# Solar wind parameters for magnetospheric magnetic field modeling

Z. Qin,[1] R. E. Denton,[1] N. A. Tsyganenko,[2] and S. Wolf[3]

Received 12 October 2006; revised 12 July 2007; accepted 12 August 2007; published 22 November 2007.

[1] Magnetospheric magnetic field models are crucial for many space weather applications. However, the latest empirical models require solar wind and IMF data, which are not always available. Data gaps are especially common for times before the launch of the WIND spacecraft at the end of 1994, but even after then there are data gaps. We present a method to interpolate the solar wind characteristics across data gaps and to evaluate the *W* parameters needed for the TS05 model (Tsyganenko and Sitnov, 2005). Within some distance from the edge of a data gap, the solar wind parameters from our method yield a better estimate of the observed magnetic field than that which could be found using average values of the parameters. Deep within data gaps (far from measured values), the interpolated parameters are reasonable, or typical values, no better or worse than average values. We have created a database of hourly data with solar wind characteristics, *G*, and *W* parameters from 1963 to 31 May 2007, which is sufficient for use in all the Tsyganenko models, including the latest TS05 model. Our comparisons of the model and observed magnetic field at geosynchronous orbit give an estimate of the error in the model field as a function of status parameters defined by the interpolation scheme. We also show that the model field is on average just as accurate using the hourly data as that based on 5 min data (at least at geosynchronous orbit).

**Citation:** Qin, Z., R. E. Denton, N. A. Tsyganenko, and S. Wolf (2007), Solar wind parameters for magnetospheric magnetic field modeling, *Space Weather*, *5*, S11003, doi:10.1029/2006SW000296.

## 1. Introduction

[2] The vector magnetic field in the magnetosphere is correlated with a number of parameters, such as the dipole tilt angle and solar wind characteristics. During the last century, many spacecraft have been launched, and in situ measurements of the magnetospheric magnetic field became available. However, these spacecraft measurements only sample a small region of space at any given time. Thus geomagnetic field models are essential for space physics, and there exist several well-known and widely used empirical approximations [*Tsyganenko*, 1989, 1996, 2002a, 2002b; *Tsyganenko et al.*, 2003; *Tsyganenko and Sitnov*, 2005, hereinafter referred to as TS2005]. Solar wind and IMF data observations have been relatively continuous since the launch of the WIND spacecraft at the end of 1994. Before that time, however, the primary source of solar wind observations came from measurements by the IMP-8 spacecraft, and there were large periods of time for which IMP-8 was not in the solar wind. During those time periods, there were data gaps in the solar wind data. Even after 1994 there have been many data gaps in the solar wind data, though they are not as frequent, and usually are shorter.

[3] In this paper we use the OMNIWEB data (http://omniweb.gsfc.nasa.gov/) for the solar wind and IMF values. The OMNIWEB data set assimilates many solar wind observations, including those by IMP-8, WIND, and ACE spacecraft, and the data taken by individual spacecraft are properly time-shifted to take into account their spatial separation with respect to Earth. The solar wind data, especially the particle density, taken simultaneously but at different locations by different spacecraft, can differ significantly. This can even be the case for different instruments on the same spacecraft, or different data analysis algorithms (i.e., moments versus nonlinear analysis of distribution functions). While a more detailed discussion of this issue will be given in section 3, here we note that using a single standard interplanetary medium data set is highly desirable for space weather studies, especially in benchmarking quantitative models. Using the OMNIWEB hourly data, we have created a database of solar wind characteristics, *G* and *W* parameters, from 1963 to 31 May

[1]Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire, USA.
[2]USRA, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[3]Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan, USA.

**Table 1.** Correlation Times

| Quantity | $\tau$, h |
|---|---|
| $B_y$ | 1.47 |
| $B_z$ | 0.74 |
| $N$ | 2.22 |
| $V$ | 13.3 |
| $P_{dyn}$ | 1.68 |

2007. We will show that hourly values are sufficient to calculate the model magnetic field in the TS05 model. However, the interpolation technique we describe here can easily be used with the 5 min data now also available at the OMNIWEB data site.

[4] In the following section 2, we will describe the current status of the OMNIWEB data and how we interpolate across data gaps. In section 3, three kinds of comparisons will be made to compare our solar wind parameters to those found using 5-min data. In section 4, we describe our database of solar wind characteristics, *G* and *W* parameters. We study statistically the agreement between the *W* parameters calculated using the hourly and 5 min data. We also examine the agreement between the model and observed magnetic field values using solar wind IMF and data and *W* parameters derived from hourly or 5 min data. Section 5 summarizes our results.

## 2. Interpolation

[5] To run the empirical magnetic field models, components of the interplanetary magnetic field (IMF) and several solar wind parameters may be needed, such as the $B_y$ and $B_z$ components of the IMF, the solar wind speed $V$, proton density $N$, and the ram pressure $P_{dyn}$. In addition, most models require Dst. The T89 model [*Tsyganenko*, 1989] requires Kp. The latest models require subsidiary parameters, calculated from the solar wind and IMF data, the *G* parameters for the T02 [*Tsyganenko*, 2002a, 2002b] and TSK03 [*Tsyganenko et al.*, 2003] models and the *W* parameters for the TS05 model [TS2005]. For a particular model, not all the mentioned parameters are needed. The TS05 model requires $B_y$ and $B_z$, $P_{dyn}$, Dst and a set of six *W* parameters. Except for the *G* and *W* parameters, hourly averaged values of all these quantities can be downloaded from OMNIWEB. The data ranges from 1963 to the present. However, there are many gaps with no data, especially during the 1960s. The *W* parameters are derived from the solar wind characteristics, which should be 5-min averages using *Tsyganenko and Sitnov*'s [2005] method. Here we describe a method to interpolate across the data gaps and calculate the *W* parameters that are required by the TS05 model.

### 2.1. Correlation Time

[6] To do the interpolation, we first calculate the correlation time for each of the solar wind quantities. Suppose $f(i)$ is the $i$th value of a solar wind quantity at time $i$, and $d(i)$ is the difference between $f(i)$ and the 20 d average of this solar wind quantity (for a description of the 20 d average, see the next to last paragraph in section 2.2). Then the correlation function [see, e.g., *Press et al.*, 1997] is defined as

$$C(j) = \frac{\sum_s \sum_{i=1}^{n-j} d(i) \times d(i+j)}{\sum_s \sum_{i=1}^{n-j} d(i) \times d(i)}, \tag{1}$$

where $j$ represents a time difference (from $i$ to $i + j$ measured in hours), and the outer sum in both the numerator and denominator is evaluated over every contiguous section of data $s$ (sections without gaps) for which the number of data points $n$ is greater than $j$. This sum over data sections $s$ extends over the entire 40 a of the OMNIWeb database. (We evaluate the correlation time using $d$ rather than $f$ because we want to find out if using a value related to the last measured value is preferable to using an average value. If the correlation function were defined using $f$, we would find out if using the last measured value is preferable to using a value of 0.) Note that $C(0) = 1$ and that if $f$ is constant (and hence $d$), $C(j) = 1$ for any $j$. For the solar wind quantities, however, $C(j)$ will decrease as $j$ increases, representing the fact that values of $f$ farther away in time will be less correlated than those closer in time. Now suppose that there is a data gap. We want to evaluate the amount of time that it would be safe to use the last measured value $f(i)$ measured at time $i$. Clearly, if $C(j) \approx 1$, it is safe to use $f(i)$ for the value at time $i + j$, but if $C(j) \approx 0$, we cannot expect that the value $f(i + j)$ should be similar to $f(i)$. We choose the correlation time $\tau$ as the time at which $C(\tau) = 0.8$ (where the choice of 0.8 is somewhat arbitrary). For each of the parameters $B_y$, $B_z$, $N$, $P_{dyn}$ and $V$, a corresponding $\tau$ (in hours) is derived as described above, and these are listed in Table 1.

