AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2015SW001321

Key Points:
- Development and validation of an operational space weather tool to forecast propagation time delay
- Use the combined MVAB-0 and cross-product technique to take tilted phase plane fronts into account
- No significant improvement is observed when implemented in a real-time operational setting

Correspondence to:
M. D. Cash,
michele.cash@noaa.gov

Citation:
Cash, M. D., S. Witters Hicks, D. A. Biesecker, A. A. Reinard, C. A. de Koning, and D. R. Weimer (2016), Validation of an operational product to determine L1 to Earth propagation time delays, *Space Weather*, *14*, 93–112, doi:10.1002/2015SW001321

Received 7 OCT 2015
Accepted 4 JAN 2016
Accepted article online 7 JAN 2016
Published online 16 FEB 2016

Validation of an operational product to determine L1 to Earth propagation time delays

M. D. Cash[1,2], S. Witters Hicks[3], D. A. Biesecker[2], A. A. Reinard[1,2], C. A. de Koning[1,2], and D. R. Weimer[4]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, USA, [2]National Oceanic and Atmospheric Administration, Space Weather Prediction Center, Boulder, Colorado, USA, [3]Principia College, Elsah, Illinois, USA, [4]Center for Space Science and Engineering Research, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

Abstract We describe the development and validation of an operational space weather tool to forecast propagation delay times between L1 and Earth using the Weimer and King (2008) tilted phase front technique. A simple flat plane convection delay method is currently used by the NOAA Space Weather Prediction Center (SWPC) to propagate the solar wind from a monitoring satellite located at L1 to a point upstream of the magnetosphere. This technique assumes that all observed solar wind discontinuities, such as interplanetary shocks and interplanetary coronal mass ejection boundaries, are in a flat plane perpendicular to the Sun-Earth line traveling in the GSE *X* direction at the observed solar wind velocity. In reality, these phase plane fronts can have significantly tilted orientations, and by relying on a ballistic propagation method, delay time errors of ±15 min are common. In principle, the propagation time delay product presented here should more accurately predict L1 to Earth transit times by taking these tilted phase plane fronts into account. This algorithm, which is based on the work of Weimer and King (2008), is currently running in real time in test mode at SWPC as part of the SWPC test bed. We discuss the current algorithm performance, and via our detailed validation study, show that there is no significant difference between the two propagation methods when run in a real-time operational environment.

1. Introduction

Spacecraft which measure solar wind and the interplanetary magnetic field (IMF) conditions upstream of Earth's magnetosphere make it possible to predict future conditions in the Earth's near-space environment. The propagation time delay for fluctuations in the solar wind to travel from an upstream monitoring satellite orbiting the first Lagrange point (L1) to the magnetosphere allows for advanced prediction of approximately 1 h for approaching solar wind features. Exact delay times vary from over 100 min to less than 30 min depending on the prevailing solar wind conditions. From a space weather prediction perspective, obtaining an accurate estimate of the time that it takes a solar wind discontinuity observed at L1 to travel to Earth is important for two reasons. First, solar wind features observed at L1 can be time ordered according to their expected arrival time at the magnetosphere, permitting improvements to the input parameters driving geospace models. Second, an accurate estimate of the propagation delay time allows for improvements in the forecasted onset time of geomagnetic sudden commencements associated with a shock arrival at Earth.

The most straightforward method for estimating the propagation delay time from L1 to Earth is to use convection delay, in which the propagation time of the solar wind is determined by dividing the distance in the GSE *X* direction from the upstream monitor to a target location near Earth by the velocity of the observed solar wind. This simple flat plane propagation technique was examined by *Collier et al.* [1998] who analyzed Wind and IMP 8 magnetometer data for 543 two hour time periods from 1995 in order to determine the timing accuracy of this propagation method. They found that when using solar wind convection speeds, timing results were accurate to within 10% of the observed time delay in about two thirds of the cases considered. Timing accuracy decreased as the separation between the spacecraft increased both parallel and perpendicular to the Sun-Earth line. The separation between the spacecraft was found to be the most significant cause of timing inaccuracy and *Collier et al.* [1998] cautioned of a high probability of very bad timing agreement due to the long tail on the probability distribution. This simple flat plane convection delay technique is the method currently used by NOAA's Space Weather Forecast Office.

Another method for estimating the propagating delay time from L1 to Earth is to assume that all phase plane normals are perpendicular to the Parker spiral direction. This method was explored by *Russell et al.* [1980],

©2016. American Geophysical Union. All Rights Reserved.

who computed correlation coefficients and lag times at maximum correlation for 55 days of IMF data. The maximum correction coefficients were found to be highly variable, with the computed lag times having large departures from the expected values. While the simple Parker spiral propagation technique appears to work well in general, the presence of large deviations from the predicted values suggests that such a method cannot be used with confidence for all cases.

To compare the accuracies associated with four commonly used propagation methods, including the convection delay and Parker spiral techniques, *Ridley* [2000] determined the average and maximum propagation delay time error as a function of satellite distance from the Sun-Earth line. They show that for each method examined, larger errors are associated with increasing separation from the Sun-Earth line. For convection delay (referred to as the *X* distance method in the *Ridley* [2000] paper), errors in arrival times of ±15 min are common, with errors over 30 min observed during periods of large off-axis distance. Results for the Parker spiral method were similar. In order to more accurately determine the arrival time of propagating discontinuities, *Ridley* [2000] concluded that a better method is needed to determine the tilt of solar wind phase planes, especially during periods of large transverse separation. They suggested that the use of minimum variance analysis (MVA) to determine the orientation of propagating solar wind fluctuations could significantly reduce the associated timing errors.

Subsequently, various forms of the MVA technique [*Sonnerup and Scheible*, 1998] have been used to compute the tilt of solar wind phase planes with respect to the Sun-Earth line [*Weimer et al*., 2003; *Bargatze et al*., 2005; *Weimer and King*, 2008; *Mailyan et al*., 2008; *Pulkkinen and Rastätter*, 2009; *Haaland et al*., 2010; *Munteanu et al*., 2013]. *Weimer et al*. [2003] demonstrated that by using the minimum variance direction to define the phase front normal, they could correct for tilted phase planes and achieve better accuracy when propagating IMF measurements from ACE to three different target satellites located Earthward of L1. In a correction by *Weimer* [2004], it was noted that *Weimer et al*. [2003] adopted a nonstandard form of the equation used to calculate the magnetic field variance matrix that composed the basis of the MVA. Based on this use of a modified variance matrix, *Bargatze et al*. [2005] evaluated the physics underlying the *Weimer et al*. [2003] propagation model. The revised interpretation suggests that phase plane angles are organized in a coordinate system whose axes are defined by the mean IMF direction and the minimum and maximum perturbation directions perpendicular to the mean field vector. In this modified coordinate system, the phase front normal (PFN) is given by the eigenvector with the minimum eigenvalue. This minimum perturbation direction lies in a plane that is orthogonal to the mean magnetic field vector, and so by including an additional factor of *N* in the standard variance equation, the results of the *Weimer et al*. [2003] implementation of the MVA technique are similar to those of a version of MVA that is constrained by the condition that the average field along the phase plane normal is zero [see *Sonnerup and Cahill*, 1967].

The similarity of the Weimer method to that of the constrained MVA method was pointed out by *Haaland et al*. [2006], who argued that the MVA variant in which the average magnetic field component along the phase plane normal is zero should be used instead of the modified variance matrix presented in *Weimer et al*. [2003]. This method, referred to as MVA with the constraint that $\langle \boldsymbol{B} \rangle \cdot \widehat{\boldsymbol{n}} = 0$, or MVAB-0 hereafter, has been described by *Sonnerup and Scheible* [1998], *Sonnerup et al*. [2004], and *Haaland et al*. [2004]. By choosing a normal vector such that variations along $\widehat{\boldsymbol{n}}$ are minimized under the constraint that $\langle \boldsymbol{B} \rangle \cdot \widehat{\boldsymbol{n}} = 0$ requires that $\widehat{\boldsymbol{n}}$ lies in the plane perpendicular to $\langle \boldsymbol{B} \rangle$. As described by *Haaland et al*. [2006] the simplest implementation of MVAB-0 involves multiplying the covariance matrix on both the right and left sides by a projection matrix. The minimum eigenvalue of the resulting matrix is exactly zero and the corresponding eigenvector points along $\langle \boldsymbol{B} \rangle$. The PFN direction, $\widehat{\boldsymbol{n}}$, is now given by the eigenvector corresponding to the intermediate eigenvalue. *Haaland et al*. [2006] demonstrate the similarities between the modified MVA method of *Weimer et al*. [2003] and MVAB-0, and the authors conclude by suggesting that further testing of the use of the MVAB-0 method as a forecasting tool is justified.

Another promising method for determining the phase front normal direction is the cross-product technique [*Horbury et al*., 2001; *Knetter et al*., 2004] in which the PFN direction is found by taking a cross product of two averaged magnetic field vectors located upstream and downstream of the discontinuity. Using 22 days of Cluster spacecraft data, *Knetter et al*. [2004] found that the cross-product normal determined at each of the four spacecraft agreed fairly well with the normal direction derived by triangulation. Furthermore, the PFN directions were found to typically point in a direction approximately perpendicular to the average magnetic field direction. The authors conclude the cross-product (CP) method should be considered a reliable technique for determining discontinuity normals when only one spacecraft is available.

Weimer and King [2008, hereafter WK08] tested these various methods for calculating the tilt angle of IMF phase fronts in order to determine which methods work best for PFN and time delay calculations. They found that when optimized, both MVAB-0 and the CP technique work equally well and further improvements to delay time predictions were obtained when the two methods were used in combination. The authors suggested that the use of a two technique method that combines MVAB-0 with the CP technique (MVCP) would present a significant improvement compared to neglecting the IMF tilt angle as is currently done.

In this paper we discuss the development and validation of an operational space weather tool to forecast propagation delay times between L1 and a target location upstream of Earth using the combined MVAB-0 and cross-product technique described by WK08. In section 2, we present the numerical methods used to determine solar wind propagation times. In subsequent sections, a detailed validation study of the use of this method for real-time propagation of solar wind data from L1 to the front of Earth's magnetosphere is presented. The goal of this work is to use the computed L1 to Earth delay time information to generate operational space weather products to meet two needs that have been identified at NOAA's Space Weather Prediction Center. The first is to improve the inputs driving geospace models. In order to meet the needs of the geospace models, the chosen propagation algorithm must run continuously and provide an accurate description of the solar wind near Earth for typical solar wind conditions. The use of the combined MVCP technique, for this purpose is addressed in section 3. The second objective is to improve the predicted arrival time of space weather events, which will result in improvements to the predicted onset timing of geomagnetic storm warnings. The second part of this validation study, presented in section 4, addresses how well the MVCP method does at predicting the arrival time of observed interplanetary shocks and discontinuities. From a space weather forecasting perspective we know that there are cases when our performance could be better, such as when the predicted arrival time of a solar wind shock or discontinuity is off by greater than 15 min. Our goal with this work is to remove these large errors in order to build confidence in the forecast, allowing space weather customers to respond in an appropriate and timely manner to predicted space weather events. The results of our validation study and recommendations for implementation to operations are presented in section 5.

