Radiation belt electrons respond to multiple solar wind inputs

E. J. Rigler,[1] M. Wiltberger,[1] and D. N. Baker[2]

[1]   The multivariate statistical basis that underlies both single- and multi-input linear prediction filter analyses is reviewed, providing context necessary to understand the full capabilities and limitations of such models. A brief reanalysis of single-input filters is conducted primarily as a contrast to subsequent analysis of multi-input linear filters, which (1) guarantee similar or better prediction capabilities than single-input linear filters and (2) reduce bias in estimated filter coefficients that is inherent to underspecified linear models when ordinary least squares algorithms are employed. The former is clearly valuable from a practical standpoint, but the latter helps build confidence in any physical interpretations of both the filter coefficients, which often emulate stable low-order dynamical response functions quite well, as well as prediction error statistics that can be used to provide a lower bound on the fractional or percent variance of radiation belt electron flux that can be attributed to each different solar wind input. We find that the solar wind bulk speed tends to be the primary driver of electron flux enhancements at magnetic L shells larger than 4, with little or no relation to flux decreases. Changes in the solar wind's magnetic field strength tend to temporarily reduce electron fluxes between $L = 4$ and $L = 8$, while enhancing it between $L = 3$ and $L = 4$. In contrast to predictions generated by single-input linear filters, multi-input filters show that solar wind plasma density only contributes weakly to electron flux variability, although it does so consistently across nearly all L shells. Finally, we studied two distinct 4-year intervals within the most recent solar cycle and found that smaller, more time-stationary prediction errors are generated by multi-input linear filters. We therefore conclude that multi-input filters more accurately reflect real dynamic relationships than any single-input linear filter alone.

1.   Introduction

[2]   The Earth's magnetosphere is a highly dynamic physical system that derives its energy predominately from the impinging solar wind [*Arnoldy*, 1971; *Perreault and Akasofu*, 1978; *Akasofu*, 1979; *Vasyliunas et al.*, 1982]. Strong correlations between various solar wind measurements and magnetospheric parameters have led to the broad application of so-called linear prediction filters as proxies for the true nonlinear coupling between these two distinct space plasma regimes [*Iyemori et al.*, 1979; *Bargatze et al.*, 1985; *Clauer*, 1986; *Nagai*, 1988]. Electron flux variations in the radiation belts in particular have long been known to exhibit a strong correlation with solar wind speed [*Paulikas and Blake*, 1979; *Baker et al.*, 1986]. The National Oceanic & Atmospheric Administration's operational Relativistic Electron Forecast Model, following a technique first sug-gested by *Baker et al.* [1990], uses real time measurements of solar wind speed taken from the Advanced Composition Explorer spacecraft as input for a linear prediction filter that forecasts the daily influence of >2 MeV electrons at GOES orbital (i.e., geostationary) altitudes to help understand, if not mitigate, the effects of charge accumulation deep in otherwise insulated materials (deep dielectric charging).

[3]   More recently, this approach has been extended to a broader range of geomagnetic L shells by using electron flux data from the Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX) Proton-Electron Telescope (PET) instrument to generate multi-output linear prediction filters for nearly the entire radiation belt [*Vassiliadis et al.*, 2002, 2005; *Rigler et al.*, 2004, 2005]. Efforts are under way to implement these models in an operational setting [*Baker et al.*, 2004a; *Vassiliadis et al.*, 2004, 2005], but these studies have already led to an improved physical understanding of the coupling between the solar wind and radiation belt electron fluxes. Specifically, three distinct categories of solar wind-magnetosphere coupling have been proposed, each based on distinctive L-shell profiles of observation-prediction correlations. The first category includes variables related most closely to the fluid-like aspects of

---

[1]High Altitude Observatory, National Center for Atmospheric Research, Boulder, Colorado, USA.
[2]Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado, USA.

magnetohydrodynamics (MHD) like pressure and viscous interactions at the outer magnetospheric boundary. The second category corresponds to the electromagnetic aspects of MHD, in particular, enhanced convection driven by dayside magnetic reconnection. A third category includes variables that do not exhibit very direct dynamic coupling with the radiation belts but rather appear to modulate electron loss rates via enhanced thermospheric and ionospheric densities. The present paper is strongly motivated by these findings.

[4] Despite their heavy use throughout the 1980s, and more limited application to both scientific analysis and space weather forecasting recently, a thorough description of the mathematical theory and techniques that underlie nearly all linear prediction filter models has not been given in the magnetospheric physics literature. Indeed, it is difficult to find a concise summary of all such information in a single source even outside this field. This is in part due to the fact that, quickly following some of the earliest work by *Wiener* [1949], the immense utility of linear filters was recognized across a broad spectrum of technical and scientific disciplines, each of which then proceeded to evolve more or less independently up to the present day. Geophysicists in general, and magnetospheric physicists specifically, were perhaps a little late to exploit these powerful modeling techniques, but after catching on, they tended to follow in the foot steps of controls and process engineers who were more interested in the potential of linear filters to describe various dynamical systems.

[5] Other less "technical" disciplines have also benefited from the early work of Wiener. Medicine, social and behavioral sciences, political science, and especially econometrics are examples of fields that are concerned with understanding relationships between large numbers of variables [*Pindyck and Rubinfeld*, 1991], often time series, whose variations are not typically constrained by any known fundamental rules of nature. As such, they make careful and deliberate use of classical multivariate statistics and linear regression techniques to extract as much useful information as possible from the available data and make the best predictions possible in the absence of well-defined governing dynamics. Equally important, a high premium is assigned to a thorough understanding of the uncertainties inherent in data-derived models, as well as the design of model structures and data collection techniques, so as to help mitigate these potential errors. Although such techniques have evolved to the point that they are considered classical within these fields, they are far from commonplace, or even considered "best practice," within the magnetospheric research community. With large and ever-growing collections of space physics data, it is past time to remedy this situation.

[6] A true dynamical system requires a feedback mechanism that describes how the system's state evolves over time. A common approach in time series forecasting is to make the future state of the system a linear function of the current and prior states. A perturbation term may be included to allow external inputs to drive the system away from equilibrium, but the fundamental dynamics of any linear filter-based model are captured primarily by its autoregressive (AR) coefficients. The single coefficient present in a first-order AR model reproduces exponential decay, the two coefficients in a second-order model can generate damped harmonic oscillations, and so forth. Several magnetospheric studies during the 1990s employed this kind of discretized dynamical model, even extending it with local-linear variants [*Vassiliadis et al.*, 1995] and nonlinear continuous-time analogs [*Klimas et al.*, 1997].

[7] To date, however, most attempts to describe empirically the magnetosphere's dynamic response to various drivers have relied on so-called finite impulse response (FIR) filters, which do not posses an AR component, and therefore no explicit feedback mechanism capable of propagating the state forward in time. The proper interpretation of any such FIR filter model is that it describes a completely nondynamical system that only evolves because of external influences. Relatively low-order and highly stable dynamical systems may however exhibit a response function that can be approximated by a FIR filter's time-lagged coefficients. The radiation belts in particular exhibit roughly first-order dynamics at daily timescales. The filters used in both scientific studies and operational forecasting of radiation belt electrons capture this dynamical behavior but also exhibit more complex structure that cannot be attributed solely to statistical uncertainties inherent to the data used to train the models (see section 4).

[8] We must note that the idea to use multiple inputs with linear filters is not completely new in magnetospheric physics research [e.g., *McPherron et al.*, 1986; *Trattner and Rucker*, 1990], but it has yet to be applied to radiation belt studies, nor has it been attempted in the statistically rigorous manner we intend here. We posit that an empirical radiation belt modeling approach that employs the proven method of multi-input linear prediction filters, but presents the theory, process, and results in a manner that relies less on an understanding of dynamical systems theory, and more on understanding the multivariate statistics common in fields like econometrics, will lead both to a better understanding of the associated radiation belt dynamics and eventually to better forecasts. We do not proceed blindly, however, and constrain our choice of filter inputs to three well-studied solar wind parameters, the magnitude of the interplanetary magnetic field (B_{imf}), the solar wind bulk speed (V_{sw}), and the solar wind plasma density (ρ_{sw}). The first two inputs correspond to the two MHD-related categories of solar wind-magnetosphere coupling proposed in the work of *Vassiliadis et al.* [2005]. More specifically, these two scalar quantities exhibit stronger correlations with SAMPEX-observed electron fluxes than do their respective vector components when daily averages are employed. The third input (ρ_{sw}) is included to help contrast single- and multi-input linear filters, both in terms of their predictive capabilities and their ability to shed light on the underlying dynamics of the radiation belts.

2. Multivariate Linear Regression

[9] This section will provide the mathematical background necessary to make what we consider to be a more appropriate interpretation of the single- and multi-input linear filters presented in subsequent sections than the dynamical system approach described previously. Specifically, the ensemble of filter coefficients should be seen as a kind of input-normalized superposed epoch analysis [e.g., *O'Brien et al.*, 2001]; one where the epoch is defined as

each discrete step in the training data time series rather than by an ad hoc criterion imposed by the analyst. We may occasionally draw on the more detailed information presented in Appendix A, B, and C which are provided for those readers interested in a more thorough and pedagogical treatment of certain aspects of modern multivariate statistics.