[7] It must be emphasized that the correlation time evaluated using hourly average data may not be equal to that found using higher resolution data. Actually, the correlation time for $V$ is large, and the value in Table 1 for the correlation time of $V$ is reliable. However, the correlation time for $B_z$ would be lower if high resolution data were available, indicating that there is normally a greater benefit to having high-resolution data for $B_z$ than for $V$. Of course, there may be large sudden jumps in $V$ as well, but these are relatively more rare.

### 2.2. Interpolation Method

[8] Depending on the correlation time $\tau$ and the length in time of a data gap, three kinds of interpolation may be used. Suppose $\Delta T$ is the length in time of the gap with $f_1$ equal to the last value measured before the gap at time $t_1$ (in hours) and $f_2$ equal to the first value measured after the gap at time $t_2$ (also in hours). Then (1) if $0 \leq \Delta T \leq 2\tau$, linear interpolation from $t_1$ to $t_2$ is used across the gap. (2) If $2\tau < \Delta T \leq 4\tau$, values at the left edge of the gap from $t_1$ to $t_1$ +

$(\Delta T - 2\tau)/2$ are set equal to $f_1$, values at the right edge from $t_2 - (\Delta T - 2\tau)/2$ to $t_2$ are set equal to $f_2$, and linear interpolation is used across the middle part of the gap of width $2\tau$. (3) If $\Delta T > 4\tau$, values at the left edge of the gap from $t_1$ to $t_1 + \tau$ are set equal to $f_1$, values at the right edge of the gap from $t_2 - \tau$ to $t_2$ are set equal to $f_2$, and values within the middle region of the gap are given by

$$f(t) = f_1 w_1 + f_1 w_2 + w_{av} f_{av}(t), \tag{2}$$

where $f_{av}(t)$ is linearly interpolated from 20 d averages of $f$, and

$$w_1 = \max\left(1 - \frac{t - (\tau + t_1)}{2\tau}, 0\right), \tag{3}$$

$$w_2 = \max\left(1 - \frac{(t_2 - \tau) - t}{2\tau}, 0\right), \tag{4}$$

and

$$w_{av} = 1 - (w_1 + w_2). \tag{5}$$

While this last formula may look complicated, the idea is really very simple. The value of $f$ will be $f_1$ near the left edge of the gap, $f_2$ near the right edge of the gap, $f_{av}(t)$ in the middle of the gap, and linear interpolated values from $f_1$ to $f_{av}(t)$ and from $f_{av}(t)$ to $f_2$ for the other two regions, respectively. On the basis of this method, a measured value can have an influence on interpolated values up to a time difference of at most three correlation times (cases 2 and 3 above), though the influence of a measured value starts dropping off to zero after the first correlation time. Therefore we can consider that a measured value influences the interpolated values up to a time difference of about two correlation times.

[9] There is one problem with the above procedure. There are sometimes gaps in the solar wind characteristics of more than 20 d, especially in the early years of OMNIWEB data (1960s). To get the 20 d averages, we use the following procedure: (1) Take the yearly average of the data from 1963 to the present, which will be used when interpolating the 20 d average; there are no gaps in the yearly average. (2) Take the 20 d average of the raw data. This average does have gaps. (To be more precise, we divide each year into 18 segments of equal length, approximately equal to 20 d, and evaluate the average within these segments.) (3) Calculate the correlation time, $\tau_{20}$, analogous to the correlation time defined above, but now using 20 d averages. (4) Do the interpolation to get the 20 d averages using a procedure analogous to that described above for the hourly data; in this case, linearly interpolated values of the yearly averages are used for $f_{av}$. Once all the 20 d averages are determined, we can use the procedure described above to fill in the gaps of the hourly data.

### 2.3. *W* Parameters

[10] *Tsyganenko and Sitnov* [2005] used 5-min data to calculate the $W$ parameters that are used in the TS05 model [TS2005]. There are six of these, each of which is used to define a different magnetospheric current system. The $W$ parameters can be expressed as

$$W(t_i) = \frac{r}{12} \sum_{k=1}^{i} S_k \exp[-r(t_i - t_k)], \tag{6}$$

where the $t_k$ values are time values available every 5 min, but still evaluated in hours (*Tsyganenko and Sitnov* [2005] used a corresponding expression with $t_k$ expressed in minutes), and

$$S_k = \left(\frac{N_k}{5}\right)^{\lambda} \left(\frac{V_k}{400}\right)^{\beta} \left(\frac{B_{sk}}{5}\right)^{\gamma}, \tag{7}$$

with $N$ expressed in particle per cm$^3$, $V$ expressed in km/s, and $B_s$ evaluated in nT. The quantity $B_s$ is a function of the $z$ component of the interplanetary magnetic field $B_z$; $B_s = 0$ if $B_z \geq 0$, and $B_s = -B_z$ if $B_z < 0$. For use in the $S$ terms only, the 20 d average of $B_z$ used in the interpolation scheme of section 2.2 was evaluated as $-\langle B_s^{\gamma}\rangle_{20}^{1/\gamma}$, where $\langle B_s^{\gamma}\rangle_{20}$ is the 20 d average of $B_s^{\gamma}$. Using this formula, the hourly values of $B_z$ will be equal to the measured values when these are available. Within gaps, $B_s^{\gamma}$ will be equal to $\langle B_s^{\gamma}\rangle_{20}$, which is what we want.

[11] The mathematical form of equation (6) suggests an iterative scheme for evaluating the $W$ parameters. Note that $W(t_i)$ can be written as

$$W(t_i) = \frac{r}{12} S(t_i) + W(t_i - 1/12) \exp[-r/12]. \tag{8}$$

We used an equation like this but with the source functions $S(t)$ evaluated every 15 min from linear interpolation of the hourly data. Thus

$$W(t_i) = \frac{r}{4}\Big[S(t_i) + S(t_i - 1/4)e^{-r/4} + S(t_i - 1/2)e^{-r/2} + S(t_i - 3/4)e^{-3r/4}\Big] + W(t_i - 1)e^{-r}. \tag{9}$$

In the TS05 model [TS2005], there are six groups of parameters, $\lambda$, $\beta$, $\gamma$ and $r$, for six different field modules, and there are six corresponding $W$ parameters, $W_1$ through $W_6$.

