Ensemble CME modelling constrained by heliospheric imager observations

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

www.reading.ac.uk/centaur

CentAUR

AGU Advances

Key Points:
- Ensemble modelling of CME arrival time is a useful space weather forecasting technique
- Heliospheric imager data can improve the skill and reduce the uncertainty of ensemble hindcasts
- This supports the idea of including a heliospheric imager on the next space weather monitor

Supporting Information:
- Original Version of Manuscript
- Peer Review History
- Authors' Response to Peer Review Comments
- First Revision of Manuscript [Accepted]

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

Ensemble CME Modeling Constrained by Heliospheric Imager Observations

L. Barnard[1] , M. J. Owens[1] , C. J. Scott[1], and C. A. de Koning[2,3]

[1]Department of Meteorology, University of Reading, Reading, UK, [2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [3]Space Weather Prediction Center, NOAA, Boulder, CO, USA

Abstract Predicting the arrival of coronal mass ejections (CMEs) is one key objective of space weather forecasting. In operational space weather forecasting, solar wind numerical models are used for this task and ensemble techniques are being increasingly explored as a means to improve these forecasts. Currently, these forecasts are not constrained by the available in situ and remote sensing observations, such as those from the heliospheric imagers (HIs) on the National Aeronautics and Space Administration's (NASA's) STEREO spacecraft, which record white-light images of solar wind and CMEs. We report case studies of four CMEs and show how HI observations can be used to improve the skill and reduce the uncertainty of ensemble hindcasts of these events. Using a computationally efficient solar wind model, we produce 200-member ensemble hindcasts, perturbing the modeled CME parameters within uniform distributions about the best estimates. By comparing the trajectory of the modeled CME flanks with HI observations, we compute a weight for each ensemble member. Weighting the ensemble distribution of CME arrival times improves the skill and reduces the hindcast uncertainty of each event. For these four events, the weighted ensembles show a mean reduction in arrival time error of $20.1 \pm 4.1\%$, and a mean reduction in arrival time uncertainty of $15.0 \pm 7.2\%$, relative to the unweighted ensembles. This technique could be applied in operational space weather forecasting, if real-time HI observations were available. Therefore, as NASA and the European Space Agency are currently planning the next space weather monitoring missions, our proof-of-concept study provides some evidence of the potential value of including HIs on these missions.

Plain Language Summary Coronal mass ejections (CMEs) are large eruptions of magnetized plasma from the Sun's atmosphere that flow out through space. CMEs that reach Earth are the main cause of severe space weather and can disrupt technology we rely on, such as satellites, communications networks, and power grids. Consequently, forecasting the arrival of CMEs at Earth is an important service performed by various national weather agencies. Improving these forecasts is an important area of space weather research, particularly measuring and improving the uncertainty of the CME arrival time predictions. We show how pictures of CMEs taken by the heliospheric imagers on NASA's STEREO spacecraft can be used to reduce the uncertainty and improve the accuracy of CME arrival time predictions. NASA and ESA are currently planning the next spacecraft missions that will observe the Sun and space for space weather forecasting. Our proof-of-concept study provides some evidence that it would be useful to include a heliospheric imager on these future missions.

1. Introduction

Coronal mass ejections (CMEs) are vast eruptions of magnetized plasma from the Sun's corona. At Earth, CMEs are the main driver of space weather, which energizes Earth's space environment and disrupts critical services provided by spacecraft, power grids, and aircraft (Cannon et al., 2013; Hapgood, 2011). Consequently, understanding the propagation of CMEs through the solar wind, and being able to estimate their expected arrival at Earth, are key research questions and objectives for space weather forecasting centers. Despite recent progress, the evolution of CMEs through the solar wind and heliosphere is still not well understood, due to historically sparse heliospheric observations and open questions regarding CME structure (Luhmann et al., 2020). Owens, Lockwood, and Barnard (2020) showed that CME arrival time forecasts are valuable for a range of hypothetical operational settings. They also demonstrated that additional value can be added to a CME forecast by including information about a CME's speed and magnetic field strength.

This highlights that although CME arrival time estimates are a necessary element of CME forecasting, they are not, and should not be, the sole focus.

Since the launch of the National Aeronautics and Space Administration's (NASA's) STEREO mission in 2007 (Kaiser et al., 2008), we have been able to routinely observe the propagation of CMEs from the solar corona, out through the inner heliosphere to Earth orbit, using the white-light coronagraph and heliospheric imager (HI) instruments in the Sun-Earth-Connection-Heliospheric-Investigation instrument package (Howard et al., 2008). Before this, CMEs could only be routinely observed close to the Sun (typically within 30 solar radii) in the white-light Large Angle and Spectrometric Coronagraph (LASCO) on NASA's SOHO mission (Brueckner et al., 1995), and measured with in situ solar wind plasma monitors, typically in near-Earth space.

Observations reveal that CMEs undergo many different dynamical processes during their propagation through the corona and heliosphere. Coronagraphs show that CMEs are accelerated and can be deflected and deformed by structures in the solar corona (Jones et al., 2017). This is expected to be due largely to gradients in magnetic pressure between features like coronal holes, active regions, and the heliospheric current sheet (Kay et al., 2015). Furthermore, interactions with the solar wind are important in determining a CME's evolution. Fast and slow solar wind streams can accelerate or decelerate a CME's speed (Tucker-Hood et al., 2015), and can also lead to distortions of a CME's structure if they flow through the boundary between fast and slow solar wind streams (Owens et al., 2017; Savani et al., 2010). Consequently, forecasting the arrival of CMEs at Earth requires a good knowledge of the ambient solar wind conditions.

The current state-of-the-art for forecasting CME arrivals at Earth uses 3-D magnetohydrodynamic (3-D MHD) models of the heliosphere (Merkin et al., 2016; Odstrcil, 2003; Riley et al., 2001; Tóth et al., 2005). However, the Riley et al. (2018) review of CME forecasting techniques (many based on 3-D MHD models) concluded that the mean absolute arrival time error is ± 10 hr (20% to 50% of the expected Sun-Earth transit time) and, more interestingly, that the accuracy of arrival time forecasts had not measurably improved since 2012. However, this significant uncertainty and stalled progress is maybe not surprising.

These models, such as ENLIL (Odstrcil, 2003) (used by the UK Met Office and NOAA's Space Weather Prediction Center (SWPC)), require an inner boundary condition only for determining the state of the solar wind. However, these inner boundary conditions cannot be observed directly, and must be estimated from observations of the Sun's photosphere and extrapolated using models of the Sun's corona; heliospheric models are currently driven by the output of coronal models that are empirically tuned to match in situ observations near Earth. Furthermore, currently for forecasting, CMEs are typically introduced to these 3-D MHD models as time-dependent perturbations of the solar wind speed and plasma parameters at the model's inner boundary (so-called "cone model" CMEs), with coronagraphs used to estimate the CME's initial state (Odstrcil, 2003). Both of these procedures have significant uncertainties that are difficult to estimate.

Therefore, initial condition ensembles are being increasingly investigated as a way to estimate the uncertainty in 3D MHD simulations of CME propagation (Mays et al., 2015; Murray, 2018; Pizzo et al., 2015). Furthermore, ensemble modeling allows a probabilistic interpretation of the forecast, which can be more valuable than a single deterministic forecast (Owens & Riley, 2017a). However, 3D MHD simulations are sufficiently computationally expensive that it is challenging to appropriately sample the relevant parameter space of CME and solar wind properties, particularly for a large number of events or for multiple interacting CMEs.

Data assimilation (DA), widely used in meteorology, is another promising tool for improving space weather forecasts from solar wind numerical models (Lang & Owens, 2019). DA provides a framework to optimally combine models with observations to provide a best estimate of the state of a dynamical system. However, despite the current and past availability of HI observations throughout the domain of the heliospheric models, difficulties in interpreting HI data in terms of the model state have meant that it has not been possible to use HI and DA techniques to provide a better representation of solar wind CME interactions (Lang et al., 2017).