2. Computing Propagation Time Delays

To compute the propagation time delays, we use the technique described by WK08 to define the tilt angles of IMF phase fronts in order to obtain more accurate arrival times for space weather events. This method is a combination of the MVAB-0 and the CP techniques. Both methods must yield a valid result and be in near agreement (see agreement angle in Table 1) for the computed IMF tilt angles to be used in the determination of the expected transit time between the monitoring satellite and target. If a valid tilt angle is not obtained, then the last valid tilt angle is used to compute the transit time unless a shock or discontinuity has been observed, in which case a standard flat plane propagation method is used to compute the transit time of the discontinuity. The two methods are described in the subsections below along with a discussion of the delay time calculation and the optimization of the parameters used within the algorithm.

2.1. MVAB-0 Technique

The minimum variance technique uses a symmetric 3 by 3 magnetic variance matrix defined as

$$\boldsymbol{M_{ij}} \equiv \left\langle B_i B_j \right\rangle - \left\langle B_i \right\rangle \left\langle B_j \right\rangle \tag{1}$$

where B_i and B_j are the Cartesian vector components of the magnetic field measurements [*Sonnerup and Scheible*, 1998]. The eigenvalues λ and the corresponding eigenvectors $\boldsymbol{v}$ of the variance matrix, $\boldsymbol{M}$, can then be solved for using the equation

$$\boldsymbol{Mv} = \lambda \boldsymbol{v} \tag{2}$$

The three eigenvectors $\boldsymbol{v}$ represent the directions of maximum, intermediate, and minimum variation of the field components along each vector.

Table 1. Range of Considered Values and Calculation Options Tested in Algorithm Optimization

Parameter	Possible Values/Options	Selected Values (1 min RT Data)
Data cadence	1 s or 1 m RT data	1 m RT data
Limiting angle	0–80°	65°
Number of points in CP average	3–7 min of data	3
Number of points in minimum variance calculation	5–11 min of data	7
Agreement angle	8–50°	40°
Minimum eigenvalue ratio	1–5 ratio of $\lambda_{max}/\lambda_{int}$	2.5
Minimum *B* change angle	2–10°	2°
Step size	2–6 min of data	2
Number of points in shock average	1–5 min of data	1
For invalid tilt angles	Use previous valid value/assume flat plane	Use previous valid value[a]

[a]Only for unstructured solar wind. If a shock or discontinuity is observed in the solar wind data and an invalid tilt angle is obtained, then the algorithm reverts to assuming a flat plane propagation until the next valid tilt angle is computed.

In MVAB-0, the variant of MVA constrained by $\langle \boldsymbol{B} \rangle \cdot \widehat{\boldsymbol{n}} = 0$, the average magnetic field along the minimum variance direction is zero. Physically, this condition implies that there is no magnetic field along the IMF phase front normal; the IMF phase fronts are viewed as tangential discontinuities and there is no flow of plasma through the discontinuity [*Haaland et al*., 2010]. In order to satisfy the condition that $\langle \boldsymbol{B} \rangle \cdot \widehat{\boldsymbol{n}} = 0$, the variance matrix $\boldsymbol{M_{ij}}$ is multiplied on both sides by a projection matrix $\boldsymbol{P}$,

$$\boldsymbol{P_{ij}} = \delta_{ij} - \widehat{\boldsymbol{e}}_i \widehat{\boldsymbol{e}}_j \tag{3}$$

where δ_{ij} is the Kronecker delta and the unit vector $\hat{\boldsymbol{e}}$ is defined by $\hat{\boldsymbol{e}} = \langle \boldsymbol{B} \rangle / B$. The projection matrix describes a projection of a vector onto a plane perpendicular to the average magnetic field direction vector [*Sonnerup and Scheible*, 1998]. Multiplying the variance matrix $\boldsymbol{M_{ij}}$ on both the right and the left side by this projection matrix gives a matrix Q

$$\boldsymbol{Q_{nk}} = \boldsymbol{P_{ni} M_{ij} P_{jk}} \tag{4}$$

To solve for MVAB-0, the eigenvalues and eigenvectors of matrix $\boldsymbol{Q}$, instead of matrix $\boldsymbol{M}$ are used. By design, the minimum eigenvalue of matrix $\boldsymbol{Q}$ is exactly zero and the corresponding eigenvector points along $\langle \boldsymbol{B} \rangle$, the direction of the average magnetic field. The PFN direction is given by the eigenvector corresponding to the intermediate eigenvalue λ_{int} and the ratio of the maximum to the intermediate eigenvalues is used to determine the quality of the result with a value near unity indicating an indeterminate result for which there is not a clearly defined normal direction. There is no definitive cutoff value for determining when to discard the result obtained via MVAB-0, although values ranging from 2 to 10 have previously been used by WK08.

2.2. Cross-Product Method

The cross-product technique involves taking the vector cross product of two averaged magnetic field vectors located upstream and downstream of a discontinuity in order to determine the PFN direction [*Burlaga*, 1969; *Horbury et al*., 2001; *Knetter et al*., 2004]. Instead of limiting the use of this technique to known discontinuities, WK08 applied the cross-product method to a continuous stream of IMF data. When the angle between the upstream and downstream magnetic field vectors (the spreading angle) is sufficiently large, the vector cross product can be used to define the PFN direction. A "sufficiently large" spreading angle was determined by WK08 to correspond to a minimum angle of 8.8°.

Additional parameters associated with the CP method that require optimization are the number of data points in the upstream and downstream magnetic field averages, as well as the number of data points separating the two vectors. The optimization of these parameters depends on the cadence of the data being used. WK08 provided optimized parameters for 16 s ACE data. When implementing this technique using real-time data, we can choose between 1 s and 1 min data and vary the number of points in the average accordingly. The optimization of these and other parameters is discussed in section 2.4.

2.3. Delay Time Calculation

Once the phase plane orientation has been determined via the two techniques described above, the computed IMF tilt angles can then be used to compute the expected transit time. The valid PFN vector $\hat{\boldsymbol{n}}$ is used to propagate the solar wind parameters from the position of the solar wind monitoring satellite to a target location, either another satellite or to a point sunward of Earth's bow shock. The predicted propagation delay time (Δt) from an L1 satellite to a target location is calculated using the formula

$$\Delta t = \frac{\hat{\boldsymbol{n}} \cdot \left(\overline{\boldsymbol{P}}_{\mathbf{TARGET}} - \overline{\boldsymbol{P}}_{\mathbf{SAT}}\right)}{\hat{\boldsymbol{n}} \cdot \overline{\boldsymbol{V}}_{\mathbf{SW}}} \tag{5}$$

where the vector position of the target is given by $\overline{\boldsymbol{P}}_{\mathbf{TARGET}}$ and the vector position of the upstream satellite is given by $\overline{\boldsymbol{P}}_{\mathbf{SAT}}$, the solar wind velocity is $\overline{\boldsymbol{V}}_{\mathbf{SW}}$, and the phase plane's normal direction is given by $\hat{\boldsymbol{n}}$. Ideally, all three components of the solar wind velocity are used to calculate the delay time in equation (5). With ACE data, since only the magnitude of the solar wind velocity is available in real time, the velocity vector is assumed to be in the GSE X direction with 29.8 km/s added to the *Y* component to account for aberration effects due to the Earth's motion around the Sun. With the recently launched Deep Space Climate Observatory (DSCOVR), all three-vector components of the magnetic field will be available in near real time at 3 s resolution. Once DSCOVR becomes operational, we plan to use all three components of the velocity vector in the computation of the propagation time. For the validation study presented in this paper, we use historic ACE solar wind data and assume a radial velocity vector.

In order to compute Δt, a valid PFN vector $\hat{\boldsymbol{n}}$ needs to be defined for each time step. A reliable determination of the phase front orientation is not always obtained using the MVAB-0 and CP methods and in these cases the invalid phase plane normal needs to be replaced with another value in order to determine the predicted propagation delay time. In WK08, interpolation between the well-determined phase planes is used; however, this is not an option for real-time applications. *Pulkkinen and Rastätter* [2009] proposed an alternative technique for stabilizing the phase plane orientations that removes the influence of the small-scale fluctuations on the calculated normals, effectively acting as a low-pass filter. *Haaland et al*. [2010] argued that the filtering should be performed on the input data rather than the computed normals, and instead of frequency filtering, *Haaland et al*. suggested the use of wavelet denoising. The effect of wavelet denoising on the timing accuracy of three propagation delay estimation methods was recently investigated by *Munteanu et al*. [2013], who found that the use of wavelet denoising improved the predictions of the propagation time delay of solar wind discontinuities. In our validation study we do not employ the use of wavelet denoising as this adds an additional computational layer that may not be appropriate for real-time applications. Simplicity and computational efficiency are valued in real-time forecasting and more complex methods must demonstrate a significant improvement before they are considered for implementation in a real-time operational setting.

2.4. Time Ordering the IMF Data

When time ordering the solar wind data, several issues need to be addressed. The first is how to handle out-of-sequence arrival times. WK08 explored several options for time ordering the IMF data and found the best results were produced by simply sorting the downstream IMF values in sequential order according to their arrival time tags. This method is straightforward and does not make an assumption as to which phase plane dominates when one phase plane overtakes another. Other approaches give priority to overtaking phase fronts or to earlier observations; however, given the difficulty in determining how overlapping phase fronts will interact, we follow the recommendation of WK08 and sort the downstream IMF values by arrival time. While more sophisticated physics-based approaches exist for propagating the observed L1 time series after accounting for the tilted phase fronts, i.e., using a 1-D hydrodynamic model or a 1-D modified kinematic model such as that described by *Arge and Pizzo* [2000], the simpler method recommend by WK08 is used here.

In this time-ordering scheme, each solar wind measurement is placed in a 1 min bin according to the predicted arrival time and if more than one solar wind parcel is expected to arrive for a given minute, then a decision must be made as to how to select a representative solar wind values for that time bin. The two obvious choices are to (1) take an average of all the solar wind parcels arriving in a given 1 min bin or (2) select the largest value and neglect all other values predicted to arrive at a similar time. The method of averaging the solar wind values in each bin was found by WK08 to produce the best test scores, and it is the method that we employ in our algorithm. We also tested the method of using the solar wind parcel with the maximum dynamic pressure as suggested by space weather forecasters; however, we found that this method did not provide any improvement over simply taking an average of all solar wind parcels arriving at a given time.

In the Predicted Solar Wind at Earth product, described in section 3, for minutes when no solar wind parcels are predicted to arrive, solar wind values from the previous minute are repeated and the data are flagged by noting that 0 points were used in the solar wind average. This method allows the user to decide if they would rather interpolate the data or use the previous solar wind value.

