An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model

M. J. Owens

Center for Space Physics, Boston University, Boston, Massachusetts, USA

C. N. Arge

Space Vehicles Directorate, Air Force Research Laboratory, Hanscom Air Force Base, Massachusetts, USA

H. E. Spence and A. Pembroke

Center for Space Physics, Boston University, Boston, Massachusetts, USA

[1]   One of the primary goals of the Center for Integrated Space Weather Modeling (CISM) effort is to assess and improve prediction of the solar wind conditions in near-Earth space, arising from both quasi-steady and transient structures. We compare 8 years of L1 in situ observations to predictions of the solar wind speed made by the Wang-Sheeley-Arge (WSA) empirical model. The mean-square error (MSE) between the observed and model predictions is used to reach a number of useful conclusions: there is no systematic lag in the WSA predictions, the MSE is found to be highest at solar minimum and lowest during the rise to solar maximum, and the optimal lead time for 1 AU solar wind speed predictions is found to be 3 days. However, MSE is shown to frequently be an inadequate "figure of merit" for assessing solar wind speed predictions. A complementary, event-based analysis technique is developed in which high-speed enhancements (HSEs) are systematically selected and associated from observed and model time series. WSA model is validated using comparisons of the number of hit, missed, and false HSEs, along with the timing and speed magnitude errors between the forecasted and observed events. Morphological differences between the different HSE populations are investigated to aid interpretation of the results and improvements to the model. Finally, by defining discrete events in the time series, model predictions from above and below the ecliptic plane can be used to estimate an uncertainty in the predicted HSE arrival times.

Citation: Owens, M. J., C. N. Arge, H. E. Spence, and A. Pembroke (2005), An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model, *J. Geophys. Res.*, *110*, A12105, doi:10.1029/2005JA011343.

1.   Introduction

[2]   Variability in the solar wind conditions impinging on the Earth's magnetosphere can have adverse effects on a number of ground- and space-based technologies [*Feynman and Gabriel*, 2000], making "space weather" prediction highly desirable. As the community moves toward detailed numerical simulations to provide such forecasts, it is imperative to track the predictive capabilities of the available models throughout their development, as was done for atmospheric weather models through the 20th century [e.g., *Siscoe et al.*, 2004]. A model's predictive accuracy is compared with the predictions of an unchanging "baseline" model for the same period, so as to compute the relative skill of the model [*Spence et al.*, 2004]. In this paper we outline the testing techniques used for the Center for Integrated Space Weather Modeling (CISM) solar wind models and suggest an event-based approach (using high-speed enhancements, HSEs) to better understand their strengths and weaknesses.

2.   Wang-Sheeley-Arge Model

[3]   CISM is using the Wang-Sheeley-Arge model (WSA) [*Arge and Pizzo*, 2000; *Arge et al.*, 2003] as the baseline model to which all other solar wind models (e.g., the MHD-based CORHEL model [*Odstrčil et al.*, 2004]) will be compared. In this paper the WSA model is both validated

and used to demonstrate a new event-based validation technique.

[4] The WSA model is a combined empirical and physics-based representation of the quasi-steady global solar wind flow. It can be used to predict the ambient solar wind speed and interplanetary magnetic field polarity at Earth. It is an improved version of the original Wang and Sheeley model [*Wang and Sheeley*, 1990]. WSA uses ground-based line-of-sight (LOS) observations of the Sun's surface magnetic field (in the form of synoptic maps) as its input. LOS field observations taken with the Mount Wilson Solar Observatory (MWO) magnetograph are used in this study, and the synoptic maps assembled from them are updated on a daily basis with new observations (i.e., with magnetograms) introduced as the Sun rotates and new regions come into the instrument's field of view. Thus the WSA predictions are updated and refined on a day-by-day basis. The updated synoptic maps are used as input to a magnetostatic potential field source surface (PFSS) model [*Schatten et al.*, 1969; *Altschuler and Newkirk*, 1969], which determines the coronal field out to 2.5 solar radii ($2.5R_S$). The output of the PFSS model serves as input to the Schatten Current Sheet (SCS) model [*Schatten*, 1971], which provides a more realistic magnetic field topology of the upper corona. The following empirical relationship (similar in form to the one by *Arge et al.* [2004]) is used to assign solar wind speed at a radius of 5 R_S in this study:

$$V(f_s, \theta_b) = 265 + \frac{1.5}{(1+f_s)^{2/7}} \left\{ 5.8 - 1.6e^{\left[1-(\theta_b/7.5)^3\right]} \right\}^{3.5} \text{ km s}^{-1}.$$

(1)

It is a function of two coronal parameters, flux tube expansion factor (f_s) and the minimum angular separation (at the photosphere) between an open field foot point and its nearest coronal hole boundary (θ_b). The empirically derived solar wind speeds and magnetic field polarity at 5 R_S from the SCS model are then fed into a one-dimensional (1-D) modified kinematic code [*Arge and Pizzo*, 2000] that propagates the solar wind out to 1 AU and accounts for stream interactions. A more comprehensive summary of the WSA model is provided by *Arge et al.* [2004].

[5] Thus WSA currently provides predictions of the solar wind radial flow speed and the magnetic polarity at 1 AU. To infer additional solar wind properties, we assume the nonradial components of the solar wind flow are zero. Proton density is then estimated by assuming constant mass flux, and proton temperature is derived from the speed-temperature relation of [*Lopez*, 1987]. To calculate a magnetic field, we assume the vector lies in the ecliptic plane with the speed used to compute a Parker spiral angle. An intensity of 5 nT is assumed. This complete set of basic solar wind magnetic field and plasma parameters allows all aspects of the numerical models to be compared to this baseline, if so required. However, as the WSA predictions are all (to greater or lesser degrees) derived from the solar wind speed prediction, this validation study will focus attention on this parameter.

3. Mean Square Error

[6] We begin by investigating the performance of the WSA model using mean square error (MSE) as a measure of

prediction accuracy. This method is extremely useful as a first-order assessment of the model prediction and is the technique that will be used to compute the standard CISM metrics (in the form of skill scores) for the models [*Spence et al.*, 2004].

[7] The L1 solar wind speed predictions of WSA are compared with in situ magnetic field and plasma measurements made by Wind (MFI [*Lepping et al.*, 1995], SWE [*Ogilvie et al.*, 1995]) and ACE (MAG [*Smith et al.*, 1998], SWEPAM [*McComas et al.*, 1998]). The observational data are available at very high time resolutions (<5 min); however, the spatial resolution of the photospheric field maps (i.e., 5° in this study) and the steady state nature of WSA lead to an effective limit on the temporal resolution of predictions at 1 AU of ∼9.6 hours. Both the WSA and observational data sets must be averaged/interpolated to be at the same time resolution. We choose a convenient resolution of 8 hours, as only large-scale structures are of interest. Figure 1 shows the 3-day advance WSA solar wind speed prediction (red) against the ACE/Wind observed solar wind speed (black) for the time period covered by this study (1995 through 2002).

