sim \mathcal{F}(h-1, h-1), \tag{A4}$$

where $\bar{x} = \frac{1}{h}\sum_{m=1}^{h} x_{i+m}$ is the sample mean over the sliding window, $\chi^2(h)$ is the Chi-square distribution with degree of freedom h, and $\mathcal{F}(n,m)$ is the F distribution with n and m degrees of freedom.

Using the statistics F_i, the *change point* is detected by statistical hypothesis testing. The null hypothesis H_0 assumes that the mean value of x_t on the i-th window, μ_i, is equal to the mean value of x_t on the $(i+1)$-th window μ_{i+1}, that is,

$$H_0 : \mu_i = \mu_{i+1} \tag{A5}$$

Since the arrival of a CME results in an upward jump in velocity, we use an upper one-tailed F-test with prescribed significance level α to detect the *change point*:

$$s = \min\{t | F_{t-h-1} > F_{1-\alpha}(h-1, h-1)\}, \tag{A6}$$

where $F_\alpha(h-1, h-1)$ is the upper percentile of an F-distribution with $h-1$ and $h-1$ degrees of freedom.

Data Availability Statement

The scripts and routines used to produce the results in this manuscript are available in the University of Michigan (UM) Library Deep Blue Data Repository at Chen et al. (2024). The SWMF source code is publicly available through Gombosi et al. (2021). The R package MaxPro is publicly available through Ba and Joseph (2015).

Acknowledgments
This work is supported by the National Science Foundation (NSF) under Grant PHY-2027555: "SWQU: NextGen Space Weather Modeling Framework Using Data, Physics and Uncertainty Quantification." The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources under the "LRAC: NextGen Space Weather Modeling Framework Using Data, Physics and Uncertainty Quantification" allocation on the Frontera supercomputer. W. Manchester was partially supported by NASA Grant 80NSSC21K1685, and is also funded by NSF Solar Terrestrial Grant 2323303 and NASA LWS Grants 80NSSC22K0892 and 80NSSC24K1104. N. Sachdeva is also supported by NASA LWS Grants 80NSSC24K1104 and 80NSSC22K0892. Z. Huang is also supported by the NSF Solar Terrestrial Grants 2323303 and the NASA R2O2R Grant 80NSSC23K0450. The SOHO/LASCO data used here are produced by a consortium of the Naval Research Laboratory (USA), Max-Planck-Institut fuer Aeronomie (Germany), Laboratoire d'Astronomie (France), and the University of Birmingham (UK). SOHO is a project of international cooperation between ESA and NASA. We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb (or CDAWeb or ftp) service, and OMNI data (Papitashvili & King, 2020). We also acknowledge use of CDAW CME Catalog. This CME catalog is generated and maintained at the CDAW Data Center by NASA and The Catholic University of America in cooperation with the Naval Research Laboratory.

References

Alobaid, K. A., Abduallah, Y., Wang, J. T. L., Wang, H., Jiang, H., Xu, Y., et al. (2022). Predicting CME arrival time through data integration and ensemble learning. *Frontiers in Astronomy and Space Sciences*, *9*, 1013345. https://doi.org/10.3389/fspas.2022.1013345

Alshehhi, R., & Marpu, P. R. (2021). Detection of coronal mass ejections using unsupervised deep clustering. *Solar Physics*, *296*(6), 104. https://doi.org/10.1007/s11207-021-01854-w

Ba, S., & Joseph, V. R. (2015). Maxpro: Maximum projection designs [Software]. *CRAN*. https://doi.org/10.32614/CRAN.package.MaxPro

Baruah, Y., Roy, S., Sinha, S., Palmerio, E., Pal, S., Oliveira, D. M., & Nandy, D. (2024). The loss of Starlink satellites in February 2022: How moderate geomagnetic storms can adversely affect assets in low-Earth orbit. *Space Weather*, *22*(4), e2023SW003716. https://doi.org/10.1029/2023SW003716

Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: Theory and application*. Prentice Hall.

