A multi-model ensemble system for the Outer Heliosphere (MMESH): solar wind conditions near Jupiter

Article

Accepted Version

Rutala, M. J. ORCID: https://orcid.org/0000-0002-1837-4057, Jackman, C. M. ORCID: https://orcid.org/0000-0003-0635-7361, Owens, M. J. ORCID: https://orcid.org/0000-0003-2061-2453, Tao, C. ORCID: https://orcid.org/0000-0001-8817-0589, Fogg, A. R. ORCID: https://orcid.org/0000-0002-1139-5920, Murray, S. A. ORCID: https://orcid.org/0000-0002-9378-5315 and Barnard, L. ORCID: https://orcid.org/0000-0001-9876-4612 (2024) A multi-model ensemble system for the Outer Heliosphere (MMESH): solar wind conditions near Jupiter. Journal of Geophysical Research: Space Physics, 129 (6). e2024JA032613. ISSN 2169-9402 doi: https://doi.org/10.1029/2024JA032613 Available at https://centaur.reading.ac.uk/116874/

To link to this article DOI: http://dx.doi.org/10.1029/2024JA032613

Publisher: American Geophysical Union

A Multi-Model Ensemble System for the outer Heliosphere (MMESH): Solar Wind Conditions near Jupiter

M. J. Rutala[1], C. M. Jackman[1], M. J. Owens[2], C. Tao[3], A. R. Fogg[1], S. A. Murray[1]

[1]School of Cosmic Physics, DIAS Dunsink Observatory, Dublin Institute for Advanced Studies, Dublin, Ireland
[2]Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK
[3]Space Environment Laboratory, National Institute of Information and Communications Technology (NICT), Koganei, Japan

Key Points:

- The performance of several existing solar wind propagation models at the orbit of Jupiter is measured for multiple spacecraft epochs.
- A flexible system is developed to generate an ensemble of multiple propagation models in order to best leverage each input model's strengths.
- Over the epoch tested, the multi-model ensemble outperforms individual input models by $7\% - 110\%$ in forecasting the solar wind flow speed.

Corresponding author: M. J. Rutala, `mrutala@cp.dias.ie`

Abstract

How the solar wind influences the magnetospheres of the outer planets is a fundamentally important question, but is difficult to answer due to the lack of consistent upstream monitoring of the interplanetary medium (IPM) and the large-scale dynamics internal to the magnetosphere. This makes it very challenging to link external drivers with the magnetospheric dynamics measured by a single orbiting spacecraft. To compensate for the relative lack of in-situ data, solar wind propagation models are often used to estimate the ambient IPM conditions at the outer planets for comparison to remote observations or in-situ measurements. This introduces another complication: the propagation of near-Earth measurements of the solar wind introduces uncertainties in both timing and magnitude which are themselves difficult to assess. Here, we present the Multi-Model Ensemble System for the outer Heliosphere (MMESH) to begin to address these issues, along with the resultant multi-model ensemble (MME) of the solar wind conditions near Jupiter as a means to assess the system. MMESH accepts as input any number of solar wind models together with contemporaneous in-situ spacecraft data. From these, the system characterizes typical uncertainties in model timing, quantifies how these uncertainties vary under different conditions and time periods, attempts to correct for systematic biases in the input model timing, and composes a multi-model ensemble (MME) with uncertainties from the results. For the case of the Jupiter-MME here, three solar wind propagation models were compared to in-situ measurements from near-Jupiter spacecraft spanning diverse spacecraft-Sun-Earth alignments and phases of the solar cycle, amounting to more than 23,000 hours over four decades. The resulting MME produces the most-probable near-Jupiter IPM conditions for times within the tested epoch. Finally, we will discuss how the work presented here can be extended towards more robust characterization of solar wind parameters and time-dependent propagation of solar wind conditions at other planetary magnetospheres.

1 Background

The solar wind is a continuous stream of plasma emanating from the Sun in all directionswhich evolves as it travels through the heliosphere, interacting with every planetary magnetosphere in the solar system along the way. Near the Earth, the typical values of the solar wind flow speed $u_{mag,\oplus}$(324−584 km/s), proton density $n_{\oplus}$(2.2−12.7 cm^{-3}), dynamic (ram) pressure $p_{dyn,\oplus}$(0.86−3.92 nPa), and interplanetary magnetic field (IMF) magnitude $B_{mag,\oplus}$(3.1−9.7) nT have all been statistically characterized by the expansive OMNI dataset (King & Papitashvili, 2005; Papitashvili & King, 2020), with values here spanning the start of OMNI2 to the start of 2023 (1963/11/27 – 2023/01/01) and characterizing 80% ($10^{th} - 90^{th}$ percentiles) of all measurements. The OMNI dataset is a composite of many near-Earth observations encompassing some 19 total spacecraft over its full time domain, including most recently *Wind* (Lepping et al., 1995; Kasper, 2002; King & Papitashvili, 2005) and ACE (D. J. McComas et al., 1998; Smith et al., 1998; King & Papitashvili, 2005).

While fewer in-situ heliospheric data are available in the outer solar system, the average solar wind conditions have still been constrained by the various spacecraft to visit the outer planets, whether during planetary flyby or approach. At Jupiter, the most-visited of the outer planets, the solar wind has been characterized during the flybys of *Pioneers 10* and *11*, *Voyagers 1* and *2*, *Ulysses*, *Cassini*, and *New Horizons* (e.g. Slavin et al., 1985; J. D. Richardson et al., 1995; Ebert et al., 2014; Hanlon et al., 2004; Ebert et al., 2010). Compared to flybys, orbiter missions, including *Galileo* and *Juno* at Jupiter, generally provide fewer in-situ data: these missions have close-in orbits to best study the planet itself, thus setting them deep inside the planet's magnetosphere and shielding them from the solar wind. As a result, they only sample the wind during the planetary approach phase prior to orbital insertion and occasional excursions into the solar wind near apoapsis. The polar solar orbiter *Ulysses* gives the best single-spacecraft characterization of the average near-Jupiter solar wind owing to its 18-year lifetime: 80% of *Ulysses* measurements span 380−520 km/s in solar wind flow speed u_{mag}, 0.05−0.55 cm^{-3} in plasma density n, $0.02-0.20$ nPa in dynamic pressure p_{dyn}, and $0.22-1.5$nT in IMF magnitude B_{IMF} (Ebert et al., 2014). Despite a large number of measurements, these reported distributions are only approximate due to the polar orbit of the *Ulysses* spacecraft; *Ulysses* samples the solar wind in the ecliptic plane periodically, and these numbers were drawn from two non-consecutive spans at different phases of the solar cycle– one with a slower, cooler, and denser average solar wind than the other (Ebert et al., 2014).

The highly dynamic nature of the solar wind is not captured by these average values. Singular events, such as the eruption of coronal mass ejections (CMEs), and their propagation through the heliosphere as interplanetary coronal mass ejections (ICMEs), are a major source of short timescale variation in the measured solar wind (Palmerio et al., 2021, and references therein). In terms of the quantities already discussed, interplanetary coronal mass ejections (ICMEs) show expansion, which manifests in measurements as an increase in u_{mag} at the leading edge and a decrease at the trailing edge, large drops in n, and an enhancement in B_{IMF} magnitude but decrease in B_{IMF} variance (Zurbuchen & Richardson, 2006; M. J. Owens, 2018). Beyond these events, the ambient solar wind is dynamic due to the presence of two different streaming plasma populations originating in different regions of the solar corona: a comparatively fast, hot, and tenuous stream and a comparatively slow, cool, and dense stream (Crooker et al., 1999). These streams are essentially bimodal during solar minimum, with fast streams originating at high heliolatitude and slow streams originating nearer the solar equator (D. J. McComas et al., 1998, 2000); during solar maximum, these streams are markedly less ordered (D. J. McComas et al., 2003). From solar cycle to solar cycle, the bulk parameters of the fast stream in particular can change dramatically (D. McComas et al., 2008; Ebert et al., 2009; D. J. McComas et al., 2013; Ebert et al., 2014). As different regions of the sun rotate underneath, corotating interacting regions (CIRs) are formed where a fast flow catches up to a slow flow; this process is common throughout the heliosphere (I. G. Richardson, 2018) and drives significant interactions with planetary magnetospheres, including at the Earth (Crooker et al., 1999; Gosling & Pizzo, 1999; Tsurutani et al., 2006; Borovsky & Denton, 2010) and at Jupiter (D. J. McComas et al., 2003; Hanlon et al., 2004; Ebert et al., 2014).

In-situ-data-driven statistical studies of the time variable solar wind at specific locations within the outer heliosphere (e.g. at Jupiter) are hampered by the limited temporal coverage of visiting spacecraft; there is no continuous composite model like OMNI for any outer planet. Such statistical studies often instead have solar wind data supplemented by solar wind propagation models, which attempt to reproduce the time-varying solar wind at one location from measurements at another location at which the solar wind is known. Many of these models have been employed in the outer heliosphere, including, but not limited to, the model of Tao et al. (2005) ("Tao+", hereafter), ENLIL (Odstrcil, 2003), mSWiM (Zieger & Hansen, 2008), HUXt (Barnard & Owens, 2022; M. Owens et al., 2020), and MSWIM2D (Keebler et al., 2022). These models all differ in their dimensionality, the simplifications made to the magnetohydrodynamics (MHD) equations underlying them, and the source of the input solar wind conditions used to initialize the model. By virtue of modelling solar wind conditions for times and locations where no in-situ spacecraft measurements are available, the outputs of these models cannot be directly compared to data in typical usage scenarios. Generally, solar wind propagation models are instead compared to in-situ spacecraft measurements at times and locations where they are available in order to approximate the model errors– generally, shock arrival time (or "timing") errors– prior to being used to supplement the data (Tao et al., 2005; Zieger & Hansen, 2008; Keebler et al., 2022). Measured timing uncertainties can be as high as ±4 days and often trends with other physical parameters of the system, such as with Target-Sun-Observer (TSO) angle (Tao et al., 2005; Zieger & Hansen, 2008; Keebler et al., 2022) or with phase of the solar cycle (Zieger & Hansen, 2008).

These resulting time-varying timing uncertainties introduce a challenge in interpreting the results of these models and performing statistical analyses, particularly because the characterizations of timing uncertainty in each propagation model are often not measured by the same methods, and thus are not directly comparable to one another. Timing uncertainties can be measured by manually identifying shocks and shock-like structures in both modeled and measured solar wind time series and comparing their occurrence times (e.g. Tao et al., 2005) or by offsetting one time series relative to the other and maximizing the resulting prediction efficiency, or Pearson correlation coefficient (e.g. Zieger & Hansen, 2008). Measuring uncertainties with the latter method implies that a single timing uncertainty characterizes the model over the full time period inspected. An alternative to this is to employ dynamic time warping to explicitly allow for time-varying timing uncertainties (e.g. Samara et al., 2022). If these model uncertainties were quantified in a cross-model-consistent manner, the time-varying uncertainties could be accounted for and partially mitigated. For instance, as propagation model output uncertainties are known to trend with physical quantities, each individual model's outputs could be de-trended with sufficient characterization of the uncertainties. Alternatively, a multi-model ensemble (MME) could be composed by cross-comparison of the models in order to mitigate uncertainties. An MME is, in essence, a weighted average of different model outputs (Murray, 2018); the weighting scheme can be adjusted based on metrics of the models performance (or "skill") during intervals where in-situ data are available (Murray, 2018; Elvidge et al., 2023). Ideally, fully-independent models would be used in an MME, so that they would be expected to have independent random errors which would thus tend to cancel, rather than add (Hagedorn et al., 2005; Riley et al., 2018). If all input models capture the same physics, outperform one another in different parameter spaces, and have independent errors, a MME of these models should describe the underlying physical system more accurately than any individual input.