[10] Let us begin with the simplest linear relationship possible between a dependent variable (Y) and an independent variable (X). These variables are usually realized as column vectors whose elements correspond to individual observations. While it is common to consider Y and X to be drawn from a given population at regular discrete times, the assumption that they are time series is not necessary to the validity of the more general statistical relationships being discussed. A time parameter "t" will therefore be dropped from this point forward unless and until it is required for clarity.

$$Y = a + bX + \varepsilon \qquad (1)$$

[11] Equation (1) is just a line in the XY plane with a slope b and a Y intercept a. ε is included for generality and describes the prediction residuals of this model. This relationship implies that Y and X vary synchronously and that the ratio of each change in Y to each change in X corresponds to the slope b. What's more, if Y and X are standardized (that is, each variable is divided by its standard deviation, creating unit-variance distributions), the slope that optimizes equation (1) in a least squares sense (which we will now call b') can be squared to provide a metric that represents the fraction of Y's variance that can be attributed to X.

[12] The linear relationship between two variables can be extended easily to include a set of multiple independent variables:

$$Y = a + b_1 X_1 + b_2 X_2 + \cdots + b_K X_K + \varepsilon \qquad (2)$$

Slopes b_k associated with the additional independent variables $X_{k=1 \to K}$ now describe the projection of a regression line onto corresponding planes in a K-dimensional space. If $\hat{Y}$ is defined as the linear prediction in the absence of errors, equation (2) may be rewritten as $Y = \hat{Y} + \varepsilon$. If all the variables are standardized to unit variance distributions, the square of the optimal coefficient for regressing Y' on $\hat{Y}'$ must fall between zero and unity and, like before, describes the fraction of Y's variance that can be attributed to a linear combination of all the different independent variables. This metric is commonly referred to as the "coefficient of multiple determination," even when applied to a linear equation using only a single independent variable. However, after defining a set $H \equiv X_{k=1 \to K}$, we will follow the example of the vast majority of statisticians and write R^2_{YH} for brevity.

[13] Interpreting individual regression coefficients is not as straightforward. Each coefficient b_k represents the ratio of a change in Y to a change in X_k when every other independent variable is held constant. Similarly, if all variables, Y and X_k, are standardized to produce unit-variance distributions, the square of each least squares optimized coefficient, b_k', represents X_k's fractional contribution to Y's variance when every other independent variable is held constant. If X_k is uncorrelated with every other

independent variable, it is a simple matter to vary this independent variable without altering the others, so $b_k'^2$ represents X_k's total contribution to the variance in Y. If all the independent variables are uncorrelated with one another, R^2_{YH} is simply equal to the summation of the squares of the standardized slopes.

[14] However, when X_k is correlated with any or all of the other independent variables, as is generally the case, such a simple summation will not equal R^2_{YH} and can in fact be either larger or smaller, depending on the signs of correlations between different independent variables. One must account for these correlations when trying to explain some fraction of Y's variance. A general relationship that relates the coefficient of determination, the standardized slopes from a multivariate linear regression, and the correlations between multiple independent variables is given in equation (3):

$$R^2_{YH} = \sum_{i=1}^{K} \sum_{j=1}^{K} b_i' b_j' r_{x_i x_j} \qquad (3)$$

Appendix A provides the derivation for these expressions, as well as for the alternative and simpler expression given in equation (4).

$$R^2_{YH} = \sum_{k=1}^{K} b_k' r_{YX_k} \qquad (4)$$

[15] As it happens, there is nothing to prevent the subscripts $k = 1 \to K$ from being modified to represent a discrete time shift in the independent variable X. This leads to the familiar convolution equation typically used to implement linear prediction filters in the time domain.

$$\hat{Y}(t) = a + b_{min} X(t - min) + b_{min+1} X(t - min - 1) + \cdots + b_{max-1} X(t - max + 1) + b_{max} X(t - max)$$

$$\text{or} \qquad (5)$$

$$\hat{Y}(t) = a + \sum_{k=min}^{max} b_k X(t - k)$$

In equation (5), "min" refers to the minimum time lag relative to the desired prediction time, "max" refers to the maximum time lag, and the indices are assumed to increase incrementally from the former to the latter. Note that here the parameter t has been reinserted for clarity, since in this instance, time is an explicit component of the model equation in the form of a relative time lag. One may also notice that there is nothing to prevent additional subscripts representing multiple inputs from being introduced, since equations (2) or (5) are purely additive.

$$\hat{Y}(t) = a + \sum_{k=min}^{max} b_{1k} X_1(t - k) + \sum_{k=min}^{max} b_{2k} X_2(t - k)$$

$$+ \cdots + \sum_{k=min}^{max} b_{ik} X_i(t - k) \qquad (6)$$

Each subscript i in equation (6) corresponds to that component of the prediction $\hat{Y}$ which is generated when only the input X_i is allowed to vary. Therefore if each

component prediction $\hat{Y}_i$ is recast as its own independent variable and standardized, the optimal $b_i'^2$s will represent $\hat{Y}_i$'s contribution to Y's total variance when every other component prediction is held constant. With no clear alternatives established in any of the statistics literature reviewed for this paper, we will refer to these coefficients describing the regression of the standardized prediction on the standardized observations as (component) prediction regression coefficients (PRCs). There is no guarantee that component predictions will be uncorrelated, even if their respective inputs are, so the squares of PRCs are not expected to add up to exactly R^2_{YH}. Therefore equations (3) and (4) will be required once again to account for cross correlations between the component predictions.

[16] Up to this point, no technique for calculating optimal slopes/regression coefficients has been discussed. Ordinary least squares (OLS) is by far the most common procedure used to determine optimal parameters that minimize the residuals of a linear model, both because it is a numerically efficient algorithm and because it guarantees a globally optimal solution due to the quadratic nature of the cost function being minimized, namely the sum of the squared prediction errors $\Sigma\varepsilon^2$. A matrix-oriented version of the technique is summarized in Appendix B, which can be modified easily to determine single or multi-input linear filter coefficients by time shifting the relevant inputs and making them serve as independent variables in a multivariate regression.

[17] Even though OLS is guaranteed to minimize prediction residuals in an averaged sense for a given set of training data, there is no guarantee that the estimated parameters are optimal for all possible data. This becomes most apparent when the magnitude of the error term ε is comparable to the predictions themselves. However, if certain assumptions about ε are met, namely that it is a zero-mean, normally distributed random sequence, it becomes possible to place bounds on the optimal parameters based solely on the results obtained from a single-sample OLS estimation. The standard error associated with a given OLS-estimated parameter is described in Appendix C and provides a confidence interval for the estimate that may be used to determine whether or not it is statistically significant.

[18] It is important to note that the uncertainty associated with these confidence intervals is not just a function of the number of observations in a given sample (N) and the variance of the prediction error ($1 - R^2_{YH}$) alone, as is the case for all single-input linear regressions, but they are scaled up by cross correlations that may exist between multiple independent variables. This phenomenon, called multicollinearity, is described in more detail in Appendix C, but it makes intuitive sense if one considers the extreme case when a perfect correlation exists between any two independent variables. In such a situation, it becomes impossible to ascertain which of two cross-correlated independent variables caused a particular change in the independent variable, the standard error becomes infinite, and there is zero statistical significance to the estimated parameter.

[19] One question that has never been adequately addressed in any magnetospheric statistical studies is what happens when prediction residuals are not composed of white Gaussian noise. This is potentially the most relevant consideration when it comes to attempts to interpret linear filter coefficients in any physical context. Appendix C explains how autocorrelated structure in a model's prediction residuals will tend to artificially inflate the statistical significance of a parameter estimate. Circumstances that lead to biased parameter estimates, regardless of the number of observations available in a given sample, are also covered. Specifically, if the investigator fails to include a relevant independent variable in the OLS regression matrix, and that missing variable is partly correlated with one or more included independent variables, the estimated parameters will be biased, thus leading to erroneous physical interpretations. This is in contrast to the situation that arises when an irrelevant variable is mistakenly included in the regression matrix, which simply results in an expected parameter value equal to 0. Multi-input linear filters are one way to reduce the inherent bias associated with OLS parameter estimation, as we see in subsequent sections of the present paper.

3. Data Sources and Preparation

[20] If it is to be independently verified, the empirical nature of this study warrants a thorough description of the data sources and, more importantly, an explanation of the steps taken to preprocess the data used in our linear filter analyses. The solar wind data were obtained from the GSFC/SPDF OMNIWeb interface on the World Wide Web at http://omniweb.gsfc.nasa.gov/ [King and Papitashvili, 2005]. This is an hourly averaged collection of over 40 types of solar wind plasma and magnetic field data. These observations were collected from a variety of satellites and ground-based instruments, cross normalized, and when appropriate, time-shifted for use in magnetospheric studies at 1 AU. This study is limited to three OMNI parameters, the magnitude of the interplanetary magnetic field (B_{imf}), the solar wind's bulk speed (V_{sw}), and solar wind plasma density (ρ_{sw}).

[21] The SAMPEX [Baker et al., 1993] PET [Cook et al., 1993] detects electrons with energies between 2 and 6 MeV. It does so from low altitudes (<600 km) but passes through a broad range of geomagnetic L shells because of its nearly polar orbit. Even though the populations sampled at these low altitudes are dominated by particles with large equatorial pitch angles, it has been shown that SAMPEX-measured fluxes track measurements made at geostationary and other altitudes along the same magnetic flux tube quite closely on timescales less than a day [Kanekal et al., 2001]. This predominately global coherence allows SAMPEX observations to serve as a proxy for radiation belt electron fluxes throughout the inner magnetosphere.

[22] As provided, the OMNI solar wind parameters exhibit very different absolute magnitudes. These differences are worse when the parameters are transformed into SI units. In order to compare directly the prediction filter coefficients associated with each input parameter, these data are normalized by dividing each by their respective standard deviations. It is also common practice to remove the mean of each input parameter from the observation time series, however, we chose to construct the regression matrix in a manner that allows the estimation of an optimal intercept coefficient with the OLS algorithm, rather than forcing it to equal 0 by removing the mean(s) from the training data [see equation (B3) in Appendix B].

