AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2016SW001437

**Key Points:**
- Confusion matrices were calculated to assess the ability of the new PDF model to predict high-speed events
- Among the positive predictions, 60.4% are correct, 91.4% of the negative predictions are correct, and 20.3% of the peaks in the speed are found
- Ensemble predictions of high-speed events by the PDF model provide the forecast community with an interval of uncertainty on the prediction

**Correspondence to:**
C. D. Bussy-Virat,
cbv@umich.edu

**Citation:**
Bussy-Virat, C. D., and A. J. Ridley (2016), Twenty-four hour predictions of the solar wind speed peaks by the probability distribution function model, *Space Weather*, *14*, 861–873, doi:10.1002/2016SW001437.

Received 30 MAY 2016
Accepted 29 AUG 2016
Accepted article online 6 SEP 2016
Published online 27 OCT 2016



# Twenty-four hour predictions of the solar wind speed peaks by the probability distribution function model

**C. D. Bussy-Virat[1] and A. J. Ridley[1]**

[1]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan, USA

**Abstract** Abrupt transitions from slow to fast solar wind represent a concern for the space weather forecasting community. They may cause geomagnetic storms that can eventually affect systems in orbit and on the ground. Therefore, the probability distribution function (PDF) model was improved to predict enhancements in the solar wind speed. New probability distribution functions allow for the prediction of the peak amplitude and the time to the peak while providing an interval of uncertainty on the prediction. It was found that 60% of the positive predictions were correct, while 91% of the negative predictions were correct, and 20% to 33% of the peaks in the speed were found by the model. This represents a considerable improvement upon the first version of the PDF model. A direct comparison with the Wang-Sheeley-Arge model shows that the PDF model is quite similar, except that it leads to fewer false positive predictions and misses fewer events, especially when the peak reaches very high speeds.

## 1. Introduction

The space weather forecasting community has a particular interest in the prediction of transitions from slow to fast solar wind [*Wright et al.*, 1995]. A strong increase in the speed leads to enhanced geomagnetic activity that affects the near-Earth space environment. Such transitions can be associated with shocks or high-speed streams, which can be a consequence of a coronal mass ejection (CME) [*MacNeice*, 2009a].

Many models have been created and refined to predict the magnitude of the solar wind speed increase and its arrival time at the Earth. The Interplanetary Shock Propagation Model (ISPM) [*Smith and Dryer*, 1990] and the Shock Time of Arrival (STOA) model [*Dryer*, 1974] estimate the propagation time of the shocks from the Sun to the Earth and are based on type II meter wave burst data. The Hakamada-Akasofu-Fry (HAF) kinematic solar wind model has the ability to predict the speed, the solar wind density, and the interplanetary magnetic field (IMF). *Fry et al.* [2001] compared the predictions of shock arrival times made by HAF with the predictions made by the ISPM and the STOA models for 36 events between 1997 and 1999. They defined the shock arrival time as the time when the ram pressure rose above the background ram pressure by a certain threshold. They calculated the contingency table for the 36 events and showed that HAF (version 1) had comparable skills to the two shock models. *Smith et al.* [2009] evaluated the skills of HAF (version 2) during the declining phase of Solar Cycle 23 (2002 to 2006) for predictions of shock arrival times and compared them with predictions during the ascending phase (1997 to 2000 [*Fry et al.*, 2003]) and the maximum of the solar cycle (2000 to 2002 [*McKenna-Lawlor et al.*, 2006]). They predicted particular events and compared the shock arrival time and the strength (given by the ratio of the maximum pressure and background pressure) of the predicted shocks to the observed shock at 1 AU. Comparing the number of hits, misses, and false positives, they concluded that the predictions were a bit improved for the ascending and maximum conditions, as opposed to the solar minimum conditions. *McKenna-Lawlor et al.* [2006] compared predictions of 173 shocks during the maximum phase of Solar Cycle 23 by STOA, ISPM, and the HAF models (version 2). The accuracy of a model can be defined as the sum of hits (i.e., the number of times the model correctly identified shocks) and correct nulls (i.e., the number of times the model correctly identified nonshock time periods) divided by the total number of predictions times 100. In that study, the STOA, ISPM, and HAF v2 models received scores of 54%, 60%, and 52%, respectively, for a hit window (size of the window in which the event was allowed to be observed) of $\pm$24 h. The number of events classed as hits decreased when the window size was $\pm$12 h, so the accuracies were lower: 44%, 54%, and 42%, respectively. For a window size larger than $\pm$24 h, the results were not relevant, because the shock might not have been related to the specific solar event. However, the accuracy is the not the best measure to assess the ability of a model in predicting rare events: a high accuracy could mainly

be due to a large number of correct nulls, which is not crucial for forecasters, who are mostly interested in hits. The threat score, defined as the number of hits divided by the sum of hits, misses (i.e., the model did not predict an event that actually occurred), and false positives (i.e., the model predicted an event that actually did not occur) are more suitable measures. Similar to the accuracy, it does not take into account the number of correct nulls [*Buizza*, 2001].

*Vandegriff et al.* [2005] introduced a model that predicted shock arrival times. The method relied on the energetic particle intensity data. It was able to predict the shocks 24 h in advance with an accuracy of 8.9 h, and 12 h in advance with an accuracy of 4.6 h.

*Kim et al.* [2010] developed a model to predict the occurrence of geomagnetic storms and their strengths, which were represented by the *Dst* index. The prediction was based on the CME parameters (southward orientation of the magnetic field in the CME source region, asymmetry of the CME shape, and CME source location). They found that the predictions of the *Dst* index were far from the observed index, but the model showed good skill at predicting when a geoeffective event may occur, based on when the *Dst* index calculated by the model became less than −50 nT.

Finally, the Wang-Sheeley-Arge (WSA) model was tested for the predictions of high-speed enhancements. These correspond to abrupt transitions from slow to fast solar wind. *MacNeice* [2009a] presented an algorithm to characterize this type of transition and tested the performance of the WSA model in predicting these events. He answered the question "[I]f the WSA model predicts or does not predict a high-speed event in the next 24 h, what is the probability that a high-speed event will or will not occur?" [*MacNeice*, 2009a] and found that 29% of the positive forecasts and 89% of the negative forecasts were accurate, showing an improvement compared with the study by *Owens et al.* [2005]. *MacNeice* [2009b] provided a similar study but used other sources of magnetograms. Averaging over the three sources of magnetograms, he found that a positive/negative forecast was correct 17%/94% of the time.

Forecasters need to be able to understand the level of confidence they can place in the predictions. A single value for the prediction of the magnitude of the speed and the time when the increase occurs needs to be associated with an estimation of the error made on the prediction. Ensemble forecasts can quantify this uncertainty. Few studies have been published regarding this issue. The WSA model, coupled to the ENLIL model, has been used to start to perform ensemble modeling of the background solar wind at the Community Coordinated Modeling Center since 2009 by taking the average of the two models' output for each of the solar wind parameters [*Taktakishvili et al.*, 2009]. *Emmons et al.* [2013] coupled the WSA, ENLIL, and CONED models to create ensemble forecasts for 15 CMEs. They compared the predicted distributions of the propagation time and the *Kp* index (obtained from 100 sets of CME cone parameters derived from the CONED model) with the observations and found that out of the 15 events, 5 (33%) had propagation times that were within the ensemble average plus or minus one standard deviation and 8 (53%) had propagation times within the range of the ensemble, while 10 (67%) had maximum *Kp* indices that were within the range of the ensemble. *Mays et al.* [2015] found that the arrival time was within the range of the ensemble for 8 out of 17 events (47%). They also showed that the accuracy of the predictions depends on the CME input parameters and that the ensembles do not sample a wide enough spread in these parameters. *Cash et al.* [2015] used an ensemble approach to study the effect of the input parameters. They considered the initial speed, the angular width, the direction, and the ambient solar wind background and put in evidence their influence on the predicted arrival time of a CME that occurred in July 2012. Goddard's Space Weather Laboratory received support in 2012 to implement such ensemble forecasts.