## 3. Comparison Between the Interpolated Data and the 5-Min Data

[12] Now we will compare the 5-min solar wind characteristics and $W$ parameters calculated using the 5-min data to the hourly solar wind characteristics and $W$ parameters calculated using the method we described above. This comparison is done using 5-min resolution values calcu-

**Table 2.** F and $\lambda$ for $W_1$ Through $W_6$

| $i$ | $F_i$ | $\lambda_i$ | $1.18^{\lambda i}$ |
|---|---|---|---|
| 1 | 1.07 ± 0.01 | 0.39 | 1.03 |
| 2 | 1.03 ± 0.01 | 0.46 | 1.08 |
| 3 | 1.25 ± 0.00 | 0.39 | 1.03 |
| 4 | 1.10 ± 0.01 | 0.42 | 1.07 |
| 5 | 0.98 ± 0.01 | 0.41 | 1.07 |
| 6 | 1.14 ± 0.10 | 1.29 | 1.23 |

lated for the year 2000. These data were measured by the ACE spacecraft, and a shift in time was implemented using the solar wind speed (as in the OMNIWEB data) to account for propagation to the Earth. There are a few gaps in the 5-min data, and we did not interpolate across these. After each gap, the $W$ parameters were set equal to the values from the hourly data in order to make a fair comparison.

[13] Our final $W$ parameters based on hourly OMNIWEB data are adjusted by multiplication by a factor $F$. This factor could arise in part from the fact that the proton density on OMNIWEB is different from that used by *Tsyganenko and Sitnov* [2005]. The OMNIWEB densities were calibrated on the basis of WIND measurements, but TS2005 in many cases used both ACE and WIND data, or only those of ACE (when WIND was too far from the Sun-Earth line; see *Tsyganenko et al.* [2003] for details). The ACE proton density is larger than the WIND proton density by up to 18%, depending on the solar wind speed [*King and Papitashvili*, 2005], and the OMNIWEB density values measured by ACE have been adjusted to make them consistent with values from WIND. We used the method of least squares to calculate the multiplier $F$, minimizing the difference between the values of $W$ calculated from the hourly data and those calculated from the 5-min data for years 1996 to 2000 during which the space magnetometer data in 37 major events were used to develop the TS05 model [*Tsyganenko and Sitnov*, 2005]. Table 2 shows these multipliers for $W_1$ through $W_6$. Note that the multipliers are all positive, and $\geq 1$ except $F_5$, consistent with an increase that would result if the densities were adjusted higher (to account for the lower ACE densities sometimes used in the database of TS2005). For comparison, we also show $\lambda_i$ (used for the calculation of $W_i$ in equation (7)) and $1.18^{\lambda i}$ in Table 2. One might expect that $F$ would be in between unity (if all WIND data were used by TS2005) and $1.18^{\lambda i}$ (if all ACE data were used by TS2005). While some of the values do lie inside these bounds, it appears that the difference between the two numbers is also dependent on other factors, presumably related to the resolution of the data.

## 3.1. Comparison During Storm Periods

[14] In 2000 there were a number of solar storms. We picked two of these and the data during these storms are displayed in Figures 1 and 2. In each figure, from top to bottom are shown Dst, the solar wind density $N$, the solar wind speed $V$, and the $z$ component of the interplanetary magnetic field $B_z$, and the $W$ parameters (which are derived from the solar wind characteristics using equations (6) for the 5-min data and (9) for the hourly data). The solid curves show the 5-min data and values of the $W$ parameters found using the 5-min data, while the dotted curves show the hourly data and values of the $W$ parameters found using the hourly data.

[15] Figure 1 shows the parameters from the day of year (DOY) 222 to 227 (9 August to 14 August) in 2000. In this time interval, there are no gaps in either the hourly data or the 5-min data. While there is some high-frequency variation of the solar wind quantities $N$, $V$, and $B_z$ which is not represented in the hourly data, the 5-min and hourly values are not greatly different and the $W$ parameters calculated using the 5-min or hourly data are also almost the same. The same result can be seen in Figure 2 for storm periods DOY 308 to 313 (3 November to 8 November). These plots show that our procedure using the hourly data from OMNIWEB is sufficient to calculate the $W$ parameters needed for the TS05 magnetic field model to a good accuracy, at least when there are no gaps in the OMNIWEB data.

## 3.2. Comparison for All of Year 2000

[16] As a further test, we plot in Figure 3 for the first 75 days of 2000 the same quantities as were plotted in Figure 1, except that the 5-min data has been averaged to hourly values after calculating the $W$ parameters from the 5-min data. (This is done so we can display a large amount of data in one plot.) The periods marked off by vertical lines with horizontal bars at the top of each panel indicate the gaps in the 5-min data (only gaps with time interval more than one day are shown). In this case, the solar wind parameters should be exactly the same, and they are (except for a few very brief time intervals). More significantly, the hourly averaged $W$ parameters calculated using the 5-min data are practically indistinguishable from the $W$ parameters calculated from the hourly average OMNIWEB data.

## 3.3. Comparison With Simulated Gaps

[17] As described above, we developed a system of interpolation in order to use the measured values of solar wind quantities in regions of data gaps that are close to the measured values, and averaged quantities in regions of gaps that are far from measured values. However, there are few data gaps in the year 2000 data. There are larger data gaps for earlier years such as 1991, during which solar wind data were derived mainly from measurements by the IMP-8 spacecraft; IMP-8 was only in the solar wind during part of its orbit. In order to test how well the interpolated data works as a representation of the solar wind characteristics and for calculating the $W$ parameters when there are large data gaps, we generate a simulated set of year 2000 data with data gaps by introducing the same data gaps

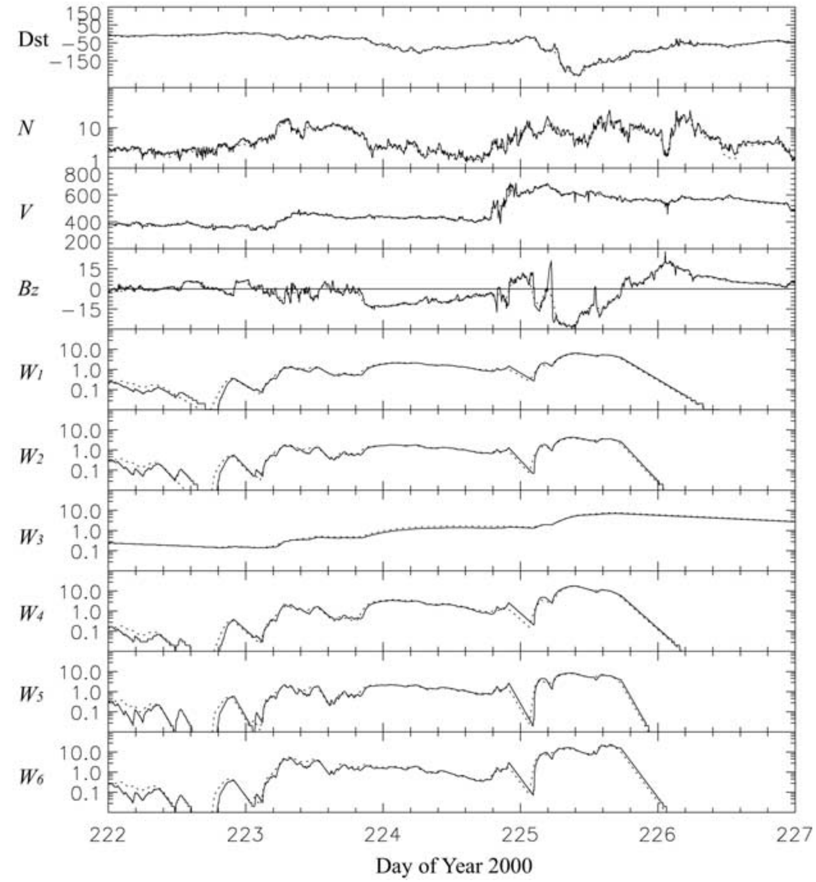


**Figure 1.** Dst, solar wind density $N$ and speed $V$, interplanetary magnetic field (GSM) component $B_z$, and *Tsyganenko and Sitnov*'s [2005] $W$ parameters from DOY 222 to 227 in year 2000. The solid (dotted) curve is the 5-min (hourly) data.

