

Journal of Geophysical Research: Space Physics

RESEARCH ARTICLE

10.1002/2016JA022711

Key Points:

- Given V_{sw} , the geosynchronous MeV electron response lag time to n_{sw} is < 24 h
- Transfer entropy can be used to detect changes in the solar wind-radiation belt system dynamics
- V_{sw} anticorrelates with n_{sw} with a lag of 14–17 h, but the lag time has yearly variation

Correspondence to:

S. Wing,
simon.wing@jhuapl.edu

Citation:

Wing, S., J. R. Johnson, E. Camporeale, and G. D. Reeves (2016), Information theoretical approach to discovering solar wind drivers of the outer radiation belt, *J. Geophys. Res. Space Physics*, 121, doi:10.1002/2016JA022711.

Received 16 MAR 2016

Accepted 25 JUL 2016

Accepted article online 29 JUL 2016

Information theoretical approach to discovering solar wind drivers of the outer radiation belt

Simon Wing¹, Jay R. Johnson^{2,3}, Enrico Camporeale⁴, and Geoffrey D. Reeves⁵

¹The Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland, USA, ²Plasma Physics Laboratory, Princeton University, Princeton, New Jersey, USA, ³Department of Engineering and Computer Science, Andrews University, Berrien Springs, Michigan, USA, ⁴Center for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands, ⁵Space Science and Applications Group, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract The solar wind-magnetosphere system is nonlinear. The solar wind drivers of geosynchronous electrons with energy range of 1.8–3.5 MeV are investigated using mutual information, conditional mutual information (CMI), and transfer entropy (TE). These information theoretical tools can establish linear and nonlinear relationships as well as information transfer. The information transfer from solar wind velocity (V_{sw}) to geosynchronous MeV electron flux (J_e) peaks with a lag time of 2 days. As previously reported, J_e is anticorrelated with solar wind density (n_{sw}) with a lag of 1 day. However, this lag time and anticorrelation can be attributed at least partly to the $J_e(t+2$ days) correlation with $V_{sw}(t)$ and $n_{sw}(t+1$ day) anticorrelation with $V_{sw}(t)$. Analyses of solar wind driving of the magnetosphere need to consider the large lag times, up to 3 days, in the (V_{sw} , n_{sw}) anticorrelation. Using CMI to remove the effects of V_{sw} , the response of J_e to n_{sw} is 30% smaller and has a lag time < 24 h, suggesting that the MeV electron loss mechanism due to n_{sw} or solar wind dynamic pressure has to start operating in < 24 h. n_{sw} transfers about 36% as much information as V_{sw} (the primary driver) to J_e . Nonstationarity in the system dynamics is investigated using windowed TE. When the data are ordered according to transfer entropy value, it is possible to understand details of the triangle distribution that has been identified between $J_e(t+2$ days) versus $V_{sw}(t)$.

1. Introduction

The Earth's radiation belt is inhabited by electrons having energies of a few hundreds of keV or higher, which are often referred to as "killer electrons" because of the potential damages when they encounter satellites. For example, the radiation belt electrons with energies of a few MeV or higher can penetrate deep into spacecraft components, while those with energies lower than 1 MeV can lodge on the surface of the spacecraft bodies, leading to devastating electrical discharges.

To explain the origin of the MeV electrons in the radiation belt, studies suggested that the storm and substorm injection process from plasma sheet into the inner magnetosphere accelerates low-energy (e.g., a few keV) electrons to a few hundred keV. Once in the inner magnetosphere, electrons interact with ultra low frequency (ULF) waves [e.g., Elkington *et al.*, 1999; Rostoker *et al.*, 1998; Ukhorskiy *et al.*, 2005; Mathie and Mann, 2000, 2001], very low frequency (VLF) waves [e.g., Summers *et al.*, 1998; Omura *et al.*, 2007; Thorne, 2010; Simms *et al.*, 2015; Camporeale, 2015; Camporeale and Zimbardo, 2015], or magnetosonic waves [e.g., Horne *et al.*, 2007; Shprits *et al.*, 2008], which can energize electrons to MeV energy range.

Enhancements of ULF waves can be associated with increased occurrences of Kelvin-Helmholtz instability (KHI) along the magnetopause flanks due to large solar wind velocity (V_{sw}) [e.g., Johnson *et al.*, 2014; Engebretson *et al.*, 1998; Vennestrøm, 1999]. Indeed, studies have shown that V_{sw} is a dominant, if not the most dominant, driver of geosynchronous relativistic electron fluxes (herein J_e refers to geosynchronous MeV electron energy flux) [e.g., Paulikas and Blake, 1979; Baker *et al.*, 1990; Li *et al.*, 2001, 2005; Vassiliadis *et al.*, 2005; Ukhorskiy *et al.*, 2004; Rigler *et al.*, 2007; Kellerman and Shprits, 2012; Reeves *et al.*, 2011]. However, the geosynchronous electron response to V_{sw} has a lag time that is energy dependent. For example, Li *et al.* [2005] finds that the average lag times of 50 keV to 1.1 MeV electrons approximately range from –4 to 25 h, respectively. For MeV electrons, a lag time of about 2 days has been consistently observed in many studies [e.g., Baker *et al.*, 1990; Vassiliadis *et al.*, 2005; Reeves *et al.*, 2011; Balikhin *et al.*, 2011; Lyatsky and Khazanov, 2008a], suggesting the time scale needed to accelerate electrons to MeV energy range.

In contrast to V_{sw} , which correlates with J_e , solar wind density (n_{sw}) anticorrelates with J_e for reasons that are not entirely clear [e.g., Lyatsky and Khazanov, 2008a; Kellerman and Shprits, 2012]. Li *et al.* [2005] suggests that an increase in n_{sw} increases solar wind dynamic pressure (P_{dyn}), which, in turn, pushes the magnetopause inward, leading to electron losses. However, Lyatsky and Khazanov [2008a] argues that the poor correlation between P_{dyn} and J_e suggests that compression of the magnetosphere is probably not the main factor. Moreover, the effectiveness of n_{sw} at influencing J_e is also not clear. Some studies found that n_{sw} has weaker effects than V_{sw} on J_e [e.g., Vassiliadis *et al.*, 2005; Rigler *et al.*, 2007; Kellerman and Shprits, 2012]. However, Balikhin *et al.* [2011] finds that J_e has the strongest dependence on n_{sw} with a lag of 1 day.

The interpretation of the relationship between n_{sw} and J_e is complicated by the anticorrelation between V_{sw} and n_{sw} [e.g., Hundhausen *et al.*, 1970]. Because J_e and V_{sw} are correlated, the anticorrelation between J_e and n_{sw} could simply be coincidence. A few studies attempted to separate the effects of n_{sw} from V_{sw} by using methods that bin the data into small intervals of V_{sw} and n_{sw} [e.g., Lyatsky and Khazanov, 2008a]. This type of analysis has offered insights into solar wind driving of J_e , but it does not address the question of how much additional information n_{sw} provides to J_e , given V_{sw} and vice versa. Other studies showed that interplanetary magnetic field (IMF) and other solar wind parameters can also contribute to J_e variations [e.g., Balikhin *et al.*, 2011; Rigler *et al.*, 2007; Vassiliadis *et al.*, 2005; Li *et al.*, 2005; Onsager *et al.*, 2007; Simms *et al.*, 2014], but presently, it is not entirely clear quantitatively given the main driver, e.g., V_{sw} (or n_{sw}), how much additional information these parameters provide to J_e . This knowledge can help radiation belt modelers decide what input parameters to consider for their models.

The solar wind-magnetosphere and solar wind-radiation belt systems have been shown to be nonlinear [e.g., Wing *et al.*, 2005a; Johnson and Wing, 2005; Reeves *et al.*, 2011; Kellerman and Shprits, 2012]. An example is presented in Figure 1, which plots $\log J_e(t+\tau)$ versus $V_{sw}(t)$ for $\tau=0, 1, 2$, and 7 days. The figure, which is similar to Figure 9 in Reeves *et al.* [2011], shows that the relationship between J_e and V_{sw} is nonlinear. For nonlinear system, qualitative linear correlational analysis can be misleading [e.g., Balikhin *et al.*, 2010, 2011]. Moreover, correlational analysis cannot establish causalities. On the other hand, as described below, information theory can help identify nonlinearities in the system and information transfer from input to output parameters.

Although information theoretical tools are still considered novel in space physics and space weather, a few studies have successfully applied these tools to solve problems in these fields. For example, Johnson and Wing [2005] applied mutual information and cumulant-based analysis to discover the nonlinear dependencies inherent in the K_p time series. Materassi *et al.* [2011] applied mutual information to characterize the influence and timing of solar wind forcing of the ionospheric scintillations. De Michelis *et al.* [2011] applied transfer entropy to establish transfer of information from AL to SYM-H indices on a time scale shorter than 10 h. More recently, Johnson and Wing [2014] applied conditional redundancy to examine the roles of internal versus external triggering in substorms.

In the present study, we investigate further the linear and nonlinear relationships between solar wind parameters and geosynchronous relativistic electrons using tools that are based on information theory. In particular, we apply mutual information [e.g., Li, 1990; Tsonis, 2001], conditional mutual information [e.g., Wyner, 1978], and transfer entropy [e.g., Schreiber, 2000] to determine the solar wind drivers of J_e and to quantify how much information is transferred from solar wind parameters to J_e .

The scatterplots of J_e versus V_{sw} in Figures 1a–1c look like a triangle, which Reeves *et al.* [2011] refers to as the triangle distribution. The mystifying part of the triangle distribution is that high values of J_e are observed for all V_{sw} conditions. The present study probes deeper at this triangle distribution using information theory.

2. Data Set

Most studies of geosynchronous MeV electrons have been performed with data having 1 day resolution [e.g., Reeves *et al.*, 2011; Balikhin *et al.*, 2011; Kellerman and Shprits, 2012]. As pointed by Reeves *et al.* [2011], because of the asymmetry of the geomagnetic field along the geosynchronous orbit, geosynchronous electron fluxes exhibit a diurnal or magnetic local time (MLT) variation as well as latitude-longitude dependence. However, these effects are reduced in daily resolution data.

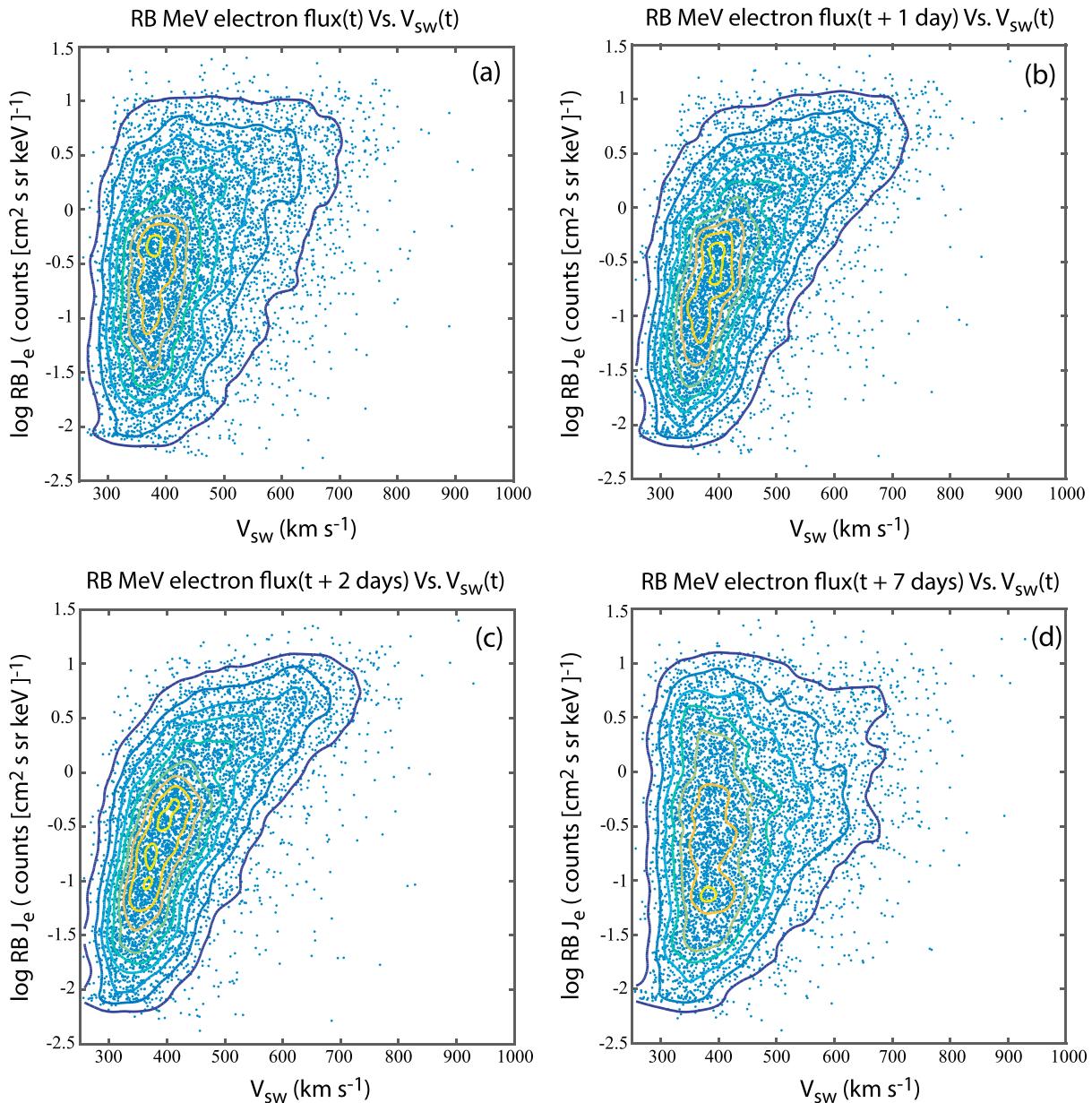


Figure 1. (a–d) Scatterplots of $\log J_e(t + \tau)$ versus $V_{sw}(t)$ for $\tau = 0, 1, 2$, and 7 days. The data points are overlaid with density contours showing the nonlinear trends. The panels show that J_e has dependence on V_{sw} for $\tau = 0, 1$, and 2 days and the dependence is strongest for $\tau = 2$ days. In Figure 1d at large τ , e.g., $\tau = 7$ days, J_e dependence on V_{sw} is very weak. The triangle distribution [Reeves et al., 2011] can be seen in Figures 1a–1c. This is essentially the same as Figure 9 in Reeves et al. [2011], except that density contours are drawn and Figure 1d plots $\tau = 7$ days instead of $\tau = 3$ days.

The present study uses the same data set in Reeves et al. [2011]. The data and format description can be found at <ftp://ftop.agu.org/apend/ja/2010ja015735>. This data set contains daily averages of electron fluxes obtained from energetic sensor for particles [Meier et al., 1996] and synchronous orbit particle analyzer [Belian et al., 1992] on board of all seven Los Alamos National Laboratory (LANL) geosynchronous satellites from 22 September 1989 to 31 December 2009. The present study only examines the fluxes of electrons with energy range of 1.8–3.5 MeV (which is referred herein as J_e). A detailed description of the data set and its processing are given in Reeves et al. [2011]. The daily and hourly averaged solar wind data 1989–2009 come from OMNI data set provided by NASA (<http://omniweb.gsfc.nasa.gov/>). The LANL and solar wind data are merged. The LANL data set has 7187 data points (days of data), out of which, 6438 data points have simultaneous solar wind observations.

3. Mutual Information, Conditional Mutual Information, and Transfer Entropy

Dependency is a key discriminating statistic that is commonly used to understand how systems operate. The standard tool used to identify dependency is cross correlation. Considering two variables, x and y , the correlation analysis essentially tries to fit the data to a 2-D Gaussian cloud, where the nature of the correlation is determined by the slope and the strength of correlation is determined by the width of the cloud perpendicular to the slope.

