

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.



RESEARCH ARTICLE

10.1029/2023SW003554

Key Points:

- An algorithm for connecting a location in the solar system to its source location using Air Force Data Assimilative Photospheric Flux Transport-Wang Sheeley Arge (ADAPT-WSA) is presented
- Multiple variations of ADAPT-WSA are collected and summarized into probability distributions using kernel density estimation
- Usage of source location modeling to inform scientific research and spacecraft mission operations are reviewed and discussed

Correspondence to:

D. E. da Silva,
Daniel.e.daSilva@nasa.gov

Citation:

da Silva, D. E., Wallace, S., Arge, C. N., & Jones, S. (2023). Ensemble forecasts of solar wind connectivity to 1 R_s using ADAPT-WSA. *Space Weather*, 21, e2023SW003554. <https://doi.org/10.1029/2023SW003554>



Received 11 MAY 2023

Accepted 15 SEP 2023

Author Contributions:

Conceptualization: D. E. da Silva, C. N. Arge
Methodology: D. E. da Silva, C. N. Arge, S. Jones
Software: D. E. da Silva
Supervision: C. N. Arge
Validation: D. E. da Silva, S. Jones
Visualization: D. E. da Silva
Writing – original draft: D. E. da Silva, S. Wallace
Writing – review & editing: D. E. da Silva, S. Wallace, C. N. Arge, S. Jones

Ensemble Forecasts of Solar Wind Connectivity to 1 R_s Using ADAPT-WSA

D. E. da Silva^{1,2,3} , S. Wallace¹ , C. N. Arge¹, and S. Jones^{1,4}

¹Solar Physics Laboratory, NASA Goddard Spaceflight Center, Greenbelt, MD, USA, ²Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO, USA, ³Goddard Planetary Heliophysics Institute, University of Maryland, Baltimore County, Baltimore, MD, USA, ⁴Catholic University of America, Washington, DC, USA

Abstract The solar wind which arrives at any location in the solar system is, in principle, relatable to the outflow of solar plasma from a single source location. This source location, itself usually being part of a larger coronal hole, is traceable to 1 R_s along the Sun's magnetic field, in which the entire path from 1 R_s to a location in the heliosphere is referred to as the solar wind connectivity. While not directly measurable, the connectivity between the near-Earth solar wind is of particular importance to space weather. The solar wind solar source region can be obtained by leveraging near-sun magnetic field models and a model of the interplanetary solar wind. In this article, we present a method for making an ensemble forecast of the connectivity presented as a probability distribution obtained from a weighted collection of individual forecasts from the combined Air Force Data Assimilative Photospheric Flux Transport-Wang Sheeley Arge (ADAPT-WSA) model. The ADAPT model derives the photospheric magnetic field from synchronic magnetogram data, using flux transport physics and ongoing data assimilation processes. The WSA model uses a coupled set of potential field type models to derive the coronal magnetic field, and an empirical relationship to derive the terminal solar wind speed observed at Earth. Our method produces an arbitrary 2D probability distribution capable of reflecting complex source configurations with minimal assumptions about the distribution structure, prepared in a computationally efficient manner.

Plain Language Summary Solar wind which arrives at Earth mostly comes from one place on the sun at a time. There is no measurement of where that place is, but scientific modeling can give us an estimate. In this article, we present one way of making that kind of estimate. Our way gives a probability distribution of where the solar wind is coming from, instead a single estimate. The primary engine of our method is the Air Force Data Assimilative Photospheric Flux Transport-Wang Sheeley Arge model. Applications of knowing where the solar wind is coming from on the Sun include helping other studies and deciding where to point solar imagers.

1. Introduction

The solar corona is comprised of complex time-dependent magnetic fields frozen into a highly ionized plasma. The solar magnetic field in this region can instantaneously be categorized into closed field lines, extending no further than a critical point and open field lines, extending far beyond the critical point out to the heliopause. Nevertheless, on long timescales the corona is well approximated by potential fields (PFs) in which field lines are either open or closed (Riley et al., 2006).

Concentrated regions of open and unipolar magnetic field at the Sun's surface correlate with dark regions observed on the Sun at extreme ultraviolet (EUV) wavelengths, defined as coronal holes. Coronal holes are the primary sources of coronal plasma outflow that forms the solar wind. Models can be used to trace the solar wind from any point in interplanetary space to the corresponding source regions back at the Sun. Except for limited scenarios involving stream interactions and inter-stream mixing, an observation of solar wind using satellite instrumentation can, in principle, be readily linked to a single source location. The solar wind connectivity is defined for the purposes of this work as the path a parcel of plasma takes from the source location at 1 R_s to a specified traveling satellite. Examples of satellites include Earth and planetary bodies, as well as spacecraft with arbitrary orbits. The properties of the solar wind observed in situ, such as number density (n_i), proton T_i and electron T_e temperature, radial solar wind speed (v_{sw}), and compositional data from mass spectroscopy, are correlated to the dynamics and physical processes acting on the plasma along the

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

path of connectivity, from the source location to the location in the corona where the solar wind parcel fully disconnects from the Sun and flows radially outward (e.g., Borovsky, 2016; Zhao et al., 2017; Zurbuchen et al., 2002).

Space weather research and spaceflight operations are aided by modeling efforts which connect the solar wind observed in situ by spacecraft or at an interplanetary body to its specific source region in the low corona. This modeling task is the focus of this article. For example, studies utilizing multi-point spacecraft observations of the solar wind at the scale of the solar system are aided by connecting each observation to a coronal hole, potentially generating a global solar outflow perspective (Dresing et al., 2012; Gómez-Herrero et al., 2011; Zhu et al., 2018). Other research studies analyzing compositional changes in the solar wind are aided by relating to multi-spectral imagery of the source region. In such a case, the multi-spectral imagery may provide supportive information regarding per-species abundances and photon-releasing chemical reactions at the source location. In another example, spaceflight missions operating narrow-view solar imagers for solar wind studies must decide where to point their instruments, and often the most useful place to point is the coronal hole linked to an observing spacecraft or body such as the Earth (e.g., for space weather applications) (Rouillard et al., 2020).

In this article, we address the modeling task using the combined Air Force Data Assimilative Photospheric Flux Transport-Wang Sheeley Arge (ADAPT-WSA) model. The output of ADAPT, which are global photospheric magnetic field maps, is used as an input to WSA. WSA is a semi-empirical model which predicts the speed and interplanetary magnetic field polarity of the solar wind speed (Arge, Harvey, et al., 2003; Arge, Odstreil, et al., 2003; Arge & Pizzo, 2000; McGregor et al., 2008; Sheeley, 2017). ADAPT (Arge et al., 2010, 2011, 2013) is a model which generates estimates of the global photospheric maps at any specified point in time by evolving the global photospheric magnetic field using flux transport modeling and performing on-going data assimilation of near-side magnetogram measurements (Hickmann et al., 2015; Schrijver & DeRosa, 2003; Worden & Harvey, 2000). The ADAPT model can assimilate this class of measurements, in this work originating from the Global Oscillation Network Group instrument network (Harvey et al., 1996; Plowman & Berger, 2020). ADAPT itself is an ensemble model and produces a range of global photospheric magnetic field maps (or realizations) that represent different samplings of convection cell parameterizations.

In this work, we use a varied set of the velocity and magnetic field output from ADAPT-WSA to create an ensemble prediction of the source location of the in situ observed solar wind. The ensemble dimension is discussed in the methodology section. The result is a probability map of where the source location is most likely to be for a given plasma parcel in the heliosphere. This probability map approach is more powerful than a single point estimate as it inherently highlights uncertainty and consensus in the model variations. The probability map used is flexible in its shape and structure, allowing it to take on non-Gaussian forms including bimodal and tailed shapes. For instance, the shape of the probability distribution relative to the edge of coronal hole can speak to the probability distribution of solar wind characteristics, such as v_{sw} (the solar wind speed), in a way a single point estimate cannot.

Previous work on source location estimation has incorporated ensemble methodology to respond to the issue of uncertainty. In the work of Badman et al. (2023), the prediction of solar wind sources for the Parker Solar Probe at $13.3 R_s$ was studied. Their methodology uses an ensemble set built from a diverse array of models, model parameters, as well physical boundary conditions (Badman et al., 2023). In the work of Koukras et al. (2022), the problem is studied for fast solar wind in particular, where ballistic backmapping is used with potential field source surface (PFSS) models, and an ensemble set is built from varied solar wind speeds, random noise added to the magnetogram, and varied model boundary locations. The work of Rouillard et al. (2020) studies a related problem of magnetic connectivity in the context of Solar Orbiter, and uses an ensemble set which perturbs an intermediate position in the backmapping algorithm.