Brueckner, G. E., Howard, R. A., Koomen, M. J., Korendyke, C. M., Michels, D. J., Moses, J. D., et al. (1995). The large angle spectroscopic coronagraph (LASCO): Visible light coronal imaging and spectroscopy. *Solar Physics*, *162*(1–2), 357–402. https://doi.org/10.1007/BF00733434

Cargill, P. J., Chen, J., Spicer, D. S., & Zalesak, S. T. (1996). Magnetohydrodynamic simulations of the motion of magnetic flux tubes through a magnetized plasma. *Journal of Geophysical Research*, *101*(A3), 4855–4870. https://doi.org/10.1029/95JA03769

Carrassi, A., Bocquet, M., Bertino, L., & Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *WIREs Climate Change*, *9*(5), e535. https://doi.org/10.1002/wcc.535

Chandran, B. D. G., Dennis, T. J., Quataert, E., & Bale, S. D. (2011). Incorporating kinetic physics into a two-fluid solar-wind model with temperature anisotropy and low-frequency Alfvén-wave turbulence. *The Astrophysical Journal*, *743*(2), 197. https://doi.org/10.1088/0004-637X/743/2/197

Chen, H., Sachdeva, N., Huang, Z., van der Holst, B., Manchester, W., Jivani, A., et al. (2024). Decent estimate of CME arrival time from a data-assimilated ensemble in the Alfvén wave solar atmosphere model [Dataset]. *University of Michigan - Deep Blue Data*. https://doi.org/10.7302/b8ah-wq93

Chierichini, S., Liu, J., Korsós, M. B., Del Moro, D., & Erdélyi, R. (2024). CME arrival modeling with machine learning. *The Astrophysical Journal*, *963*(2), 121. https://doi.org/10.3847/1538-4357/ad1cee

Chipman, H. A., George, E. I., & McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *Annals of Applied Statistics*, *4*(1). https://doi.org/10.1214/09-AOAS285

Coleman, G., & Andrews, H. (1979). Image segmentation by clustering. *Proceedings of the IEEE*, *67*(5), 773–785. https://doi.org/10.1109/PROC.1979.11327

Dere, K. P., Landi, E., Mason, H. E., Monsignori Fossi, B. C., & Young, P. R. (1997). CHIANTI - An atomic database for emission lines: I. Wavelengths greater than 50 Å. *Astronomy & Astrophysics Supplement Series*, *125*(1), 149–173. https://doi.org/10.1051/aas:1997368

Emmons, D., Acebal, A., Pulkkinen, A., Taktakishvili, A., MacNeice, P., & Odstrcil, D. (2013). Ensemble forecasting of coronal mass ejections using the WSA-ENLIL with CONED Model. *Space Weather*, *11*(3), 95–106. https://doi.org/10.1002/swe.20019

Evensen, G. (2009). The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems*, *29*(3), 83–104. https://doi.org/10.1109/MCS.2009.932223

Evensen, G., Dee, D. P., & Schröter, J. (1998). Parameter estimation in dynamical models. In E. P. Chassignet & J. Verron (Eds.), *Ocean modeling and parameterization* (pp. 373–398). Springer Netherlands. https://doi.org/10.1007/978-94-011-5096-5_16

Fisk, L. A. (1996). Motion of the footpoints of heliospheric magnetic field lines at the Sun: Implications for recurrent energetic particle events at high heliographic latitudes. *Journal of Geophysical Research*, *101*(A7), 15547–15553. https://doi.org/10.1029/96JA01005

Fisk, L. A., & Schwadron, N. A. (2001). The behavior of the open magnetic field of the sun. *The Astrophysical Journal*, *560*(1), 425–438. https://doi.org/10.1086/322503

Ghil, M., & Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography. In *Advances in geophysics* (Vol. 33, pp. 141–266). Elsevier. https://doi.org/10.1016/S0065-2687(08)60442-2

Gibson, S. E., & Low, B. C. (1998). A time-dependent three-dimensional magnetohydrodynamic model of the coronal mass ejection. *The Astrophysical Journal*, *493*(1), 460–473. https://doi.org/10.1086/305107

Gombosi, T. I., Chen, Y., Glocer, A., Huang, Z., Jia, X., Liemohn, M. W., et al. (2021). What sustained multi-disciplinary research can achieve: The space weather modeling framework. *Journal of Space Weather and Space Climate*, *11*, 42. https://doi.org/10.1051/swsc/2021020

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

Gopalswamy, N., Yashiro, S., Michalek, G., Stenborg, G., Vourlidas, A., Freeland, S., & Howard, R. (2009). The SOHO/LASCO CME catalog. *Earth, Moon, and Planets*, *104*(1–4), 295–313. https://doi.org/10.1007/s11038-008-9282-7

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. https://doi.org/10.1093/biomet/57.1.97