Alternatively, in recent publications, there is growing interest in using HI observations to weight ensemble members as a means of providing improved estimates of solar wind-CME interactions and CME arrival times at Earth (Harrison et al., 2017; Murray, 2018; Wharton et al., 2019). Here we report case studies of four CMEs

Table 1
Cone CME Properties

Event	CME-1	CME-2	CME-3	CME-4
Time at $21.5r_\odot$	2012-08-31T22:46	2012-09-28T03:49	2012-10-05T08:47	2012-11-20T17:40
Longitude (deg)	−30	20	9	22
Latitude (deg)	0	4	−24	20
Width (deg)	66	110	84	94
Speed (km s^{-1})	1,010	872	698	664
Predicted arrival	2012-09-03T17:00	2012-09-30T15:00	2012-10-08T15:00	2012-11-23T17:00
Observed arrival	2012-09-03T11:23	2012-09-30T22:13	2012-10-08T04:31	2012-11-23T21:12

that highlight the potential utility of such an approach, enabled through computationally efficient solar wind modeling. Using the HUXt solar wind model (Owens, Lang, et al., 2020; Riley & Lionello, 2011), we consider the evolution of the CMEs launched on 31 August, 28 September, 10 October, and 20 November in 2012. With HUXt, we produce an ensemble hindcast of these events and demonstrate that HI observations of these CMEs can be used to appropriately weight the ensemble members. This leads to improved CME arrival time estimates and reduced ensemble spread, and thus improved hindcast performance. The method presented here could be readily integrated into an existing forecasting system, assuming real-time availability of HI data, and so might be a valuable tool for operational space weather forecasting centers.

2. Methods and Data

2.1. CME Events

Here we study four CMEs that were previously analyzed by Barnard et al. (2017). For each event, NOAA's SWPC produced a forecast of the CME using the standard WSA-ENLIL Cone modeling system. Using the available coronagraph data, SWPC calculated estimates of the required cone CME parameters, specifically the CME source longitude and latitude, width, speed, and the time at a height of $21.5r_\odot$; these are given in Table 1. Verification of these forecasts using in situ solar wind plasma observations from the ACE spacecraft at the L1 point subsequently established their arrival times. Over this period, the STEREO-A and STEREO-B spacecraft were at HEEQ longitudes of ~125° and 240°, and so were separated from Earth by ~125° and 120°, respectively. In the following case studies, we use the same cone CME parameters as used by SWPC, except for adjusting the initiation time, to account for the fact the inner boundary of HUXt is at $30r_\odot$, rather than $21.5r_\odot$. For the rest of the article, these events will be referred to as CME-1, CME-2, CME-3, and CME-4.

2.2. Solar Stormwatch CME Tracking

As part of the study of Barnard et al. (2017), these CMEs were tracked by the Solar Stormwatch (SSW) project. Many citizen scientists tracked the events through the HI1 field of view (FOV) on both the STEREO-A and STEREO-B spacecraft (referred to as HI1A and HI1B, respectively). The HI1 instruments are white-light imagers that observe sunlight Thomson scattered from solar wind electrons (Howard et al., 2008). Figure 1 shows examples of running differenced images for HI1A and HI1B, which highlight transient features such as CMEs. The HI1 FOV spans 20° and is centered at 14° in the ecliptic plane. The nominal cadence of HI1 science images is 40 min with a binned pixel size of 70 arcsec. The location of features in the HI FOV will be discussed in terms of Helioprojective-Radial-Coordinates: position angle (PA), the anticlockwise angle from solar north, and elongation (ϵ), the angular distance from Sun center.

Barnard et al. (2017) developed a procedure to produce a statistical consensus estimate of the CME front location for each HI1A and HI1B image, which allows the time-elongation profile of the CME front to be computed at any position angle spanned by the CME. Examples of these consensus profiles are presented in Figure 1. They also demonstrated that the resulting time-elongation profiles had better uncertainty estimates and more stable feature tracking than those obtained through the more commonly used J-map technique (Davies et al., 2009). Here we use the SSW profiling of these events to extract the time-elongation profile of the CME flanks along the latitudinal plane corresponding to the mean HEEQ latitude of Earth during each event. Here we define the CME flank as the maximum elongation of a CME front viewed from a particular vantage point, along a specified position angle direction.
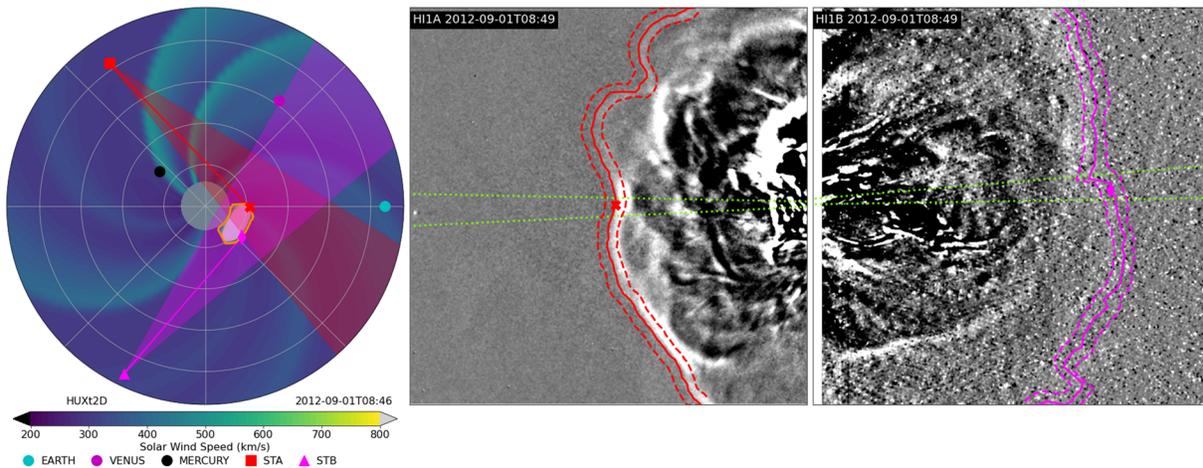
Figure 1. (left) The HUXt model solution along Earth's latitudinal plane for the deterministic (i.e., "best estimate") run of CME-1 at 2012-09-01T08:46. The orange line marks the boundary of the CME region. The shaded red and pink regions mark the HI1A and HI1B fields-of-view. The red-cross and pink-diamond mark the flanks of the modeled CME from the HI1A and HI1B perspective. (middle) The meridional plane of the same CME differenced imaged by HI1A, at the closest time to the model snapshot. The solid-red line marks the SSW consensus profile of the CME front, while the dashed lines mark the uncertainty in this profile. The dashed-green lines mark a 4° position angle band around the solar equatorial plane. The red cross marks the elongation of the CME front in the equatorial plane. (right) HI1B differenced image of this CME, marked in the same way as the HI1A image.