2.5. Optimization of Parameters

There are several parameters that require optimization within the MVCP method. The optimization of these parameters is discussed in WK08 for 16-s ACE science data. We use 1 s and 1 min real-time ACE beacon data stored at SWPC for our analysis and have modified the code to run in near-real time (NRT); thus additional optimization is necessary. For each adjustable parameter within the MVCP algorithm, we performed a systematic optimization for both 1 s and 1 min data. The range of values that we considered, along with the optimized values for the 1 min cadence data, are shown in Table 1. No significant difference was observed between the algorithm performance using 1 s versus 1 min cadence data, and we opt to use the 1 min data in the determination of the propagation time delays. A coarse optimization was performed to determine these values, but potentially more refined values exist.

Modifying the MVCP code to run in NRT excludes the use of some of the preferred methods identified in WK08, such as interpolating between valid tilt angles. For the purposes of running this algorithm in NRT for use as an operational forecast tool, we are limited to two different options for handling the times when a valid tilt angle is not obtained. The first option is to use the last valid tilt angle until another valid tilt angle is acquired or until a specified amount of time has passed. The second option is to revert to using the standard flat plane propagation method. Both methods have been examined and using the last valid tilt angle was found to produce the best results for unstructured solar wind. If a shock or discontinuity is observed in the solar wind data and an invalid tilt angle is obtained, then the algorithm reverts to assuming a flat plane propagation until the next valid tilt angle is computed as this was determined to produce better results during the optimization process. Using the optimized parameters specified in Table 1, valid tilt angles are obtained 41% of the time, with a standard deviation of 14%. The maximum number of valid tilt angles obtained in a given day was 62% and the minimum number of valid tilt angles obtained in a given day was 14%. Invalid tilt angles are replaced with the last valid tilt angle the majority of the time with less than 1% of the invalid tilt angles replaced with a flat phase plane.

3. Continuous Prediction of the Solar Wind at Earth

Using the output of the MVCP algorithm, two distinct products are generated. The first is the Predicted Solar Wind at Earth product, which aims to improve the inputs driving geospace models by continuously estimating the solar wind at a point upstream of Earth's magnetosphere. Using the methods described above, this product propagates real-time solar wind data from an L1 orbiting satellite to a target location and orders the data according to the predicted arrival times. A target location 30 R_E sunward of the Earth and located on the Sun-Earth line was selected as this is the location of the inner boundary of the geospace model [*Ridley et al.*, 2002] that NOAA/SWPC is currently in the process of transitioning to operations.

3.1. Validation of Time-Ordered Data

To determine how well the combined MVCP technique performs requires the use of two in situ spacecraft: a NRT solar wind monitor located near L1 provides the initial observations, and a satellite located near Earth's bow shock serves as a ground truth. The ACE spacecraft, which has been located in an L1 Lissajous orbit since early 1998, is used to provide the input solar wind data. In order to evaluate how well the propagated solar wind data compares to what was actually observed near Earth, a spacecraft positioned upstream of Earth's bow shock is required. Previous studies have used Wind data [*Horbury et al.*, 2001] and Cluster data for this purpose [*Mailyan et al*., 2008; *Munteanu et al*., 2013]. In this study, we use Wind data from the second half of 1998 when the Wind spacecraft was in Earth orbit and was often located in the solar wind. In addition, we use a handful of days from 1999 to 2004 when Wind was located approximately 30 R_E sunward of Earth and near the Sun-Earth line. Using the days when the Wind spacecraft was located in such an ideal location, we compare ACE data that has been propagated to the location of the Wind spacecraft to the observed Wind data.

To select the dates used in this validation study, we used the following criteria: (1) Wind was located in the solar wind and not in the foreshock, magnetosheath, or magnetosphere; (2) Wind was located less than 100 R_E sunward of Earth and within 30 R_E of the Sun-Earth line—this removes the times when Wind was located too close to ACE and too far from the Earth, when it would be a poor proxy for Earth's bow shock; (3) There was not a data gap in either the ACE or Wind data of more than 4 h. These selection criteria are similar to those used by *Horbury et al*. [2001] in their study predicting the Earth arrival times of ACE-observed IMF southward turnings. To optimize the algorithm, we used a subset of 15 days of data from 1999 to 2004. These days were selected to represent a range of dates and solar wind conditions while the modest sample size allowed for computational speed while stepping through numerous permutations of input parameters for optimization. An independent set of 24 days of data from the second half of 1998 were used for the validation study. These days were chosen based on the location of the Wind spacecraft upstream of the magnetopause and located near the Sun-Earth line. When computing all test statistics discussed below, only data from each day in question are used and we do not include data from the adjacent days. Thus, for each day there are slightly less than 1440 samples due to averaging windows for the MVAB-0 (seven samples) and CP methods (three samples). The use of two independent data sets for optimization and validation prevents over specifying the data set and therefore biasing the results.

For the days used in the validation study, the separation between the two spacecraft in Earth radii, and the average tilt angle for that 24 h period are given in Table 2. For each day, we compute continuous predictant skill scores [*Wilks*, 2011] to assess the performance of these two predictive models and use paired *t* test scores to determine if the differences between the two methods are statistically significant. The skill of a forecast refers to the relative accuracy of the forecast with respect to a standard reference forecast, and a skill score is used to provide the percentage improvement over the reference forecast. In this case, the standard reference forecast is the flat plane propagation technique and the MVCP method is the forecast to be evaluated. To compute the scalar measures of forecast accuracy, we use the mean absolute error (MAE) as the underlying accuracy statistic,

$$\mathrm{MAE} = \frac{1}{n}\Sigma_{k=1}^{n}|y_k - o_k| \tag{6}$$

where (y_k, o_k) is the *k*th of *n* pairs of forecasts and observations [see *Wilks*, 2011, chapter 8] and is used to compute the skill score of the forecast under evaluation (MVCP technique) compared to a standard reference forecast (convection delay method). For each minute of data we compute the Skill Score (SS) using the MAE,

$$\mathrm{SS}_{\mathrm{con}} = 100\times\left(1 - \frac{\mathrm{MAE}}{\mathrm{MAE}_{\mathrm{con}}}\right) \tag{7}$$

If the MAE is zero then the forecast is perfect, and the MAE increases as the discrepancies between the forecasts and observations become larger. We compute a separate SS for B_x, B_y, and B_z, as well as an overall SS which incorporates all three values. In Table 2 the results have been sorted by the overall SS with a horizontal black line indicating days for which a >5% improvement is observed when comparing the MVCP technique to convection delay. For 8 out of the 24 days in this validation study or 33% of the time, a skill score of >5% is obtained.

Next, to determine if the difference between the two methods is statistically significant, we use a two-sample *t* test for paired data [see *Wilks*, 2011, chapter 5]. The tilted phase plane method and the convection delay method provide the paired data (y_k, x_k), which is analyzed by taking the differences ($\Delta_k = y_k - x_k$) between the corresponding values (in nanotesla) generated by each propagation method for every time step *k*. By using the differences between the corresponding pairs, we can transform the two-sample problem into the familiar one-sample *t* test, in which the *z* value is computed as follows:

$$z = \frac{\overline{\Delta}\sqrt{n}}{s_\Delta} \tag{8}$$

where $\overline{\Delta}$ is the mean difference between the MVCP and convection delay pairs of forecasts, s_Δ is the standard

Table 2. Validation of Continuous Solar Wind Data Propagated From ACE to Wind

	Spacecraft Separation			Tilt Angle		Skill Scores				
Date (yyyy-mm-dd)	X (R_E)	Y (R_E)	Z (R_E)	Phi (deg)	Theta (deg)	B_x	B_y	B_z	Overall[a]	Z Value[b]
1998-07-27	196	59	17	−4.8	10.6	16.7	15.7	22.1	18.2	0.23
1998-07-30	166	49	20	11.7	−23.1	11.6	17.9	6.8	12.1	0.14
1998-08-06	152	19	23	11.8	−5.1	5.6	11.1	13.5	10.1	0.62
1998-08-07	155	15	24	19.8	0.3	3.3	13.7	10.6	9.2	0.43
1998-07-05	182	45	0	21.4	−4.3	1.6	7.1	16.0	8.2	0.30
1998-07-08	157	27	3	29.2	−13.0	1.0	10.7	11.2	7.6	1.28
1998-07-29	174	53	19	−21.8	−4.2	2.3	9.5	11.0	7.6	0.56
1998-07-28	183	56	18	−1.5	3.7	5.1	9.2	2.4	5.6	1.90
1998-07-11	149	12	6	19.5	−8.2	2.8	2.8	8.4	4.7	0.56
1998-07-10	150	17	5	15.7	−0.7	4.4	4.5	3.0	4.0	0.55
1998-07-06	171	39	1	−23.3	0.1	−1.1	−0.1	9.7	2.8	2.29
1998-08-04	150	28	23	20.4	−19.2	−1.6	6.8	2.9	2.7	1.42
1998-08-10	170	3	24	6.1	−5.4	7.4	2.5	−5.6	1.4	3.11
1998-07-07	163	32	2	−22.9	−15.5	0.3	2.9	0.8	1.3	0.96
1998-08-14	219	6	23	10.5	−19.4	1.3	0.5	−0.2	0.5	0.31
1998-07-31	160	45	20	−2.8	−0.8	4.4	−2.8	−0.1	0.5	0.29
1998-08-08	159	11	24	12.3	−34.6	−1.5	2.6	0.0	0.4	0.15
1998-07-12	148	8	7	28.7	1.9	−0.5	0.5	1.1	0.4	0.12
1998-08-09	164	7	24	42.1	3.4	−6.1	6.8	−1.5	−0.3	1.85
1998-08-11	179	0	24	0.8	−21.1	−0.2	−0.9	−1.1	−0.7	1.09
1998-07-09	153	22	4	7.1	−0.9	3.0	−4.3	−1.0	−0.8	0.48
1998-08-12	189	3	24	38.2	−4.8	−1.6	−0.9	−0.5	−1.0	0.93
1998-08-03	151	32	22	6.1	0.6	2.3	−3.3	−2.6	−1.2	1.44
1998-08-02	153	36	22	30.2	12.8	−1.4	−22.1	−6.3	−9.9	0.34

[a]The overall skill score is the average of the B_x, B_y, and B_z skill scores for the indicated day.
[b]The absolute value of the two-sample t test scores are listed. Values above 1.96 suggest a 95% confidence level of a significant difference between the two methods.

deviation of Δ, and n is the number of data points in the sample. To compute the z value, we treat the B_x, B_y, and B_z data sets independently, computing a mean difference $\overline{\Delta}$ and standard deviation for each vector component. In order to arrive at a single z score value for a given day, we combine these three data sets using conflation [*Hill*, 2011; *Hill and Miller*, 2011] and account for persistence in the solar wind values using a correction based on the lag-1 autocorrelation [*Wilks*, 2011, chapter 3] which reduces the effective sample size. Once the test statistic z is obtained, a t distribution critical values table can be used to quantify the likelihood that the two methods are significantly different. A confidence level of $>$95% is considered to be a significant difference in this analysis and that corresponds to a z value of 1.960.