into the year 2000 hourly OMNIWEB data as were present in the year 1991 data.

[18] Figure 4 shows for the first 75 days of 2000 the hourly average of the 5-min solar wind data and hourly average of the $W$ parameters calculated from the 5-min solar wind data (solid curves) and the simulated OMNIWEB data with gaps and $W$ parameters calculated from these data (dotted curves). The periods marked off by vertical lines with horizontal bars at the top of the $N$, $V$, and $B_z$ panels indicate the periods of the simulated gaps in the OMNIWEB data and those in the panels for the $W$ parameters indicate the regions in which the status variable for $W$ is 1 or 0, indicating regions where the values of $W$ are not as reliable as in other regions. (How we define the status variable for $W$ is explained in the second paragraph of section 4.) Only gaps with a time interval more than one day are shown. In the gaps, the interpolated $B_z$ values (dotted curve) quickly go to the average value, since the correlation time for $B_z$ is short (Table 1), whereas, the interpolated $V$ values tend to better follow the hourly average of the 5-min data (solid curve).

[19] Since $B_z$ plays such an important role in the calculation of the $W$ parameters, the resulting $W$ parameters found from the interpolated data do not precisely follow the hourly average of the 5-min data. They do, however, yield values of $W$ parameters which are typical. Considering for instance the data gap from DOY 5 to 9 (5 January to 9 January) in Figure 4, the $B_z$ from the interpolated hourly data rapidly goes to a nearly constant value. For the purposes of calculating the $W_i$ terms, each $B_{zi}$ is slightly negative $(=-(\langle B_s^\gamma \rangle_{20})^{1/\gamma}$ as described in section 2.3) and represents the effect of typical fluctuations. The $V$ interpolated from the hourly data (dotted curve) has a constant value near the left edge of the gap (DOY 5–6), then decreases to the value at the right edge of the gap, where it stays constant from DOY 8 to 9. The interpolated $V$ tracks the hourly average of the 5-min data (solid curve) fairly well and this is often the case. The density $N$ interpolated from hourly data, with a shorter correlation time (Table 1), stays at the average value over most of the gap (dotted curve) and this average value is higher than that of the 5-min data. The $W$ values calculated from the interpolated hourly

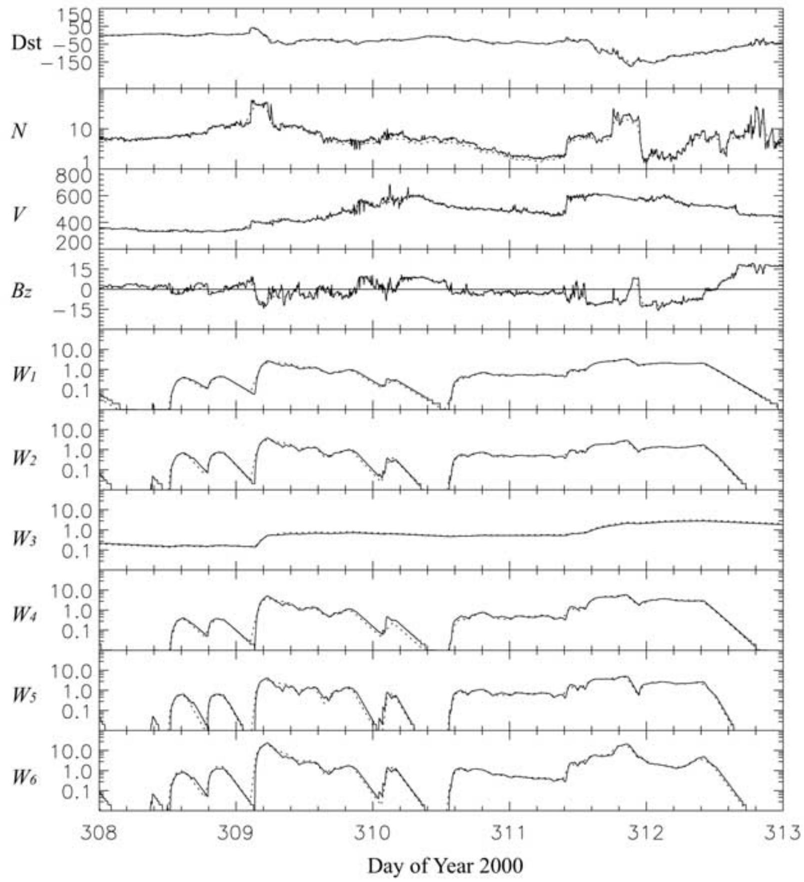


**Figure 2.** Same as Figure 1, but for DOY 308 to 313 of year 2000.

data show a decrease across the gap, responding to the increase in $V$, but the $W$ values from the hourly data are somewhat higher than those found using the 5-min data. This is due in part to the higher average value of $N$. The values of the $W$ parameters in the gaps therefore are typical values which may show some features consistent with the real data but which do not precisely match the real data.

## 4. Database of Hourly Magnetic Field Input Parameters

[20] We have created a file with interpolated values of hourly data for 1963 to 31 May 2007. This file is available at http://www.dartmouth.edu/~rdenton/magpar and will be available through the Virtual Radiation Belt Observatory (VIRBO, http://virbo.org) in January, 2008. It may be freely used subject to the rules of the road listed in the header of the file. This file contains solar wind characteristics from OMNIWEB, the IMF $B_y$ and $B_z$, the proton density $N$, the solar wind speed $V$, and the dynamic pressure $P_{\rm dyn}$, geomagnetic indices Kp, Kp3 (an average value of Kp at time $t$ weighing the values at preceding time $t_p$ with a factor $\exp(-(t - t_p)/(3 \text{ days}))$ [*Gallagher et al.*, 1988]), and Dst, the $G$ parameters used in the T02 [*Tsyganenko*, 2002a, 2002b] and TSK03 [*Tsyganenko et al.*, 2003] models, and the $W$ parameters used in the TS05 model [TS2005].