2.3. HUXt

HUXt (Owens, Lang, et al., 2020) is a numerical model of the solar wind that uses a reduced physics approach, treating the solar wind as a 1-D incompressible hydrodynamic flow. This allows very efficient computational solutions, being ~1,000 times faster than comparable 3-D MHD solar wind models. Despite this reduced physics approach, HUXt has been shown to closely emulate full 3-D MHD models; a 40-year validation test of ambient solar wind from HelioMas (Riley et al., 2001) was reproduced to within 7% of the HelioMAS solar wind speeds throughout the entire model domain (Owens, Lang, et al., 2020; Riley & Lionello, 2011). Therefore, HUXt can serve as an effective surrogate in situations where full 3-D MHD simulations are too computationally expensive. In particular, HUXt ensembles of model solutions that more completely explore the full range of uncertainties in the initial conditions. HUXt only requires the solar wind speed at the model inner boundary to be specified and so can work with the output of any model that can provide this. Here we use output from the MAS coronal model (Riley et al., 2001), but it can also operate with output from, for example, the Wang-Sheely-Arge model (Arge & Pizzo, 2000). In this work, HUXt is run in the latitudinal plane corresponding to Earth's HEEQ latitude at the initialization time of each cone CME.
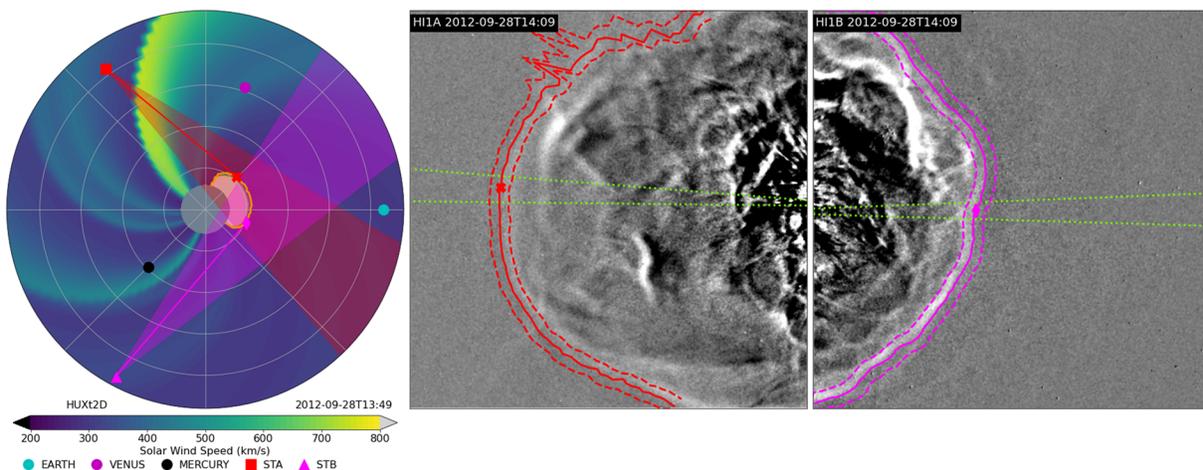
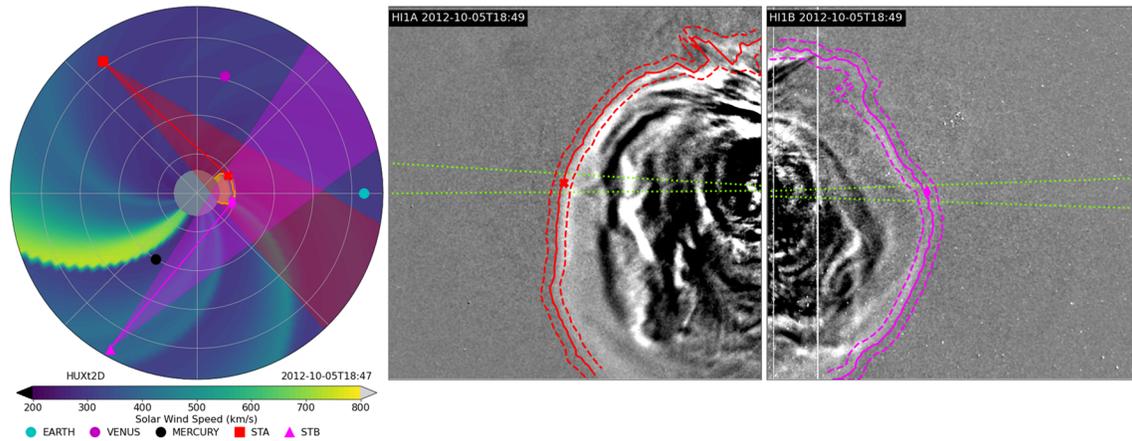**Figure 2.** The same structure as Figure 1 but for CME-2.

Figure 3. The same structure as Figure 1 but for CME-3.

Within HUXt, CMEs are parameterized as cone CME perturbations to the solar wind speed. Spatially, the cone CME consists of two hemispheres connected by a cylinder; at one extreme a CME is spherical, while more generally it is sausage shaped. The axis of the cone CME is directed radially and located by the CME's source longitude and latitude. The CME's width is used to parameterize the angular extent of the hemispheres, while the "thickness" sets the length of the cylindrical portion. This structure is advected through the model inner boundary at the CME's speed, and anywhere on the boundary within the cone CME domain is assigned the CME speed. This perturbation then propagates hydrodynamically through the model solution. CMEs are traced through the HUXt solution by computing the difference between the ambient and CME solutions, identifying and tracking areas where the speed differences are >20 km s^{-1}. In this work we must also calculate the flank of the CME from the HI1A and HI1B perspectives. This is achieved by computing the elongation of each point on the CME boundary and finding the point with maximum elongation, for both the HI1A and HI1B field of view.

For each event, we first use HUXt to produce a deterministic hindcast, using the same cone model parameters as was used in the SWPC WSA-ENLIL forecast. The HUXt inner boundary solar wind speeds at $30r_\odot$ are set by the MAS coronal model, driven by Carrington maps of the photospheric magnetic field provided by the HMI instrument on NASA's Solar Dynamics Observatory. The MAS solutions are available at this site (https://www.predsci.com/mhdweb/home.php). Thus, we must estimate the time at which the cone CME enters the model domain. We linearly extrapolate the WSA-ENLIL CME initiation time at $21.5r_\odot$, assuming the the CME speed remains constant. An example of the output of HUXt for these hindcasts are shown in Figures 1 to 4. HUXt returns the boundary of the CME region, and from this we compute the maximum elongation of the CME flank in the HI1A and HI1B fields of view.
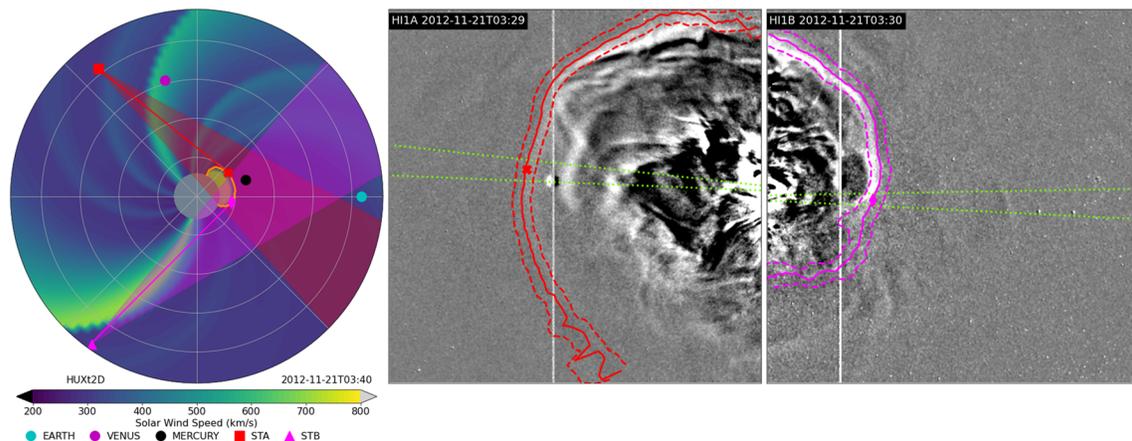

Figure 4. The same structure as Figure 1 but for CME-4.

Further to this, we also produce 200-member ensemble hindcasts of each event, where the CME cone model parameters are randomly perturbed. Ideally, we would sample from the empirically determined uncertainty distribution of each parameter. However, these distributions are not accurately known. Therefore, for this proof-of-concept study, we simply assume that the parameters uncertainty distributions are uniform. Each parameter is perturbed by sampling from the random uniform distribution, with the following limits: speed is ± 200 km/s; longitude is $\pm 20°$; latitude is $\pm 20°$; width is $\pm 20°$; radial thickness is $\pm 2\ r_\odot$; and initial height is $\pm 3\ r_\odot$. We consider these uncertainty limits as at the larger end of plausibility, but they are comparable to those explored in Pizzo et al. (2015). Additionally, the time that the cone CME enters the model domain is also updated to account for the perturbed CME speed and initial height.

3. Results

From both the deterministic and ensemble HUXt runs, we compute the expected arrival of a CME at Earth, and the time-elongation profile of a CMEs flank that would be observed by HI1A and HI1B, as shown schematically in Figures 1–4. Figure 5 shows the observed time elongation profiles for HI1A and HI1B, as well as the modeled time elongation profiles for both the deterministic and ensemble HUXt runs, for each of CME-1 to CME-4. It should be noted that the comparison between the observed and modeled time elongation profiles is limited by idealized nature of the cone CME paramterization and that we do not forward model the HI observations from the HUXt solutions. However, for the purposes of this proof-of-concept study, we consider the comparison of these observed and modeled time elongation profiles sufficient.

For these four events, there is some agreement between the observed and modeled time elongation profiles. We note that the agreement is typically better for the HI1B SSW observations, rather than HI1A, especially for CME-2, CME-3, and CME-4; these SSW profiles are within the ensemble spread and track closely to the deterministic run and other ensemble members.

This is less the case for the HI1A SSW profiles, which are typically closer to the edge of the ensemble distribution and in some cases outside it, particularly for CME-2 and CME-4. For CME-2 the HI1A SSW profile follows an obviously different trajectory from the deterministic HUXt run and is outside of the spread of all ensemble members. The cause of the disagreement in HI1A is not immediately clear. Both the SSW tracking of the CME in HI1A images, and the representation of the CMEs evolution within HUXt could be factors. It is also possible that the best estimates of the CME parameters were far from the truth, or that the ensemble spread was not appropriate for this event. We investigated this by repeating the ensemble analysis with twice the range in the cone CME parameter. Even with double the range, the SSW HI1A profile for CME-2 remained outside the ensemble spread, while for CME-4 it was at the very edge of the ensemble. However, the CME front is clearly defined in both the HI1A and HI1B differenced images, and the uncertainty in the SSW profiling of this event is modest. Therefore, on balance, we expect that the representation of the CME with the cone model in HUXt is the main factor. In fact, it is plausible that CME front could have been deformed at heliospheric distances below the inner boundary of HUXt at $30r_\odot$ (Kay et al., 2015), which might explain these differences. We are experimenting with moving the inner boundary of HUXt to lower altitudes, and the impacts this has will form the basis of a future study.