Here we present the Multi-Model Ensemble System for the outer Heliosphere (MMESH): a framework to quantify and mitigate timing uncertainties in solar wind propagation models and produce a single prediction by combining all of these approaches. This system allows for the automatic quantification of model timing uncertainties, trending of timing uncertainties with physically relevant parameters, de-trending of the original model timing, and combination of distinct models into a single MME. MMESH is designed to flexibly compare any combination of input solar wind propagation models and contemporaneous in-situ data in order to create an MME. To demonstrate this concretely, here we construct an MME of the solar wind conditions at Jupiter during the *Juno* era.

Thus prior to discussing MMESH itself, we first discuss the in-situ spacecraft datasets to be used for comparison (Section 2.1) and give some introduction to the specific solar wind propagation models considered here (Section 2.2). We then introduce the MMESH framework in Section 3, beginning with a description of the statistical techniques and tools used to compare models, including the MME, to contemporaneous data and measure their performance (Section 3.1. In Section 3.2 we discuss the methods available to characterize the model timing uncertainties relative to the in-situ time series: constant time offsetting (Section 3.2.1) and dynamic time warping (DTW, Section 3.2.2). We then proceed to describe how trends in the empirical timing uncertainties are characterized and estimated for epochs without contemporaneous in-situ data (Section 3.3) before discussing the composition (Section 3.4) and performance (Section 3.4.1) of the multi-epoch MME composed of the de-trended models. Having described MMESH, we then present the MME of the solar wind conditions at Jupiter for the first 7 years of the *Juno* mission, spanning 2016/7/4 – 2023/7/4, for use in future statistical analyses (Section 4), prior to concluding.

Table 1. In-situ measurements of solar wind parameters near Jupiter's orbit.

Mission	Coverage [yyyy/mm/dd]	Range [AU]	Heliolatitude [deg]	Measurements [hr]
Ulysses	1991/12/08 – 1992/02/02	4.90 – 5.41	−6.10 – +6.10	1,344
	1997/08/14 – 1998/04/16	4.90 – 5.41	−6.10 – +6.10	5,878
	2003/10/24 – 2004/06/22	4.90 – 5.41	−6.10 – +6.10	5,801
Juno	2016/05/15 – 2016/06/29	5.27 – 5.44	−5.76 – −5.23	1,080

2 Inputs

2.1 Solar Wind Data

The present aim for the MME framework discussed here is to find the most accurate combination of solar wind models in the near-Jupiter region of the outer heliosphere. As such, limiting the data included for comparison to the input and ensemble models to that which is representative of conditions at Jupiter is essential. Including too large a range of radial or helio-latitudinal in-situ measurements risks including different regimes of solar wind properties which the models are not, and should not be, expected to reproduce. This is particularly an issue in choosing a useful range of heliolatitude– too narrow a range and the amount of data available shrinks, but too large a range and the faster solar wind flows at higher heliolatitudes are included erroneously. This issue primarily relates to data acquired by the *Ulysses* spacecraft, which is a solar polar orbiter. Previous *Ulysses* measurements show that, during solar minimum when the latitudinal structure of the solar wind is well-ordered, the equatorial slow solar wind zone may extend to $\pm 20° - \pm 30°$ about the solar equator (D. J. McComas et al., 2003). Ebert et al. (2014) further restricts this range in surveying near-Jupiter solar wind conditions measured with *Ulysses* and selects for data $\pm 10°$ about the solar equator.

Here, the near-Jupiter outer heliosphere is defined as the region of the heliosphere spanning 4.9 AU $< r <$ 5.5 AU for spherical distance from the Sun r and $-6.1° \leq \theta \leq 6.1°$ for heliolatitude θ. Jupiter's perihelion and aphelion (5.04 and 5.37 AU, respectively) fit entirely within this range, which includes padding of ~ 0.15 AU, or approximately 50%, on either end to increase the number of observations included. The heliolatitude range selected represents the maximal range of Jupiter's location in heliolatitude without any padding in order to avoid unrealistic sampling of the high latitude fast solar wind flows.

Several spacecraft have transited this region, including *Pioneers 10* and *11*, *Voyagers 1* and *2*, *Ulysses*, *Galileo*, *Cassini*, *New Horizons* and *Juno*. Here, just data from just the *Ulysses* and *Juno* missions are used; the remaining spacecraft are not used in this analysis either due to being discontinuous at temporal resolutions of 1 hour (*Galileo*,*Cassini*, and *New Horizons*) or due to a lack of coverage in all or some of the models to be discussed in Section 2.2 (*Pioneers 10* and *11* and *Voyagers 1* and *2*). A brief overview of the used spacecraft trajectories and data is included in Table 1 and the durations of the visits of these spacecraft to the near-Jupiter outer heliosphere is illustrated in Figure 1 relative to the solar cycle, as measured by F10.7 radio flux derived from observations at the Dominion Radio Astrophysical Observatory (DRAO) and adjusted to account for variations in the Earth's distance from the Sun. While the majority of these spacecraft passed near Jupiter, the *Ulysses* spacecraft, as a polar orbiter, transits through the near-Jovian outer heliosphere away from the planet itself after its initial Jupiter flyby. The relevant orbital components for all the spacecraft in Table 1 are shown in Figure 2, which highlights the rarity of near-Jupiter outer heliosphere measurements made far from Jupiter itself and the comparative evenness of coverage in Target-Sun-Earth angle.

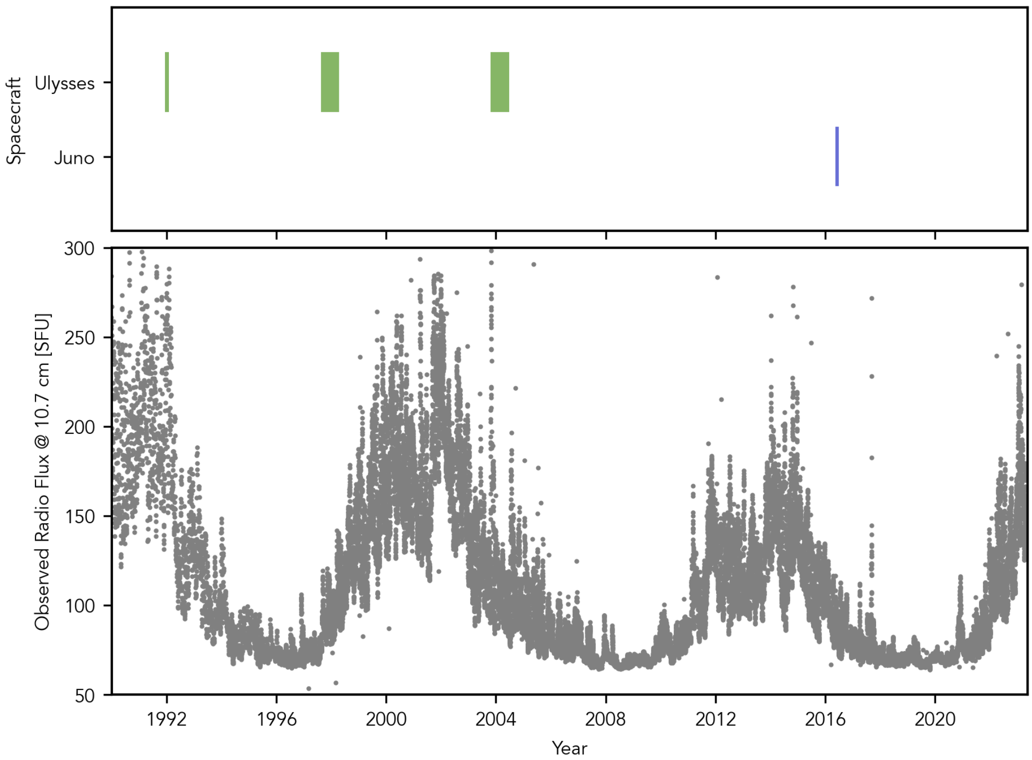

Figure 1. The (a) spans during which each spacecraft used in this analysis was measuring the near-Jupiter solar wind compared with the (b) solar F10.7 cm radio flux, a proxy for the phase of the solar cycle, over the period 1990-2023. Spacecraft coverage spans the ascending and descending phases of the solar cycle, but largely excludes solar minimum and solar maximum. These spacecraft have been selected for the frequency of their plasma and magnetic field measurements, which are generally hourly or better.

All of the spacecraft referenced in Table 1 have both magnetometers and plasma instruments, and thus provide sampling of the interplanetary magnetic field (IMF) B_{IMF}, the solar wind ion number density n, and the magnitude of the solar wind flow speed u_{mag}, which is itself dominated by the radial component of the outward flow of the solar wind. As the proton density n_p is measured in all cases and protons are the dominant ion component of the solar wind (Ebert et al., 2014, e.g.), the total density of the solar wind is approximately equal to the proton density ($n \approx n_p$) and is assumed to be exactly equal in calculating the solar wind dynamic pressure $p_{dyn} = m_p n u_{mag}^2$, where m_p is the proton mass. Detailed descriptions of these instruments, including their heritages, limitations, and data products, are discussed in their respective instrument papers (Balogh et al., 1992; Bame et al., 1992; Connerney et al., 2017; D. J. McComas et al., 2017). Pre-processed data was obtained from the Goddard Space Flight Center (GSFC) Space Physics Data Facility (SPDF) COHOWeb archives, with the exception of *Juno* plasma data which was instead obtained from Wilson et al. (2018).

2.2 Solar Wind Models

While several solar wind propagation models for the outer heliosphere are available, three were chosen for detailed study and inclusion in the MME: the Tao+ (Tao et al., 2005), ENLIL (Odstrcil, 2003), and HUXt (M. Owens et al., 2020; Barnard & Owens,

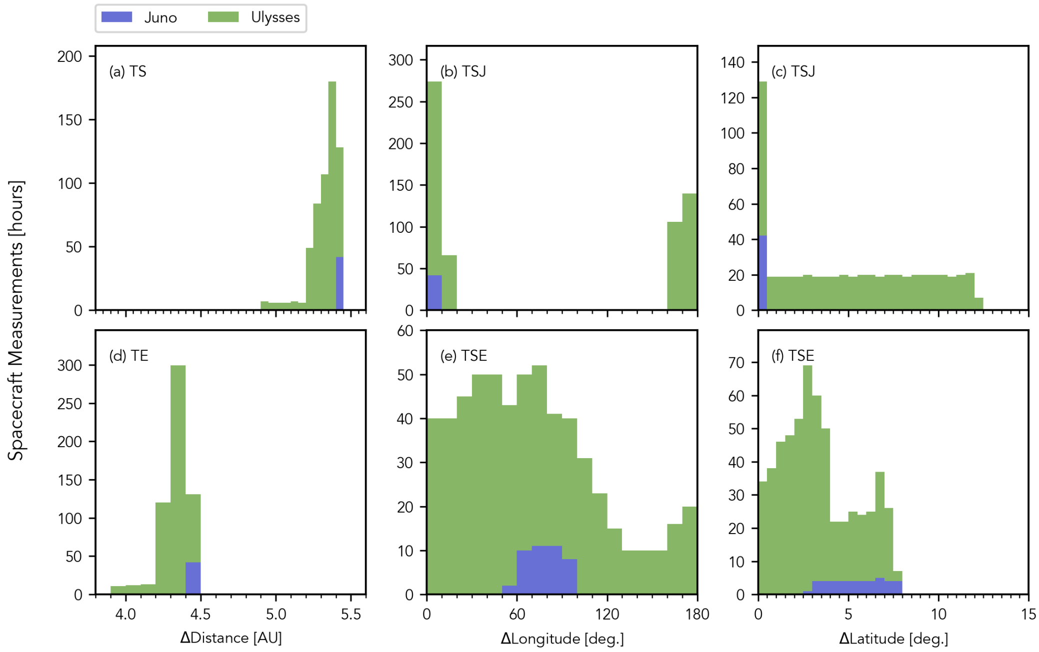

Figure 2. Histograms showing the spatial coverage of all the spacecraft used here, including the (a) Target-Sun (TS) distance, (b) Target-Sun-Jupiter (TSJ) longitude angle, (c) TSJ latitude angle, (d) Target-Earth (TE) distance, (e) Target-Sun-Earth (TSE) longitude angle, and (f) TSE latitude angle. The angles are measured in the Sun's inertial reference frame, such that longitude measures distance along the solar equator and latitude measures perpendicular distances along the sphere of the Sun. The majority of spacecraft measurements occur very near Jupiter, with minimal separation in TSJ longitude or latitude angles. The unique coverage of the polar-orbiting *Ulysses* spacecraft stands out, and provides even coverage across TSJ, and to a lesser extent TSE, latitudes. Taking all spacecraft into consideration, the spatial coverage relative to the Earth's location is fairly even.