[8] It is instructive to first verify that the choice of time resolution (8 hours) does not significantly affect the calculated MSE. The open circles in Figure 2 show how the normalized root mean square (RMS) error between the observed solar wind speed and the 3-day advance WSA prediction varies when using different time resolutions, ranging from 5 min up to 72 hours (the data cover 1995 through 2002). It is clear that when dealing with solar wind speed, the measured MSE is not significantly influenced by the convenient choice of resolution (shown as the solid vertical line). However, this is not necessarily true for solar wind properties with shorter coherence times (such as B_Z, shown as crosses in Figure 2). Thus the time resolution for metric calculations must be chosen carefully for each individual property, based upon the structures and factors that are deemed important for space weather predictions.

[9] As WSA uses daily updated synoptic maps of the photospheric field as its input, predictions can be made many days in advance and then refined as the Sun rotates and new magnetograms become available. Table 1 shows the RMS error for 1 to 7 day advance predictions for the period 1995 through 2002, as well as the average across all predictions and all years. The left-hand panel of Figure 3 shows the same data, but in the form of a "skill" score:

$$Skill = \left(1 - \frac{MSE}{MSE_{REF}} \right) \times 100,$$

(2)

where MSE_{REF} is a reference MSE, usually that of the baseline model for the same time period. However, in this case MSE_{REF} is taken to be the average MSE over all predictions for that year. Figure 3 also shows the average skill over all years: skill increases from the 7-day advance prediction as newer synoptic maps are used, but only as far as 3-days advance, after which using newer maps causes the skill gradually to drop again. This is not unexpected, as the solar wind transit time is ∼3−4 days, meaning the important factor in the creation of synoptic maps for prediction purposes is to correctly reproduce the magnetic structure at the solar central meridian at the time the solar wind left the
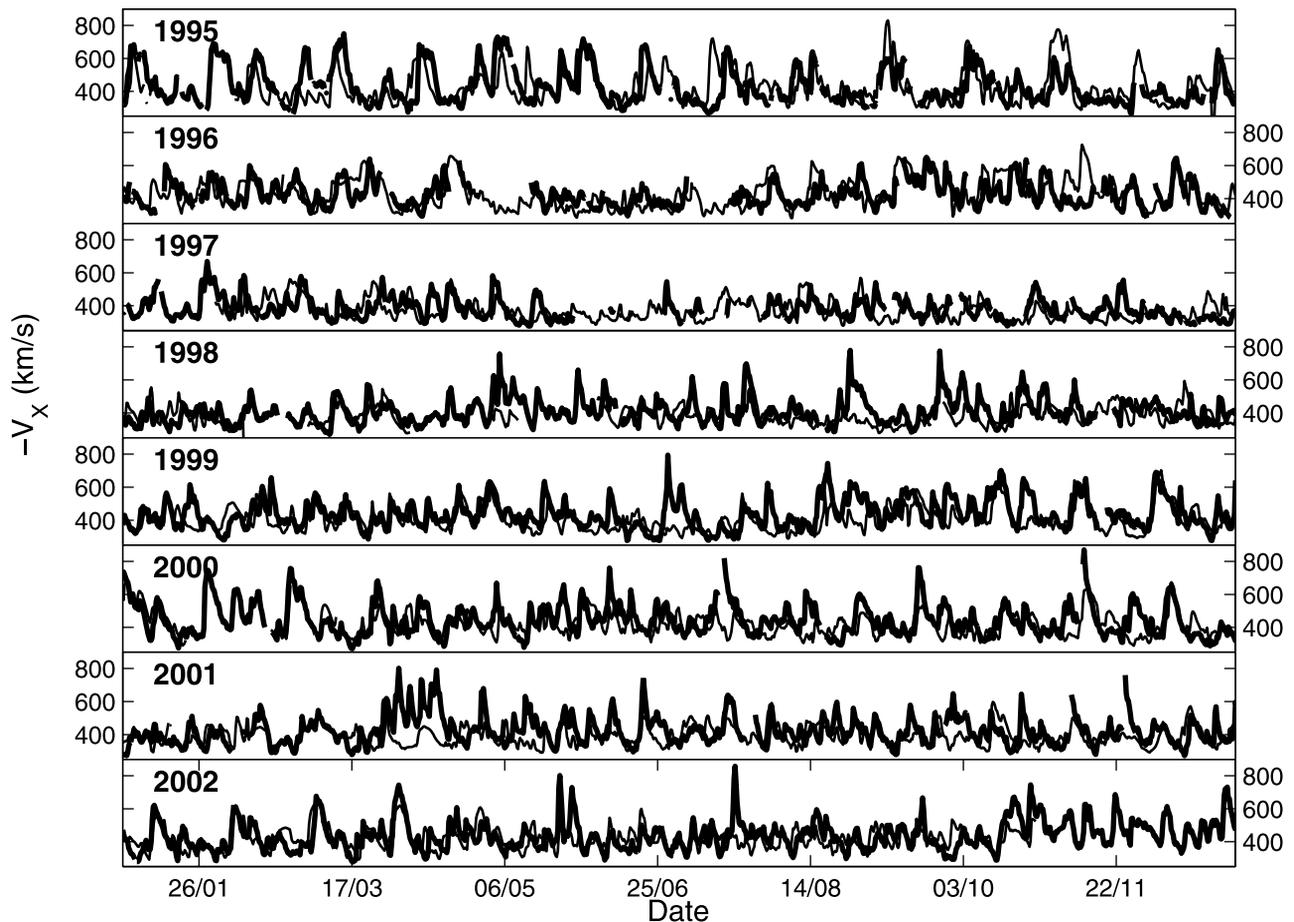
Figure 1. Yearly time series plots of the 3-day advance WSA solar wind speed prediction (red) against the ACE/Wind observed solar wind speed (black) for the time period covered by this study (1995 through 2002). See color version of this figure in the HTML.

Sun (i.e., reproduce the solar wind launch location). The accuracy of predictions made with lead times longer than 5–6 days suffer greatly, placing a practical limit on the lead time of solar wind forecasts made with WSA.