By nature, the response of the radiation belts to solar wind variables is nonlinear [Reeves *et al.*, 2011; Kellerman and Shprits, 2012] as evidenced by the triangle distribution in J_e versus V_{sw} seen in Figures 1a–1c. Such a distribution is not well described by a Gaussian cloud of points and is not well characterized by a slope. For such distributions, it is better to use a statistical-based measure such as mutual information (MI) [Tsonis, 2001; Li, 1990; Darbellay and Vajda, 1999]. Mutual information between two variables, x and y , compares the uncertainty of measuring variables jointly with the uncertainty of measuring the two variables independently. The uncertainty is measured by the entropy. In order to construct the entropies, it is necessary to obtain the probability distribution functions, which in this study are obtained from histograms of the data based on discretization of the variables (i.e., bins).

Suppose that two variables, x and y , are binned so that they take on discrete values, \hat{x} and \hat{y} , where

$$x \in \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n\} \equiv \aleph_1; \quad y \in \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m\} \equiv \aleph_2. \quad (1)$$

The variables may be thought of as letters in alphabets \aleph_1 and \aleph_2 , which have n and m letters, respectively. The extracted data can be considered as sequences of letters. The entropy associated with each of the variables is defined as

$$H(x) = - \sum_{\aleph_1} p(\hat{x}) \log p(\hat{x}); \quad H(y) = - \sum_{\aleph_2} p(\hat{y}) \log p(\hat{y}) \quad (2)$$

where $p(\hat{x})$ is the probability of finding the word \hat{x} in the set of x data and $p(\hat{y})$ is the probability of finding word \hat{y} in the set of y data. To examine the relationship between the variables, we extract the word combinations (\hat{x}, \hat{y}) from the data set. The joint entropy is defined by

$$H(x, y) = - \sum_{\aleph_1 \aleph_2} p(\hat{x}, \hat{y}) \log p(\hat{x}, \hat{y}) \quad (3)$$

where $p(\hat{x}, \hat{y})$ is the probability of finding the word combination (\hat{x}, \hat{y}) in the set of (x, y) data. The mutual information is then defined as

$$MI(x, y) = H(x) + H(y) - H(x, y) \quad (4)$$

In the case of Gaussian distributed data, the mutual information can be related to the correlation function; however, it also includes higher-order correlations that are not detected by the correlation function. Hence, MI is a better measure of dependency for variables having a nonlinear relationship [Johnson and Wing, 2005].

While MI is useful to identify nonlinear dependence between two variables, it does not provide information about whether the dependence is causal or coincidental. Herein, we use the working definition that if there is a transfer of information from x to y , then x causes y . In this case, it is useful to consider conditional dependency with respect to a conditioner variable z that takes on discrete values, $\hat{z} \in \{z_1, z_2, \dots, z_n\} \equiv \aleph_3$. The conditional mutual information [Wyner, 1978]

$$CMI(x, y | z) = \sum_{\aleph_1 \aleph_2 \aleph_3} p(\hat{x}, \hat{y}, \hat{z}) \log \frac{p(\hat{x}, \hat{y} | \hat{z})}{p(\hat{x} | \hat{z}) p(\hat{y} | \hat{z})} = H(x, z) + H(y, z) - H(x, y, z) - H(z) \quad (5)$$

determines the mutual information between x and y given that z is known where $p(\hat{x} | \hat{z})$ is the probability of finding the word \hat{x} in the set of x -data given \hat{z} . In the case where z is unrelated, $CMI(x, y | z) = MI(x, y)$, but in the case that x or y is known based on z , then $CMI(x, y | z) = 0$. CMI therefore provides a way to determine how much additional information is known given another variable. CMI can be seen as a special case of the more general conditional redundancy that allows the variable z to be a vector [e.g., Prichard and Theiler, 1995; Johnson and Wing, 2014].

A common method to establish causal-relationships between two time series, e.g., $[x_t]$ and $[y_t]$, is to use a time-shifted correlation function [e.g., Borovsky *et al.*, 1998]

$$r(\tau) = \frac{x_t y_{t+\tau} - \bar{x} \bar{y}}{\sqrt{x^2 - \bar{x}^2} \sqrt{y^2 - \bar{y}^2}} \quad (6)$$

where r = correlation coefficient and τ = lag time. The results of this type of analysis may not be particularly clear when the correlation function has multiple peaks or there is not an obvious asymmetry. Additionally, correlational analysis only detects linear correlations. If the feedback involves nonlinear processes, its usefulness may be seriously limited.

Alternatively, time-shifted mutual information, $MI(x(t), y(t+\tau))$, can be used to detect causality in nonlinear systems, but this, too, suffers from the same problems as time-shifted correlation when it has multiple peaks and long-range correlations.

A better choice for studying causality is the one-sided transfer entropy [Schreiber, 2000]

$$TE_{x \rightarrow y}(\tau) = \sum_t p(y_{t+\tau}, yp_t, x_t) \log \left(\frac{p(y_{t+\tau} | yp_t, x_t)}{p(y_{t+\tau} | yp_t)} \right) \quad (7)$$

where $yp_t = [y_t, y_{t-\Delta}, \dots, y_{t-k\Delta}]$, $k+1$ = dimensionality of the system, and Δ = first minimum in MI. Transfer entropy (TE) can be considered as a specialized case of conditional mutual information:

$$TE_{x \rightarrow y}(\tau) = CMI(y(t+\tau), x(t) | yp(t)) \quad (8)$$

where $yp(t) = [y(t), y(t-\Delta), \dots, y(t-k\Delta)]$. The transfer entropy can be considered as a conditional mutual information that detects how much average information is contained in an input, x , about the next state of a system, y , that is not contained in the past history, yp , of the system [Prokopenko *et al.*, 2013]. In the absence of information flow from x to y , $TE(x \rightarrow y)$ vanishes. Also, unlike correlational analysis and mutual information, transfer entropy is directional, $TE(x \rightarrow y) \neq TE(y \rightarrow x)$. The transfer entropy accounts for static internal correlations, which can be used to determine whether x and y are driven by a common driver or whether x drives y or y drives x .

4. Applying Information Theory to Radiation Belt MeV Electron Data

4.1. Radiation Belt MeV Electron Flux Versus V_{sw}

A good starting point for our analysis is Figure 9 in Reeves *et al.* [2011], which is replotted in Figure 1 with some modifications. As in Reeves *et al.* [2011], the present paper uses the convention that V_{sw} is positive in the antisunward direction. Consistent with Reeves *et al.* [2011], Figure 1 shows that (1) the correlation is best at $\tau = 2$ days (Figure 1c); (2) the relationship between $\log J_e$ and V_{sw} is nonlinear, which can be seen more clearly in the data density contours; and (3) the data point distribution looks like a triangle. This so called triangle distribution is discussed further in section 5.4.

The blue curve in Figure 2a shows the correlation coefficient of $[\log J_e(t+\tau), V_{sw}(t)]$. Note that herein, unless otherwise stated, all linear and nonlinear analyses performed with J_e uses $\log J_e$ values. Figure 2a shows that the linear correlation coefficient peaks at $\tau_{max} = 2$ days with $r = 0.63$. There is a smaller peak at $\tau = 29$ days ($r = 0.42$), which can be attributed to the 27 day synodic solar rotation. The red curve shows the correlation coefficient of $[J_e(t), V_{sw}(t+\tau)]$. The red curve has a small peak at $\tau = 25$ days ($r = 0.39$) because of the 27 day solar rotation. That is, $J_e(t)$ correlates best with $V_{sw}(t-2$ days), but $V_{sw}(t-2$ days) correlates with $V_{sw}([t-2$ days] + 27 days). Because of the large number of data points ($n > 5772$), the three peak correlation coefficients are highly significant with $P < 0.01$ (the probability of two random variables giving a correlation coefficient as large as r is < 0.01).

However, the relationship between J_e and V_{sw} is nonlinear, and hence, linear cross correlation may not capture the full extent of the relationship, as described in section 3. Although the correlation coefficient may give some indication about the sign and strength of the relationship, it is not quantitatively precise [Reeves *et al.*, 2011]. In order to take into account the nonlinearities in the relationship, we apply mutual information and transfer entropy.

Figure 2b plots the mutual information (MI) of $[J_e(t+\tau), V_{sw}(t)]$ (blue curve) and $[J_e(t), V_{sw}(t+\tau)]$ (red curve) as well as transfer entropy (TE) of $[J_e(t+\tau), V_{sw}(t)]$ (yellow curve) and $[J_e(t), V_{sw}(t+\tau)]$ (purple curve). For simplicity,

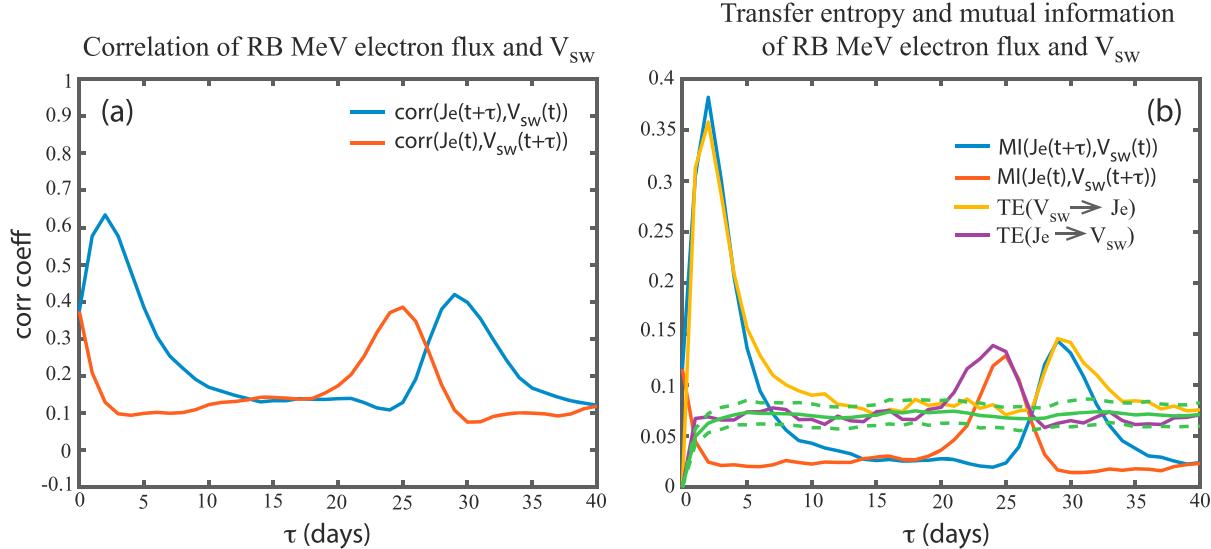


Figure 2. (a) Correlation coefficient of $[J_e(t + \tau), V_{sw}(t)]$ (blue) and $[J_e(t), V_{sw}(t + \tau)]$ (red). (b) $\text{MI}[J_e(t + \tau), V_{sw}(t)]$ (blue), $\text{MI}[J_e(t), V_{sw}(t + \tau)]$ (red), $\text{TE}[J_e(t + \tau), V_{sw}(t)]$ (yellow), and $\text{TE}[J_e(t), V_{sw}(t + \tau)]$ (purple). The green solid and dashed curves are the mean and the 3σ from the mean of the noise. The transfer of information from V_{sw} to J_e ($V_{sw} \rightarrow J_e$) peaks at $\tau_{\max} = 2$ days.

we assume $k=0$ in equations (7) and (8). As described in section 3, mutual information provides a measure of nonlinear correlation between the two parameters, while the transfer entropy provides a measure of transfer of information from one variable to another. Herein, we adopt the convention that forward direction is the direction for information transfer from a solar wind parameter to J_e and backward direction is the opposite. So blue and yellow curves plot the forward MI and TE, while red and purple curves plot the backward MI and TE, respectively.

The forward and backward mutual information peak at $\tau_{\max} = 2$ and 25 days, respectively. Also, the forward mutual information has a secondary peak at $\tau = 29$ days. The results suggest that the response may be dominated by the linear dynamics in this case, although, in general, this is not necessarily the case. The forward transfer entropy from V_{sw} to J_e , $\text{TE}(V_{sw} \rightarrow J_e)$, peaks at $\tau_{\max} = 2$ days (yellow curve), suggesting that the transfer of information from V_{sw} to geosynchronous MeV electrons has a 2 day delay. Similar to MI and correlational analysis, $\text{TE}(V_{sw} \rightarrow J_e)$ has a small peak at $\tau = 29$ days. The backward TE, $\text{TE}(J_e \rightarrow V_{sw})$, has a peak at $\tau_{\max} = 24$ days (purple). The backward peak in the transfer entropy raises some questions about how well TE is able to eliminate the self-correlation of the solar wind. The self-correlations are probably better eliminated by using a two element vector of V_{sw} that includes $V_{sw}(t)$ and $V_{sw}(t - \Delta)$ to capture more of the dynamics of V_{sw} .

In order to get a measure of the significance of $\text{TE}(V_{sw} \rightarrow J_e)$, we calculate noise = $\text{TE}[\text{sur}(V_{sw}) \rightarrow J_e]$ where $\text{sur}(V_{sw})$ is the surrogate data of V_{sw} , which is obtained by randomly permuting the order of the time series array V_{sw} . The mean and standard deviation of the noise are calculated from an ensemble of 100 random permutations of $\text{TE}[\text{sur}(V_{sw}) \rightarrow J_e]$. The mean noise and 3σ (standard deviation) from the mean noise are plotted with green solid and dashed curves, respectively, in Figure 2b. The maximum TE, $\text{TE}[J_e(t + 2 \text{ days}), V_{sw}(t)]$ has peak information transfer (it_{\max}) = 0.30, signal-to-noise ratio (snr) = 5.7, and significance = 94σ where it_{\max} = peak – mean noise, snr = peak/mean noise, and significance = $it_{\max}/\sigma(\text{noise})$. From the snr, it_{\max} , and significance, we conclude that there is a significant transfer of information from V_{sw} to J_e with a 2 day delay. Note that the linear correlation, MI, and TE analyses are consistent with the previous studies [e.g., Baker et al., 1990; Vassiliadis et al., 2005; Reeves et al., 2011; Balikhin et al., 2011; Lyatsky and Khazanov, 2008a].