This study presents an improvement of the probability distribution function (PDF) model [*Bussy-Virat and Ridley*, 2014] for predictions of transitions from slow to fast solar wind. The older model did not have this ability as it was based on all solar wind data over 15 years rather than focused only on solar wind speed increases. The model now is able to predict these transitions when there is an increase in the slope of the solar wind speed and performs better when transitions are associated with strong variations in the magnitude of the IMF or the density. This includes predictions of shocks. The way that this study characterized such transitions was very similar to the one presented in *Owens et al.* [2005] and *MacNeice* [2009a]. The model presented here predicts several features of the peak in speed, including the amplitude of the peak as well as the time when the maximum occurs. It also provides ensembles of predictions of each peak, which enables forecasters to have an estimation of the uncertainty in the prediction, or the possible spread in the characteristics. We compare the predictions with the observations made by the Advanced Composition Explorer satellite and compare the

accuracy of the model with the previous studies that predicted such transitions, in particular those made by the WSA model [*MacNeice*, 2009a, 2009b; *Owens et al.*, 2005, 2008; *Emmons et al.*, 2013].

## 2. Methodology

### 2.1. Old PDF Model

In *Bussy-Virat and Ridley* [2014], the PDF model was introduced as a combination of two models. The first one uses the speed approximately one solar rotation ago (OSRA), the optimum lag being calculated using a temporally weighted normalized root-mean-square comparison (OSRA model). The second model uses PDFs based on the current speed and the slope in the 12 previous hours. The predicted speed $i$ hours from "now" is calculated as

$$v_{\mathrm{pred},i} = a * v_{\mathrm{PDF},i} + b * v_{\mathrm{OSRA},i} \tag{1}$$

where $v_{\mathrm{PDF},i}$ is the predicted speed using the PDFs and $v_{\mathrm{OSRA},i}$ is the predicted speed using the data one solar rotation ago. $a$ and $b$ are linear weights that can vary to optimize the prediction ability and throughout the prediction ($a + b = 1$).

The OSRA model relies on the lag used to find the optimum time when the previous solar rotation started. It was found in *Bussy-Virat and Ridley* [2014] that this lag is highly variable and hard to predict.

The main flaw of the PDF model is that it was not able to predict transitions from slow to high solar wind speed. The cause was that $v_{\mathrm{PDF},i}$ was based only on the current speed and the slope in the 12 previous hours. Statistically, the PDF model showed that the majority of times when the slope of the solar wind speed was positive the speed would start to decrease after a few hours. What the PDF model missed is that when the slope was above certain levels, the speed would rise to a peak above 500 km/s and be considered a high-speed event. Therefore, when a shock reached the Earth, the PDF model did not predict the speed reaching above the high-speed event cutoff of 500 km/s. Since CMEs often happen randomly and are not often associated with a similar event that occurred roughly 27 days ago, the OSRA model should not be used for the prediction of such transitions. Indeed, even when an event repeats over one rotation, *Bussy-Virat and Ridley* [2014] showed that the time between the two events is not exactly one solar rotation but that there is an unpredictable and highly variable lag. Consequently, the present study focuses on predictions using only PDFs, which means that $a = 1$ and $b = 0$ in equation (1).

### 2.2. Identification of High-Speed Events

An algorithm was created to identify transitions from slow to fast solar wind, named high-speed enhancements (or high-speed events). It is similar to the one presented in *MacNeice* [2009a] because it detects any sharp transition above a certain speed gradient. Figure 1 (left) shows an example of a single peak and key times within the peak for reference. The peak-finding algorithm can be summarized as follows.

1. The solar wind speed is averaged with an 11 h running average to filter out the high-frequency variability of the solar wind.
2. The algorithm flags any increase in speed by more than 50 km/s in less than 24 h, noting the time in which this occurs as $t_{\mathrm{beginning}}$.
3. The end of the peak is identified as the time when the speed decreased by 50 km/s in less than 36 h and is noted as $t_{\mathrm{end}}$. This less sharp gradient is a consequence of a study (not shown here) that showed that the slope is typically less sharp in the descending phase of the peak than in its ascending phase.
4. The time of the maximum of the peak is noted as $t_{\mathrm{peak}}$.

The algorithm allows for any modification in the choice of these gradients (50 km/s in 24 h can be changed to 70 km/s in 30 h, for example) as well as setting a threshold on the minimum value of the maximum speed. Examples of peaks found by the automated algorithm are shown in Figure 1 (right). Vertical red lines are drawn when peaks are found by the algorithm using a minimum gradient of 50 km/s in 24 h and adding a threshold at 500 km/s on the minimum value of the peak. We can notice that the algorithm finds every peak in this 3 month period.

For this study, peaks with gradients larger than 50 km/s in 24 h were investigated.

### 2.3. New PDF Model

Forecasters need to know when the high-speed events will occur as well as the magnitudes. In this regard, several approaches were investigated. The first idea was to relate the amplitude of the peak (defined as the

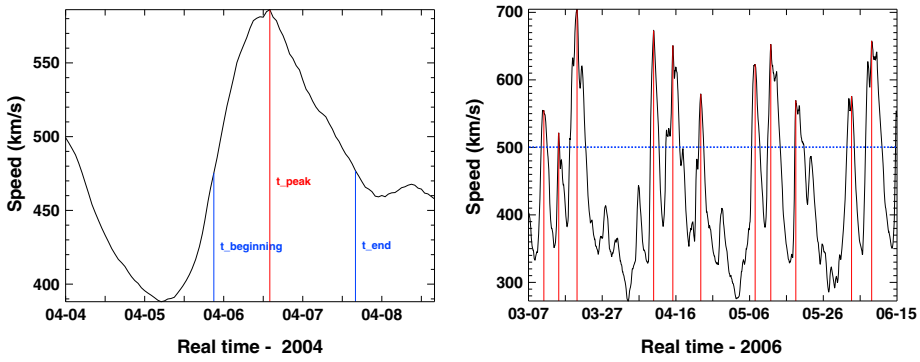


**Figure 1.** (left) Example of the detection of a peak in April 2004 using a minimum gradient of 70 km/s in 24 h. (right) Example of the detection of peaks in March–June of 2006 adding a threshold at 500 km/s for the maximum value of the peaks.

difference between the minimum and maximum speed in the peak) to the maximum slope in the ascending phase of the peak (noted $s_{max}$). Figure 2 shows the distribution of the amplitude of 761 peaks detected by the algorithm as a function of the maximum slope. The slope is calculated with a step of 2 h

$$\text{slope}_i = \frac{v_{i+1} - v_{i-1}}{2} \tag{2}$$

$$s_{max} = \max([\text{slope}_0, \text{slope}_1, \text{slope}_2, \ldots, \text{slope}_n]) \tag{3}$$

where $\text{slope}_i$ is the slope $i$ hours after the beginning of the peak (in km/s/h), $n$ is the number of hours until the maximum of the peak, and $v_{i\pm1}$ is the speed $i \pm 1$ h after the beginning of the peak.