2022) models. These models in particular are ideal for inclusion in a MME due to their differing input parameters, dimensionality, and approaches to propagating the solar wind beyond the Earth, as is summarized in Table 2 and will be discussed further here.

Fundamentally, most models propagate solar wind conditions outwards by solving the system of equations which constitute MHD, these being: the continuity equation, the momentum equation, the equation of state, and several physical laws necessary to close the system (Faraday's, Ohm's, and Ampère's laws). Propagation models differ primarily in their treatment of the momentum equation. For a single-species plasma composed of protons, this is:

$$\frac{\partial(m_p n\vec{u})}{\partial t} + \rho(\vec{u}\cdot\nabla)\vec{u} = -\underbrace{\nabla p}_{pressure} + \underbrace{\vec{j}\times\vec{B}}_{Lorentz} - \underbrace{\frac{GM_\odot \rho}{r^2}\hat{r}}_{gravity} + \underbrace{\nu\nabla^2\vec{u}}_{collision} \tag{1}$$

where m_p is the proton mass, n is the plasma number density, $\vec{u}$ is the plasma flow velocity, p is the total plasma pressure, $\vec{j}$ is the plasma current density, $\vec{B}$ is the ambient magnetic field, G is the gravitational constant, $M_\odot$ is a solar mass, r is the radial distance in a heliocentric spherical frame with the $\hat{r}$ direction pointing radially outward, and ν is a collisional frequency. In Equation 1, the right-hand-side terms are labelled cor-

Table 2. Descriptive parameters of solar wind propagation models as used in this study.

Model	Type	Inner Boundary[a] [AU]	MHD Terms[b] (P, L, G, C)	Input[a] type (source)	Output[c] $(n, u_{mag}, p_{dyn}, B_{IMF})$
ENLIL	3D MHD	∼0.1	(P, L, G)	remote (WSA)	$(n, u_{mag}, p_{dyn}, B_{IMF})$
HUXt	1D HD	∼1	–	in-situ (OMNI)	(u_{mag})
Tao+	1D MHD	∼1	(P, L, G)	in-situ (OMNI)	$(n, u_{mag}, p_{dyn}, B_{IMF})$

[a] Inner boundaries and input types are reported for the versions of the models used here. The models are not necessarily limited to these inner boundaries and input types only, as described in the text.
[b] The (P)ressure, (L)orentz, (G)ravitational, and (C)ollisional terms of the governing MHD momentum equation (Eqn. 1).
[c] Components of the solar wind: plasma density (n), plasma flow speed (u_{mag}), plasma dyamic pressure (p_{dyn}), or IMF (B_{IMF}).

responding to the physical forces they represent, these being the (gradient) pressure, Lorentz, gravitational, and collisional forces, respectively. As summarized in Table 2, solar wind propagation models differ in which terms of the momentum equation they assume are insignificant in the solar wind. Most propagation models, including all those discussed here, do not consider collisional forces within the solar wind plasma. Both ENLIL and Tao+ keep all the remaining terms shown in Equation 1 (Tao et al., 2005; Odstrcil, 2003). HUXt assumes that all forces are negligible compared to the magnitude of the left-hand-side momentum terms in Equation 1, and thus does not consider any force terms (M. Owens et al., 2020).

The variables propagated by each model are directly related to the force terms that they consider in Equation 1, and are listed in Table 2 for the three models discussed here. The dimensionality of each model changes which components of the vector terms in Equation 1 can be propagated; for cross-model consistency, we therefore compare solar wind parameter magnitudes rather than vector components, where each magnitude is calculated as the root-sum-square of available components. The solar wind flow speed u_{mag} is thus available from all three propagation models considered here. This is the only parameter available from HUXt; none of the solar wind density n, temperature T, or IMF strength B_{IMF} are propagated as these variables are eliminated from the version of the momentum equation used. These parameters– density n, temperature T, and IMF strength B_{IMF} of the propagated solar wind– are available from both ENLIL and Tao+.

Each of these models has an inner boundary at which the conditions of the solar wind are input and continuously updated over the course of the model run. The location of this inner boundary and the sources from which the input solar wind conditions are drawn vary between models and are summarized in Table 2. ENLIL takes as input a 3-dimensional description of the solar corona and near-sun environment, here supplied by the Wang-Sheeley-Arge (WAS) model (Arge & Pizzo, 2000) which itself takes remote observations of the sun as input. For this study, solar magnetograms from the Kitt Peak Observatory are used, with gaps in observations filled in by those from the Mount Wilson Observatory. This sort of boundary is unique amongst the models considered here: HUXt and Tao+ instead take in-situ spacecraft measurements, or proxies thereof, as inputs. In this study, both models take OMNI measurements at ∼1 AU as inputs, although they both have the functionality to be run at any other location in the solar system, provided there are sufficient in-situ solar wind data available (e.g. Sanchez-Diaz et al., 2016; Barnard & Owens, 2022). Accuracy in these input solar wind conditions are the single largest factor in determining the propagated solar wind accuracy (Riley et al., 2018), and as such including a variety of inputs is beneficial to the final MME.

The input solar wind conditions used here are assumed to be sampled from the background solar wind. This means that coronal mass ejections (CMEs) sampled at the model inner boundary are not propagated using the standard cone model (Zhao et al., 2002; Xie et al., 2004) but are instead interpreted as fast solar wind flows; rather than propagate CMEs as radially-expanding regions of constant angular size, they are treated by the same fluid description used by each model to describe the rest of the solar wind flow. This introduces an intrinsic error into the background solar wind parameters in all of the models. Future studies could mitigate this additional source of error by subtracting CMEs from the input data prior to propagation, then simultaneously propagating the quiescent solar wind and the CME using the cone model, but such an involved change to the modeling is ultimately beyond the current scope of this project.

These three models each run at different spatial and temporal resolutions which are directly related to their dimensionality and domains within the heliosphere, and which directly impact the small-scale shape of their output propagated solar wind estimates. ENLIL covers three spatial dimensions, spanning $0.1-10$ AU radially at 0.02 AU resolution, 360^o in longitude at 2^o resolution, and $\pm 60^o$ in latitude at 2^o resolution, with a temporal resolution of 1 hour. HUXt is physically a one-dimensional radial model, but in practice here it is run in its two-dimensional form in order to more easily sample the model at the spacecraft position. Functionally, the two-dimensional form of HUXt is a series of independent one-dimensional models spanning $1-6$ AU radially with a resolution of 0.007 AU, 360^o in longitude at $\sim 2.8^o$ resolution, and an intrinsic temporal resolution of 17.4 minutes in the version of the model used here. Tao+ spatial dimension, ranging from $1-8$ AU at a resolution of 1/300 AU, with an intrinsic temporal resolution of 10 s. The outputs of both HUXt and Tao+ have been downsampled to a resolution of 1 hour to better match the spacecraft data and other models for use in this study.

Figure 3 shows the model-propagated solar wind flow speed u_{mag} during the *Juno* cruise towards Jupiter compared with contemporaneous JADE in-situ measurements from Wilson et al. (2018) for each of the models detailed here. While these models are all able to propagate solar wind conditions during the other spacecraft epochs shown in Table 1, and both ENLIL and Tao+ are able to propagate parameters other than u_{mag}, here we have chosen to show just a single-spacecraft and single-parameter comparison for illustrative purposes. The agreement between each model and the data in general form is clear, but significant deviations in the arrival time of large-scale shocks and smaller-scale increases in flow speed between the models and data are evident. These temporal lags, which represent single measurements of the full distribution of model timing uncertainties, appear to be of the same sign for Tao+ and HUXt but are substantially different for ENLIL. Characterizing these differences in arrival time is critically important to understanding the accuracy of these models in propagating the solar wind, and will be further explored here.

3 Description of MMESH

The clear disagreements between the propagation models and in-situ data in both the modeled arrival time and magnitude, as illustrated in Figure 3, makes the need for careful consideration of uncertainties and new statistical approaches in solar wind propagation modeling evident. MMESH has been designed as a framework to tackle these issues. After briefly introducing the statistical metrics used in quantifying model performance (Section 3.1), the MMESH framework will be described. This system allows any number of solar wind propagation models to be compared to simultaneous in-situ data; from this comparison, timing uncertainties are characterized either as a constant value over the full duration of each model (i.e. as a bias, as explored in Section 3.2.1)

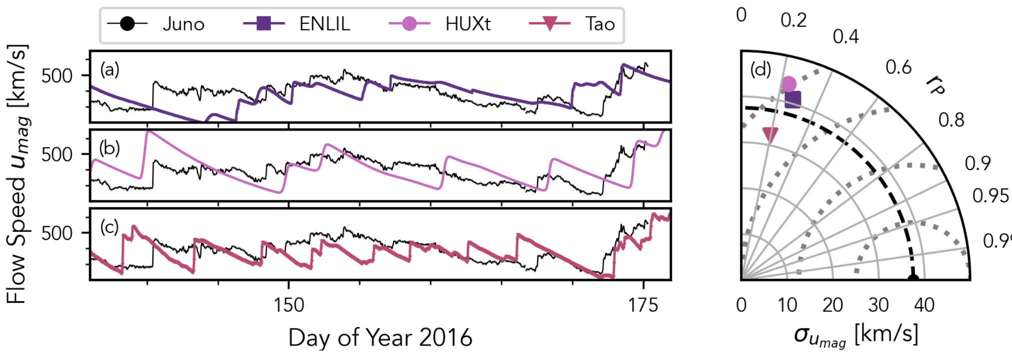

Figure 3. Measured solar wind flow speed u_{mag} from *Juno* JADE moments (Wilson et al., 2018) with the same from the (a) ENLIL, (b) HUXt, and (c) Tao+ models, as labeled, with (d) a Taylor diagram illustrating the performance of each model relative to the data, as discussed in Section 3.1 The flow speed is referenced as the root-mean-square of all velocity components, where components are available. Temporal lags in the timing of the modeled solar wind flow speed u_{mag} are apparent in all models, and are made evident by the Taylor diagram.

or as a dynamic value (Section 3.2.2). This framework fundamentally supports a multi-epoch analysis, in which the same timing uncertainties are quantified over multiple spacecraft epochs, each with one set of in-situ data and multiple models, in order to better characterize the model timing uncertainties, including any timing biases (Section 3.3). From this characterization, the solar wind propagation models can then have any identified biases in timing removed before being assemble into an MME (Section 3.4).