[10] The right-hand panel of Figure 3 shows how the RMS error for the various prediction lead times varies with time (the 3-day advance prediction, deemed to be the most accurate, is shown as the solid bold line). There does not appear to be any obvious correlation of MSE with the solar activity cycle. Indeed, the greatest RMS error occurs during solar minimum (1995), with most accurate predictions during the rise to solar maximum (1997). However, visual inspection of the WSA predictions suggests the model is tracking the large-scale behavior of the solar wind speed much better during 1995 than during more active periods (e.g., 2000). This is especially evident in Figure 1.

[11] MSE analysis is also used to look for any systematic offset in the WSA predictions, as *McPherron et al.* [2004] reported a 12-hour discrepancy when using WSA predictions of the solar wind speed during 1995 to drive *Ap* index predictions. In this study a time lag (ranging from +2 days to −2 days in 10 min increments) was added to the 8 years of WSA predictions, the time series was reinterpolated, and the MSE was recalculated. We find the error to be lowest for zero offset, increasing steadily with both positive and negative temporal offsets, indicating no systemic lead/lag time in the WSA predictions (however, see also section 4.4).

4. An Event-Based Approach

[12] While it is extremely useful to have a single number (i.e., MSE or skill score) to monitor performance of a predictive model during its development, it gives little information about why the predictive accuracy has changed. Furthermore, there do exist circumstances where point-by-point data/model time series comparison techniques such as MSE (and linear or Spearman correlation coefficients, r_L and r_S) can be misleading. For example, if the timing of a fast solar wind stream arrival and/or duration differs slightly between model and data, the MSE can be larger than if the model had predicted no fast stream at all. This is illustrated by the hypothetical situation in Figure 4, where model B (the red line) is capturing the large-scale behavior of the observed solar wind speed (the solid line) better than model A (the blue line). However, the RMS error is smaller for model A (72.8 km/s) than model B (94.0 km/s), and the correlation coefficients for model A are significantly higher than for model B.

[13] Using MSE for assessing the accuracy of solar wind speed predictions does not, by definition, emphasize the
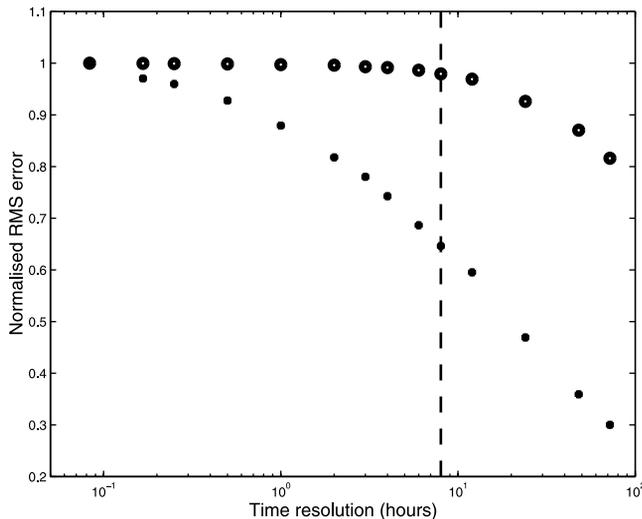
Figure 2. A plot of the RMS error between observed and WSA predicted solar wind properties as a function of the time resolution of the data, for the years 1995 through 2002. The errors have been normalized to the RMS error at a 5-min time resolution. Open circles (crosses) represent the plasma flow speed (B_Z). Note the logarithmic scale on the x axis: The time resolution varies from 5 min to 72 hours, with the dashed line denoting the 8-hour resolution used in the remainder of this study.

solar wind features forecasters/model developers are necessarily interested in. Both small-scale differences over long time periods and large-scale differences with small timing errors can lead to poor skill scores, when in fact these effects may be far more tolerable to forecasters than a model completely failing to capture the large-scale structure. For these reasons MSE is often an inadequate "figure of merit," and in this study we propose an event-based approach using key features in the solar wind speed time series, namely high-speed-enhancements (HSEs). This large-event focused approach has been employed for both shock arrival [*Smith et al.*, 2004] and magnetospheric response [*Weigel et al.*, 2004] as it gives a quantification of a specific aspect of a model's capability as a forecast tool.

[14] In the remainder of this paper we use WSA predictions to outline methods for defining a HSE, systematically selecting and associating HSEs in model and spacecraft data, and interpreting the results. It is hoped such an analysis could be pragmatically applied to model predictions to enable subjective comparison of their performance.

4.1. Defining HSEs

[15] The method and criterion used to select HSEs from the data effectively defines the features in the solar wind speed profile in which we are interested. Operationally, any large change in the bulk behavior of the solar wind speed is important. Thus we do not require an HSE to necessarily be a transition from slow, helmet-streamer-associated solar wind to pure, coronal fast wind. This also makes the selection process simpler, as systematically differentiating between these two solar wind populations can be difficult for observational in situ data, and it is unclear how these definitions would translate to different model outputs.

Additionally, high-speed solar wind flows resulting from fast ICMEs are also included. This approach allows models that successfully include transient solar wind events to improve their prediction accuracy, whether it be characterized by MSE or HSE analysis. Again, this also makes the process of HSE selection simpler, as completely removing ICME-driven HSEs from data could be problematic: frequently, only the disturbance to the ambient solar wind (such as the flank of the shock wave) is observed and not the actual ejecta itself, complicating their systematic identification and subsequent removal from the data.

[16] As the nature of the speed enhancement is not of interest, a selection method independent of the form of the speed profile can be used, and we simply require that a speed gradient threshold be exceeded for a minimum duration. This makes the HSE index easy to both implement and use. By a process of trial and error, we found that with the smoothed 8-hour resolution data, a speed increase (dV) of 100 km/s over a period (dT) of 2 days are reasonable criteria for selecting "significant" events (though, of course, the selection criteria are in fact defining what we regard as significant). Thus a HSE is defined as any region of solar wind, 2 days or longer, in which the net solar wind speed increase is 100 km/s or more. The HSE continues to last as long as these criteria continue to be met, but we define the characteristic time of a HSE as the time of maximum gradient. We find that even with the highly smoothed data, short duration speed decreases can still cause a single speed enhancement (as picked out by eye) to be interpreted as multiple events. Thus any HSEs occurring with 2 days of each other are regarded as a single event. These parameters could be modified for specific purposes.

4.2. Associating HSEs

[17] Applying the selection criterion to the observation and model time series produces two lists of HSEs. The next step is to unambiguously associate HSEs between the different data sets. First, potential associations are made if the temporal boundaries of two HSEs overlap. A margin for error is also included, in that HSEs with boundaries within T_{GAP} of each other are also potentially associated (a constant value of 2 days is used for T_{GAP}, but section 5 describes how out of ecliptic observations can be used to put event-specific estimates on the timing errors of HSEs).