The correlation between the maximum slope and the amplitude was 0.71 with a mean absolute error between the data and a linear fit of 55 km/s. Distribution functions of the amplitude were made for each bin representing a range of slopes of 2 km/s/h. Note that the minimum value on the $x$ axis of Figure 2 is 4.77 km/s/h: every peak found by the algorithm between 1995 and 2012 had a maximum slope greater than 4.77 km/s/h. The PDF model uses these distributions for slopes below 10 km/s/h. Although the exact value of 10 km/s/h is arbitrary, for slopes above 10 km/s/h, events get rarer so not enough data can be used to make such accurate distributions. Therefore, the PDF model uses Gaussian distributions with a mean value equal to the amplitude given by the linear fit (blue curve in Figure 2) and a full width at half maximum of 110 km/s. Figure 3 shows the distribution of the peak amplitudes given a maximum slope of 6 km/s/h (left) and 20 km/s/h (right). From these distributions, the median and the deviations from this median (10%, 25%, 75%, and 90%) were determined.

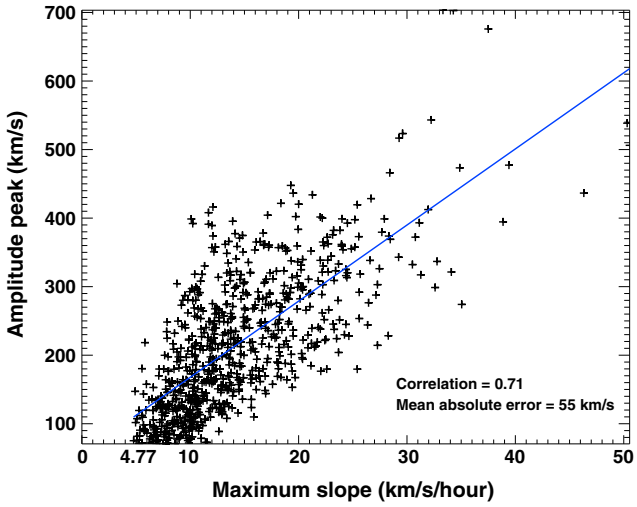


**Figure 2.** Amplitude of a peak in function of the maximum slope.

In the first case, the distribution of the predicted amplitudes has a median at 88 km/s, and in the second case the median is at 275 km/s. This higher value can be explained by the fact that the maximum slope detected in the ascending phase is almost 4 times higher than in the first case. When the model detects that a peak is occurring, the maximum slope is determined and the predicted peak amplitude (noted $\Delta v$) of the solar wind speed is derived, given Figure 2 and the distribution functions described above, to specify the uncertainty on the predicted amplitude. This method is used to determine the speed of the peak ($v_{peak}$) and the uncertainty in the speed of the peak.

The time of the peak is important and can be deduced in two different ways. The first

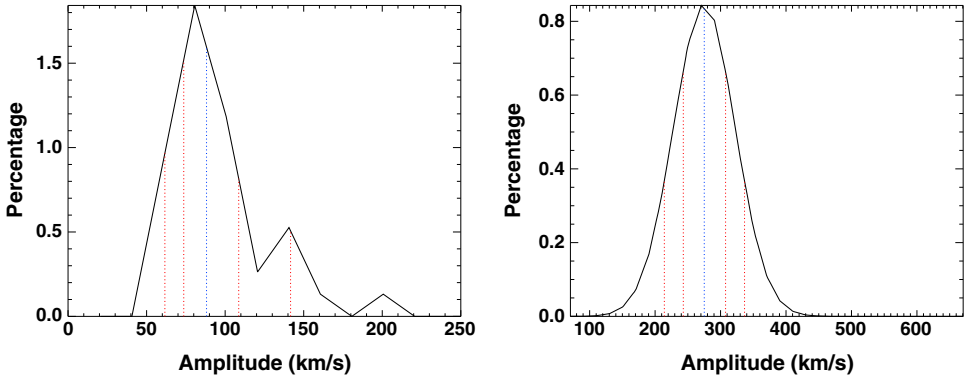


**Figure 3.** Example of a (left) PDF and a (right) Gaussian distribution of the amplitude of the peak for a maximum slope of 6 km/s/h (Figure 3, left) and 20 km/s/h (Figure 3, right): the red dash lines represent the 10%, 25%, 75%, and 90% quantiles, and the blue dash line represents the median.

relies on the solar wind speed continuing to increase at the maximum measured rate to the peak amplitude, as described above. The time (further noted $t_{75\%}$, the meaning of which will be described below) can thus be derived by dividing the amplitude by the slope $s_{max}$: $t_{75\%} = t_{current} + (v_{peak} - v_{current})/s_{max}$. However, it was noted that this time occurs earlier than the actual time of the peak. The speed typically increases quickly, and then the gradient decreases significantly, resulting in a peak that lasts longer than expected if only the maximum slope is considered. Therefore, a second method was explored to find the actual time of the peak.

This method relies on the PDF of the time it takes for the speed to increase from the beginning of the peak (at $t_{beginning}$) to the maximum of the peak (at $t_{peak}$). A study not presented here showed that although there is a relationship between the amplitude of the peak and the maximum slope (Figure 2), there is not a clear relationship between $t_{peak} - t_{beginning}$ and the maximum slope. The distribution of times for the 761 peaks is shown in Figure 4. It is sharp enough, and the most probable value is high enough for the prediction based on this method to be valuable. The median of the distribution is 23 h, meaning that 50% of the time, the speed will take less than 23 h to increase from the speed at $t_{beginning}$ to the speed at $t_{peak}$. Therefore, the predicted time from the beginning to the maximum of the peak (noted $\Delta t$) is specified as 23 h, with an uncertainty distribution as specified in Figure 4. This method is used to determine the time of the peak ($t_{peak}$) and the uncertainty in the time of the peak.

Combining the two methods described above and in Figure 5, $v_{peak}$ ($= v_{current} + \Delta v$) occurs at $t_{peak}$ ($= t_{current} + \Delta t$). When the rapid rise ends (at $t_{75\%}$), the peak is assumed to have reached 75% of its maximum speed: $v_{75\%} = v_{peak} - 0.25 * (v_{peak} - v_{current})$. Two polynomial expressions of order 3 link the speeds $v_{current}$ to $v_{75\%}$ (at $t_{75\%}$) and $v_{75\%}$ to $v_{peak}$ (at $t_{peak}$), so that the slopes before and after $t_{75\%}$ are continuous and the slope at $t_{peak}$ is 0, which is the definition of a maximum. Specifically, the following algorithm is used:

1. The maximum slope $s_{max}$ is calculated in the ascending phase (from $v_{data,min}$ to $v_{current}$). Using Figure 2 and the distribution functions described in the first method, the PDF model derives $\Delta v$ to find $v_{peak}$: $v_{peak} = v_{data,min} + \Delta v$.
2. The time and speed when the rapid rise ends are calculated from $v_{peak}$ and $s_{max}$: $t_{75\%} = t_{current} + (v_{peak} - v_{current})/s_{max}$ and $v_{75\%} = v_{peak} - 0.25 * (v_{peak} - v_{current})$. The first polynomial expression calculates the speeds that occur between $t_{current}$ and $t_{75\%}$ (blue dashed line in Figure 5).
3. The time $t_{peak}$ when $v_{peak}$ is reached is derived from $t_{beginning}$ (increase by 50 km/s in less than 24 h) and the PDF of Figure 4. The second polynomial expression calculates the speeds that occur between $t_{75\%}$ and $t_{peak}$ (red dashed line in Figure 5).
4. Once the ascending phase has been predicted, the model uses the methodology described in *Bussy-Virat and Ridley* [2014] to predict the descending phase of the peak (black dashed line).