3.1 Performance Metrics

The correlation coefficient$r$, as a robust measure of model goodness-of-fit, is a good metric to be maximized in optimizing the alignment of solar wind model to data, as will be discussed in Section 3.2.1. For simple methods of aligning the model and data, the correlation coefficient r is sufficient alone as a metric. More complex methods of alignment, such as discussed in Section 3.2.2, are better optimized while considering some penalty against increasing complexity, in order to maintain physical realism and interpretability. In this case, a statistic determined by both the correlation coefficient and some measure of the width of the distribution of timing uncertainties is preferred for optimization. Such a statistic is less likely to reach its maximum value when a large range of timining uncertainties are predicted, thus preventing unphysical alignment of a model with, for instance, a shock-like structure from a previous Carrington rotation. Within MMESH, we define σ_T to be the half-width containing 34% of the distribution of timing offsets, such that it would reduce to one standard deviation in a normal distribution. The optimization metric for these cases is then defined as $r+(1-\sigma_T/\Delta T)$, where ΔT represents half the largest allowed magnitude of a timing uncertainty, such that the statistic varies between $0-2$, with the former corresponding to the worst performance and the latter corresponding to the best.

Neither the correlation coefficient r nor the statistic r/σ_T as single numbers fully characterize how closely a model matches data on different scales. Combining the correlation coefficient r with the overall standard deviation of both the time series and the model residuals forms the basis for a more complete multi-scale comparison of model and data summarized by the Taylor diagram (Taylor, 2001), illustrated in Figure 6. This type of plot relates the standard deviation of the modeled time series, the correlation coefficient of the modeled time series relative to the measured time series, and the centered root-mean-square difference between the modeled and measured time series to one another by analogy with the law of cosines, allowing all three quantities to be displayed as a single point on the diagram. This is particularly useful for comparing the performance of different models to one another on the same axes. All-around better models– those with high correlation coefficients, small residuals when compared with the data, and similar intrinsic variances– appear graphically closer to the point representing the data time series along the x-axis.

3.2 Characterization of Propagation Model Performance

The arrival time of shocks is of particular interest in statistical studies both at Jupiter and elsewhere in the outer heliosphere; the arrival of a shock is expected to compress the magnetosphere, directly impacting plasma and magentic flux transport and auroral activity (Southwood & Kivelson, 2001; Cowley et al., 2003; Vogt et al., 2019; Nichols et al., 2019; Kita et al., 2019). While individual models typically quote some uncertainty in modeled arrival times (Tao et al., 2005; Zieger & Hansen, 2008; M. Owens et al., 2020), these uncertainties are often characterized relative to different standards and using different methods, making cross-model comparisons difficult.

To allow direct comparisons of outer heliosphere solar wind models, independent quantification of modeled arrival time uncertainty can be performed with MMESH, as is common for near-Earth solar wind modeling (Gressl et al., 2014; Riley et al., 2018). The goal in quantifying the arrival time uncertainty is twofold: understanding the error intrinsic to each model is necessary to give context to its forecasts, and characterizing these errors can give clues as to which aspects of the solar wind system an individual model may not be capturing sufficiently. For both of these reasons, here we explore two methods available in MMESH of quantifying the arrival time uncertainties in the previously discussed models. These comparisons and uncertainty characterization are performed identically for every combination of spacecraft and model previously discussed; to keep illustrations of these informative and uncluttered, the *Juno* in-situ solar wind flow speed u_{mag} measurements will again be used alone.

3.2.1 Constant Time Offsetting

A simple metric to characterize the performance of a propagation model is to calculate the prediction efficiency, or correlation coefficient, between the propagated time series and an in-situ measurement of the same quantity (Zieger & Hansen, 2008; Keebler et al., 2022, e.g.). This offers a straightforward method to determine systematic, spacecraft-epoch-wide propagation model errors in the arrival time of shocks and other solar wind structures. The time span covered by the model can be shifted off that of the measured data by an offset time Δt both forward (i.e. later) and backward (i.e. earlier) in time, then the correlation coefficient between this offset model propagated time series and the in-situ measurements can be calculated and compared to the original.

Performing this $2n+1$ times for temporal offsets spanning the values $[-n, -n+\Delta t, ..., n-\Delta t, n]$ for a realistic maximum offset time of $n \approx 4$ days (Tao et al., 2005; Zieger & Hansen, 2008) yields the correlation coefficient as a function of constant temporal offsets, $r(\Delta t)$, with positive temporal offsets indicating that the un-offset model leads the data and negative offsets indicating that the un-offset model lags the data. Maximizing the correlation coefficient $r(\Delta t)$ thus gives a constant temporal offset which best aligns the propagation model with the measured time series; equivalently, this offset represents a systematic error in the arrival time of the original model. There are two drawbacks to this method of accounting for temporal offsets in the model: first, it can only account for a constant temporal offset Δt, rather than a distribution of uncertainties or a time-varying offset; second, this metric conflates the temporal alignment of the time series with the magnitudes of their predicted values, and thus does not necessarily characterize the model lag/lead time alone. Nonetheless, constant temporal offsetting is frequently used as a method to simply and quickly estimate model uncertainties, and as such remains available in MMESH.

3.2.2 Dynamic Time Warping

The performance of a solar wind propagation model can be decomposed into two components: the performance in modeling the arrival time and the performance in modeling the magnitude of the solar wind time series. These two are essentially represented by the abcissa and ordinate pairs of a propagated time series, respectively. Theoretically, differences between the propagation model and data time series should be decomposable by first optimizing the alignment of the model relative to the data to characterize the performance in arrival time, then secondly measuring the residuals between the aligned model and data time series to characterize the performance in magnitude. Aligning the model to the data in this way is often done by manually identifying patterns of shocks in both time series and calculating the difference in their observation times (e.g. Tao et al., 2005).

In practice, characterizing model performance in arrival time alone is not so straightforward, as the identification of patterns of shocks and shock-like structures in the solar wind data is often subjective. To more objectively define such structures, here we have "binarized" both the in-situ and propagation model time series data to identify extrema in both. The binarization process developed here involves taking the standard score (z-score) of the time derivative of a boxcar-smoothed time series and threshholding the result at a given significance level. This process has the end effect of identifying and isolating steep gradients in the time series of a given parameter, as would be expected in a shock, and is described in more detail in Appendix A and illustrated in A1. The binarization process was applied to the solar wind flow speed u_{mag} time series in both the model-propagated and in-situ data sets. The boxcar-smoothing-widths used for each time series and in each epoch were found dynamically and are listed in Table A1. Here a constant significance level of 3σ, measured across the full duration of each time series, has then been used for binarization.

Identifying shocks and shock-like structures in the now-binarized time series is trivial; aligning the patterns of structures found in the model and data time series is not, and remains subjective if performed manually. For reproducability, an here we employ an objective, automated method of aligning the two binarized time series based on the class of algorithms collectively known as dynamic time warping (DTW). Qualitatively, the aim of DTW is to locally shift, stretch, and compress one time series to better resemble another. DTW has only recently been applied to space weather modeling problems; the calculated net distance has been suggested as a useful, multi-scale metric for measuring the performance of solar wind models by Samara et al. (2022), and the resulting alignments have been used to create more accurate boundary conditions for solar wind propagation models by M. J. Owens and Nichols (2021). Within MMESH, the `dtw-python` package for the Python programming language developed by Giorgino (2009) is employed to warp the modeled time series to more closely resemble the in-situ data. The recommended usage, and that which will be followed in this discussion, is to use DTW to align the binarized model solar wind flow speed to the binarized measured flow speed, as the flow speed generally shows the clearest signatures of shocks and shock-like structures after binarization. Both the binarization and DTW methods within MMESH can, however, be applied independently to any of the propagated solar wind quantities (i.e. n, u_{mag}, p_{dyn}, or B_{IMF}) at the discretion of the user.

An overview of the two-series implementation of DTW is illustrated in Figure 4 and described here. This approach involves calculating the Euclidean distance between every permutation of the elements of each series, resulting in a two-dimensional matrix;

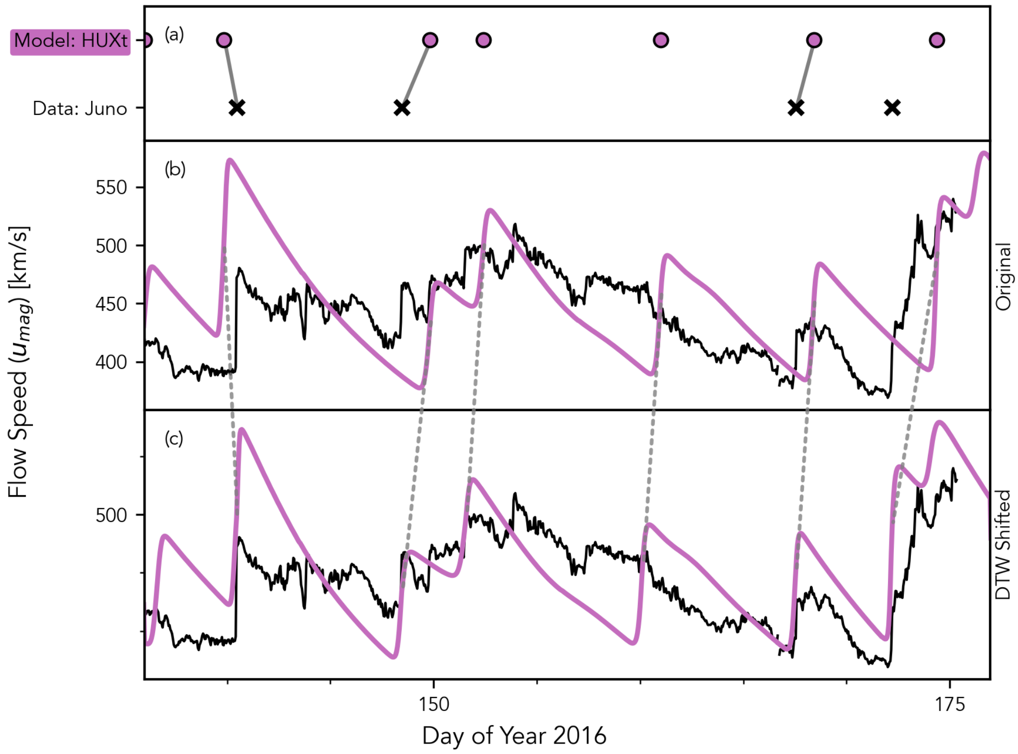

Figure 4. A composite diagram offering an overview of the dynamic time warping (DTW) process used to characterize model arrival time uncertainties. The (a) binarized model and data are shown as points representing the calculated extrema, with lines connecting model and data features which were identified to map to one another in the DTW process. The (b) original model and data time series are plotted to show the original alignment and may be compared to the (c) alignment of the warped model to the unchanged data, which demonstrates significantly reduced arrival time uncertainties. Dashed lines (b-c) connect the extrema identified in the original model to the same in the warped model; the horizontal component of these lines represents the offsets δt used to warp the model to best match the data. These δt are then taken as the distribution of arrival time uncertainties for the model.

a path, or alignment curve, through this matrix is then computed which minimizes the net distance, and this serves to effectively align the input modeled time series to best match the data by reindexing the former. DTW is here applied to the binarized time series data (Figure 4a) in order to eliminate the effect that each series' amplitude may have on the alignment calculation. From the aligned time series, tie points connecting the model time series to the data are then chosen from the alignment curve for each matching pair of model-data extrema (Figure 4a). The original model time series (Figure 4b) is then warped according to a linear interpolation of these tie points, which represents both the offsets of the matched extrema and the linear interpolations at each abcissa between these. The result is a warped time series which is better aligned with the spacecraft data (Figure 4c). While this process uses the binarized solar wind flow speed to compute the alignment, every parameter within a given model can then have the same warping applied to it. This allows for better alignment between all parameters, not just u_{mag}, by implicitly assuming that the input model parameters are aligned correctly with one another, and misaligned only relative to the measured data.