[23] The distribution of SAMPEX electron flux data is highly skewed and has a log normal distribution, so a $\log_{10}$

transform was applied to make it more symmetric. This is not absolutely necessary for OLS parameter estimation, but it does help to reduce bias error that arises from the fact that we are attempting to reproduce the dynamics of a nonlinear system with linear models. As a result, any estimated filters will tend to make predictions that are most consistent with average radiation belt conditions, while outliers in the training data will have relatively little effect on the final values of the filter coefficients. Unlike the OMNI data, we did not standardize electron flux data, which means a unit-valued filter coefficient represents an order of magnitude increase in electron flux resulting from a single standard deviation increase in the respective time-lagged solar wind input.

[24] Finally, all data used in this study were initially provided at a subdaily resolution and subsequently reorganized into daily averaged bins that begin and end at noon Universal Time (UT). Subsequently, all time stamps presented in this paper are assumed to fall in the middle of this sample period, or midnight UT. The entire data set used to train and validate the models in this study is presented in Figure 1.

4. Single-Input Radiation Belt Linear Predictions (Revisited)

[25] As noted in section 1, a primary motivation for our particular choice of solar wind parameters is that they represent at least two of the three fundamental coupling mechanisms between the solar wind and the Earth's electron radiation belt that were addressed in detail by *Vassiliadis et al.* [2005]. We begin our own analysis by taking a closer look at the single-input response of SAMPEX electrons to each of the solar wind inputs described above in order to point out one or two aspects of these response functions that were not covered previously.

[26] Figure 2 contains representative single-input linear filter coefficients derived from SAMPEX observations and the solar wind parameters described previously. Each coefficient is plotted as a function of the number of days since a hypothetical day-long "impulse" in its corresponding input equal to 1 standard deviation of that input (see section 3). One-sigma error bars are also included. Because of a large number of observations (nearly 3000) that comprised this training sample, the error bars are small relative to the absolute value of most coefficients. While this is not strictly a formal indication of statistical significance, it is information that could be used to construct confidence intervals and conduct formal hypothesis tests using standard F-score tables if we were concerned primarily with the individual filter coefficients. We are not. It is worth noting, however, that the V_{sw} coefficients' standard errors are larger than those for the other filters, a phenomenon which arises from the high degree of autocorrelation found in the this particular input time series. Even after correcting for this multi-collinearity, the standard error bars remain small relative to the dominant response, and most coefficient estimates would still be statistically significant if our assumptions about ε are valid. We shall see later that this is not necessarily the case.
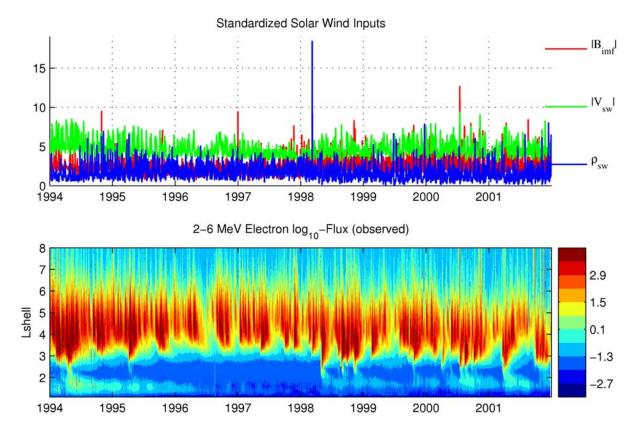
[27] Assuming for the moment that the coefficients are significant, it should also be noted that those corresponding to negative time lags are for the most part near 0. This implies

that changes in flux do not precede changes in a particular solar wind input, validating our initial assumptions about causality. The small deviations from 0 seen just prior to the input impulse may be indicative of some sort of preconditioning of the radiation belts, although it is impossible to determine from the results here whether this is due to internal dynamics or the existence of an additional input that was not considered when training the model parameters with OLS.

[28] The gross features of the response to B_{imf} and V_{sw} are similar. A drop in flux occurs just after an input impulse at a zero-day time lag, followed by a rapid flux increase, and ending with a relatively slow decay to the background state. At first glance, the response to ρ_{sw} seems to have more in common with the response to B_{imf}, except that it has no positive deflection at later time lags. However, when the large negative correlation (<-0.75) between V_{sw} and ρ_{sw} is taken into account, a phenomenon that is due to the pressure balance required in incompressible fluid flows, it is reasonable to imagine that the coefficients might be flipped about the abscissa, giving a response that is actually more similar to that for bulk speed.
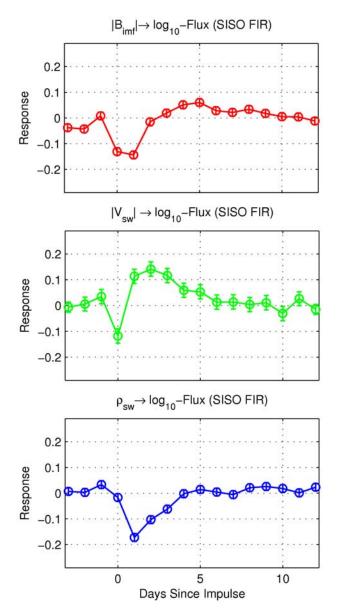
[29] Similarities in all responses can be seen to extend across a wide range of L shells in the profiles shown in the left-hand column of Figure 3, along with additional common features. After one considers the strong anticorrelation between bulk speed and density, there is a significant flux enhancement seen in all three single-input responses that occurs between $L \sim 3$ and $L \sim 4$. Also, there appears to be a general tendency for flux levels to be slightly depressed just prior to the input impulses between $L \sim 4$ and $L \sim 8$, possibly indicating some kind of preconditioning of the radiation belt prior to significant flux enhancements. The latter was identified by *Vassiliadis et al.* [2005] and described as a "precursor" response caused by earlier interplanetary disturbances that are not causally connected to the single input being studied. Such commonality is mostly a consequence of the fact that the same output data were used to train separately each of the single-input filters. These are observations of a radiation belt state that is highly persistent in time when it is not perturbed by changes in the solar wind drivers. Since FIR filters do not include internal dynamics that might account for this persistence, OLS compensates by changing the magnitude of time-lagged filter coefficients to mimic an extended internal response to external stimuli.

[30] There remain notable differences in the response profiles. First, the IMF magnitude responses appear more distributed in time than those for plasma bulk speed and density; the contours of the former almost appear to be a stretched and smoothed version of the latter. This is expected and partly related to the similarities just noted. OLS's compensation for a lack of internal dynamics in the FIR model structure requires that the typically impulsive nature of changes in the IMF magnitude be smoothed over time to reproduce the relatively long timescales associated with the decay of electron flux in the radiation belts. Since V_{sw}, and the anticorrelated portion of ρ_{sw}, vary on timescales similar to the electron fluxes, the responses appear more impulsive. In other words, the area under the curve produced by convolving a filter with its respective input must roughly match the area generated by convolutions of other filters and inputs designed to predict the same output time series.

Figure 1. Eight years of normalized solar wind inputs taken from NASA's OMNI database are shown in the top panel. The $\log_{10}$ of $2-6$ MeV electron fluxes measured by the SAMPEX PET instrument is shown in the bottom panel as a function of L shell in units of Earth radii (R_E).

[31] Also, the minimum observed in the response to V_{sw} occurs about 1 day sooner than the minimum seen in the B_{imf} response. This is explained by known temporal relationships between these two solar wind parameters. In particular, empirical [*Hruska and Hruska*, 1989] and theoretical [*Pizzo*, 1983] studies show that strong enhancements in IMF magnitude tend to precede significant increases in solar wind speed near corotating interaction regions (CIRs) by a day or so. The same Pizzo study, as well as empirical studies as far back as the study of *Gosling et al.* [1972], also showed that sharp increases in density occur at the leading edge of CIRs. However, at a 1-day time resolution, the prolonged density depression that occurs as one enters the rarefaction behind the CIR statistically overwhelms the shock-like enhancement leading the CIR, resulting in a FIR profile that nearly mirrors that for bulk speed. Finally, the minimum of this depression tends to occur $1-2$ days after the peak in bulk speed, explaining the earlier antipeak seen in the density response profile.

[32] While response function profiles certainly provide a useful visual tool, more robust error statistics are required if any confidence is to be assigned to a particular interpretation of the filter coefficients. Standard error bars help in this regard, but it can be difficult to interpret their meaning when a large number of coefficients are associated with a single prediction. Also, standard error bars fail to take into account the effects of correlated prediction errors (see Appendix C). It can be illuminating to examine net prediction error statistics for a particular filter since doing so actually measures the dynamic capabilities of the model in question.

[33] The single-input prediction regression coefficient (PRC, see section 2) profiles shown in the right-hand column of Figure 3 are comparable to the correlation coefficient profiles used by *Vassiliadis et al.* [2002, 2004, 2005], except that they have been constructed using slightly different data sets. They indicate the square root of the fractional variance in the output captured by a model driven by a single solar wind input. In addition to calculating PRCs for the entire 1994–2001 sample, we determined PRCs for two subsets of this interval. The first includes observations between 1994 and the end of 1997, roughly corresponding to the last fully observed minimum in solar activity. The second interval includes observations between 1998 and the end of 2001, corresponding to the last solar maximum.