Examples of predictions made by the model are shown in Figure 6. For example, in the top left graph the actual speed (black line) increased from 440 km/s to 600 km/s in less than 15 h. The peak was predicted (blue line) to occur in 24 h from the start time with an amplitude of 110 km/s (above the current solar wind speed), which corresponded to an underestimation of the peak speed. The descending phase of the actual speed followed the predicted descending phase very closely. Moreover, the actual speed was in between the 25% and 75%

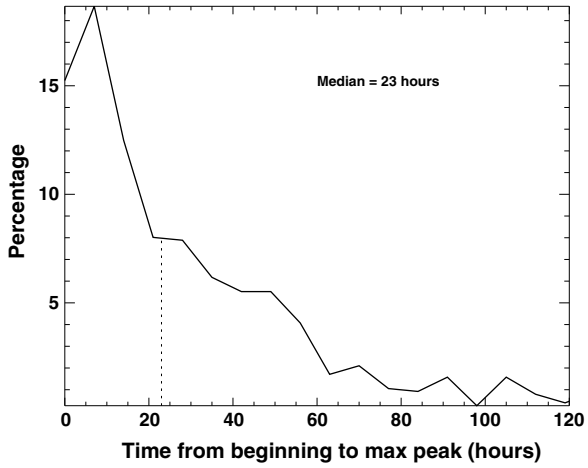


**Figure 4.** Histogram of the time from the beginning of the peak to its maximum.

quartiles 108 h out of the 120 h prediction horizon and in between the 10% and 90% deciles 115 h out of the 120 h. The other examples illustrate the variety of results that the PDF model predicts, as well as the actual solar wind speed data.

### 2.4. Comparison With Peaks in the IMF and the Density

A sudden large increase in the speed could be due to either a high-speed stream or could be the result of a shock in front of a CME that reaches the Earth environment. In such events, the IMF magnitude and density should also increase before the velocity peaks, since speed increases can act as a "snow plow" on the density and IMF. The same algorithm that characterized a peak in the speed was also applied to find peaks in the IMF magnitude and the density. The thresholds on the IMF and density gradients were 4 nT/d and 8 $cm^{-3}$/d, respectively. A threshold was added on the minimum value for the peaks in IMF and density: 6 nT and 10 $cm^{-3}$, respectively (i.e., any peak under 6 nT and 10 $cm^{-3}$ was not considered). Once peaks in the solar wind density and IMF magnitude were found, they were compared with peaks in the solar wind speed.

Figure 7 (left) shows the histogram of the lags between the maxima of a peak in the IMF magnitude and the maxima in the speed, with a bin size of 3 h. The maximum in the speed typically occurs 10–30 h after the maximum of the peak in IMF magnitude. These results imply that if the algorithm finds a peak in the IMF magnitude, then it can be expected that it will find a peak in the speed several hours after. Specifically, the most probable time delay is 11 h, while the median delay is 20 h. A similar analysis shows that peaks in density occur before peaks in speed as well.

Figure 7 (left) also indicates that peaks in the speed can be found more than 40 h after the peak in the IMF magnitude. However, this might be due to random peaks that are not related to the shock. To reduce the effect of limiting the peak in speed to being above a given threshold, the times when both peaks (IMF and speed) start to increase can be compared by determining the lag $t_{beginning,speed} - t_{beginning,IMF}$, where $t_{beginning}$ is the time in which the gradient in IMF magnitude or speed has crossed a given threshold. The histogram of these time delays is shown in Figure 7 (right). The median value of the histogram is now at 10 h, but the curve drops at 15 h such that 80% of events have delays less than 30 h. This means that most of the time, a rapid increase in the IMF magnitude will be followed by a rapid increase in the speed in the following day (most likely in the next 10 h). Any increase in speed beyond 1 day most likely cannot be attributed to the shock but may correspond to a random increase in speed.

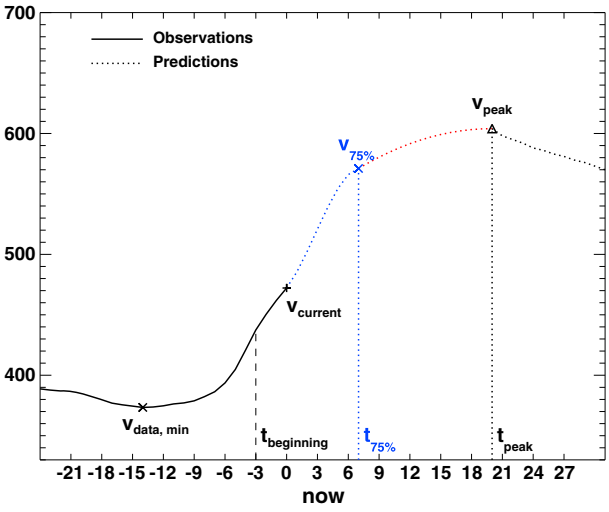


**Figure 5.** Calculation of the speeds from the current time to the time the speed is predicted to reach its maximum value.

Figure 8 (left) shows an example of a high-speed event that occurred after a peak in the IMF. At $t_{beginning,speed}$ (marked by a vertical blue line), the IMF magnitude (solid red line) had already increased by large values. However, Figure 8 (right) shows an example