[18] Once a list of potentially associated HSEs has been compiled, it is necessary to further require that each model

Table 1. Root MSE (in km/s) Between the Observed Solar Wind Speed and the 1- to 7-Day Advanced WSA Predictions for the Years 1995 Through 2002[a]

	Days Advance Prediction Was Made							
Year	1	2	3	4	5	6	7	Average
1995	108.55	109.53	112.40	108.57	112.40	110.43	115.60	111.09
1996	86.99	87.35	84.35	83.53	82.25	82.04	84.13	84.40
1997	75.65	73.91	74.33	74.71	73.36	77.02	75.47	74.93
1998	96.58	92.45	90.17	90.49	95.34	99.77	100.43	95.11
1999	89.33	90.08	90.08	93.37	95.70	95.18	94.35	92.62
2000	94.07	93.36	91.85	93.85	94.88	97.52	96.24	94.55
2001	102.07	101.73	99.89	101.64	103.70	100.82	102.64	101.79
2002[b]	84.10	85.73	82.92	85.06	85.64	86.27	88.42	85.46
Average	92.68	92.31	91.38	91.94	93.62	94.20	95.35	

[a]The average values over all years and all predictions are also shown.
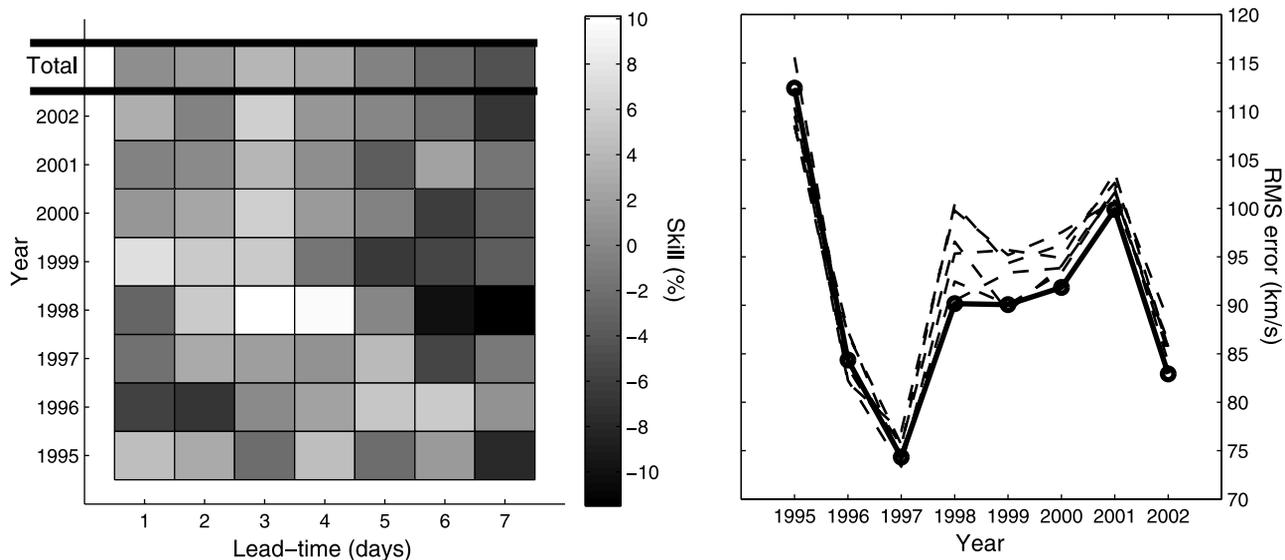[b]Note that the predictions for 2002 only extend to 25 October.

Figure 3. The left-hand panel shows the skill of the 1- to 7-day advance predictions of solar wind speed for the years 1995 to 2002 (i.e., the MSE of a prediction for a certain year, normalized by the average MSE over all predictions for that year). The top row shows the average skill over all years (i.e., the MSE of a prediction over all years, normalized by the average MSE over all predictions and all years): The 3-day advance prediction has the highest skill, whereas the prediction using a 7-day lead time suffers significantly. The right hand panel shows the RMS error of the 1- to 7-day predictions as a function of time, with the 3-day prediction as the solid bold line.

HSE be associated with no more than one observed HSE, and vice versa. When multiple associations are possible, a number of steps are used to determine the best pairwise associations, as shown in Figure 5. First, the numbers of unique model and observed HSEs in a group of multiple pairings are checked. If equal, a solution may be possible by simply maximizing the number of pairwise associations (i.e., finding a single set of pairings that uses each model and observed HSE once only). If this "pair logic" approach fails to resolve the problem, any HSE pairings that can be associated with data gaps in the opposing data set are removed (it should be noted that due to the high degree of smoothing of the data, data gaps are infrequent). If any HSE pairings were eliminated, we return to the "pair logic" test, otherwise the worst pairing based on timing differences

is removed. This process is continued until a unique set of pairings is achieved (see also Figure 5).

[19] The final, unique associations between model and observed HSEs are termed "hits." Any remaining model (observed) HSE that cannot be associated with an observed (model) HSE or a data gap is labeled a "false" ("missed") HSE.

4.3. Assessing the Model

[20] Comparison of discrete events rather than continuous time series allows the use of several new methods for quantitatively characterizing the performance of a model. Perhaps the most fundamental measures are the numbers of true positive, false positive, and false negative predictions (the hit, missed, and false HSEs, respectively). We do not
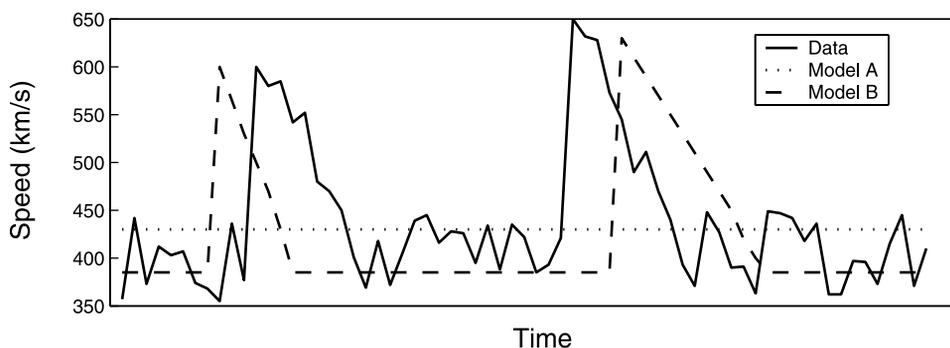