Hollweg, J. V. (1986). Transition region, corona, and solar wind in coronal holes. *Journal of Geophysical Research*, *91*(A4), 4111–4125. https://doi.org/10.1029/JA091iA04p04111

Huang, Z., Tóth, G., Sachdeva, N., & Van Der Holst, B. (2024). Solar wind driven from gong magnetograms in the last solar cycle. *The Astrophysical Journal*, *965*(1), 1. https://doi.org/10.3847/1538-4357/ad32ca

Huang, Z., Tóth, G., Sachdeva, N., Zhao, L., Van Der Holst, B., Sokolov, I., et al. (2023). Modeling the solar wind during different phases of the last solar cycle. *The Astrophysical Journal Letters*, *946*(2), L47. https://doi.org/10.3847/2041-8213/acc5ef

Hundhausen, A. J. (1993). Sizes and locations of coronal mass ejections: SMM observations from 1980 and 1984-1989. *Journal of Geophysical Research*, *98*(A8), 13177–13200. https://doi.org/10.1029/93JA00157

Iwai, K., Shiota, D., Tokumaru, M., Fujiki, K., Den, M., & Kubo, Y. (2021). Validation of coronal mass ejection arrival-time forecasts by magnetohydrodynamic simulations based on interplanetary scintillation observations. *Earth Planets and Space*, *73*(1), 9. https://doi.org/10.1186/s40623-020-01345-5

Jin, M., Manchester, W. B., van der Holst, B., Sokolov, I., Tóth, G., Mullinix, R. E., et al. (2017b). Data-constrained coronal mass ejections in a global magnetohydrodynamics model. *The Astrophysical Journal*, *834*(2), 173. https://doi.org/10.3847/1538-4357/834/2/173

Jin, M., Manchester, W. B., van der Holst, B., Sokolov, I., Tóth, G., Vourlidas, A., et al. (2017a). Chromosphere to 1 au simulation of the 2011 March 7th event: A comprehensive study of coronal mass ejection propagation. *The Astrophysical Journal*, *834*(2), 172. https://doi.org/10.3847/1538-4357/834/2/172

Jivani, A., Sachdeva, N., Huang, Z., Chen, Y., Van Der Holst, B., Manchester, W., et al. (2023). Global sensitivity analysis and uncertainty quantification for background solar wind using the Alfvén wave solar atmosphere model. *Space Weather*, *21*(1), e2022SW003262. https://doi.org/10.1029/2022SW003262

Joseph, V. R., Gul, E., & Ba, S. (2015). Maximum projection designs for computer experiments. *Biometrika*, *102*(2), 371–380. https://doi.org/10.1093/biomet/asv002

Kapelner, A., & Bleich, J. (2016). bartMachine: Machine learning with Bayesian additive regression trees. *Journal of Statistical Software*, *70*(4). https://doi.org/10.18637/jss.v070.i04

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society - Series B: Statistical Methodology*, *63*(3), 425–464. https://doi.org/10.1111/1467-9868.00294

Kirnosov, V., Chang, L., & Pulkkinen, A. (2015). Automatic CME front edge detection from STEREO white-light coronagraph images. *Space Weather*, *13*(8), 469–483. https://doi.org/10.1002/2015SW001190

Lang, M., Browne, P., Van Leeuwen, P. J., & Owens, M. (2017). Data assimilation in the solar wind: Challenges and first results. *Space Weather*, *15*(11), 1490–1510. https://doi.org/10.1002/2017SW001681

Liu, J., Ye, Y., Shen, C., Wang, Y., & Erdélyi, R. (2018). A new tool for CME arrival time prediction using machine learning algorithms: CAT-PUMA. *The Astrophysical Journal*, *855*(2), 109. https://doi.org/10.3847/1538-4357/aaae69

Lloveras, D. G., Vásquez, A. M., Nuevo, F. A., & Frazin, R. A. (2017). Comparative study of the three-dimensional thermodynamical structure of the inner corona of solar minimum Carrington rotations 1915 and 2081. *Solar Physics*, *292*(10), 153. https://doi.org/10.1007/s11207-017-1179-z

Lloveras, D. G., Vásquez, A. M., Nuevo, F. A., Frazin, R. A., Manchester, W., Sachdeva, N., et al. (2022). Three-dimensional structure of the corona during WHPI campaign rotations CR-2219 and CR-2223. *Journal of Geophysical Research: Space Physics*, *127*(6), e2022JA030406. https://doi.org/10.1029/2022JA030406