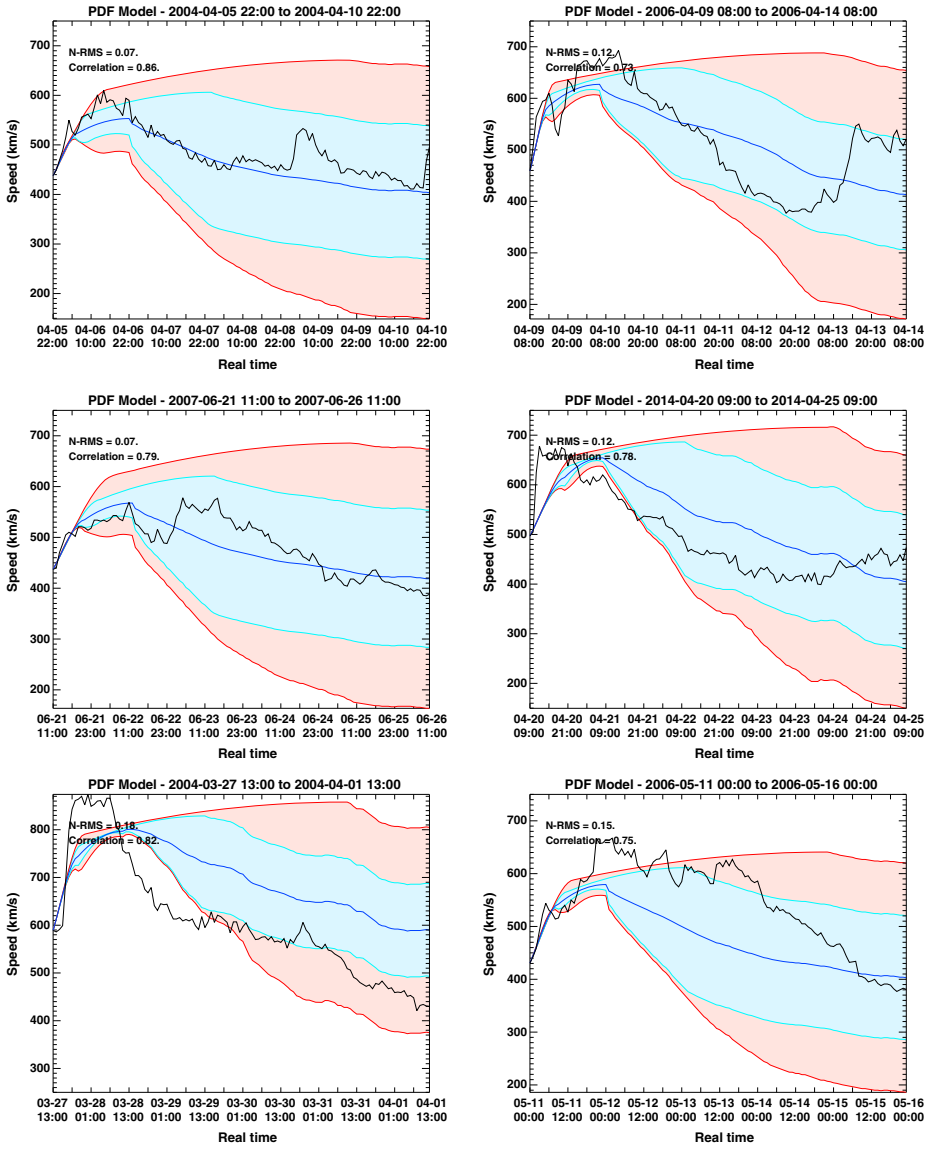


**Figure 6.** Examples of predictions with the new PDF model. The black line is the actual speed; the blue line is the median of the PDFs; the light blue lines are the 25% and 75% quartiles (interquartile range); and the red lines are the 10% and 90% deciles.

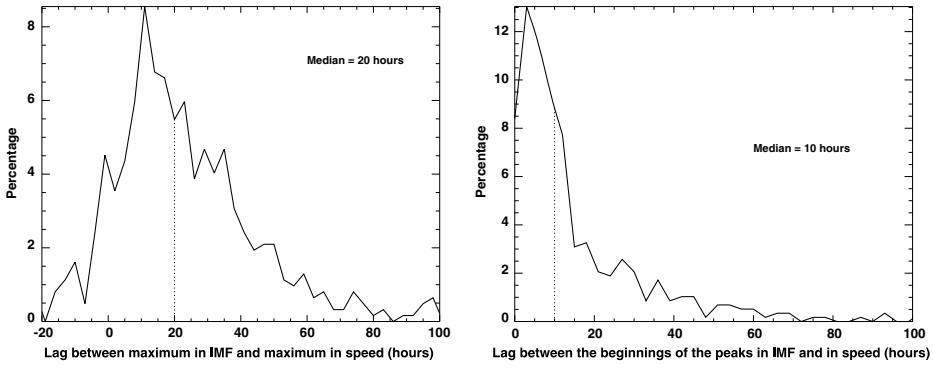


**Figure 7.** (left) Histogram of lag between a peak in IMF and a peak in the speed. (right) Histogram of lag between the beginning of a peak in the IMF and the beginning of a peak in the speed.

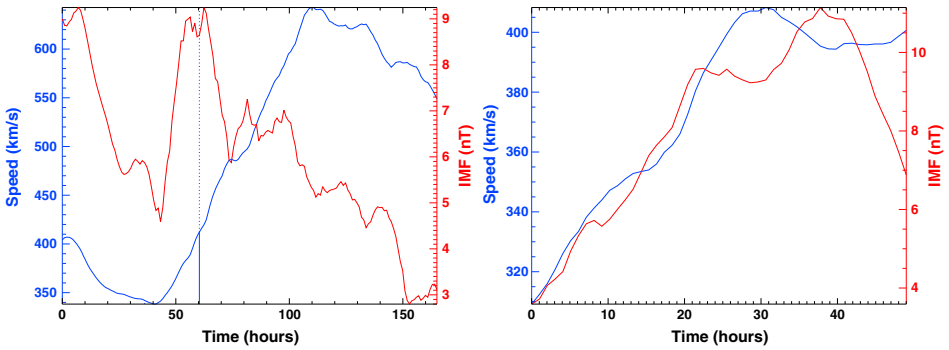


**Figure 8.** (left) Example of a peak in the speed after a peak in the IMF; (right) the peak in IMF occurs after the peak in the speed.

where the peak in the IMF occurred after the peak in the speed. As illustrated in Figure 7 (left), this situation happens much more rarely.

### 2.5. Assessment of the New PDF Model

To assess the ability of the new PDF model to predict high-speed events, contingency tables were calculated. A typical contingency table presents the number of events that were correctly predicted or true positives (i.e., the model predicted an event and there actually was an event), the number of false positives (i.e., the model predicted an event that actually did not occur), the number of missed events or pure misses (i.e., the model did not predict an event that actually occurred), and the number of correct nonevents or true negatives (i.e., the model predicted there would be no event and there actually was no event).

The sensitivity, also called the probability of detection, of a model represents the probability that an event that actually occurred was correctly predicted by the model and is defined as

$$\text{sensitivity} = \frac{\#\text{ true positives}}{\#\text{ events}} = \frac{\#\text{ true positives}}{\#\text{ true positives } + \#\text{ pure misses}} \tag{4}$$

The positive predicted value (PPV) is the probability that a predicted event actually occurred and is defined as

$$\text{PPV} = \frac{\#\text{ true positives}}{\#\text{ predictions of an event}} = \frac{\#\text{ true positives}}{\#\text{ true positives } + \#\text{ false positives}} \tag{5}$$

The negative predicted value (NPV) is the probability that a negative forecast was correct and is defined as

$$\text{NPV} = \frac{\#\text{ true negatives}}{\#\text{ predictions of a nonevent}} = \frac{\#\text{ true negatives}}{\#\text{ true negatives} + \#\text{ pure misses}} \tag{6}$$

For the three numbers, the ideal value is 100%.

To get a complete view of the quality of the model, all three of these numbers should be calculated and compared. For example, a model that does not ever predict any peaks never leads to any false positives but has 0 true positive so the sensitivity is 0%. A model that always predicts a peak in the next 24 h would have a sensitivity of 100%, but its PPV would tend to 0% since it would have a very large number of false positives. This is why forecasters are interested in models that present a good balance between these numbers so that they are able to give correct predictions (positive and negative) as often as possible. Finally, the false alarm rate (= 1 − PPV) is a measure that forecasters are also interested in, as it corresponds to the probability that a predicted event did not actually occur.