[21] For the solar wind characteristics, the $G$ parameters, and the $W$ parameters, there is a status indicator. For the solar wind values, the status indicator is 2 if the value is measured, 1 if it is within 2 correlation times of the measured values (so that the value is significantly affected by measured values), and otherwise 0 (indicating that the value is mostly determined by average values). The $G$ parameters were determined where the solar wind data was measured and were similarly interpolated. The $W$ parameters were determined as described above in section 2.3. Recall that the $W$ parameters are found from a weighted average of the source term $S$ over the preceding time (equation (7)). Thus the status variables of the $W$ parameters also need to be averaged in the same way using the source term. For the purposes of averaging the status variables, one of the status values of the solar wind characteristics was used at each preceding time, equal to

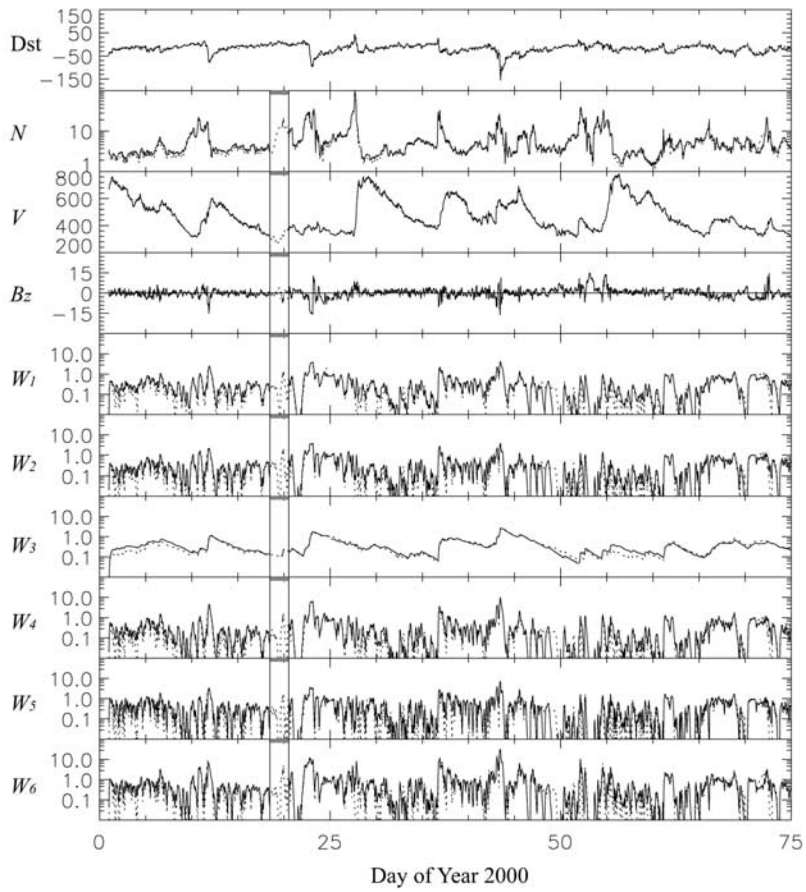


**Figure 3.** Same quantities as were plotted in Figure 1, except that the 5-min data has been averaged over an hour (after calculating the $W$ parameters using the 5-min data). The solid (dotted) curve is the hourly average of the 5-min data (hourly average from the OMNIWEB data) for DOY 0 to 75 of year 2000. The periods marked off by vertical lines with horizontal bars at the top of each panel indicate the gaps in the 5-min data.

the least value of the status variables for $B_z$, $N$ and $V$, let us say, $I_l$. Here we define another quantity,

$$\tilde{I}(t_i) = \frac{r}{4}\Big[S(t_i) + S(t_i - 1/4)e^{-r/4} + S(t_i - 1/2)e^{-r/2} + S(t_i - 3/4)e^{-3r/4}\Big] I_l(t_i) + \tilde{I}(t_i - 1)e^{-r}, \quad (10)$$

which is the average status variable weighted by the source terms used to calculate $W$. The average status of the $W$ parameter is then $\tilde{I}(t_i)/W(t_i)$. In our data file, we use 2 for an average status of the $W$ parameter greater than 1.5, 1 for an average status from 0.5 to 1.5, and 0 for an average status less than 0.5. The gaps indicated for the $W$ parameters (vertical lines with gray bars) in Figure 4 represents the regions where the status variable is 1 or 0. Outside of these regions where the status variable is equal to 2, the values of $W$ found from the hourly data agree well with those from the 5-min data.

[22] In Table 3, we compare statistically the $W$ parameters calculated from OMNIWEB hourly data with simulated gaps to those calculated from the 5 min solar wind data with no gaps. For each $W$ parameter (group of four rows divided by solid lines), the data is broken up into categories (columns) for all values of the status variable, and for values of 2, 1, and 0. Within each group, values are given for the maximum value of the $i$th $W$ parameter $W_{i,\max}$, the average value of the $W$ parameter and its standard deviation $\langle W_i \rangle \pm \sigma_{W_i}$, the standard difference between the $W$ parameter calculated using the 5 min data and that calculated using the hourly data $\delta W_i$, and the number of hourly entries in each category $N_{i,\mathrm{status}}$. Table 3 shows that for each $W$ parameter both $\sigma_W$ and $\delta W$ are much less than $W_{\max}$. For the status value of 2, $\delta W$ is significantly less than $\sigma_W$ (factor of 3 or more), and this shows that values of $W$ calculated using the hourly data are significantly better than an average value $\langle W \rangle$. For a status value of 1, $\delta W$ is also smaller

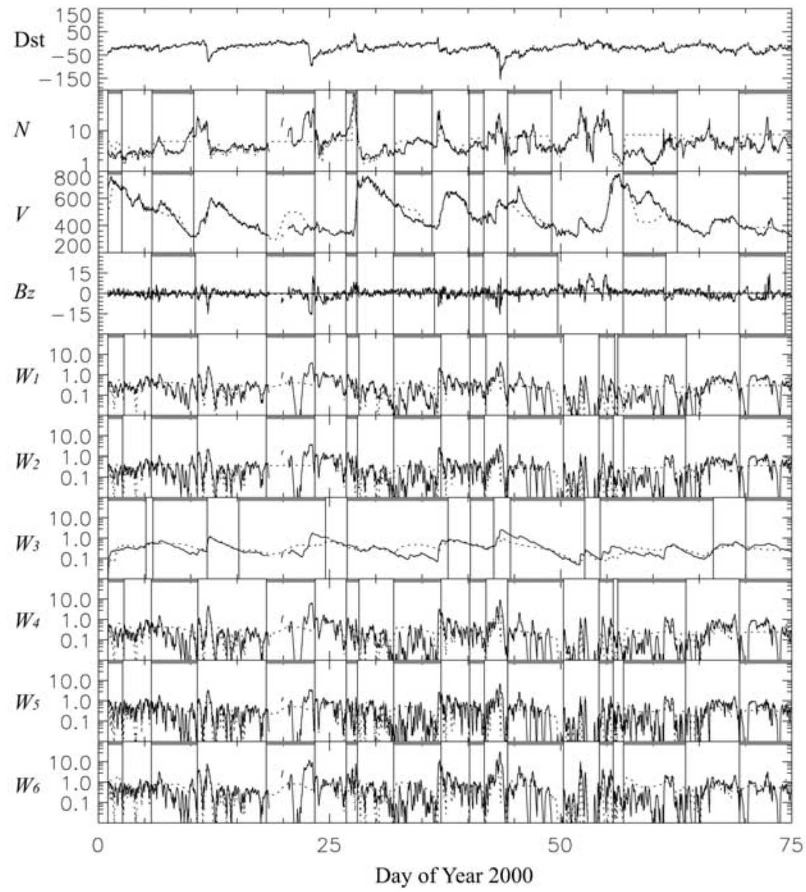


**Figure 4.** This plot is like Figure 3, except that gaps have been inserted into the hourly data (before interpolation or calculation of the $W$ parameters) using the interval of solar wind data missing in the OMNIWEB database in 1991. The horizontal bars at the top of panels 2–4 indicate the simulated gaps and those of panels 5–10 indicate regions where the values of $W$ are not as reliable as in other regions (status variable = 0 or 1 as described in section 4). Again, the solid (dotted) curve corresponds to the hourly average of the 5-min data (hourly data from OMNIWEB with data gaps) for DOY 0 to 75 in 2000.