Furthermore, we note that in all instances the SSW profiles lead HUXt, both in the deterministic form and the majority of the ensemble members. We interpret this as evidence there might be a bias in HUXt, such that CMEs slightly lag the observations. However, it is not obvious what the source of such a bias would be, and could involve many factors in the model's implementation. This too requires further investigation in a future study. But we are nevertheless able to compare weighted and unweighted HUXt ensembles, which are subject to the same biases.

To be able to use the SSW profiles to constrain the ensemble hindcast, it is first necessary to show that it is possible to weight each ensemble member according to its agreement with the SSW time-elongation profile. Here we test using a weight based on the root-mean-square difference between the SSW and ensemble time-elongation profiles. To do this, we linearly interpolate the HUXt time-elongation profile onto the SSW profile, and then compute the root-mean-square difference between the HUXt and SSW profiles. The inverse of the root-mean-square difference is then used to weight each ensemble member. We also compute a HI1A and HI1B combined weight using the mean of the HI1A and HI1B weights. Figure 6 shows scatter plots of the absolute arrival time error for each ensemble member as a function of the weight computed from the HI1A, HI1B, and combined HI1A + HI1B data, respectively.
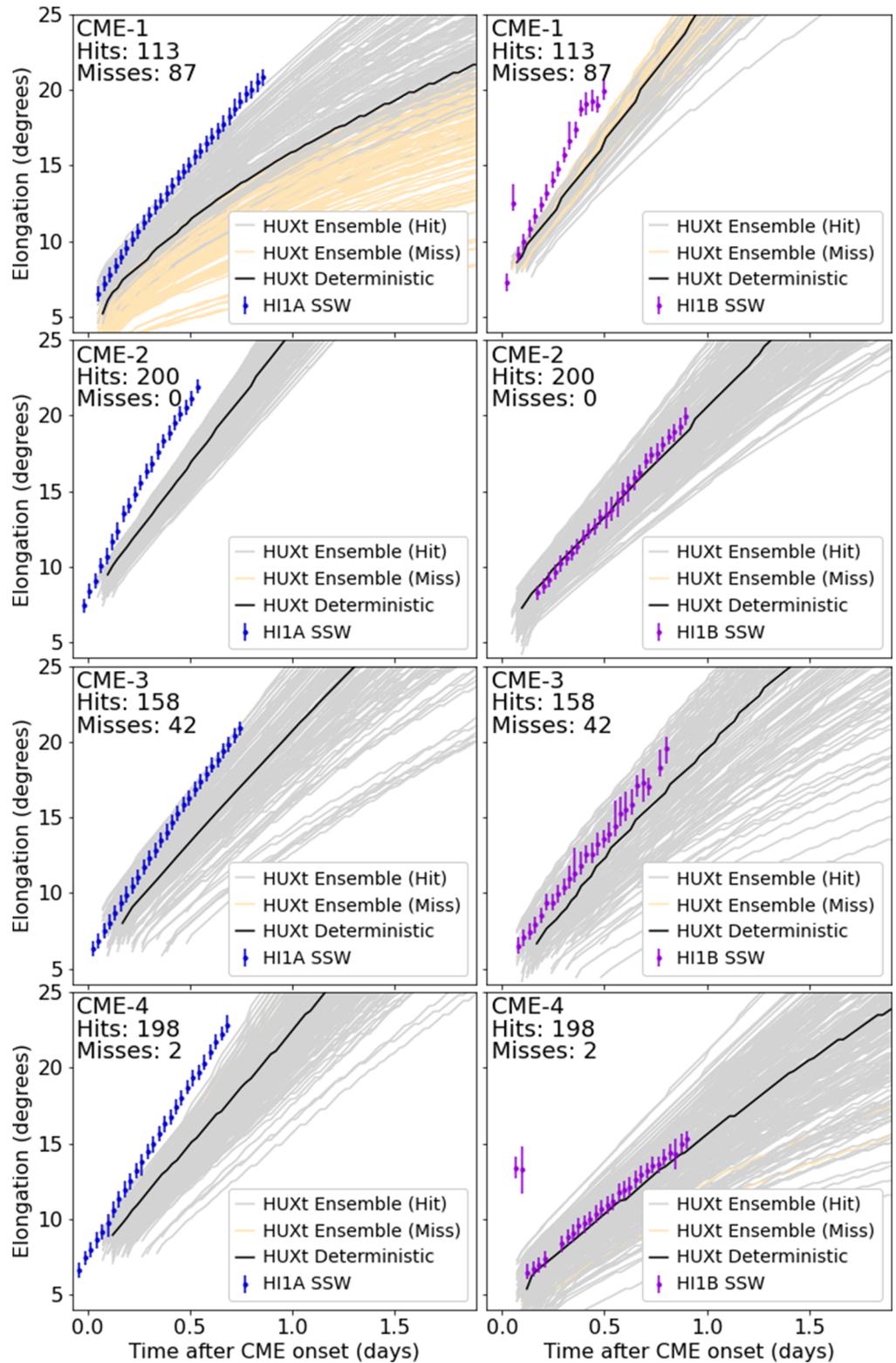
Figure 5. Time elongation profiles of the modeled and observed CME flanks, from the HI1A and HI1B perspectives, for each CME (top to bottom). (left column) The blue points show the HI1A SSW consensus profile along Earth's latitudinal plane. The black line shows the modeled flank from the HI1A perspective for the deterministic HUXt run. The cluster of gray and orange lines show the modeled flank for the 200 ensemble members; gray lines mark ensemble members that hit Earth; orange lines mark ensemble members that miss Earth. (right column) This panel shows the observed and modeled data from the HI1B perspective.
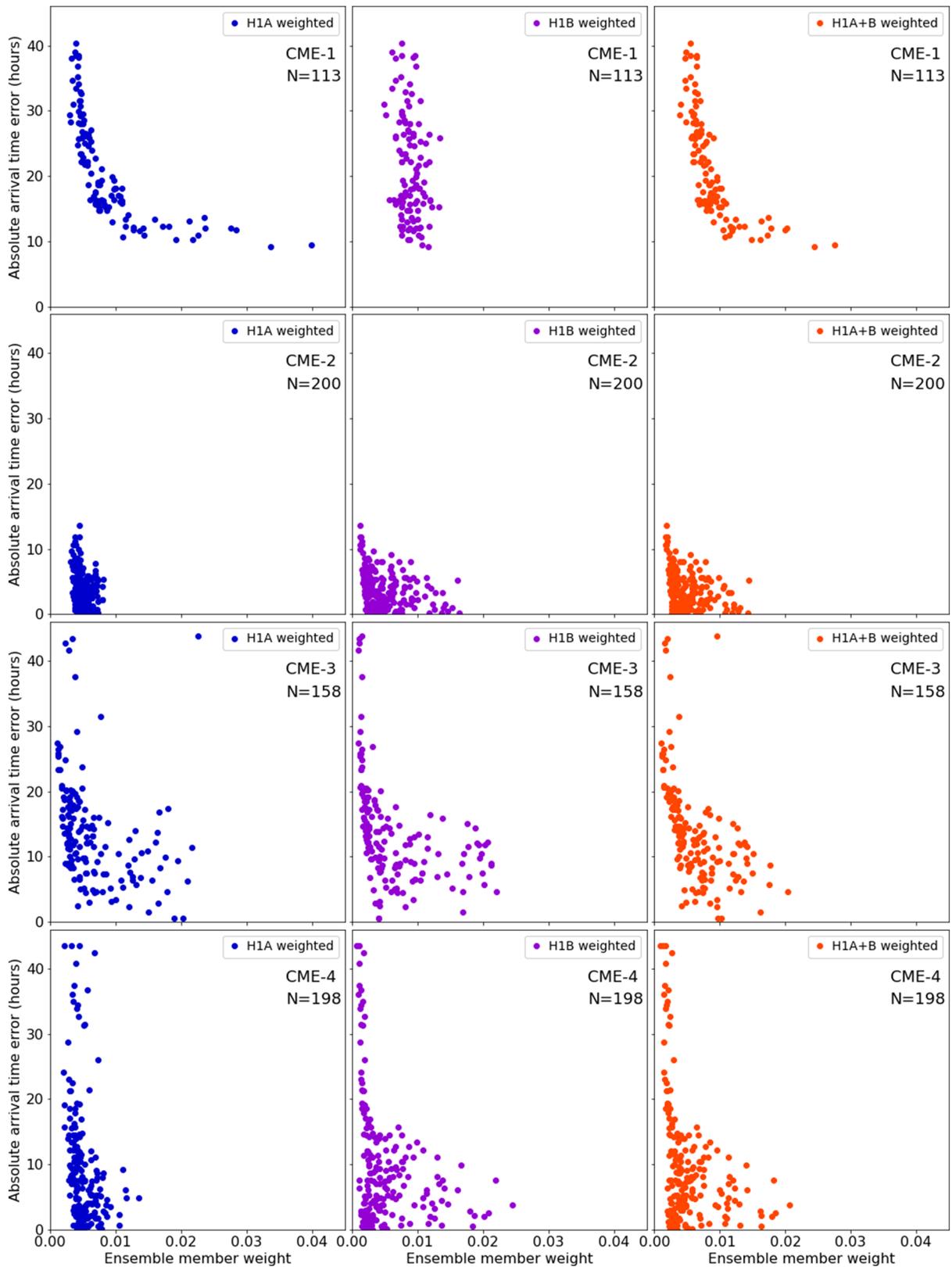
Figure 6. Panels show scatter plots of the HUXt ensemble member weights versus the absolute arrival time error, for each of CME-1 to CME-4 (top to bottom). The ensemble member weight is calculated from the comparison of the observed and modeled CME flanks. Only ensemble members that hit Earth (and so have an arrival time error) are displayed. Left, middle, and right columns show the weightings for the HI1A, HI1B, and combined HI1A and HI1B observations, respectively.
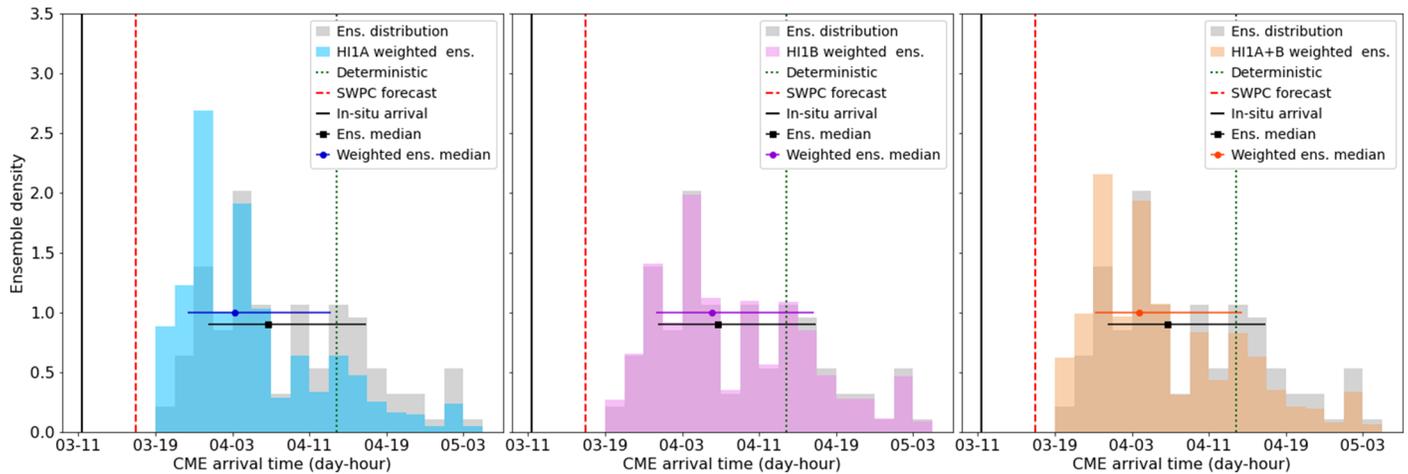
Figure 7. Histograms of the distribution of modeled CME arrival time for CME-1. In each panel, the gray histogram shows the distribution of transit times for the basic ensemble of HUXt runs, and the black square marks the ensemble median and spread of the distribution. The dashed red line shows the predicted CME transit time from SWPC's WSA-ENLIL forecast. The solid black line marks the observed CME transit time, and the dotted green line marks the transit time of the deterministic HUXt run. The blue, pink, and orange histograms show the distribution of transit times for the ensemble members weighted by their agreement with the HI1A, HI1B, and HI1A + HI1B SSW observations of the CME. The blue, pink, and orange circles mark the weighted medians and spread of the transit time distributions.