## 3. Results and Discussion

The improved PDF model was run between 1995 and 2012 to make 1 day forward predictions, moving by steps of 3 h between each of the predictions. This enabled the testing of the ability of the model to predict the speed at any phase of the high-speed event. For each prediction, a similar question was asked as the one by *MacNeice* [2009a, 2009b]: "[I]f the model predicts/does not predict a high-speed event in the next 24 h,

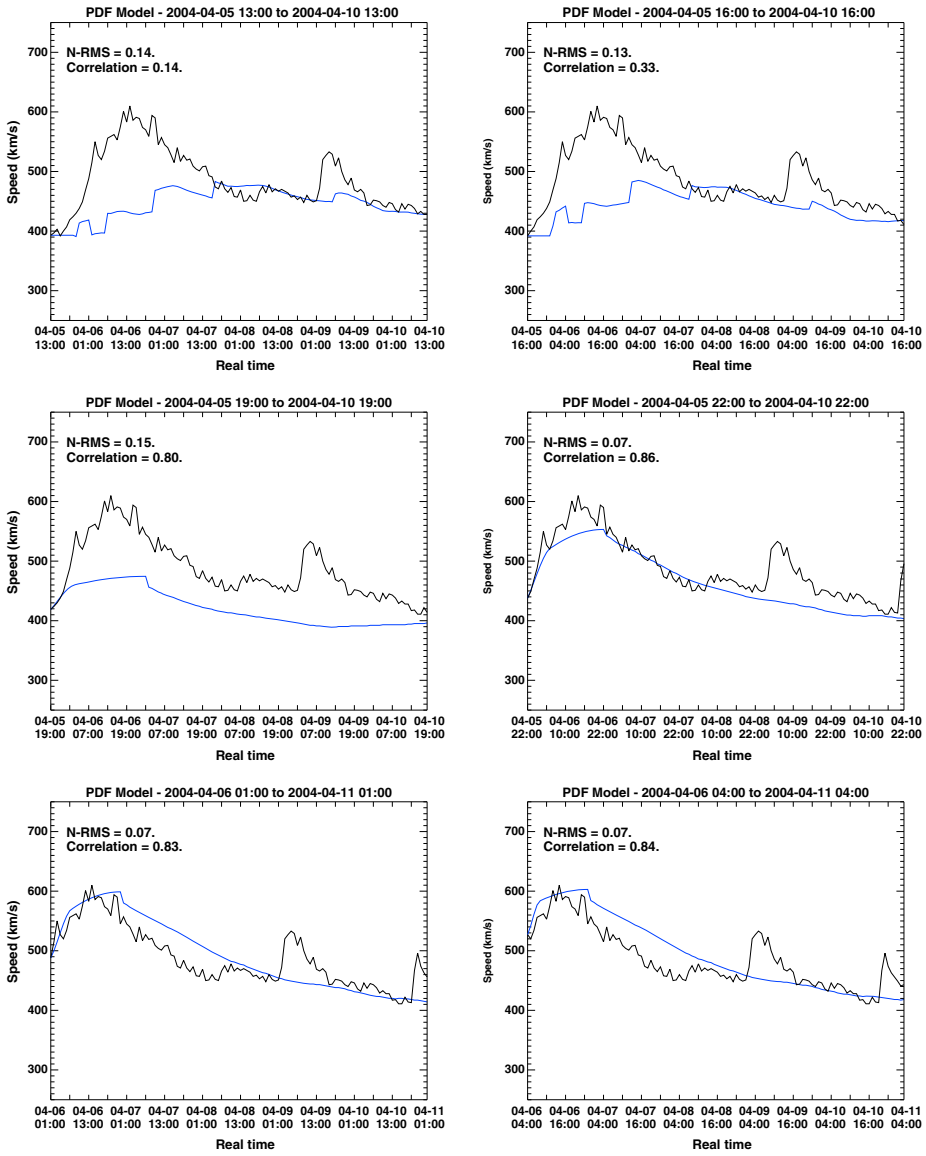


**Figure 9.** Example of the assessment of the ability of the PDF model to predict high-speed events for consecutive forecasts. The black line represents the actual speed and the blue line the predictions made by the PDF model.

does a high-speed event actually occur?" However, answering this question gives only the PPV and the NPV. To determine the sensitivity, a further question must be asked: "[I]f there is actually a high-speed event, what is the probability that the model predicted it?" To answer these questions, we used the definition of the high-speed event given in the previous section. Specifically, an event, predicted ($v_{pred}$) or observed ($v_{data}$), is categorized as a high-speed event if the speed now or in the next 24 h has a gradient greater than 50 km/s/d and goes over 500 km/s. If one of these two conditions is not met, then it is not a high-speed event.

Figure 9 illustrates how this analysis was made. The six graphs represent predictions of the same (observed) high-speed event at different times, moving by steps of 3 h between each prediction: 13 UT, 16 UT, 19 UT, 22 UT, 01 UT (next day), and 04 UT (next day). At 13 UT and 16 UT, the event had not yet started so the slope was not higher than the threshold at 4.77 km/s/h. According to the distributions from Figure 2, every peak found by the algorithm between 1995 and 2012 had a maximum slope greater than 4.77 km/s/h. Therefore, no peak was predicted and the two predictions were pure misses (bottom left of the contingency table). At 19 UT, the current speed had increased sufficiently for the PDF model to predict a peak, but the gradient was not high enough for the prediction of the maximum speed to be close to the actual peak. Therefore, the PDF model, although predicting a peak, underestimated the amplitude of the increase and was a pure miss as well.

**Table 1.** Contingency Table for Any Types of High-Speed Events (HSE: High-Speed Event)

| | | Actual Solar Wind | |
|---|---|---|---|
| | | HSE | NO HSE |
| PDF | HSE | 798 | 523 |
| Predictions | NO HSE | 3,125 | 33,401 |

At 22 UT, 01 UT (next day), and 04 UT (next day), the PDF model correctly predicted the event: both the amplitude and the time of the peak closely match the actual speed, and the descending phase follows the shape of the actual speed. However, recall that the event is categorized as a high-speed event if the current speed is below 500 km/s and increase to a value above 500 km/s. This implies that only the predictions at 22 UT and 01 UT (next day) are taken into account in the contingency table and counted as true positives (top left of the contingency table). Ignoring predictions when the current speed is already higher than this threshold ensures that the assessment of the model reflects a meaningful forecast quality and not simply observational assist. To conclude, of the six times, three were pure misses, two were true positives, and the last event was ignored because the current speed had already increased at a value higher than 500 km/s so the forecast was not meaningful.

### 3.1. Predictions of High-Speed Events

The contingency table for the results of the predictions from 1995 to 2012 data is presented in Table 1. Out of the 3923 events that occurred, 798 were predicted by the PDF model (note that several predictions can correspond to the same high-speed event but at different times of the ascending phase of the peak, as shown in Figure 9). This corresponds to a sensitivity of 20.3%. There were 523 false positives, which implies that the PPV was 60.4%. The 3125 of the 36,526 negative predictions were incorrect, which means that the NPV was 91.4%, so 8.6% were pure misses. If the past performance is an indication of future performance, a positive forecast by the PDF model will be correct 60.4% of the time, while a negative forecast will be correct 91.4% of the time. The PDF model should be able to predict 20.3% of the high-speed events.