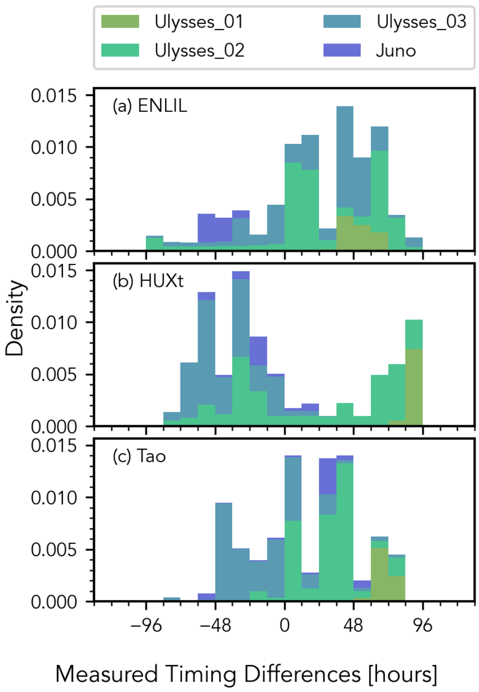

Figure 5. Histograms showing the distribution of total temporal shifts needed to best align each model with the in-situ data for the all spacecraft epochs, as found using dynamic time warping (DTW) as described in the text. The means of these distributions are equivalent to intrinsic shock arrival time errors, or timing biases, and the widths are representative of timing uncertainties.

For this demonstration of MMESH, DTW was used to align the binarized solar wind plasma flow speeds from each model to that of the data in each spacecraft epoch. Limits were placed on the DTW algorithm to ensure the resulting warped time series was physically meaningful: the maximum offsets allowed were ± 4 days (± 96 hours), chosen to be representative of the maximum temporal offsets measured in other studies (Tao et al., 2005; Zieger & Hansen, 2008). The first value of the modeled time series is forced to align to the first value of the measured time series by the DTW algorithm used here, as is the final value of the modeled time series to that of the measured. To account for this, the DTW process was applied to the same $2n + 1$ models with constant temporal offsets in the range $[-n, n]$ and with step size Δt as was previously discussed in Section 3.2.1. The optimal alignment within these $2n+1$ DTW results was found by maximizing the correlation coefficient of the warped model plasma flow speed u_{mag} to the data divided by the quasi-1σ half-width of the distribution of total temporal offsets r/σ_P (i.e., both constant and dynamic temporal offsets combined). The total distributions of temporal offsets in each model are illustrated in Figure 5 for reference. These distributions are not normally distributed, suggesting that the uncertainties in the modeled solar wind arrival times are not random, and are not centered at zero, indicating biases in the modeled arrival times.

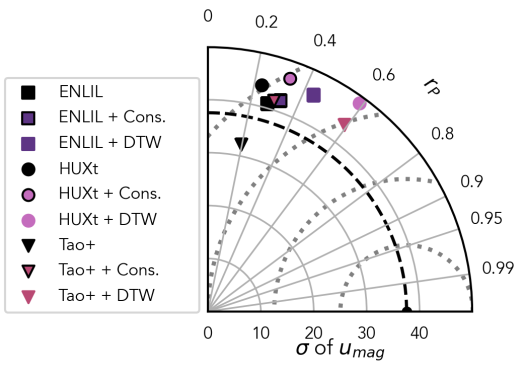

Figure 6. A Taylor Diagram showing the performance of each model, before and after temporal shifting, relative to the in-situ *Juno* solar wind data. The unshifted models (black symbols) all have correlation coefficient r in the range 0.2$\sim$0.3. Both constant time offsetting (outlined symbols) and DTW (full color symbols) improve the correlation coefficients of all models, but DTW improves the correlation coefficient more (r between 0.3$\sim$0.4 compared to r between 0.4$\sim$0.6, respectively). Employing time-varying temporal shifts is beneficial to matching the models to the data more closely.

3.3 Prediction of Time-Varying Model Timing Uncertainties

The cross-model consistent characterization of systematic timing biases and uncertainties already discussed allows the performance of the solar wind models to be quantitatively compared to one another. As the methods discussed in Section 3.2 rely on direct comparison to contemporaneous data, however, the timing biases and uncertainties cannot be empirically quantified in the absence of in-situ data– the main use case for solar wind modeling. To circumvent this, the distribution of timing uncertainties, as illustrated in Figure 5, could be considered invariant in time and propagated as such; this method of propagating timing uncertainties is supported by MMESH. As these timing uncertainties and biases are known to vary in time, as can be seen by the different space-

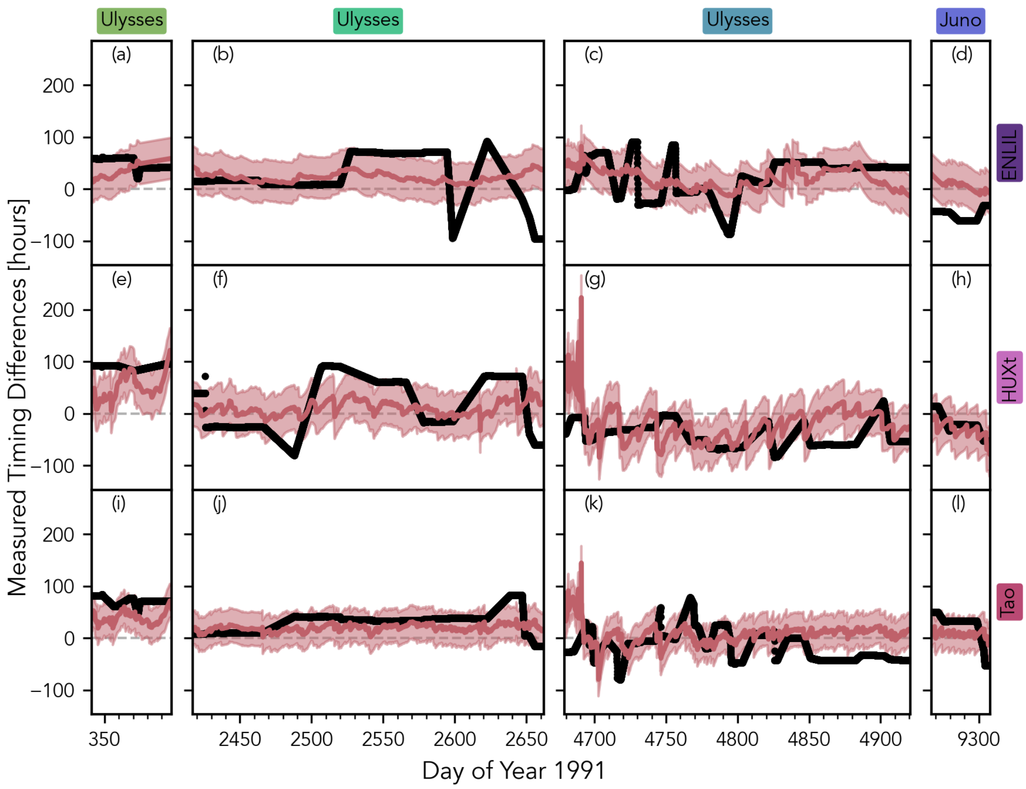

Figure 7. Plots of the measured temporal offsets (black lines) from DTW for each model-spacecraft-epoch set (e.g., a-d for ENLIL, e-h for HUXt, and i-l for Tao+), along with the multiple linear regression (MLR) fit to the temporal offsets found by fitting the offset time series with the parameters described in the text (red lines). While the independent parameters add significant variation in time, they nonetheless describe the emprirical timing uncertainties and systematics fairly well. The 1σ prediction uncertainities in the MLR fit (shaded red regions) are also plotted.

craft epochs covered in Figure 5, this method has the drawback of explicitly overestimating the uncertainties at any given time.

Alternatively, MMESH also supports a simple- or multiple-linear regression model description of the timing uncertainties. Multiple linear regression models are simple models which describe one continuous target variable as a linear combination of multiple continuous predictor variables; simple linear regression refers to the special case of a single predictor variable. The coefficients calculated for each predictor variable thus describe the contributions of each to the target variable. Similarly, the estimated standard deviation on these coefficients gives a sense of the relative importance of each predictor: relative to the coefficient value, a large standard deviation denotes a less significant predictor, with the opposite being true for a relatively small standard deviation. The linear regression method thus allows the propagation model timing uncertainties to be estimated even in the absence of in-situ data for comparison, provided the values of each predictor variable are known.

As the input propagation models do not propagate measurement error, the arrival time uncertainties characterized previously are present due to the limitations of these MHD-based models, each of which makes different simplifications of the physics describing the solar wind. These simplification give rise to correlations between the timing uncertainties in these models and other physical parameters describing the solar wind environment. Timing uncertainties in models with an inner boundary set by near-Earth measurements often trend with target-Sun-observer (TSO) angle in heliolongitude, in at least magnitude if not also in sign (Tao et al., 2005; Zieger & Hansen, 2008). Physically, this trend represents increasing uncertainty in the solar wind conditions as separation in heliolongitude (or Carrington longitude) increases away from the measurement point. While less commonly used, the offsets are expected to trend with the TSO angle in heliolatitude in a similar way, as the solar wind flow speed is known to be strongly ordered in heliolatitude during solar minimum (D. J. McComas et al., 2003; D. McComas et al., 2008). This well-ordered structuring with heliolatitude breaks down during solar maximum (D. J. McComas et al., 2003), which further suggests a physical connection between the offsets and the 11-year solar cycle. A final reasonable expectation is that the timing systematics and uncertainties are related to the models solar wind flow speed u_{mag}. This comes from the assumption that the propagation model is more likely to lag the data when underestimating the solar wind flow speed and more likely to lead when overestimating; if the underestimates tend to have lower magnitudes and overestimates tend have larger magnitudes, then a trend between modeled solar wind flow speed and temporal offset is expected.

These physical relationships between total, time-variable model offsets and descriptive parameters about the state of the solar wind can be leveraged to estimate the model offsets in the absence of simultaneous in-situ data. Here, multiple linear regression has been employed to use all of these physical parameters (i.e. TSO angle in heliolongitude and heliolatitude, solar cycle phase, and modeled u_{mag}) as predictors of the time-variable timing uncertainties and biases by fitting the predictors to the combined spacecraft epochs during which simultaneous in-situ measurements are available, as illustrated in Figure 7. Despite its simplicity, the multiple linear regression technique matches the known temporal offsets well. The combination of parameters used here accounts for 12% of the variation in the measured timings for the ENLIL model (i.e., $R^2 = 0.12$), 37% in the HUXt model, and 20% in the Tao+ model.

