Probabilistic Forecasting of the Dst Index

Robert L. McPherron

Institute of Geophysics and Planetary Physics, University of California Los Angeles, Los Angeles California

George Siscoe and Nancy U. Crooker

Center for Space Physics, Boston University, Boston, Massachusetts

Nick Arge

Phillips Lab, Hanscom Air Force Base, Boston, Massachusetts

The Dst index can be predicted by an autoregressive moving average filter that transforms the solar wind electric field and dynamic pressure into changes in the index. Integration of these changes gives Dst. The maximum lead time possible is less than 60 minutes. Longer lead times are probably unobtainable due to the stochastic nature of IMF Bz. In this paper we use photospheric observations of the solar magnetic field to predict Dst. Our technique uses the Wang-Sheeley-Arge (W-S-A) model to forecast the time variations of solar wind speed, IMF polarity, and IMF strength at Earth four days in advance. Since the waveform of Bz is incalculable from solar observations we make use of air mass climatology. In this technique the cumulative probability distribution is determined as a function of time relative to a stream interface. During the declining phase of the solar cycle persistent coronal holes create high speed streams. The interface between these streams and slower speed solar wind ahead provides a reference time that can be used to organize the cumulative probability distributions determined from historic data (climatology). We then use the predicted arrival of an interface and the associated cumulative distributions to predict the probability that $|\text{Dst}|$ will fall within a range of values. This technique works best for small to moderate storms occurring during the declining phase of the solar cycle. CMEs are not predictable by this model hence we do not expect this approach to work as well at solar maximum.

1. INTRODUCTION

The Dst index is the primary indicator of the state of the inner magnetosphere. Traditionally it has been associated with the strength of the symmetric ring current. The DPS theorem [Dessler and Parker, 1959; Sckopke, 1966] establishes a relation between the total kinetic energy of ring current particles and the magnetic perturbation they cause at the center of the Earth. The Dst index is an approximation to this perturbation calculated from a local time average of disturbances in the horizontal component of the magnetic field on the Earth’s surface. More recently it has been recognized that there are many magnetospheric currents that contribute to the Dst index [Campbell, 1996] and that throughout the main phase of a storm it is unlikely that the ring current is symmetric [Kamide et al., 1997; Sharma et al., 2002]. None-the-less as shown by [Carovillano
2 PROBABILISTIC FORECASTING OF Dst

and McGuire, 1968] the relationship holds approximately for asymmetric currents. If it were possible to predict this index some time in advance it would provide useful warnings of hazards to technological systems dependent on the space environment.

The Dst index has been found to be remarkably easy to forecast empirically on a short-term [McPherron and O’Brien, 2001; Temerin and Li, 2002]. The latter authors claim nearly 95% prediction efficiency using a scheme involving about 35 terms that transforms solar wind observations into the Dst index. Unfortunately, all such schemes require the solar wind velocity and interplanetary magnetic field (IMF-Bz) to drive the prediction model. At the present time the furthest upstream of the Earth that we can measure these parameters provides only 30-60 minutes advance warning. Any attempt to utilize a monitor further upstream is unlikely to be successful. The IMF Bz is a turbulent variable with a median time between zero crossings of about six minutes. The fluctuations in Bz are produced by a variety of processes in the solar wind including large-scale structures, stream interactions, and Alfven waves. As the solar wind flows from the upstream monitor towards the Earth waves enter or leave the stream, and stream interactions change Bz from what was measured at the monitor.

Even though the deterministic approach is likely to fail it is still possible to use probabilistic methods. In particular we have advocated the use of air mass climatology [McPherron and Siscoe, 2004]. In this approach one utilizes predictable properties of the solar wind (the air mass) to organize probability distribution functions for solar wind and magnetospheric variables. For example, during the declining phase of the solar cycle the solar wind is dominated by recurrent high speed streams. These streams overtake slow solar wind ahead of them and create an interface responsible for disturbances in the magnetosphere. It is possible to use past observations of the type and level of activity caused by these interfaces to construct probability distributions for magnetic indices as a function of time relative to the interface. Then if some method exists to predict the arrival of the interface some time in advance one can use the distributions to predict the range of activity expected at various times relative to the interface.