The methodology is split into several sections. In Section 2, we describe the methods used on a single set of WSA velocity/magnetic fields to connect a position in space to a source location at $1 R_s$. In Section 3, we analyze the theory of the uncertainties involved, how that informs the construction of a representative ensemble set, and how the connectivity algorithm applied to each member of the ensemble set is aggregated into a smooth probability distribution. Section 4 presents the results of studies designed to gauge the effectiveness of the model through analysis of secondary characteristics. Finally, Section 5 summarizes the method presented and reviews applications.

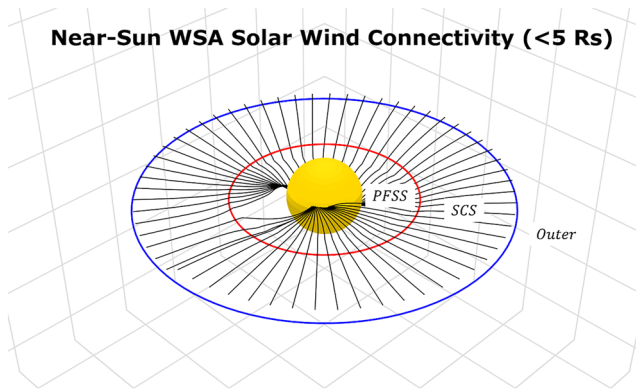


Figure 1. Connections between the subsatellite points for Earth at the outer boundary (blue) to the surface of the sun, showing that subsatellite points may be connected to one of several coronal holes. These connections were obtained by tracing the magnetic fields generated by Wang Sheeley Arge (WSA) starting at each subsatellite point at the outer boundary ($5 R_s$). The red line represents the interface boundary ($2.5 R_s$) between the Schatten Current Sheet (SCS) sub-domain and the potential field source surface (PFSS) sub-domain within WSA.

2. Connectivity Algorithms

In this section, we describe the algorithm for determining connectivity between a location in the solar system and $1 R_s$. The methodology of this individual section is identical to that which exists in WSA v5.2, as deployed at the Community Coordinate Modeling Center.

The connectivity computed is driven by a single global coronal magnetic field solution over the WSA domain ($1 R_s \leq r < 5.0 R_s$). This method is repeated for each of the magnetic field solutions in the varied set generated by the different ADAPT map realizations. This algorithm associates a traveling point in the solar system, discussed here as the satellite, with a source location in heliographic latitude and longitude at $1 R_s$.

WSA is comprised of two PF type models. The inner model is a traditional PFSS model (Altschuler & Newkirk, 1969; Schatten et al., 1969; Y.-M. Wang & Sheeley, 1992) with inner boundary at the $1 R_s$, the photosphere, and source surface at $2.51 R_s$. Every magnetic field line that reaches the source surface is open by construction, and this boundary effectively serves as the critical point. The Schatten current sheet (SCS; Schatten, 1971) model is used to specify the magnetic field of the outer coronal region and has an inner boundary at $2.49 R_s$ and outer boundary is located at infinity, however the model solution is normally used only out to $5 R_s$, though different outer

boundary radii can be used. A small overlap region between the PFSS and SCS is used in the model coupling. The PFSS magnetic field values at $2.49 R_s$ are used as the inner boundary to the SCS. Because the WSA outer boundary is outside the source surface, and we make the assumption that outside the outer boundary the solar velocity can be approximated by $\vec{v}_{sw} = v_r \hat{r}$ while inside the model domain the solar wind must be considered in theory as a three-component vector $\vec{v}_{sw} = v_r \hat{r} + v_\theta \hat{\theta} + v_\phi \hat{\phi}$ (Schatten et al., 1969). The three-component vector v_{sw} within the WSA domain is supplied directly by the magnetic field solution. The plasma beta in the corona is low (magnetically dominated) and therefore the plasma flows along magnetic field lines without strongly perturbing them. Within the WSA coronal domain, there are therefore two sub-domains separate by an interface radius ($2.49 R_s$), known as the PFSS and SCS domains. For more information on these sub-domains, the reader is pointed toward Mayer (2012) and McGregor et al. (2005, 2008).

The approach we will use is to break the path determining the magnetic field connecting the satellite to the source location into three sub-paths: from the satellite to the model outer boundary (P_{outer}), from the outer boundary to the interface radius (P_{SCS}), and from the interface radius to the solar surface (P_{PFSS}). In this decomposition, each starts where the previous one ends.

Decomposing the Full Connectivity Path into Three Sub-paths

$$P_{connectivity} = P_{outer} \cup P_{SCS} \cup P_{PFSS} \quad (1)$$

In Figure 1 an illustration of P_{SCS} and P_{PFSS} is given in the WSA domain. In this plot, the blue line is the model outer boundary, and the red line the interface radius. In this plot, field lines are generated from the satellite position projected onto the outer boundary (henceforth known as “subsatellite points”) over one Carrington rotation. This plot shows how all paths end up in one of several coronal holes, and the field becomes increasingly more radial in the SCS sub-domain until the radial assumption is fully satisfied at the outer boundary. We note that the red and blue lines are drawn on the ecliptic, where-as the subsatellite points are slightly off the ecliptic, which may make it appear that the connection do not continue out to the outer boundary when they in fact do.

To solve for P_{outer} , we start by utilizing ephemeris data of the satellite. The ephemeris data holding the satellite's position is projected down to the outer boundary of the WSA model (i.e., the subsatellite points are identified). This is done up to a Carrington rotation into the future for the satellite's position at the grid resolution of the model (in this case 2°). This produces a ring on the outer boundary of subsatellite points, and 2nd and 3rd versions of this ring are created by perturbing the latitude by half the grid resolution (in this case $\pm 1^\circ$).

WSA uses an empirical relationship to estimate the solar wind speed at the model outer boundary. The current variation of the equation is given by,

WSA Empirical Velocity Relationship

$$v_{sw}(f_s, \theta_b) = 285 + \frac{625}{(1 + f_s)^{0.22222}} \left\{ 1.0 - 0.8e^{-\left(\frac{\theta_b}{2}\right)^2} \right\}^3 \text{ km/s} \quad (2)$$

where f_s is the magnetic field expansion factor and θ_b is the minimum angular distance that an open field footprint from the nearest coronal hole boundary (Arge et al., 2004; Riley et al., 2015; Wallace et al., 2020).

Radially outflowing point-parcels starting at the outer boundary are then simulated as originating from these three rings using velocity $\vec{v}_{sw}^i = v_r^i \hat{r}$ where v_r^i is obtained from the WSA empirical velocity relationship. These radially outflowing point-parcels are then propagated out to the radius of the satellite position. The propagation is ballistic except for a simple ad hoc methodology to prevent fast particles from overtaking slow ones, which results in stream interaction regions reminiscent of the real solar wind (Arge & Pizzo, 2000). Stream rarefactions are also produced. This simple point-parcel simulation is extremely fast due to its simplicity and provides a good-enough solution while avoiding overhead (from both operational and computational perspectives) from coupling with an magnetohydrodynamic (MHD) model such as Enlil.

The output of the point-parcel simulation is P_{outer} , providing the connectivity of the satellite with the set of subsatellite points at the outer boundary, which are deemed as connected to the satellite at an associated time in the future. The next steps are to find P_{SCS} and P_{PFSS} . The sub-path for P_{SCS} is found by tracing the magnetic field line inward from the outer boundary starting at each subsatellite point to the interface radius (across the SCS sub-domain), and then the sub-path for P_{PFSS} is found by tracing the field line inward from the interface radius to the model inner boundary, which is the solar surface. Because the solar wind velocity and the magnetic field are aligned (parallel or anti-parallel) in this region by the plasma β , tracing the magnetic field inward is sufficient for determining the solar wind trajectory. The field line tracing methodology used is an implementation of Runge-Kutta 45, where the field at each step is determined by calculating the magnetic field, using spherical harmonics, at the four grid cell centers (at fixed R) bounding the end point position of the last step inward and then interpolating these values to the required position in latitude and longitude. To perform the field line trace, an iterative algorithm steps along the field line in the direction of $a\hat{B}$, solving the equation $\frac{d\mathbf{r}}{dt} = a\hat{B}$ where $\mathbf{r}(t)$ is the position along the trace and a is either +1 or −1. The value of a is chosen before iteration begins and is selected based on the polarity of the field at the outer boundary, specifically to force the trace move toward the sun even if \hat{B} is pointed outwards.