3.4 Multi-Model Ensemble

An MME is now created by the combination of the propagation models. MMESH supports the creation of MMEs from input propagation models alone, from propagation models with characterized timing uncertainties, whether through constant time offsetting or dynamic time warping, and from propagation models de-trended (i.e. warped) to account for timing biases with propagated uncertainties. Here, this final type of MME is created from the ENLIL, HUXt, and Tao+ solar wind propagation models warped according to the timing biases estimated by multiple linear regression to the multi-epoch in-situ dataset, and timing uncertainties propagated through.

For simplicity, an equal weights average of the each input model is taken. While there is some evidence that carefully-chosen weighting schemes may improve model performance (Guerra et al., 2020), more complicated weighting schemes may also decrease model performance compared to the equal weights, making equal weighting the more robust choice (Genre et al., 2013). Thus, the only improvement on the simple equal-weights averaging scheme we impose is to set the weight to 0 when a model does not yield an output at a given time step, whether due to the model's design (e.g. the lack of parameters other than solar wind flow speed in HUXt) or a lack of access to more recent models. The resulting MME of solar wind flow speed is shown in Figure 8a, superimposed on the in-situ measurements of the *Juno* spacecraft during the missions's cruise phase (cf. Figure 3).

3.4.1 Model Performance

The performance of this multi-model, multi-epoch ensemble is summarized in Figure 8, which shows that the ensemble has improved prediction efficiency of the solar wind flow speed u_{mag} during the *Juno* cruise epoch compared to any individual input model. All of the model time series, including that of the ensemble, show decreased standard deviations in Figure 8. This results from considering the distribution of timing uncertainties in calculating the mean values for each time series: when the distribution of timing uncertainties is measured or predicted to be large, the shifted fore-shocks in the solar wind appear more 'smoothed out'.

Despite this decreased standard deviation, the predicted flow speed u_{mag} of the MME ($r = 0.49$) outperforms ENLIL by 110% ($r = 0.23$), HUXt by 7% ($r = 0.46$), and Tao+ by 51% ($r = 0.32$) in correlation coefficient and achieves a centered root-mean-square difference (RMSD= 32.8) 28% lower than ENLIL (RMSD= 45.9), 14% lower than HUXt (RMSD= 38.1), and 9.1% lower than Tao+ (RMSD= 36.1). As HUXt does not contribute to parameters in the MME other than u_{mag}, and the performance of ENLIL beyond u_{mag} is poor here (i.e., ENLIL is evidently anticorrelated with the data in Figure 8f-h), the MME underperforms Tao+ in n_{tot}, p_{dyn}, and B_{IMF} by 12%−24% in correlation coefficient with $5\% - 8\%$ larger RMSD. These shortcomings of the MME are thus slight, and would likely be reduced further or eliminated in epochs where ENLIL performs more similarly to Tao+; alternatively, adding new solar wind propagation models to the MME discussed here would be expected to have a similar effect.

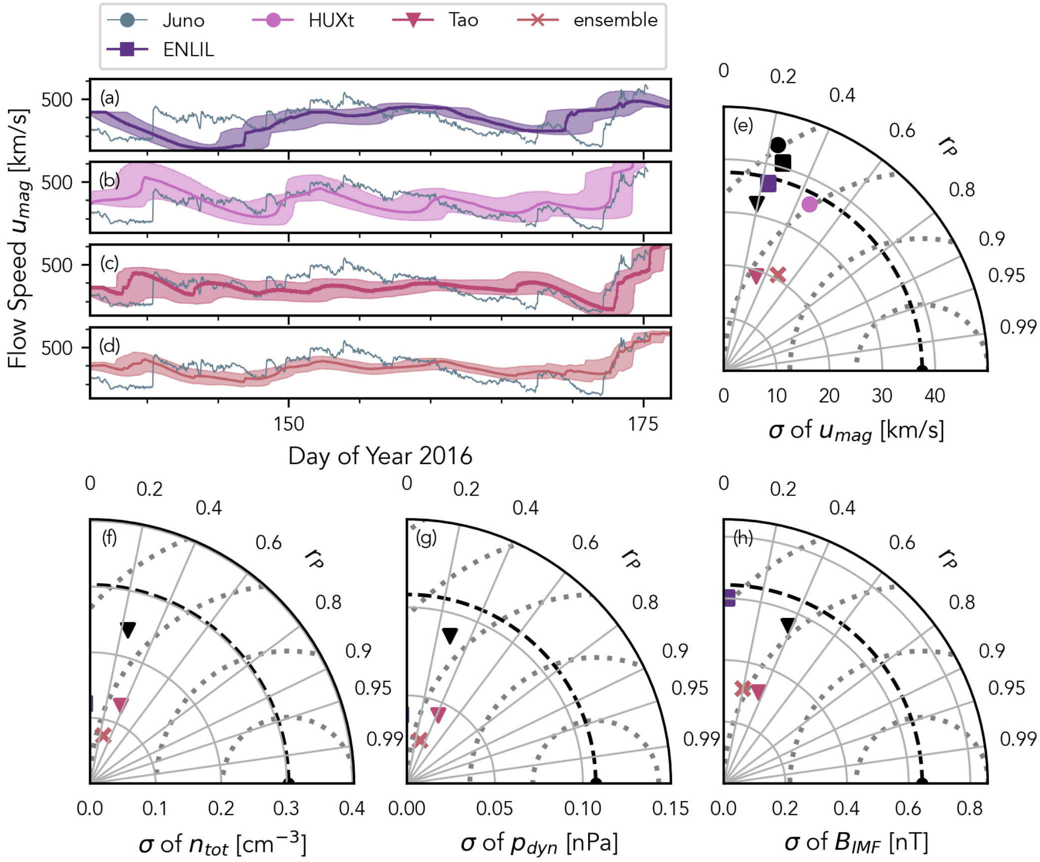

Figure 8. The solar wind flow speed u_{mag}, with timing uncertainties characterized by DTW and MLR applied over all spacecraft epochs, for (a) ENLIL, (b) HUXt, (c) Tao, and (d) the MME, compared to in-situ *Juno* data in each. The performance of the MME is summarized in the (e) Taylor diagram for u_{mag}, which illustrates that the MME outperforms all input models for this parameter; the Taylor diagram includes both the multi-epoch MLR-adjusted input models (colored symbols) and the original input models (black symbols) for comparison. Additional Taylor diagrams for (f) the total solar wind density n_{tot}, (g) the solar wind dynamic pressure p_{dyn}, and (h) the IMF magnitude B_{IMF} are included to show the performance of the MME in these parameters. As HUXt does not contribute to these parameters, the MME slightly underperforms Tao+.

4 *Juno*-epoch Solar Wind MME for Jupiter

Now that the multi-model epoch system has been fully described, all that remains is to generate MMESH-propagated solar wind for a useful epoch. Here we have chosen to run the ensemble for Jupiter contemporaneously with the *Juno* mission, beginning before the spacecraft entered the planet's magnetosphere (2016/05/15) and continuing seven years through mid-2023 (2023/05/15), in order to provide valuable context for the upstream conditions near Jupiter during *Juno*'s mission. A subset of the ensemble model results are shown in Figure 9, along with the results of the component models, spanning the first 6 months of coverage provided by this MME. The *Juno* in-situ measurements prior to entering Jupiter's magnetosphere are shown in Figure 9 for context, but as the MME here is for Jupiter's location, rather than that of *Juno*, the two timeseries are not expected to align as well as in Figure 9. The results of this specific MME are available at `https://zenodo.org/link-to-specific-results`; more generally, the results of this Jupiter MME along with any future updates to improve its predictive power or extend the temporally coverage will be available, and documented, at `https://zenodo.org/link-to-all-results`.

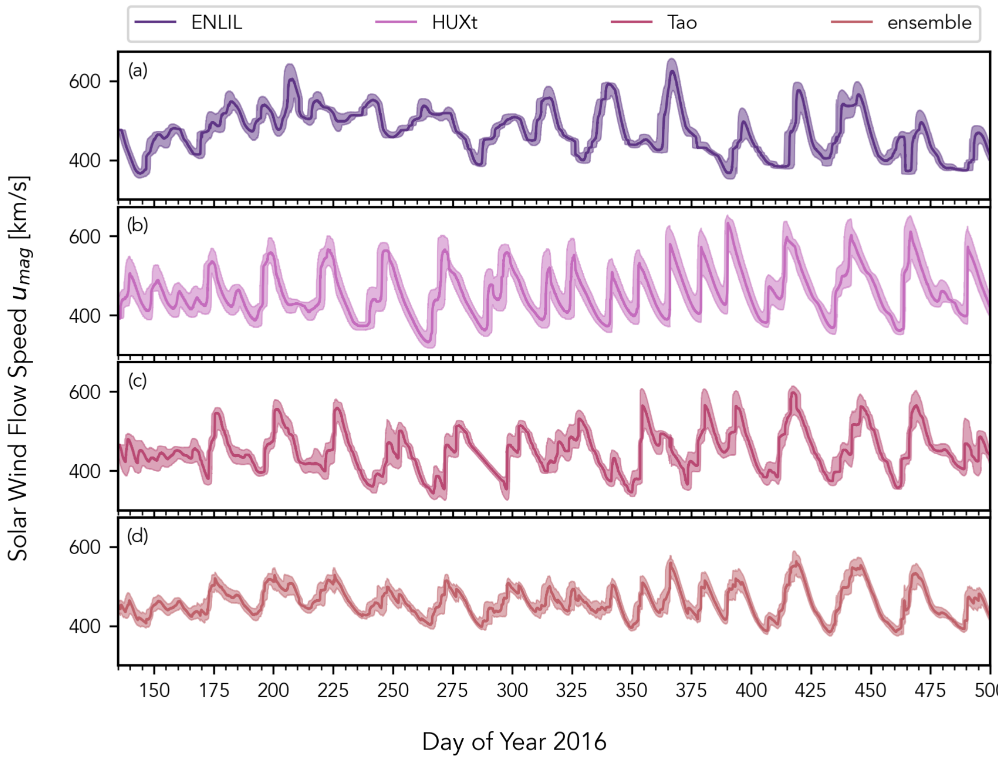

Figure 9. A 12-month subset of the *Juno*-era solar wind flow speed u_{mag} results, adjusted for timing biases measured using DTW and characterized using MLR, for the (a) ENLIL, (b) HUXt, (c) Tao+, and (d) MME, presented here starting during *Juno*'s approach to Jupiter in May 2016. The 1σ uncertainties in the solar wind flow speed u_{mag} are shown in each panel (shaded regions). Based on the results discussed here, the MME is expected to significantly outperform each of the component models in predicting the solar wind flow speed u_{mag}.

5 Summary and Conclusions

Here we have introduced MMESH, a Multi-Model Ensemble System for the Heliosphere, and described one use-case of this system to create a multi-model ensemble of the outer heliosphere solar wind near Jupiter through the first 7 years of *Juno* mission, spanning 2016/07/04 – 2023/07/04.

MMESH provides a framework with two central objectives: first, to allow easy characterization of solar wind propagation model performance; and second, to create multi-model ensembles of the solar wind. The first objective is crucial to statistically evaluating the strengths of the various solar wind propagation models available, as the original discussions of the performance of these models often quote different statistics or span non-overlapping epochs of the solar wind and thus cannot be compared one-to-one. Further, characterization of model performance yields an estimate of the model uncertainty, a quantity which is not provided internally by any model discussed here but which is essential for statistical analyses. With the second objective, we aim to create reliable composite models of the solar wind by combining physics-based solar wind propagation models with their estimated variances to be used in statistical analyses of solar-wind-magnetosphere interactions throughout the solar system. The strength of ensemble modeling lies in leveraging the different strengths of the constituent models, and so these two objectives are closely intertwined.

