The $D_{st}$ geomagnetic response as a function of storm phase and amplitude and the solar wind electric field

D. Vassiliadis, A. J. Klimas, J. A. Valdivia, and D. N. Baker

Abstract. We examine the dependence of the $D_{st}$ timescales on storm conditions and its implications for the storm-substorm relationship. The growth, decay and oscillation timescales are expressed as functions of the storm magnitude and phase, and the solar wind electric-field input $V_{i1}$. Nonlinear, second-order autoregressive moving average (ARMA) models fit to 5-min data and yield two timescales, an exponential decay with an average e-folding time $\tau_1 = -4.69$ hours (7.26 hours for the pressure-corrected $D_{st}(0)$) and an inductive time $\tau_2 = -0.81$ hours (0.05 hours for $D_{st}(0)$). Around these average values there is a systematic variation: (1) For most of the storm duration, $\tau_1$ is negative representing the rapid adjustment of the inner magnetosphere to the imposed electric field. (2) In the early main phase, however, $\tau_1 = 5.29$ hours (1.76 hours for $D_{st}(0)$), so the disturbance grows as a slow exponential. (3) During commencement and main phase, the timescales are complex conjugate and the response is oscillatory. Fast oscillations during storm commencement (period 1.13 hours: 8.48 min for $D_{st}(0)$) are a "ringing" response to interplanetary pressure enhancements. Slow oscillations in the main phase have an average period of 1.96 hours (1.55 hours for $D_{st}(0)$) and coincide with $AL$ intensifications. The main phase can be separated into periods of oscillatory, fast decay (coincident with $AL$ activity and probably due to injections) and monotonic slow decay (regular convection). (4) All timescales decrease with increasing interplanetary activity because high activity involves acceleration and loss of heavy ions with shorter lifetimes than protons. (5) Also, decay times are about twice as long during recovery than during main phase. (6) Similar dependences are found for the solar wind coupling coefficients. The models are similar to linear models in predictability and are stable with respect to perturbation in the initial conditions.

1. Introduction

The ring current is one of the permanent, basic current systems of the magnetosphere (I. A. Dalgis et al., "The terrestrial ring current: origin, formation, and decay," Rev. Geophys., in print, 1999), and its variations lead to magnetic storms [Tsurutani et al., 1997; see also special issue of the Journal of Geophysical Research, 104 (A7) 1997]. Its ground geomagnetic effects in particular are conventionally measured by the $D_{st}$ geomagnetic index, with some limitations imposed by the index construction process which are discussed in section 2. Many studies have estimated physical and dynamical properties of the current by measuring the $D_{st}$ "signature." However, one of the most physically interesting properties, the ring current decay rate, which is closely associated with the dynamics of the charge carriers, is still not measured to a high accuracy. Our nonlinear analysis yields additional timescales and coupling strengths providing physical insights into ring current development and reduction. One of these timescales is an oscillation period, which we will argue shows an effect of substorms on $D_{st}$, usually during storm main phase.

Continuing earlier research [Vassiliadis et al., 1990], we measure the dependence of characteristic timescales of decay, growth, and oscillation on storm amplitude and phase, and the interplanetary activity level. Since the ring current is determined primarily by the ion drift, these timescales are associated with ion acceleration and loss processes. Therefore measuring the time parameters provides constraints to ring current simulation models [e.g., Fok et al., 1995; Kosra et al., 1998]. To improve on earlier measurements [e.g., as reviewed by Feldstein (1992)], we use nonlinear dynamical methods which encompass earlier methods of correlation or superposed-epoch analysis and are applicable over a wider range of parameters. Furthermore, in using time series models to measure physical parameters, we demonstrate the nonlinear models' usefulness in areas other than those they have been most known for, namely prediction and forecasting.

Many previous studies [Feldstein, 1992] have used the standard 1-hour index, which, however, is too coarse for realistic modeling needs. Also, most studies have examined rather small numbers of storm intervals. Then a usual procedure has been to consider intervals for which the solar wind input is zero and to assume an exponential decay

$$D_{st}(t) = D_{st}(t_0) \exp \left( \frac{(t - t_0)}{\tau} \right)$$

(1)

from which the decay time $\tau (< 0)$ can be obtained. In reality
the decay does not asymptote to zero, however, but to a value imposed by the solar wind electric field. Therefore we model the $D_s$ activity for a continuous range of activity levels and include the solar wind input explicitly. In previous models the choice of data sets and fitting parameters and division into storm phases have produced a range of timescales which remains wide and ambiguous. The estimates for the decay rate during main phase range from 2 to 24 h, while for the recovery the estimates are from 3 to 50 h [Feldstein, 1992, Table I]. Some general conclusions can be gleaned, however (paraphrased here from the Feldstein [1992] review):

1. The decay timescale $\tau$ is lower for main phase than for recovery. It is not a good approximation to use a single timescale for both phases as many early models do.
2. The decay timescale decreases during injections.
3. It increases with time during recovery phase as $D_s$ increases back towards zero.
4. It is higher in strong storm recoveries compared with moderate storm recoveries.
5. The variation of $\lambda$ must be accounted for in studies of magnetosphere input-output energetics. For example, the rate of increase of $D_s$ during the recovery phase has been expressed as a function of the total magnetospheric energy [Vasyliunas, 1987], in qualitative agreement with observations.

One main goal of this paper is to make the first four statements above more precise (statement 5 is one of the bases of our analysis). Instead of the exponential decay (1), we include the input explicitly in a way similar to previous nonlinear models [Valdivia et al., 1996; Klimas et al., 1991, 1998]. Also, instead of examining isolated storm intervals, we use a continuous long data set of $D_s$ and interplanetary parameters, at a 5-min resolution.

The triplet of variables ($D_s$, $dD_s/dt$, VB) gives an efficient description and physical interpretation of the basic storm dynamics. We use $D_s$ and its derivative to identify four storm phases: (1) early storm commencement, or EC ($D_s$ and $dD_s/dt > 0$), (2) late commencement, or LC ($D_s$ positive and decreasing), (3) main phase, or M (both negative), (4) recovery phase, or R ($D_s$ negative and increasing). The phases correspond to the quadrants on the ($D_s$, $dD_s/dt$) plane shown in Figure 1. The dynamics of $D_s$ is associated with the variation of the drift kinetic energy of the ring current particles [Dessler and Parker, 1959; Schopke, 1966; Gonzalez et al., 1994; Valdivia et al., 1999]. Ring current particles are lost mainly through convection across the magnetopause, and charge exchange [Fok et al., 1995]. The variation of the $D_s$ timescales during recovery has been related to the acceleration and loss of different ion species, in particular protons versus oxygen and other heavy ions [Roeder et al., 1995; Kozyra et al., 1998]. In addition to $D_s$ and its derivative, since the overall storm intensity is related to the energy made available, the solar wind input is used as a third variable.

Section 2 presents the data and the preprocessing methods. We discuss our and previous models in section 3 and present the results in sections 4 and 5.

2. Data Set and Preprocessing

2.1. Effects of Various Current Systems on $D_s$

Before interpreting midlatitude measurements in terms of ring current dynamics, one needs to remove the effects of other current systems. Some steps to that effect are taken in the compilation of $D_s$ [Mayaud, 1980]; however, the procedure does not remove all other contributing currents [Campbell, 1996], and needs improvement.

Since the index is derived from midlatitude magnetometers at geomagnetic latitudes $20^\circ < |\lambda| < 35^\circ$, the current systems that affect it are the azimuthally symmetric and asymmetric (partial) ring current, the magnetopause current, the tail current and the substorm current wedge, the image currents flowing in the Earth’s crust, the quiet ionospheric currents including the equatorial electrojet, and other low-intensity systems [McPherron, 1997]. The azimuthally averaged value of the corrected geomagnetic disturbance is the $D_s$ index, while the range of the residual disturbance defines the ASYM index [Mayaud, 1980]. The partial ring current is estimated to have a 20% effect on $D_s$ [Siscoe and Crooker, 1974]. More generally, the partial ring current is a significant part of the total disturbance, as clearly shown in a spatiotemporal analysis of magnetometer data [Valdivia et al., 1999]. The magnetopause current increases mostly because of solar wind pressure enhancements, and that effect is corrected by using a pressure-balance relation to enhance the magnitude of $D_s$ (see section 2.3). The substorm current wedge affects both $D_s$ [Iyemori and Rao, 1996; Rostoker et al., 1997; Friedrich
et al., 1999] and ASYM [Clauer et al., 1983], but its effect on $D_t$ is typically less than 20 nT.