Supplementary visualizations for understanding of the point-parcel simulation for P_{outer} and the magnetic field tracing for P_{SCS} and P_{PFSS} are found in Figure 2. The top panel shows links between points on the first subsatellite ring and the end points on the solar surface residing in coronal holes (colored by their asymptotic final speed per the model). The many points along the ring correspond to very different solar speeds (from about 350 to 700 km/s). This justifies the need for the point-parcel simulation, since the transit time to, for example, Earth, accordingly, varies by a factor of 2× and the corresponding outer boundary location is not obvious.

The bottom panel of Figure 2 shows a map of the solar wind velocities over the outer boundary sphere determined by using the WSA empirical relationship. This demonstrates the complex longitude-dependence of v_{sw} on the outer boundary, which is most variable near the equator and least variable at high latitudes (Riley et al., 2001, 2015). At this stage, the major uncertainties lie in the ability to estimate the solar wind speed at the model outer boundary (performed through the WSA empirical relationship), and the ability to model parcel's transit between the model outer boundary and the satellite position (performed through the point-parcel simulation).

3. Ensemble Modeling

The previous section describes how the connectivity endpoints are determined for individual magnetic field solutions. In this section, we discuss how we turn a collection of forecasts of connectivity endpoints into a probabilistic ensemble forecast for the connectivity endpoint as a random variable, represented as a probability density

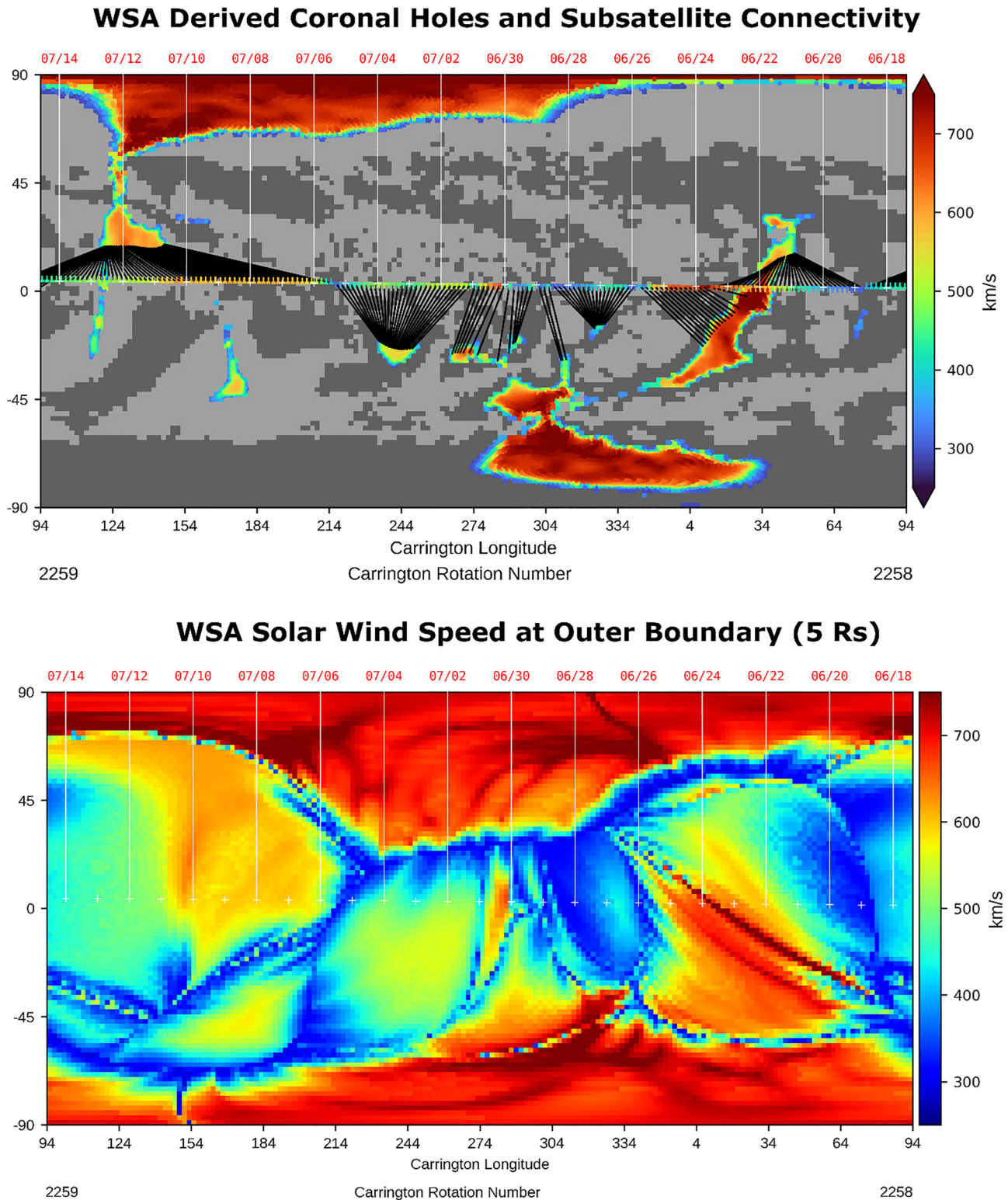


Figure 2. (Top panel) Wang Sheeley Arge (WSA)-derived coronal holes and subsatellite point connectivity. The black lines show the connectivity between subsatellite points (for Earth) at the model outer boundary ($5 R_s$) and their source location within a coronal hole. The coronal holes are colored by the outflow velocity at each location; with the fastest solar wind originating farthest from the edge. (Bottom panel) The solar wind velocity estimated by WSA at model outer boundary. We see a complex velocity field constructed from the combined coronal hole sources, where-in each patch in a coronal hole becomes mapped to a much larger region on the outer boundary surface.

Table 1

Sources of Variation in Ensemble Varied Set

Variable	Name	Description	Multiplier
$N_{\text{realizations}}$	ADAPT ensemble variation	Number of ADAPT realizations	12 (as of publication)
$\theta_{\text{subsat}} \pm \frac{\Delta\theta}{2}$	Perturbed physics	The subsatellite points on the model outer boundary used in the point-parcel simulation is perturbed by plus or minus half grid cell, equal to 1° (see Section 2)	3
Δt_{window}	Window of acceptable parcel arrivals times around target time	Because solar wind results often show structure delayed or too soon, accept parcel intersections occurring within ± 12 hr	Not consistent due to varying v_{sw} , but on average around 7

function (PDF) over space. While past works have made use of ensemble sets to represent uncertainty (Badman et al., 2023; Macneil et al., 2022), our methodology differs in the final presentation PDF using kernel density estimation.

Using the connectivity algorithm, we can provide a collection of connectivity solar surface endpoints \bar{x}_i corresponding to times t_i within a given time window to be discussed below. To create a broad representative collection of connectivity endpoints, which we call the varied set, we utilize multiple solutions originating from the following sources (Table 1). Overall, the varied set has about 250–350 members.

The ADAPT model provides a natural source of variation because it provides an ensemble of global photospheric field maps. The ensemble is generated by varying one or more parameters in model. The parameter generally varied in ADAPT corresponds to unobservable supergranular convection cells, which affect the magnetic flux transport, on the far-side and high-latitude regions outside the area of assimilated inputs (Hickmann et al., 2015). As of publication, the ADAPT model provides a set of 12 maps (or realizations) and corresponding to an increase of the size of the varied set by a factor of 12.

Furthermore, as mentioned in Section 2, we perturb the subsatellite points on the model outer boundary used as starting points for the point-parcel simulation. The perturbation is simply plus or minus half a grid cell, which comes out to 1° at the current grid resolution., which is 2° This increases the size of the varied set by $3\times$.

Finally, we consider parcels arriving at the satellite within ± 12 hr around the target time. This is justified because solar wind models have been identified in practice to produce correct structures (such as current sheet crossings or high-speed streams) but often offset in time with what is observed at L_1 . Therefore, considering connections from parcels arriving slightly in the past and the future gives the opportunity to catch events that arrive too soon or that are delayed. Because variation exists in the solar wind speed, parcels will arrive at an unsteady rate, but at model resolution of 2° , about ~ 7 parcels arrive at earth in a 24-hr interval, which increases the size of the varied set by the same factor.

The next step is to transform our discrete set of point-estimates into a smooth probability distribution for ease of visualization and interpretation. The method used is the Kernel Density Estimate with a Gaussian Kernel (Hastie et al., 2009). This method is simply akin to placing a Gaussian distribution at the location of each point estimate, and then adding together all the Gaussians to form a final distribution. An example of this in one dimension is shown in Figure 3. Conceptually, where there are many point-estimates close by, many Gaussians will super-impose together to create a high-point of the probability distribution, and when few point-estimates lie within a region, the probability distribution will lack Gaussian contributions and remain low.