On the other hand, the performance of the PDF model changes in the presence of IMF and density peaks. While the model itself does not change at all, if an operator of the model sees a peak in the IMF and density, the operator can be more confident in the PDF model. This is shown by considering only time periods that include peaks in both the IMF and the density. The results of these events are shown in the contingency table in Table 2.

A positive forecast was correct 69.5% of the time, while a negative forecast was correct 72.0% of the time. The PDF model was able to predict 33.6% of the high-speed events. Although the NPV decreased (91.4% to 72.0%), the PDF model missed fewer high-speed events (the sensitivity increased from 20.3% to 33.6%) and led to fewer false positives under these conditions (the PPV increased from 60.4% to 69.5%).

### 3.2. Direct Comparison With the WSA Model

*MacNeice* [2009b] provided results of high-speed event predictions made by the WSA model. He found that a positive forecast was correct 17% (= PPV) of the time and that a negative forecast was correct 94% (= NPV) of the time. However, we chose to make a direct comparison between the PDF and the WSA models using the same definition of the high-speed event described earlier in the text for both models, since it was unclear if exactly the same things were being compared in the *MacNeice* [2009b] study as here. In addition to comparing the PPV and the NPV of the two models, their sensitivities were also compared, which is of equal importance for the space weather forecasting community as the PPV and the NPV. The WSA model was run for 1 year

**Table 2.** Contingency Table for High-Speed Events Associated With IMF and Density Peaks (HSE: High-Speed Event)

| | | Actual Solar Wind | |
|---|---|---|---|
| | | HSE | NO HSE |
| PDF | HSE | 137 | 60 |
| Predictions | NO HSE | 271 | 696 |

**Table 3.** Comparison With the WSA Model for Different Sets of Values for *gd* and $v_{th}$[a]

| | 50/500 | | 70/500 | | 50/600 | | 70/600 | |
|---|---|---|---|---|---|---|---|---|
| | PDF | WSA | PDF | WSA | PDF | WSA | PDF | WSA |
| Sensitivity (in %) | 19.1 | 22.0 | 15.5 | 17.8 | 20.9 | 10.5 | 18.8 | 10.0 |
| PPV (FAR) (in %) | 66.7 (33.3) | 17.9 (82.1) | 62.3 (37.7) | 13.6 (86.4) | 31.0 (69) | 12.7 (87.3) | 29.4 (70.6) | 11.4 (88.6) |
| NPV (in %) | 91.1 | 90.5 | 91.9 | 91.2 | 97.3 | 97.0 | 97.5 | 97.2 |

[a]The first line corresponds to the value of *gd* followed by the value of $v_{th}$: $gd/v_{th}$ (FAR: False Alarm Rate).

in 2011 to make 1 day ahead predictions at a temporal resolution of 24 h. To link the current speed to the speed 24 h ahead and to provide the same temporal resolution as the PDF model, a cubic spline interpolation was used.

Multiple parameters that characterize a high-speed event were used to explore their influence on the model results: the gradient in 24 h (noted *gd*, equal to 50 km/s/d so far) and the threshold (noted $v_{th}$, equal to 500 km/s so far) between slow and fast solar wind. Four sets of values were chosen from different combinations of $v_{th}$ and *gd*: $v_{th} = 500$ km/s and $v_{th} = 600$ km/s with $gd = 50$ km/s/d and $gd = 70$ km/s/d. As *MacNeice* [2009a] concludes, the algorithm that defines a high-speed event can have a significant influence on the results. It is therefore important to apply the selection criteria and evaluation matrix to both models.

The sensitivity, the PPV, and the NPV for both the WSA and the newly modified PDF model are presented in Table 3. *MacNeice* [2009b] used a value for *gd* of 50 km/s/d and a value for $v_{th}$ of 500 km/s. For the corresponding column of Table 3, a sensitivity of 22%, a PPV of 17.9% (corresponding to a false alarm rate of 82.1%), and an NPV of 90.5% for the WSA model were found, very similar to values found in *MacNeice* [2009b]. For the PDF model, the sensitivity, the PPV, the false alarm rate, and the NPV were 19.1%, 66.7%, 33.3%, and 91.1%, respectively, close to the values found for the study over 17 years. The sensitivity of the WSA model (22%) is slightly better than the sensitivity of the PDF model (19.1%). However, the WSA model leads to 2 to 3 times as many false positives than the PDF model (i.e., the false alarm rate is 2.5 times higher for the WSA model). The PDF model leads to 4 times more true positives than the WSA model (i.e., the PPV is 4 times higher for the PDF model). Both NPVs are comparable, which means that when either model says that there is not an event, it is true about 91% of the time.

Changing *gd* to 70 km/s/d (leaving $v_{th}$ as 500 km/s) does not change the results by much. It decreases the sensitivity and increases the false alarm rate by about four percentage points for both models, implying that both models have slightly more false positives.

However, when changing $v_{th}$ from 500 km/s to 600 km/s (keeping *gd* at 50 km/s/d, for example), the sensitivity of the WSA model drops by a factor of 2. The false alarm rate of both models is also affected, particularly for the PDF model, in which this increased $v_{th}$ leads to twice as many false positives (i.e., the false alarm rate increases by a factor of 2). The NPV is better for both models. This may be due to the fact that since $v_{th}$ has increased, there are fewer events that are counted as high-speed events in the observations, so the number of true negatives increases for the same number of pure misses. These results indicate that the PDF model most likely overpredicts the solar wind speeds during a peak.

Whatever the value of *gd* and $v_{th}$ are, the PDF model appears to be more reliable than the WSA model for predictions of high-speed events. Because the false alarm rate is smaller for the PDF model, the PDF model leads to fewer false positives than the WSA model. If the threshold for the fast solar wind speed is set at 600 km/s, then the number of misses is also smaller for the PDF model. This can be seen by a higher (by a factor $\sim 2$) sensitivity. The NPV, which represents the ability of a model to predict correctly the absence of an event, is comparable for each set of *gd* and $v_{th}$. Finally, the PPV and NPV found in this study are consistent with those described in *MacNeice* [2009a, 2009b].

### 3.3. Ensemble Predictions of High-Speed Events

Figure 6 shows examples of ensemble predictions of high-speed events made by the PDF model. For example, the top left graph shows that the actual speed (black solid line) stays in between the 25% and 75% quartiles for about 108 h and in between the 10% and 90% deciles for about 115 h. This means that most of the time, the PDF model manages to predict the uncertainty interval associated with the main prediction. However, these intervals by themselves are not sufficient to judge if they are relevant. For example, if the widths of the

probability distributions the model uses were of the order of 1000 km/s, then the actual speed would always be in between the 25% and 75% quartiles, but the prediction of the uncertainty would not be useful for the forecasters because of the huge uncertainty interval (± 500 km/s) associated with the prediction. Therefore, it is important to look at the width of these uncertainty intervals. On the six different predictions shown in Figure 6, the half widths are, on average, about 100 km/s for the 25%–75% quartiles and 150 km/s for the 10%–90% deciles.

The average time the observations last between the 25% and 75% quartiles, and the 10% and 90% deciles were calculated, as well as the average 25%–75% and 10%–90% widths over all the predictions made between 1995 and 2012. The results were classified according to whether the prediction was a true positive, a false positive, or a pure miss.