Based on random selections of about 100 noise ensembles taken at various τ in our TE and CMI calculations throughout this study, the noise (surrogates) statistics are reasonably characterized by a normal distribution. This justifies our definition of significance in terms of the mean and σ characterizing the surrogate distribution. The present study finds that using 100 surrogates in each ensemble is adequate to characterize its

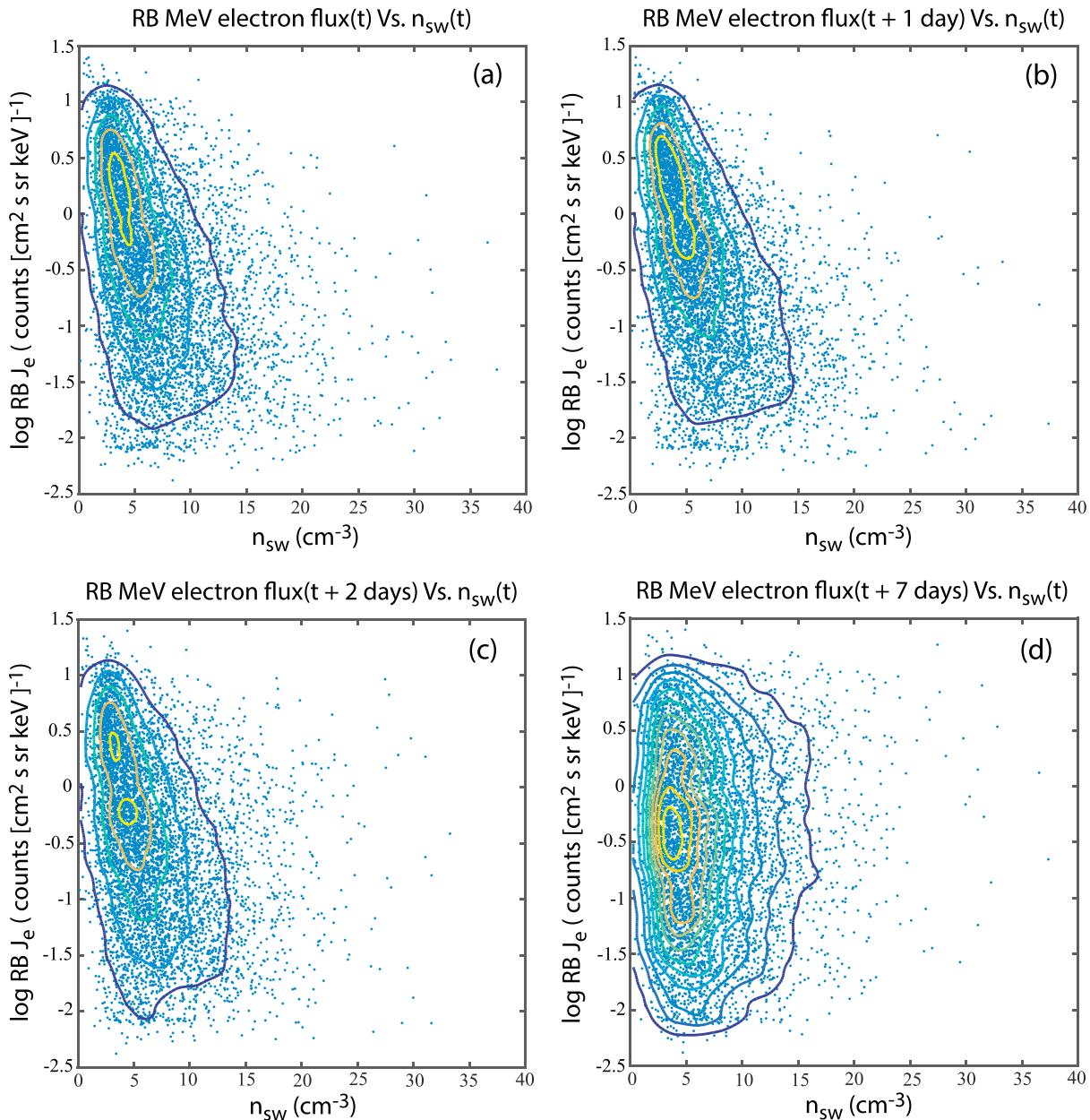


Figure 3. J_e anticorrelates with n_{SW} . (a–d) Scatterplots of $\log J_e(t + \tau)$ versus $n_{\text{SW}}(t)$ for $\tau = 0, 1, 2$, and 7 days. The data points are overlaid with density contours, which show the trends. The panels show that J_e has dependence on n_{SW} for $\tau = 0, 1$, and 2 days. The slope of the contours is most negative for $\tau = 1$ day, suggesting strongest dependence on n_{SW} at $\tau = 1$ day. In Figure 3d at large τ , e.g., $\tau = 7$ days, J_e dependence on n_{SW} vanishes.

distribution (mean and σ), especially with the choice of mean + 3σ as a threshold for noise level. Using a larger number of surrogates in each ensemble would result in a smoother distribution that better approximates a normal distribution, but it does not significantly change the mean and σ presented in the present study because the surrogates are drawn from a normal distribution. The mean and σ of a normal distribution are nearly independent of the number of surrogates as long as we take a large enough sample size.

The TE($V_{\text{sw}} \rightarrow J_e$) (yellow) curve shows that V_{sw} has little influence on the geosynchronous MeV electrons after a delay of 7–10 days, which is essentially the prediction or information horizon. This result is consistent with Figure 1d, which shows poor correlation in $\log J_e(t + 7 \text{ days})$ versus V_{sw} distribution.

In applying our information theoretical tools, the number of bins (n_b) needs to be chosen appropriately. Sturges [1926] proposes that for a normal distribution, optimal $n_b = \log_2(n) + 1$ and bin width (w) = range/ n_b , where

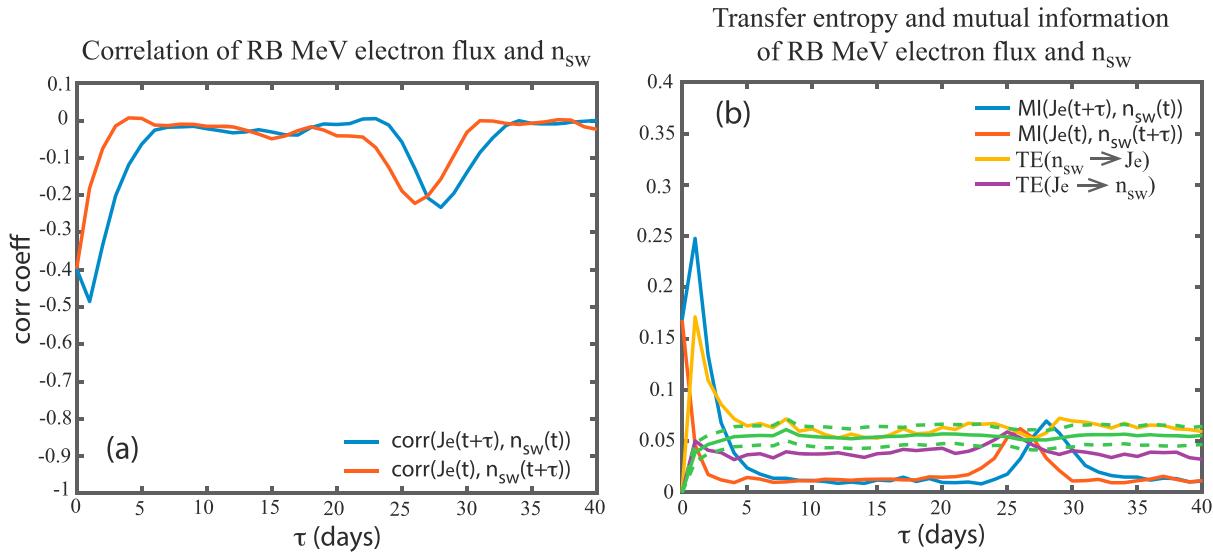


Figure 4. (a) Correlation coefficient of $[J_e(t + \tau), n_{sw}(t)]$ (blue) and $[J_e(t), n_{sw}(t + \tau)]$ (red). (b) MI $[J_e(t + \tau), n_{sw}(t)]$ (blue), MI $[J_e(t), n_{sw}(t + \tau)]$ (red), TE $[J_e(t + \tau), n_{sw}(t)]$ (yellow), and TE $[J_e(t), n_{sw}(t + \tau)]$ (purple). The green solid and dashed curves are the mean and 3σ from the mean of the noise. The transfer of information from n_{sw} to J_e ($n_{sw} \rightarrow J_e$) peaks at $\tau_{\max} = 1$ day.

n = number of points in the data set and range = maximum value – minimum value of the points. In practice, there is usually a range of n_b that would work. Using Sturges [1926] formula, with roughly 6400 points, $n_b \sim 13.6$. For the present study, we find that $n_b = 10$ to 15 would work well. Having too few bins would lump too many points into the same bin, leading to loss of information. Conversely, having too many bins would leave many bins with 0 or a few number of points, which also leads to loss of information. For the present study, we choose $n_b = 10$.

4.2. Radiation Belt MeV Electron Flux Versus n_{sw}

We repeat the above analyses for J_e versus n_{sw} . Figure 3 plots $J_e(t + \tau)$ versus n_{sw} for $\tau = 0, 1, 2$, and 7 days. It shows that (1) J_e anticorrelates with n_{sw} and (2) $J_e(t + 1$ day) versus n_{sw} (Figure 3b) has the most negative slope, suggesting the best anticorrelation. The anticorrelation is shown more clearly in Figure 4a, which plots $\text{corr}[J_e(t + \tau), n_{sw}(t)]$ (blue curve, forward direction) and $\text{corr}[J_e(t), n_{sw}(t + \tau)]$ (red curve, backward direction). The blue curve shows $\tau_{\min} = 1$ day ($r = -0.40$) and a secondary minimum at $\tau = 28$ days ($r = -0.23$). The latter can be attributed to solar rotation. The red curve shows $\tau_{\min} = 26$ days ($r = -0.22$), which again can be attributed to solar rotation. Although the correlation coefficients are smaller, all three of them are still highly significant ($P < 0.01$) due to the large number of data points.

Figure 4b is similar to Figure 2b, except that it shows MI and TE for (J_e, n_{sw}) . Forward MI and TE, both have $\tau_{\max} = 1$ day, which is consistent with Figure 4a. Note that unlike correlational analysis, MI and TE only give positive values for both correlations and anticorrelations. The forward TE, $\text{TE}(n_{sw} \rightarrow J_e)$, is not as large as $\text{TE}(V_{sw} \rightarrow J_e)$ shown in Figure 2b, but TE at $\tau = 1$ day is still significant. $\text{TE}[n_{sw}(t) \rightarrow J_e(t + 1 \text{ day})]$ has $it_{\max} = 0.13$, $snr = 4.4$, and significance = 42σ . This result suggests that there is a transfer of information from n_{sw} to geosynchronous MeV electrons with 1 day delay.

Besides lag time and peak size, there are other differences between Figures 2b and 4b that are noteworthy. First, in the forward direction, there is little information transfer from n_{sw} to J_e after 4 days. That is, $\text{TE}[J_e(t + \tau), n_{sw}(t)]$ for $\tau > 4$ days is in the noise level, which is consistent with Figure 3d for $\tau = 7$ days. Second, the TE peaks due to solar rotation in the forward and backward directions are a lot smaller and, in fact, are at the noise level. This result suggests that unlike V_{sw} , there is little information transfer from the n_{sw} to J_e one solar rotation later. Practically, there is no information flow in the backward direction, from J_e to n_{sw} at any lag because $\text{TE}[J_e(t), n_{sw}(t + \tau)]$ (purple curve) is small and within noise level for all τ . In contrast, $\text{corr}[J_e(t), n_{sw}(t + \tau)]$ and $\text{MI}[J_e(t), n_{sw}(t + \tau)]$ are still significant for small τ and $\text{corr}[J_e(t), n_{sw}(t + \tau)]$ has a significant

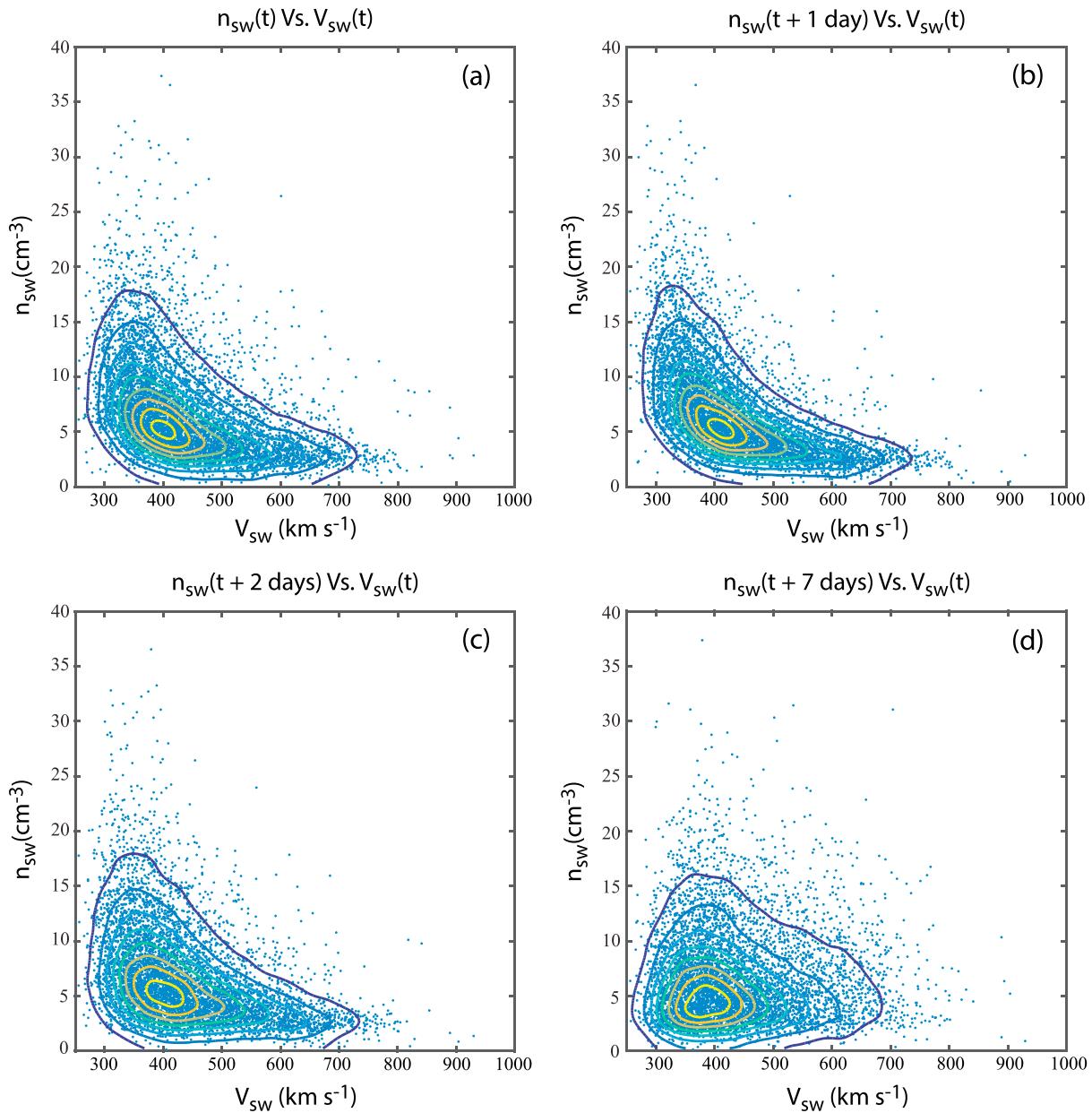


Figure 5. n_{sw} anticorrelates with V_{sw} . (a–d) Scatterplots of $n_{sw}(t + \tau)$ versus $V_{sw}(t)$ for $\tau = 0, 1, 2$, and 7 days. The data points are overlain with density contours showing strong nonlinear trends. The panels show that n_{sw} has dependence on V_{sw} for $\tau = 0, 1$, and 2 days and the dependence is strongest for $\tau = 1$ day. In Figure 5d at large τ , e.g., $\tau = 7$ days, n_{sw} dependence on V_{sw} is very weak.

value that is due to solar rotation. This illustrates that TE does not always give the same results as correlation and MI.

4.3. Anticorrelation of V_{sw} and n_{sw} and Its Effect on Radiation Belt

In section 4.1, we show that $J_e(t + 2 \text{ days})$ linearly and nonlinearly correlate with $V_{sw}(t)$. It is well known that n_{sw} anticorrelates with V_{sw} [e.g., Hundhausen et al., 1970]. However, if the anticorrelation were instantaneous, that is, $n_{sw}(t + 0 \text{ day})$ anticorrelates with $V_{sw}(t)$, then we would expect that $J_e(t + 2 \text{ days})$ to anticorrelate with $n_{sw}(t)$. However, in section 4.2, we show that $J_e(t + 1 \text{ day})$ linearly and nonlinearly anticorrelate with $n_{sw}(t)$, suggesting that other factors may be involved.

To investigate this, we plot in Figure 5 $n_{sw}(t + \tau)$ versus $V_{sw}(t)$ in the same format as in Figures 1 and 3. Figure 5 suggests that n_{sw} anticorrelates with V_{sw} and the relationship is not linear.

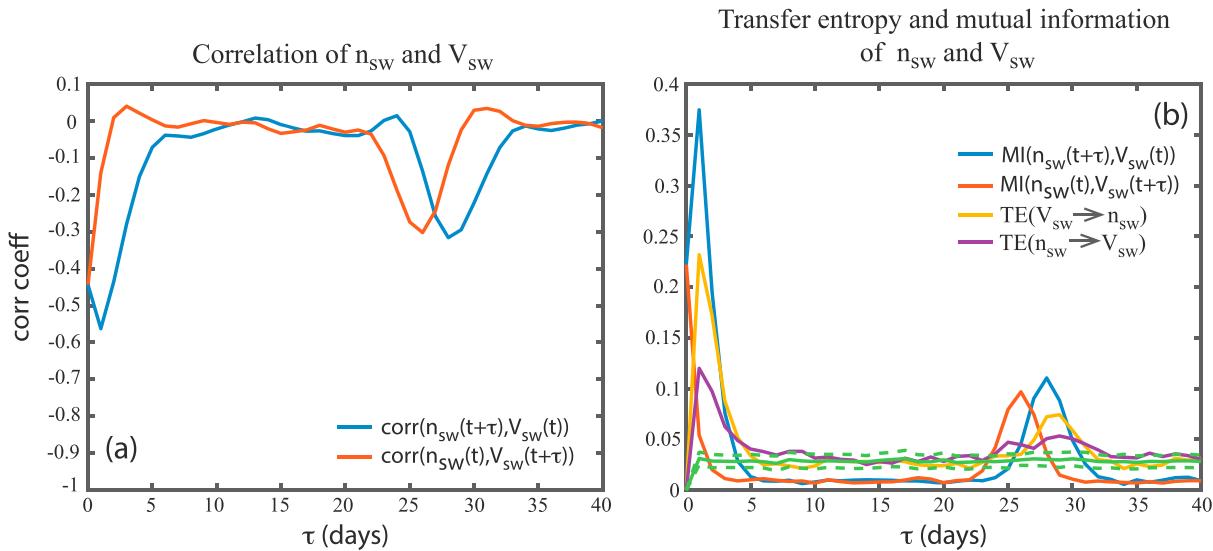


Figure 6. (a) Correlation coefficient of $[n_{SW}(t + \tau), V_{SW}(t)]$ (blue) and $[n_{SW}(t), V_{SW}(t + \tau)]$ (red). (b) MI($n_{SW}(t + \tau), V_{SW}(t)$) (blue), MI($n_{SW}(t), V_{SW}(t + \tau)$) (red), TE($n_{SW}(t + \tau), V_{SW}(t)$) (yellow), and TE($n_{SW}(t), V_{SW}(t + \tau)$) (purple). The green solid and dashed curves are the mean and 3σ from the mean of the noise. The transfer of information from n_{SW} to J_e ($n_{SW} \rightarrow V_{SW}$) peaks at $\tau_{\max} = 1$ day.

Figure 6a plots $\text{corr}[n_{SW}(t + \tau), V_{SW}(t)]$ (blue curve) and $\text{corr}[n_{SW}(t), V_{SW}(t + \tau)]$ (red curve). The blue curve has a minimum at $\tau_{\min} = 1$ day ($r = -0.56$) and a secondary minimum at $\tau = 28$ days ($r = -0.32$). The red curve has a minimum at $\tau_{\min} = 26$ days ($r = -0.30$).

Figure 6b plots TE and MI for (n_{SW}, V_{SW}) in a similar manner as in Figures 2b and 4b. Both, MI and TE for $[n_{SW}(t + \tau), V_{SW}(t)]$, blue and yellow curves, respectively, show a peak at $\tau_{\max} = 1$ day, which is consistent with the linear correlational analysis. TE($n_{SW}(t + 1$ day), $V_{SW}(t)$) has $it_{\max} = 0.20$, snr = 7.4, and significance = 95σ .

From the considerations of the lag times, it is entirely possible that anticorrelation of $[J_e(t + 1$ day), $n_{SW}(t)]$ is caused by $[J_e(t + 2$ days), $V_{SW}(t)]$ correlation and the anticorrelation of $[n_{SW}(t + 1$ day), $V_{SW}(t)]$. Note by correlation here we mean both linear and nonlinear correlations. However, we cannot rule out that n_{SW} may also influence J_e independently of V_{SW} . To investigate this, we perform CMI calculation as described in section 4.4.

So far, we have determined the lag times at daily resolution because we use daily solar wind and LANL data. The LANL higher time resolution data are not yet available, but the OMNI solar wind data are available at hourly resolution. Hence, we can investigate the $\text{corr}(n_{SW}, V_{SW})$ at hourly resolution.

Figure 7a plots the $\text{corr}[n_{SW}(t + \tau), V_{SW}(t)]$ for $\tau = 0$ –100 h (solid curve). It shows that the correlation reaches a minimum at $\tau_{\min} = 14$ h for the data interval used in the present study 1989–2009. It also shows that n_{SW} anticonrelation with V_{SW} has a broad minimum. To quantify the width of the minimum, we draw a dashed horizontal line that intersects the solid curve at $\tau = 0$ h and show that the anticonrelation does not worsen than that at $\tau = 0$ h until $\tau = 36$ h. However, τ_{\min} , the correlation coefficient at τ_{\min} , and the width of the minimum are time dependent. As an example, Figure 7b shows that for the period 2000–2014, $\tau_{\min} = 17$ h and the width of the minimum using the above criterion is about 46 h. Moreover, Figure 6b shows that TE($n_{SW}(t + \tau), V_{SW}(t)$) does not reach the noise level until $\tau > 3$ days, suggesting a rather long period when V_{SW} affects the trailing density, n_{SW} .

4.4. Ranking of Solar Wind Parameters Based On Information Transfer to Radiation Belt Electrons

From our analysis above, V_{SW} is a stronger driver of J_e than n_{SW} , i.e., V_{SW} transfers more information to J_e than n_{SW} . For example, TE($V_{SW}(t) \rightarrow J_e(t + 2$ days)] has $it_{\max} = 0.30$ and snr = 5.7, while TE($n_{SW}(t) \rightarrow J_e(t + 1$ days)] has $it_{\max} = 0.13$ and snr = 4.4. Because V_{SW} anticorrelates with n_{SW} , there is some embedded dependence, so it is necessary to use conditional mutual information (CMI) to determine how much information passes from n_{SW} to J_e , given V_{SW} and vice versa.

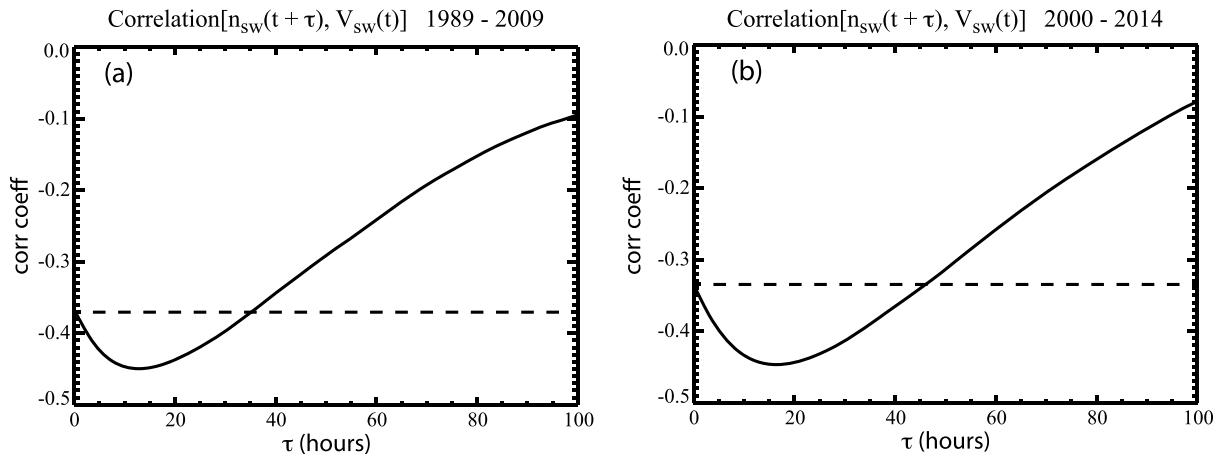


Figure 7. Correlation coefficient of $[n_{sw}(t+\tau), V_{sw}(t)]$ for (a) 1989–2009 and (b) 2000–2014 at hourly resolution. The anticorrelation improves with increasing τ , reaching minimum at $\tau_{min} = 14$ and 17 h in Figures 7a and 7b, respectively. The correlation coefficient finally reaches the same value as that at $\tau = 0$ h (the dashed line) at $\tau = 36$ and 46 h in Figures 7a and 7b.

To calculate how much information flows from n_{sw} to J_e , given V_{sw} , we calculate $CMI[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$, which is plotted as blue curve in Figure 8a. Using a similar approach as for TE, we determine the noise level of the surrogates: $CMI[J_e(t+\tau), \text{sur}[n_{sw}(t)]|V_{sw}(t)]$. The mean and σ of the noise are calculated in the same manner as TE (described in section 4.1) and used to determine the significance of the results. The mean noise and 3σ are plotted as green solid and dashed curves, respectively. Figure 8a shows that $CMI[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$ peaks at $\tau_{max} = 0$ day with $it_{max} = 0.091$ and $snr = 3.2$. The $\tau_{max} = 0$ day suggests that J_e response lag time to n_{sw} is less than 24 h.

We can now revisit the J_e response lag times to V_{sw} and n_{sw} . Earlier, we establish that $J_e(t+2$ days) correlates with $V_{sw}(t)$ (Figure 2), $J_e(t+1$ day) anticorrelates with $n_{sw}(t)$ (Figure 4), but $n_{sw}(t+1$ day) anticorrelates with $V_{sw}(t)$ (Figure 6). However, our CMI analysis (Figure 8a) shows that given V_{sw} , J_e response lag time to n_{sw} is 0 day (<24 h). This suggests that the $J_e(t+1$ day) anticorrelation with $n_{sw}(t)$ seen in Figure 4 at least partly comes from the

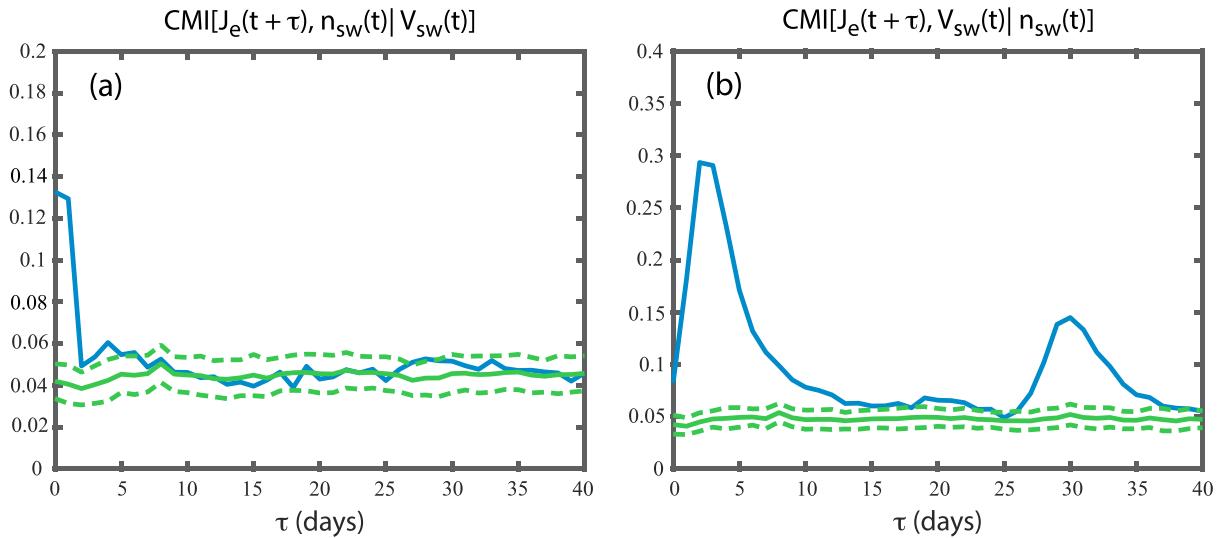


Figure 8. Blue curve showing $CMI[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$ and $CMI[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$. The green solid and dashed curves are the mean and 3σ from the mean of the noise. (a) Unlike $TE[J_e(t+\tau), n_{sw}(t)]$, which peaks at $\tau_{max} = 1$ day, $CMI[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$ peaks at $\tau_{max} = 0$ day ($it_{max} = 0.091$). The smaller τ_{max} comes about because CMI removes the effect of V_{sw} on J_e (see text). (b) The peak in $CMI[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$ ($it_{max} = 0.25$) is broader and has slightly higher snr than that of $TE[J_e(t+\tau), V_{sw}(t)]$ in Figure 2b because CMI removes the effect of n_{sw} , which anticorrelates with J_e . V_{sw} transfers about 2.7 times more information to J_e than n_{sw} .

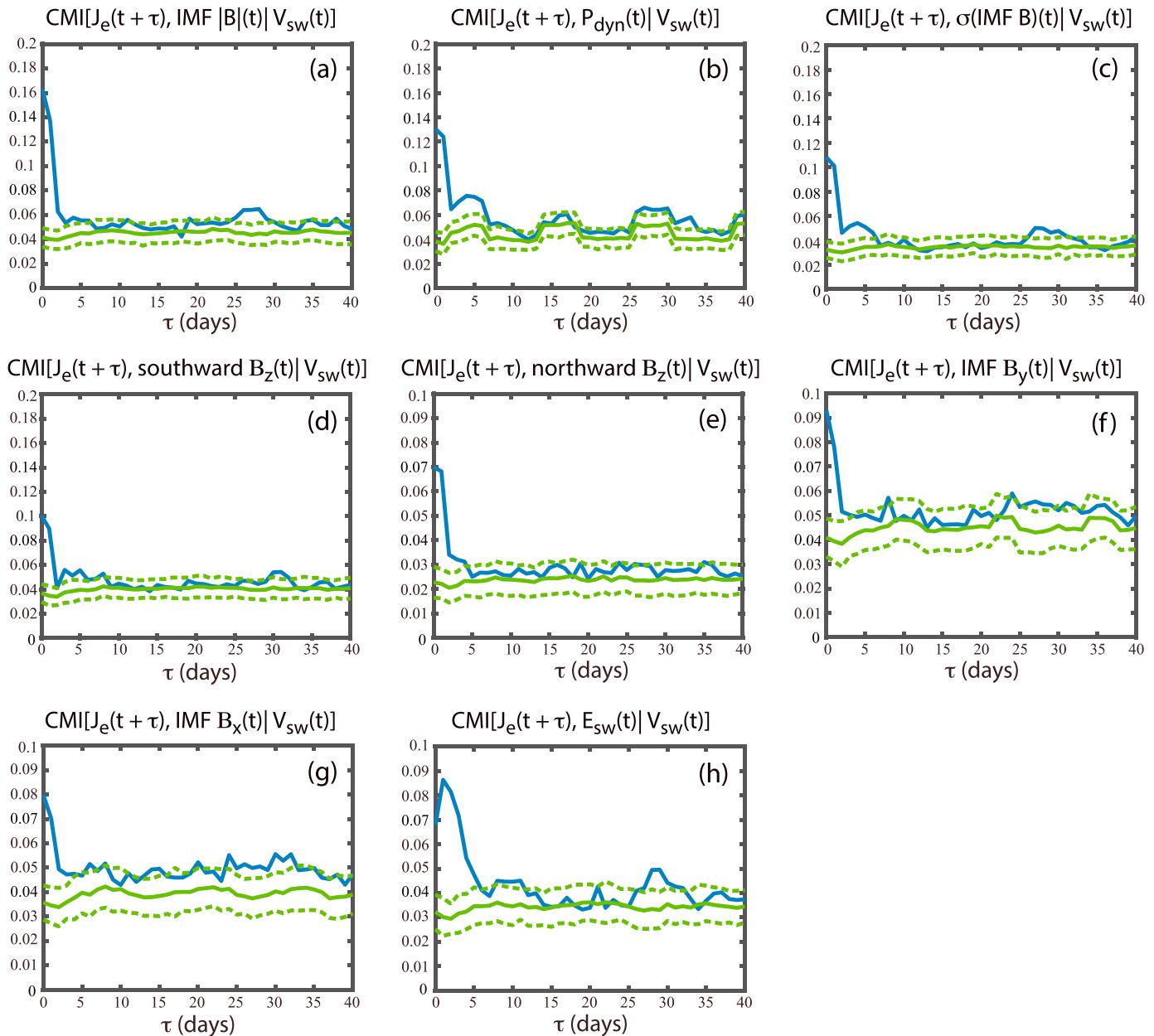


Figure 9. (a–h) CMI solar wind parameter with J_e , given V_{sw} for IMF $|B|$, P_{dyn} , $\sigma(\text{IMF } B)$, southward IMF B_z , northward IMF B_z , IMF B_y , IMF B_x , and E_{sw} respectively, as blue solid curves. The green solid and dashed curves are the mean and 3σ from the mean of the noise. The relationships are summarized in Table 1. Note that the scale of the y axis is 0–0.2 for Figures 9a–9d and 0–0.1 for Figures 9e–9h.

$J_e(t+2 \text{ days})$ correlation with $V_{sw}(t)$ and $V_{sw}(t)$ anticorrelation with $n_{sw}(t+1 \text{ day})$. However, Figure 8a shows that the peak is rather broad, suggesting that the J_e response is still significant at $\tau = 1 \text{ day}$.

We also calculate $\text{CMI}[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$, which is plotted in Figure 8b as blue solid curve. The blue curve peaks at $\tau = 2 \text{ days}$ with $it_{max} = 0.25$ which is about 2.7 times larger than the it_{max} of 0.091 for $\text{CMI}[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$. Thus, V_{sw} transfers more information to J_e than n_{sw} does.

Interestingly, the peak in $\text{CMI}[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$ (Figure 8b) is broader than the peak in $\text{TE}[J_e(t+\tau), V_{sw}(t)]$ (Figure 2b). The former also has slightly higher snr (6.6) than the latter (5.7). Removing the effect of n_{sw} , which anticorrelates with J_e , has the effects of broadening the peak, lowering the noise, and increasing the snr. These effects are discussed in section 5.3.

Table 1. Ranking of the Importance of the Solar Wind Parameters Based on Information Transfer to Geosynchronous MeV Electron Flux (J_e) at τ_{\max} , Where τ_{\max} is the Lag Time When the Information Transfer Peaks^a

Rank	Solar Wind Parameters	Peak Information Transfer ($i_{t_{\max}}$)	Signal-to-Noise Ratio at τ_{\max}	Significance at τ_{\max} (σ)	τ_{\max} (days)	Prediction Horizon (days)
1	V_{sw}	0.25	6.6	94	2	10 ^b
2	IMF $ \mathbf{B} $	0.12	3.9	48	0	2
3	P_{dyn}	0.092	3.4	35	0	2
3	n_{sw}	0.091	3.2	34	0	2
4	$\sigma(\text{IMF } B)$	0.075	3.9	48	0	2
5	IMF $B_z < 0$	0.064	2.7	26	0	2
6	E_{sw}	0.056	2.9	22	1	5
7	IMF B_y	0.052	2.3	20	0	2
8	IMF $B_z > 0$	0.048	3.1	22	0	2
9	IMF B_x	0.044	2.2	19	0	2

^aParameters 2–9 are calculated from $CMI[J_e(t + \tau), x(t)|V_{sw}(t)]$, where x = parameters 2–9, whereas parameter 1 is calculated from $CMI[J_e(t + \tau), V_{sw}(t)|n_{sw}(t)]$. The peak information transfer ($i_{t_{\max}}$) = peak – mean noise, the signal-to-noise ratio = peak/noise, and significance = $i_{t_{\max}}/\sigma(\text{noise})$. Noise is calculated from surrogate data (see section 4.1). The prediction horizon gives the lag time when there is no information transfer from the solar wind parameter to J_e . Note that n_{sw} and P_{dyn} are both ranked at number 3 because they have similar $i_{t_{\max}}$ (the effect of V_{sw} has been removed [see section 5.3]). Northward IMF has slightly higher snr than southward IMF because northward IMF has lower noise level than southward IMF.

^bExcluding the effect of solar rotation.

The above analysis suggests that V_{sw} is the major driver of J_e . Next, we investigate whether other solar wind parameters also contribute to J_e . We calculate the information transfer from IMF $|\mathbf{B}|$, P_{dyn} , $\sigma(\text{IMF } B)$, southward IMF B_z , northward IMF B_z , IMF B_y , IMF B_x , and solar wind electric field (E_{sw}) to J_e , given V_{sw} . The northward (southward) IMF B_z is calculated from the daily average of the hourly IMF B_z when $\text{IMF } B_z > 0$ ($\text{IMF } B_z < 0$). The results are plotted in Figure 9. Table 1 gives the ranking based on the $i_{t_{\max}}$ of various solar wind parameters. Thus, the ranking gives the importance of each solar wind parameter based on the information transfer to J_e . Table 1 also lists τ_{\max} for the curves in Figures 8 and 9, which signifies the lag time when information transfer to J_e maximizes.

Note that the ranking in Table 1 is obtained with daily resolution data. It is possible that the ranking of some parameters may change if the data are analyzed at higher time resolution. For example, some studies showed that southward IMF B_z can influence J_e [e.g., Li *et al.*, 2005; Onsager *et al.*, 2007; Miyoshi and Kataoka, 2008], but southward IMF B_z is only ranked number 5 in Table 1. IMF fluctuates with periods of northward and southward IMF at minutes or tens of minutes time scale. Thus, the low ranking of the southward IMF B_z most likely results from the fluctuations of IMF B_z within 1 day period [e.g., Li *et al.*, 2001; Balikhin *et al.*, 2011; Reeves *et al.*, 2011]. Consistent with our result, Li *et al.* [2001] found that IMF B_z is poorly correlated with J_e at daily resolution. Interestingly, although southward IMF B_z has higher $i_{t_{\max}}$ than northward IMF B_z , northward IMF B_z has lower noise level and hence higher snr than southward IMF B_z . The τ_{\max} for E_{sw} is 1 day, which may be the average of $\tau_{\max} = 2$ days for V_{sw} and $\tau_{\max} = 0$ day for IMF B_z or IMF $|\mathbf{B}|$.

4.5. Detecting Changes in the System Dynamics

As described in section 3, transfer entropy from x to y , $TE(x \rightarrow y)$, gives a measure of information transfer from variable x to y . In the solar wind-magnetosphere system, the solar wind driving of the magnetosphere is not constant, depending on the strength of the driver and internal dynamics [e.g., Wing *et al.*, 2005a; Johnson and Wing, 2005]. So the system dynamics may not be stationary. The dynamics of the system can be detected by applying TE to a sliding window of data. Figure 10 shows the behavior of windowed $TE[V_{sw}(t) \rightarrow J_e(t + 2 \text{ days})]$ over the course of 0–2500 days since 1 January 1989 (a sliding 50 day window is used). One of the key features of the figure is the variation in TE over the course of 7 years, indicative of nonstationary dynamics. There are periods when TE has higher values, suggesting stronger solar wind-radiation belt coupling and vice versa.

With windowed TE, we are limited to a small number of points, which can pose difficulties for noise calculation from an ensemble of surrogates. A small number points can lead to a rather large noise and uncertainty

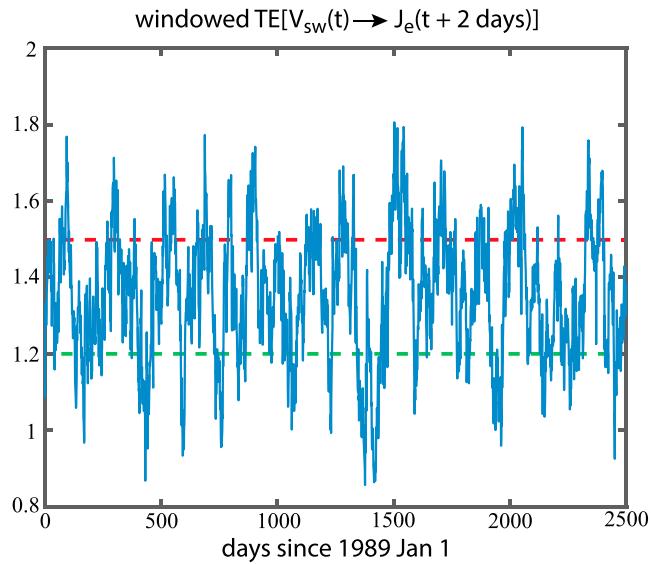


Figure 10. Blue curve showing windowed $TE[J_e(t+2 \text{ days}), V_{sw}(t)]$ over the course of 0–2500 days after 1 January 1989. The dynamics in the solar wind–outer radiation belt system changes with time, showing periods of high and low TEs.

accelerate low-energy electrons from a few keV to \sim 100 keV and once in the inner magnetosphere, wave-electron interactions accelerate the electrons further to several MeV [e.g., Lyatsky and Khazanov, 2008a; Baker and Kanekal, 2008]. The mechanisms for accelerating the electrons to MeV energy range generally fall into two categories. In the first mechanism, electron interactions with global ULF waves can increase radial diffusion or nonadiabatic transport of electrons resulting in acceleration [e.g., Baker et al., 1998; Li and Temerin, 2001; Li et al., 2005; Elkington et al., 1999; Rostoker et al., 1998; Ukhorskiy et al., 2005; Mathie and Mann, 2000, 2001; Reeves, 2007; Shprits et al., 2009; Green and Kivelson, 2004; Kellerman and Shprits, 2012]. The second mechanism is often referred to as local acceleration where low-energy electrons interact with locally grown waves such as VLF whistler-mode waves [Summers et al., 1998, 2007; Omura et al., 2007; Horne et al., 2005], or fast magnetosonic waves [e.g., Horne et al., 2007; Shprits et al., 2008]. These two mechanisms are not necessarily mutually exclusive.

Figure 1 shows that the relationship between J_e and V_{sw} is nonlinear, and hence, it is necessary to use information theoretical tools to discover the full extent of the relationships between these two parameters. Previous correlational analyses show that $J_e(t+2 \text{ days})$ correlates best with $V_{sw}(t)$ [e.g., Reeves et al., 2011; Lyatsky and Khazanov, 2008a], but correlational analysis only establishes linear correlation and does not establish causality. The present study establishes that $J_e(t+2 \text{ days})$ and $V_{sw}(t)$ are nonlinearly correlated. Using TE, we establish that indeed, information transfer or causality from V_{sw} to J_e peaks at $\tau = 2 \text{ days}$. However, n_{sw} anticorrelates with V_{sw} . Removing the effects of n_{sw} , $CMI[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$ shows that the information transfer peaks also at $\tau = 2 \text{ days}$, but the peak is broader. For example, the information transfer at $\tau = 3 \text{ days}$ is only slightly lower than that at $\tau = 2 \text{ days}$, as shown in Figure 8b.

Our result is consistent at least with the first electron acceleration mechanism mentioned above. Large V_{sw} can increase the occurrences of KHI along the magnetopause flanks [e.g., Fairfield et al., 2000; Johnson et al., 2014; Wing et al., 2005b], leading to enhancements of ULF waves within the magnetosphere [e.g., Engebretson et al., 1998; Vennerstrøm, 1999] and electron acceleration. Thus, the process to accelerate the electrons to MeV energy range takes 2 days, as previously suggested [e.g., Kellerman and Shprits, 2012; Reeves et al., 2011]. V_{sw} may also be tied to the local acceleration mechanism through substorm particle injections [e.g., Baker and Kanekal, 2008; Kissinger et al., 2011; Tanskanen, 2009; Kellerman and Shprits, 2012; Newell et al., 2016].

(large $\sigma(\text{noise})/\text{mean noise}$), rendering the noise to be unreliable. Therefore, rather than using it ($TE - \text{mean noise}$), we just use TE to characterize the dynamics of the system. Generally, periods of low TE can be considered a baseline and deviations from the baseline may indicate the presence of significant information transfer in the dynamics. The value of this approach can ultimately be measured by how well it detects dependencies and changes in system dynamics. In section 5.4, we show that even with just using TE, we are able to dissect the triangle distribution [Reeves et al., 2011] quite well and in section 5.5, we discuss applications to modeling.

5. Discussion

5.1. Solar Wind Velocity Driving Geosynchronous MeV Electron Flux

Studies suggested that substorm or storm injection process from plasma sheet into the inner magnetosphere

5.2. n_{sw} and V_{sw} Anticorrelation

The anticorrelation of n_{sw} and V_{sw} is well known [e.g., Hundhausen *et al.*, 1970], but the long lag time for this anticorrelation is relatively unknown. The anticorrelation may result from the solar wind high-speed streams that originate from the coronal holes, which have higher velocities and lower densities than the background solar wind. Surprisingly, the anticorrelation peaks at $\tau = 14\text{--}17$ h, depending on the year. It is not clear what causes the lag time to peak at 14–17 h. The solar wind high-speed stream may originate from the coronal holes at high latitudes of the Sun whereas the background slower solar wind may originate from lower latitudes [e.g., Kreiger *et al.*, 1973; Schwenn *et al.*, 1978]. This lag time may result from the compression of the leading edge of the high-speed stream structure when it encounters the denser background solar wind. Such compression may create slower and denser structure at the leading edge of the high-speed stream [e.g., Gosling *et al.*, 1972]. As a result, the anticorrelation at the leading edge of the high-speed stream is not as good as that at the trailing edge, which may preserve better the high-speed, low-density structure. Figure 6b suggests that there is information transfer from V_{sw} to n_{sw} up to 4–5 days, suggesting perhaps that the longevity of the high-speed stream structure is about 4–5 days. This is consistent with Gosling *et al.* [1972], which reports that the widths of high-speed streams are about 4 days. This property needs to be further investigated.

The correlational analyses of $[n_{sw}(t + 1 \text{ day}), V_{sw}(t)]$ and $[J_e(t + 2 \text{ day}), V_{sw}(t)]$ return correlation coefficients of -0.56 and 0.63, respectively. So the former has a slightly lower correlation than the latter (the two set of data have similar size). However, the scatterplots in Figures 5b and 1c show that both sets of data exhibit nonlinear behaviors. Hence, the linear correlational analysis may not capture the full extent of their relationships. Indeed, analyses with transfer entropy reveal that $\text{TE}[n_{sw}(t + 1 \text{ day}), V_{sw}(t)]$ has comparable significance (95σ) to that of $\text{TE}[J_e(t + 2 \text{ days}), V_{sw}(t)]$ (94σ).

The significant transfer of information from V_{sw} to n_{sw} has implications to the studies of solar wind driving of the magnetosphere that involve n_{sw} and V_{sw} . These studies should take into account the strong anticorrelation between n_{sw} and V_{sw} that can persist even at large lag times. For example, any attempt to isolate the effects of $n_{sw}(V_{sw})$ on the magnetosphere would need to effectively remove the effects of $V_{sw}(n_{sw})$ using CMI or similar methods.

5.3. Solar Wind Density Driving Geosynchronous MeV Electron Flux

Balikhin *et al.* [2011] investigated the control of solar wind parameters on J_e and found that the most dominant solar wind parameter is n_{sw} , which controls about 78% of the variance of J_e with 1 day lag, while V_{sw} only controls 11% of the variance. On the other hand, Vassiliadis *et al.* [2005] examined the geoeffectiveness of 17 solar wind and magnetospheric parameters and found that n_{sw} is weakly linked to J_e in the outer radiation belts. Other studies found that V_{sw} is the most dominant driver of J_e [e.g., Li *et al.*, 2001; Vassiliadis *et al.*, 2005; Kellerman and Shprits, 2012; Ukhorskiy *et al.*, 2004].

The present study finds that $J_e(t + 1 \text{ day})$ anticorrelates with $n_{sw}(t)$. The lag time of 1 day is consistent with that found in Balikhin *et al.* [2011]. However, $n_{sw}(t + 1 \text{ day})$ anticorrelates with $V_{sw}(t)$. Moreover, $\text{CMI}[J_e(t + \tau), n_{sw}(t)|V_{sw}(t)]$ peaks at $\tau = 0$ day, suggesting that given V_{sw} , J_e responds to n_{sw} in < 24 h. Hence, J_e response lag time of 1 day to n_{sw} in Figure 4 and in Balikhin *et al.* [2011] can be attributed at least partly to $J_e(t + 2 \text{ days})$ correlation with $V_{sw}(t)$, and $V_{sw}(t)$ anticorrelation with $n_{sw}(t + 1 \text{ day})$. Figure 8 and Table 1 show that V_{sw} is by far the dominant driver of J_e , transferring 2.7 times more information to J_e than n_{sw} does.

Lyatsky and Khazanov [2008b] assumed that the anticorrelation of V_{sw} and n_{sw} is weak in their analysis of the effect of K_p and n_{sw} on J_e . They concluded that n_{sw} has a strong effect on J_e within 2 days before geomagnetic disturbances. However, some of the effects attributed to n_{sw} may be due to V_{sw} in their study. Figures 6 and 7 show that the anticorrelation of V_{sw} and n_{sw} is not trivial and any attempt to interpret the effects of n_{sw} and V_{sw} on J_e should take into account the anticorrelation of V_{sw} and n_{sw} . To illustrate, $\text{TE}[n_{sw}(t) \rightarrow J_e(t + 1 \text{ day})]$ has $it_{\max} = 0.13$, but removing the effects of V_{sw} , the it_{\max} drops $\sim 30\%$ to 0.091 [it_{\max} of $(\text{CMI}[J_e(t + 0 \text{ day}), n_{sw}(t)|V_{sw}(t)]$ is 0.091)].

Conversely, some of the effects attributed to V_{sw} may be due to n_{sw} , but the effects of n_{sw} are smaller. For example, $\text{TE}[V_{sw}(t) \rightarrow J_e(t + 2 \text{ days})]$ has $it_{\max} = 0.30$, but removing the effects of n_{sw} , the it_{\max} drops only $\sim 17\%$ to 0.25 [it_{\max} of $\text{CMI}[J_e(t + 2 \text{ days}), V_{sw}(t)|n_{sw}(t)] = 0.25$]. Interestingly, Figure 8b shows that the $\text{CMI}[J_e(t + \tau), V_{sw}(t)|n_{sw}(t)]$ peak is broader than that of $\text{TE}[V_{sw}(t) \rightarrow J_e(t + \tau)]$ as shown in Figure 2b. For example,

the CMI at $\tau = 3$ days is only slightly smaller than that at $\tau = 2$ days. The reason for the broader peak is that as shown in Figure 8a, there is a significant information transfer from n_{sw} to J_e at $\tau = 0$ –1 day, but it falls off rapidly at larger τ . Because the anticorrelation between V_{sw} and n_{sw} has 1 day lag, removing the effects of n_{sw} would lower information transfer from V_{sw} to J_e (i.e., $TE[V_{sw}(t) \rightarrow J_e(t+\tau)]$) at $\tau = 1$ –2 days. So it for $TE[V_{sw}(t) \rightarrow J_e(t+\tau)]$ at $\tau = 1, 2$, and 3 is $0.26, 0.30$, and 0.23 , respectively, whereas the corresponding values for $CMI[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$ are $0.14, 0.25$, and 0.24 , respectively. Note that, at $\tau = 1$ and 2 , there are reductions in information transfer, while at $\tau = 3$, the information transfer is more or less the same (the difference is within 1σ [~ 0.01]). This leads to a broader peak in the $CMI[J_e(t+\tau), V_{sw}(t)|n_{sw}(t)]$ curve in Figure 8b than that in the $TE[V_{sw}(t) \rightarrow J_e(t+\tau)]$ curve in Figure 2b.

An increase in n_{sw} would increase solar wind dynamic pressure (P_{dyn}), which, in turn, would push the magnetopause inward, leading to electron losses at the high L shell [e.g., Li *et al.*, 2001]. Furthermore, the magnetopause compression would drive ULF waves [e.g., Korotova and Sibeck, 1995; Kepko and Spence, 2003; Claudepierre *et al.*, 2010] leading to fast radial diffusion, which redistributes the losses to the magnetopause to lower L shells, including at geosynchronous orbit [Shprits *et al.*, 2006; Kellerman and Shprits, 2012; Turner *et al.*, 2012]. Ukhorskiy *et al.* [2006] uses a test particle simulation to demonstrate this scenario, which is known as magnetopause shadowing. However, Lyatsky and Khazanov [2008a] reports that the correlation between P_{dyn} and J_e is poor and suggests that compression of the magnetosphere is probably not the main factor for the electron losses. Kellerman and Shprits [2012] examines various mechanisms for the losses and suggest that further investigation is needed. Our result suggests that based on information transfer from n_{sw} to J_e , any mechanism for n_{sw} anticorrelation with J_e has to operate or start operating within < 24 h (Figure 8a). We note that the magnetospheric compression due to an increase in n_{sw} or P_{dyn} would be nearly instantaneous, although it is not clear how long it would take for the electron losses to redistribute radially and how long the loss process would continue. Although the J_e response to n_{sw} peaks at $\tau = 0$, at $\tau = 1$ day, there is still a significant amount of information transferred from n_{sw} to J_e , as shown in Figure 8a.

The effect of P_{dyn} on J_e is complex. Previous studies showed a weak linear correlation between P_{dyn} and J_e [e.g., Lyatsky and Khazanov, 2008a], but such finding does not necessarily mean that the effect of P_{dyn} is insignificant or negligible. The poor correlation results from the complex response of J_e to both n_{sw} and V_{sw} on different time scales as well as the anticorrelation of these two variables. For example, the strong response of J_e to V_{sw} is masked when it is multiplied by n_{sw} . To examine the effect of P_{dyn} on J_e , it may be useful to consider a scenario in which n_{sw} changes while V_{sw} remains constant. Such scenario is not too outlandish, and in fact, observationally, an abrupt change in P_{dyn} is sometimes or even regularly associated with a change in n_{sw} , while V_{sw} changes relatively more slowly. In order to see how P_{dyn} affects J_e in this scenario, we can calculate $CMI[J_e(t+\tau), P_{dyn}(t)|V_{sw}(t)]$, but we might expect that $CMI[J_e(t+\tau), P_{dyn}(t)|V_{sw}(t)] \sim CMI[J_e(t+\tau), (n_{sw}(t)V_{sw}(t)^2)|V_{sw}(t)] \sim CMI[J_e(t+\tau), n_{sw}(t)|V_{sw}(t)]$. So the transfer of information from P_{dyn} to J_e , given V_{sw} , might be expected to be similar to that from n_{sw} to J_e , given V_{sw} . This is indeed seen in our analysis as shown in Table 1, which lists that n_{sw} and P_{dyn} have virtually the same ranking based on information transfer. The difference in their it_{max} is smaller than σ . Just to confirm that there is a significant information transfer from P_{dyn} to J_e , $TE[P_{dyn}(t) \rightarrow J_e(t+\tau)]$ is calculated and the result shows that the amount of information transfer from P_{dyn} to J_e is similar to that found in $CMI[J_e(t+\tau), P_{dyn}(t)|V_{sw}(t)]$. We will examine the effect of P_{dyn} on J_e in more details and in depth in a future study.

5.4. Revisiting the Triangle Distribution

Reeves *et al.* [2011] is the first to note the right triangle distribution exhibited in Figure 1a. Figure 1a plots $J_e(t+\tau)$ versus $V_{sw}(t)$ with no delay, $\tau = 0$. However, we note that as shown in Figures 1b and 1c, even with $\tau = 1$ or 2 days, respectively, the triangle distribution is still evident albeit not as prominent as for $\tau = 0$. For example, the triangle distribution can still be seen in Figure 1c, which is replotted in Figure 11a without the contour overlays. Reeves *et al.* [2011] notes that the left-hand side of the triangle forms because V_{sw} rarely goes below 300 km s^{-1} . The hypotenuse of the triangle suggests that the lower limit of J_e more or less increases with V_{sw} . The topside of the triangle suggests that J_e saturates, which can be attributed to local instabilities [Kennel and Petschek, 1966]. Reeves *et al.* [2011] considers this and other possible explanations for the J_e saturation. As noted by Reeves *et al.* [2011], the most interesting and perhaps mystifying

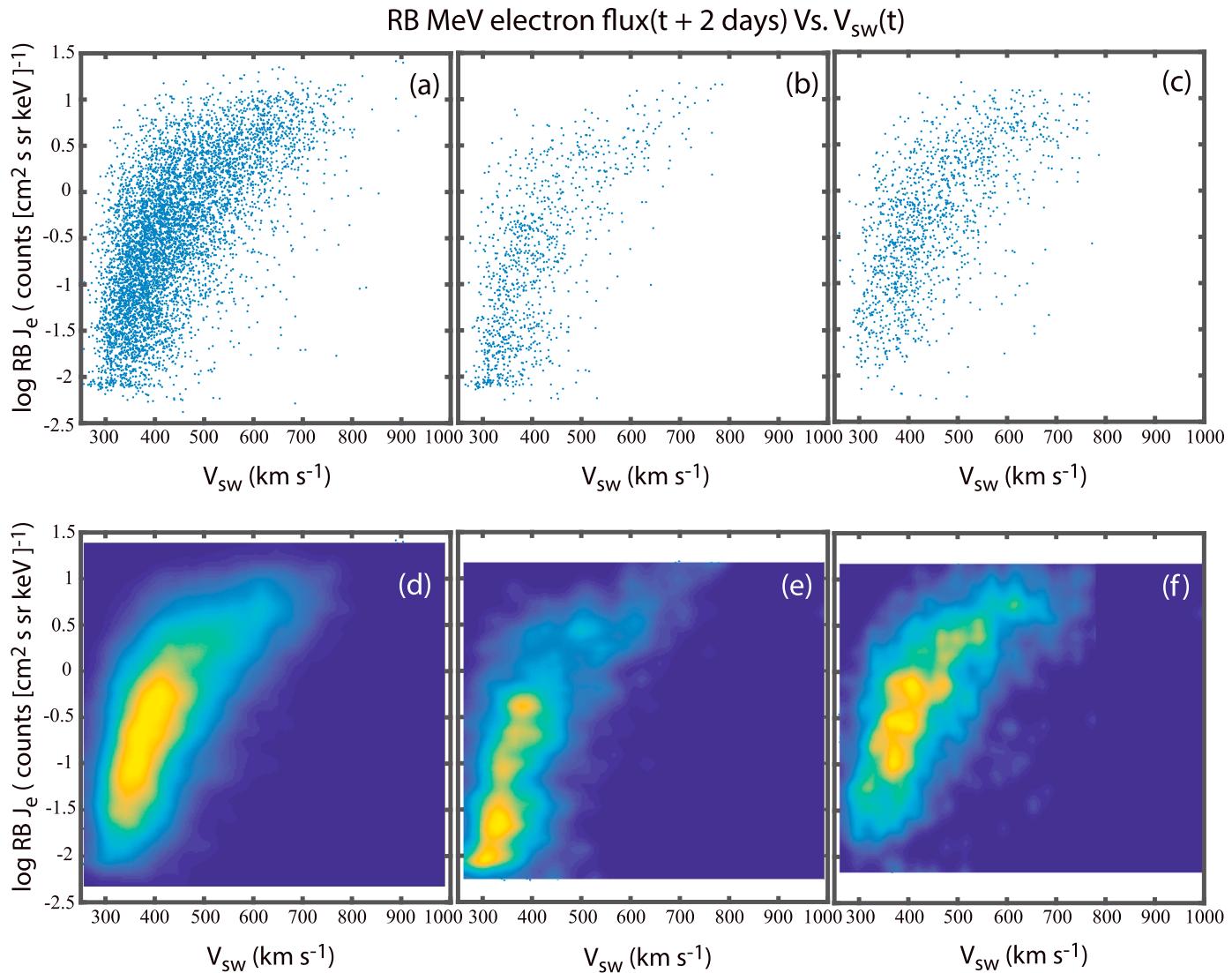


Figure 11. (a) Scatterplot of $\log J_e(t+2 \text{ days})$ versus $V_{sw}(t)$ showing the triangle distribution. This is similar to Figure 1c, but without the density contours. (b) The points in Figure 11a are plotted when $\text{TE}[J_e(t+2 \text{ days}), V_{sw}(t)] < 1.2$ (below the green dashed line in Figure 10). (c) The points in Figure 11a are plotted when $\text{TE}[J_e(t+2 \text{ days}), V_{sw}(t)] > 1.5$ (above the red dashed line in Figure 10). The distributions in Figures 11b and 11c differ and both do not have the same triangle distribution as in Figure 11a. Figures 11d–11f show the data density maps of the data in Figures 11a–11c, respectively. Figures 11e and 11f reinforce the differences in the data distributions in Figures 11b and 11c.

aspect of the triangle distribution is that high J_e is observed for all V_{sw} conditions and the variability of J_e at lower V_{sw} is much larger than that at higher V_{sw} .

Reeves *et al.* [2011] notes that the triangle distribution appears in the declining phase of the solar maximum, but it vanishes during solar maximum (although high J_e and low V_{sw} points still appear during solar maximum). This dependence of solar cycle suggests that perhaps the mode in which the solar wind couples to magnetosphere/radiation belt can be a factor. We probe the possible effects of solar wind-radiation belt coupling further with information theoretical tools. Particularly, we separate the points in Figure 11a based on the information transfer from V_{sw} to J_e .

Figure 11b shows $J_e(t+2 \text{ days})$ versus $V_{sw}(t)$ when $\text{TE}[V_{sw}(t) \rightarrow J_e(t+2 \text{ days})]$ is below 1.2, below the dashed green line in Figure 10. Figure 11c plots the points when TE is above 1.5, above the dashed red line in Figure 10. It is clear that the data distribution in Figure 11b looks different than that in Figure 11c. Also, the triangle distribution in Figure 11a is not reproduced in Figure 11b nor in Figure 11c. Although

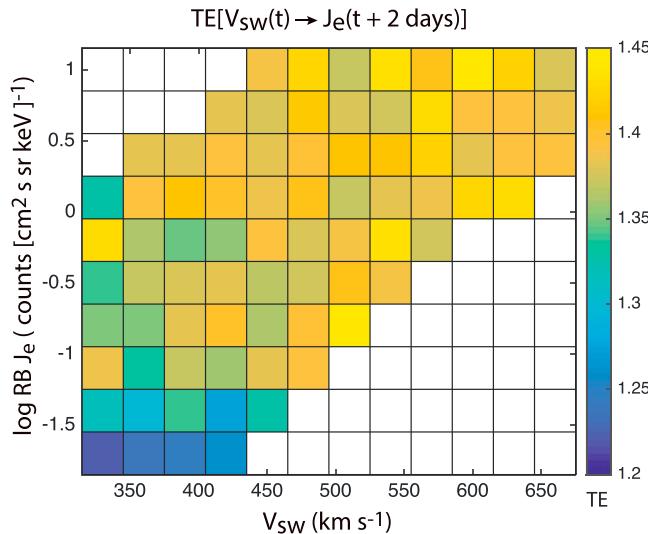


Figure 12. Mean $\text{TE}[V_{\text{sw}}(t) \rightarrow J_e(t + 2 \text{ days})]$ of each bin in $J_e(t + 2 \text{ days})$ versus $V_{\text{sw}}(t)$ distribution shown in Figure 11a. The bin size is 0.3 counts ($\text{cm}^2 \text{ s sr keV}^{-1}$) $\times 30 \text{ km s}^{-1}$. Figure 11a shows that at $V_{\text{sw}} < 500 \text{ km s}^{-1}$, J_e has a large variance as previously shown, but it turns out that these points are well ordered by TE. Low J_e corresponds to low TE and vice versa.

and (4) for $V_{\text{sw}} \geq 600 \text{ km s}^{-1}$, the higher limit of J_e saturates as in Figure 11a. In general, Figure 11c shows stronger dependence of J_e on V_{sw} than Figures 11a or 11b. The large spread of J_e at lower V_{sw} in the parent triangle distribution in Figure 11a appears smaller in Figure 11c. Thus, during the periods when TE is large, there is a large information transfer from V_{sw} to J_e , and we can indeed see that there is a stronger dependence of J_e on V_{sw} .

Figures 11d–11f show the data density maps of Figures 11a–11c, respectively. Figures 11e and 11f help draw sharper contrasts between the distributions in Figures 11b and 11c. There are proportionally more points with higher J_e and stronger dependency of J_e on V_{sw} in Figure 11f than in Figure 11e.

We investigate further the large spread of J_e at lower V_{sw} that can be seen in Figure 11a. The $\log J_e$ versus V_{sw} data in Figure 11a are binned in 0.3 counts ($\text{cm}^2 \text{ s sr keV}^{-1}$) $\times 30 \text{ km s}^{-1}$ bins. From Figure 10, we have calculated windowed TE for each point in the data set. We then assign the windowed TE for each point in the $\log J_e$ versus V_{sw} bins. Figure 12 shows the mean TE in each bin. Bins with fewer than 15 points are not displayed. The figure shows that for $V_{\text{sw}} < 500 \text{ km s}^{-1}$, there is a large spread of J_e . However, these J_e s are well ordered by TE. Large TE corresponds to large J_e and conversely small TE corresponds to small J_e . This suggests for $< 500 \text{ km s}^{-1}$, when there is small information transfer from V_{sw} to J_e , J_e is small and vice versa. We have also binned the data in Figure 11b in similar manner and obtained similar result, albeit with higher noise due to lower statistics in the bins.

Balikhin et al. [2011] suggests that the triangle distribution can be attributed to n_{sw} , and *Kellerman and Shprits [2012]* suggests the saturation of J_e in the triangle distribution can be attributed to n_{sw} . Our analysis in section 4.4 certainly supports the argument that n_{sw} has a significant effect on J_e . We investigate further the effect of n_{sw} on the triangle distribution. We assign $n_{\text{sw}}(t)$ for each point in the $\log J_e(t + 2 \text{ days})$ versus $V_{\text{sw}}(t)$ scatterplot in Figure 11a. These points are then binned using the same bin size as in Figure 12. Figure 13a shows the mean n_{sw} of each bin. As in Figure 12, bins with fewer than 15 points are not displayed. The most prominent trend in Figure 13a is a strong density gradient in the x direction because n_{sw} anticorrelates with V_{sw} .

However, our analysis and Figure 8a suggest that the maximum transfer of information from $n_{\text{sw}}(t)$ to $J_e(t + \tau)$ occurs at $\tau = 0 \text{ day}$ (<24 h). Hence, instead of assigning $n_{\text{sw}}(t)$ to each point in the $J_e(t + 2 \text{ days})$ versus $V_{\text{sw}}(t)$ plot, we assign $n_{\text{sw}}(t + 2 \text{ days})$ so that J_e is not time shifted with respect to n_{sw} . We repeat the same procedure

Figure 11c still shows a triangle, the triangle, which is not a right triangle, shows a different characteristic than that in Figure 11a.

Figures 11b and 11c contrast the differences between low and high TE cases. Figure 11b shows that (1) most of the points tend to have $V_{\text{sw}} < 500 \text{ km s}^{-1}$ and (2) excluding points with $V_{\text{sw}} > 500 \text{ km s}^{-1}$, J_e tends to have only weak dependency on V_{sw} and the distribution looks more like a rectangle than a triangle. On the other hand, Figure 11c shows that (1) the points tend to have more uniform distribution in velocity; (2) the hypotenuse of the triangle in Figure 11a that shows the lower limit of J_e increases with V_{sw} can still be seen; (3) unlike in Figure 11a where the left side of the triangle is nearly parallel to the y axis, the left side of the triangle now has a positive slope, suggesting that J_e increases with V_{sw} ;

solar wind density effect on the triangle distribution

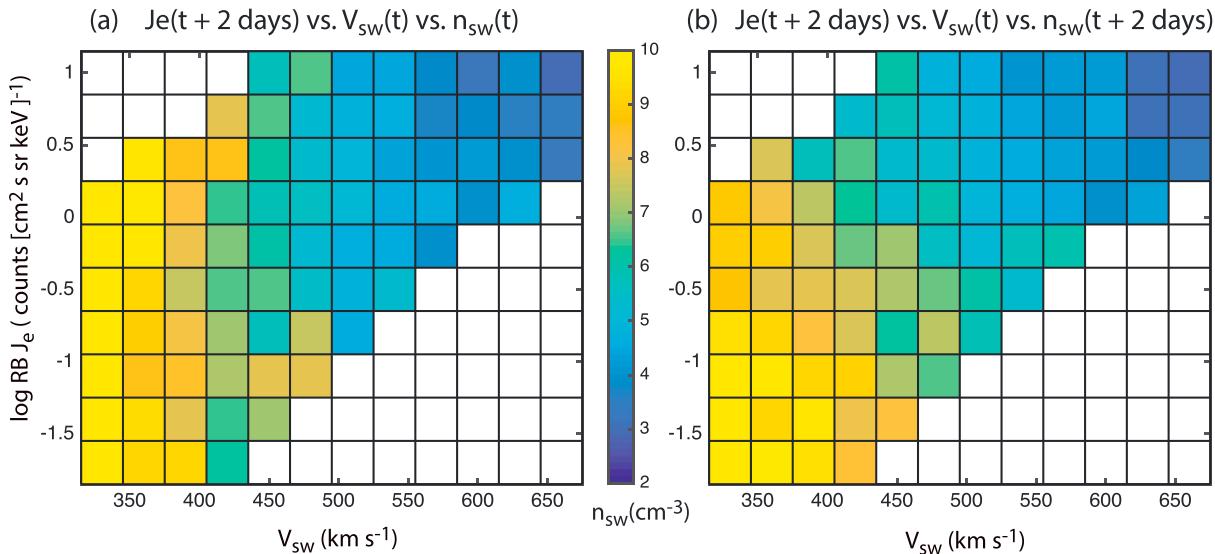


Figure 13. Points in $J_e(t + 2 \text{ days})$ versus $V_{\text{sw}}(t)$ distribution in Figure 11a are binned in 0.3 counts ($\text{cm}^2 \text{s} \text{sr} \text{keV}^{-1}$) $\times 30 \text{ km s}^{-1}$ bins. Each point is assigned its $n_{\text{sw}}(t)$ and $n_{\text{sw}}(t + 2 \text{ days})$ values. The latter has no time shift with respect to J_e such that information transfer from n_{sw} to J_e maximizes. (a) Mean $n_{\text{sw}}(t)$ and (b) mean $n_{\text{sw}}(t + 2 \text{ days})$ of each bin. In Figure 13a, the density gradient is mainly in the x direction due to the anticorrelation between n_{sw} and V_{sw} . However, in Figure 13b, there are density gradients in x and y directions. The latter can be attributed to P_{dyn} and magnetopause shadowing.

done for Figure 13a, and the result is shown in Figure 13b. Now, there are density gradients in both x and y directions. As in Figure 13a, the density gradient in the x direction is due to the anticorrelation of n_{sw} with V_{sw} . Figure 13b clearly shows that for $V_{\text{sw}} < 500 \text{ km s}^{-1}$, larger n_{sw} hence larger P_{dyn} can be associated with lower J_e and vice versa. This density gradient in the y direction may be attributed to the magnetopause shadowing effect, which rapidly depletes radiation belt fluxes when solar wind pressure is increased, as discussed in section 5.3.

Figure 13b shows that large V_{sw} can be associated with large J_e and small n_{sw} . The latter can be mostly attributed to the anticorrelation of n_{sw} with V_{sw} . Figure 13b also shows that large n_{sw} can decrease J_e , consistent with our analysis in section 5.3, but it is not clear if n_{sw} alone can explain why small V_{sw} ($< 500 \text{ km s}^{-1}$) can lead to high J_e and saturation of J_e for small n_{sw} . The high J_e and the saturation of J_e when $V_{\text{sw}} < 500 \text{ km s}^{-1}$ can probably be attributed to the strong solar wind-radiation belt coupling as suggested by the high TE in Figure 12.

5.5. Improving Models With Information Theory

Tools based on information theory can be used to improve modeling. Several ideas are discussed below.

5.5.1. Selecting Input Parameters

Often the first step in developing a parametric forecasting model is to decide which parameters should be used as inputs to the model. Using TE and CMI, one can determine the ranking of each parameter based on information transfer from the input to the output parameters. For example, Table 1 shows the ranking of solar wind parameters for solar wind-outer radiation belt system at daily resolution.

5.5.2. Detecting Nonstationarity in System Dynamics

As shown in Figure 10, TE can be used to detect changes in the system dynamics, e.g., nonstationarity of the system dynamics. Higher TE suggests that the solar wind-radiation belt system is more tightly coupled and vice versa. Figures 11b and 11c show visually the differences in the dynamics for low and high TE cases in J_e versus V_{sw} . In this case, TE can help decompose the triangle distribution into something that can be more easily understood. This kind of information can help modelers. For example, modelers may want to create a model that varies the coupling function strength within the model, depending on the value of TE. Alternatively, two separate models may be developed: one for low TE and one for high TE.

5.5.3. Prediction Horizon

TE shows how much information is transferred from the input time series x to output time series y . When TE is significantly above the noise level, it suggests that there is hope for the model to predict parameter y . Conversely, when TE is at noise level, there is not much information transfer from x to y and hence it can be expected that x would not be able to predict y accurately.

For example, Figure 8a suggests that n_{sw} transfers the most information to J_e at $\tau_{max}=0$ day, but the blue curve falls rapidly at $\tau > 2$ days, suggesting that little information is transferred from n_{sw} to J_e after 2 day lag time. Hence, the prediction horizon for using n_{sw} to predict J_e is about 1 day. Here we use mean noise $+3\sigma$ as the noise threshold, but using different threshold criterion would yield different prediction horizon. Likewise, Figure 8b suggests that the prediction horizon for using V_{sw} to predict J_e is about 7–10 days. Table 1 lists the prediction horizon for the parameters considered in the present study.

6. Summary

The present study applies information theoretical tools to investigate the solar wind drivers of the geosynchronous MeV electron fluxes. The following lists the summary of our results.

1. V_{sw} , n_{sw} , IMF $|\mathbf{B}|$, P_{dyn} , $\sigma(\text{IMF } B)$, E_{sw} , IMF B_x , IMF B_y , and IMF B_z (southward and northward) are causally related to J_e , but the amount of information transfer from each of this parameter to J_e differs. The ranking of these 10 parameters in terms of information transfer is given in Table 1.
2. V_{sw} is the most dominant driver, and the transfer of information from time series $V_{sw}(t)$ to $J_e(t+\tau)$ peaks at $\tau_{max}=2$ days. V_{sw} transfers 2.7 times more information to J_e than n_{sw} .
3. Although the anticorrelation between n_{sw} and V_{sw} is perhaps well known, the large and persistent lag times for this anticorrelation is relatively unknown. $n_{sw}(t+\tau)$ anticorrelates with $V_{sw}(t)$ with $\tau_{max}=14\text{--}17$ h, but the exact τ_{max} has time dependence. It is not clear what causes τ_{max} at 14–17 h. This may be due to the compression of the leading edge of the high-speed stream when it encounters the denser background solar wind. Analyses of solar wind driving of the magnetosphere involving V_{sw} and n_{sw} should take into account this anticorrelation that can persist even at large lag times, up to 4–5 days. For example, the information transfer from n_{sw} to J_e drops 30% after the effects of V_{sw} are removed.
4. $J_e(t+1$ day) anticorrelates with $n_{sw}(t)$, but the 1 day lag and the anticorrelation are at least partly due to (1) $J_e(t+2$ days) correlation with $V_{sw}(t)$ and (2) $V_{sw}(t)$ anticorrelation with $n_{sw}(t+1$ day). Given V_{sw} , the transfer of information from $n_{sw}(t)$ to $J_e(t+\tau)$ peaks at $\tau_{max}=0$ day (<24 h), suggesting the loss mechanism due to n_{sw} or P_{dyn} has to start operating in <24 h. However, the loss mechanism or mechanisms can operate for a long duration because at $\tau=1$ day, there is still significant information transfer from n_{sw} to J_e .
5. The triangle distribution in J_e versus V_{sw} plot shows a large variability of J_e for $V_{sw} < 500 \text{ km s}^{-1}$. However, these points are well ordered by their TE values: high TE corresponds to high J_e and vice versa. The triangle distribution can be decomposed to low and high TE cases. In the low TE case, the distribution looks more like a rectangle for $V_{sw} < 500 \text{ km s}^{-1}$, suggesting that V_{sw} has little influence on J_e in these conditions. In the high TE case, the lower and upper limits of J_e increase with V_{sw} for $V_{sw} < 600 \text{ km s}^{-1}$, but for $V_{sw} \geq 600 \text{ km s}^{-1}$, the higher limit of J_e saturates.
6. TE and CMI can be used effectively to improve modeling by (1) selecting model input parameters, (2) detecting changes in the dynamics of the system, and (3) determining prediction horizon. Table 1 gives this information for solar wind-outer radiation belt system.

The present study uses daily resolution LANL data. Reeves *et al.* [2013] investigated longer-term relationships between J_e and V_{sw} . They found that longer-term, 1 month to 1 year, averages show much stronger correlations than 1 day averages. They showed that this is not just because there is more “noise” superposed on a linear distribution. The distribution of $\log J_e$ around the baseline (the yearly mean) is very stable throughout the solar cycle. It would be interesting to apply our information theoretical tools to this normalized (rebaseline) data. It would also be interesting to apply our tools to higher resolution data, e.g., hourly resolution data, that are binned by D_{st} , P_{dyn} , and MLT. It would also be interesting to consider solar wind dynamic pressure variance in the Pc5 band, which has been shown to highly correlate with the Pc5 magnetic pulsations at geosynchronous orbit, which, in turn, can play a significant role in the radial transport of the radiation belt electrons [Takahashi and Ukhorskiy, 2007].

Acknowledgments

All the derived data products in this paper are available upon request by email (simon.wing@jhuapl.edu). Simon Wing acknowledges support from JHU/APL Janney Fellowship, NSF grant AGS-1058456, and NASA grants NNX13AE12G, NNX15AJ01G, NNX16AR10G, and NNX16AQ87G. Jay R. Johnson acknowledges support from NASA grants NNN11AR07I, NNX14AM27G, NNN14AY20I, NNX16AR10G, and NNX16AQ87G. NSF grants ATM0902730 and AGS-1203299, and DOE contract DE-AC02-09CH11466.