This has advantages over a histogram in that no “pixelization” effect occurs, as the probability distribution is smooth between points. Compared to fitting a single shape to the data, the kernel density estimate can take on multiple shapes, from bimodal, to tailed and curved, to adapt to the data. This method has been previously applied to ensemble forecasts in the context of weather simulation (X. Wang & Bishop, 2005; Wilks, 2011). While previous work has predicted source location probability distributions using prescribed shapes such as Gaussian Mixture Clustering (Koukras et al., 2022), this is the first work which predicts a source location probability distribution without any a priori assumptions about its structure.

The equation for the kernel density estimate in two dimensions is given in Equation 2. We note that we also utilize a weighting scheme for the Gaussians, so-in that some Gaussian contribute more than others based on our confidence in that point-estimate. The weighting scheme used here is a “time-decay” weighting scheme based on the Δt_{window} component of the varied set. That is, if we are making a probability map for a target time, solar wind parcels in the acceptable window are weighted in an exponentially decreasing fashion based on their time-difference from the target time.

Kernel Density Estimate and Time-Decay Weighting

$$f(\mathbf{x}) = \frac{1}{\sum_{k=1}^N w_i} \sum_{k=1}^N w_i G(d(\mathbf{x}, \mathbf{x}_i); \sigma) \quad (3)$$

$$w_i = e^{-|t_i - t_{\text{target}}|/T}$$

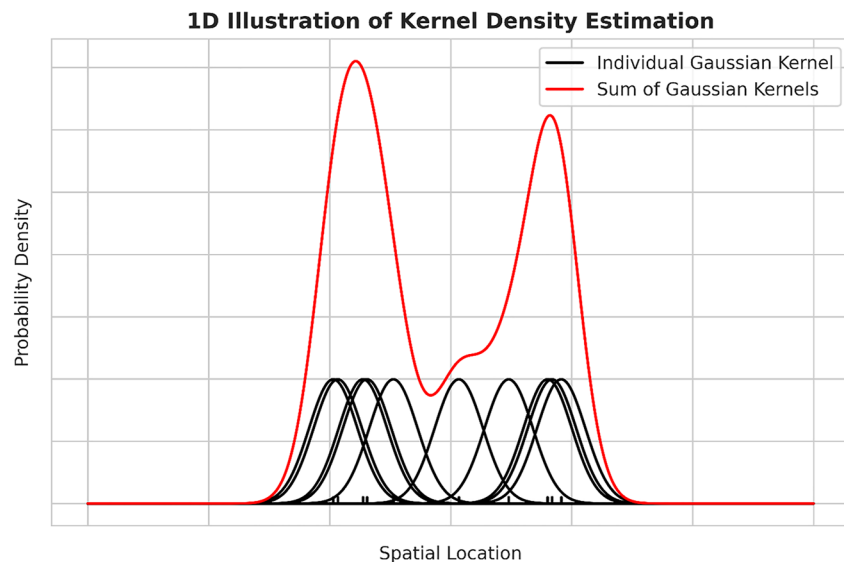


Figure 3. 1D Illustration of how the Kernel Density Estimate algorithm works to turn a collection of point estimates into a smooth probability distribution. In the method, a Gaussian is placed at each of the point estimates, and final smooth probability distribution is the sum of the Gaussians. By adding these Gaussians, the result will be taller where many points are clustered together and shorter where fewer points occur, with little constraint on its shape. The standard deviation of the Gaussians is a parameter which is calibrated in Appendix A.

In Equation 3, $f(\mathbf{x})$ is the value of the PDF at the location \mathbf{x} on the surface of the Sun, nominally given by (θ, ϕ) . The values w_i are the weights of each Gaussian. The function $d(\mathbf{x}, \mathbf{x}_i)$ is the great-circle distance (haversine distance) metric between the surface locations \mathbf{x} and \mathbf{x}_i , in units of degrees. The function $G(\cdot; \sigma)$ is the Gaussian PDF with standard deviation σ . For more information about how the σ parameter is calibrated, see Appendix A. In short, σ is chosen to compromise between over and under-smoothing via a maximum likelihood test for the full ensemble. In the definition for the weights w_i , the variable t_i is the corresponding time of parcel arrival, t_{target} is the target time for the forecast, and T is a scale time (comparable to a scale height) which defines the rate of exponential decay. A scale time of 3 hr was used so that when t_i is 12 hr away from t_{target} , the weight drops to 2% of what it would be if $t_i = t_{\text{target}}$.

An example ensemble forecast for the solar wind connectivity source locations at Earth in the form of a PDF can be found in Figure 4. The PDF is displayed red/orange/yellow normalized to the peak (red), where it fades to transparent when the value is very close to zero. The background of this map is the photospheric magnetic field B_r from a single ADAPT ensemble member. Since the ensemble set of maps generally do not appear significantly different except at the finer scales, selecting any map for display purposes is sufficient. The green regions identify the coronal holes based on coronal predictions by the WSA model. The WSA model is run for each ADAPT map used in the 24-time interval, and they are then all weighted together and normalized (i.e., 1 implies 100% agreement that a pixel is within a coronal hole, while 0 indicates that none of the solutions predicts the pixel resides within a coronal hole). The transparency level is set by the consensus across the varied set that a coronal hole exists at that location.

The PDF shows a bimodal and tailed shape, collectively stating that the connectivity is expected to come from one of two nearby (but distinct) low-latitude coronal holes. For the more western coronal hole, the PDF describes the source location most likely being in the eastern edge of that hole, and for the more eastern hole, spread somewhere across the length (as a tailed distribution). The nominal boundaries of each PDF region (teal lines) are found by searching for closed loops of isolines of a particular very low PDF value (taken here as 10^{-7}). Once a boundary is found for a region, the PDF can be integrated within the closed loop to summarize the total enclosed probability within that region (yellow text). In this example, we see that each coronal hole is represented about equally in the varied set (51% vs. 49%).

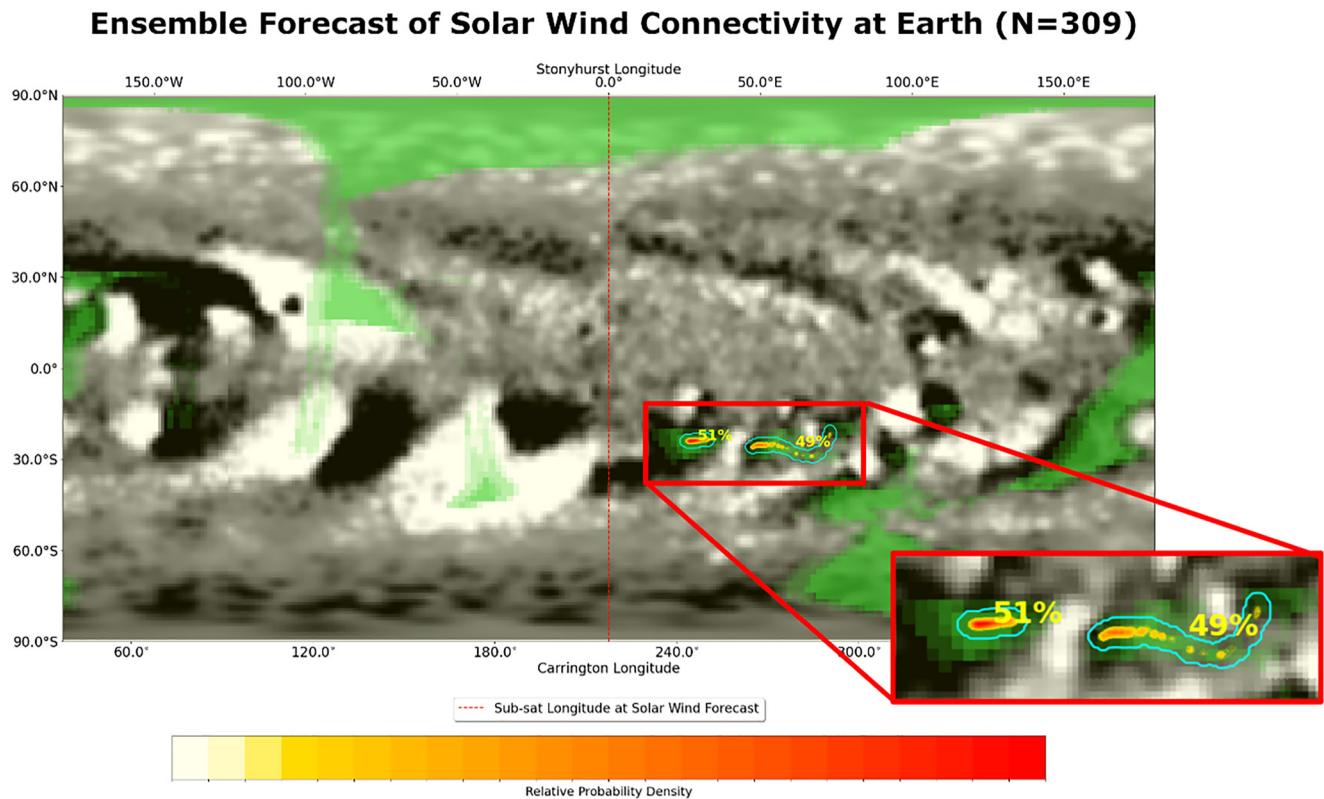


Figure 4. An ensemble forecast of the solar wind connectivity at Earth using the method presented in this article. This utilizes the conversion of a set of point estimates of the solar wind source location ($N = 309$) to a smooth probability distribution using kernel density estimation with a Gaussian kernel. The background coronal holes are colored in green and averaged between the varied set, with the opacity proportional to the consensus that a coronal hole exists at that location. The background magnetogram is taken from a single realization for reference. The probability distribution is normalized to the peak, with red representing the most probable location and yellow representing declining likelihood.