than $\sigma_W$, though not as much smaller as was the case for a status variable of 2. For a status value of 0, $\delta W$ is close to $\sigma_W$, showing that the $W$ values are as good as but not significantly better than an overall average. Table 3 also shows that there are a significant number of points with the status values of 2 and 1 (values that are superior to using an average value).

[23] Finally, in Table 4 we statistically compare magnetic field values found from the TS05 model with the observed magnetic field measured by the GOES 8 geosynchronous spacecraft during year 2000. An average value of the GOES 8 magnetic field is used for each 5 min interval during the year (actually the first 365 days of the year). Five minutes is the minimum interval of time that could be expected to respond to changes in the TS05 magnetic field model (since the model was developed using 5 min resolution data). We define an error parameter

$$d\bar{B} \equiv \frac{|\mathbf{B}_{\rm model} - \mathbf{B}_{\rm observed}|}{B_{\rm observed}}, \quad (11)$$

equal to the fractional error of the vector magnetic field. We calculate the average and standard deviation of this error parameter for various subsets of the data. At each time, values of the solar wind and $W$ factors are found from a linear interpolation of the values from our solar wind parameter database. Each row in Table 4 represents a set of data based on the status variable values for the hourly data with simulated gaps; for each row, the status value for the solar wind variables $B_y$, $B_z$, and $P_{\rm dyn}$ has the value given in the column labelled

Table 3. Statistic Comparison of W Parameters[a]

| | All[c] | Status Value[b] 2 | 1 | 0 |
|---|---|---|---|---|
| $W_{1,max}$[d] | 11.28 | 6.35 | 4.31 | 11.28 |
| $\langle W_1\rangle \pm \sigma_{W1}$[e] | 0.38 ± 0.60 | 0.44 ± 0.56 | 0.33 ± 0.52 | 0.36 ± 0.63 |
| $\delta W_1$[f] | 0.49 | 0.13 | 0.35 | 0.63 |
| $N_{1,status}$[g] | 8760 | 2757 | 1249 | 4754 |
| $W_{2,max}$ | 6.55 | 4.19 | 3.74 | 6.55 |
| $\langle W_2\rangle \pm \sigma_{W2}$ | 0.39 ± 0.53 | 0.42 ± 0.51 | 0.35 ± 0.47 | 0.37 ± 0.56 |
| $\delta W_2$ | 0.43 | 0.13 | 0.32 | 0.55 |
| $N_{2,status}$ | 8760 | 2883 | 1095 | 4782 |
| $W_{3,max}$ | 15.87 | 6.99 | 5.68 | 15.87 |
| $\langle W_3\rangle \pm \sigma_{W3}$ | 0.55 ± 0.94 | 0.66 ± 0.68 | 0.39 ± 0.55 | 0.60 ± 1.12 |
| $\delta W_3$ | 0.84 | 0.20 | 0.22 | 1.25 |
| $N_{3,status}$ | 8760 | 2011 | 2911 | 3838 |
| $W_{4,max}$ | 49.00 | 17.47 | 9.05 | 49.00 |
| $\langle W_4\rangle \pm \sigma_{W4}$ | 0.44 ± 1.36 | 0.48 ± 1.02 | 0.31 ± 0.67 | 0.45 ± 1.62 |
| $\delta W_4$ | 1.21 | 0.18 | 0.35 | 1.61 |
| $N_{4,status}$ | 8760 | 2728 | 1200 | 4832 |
| $W_{5,max}$ | 23.92 | 8.29 | 6.63 | 23.92 |
| $\langle W_5\rangle \pm \sigma_{W5}$ | 0.45 ± 0.82 | 0.49 ± 0.65 | 0.37 ± 0.62 | 0.46 ± 0.95 |
| $\delta W_5$ | 0.71 | 0.17 | 0.31 | 0.95 |
| $N_{5,status}$ | 8760 | 2880 | 1218 | 4662 |
| $W_{6,max}$ | 87.80 | 27.32 | 38.14 | 87.80 |
| $\langle W_6\rangle \pm \sigma_{W6}$ | 1.01 ± 3.16 | 0.91 ± 1.84 | 0.85 ± 2.05 | 1.11 ± 3.91 |
| $\delta W_6$ | 2.92 | 0.63 | 0.98 | 3.90 |
| $N_{6,status}$ | 8760 | 2916 | 1055 | 4789 |

[a]We have given here the average values of the $W$ parameters, $\langle W_i\rangle$. For those interested in other average values (to use, for instance, as a substitute for the method we have outlined if high accuracy is not required), here are the average values for the other solar wind related quantities averaged over the entire period from 1960 to 2004: $B_y = 0 \pm 4.3$ nT, $B_z = 0 \pm 3.3$ nT, $P_{dyn} = 2.5 \pm 2.0$ nPa, $N = 7.1 \pm 5.8$ cm$^{-3}$, $V = 443 \pm 106$ km/s, $G_1 = 2.6 \pm 6.4$, $G_2 = 2.5 \pm 4.5$, and $G_3 = 1.9 \pm 5.7$. (The standard deviation for the $G$ parameters would be better expressed as a log error, since the values should be positive.)
[b]Subset of data with $W_i$ having the listed status value.
[c]All data for the year 2000.
[d]Maximum value of $W_i$ within the data set.
[e]Average value of $W_i$ and the corresponding standard deviation.
[f]Standard difference between the W parameter calculated from 5 min solar wind data and that calculated from hourly solar wind data.
[g]Number of data points for $W_i$ with the given status value.

"SW" (under "Status Value"), and all six $W$ parameters have the status value listed in the column labelled "W" (under "Status Value"). The total number of data points in each group of data is listed under the column labelled "$N_{status}$". The total number of 5 min intervals in a year is about 105,000, while the total number of data points represented in all the groups listed in Table 4 is about 46,000. Therefore it is clear that at most times $B_y$, $B_z$, and $P_{dyn}$ do not have the same status value at the same time that all the $W$ parameters have the same status value. Furthermore, at the majority of times, at least one of the status variables has a value of 1 (since the cases where all the status values are 0 and where all the status values are 2 are included in the Table). Nevertheless, the groups in Table 4 can give us a rough idea of how the accuracy of the magnetic field model is affected by the accuracy of the solar wind or $W$ parameters as indicated by the status values. (A more complete test would find which particular solar wind variables and $W$ parameters are most important for accurately modeling the magnetic field.)