In most cases, it is clear that members with larger weights typically have smaller arrival time errors. However, it is also apparent that the relationship between ensemble member weight and arrival time error is not linear and there is significant scatter. A detailed understanding of this requires further investigation, but we suggest that this is possibly due to the fact the observed time-elongation profiles of the CME flanks are projections onto a plane of the true CME evolution. Consequently, there is no unique mapping of a CMEs parameters to the observed time elongation profiles, and many profiles are degenerate for different combinations of CME parameters, both modeled and observed. Nonetheless, Figure 6 does show that the computed weights do contain useful information about the expected arrival of the studied CMEs. On this basis, it is reasonable to test if these weights can be used to improve the expected CME arrival times from the ensemble hindcasts.

In the following, we analyze the CME arrival time distributions for the ensemble hindcasts and compare them with weighted arrival time distributions, where each ensemble member is weighted according to their agreement with the SSW time elongation profile, as presented in Figure 6. We note that weighting ensembles with one source of relevant information is expected to lead to a reduction in ensemble spread. Therefore, although we expect that the HUXt ensemble spread will decrease after weighting with the SSW profiles, it is necessary to quantify by how much, and whether the reduced spread remains consistent with the observations.

3.1. CME-1 Arrival Time Distribution

Figure 1 shows that for the deterministic run this CME hits Earth with a glancing blow, with Earth intersecting the western flank. Consequently, when perturbing the cone CME initial conditions, many of the ensemble members subsequently miss Earth. Of the 200 ensemble members, there are 113 hits, and 87 misses. Figure 7 presents histograms of the distribution of the CME arrival time for the HUXt ensemble members that hit Earth. In each panel, the gray histogram shows the distribution of the full ensemble of arrival times. The solid black vertical line shows the observed arrival time, the dashed red line marks the SWPC forecast arrival time, and the dotted green line marks the arrival time of the deterministic HUXt run. The black square shows the ensemble median, with the error bar marking the range between the 0.16 and 0.84 quantiles of the arrival time distribution (an approximate $1 - \sigma$ error). The full range of the ensemble arrival time distribution spans 31.2 hr, which is significantly larger than the average arrival time uncertainty over many different CME forecasts of ±10 hr (Riley et al., 2018). This highlights the need for event-specific uncertainty estimates.

Table 2
CME Arrival Time Forecast Statistics

Event	Type	Arrival time	Arrival uncertainty (hr)	Arrival error (hr)
	Deterministic	2012-09-04T13:51	—	26.5
	Ensemble	2012-09-04T06:41	16.4 (−6.2, +10.2)	19.3
CME-1	HI1A weighted	2012-09-04T03:12	14.9 (−4.8, +10.0)	15.8
	HI1B weighted	2012-09-04T06:06	16.3 (−5.8, +10.5)	18.7
	HI1A/B weighted	2012-09-04T03:47	15.2 (−4.6, +10.6)	16.4
	Deterministic	2012-09-30T23:31	—	1.3
	Ensemble	2012-10-01T00:24	8.6 (−4.1, +4.5)	2.2
CME-2	HI1A weighted	2012-09-30T23:49	8.7 (−4.1, +4.6)	1.6
	HI1B weighted	2012-09-30T23:49	7.5 (−3.5, +4.1)	1.6
	HI1A/B weighted	2012-09-30T23:49	7.8 (−3.8, +4.1)	1.6
	Deterministic	2012-10-08T16:32	—	12.0
	Ensemble	2012-10-08T16:47	13.5 (−5.7, +7.8)	12.3
CME-3	HI1A weighted	2012-10-08T14:54	12.4 (−5.6, +6.8)	10.4
	HI1B weighted	2012-10-08T14:19	9.0 (−4.1, +4.9)	9.8
	HI1A/B weighted	2012-10-08T14:28	10.7 (−4.8, +5.9)	10.0
	Deterministic	2012-11-24T01:16	—	4.1
	Ensemble	2012-11-24T03:45	16.4 (−7.3, +9.1)	6.6
CME-4	HI1A weighted	2012-11-24T01:42	16.2 (−6.5, +9.7)	4.5
	HI1B weighted	2012-11-24T02:43	11.9 (−6.1, +5.8)	5.5
	HI1A/B weighted	2012-11-24T02:26	12.6 (−5.8, +6.8)	5.2

The weighted ensemble distributions are also shown in Figure 7. The blue, pink, and orange histograms show the arrival time distributions weighted by the HI1A, HI1B, and combined HI1A and HI1B observations, respectively. Similarly, the blue, pink, and orange circles mark the medians and approximate $1 - \sigma$ spread of the weighted distributions. Details of the modeled arrival times are provided in Table 2.