In this paper we describe a procedure for medium-term prediction (2-4 days in advance) of the Dst index. The technique utilizes observations of the photospheric magnetic field of the Sun to predict the solar wind velocity on a source surface around the Sun [Wang and Sheeley, 1990]. This velocity is then propagated to the Earth [Arge and Pizzo, 2000] and processed to detect stream interfaces. Knowing the time of arrival of an interface, probability distributions for Dst are used to forecast the expected level of activity. We apply this technique to data acquired during the last solar minimum in 1995 and discuss the problems found in implementing the technique.

2. HIGH-SPEED SOLAR WIND STREAMS

Two typical high-speed streams observed in January of 1995 are presented in Figure 1. Four panels show the velocity, density, temperature and azimuthal flow direction of the solar wind plasma. Labeled vertical dashed lines show the times of three interfaces between high and low velocity streams. The interface is clearly defined by the behavior of the four variables. The velocity increases rapidly from a low value of order 300 km/s to a high value of order 700 km/s. The density changes in the opposite direction increasing slowly from moderate to high value before the interface. Then, just at the interface, it drops rapidly to a low value. As the density drops the temperature rises. In the bottom panel the azimuthal flow angle switches from positive values, an outward flow westward of radial (in the direction of solar rotation), to an eastward value. The time most easily identified on higher resolution plots is the zero crossing of the flow angle. Below we use this time as the reference time in a superposed epoch analysis of solar wind and magnetospheric variables.

Other solar wind variables including the IMF Bz are organized by stream interfaces. Below we show that during 1995 the IMF Bz was biased negative and had large fluctuations for a day centered on nearly every interface. After the interface the combination of high velocity and occasional strong negative Bz drives geomagnetic activity and produces moderate magnetic storms.

3. SUPERPOSED EPOCH ANALYSIS

The behavior of the solar wind velocity for ten days centered on a stream interface is shown in Figure 2. The figure presents the results of a superposed epoch analysis of the solar wind velocity measured by Wind spacecraft during 1995. Epoch zero in this analysis was defined by an interactive procedure in which traces of the velocity, density, and azimuthal flow angle were displayed on a high resolution plot. First the plot of velocity was examined for a transition from a low to high speed solar wind stream. Then the density trace was examined near the time of the velocity transition to identify intervals of slow increase in density. Finally, the azimuthal flow angle was examined just after a rapid decrease in density. A positive to negative zero crossing of the flow angle was used to precisely define the stream interface. Ten-day segments of data centered on the time of each stream interface were then selected and stored as rows of an ensemble array. The figure illustrates the results obtained for velocity in the year 1995. Each row of
beginning of the transition from low to high speed at the time of most rapid increase in velocity.

4. CUMULATIVE PROBABILITY DISTRIBUTIONS FOR INDICES RELATIVE TO INTERFACE

The stream interfaces organize solar wind properties in such a way that geomagnetic activity is peaked near the time of the interface. Results for the Dst index are plotted in Figure 3. The figure shows that significant disturbances are most likely to occur in the six hours after the interface. Occasionally these storms are as large as $-150$ nT, but this happens less than $\sim 10\%$ of the time. The median Dst index at the peak of these storms is only $-35$ nT! Such events would not normally be classified as magnetic storms. The initial phase of these storms is very weak ($\sim 5$ nT) and would be undetectable in the original magnetograms. In contrast with stronger storms, the average recovery phase of these weak storms is nearly linear rather than exponential. Note also that storms can occur randomly with respect to the stream interface. These storms are caused either by multiple stream interfaces or by CMEs. The CME storms are often as large or larger than the interface storms. These CME storms are not predictable by the method described in this paper.

**Figure 1.** Three solar wind stream interfaces observed in January 1995. From the top down the traces include velocity, density, temperature, and azimuthal flow angle.