Averaging over several midlatitude stations at different LT reduces the effects of the currents of limited azimuthal extent such as tail and substorm current wedge on the nightside and the equatorial electrojet on the dayside. To obtain the $B$ field component due to a near-equatorial component of the ring current it is useful to project the measured North-South component on the direction of the dipole axis. This is accomplished by multiplying with the cosine of the magnetic latitude which produces the “axial field.” Finally, one can smooth out the unwanted effects of minor currents by going to a low time resolution, but this choice often removes essential details of other important magnetic effects as well, as will be seen in section 4 below.

We used a data set covering the interval of January 1 to June 30, 1979, at a 5-min resolution. The index was produced from the four standard midlatitude stations plus two more (Bouliana and Tashkent). To reduce high-frequency effects, we smoothed the $D_t$ with a 25-min running average.

2.2. Solar Wind Propagation and Ring Current Lag Time

The solar wind and interplanetary magnetic field (IMF) were measured by the ISEE 3 spacecraft at the Lagrangian point L1. Those measurements were propagated to the sub-solar magnetopause in a planar ballistic propagation scheme [Zwickl et al., 1980; related remarks in Hirston and Heelis, 1996] that uses the solar wind velocity at the spacecraft. To reduce the non-geoeffective high frequencies we smooth the propagated solar wind with a 25-min moving-average filter.

In addition to the propagation delay, there is a lag in $D_t$ response after the solar wind input reaches the dayside magnetopause related to the adjustment of the inner magneto-ospheric plasma and electric field to a change at the magnetopause boundary. In the Durton et al. [1975] model (hereafter referred to as B75) this delay is 25 min on the average, and we time shift the solar wind input relative to $D_t$ by that amount. Including the lag increases the decay time in better agreement with previous studies and improves $D_t$ prediction by a small percentage.

2.3. Solar Wind Pressure Correction

The largest effect on $D_t$ after the ring-current variations is due to a diamagnetic effect occurring when increased solar wind pressure, usually at the beginning of a geoeffective solar wind structure (interplanetary shock, magnetic cloud, or high-speed stream), moves the magnetopause closer to Earth and induces a stronger magnetopause current. In order to remove the effect, one usually considers a time-independent pressure balance between solar wind ram pressure (ignoring thermal and IMF pressures) and magnetospheric $B$ field pressure (ignoring the contribution of the tenuous plasma). Close to equilibrium the magnetopause current effect is proportional to the square root of the pressure

$$D_{st}^{(0)} = D_t - b \sqrt{P_{SW}} - c$$

where $b$ is the effective coupling to a pressure increase, and $c$ is predominantly the geomagnetic effect of the quiet ring current. Representative values are $b = 0.2$ nT/$\sqrt{V_{SW}}$ and $c = -20$ nT (R75). However, both $b$ and $c$ vary seasonally depending on the dipole tilt angle from the Earth-Sun line [McPherron, 1995]. Also, the quiet current has a solar cycle dependence [Vennerstrom and Friis-Christensen, 1996]. Therefore we calculate $b$ and $c$ as functions of activity and season using (2) with data that satisfy the following constraints:

1. Low solar wind activity, ideally $P_{SW} \leq 0$ over an interval of several tens of minutes. In practice we use $P > 1$ nPa for the instantaneous $D_t$.

2. Low geomagnetic activity, $D_t > D_{t,min}$. In principle one could use $D_{t,min} = 0$ nT, since the effect of a magnetopause compression is a positive increase, $\Delta D_t$. However, this increase is added to the quiet current effect, $c$, which is not zero, but negative. In order to bracket $c$ from below one needs to have $D_{t,min} \leq c$ and therefore also need data with $D_t < 0$. We choose $D_{t,min} = -40$ nT.

3. Low pressure. For large $P_{SW}$ the equilibrium (2) is not time-independent. We use data with $P > P_{min} = 2$ nPa.

4. To ensure the time-independence of (2), we use data which satisfy $|dP/dt| < 0.1$ nPa/$\Delta t$, and $|dD_t/dt| < 1$ nT/$\Delta t$.

The coefficients can be calculated in the following ways:

1. Least squares: A linear regression can be applied to all data that satisfy 1-4 above.

2. Linear envelope: Positive intensifications of $D_t$ form an easily distinguishable upper envelope in the $(D_t, \sqrt{P_{SW}})$ plane [Araki et al., 1993, Figures 1, 2, and 4]. Data points on the envelope correspond to rapid increases in pressure at times when $D_{t,min}^{(0)} \approx 0$, approximating satisfying equation (2). One can calculate the equation of the envelope using a generalization of least squares (appendix B). Data with small $P_{SW}$ should be used, because for large values the envelope is not linear.

3. Probability maximum: The coefficient $c$ can be computed independently of $b$. Since storm intervals are rather infrequent events and $D_t$ relaxes to its quiet value exponentially fast after any external disturbance, one can consider the quiet-time $D_t$ data distribution (with stormtime and high-pressure intervals removed) and calculate $c$ as the most probable value in the distribution.

4. Ratio of differences: $b$ can be calculated independently of $c$, as the ratio $\Delta D_t/\Delta V_P$ from a set of quiet-activity data [B75, Smith et al., 1986].

5. Use of $P_{SW}$ as a second input in a dynamic model for $D_t$, as implemented in multi-channel linear prediction filtering [McPherron et al., 1984; Travers et al. 1990] and neural networks [Wu and Landstadt, 1996, 1997].

We have calculated $b$ and $c$ with method 2 and $c$ with method 3 for a sequence of successive overlapping intervals within the data set. Each interval is 1 month long, while the centers of successive intervals are separated by 2 days. The $b$ and $c$ associated with the centers of the intervals are interpolated in time (Figure 2a-2b). At the ends of the interval, $b(t)$ and $c(t)$ are tapered to the outermost calculated values. Method 3 gives slowly varying and less-extreme values for $c$ (not shown). Both parameters show an increasing trend from January to June modulated by the 27-day solar rotation. The quiet ring current, $c$, reaches its highest level before the spring equinox. The two parameters are well correlated (Figure 2c).

The pressure correction measures up to 86 nT and for about 4% of the time it is above 20 nT (Figure 2d). Thus solar
wind pressure enhancements can obscure a significant part of stormtime activity. For a statistical study this is of minor importance since the large pressure increases are statistically insignificant. For a dynamical study, however, the correction can significantly modify the model timescales and coefficients in individual phase space regions.

3. Model Estimation

3.1. The Burton et al. [1975] Model

This is a basic model for $D_{st}$ whose form has been adopted by subsequent models [e.g., Valdivia et al., 1996; Klimas et al., 1997, 1998]. The pressure-corrected index $D_{st}^{(0)}$ is driven by an injection function $F$, which depends on the solar wind input. The current adjusts to the input level with a timescale $\tau = 1/a = 7.7$ hours so that

$$\frac{dD_{st}^{(0)}}{dt} = F(t) - aD_{st}^{(0)}$$  (3)

B75 used 2.5–min averages of interplanetary and geomagnetic data to calculate the parameter $a$ and pressure-corrected parameters $b$ and $c$. The $D_{st}$ index was calculated as the average of the axial field at 12 midlatitude stations. The injection function $F$ was set proportional to a rectified component of the solar wind $\mathbf{v} \times \mathbf{B}$ electric field, $E \propto \nabla \times (\mathbf{v} \times \mathbf{B}) \times 0.5$ m/s, with $0$ being the step (Heaviside) function. Here $V$ is the $V_e$ component of the solar wind velocity, and $\mathbf{B}$ is the IMF, both in geocentric solar-magnetic (GSM) coordinates. The solar wind input $E$ was delayed by 25 min relative to $D_{st}$ and smoothed with a low-pass filter of a corner frequency $1/(0.5$ hours) and an attenuation of $6$ dB per octave. To measure the decay time $\tau = 1/a$, B75 identified 19 1-hour-long stormtime intervals with $D_{st} < -20$ nT and $B_t > 0$ so that approximately $F = 0$, and fit (3). Effectively, they measured $\tau$ for moderate/strong activity levels during the recovery phase. As a matter of fact each parameter was measured in a region of phase space (cf. regions in Figure 1) where its effect would be most clearly seen, and then it was assumed that the same values could be used in all phase space regions. In spite of this simplification the resulting model is an effective predictor of $D_{st}$ with out of sample correlations between predicted and observed time series reaching as high as $90\%$.