For this event, all the ensemble members arrive significantly later than observed CME at L1. Consequently, the observed arrival is outside the uncertainty region of the unweighted and weighted ensemble medians, and the corresponding errors of the ensemble medians span 15.8 to 19.3 hr. The medians of the weighted distributions are closer to the observed arrival time than the unweighted ensemble median, although these differences are smaller than the ensemble spread in each case. Additionally, the spread of the arrival time distributions is slightly less for each of the weighted ensembles, indicating smaller uncertainty (~1 hr) in the expected arrival time. Focusing on the HI1A/HI1B weighted ensemble, there is a 15% reduction in the arrival error, and a 7% reduction in the arrival uncertainty, relative the the unweighted ensemble.

3.2. CME-2 Arrival Time Distribution

For CME-2, all 200 of the ensemble members hit Earth, and the distribution of arrival times is shown in Figure 8. The full range of the ensemble arrival time distribution spans 19.5 hr; this is a much smaller ensemble spread than any of the other studied CMEs. In this instance, the observed arrival time is within the uncertainty band for each of the weighted and unweighted ensembles, and the arrival errors are all small, between 1.6 and 2.2 hr. Again, the medians of the weighted distribution are closer to the observed arrival time than the unweighted ensemble median, but these differences are smaller than the ensemble spread in each case. The spread of the arrival time distributions is less for each of the weighted ensembles, indicating slightly smaller uncertainty (~1 hr) in the expected arrival time. For the HI1A/HI1B weighted ensemble, there is a 27% reduction in the arrival error, and a 9% reduction in the arrival uncertainty, relative the the unweighted ensemble.

3.3. CME-3 Arrival Time Distribution

For CME-3, 158 of the ensemble members hit Earth, and the distribution of arrival times is shown in Figure 9. The full range of the ensemble arrival time distribution spans 50 hr. In this instance, the observed arrival time
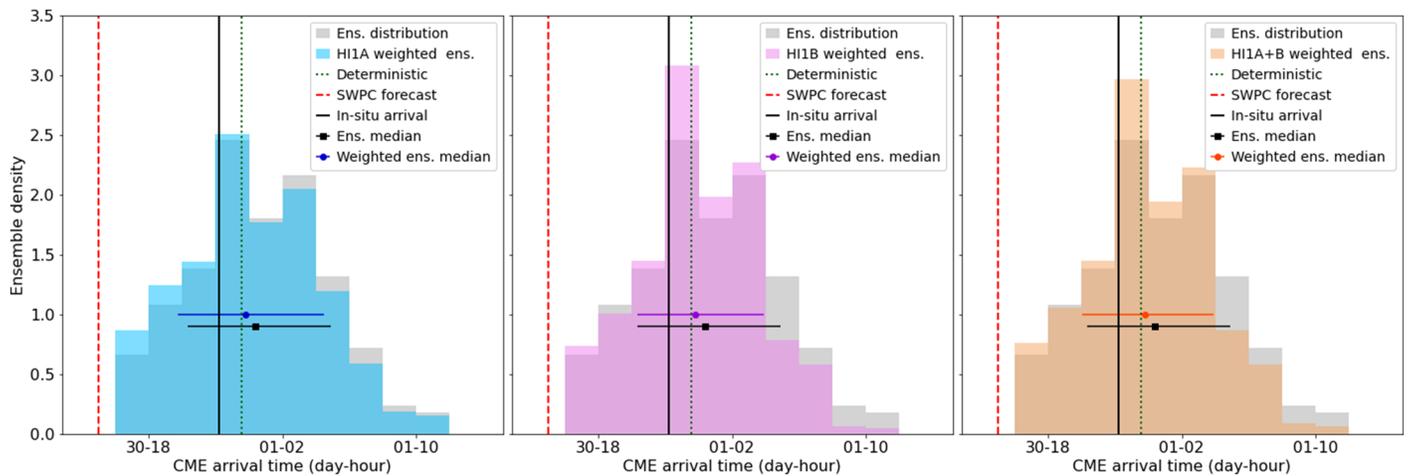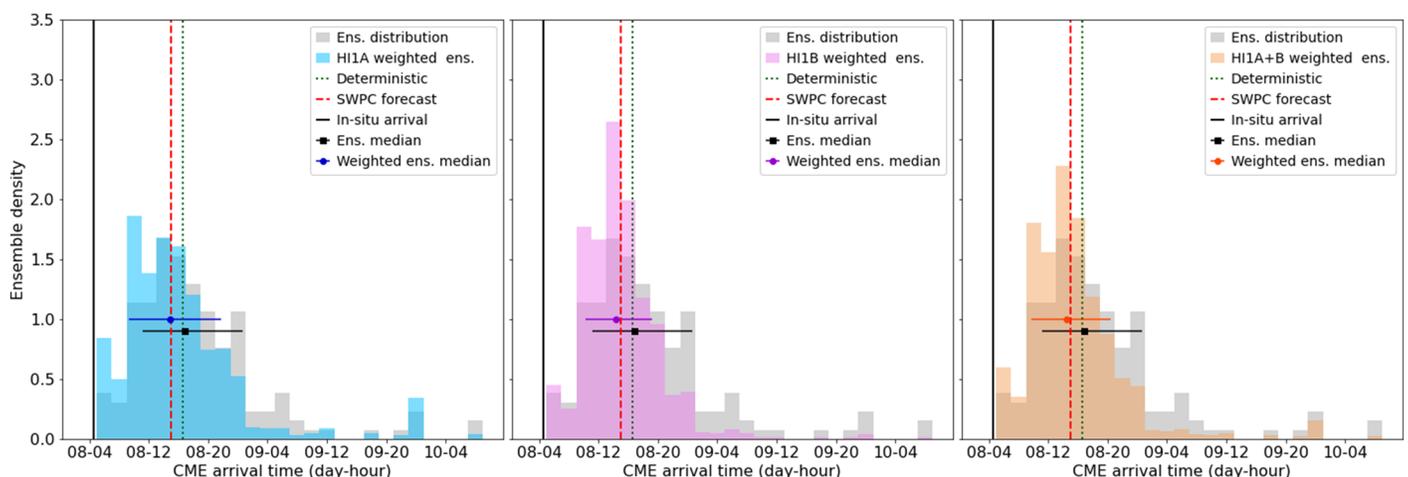
Figure 8. Histograms of the distribution of modeled CME arrival time for CME-2, in the same format as Figure 7.

leads all the ensemble members, being on the edge of the ensemble distribution, but outside the uncertainty band for each of the weighted and unweighted ensembles. The arrival errors of the ensemble medians are between 10 and 12.3 hr. Again, the medians of the weighted distribution are closer to the observed arrival time than the unweighted ensemble median, but these differences are smaller than the ensemble spread in each case. The spread of the arrival time distributions is reduced for each of the weighted ensembles, indicating smaller uncertainty (∼3 hr) in the expected arrival time. For the HI1A/B weighted ensemble, there is a 19% reduction in the arrival error, and a 21% reduction in the arrival uncertainty, relative the the unweighted ensemble.

3.4. CME-4 Arrival Time Distribution

For CME-4, 198 of the ensemble members hit Earth, and the distribution of arrival times is shown in Figure 10. The full range of the ensemble arrival time distribution spans 51 hr. In this instance, the observed arrival time leads many of the ensemble members, but is close to the center of the ensemble distribution, and inside the uncertainty band for each of the weighted and unweighted ensembles. The arrival errors of the ensemble medians are between 4.5 and 6.6 hr. Again, the medians of the weighted distribution are closer to the observed arrival time than the unweighted ensemble median, but these differences are smaller than the ensemble spread in each case. The spread of the arrival time distributions is less for each of the weighted ensembles, indicating a smaller uncertainty (∼4 hr) in the expected arrival time. For the HI1A/B weighted ensemble, there is a 20% reduction in the arrival error, and a 23% reduction in the arrival uncertainty, relative the the unweighted ensemble.