References

- Baker, D. N., and S. G. Kanekal (2008), Solar cycle changes, geomagnetic variations, and energetic particle properties in the inner magnetosphere, *J. Atmos. Sol.-Terr. Phys.*, **70**, 1950–206, doi:10.1016/j.jastp.2007.08.031.
- 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**(A9), 15,133–15,140, doi:10.1029/JA095iA09p15133.
- Baker, D. N., X. Li, J. B. Blake, and S. Kanekal (1998), Strong electron acceleration in the Earth's magnetosphere, *Adv. Space Res.*, **21**, 609–613, doi:10.1016/S0273-1177(97)00970-8.
- Balikhin, M. A., R. J. Boynton, S. A. Billings, M. Gedalin, N. Ganushkina, D. Coca, and H. Wei (2010), Data based quest for solar wind-magnetosphere coupling function, *Geophys. Res. Lett.*, **37**, L24107, doi:10.1029/2010GL045733.
- Balikhin, M. A., R. J. Boynton, S. N. Walker, J. E. Borovsky, S. A. Billings, and H. L. Wei (2011), Using the NARMAX approach to model the evolution of energetic electrons fluxes at geostationary orbit, *Geophys. Res. Lett.*, **38**, L18105, doi:10.1029/2011GL048980.
- Belian, R. D., G. R. Gisler, T. Cayton, and R. Christensen (1992), High-Z energetic particles at geosynchronous orbit during the Great Solar Proton Event Series of October 1989, *J. Geophys. Res.*, **97**(A11), 16,897–16,906, doi:10.1029/92JA01139.
- Borovsky, J. E., M. F. Thomsen, and R. C. Elphic (1998), The driving of the plasma sheet by the solar wind, *J. Geophys. Res.*, **103**, 17,617–17,640, doi:10.1029/97JA02986.
- Camporeale, E. (2015), Resonant and nonresonant whistlers-particle interaction in the radiation belts, *Geophys. Res. Lett.*, **42**, 3114–3121, doi:10.1002/2015GL063874.
- Camporeale, E., and G. Zimbardo (2015), Wave-particle interactions with parallel whistler waves: Nonlinear and time-dependent effects revealed by Particle-in-Cell simulations, *Phys. Plasmas*, **22** 092104, doi:10.1063/1.4929853.
- Claudepierre, S. G., M. K. Hudson, W. Lotko, J. G. Lyon, and R. E. Denton (2010), Solar wind driving of magnetospheric ULF waves: Field line resonances driven by dynamic pressure fluctuations, *J. Geophys. Res.*, **115**, A11202, doi:10.1029/2010JA015399.
- Darbellay, G. A., and I. Vajda (1999), Estimation of the information by an adaptive partitioning of the observations space, *IEEE Trans. Inf. Theory*, **45**, 1315–1321, doi:10.1109/18.761290.
- De Michelis, P., G. Consolini, M. Materassi, and R. Tozzi (2011), An information theory approach to the storm-substorm relationship, *J. Geophys. Res.*, **116**, A08225, doi:10.1029/2011JA01635.
- Elkington, S. R., M. K. Hudson, and A. A. Chan (1999), Acceleration of relativistic electrons via drift-resonant interaction with toroidal-mode Pc-5 ULF oscillations, *Geophys. Res. Lett.*, **26**, 3273–3276, doi:10.1029/1999GL003659.
- Engebretson, M., K.-H. Glassmeier, M. Stellmacher, W. J. Hughes, and H. Lühr (1998), The dependence of high-latitude Pc5 wave power on solar wind velocity and on the phase of high-speed solar wind streams, *J. Geophys. Res.*, **103**(A11), 26,271–26,283, doi:10.1029/97JA03143.
- Fairfield, D. H., A. Otto, T. Mukai, S. Kokubun, R. P. Lepping, J. T. Steinberg, A. J. Lazarus, and T. Yamamoto (2000), Geotail observations of the Kelvin-Helmholtz instability at the equatorial magnetotail boundary for parallel northward fields, *J. Geophys. Res.*, **105**(A9), 21,159–21,173, doi:10.1029/1999JA000316.
- Gosling, J. T., A. J. Hundhausen, V. Pizzo, and J. R. Asbridge (1972), Compressions and rarefactions in the solar wind: Vela 3, *J. Geophys. Res.*, **77**(28), 5442–5454, doi:10.1029/JA077i028p05442.
- Green, J. C., and M. G. Kivelson (2004), Relativistic electrons in the outer radiation belt: Differentiating between acceleration mechanisms, *J. Geophys. Res.*, **109**, A03213, doi:10.1029/2003JA010153.
- Horne, R. B., R. M. Thorne, S. A. Glaert, J. M. Albert, N. P. Meredith, and R. R. Anderson (2005), Timescale for radiation belt electron acceleration by whistler mode chorus waves, *J. Geophys. Res.*, **110**, A03225, doi:10.1029/2004JA010811.
- Horne, R. B., R. M. Thorne, S. A. Glaert, N. P. Meredith, D. Pokhotelov, and O. Santolik (2007), Electron acceleration in the Van Allen radiation belts by fast magnetosonic waves, *Geophys. Res. Lett.*, **34**, L17107, doi:10.1029/2007GL030267.
- Hundhausen, A. J., S. J. Bame, J. R. Asbridge, and S. J. Sydoriak (1970), Solar wind proton properties: Vela 3 observations from July 1965 to June 1967, *J. Geophys. Res.*, **75**(25), 4643–4657, doi:10.1029/JA075i025p04643.
- Johnson, J. R., and S. Wing (2005), A solar cycle dependence of nonlinearity in magnetospheric activity, *J. Geophys. Res.*, **110**, A04211, doi:10.1029/2004JA010638.
- Johnson, J. R., and S. Wing (2014), External versus internal triggering of substorms: An information-theoretical approach, *Geophys. Res. Lett.*, **41**, 5748–5754, doi:10.1002/2014GL060928.
- Johnson, J. R., S. Wing, and P. A. Delamere (2014), Kelvin Helmholtz instability in planetary magnetospheres, *Space Sci. Rev.*, **184**, 1–31, doi:10.1007/s11214-014-0085-z.
- Kellerman, A. C., and Y. Y. Shprits (2012), On the influence of solar wind conditions on the outer-electron radiation belt, *J. Geophys. Res.*, **117**, A05217, doi:10.1029/2011JA017253.
- Kennel, C. F., and H. E. Petschek (1966), Limit on stably trapped particle fluxes, *J. Geophys. Res.*, **71**(1), 1–28, doi:10.1029/JZ071i001p00001.
- Kepko, L., and H. E. Spence (2003), Observations of discrete, global magnetospheric oscillations directly driven by solar wind density variations, *J. Geophys. Res.*, **108**(A6), 1257, doi:10.1029/2002JA009676.
- Kissinger, J., R. L. McPherron, T.-S. Hsu, and V. Angelopoulos (2011), Steady magnetospheric convection and stream interfaces: Relationship over a solar cycle, *J. Geophys. Res.*, **116**, A00119, doi:10.1029/2010JA015763.
- Korotova, G. I., and D. G. Sibeck (1995), A case study of transient event motion in the magnetosphere and in the ionosphere, *J. Geophys. Res.*, **100**(A1), 35–46, doi:10.1029/94JA02296.
- Kreiger, A. S., A. F. Timothy, and E. C. Roelof (1973), A coronal hole and its identification as the source of a high velocity solar wind stream, *Sol. Phys.*, **29**, 505–525, doi:10.1007/BF00150828.
- Li, W. (1990), Mutual information functions versus correlation functions, *J. Stat. Phys.*, **60**, 823–837, doi:10.1007/BF01025996.
- Li, X., and M. A. Temerin (2001), The electron radiation belt, *Space Sci. Rev.*, **95**, 569–580, doi:10.1023/A:1005221108016.
- Li, X., M. Temerin, D. Baker, G. Reeves, and D. Larson (2001), Quantitative prediction of radiation belt electrons at geostationary orbit based on solar wind measurements, *Geophys. Res. Lett.*, **28**(9), 1887–1890, doi:10.1029/2000GL012681.
- Li, X., D. N. Baker, M. Temerin, G. Reeves, R. Friedel, and C. Shen (2005), Energetic electrons, 50 keV to 6 MeV, at geosynchronous orbit: Their responses to solar wind variations, *Space Weather*, **3**, S04001, doi:10.1029/2004SW000105.
- Lyatsky, W., and G. V. Khazanov (2008a), Effect of solar wind density on relativistic electrons at geosynchronous orbit, *Geophys. Res. Lett.*, **35**, L03109, doi:10.1029/2007GL032524.
- Lyatsky, W., and G. V. Khazanov (2008b), Effect of geomagnetic disturbances and solar wind density on relativistic electrons at geostationary orbit, *J. Geophys. Res.*, **113**, A08224, doi:10.1029/2008JA013048.
- Materassi, M., L. Ciraolo, G. Consolini, and N. Smith (2011), Predictive Space Weather: An information theory approach, *Adv. Space Res.*, **47**, 877–885, doi:10.1016/j.asr.2010.10.026.

- Mathie, R. A., and I. R. Mann (2000), A correlation between extended intervals of ULF wave power and storm-time geosynchronous relativistic electron flux enhancements, *Geophys. Res. Lett.*, 27, 3621–3264, doi:10.1029/2000GL003822.
- Mathie, R. A., and I. R. Mann (2001), On the solar wind control of Pc5 ULF pulsation power at mid-latitudes: Implications for MeV electron acceleration in the outer radiation belt, *J. Geophys. Res.*, 106(A12), 29,783–29,796, doi:10.1029/2001JA000002.
- Meier, M. M., R. D. Belian, T. E. Cayton, R. A. Christensen, B. Garcia, K. M. Grace, J. C. Ingraham, J. G. Laros, and G. D. Reeves (1996), *The Energy Spectrometer for Particles (ESP): Instrument Description and Orbital Performance*, AIP Conf. Proc., vol. 383, pp. 203–210, Ameri. Inst. Phys., New York.
- Miyoshi, Y., and R. Kataoka (2008), Flux enhancement of the outer radiation belt electrons after the arrival of stream interaction regions, *J. Geophys. Res.*, 113, A03S09, doi:10.1029/2007JA012506.
- Newell, P. T., K. Liou, J. W. Gjerloev, T. Sotirelis, S. Wing, and E. J. Mitchell (2016), Substorm probabilities are best predicted from solar wind speed, *J. Atmos. Sol. Terr. Phys.*, 146, 28–37, doi:10.1016/j.jastp.2016.04.019.
- Omura, Y., N. Furuya, and D. Summers (2007), Relativistic turning acceleration of resonant electrons by coherent whistler mode waves in a dipole magnetic field, *J. Geophys. Res.*, 112, A06236, doi:10.1029/2006JA012243.
- Onsager, T. G., J. C. Green, G. D. Reeves, and H. J. Singer (2007), Solar wind and magnetospheric conditions leading to the abrupt loss of outer radiation belt electrons, *J. Geophys. Res.*, 112, A01202, doi:10.1029/2006JA011708.
- Paulikas, G. A., and J. B. Blake (1979), Effects of the solar wind on magnetospheric dynamics: Energetic electrons at the synchronous orbit, in *Quantitative Modeling of Magnetospheric Processes*, *Geophys. Monogr. Ser.*, vol. 21, pp. 180–202, AGU, Washington, D. C.
- Prichard, D., and J. Theiler (1995), Generalized redundancies for time series analysis, *Phys. D*, 84, 476–493, doi:10.1016/0167-2789(95)00041-2.
- Prokopenko, M., J. T. Lizer, and D. C. Price (2013), On thermodynamic interpretation of transfer entropy, *Entropy*, 15(2), 524–543, doi:10.3390/e15020524.
- Reeves, G. D. (2007), Radiation belt storm probes: A new mission for space weather forecasting, *Space Weather*, 5, S11002, doi:10.1029/2007SW000341.
- Reeves, G. D., S. K. Morley, R. H. W. Friedel, M. G. Henderson, T. E. Cayton, G. Cunningham, J. B. Blake, R. A. Christensen, and D. Thomsen (2011), On the relationship between relativistic electron flux and solar wind velocity: Paulikas and Blake revisited, *J. Geophys. Res.*, 116, A02213, doi:10.1029/2010JA015735.
- Reeves, G., S. Morley, and G. Cunningham (2013), Long-term variations in solar wind velocity and radiation belt electrons, *J. Geophys. Res. Space Physics*, 118, 1040–1048, doi:10.1002/jgra.50126.
- Rigler, E. J., M. Wiltberger, and D. N. Baker (2007), Radiation belt electrons respond to multiple solar wind inputs, *J. Geophys. Res.*, 112, A06208, doi:10.1029/2006JA012181.
- Rostoker, G., S. Skone, and D. N. Baker (1998), On the origin of relativistic electrons in the magnetosphere associated with some geomagnetic storms, *Geophys. Res. Lett.*, 25, 3701–3704, doi:10.1029/98GL2801.
- Schreiber, T. (2000), Measuring information transfer, *Phys. Rev. Lett.*, 85, 461–464, doi:10.1103/PhysRevLett.85.461.
- Schwenn, R., M. D. Montgomery, H. Rosenbauer, H. Middendorf, K. H. Mühlhäuser, S. J. Bame, W. C. Feldman, and R. T. Hansen (1978), Direct observation of the latitudinal extent of a high-speed stream in the solar wind, *J. Geophys. Res.*, 83(A3), 1011–1017, doi:10.1029/JA083iA03p01011.
- Shprits, Y. Y., R. M. Thorne, R. Friedel, G. D. Reeves, J. Fennell, D. N. Baker, and S. G. Kanekal (2006), Outward radial diffusion driven by losses at magnetopause, *J. Geophys. Res.*, 111, A11214, doi:10.1029/2006JA011657.
- Shprits, Y. Y., D. A. Subbotin, N. P. Meredith, and S. R. Elkington (2008), Review of modeling of losses and sources of relativistic electrons in the outer radiation belt. II: Local acceleration and loss, *J. Atmos. Sol.-Terr. Phys.*, 70, 1694–1713, doi:10.1016/j.jastp.2008.06.014.
- Shprits, Y. Y., D. Subbotin, and B. Ni (2009), Evolution of electron fluxes in the outer radiation belt computed with the VERB code, *J. Geophys. Res.*, 114, A11209, doi:10.1029/2008JA013784.
- Simms, L. E., V. Pilipenko, M. J. Engebretson, G. D. Reeves, A. J. Smith, and M. Clilverd (2014), Prediction of relativistic electron flux at geostationary orbit following storms: Multiple regression analysis, *J. Geophys. Res. Space Physics*, 119, 7297–7318, doi:10.1002/2014JA019955.
- Simms, L. E., M. J. Engebretson, A. J. Smith, M. Clilverd, V. Pilipenko, and G. D. Reeves (2015), Analysis of the effectiveness of ground-based VLF wave observations for predicting or nowcasting relativistic electron flux at geostationary orbit, *J. Geophys. Res. Space Physics*, 120, 2052–2060, doi:10.1002/2014JA020337.
- Sturges, H. A. (1926), The choice of a class interval, *J. Am. Stat. Assoc.*, 21, 65–66, doi:10.1080/01621459.1926.10502161.
- Summers, D., R. M. Thorne, and F. Xiao (1998), Relativistic theory of wave-particle resonant diffusion with application to electron acceleration in the magnetosphere, *J. Geophys. Res.*, 103(A9), 20,487–20,500, doi:10.1029/98JA01740.
- Summers, D., B. Ni, and N. P. Meredith (2007), Timescales for radiation belt electron acceleration and loss due to resonant wave-particle interactions: 1. Theory, *J. Geophys. Res.*, 112, A04206, doi:10.1029/2006JA011801.
- Takahashi, K., and A. Y. Ukhorskiy (2007), Solar wind control of Pc5 pulsation power at geosynchronous orbit, *J. Geophys. Res.*, 112, A11205, doi:10.1029/2007JA012483.
- Tanskanen, E. I. (2009), A comprehensive high-throughput analysis of substorms observed by IMAGE magnetometer network: Years 1993–2003 examined, *J. Geophys. Res.*, 114, A05204, doi:10.1029/2008JA013682.
- Thorne, R. M. (2010), Radiation belt dynamics: The importance of wave-particle interactions, *Geophys. Res. Lett.*, 37, L22107, doi:10.1029/2010GL044990.
- Tsonis, A. A. (2001), Probing the linearity and nonlinearity in the transitions of the atmospheric circulation, *Nonlinear Processes Geophys.*, 8, 341–345, doi:10.5194/npg-8-341-2001.
- Turner, D. L., Y. Shprits, M. Hartinger, and V. Angelopoulos (2012), Explaining sudden losses of outer radiation belt electrons during geomagnetic storms, *Nat. Phys.*, 8, 208–212, doi:10.1038/nphys2185.
- Ukhorskiy, A. Y., M. I. Sitnov, A. S. Sharma, B. J. Anderson, S. Ohtani, and A. T. Y. Lui (2004), Data-derived forecasting model for relativistic electron intensity at geosynchronous orbit, *Geophys. Res. Lett.*, 31, L09806, doi:10.1029/2004GL019616.
- Ukhorskiy, A. Y., K. Takahashi, B. J. Anderson, and H. Korth (2005), Impact of toroidal ULF waves on the outer radiation belt electrons, *J. Geophys. Res.*, 110, A10202, doi:10.1029/2005JA011017.
- Ukhorskiy, A. Y., B. J. Anderson, P. C. Brandt, and N. A. Tsyganenko (2006), Storm time evolution of the outer radiation belt: Transport and losses, *J. Geophys. Res.*, 111, A11S03, doi:10.1029/2006JA011690.
- 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.
- Vennestrøm, S. (1999), Dayside magnetic ULF power at high latitudes: A possible long-term proxy for the solar wind velocity? *J. Geophys. Res.*, 104(A5), 10,145–10,157, doi:10.1029/1999JA900015.
- Wing, S., J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman, K. Costello, M. Balikhin, and K. Takahashi (2005a), *Kp* forecast models, *J. Geophys. Res.*, 110, A04203, doi:10.1029/2004JA010500.
- Wing, S., J. R. Johnson, P. T. Newell, and C.-I. Meng (2005b), Dawn-dusk asymmetries, ion spectra, and sources in the northward interplanetary magnetic field plasma sheet, *J. Geophys. Res.*, 110, A08205, doi:10.1029/2005JA011086.
- Wynner, A. D. (1978), A definition of conditional mutual information for arbitrary ensembles, *Info. Control*, 38, 51–59, doi:10.1016/S0019-9958(78)90026-8.