Lloveras, D. G., Vásquez, A. M., Nuevo, F. A., Mac Cormack, C., Sachdeva, N., Manchester, W., et al. (2020). Thermodynamic structure of the solar corona: Tomographic reconstructions and MHD modeling. *Solar Physics*, *295*(6), 76. https://doi.org/10.1007/s11207-020-01641-z

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Lugaz, N., Manchester Iv, W. B., & Gombosi, T. I. (2005). The evolution of coronal mass ejection density structures. *The Astrophysical Journal*, *627*(2), 1019–1030. https://doi.org/10.1086/430465

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Volume 1: Statistics* (Vol. 5(1), pp. 281–298). University of California Press. Retrieved from https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992

Maharana, A., Scolini, C., Schmieder, B., & Poedts, S. (2023). Rotation and interaction of the CMEs of September 8 and 10, 2014, tested with EUHFORIA. *Astronomy & Astrophysics*, *675*, A136. https://doi.org/10.1051/0004-6361/202345902

Manchester, W. B., Gombosi, T. I., Roussev, I., Ridley, A., De Zeeuw, D. L., Sokolov, I. V., et al. (2004). Modeling a space weather event from the Sun to the Earth: CME generation and interplanetary propagation. *Journal of Geophysical Research*, *109*(A2), 2003JA010150. https://doi.org/10.1029/2003JA010150

Manchester, W. B., Kilpua, E. K. J., Liu, Y. D., Lugaz, N., Riley, P., Török, T., & Vršnak, B. (2017). The physical processes of CME/ICME evolution. *Space Science Reviews*, *212*(3–4), 1159–1219. https://doi.org/10.1007/s11214-017-0394-0

Manchester, W. B., Van Der Holst, B., & Lavraud, B. (2014). Flux rope evolution in interplanetary coronal mass ejections: The 13 May 2005 event. *Plasma Physics and Controlled Fusion*, *56*(6), 064006. https://doi.org/10.1088/0741-3335/56/6/064006

Manchester, W. B., Vourlidas, A., Tóth, G., Lugaz, N., Roussev, I. I., Sokolov, I. V., et al. (2008). Three-dimensional MHD simulation of the 2003 October 28 coronal mass ejection: Comparison with LASCO coronagraph observations. *The Astrophysical Journal*, *684*(2), 1448–1460. https://doi.org/10.1086/590231

Meier, M. M., Copeland, K., Klöble, K. E. J., Matthiä, D., Plettenberg, M. C., Schennetten, K., et al. (2020). Radiation in the atmosphere—A hazard to aviation safety? *Atmosphere*, *11*(12), 1358. https://doi.org/10.3390/atmos11121358

Meng, X., Van Der Holst, B., Tóth, G., & Gombosi, T. I. (2015). Alfvén wave solar model (AWSoM): Proton temperature anisotropy and solar wind acceleration. *Monthly Notices of the Royal Astronomical Society*, *454*(4), 3697–3709. https://doi.org/10.1093/mnras/stv2249

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Morgan, H., Byrne, J. P., & Habbal, S. R. (2012). Automatically detecting and tracking coronal mass ejections. I. Separation of dynamic and quiescent components in coronagraph images. *The Astrophysical Journal*, *752*(2), 144. https://doi.org/10.1088/0004-637X/752/2/144

Navon, I. M. (2009). Data assimilation for numerical weather prediction: A review. In S. K. Park & L. Xu (Eds.), *Data assimilation for atmospheric, oceanic and hydrologic applications* (pp. 21–65). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-71056-1_2

Olmedo, O., Zhang, J., Wechsler, H., Poland, A., & Borne, K. (2008). Automatic detection and tracking of coronal mass ejections in coronagraph time series. *Solar Physics*, *248*(2), 485–499. https://doi.org/10.1007/s11207-007-9104-5

Owens, M. J., Cargill, P. J., Pagel, C., Siscoe, G. L., & Crooker, N. U. (2005). Characteristic magnetic field and speed properties of interplanetary coronal mass ejections and their sheath regions. *Journal of Geophysical Research*, *110*(A1), 2004JA010814. https://doi.org/10.1029/2004JA010814

Papitashvili, N. E., & King, J. H. (2020). Omni hourly data set [Dataset]. *NASA Space Physics Data Facility*. https://doi.org/10.48322/1SHR-HT18