Figure 9. Histograms of the distribution of modeled CME arrival time for CME-3, in the same format as Figure 7.
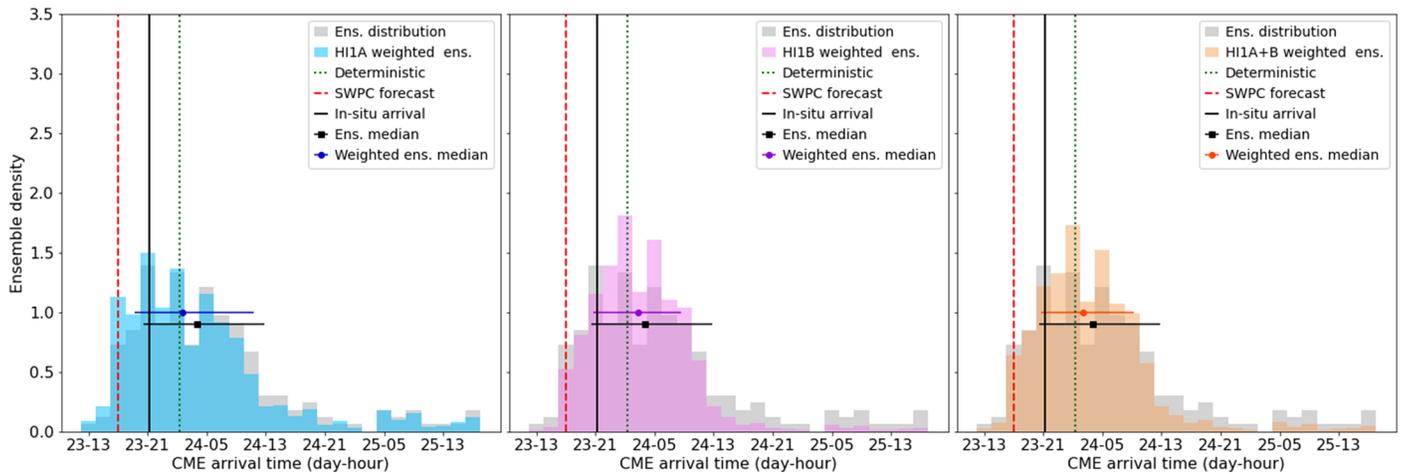
Figure 10. Histograms of the distribution of modeled CME arrival time for CME-4, in the same format as Figure 7.

For each of these four events, the weighted ensemble medians are closer to the observed arrival time than the unweighted ensemble and also have less uncertainty (with the exception of CME-2, for HI1A only). The results are similar when weighting by either HI1A, HI1B, or HI1A/HI1B combined. For the HI1A/HI1B weighting, there is a mean reduction in the arrival error of $20.1 \pm 4.1\%$, and a reduction in the arrival uncertainty of $15.0 \pm 7.2\%$, relative to the unweighted ensemble. These statistics are similar for the HI1A and HI1B weightings, but we use the HI1A/HI1B combined weighting to illustrate this as we think the combined weighting is probably more robust than weighting by a single perspective. We interpret this as evidence that the inclusion of the HI1 tracking data has improved the performance of the ensemble hindcast, relative to the unweighted ensemble.

Note that availability of photospheric magnetic field data, and hence the ambient solar wind speed input, were significantly different between the HUXt hindcasts and the SWPC genuine forecast. Thus, it isn't useful to directly compare their performance. The key conclusions we can draw are that HUXt can be used to perform a skilful hindcast of CME arrival time and that weighting the ensemble by comparison with HI observations can improve the skill and reduce the uncertainty.

4. Conclusions

We have used the HUXt solar wind model to produce an ensemble hindcast of four CMEs from 2012. By comparing the modeled and observed locations of the CMEs flanks, we have shown how HI observations can be used to weight ensemble members to provide an improved estimate of CME arrival time.

We have shown that HI observations can be used to constrain the ensemble hindcasts of these four CMEs and that this leads to reduced uncertainty in the CME arrival time hindcasts. However, it is important to note that uncertainty estimates are only useful if they are, on average, consistent with the observed arrival times. Of the four CMEs we study, the hindcasts of CME-2 and CME-4 are consistent with the observed L1 arrival, and so it is fair to say that the weighted ensemble has improved these hindcasts. For CME-1 and CME-3, the observed arrival leads all the ensemble members, and so the hindcast uncertainty is not consistent with the observed arrival, and the weighting cannot improve this situation. It is not clear why the observed arrival is outside the ensemble spread for these two events, especially CME-1 which arrives significantly earlier. We repeated these hindcasts with double the spread in the perturbed cone CME parameters and it did not significantly improve the situation. We are confident that the Solar Stormwatch identifications are robust and consistent with the CME identified by SWPC. Therefore, this suggests to us that HUXt isn't correctly capturing the evolution of this CME due to either the background solar wind environment or a breakdown of the assumptions of the cone CME parameterization, or perhaps both factors. Additionally, it is possible that the incorrect signature has been identified as this event in the L1 in situ data. Clearly, these issues need further investigation so that the technique proposed here can be developed into a robust tool for hindcasting, and potentially forecasting.

Therefore, following this work, several lines of research should be taken. First, we must test this method on more events to build a robust assessment of its performance. Solar Stormwatch is currently tracking over 1,000 CMEs in the same manner as Barnard et al. (2017), and these will be ready for analysis in 2020. Additionally, the time elongation profiles provided by the HELCATS project could be used (https://doi.org/10.6084/m9.figshare.5803152.v1); however, these do not track features as consistently as the Solar Stormwatch profiles.

Furthermore, a similar experiment could be performed with the WSA-ENLIL ensemble forecast runs. The approach advocated in Harrison et al. (2017) was an "Ensemble pruning" technique, where HI data are used to exclude ensemble members considered unrealistic compared to the HI observations. Given the low computational expense of HUXt, it might be beneficial to run a large ensemble using HUXt to compare against the HI observations, from which a targeted ensemble of the most plausible CME parameters could be run in WSA-ENLIL. Similarly, this approach could also be useful within the ADAPT-WSA-ENLIL modeling framework (Adamson et al., 2019).

The initial condition ensemble used here only examined uncertainty in the cone model CME parameters. However, there is also uncertainty in the ambient solar wind speed solution. Perturbation of the ambient solar wind solution, using a method similar to Owens and Riley (2017b), could help more fully quantify all sources of uncertainty; this is a practical option with a computationally efficient model such as HUXt.

It is also possible more advanced comparisons between the HI observations and modeled CMEs could be beneficial. Here we have limited comparison to the forward flank of the CME from the HI1A and HI1B perspectives. This consequently ignores much of the extra information regarding both the modeled and observed CME. The advantage of our approach is that there are several well established tools and processes for identifying the forward flank of a CME in HI data (Barnard et al., 2017; Davies et al., 2009). However, it is possible that further improvements could be gained by also tracking the rear flank of the CME, as per Savani et al. (2009), or other features such as the "ghost fronts" reported in Scott et al. (2019). Physics-based forward modeling of the HI images from the modeled solar wind and CME is also worth investigation. A Thomson scattering simulation of the modeled inner heliosphere could allow a more direct comparison with HI data, that doesn't require manually identifying CME properties. However, due to the nature of the cone model CME parameterization, this approach could be particularly challenging. Cone model CMEs are a hydrodynamic perturbation that ignores the magnetic and density structure of observed CMEs. This parameterization works well for estimating the front of CMEs in numerical solar wind models, but could be misleading when comparing real and forward modeled HI images. Therefore, it is quite plausible this approach would require more complex CME parameterizations to be effective.

The next operational space weather missions are currently being developed by NASA and ESA. Of particular note is the proposed Lagrange mission to the L5 point. Our results provide support for including a HI on such a mission, as we have demonstrated a simple way that these data can be used to improve the prediction of CME arrival time at Earth.

Conflict of Interest

Luke Barnard received a grant from NERC (Grant NE/P016928/1/) and personal fees from University of Reading for his work as a postdoctoral research assistant during the conduct of this study. Mathew Owens received a grant from NERC (NE/P016928/1/) during the conduct of this study..

Data Availability Statement

The STEREO HI data used in this study were obtained from the UK Solar System Data Centre (UKSSDC, https://www.ukssdc.ac.uk/solar/stereo/data.html). The HUXt model, including the boundary conditons used here, was obtained from this site (https://github.com/University-of-Reading-Space-Science/HUXt). The Solar Stormwatch classifications of the CME flank and the analysis code used in this study are made available at this site (https://github.com/University-of-Reading-Space-Science/HIEnsembleHindcast).