Out of 120 h, the observations were in between the 25% and 75% quartiles for 58 h for true positives, 68 h for false positives, and 26 h for pure misses. These numbers increased to, respectively, 88 h, 86 h, and 55 h for 10%–90% deciles. The smaller time for the pure miss in both cases may be explained by the fact that the model did not predict a high-speed event for these times, so the associated speed distributions from the PDF model were narrower. This is confirmed by the measurements of the half widths: 53 km/s and 95 km/s for the 25%–75% and 10%–90% quantiles, respectively, when the prediction was a miss. When it was a true positive and a false positive, the average widths of the 25%–75% and 10%–90% quantiles were around 145 km/s and 230 km/s, respectively.

## 4. Conclusion

An algorithm was created to identify transitions from slow to fast solar wind as any increase in the speed by more than 50 km/s in less than 24 h. This enabled the construction of new PDFs for the prediction of high-speed events. To predict the peak value of the speed, PDFs of the amplitude of the peak were calculated as a function of the maximum slope before the peak. The distribution of the time it takes for the speed to increase from the beginning of the peak to the maximum of the peak was used to predict the time of the peak. From these, predictions of the solar wind speed including peaks could be made.

Contingency tables were calculated to assess the ability of the new PDF model to predict high-speed events by answering the following question: "[I]f the model predicts/does not predict a high-speed event in the next 24 h, does a high-speed event actually occur?" It was found that 60.4% of the positive predictions were correct, while 91.4% of the negative predictions were correct. The 20.3% of the peaks in the speed are found by the model. The percentage increases to 33.6% when there is an associated peak in both the solar wind density and IMF magnitude before the increase in the solar wind speed.

A direct comparison with the WSA model showed that the number of false positives is more than 3 times smaller for the PDF model compared with the WSA model. The sensitivity (probability that an actual occurring event is correctly predicted by the model) is similar in both models, as is the NPV (probability that a negative forecast is correct). However, when the peak reaches very high speeds (> 600 km/s), the PDF model misses fewer events than the WSA model but tends to have more false positives than when the speed is lower.

Finally, ensemble predictions of high-speed events by the PDF model provides the forecast community with an interval of uncertainty on the prediction. The study showed that, on average, the observations were in between the 25% and 75% quartiles ~ 60 h out of 120 h and in between the 10% and 90% deciles ~ 85 h out of 120 h.

**Acknowledgments**
This research was funded by Air Force Office of Scientific Research grant FA9550-12-1-0265 and NASA grant NNL13AQ00C. We acknowledge the use of NASA/GSFC's Space Physics Data Facility's OMNIWeb service and OMNI data. We are extremely grateful to Peter MacNeice for providing WSA simulation results for 2011. Real-time predictions of the solar wind speed are available at http://vmr.engin.umich.edu/Model/_pdf/.

## References

Buizza, R. (2001), Accuracy and potential economic value of categorical and probabilistic forecasts of discrete events, *Mon. Weather Rev.*, *129*, 2329–2345, doi:10.1175/1520-0493(2001)129<2329:AAPEVO>2.0.CO;2.

Bussy-Virat, C., and A. Ridley (2014), Predictions of the solar wind speed by the probability distribution function model, *Space Weather*, *12*, 337–353, doi:10.1002/2014SW001051.

Cash, M. D., D. A. Biesecker, V. Pizzo, C. A. de Koning, G. Millward, C. N. Arge, C. J. Henney, and D. Odstrcil (2015), Ensemble modeling of the 23 July 2012 coronal mass ejection, *Space Weather*, *13*, 611–625, doi:10.1002/2015SW001232.

Dryer, M. (1974), Interplanetary shock waves generated by solar flares, *Space Sci. Rev.*, *15*, 403–468.

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

Fry, C., W. Sun, C. Deehr, M. Dryer, Z. Smith, S.-I. Akasofu, M. Tokumaru, and M. Kojima (2001), Improvements to the HAF solar wind model for space weather predictions, *J. Geophys. Res.*, *106*, 20,985–21,001.

Fry, C., M. Dryer, Z. Smith, W. Sun, C. Deehr, and S.-I. Akasofu (2003), Forecasting solar wind structures and shock arrival times using an ensemble of models, *J. Geophys. Res.*, *108*(A2), 1070, doi:10.1029/2002JA009474.

Kim, R., K. Cho, Y. Moon, M. Dryer, J. Lee, Y. Yi, K. Kim, H. Wanga, Y. Park, and Y. H. Kim (2010), An empirical model for prediction of geomagnetic storms using initially observed CME parameters at the Sun, *J. Geophys. Res.*, *115*, A12108, doi:10.1029/2010JA015322.

MacNeice, P. (2009a), Validation of community models: Identifying events in space weather model timelines, *Space Weather*, *7*, S06004, doi:10.1029/2009SW000463.

MacNeice, P. (2009b), Validation of community models: 2. Development of a baseline using the Wang-Sheeley-Arge model, *Space Weather*, *7*, S12002, doi:10.1029/2009SW000489.

Mays, M. L., et al. (2015), Ensemble modeling of CMEs using the WSA-ENLIL+Cone model, *Sol. Phys.*, *290*, 1775–1814, doi:10.1007/s11207-015-0692-1.

McKenna-Lawlor, S. M. P., M. Dryer, M. D. Kartalev, Z. Smith, C. D. Fry, W. Sun, C. S. Deehr, K. Kecskemety, and K. Kudela (2006), Near real-time predictions of the arrival at Earth of flare-related shocks during Solar Cycle 23, *J. Geophys. Res.*, *111*, A11103, doi:10.1029/2005JA011162.

Owens, M., C. Arge, H. Spence, and A. Prembroke (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.

Owens, M. J., H. E. Spence, S. McGregor, W. J. Hughes, J. M. Quinn, C. N. Arge, P. Riley, J. Linker, and D. Odstrcil (2008), Metrics for solar wind prediction models: Comparison of empirical, hybrid, and physics-based schemes with 8 years of L1 observations, *Space Weather*, *6*, S08001, doi:10.1029/2007SW000380.

Smith, Z., and M. Dryer (1990), MHD study of temporal and spatial evolution of simulated interplanetary shocks in the ecliptic plane within 1 AU, *Sol. Phys.*, *1291*, 387–405.

Smith, Z., M. Dryer, S. McKenna-Lawlor, C. Fry, C. Deehr, and W. Sun (2009), Operational validation of HAFv2's predictions of interplanetary shock arrivals at Earth: Declining phase of Solar Cycle 23, *J. Geophys. Res.*, *114*, A05106, doi:10.1029/2008JA013836.

Taktakishvili, A., P. J. MacNeice, and M. Hesse (2009), Ensemble modeling of the inner heliosphere using WSA-ENLIL and SWMF models, *Eos Trans. AGU*, *90*(52), Fall Meet. Suppl., Abstract SH13B–1533.

Vandegriff, J., K. Wagstaff, G. Ho, and J. Plauger (2005), Forecasting space weather: Predicting interplanetary shocks using neural networks, *Adv. Space Res.*, *36*, 2323–2327, doi:10.1016/j.asr.2004.09.022.

Wright, J. M., T. J. Lennon, R. W. Corell, N. A. Ostenso, W. T. Huntress, J. F. Devine, P. Crowley, and J. B. Harrison (1995), The National Space Weather Program—Strategic plan, in *National Space Weather Program Council*, pp. FCM–P30–1995, Office of the Federal Coordinator for Meteorological Services and Supporting Research.