[24] First of all, we note from Table 4 that the standard deviation of $d\bar{B}$ is on the order of the average value, indicating that there is a large variation in the accuracy of the model (depending on factors other than the input parameters). Second, we see that when all the status variables are 0 (first row of Table 4), $d\bar{B} = 0.16 \pm 0.13$, whereas when all the status variables are 2 (last row of Table 4), $d\bar{B} = 0.11 \pm 0.09$. This shows that the average error of the model using typical values for the W parameters (like the $\langle W_i\rangle$ values listed in Table 3) is statistically only about 50% larger than that found when the solar wind parameters are well known. For an order of magnitude estimate of the magnitude of the magnetic field, either is good enough. However, if one is tracing field lines to other regions, a small improvement in the model is likely to have a greater impact on the results.

[25] From Table 4, one can see that if the status values for the raw solar wind parameters ($B_y$, $B_z$, and $P_{dyn}$) are 1, a higher value for the status value of the $W$ parameters leads to a more accurate model (lower $d\bar{B}$). However, when the status values for the raw solar wind parameters are 2, the values for $d\bar{B}$ listed in the fourth column of Table 4 seem to indicate that the status values of the $W$ parameters do not significantly affect the accuracy of the model. However, there may be other factors outside of the inputs to the magnetic field model that lead to a better agreement between the model and observed fields for one group of data than for another (considering the limited number of data points with some combinations of status values). In order to test whether this is so, we also calculate $d\bar{B}$ for the same sets of times (based on the status variables for the hourly data with simulated gaps), but calculating the values of the model field (and thus $d\bar{B}$) using the hourly data without gaps and $W$ parameters calculated using the hourly data without gaps (fifth column in Table 4) and also calculated using the 5 min data without gaps and the $W$ parameters found from the 5 min data without gaps (7th column in Table 4). That is, the $d\bar{B}$ values are calculated using the accurate solar wind and $W$ parameters, but for the sets of data based on the status variables for the data with simulated gaps.

[26] There is not a great variation in $d\bar{B}$ for these data sets, though there is some. On the basis of the largest groups of data (all status variables equal to 0 or 2 in the data with simulated gaps), it is clear that the average value of $d\bar{B}$ is 0.10 when the model values are found using either the hourly or 5 min data. Therefore it appears that the hourly data is good enough (when linearly interpolated to a particular time) for calculating the model magnetic field. For comparison, using the same GOES 8 data during 2000, $d\bar{B} = 0.25 \pm 0.19$ for

**Table 4.** Error Parameter $d\bar{B}$ for Year 2000 GOES 8 Data[a]

| Status Value | | | Hourly Data With Gaps | Hourly Data Without Gaps | | 5 Min Data Without Gaps | |
|---|---|---|---|---|---|---|---|
| SW[b] | $W$[c] | $N_{status}$[d] | $d\bar{B}$ | $d\bar{B}$ | $d\bar{B}_{gaps}/d\bar{B}$[e] | $d\bar{B}$ | $d\bar{B}_{gaps}/d\bar{B}$[e] |
| 0 | 0 | 28843 | 0.16 ± 0.13 | 0.10 ± 0.11 | 1.57 | 0.10 ± 0.11 | 1.60 |
| 1 | 0 | 157 | 0.13 ± 0.08 | 0.10 ± 0.07 | 1.36 | 0.09 ± 0.08 | 1.43 |
| 1 | 1 | 1622 | 0.12 ± 0.08 | 0.11 ± 0.08 | 1.11 | 0.11 ± 0.08 | 1.10 |
| 1 | 2 | 457 | 0.08 ± 0.05 | 0.07 ± 0.05 | 1.14 | 0.07 ± 0.04 | 1.13 |
| 2 | 0 | 349 | 0.11 ± 0.06 | 0.09 ± 0.06 | 1.16 | 0.08 ± 0.05 | 1.34 |
| 2 | 1 | 480 | 0.08 ± 0.05 | 0.08 ± 0.05 | 1.07 | 0.09 ± 0.05 | 0.97 |
| 2 | 2 | 14017 | 0.11 ± 0.09 | 0.10 ± 0.09 | 1.04 | 0.10 ± 0.09 | 1.04 |

[a]Note carefully: The data groups here are based on the status values of the hourly data with gaps. The $d\bar{B}$ values for the hourly data without gaps and for the 5 min data without gaps are calculated using the real (no gap) solar wind data and $W$ parameters.

[b]The status variable for all three of the solar wind variables $B_y$, $B_z$, and $P_{dyn}$.

[c]The status variable for all six of the $W_i$.

[d]The number of 5 min intervals in 2000 having all of $B_y$, $B_z$, and $P_{dyn}$ with the listed status value, and all of the $W_i$ with the listed status value.

[e]The average $d\bar{B}$ from the hourly data with gaps divided by the average $d\bar{B}$ for this data without gaps. This is error degradation factor for not having measured solar wind characteristics or $W$ factors based on solar wind characteristics.

the dipole field model, showing that the TS05 model does significantly better on average at describing the observed magnetic field at geosynchronous orbit. If we limit the data to times when Dst $\leq -100$ nT, the average value of $d\bar{B}$ is 0.25 ± 0.25 using the TS05 model with the hourly input data without gaps, 0.24 ± 0.28 using the TS05 model with 5 min input data without gaps, and 0.36 ± 0.28 for a dipole field model. This comparison shows that for storm conditions the TS05 model is significantly less accurate than for quiet conditions, though still better than the dipole field model. Again (as was the case for all conditions (Table 4)) the average value of $d\bar{B}$ for Dst $\leq -100$ was not significantly different when using the TS05 model with hourly or 5 min input data, and this indicates that the hourly data is sufficient as an input to the TS05 model for storm conditions also.

[27] In order to factor out the effect of extra factors not accounted for by the model field input parameters, we divide the average value of $d\bar{B}$ based on the hourly data with gaps, $d\bar{B}_{gaps}$, by the more accurate values found using the data without gaps. These values are listed in columns 6 and 8 of Table 4 for the hourly data without gaps (column 6) and the 5 min data without gaps (column 8). These ratios represent the factor by which the model accuracy degrades if the values of the solar wind and $W$ parameters are not well known. The results are fairly similar for hourly and 5 min data. The degradation factor is about 1 if the status values of the raw solar wind parameters are 2 and the status values of the $W$ parameters are 1 or 2. (The degradation factor should certainly be 1 if all the status values are 2, since in both the data with or without gaps, the model input parameters are well known.) If the status values of the raw solar wind parameters are 2, but those of the $W$ parameters are 0 (such as would occur immediately after a data gap, since the $W$ parameters are evaluated using an integral over previous times), then the degradation factor is 1.34. If the status values of the raw solar wind variables is 1, but the status values of the $W$ parameters is 1 or 2, the degradation factor is about 1.1, while if the status values of the $W$ parameters is 0, the degradation factor is 1.43. If the status values of all the parameters is 0, the degradation factor is 1.6. These results suggest that our definition of the status variable for the $W$ parameters may be too conservative. When the status values for the $W$ parameters are all equal to 1, the accuracy of the model is about as good as that when the status values for the $W$ parameters are 2 (better based on the numbers in Table 4). At any rate, this comparison shows that the model magnetic field in the TS05 model is improved not only if the solar wind parameters from our database have a status value of 2 (corresponding to measured values for solar wind parameters or values mostly based on measured parameters for the $W$ parameters), but also for a status value of 1 (corresponding to interpolated values within about two correlation times of measured values for the solar wind parameters or on an average largely based on such interpolated values for the $W$ parameters).