Figure 4. Artificially generated observation and model solar wind speed time series to illustrate a potential problem with interpreting MSE. While model B (the red line) captures the large-scale solar wind structure much better than model A (the blue line), the errors in the model B timings of the fast stream arrivals mean that $MSE_A < MSE_B$. See color version of this figure in the HTML.
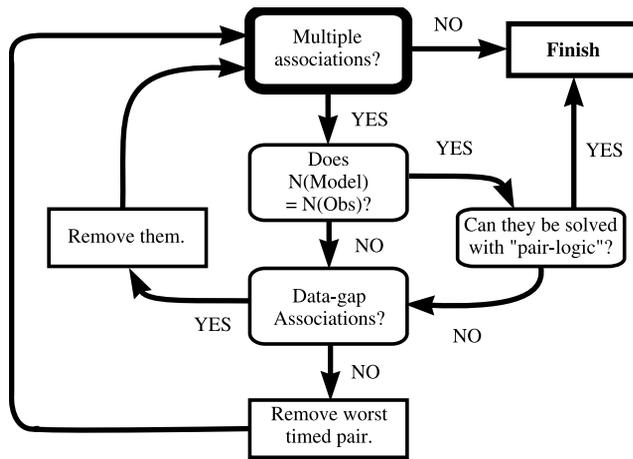
Figure 5. A flow-chart showing the method by which multiple associations between HSEs are sorted into unique pairings. When multiple associations are possible between model and observed HSEs, we first attempt look for a unique solution that maximizes the number of pairs ("pair logic"). If this fails, any HSE pairings that can be associated with data gaps in the opposing data set are removed and pair logic reattempted. Otherwise, the pair with the largest timing error is removed. This process is repeated until a unique set of pairings is achieved.

consider the "true negative" events (i.e., both model and observation showing no event) in this study. Such statistics are best presented in the form of a contingency table (see Table 2), allowing the user to convolve the data as they see fit. However, there are several established "indices" (particularly within meteorological forecast community) that can provide useful insight performance of the model (see *Murphy* [1993] for more detail). The "threat score" (TS, or "critical success index") provides a measure of the accuracy of the prediction when true negatives are not important:

$$TS = \frac{N_{HIT}}{N_{HIT} + N_{MISS} + N_{FALSE}}. \tag{3}$$

The threat score ranges from 1 for perfect skill to 0 for no skill and provides a statistical measure of a model's ability to predict events. An example of a diagnostic index is "bias," which ranges from 0 to infinity (with 1 as a perfect score) and is a measure of a model's tendency to underforecast (<1) or overforecast (>1) events:

$$BIAS = \frac{N_{HIT} + N_{FALSE}}{N_{HIT} + N_{MISS}}. \tag{4}$$

Obviously, different indices with different weightings to hit, miss, and false HSEs could be used as different factors need to be emphasized for various scientific/operational purposes.

4.4. HSEs in WSA

[21] We now apply the HSE analysis described in the previous sections to the WSA predictions. Using HSE selection criterion of dV = 100 km/s and dT = 2 days, we find 246 (208) HSEs in the ACE/Wind (WSA) data set for the years 1995 through 2002. An association time of T_{gap} = 2 days gives 166 "hits," with 80 (42) observed (model) HSEs with no companion in the model (observational) dataset. After data gap considerations, these numbers reduce to 64 (36) true missed (false) HSEs (see Table 2). The mean maximum speed, $\langle V_{MAX} \rangle$, of the model (observed) hit HSEs is 567 (530) km/s, while $\langle V_{MAX} \rangle$ of the missed (false) HSEs is 546 (485) km/s. For the hit HSEs, the average timing difference between the observed and model HSEs is 1.68 days.

[22] For 1995–2002, WSA would have a threat score (equation (3)) of 0.63. Such a number can be directly compared with other model predictions for the same period to assess their relative ability to correctly predict HSEs. It is hoped that such a measure will be useful, both scientifically and operationally, in assessing the strengths and weakness of solar wind speed models. WSA has a bias (equation (4)) of 0.88, indicating it slightly underforecasts HSEs (due to the absence of transients in the model. See also section 6).

[23] Figure 6 shows the details of the observed and model HSEs. The left and middle panels show properties of the HSEs that were captured in both data sets (i.e., the "hits"): The left (middle) panel shows a histogram of the difference between the observed and model arrival times (maximum speeds), with the solid vertical line representing the mean value of 0.44 days (36.7 km/s). There is some suggestion of a half day time lag in the WSA prediction of HSE arrival times even though the MSE analysis clearly showed no such effect. However, the distribution of dT (left panel of Figure 6) does seem symmetric about zero, and we also note that the time resolution of the predictions (8 hours) and the average error in the HSE arrival time prediction (i.e., $\langle |dT| \rangle$ = 1.68 days) are comparable and significantly larger, respectively, than the suggested lag. Hence it is not currently possible to say conclusively whether any time lag is present in the WSA prediction of HSEs.

[24] The distribution of dV (middle panel of Figure 6) does not appear to be symmetric about zero: Although the peak of the distribution is close to zero, the tail decays less rapidly in the positive dV direction than in the negative. This is reflected in the nonzero mean of the distribution (36.7 km/s), though this value is well below the average error in the HSE V_{MAX} (i.e., $\langle |dV| \rangle$ = 79.6 km/s).

[25] The right panel of Figure 6 summarizes the distribution of V_{MAX} of all HSE populations: Solid (dashed) blue lines are missed (false) HSEs, while solid (dashed) red lines are the observed (model) values for the hit HSEs. There are a several points to note in this plot:

[26] 1. The number of hits (i.e., area under the red lines) is much larger than both the number of missed and false HSEs (the areas under the blue dashed and solid lines, respectively).

Table 2. A Contingency Table Showing the Numbers of Hit, Miss, and False HSEs Using the WSA 3-Day Advanced Predictions[a]