The discrete-time version of (3) is

$$D_{st}(t + \Delta t) = a_0D_{st}(t) + b_0F(E(t))$$  (4)

where $a_0 = 1 - a\Delta t$ and $b_0 = F\Delta t$. This will be useful in comparisons with the models of 3.3 and 3.4.

3.2. Other $D_{st}$ Models

The high correlation between $D_{st}$ and solar wind parameters has led to a large variety of models [reviewed by Detman and Vassiliadis, 1997]. For example, the index correlates well with the electric field $VB_p$ proportional to the reconnection rate on the dayside, and effectively a solar wind energy input rate to the magnetosphere. The component $B$, is the negative part of the IMF $B$.

Linear prediction filters [Clauer, 1986] using a $VB_p$ input produce an output which is highly correlated with $D_{st}$ (90%) at the 1-hour resolution [Iyemori et al., 1979; Trattner and Rucker, 1990]. Using multichannel filters McPherron et al. [1984] were able to reproduce 70% of the variance of the observed 5-min $D_{st}$. The impulse response of these filters increases in the first hour and decays exponentially with a timescale of 6–8 hours. A prolonged solar wind input results in a main phase of a few hours and a recovery of many hours to days. These studies show that the response to an idealized delta-function-like solar wind input (i.e., the impulse response function) would be a rapid (few-minute) increase and slow exponential decay of $D_{st}$. For more complicated inputs the response can be constructed from superposing these elementary responses weighted by the solar wind at a given instant. The exponential decay is interpreted to mean that the ring current strength decays exponentially so that at equilibrium it balances the flux of particles convecting or injected due to the electric field.

By extending the B75 model to activity-dependent (nonlinear) coefficients [Valdivia et al., 1996] and to higher order
one can increase the prediction capability. Finally, neural networks provide empirical models for the solar wind $- D_{st}$ coupling and have been used to optimize the solar wind input function [Freeman et al., 1994; Wu and Lundstedt, 1996].

3.3. The Linear ARMA Model

The time resolution and length of the data allows us to go to a higher-order differential equation than B75,

$$\frac{d^2 D_{st}}{dt^2} + \alpha_1 \frac{dD_{st}}{dt} + \alpha_0 D_{st} = \beta_1 \frac{dV B_s}{dt} + \beta_0 V B_s$$

(5)

an “oscillator” model for $D_{st}$ [Klimas et al., 1997, 1998]. When we discretize it, we obtain an autoregressive moving-average (AR(2)/MA(2)) model with two autoregressive (AR) terms coupled to $D_{st}$ and two moving-average (MA) terms coupled to $V B_s$:

$$D_{st}(t + \Delta t) = a_0 D_{st}(t) + a_1 D_{st}(t - \Delta t) + b_0 V B_s(t) + b_1 V B_s(t - \Delta t)$$

(6)

The model can be seen as an extension of B75 (equation 4) in two ways:

1. We increase the range resolution in $D_{st}$ and other variables, so that the model covers a sequence of activity levels and all storm phases. The coefficients will vary with activity level, whereas the B75 model is linear and time-independent.

2. Similarly to recent models by Klimas et al. [1997, 1998], we introduce a second timescale by going to a higher-order differential equation. This is justified by reports of a secondary peak in the response function [Iyemori et al., 1991; McPherron et al., 1984]. Also, rapid processes of ring current formation and loss can be manifested as additional model timescales. The B75 model made use of 2.5-min data, but the coefficients were obtained by selecting appropriate intervals for each parameter, calculating the parameter value for 1-hour intervals and then regressing the secondary values. (In B75, each model parameter was fitted in a different phase space region and then those values were used for all regions.) Other $D_{st}$ models developed from 1-hour data can only explain the long timescale features (hours/days) of storm phases [Feldstein, 1992].

The first two autoregressive terms on the right-hand side of (6) model the capacitive and inductive properties of the current. The next two moving-average terms represent the coupling to the $V B_s$ electric field and have dimensions of inverse impedance-per-unit-length. Since $D_{st}$ is proportional to the total energy density of the ring current, the $b_i$ terms represent magnetospheric currents that the solar wind input couples to in order to transfer energy to the magnetosphere, namely the magnetopause current and the tail current. If we neglect the AR feedback terms (which is justified if the current decay time is so small that the index adjusts quickly to the $V B_s$ variation), the remaining MA terms form the linear prediction filters [Fay et al., 1986] relating the input directly to the output. The impulse response functions of these filters have lengths of several hours.

While the moving-average coefficients $b_i$ can be calculated from (6) and the current intensity associated with them obtained directly, in order to get the timescales associated with $a_i$ we need to take a more circuitous route. If we dealt with linear prediction filters the timescales would be obtained quasempirically as the locations of the peaks of the impulse response function. With ARMA filters we calculate them from the $a_i$ values.

The relation between the $D_{st}$ timescales and the $a_i$ values is best seen in the frequency domain. We apply a Laplace transform on $\phi$ for the discrete-time system we use the z transform [Kailath, 1980] and obtain

$$\tilde{D}_{st}(\omega) = \tilde{H}(\omega)\tilde{V}B_s(\omega) + I_0$$

(7)

The transfer function $\tilde{H}(\omega)$ represents the solar wind-magnetosphere coupling between the spectral components $\tilde{D}_{st}(\omega)$ and $\tilde{V}B_s(\omega)$. A tilde represents the Laplace transform of a given quantity. The term $I_0$ depends on the initial conditions and can be set to zero if the geomagnetic activity starts from a quiet level ($D_{st} \approx -23$ nT) which is what we assume. For an ARMA model with $m$ autoregressive and $l$ moving average terms, the transfer function can be calculated analytically as a rational function of the model coefficients fit to the time series [Kailath, 1980]:

$$\tilde{H}(\omega) = \frac{\sum_{i=0}^{m-1} b_i p^{-i}}{1 - \sum_{i=0}^{l-1} a_i p^{-i}} = \frac{\sum_{i=0}^{m-1} b_i p^{-i}}{p^m - \sum_{i=0}^{l-1} a_i p^{-i}}$$

(8)

where $p \equiv \exp(\omega \Delta t)$. In equation 6 we have $m = 2 = l$.

The timescales of $D_{st}$ are calculated by setting the denominator polynomial $p^m - \sum_{i=0}^{l-1} a_i p^{-i} = 0$, solving for $p$ and then using $\nu_i \equiv 1/\text{Re}(\omega_i)$. The roots $p = p_i$ of the polynomial represent the lowest-order variations of $D_{st}$ which are called modes. The roots are the poles of the transfer function (i.e., $\tilde{H}(\omega)$ diverges for those values). The timescales are generally complex with a real part $\text{Re}(\omega_i) \equiv \text{Re}(\ln p_i)/\Delta t$ representing the decay/growth of $D_{st}$ and the imaginary ($\text{Im}(\omega_i)$) representing an oscillation. For the standard 1-hour $D_{st}$ the poles, and therefore the decay rates, are found to be real-valued. Even for the 5-min data the first timescale $\tau_1 \equiv 1/\text{Re}(\omega_i)$ is usually of the order of several hours, while the second “inductive” rise time $\tau_2 \equiv 1/\text{Re}(\omega_2)$ is much smaller. Both are negative, corresponding to decay of the disturbance towards an equilibrium determined by the solar wind.

Now that we have calculated the timescales from the $a_i$ coefficients in the frequency domain, we can return in the time domain and identify them in the impulse response function, as we would if we had been calculating linear prediction filters. The impulse response function represents the collective effect of all processes that are energized by the solar wind input. We apply the inverse Laplace transform to the transfer function in 7 and obtain:

$$D_{st}(k\Delta t) = \sum_{i=1}^{m} a_i e^{\nu_i k \Delta t} + \sum_{j=0}^{l} H(j\Delta t)V B_s[(k-j)\Delta t]$$

(9)

where the impulse response function of the ARMA model is

$$H(t) = H(k\Delta t) = \sum_{i=1}^{m} \frac{\sum_{j=0}^{m-1} b_i p_j^{-i-j}}{p_i^{m} - \sum_{j=0}^{l} a_j p_i^{-j}}$$

(10)
a. Characteristic Frequencies  

First-Order Model:  
[Burton et al., 1975]

- **One Frequency**
- **Exponential Decay**

Second-Order Model:

- **Two Real Frequencies**
- **Rise-and-Decay Response**

- **Two Complex Frequencies**
- **Exponent-Decaying Oscillation**

**Figure 3.** (a) Characteristic frequencies in the Burton et al. [1975] (equation (4)) and second-order model (equation (6)) are plotted on the complex plane. The second-order model has either two real-valued frequencies or two complex-conjugate frequencies. (b) The frequencies determine the response of $D_{st}$ following a solar wind $VB$ impulse. In the case of complex conjugate frequencies, the response is oscillatory.

The impulse response contains, in a compressed form, the information on all the physical processes given by the timescales and coupling coefficients. Figure 3 shows the characteristic frequencies and impulse response functions for a first-order model (like B75) and a second-order model.

### 3.4. Nonlinear ARMA Models

If the linear analysis is performed for subsets of the data associated with different activity level and storm phases rather than the whole data set, then the coefficients generally change with activity level, and become functions of the three variables:

$$ a_i = a_i(D_{st}, dB_{st}/dt, VB) $$

Small, statistically significant variations show the nonlinearity of the storm dynamics that shape $D_{st}$. More importantly, if the variations are significant or qualitative, they are evidence of different processes in each phase of the ring current development. For example, in some storm phases the timescales will be found to be complex conjugate, so the $D_{st}$ response changes from monotonic to oscillatory with period $1/Im(\omega_1)$ (Figure 3). In other parts of the storm evolution the timescales are positive corresponding to a slow exponential growth.

We divide the phase space in cells (see Figure 1b), and the regression (6) is carried out separately in each cell. The coefficients $a_i, b_i$, frequencies $\omega_i$, and response functions $H(t)$ are parametrized by the value of $(D_{st}, dB_{st}/dt, VB, \gamma)$ at the center of the cell. In earlier work, we had instead followed the "trajectory" of $D_{st}$ in the phase space and defined a local linear ARMA model in a small region of the phase space at each time step. While that method is quite good in prediction [Vassiliadis et al., 1995; Valdivia et al., 1996], the local region often may not contain a sufficient quantity or quality of data points for calculating model coefficients. Here we overcome this problem by subdividing only in a few large cells and thereby averaging over the large number of data in each cell, but the price for that is a lower prediction accuracy. Other ways to solve the problem are Kalman or adaptive filtering techniques.

On the basis of this division in phase space cells we discuss the following four models:

1. **Model 1** is the linear ARMA model of equation 6. There is no division of the phase space in cells, so in a sense this is a simpler model than B75 in that it averages over all storm phases.

2. For the **2nd model** we divide the space into $2 \times 2 \times 2 = 8$ cells. The use of $D_{st}$ and $dB_{st}/dt$ as phase space variables divides the data according to the storm activity level and phase, respectively, while in $VB$, direction the data are divided depending on the interplanetary activity level (less or greater than 1 nV/m). This model is an extension of B75, in the sense that the B75 would calculate the decay rate parameter $a$ from one of the two R cells (namely the cell $(D_{st} < 0 nT, dB_{st}/dt > 0 nT/min, VB < 1 mV/m)$), the pressure correction $b$ and $c$ from two of the EC and LC cells $(D_{st} > 0 nT, VB < 1 mV/m)$, and similarly for other parameters.

3. **Model 3** is a finer subdivision of the space into 4 cells per coordinate. For each coordinate, the 5 boundaries $g_i$ are defined so that each cell contains an equal number of data (1/4 of the total), when the data are projected on that axis. For example for $D_{st}$ if $P(D_{st})$ is the probability density function, the boundaries $g_i$ are calculated from

$$ g_{i+1} = \int_{g_i}^{g_{i+1}} P(D_{st}) dD_{st} = \frac{1}{4} \quad i = 1, \ldots, 4 \quad (12) $$

However, we need to separate positive from negative values so we set $g_1 = 0$. Similarly for $dB_{st}/dt$, we must set $g_1 = 0$. See Table 1.

4. **Model 4** is constructed in a way analogous to that of model 3. Again for $D_{st}$ and $dB_{st}/dt$ $g_i$ and $g_s$, respectively, are set to zero so as to separate positive from negative values.

### 4. Storm Timescales

#### 4.1. Nonlinear Decay and Growth of $D_{st}$

First we show how the model coefficients of (6) vary with activity level and storm phase. The dependence becomes more apparent as the phase space resolution increases, but this is eventually masked by another effect, the increase of
Table 1. Boundaries of $D_{st}$, $dD_{st}/dt$, and $VB_s$ Cells Used in Model 4\(^3\).

<table>
<thead>
<tr>
<th>$g_1$</th>
<th>$g_2$</th>
<th>$g_3$</th>
<th>$g_4$</th>
<th>$g_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-227.90</td>
<td>-39.19</td>
<td>-20.06</td>
<td>0.00</td>
<td>94.79</td>
</tr>
<tr>
<td>-18.06</td>
<td>-0.35</td>
<td>0.00</td>
<td>0.41</td>
<td>11.57</td>
</tr>
<tr>
<td>0.00</td>
<td>0.61</td>
<td>1.21</td>
<td>2.08</td>
<td>26.2</td>
</tr>
</tbody>
</table>

the statistical uncertainty in the coefficients, $\Delta a_i$ and $\Delta b_i$. Figure 4a shows both effects in the coefficient $a_0$ for models 1\(^3\), 2\(^3\), and 4\(^3\). The graphs are arranged from lower solar wind activity level (bottom) to higher (top), and from low resolution model (left side) to high (right side). Note that as we go from the low-resolution to high-resolution models, the coefficient is initially constant and numerically close to the B75 values (model 1\(^3\)), then varies somewhat systematically depending on the phase and activity level for model 2\(^3\), and then varies more randomly for model 4\(^3\). Going to even higher resolution (8\(^3\), not shown) produces phase space cells that contain so few data that the solution of (6) becomes numerically singular. The numerical singularities lead to random variations of the coefficients with phase and activity level. Similar effects are shown in Figure 4b for coefficient $b_0$ (the current intensity that the solar wind electric field couples to).

![ARMA Filter Coefficient $a_0$](image)

**Figure 4a.** Model coefficients, here $a_0$ (in equation (5)), change systematically with storm phase and solar wind activity level. In the leftmost graph the coefficient is constant for the linear model 1\(^3\). For higher resolution (models 2\(^3\), 4\(^3\)) $a_0$ depends on the storm phase: early and late commencement (FC, LC), main phase (M), and recovery (R). At the same time, the measurement uncertainty grows as the average number of data per cell goes down, and this limits the maximum resolution. Error bars denote the uncertainty $\Delta a_i$ from solving equation (6) for $a_i$, $b_i$. The $\Delta a_i$ increases with the resolution; in the case of the 8\(^3\) model (not shown) the uncertainties are too large for any conclusions to be drawn.

![ARMA Filter Coefficient $b_0$](image)

**Figure 4b.** Same as Figure 4a, but for coefficient $b_0$ which represents the coupling to the solar wind $VB_s$. 
Figure 4c. Same as Figure 4a, but for the first timescale (decay/growth) \( \tau_1 = 1 / Re(\omega_1) \). Most values are negative, indicating an exponential decay of \( D_\alpha \). However, in several cases in Model 4^3 positive values appear associated with exponential growth. In some cases numerical roundoff error can make the frequency close to zero, so that the timescale value diverges. In order to keep the vertical scale reasonable, we limit the value of \(|\tau_1|\) to no higher than 15 hours.

The uncertainties in the coefficients, \( \Delta \alpha \), and \( \Delta \beta \), increase dramatically as the phase space is divided in smaller cells (Figures 4a and 4b). Both \( \Delta \alpha \) and \( \Delta \beta \) scale inversely proportional to the number of data per cell, and that number decreases by approximately an order of magnitude from one model to the next. This limits the phase-space resolution of these models. For instance, the increase of \( \Delta \alpha \) (Figure 4a) shows the difference between using the ARMA models for calculating timescales and for predicting \( D_\alpha \). \( \Delta \alpha \) grows roughly inversely proportional to the number of data per cell. For models 1^3 and 2^3, \( \Delta \alpha < |\alpha| \), so \( \alpha \), \( \beta \) and the resulting timescales vary smoothly with storm phase and activity level. Because the phase space cells are so large, however, these two models are close to linear models and their predictions are not significantly more accurate than B75 or other linear models. For 4^3 the uncertainty is comparable to the value of the coefficient, while for model 8^3 the number of data per cell makes \( \Delta \alpha > |\alpha| \), and timescales of very rapid growth and decay appear at all storm phases. At the other end, if the objective is to predict \( D_\alpha \) rather than calculate the timescales, it is best to use small phase space cells with few data points (the so-called “nearest neighbor” points by, e.g., Vassiliadis et al. [1995] and Vallis et al. [1996]). However, at such high resolution, the uncertainty \( \Delta \alpha \) is too large for the \( \alpha \) to show any systematic variation with storm phase and activity level.

A similar effect can be seen for the \( D_\alpha \) timescales which are derived from the \( \alpha \), \( \beta \) coefficients. Model 1^3 has one pair of real-valued timescales (Table 2). Each one produces a decaying exponential, \( \exp(t/\tau_c) \) with the “inductive” timescale \( \tau_c \) dominating \( H(t) \) before \( t = (ln(\omega_1/\omega_2))/(\omega_1 - \omega_2) \), and the longer timescale \( \tau_l \) taking over afterwards (compare with Eq. (10) and Figure 3).

In model 2^3 the dominant timescale \( \tau_l \) is always negative, but the value changes with activity level (Table 3). The magnitude of the second timescale is always less than 1 hour, and this might justify approximating the response with an exponential decay for low-resolution \( D_\alpha \) data as suggested in the B75 model. However, as the phase-space resolution increases (which is what happens with model 4^3 and its 64

Figure 4d. Same as Figure 4c, but for the second timescale (usually an “inductive” rise time) for the \( D_\alpha \) response, \( \tau_2 = 1/Re(\omega_2) \) (see Figure 3). The \( |\tau_2| \) values are limited to no higher than 2.5 hours.
Table 2. Timescales of Model 1 for $D_N^{(0)}$

| Decay time, $\tau_1$ | -6.05 |
| Rise time, $\tau_2$  | -0.80 |

Timescales are given in hours.

Pairs of timescales it becomes possible to resolve their variation with position in the phase space (Figures 4c and 4d). While in some cases we still find $\tau_2 \ll \tau_1$, in most situations the timescales are comparable in size. Further, in some of the phase-space cells that correspond to the early main phase, one frequency may be small and positive so $D_n$ has a slow exponential growth. Evidently the ring current dynamics cannot be represented by a single timescale independently of storm phase and activity.

4.2. Impulse Response Functions and Coupling to the Solar Wind

The impulse response function (10) depends on the frequencies $\omega_{1,2}$; therefore it too will vary with phase space region. Figure 5 shows the eight impulse responses for model 2 with the single response function of model 1 superposed as the heavy solid line. High solar wind input ($>1$ mV/m) is indicated with a solid line and low or zero input is indicated with a dashed line. In the recovery phase the impulse response is almost identical to the average response, which is not surprising since the recovery phase contains most of the data points. The main phase, however, has a faster response than the average for high input and a much slower response for low. Further, during storm commencement the response changes qualitatively. In late commencement (LC) it is underdamped with an oscillation period of 16 min for $D_N^{(0)}$. In both the early and late commencement phases the response is significantly faster than during main or recovery phases. This is due to the faster timescales of the magnetopause current response and adjustment of the inner magnetosphere (section 4.3). The qualitative change in the impulse response and dependence on storm phase and activity level is even more pronounced in the higher resolution models, 4 and 8. There complex-valued frequencies appear in almost all regions of the phase space and levels of activity. We find underdamped responses with few-minute oscillations during storm commencement and few-hour oscillations in the main and, less frequently, in the recovery phase. For simplicity we call “fast” the oscillations with $\text{Im}(\omega_1) > 1$ hour$^{-1}$ and “slow” the ones with smaller values of $\text{Im}(\omega_1)$. Their spectral characteristics are different, and they arise because of different magnetospheric processes as is discussed next.

Table 3. Timescales of Model 2 for $D_N^{(0)}$ st

<table>
<thead>
<tr>
<th></th>
<th>VB $&lt; 1$ mV/m</th>
<th>VB $&gt; 1$ mV/m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>$\tau_2$</td>
</tr>
<tr>
<td>Early commencement</td>
<td>-2.83</td>
<td>-0.031</td>
</tr>
<tr>
<td>LC</td>
<td>-1.22</td>
<td>0.113</td>
</tr>
<tr>
<td>Main phase</td>
<td>-24.07</td>
<td>-0.046</td>
</tr>
<tr>
<td>Recovery</td>
<td>-8.32</td>
<td>-0.050</td>
</tr>
</tbody>
</table>

Timescales are given in hours.
4.3. Fast ( < 1 hour) Oscillations: Magnetopause Vibrations

For model 2³, there are oscillations only in the late commencement phase, with a period of 16 min for Dst (0) (Figure 5). Correcting for solar wind pressure is important. The uncorrected Dst data do not show an oscillatory response. In model 4³ we find the fast oscillations occurring in the late commencement phase for solar wind levels 0–2, and early commencement phase for the highest solar wind activity level, 3 (Figure 6a). The average value is 8.48 min for Dst (0) and 1.13 hours for Dst.

The fast oscillations are produced by variations in the magnetopause current intensity and altitude as well as changes in the inner magnetosphere population. During the commencement phase that follows the pressure increase, the magnetopause adjusts its position and current strength. The pressure pulse launches waves from the magnetopause surface inwards with speeds exceeding 100 km/sec. These waves are seen as a brief broadband response in the dynamic spectrum of Dst (examples given in section 4.5). Fast oscillations are also related to modulations of the energetic particle population in the inner magnetosphere. Storm commencement and more generally sudden impulses are correlated with VLF wave intensifications at timescales of 10 min [Nishida, 1978]. Magnetosphere compression and an increase in energy anisotropy excite the electron cyclotron instability which generates the waves. The 10-min timescale should not be confused with the time it takes a disturbance to propagate radially to the ionosphere at the fast mode speed (∼ 1000 km/sec) which is roughly 1 min.

4.4. Slow ( > 1 hour) Oscillations: Injections and Drift Echoes

At different frequency regime we find oscillations during the main phase and, to a lesser extent, the recovery phase. They are not seen in model 2³ because of its low phase space resolution, but appear in the response functions of model 4³. Figure 6a shows the regions in the 4³ phase.
We select four phase space cells in the main phase, marked with 'X', and we measure the local impulse response functions which are shown in Figure 6b. In three out of the four cases the response is a damped oscillation as was illustrated in Figure 3.

In the case of model oscillatory responses appear in virtually all phase space regions. Correcting for solar wind pressure reduces the oscillations' amplitude and increases their frequency. However, the uncertainty in the coefficients is comparable to their values, so for this model the oscillatory responses may be partly due to numerical errors.

4.5. An Aspect of the Storm-Substorm Relationship

In order to identify the source of the oscillations we searched the $D_s$ and other time series for frequencies similar to those in the ARMA response functions. We will argue that these oscillations are due to changes in the ring current population rather than due to other current systems.