[34] There is a clear lack of time-stationarity in these PRC profiles. Noting that the variance of the electron flux observations between the two intervals is relatively constant, and assuming that the observation errors are similarly time stationary, the only reason for the change in error statistics is that a relevant input variable was left out of the regression. If one imagines that the included input is statistically independent of any such missing variables, it follows that the interval over which the included input exhibited the largest variance will correspond to the largest PRCs. This partly explains why, in general, B_{imf} predicts better during solar maximum, and V_{sw} predicts better during the solar minimum interval; the variance of B_{imf} is larger during solar max, and the variance of V_{sw} is larger during solar min.

[35] This does not explain why ρ_{sw} does a better job predicting electron flux variations during solar minimum

Figure 2. Time-lagged regression coefficients and their respective $1-\sigma$ error bars are shown for each of three single-input linear filters designed to "predict" the $\log_{10}$ of 2–6 MeV electron fluxes when convolved with their respective solar wind inputs at a single representative L shell.

since, just like B_{imf}, the variance of ρ_{sw} is actually larger during solar max. However, as we already noted, plasma density is highly anticorrelated with solar wind speed. If optimized coefficients in the single-input filters driven by ρ_{sw} are just compensating for the fact that a relevant and correlated input variable was not included in the OLS procedure, it is possible that this overwhelms the effects of nonstationary variances in ρ_{sw}. A similar phenomenon occurs with B_{imf} predictions made at the highest L shells, which appear slightly better during solar minimum than solar max. The next phase of our study moves beyond single-input linear filters and considers the simultaneous effects of different solar wind inputs on radiation belt flux

variations; in particular, what happens when those inputs are correlated with one another.

5. Multi-Input Radiation Belt Predictions

[36] Figure 4 contains representative components of a three-input linear filter and is analogous, if not directly comparable, to Figure 2. However, whereas the OLS estimation of the coefficients for each filter described previously was designed with the a priori assumption that there were no other relevant inputs, here the regression matrix was expanded to include observations of all the solar wind inputs being considered in this study. As before, the standard errors for the solar wind bulk speed associated coefficients are larger than the other inputs due to the high degree of autocorrelation in the V_{sw} time series. Correlations between the separate inputs slightly increase the standard errors of all the coefficients, but the general improvement in predictive ability helps compensate for this increase in uncertainty.

[37] The small standard error bars should imply that multi-input filter coefficients must not differ much from those estimated using a single input only. However, they clearly do differ from their single-input counterparts, often by factors of several to tens of standard deviations, a fact that becomes even more apparent in Figure 5. One obvious difference is the absence of a flux dropout between $L \sim 4$ and $L \sim 8$ immediately following increases in V_{sw} when other inputs are considered. The dropout remains in the B_{imf} component filter profile, however. Also, the brief enhancement seen between $L \sim 3$ and $L \sim 4$ in the single-input V_{sw} response profile disappears when additional inputs are considered, while the relatively prolonged response to B_{imf} remains. Considered together, these results are highly suggestive that the single input V_{sw} filter coefficients compensate for dynamics normally associated with changes in the B_{imf} when the latter is not considered during OLS parameter estimation.

[38] A less obvious but possibly more relevant difference can be seen in the ρ_{sw} response profile. If one compares the model output from the single-input ρ_{sw} filter with the predictions made using the ρ_{sw} component of the multi-input filter alone (neither shown here), they will see that enhancements predicted using the single-input filter occur when plasma density drops, while flux enhancements predicted using the ρ_{sw} component of the multi-input filter correspond to increases in plasma density. Even though the negative response between $L \sim 4$ and $L \sim 8$ initially appears to remain intact when switching from the single-input to the ρ_{sw} component of our multi-input filter, it can no longer be considered just a consequence of the anticorrelation between ρ_{sw} and V_{sw}. Indeed, it appears to be a genuine response to changes in the plasma density. We conclude from these results that single-input ρ_{sw} filters provide a highly biased representation of the true radiation belt dynamic response to solar wind changes when much more relevant input variables, especially the solar wind bulk speed, are neglected during OLS estimation.

[39] The multi-input component PRCs shown in the right-hand column of Figure 5 provide additional dynamical insights into the system being studied, especially when contrasted with their single-input counterparts. Recall that
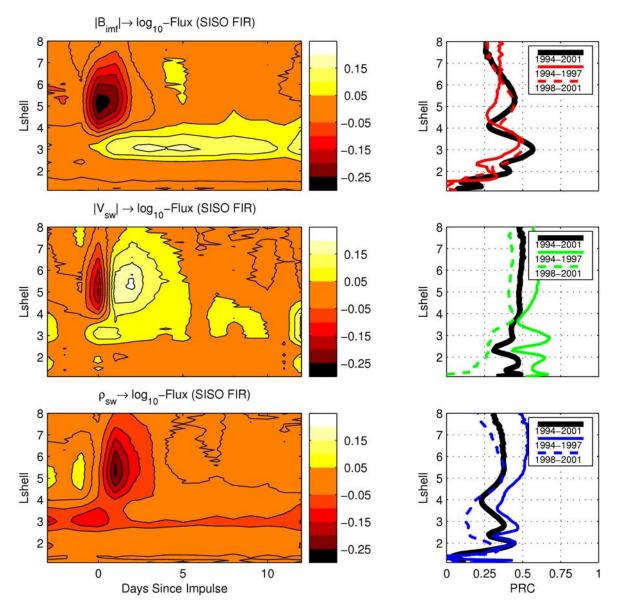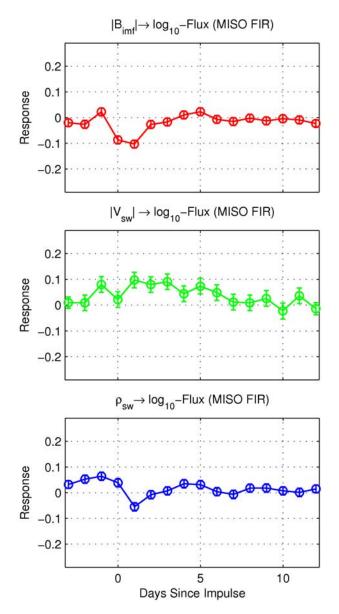
Figure 3. Single-input, L-shell-dependent FIR linear filter profiles are shown on the left. Their corresponding PRC profiles are shown on the right and represent the square root of the fractional variance of the $\log_{10}$ of $2-6$ MeV electron fluxes attributable to the single input if no other inputs are considered.

these may be considered the square root of a fraction of the variance of observed radiation flux that can be attributed to each input when every other input is held constant. Since we have already acknowledged that our inputs are correlated, this is not a physically realistic scenario, but this is precisely what we attempted to do in the single-input case by effectively fixing all other relevant inputs equal to zero. If the additional inputs are uncorrelated with the input being considered, there should be no difference between single- and multi-input profiles when real observations of additional inputs are considered. A comparison of single-input PRCs with component PRCs from the multi-input filter should therefore help identify where single-input filters compensate for missing but correlated inputs.

[40] First, the ρ_{sw} component of the multi-input filter appears to have little predictive ability at any but the lowest L shells when it is considered in conjunction with the other two inputs, and this is mostly a consequence of the high

dynamic stability of the inner electron belt. Combined with the fact that single input ρ_{sw} filters predict better during solar minimum, despite the fact that plasma density exhibits more variability during solar max (see section 4), it becomes clear that ρ_{sw}-driven predictions mostly compensate for the effects of a neglected V_{sw} input in the form of biased filter coefficients. However, as explained in section 2 and Appendix C, extraneous inputs will not adversely affect the optimal estimation of filter coefficients for the more relevant inputs. The fact that the ρ_{sw} filters retain at least a little predictive power means that this input variable should not be dropped from the model unless the OLS estimation becomes an unreasonable computational burden.
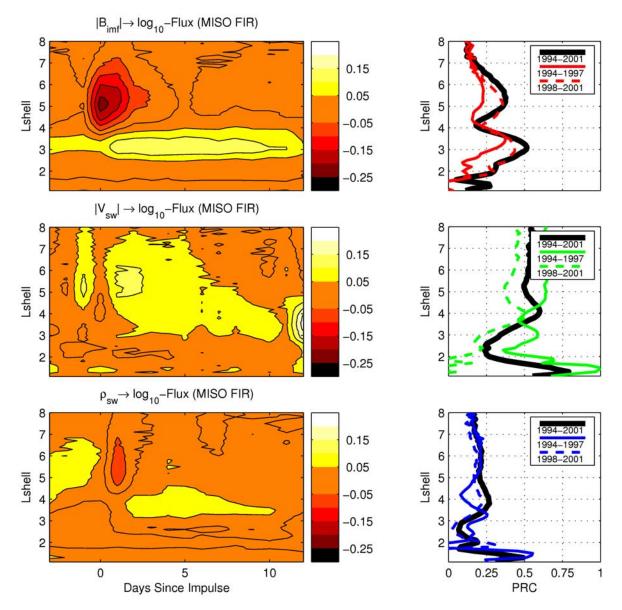
[41] Next, the PRC profiles for the B_{imf} component predictions are largely unchanged from their single-input counterparts, whether determined from the entire 1994–2001 sample or one of the two subintervals. This suggests that the time variability of the B_{imf} PRCs is mostly a

Figure 4. Time-lagged regression coefficients, and their respective $1 - \sigma$ error bars, are shown for each of three components of a multi-input linear filter designed to "predict" the $\log_{10}$ of 2–6 MeV electron fluxes when convolved with their respective solar wind inputs independently at a single representative magnetic L shell.

function of the changing variance of B_{imf} itself, since the variance of the electron flux is relatively constant from one subinterval to the next, not cross correlations with other inputs. One difference between single- and multi-input PRCs is seen easily at the highest L shells, where the influence of B_{imf} during the solar minimum interval drops to near zero when additional inputs are considered. This region was noted previously, when we hinted that optimized single-input B_{imf} filter coefficients might compensate for a missing input because time-dependent single-input PRCs were not consistent with changes in the variance of B_{imf} between subintervals. We conclude that single-input B_{imf} filters associated with these altitudes contain biased coefficients.

[42] Finally, while the V_{sw} component PRC profiles are largely unchanged above $L \sim 5$, there is a substantial drop seen in the influence of solar wind bulk speed on electron flux between $L \sim 2$ and $L \sim 3$ when additional inputs are considered. In contrast, while the tendency for B_{imf} to drive flux variations in the same region is not very strong, there is little difference between single- and multi-input filters. Therefore, of the solar wind inputs considered here, we conclude that B_{imf} is the biggest contributor to flux variations in the so-called slot region. Researchers exploring connections between plasmapause position and strong electron enhancements at low L shells [e.g., *Baker et al.*, 2004b; *Goldstein et al.*, 2005] may find this result intriguing, but a first-principles physical explanation of this phenomenon is beyond the scope of the present paper.

[43] As additional relevant inputs are included in the linear regression, improvements in the cumulative predictive ability of the model are expected because of the fact that the statistically independent portions of partly correlated input data will either contribute nothing or add positively to the total predictable variance in the electron flux time series; they cannot subtract from the model's predictive ability if optimized using OLS [*Cohen and Cohen*, 1983]. This may be seen in Figure 6 as a significant drop in flux in mid-1997, a major enhancement lasting through the final months of 1999 and into early 2000, and an impulsive increase extending to relatively low L shells early in 2001, all of which correspond to similar episodes in the real data presented in Figure 1. These changes are simply not apparent in single-input filter predictions, which have not been presented here for the simple reason that they do not exhibit any interesting time-variability. Clearly, our model misses some major events, especially during 1998. However, 1998 was characterized by a number of strong CME-driven magnetic storms that exhibit a highly nonlinear influence on the radiation belts, something our filters, and indeed any linear model, is not capable of reproducing with time-stationary coefficients.

[44] Perhaps more important than visually "interesting" time series is the fact that cumulative prediction error statistics will become more time stationary as additional inputs are incorporated into the multi-input filter if it is assumed that observation uncertainty remains constant in time. This can be seen at all but the lowest L shells in Figure 7. This cumulative PRC profile is directly comparable to each of the profiles presented in section 4 because the statistics describe the predictive power of the model as a whole. Regions where cumulative PRCs still vary substantially with time are indicative of missing relevant inputs that are partly correlated with those already included in the model. These missing independent variables might be a nonlinear transformation of an input already used or a completely different variable, including time-lagged versions of the electron flux itself. We performed a brief survey of some typical nonlinear transformations, including logarithms and low-order polynomial expansions, and concluded that the trivial improvements seen in our prediction statistics meant that the relatively raw versions of our three solar wind parameters were adequate for our needs.

[45] We also tried several new variables, including nonlinear combinations of our three inputs. These gave small, but nontrivial, improvements beyond $L \sim 4$ that may be
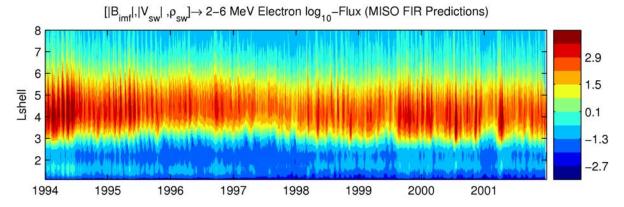
Figure 5. Multi-input, L-shell-dependent FIR linear filter profiles are shown on the left. Their corresponding PRC profiles are shown on the right and represent the square root of the fractional variance of the $\log_{10}$ of $2-6$ MeV electron fluxes if every other input considered in the multi-input model were held constant.

studied more closely in the future, but they failed to account for the large temporal discrepancy seen in the PRC profile below $L \sim 2$. The latter region is characterized by extremely autocorrelated prediction errors, with time constants on the order of the sampling period. According to statements made in section 2, and more importantly the demonstrations provided in Appendix C, this means that normal standard error bars will not provide a realistic bound on filter coefficients estimated for this region. The large difference between PRCs calculated during solar minimum and solar max may very well be consistent with the true uncertainties for this region.

[46] We found that if a one-day time-lagged measurement of the electron flux itself was added to the regression matrix, prediction statistics improve substantially, especially at the lowest L shells. The addition of this autoregressive (AR) component to our model creates what is most commonly referred to as an autoregressive filter with exogenous inputs (ARX) or, in some older literature, an autoregressive moving average filter. The AR component does so much to improve prediction statistics at lower L shells because the inner electron radiation belt is highly persistent, exhibiting e-folding times on the order of months, if not years. The 16 one-day time lags used in our FIR filters cannot mimic this entire response, leading to strongly biased coefficients below $L \sim 2$. The single AR term actually serves as a discretized form of the first-order linear dynamics that dominate the inner belt, while the remaining FIR coefficients now account for solar wind perturbations and higher-order dynamics only.

[47] There are problems associated with the naive use of AR filters. If the flux observations on which the AR model

Figure 6. Cumulative multi-input predictions of the $\log_{10}$ of $2-6$ MeV electron fluxes as a function of L shell.

operates are rewritten as $Y \equiv \hat{Y} + \varepsilon$, and the AR model is composed of a single linear coefficient b_{ar}, the predicted output will be $\hat{Y}_t = b_{ar} \hat{Y}_{t-1} + b_{ar}\varepsilon_{t-1}$. In other words, the same linear dynamics govern both the model's predicted output $\hat{Y}$ and the prediction error ε. The coefficient b_{ar} is guaranteed to be unbiased, and therefore represent the true linear first-order response of the system, if and only if the errors are white Gaussian noise sequences with a mean of zero, since these cannot influence the OLS estimation of the AR coefficient. If ε is not distributed like white Gaussian noise, the OLS-estimated b_{ar} may still be unbiased, but only if the assumption that system and error dynamics are identical holds true. This will not be the case in general, so b_{ar} will be biased in order to compensate for correlated structure in the error term that is unrelated to the system dynamics. This can lead to a severe misrepresentation of the true system dynamics.

[48] This problem is inherent to the OLS estimation procedure [*Ljung*, 1999; *Nelles*, 2001]. One solution to this dilemma is to drop OLS and use a more sophisticated maximum likelihood solver that guarantees unbiased estimates. This usually involves some sort of iterative algorithm and can be computationally burdensome. Aside from the computational requirements, however, there is generally no guarantee that these algorithms will converge to a globally optimal set of filter coefficients, since the recursive nature of the integration required to generate $\hat{Y}$ results in an inherently nonlinear estimation problem. For these reasons, we will save the addition of an AR term to our model for a future study and accept the poor error statistics generated by FIR filters in the inner electron radiation belt.
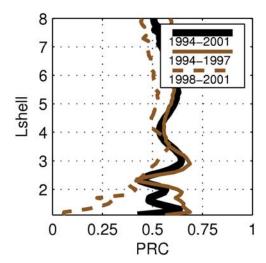
6. Summary and Conclusions

[49] As discussed in this paper, so-called finite impulse response (FIR) linear filters have been used for decades to help investigators of radiation belt dynamics reproduce variations in the real world observations, even when a full understanding of the underlying system dynamics was absent. These models do not constitute a real dynamical system in the strictest sense, since they do not possess feedback mechanisms. Even single-input filters are nothing more than sophisticated multivariate linear regression models whose independent variables happen to be composed of time-lagged versions of one relevant input. This, however, means that such models are in fact more flexible than their name

would suggest. In essence, they constitute a normalized superposed epoch representation of a system's average response, which may or may not be linear, that can be used to make robust predictions at relatively little computational expense.

[50] The most common type of FIR models used to date have been single-input linear prediction filters. These are relatively easy to generate from a limited set of training data, and they accurately predict all but the most extreme modulations observed in the magnetosphere because this system is driven strongly by external perturbations, namely the solar wind. In addition, single-input FIR filters have been able to provide valuable physical insights about the Earth's radiation belts because this is a system characterized by low-order dynamics, at least at global scales, and the time-lagged filter coefficients do indeed resemble the true dynamical response of the radiation belts. However, it was clear early on that variations in the radiation belts occurred that were not related to the single input used to drive the model.

[51] *Vassiliadis et al.* [2005], often cited in this paper, represented the first systematic survey of the dynamic response of the electron radiation belts to a variety of inputs in an effort to better understand both the temporal and

Figure 7. L-dependent PRCs based on the cumulative output of the multi-input filter. These represent the square root of the total fractional variance, $R^2_{\hat{Y}H}$, captured by the model, and are directly comparable to the single-input PRCs presented in Figure 3.

spatial coupling between these inputs and changes in energetic electron flux within the Earth's magnetosphere. Even though this study continued to use single-input linear filters, a careful analysis of the prediction error statistics associated with each of these models led to a statistically robust demonstration that the radiation belts not only exhibit a complicated internal dynamical structure but that different solar wind inputs impact the radiation belts in very distinctive ways.

[52] The original intent of the study detailed in the present paper was simply to document a model that was designed to incorporate this improved understanding of radiation belt dynamics into a better prediction tool. In the process, however, we came to appreciate many nuances associated with multivariate linear regression, many of which have been detailed in section 2 of this paper, as well as Appendix C. Perhaps the most important conclusion to come from this is that underspecified models will invariably possess biased coefficient estimates. In other words, if all the relevant inputs are not included in the OLS estimation, coefficients associated with the inputs that are included will adjust to compensate. At best, this improved understanding of multivariate linear regression complicates any physical interpretations that have been made previously using single-input linear filters; at worst, it might completely invalidate them. Fortunately, well-considered error statistics are fairly robust to these issues, and we find no cause to contradict most of the conclusions reached by *Vassiliadis et al.* [2005].

[53] We briefly reviewed our own single-input filter results from the viewpoint of multivariate linear regression, partly to nudge the reader away from the more traditional dynamical systems interpretation of linear filter output but also to suggest insights that may not have been addressed in previous magnetospheric studies. First, we noted that the L-dependent response profiles for single-input models tend to share many common features, a phenomenon that arises because the FIR filter coefficients corresponding to different solar wind inputs are mostly mimicking the low-order internal dynamics of the electron radiation belts. This is necessary because the feedback mechanism necessary for a proper dynamical response is not included in the FIR model structure. In one sense, these FIR coefficients are all severely biased because the most relevant input, the previous day's electron flux, was left out of the regression. We accept this bias because the FIR filters used here nearly all had a sufficient number of time lags to capture the radiation belt's initial response, peak, and decay back to zero.

[54] Some similarities between the single-input response profiles began to fade as we moved from an analysis of the filter coefficients themselves and began to look more closely at the prediction error statistics in the form of PRC profiles as a function of L shell. These fading similarities were especially obvious when comparing the V_{sw} and ρ_{sw}-driven predictions with those derived from B_{imf} and its corresponding filters. The predictive ability of the former was relatively high across a broad range of higher ($>4R_E$) L shells, while the latter exhibited two distinct peaks, one near $L \sim 3$ and the other near $L \sim 5$.

[55] There was significant time-variability in these PRC profiles as well. The B_{imf} and V_{sw} PRCs appear to vary as functions of solar cycle in a manner consistent with the

variances of their respective observation time series (that is, B_{imf} PRCs are higher at solar max, when the variance of the interplanetary magnetic field is strongest, while V_{sw} PRCs are higher at solar minimum, when the variations in observed solar wind speed are most pronounced). In contrast, the time-variability of the ρ_{sw} PRC profiles roughly matched that of the V_{sw} PRCs, even though the variance in ρ_{sw} is largest during solar maximum. This turns out to be a consequence of the large anticorrelation that exists between V_{sw} and ρ_{sw}, which generates a fairly strong bias in the single-input filter coefficients for ρ_{sw} when no other inputs are considered.

[56] Our multivariate statistical approach led naturally to a comparison of single-input linear filters with their analogous component filters in multi-input linear filters, as well as single-input with cumulative multi-input predictions. Several additional insights arose from this comparison. First, most of the filter coefficients vary significantly from their single-input counterparts, often by several standard errors, implying that the single-input filters are composed of biased coefficients. Such changes indicate that a flux decrease seen between $L \sim 4$ and $L \sim 8$ that was common between the B_{imf} and V_{sw} single-input filter profiles is due almost exclusively to B_{imf}, since the dropout disappeared from the V_{sw} component of the multi-input response profile entirely. Also, the quick positive response to solar wind speed between $L \sim 3$ and $L \sim 4$ disappeared in the multi-input filter profile, while the corresponding flux enhancement following increases in B_{imf} remained when other inputs were considered. This implies that the single-input V_{sw} response profiles contain biased filter coefficients. In addition to a substantial drop in the average magnitude of the filter coefficients, we noted that while flux enhancements were associated mostly with drops in ρ_{sw} when the single-input filter is used, it is the increases in ρ_{sw} that tend to correspond to flux enhancements when the multi-input filter is used. We interpreted this to mean that single-input ρ_{sw} filter coefficients are severely biased as a consequence of neglecting the highly anticorrelated solar wind bulk speed in the OLS estimation.

[57] Next, multi-input PRC profiles pointed toward additional regions where the single-input linear filter coefficients were likely biased as a result of missing relevant inputs. Single-input ρ_{sw} predictions in particular seem to be largely a consequence of biased coefficients, since the ρ_{sw} component of the multi-input filter captured almost none of the observed variance in the electron flux measurements when other inputs were considered. PRC profiles for B_{imf} changed at higher L shells, while PRC profiles for V_{sw} changed between $L \sim 2$ and $L \sim 3$, when comparisons were made between single- and multi-input component filters. This indicates that V_{sw} is the principal solar wind driver at the highest radiation belt altitudes, and B_{imf} is the principal driver in the slot region. Finally, the PRC profile associated with the cumulative multi-input predictions showed how adding relevant inputs to the linear regression not only improves the overall predictive ability of the model with each added input but also how multi-input linear filters reduce the time variability of the prediction error statistics relative to single-input filters.

[58] A substantial fraction of the overall electron flux variability can be attributed to the one-day time-lagged electron flux itself, especially at the lowest L shells, sug-

gesting that an ARX model might better describe these system dynamics. However, we described a well-known shortcoming of the OLS estimation procedure that increases further the potential for biased filter coefficients when time-lagged observations of the desired system output are incorporated into the estimation procedure. Any useful model of the Earth's electron radiation belts should probably include some sort of AR term, but we left the more sophisticated estimation algorithm necessary to do this properly to be described and implemented in a future study.

[59] All of this is not to say that the best approach to modeling magnetospheric systems like the electron radiation belts is to throw every conceivable input at the problem and expect physically meaningful coefficients to come from the OLS procedure. Even though OLS should in principle weigh the true relevant input to the model higher than another variable that happens to be correlated with the true input, a more likely scenario involves two variables, neither of which are the true input, but are both correlated to some degree with the true input. It remains the researcher's responsibility to study the filter coefficients and, perhaps more importantly, the general statistical characteristics of all the inputs, the output, and the prediction errors, in order to deduce the true relevant input and the physics that govern how it couples with the system being studied. Multi-input linear filters facilitate this process.

Appendix A: Estimating R^2

[60] Equations (3) and (4) in section 2 describe two different relationships between R^2, or the explained variance of a dependent variable, standardized multivariate linear regression coefficients, and any linear correlations that exist between the dependent variable and/or separate independent variables. Here we present proofs that lead to these relationships in the case of two independent variables. Proofs relating to an arbitrary number of independent variables are a simple, if somewhat algebraically tedious, extension of the following.

[61] First, consider a two independent variable linear regression with zero-mean errors ε:

$$Y = b_1 X_1 + b_2 X_2 + \varepsilon \tag{A1}$$

There are two ways to proceed, the first leads to equation (3) in section 2. First, square each side of equation (A1). Assuming the coefficients are constant, applying expectations leads to:

$$E(Y^2) = E\left([b_1 X_1 + b_2 X_2 + \varepsilon]^2\right)$$

$$E(Y^2) = E\left(b_1^2 X_1^2 + 2b_1 b_2 X_1 X_2 + b_2^2 X_2^2 + 2b_1 \varepsilon X_1 + 2b_2 \varepsilon X_2 + \varepsilon^2\right)$$

$$E(Y^2) = b_1^2 \sigma_{X_1}^2 + b_2^2 \sigma_{X_2}^2 + 2b_1 b_2 \sigma_{X_1 X_2} + 2b_1 \sigma_{\varepsilon X_1} + 2b_2 \sigma_{\varepsilon X_2} + \sigma_\varepsilon^2 \tag{A2}$$

Since the intercept is 0 by construction, the expected value of Y^2 is just its variance σY^2. Now if we assume that the errors are white noise, and therefore unrelated to the independent variables in any way, and we assume that we

defined the dependent and all independent variables such that $\sigma_Y^2 = \sigma_{X_k}^2 = 1$, the right-hand side of equation (A2) simplifies to:

$$1 = b_1'^2 + b_2'^2 + 2b_1' b_2' \sigma_{X_1 X_2} + \sigma_\varepsilon^2$$

$$1 = \sum_{i=1}^{K} \sum_{j=1}^{K} b_i' b_j' \sigma_{X_i X_j} + \sigma_\varepsilon^2$$

$$1 = R^2 + \sigma_\varepsilon^2$$

... or ...

$$1 = \text{explained \% variance} + \text{unexplained \% variance} \tag{A3}$$

[62] It would have been just as valid to multiply both sides of equation (A1) by the variable Y and then apply expectations. Again, assuming the coefficients are constant, we get:

$$E(YY) = b_1 E(X_1 Y) + b_2 E(X_2 Y) + E(\varepsilon Y)$$

$$E(YY) = b_1 E(X_1 Y) + b_2 E(X_2 Y) + E(\varepsilon [b_1 X_1 + b_2 X_2 + \varepsilon])$$

$$E(YY) = b_1 \sigma_{X_1 Y} + b_2 \sigma_{X_2 Y} + b_1 \sigma_{\varepsilon X_2} + \sigma_\varepsilon^2 \tag{A4}$$

Here the terms $\sigma_{X_k} Y$ represent cross correlations between the dependent variable Y and each independent variable X_k. Likewise, $\sigma_{\varepsilon X_k}$ represents cross correlations between the errors and the independent variables, which should be zero if the errors are indeed white noise. Finally, once again assuming that $\sigma_Y^2 = \sigma_{X_k}^2 = 1$ and simplifying the left-hand side gives:

$$1 = b_1' \sigma_{X_1 Y} + b_2' \sigma_{X_2 Y} + \sigma_\varepsilon^2$$

$$1 = \sum_{k=1}^{K} b_k' \sigma_{X_k Y} + \sigma_\varepsilon^2$$

$$1 = R^2 + \sigma_\varepsilon^2$$

... or ...

$$1 = \text{explained \% variance} + \text{unexplained \% variance} \tag{A5}$$

This proof corresponds to the relationship given in equation (4).

Appendix B: Ordinary Least Squares (OLS)

[63] The OLS algorithm minimizes a quadratic function, the sum of the squared errors, and so possesses a single global minimum that can be solved for analytically. This linear estimation algorithm has origins that can be traced back in time to Carl Friedrich Gauss and has undergone countless refinements since, so we will not provide a full derivation here. We do briefly summarize how it is usually implemented in practice using matrix notation. As we already noted, Y and $X_{k=1 \to K}$ represent a single dependent and a set of independent variables used for our linear regression model, respectively. If our training data sample includes N observations, it is possible to write out the system of equations as follows:

$$Y_1 = a + b_1 X_{1,1} + b_2 X_{2,1} + \cdots + b_K X_{K,1} + \varepsilon_1$$
$$Y_2 = a + b_1 X_{1,2} + b_2 X_{2,2} + \cdots + b_K X_{K,2} + \varepsilon_2$$
$$\vdots$$
$$Y_N = a + b_1 X_{1,N} + b_2 X_{2,N} + \cdots + b_K X_{K,N} + \varepsilon_N \tag{B1}$$

This set of equations can be expressed in the following matrix formulation:

$$\mathbf{Y} = \mathbf{XB} + \varepsilon \tag{B2}$$

where the matrices $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{B}$, and ε are defined as:

$$\mathbf{Y} \equiv \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} \mathbf{X} \equiv \begin{bmatrix} 1 & X_{1,1} & X_{2,1} & \cdots & X_{K,1} \\ 1 & X_{1,2} & X_{2,2} & \cdots & X_{K,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{1,N} & X_{2,N} & \cdots & X_{K,N} \end{bmatrix}$$

$$\mathbf{B} \equiv \begin{bmatrix} a \\ b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} \varepsilon \equiv \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix} \tag{B3}$$

Assuming that the so-called regression matrix $\mathbf{X}$ has rank K, implying that no independent variable is perfectly correlated with any other, and assuming that K is less than the number of observations N, the optimal set of coefficients minimizing the sum of the squared errors $\varepsilon^{\mathrm{T}}\varepsilon$ is determined by calculating the pseudo-inverse of the nonsquare matrix $\mathbf{X}$ and multiplying it by the vector of dependent variables $\mathbf{Y}$, as shown in equation (B4).

$$\hat{\mathbf{B}} = \left[\mathbf{X}^{\mathrm{T}}\mathbf{X}\right]^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \tag{B4}$$

Appendix C: Uncertainty and Bias in Parameter Estimates

[64] If two independent and finite samples of observations are drawn from a theoretically infinite population, then used to calculate a set of regression coefficients using OLS, the two estimates will differ because of errors ε. As more samples are drawn, and additional estimates are calculated, the Central Limit Theorem dictates that a Gaussian distribution of estimates will evolve whose mean equals the expected value of the estimate in the absence of errors. An estimate distribution with a relatively small standard deviation is said to be "efficient." An efficient estimate is unlikely to differ significantly from the expected value and can usually be trusted. Conversely, an estimate distribution with a large standard deviation indicates that any single parameter estimate is likely to differ from the population's expected value and should be used with caution.

[65] The so-called "standard error" of a regression coefficient is really just a convenient means of determining the standard deviation of an estimate distribution using only the parameter estimate from a single sample and its associated prediction errors. It does this by making the assumption that ε is a zero-mean, perfectly random sequence (white noise). Our confidence (or lack thereof) in the kth estimated coefficient can then be calculated using equation (C1).
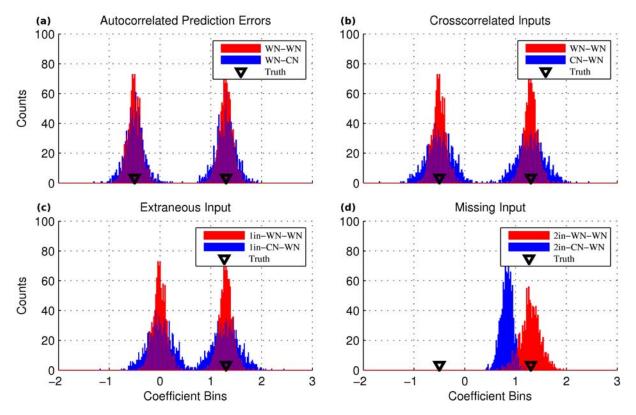
$$S_{b_k} = \sqrt{\frac{1 - R_{YH}^2}{\left(1 - R_{X_k G_k}^2\right)(N - K - 1)}\left(\frac{\sigma_Y}{\sigma_{X_k}}\right)} \tag{C1}$$

In this relationship, N represents the size of the sample, and K represents the number of independent variables, while σ_Y and σ_{X_k} represent the standard deviations of the dependent and each independent variable, respectively. The $R_{X_k G_k}^2$ term in the denominator of the radical is similar to the coefficient of determination described in section 2 and Appendix A. Here, however, each X_k serves as the dependent variable, while G_k is defined as the set of all independent variables other than X_k. This is known as multicollinearity and describes how confidence in a particular estimate will be reduced if the independent variables are correlated with one another. In the extreme case when a perfect correlation exists between two independent variables, it is impossible to ascertain which one actually caused a change in the dependent variable, and accordingly, the standard error becomes infinite.

[66] Figure C1b demonstrates this phenomenon using a virtual population constructed by simulating a two-input system whose linear coefficients equal 1.3 and −0.5, respectively. Two scenarios were then explored, one in which the two inputs were uncorrelated with one another, and the other in which the two inputs exhibited a finite correlation equal to ∼0.9. The simulated output from each model was corrupted with white noise to ensure a nonzero standard deviation for each parameter's estimate distribution. The standard deviation of the distributions of OLS estimates made from samples of the population constructed with uncorrelated inputs are approximately equal to 0.10, which is exactly the value of the standard errors calculated using equation (C1). The standard deviation of the distributions of estimates made using correlated inputs is equal to 0.24, also identical to the standard error, and indicative of the higher level of uncertainty in those estimates that arises from correlations between independent variables.

[67] The assumption that prediction errors comprise a zero mean, white noise sequence rarely holds true in practice. In particular, it is often the case in time series analysis that the error term exhibits first-order, positive autocorrelated structure [*Pindyck and Rubinfeld*, 1991]. If the interval over which a sample is taken is characterized by significant errors, and that interval is less than, or comparable to, the correlation time of the errors, the estimated parameters will deviate significantly from their expected value. However, because the errors were positively correlated with themselves, the spread of these residuals will actually be relatively small, even though the actual amplitudes might be quite large. If another sample is taken, the autocorrelated nature of the residuals imply that it is equally likely that the residuals are small or even that they are significant but of opposite sign to the first scenario. The end result is that S_{b_k} for any single sample will be less than the standard deviation determined by calculating a large number of parameter estimates using many independent samples from a common population to build up a distribution of these values. In other words, the presence of a first-order positive autocorrelation in the error term will lead to overconfidence in the regression coefficients.

[68] To demonstrate this particular phenomenon, another virtual population was constructed similar to the first, except that this time it was corrupted with noise exhibiting a first-order positive autocorrelation with a correlation time equal to approximately 100 simulation steps. Figure C1a shows the

Figure C1. The spectral characteristics of both inputs and observation noise may influence estimate uncertainty and/or bias. (a) Comparison of distributions of parameter estimates made using observations of simulated populations corrupted by white (-WN) and correlated (-CN) noise, respectively. (b) Comparison of estimate distributions made from populations corrupted by white noise only but whose inputs may (CN-) or may not (WN-) be correlated with one another. (c) Comparison of estimate distributions made from populations corrupted by white noise only and whose inputs may or may not be correlated with one another, when an extraneous input is included in the OLS estimation procedure. (d) Comparison of estimate distributions made from populations corrupted by white noise only and whose inputs may or may not be correlated with one another, when a relevant input is left out of the regression.

estimate distributions for the parameters for both the white noise corrupted simulation and the colored noise corrupted simulation. The standard deviations for estimates derived from both scenarios are equal to approximately 0.10, however when one uses equation (C1) to determine the standard error, it turns out to be approximately 0.05 for the colored noise scenario. This is significantly smaller than the real uncertainty, represented accurately by the blue histograms in Figure C1a, and may lead the unwary investigator to assign more confidence to the parameter estimates than they should.

[69] There are circumstances when even though the efficiency may be high, the population's expected value does not equal the true parameter value. This is called bias error and has two primary causes, (1) a correlation exists between the error term and one or more of the independent variables and (2) the model structure is under specified (that is, one or more relevant inputs are disregarded during parameter estimation). The former is very similar to the problem of an autocorrelated error term, except that the errors do not have a mean equal to 0 for the entire population, and skewed parameter estimates will never cancel out, regardless of how many independent samples are taken or how large the sample size is. A familiar real-world example might involve an instrument whose meas-

urements always exhibit an offset equal to a linear function of its target's true value. This might produce an output that is clearly nonphysical to a trained expert, but the OLS algorithm has no way of differentiating this offset from what should be the measurement's true value. There is very little an investigator can do to alleviate this type of bias except to remain vigilant during the instrument design and calibration processes.

[70] The second source of bias error, model under specification, is pertinent to multivariate linear parameter estimation, but before we address this issue, let us first consider a model that has been over-specified. Imagine yet another virtual population that has been constructed using only the first of the two true coefficients discussed previously. The investigator has no a priori knowledge that this is the case however, so they choose a second, and ultimately irrelevant, independent variable to include in the OLS procedure. While this second independent variable exhibits no relationship to the dependent variable, it is still potentially correlated with the first independent variable. Estimates of the first regression coefficient remain exactly the same as they were previously, as seen in Figure C1c. Estimates for the second regression coefficient, however, shift to the right, resulting in a zero-centered distribution. This is not bias.