MMESH additionally includes a method to compare biases and variances in the model timing to physical parameters across disparate epochs prior to creating an ensemble. The objective of this multi-epoch method is to de-trend biases in the model timing which may arise from the various assumptions and simplifications made by each model. De-trending is performed here through multiple linear regression (MLR) of the measured model timing biases with a subset of the physically reasonable parameters with which model performance is expected to vary. The phase of the solar cycle, difference in heliolongitude and heliolatitude between the model target and the observer, and the modeled solar wind flow speed are all reasonable and considered here. As estimation of the model timing biases and variances is only possible when contemporaneous in-situ data are available for comparison, the spans over which the MLR de-trending can be performed are limited. The MLR de-trending is made more robust by considering multiple disparate epochs during which spacecraft data are available.

Using all of these methods, a multi-model ensemble of the solar wind conditions at Jupiter during the *Juno*-epoch has been created by combining three physics-based solar wind propagation models (ENLIL, HUXt, and Tao+); the version of this ensemble discussed here is available at `https://zenodo.org/link-to-specific-results` and the latest release of is available at `https://zenodo.org/link-to-all-results`. Biases and variances in each models timing were characterized for four epochs during which *Ulysses* or *Juno* data were available for comparison, spanning in total from 1991/12/08 – 2016/06/29. The model timing biases were then de-trended using MLR to the heliolatitude and modeled flow speed, which were determined to provide the best balance between describing the timing biases and overfitting. The biases in the three constituent solar wind models were corrected according to the MLR equation for the full MME span of 2016/07/04 – 2023/07/04 and combined. The resulting ensemble model outperforms all of the constituent models relative to the *Juno* cruise data immediately preceding this epoch; the ensemble has a correlation coefficient of 0.41 (78% increase over ENLIL, 32% increase over HUXt, and 86% increase over Tao+, after accounting for timing offsets in each). The improved upstream solar wind monitoring capabilities demonstrated by this MME are available to be downloaded and used immediately, and should prove crucial to ongoing and future in-situ studies of the Jovian magnetosphere using *Galileo*, *Juno*, JUICE, and *Europa Clipper*, as well as remote sensing studies using observatories such as JWST, HST, and Chandra.

Appendix A Time Series Binarization

Here, the measured and modeled magnitude of the solar wind flow speed u_{mag} is post-processed by first smoothing the series, then taking the standard score of its time derivative. Smoothing is accomplished by taking a rolling boxcar average of the flow speed u_{mag}. Smoothing in this way serves as a low-pass filter, allowing the recovery of the large-scale shape of the time series while ignoring small-scale fluctuations, which may dominate in in-situ spacecraft measurements. The time derivative of the flow speed time series $u_{mag}(t)$ is chosen in order to better identify the transition of a spacecraft or model trajectory through a slow-fast wind interface; these increases in solar wind flow speed occur over timescales less than 1hour and are more easily identifiable than changes in other solar wind parameters, which typically occur over longer timescales. The standard-score of the time series, or the time series normalized to its own standard deviation, allows for direct comparison of the relative changes between different time series which may have widely varying mean values.

Binarization requires subjective input of a boxcar-smoothing-width and significance level for each time series, however these parameters are partially degenerate with one another– a smaller smoothing window and a higher significance level will yield similar results to a larger window with lower significance level. To limit subjectivity, boxcar-smoothing-widths are found for each time series within a given epoch as the smallest width which, when applied to each time series before the derivative is taken, results in an equal series standard deviation to the smallest such standard deviation in the epoch. Qualitatively, this is the boxcar-smoothing-width required to make each time series look as "smooth" as the "smoothest" time series of the epoch. The boxcar-smoothing-widths used for each time series and in each epoch are listed in Table A1.

Table A1. Boxcar-smoothing widths for binarization, in hours

Source	Epoch			
	Ulysses 1	Ulysses 2	Ulysses 3	Juno
in-situ	5	15	2	7
ENLIL	8	1	4	5
HUXt	1	9	1	1
Tao+	6	11	2	10

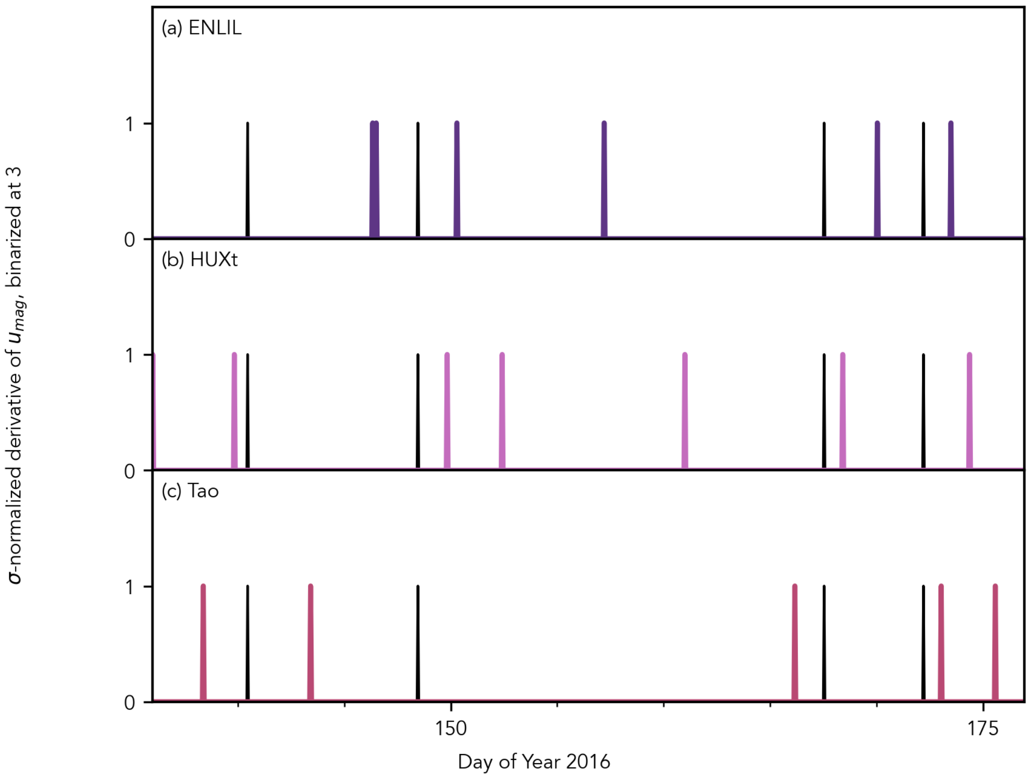

Figure A1. Binarized time series of the solar wind flow speed u_{mag} for the (a) ENLIL, (b) HUXt, and (c) Tao+ solar wind propagation models, with the binarized time series of the in-situ *Juno* data superimposed on each (black lines). The time-derivatives of all these series have been binarized at a value of 3σ, such that each 'spike' represents a change in the time-derivative of 3σ or larger.

Open Research

The results presented in this document rely on data collected by the Solar Radio Monitoring Program (`https://www.spaceweather.gc.ca/forecast-prevision/solar-solaire/solarflux/sx-en.php`) with additional processing by the NOAA National Centers for Environmental Information (`https://www.ncei.noaa.gov/`). These data were accessed via the LASP Interactive Solar Irradiance Datacenter (LISIRD) (`https://lasp.colorado.edu/lisird/`). Ephemeris information was obtained by use of the NASA Navigation and Ancillary Information Facility (NAID) SPICE toolkit.

Simulation results for the ENLIL solar wind propagation model (version 2.8f) have been provided by the Community Coordinated Modeling Center (CCMC) at Goddard Space Flight Center through their publicly available simulation services (`https://ccmc.gsfc.nasa.gov`). The ENLIL Model was developed by Dusan Odstrcil at George Mason University. Spacecraft data were acquired from the Goddard Space Flight Center Space Physics Data Facility (SPDF) COHOWeb service, except for the *Juno* in-situ data, which were instead acquired from Wilson et al. (2018) (plasma data) and the Automated Multi-Dataset Analysis web tool hosted at `https://amda.irap.omp.eu/`.

The MMESH code is available at `https://github.com/mjrutala/MMESH`, and includes the routines used to create the figures shown here. The *Juno*-epoch MME presented here is available at `https://zenodo.org/link-to-specific-results`, and future updates to this MME will be accessible from `https://zenodo.org/link-to-all-results`.

Acknowledgments

MJR's and CMJ's work at DIAS was supported by Science Foundation Ireland Grant 18/FRL/6199.

References

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: Space Physics*, *105*(A5), 10465–10479. doi: 10.1029/1999JA000262

Balogh, A., Beek, T. J., Forsyth, R. J., Hedgecock, P. C., Marquedant, R. J., Smith, E. J., . . . Tsurutani, B. T. (1992, January). The magnetic field investigation on the ULYSSES mission - Instrumentation and preliminary scientific results. , *92*(2), 221–236.

Bame, S. J., McComas, D. J., Barraclough, B. L., Phillips, J. L., Sofaly, K. J., Chavez, J. C., . . . Sakurai, R. K. (1992, January). The ULYSSES solar wind plasma experiment. , *92*(2), 237–265.

Barnard, L., & Owens, M. (2022, October). HUXt—An open source, computationally efficient reduced-physics solar wind model, written in Python. *Frontiers in Physics*, *10*, 1005621. doi: 10.3389/fphy.2022.1005621

Borovsky, J. E., & Denton, M. H. (2010). Solar wind turbulence and shear: A superposed-epoch analysis of corotating interaction regions at 1 AU. *Journal of Geophysical Research: Space Physics*, *115*(A10). doi: 10.1029/2009JA014966

Connerney, J. E. P., Benn, M., Bjarno, J. B., Denver, T., Espley, J., Jorgensen, J. L., . . . Smith, E. J. (2017, November). The juno magnetic field investigation. , *213*(1-4), 39–138. doi: 10.1007/s11214-017-0334-z

Cowley, S. W. H., Bunce, E. J., & Nichols, J. D. (2003). Origins of Jupiter's main oval auroral emissions. *Journal of Geophysical Research: Space Physics*, *108*(A4). doi: 10.1029/2002JA009329

Crooker, N. U., Gosling, J. T., Bothmer, V., Forsyth, R. J., Gazis, P. R., Hewish, A., … Wimmer-Schweingruber, R. F. (1999, July). CIR morphology, turbulence, discontinuities, and energetic particles. , *89*, 179–220. doi: 10.1023/A:1005253526438

Ebert, R. W., Bagenal, F., McComas, D., & Fowler, C. (2014, September). A survey of solar wind conditions at 5 AU: A tool for interpreting solar wind-magnetosphere interactions at Jupiter. *Frontiers in Astronomy and Space Sciences*, *1*, 4. doi: 10.3389/fspas.2014.00004

Ebert, R. W., McComas, D. J., Bagenal, F., & Elliott, H. A. (2010, December). Location, structure, and motion of Jupiter's dusk magnetospheric boundary from ~1625 to 2550 R_J. *Journal of Geophysical Research (Space Physics)*, *115*(A12223), A12223. doi: 10.1029/2010JA015938

Ebert, R. W., McComas, D. J., Elliott, H. A., Forsyth, R. J., & Gosling, J. T. (2009, January). Bulk properties of the slow and fast solar wind and interplanetary coronal mass ejections measured by Ulysses: Three polar orbits of observations. *Journal of Geophysical Research (Space Physics)*, *114*(A1), A01109. doi: 10.1029/2008JA013631