The final column in Table 2 lists the z value comparing the MVCP technique to convection delay. From this column it becomes immediately apparent that for the majority of days considered in this study or ~80% of the time, the differences between the two propagation methods are not statistically significant, even for cases with high skill scores. The average skill score for all 24 days is $3.5 \pm 5.7\%$ while that average z value is 0.89 ± 0.78. Both metrics indicate that on average there is not a significant difference between the two propagation methods considered in this study.

3.2. Examples of Solar Wind Propagated From ACE to Wind

The results of the propagated ACE data compared to the observed Wind data can be viewed graphically as shown in Figure 1 for 02 July 1999. In Figure 1 and in the subsequent figures, the three components of the IMF values measured at the Wind spacecraft are plotted in black. ACE data that has been shifted in time assuming a flat plane propagation at the solar wind velocity (convection delay) is shown in red, and the blue lines represent the ACE measurements that have been shifted in time according to the tilt angles from the MVCP method. This figure is similar to Figure 4 in WK08 of propagated data using the two different propagation methods and demonstrates that our algorithm using 1 m beacon data shows reasonable agreement with

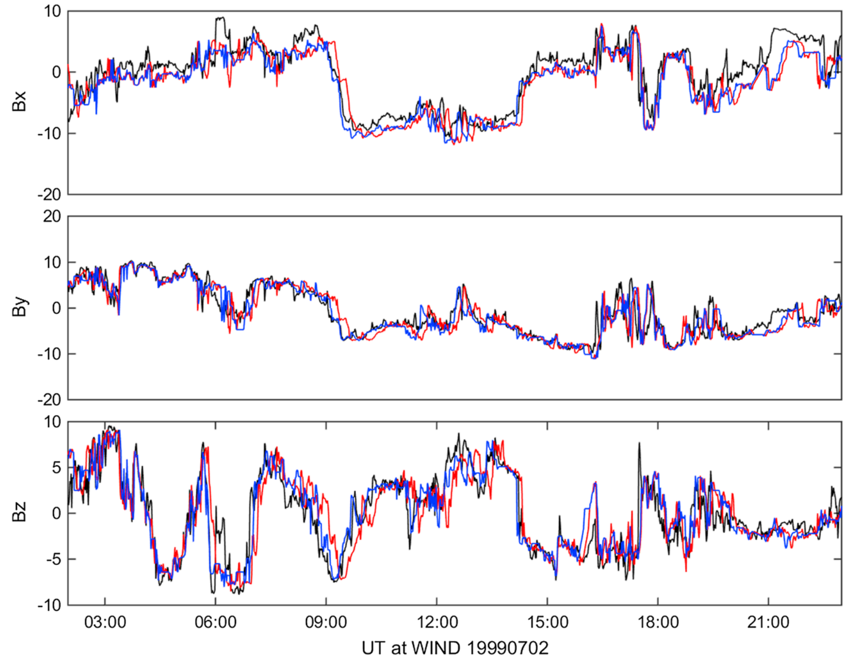

Figure 1. Comparison of the IMF measurements from both the Wind and ACE satellites, taken on 2 July 1999. Panels show the three components of the IMF as measured at Wind (black lines), ACE data that has been shifted in time according to a flat plane propagation at the solar wind velocity (red lines), and ACE measurements that have been shifted in time according to the tilt angles from the MVCP method (blue lines). This date was selected for comparison to Figure 4 in WK08. The overall skill score for this event is 10.7% with a *z* value of 1.40.

the results published in WK08 using 16 s science data. As can be seen in Figure 1 at time 0930 UT, much better agreement is observed when using the MVCP technique (blue line) versus convection delay (red line), particularly in B_y. In this example, the vector separation between the two spacecrafts in GSE coordinates was [28, 61, 1.2] R_E and the average tilt was $-14.6°$ in phi (the angle between the PFN projection in the ecliptic plane and the Sun-Earth line) and 15.0° in theta (the angle between the PFN vector and the ecliptic plane). The overall SS was 10.7% with a *z* value of 1.40, meaning that while there is an observed improvement using the MVCP technique, this difference is not statistically significant. In fact, when comparing the red and blue lines in Figure 1, results from both propagation methods are quite similar the majority of the time.

Magnetic field data for the date with the largest skill score, 27 July 1998, is shown in Figure 2. On this date the overall SS is 18.2% with a *z* value of 0.23. While an improvement in the skill score is computed using the MVCP technique compared to convection delay, this difference is not statistically significant. The data on 27 July 1998 was quite noisy with little in the way of coherent structure. This noisy data give rise to an artificially high skill score which does not accurately reflect the performance of the MVCP technique compared to convection delay. Overall neither method does a good job at reproducing the solar wind magnetic field as observed by Wind. Another day in which a large skill score is obtained (12.1% improvement) but for which no significant difference between the two propagation methods is observed is 30 July 1998 (Figure 3). On this day the solar wind was quite featureless and the *z* value was 0.14. Again neither method performed significantly better than the other.

Only 2 days showed a significant difference between the tilted phase plane technique and convection delay when the two-sample *t* test was applied. *Z* values of 2.29 and 3.11 were obtained for 06 July 1998 and 10 August 1998, respectively, which correspond to a confidence level of >99% that the two propagation techniques are significantly different. However, as can be seen in Figures 4 and 5, only minimal differences are observed between the two propagation methods on these days and neither method represents the IMF values observed by the Wind spacecraft well.

Examples of two days from Table 2 when an improvement can be identified in the graphical data when using the MVCP technique compared to convection delay are shown in Figures 6 and 7 for 7 and 12 July 1998,

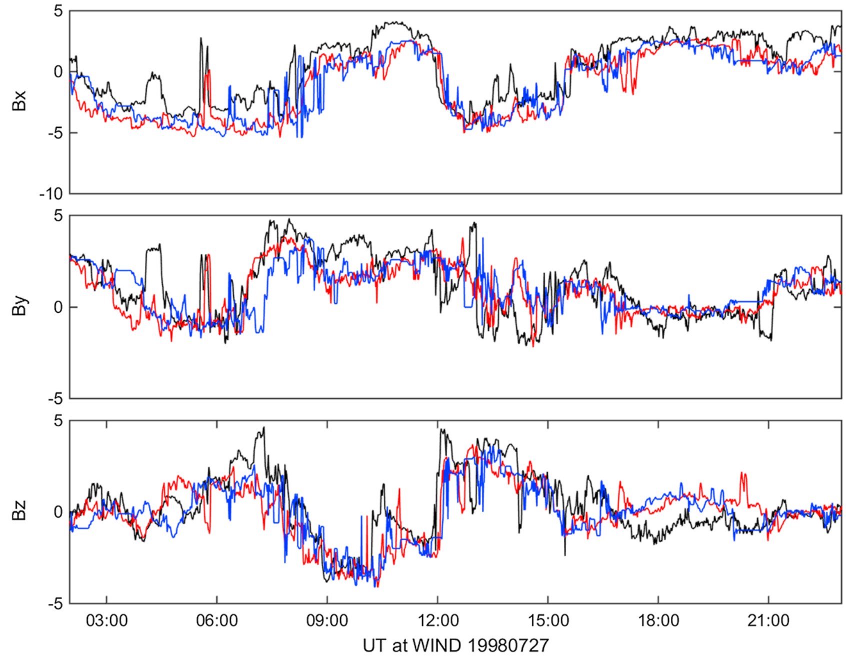

Figure 2. Similar to Figure 1 but for the solar wind on 27 July 1998. This day had an overall skill score of 18.2% with a *z* value of 0.23, indicating that no significant difference is observed between the two propagation methods.

respectively. The differences between the two methods of propagating the data from ACE to Wind are most easily observed in discontinuities in the solar wind magnetic field, such as at 1400 UT in Figure 6 and in the B_z component of the IMF at 0600 UT in Figure 7. In both examples, good agreement is seen between the ACE solar wind data propagated using the MVCP technique (blue line) to the values observed by Wind (black line). For the ACE solar wind data propagated to the location of Wind using convection delay, the discontinuity

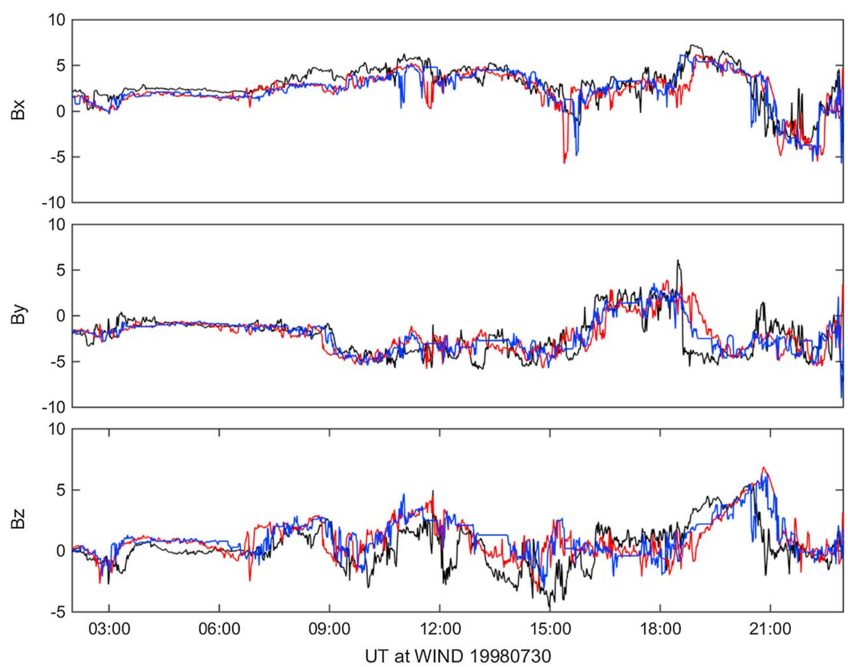

Figure 3. Similar to Figure 1 but for the solar wind on 30 July 1998. This day had an overall skill score of 12.1% with a *z* value of 0.14, indicating that no significant difference is observed between the two propagation methods.

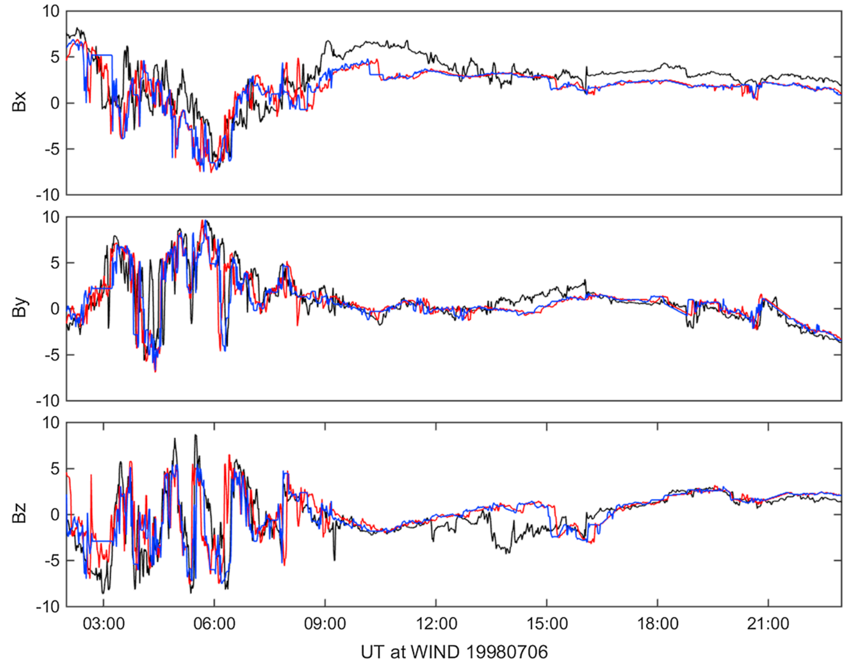

Figure 4. Similar to Figure 1 but for the solar wind on 06 July 1998. This day had an overall skill score of 2.8% with a *z* value of 2.29. While such a large *z* value would suggest a significant difference between the two propagation techniques, overall there are only minimal differences between the two propagation methods observed on this day.

is predicted to arrive at Wind 11 min later than was actually observed for 7 July 2012 and 8 min earlier than was observed on 12 July 1998. However, the *z* values for both dates indicate that this improvement is not significant.

The date with the lowest skill score, 2 August 1998, is plotted in Figure 8. The skill score for this date is −9.9%, which is the worst skill score of all the dates in the validation study. During this day it difficult to tell which

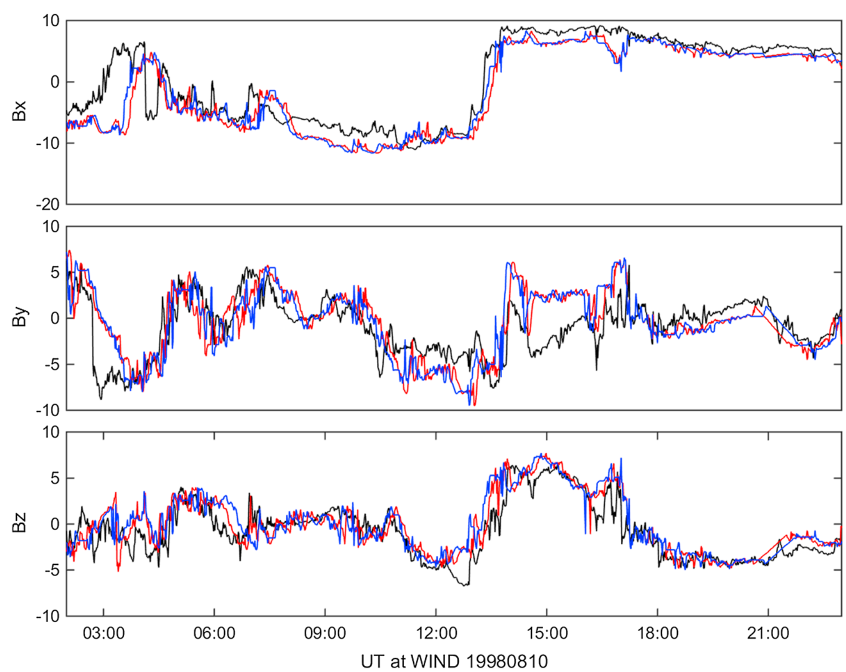

Figure 5. Same form as Figure 1 but for the solar wind on 10 August 1998. This day had an overall skill score of 1.4% with a *z* value of 3.11. While such a large z score would suggest a significant difference between the two propagation techniques, neither method does well in reproducing the solar wind as observed by the Wind spacecraft.

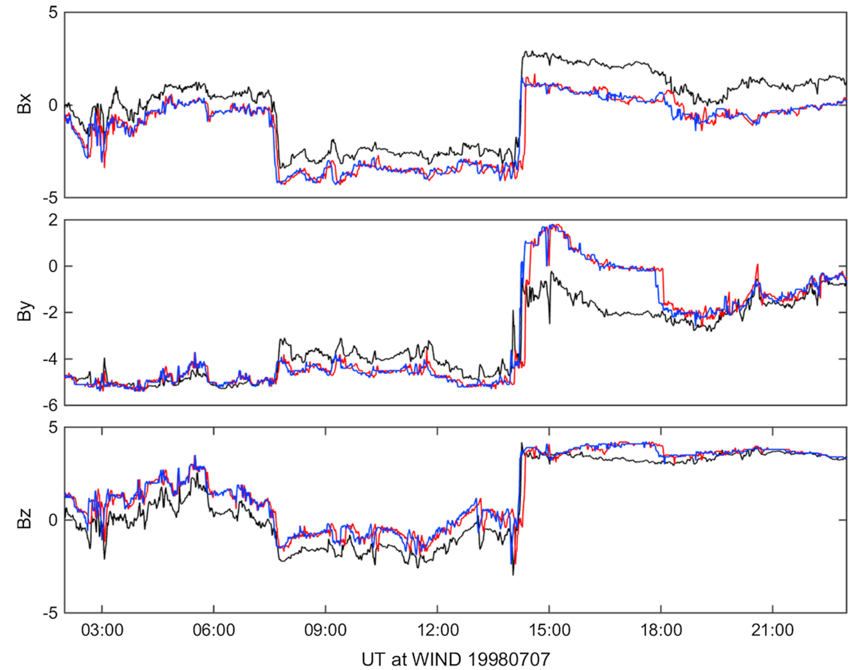

Figure 6. Same form as Figure 1 but for 7 July 1998. The overall skill score for this event is 1.3% with a *z* value of 0.96. While not a statistically significant improvement, good agreement is seen between the ACE solar wind data propagated using the MVCP technique (blue line) to the values observed by Wind (black line) for the discontinuity at 1400 UT. For the ACE solar wind data propagated to the location of Wind using convection delay, the discontinuity is predicted to arrive at Wind 11 min later than was actually observed.

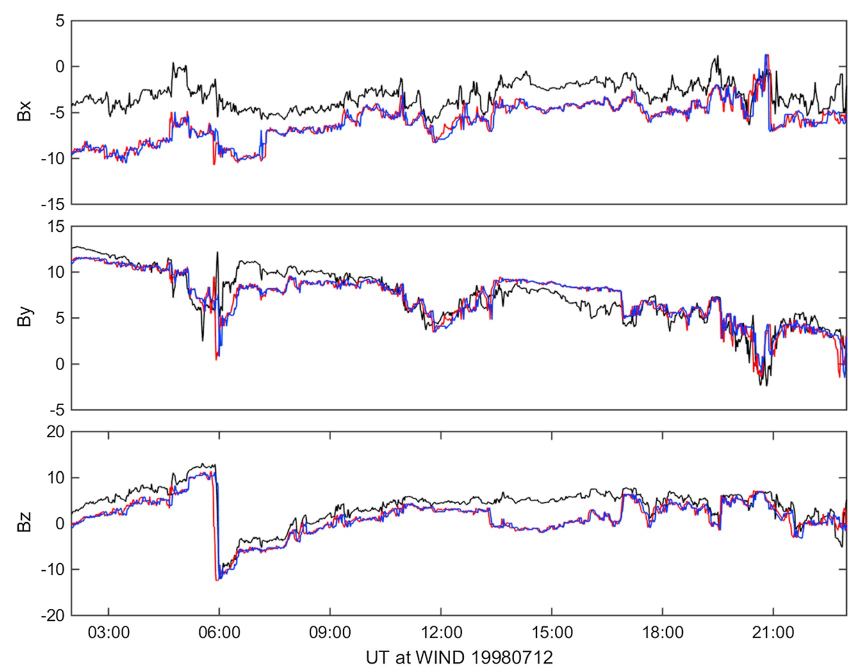

Figure 7. Another example like Figure 1 but this time for 12 July 1998. The differences between the two propagation methods are minimal except for the discontinuity observed in B_z at 0600 UT, for which better agreement is observed between the Wind spacecraft solar wind data (black line) and the MVCP technique propagated solar wind data (blue line) than the data propagated using convection delay (red line). The overall skill score for this event is 0.4% with a *z* value of 0.12.

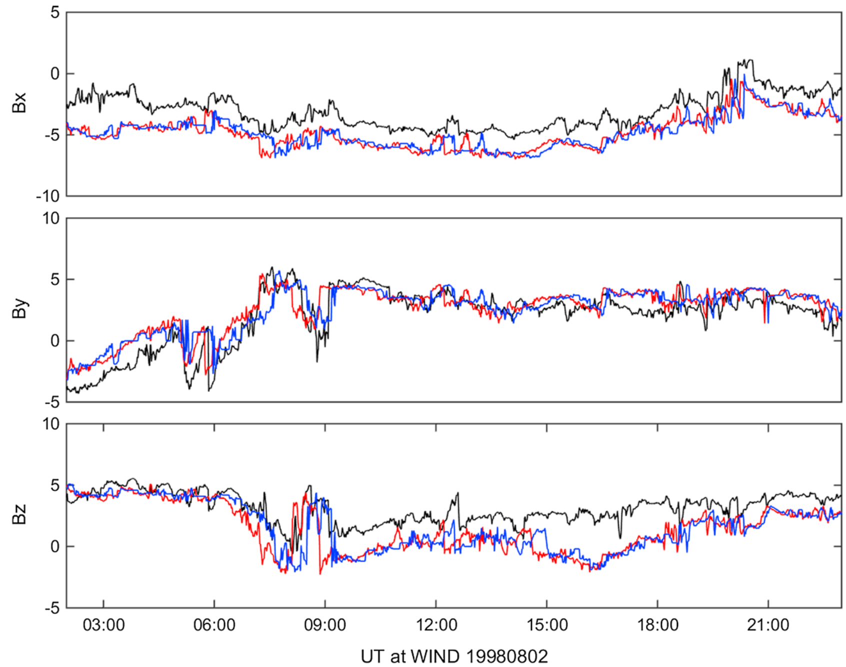

Figure 8. Same form as Figure 1 but for 2 August 1998. The skill score for this date is −9.9%, which is the worst skill score for all the dates in the validation study. In this example neither method does particularly well at reproducing the Wind data (black line). The *z* value for this date is 0.34, indicating that no significant difference is observed between the two propagation methods.

method is performing best overall, and indeed, neither method does particularly well at reproducing the Wind data (black line). The *z* value for this date is 0.34, indicating that no significant difference is observed between the two propagation methods.

Overall, we find that the tilted phase planes method does not show a statistically significant improvement in most situations. When there is not much separation from the Sun-Earth line or the solar wind phase plane normal is predicted to be approximately radial, then no significant deviation from convection delay is expected or observed. Improvements in the computed skill scores between the convection delay method and the tilted phase plane method were quite variable with the largest improvement of 18.2% observed on 27 July 1998 and the worst results, 9.9% worse than convection delay, observed on 2 August 1998. However, as can be seen via the examples discussed above, skill scores and two-sample *t* test scores do not provide a definitive assessment of which method performs best. Often a lack of solar wind structures or spacecraft separation distances greater than the solar wind-scale length confuse the results. While it makes intuitive sense to state that using the MVCP technique would produce better results during times when the tilt is large and there is a large separation between the two spacecrafts, this is not always what is observed. As demonstrated in Figures 6 and 7, for discontinuities, some improvement may be gained by using the MVCP technique over convection delay, and this topic in explored more fully in the next section of this paper.

4. Forecasting Sudden Impulse Arrival Times

The second product generated using the MVCP algorithm is the Solar Wind Transit Time product, which aims to predict the arrival time of a sudden impulse (SI) associated with the arrival of a shock or discontinuity within the solar wind. This product uses the same methods described previously to continuously propagate real-time solar wind data from an L1 orbiting satellite to a target location upstream of Earth, and for each solar wind parcel, an expected transit time is given. When a shock or discontinuity is detected in the solar wind, the predicted arrival time at the target location is specified. A space weather forecaster can then use the arrival time information to predict the onset time of an anticipated geomagnetic storm.

The first step in predicting the arrival time of a sudden impulse is to identify a potential shock in the solar wind data. To do this, we use the automatic shock detection algorithm described in *Cash et al*. [2014], which

Table 3. Validation of Sudden Impulse Arrival Time Predictions

Date (yyyy-mm-dd)	SI Observed at Earth[a]	Predicted Arrival Time[b]		Arrival Time Error (min)		Valid Tilt Computed[c]
		MVCP	Convection	MVCP	Convection	
1998-05-04	3:01	3:03:04	3:02:53	2.07	1.88	Yes
1998-08-26	6:51	7:06:35	7:06:35	15.6	15.6	Yes
1998-09-24	23:45	23:58:54	0:02:39	13.9	17.7	Yes
1998-11-08	4:51	4:52:50	4:55:24	1.8	4.4	Yes
1999-09-22	12:22	12:28:35	12:38:07	6.6	16.1	Yes
1999-10-21	2:25	2:23:15	2:24:58	−1.8	0.0	Yes
2000-08-11	18:45	18:56:09	18:56:09	11.2	11.2	Yes
2000-10-05	3:26	3:30:31	3:26:21	4.5	0.4	Yes
2000-11-06	9:48	9:49:17	9:49:16	1.3	1.3	Yes
2001-03-31	0:52	0:58:08	1:00:46	6.1	8.8	Yes
2001-04-11	13:43	13:48:49	13:48:49	5.8	5.8	No
2001-10-21	16:48	16:50:47	16:48:49	2.8	0.8	Yes
2001-10-28	3:19	3:21:57	3:21:57	3.0	3.0	No
2002-03-18	13:22	13:21:34	13:21:34	−0.4	−0.4	Yes
2002-03-20	13:28	13:43:27	13:42:07	15.5	14.1	Yes
2002-03-23	11:37	11:39:02	11:39:02	2.0	2.0	No
2002-04-17	11:07	11:06:25	11:05:27	−0.6	−1.6	Yes
2002-04-19	8:35	8:41:21	8:41:21	6.4	6.4	Yes
2002-04-23	4:50	4:50:55	4:50:55	0.9	0.9	Yes
2002-05-10	11:24	11:24:45	11:22:39	0.8	−1.4	Yes
2002-05-11	10:14	10:14:16	10:15:02	0.3	1.0	Yes
2002-05-18	20:07	20:08:35	20:08:35	1.6	1.6	Yes
2002-05-20	3:40	3:41:17	3:41:03	1.3	1.1	Yes
2002-05-21	22:05	22:23:20	21:55:52	18.3	−9.1	Yes
2002-05-23	10:50	11:00:23	10:55:32	10.4	5.5	Yes
2002-11-26	21:50	21:53:03	21:53:03	3.1	3.1	No
2003-03-20	4:40	4:47:00	4:46:50	7.0	6.8	Yes
2003-05-29	18:59	18:59:52	19:00:28	0.9	1.5	Yes
2003-08-17	14:21	14:25:38	14:24:37	4.6	3.6	Yes
2003-10-24	15:24	15:13:35	15:24:57	−10.4	1.0	Yes
2003-11-04	6:27	6:32:45	6:29:03	5.8	2.1	Yes
2004-01-22	1:37	1:37:35	1:38:03	0.6	1.1	Yes
2004-07-16	21:58	21:59:05	21:58:34	1.1	0.6	Yes
2004-07-22	10:36	10:38:13	10:40:12	2.2	4.2	Yes
2004-07-24	6:13	6:15:59	6:15:59	3.0	3.0	No
2004-07-26	22:49	22:49:46	22:49:44	0.8	0.7	Yes
2004-09-13	20:03	20:10:15	20:10:15	7.3	7.3	Yes
2004-09-22	6:37	6:33:26	6:33:26	−3.6	−3.6	No
2004-11-07	18:27	18:44:08	18:31:18	17.1	4.3	Yes
2004-11-09	9:31	9:44:59	9:44:59	14.0	14.0	Yes
2004-11-11	17:10	17:13:59	17:22:51	4.0	12.9	Yes
2004-12-05	7:47	7:45:13	7:45:13	−1.8	−1.8	Yes
2005-01-21	17:11	17:11:51	17:11:45	0.9	0.8	Yes
2005-05-28	4:36	4:45:37	4:48:12	9.6	12.2	Yes
2005-05-29	9:52	9:55:48	9:55:55	3.8	3.9	Yes
2005-06-14	18:35	18:38:41	18:38:41	3.7	3.7	Yes
2005-06-16	8:47	8:52:07	8:52:07	5.1	5.1	No
2005-07-10	3:37	3:38:44	3:38:44	1.7	1.7	No
2005-07-17	1:34	1:38:23	1:38:23	4.4	4.4	Yes
2005-08-01	6:41	6:44:47	6:44:47	3.8	3.8	Yes
2005-08-24	6:13	6:15:52	6:16:32	2.9	3.5	Yes
2005-09-02	14:19	14:19:28	14:19:28	0.5	0.5	Yes
2005-09-09	13:59	14:00:32	14:00:32	1.5	1.5	Yes
2005-09-12	6:24	6:29:59	6:29:59	6.0	6.0	No
2005-09-15	9:07	9:06:36	9:06:36	−0.4	−0.4	No
2006-01-01	14:06	14:02:07	14:12:20	−3.9	6.3	Yes
2006-07-09	21:36	21:38:10	21:35:29	2.2	−0.5	Yes
2006-12-14	14:14	14:18:27	14:17:32	4.5	3.5	Yes
2006-12-16	17:55	17:57:03	17:57:03	2.1	2.1	No
2007-11-19	18:11	18:08:26	18:08:28	−2.6	−2.5	Yes

Table 3. (continued)

Date (yyyy-mm-dd)	SI Observed at Earth[a]	Predicted Arrival Time[b]		Arrival Time Error (min)		Valid Tilt Computed[c]
		MVCP	Convection	MVCP	Convection	
2007-12-17	3:00	2:55:16	2:55:16	−4.7	−4.7	No
2008-11-24	23:50	0:00:12	23:59:44	10.2	9.7	Yes
2009-02-03	20:00	20:14:13	20:12:13	14.2	12.2	Yes
2010-04-05	8:26	8:27:35	8:27:35	1.6	1.6	Yes
2010-04-11	13:04	13:02:53	13:04:15	−1.1	0.3	Yes
2010-08-03	17:41	17:38:37	17:42:37	−2.4	1.6	Yes
2011-02-18	1:30	1:35:03	1:34:18	5.1	4.3	Yes
2011-03-29	16:02	16:00:53	16:00:48	−1.1	−1.2	Yes
2011-06-10	8:55	8:51:49	8:50:49	−3.2	−4.2	Yes
2011-08-05	17:51	18:08:17	18:08:17	17.3	17.3	No
2011-09-17	3:43	3:46:50	3:46:50	3.8	3.8	Yes
2011-09-26	12:34	12:38:55	12:39:35	4.9	5.6	Yes
2011-10-05	7:36	7:37:20	7:35:27	1.3	−0.6	Yes
2011-10-24	18:31	18:33:45	18:36:18	2.8	5.3	Yes
2011-11-28	21:50	21:56:49	21:58:28	6.8	8.5	Yes
2012-01-21	5:01	5:00:07	5:06:55	−0.9	5.9	Yes
2012-01-22	6:12	6:20:25	6:12:59	8.4	1.0	Yes
2012-03-15	13:07	13:04:41	13:09:09	−2.3	2.2	Yes
2012-04-23	3:20	3:20:28	3:20:38	0.5	0.6	Yes
2012-06-16	20:20	20:20:55	20:19:14	0.9	−0.8	Yes
2012-09-03	12:13	12:16:56	12:16:58	3.9	4.0	Yes
2012-09-30	11:31	11:29:03	11:29:09	−2.0	−1.9	Yes
2012-09-30	23:05	23:08:42	23:06:42	3.7	1.7	Yes
2012-10-31	15:39	15:40:03	15:40:04	1.1	1.1	Yes
2012-11-12	23:11	23:20:01	23:12:03	9.0	1.1	Yes
2012-11-26	5:12	5:03:36	5:15:13	−8.4	3.2	Yes
2013-02-16	12:09	12:08:58	12:08:58	0.0	0.0	Yes
2013-03-17	5:59	5:57:16	6:02:07	−1.7	3.1	Yes
2013-06-27	14:38	14:43:10	14:42:55	5.2	4.9	Yes
2013-07-09	20:49	20:55:02	20:53:42	6.0	4.7	Yes
2013-10-08	20:21	20:25:25	20:25:25	4.4	4.4	Yes
2014-02-07	17:05	17:10:27	17:12:33	5.5	7.6	Yes
2014-02-15	13:17	13:25:16	13:24:27	8.3	7.5	Yes
2014-02-20	3:20	3:14:34	3:29:35	−5.4	9.6	Yes
2014-04-20	10:56	11:01:48	10:58:21	5.8	2.4	Yes
2014-06-07	16:52	17:05:43	17:05:43	13.7	13.7	No
2014-06-23	23:08	23:07:13	23:03:59	−0.8	−4.0	Yes

[a]SI time as reported in the SWPC database of Geomagnetic Sudden Impulse Alerts.
[b]The "target location" for the predicted arrival time calculation was (30, 0, 0) R_E. Times are in hh:mm:ss format.
[c]If the computed tilted angle was invalid, then a flat phase front was assumed and both the MVCP technique and convection delay will predict the same arrival time.

looks for concurrent discontinuities in four solar wind parameters. Once a shock has been identified, the time for that phase front to reach Earth is computed using the MVCP technique with all of the optimization parameters the same as those shown in Table 1. Overtaking solar wind parcels occurring within the subsequent 3 min after the observed discontinuity are also considered, with the earliest arrival time selected as the predicted shock arrival time.

4.1. Validation of SI Arrival Time Predictions

To compare how well the MVCP technique does at predicting geomagnetic storm onset times compared to the standard convection delay method, we consider 97 events for which a sudden impulse is observed at Earth. Events were selected from the SWPC database of Geomagnetic Sudden Impulse Warnings and Alerts, which lists the time the shock was observed at ACE, the time the SI was observed at a ground-based magnetometer as determined by the forecaster on duty, and the deviation in the Earth's magnetic field in nanotesla observed at the magnetometer station. For each event used in this study, we required that the shock associated with the SI was also recorded on either the ACE Science Center's List of Disturbances and Transients (http://www.ssg.sr.unh.edu/mag/ace/ACElists/obs_list.html) or the Richardson and Cane list of

Near-Earth Interplanetary Coronal Mass Ejections (http://www.srl.caltech.edu/ACE/ASC/DATA/level3/icmetable2.htm), in addition to being listed in the SWPC database of Geomagnetic Sudden Impulse Warnings. The SI events included in this study occurred between May 1998 and September 2014. For each event, the predicted transit time from the location of the ACE spacecraft to a target location 30 R_E ahead of the Earth was computed for both the tilted phase plane method and convection delay. The target location was set to [30, 0, 0] R_E as this is the same location used as the input to the geospace model. Choosing a target location slightly ahead of Earth's magnetopause could result in early arrival time predictions; however, with both methods we find that the average arrival time is ~3 min later than the observed SI time. Also, as our main objective is to compare the performance of the tilted phase plane method to convection delay, the choice of the target location is not as critical as long as the delay is calculated to the same location for both techniques and no systematic differences between the two methods are observed.

Predicted arrival times and the error in the predicted arrival time for all 97 events are listed in Table 3. Out of the 97 events considered in this study, for 33 events (~34% of the time) both methods yielded the same predicted arrival time. For 28 events, the MVCP method performed better than convection delay, and for the remaining 36 events convection delay performed better than the MVCP technique. Within the error bars, both methods perform equally well. The mean absolute error for the MVCP method is 4.7 ± 4.4 min while the mean absolute error when using convection delay is 4.4 ± 4.3 min, indicating that the difference between the two propagation methods is not statistically significant.

Figure 9 shows the absolute error in the predicted transit time for each method for the 97 observed SI events. The absolute error associated with the convection delay method is given on the *x* axis and the *y* axis shows the absolute error associated with the MVCP method. The black line is a linear fit through the data. If both methods performed equally well then a line with a slope of 1.0 and an intercept of 0.0 would be expected (middle dashed blue line) and if the MVCP technique performs better, then we would expect a slope much less than one. For these 97 events a slope of 0.78 with a *y* intercept of 1.27 is obtained, indicating that a marginal improvement is obtained when using the MVCP technique instead of convection delay. The nonzero intercept could suggest a bias in the MVCP method versus convection delay; however, the large uncertainties associated with each method, as illustrated by the error bars plotted on the point in the upper left, indicate that this nonzero intercept is within the error bars and does not signal the presence of a bias in either method. These results suggest that on average there is not a significant difference between the two propagation methods considered in this study and within the error bars, both methods perform equivalently.

Another way to view the data is shown in Figure 10, in which the observed transit time is plotted versus the predicted transit time. Results for the tilted phase plane method are noted with blue diamonds, while red triangles mark results for the convection delay method. Error bars for each method are shown in black. A linear fit through the MVCP data yields a slope of 0.87 and a *y* intercept of 2.4, while a linear fit through the flat plane propagation data give a slope of 1.00 and a *y* intercept of -3.4. Given the ~4 min error bars for each method, there is no statistical difference between these two methods. However, it is interesting to note that the convection delay method has a slope of 1.00 and that by subtracting 3.5 min from each arrival time prediction, this method would agree reasonably well with the observed SI time, on average. One possible explanation for the late arrival time predictions could be using a too slow velocity in the transit time calculation (equation (5)). For this study, we use the solar wind velocity at the time of the observed discontinuity; however, computing the actual shock speed could potentially result in faster speeds and earlier arrival time predictions. Computing the shock speed in a real time, automated process is difficult but has been done recently [see *Vorotnikov et al.*, 2011] and could be incorporated into a propagation delay time algorithm such as those presented here. The error in the forecasted transit time for each method was also plotted versus the observed transit time with no discernable trend observed between the forecast errors and the transit time for either method.

The data were also sorted by tilt angle and separation distance from the Sun-Earth line in order to determine if a statistically significant improvement was observed under conditions when one would expect the tilted phase plane method to perform better—when a large tilt angle is present and when the separation from the Sun-Earth line is large. Table 4 gives the average absolute mean and standard deviation for three subsets of the data: (1) times when the spacecraft separation from the Sun-Earth line was $> 40\ R_E$, (2) times when the computed tilt angle was $> 40°$, and (3) times when both conditions were met. For each of these situations, there is no statistically significant difference between the two methods under investigation.

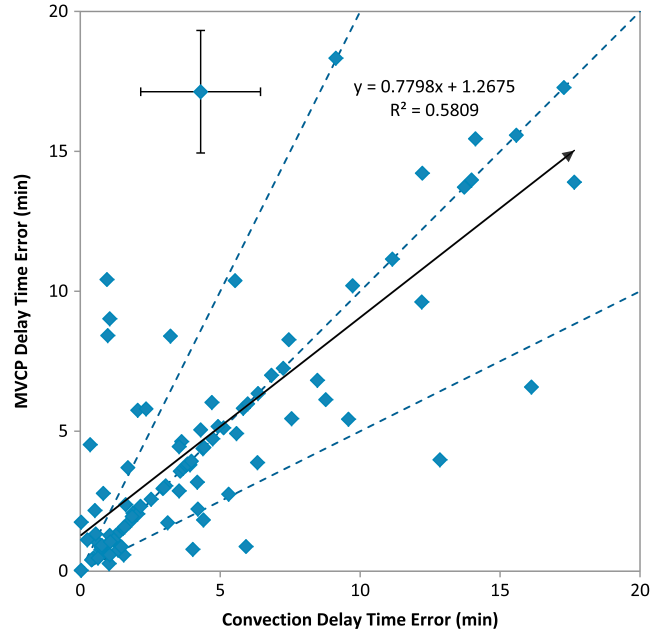

Figure 9. Absolute value of the error in the predicted transit time for 97 observed SI events. The *x* axis gives the error associated with the convection delay method, and the *y* axis shows the error associated with the MVCP method. The black line is a linear fit through the data with slope 0.78, a *y* intercept of 1.27 and an R^2 value of 0.58. The three dashed blue lines indicate slopes of 2, 1, and ½. Representative error bars based on the sample standard deviation are shown on the point in the upper left. The standard deviation for the MVCP method is 4.4 min and the standard deviation for the convection delay method is 4.3 min. Within the error bars, both methods perform equivalently.

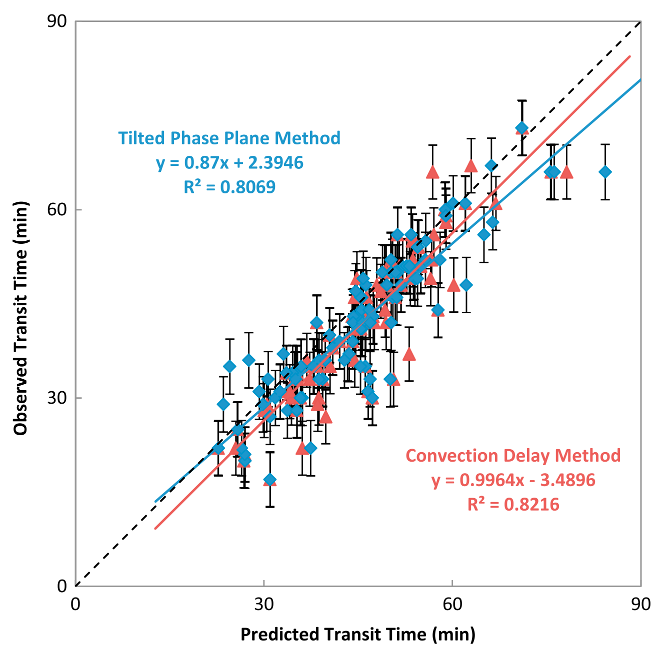

Figure 10. Observed transit time (from time the shock is observed at ACE to time that the sudden impulse was observed at Earth) versus the predicted transit time (from the time the shock is observed at ACE to the time the shock is expected to arrive at 30 R_E) for the tilted phase plane method (blue diamonds) and the convection delay method (red triangles). Error bars for each method are shown in black. The dashed black line indicates a slope of 1. The blue line shows a linear fit to the tilted phase plane data and the red line shows a linear fit to the flat plane propagation data.

Table 4. Average Absolute Error in Sudden Impulse Arrival Time Prediction

	Convection Error	MVCP Error
All 97 events	4.4 ± 4.3 min	4.7 ± 4.4 min
Events with spacecraft separation from Sun-Earth >40 R_E (21 events)	3.3 ± 3.4 min	4.0 ± 4.5 min
Events with computed tilt angle>40° (33 events)	6.0 ± 4.6 min	6.1 ± 5.6 min
Events with spacecraft separation from Sun-Earth >40 R_E and computed tilt angle >40° (five events)	5.0 ± 4.0 min	5.8 ± 7.2 min

4.2. Discussion of SI Arrival Time Prediction Results

In this study we have assumed that solar wind shocks have a flat phase front and the tilt angle is constant over the scale lengths separating ACE and the target location in front of Earth's magnetopause. However, if shock fronts have a more featured, wavy surface, this would cause the observed surface normal angle to vary along the shock surface, reducing the timing accuracy from a one-point measurement. Previous studies have indicated that this may be the case [*Heinemann and Siscoe*, 1974; *Chao*, 1984; *Gonzalez-Esparza and Bravo*, 1998]. *Heinemann and Siscoe* [1974] stated that the large-scale shape of the shocks fronts should suffer large deviations due to variations in the ambient solar wind and *Gonzalez-Esparza and Bravo* [1998] found that shocks and ejecta suffer significant deformations due to their interactions with different ambient solar wind streams. This could potentially explain why the MVCP method does not perform as well as one would expect, especially with increasing distance from the Sun-Earth line

In the future, additional methods could be employed to more accurately evaluate the performance of these two methods in computing the predicted transit time of a shock or discontinuity from the location of ACE or DSCOVR to Earth. Here we have assumed that sudden impulses are initiated by the interaction of a shock or discontinuity with the nose of the magnetopause; however, since a tilted phase front could initially impact the three-dimensional magnetopause at a location other than the subsolar point, depending on the particular tilt angle as well as on the monitoring spacecraft's location, it could be more accurate to calculate the arrival time with a 3-D magnetopause shape. Incorporating a 3-D magnetopause within the model would not influence the predicted arrival time of the flat plane convection delay but could potentially result in better agreement when using the MVCP technique. In addition, it could be worthwhile to consider the additional magnetosphere response time from the time when a shock is predicted to arrive at 30 R_E to the time when an SI is observed by a ground magnetometer station as this could introduce timing errors on the order of minutes.

Alternatively, instead of comparing the predicted arrival time to the observed SI time, one could use other spacecraft such as Cluster or Themis to determine the actual arrival time at a specified target location, as was done by *Mailyan et al*. [2008] and *Munteanu et al*. [2013] using Cluster data. *Mailyan et al*. found that the best predictions of the arrival times of discontinuities at Earth's magnetopause were obtained using the MVAB-0 method described by WK08. They found that 65% of the 198 events considered had a timing accuracy of ±5 min or betters and more than 30% had an arrival accuracy of ±2 min or less. This agrees with our findings in which 62% of the cases have a timing accuracy of ±5 min or better and more than 31% have an arrival accuracy of ±2 min or less.

5. Summary and Conclusion

The aim of the operational algorithm under investigation is to improve arrival time predictions, including improvements to the inputs driving geospace models and improvements to the predicted onset timing of geomagnetic storm warnings. In this paper we have presented a detailed validation study of the use of the MVCP technique as described by WK08 for real-time propagation of solar wind data from L1 to Earth, both for the case of propagating the continuous solar wind as well as for estimating the predicted arrival time of a solar wind discontinuity at Earth. For both of these use cases, we find no significant difference between the two propagation methods included in this study. Given the lack of a significant improvement, it is difficult to recommend a change to the use of the MVCP technique over the currently employed convection delay method, which provides a more straightforward determination of the predicted solar wind at Earth and works reasonably well in most situations.

These results agree with the findings published in two earlier papers which also explored the question of the best method to use when calculating the solar wind propagation delay [*Mailyan et al.*, 2008; *Pulkkinen and Rastätter*, 2009]. Via an investigation of four different methods for calculating the solar wind propagation time, including convection delay (flat delay), CP, and MVAB-0, *Mailyan et al.* [2008] found that while taking the orientation of the phase front and the separation between the solar wind monitor and target into account gives a more precise time delay estimation in most cases, the marginal improvement is obtained by a much more computational complex calculation compared to a simple flat plane calculation "which may be good enough." Our results are also corroborated by the *Pulkkinen and Rastätter* [2009] study, in which the tilted phase plane method of propagation described in WK08 was used to drive real-time global magnetohydrodynamic (MHD) simulations for the October 2003 Halloween storm event. The authors found that when using the phase plane-based propagation technique "the improvement is so modest that from the statistical viewpoint, the benefit over using the simpler propagation technique vanished for the studied storm period when the information is passed through real-time global MHD modeling process." The reason why MHD model runs don't see a large difference could be that the IMF can be relatively steady for 3 to 5 h, with only a few minutes timing error at the transitions that make up a small portion of the overall time period, as shown in Figure 1. Thus over long time periods, the more steady IMF mostly overwhelms the smaller influence of the timing errors at the edges.

The results from these two previous studies, combined with the in-depth validation study presented here, which applied the promising MVCP technique to the real-time propagation of solar wind data from L1 to near Earth, indicate that this method does not offer a significant improvement over the simple method of assuming a flat plane propagation time delay when implemented in a real-time operational setting. The standard method used by NOAA's Space Weather Forecast Office of assuming a flat plane propagation delay may continue to be the best option when accounting for simplicity and consistency in the observed results. Other methods, such as wavelet denoising [*Haaland et al.*, 2010; *Munteanu et al.*, 2013], have shown promising results and future efforts could focus on evaluating how well such a technique works when implemented in a real-time operational setting with a continuous stream of real-time solar wind data.

Acknowledgments
S.W.H. gratefully acknowledges funding through the NSF REU grant 11507020 and the University of Colorado as part of the LASP 2015 Research Experience for Undergraduates Program in Solar and Space Physics. We thank R. Steenburgh (SWPC) for constructive comments. ACE and Wind data were obtained from the Coordinated Data Analysis Web (CDAWeb: http://cdaweb.gsfc.nasa.gov). Solar wind shock information was obtained from the ACE Science Center's List of Disturbances and Transients (http://www.ssg.sr.unh.edu/mag/ace/ACElists/obs_list.html) and the Richardson and Cane list of Near-Earth Interplanetary Coronal Mass Ejections (http://www.srl.caltech.edu/ACE/ASC/DATA/level3/icmetable2.htm).

References

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*(A5), 10,465–10,479, doi:10.1029/1999JA000262.

Bargatze, L. F., R. L. McPherron, J. Minamora, and D. Weimer (2005), A new interpretation of Weimer et al.'s solar wind propagation delay technique, *J. Geophys. Res.*, *110*, A07105, doi:10.1029/2004JA010902.

Burlaga, L. F. (1969), Directional discontinuities in the interplanetary magnetic field, *Sol. Phys.*, *7*, 54–71.

Cash, M. D., J. S. Wrobel, K. C. Cosentino, and A. A. Reinard (2014), Characterizing interplanetary shocks for development and optimization of an automated solar wind shock detection algorithm, *J. Geophys. Res. Space Physics*, *119*, 4210–4222, doi:10.1002/2014JA019800.

Chao, J. K. (1984), Some characteristics of propagation of flare- or CME-associated interplanetrary shock waves, *Adv. Space Res.*, *4*, 327–330.

Collier, M. R., J. A. Slavin, R. P. Lepping, A. Szabo, and K. Ogilvie (1998), Timing accuracy for the simple planar propagation of magnetic field structures in the solar wind, *Geophys. Res. Lett.*, *25*, 2509–2512, doi:10.1029/98GL00735.

Gonzalez-Esparza, J. A., and S. Bravo (1998), Two spacecraft observations of transient shocks and ejecta in the interplanetary medium, *J. Geophys. Res.*, *103*, 29,643–29,650, doi:10.1029/98JA02824.

Haaland, S., et al. (2004), Four-spacecraft determination of magnetopause orientation, motion and thickness: Comparison with results from single-spacecraft methods, *Ann. Geophys.*, *22*, 1347–1365, doi:10.5194/angeo-22-1347-2004.

Haaland, S., G. Paschmann, and B. U. Ö. Sonnerup (2006), Comment on "'A new interpretation of Weimer et al.'s solar wind propagation delay technique" by Bargatze et al, *J. Geophys. Res.*, *111*, A06102, doi:10.1029/2005JA011376.

Haaland, S., C. Munteanu, and B. Mailyan (2010), Solar wind propagation delay: Comment on "Minimum variance analysis-based propagation of the solar wind observations: Application to real-time global magnetohydrodynamic simulations" by A. Pulkkinen and L. Raststätter, *Space Weather*, *8*, S06005, doi:10.1029/2009SW000542.

Heinemann, M. A., and G. L. Siscoe (1974), Shapes of strong shocks fronts in an inhomogeneous solar wind, *J. Geophys. Res.*, *79*, 1349–1355, doi:10.1029/JA079i010p01349.

Hill, T. P. (2011), Conflations of probability distributions, *Trans. Am. Math. Soc.*, *363*, 3351–3372, doi:10.1090/S0002-9947-2011-05340-7.

Hill, T. P., and J. Miller (2011), How to combine independent data sets for the same quantity, *Chaos*, *21*, 033102, doi:10.1063/1.3593373.

Horbury, T. S., D. Burgess, M. Fränz, and C. J. Owen (2001), Prediction of Earth arrival times of interplanetary southward magnetic field turnings, *J. Geophys. Res.*, *106*(A12), 30,001–30,009, doi:10.1029/2000JA002232.

Knetter, T., F. M. Neubauer, T. Horbury, and A. Balogh (2004), Four-point discontinuity observations using Cluster magnetic field data: A statistical survey, *J. Geophys. Res.*, *109*, A06102, doi:10.1029/2003JA010099.

Mailyan, B., C. Munteanu, and S. Haaland (2008), What is the best method to calculate the solar wind propagation delay?, *Ann. Ge-phys.*, *26*, 2383–2394, doi:10.5194/angeo-26-2383-2008.

Munteanu, C., S. Haaland, B. Mailyan, M. Echim, and K. Mursula (2013), Propagation delay of solar wind discontinuities: Comparing different methods and evaluating the effect of wavelet denoising, *J. Geophys. Res. Space Physics*, *118*, 3985–3994, doi:10.1002/jgra.50429.
Pulkkinen, A., and L. Rastätter (2009), Minimum variance analysis-based propagation of the solar wind: Application to real-time global magnetohydrodynamic simulations, *Space Weather*, *7*, S12001, doi:10.1029/2009SW000468.
Ridley, A. J. (2000), Estimations of the uncertainty in timing the relationship between magnetospheric and solar wind processes, *J. Atmos. Sol. Terr. Phys.*, *62*, 775–771, doi:10.1016/S1364-6826(00)00057-2.
Ridley, A. J., K. C. Hansen, G. To´th, D. L. De Zeeuw, T. I. Gombosi, and K. G. Powell (2002), University of Michigan MHD results of the Geospace Global Circulation Model metrics challenge, *J. Geophys. Res.*, *107*(A10), 1290, doi:10.1029/2001JA000253.
Russell, C. T., G. L. Siscoe, and E. J. Smith (1980), Comparison of ISEE-1 and -3 interplanetary magnetic field observations, *Geophys. Res. Lett.*, *7*, 381–384.
Sonnerup, B. U. Ö., and L. J. Cahill Jr. (1967), Magnetopause structure and attitude from Explorer 12 observations, *J. Geophys. Res.*, *72*(1), 171–183, doi:10.1029/JZ072i001p00171.
Sonnerup, B. U. Ö., and M. Scheible (1998), Minimum and maximum variance analysis, in *Analysis Methods for Multi-Spacecraft Data*, edited by G. Paschmann and P. W. Daly, pp. 185–220, Int. Soc. for Solid-State Ionics, Bern, Switzerland.
Sonnerup, B. U. Ö., S. Haaland, G. Paschmann, B. Lavraud, M. W. Dunlop, H. Rème, and A. Balogh (2004), Orientation and motion of a discontinuity from single-spacecraft measurements of plasma velocity and density: Minimum mass flux residue, *J. Geophys. Res.*, *109*, A03221, doi:10.1029/2003JA010230.
Vorotnikov, V. S., C. W. Smith, C. J. Farrugia, C. J. Meredith, Q. Hu, A. Szabo, R. M. Skoug, C. M. S. Cohen, A. J. Davis, and K. Yumoto (2011), Use of single-component wind speed in Rankine-Hugoniot analysis of interplanetary shocks, *Space Weather*, *9*, S04001, doi:10.1029/2010SW000631.
Weimer, D. R. (2004), Correction to "Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique,", *J. Geophys. Res.*, *109*, A12104, doi:10.1029/2004JA010691.
Weimer, D. R., and J. H. King (2008), Improved calculations of interplanetary magnetic field phase front angles and propagation time delays, *J. Geophys. Res.*, *113*, A01105, doi:10.1029/2007JA012452.
Weimer, D. R., D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness, C. W. Smith, and J. Watermann (2003), Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique, *J. Geophys. Res.*, *108*(A1), 1026, doi:10.1029/2002JA009405.
Wilks, D. S. (2011), *Statistical Methods in the Atmospheric Sciences*, *Int. Geophys. Ser.*, vol. 100, 3rd ed., Academic Press, Oxford.