3.1. Uncertainty and Limitations

There is a fair amount of uncertainty and limitations in virtually all coronal and solar wind models (MacNeice, 2009a, 2009b; MacNeice et al., 2011, 2018; Norquist, 2013; Norquist & Meeks, 2010; Owens et al., 2005, 2008), due to the lack of knowledge of the specific mechanism(s) heating the corona, as well as those generating and accelerating the solar wind. As a result, both potential-field type models such as WSA, and more advanced MHD codes account for these processes by relying on free parameters that are tuned empirically. PF based models are magnetostatic, and as such, the coronal field is derived based on the assumption that the corona is nearly potential. The coronal field is then forced (often prematurely) to be open and radial beyond a certain height (typically $2.5 R_s$), defined as the source surface. When the source surface (an artificial construct) is raised, PFSS models produce smaller coronal holes and less total open area, while the opposite holds true when this height is lowered. To improve on the original PFSS model, WSA additionally derives the outer coronal field with the SCS model, extending the coronal field to anywhere from 5 to $21.5 R_s$ before it becomes radial (Arge et al., 2004; McGregor et al., 2008; Y.-M. Wang & Sheeley, 1995), providing a more realistic magnetic field topology of the upper corona.

Several studies suggest that the source surface height is dependent on solar cycle, as well as input magnetogram data (e.g., Arden et al., 2014; Nikolić, 2019). In the results shown here, we do not vary the source surface height, as the traditional value has been shown to produce good agreement between WSA-derived coronal holes and those observed in He 10,830 nm and EUV on three Carrington rotation averages (Wallace et al., 2019). Thus, one would only need to adjust the source surface height on a case-by-case basis. When using our tool, we advise that any adjustments made to the source surface height be made to within reasonable thresholds as defined by the current literature, and that any model output be validated comparing the model-derived v_{sw} and Interplanetary Magnetic Field (IMF) with that observed in situ.

On the other hand, MHD models use a time-dependent energy equation to generate the coronal magnetic field, where the amount of heating can be modulated to open and close down flux. While it is necessary to use a time-dependent MHD model for certain applications (e.g., identifying constraints for coronal heating processes, modeling transient events), a PFSS-based model is better suited for identifying the source regions of the ambient solar wind observed in situ as described in this paper for at least the three following reasons. First, MHD models are not currently capable of deriving the spacecraft magnetic connectivity to the solar surface in a way that is computationally efficient enough to provide routine predictions of the source locations of the observed solar wind. Thus, PFSS models are preferred for applications such as our tool where we are deriving the spacecraft connectivity to the solar surface. While static PFSS-based models have their limitations, using an ensemble of updated synchronic-like photospheric field maps (such as those provided by ADAPT) to drive WSA both minimizes and helps to quantify the uncertainty in the model solution. Second, while MHD and force free models can more accurately determine the coronal magnetic field configuration in non-potential regions such as active regions, the input boundary conditions in these regions are usually the most suspect (e.g., most photospheric field maps make use of the line-of-sight magnetic field measurements along with the radial field assumption, which is known to be a very poor assumption in active regions). It is therefore unclear if using more advanced models actually improves magnetic connectivity determinations over that of the PFSS model, given the large uncertainty in the input boundary conditions. Third, both PFSS and MHD solutions have been shown to produce relatively similar coronal topologies at solar minimum (Riley et al., 2006). This is likely due to the lack of knowledge in the detailed physics driving coronal heating and the acceleration of the solar wind, and the empirical nature of both types of models.

Further, a primary source of uncertainty for all coronal models are the large uncertainties in the global photospheric magnetic field maps, which serve as a key driver to these models (see Posner et al., 2021 for a detailed review of this problem). One major issue is that polar fields are not observed, yet they are a crucial component to the global coronal field (Arge & Pizzo, 2000; Riley et al., 2019). Similarly, global coronal and solar wind solutions can be significantly impacted by the lack of far-side observations (Arge et al., 2013; Cash et al., 2015; Wallace et al., 2022). This can occur when a new active region emerges on the Sun's far-side and therefore is not included in the photospheric field map. Coronal solutions can also be adversely impacted when these new active regions rotate onto the solar near-side and “suddenly” appear in the maps. The effects of this can be mitigated in retrospective analyses, where we can allow for new far-side emergence to assimilate into the solution for 1–2 days after a period of interest, which has been shown in previous work to stabilize the coronal solution (Wallace et al., 2022). We also validate our model results with v_{sw} and B_r observations in retrospective studies. While this issue cannot be mitigated for forecasting, our tool helps improve forecasts which suffer the most from the lack of global photospheric field observations, by quantifying the uncertainty in the source location of the solar wind observed at various spacecraft or satellites through a varied ensemble set of estimates.

Lastly, it is important to note that it is not possible for PFSS models to capture the Sun's time dependent phenomena associated with the opening and closing of magnetic flux (e.g., magnetic reconnection, transient eruptions such as coronal mass ejections). While we can account for time-dependent photospheric phenomena with ADAPT, WSA only derives the magnetic connectivity between an observing spacecraft and model-derived field lines that are open. Similarly, WSA cannot provide information regarding how long a particular field line has been open. Therefore, when the model predicts that a spacecraft measured plasma near the time-dependent magnetic open-closed boundary, the two physical scenarios that are possible are 1) the plasma originated from that particular open field line, or 2) the plasma originated on closed field that was recently opened via interchange reconnection, whereas WSA cannot make the distinction between the two possible scenarios. Reliably determining how long the magnetic field has been open at the solar wind source region is currently beyond the capability of *all* global coronal models due to the inaccuracies of the input boundary conditions (i.e., photospheric field maps). However, the WSA model can reliably distinguish when the in situ observed solar wind originates from intermittently open field at the magnetic open-closed boundary (i.e., roughly the coronal hole boundary), or from sources that are deep inside coronal holes where the field is continuously open. This information is critical for studies on how the solar wind forms, and for characterizing the in situ properties of solar wind from these two distinct sources.

