

Andrews University

Digital Commons @ Andrews University

Faculty Publications

1-23-2018

Transfer Entropy and Cumulant Based Cost as Measures of Nonlinear Causal Relationships in Space Plasmas: Applications to D St.

Jay R. Johnson
Andrews University, jrj@andrews.edu

Simon Wing
Johns Hopkins University

Enrico Camporeale
Center for Mathematics and Computer Science, Amsterdam, Netherlands

Follow this and additional works at: <https://digitalcommons.andrews.edu/pubs>



Part of the [Astrophysics and Astronomy Commons](#)

Recommended Citation

Johnson, Jay R.; Wing, Simon; and Camporeale, Enrico, "Transfer Entropy and Cumulant Based Cost as Measures of Nonlinear Causal Relationships in Space Plasmas: Applications to D St." (2018). *Faculty Publications*. 611.

<https://digitalcommons.andrews.edu/pubs/611>

This Article is brought to you for free and open access by Digital Commons @ Andrews University. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Digital Commons @ Andrews University. For more information, please contact repository@andrews.edu.



1 **Transfer entropy and cumulant based cost as**
2 **measures of nonlinear causal relationships in space**
3 **plasmas: applications to D_{st}**

Jay R. Johnson

4 Andrews University, Berrien Springs, MI, USA

Simon Wing

5 The Johns Hopkins University, Applied Physics Laboratory, Laurel,

6 Maryland, USA

Enrico Camporeale

7 Center for Mathematics and Computer Science (CWI), Amsterdam, The

8 Netherlands

Jay R. Johnson, Andrews University, Berrien Springs, MI, 49104, USA. (jrj@andrews.edu)

Simon Wing, The Johns Hopkins University, Applied Physics Laboratory, Laurel, Maryland,
20723, USA (simon.wing@jhuapl.edu)

Enrico Camporeale, Center for Mathematics and Computer Science (CWI), Amsterdam, The
Netherlands (e.camporeale@cw.nl)



9 **Abstract.** It is well known that the magnetospheric response to the so-
10 lar wind is nonlinear. Information theoretical tools such as mutual informa-
11 tion, transfer entropy, and cumulant based analysis are able to characterize
12 the nonlinearities in the system. Using cumulant based cost, we show that
13 nonlinear significance of D_{st} peaks at 3 – 12 hours lags that can be attributed
14 to VBs , which also exhibit similar behavior. However, the nonlinear signif-
15 icance that peaks at lags 25, 50, and 90 hours can be attributed to internal
16 dynamics, which may be related to the relaxation of the ring current. These
17 peaks are absent in the linear and nonlinear self-significance of VBs . Our
18 analysis with mutual information and transfer entropy show that both meth-
19 ods can establish that there are a strong correlation and transfer of infor-
20 mation from V_{sw} to D_{st} at a time scale that is consistent with that obtained
21 from the cumulant based analysis. However, mutual information also shows
22 that there is a strong correlation in the backward direction, from D_{st} to V_{sw} ,
23 which is counterintuitive. In contrast, transfer entropy shows that there is
24 no or little transfer of information from D_{st} to V_{sw} , as expected because it
25 is the solar wind that drives the magnetosphere, not the other way around.
26 Our case study demonstrates that these information theoretical tools are quite
27 useful for space physics studies because these tools can uncover nonlinear
28 dynamics that cannot be seen with the traditional analyses and models that
29 assume linear relationships.



1. Introduction

30 One of the most practically important concepts in dynamical systems is the notion of
31 causality. It is particularly useful to organize observational datasets according to causal
32 relationships in order to identify variables that drive the dynamics. Understanding causal
33 dependencies can also help to simplify descriptions of highly complex physical processes
34 because it constrains the coupling functions between the dynamical variables. Analysis
35 of those coupling functions can lead to simplification of the underlying physical processes
36 that are most important for driving the system. It is particularly useful from a practi-
37 cal standpoint to understand causal dependencies in systems involving natural hazards
38 because monitoring of causal variables is closely linked with warning.

A common method to establish causal dependencies in a data stream of two variables, e.g., $[a(t)]$ and $[b(t)]$, is to apply linear correlation studies such as *Strangeway et al.* [2005], which showed the relationship between downward Poynting flux and ion outflows. Causal relationships are typically identified by considering a time-shifted correlation function

$$\lambda_{ab}(\tau) \triangleq \frac{\langle a(t)b(t+\tau) \rangle - \langle a \rangle \langle b \rangle}{\sqrt{\langle a^2 \rangle - \langle a \rangle^2} \sqrt{\langle b^2 \rangle - \langle b \rangle^2}} \quad (1)$$

39 where $\langle \dots \rangle$ is an ensemble average obtained by drawing samples at a set of measurement
40 times, $\{t_0, t_1, \dots, t_N\}$. For example, [*Borovsky et al.*, 1998] used such a method to iden-
41 tify relationships between solar wind variables and plasma sheet variables. The causal
42 dependency that the plasma sheet responds to changes in the solar wind can be identified
43 from the time-shift of the peak of the cross correlation indicating a response time. From
44 this type of analysis it can be found that the plasma sheet generally responds from the



45 tail to the inner magnetosphere consistent with the notion of earthward convection. Such
46 analysis has been particularly useful to help understand plasma sheet transport.