		Observed:	
		HSE	No HSE
Model:	HSE	166 (166)	36 (42)
	No HSE	64 (80)	-

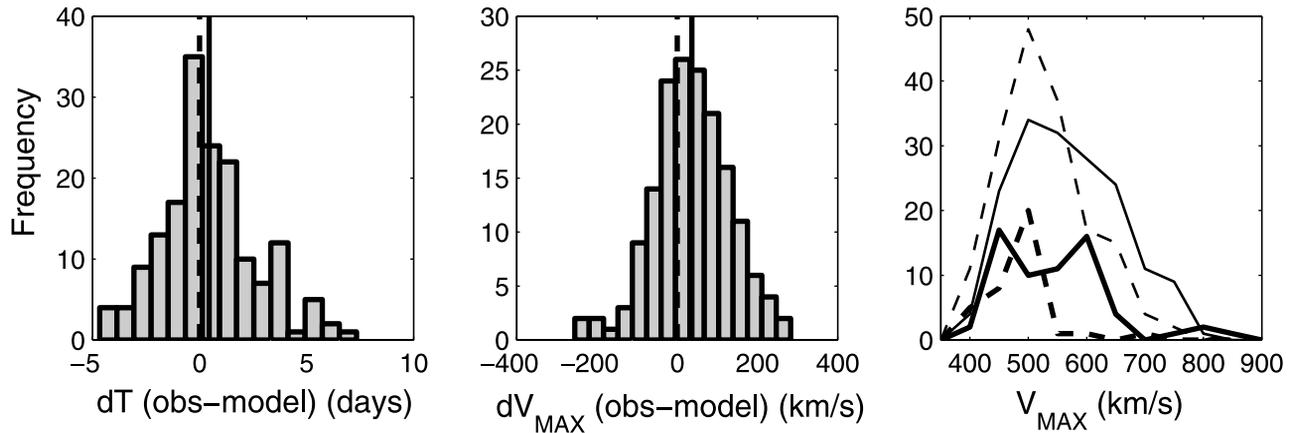[a]The values in parentheses represent the statistics before data gap considerations are made.

Figure 6. Three plots showing the properties of the observed and model HSEs. The left and middle panels show properties of the HSEs appearing in both data sets (i.e., the "hits"); the left (middle) panel shows a histogram of the difference between the observed and model arrival times (maximum speeds), with the solid vertical line representing the mean value of 0.44 days (36.7 km/s). The right panel summarizes the distributions of the maximum speeds of HSEs; solid (dashed) blue lines are missed (false) HSEs, while solid (dashed) red lines are the observed (model) values for the hit HSEs. See color version of this figure in the HTML.

[27] 2. While the observational and model V_{MAX} distributions for the hit events (the red lines) peak at similar values, the observational distribution has a broader tail in the high-speed direction, resulting in the positive offset of $\langle dV \rangle$.

[28] 3. False HSEs (the dashed blue line) are limited to smaller events (i.e., V_{MAX} peaking at <500 km/s and dropping off by 550 km/s).

[29] 4. Missed HSEs (the solid blue line) have a similar peak in the V_{MAX} distribution corresponding to smaller events; however, there is also a second peak at higher speeds (peaking at 600 km/s, cutting off by 700 km/s).

[30] The false and small magnitude missed HSEs are probably events for which the counterpart in the opposing data set was too small to be picked up by the simple selection criterion. However, the larger-magnitude missed HSEs cannot be explained in this manner. It is likely that

they are the result of fast ICME-driven disturbances which are not expected to be captured by WSA.

[31] An alternative method of displaying the differing characteristics of the various types of HSE is by a multiposed epoch analysis. Figure 7 shows the observed and predicted HSEs (using the same identifiers for hit/miss/false as Figure 6) with the maximum speed gradient taken to be the epoch time. The model density and temperatures were calculated assuming constant mass flux and a speed-temperature relation, respectively (see section 1). Error bars are not shown for reasons of clarity.

[32] As expected, all HSEs show a sudden jump in speed at $t = 0$, decaying back down to ambient slow solar wind speeds within ∼5 days. The observed HSEs show a significant spike in density and temperature at the epoch time.
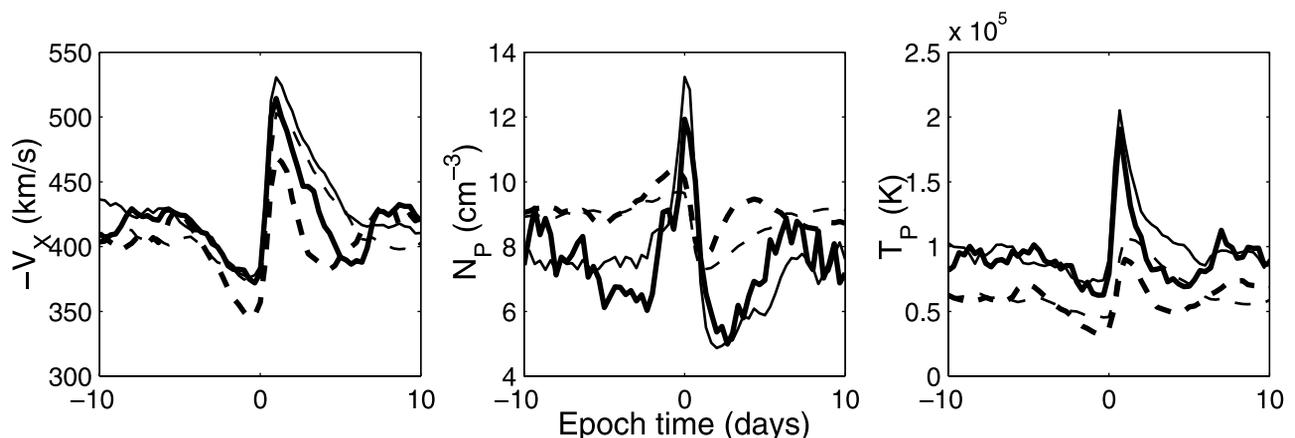

Figure 7. A mulitposed epoch analysis (using the time of maximum speed gradient as $t = 0$) for the various groups of HSE (using the same format as Figure 6). Model density and temperatures were calculated assuming constant mass flux and a speed-temperature relation, respectively (see section 1). While the speed profile is well captured by the WSA model, the empirical relations used to derive density and temperature (not part of the standard WSA model) are clearly inadequate for forecasting purposes. See color version of this figure in the HTML.
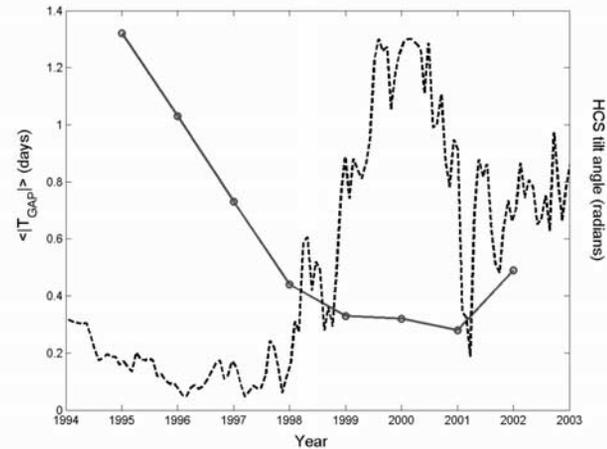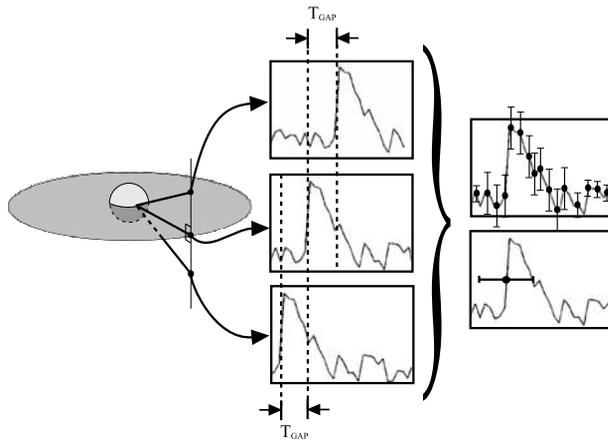
Figure 8. The figure on the left shows how WSA predictions above/below the ecliptic at the L1 point are currently used to estimate uncertainties in predicted speed. However, by describing discrete events (HSEs), uncertainties in the timing of events (T_{GAP}) can also be estimated. The figure on the right shows $\langle |T_{GAP}| \rangle$ as a function of time (the solid red line). The dashed line shows the HCS tilt angle (in radians) computed by a potential field solution to the WSO data. There is a clear anticorrelation between these two parameters. See color version of this figure in the HTML.