Pirjola, R., Viljanen, A., Pulkkinen, A., & Amm, O. (2000). Space weather risk in power systems and pipelines. *Physics and Chemistry of the Earth - Part C: Solar, Terrestrial & Planetary Science*, *25*(4), 333–337. https://doi.org/10.1016/S1464-1917(00)00027-1

Powell, K. G., Roe, P. L., Linde, T. J., Gombosi, T. I., & De Zeeuw, D. L. (1999). A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *Journal of Computational Physics*, *154*(2), 284–309. https://doi.org/10.1006/jcph.1999.6299

Riley, P., Mays, M. L., Andries, J., Amerstorfer, T., Biesecker, D., Delouille, V., et al. (2018). Forecasting the arrival time of coronal mass ejections: Analysis of the CCMC CME scoreboard. *Space Weather*, *16*(9), 1245–1260. https://doi.org/10.1029/2018SW001962

Robbrecht, E., Berghmans, D., & Van Der Linden, R. A. M. (2009). Automated LASCO CME catalog for solar cycle 23: Are CMEs scale invariant? *The Astrophysical Journal*, *691*(2), 1222–1234. https://doi.org/10.1088/0004-637X/691/2/1222

Sachdeva, N., Holst, B. V. D., Manchester, W. B., Tóth, G., Chen, Y., Lloveras, D. G., et al. (2019). Validation of the Alfvén wave solar atmosphere model (AWSoM) with observations from the low corona to 1 au. *The Astrophysical Journal*, *887*(1), 83. https://doi.org/10.3847/1538-4357/ab4f5e

Sachdeva, N., Tóth, G., Manchester, W. B., Van Der Holst, B., Huang, Z., Sokolov, I. V., et al. (2021). Simulating solar maximum conditions using the Alfvén wave solar atmosphere model (AWSoM). *The Astrophysical Journal*, *923*(2), 176. https://doi.org/10.3847/1538-4357/ac307c

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Shan, J.-H., Feng, L., Yuan, H.-Q., Zhang, Y., Zhong, X., Gan, W.-Q., et al. (2020). Automatic coronagraph image classification with machine learning methods. *Chinese Astronomy and Astrophysics*, *44*(4), 507–518. https://doi.org/10.1016/j.chinastron.2020.11.006

Shiota, D., & Kataoka, R. (2016). Magnetohydrodynamic simulation of interplanetary propagation of multiple coronal mass ejections with internal magnetic flux rope (SUSANOO-CME). *Space Weather*, *14*(2), 56–75. https://doi.org/10.1002/2015SW001308

Singh, T., Benson, B., Raza, S. A. Z., Kim, T. K., Pogorelov, N. V., Smith, W. P., & Arge, C. N. (2023). Improving the arrival time estimates of coronal mass ejections by using magnetohydrodynamic ensemble modeling, heliospheric imager data, and machine learning. *The Astrophysical Journal*, *948*(2), 78. https://doi.org/10.3847/1538-4357/acc10a

Singh, T., Kim, T. K., Pogorelov, N. V., & Arge, C. N. (2022). Ensemble simulations of the 2012 July 12 coronal mass ejection with the constant-turn flux rope model. *The Astrophysical Journal*, *933*(2), 123. https://doi.org/10.3847/1538-4357/ac73f3

Smith, P. J., Dance, S. L., Baines, M. J., Nichols, N. K., & Scott, T. R. (2009). Variational data assimilation for parameter estimation: Application to a simple morphodynamic model. *Ocean Dynamics*, *59*(5), 697–708. https://doi.org/10.1007/s10236-009-0205-6

Soille, P. (2004). Erosion and dilation. In P. Soille (Ed.), *Morphological image analysis: Principles and applications* (pp. 63–103). Springer. https://doi.org/10.1007/978-3-662-05088-0_3

Sokolov, I. V., Holst, B. V. D., Manchester, W. B., Su Ozturk, D. C., Szente, J., Taktakishvili, A., et al. (2021). Threaded-field-line model for the low solar corona powered by the Alfvén wave turbulence. *The Astrophysical Journal*, *908*(2), 172. https://doi.org/10.3847/1538-4357/abc000

Sokolov, I. V., Van Der Holst, B., Oran, R., Downs, C., Roussev, I. I., Jin, M., et al. (2013). Magnetohydrodynamic waves and coronal heating: Unifying empirical and MHD turbulence models. *The Astrophysical Journal*, *764*(1), 23. https://doi.org/10.1088/0004-637X/764/1/23