## 5. Summary

[28] Accurate magnetic field models are crucial for many space weather applications, but modern models depend on solar wind characteristics, which are not always known, especially for dates before 1995. We have shown that hourly data from OMNIWEB with interpolation to smaller intervals (15 min in our study) can be used to calculate the $W$ parameters used in the latest TS05 model [TS2005] with sufficient accuracy. We have also presented a system to evaluate values of the solar wind characteristics and $W$ parameters within data gaps where measured values are not available. Within about two correlation times from the edge of a data gap, the solar wind parameters from our method yield a better estimate of the observed magnetic field than that which could be found using average values of the parameters. Far away from measured values, the values of the solar wind characteristics and $W$ parameters

are reasonable, or typical, though not necessarily accurate, and no better or worse than average values. We have used our techniques to create a file with solar wind characteristics, $G$ parameters [*Tsyganenko*, 2002a, 2002b; *Tsyganenko et al.*, 2003], and $W$ parameters [TS2005] for 1963 to 1 June 2007. The file contains status parameters that indicate the quality of the interpolated values in the file. When the status variable of the $W$ parameters is 2 or 1, the values in our data file are significantly better (on average) than average values. This file is available to the space physics community at http://www.dartmouth.edu/~rdenton/magpar and will be available through VIRBO (http://virbo.org) in January, 2008. It will be be a great help for magnetic field modeling. The database is not intended to replace the original (measured and averaged) solar wind data provided through OMNIWEB and recent changes to the OMNIWeb database may or may not be incorporated into this file. Our current goal is to automate the process of calculating the input parameters through VIRBO, so that our input parameters will reflect future modifications to the OMNIWEB data and be available for dates after 31 May 2007.

[29] The method of interpolation we have outlined is relatively simple (essentially linear interpolation to average values) but is not necessarily optimal. As mentioned previously, the definition of the correlation time as the time difference yielding a correlation coefficient of 0.8 was somewhat arbitrary, and other choices could be made. (We adjusted this value slightly from the first value we tried, 0.6, in order to make sure that parameters with a status value of 1 led to significantly more accurate values for the calculated $W$ parameters than those with a status value of 0.) This value might ideally be different for different solar wind parameters. One might also prefer to alter the method so that within large data gaps, the $W$ parameters become equal to their average values, rather than integrals of average solar wind values. *O'Brien* [2005] outlines a method of determining maximum likelihood estimates that might be preferable to our method.

[30] One minute resolution solar wind data are now available on OMNIWEB for times from 1995 to the present [*Papitashvili and King*, 2006]. Minimum variance analysis is used to establish wave front planes and reduce the error involved in propagating the solar wind from the monitor to the Earth. One could easily use the techniques described here to create a file of 5 min resolution solar wind parameters. On the other hand, the results presented in section 4 suggest that the hourly data may be sufficient since the accuracy of the magnetic field model at geosynchronous was not improved on average when 5 min data was used.

[31] **Acknowledgments.** We thank Joseph King, Jim LaBelle, and Bob Weigel for helpful conversations. We are grateful to the Goddard Space Flight Center (GSFC) Space Physics Data Facility (SPDF) for making the IMP-8 (GSFC and LANL), WIND (courtesy of R. P. Lepping and K. W. Ogilvie), and ACE (courtesy of L. Burlaga and D.J. McComas) solar wind data, Dst (Kyoto University), and Kp (NOAA) data available from the OMNIWeb interface at http://omniweb.gsfc.nasa.gov. We thank NASA CDAWeb for making GOES 8 magnetometer data (courtesy of H.J. Singer and NOAA) available. Work at Dartmouth was supported by NSF grants ATM-0632740 and ATM-0120950 (Center for Integrated Space Weather Modeling, CISM, funded by the NSF Science and Technology Centers Program) and NASA grants NNG05GJ02G and NNX06AD81G. NAT acknowledges support by NSF grants ATM-0201991 and ATM-0207248 as well as NASA's grants NAG5-12185 and NNG04GE92G.

## References

Gallagher, D. L., P. C. Craven, and R. H. Comfort (1988), An empirical model of the Earth's plasmasphere, *Adv. Space Res.*, *8*, 15.

King, J. H., and N. E. Papitashvili (2005), Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data, *J. Geophys. Res.*, *110*, A02104, doi:10.1029/2004JA010649.

O'Brien, T. P. (2005), A framework for next-generation radiation belt models, *Space Weather*, *3*, S07B02, doi:10.1029/2005SW000151.

Papitashvili, N. E., and J. H. King (2006), A draft high resolution OMNI data set, *Eos Trans. AGU*, *87*(36), Jt. Assem. Suppl., Abstract SM33A-02.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1997), *Numerical Recipes in Fortran 77: The Art of Scientific Computing, Fortran Numerical Recipes*, vol. *1*, 994 pp., Cambridge Univ. Press, New York.

Tsyganenko, N. A. (1989), A magnetospheric magnetic field model with a warped tail current sheet, *Planet. Space Sci.*, *37*, 5–20.

Tsyganenko, N. A. (1996), Effects of the solar wind conditions on the global magnetospheric configuration as deduced from data-based field modelse, *Eur. Space Agency Spec. Publ.*, *ESA SP-389*, 181.

Tsyganenko, N. A. (2002a), A model of the near magnetosphere with a dawn-dusk asymmetry: 1. Mathematical structure, *J. Geophys. Res.*, *107*(A8), 1179, doi:10.1029/2001JA000219.

Tsyganenko, N. A. (2002b), A model of the near magnetosphere with a dawn-dusk asymmetry: 2. Parametrization and fitting to observations, *J. Geophys. Res.*, *107*(A8), 1176, doi:10.1029/2001JA000220.

Tsyganenko, N. A., and M. I. Sitnov (2005), Modeling the dynamics of the inner magnetosphere during strong geomagnetic storms, *J. Geophys. Res.*, *110*, A03208, doi:10.1029/2004JA010798.

Tsyganenko, N. A., H. J. Singer, and J. C. Kasper (2003), Storm-time distortion of the inner magnetosphere: How severe can it get?, *J. Geophys. Res.*, *108*(A5), 1209, doi:10.1029/2002JA009808.

R. E. Denton and Z. Qin, Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, NH 03755, USA. (richard.e.denton@dartmouth.edu; zhengrui.qin@dartmouth.edu)

N. A. Tsyganenko, USRA, NASA Goddard Space Flight Center, Code 612.3, Greenbelt, MD 20771, USA.

S. Wolf, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA.