Solar, interplanetary, and magnetospheric parameters for the radiation belt energetic electron flux

D. Vassiliadis, S. F. Fung, and A. J. Klimas
NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Received 17 February 2004; revised 4 October 2004; accepted 24 November 2004; published 1 April 2005.

[1] In developing models of the radiation belt energetic electron flux, it is important to include the states of the interplanetary medium and the magnetosphere, as well as the solar activity. In this study we choose the log flux \( j_e(t;L;E) \) at 2–6 MeV, as measured by the Proton-Electron Telescope (PET) on SAMPEX in the period 1993–2002, as a representative flux variable and evaluate the usefulness of 17 interplanetary and magnetospheric (IP/MS) parameters in its specification. The reference parameter is the solar wind velocity, chosen because of its known high geoeffectiveness. We use finite impulse response filters to represent the effective coupling of the individual parameters to the log flux. We measure the temporal and spatial scales of the coupling using the impulse response function and the input’s geoeffectiveness using the data-model correlation. The correlation profile as a function of L is complex, and we identify its peaks in reference to the radial regions \( P_0 (L = 3.1–4.0, \) inner edge of the outer belt), \( P_1 (4.1–7.5, \) main outer belt), and \( P_2 (>7.5, \) quasi-trapped population), whose boundaries are determined from a radial correlative analysis (Vassiliadis et al., 2003b). Using the profiles, we classify the IP/MS parameters in four categories: (1) For the solar wind velocity and pressure the correlation is high and largely independent of L across \( P_0 \) and \( P_1 \), reaching its maximum in \( L = 4.8–6.1, \) or the central part of \( P_1. \) (2) The IMF \( B_{South} \) component and related IP/MS parameters have a bimodal correlation function, with peaks in region \( P_0 (L = 3.0–4.1, \) and the geosynchronous orbit region within \( P_1. \) (3) The IMF \( B_{North} \) and four other interplanetary or solar irradiance parameters have a minimum correlation in \( P_1, \) while the highest correlation is in the slot–outer belt boundary (\( L = 2.5. \)) (4) Finally, the solar wind density has a unique correlation profile, which is anticorrelated with that of the solar wind velocity for certain L shells. We verify this classification using more complex filtering methods as well as standard correlation analysis. The categories correspond to four types of solar-terrestrial interactions, namely, viscous interaction, magnetic reconnection, effects of ionospheric heating, and effects of high solar wind density. The response to these interactions produces the observed inner magnetospheric coherence. In each category the L dependence of the correlation profile helps explain why geoeffective solar wind structures are followed by electron acceleration in some L ranges but not in others.

Ma, 2000], and others. In expressing the electron scattering and diffusion in terms of pitch angle, energy, and L shell, various authors have found it useful to include the interplanetary or magnetospheric activity using one or several IP/MS parameters. To complicate the issue, a particular IP/MS parameter is generally important in the description of more than one process, that is, has a nonunique role. Under these conditions, identification of input parameters can be challenging.

[4] Nevertheless, certain variables have been clearly associated with changes in the electron flux content of the belts in several case and statistical studies [e.g., Blake et al., 1997; Fung and Tan, 1998; O’Brien et al., 2001]. Most studies tend to focus on flux variations at the geosynchronous orbit, so the radial extent of the geoeffectiveness is usually unknown. Instead, this paper examines time variations of the flux in the entire inner magnetosphere (L < 10).

Among IP parameters, solar wind velocity is the most important geoeffective one in determining flux variations. High-speed solar wind streams are followed by enhanced electron fluxes, especially in the descending phase of the solar cycle [e.g., Paulikas and Blake, 1979; Baker et al., 1990; Fung and Tan, 1998]. Fast-moving magnetic clouds and ejecta produce rapid response acceleration, especially before and during solar maximum [Baker et al., 1998; Reeves et al., 1998]. More recently, Li et al. [2001] have developed a diffusion model parametrized by the solar wind velocity and its fluctuations, to represent time variations of the geosynchronous flux. Solar wind density $\rho_{SW}$ and ram pressure $p_{Ram}V_{SW}^2$ are also geoeffective as surmised from the flux growth following the passage of interplanetary shocks [Li et al., 1993].

[5] IMF variations are also geoeffective, in particular those of the $B_z$ component. Magnetic reconnection on the dayside magnetopause drives the magnetospheric convection cycle. Moreover, the rate of reconnection depends on the interplanetary electric field (IEF) vector $V_{SW} \times B$, and therefore on both the reconnecting IMF as well as the solar wind speed. Prolonged increases in the IEF dusk-to-dawn component result in storms and substorms, which produce a “seed” population of 10–300 keV electrons.

[6] There are several MS parameters that represent either the level of global magnetospheric activity, or specific current systems, and their respective effects on the electron flux. Most of these parameters are geomagnetic indices or measures of geomagnetic activity. First, the flux at geosynchronous orbit is closely related to the previous global geomagnetic activity as parametrized by the sum of the $Kp$ magnetic index [Nagai, 1988]. In fact, radial diffusion in the outer belt, including the geosynchronous orbit is generally assumed to be a function of magnetospheric activity and has often been parametrized by the planetary geomagnetic index $Kp$. An example is a radial diffusion model computed from CRRES measurements of the flux [Brautigam and Albert, 2000]. The same index has been used for parametrizing multimode diffusion, notably in the Salammbro model [Boscher et al., 1996]. Second, the $Dst$ index is a conventional proxy for the ground geomagnetic perturbation because of the ring current (although the index is also affected by several other magnetospheric and ionospheric currents). Intensification and displacement of the ring current may adiabatically displace and/or scatter the local trapped electron population producing the “Dst effect” on the flux time histories [Green and Kivelson, 2001]. Comparisons of peak $Dst$ to the storm-time flux shows, however, that storms are not consistently associated with electron acceleration [Reeves et al., 2003]. The interpretation of this finding is that a delicate balance between acceleration and loss processes determines the flux level. Third, low-frequency MHD wave amplitudes have been associated with subsequent electron flux increases [Rostoker et al., 1998; O’Brien et al., 2003]. Wave-particle interactions can boost the energy of seed electrons to relativistic levels, or decrease it by orders of magnitude, via different mechanisms [Hudson et al., 1999; Elkingston et al., 2003; Liu et al., 1999; Tan et al., 2004]. Fourth, in other cases, strong dipolarizations of stormtime substorms can inject and simultaneously accelerate electrons, and may directly contribute to the radiation belt population [Ingraham et al., 2001]. In this case, however, substorm dipolarizations are measured as changes in the local field rather than by an index. Finally, interaction with the plasmasphere and high-latitude ionosphere typically results in electron scattering and loss [e.g., Abel and Thorne, 1998]. The plasmaspheric dynamics is closely correlated with auroral and inner magnetospheric geomagnetic activity [O’Brien and Moldwin, 2003]. The ionospheric activity is often represented by auroral electrojet and other high-latitude geomagnetic indices. Increases in ionospheric conductance lead to heating and expansion of the $F$ region thereby increasing the scattering and precipitation of energetic electrons.

[7] While it is important to integrate the above processes into a comprehensive framework [Friedel et al., 2002], the effects produced by individual IP/MS parameters can be understood by answering the questions of the first paragraph. In addition, this effort is expected to produce an improved specification of radiation belt models [Fung, 1996]. Traditional empirical models of the belts are static, representing the flux as a function of L shell, energy, sometimes pitch angle, and climatological/seasonal parameters [Vampola et al., 1982; Vette, 1991]. However, forecasting and nowcasting require dynamic models, which can ingest real-time measurements of the controlling IP/MS parameters [Fung, 2004]. Some of the parameters mentioned above have been used as inputs to dynamic models of the flux at the geosynchronous region [e.g., Nagai, 1988; Baker et al., 1990; Koons and Gorney, 1991; Itsukai et al., 1999].

[8] Below we analyze the electron flux at a key energy channel (2–6 MeV) in conjunction with a comprehensive data set of IP/MS measurements. The study generalizes the modeling of the interaction with the solar wind velocity, $V_{SW}$ [Vassiliadis et al., 2002]. We next discuss the measurements and the spatial range of the analysis.

1.1. Flux Measurements

[9] The polar-orbiting Solar, Anomalous, and Magnetospheric Explorer (SAMPEX) [Baker et al., 1993] provides a synoptic coverage of the radiation belts in time and McIlwain L shell. The directional flux $J_d(L;E = \text{const})$ at 2–6 MeV is measured by the Proton-Electron Telescope (PET) in the interval 1993–2000 [Cook et al., 1993]. While SAMPEX is in low Earth orbit with an apogee of 600 km, the measured
flux corresponds closely to flux at higher altitudes up to the equatorial plane [Kanekal et al., 1999, 2001]. In this work we use daily averages of the log flux $j_{\|}(t;L)$ in the range $L = 1-10$. A discussion of the advantages and limitations of the SAMPEX/PET data set, relevant to its analysis below, is found in [Vassiliadis et al., 2002].

### 1.2. Radial Structure of the Outer Belt Flux

[10] In the analysis of the next sections, we parametrize the effective interaction of the flux at a fixed L shell with the IP/MS parameters. From previous research we know that the outer belt is divided in three ranges of L, each one with a distinct flux variation [Vassiliadis et al., 2003b] and response to solar wind velocity [Vassiliadis et al., 2002, 2003a]. The results below demonstrate that the radial regions are useful in organizing and interpreting the spatial features of the flux interaction with most IP/MS variables. Therefore we briefly review them here:

1. P0: this innermost ($L = 3.1-4.0$), high-flux region partly overlaps with the outer plasmasphere. While it shares certain dynamical similarities with the main region (P1), its response is much more rapid (<1 day). The region is characterized by a high acceleration, which often occurs during the passage of magnetic clouds and more generally interplanetary coronal mass ejections (ICMEs) such as the January 1997 cloud [Baker et al., 1998; Reeves et al., 1998], and other solar wind ejecta.

2. P1: the main region ($L = 4.0-7.5$) of the outer belt has by far the largest trapping capacity among the P1 regions. Its response peaks at 2–3 days following an interplanetary disturbance and lasts for several days. High-speed streams are particularly geoeffective in P1, especially during the descending phase of the solar cycle. The response of the flux to $V_{SW}$ in P1 [Vassiliadis et al., 2002] is qualitatively identical to that of the geosynchronous orbit [Baker et al., 1990]. As will be seen below, however, the response to an arbitrary IP/MS variable generally varies as a function of L within P1.

3. P2: the outermost region ($L > 7.5$) overlaps with the inner plasma sheet and contains marginally trapped, low-amplitude fluxes. Its response is fast (1 day), but is anticorrelated with that of P0 and P1 [Vassiliadis et al., 2003a].

1.3. IP/MS Parameters

[16] In this paper we will estimate the geoeffectiveness of the following 17 variables:

- Solar wind plasma density $\rho_{SW}$, velocity $V_{SW}$, and ram pressure $P_{SW} = \rho_{SW}V_{SW}^2$. They represent the hydrodynamic coupling between solar wind and magnetosphere.
- IMF components $B_{South}$ and $B_{North}$, IEF components $VB_{South}$ and $VB_{North}$, and the solar wind quasi invariant (QI; see definition below). The two rectified components of IMF $B_z$, $B_{South}$ and $B_{North}$, represent the transfer of magnetic flux to the magnetosphere during reconnection on the dayside and poleward of the cusp, respectively. The corresponding IEF components include the rate at which the magnetic flux is convected to the magnetopause.
- The quasi invariant is a function of the magnetic Mach number $M_A$ [Osherovich et al., 1999]

$$QI \equiv (B^2/8\pi)/(P_{SW}/2) = M_A^{-2}$$

and is used as an index of solar wind magnetic pressure. There is a high correlation between QI and the sunspot number (SSN) for timescales of a month or longer, indicating the magnetic flux emission from active regions. At shorter timescales (2–3 days), QI has been proposed as a geoeffectiveness parameter for magnetic clouds [Osherovich et al., 2003].

Geomagnetic indices $Kp$, $Ap$, $C9$, $Dst$, $AL$, $AU$, and $PC$ represent the ground geomagnetic activity produced by global and regional, magnetospheric and ionospheric current systems. That activity is sampled by a number of magnetometers from whose measurements the indices are compiled. The currents intensify following the transfer of momentum and energy from the interplanetary medium to the magnetosphere, for example, via compression or magnetic reconnection.

[21] We also examine the solar parameters $F10.7$ and SSN: The $F10.7$ radio flux is strongly correlated with solar UV irradiance. The emitted radio flux is highly correlated with the sunspot number [Jursa, 1985], and both are conventional proxies of solar cycle phase.

Interplanetary and heliospheric drivers (a), (b), and (d), and magnetospheric parameters (c) represent the solar, interplanetary, and magnetospheric states [Fung, 1996].

### 2. Analysis

[21] Electron acceleration in a realistic plasma environment under time-dependent conditions is a highly complex process, and this is reflected in the relation between the “output” flux $j_{\|}(t;L)$ to an IP/MS “input” parameter $I(t)$. To measure the effectiveness of IP/MS parameters, we make the following approximations: only one input is modeled at a time while the effect of other variables is
The impulse response function \( H^{(0)}(\tau;L) \) depends on the lag time \( \tau \) and the L shell and its peaks indicate the salient timescales of the coupling. We set the memory time \( T \) to 20 days so it is comparable to the solar rotation period, and use a small number of negative lags \( T_n \) for numerical stability. After solving equation (1) for \( H^{(0)}(\tau;L) \), we can use it to predict the flux during the same or a different interval. The predicted log flux is denoted as \( j_e(\tau;L) \).

\[
j_e(\tau;L) = \int_{-T_0}^{T_0} H^{(0)}(\tau;L)I(t - \tau)d\tau. \tag{1}
\]

The impulse response function \( H^{(0)}(\tau;L) \) is represented by the finite impulse response (FIR) filter. This is a general linear input-output relation \( \mathbf{y} = \mathbf{A} \mathbf{x} + \mathbf{d} \), where \( \mathbf{y} \) is the output, \( \mathbf{x} \) is the input, \( \mathbf{A} \) is the system matrix, and \( \mathbf{d} \) is the disturbance. In this paper we evaluate \( H^{(0)}(\tau;L) \) for the same interval from which \( H^{(0)}(\tau;L) \) was calculated. A correlation value of \( C(L) = c \) means that \( j_e(\tau;L) \) accounts for \( c^2 \) of the log flux variance [e.g., Bevington and Robinson, 1992]. The quantity \( c^2 \) is approximately equal to the traditional measure of “prediction efficiency.”

3. Results

3.1. Impulse Response Functions

[25] We measure the goodness of fit of model (1) using the correlation coefficient between observed and model flux, or data-model correlation:

\[
C_{j_e,j_e}(L) = \frac{1}{T_s \sigma_{j_e} \sigma_{j_e}} \int_0^{T_s} (\langle j_e(\tau;L) \rangle - \langle j_e(\tau;L) \rangle) (j_e(\tau;L) - \langle j_e(\tau;L) \rangle) d\tau,
\]

where \( T_s \) is the length of the training set, and \( \sigma_{j_e} \) is the standard deviation of observed log flux \( j_e(\tau;L) \). Clearly, the correlation depends on \( I(t) \) through \( j_e(\tau;L) \). In this paper we evaluate \( C_{j_e,j_e}(L) \) for the same interval from which \( H^{(0)}(\tau;L) \) was calculated. A correlation value of \( C(L) = c \) means that \( j_e(\tau;L) \) accounts for \( c^2 \) of the log flux variance [e.g., Bevington and Robinson, 1992]. The quantity \( c^2 \) is approximately equal to the traditional measure of “prediction efficiency.”

3.2. Solar Wind Velocity

[26] We first review the impulse response function \( H^{(0)}(\tau;L) \) of the log flux to the solar wind speed [Vassiliadis et al., 2002], and use it as a reference to compare the responses \( H^{(0)}(\tau;L) \) for other IP/MS parameters. For most parameters, \( H^{(0)}(\tau;L) \) has well defined peaks, showing that the flux response strongly depends on L shell, and indicating that it may be due to acceleration types with very different timescales. In section 3.2 the responses are classified in four categories on the basis of their predictive capability, so here we discuss the responses to four representative parameters: the solar wind ram pressure \( P_{SW} \) and density \( \rho_{SW} \), the geomagnetic index \( Kp \), and the F10.7 flux. Following the description of a response function, we give a brief interpretation.

3.1.1. Solar Wind Velocity

[27] Figure 1 shows the response \( H^{(0)}(\tau;L) \) composed from fixed-L impulse response functions. The lag time \( \tau \) is measured from the solar wind arrival at the magnetopause. Horizontal lines divide the L range in the three regions \( P_1 \) of the outer belt, the slot, and the inner belt as estimated from the radial correlation analysis [Vassiliadis et al., 2003b]. The response has three extrema, \( P_0 \), \( P_1 \), and \( V_F \).

[28] In the \( P_1 \) region, the \( H^{(0)}(\tau;L) \) response is qualitatively similar over a wide range of radial distances, \( 3.9 < L < 8.7 \). For these L shells, the response peaks simultaneously at \( \tau = 2 \) days, demonstrating the coherence of this radial region. The coherence is an indication that similar coupling processes operate in this range, including the geosynchronous orbit. Thus \( \tau = 2 \) is the average time it takes for \( V_{SW} \)-driven acceleration processes to boost low-energy electrons to MeV energies [Vassiliadis et al., 2002]. The \( P_1 \) peak is best seen during solar minimum conditions following solar wind high-speed streams, and corresponds to the 2–3 day response to \( V_{SW} \) at the geosynchronous orbit [Baker et al., 1990]. For comparison, we note that the ULF wave peak response occurs at 1 day [Rostoker et al., 1998]. Region \( P_2 \) is not distinct from \( P_1 \), but instead shares several
Following the \( P_1 \) peak, the flux decays with a timescale \( \Delta \sigma / C_{17} \). We approximate the timescale as the time at which the response reaches a given amplitude (contour level, cf. Figure 1). Over a significant range \( \Delta L \), the decay timescale varies linearly as follows:

\[
\Delta L / \Delta \tau \equiv v = -1.12 \pm 0.04 \text{ L/days.}
\]
Thus postpeak flux decays more slowly at lower L shells. The timescale dependence on L is consistent with Earthward radial diffusion and gives an estimate of the diffusion speed.

At low L shells, $3 < L < 4$, $H^{(\text{VSW})}(\tau; L)$ has a second peak, $P_0$, at $\tau = 0$ days. Since the time resolution is 1 day, the peak represents a rapid (several hour) flux increase. Later on, at $\tau = 1–2$ days, the peak decays rapidly at $L = 3.5$ while the flux increases at higher L shells.

The $P_0$ peak is consistent with the rapid acceleration, typically below $L = 4$, which occurs during ICME passages such as the January 1997 magnetic cloud [Baker et al., 1998; Reeves et al., 1998]. Subsequently, at $\tau = 1–2$ days, the flux increases at higher L shells, $L \sim 4$, indicating an outward diffusion toward the region of long-term trapping, $P_1$.

In the same radial range as $P_1$, a negative peak $V_1$ at $\tau < 0$ represents a flux decrease. Since the peak occurs at negative lags it indicates precursor activity, that is, flux changes that occur prior to the arrival of the main solar wind velocity pulse at $\tau = 0$. These changes are produced by earlier interplanetary disturbances. Clearly, precursor information is useful in extending the prediction horizon. Studies on individual events show that the $V_1$ peak appears as far as $\tau = \sim 5$ days prior to the arrival of geoeffective velocity, so it characterizes much earlier changes in the flux at those L shells. Note that the $H^{(\text{VSW})}(\tau; L)$ amplitude at $V_1$ is smaller than at $P_1$, and therefore the flux decrease is temporary so $P_1$ is a region of net flux growth.

We interpret the negative peak $V_1$ as the response due to the ring current which intensifies near $L = 5.5$ during storms [Vassiliadis et al., 2002]. During storm main phase, the disturbed field significantly modifies the electron drift paths and displaces them away from the ring current location. This $Dst$ effect is mostly adiabatic, but can be accompanied by scattering and precipitation [Kim and Chan, 1997; Green and Kivelson, 2001]. The electron depletion is clearly seen in the $H^{(\text{VSW})}(\tau; L)$ function as a negative flux response.

The impulse response to $V_{\text{SW}}$ leads to two main conclusions: First, the broad peaks of $H^{(\text{VSW})}(\tau; L)$ support the notion of the inner magnetospheric coherence [Kanekal et al., 1999]. The simultaneity of the flux response is the signature of processes that respond rapidly and over long radial ranges compared to diffusive timescales and the dimensions of the trapping region.

Second, the response of the flux near the geosynchronous orbit to $V_{\text{SW}}$ is qualitatively similar to the response over a much wider region, namely $P_1$. Thus, at times when the $V_{\text{SW}}$ is dominant, the flux variations at geosynchronous orbit can be scaled to represent the variations in $P_1$. We will see that this is not the case for the response of the geosynchronous flux to different inputs.

Next we compare $H^{(\text{VSW})}(\tau; L)$ to the responses of four representative IP/MS parameters.

### 3.1.2. Solar Wind Ram Pressure

The response $H^{(\text{VSW})}(\tau; L)$ to the ram pressure is shown in Figure 2a. As with $V_{\text{SW}}$, the lag time $\tau$ is measured from the solar wind arrival at the magnetopause. Horizontal lines divide the L range in (from top to bottom) the three regions $P_1$, of the outer belt, the slot, and the inner belt [Vassiliadis et al., 2003b]. The response has three peaks, $P_0$, $P_1$, and $V_1$, similar to those in $H^{(\text{VSW})}(\tau; L)$, but displaced to later lag times:

In the $P_1$ region, the $H^{(\text{VSW})}(\tau; L)$ response is qualitatively independent of $L$ over a wide range of radial distances, $3.6 < L < 8$. The responses peak simultaneously at $\tau = 4$ days, showing the coherence of the region. The time for pressure-driven acceleration processes to boost electrons to MeV energies is longer than the 2–3 day growth time following a $V_{\text{SW}}$ impulse. We also note that the flux response at geosynchronous orbit is representative of the flux response at all L shells in $P_1$.

The flux decay following the $P_1$ peak depends on L as

$$v = -0.59 \pm 0.05 L/\text{days},$$

significantly slower than the scaling for $H^{(\text{VSW})}(\tau; L)$. Responses to pressure decay on average more slowly compared to responses to velocity. Again, the scaling (4) gives a speed consistent with Earthward diffusion from the outer part of the trapping region, or the $P_1$–$P_2$ boundary. Region $P_2$ is not distinct from $P_1$ as far as the response to $P_{\text{SW}}$ is concerned.

At $3 < L < 4$, the second peak, $P_0$, occurs at $\tau = 1$ day, about 1 day later than the response to $V_{\text{SW}}$ in that region. As discussed above, peaks $P_0$ and $P_1$ are representative of two different modes of flux increases, following high-speed streams and ICMEs, respectively.

The negative peak $V_1$ indicates a flux decrease, but its amplitude is smaller than that of $P_1$, and therefore the decrease is temporary, and the flux has a net growth. The $V_1$ peak is qualitatively similar to the corresponding minimum for $H^{(\text{VSW})}(\tau; L)$ with two differences: the duration of the minimum is much shorter for the pressure: 1–2 days as opposed to $>3$ days for the velocity. Second, and more important, the minimum is at $\tau = 1$ day, whereas the corresponding minimum for $V_{\text{SW}}$ is at negative time lags.

The interpretation for this minimum is again in terms of a quasi-adiabatic response to a ring current intensification (and resulting flux decrease). However, the intensification simply follows, and is probably due to the compression pulse. In the case of $V_{\text{SW}}$ the intensification precedes the increase in the velocity and therefore represents the effect of earlier interplanetary activity, which may be similar to preconditioning of the radiation belts. An increase in $P_{\text{SW}}$, on the other hand, contributes to the flux decrease, presumably through by compressing the inner magnetosphere and intensifying the ring current, but only for a brief time and following the arrival of the solar wind at the magnetopause.

Overall, the impulse response to $P_{\text{SW}}$ and $V_{\text{SW}}$ are similar. They demonstrate a coherent, cross-L response that can result in magnetospheric coherence. Flux observations from the geosynchronous orbit can be used to characterize the entire $P_1$ region.

#### 3.1.3. Solar Wind Density

The response to the density is shown in Figure 2b. The $P_1$ peak has disappeared which is consistent with its interpretation as the response to high-speed streams. Density and its fluctuations are low during streams. Negative peak $V_1$ occurs at a similar, positive lag time and over similar L range as the pressure. However, its duration can be as high as 4 days at high L shells, compared to 1–2 days for...
the corresponding minima for $V_{SW}$ and $P_{SW}$. A small positive peak, $P_3$, occurs in $L = 4–5.5$ at $\tau = -1$ day, but its effect is small compared to the decrease during $V_1$. Since $P_1$ is weak and $V_1$ is intense and of long duration, the net geoeffectiveness of the solar wind density in region $P_1$ is expected lower than that of the other two parameters in $L = 3.6–10$.

[45] The $P_0$ peak at $\tau = 1–2$ days is the dominant peak in this response. Interestingly, the beginning of $P_0$ coincides in time with the lowest response in $V_1$. At later times, $\tau > 2$ days, the flux intensifies at higher $L$ shells indicating antearthward transport. The corresponding timescale is

$$v = 0.28 \pm 0.10L/\text{days},$$

which is consistent with slow radial diffusion.

[46] Thus increases in solar wind density are linked to weak or no flux increases in $P_1$ and $P_2$, but some increase in $P_0$. It is interesting to consider how the density contributes to the geoeffectiveness of the pressure. A comparison of the responses to $V_{SW}$ and $P_{SW}$ shows that they are virtually identical, indicating that the geoeffectiveness of $P_{SW}$ relative to $V_{SW}$ will be determined by the geoeffectiveness of the density. The geoeffectiveness of these parameters will be discussed in section 3.1.4.

3.1.4. $Kp$ Index

[47] The impulse response $H^{(K_p)}(\tau;L)$ (Figure 2c) corresponds more closely to that of $V_{SW}$ than the two other responses do. It is representative of the response of the flux to several geomagnetic indices. The flux in peak, $P_1$, responds simultaneously over the range $L = 3.6–8.4$ and reaches the peak at $\tau = 2.5–3$ days, similar to $V_{SW}$.

[48] Following the peak, the flux decays with a timescale which scales

$$v = -0.83 \pm 0.055L/\text{days}.$$  

Thus flux variations associated with changes in the major geomagnetic currents decay on average about as fast as pressure-induced variations. At geosynchronous orbit, $H^{(K_p)}(\tau;L = 6.6)$ peaks at $\tau = 2–3$ days after solar wind arrival, as expected [Nagai, 1988; Baker et al., 1990]. The geosynchronous orbit response is characteristic for the entire $P_1$ region.

[49] The other two peaks, $P_0$ and $V_1$, are comparable to those of the $V_{SW}$ response. The $P_0$ peak at $L = 3–4$ occurs at $\tau = 1$ day and its amplitude is much lower than that of $P_1$. The $V_1$ peak is a manifestation of the $Dst$ effect.

3.1.5. F10.7 Flux

[50] The impulse response to F10.7 is different from all responses discussed so far (Figure 2d). Its amplitude is more rapidly varying in $L$ shell and lag time than others, and close to zero. To reduce the amount of numerical noise, we smooth $H^{(F10.7)}(\tau;L)$ with a 2-day running average window.

[51] For a range of $L > 6$, the response is close to zero or negative; a negative $H^{(F10.7)}(\tau;L)$ means that an increase in F10.7 is followed by a decrease in the electron flux. Furthermore, there are no peaks corresponding to $P_1$ or $V_1$. Their absence indicates that processes modulated by the F10.7 flux are unrelated to the $Dst$ effect and the accelerating processes in that region. Third, in the range $L = 2.2–5.8$ the response is uniformly positive. The lower part, $L = 2.2–3.0$, is part of the slot region, and therefore a positive peak in that region will produce only a temporary increase in the electron flux. However, the F10.7 level may modulate the response in region $P_0$ and the lower part of $P_1$. Its effect on the flux amplitude is expected to be small (as will be checked with a correlation analysis below), and have a clear seasonal variation.

[52] Thus the interpretation is that $H^{(F10.7)}(\tau;L)$ is not directly associated with any of the known acceleration processes. Increases due to F10.7 are not related to interplanetary activity commonly associated with storms, such as high-speed streams or magnetic clouds/ICMEs, although F10.7 may weakly modulate their geoeffectiveness. In the next section, this conclusion will be generalized after a comparison of F10.7 with other inputs.

3.1.6. Summary

[53] Impulse responses $H^{(F10)}(\tau;L)$ measure the relative timing and radial location of the coupling for the five IP/MS parameters we have examined. In fact the five responses are representative of the remaining 12 parameters (not shown in detail), and suggest that these variables are involved in the similar types of acceleration and transport processes. Five variables, including the F10.7 discussed above, have particularly noisy impulse responses and may play a minor role in determining the flux variation. Responses may vary significantly with $L$ shell, and their peaks indicate altitude ranges where acceleration and transport occur at a global scale.

[54] Differences in timing between responses to two different variables mean that the coupling occurs at different times in the passage of a solar wind structure (high-speed stream, CME, or shock). Therefore each part of such a structure has a specific geoeffectiveness. As an example, we consider the interaction of a high-speed stream with the magnetosphere: several days before the peak flux, the high plasma flow speed produces an average increase in the reconnection rate, which intensifies various current systems (increase in $Kp$). The same increase in dayside reconnection leads to magnetospheric plasma convection and intensification of the ring current, producing the $Dst$ effect ($V_1$ peak). At the same time reconnection transports seed electrons into the radiation belts which are accelerated by ULF waves in the region $L = 5–7$ ($P_1$ region). The waves are excited by compressive or shear effects of the high-speed structure on the magnetosphere flanks. If the solar wind structure is a magnetic cloud on the other hand, a different type of flux increase occurs (much faster, $\tau < 1$ day, and at lower $L$ shells).

[55] Differences in response amplitude for different IP/MS parameters, on the other hand, are not directly comparable since the responses are measured in different units. We resolve this issue next, by comparing the data-model correlations rather than the responses themselves.

3.2. Degree of Geoeffectiveness and Identification of Interaction Types

3.2.1. Measuring Geoeffectiveness

[56] A model’s accuracy, in, for example, reproducing the amplitude and timing of electron flux variations, is a measure of its capability to represent the effects of the solar wind-magnetosphere interaction on the flux. A standard
A metric to measure predictive capability is the data-model correlation \((2)\). We will use the correlation to measure the geoeffectiveness of input \(I(t)\). To see that, consider a fixed input sequence \(I(t)\) which is used to drive models of different degrees of sophistication. As models become more accurate, the correlation increases until it eventually approaches a limiting value (generally less than 1.00, ...}

**Figure 3.** Data-model correlation \(C(L)\) from equation (2) for 17 interplanetary/magnetospheric variables. (top) Solar wind plasma parameters \(V_{SW}\) (red thick line), \(P_{SW}\) (red thin line), and \(P_{SW}\) (orange dashed line) represent the hydrodynamic aspects of the solar wind–magnetosphere interaction. (middle) Interplanetary inputs \(B_{South}\) (blue thick line) and \(V_{SW}B_{South}\) (blue thin line) and magnetic indices \(Kp, Ap, C9, Dst, AL, AU, PC\) (light blue lines) represent the effects of magnetic reconnection. \(AL\) and \(AU\) are denoted by dashed lines. (bottom) Interplanetary inputs \(B_{North}\) (green thick line), \(V_{SW}B_{North}\) (green thin line), and solar wind quasi invariant QI as well as solar variables F10.7 and SSN (cyan lines) show the effect of electron losses.
depending on the interval). This asymptotic accuracy level represents the actual geoeffectiveness of input I(t). It is related to the maximum variance of the flux that can be explained by I(t). Thus a data-model correlation of a specific model, such as equation (1), is a lower bound of the actual geoeffectiveness. Nevertheless, it may be indicative of the correlation achieved with more accurate models.

The \( C^{(d)}(L) \) correlation for all IP/MS parameters is shown in Figure 2 as a function of L. We will classify the parameters in four categories on the basis of the L profiles. After describing each category, we discuss its physical interpretation.

### 3.2.3. Solar Wind Density

The data-model correlations for \( V_{SW} \) and \( P_{SW} \) are shown in Figure 3 (thick and thin red lines, respectively). For \( V_{SW} \), the maximum correlation is moderately high at 52%. For the long correlation interval used here (1993–2000), there is little variation in L over \( P_0 \) and \( P_1 \). For shorter correlation intervals (e.g., of the order of a year), however, the correlation generally increases, and two peaks appear close to the centers of \( P_0 \) and \( P_1 \) [Vassiliadis et al., 2002]. Beyond the geosynchronous orbit, the correlation steadily declines with L eventually reaching zero at L = 10. At the other end, the correlation is negative in the slot while in the inner belt (L < 2.0) the correlation is positive in the narrow range L = 1.3–1.6.

The pressure-driven model has a similar correlation as that for the solar wind velocity, but lower in value (with a maximum about 40%). Models driven by \( V_{SW} \) and \( V_{SW} \) give virtually identical data-model correlations, so the difference between \( C^{(v_{SW})}(L) \) and \( C^{(P_{SW})}(L) \) is due to the variation of the density term. Here, too, the correlation is essentially independent of L over \( P_0 \) and \( P_1 \).

The interpretation of \( C^{(v_{SW})}(L) \) and \( C^{(P_{SW})}(L) \) being so similar is that in the long-term the coupling of the flux to the velocity and pressure is fairly similar, and dominated by the response to the velocity variations. The moderately high correlation in \( P_0 \) and \( P_1 \) shows that the solar wind is geoeffective in those regions, although it drives or modulates very different types of mechanisms in each region. Finally, the negative correlation in the slot indicates that an increase in solar wind speed is followed by a decrease in that region’s flux.

### 3.2.4. IMF \( B_{South} \) Component and Geomagnetic Indices

The middle plot of Figure 3 shows the data-model correlations for \( B_{South} \), \( V_{SW}B_{South} \), and seven magnetic indices. The thick blue line denotes \( C^{(B_{South})}(L) \) which has two well-defined peaks, one in \( P_0 \) (L = 3–4) and one at the geosynchronous region. The first peak is centered at L = 3.3 and extends to the slot and \( P_1 \). The peak correlation is 36% so \( B_{South} \) accounts for 13% of the flux variance at that L shell. The second peak occurs exactly at geosynchronous orbit. The correlation profile due to IEF \( V_{SW}B_{South} \) (thin blue line) is higher, but essentially identical to that of \( B_{South} \), since the solar wind velocity varies slowly compared to the IMF.

On the other hand, the geomagnetic indices also have similar \( C^{(d)}(L) \) profiles to \( C^{(B_{South})}(L) \), but correlation values are significantly higher. For instance, the peak correlation at \( P_0 \) is in the range 60–75% for a number of indices, so that an index can explain 36–56% of the variance in the L shells of \( P_0 \). For the geosynchronous orbit peak, the corresponding correlation is 45–52%. The highest correlation among all indices comes from the \( Kp \)-driven model, confirming the usefulness of that index for the specification of the outer belt state [Nagai, 1988].

The \( AL \) and \( AU \) indices (dashed lines) are proxies for the disturbance amplitude due to the aural electrojets, which are linked to convection as well as substorm activity. Note that the \( AL \) and \( AU \) records are not available for the entire period 1993–2000 therefore only 2 years (1993–1994) are used instead. Generally, shorter interval lengths result in higher correlations, and therefore the \( C^{(AU,AL)}(L) \) profiles are higher than they would be if data from all years were available. When we compare all IP/MS parameters in the 1993–1994 interval, however, correlations for \( AL \) and \( AU \) are still the highest in region \( P_0 \) and the lowest (most negative) in S, but lower than all indices in \( P_1 \) and \( P_2 \).

The interpretation for the middle plot of Figure 3 is that the similarity of the interplanetary and magnetic index
profiles is reasonable, since increases in $B_{South}$ are followed by reconnection, energy/momentum transfer, and intensification of various magnetospheric/ionospheric current systems. The $B_{South}$ component rate represents the magnetic flux available for equatorial dayside magnetic reconnection, while $V_{SW}B_{South}$ controls the rate at which this magnetic flux is made available. Reconnection transfers momentum and energy to the inner magnetosphere in a variety of pathways. A prolonged $B_{South}$ component results in storms and/or substorms, which radially transport and accelerate electrons from the plasma sheet (including the $P_2$ region) to the keV energy seed population. In a second pathway, stormtime substorms produce direct acceleration [Ingraham et al., 2001]. The peak at $L = 6.6$ is indicative of these two pathways in which energy is imparted to the trapped-flux region of $P_1$. After flux at geosynchronous has increased, it is expected to diffuse radially as well as in energy. However, the correlation at $L$ above and below geosynchronous is zero, therefore $B_{South}$ is not controlling the radial diffusion processes in the rest of $P_1$. This is an example where flux variations at geosynchronous are not directly related to flux variations at lower $L$ shells.

A third pathway in which $B_{South}$ transfers energy directly to the inner magnetosphere may be related to the response in region $P_0$, which is excited by magnetic clouds and other ICMEs [Baker et al., 1998; Reeves et al., 1998]. The peak of $C(B_{South}) (L)$ at $P_0$ is possibly a signature of local acceleration rather than transport. The relative significance of acceleration versus transport must be determined through analysis with multiple-energy flux measurements. In any case, the correlation in $P_0$ is higher than at geosynchronous, therefore, according to the linear model (1), $B_{South}$ is more geoeffective per unit $L$ shell in $P_0$ than at the geosynchronous region.

In summary, the $L$ profiles are qualitatively different from those of the hydrodynamic parameters.

3.2.5. IMF $B_{North}$ Component and Solar UV Flux

The final group of IP/MS parameters includes $F10.7$, SSN, QI, $B_{North}$, and $V_{BNorth}$ (Figure 3). The thick line in the bottom plot of Figure 3 denotes $C(B_{North}) (L)$ for the IMF $B_{North}$-driven model and has three extrema: a minimum in region $P_0$, and two maxima at the S-$P_0$ and $P_1$–$P_2$ boundaries. In $P_1$, the minimum means that an increase in $B_{North}$ results in a decrease in the flux $j_d(t;P_1)$. The similarity of the $B_{North}$ and $V_{SW}B_{North}$ profiles (thick and thin green lines, respectively) is straightforward. The $C(j) (L)$ profiles of the other three variables (in cyan) are virtually identical to $C(B_{North}) (L)$, except that their correlation values are somewhat higher. The relation between $F10.7$ and sunspot number is well known [Jursa, 1985] and the quasi invariant QI tracks the sunspot number well both being measures of the hemispherically asymmetric solar activity [Osherovich et al., 1999].

The $H^{j}(\tau;L)$ responses to the five parameters are similar to each other only for those $L$ shells where $|C(j)(L)| > 0.20$, but vary significantly elsewhere. The similarity of the responses suggests that in the higher $|C(j)(L)|$ range, the impulse response is indicative of temporal and spatial scales, but in the $L$ range where $|C(j)(L)|$ is close to 0, the responses do not provide any information.

The interpretation of $C(B_{North}) (L)$ in $P_1$ is that increases in $B_{North}$ or $V_{SW}B_{North}$ tend to be associated with particle loss rather than acceleration in that region. Second, the maximum at $L = 3.0$ does not characterize the slot or region $P_0$, but probably the boundary between the two. The maximum correlation may correspond to the transport from the high-flux, rapid-response region $P_0$ to the slot, which acts as a particle sink because of strong scattering and rapid precipitation. Third, at $L = 6.6$, the correlation $C(B_{South})(L)$ becomes positive indicating that $B_{North}$ is geoeffective in that low-flux region. During periods of prolonged $B_{South}$, there is reconnection at high latitudes, antisunward of and close to the cusp region. The cusp has been proposed as an effective energetic electron source [e.g., Sheldon et al., 1998], and its effect is best seen during $B_{South}$ zero or positive.

High levels of $F10.7$ (or SSN) produce electron depletions as well, but over much longer timescales than variations in $B_{South}$. The UV illumination determines the ionospheric conductance and heating. Significant heating leads to expansion of the neutral atmosphere and ionosphere, and precipitation of the trapped electron population. Over long intervals the solar wind quasi invariant is an interplanetary measure of coronal hole activity and therefore sunspot number.

Thus these two very different types of parameters are associated with low acceleration efficiency or outright precipitation and loss. Northward reconnection corresponds to a significant decrease in seed electron flux while increases in ionospheric conductance are associated with increases in precipitation.

3.3. Maximizing Prediction Capability

We next consider which IP/MS parameter best specifies the flux at a given $L$ shell, an important selection for a predictive radiation belt model. On the basis of the discussion at the outset of the previous section, such a parameter is statistically the most geoeffective for that $L$ shell; that is, its variations have the greatest impact on the variations of the flux. In terms of the four interaction types, one can understand in which $L$ shell range each type of interaction has its greatest effect.

Figure 4 shows the correlation $C(j)(L)$ maximized over all IP/MS parameters and plotted as a function of $L$. The three colors represent the classification of Figure 2: hydrodynamic coupling due to the three plasma parameters (red), magnetic reconnection coupling (blue), and loss processes (green). The correlation $C(j)(L)$ of the solar wind velocity is included for reference (dashed line). In most $L$ shells, the maximum correlation is achieved by coupling to either the solar wind velocity (as in the central part of $P_1$) or to a magnetic reconnection parameter. The parameter is $Kp$ or $Ap$ for the rest of $P_1$ including the geosynchronous orbit, and $V_{BNorth}$ for $P_2$. For the slot region, clearly the highest correlation is achieved by coupling to the solar wind density. Finally, the $F10.7$-group variables are the best predictors only for inner belt ($L < 2$) fluxes.

Note that, for almost all parameters and regions, the maximum data-model correlations are low and thus the prediction efficiency is typically <50%. Normally such values would be questionable from a statistical significance point of view; the fact that they are not is because $H(\tau;L)$ and $C(L)$ are found to vary little for a wide variety of inputs.
and intervals. In any case, the correlation is used as a discriminating statistic for the IP/MS parameters, rather than for actual prediction and forecasting.

[79] The low \(|C(L)|\) values occur for two reasons: first, we use linear filters to model a nonlinear response so it is expected that they will explain a rather small percentage of the variance. We have increased the prediction efficiencies significantly (>90% for the in-sample correlation of equation (2)) by going to more sophisticated models for \(J_{\text{eq}}\) than the FIR model (1), such as autoregressive-moving average (ARMA) models. These correlation profiles \(C(L)\) for ARMA models fall in the same four categories as for FIR models discussed above, but correlation values are so high that the differences between categories are much smaller. Therefore ARMA models are not as useful as FIR models in distinguishing categories and interpreting in terms of physical processes.

[80] Second, in this study we have mostly trained the models on fairly long intervals compared to the typical length of a storm, thereby using the same model to explain very different types of activity. When we use shorter intervals, \(|C(L)|\) grows significantly (cf. discussion by Vassiliadis et al. [2002, 2004]).

4. Summary and Discussion

[81] The analysis shows that the electron flux at a given L shell is a function of the interplanetary and magnetospheric states as expressed by four types of parameters. For each category, the effects of the solar wind–magnetosphere interaction on the electron flux vary significantly with radial region: the slot S, the \(P_1\) regions (\(i = 1–3\)), and layer-like regions such as the S-P0 boundary and the geosynchronous orbit. There seem to be 4 distinct types of interaction between interplanetary inputs and the magnetosphere that have an effect on the flux at a given L. Since these interactions are fundamentally different, a dynamic radiation belt model is expected to need at least one input variable from each IP/MS category to accurately represent the time variations of the electron flux.

[82] Clearly, a key consideration in this study is the model of the interaction. This study has used linear, single-input filters, which are well known and can be computed accurately. Their limitations can be overcome by developing nonlinear, multi-input models, which is the subject of an ongoing study.

[83] Acknowledgments. We thank S. Kanekal for providing the SAMPEX/PET data set and for discussions on energetic particle measurement techniques. We thank D. Baker, R. Weigel, and J. Rigler for discussions on flux modeling, and V. Osherovich and J. Fainberg for discussions on the solar wind quasi invariant. Interplanetary and index data were provided by NSSDC, NGDC, and World Data Centers WDC-1 and WDC-2. This work was funded by NASA RTOP grant 784-50-51-02. [84] Lou-Chuang Lee thanks Reiner Friedel and Xinlin Li for their assistance in evaluating this paper.

References


Vassiliadis, D., R. S. Weigel, A. J. Klimas, S. G. Kanekal, and R. A. Mewaldt (2003a), Modes of energy transfer from the solar wind to the inner magnetosphere, Phys. Plasmas, 10(2), 463–473.