To detect the oscillations in the $D_s$ index, we use spectrograms (dynamic spectra). Currents that produce the oscillations can be induced by causes either external or internal to the magnetosphere. An external cause can be a pressure increase, as in an interplanetary shock; an internal one may be detectable in other magnetospheric activity measures such as $AE$ indices (which are of course more closely related to substorms than storms). In order to identify these effects we compared the $D_s$ spectrogram to those of $P_w$ and the $AL$ index. The DC component of the two indices is usually much larger than any individual other frequency, so we examined spectrograms of the derivatives of these quantities. For the power spectrum estimation we used a 4-hour (48-step) Hanning window [Press et al., 1993]. For $AL$ the results do not differ significantly whether we use the 1 min data or reduce the resolution to 5 min.

Plate 1 shows an example of such oscillations occurring on January 2, 1979, during the main phase of a moderate storm. From upper to lower, the four panels show the spectrograms of the derivatives of the solar wind pressure, the rectified interplanetary electric field, and the midlatitude and electrojet indices. The color contours are in log base-10 scale. Superposed on each spectrogram are the time series (lower solid line) and its derivative (upper). The range of the time series is given on the right of the spectrogram. The five dotted horizontal lines separate the first 5 frequencies in the spectrogram: the lowest one is the DC component, the second the fundamental at 1/(4 hours), the next the first harmonic at 1/(2 hours), etc. At those times when the power in the range 0.5–1 hours $^{-1}$ (3 bins) is greater than the power for frequencies in the range 1–2.5 hours $^{-1}$, we mark the time series with white dots. (We do not include the DC component or the 4-hour bin which may contain aliased power or too long timescales.)

The oscillation in $D_s$ (third panel) starts close to 1040 UT and its peak frequency grows with time from 1030 to 1300 UT. The oscillations in the index reappear after 3.5 and 7 hours. Also at those times the program has marked the $D_s$ time series with white dots, showing that the power of the 1–2-hour oscillations dominated over the high-frequency part of the spectrum. No such characteristic is seen in the interplanetary electric field or pressure. However, shortly before the oscillation appears the $AL$ activity has started increasing in a series of activations (probably substorms). We interpret the oscillations as an effect of substorm activity.
Plate 1. A slow oscillation event in the main phase of a moderate storm can be seen in the dynamic spectra of the derivatives of (from upper to lower): $P_{sw}$, $E_{sw} = VB_D$, $D_s$, and AL for the day of January 17, 1979. The interplanetary data have been propagated from ISEE 3 to the subsolar magnetopause. All spectrograms are shown in a base-10 logarithmic scale. The time series and its derivative are superposed on each spectrogram for reference. The range of the time series is given on the right (derivatives are not to scale). Horizontal lines separate the DC component, the fundamental of 1/4 hours and the first 3 harmonics up to 1 hour$^{-1}$. White dots on the time series mark the times when the spectral power of 1-2 hours is dominant over that of all higher frequencies. Note (1) the oscillation in $D_s$ drifting from low frequencies to high, (2) the onset of an AL intensification immediately preceding the oscillation, and (3) no spectral content similar to $D_s$ either in the pressure or the electric field. Grey boxes denote data gaps. For more details see text.

Plate 2. Same as Plate 1, but showing several oscillation intervals in a series of minor storms over 3.5 days starting from January 15, 1979. The minimum of $D_s$ is -37 nT and that of AL is -858 nT. The 1-2 hour oscillations are narrowband and appear during the early main phase. In the 4 oscillation events marked by black arrows the peak frequency drifts from low to higher frequency. From these and other features they are interpreted as due to particle injections rather than modulations of high-latitude currents. Higher-frequency oscillations (marked by red arrows) occur after pressure increases and are broadband in spectrum.
on the ring current, rather than interplanetary effects or the growth of the magnetopause current. The activity may be caused by high-latitude ionospheric currents or lower-latitude injected particles.

Plate 2 shows several examples of the slow oscillations over an interval of 3.5 days. They occur at \( t = 11, 18, 58, \) and 64 hours (marked by black arrows), in the beginning of main phases of small storms and coinciding with AL intensifications. The peak frequencies start from near DC and increase with time as was the case for the oscillation of Plate 1. Incidentally there are also broadband oscillations (at \( t = 14 \) and 45 hours, as marked by red arrows) as responses to increased solar wind pressure, which correspond to the fast oscillations of section 4.3. In those cases the peak frequency appears initially at high frequencies and in a few minutes spreads to lower ones.

We have also identified the 1–2–hour and slower oscillations in the axial field magnetograms of midlatitude stations. Spectrograms of \( dh/dt \) show that during main phase the peak frequency moves from the DC component to a spectrum with a peak close to 1–2 hours. The high amplitudes of the magnetogram oscillations are reduced in the averaging procedure that produces \( D_s \).

What is the interpretation of the oscillations? Since they occur during the storm main phase, coincide with AL activity and in each instance peak around well-defined periods, our explanation is that they are related to particle injections during substorms. After such an injection the trapped particles are detected as drift echoes by spacecraft in the inner magnetosphere [Mauk and Meng, 1983]. The echo period is inversely proportional to the drift speed. The correlation between storm peak \( D_s \) and injected relativistic particles has been measured in several ways [Reeves, 1998, and references therein], similar studies have been carried out for less energetic particles. If the injection occurs in a narrow range of \( L \) shell and azimuth, the resulting particle population will drift as a beam and will induce a narrow range of geomagnetic oscillations rather than a broadband signal, this is what is observed in Plates 1 and 2. The peak frequency is consistent with the drift period of particles with energies close to 100 keV (2 hours) or 200 keV (1 hour). All ground stations measure similar oscillations within fractions of an hour, which is consistent with the drift period. Finally, in comparisons made when this paper was in review, geosynchronous injections measured by LANL spacecraft occur at the same time as the oscillations.

An alternative explanation is that the oscillations are due to an intense substorm current wedge on an expanded oval. This explanation can be rejected since (1) the geomagnetic effect appears even for very low AL intensifications which are not consistent with an expanded oval, (2) the spectrum is narrowband rather than broadband (Plates 1 and 2), and (3) within one hour it is observed at all midlatitude magnetometers, not only those on the nightside.

### 4.6. Dependence of Decay Rates and Oscillation Frequencies on Storm Amplitude and Phase, and Solar Wind Activity

To obtain a scaling relation of the timescales with activity and storm phase we use the coefficients of model 4\(^2\) (the uncertainty in the coefficients of model 8\(^3\) is higher due to the small number of data in the phase space cells of that model). Figure 7 shows the scaling of the average decay rate and corresponding timescale \( \tau_i = 1/\text{Re}(\omega_i) \) with each of the three variables. We used data from those phase space cells that satisfy \(-2 \text{ hours} < \tau_i < -20 \text{ hours} \); we considered other measurements to contain numerical errors. We make the following observations:

1. As shown in Figure 7a the rate increases with \( VB_i (\omega_i) \) (averaged over \( D_s \) and \( dD_s/dt \)). (Note that the abscissa values are not equidistant: They are the values at the center of the cells where the averaging has been performed.) This is expected since stronger solar wind input imposes a stronger field and also accelerates greater numbers of plasma sheet particles which either convect or are injected into the ring current. As discussed in sections 4.2–4.5, this increase occurs for \( \text{Re}(\omega_i) \) and \( \text{Im}(\omega_i) \). However, the decay is fastest for intermediate inputs \( (VB_i = 1.7 \text{ mV/m}) \) and shown down for higher inputs.

2. Figure 7b compares \( \tau_i \) during recovery and main phase and during early and late commencement (the averaging is now performed over \( D_s \) and \( VB_i \)). The average decay time is longer during recovery, with \( \tau_i = 5.8 \) versus 4.9 hours during the main phase. The main phase is dominated by a fast response from particles being accelerated by the electric field, whereas during most of the recovery time the dominant effect is a slower energy diffusion mainly due to charge exchange of the trapped population [Feldstein, 1992]. In some less frequent situations where \( |dD_s/dt| > 1 \text{ nT/5 min,} \) corresponding to high rates of either injection or diffusion, the time \( \tau_i \) is much smaller (\( 2.4–5.3 \text{ hours} \)) for either storm phase. In those cases the current system reduces any excess energy quickly and returns to the normal slowly evolving state. Also, the decay rate is faster in early commencement than in late commencement.

3. The decay time is more complex as a function of \( D_s \) (Figure 7c). In the main phase (solid line) the decay time decreases from 6.8 hours for \( D_s \) close to \(-10 \text{ nT} \) down to \(-3.9 \text{ hours} \) for \( D_s \) close to \(-134 \text{ nT} \). In the recovery phase (dotted line) the recovery is slower than in the main phase; as recovery progresses the decay rate changes from \(-4.9 \text{ hours} \) to \(-8.9 \text{ hours} \) (compare with Figure 8d). Note, however, that for small storms (\( D_s \) close to \(-10 \text{ nT} \)) the decay is much faster. For EC it is 3 hours, while for late commencement it is about 7.1 hours. These results are consistent with the statements by Feldstein [1992] given in section 1.

4. When the response is oscillatory (\( \text{Im}(\omega_i) \neq 0 \)), the decay of \( D_s \) is faster. In Figure 7d we plot the decay rate \( \text{Re}(\omega) \) measured in phase space region M. We divide the data in two groups depending on whether \( \text{Im}(\omega) \) equals zero or not. The decay rate is between 2 and 4 times higher when the response is oscillatory (solid curve) than when it is monotonic (dotted curve).

In a dynamical example, Figure 8a shows how the decay rate and other variables change during a storm. The decay rate is plotted as a color contour on a \( (D_s, dD_s/dt) \) cross section of the phase space. As the ring current evolves the point \( (D_s, dD_s/dt) \) traces out a trajectory in the space. Figures 8b–c show the variation of \( D_s \) and \( dD_s/dt \), and 8d shows the time series of \( \tau \). Note the increase of \( \tau \) during recovery phase as \( D_s \) increases back to zero (compare with statement 3 in section 1).

These properties quantify the statements 1–4 in section 1. In addition they demonstrate that although the midlatitude disturbances are often interpreted as the result of linear dy
Figure 7. Average decay rate $\text{Re}(\omega_1)$ and its inverse, $\tau_1$, for Model 4 as a function of (a) $D_n$, (b) $dD_n/dt$, and (c) $VB_s$. Data outside the range $-2 \text{ hours} > \tau_1 > -20 \text{ hours}$ have been excluded. The abscissa values are not equidistant. They are the values at the center of the phase space cells in which the averaging was performed. (d) The main phase decay rate is plotted versus $VB_s$ when oscillations are present ($\text{Im}(\omega_1) > 0$) or absent. During the early main phase (small $VB_s$), the decay is much faster when $\text{Im}(\omega) > 0$ than when it is zero. Therefore the large-scale changes in the geomagnetic disturbance occur during the oscillatory part of the main phase which we associate with substorm injections. Again the abscissa values indicate centers of phase space cells.
5. Prediction of $D_{st}$ and Model Stability

We have used the ARMA models to derive timescales and coupling strengths related to ring current dynamics, but they can also be used to predict $D_{st}$. Given a set of model coefficients for each phase space region we predict the index one-step ahead. The prediction is iterated over several steps by using the prediction at step $k$, $\tilde{D}_{st}(t+k\Delta t)$, and the observed input $V_B(t+k\Delta t)$ to obtain $\tilde{D}_{st}(t+(k+1)\Delta t)$. Prediction capability is measured by the average absolute prediction error [Vassiliadis et al., 1990], or by the correlation coefficient between prediction and observation. In addition it is useful to consider prediction skill. This is the improvement of the prediction of a given model relative to the prediction of a reference model. Here we compare the models $1^3$, $2^3$, and $4^3$ to each other.

While the predictions with the ARMA models are quite ac-
accurate, they should be compared with simpler models whose
description capabilities are similarly high (again the reason is
that these models have been optimized for coefficient calcula-
tions rather than for prediction as they have been elsewhere
[Valdivia et al., 1996]). For example, linear prediction filters
can reproduce 70% of the variance in 5-min $D_n$ data [McPher-
son et al., 1984]. Better predictions can be achieved by using
multichannel filters [Trattner and Rucker, 1990], nonlinear
prediction filters and ARMA models, and neural networks.
The techniques and results have been reviewed recently by
Detman and Vassiliadis [1997] and Lundstedt [1997]. Here we
measure prediction capability in several ways. First we repro-
duce the data that the model was calculated from (i.e., the
"in-sample" data) as a consistency check. This is a one-step
prediction. Second, the same model can be driven with the
same solar wind data and run over the data set (in-sample
multistep prediction). In the third test, we use the model
to make a multistep prediction of a different interval (out-of-
sample multistep prediction).

The performance of models $1^3$-$4^3$ in reproducing the data
is measured in terms of the average absolute error and cor-
relation (Figure 9 and Table 4). Model $4^3$ generally has the
lowest error for in-sample predictions. However the difference
is small.

An issue related to prediction is the models' robustness
or sensitivity to initial conditions, inaccuracies or gaps in
the solar wind data. Here we examine the first issue by
perturbing the initial condition by small amounts, essentially
performing a stability analysis. An instantaneous change in
initial condition can be considered as the effect of an impulsive
component in the solar wind input. It is then expected that
the resulting $D_n$ perturbation should have the form of the
impulse response function (10) and decay similarly, i.e., at a
rate $\sim \Re(\omega_1)$.

Figure 10 shows such a pair of perturbations in the initial
conditions with amplitude $\pm 10$ nT. The perturbations are
applied to the models $2^3$ and $4^3$ (Figures 10a and 10b, re-
spectively). After about 20 hours for $2^3$ and 40 hours for Model
$4^3$, the perturbed trajectories converge on the unperturbed
one within 1 nT. The prediction error drops exponentially
(Figure 10c) and reaches the computer program's numerical
precision. For model $4^3$ the error time profile is more compli-
cated, since in some phase space regions the solutions grow
exponentially with time ($\Re(\omega_1) > 0$).

6. Summary and Discussion

We measured the $D_n$ timescales as a function of activity
and storm phase. This approach leads to a nonlinear model
for $D_n$ which depends on changes in the solar wind pressure
and ring current composition. We discovered oscillations of
period $> 1$ hour in $D_n$ which we interpret as due to particle
injections from the plasma sheet into the ring current as well

<table>
<thead>
<tr>
<th>Regression</th>
<th>Av. Abs. Error [nT]</th>
<th>Model 1$^3$</th>
<th>Model 2$^3$</th>
<th>Model 4$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00767</td>
<td>0.00766</td>
<td>0.00765</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>In-Sample Run</th>
<th>Av. Abs. Error</th>
<th>Model 1$^3$</th>
<th>Model 2$^3$</th>
<th>Model 4$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6024</td>
<td>0.6025</td>
<td>0.5849</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>60.60</td>
<td>59.94</td>
<td>63.93</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Out-of-Sample Run</th>
<th>Av. Abs. Error</th>
<th>Model 1$^3$</th>
<th>Model 2$^3$</th>
<th>Model 4$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>26.05</td>
<td>25.03</td>
<td>25.83</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>75.27</td>
<td>71.42</td>
<td>70.11</td>
<td></td>
</tr>
</tbody>
</table>
as subsequent drift echoes. (After the paper was submitted, a comparison of $D_s$ with electron and ion injections from LANL geosynchronous spacecraft showed that the times of particle injections coincided with the oscillations.)

In the study the $D_s$ response was modeled as a nonlinear damped oscillator driven by the solar wind electric field as discussed earlier [Valdivia et al., 1996; Klimas et al., 1997, 1998; Vassiliadis et al., 1999]. Most of the time the $D_s$ decays exponentially at an average timescale of 4.69 consistent with previous studies [Burton et al., 1975; McPherron, 1995]. Under those conditions the inductive timescale $\tau_2$ is smaller than the decay time $\tau_1$ and can be neglected when working with 1-hour data. However, 5-min data show that the decay rate varies systematically with storm amplitude, phase, and solar wind activity. The response is exponentially growing at the beginning of the storm main phase, rapidly oscillatory during storm commencement and slowly oscillatory during main phase. The rapid oscillations (8 min), clearly seen in the pressure-corrected $D_s^{(0)}$, occur after the pressure has reached the peak value and while it decreases. These oscillations are consistent with magnetopause and inner magnetosphere vibrations during solar wind pressure enhancements. They are often broadband as would be interpreted by a step-like pulse inducing responses at all frequencies. The peak frequency decreases with time.

However, more interesting for the storm-substorm relationship are slow oscillations which usually appear during the main phase and coincide with $AL$ intensifications. These are found in all midlatitude magnetometers. They are confined in the >1-hour timescale range and often the peak frequency increases with time. Therefore we associate them with substorm injections and drift echoes. The spectral features of the oscillations resemble the spectra of particle energies measured at geostationary orbit [Mauk and Meng, 1983].

If the oscillations correspond to injections, we should use a modified "injection function" (3), which so far has been represented only by the solar wind input. We can write $D_s$ as a complex number $D_{st} = A e^{i\omega t} e^{i\omega t} = |D_{st}| e^{i\omega t}$ where the absolute value is the exponential monotonic decay, and the remainder represents the oscillatory response. After inserting it in (5) we obtain two equations. The one for the absolute value is

$$\frac{d^2}{dt^2} |D_{st}| + \alpha_1 \frac{d}{dt} |D_{st}| + (\alpha_0 - \omega^2) |D_{st}| = \left( \beta_1 \frac{dV}{dt} + \beta_0 V \right) \cos(\omega t)$$

(13)
Table 5. Coefficients of Model $1^3$

<table>
<thead>
<tr>
<th></th>
<th>$D_{st}$</th>
<th>$D_{st}^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>1.885±0.005</td>
<td>1.220±0.002</td>
</tr>
<tr>
<td>$a_1$</td>
<td>-0.886±0.005</td>
<td>-0.228±0.003</td>
</tr>
<tr>
<td>$b_0$</td>
<td>-0.304±0.026</td>
<td>-0.977±0.029</td>
</tr>
<tr>
<td>$b_1$</td>
<td>0.273±0.030</td>
<td>0.814±0.027</td>
</tr>
</tbody>
</table>

The injections, represented by $\cos(\omega_i t)$, modulate the solar wind input function. Substorms may play that role, since they change the total cross-tail electric field which determines how effectively plasma sheet particles convect into the inner magnetosphere. Another effect is the decrease of the oscillation frequency to $\alpha_o = \omega_i^2$.

Reproducing the geomagnetic oscillations seen in the main phase can be a test for ring-current models during substorms as well as injection models [Fok et al., 1999]. We will gain a better understanding for the oscillations when we compare the ground geomagnetic signature, and especially its spatiotemporal development [Valdina et al., 1999], to in situ data from ring current crossings.

Appendix A. Models $1^3$ and $2^3$

The following is a summary of the ARMA model construction procedure: (1) propagation of the upstream solar wind data to the subsolar magnetopause, (2) additional timeshift of the solar wind data relative to the $D_{st}$ by 25 min, (3) smoothing of $D_{st}$ and $V_{Bz}$, with a 25-min moving-average filter, (4) pressure correction. We used slowly-varying coefficients $b(t)$ and $c(t)$ for the pressure correction of Eq. (2) since both vary seasonally and in the course of a solar cycle. One may use simply the median values of that distribution: $b_{median} = 14.7 nT/\sqrt{\text{Pa}}$ and $c_{median} = -10.8 nT$. For comparison the B75 values are $b = 15.8 nT/\sqrt{\text{Pa}}$ and $c = -20 nT$. (5) Coefficients of model $1^3$ with and without pressure correction are shown in Table 5, and coefficients of model $2^3$ are shown in Table 6.

Table 6. Coefficients of Model $2^3$

<table>
<thead>
<tr>
<th>Storm phase</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$b_0$</th>
<th>$b_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_Bs &lt; 1 \text{ mV/m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>1.818</td>
<td>-0.820</td>
<td>-0.0199</td>
<td>0.0739</td>
</tr>
<tr>
<td>LC</td>
<td>1.876</td>
<td>-0.877</td>
<td>0.0672</td>
<td>-0.0782</td>
</tr>
<tr>
<td>M</td>
<td>1.874</td>
<td>-0.876</td>
<td>-0.0426</td>
<td>0.0101</td>
</tr>
<tr>
<td>R</td>
<td>1.905</td>
<td>-0.906</td>
<td>0.0380</td>
<td>-0.0356</td>
</tr>
<tr>
<td>$V_Bs &gt; 1 \text{ mV/m}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>1.801</td>
<td>-0.807</td>
<td>-0.300</td>
<td>0.300</td>
</tr>
<tr>
<td>LC</td>
<td>1.872</td>
<td>-0.875</td>
<td>-0.0192</td>
<td>-0.00223</td>
</tr>
<tr>
<td>M</td>
<td>1.886</td>
<td>-0.888</td>
<td>0.00862</td>
<td>-0.0545</td>
</tr>
<tr>
<td>R</td>
<td>1.882</td>
<td>-0.884</td>
<td>-0.0259</td>
<td>0.00317</td>
</tr>
</tbody>
</table>

Appendix B. Calculation of the Pressure Correction

Section 2 discussed the solar wind pressure on $D_{st}$ and the linear envelope formed in the ($\phi/\sqrt{P_{sw}}$, $D_{st}$) scatter plot [Araki et al., 1993]. The slope and intercept of the line correspond to the effect of the magnetopause current on the index in the absence of storms, and to the quiet ring current. For an interval short enough compared to the seasonal or solar cycle variability (e.g., 1 month long) these quantities can be calculated by modifying the standard least squares algorithm to an iterative, weighted procedure. Starting with a distribution of points $(x_i, y_i)$ and a line $y = bx + c$, the sum of squares of the distances from the line is

$$E(b, c) = \sum_{i} d_i^2 = \sum_{i} (y_i - bx_i - c)^2$$  \hspace{1cm} (B1)

The least squares algorithm adjusts $b$ and $c$ to find the minimum of the “energy” $E$. The algorithm weighs equally the data points above and below the line. If we want to approximate an upper envelope, below which are the majority of the distribution of points, and above which are only a few outliers, we weigh the points above more than the points below, so the new function to be minimized is

$$E'(b, c) = W_1 \sum_{\text{above}} d_i^2 + W_2 \sum_{\text{below}} d_i^2$$  \hspace{1cm} (B2)

where $0.5 \leq W_1 < 1$, $W_2 = 1 - W_1$. The minimization of $E'$ is a new line above the least squares line. By letting $W_1$ be close enough to 1, we obtain a good approximation of the envelope.

The solution can be found by integrating the dynamical system

$$\begin{align*}
  b_{n+1} &= b_n - \frac{\partial E'}{\partial b} \\
  c_{n+1} &= c_n - \frac{\partial E'}{\partial c}
\end{align*}$$

In practice we set $W_1 = 0.85$, and use the sum of absolute values $|d_i|$, rather than that of the squares $d_i^2$, in (B2).
Acknowledgments. The high-time-resolution $D_s$ data were provided by Y. Kamide and S. Sharma. We are grateful to R. Nakamura for preparing the $D_r$ index, magnetogram, and ISSE 3 data sets. N. Tsayagennko provided the 1-min $AL$ index. We thank R. McPherron, I. Dogli, T. Detman, M.-C. Fok, and W. Gonzalez for discussions.

Janet G. Luhmann thanks John W. Freeman and Walter D. Gonzalez for their assistance in evaluating this paper.

References


---

D. N. Baker, Laboratory of Atmospheric and Space Physics, Space Technology Building, Campus Box 590, Boulder, CO 80309-0590. (baker@lyux.colorado.edu)

A. J. Klimas, Code 692, Laboratory for Extraterrestrial Physics, NASA/Goddard Space Flight Center, Greenbelt, MD 20771. (alex@bokah.gsfc.nasa.gov)

J. A. Valdivia and D. Vassiliadis, Universities Space Research Association, Code 692, NASA/Goddard Space Flight Center, Greenbelt, MD 20771. (alej@roselott.gsfc.nasa.gov and vassi@lepst.gsfc.nasa.gov)

(Received January 4, 1999; revised March 25, 1999; accepted April 14, 1999.)