**Figure 2.** Superposed epoch analysis of solar wind velocity for 26 stream interfaces observed in 1995. Heavy lines are quartiles of the distribution. The ensemble was plotted with a different color. Heavy white lines denote the quartiles of the cumulative probability distribution at each sample point relative to epoch zero. It is evident that the stream interface is located a little after the beginning of the transition from low to high speed at the time of most rapid increase in velocity.
The quartiles plotted in Figure 3 define a simple pattern characteristic of a classic magnetic storm including initial phase, main phase, and recovery phase. However, this pattern is an artifact of superposed epoch analysis. An examination of the 26 individual storms contributing to the quartiles shows that only 6 storms behave approximately in the manner suggested by this pattern. Four of the storms were “double dip” storms in which there was a second decrease in Sym-H before the preceding storm had recovered. Ten of the storms were actually multiple storms in which a second storm followed the first within the 10-day interval of the plot. Several of the storms never recovered in the 10-day interval or remained steadily depressed until a new storm started. We have selected intervals of recovery for each storm and performed both linear and exponential fits to the recovery phase. The quality of the two types of fits are indistinguishable both averaging about 48% prediction efficiency. The apparent linear recovery time given by $T = D_0/(em)$ where $D_0$ is Sym-H at minimum, $m$ is the slope of the linear recovery, and $e = 2.7183$ averages 2.3 ± 1.4 days. The corresponding average $e$-folding time from the exponential fit is 3.6 ± 3.5 days. In either case these apparent recovery rates are much longer than expected from either convection (a few hours) or charge exchange during northward field (16 hours) [O’Brien and McPherron, 2000]. This fact, and the frequent occurrence of double dip or second storms suggests that the IMF must turn southward rather frequently in the interval following a stream interface, maintaining a depressed Sym-H by frequent injections that decrease in intensity with time. In fact, in selecting stream interfaces we found it rather common to have multiple upward steps in solar wind velocity and zero crossings of the azimuthal flow angle. In such cases we selected only the first in a sequence of zero crossings. Clearly this structure poses a problem for probabilistic forecasting unless some method is developed to resolve these multiple interfaces in the solar observations.

We have analyzed a variety of other indices in a similar manner and present the quartiles of the distribution function in Figure 4. The top panel shows the velocity quartiles from Figure 2. The second panel shows quartiles of the GSM dawn-dusk electric field. The distribution is asymmetric near the interface with both the median and lower quartile negative. It was this effect of the stream interface on the IMF Bz that caused moderate activity at the interface during 1995. In 1996 this bias in Bz was absent but fluctuations were still

Figure 3. Superposed epoch analysis of the Dst index for 56 stream interfaces observed in 1995 and 1996. Heavy lines are quartiles of the distribution.

Figure 4. Quartiles of the 1995 ensembles of stream interfaces for velocity, VBs, ap, Sym H, Log Pc 5 Power, and log flux of relativistic electrons at noon synchronous orbit.
enhanced. As a consequence interface storms in 1996 were weaker than in 1995. Note also in Figure 4 that for several days after the interface the median electric field was slightly negative and the separation of the quartiles was larger than before the interface. This is primarily due to amplification of Bz fluctuations by the higher velocity occurring at this time. The consequence of this behavior is evident in the third panel which shows quartiles of the 3-hr ap index ensemble. Magnetic activity is strongest at the interface but decays slowly for many days. The fourth panel shows the quartiles of sym-H already discussed. It is apparent that the electric field around the interface produces the main phase of these weak magnetic storms, and then the prolonged decay of the electric field after the interface produces the linear recovery. Apparently the ring current is in quasi-equilibrium with injection nearly matching decay during the storm recovery phase. The fourth panel presents a measure of Pc 5 power injection nearly matching decay during the storm recovery. Apparently the ring current is in quasi-equilibrium with the stream interface detector described next.

5. A STREAM INTERFACE DETECTOR

To predict magnetospheric activity associated with storms two things are needed. The first is a method of predicting the solar wind velocity profile several days in advance of its arrival at the Earth. The second is a method for identifying the stream interface from the predicted velocity profile. The technique for predicting the solar wind velocity at 1 AU 2-4 days in advance has been described elsewhere [Arge and Pizzo, 2000; Wang and Sheeley, 1990] and because of space limitations we do not repeat this here. For our purposes we assume that a real time procedure generates the solar wind velocity time series about three days in advance. We process this series with the stream interface detector described next.

The detector is based on simple pattern recognition. We start with the shape of the median velocity profile for two days centered on the interface. We remove the mean and normalize the patterns so the sum of the absolute value of the points is 1.0. Then we standardize the predicted time series by removing the mean and dividing by the standard deviation. We next convolve the pattern with the standardized velocity. When the shape of the two curves is similar the output of the convolution approaches 1.0. We take the center of the time interval in which the response is greater than 0.5 as the time of the interface. In practice the predictions are made approximately every 8 hours so we resample both series to precisely 8-hour cadence. This reduces the interface pattern to seven points. Note that at least one day beyond the interface is required to identify the position of the interface. This reduces the advance warning by one day.

The 7-point pattern appears to work well identifying a large number of stream interfaces at a point midway along the rising ramp of predicted velocity. However, a comparison with the interfaces detected manually in high resolution data shows that the W-S-A interfaces are systematically early. Statistically the bias is −0.25 days. This appears to be a characteristic of the predicted time series and not a bias of the detector. Apparently additional work is needed to optimize the W-S-A predictions. In 1995 there were eight manual interfaces not detected by our automatic procedure. Four of these appear to be the consequence of missing data. Another seems to be a poor manual choice of an interface. At least one miss is a result of an outright failure of the solar wind model to predict a rise in velocity in early March of 1995.

6. EXAMPLE PREDICTIONS

Using the stream interface detector we predict the arrival of a stream interface two days in advance. Knowing this we use climatology of the Dst index relative to the interface to predict the quartiles of expected activity. The simplest climatology for the years 1995-1996 depending only on time relative to the interface is illustrated in Figure 5.

Since the time resolution of the W-S-A predictions is roughly 8 hours we have chosen to create cumulative probability distributions (cdfs) averaged over nine hour intervals. To provide some smoothing we overlap these estimates by three hours - the resolution of the ap index. Using the 16 day ensemble shown in Figure 3 we produced 125 distribution functions centered about epoch zero. There is a significant difference between the distributions two days before the interface (heavy line on right side) and ½ day after the interface (heavy line on left side). For example, the probability that Sym-H will be below −50 nT two days before an interface is zero! In contrast, there is a 4% probability that it will be below −100 nT ½ day after the interface. These distributions provide the “climatology” used to make a probabilistic forecast of the Sym-H index. In the following we have chosen the upper and lower quartiles of these distributions as a function of time relative to an interface to bracket the expected value of Sym-H.
A data segment showing our prediction of the quartiles of Sym-H in 1995 is presented in Figure 6. The top panel compares the observed velocity (thin gray curve) at 90-second resolution with the predicted velocity (dark dotted curve) at 8-hour resolution. The second panel compares the observed Sym-H index (thin gray lines) with our prediction of the upper and lower quartiles expected from climatology (heavy black lines). Based on Figure 3 we expect Sym-H to become more negative at the times of stream interfaces. The predicted times of these interfaces is shown by crosses at the top of the bottom panel. The actual times are shown by dotted circles. There were two interfaces detected manually (04/23 and 05/16) that were not detected automatically. In both cases this was a result of missing data in the predicted velocity curve. Four other interfaces were properly identified hence the envelope defined by the quartiles of the cdf drops indicating a magnetic storm is expected. The thin gray lines corresponding to observations also drops at these times indicating a successful prediction. Note that every storm is predicted to have the same magnitude since the only climatology used is time relative to the interface. A measure of quality of the predictions can be obtained by calculating how frequently the observations lie above the upper quartile or below the lower quartile. Statistically we expect the result to be 25% in each case. The results for the entire year are 25% above and 20% below, close to expectations.

7. DISCUSSION

In this paper we have presented preliminary results describing a method of forecasting the probability of magnetic storms about 2 days in advance. The technique utilizes Earth-based observations of the photospheric solar magnetic field. The field is expanded to a surface around the Sun and converted to “flux tube expansion factors”. The velocity of solar wind emitted from the sub-Earth point on the Sun is inversely proportional to this factor [Wang and Sheeley, 1990]. This wind is propagated to the Earth taking into account compression of the slower wind that precedes fast streams [Arge and Pizzo, 2000]. Predictions of the speed at 1 AU are typically made 3-4 days in advance of arrival of the wind at the Earth. The known pattern of the velocity at a stream interface is convolved with the predicted time series obtaining the expected time of arrival of the interface. At least one day warning is lost since our detector requires one day of data after the interface. Cumulative distribution functions determined from historical data as a function of time relative to the interface are then used to predict the values of Sym-H which are likely to bound the observed index with a given probability.

The quality of the predictions is dependent on the accuracy of the W-S-A model for the solar wind velocity. This model works particularly well during the declining phase of the solar cycle when large coronal holes near the Sun’s equator produce high-speed streams that subsequently interact with slow-speed wind emitted earlier from regions to the west of the coronal hole. We have found two problems in the current W-S-A predictions. First there was a bias in the velocity profile for 1995 so that the predicted interfaces...
arrived 0.25 days early. We have not removed this bias in the predictions because it is less than a single sample interval. In addition there is a random error in the predicted arrival times of about 1.3 days. This is significant to the potential users of the forecasts. Also this model does not predict the occurrence of coronal mass ejections (CME) that are responsible for much larger magnetic storms than those caused by the recurrent high-speed streams. Since CMEs also occur at solar minimum it is inevitable that there will be events not anticipated by our procedure.

It is possible to make a number of improvements in this procedure. First, data from more than one solar observatory could be used to eliminate problems caused by missing data at one location. Also this would have the advantage of using features in the solar field that are seen at all observatories rather than just one. Second it is possible to improve the quality of the predicted speed on the solar source surface using a more complex model that includes location of the sub-Earth point relative to a coronal hole as well as the flux tube expansion factor [Arge and Pizzo, 2000]. It is also likely that the kinematic model for the propagation and interaction of high speed streams can be improved by parameterizing the model and then optimizing these parameters with historical data.

It may also be possible to improve the detection of the stream interfaces in the predicted velocity. We have used a centered pattern to convolve with the series. A pattern containing only data before the interface might add several 9-hour samples of advance warning. Another possible improvement would use precursors in the measured solar wind. For example the solar wind density (not shown here) begins to increase about two days before a stream interface, and is obviously elevated one day before. About 12 hours before the interface a variety of indicators including density, azimuthal flow angle, and magnetic field strength all show significant perturbations. A probabilistic stream interface detector utilizing all of these parameters could be developed. Such a detector could be used to refine the W-S-A predicted interface arrival times.

Another way to improve predictions is to include two other parameters predictable by W-S-A. These are the polarity of the IMF either towards or away from the Sun, and the strength of the IMF (not implemented). The polarity is important for the Russell-McPherron effect [Russell and McPherron, 1973] which increases the strength of activity significantly when the IMF satisfies the rule “spring to and fall away”.

8. REFERENCES


Robert L. McPherron, Institute of Geophysics and Planetary Physics, University of California Los Angeles, Los Angeles, CA 90095-1567, USA, e-mail:rmcperron@igpp.ucla.edu

George Siscoe, Center for Space Physics, Boston University, Boston, Massachusetts, 02215, USA, siscoe@bu.edu

Nancy Crooker, Center for Space Physics, Boston University, Boston, Massachusetts, 02215, USA, siscoe@bu.edu

Nick Arge, AFRL/VSBXS29 Randolph Rd, Hanscom Afb, MA 01731-3010, USA, nick.arge@hanscom.af.mil