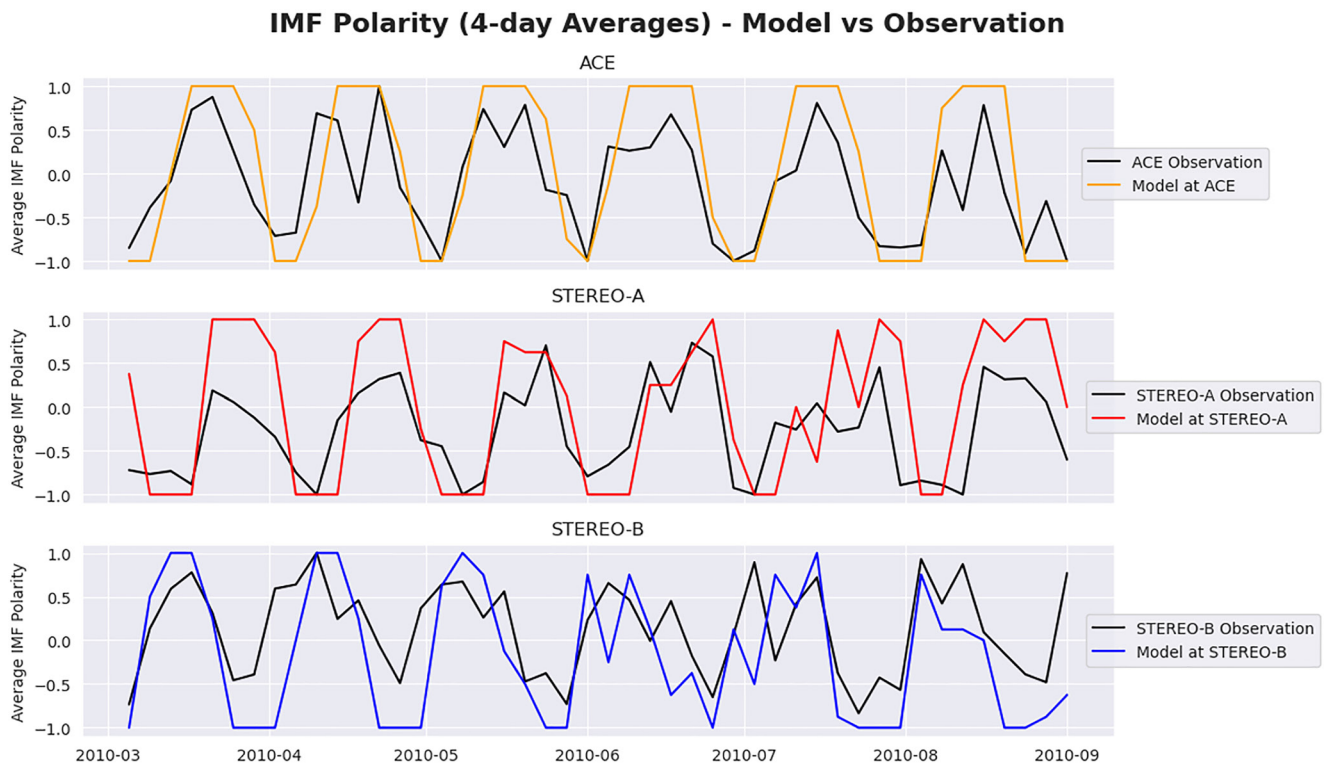


Figure 5. Comparison of average of model distribution of Interplanetary Magnetic Field (IMF) polarities versus the average of the observed IMF polarities, for three locations in the solar system with corresponding in situ observations. The size of the averaging window is 4 days. In this plot, we notice the broad agreement as to when the current sheet is crossed and the polarity switches. This gives us confidence (though not absolute guarantee) that the model is accurately predicting the correct coronal hole.

4. Validation

Validating a connectivity forecast is difficult due to the lack of a ground truth data set (<https://www.sciencedirect.com/science/article/pii/S0273117723006877?via%3Dihub>). It is possible to compare to other models, however without the ability to verify those other models against ground truth, the comparison can at best conclude both models in agreement without claim that either is correct. However, it is possible to use secondary features to compare a forecast with observation, not giving a full guarantee but still determining a baseline of validity. Overall, such a comparison can only give modest confidence and provide a sanity-check; it cannot provide conclusive proof that the model is valid. We note that on a case-by-case basis, a number of secondary features can be used to support validation. Such secondary features include the darkness of the source location EUV imagery, whether the observed solar wind is fast/slow, and observed solar wind composition characteristics such as first ionization potential (Young, 2005, 2018).

The secondary features we will compare are the distribution of the IMF polarity at the photosphere (-1 or $+1$) provided by our model, and the similar IMF polarity observed at the spacecraft using in situ instrumentation. Until periods of current sheet crossings, the distribution of IMF polarities should begin dominated by one polarity, then split during the current sheet crossing, and return to being dominated a single polarity. A similar pattern should be visible in the satellite-observed IMF polarity. While one could in principle do a similar comparison between model and observation with the solar wind speed, in practice this is difficult because model solar wind speed forecasts are prone to large errors (i.e., Owens et al., 2008) and one can forecast the correct speed while getting the source location wrong. While consistency between model predictions of speed and IMF polarity with observations provides greater confidence that the model is representing the overall state of the corona and solar wind well, IMF polarity is a much more reliable and definitive indicator as to whether the model is not magnetically connected to the correct region (i.e., if the model predicts the wrong polarity, one knows the model forecast is incorrect).

In Figure 5, we compare the distribution of IMF polarities with data from Advanced Composition Explorer (ACE), Solar Terrestrial Relations Observatory Spacecraft A (STEREO-A), and Solar Terrestrial Relations Observatory Spacecraft B (STEREO-B) during the period of March 2010 to August 2010. This provides a multi-point perspective on the model quality for three locations at roughly 1 AU, with each location experiencing current sheet crossings at different times. In this plot, the model line (colored) shows (a) the IMF polarity obtained from the varied sets, averaged over a 4-day window, and (b) the IMF polarity observed using in situ instrumentation on the corresponding spacecraft averaged over a 4-day window.

In this plot, we can see that there is broad agreement as to when a current sheet is crossed and when the polarity switches. Between the three locations (ACE, STEREO-A, and STEREO-B), the approximate times between polarity switches is approximately half a Carrington rotation for each, though the times of the polarity switches are not in phase due to their differing locations in the solar system. This is largely consistent with the structure of the heliospheric current during this period, which is largely that of tilted dipole. This gives us confidence that the model is accurately predicting the correct coronal hole (at least as far as to distinguish between positive and negative polarity coronal holes) and switching its prediction at the correct times. Should the model be picking a coronal hole at random, or in a way which is minimally correlated with the true coronal hole, it would be statistically unlikely to reproduce the level of correlation shown here.

5. Conclusion

In this work, we present a method for preparing a set of ensemble predictions of solar wind connectivity in the form of a probability distribution. The method uses near-sun magnetic fields obtained from the ADAPT-WSA model and combines those fields with a parcel-propagation method to simulate the arrival of solar wind at a target in the solar system. This produces an ensemble of point-estimates of the solar wind source location, which is then post-processed into a smooth probability distribution using a method known as kernel density estimation.

Estimation of the solar wind connectivity has applications for both scientific investigation and operations. For research applications, tools such as the one presented here are the only way to connect in situ solar wind observations to their specific source at the Sun observed remotely. Remote sensing of the corona, through multispectral imagery and spectroscopy techniques, reveals both the source type (e.g., active region or quiet sun at the magnetic open-closed boundary, or deep inside a coronal hole), as well as its associated properties, including composition and temperature. This information is vital to answering open questions in Heliophysics, including how the solar wind is formed, from its origin, to its subsequent release, heating, and acceleration (Viall & Borovsky, 2020). In operations, solar imagers with narrow field of view must choose where to point their instruments, and having knowledge of the upcoming solar wind source region allows such instruments to point in the most useful direction.

The task of estimating solar wind connectivity is a challenge for the modeling community because there is no ground truth to compare against. The major uncertainties are (a) the ability to estimate the near-sun magnetic field, (b) the ability to accurately estimate the solar wind speed at the outer boundary of the coronal model, and (c) the ability to accurately model the transit of solar wind (including stream interactions) from the outer corona to the target. In this work, we tackle (a) and (b) using the ADAPT-WSA model previously established in the community and present a method for approaching (c) which balances computational efficiency with reasonable detail.

Future avenues for research include assessing expectations of modeled source location uncertainty in relation to key parameters. A central question to answer is: how do measures of uncertainty change in relation to key parameters such as (a) position in the solar cycle, (b) coronal hole longitude/latitude, and (c) physical distance between the final location and the sun? By addressing these questions, we acquire insight regarding fundamental qualities of PF models and gain physical intuition about the nature of coronal holes themselves. Our method, which is powerful in its ability to represent complex source configurations with minimal assumptions about the probability distribution structure, is uniquely well suited to studying these questions.

Appendix A: Calibration of the Gaussian Kernel Width (σ)

Equation 3 encodes the model of the kernel density estimation which is used to convert N point estimates of the solar wind source location (where N varies approximately between 250 and 350) into a smooth probability distribution. The method works by summing together N small Gaussian distributions, each one centered at one of the

point estimates. A free variable in this equation is σ , the standard deviation used for the Gaussian distributions. This variable is calibrated using data, and the method for doing so is the focus of this appendix section. The value of this variable was estimated to be 0.31° .

The calibration of σ can be thought of in terms of in the extremes where σ is much too small, and when σ is much too large. As σ becomes very small, the resultant probability distribution in Equation 3 approaches a set of Dirac deltas centered on each point in the varied set. When this happens, the distribution loses its ability to interpolate a probability value in between elements of the varied set. In this sense, it is under-smoothed. When σ becomes much too large, the opposite happens, and the resultant distribution becomes over-smoothed. If one imagines the case where σ is very large (say, 360°), the resultant distribution becomes roughly “flat everywhere” and approaches a uniform distribution. Neither the under-smoothed case nor the over-smoothed case is desirable, but logically there should exist a middle-ground σ which meets the intention of being appropriately-smoothed.