[33] Hit HSEs (solid and dashed red lines for observed and model HSEs, respectively) show a similar speed profile but with model peak speed slightly lower than observed. Conversely, the simple empirical relations for model density and temperature do not track well with the observations and are obviously not adequate to capture the interaction at the boundary between the two plasma regimes. It is hoped that using the WSA coronal solution to drive a MHD simulation of the heliosphere will provide a better representation of the plasma buildup and heating at the stream interface than the current method of ballistically mapping WSA solar wind speeds from the source surface to 1 AU (see section 2). HSE analysis would allow this to be directly tested.

[34] The two observed HSE populations (i.e., the missed and observed hit HSEs, shown as solid blue and red lines, respectively) appear very similar in speed, density, and temperature profiles. However, observed hit streams exhibit a smooth transition back to ambient solar wind speeds and temperatures from $t = 0$, whereas the missed streams show a sharper dropoff to lower than ambient values before a jump to background levels at $t \sim 5$ days. Thus the observed hit streams are likely to be the result of corotating interaction regions (CIRs), whereas the missed stream profiles show some additional features (such as the speed overshoot) that could be interpreted in terms of expanding, transient events (i.e., ICMEs). As WSA does not include such transient structures, such HSEs would be expected to result in missed events.

[35] Finally, we examine the false HSEs (dashed blue lines). The missed events have much smaller peak speeds than the other three categories of HSE and have a larger decrease in solar wind speed just prior to $t = 0$. Thus false events are likely to be mainly the result of the associated observed event being too small to be recognized as an HSE, as also suggested by Figure 6. It should also be noted that the false HSE speed profile for $t > 0$ follows a similar shape to that of the missed HSEs, suggesting the observed features

may be the result of the inclusion of a large fraction of small events, rather than being the effect of expanding transient structures.

5. Out-of-Ecliptic Predictions

[36] Finally, we consider how to incorporate out-of-ecliptic predictions into the event-based approach adopted in this study. Currently, predictions from 2.5° above and below the L1 (or sub-Earth) point (with respect to the ecliptic plane) are used to calculate the latitudinal variation in speed across a WSA grid cell, which can be used as a proxy of the uncertainty in the predicted solar wind speed. When the HCS is relatively flat, large out-of-ecliptic speed gradients at 1 AU occur more frequently than when it is highly warped, which translates into large uncertainties in the predicted speed (see also Figure 8). This is because the transition region from slow to fast wind is typically very narrow, and when the current sheet is relatively flat, the sub-Earth point lies near this region more often than when it is highly inclined. However, our approach allows us to identify discrete events above, below, and in-ecliptic, allowing an estimation of the prediction uncertainty in the arrival times of HSEs. This value can be used to derive individual event association times (T_{GAP}).

[37] Figure 8 shows how the average value of the uncertainty in the HSE timing ($\langle |T_{GAP}| \rangle$) varies with solar cycle. Around solar minimum (1995, 1996) the uncertainty in HSE timing is large, whereas at solar maximum (2000, 2001) the uncertainty is significantly lower. The dashed line in Figure 8 shows the heliospheric current sheet (HCS) tilt angle (in radians) computed using a potential field model of Wilcox Solar Observatory data (as provided by http://soi.stanford.edu/wso/Tilts.html). As expected, there is an anticorrelation between $\langle |T_{GAP}| \rangle$ and the HCS tilt angle (i.e., a flat heliospheric current sheet is associated with large expected errors in the arrival times of HSEs). These HSE

timing errors are the most probable explanation for the large MSE between predicted and observed solar wind speed at solar minimum (section 3). If these timing errors become too large, they could produce missed and false events that are not associated with transient structures. Conversely, during periods of high solar activity (e.g., 1999–2001), the expected timing errors in the predicted HSEs due to HCS inclination are very small, but the MSE will still rise due to an increased number of false and missed HSEs arising from coronal complexity and transient structures. It is the combined MSE and HSE analysis approach that allows both quantification of prediction errors and insight into their cause, which can hopefully lead to improvements in future models.

[38] It should also be noted that during solar minimum the tilt angle and position of the HCS is largely determined by the strength of the polar magnetic fields, which are difficult to accurately measure.

6. Discussion

[39] This study has attempted to serve two purposes: testing of the Wang-Sheeley-Arge solar wind speed prediction scheme (which is to be used as the CISM baseline solar wind model) and demonstration of an event-based validation technique.

[40] The first half of the study used standard MSE analysis to characterize the predictive ability of WSA. Using 8-hour averaged data (which was shown to be adequate for analyzing solar wind speed), it was demonstrated that synoptic maps generated three days prior to the required L1 solar wind prediction produce optimal results. The MSE between WSA and data does not follow the solar cycle variation; maximum MSE occurs at solar minimum, with minimum MSE on the rise to solar maximum. However, visual inspection of the predicted and observed time series seems to contradict the MSE results; WSA captures the large-scale structure better at solar minimum than at solar maximum.