Elvidge, S., Granados, S. R., Angling, M. J., Brown, M. K., Themens, D. R., & Wood, A. G. (2023). Multi-model ensembles for upper atmosphere models. *Space weather : the international journal of research & applications*, *21*(3), e2022SW003356. doi: 10.1029/2022SW003356

Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, *29*(1), 108–121. doi: 10.1016/j.ijforecast.2012.06.004

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, *31*(7), 1–24. doi: 10.18637/jss.v031.i07

Gosling, J., & Pizzo, V. (1999, July). Formation and Evolution of Corotating Interaction Regions and their Three Dimensional Structure. *Space Science Reviews*, *89*(1), 21–52. doi: 10.1023/A:1005291711900

Gressl, C., Veronig, A. M., Temmer, M., Odstrčil, D., Linker, J. A., Mikić, Z., & Riley, P. (2014, May). Comparative study of MHD modeling of the background solar wind. , *289*(5), 1783–1801. doi: 10.1007/s11207-013-0421-6

Guerra, J. A., Murray, S. A., Bloomfield, D. S., & Gallagher, P. T. (2020, July). Ensemble forecasting of major solar flares: Methods for combining models. *Journal of Space Weather and Space Climate*, *10*(38), 38. doi: 10.1051/swsc/2020042

Hagedorn, R., Doblas-Reyes, F. J., & Palmer, T. N. (2005, January). The rationale behind the success of multi-model ensembles in seasonal forecasting — I. Basic concept. *Tellus Series A*, *57*(3), 219–233. doi: 10.3402/tellusa.v57i3.14657

Hanlon, P. G., Dougherty, M. K., Forsyth, R. J., Owens, M. J., Hansen, K. C., Tóth, G., … Young, D. T. (2004, September). On the evolution of the solar wind between 1 and 5 AU at the time of the Cassini Jupiter flyby: Multi-spacecraft observations of interplanetary coronal mass ejections including the formation of a merged interaction region. *Journal of Geophysical Research (Space Physics)*, *109*(A09S03), A09S03. doi: 10.1029/2003JA010112

Kasper, J. C. (2002). *Solar Wind Plasma: Kinetic Properties and Micro-Instabilities* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA..

Keebler, T. B., Tóth, G., Zieger, B., & Opher, M. (2022, June). MSWIM2D: Two-dimensional outer heliosphere solar wind modeling. , *260*(2), 43. doi: 10.3847/1538-4365/ac67eb

King, J. H., & Papitashvili, N. E. (2005, February). Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data. *Journal of Geophysical Research (Space Physics)*, *110*(A02104), A02104. doi: 10.1029/2004JA010649

Kita, H., Kimura, T., Tao, C., Tsuchiya, F., Murakami, G., Yamazaki, A., … Fujimoto, M. (2019, December). Jovian UV Aurora's Response to the Solar Wind: Hisaki EXCEED and Juno Observations. *Journal of Geophysical Research: Space Physics*, *124* (12), 10,209–10,218. doi: 10.1029/2019JA026997

Lepping, R. P., Acũna, M. H., Burlaga, L. F., Farrell, W. M., Slavin, J. A., Schatten, K. H., … Worley, E. M. (1995, February). The wind magnetic field investigation. , *71* (1-4), 207–229. doi: 10.1007/BF00751330

McComas, D., Allegrini, F., Bagenal, F., Casey, P., Delamere, P., Demkee, D., … Weidner, S. (2008, October). The solar wind around pluto (SWAP) instrument aboard new horizons. , *140* (1-4), 261–313. doi: 10.1007/s11214-007-9205-3

McComas, D. J., Alexander, N., Allegrini, F., Bagenal, F., Beebe, C., Clark, G., … White, D. (2017, November). The Jovian Auroral Distributions Experiment (JADE) on the Juno Mission to Jupiter. *Space Science Reviews*, *213* (1-4), 547–643. doi: 10.1007/s11214-013-9990-9

McComas, D. J., Angold, N., Elliott, H. A., Livadiotis, G., Schwadron, N. A., Skoug, R. M., & Smith, C. W. (2013, December). Weakest solar wind of the space age and the current "Mini" solar maximum. , *779* (2), 2. doi: 10.1088/0004-637X/779/1/2

McComas, D. J., Bame, S. J., Barraclough, B. L., Feldman, W. C., Funsten, H. O., Gosling, J. T., … Neugebauer, M. (1998, January). Ulysses' return to the slow solar wind. , *25* (1), 1–4. doi: 10.1029/97GL03444

McComas, D. J., Elliott, H. A., Schwadron, N. A., Gosling, J. T., Skoug, R. M., & Goldstein, B. E. (2003, May). The three-dimensional solar wind around solar maximum. , *30* (10), 1517. doi: 10.1029/2003GL017136

McComas, D. J., Gosling, J. T., & Skoug, R. M. (2000, August). Ulysses observations of the irregularly structured mid-latitude solar wind during the approach to solar maximum. , *27* (16), 2437–2440. doi: 10.1029/2000GL000052

Murray, S. A. (2018). The importance of ensemble techniques for operational space weather forecasting. *Space weather : the international journal of research & applications*, *16* (7), 777–783. doi: 10.1029/2018SW001861

Nichols, J. D., Kamran, A., & Milan, S. E. (2019). Machine learning analysis of jupiter's far-ultraviolet auroral morphology. *Journal of Geophysical Research: Space Physics*, *124* (11), 8884–8892. doi: 10.1029/2019JA027120

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

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

Owens, M. J. (2018, August). Solar Wind and Heavy Ion Properties of Interplanetary Coronal Mass Ejections. *Solar Physics*, *293* (8), 122. doi: 10.1007/s11207-018-1343-0

Owens, M. J., & Nichols, J. D. (2021, December). Using in situ solar-wind observations to generate inner-boundary conditions to outer-heliosphere simulations - I. Dynamic time warping applied to synthetic observations. , *508* (2), 2575–2582. doi: 10.1093/mnras/stab2512

Palmerio, E., Nieves-Chinchilla, T., Kilpua, E. K. J., Barnes, D., Zhukov, A. N., Jian, L. K., … Sánchez-Cano, B. (2021). Magnetic structure and propagation of two interacting CMEs from the sun to saturn. *Journal of Geophysical Research: Space Physics*, *126* (11), e2021JA029770. doi: 10.1029/2021JA029770

Papitashvili, N. E., & King, J. H. (2020). *OMNI Hourly Data.* NASA Space Physics Data Facility. doi: 10.48322/1shr-ht18

Richardson, I. G. (2018, January). Solar wind stream interaction regions throughout the heliosphere. *Living Reviews in Solar Physics*, *15* (1), 1. doi: 10.1007/s41116-017-0011-z

Richardson, J. D., Paularena, K. I., Lazarus, A. J., & Belcher, J. W. (1995, February). Radial evolution of the solar wind from IMP 8 to Voyager 2. , *22*(4), 325–328. doi: 10.1029/94GL03273

Riley, P., Mays, M. L., Andries, J., Amerstorfer, T., Biesecker, D., Delouille, V., … Zhao, X. (2018). Forecasting the arrival time of coronal mass ejections: Analysis of the CCMC CME scoreboard. *Space weather : the international journal of research & applications*, *16*(9), 1245–1260. doi: 10.1029/2018SW001962

Samara, E., Laperre, B., Kieokaew, R., Temmer, M., Verbeke, C., Rodriguez, L., … Poedts, S. (2022, March). Dynamic time warping as a means of assessing solar wind time series. , *927*(187), 187. doi: 10.3847/1538-4357/ac4af6

Sanchez-Diaz, E., Rouillard, A. P., Lavraud, B., Segura, K., Tao, C., Pinto, R., … Plotnikov, I. (2016). The very slow solar wind: Properties, origin and variability. *Journal of Geophysical Research: Space Physics*, *121*(4), 2830–2841. doi: 10.1002/2016JA022433

Slavin, J. A., Smith, E. J., Spreiter, J. R., & Stahara, S. S. (1985, July). Solar wind flow about the outer planets: Gas dynamic modeling of the jupiter and saturn bow shocks. , *90*(A7), 6275–6286. doi: 10.1029/JA090iA07p06275

Smith, C. W., L'Heureux, J., Ness, N. F., Acuña, M. H., Burlaga, L. F., & Scheifele, J. (1998, July). The ACE magnetic fields experiment. , *86*, 613–632. doi: 10.1023/A:1005092216668

Southwood, D. J., & Kivelson, M. G. (2001). A new perspective concerning the influence of the solar wind on the Jovian magnetosphere. *Journal of Geophysical Research: Space Physics*, *106*(A4), 6123–6130. doi: 10.1029/2000JA000236

Tao, C., Kataoka, R., Fukunishi, H., Takahashi, Y., & Yokoyama, T. (2005, November). Magnetic field variations in the Jovian magnetotail induced by solar wind dynamic pressure enhancements. *Journal of Geophysical Research (Space Physics)*, *110*(A11), A11208. doi: 10.1029/2004JA010959

Taylor, K. E. (2001, April). Summarizing multiple aspects of model performance in a single diagram. , *106*(D7), 7183–7192. doi: 10.1029/2000JD900719

Tsurutani, B. T., Gonzalez, W. D., Gonzalez, A. L. C., Guarnieri, F. L., Gopalswamy, N., Grande, M., … Vasyliunas, V. (2006). Corotating solar wind streams and recurrent geomagnetic activity: A review. *Journal of Geophysical Research: Space Physics*, *111*(A7). doi: 10.1029/2005JA011273

Vogt, M. F., Gyalay, S., Kronberg, E. A., Bunce, E. J., Kurth, W. S., Zieger, B., & Tao, C. (2019, December). Solar wind interaction with jupiter's magnetosphere: A statistical study of galileo in situ data and modeled upstream solar wind conditions. *Journal of Geophysical Research (Space Physics)*, *124*(12), 10,170–10,199. doi: 10.1029/2019JA026950

Wilson, R. J., Bagenal, F., Valek, P. W., McComas, D. J., Allegrini, F., Ebert, R. W., … Thomsen, M. F. (2018, April). Solar Wind Properties During Juno's Approach to Jupiter: Data Analysis and Resulting Plasma Properties Utilizing a 1-D Forward Model. *Journal of Geophysical Research: Space Physics*, *123*(4), 2772–2786. doi: 10.1002/2017JA024860

Xie, H., Ofman, L., & Lawrence, G. (2004, March). Cone model for halo CMEs: Application to space weather forecasting. *Journal of Geophysical Research (Space Physics)*, *109*(A03109), A03109. doi: 10.1029/2003JA010226

Zhao, X. P., Plunkett, S. P., & Liu, W. (2002). Determination of geometrical and kinematical properties of halo coronal mass ejections using the cone model. *Journal of Geophysical Research: Space Physics*, *107*(A8), SSH 13-1-SSH 13-9. doi: 10.1029/2001JA009143

Zieger, B., & Hansen, K. C. (2008, August). Statistical validation of a solar wind propagation model from 1 to 10 AU. *Journal of Geophysical Research (Space Physics)*, *113*(A8), A08107. doi: 10.1029/2008JA013046

Zurbuchen, T. H., & Richardson, I. G. (2006, March). In-Situ Solar Wind and Magnetic Field Signatures of Interplanetary Coronal Mass Ejections. *Space Science*

Reviews, *123*(1), 31–43. doi: 10.1007/s11214-006-9010-4