The method used for estimating σ is maximum likelihood estimation (MLE; Hastie et al., 2009). To use MLE, we designed an experiment where we split the varied set for a given map into a training set (90% size) and a test set (10%) size. The logic used is to test multiple candidates for σ to find the one which maximizes the likelihood of the resultant distribution predicting the elements of the test set. This likelihood is encoded in Equation A1 below, where $L(\sigma)$ is the likelihood function to be maximized, Test is the set of test points, and $f(\mathbf{x}_i; \sigma)$ is the value of the resultant distribution PDF at the location \mathbf{x}_i on the surface of the Sun parameterized by candidate parameter σ (see Equation 2). The product in Equation A1 originates from the notion that we are seeking to quantify the probability of finding \mathbf{x}_1 and \mathbf{x}_2 and \mathbf{x}_3 and so on.

Likelihood Function to Find σ

$$L(\sigma) = \prod_{\mathbf{x}_i \in \text{Test}} f(\mathbf{x}_i; \sigma) \quad (\text{A1})$$

To illustrate the behavior of $L(\sigma)$, we will go back to our descriptions of under-smoothed (σ too small) and over-smoothed (σ too large) distributions, if the resultant distribution is drastically under-smoothed, then it will predict probabilities too low at each point in the test set, because the Gaussian has already dropped off. If the distribution is drastically over-smoothed, it will predict too low as well, because each Gaussian is normalized to have an area under the curve of 1, and in order to span a large area it must have lower probability density at each point in that area.

In our experiment, we performed this MLE using varied set once every 6 hr for predictions at ACE, STEREO-A, and STEREO-B between January 2010 and March 2010. Each varied set produced a single estimate of σ , which we collected and analyzed in aggregate. This produced many estimates of σ , ranging between 0.06° up to 1.00° , with a mean of 0.31° and a standard deviation of 0.23° . The distribution had a positive skew of 1.15 (tail extending to the right). To use a single estimate for operational efficiency, the mean of 0.31° was selected as the outcome of the experiment.

Data Availability Statement

Python code which implements the methodology of this paper and generates plots similar to Figure 4 can be found open-source on Zenodo at: <https://doi.org/10.5281/zenodo.8384747>. Assimilated Global Oscillation Network Group measurements in the form of ADAPT maps which describe the radial magnetic field at every point around the sun can be obtained directly for no charge from the National Solar Observatory (NSO) at <https://gong.nso.edu/>. Runs of ADAPT-WSA can be performed for no charge through the Community Coordinated Modeling Center's (CCMC's) Run-on-Request system at <https://ccmc.gsfc.nasa.gov/models/WSA~v.5.2/>. Data from the ACE, STEREO-A, and STEREO-B satellites can be obtained from NASA's Space Physics Data Facility at <https://spdf.gsfc.nasa.gov/>.

Acknowledgments

The authors would like to thank the following people who provided useful feedback and interesting conversation, in alphabetical order: Eric Adamson (NOAA Space Weather Prediction Center), Craig DeForest (Southwest Research Institute), Sarah Gibson (High Altitude Observatory), Andrew Marble (NOAA Space Weather Prediction Center), and Mark Miesch (NOAA Space Weather Prediction Center).

References

- Altschuler, M. D., & Newkirk, G. (1969). Magnetic fields and the structure of the solar corona: I: Methods of calculating coronal fields. *Solar Physics*, 9(1), 131–149. <https://doi.org/10.1007/bf00145734>
- Arden, W. M., Norton, A. A., & Sun, X. (2014). A “breathing” source surface for cycles 23 and 24. *Journal of Geophysical Research: Space Physics*, 119(3), 1476–1485. <https://doi.org/10.1002/2013ja019464>
- Arge, C. N., Harvey, K. L., Hudson, H. S., & Kahler, S. W. (2003). Narrow coronal holes in Yohkoh soft X-ray images and the slow solar wind. In *AIP Conference Proceedings* (Vol. 679, No. 1, pp. 202–205). American Institute of Physics.
- Arge, C. N., Henney, C. J., Hernandez, I. G., Toussaint, W. A., Koller, J., & Godinez, H. C. (2013). Modeling the corona and solar wind using ADAPT maps that include far-side observations. In *AIP Conference Proceedings* (Vol. 1539, No. 1, pp. 11–14). American Institute of Physics.
- Arge, C. N., Henney, C. J., Koller, J., Compeau, C. R., Young, S., MacKenzie, D., et al. (2010). Air force data assimilative photospheric flux transport (ADAPT) model. In *AIP Conference Proceedings* (Vol. 1216, No. 1, pp. 343–346). American Institute of Physics.
- Arge, C. N., Henney, C. J., Koller, J., Toussaint, W. A., Harvey, J. W., & Young, S. (2011). Improving data drivers for coronal and solar wind models. In *5th International Conference of Numerical Modeling of Space Plasma Flows (ASTRONUM 2010)* (Vol. 444, p. 99).
- Arge, C. N., Luhmann, J., Odstrcil, D., Schrijver, C., & Li, Y. (2004). Stream structure and coronal sources of the solar wind during the May 12th, 1997 CME. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(15–16), 1295–1309. <https://doi.org/10.1016/j.jastp.2004.03.018>
- Arge, C. N., Odstrcil, D., Pizzo, V. J., & Mayer, L. R. (2003). Improved method for specifying solar wind speed near the Sun. In *AIP Conference Proceedings* (Vol. 679, No. 1, pp. 190–193). American Institute of Physics.
- 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), 10465–10479. <https://doi.org/10.1029/1999ja000262>
- Badman, S. T., Riley, P., Jones, S. I., Kim, T. K., Allen, R. C., Arge, C. N., et al. (2023). Prediction and verification of Parker solar probe solar wind sources at 13.3 R_{\odot} . *Journal of Geophysical Research: Space Physics*, 128(4), e2023JA031359. <https://doi.org/10.1029/2023ja031359>
- Borovsky, J. E. (2016). The plasma structure of coronal hole solar wind: Origins and evolution. *Journal of Geophysical Research: Space Physics*, 121(6), 5055–5087. <https://doi.org/10.1002/2016ja022686>
- Cash, M. D., Biesecker, D. A., Pizzo, V., de Koning, C. A., Millward, G., Arge, C. N., et al. (2015). Ensemble modeling of the 23 July 2012 coronal mass ejection. *Space Weather*, 13(10), 611–625. <https://doi.org/10.1002/2015sw001232>
- Dresing, N., Gómez-Herrero, R., Klassen, A., Heber, B., Kartavykh, Y., & Dröge, W. (2012). The large longitudinal spread of solar energetic particles during the 17 January 2010 solar event. *Solar Physics*, 281(1), 281–300. <https://doi.org/10.1007/s11207-012-0049-y>
- Gómez-Herrero, R., Malandraki, O., Dresing, N., Kilpua, E., Heber, B., Klassen, A., et al. (2011). Spatial and temporal variations of CIRs: Multi-point observations by STEREO. *Journal of Atmospheric and Solar-Terrestrial Physics*, 73(4), 551–565. <https://doi.org/10.1016/j.jastp.2010.11.017>
- Harvey, J. W., Hill, F., Hubbard, R. P., Kennedy, J. R., Leibacher, J. W., Pintar, J. A., et al. (1996). The global oscillation network group (GONG) project. *Science*, 272(5266), 1284–1286. <https://doi.org/10.1126/science.272.5266.1284>
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). Springer.
- Hickmann, K. S., Godinez, H. C., Henney, C. J., & Arge, C. N. (2015). Data assimilation in the ADAPT photospheric flux transport model. *Solar Physics*, 290(4), 1105–1118. <https://doi.org/10.1007/s11207-015-0666-3>
- Koukras, A., Laurent, D., & Keppens, R. (2022). Estimating uncertainties in the back-mapping of the fast solar wind. *arXiv preprint arXiv:2212.11553*. Retrieved from <https://ui.adsabs.harvard.edu/abs/2022EGUGA..2410570K%2Fabstract>
- MacNeice, P. (2009a). Validation of community models: Identifying events in space weather model timelines. *Space Weather*, 7(6), S06004. <https://doi.org/10.1029/2009sw000463>
- MacNeice, P. (2009b). Validation of community models: 2. Development of a baseline using the Wang-Sheeley-Arge model. *Space Weather*, 7(12), S12002. <https://doi.org/10.1029/2009sw000489>
- MacNeice, P., Elliott, B., & Acebal, A. (2011). Validation of community models: 3. Tracing field lines in heliospheric models. *Space Weather*, 9(10), S10003. <https://doi.org/10.1029/2011sw000665>
- MacNeice, P., Jian, L. K., Antiochos, S. K., Arge, C. N., Bussy-Virat, C. D., DeRosa, M. L., et al. (2018). Assessing the quality of models of the ambient solar wind. *Space Weather*, 16(11), 1644–1667. <https://doi.org/10.1029/2018sw002040>
- Macneil, A. R., Owens, M. J., Finley, A. J., & Matt, S. P. (2022). A statistical evaluation of ballistic backmapping for the slow solar wind: The interplay of solar wind acceleration and corotation. *Monthly Notices of the Royal Astronomical Society*, 509(2), 2390–2403. <https://doi.org/10.1093/mnras/stab2965>
- Mayer, L. R. (2012). A new method for coupling the potential field source surface and the Schatten current sheet models. Colorado University at Boulder Cooperative Institute for Research in Environmental Sciences.
- McGregor, S. L., Hughes, W., & Arge, C. (2005). Investigation of the PFSS and Schatten current sheet interface in the WSA Model. *AGU Fall Meeting Abstracts*, 2005.
- McGregor, S. L., Hughes, W. J., Arge, C. N., & Owens, M. J. (2008). Analysis of the magnetic field discontinuity at the potential field source surface and Schatten Current Sheet interface in the Wang-Sheeley-Arge model. *Journal of Geophysical Research*, 113(A8), A08112. <https://doi.org/10.1029/2007ja012330>
- Nikolić, L. (2019). On solutions of the PFSS model with GONG synoptic maps for 2006–2018. *Space Weather*, 17(8), 1293–1311. <https://doi.org/10.1029/2019sw002205>
- Norquist, D. C. (2013). Forecast performance assessment of a kinematic and a magnetohydrodynamic solar wind model. *Space Weather*, 11(1), 17–33. <https://doi.org/10.1029/2012sw000853>
- Norquist, D. C., & Meeks, W. C. (2010). A comparative verification of forecasts from two operational solar wind models. *Space Weather*, 8(12), S12005. <https://doi.org/10.1029/2010sw000598>
- Owens, M. J., Arge, C. N., Spence, H. E., & Pembroke, A. (2005). An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model. *Journal of Geophysical Research*, 110(A12), A12105. <https://doi.org/10.1029/2005ja011343>
- Owens, M. J., Spence, H. E., McGregor, S., Hughes, W. J., Quinn, J. M., Arge, C. N., et al. (2008). Metrics for solar wind prediction models: Comparison of empirical, hybrid and physics-based schemes with 8-years of L1 observations. *Space Weather*, 6(8), S08001. <https://doi.org/10.1029/2007SW000380>
- Plowman, J. E., & Berger, T. E. (2020). Calibrating GONG magnetograms with end-to-end instrument simulation I: Background, the GONG instrument, and end-to-end simulation. *Solar Physics*, 295(10), 1–21. <https://doi.org/10.1007/s11207-020-01682-4>
- Posner, A., Arge, C. N., Staub, J., StCyr, O. C., Folta, D., Solanki, S. K., et al. (2021). A multi-purpose heliophysics L4 mission. *Space Weather*, 19(9), e2021SW002777. <https://doi.org/10.1029/2021sw002777>