Stanzione, D., West, J., Evans, R. T., Minyard, T., Ghattas, O., & Panda, D. K. (2020). Frontera: The evolution of leadership computing at the national science foundation. In *Practice and experience in advanced research computing* (pp. 106–111). ACM. https://doi.org/10.1145/3311790.3396656

Sudar, D., Vršnak, B., & Dumbović, M. (2016). Predicting coronal mass ejections transit times to Earth with neural network. *Monthly Notices of the Royal Astronomical Society*, *456*(2), 1542–1548. https://doi.org/10.1093/mnras/stv2782

Temmer, M., Scolini, C., Richardson, I. G., Heinemann, S. G., Paouris, E., Vourlidas, A., et al. (2023). CME propagation through the heliosphere: Status and future of observations and model development. *Advances in Space Research*. https://doi.org/10.1016/j.asr.2023.07.003

Titov, V. S., & Démoulin, P. (1999). Basic topology of twisted magnetic configurations in solar flares. *Astronomy and Astrophysics*, *351*, 707–720.

Tóth, G., De Zeeuw, D. L., Gombosi, T. I., Manchester, W. B., Ridley, A. J., Sokolov, I. V., & Roussev, I. I. (2007). Sun-to-thermosphere simulation of the 28–30 October 2003 storm with the space weather modeling framework. *Space Weather*, *5*(6), 2006SW000272. https://doi.org/10.1029/2006SW000272

Tóth, G., Van Der Holst, B., Sokolov, I. V., De Zeeuw, D. L., Gombosi, T. I., Fang, F., et al. (2012). Adaptive numerical algorithms in space weather modeling. *Journal of Computational Physics*, *231*(3), 870–903. https://doi.org/10.1016/j.jcp.2011.02.006

Turner, H., Lang, M., Owens, M., Smith, A., Riley, P., Marsh, M., & Gonzi, S. (2023). Solar wind data assimilation in an operational context: Use of near-real-time data and the forecast value of an l5 monitor. *Space Weather*, *21*(5), e2023SW003457. https://doi.org/10.1029/2023SW003457

van der Holst, B., Huang, J., Sachdeva, N., Kasper, J. C., Manchester Iv, W. B., Borovikov, D., et al. (2022). Improving the Alfvén wave solar atmosphere model based on parker solar probe data. *The Astrophysical Journal*, *925*(2), 146. https://doi.org/10.3847/1538-4357/ac3d34

Van Der Holst, B., Manchester, W. B., Frazin, R. A., Vásquez, A. M., Tóth, G., & Gombosi, T. I. (2010). A data-driven, two-temperature solar wind model with Alfvén waves. *The Astrophysical Journal*, *725*(1), 1373–1383. https://doi.org/10.1088/0004-637X/725/1/1373

van der Holst, B., Sokolov, I. V., Meng, X., Jin, M., Manchester, W. B., IV, Tóth, G., & Gombosi, T. I. (2014). Alfvén wave solar model (AWSoM): Coronal heating. *The Astrophysical Journal*, *782*(2), 81. https://doi.org/10.1088/0004-637X/782/2/81

Vršnak, B., Ruždjak, D., Sudar, D., & Gopalswamy, N. (2004). Kinematics of coronal mass ejections between 2 and 30 solar radii: What can be learned about forces governing the eruption? *Astronomy & Astrophysics*, *423*(2), 717–728. https://doi.org/10.1051/0004-6361:20047169

Vršnak, B., & Žic, T. (2007). Transit times of interplanetary coronal mass ejections and the solar wind speed. *Astronomy & Astrophysics*, *472*(3), 937–943. https://doi.org/10.1051/0004-6361:20077499

Wang, P., Zhang, Y., Feng, L., Yuan, H., Gan, Y., Li, S., et al. (2019a). A new automatic tool for CME detection and tracking with machine-learning techniques. *The Astrophysical Journal - Supplement Series*, *244*(1), 9. https://doi.org/10.3847/1538-4365/ab340c

Wang, Y., Liu, J., Jiang, Y., & Erdélyi, R. (2019b). CME arrival time prediction using convolutional neural network. *The Astrophysical Journal*, *881*(1), 15. https://doi.org/10.3847/1538-4357/ab2b3e

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on image segmentation techniques. *Procedia Computer Science*, *65*, 797–806. https://doi.org/10.1016/j.procs.2015.09.027

Zhao, X., & Dryer, M. (2014). Current status of CME/shock arrival time prediction. *Space Weather*, *12*(7), 448–469. https://doi.org/10.1002/2014SW001060