References

Adamson, E. T., Mays, M. L., Pizzo, V. J., Biesecker, D. A., MacNeice, P. J., Taktakishvili, A., et al. (2019). Update on NOAA SWPC / CCMC partnership to validate heliospheric models. In *AGU Fall Meeting Abstracts*, *2019*, pp. SM22B–08.

Arge, C. N., & Pizzo, V. J. (2000). Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates. *Journal of Geophysical Research*, *105*(A5), 10465. https://doi.org/10.1029/1999JA000262

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

Barnes, W. T., Bobra, M. G., Christe, S. D., Freij, N., Hayes, L. A., Ireland, J., et al. (2020). The SunPy project: Open source development and status of the version 1.0 core package. *The Astrophysical Journal*, *890*(1), 68. https://doi.org/10.3847/1538-4357/ab4f7a

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). *Solar Physics*, *162*(1-2), 357–402. https://doi.org/10.1007/BF00733434

Cannon, P., Angling, M., Barclay, L., Curry, C., Dyer, C., Edwards, R., et al. (2013). Extreme space weather: Impacts on engineered systems and infrastructure: Royal Academy of Engineering.

Davies, J., Harrison, R., Rouillard, P., Sheeley, N. R., Perry, C. H., Bewsher, D., et al. (2009). A synoptic view of solar transient evolution in the inner heliosphere using the heliospheric imagers on STEREO. *Geophysical Research Letters*, *36*, L02102. https://doi.org/10.1029/2008GL036182

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

Harrison, R. A., Davies, J. A., Biesecker, D., & Gibbs, M. (2017). The application of heliospheric imaging to space weather operations: Lessons learnt from published studies. *Space Weather*, *15*, 985–1003. https://doi.org/10.1002/2017SW001633

Howard, R., Moses, J. D., Vourlidas, A., Newmark, J. S., Socker, D. G., Plunkett, S. P., et al. (2008). Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI). *Space Science Reviews*, *136*(1-4), 67–115. https://doi.org/10.1007/s11214-008-9341-4

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jones, S. R., Barnard, L., Scott, C. J., Owens, M. J., & Wilkinson, J. (2017). Tracking CMEs using data from the Solar Stormwatch project; observing deflections and other properties. *Space Weather*, *15*, 1125–1140. https://doi.org/10.1002/2017SW001640

Kaiser, M. L., Kucera, T. A., Davila, J. M., St. Cyr, O. C., Guhathakurta, M., & Christian, E. (2008). The STEREO mission: An introduction. *Space Science Reviews*, *136*(1-4), 5–16. https://doi.org/10.1007/s11214-007-9277-0

Kay, C., Opher, M., & Evans, R. M. (2015). Global trends of CME deflections based on CME and solar parameters. *The Astrophysical Journal*, *805*(2), 168. https://doi.org/10.1088/0004-637X/805/2/168

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

Lang, M., & Owens, M. J. (2019). A variational approach to data assimilation in the solar wind. *Space Weather*, *17*, 59–83. https://doi.org/10.1029/2018SW001857

Luhmann, J. G., Gopalswamy, N., Jian, L. K., & Lugaz, N. (2020). ICME Evolution in the inner heliosphere. *Solar Physics*, *295*(4), 61. https://doi.org/10.1007/s11207-020-01624-0

Mays, M. L., Taktakishvili, A., Pulkkinen, A., MacNeice, P. J., Rastätter, L., Odstrcil, D., et al. (2015). Ensemble modeling of CMEs using the WSAENLIL+Cone model. *Solar Physics*, *290*(6), 1775–1814. https://doi.org/10.1007/s11207-015-0692-1

Merkin, V. G., Lyon, J. G., Lario, D., Arge, C. N., & Henney, C. J. (2016). Time-dependent magnetohydrodynamic simulations of the inner heliosphere. *Journal of Geophysical Research: Space Physics*, *121*, 2866–2890. https://doi.org/10.1002/2015JA022200

Mumford, S. J., Freij, N., Christe, S., Ireland, J., Mayer, F., Hughitt, V. K., et al. Naman9639, et al. (2020). SunPy. https://doi.org/10.5281/ZENODO.3779284

Murray, S. A. (2018). The importance of ensemble techniques for operational space weather forecasting. *Space Weather*, *16*, 777–783. https://doi.org/10.1029/2018SW001861

Odstrcil, D. (2003). Modeling 3-D solar wind structure. *Advances in Space Research*, *32*(4), 497–506. https://doi.org/10.1016/S0273-1177(03)00332-6

Owens, M. J., Lang, M., Barnard, L., Riley, P., Ben-Nun, M., Scott, C. J., et al. (2020). A computationally efficient, time-dependent model of the solar wind for use as a surrogate to three-dimensional numerical magnetohydrodynamic simulations. *Solar Physics*, *295*(3), 43. https://doi.org/10.1007/s11207-020-01605-3

Owens, M. J., Lockwood, M., & Barnard, L. (2017). Coronal mass ejections are not coherent magnetohydrodynamic structures. *Scientific Reports*, *7*(1), 4152. https://doi.org/10.1038/s41598-017-04546-3

Owens, M. J., Lockwood, M., & Barnard, L. A. (2020). The value of CME arrival time forecasts for space weather mitigation. *Space Weather*, *18*, e2020SW002507. https://doi.org/10.1029/2020SW002507

Owens, M. J., & Riley, P. (2017a). Probabilistic solar wind forecasting using large ensembles of near-sun conditions with a simple one-dimensional "upwind" scheme. *Space Weather*, *15*, 1461–1474. https://doi.org/10.1002/2017SW001679

Owens, M. J., & Riley, P. (2017b). Probabilistic solar wind forecasting using large ensembles of near-Sun conditions with a simple one-dimensional "upwind" scheme. *Space Weather*, *15*, 1461–1474. https://doi.org/10.1002/2017SW001679

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

Price-Whelan, A. M., Sipcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., et al. (2018). The astropy project: Building an open-science project and status of the v2.0 core package. *The Astronomical Journal*, *156*(3), 123. https://doi.org/10.3847/1538-3881/aabc4f

Riley, P., Linker, J. A., & Mikić, Z. (2001). An empirically-driven global MHD model of the solar corona and inner heliosphere. *Journal of Geophysical Research*, *106*(A8), 15,889–15,901. https://doi.org/10.1029/2000JA000121

Riley, P., & Lionello, R. (2011). Mapping solar wind streams from the Sun to 1 AU: A comparison of techniques. *Solar Physics*, *270*(2), 575–592. https://doi.org/10.1007/s11207-011-9766-x

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*, 1245–1260. https://doi.org/10.1029/2018SW001962

Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., et al. (2013). Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, *558*, A33. https://doi.org/10.1051/0004-6361/201322068

Savani, N. P., Owens, M. J., Rouillard, A. P., Forsyth, R. J., & Davies, J. A. (2010). Observational evidence of a coronal mass ejection distortion directly attributable to a structured solar wind. *The Astrophysical Journal*, *714*(1), L128–L132. https://doi.org/10.1088/2041‐8205/714/1/L128

Savani, N. P., Rouillard, A. P., Davies, J. A., Owens, M. J., Forsyth, R. J., Davis, C. J., & Harrison, R. A. (2009). The radial width of a coronal mass ejection between 0.1 and 0.4 AU estimated from the heliospheric imager on STEREO. *Annales Geophysicae*, *27*(11), 4349–4358. https://doi.org/10.5194/angeo-27-4349-2009

Scott, C. J., Owens, M. J., Koning, C. A., Barnard, L., Jones, S. R., & Wilkinson, J. (2019). Using ghost fronts within STEREO heliospheric imager data to infer the evolution in longitudinal structure of a coronal mass ejection. *Space Weather*, *17*, 539–552. https://doi.org/10.1029/2018SW002093

Tóth, G., Sokolov, I. V., Gombosi, T. I., Chesney, D. R., Clauer, C. R., De Zeeuw, D. L., et al. (2005). Space weather modeling framework: A new tool for the space science community. *Journal of Geophysical Research*, *110*, A12226. https://doi.org/10.1029/2005JA011126

Tucker-Hood, K., Scott, C. J., Owens, M. J., Jackson, D., Barnard, L., Davies, J. A., et al. (2015). Validation of a priori CME arrival predictions made using real-time heliospheric imager observations. *Space Weather*, *13*, 35–48. https://doi.org/10.1002/2014SW001106

Wharton, S. J., Millward, G. H., Bingham, S., Henley, E. M., Gonzi, S., & Jackson, D. R. (2019). Incorporation of heliospheric imagery into the CME analysis tool for improvement of CME forecasting. *Space Weather*, *17*, 1312–1328. https://doi.org/10.1029/2019SW002166