- Riley, P., Downs, C., Linker, J. A., Mikić, Z., Lionello, R., & Caplan, R. M. (2019). Predicting the structure of the solar corona and inner heliosphere during Parker Solar Probe's first perihelion pass. *The Astrophysical Journal Letters*, 874(2), L15. <https://doi.org/10.3847/2041-8213/ab0ec3>
- 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), 15889–15901. <https://doi.org/10.1029/2000ja000121>
- Riley, P., Linker, J. A., Mikić, Z., Lionello, R., Ledvina, S. A., & Luhmann, J. G. (2006). A comparison between global solar magnetohydrodynamic and potential field source surface model results. *The Astrophysical Journal*, 653(2), 1510–1516. <https://doi.org/10.1086/508565>
- Riley, P., Linker, J. A., & Nick Arge, C. (2015). On the role played by magnetic expansion factor in the prediction of solar wind speed. *Space Weather*, 13(3), 154–169. <https://doi.org/10.1002/2014sw001144>
- Rouillard, A. P., Kouloumvakos, A., Vourlidas, A., Kasper, J., Bale, S., Raouafi, N. E., et al. (2020). Relating streamer flows to density and magnetic structures at the Parker solar probe. *The Astrophysical Journal – Supplement Series*, 246(2), 37. <https://doi.org/10.3847/1538-4365/ab579a>
- Schatten, K. H. (1971). Current sheet magnetic model for the solar corona (No. X-692-71-132).
- Schatten, K. H., Wilcox, J. M., & Ness, N. F. (1969). A model of interplanetary and coronal magnetic fields. *Solar Physics*, 6(3), 442–455. <https://doi.org/10.1007/bf00146478>
- Schrijver, C. J., & DeRosa, M. L. (2003). Photospheric and heliospheric magnetic fields. *Solar Physics*, 212(1), 165–200. <https://doi.org/10.1023/a:1022908504100>
- Sheeley, N. R., Jr. (2017). Origin of the Wang–Sheeley–Arge solar wind model. *History of Geo-and Space Sciences*, 8(1), 21–28. <https://doi.org/10.5194/hgss-8-21-2017>
- Viall, N. M., & Borovsky, J. E. (2020). Nine outstanding questions of solar wind physics. *Journal of Geophysical Research: Space Physics*, 125(7), e2018JA026005. <https://doi.org/10.1029/2018ja026005>
- Wallace, S., Arge, C. N., Pattichis, M., Hock-Mysliwiec, R. A., & Henney, C. J. (2019). Estimating total open heliospheric magnetic flux. *Solar Physics*, 294(2), 1–20. <https://doi.org/10.1007/s11207-019-1402-1>
- Wallace, S., Arge, C. N., Viall, N., & Pihlström, Y. (2020). On the relationship between magnetic expansion factor and observed speed of the solar wind from coronal pseudostreamers. *The Astrophysical Journal*, 898(1), 78. <https://doi.org/10.3847/1538-4357/ab98a0>
- Wallace, S., Jones, S. I., Arge, C. N., Viall, N. M., & Henney, C. J. (2022). New insights into the first two PSP solar encounters enabled by modeling analysis with ADAPT-WSA. *The Astrophysical Journal*, 935(1), 24. <https://doi.org/10.3847/1538-4357/ac731c>
- Wang, X., & Bishop, C. H. (2005). Improvement of ensemble reliability with a new dressing kernel. *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, 131(607), 965–986. <https://doi.org/10.1256/qj.04.120>
- Wang, Y.-M., & Sheeley, N. R., Jr. (1992). On potential field models of the solar corona. *Astrophysical Journal, Part 1*, 392(1), 310–319. <https://doi.org/10.1086/171430>
- Wang, Y.-M., & Sheeley, N. R., Jr. (1995). Solar implications of Ulysses interplanetary field measurements. *The Astrophysical Journal*, 447(2), L143. <https://doi.org/10.1086/309578>
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic Press.
- Worden, J., & Harvey, J. (2000). An evolving synoptic magnetic flux map and implications for the distribution of photospheric magnetic flux. *Solar Physics*, 195(2), 247–268. <https://doi.org/10.1023/a:1005272502885>
- Young, P. R. (2005). The element abundance FIP effect in the quiet Sun. *Astronomy & Astrophysics*, 439(1), 361–366. <https://doi.org/10.1051/0004-6361:20052963>
- Young, P. R. (2018). Element abundance ratios in the quiet sun transition region. *The Astrophysical Journal*, 855(1), 15. <https://doi.org/10.3847/1538-4357/aaab48>
- Zhao, L., Landi, E., Lepri, S. T., Gilbert, J. A., Zurbuchen, T. H., Fisk, L. A., & Raines, J. M. (2017). On the relation between the in situ properties and the coronal sources of the solar wind. *The Astrophysical Journal*, 846(2), 135. <https://doi.org/10.3847/1538-4357/aa850c>
- Zhu, B., Liu, Y. D., Kwon, R. Y., & Wang, R. (2018). Investigation of energetic particle release using multi-point imaging and in situ observations. *The Astrophysical Journal*, 865(2), 138. <https://doi.org/10.3847/1538-4357/aada80>
- Zurbuchen, T. H., Fisk, L. A., Gloeckler, G., & von Steiger, R. (2002). The solar wind composition throughout the solar cycle: A continuum of dynamic states. *Geophysical Research Letters*, 29(9), 66-1–66-4. <https://doi.org/10.1029/2001gl013946>