[41] Owing to potential problems with interpreting MSE, the second half of the study outlines an alternative, event-based validation technique. Methods for selecting and associating high-speed events (HSEs) from the observed and predicted time series were outlined. The hit, false, and missed HSEs can then be analyzed for timing, magnitude, and morphological differences. We found no significant offset in the predicted timing of HSEs but a slight bias for underestimating their magnitude, probably the result of fast ICMEs (not present in the WSA model) embedded within fast solar wind streams. Using this validation technique, the WSA model successfully predicted, 3 days in advance, 166 (~72%) of the 230 HSEs identified in the solar wind observations. As mentioned earlier, the missed events consist of two separate populations: a small-magnitude set that are probably too small to be picked up by the selection criterion along with a high-speed population that can be attributed to fast transient solar wind structures and which the WSA model is not expected to capture. The 36 falsely predicted HSEs are predominantly small-magnitude events similar to the nontransient set of missed events. Analysis of the morphology of the different HSE populations agrees with these interpretations. The simple empirical relations

used to derive proton density and temperature employed here (not part of the general WSA model) reproduce the structure of HSEs poorly, as the plasma dynamics at the stream interface is not captured. An MHD solar wind simulation using the WSA coronal solution as input may be able to better reproduce the observed signatures. Such improvements could be easily tested and quantified with the HSE analysis technique.

[42] It should be noted that in general, the uncertainties in the photospheric field observations are at least as large as those associated with the details of the models of the corona. Improvements in this model boundary condition, both in terms of the methods used to construct the synoptic maps and the corrections used for line saturation and polar fields, can have an enormous effect on the accuracy of coronal models, both empirical and physics based.

[43] Finally, out-of-ecliptic observations were used to calculate the uncertainty in the predicted HSE arrival time. The expected timing error is anticorrelated with the HCS tilt angle. Thus at solar minimum, small angular discrepancies in the coronal solution of WSA can lead to large differences in the predicted timings of HSEs at 1 AU. It is likely that such timing errors are the cause of the large MSE in solar wind speed at solar minimum. The use of higher spatial resolution magnetograms may help to better resolve the coronal structures, and more sophisticated corrections to the observationally challenging polar fields would significantly improve solar wind predictions at these times. The uncertainties in the HSE times reduce as the solar cycle progresses and the tilt of the HCS with respect to the ecliptic increases and the polar fields exert less control over the its position and morphology. However, additional factors affecting MSE must also be taken into consideration at periods close to solar maximum; the large-scale coronal magnetic field changes on shorter timescales, meaning a steady-state, potential field solution may not be sufficient. Furthermore, the increased occurrence of fast ICMEs at L1 will also cause the solar wind speed MSE to rise, as such transient events are not currently incorporated into WSA.

References

Altschuler, M. A., and G. Newkirk Jr. (1969), Magnetic fields and the structure of the solar corona, *Sol. Phys.*, *9*, 131.

Arge, C. N., and V. J. Pizzo (2000), Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates, *J. Geophys. Res.*, *105*, 10,465.

Arge, C. N., D. Odstrcil, V. J. Pizzo, and L. R. Mayer (2003), Improved method for specifying solar wind speed near the Sun, in *Solar Wind Ten*, edited by M. Velli, R. Bruno, and F. Malara, *AIP Conf. Proc.*, *679*, 190.

Arge, C. N., J. G. Luhmann, D. Odstrcil, C. J. Scrijver, and Y. Li (2004), Stream structure and coronal sources of the solar wind during the May 12th, 1997 CME, *J. Atmos. Sol. Terr. Phys.*, *66*, 1295.

Feynman, J., and S. B. Gabriel (2000), On space weather consequences and predictions, *J. Geophys. Res.*, *105*, 10,543.

Lepping, R. P., et al. (1995), The Wind magnetic-field investigation, *Space Sci. Rev.*, *71*, 207.

Lopez, R. E. (1987), Solar-cycle invariance in the solar wind proton temperature relationships, *J. Geophys. Res.*, *92*, 11,189.

McComas, D. J., S. J. Bame, S. J. Barker, W. C. Feldman, J. L. Phillips, P. Riley, and J. W. Griffee (1998), Solar wind electron proton alpha

monitor (SWEPAM) for the Advanced Composition Explorer, *Space Sci. Rev.*, *86*, 563.

McPherron, R. L., G. Siscoe, and C. N. Arge (2004), Probabilistic forecasting of the 3-h *Ap* index, *IEEE Trans. Plasma Sci.*, *32*, 1425.

Murphy, A. H. (1993), What is a good forecast?: An essay on the nature of goodness in weather forecasting, *Weather Forecasting*, *8*, 281.

Odstrčil, D., V. Pizzo, J. A. Linker, P. Riley, R. Lionello, and Z. Mikic (2004), Initial coupling of coronal and heliospheric numerical magneto-hydrodynamic codes, *J. Atmos. Sol. Terr. Phys.*, *66*, 1311.

Ogilvie, K. W., et al. (1995), SWE, a comprehensive plasma instrument for the Wind spacecraft, *Space Sci. Rev.*, *71*, 55.

Schatten, K. H. (1971), Current sheet magnetic model for the solar corona, *Cosmic Electrodyn.*, *2*, 232.

Schatten, K. H., J. M. Wilcox, and N. F. Ness (1969), A model of interplanetary and coronal magnetic fields, *Sol. Phys.*, *9*, 442.

Siscoe, G., D. Baker, R. Weigel, J. Hughes, and H. Spence (2004), Roles of empirical modeling within CISM, *J. Atmos. Solar Terr. Phys.*, *66*, 1469.

Smith, C. W., J. L'Heureux, N. F. Ness, M. H. Acuna, L. F. Burlaga, and J. Scheifele (1998), The ACE magnetic fields experiment, *Space Sci. Rev.*, *86*, 613.

Smith, Z. K., T. R. Detman, M. Dryer, C. D. Fry, C. C. Wu, W. Sun, and C. S. Deehr (2004), A verification method for space weather forcasting models using solar data to predict arrivals of interplanetary shocks at Earth, *IEEE Trans. Plasma Sci.*, *32*, 1488.

Spence, H., D. Baker, A. Burns, T. Guild, C.-L. Huang, G. Siscoe, and R. Weigel (2004), Center for integrated space weather modeling metrics plan and initial model validation results, *J. Atmos. Sol. Terr. Phys.*, *66*, 1491.

Wang, Y. M., and N. R. Sheeley Jr. (1990), Solar wind speed and coronal flux-tube expansion, *Astrophys. J.*, *355*, 726.

Weigel, R. S., D. N. Baker, J. Rigler, and D. Vassiliadis (2004), Predictability of large geomagnetic disturbances based on solar wind conditions, *IEEE Trans. Plasma Sci.*, *32*, 1506.

———————————

C. N. Arge, Space Vehicles Directorate, Air Force Research Laboratory, Hanscom Air Force Base, MA 01731, USA.

M. J. Owens, A. Pembroke, and H. E. Spence, Center for Space Physics, Boston University, Boston, MA 02215, USA. (mjowens@bu.edu)