Even though the virtual population was constructed with the assumption that there was no second independent variable, the final result is identical to a situation in which there actually was a second independent variable, but the second coefficient was defined to be zero.

[71] Figure C1d tells a very different story. The same virtual populations used for Figures C1a and C1b were used in this scenario, but the investigator mistakenly assumed that only a single independent variable was relevant. Using OLS to estimate 1000 single coefficient estimates from hundred-observation samples generates very different distributions than what was expected had the number of independent variables been chosen correctly. In the case of uncorrelated independent variables, OLS simply regards the second independent variable as an additional source of error, thus decreasing the efficiency. Because these errors are uncorrelated with errors introduced intentionally during construction of the virtual population, there is no bias error. On the other hand, when there exists a correlation between the two true independent variables, but only one is used to try to explain variations in the dependent variable, the mean of the estimated parameter distribution is shifted significantly from its true value. This bias will remain even as the sample size tends toward infinity.

[72] This tendency for bias error when the linear model is underspecified is also pertinent when a linear model is used to approximate nonlinear system dynamics. A simple example involves a standard polynomial model.

$$Y = a + b_1 X_1 + b_2 X_1^2 + \cdots + b_N X_1^N + \varepsilon \qquad (C2)$$

[73] If equation (C2) represents the true system, but equation (1) is used as an approximation of these dynamics, OLS will give a biased estimate of the intercept and single regression coefficient b_1 because we have failed to include all relevant variables. The same conclusion holds true even when the polynomial is only an approximation of an inherently nonlinear equation.

References

Akasofu, S.-I. (1979), Interplanetary energy flux associated with magnetospheric substorms, *Planet. Space Sci.*, *27*, 425–431.

Arnoldy, R. L. (1971), Signature in the interplanetary medium for substorms, *J. Geophys. Res.*, *76*, 5189–5200.

Baker, D. N., J. B. Blake, R. W. Klebesadel, and P. R. Higbie (1986), Highly relativistic electrons in the earth's outer magnetosphere 1: Lifetimes and temporal history 1979–1984, *J. Geophys. Res.*, *91*, 4265–4276.

Baker, D. N., R. L. McPherron, T. E. Cayton, and R. W. Klebesadel (1990), Linear prediction filter analysis of relativistic electron properties at 6.6 R$_E$, *J. Geophys. Res.*, *95*, 15,133–15,140.

Baker, D. N., G. M. Mason, O. Figueroa, G. Colon, J. G. Watzin, and R. M. Aleman (1993), An overview of the Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX) mission, *IEEE Trans. Geosci. Remote Sens.*, *31*, 531–541.

Baker, D. N., R. S. Weigel, E. J. Rigler, R. L. McPherron, D. Vassiliadis, C. N. Arge, G. L. Siscoe, and H. E. Spence (2004a), Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical models, *J. Atmos. Sol.-Terr. Phys.*, *66*, 1491–1497, doi:10.1016/j.jastp.2004.04.011.

Baker, D. N., S. G. Kanekal, X. Li, S. P. Monk, J. Goldstein, and J. L. Burch (2004b), An extreme distortion of the Van Allen Belt arising from the 'Hallowe'en' Solar Storm in 2003, *Nature*, *432*, 878–881.

Bargatze, L. F., D. N. Baker, R. L. McPherron, and E. W. Hones Jr. (1985), Magnetospheric impulse response for many levels of geomagnetic activity, *J. Geophys. Res.*, *90*, 6387–6394.

Clauer, C. R. (1986), The technique of linear prediction filters applied to studies of solar wind-magnetosphere coupling, in *Solar Wind-Magnetosphere Coupling*, pp. 39–57, edited by Y. Kamide and J. A. Slaving, Terra Sci., Tokyo.

Cohen, J., and P. Cohen (1983), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Second Edition*, Lawrence Erlbaum Associates, Florence, Kentucky.

Cook, W. R., et al. (1993), PET: A proton/electron telescope for studies of magnetospheric, solar, and galactic particles, *IEEE Trans. Geosci. Remote Sens.*, *31*, 565–571.

Goldstein, J., S. Kanekal, D. N. Baker, and B. R. Sandel (2005), Dynamic relationship between the outer radiation belt and the plasmapause during March–May 2001, *Geophys. Res. Lett.*, *32*, L15104, doi:10.1029/2005GL023431.

Gosling, J. T., A. J. Hundausen, V. Pizzo, and J. R. Asbridge (1972), Compressions and rarefactions in the solar wind: Vela 3, *J. Geophys. Res.*, *77*, 5442–5454.

Hruska, A., and J. Hruska (1989), Solar wind modulation of the auroral zone geomagnetic activity when the interplanetary magnetic field has a strong northward component, *J. Geophys. Res.*, *94*, 5479–5484.

Iyemori, T. H. Maeda, and T. Kamei (1979), Impulse response of geomagnetic indices to interplanetary magnetic field, *J. Geomagn. Geoelectr.*, *31*, 1–9.

Kanekal, S. G., D. N. Baker, and J. B. Blake (2001), Multisatellite measurements of relativistic electrons: Global coherence, *J. Geophys. Res.*, *106*, 29,721–29,732.

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

Klimas, A. J., D. Vassiliadis, and D. N. Baker (1997), Data-derived analogs of the magnetospheric dynamics, *J. Geophys. Res.*, *107*, 26,993–27,009.

Ljung, L. (1999), *System Identification: Theory for the User*, Prentice Hall, Upper Saddle River, N. J.

McPherron, R. L., D. N. Baker, and L. F. Bargatze (1986), Linear filters as a method of real time prediction of geomagnetic activity, in *Solar Wind-Magnetosphere Coupling*, pp. 85–92, edited by Y. Kamide and J. A. Slavin, Terra Sci., Tokyo.

Nagai, T. (1988), "Space Weather Forecast": Prediction of relativistic electron intensity at synchronous orbit, *Geophys. Res. Lett.*, *15*, 425–428.

Nelles, O. (2001), *Nonlinear System Identification*, Springer, New York.

O'Brien, T. P., R. L. McPherron, D. Sornette, G. D. Reeves, R. Friedel, and H. J. Singer (2001), Which magnetic storms produce relativistic electrons at geosynchronous orbit?, *J. Geophys. Res.*, *106*, 15,533–15,544.

Paulikas, G. A., and J. B. Blake (1979), Effects of the solar wind on magnetospheric dynamics: energetic electrons at the synchronous orbit, in *Qualitative Modeling of Magnetospheric Processes*, *Geophys. Monogr. Ser.*, pp. 180–202.

Perreault, P., and S.-I. Akasofu (1978), A study of geomagnetic storms, *Geophys. J. R. Astron. Soc.*, *54*, 547–573.

Pindyck, R. S. and D. L. Rubinfeld (1991), *Econometric Models and Economic Forecasts*, McGraw-Hill, New York.

Pizzo, V. (1983), Quasi-Steady Solar Wind Dynamics, in *Solar Wind 5, NASA Conf. Publ., CP-2280*, pp. 675–691, edited by M. Neugebauer.

Rigler, E. J., D. N. Baker, R. S. Weigel, D. Vassiliadis, and A. J. Klimas (2004), Adaptive linear prediction of radiation belt electrons using the Kalman filter, *Space Weather*, *2*, S03003, doi:10.1029/2003SW000036.

Rigler, E. J., D. N. Baker, R. S. Weigel, and D. Vassiliadis (2005), Solar wind-driven electron radiation belt response functions at 100-min time scales, *Adv. Space Res.*, *36*, 2401–2406, doi:10:1016/j.asr.2003.09.070.

Trattner, K. J., and H. O. Rucker (1990), Linear prediction theory in studies of solar wind-magnetosphere coupling, *Anales Geophys.*, *8*, 733–738.

Vassiliadis, D., D. N. Baker, A. J. Klimas, and D. A. Roberts (1995), A description of the solar wind-magnetosphere coupling based on nonlinear filters, *J. Geophys. Res.*, *100*, 3495–3512.

Vassiliadis, D., A. J. Klimas, S. G. Kanekal, D. N. Baker, and R. S. Weigel (2002), Long-term-average solar cycle, and seasonal response of magnetospheric energetic electrons to the solar wind speed, *J. Geophys. Res.*, *107*(A11), 1383, doi:10.1029/2001JA000506.

Vassiliadis, D., R. S. Weigel, D. N. Baker, S. G. Kanekal, and A. J. Klimas (2004), Probing the solar wind-inner magnetospheric coupling: Validation of relativistic electron flux models, *J. Atmos. Sol.-Terr. Phys.*, *66*, 1399–1409, doi:10.1016/j.jast.2004.03.025.

Vassiliadis, D., S. F. Fung, and A. J. Klimas (2005), Solar, interplanetary, and magnetospheric parameters for the radiation belt energetic electron flux, *J. Geophys. Res.*, *110*, A04201, doi:10.1029/2004JA010443.

Vasyliunas, V. M., J. R. Kan, G. L. Siscoe, and S.-I. Akasofu (1982), Scaling relations governing magnetospheric energy transfer, *Planet. Space Sci.*, *30*, 359–365.

Wiener, N. (1949), *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*, John Wiley and Sons, Hoboken, N. J.

————————————

D. N. Baker, Laboratory for Atmospheric and Space Physics, University of Colorado, 1234 Innovation Drive, Boulder, CO 80303, USA.

E. J. Rigler and M. Wiltberger, High Altitude Observatory, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA. (jrigler@hao.ucar.edu)