47 However, the procedure of detecting causal relationships based on linear cross-
48 correlation suffers from a number of limitations. First it should be noted that the statisti-
49 cal accuracy of the correlation function is limited by the resolution and length of the data
50 stream. Second, the linear time series analysis ignores nonlinear correlations, which may
51 be important for energy transfer in the magnetospheric system. For example, substorms
52 are believed to involve storage and release of energy in the magnetotail, which is a highly
53 nonlinear response. Similarly, magnetosphere-ionosphere coupling may also be highly non-
54 linear involving the nonlinear development of accelerating potentials along auroral field
55 lines and nonlinear current-voltage relationships. Third, the cross-correlation may not
56 be a particularly clear measure when there are multiple peaks or if there is little or no
57 asymmetry in the forward [i.e., $\lambda_{ab}(\tau)$] and backward directions [i.e., $\lambda_{ba}(\tau) = \lambda_{ab}(-\tau)$].
58 Finally, the cross-correlation does not provide any way to clearly distinguish between two
59 variables that are passively correlated because of a common driver rather than causally
60 related.

61 In the remainder of this paper, we will discuss other methods to identify causal rela-
62 tionships based on entropy based discriminating statistics such as mutual information and
63 transfer entropy. We will also discuss the cumulant-based method. We will illustrate the
64 shortcomings and strengths of the various methods for studying causality with examples
65 from nonlinear dynamics and space physics.



2. Linear vs Nonlinear Dependency

66 It is well known that the magnetosphere responds to variation in the solar wind param-
67 eters [*Clauer et al.*, 1981; *Baker et al.*, 1983; *Crooker and Gringauz*, 1993; *Papitashvili*
68 *et al.*, 2000; *Wing and Johnson*, 2015; *Johnson and Wing*, 2015; *Wing et al.*, 2016], and
69 it has been established that the magnetosphere has a significant linear response to the
70 solar wind. However, it is also expected that the magnetosphere has a nonlinear behavior
71 due to internal dynamics [*Wing et al.*, 2005; *Johnson and Wing*, 2005]. For example, the
72 internal dynamics associated with loading and unloading of magnetic energy associated
73 with storms and substorms is nonlinear [e.g., *Johnson and Wing*, 2014, and references
74 therein]. Indeed, the data analysis of *Bargatze et al.* [1985] indicated that the dynamical
75 response of the magnetosphere to solar wind input could not be entirely understood using
76 linear prediction filters.

Suppose that we consider a set of variables \mathbf{a} and \mathbf{b} which could be vectors of variables measured in time and we would like to measure their dependency. Instead of considering the covariance matrix/correlation function, we consider a more general measure of dependency between an input and output is obtained by considering whether

$$P(\mathbf{a}, \mathbf{b}) \stackrel{?}{=} P(\mathbf{a})P(\mathbf{b}). \quad (2)$$

77 where $P(\mathbf{a}, \mathbf{b})$ is the joint probability of input \mathbf{a} and output \mathbf{b} while $P(\mathbf{a})$ and $P(\mathbf{b})$ are
78 the probability of \mathbf{a} and \mathbf{b} respectively. If the relationship holds, then the variables \mathbf{a}
79 and \mathbf{b} are independent. For all other cases, there is some measure of dependency. In the
80 case where the system output is completely known given the input, $P(\mathbf{a}, \mathbf{b}) = P(\mathbf{a})$. The
81 advantage of considering Equation 2 is that it is possible to detect the presence of higher



82 order nonlinear dependencies between the input and output even in the absence of linear
83 dependencies [Gershenfeld, 1998].

2.1. Mutual Information and Cumulant based cost

84 Mutual information and cumulant-based cost are two useful measures that quantify
85 Eq. 2. Mutual information has the advantage that in the limit of Gaussian joint proba-
86 bility distributions, it may be simply related to the correlation coefficient $C_{ab}(\tau)$ defined
87 in equation 1 [Li, 1990]. Cumulants have the advantage of good statistics for limited
88 datasets and noisy systems [Deco and Schürmann, 2000]. Moreover, for high-dimensional
89 systems it is more efficient to compute moments of the data rather than try to construct
90 the probability density function.

Correlation studies also only detect linear correlations, so if the feedback involves non-
linear processes (highly likely in this case) then their usefulness may be seriously lim-
ited. Alternatively, entropy-based measures such as mutual information [Prichard and
Theiler, 1995] and cumulants [Johnson and Wing, 2005] are useful for detecting linear
as well as nonlinear correlations. The mutual information is constructed from the proba-
bility distribution function of the variables and may be computed using an quantization
procedure where data is binned such that the samples $[a(t)]$ are assigned discrete values
 $\hat{a} \in \{a_1, a_2, \dots, a_n\}$ of an alphabet \aleph_1 and $[b(t)]$ is assigned discrete values $\hat{b} \in \{b_1, b_2, \dots, b_m\}$
of an alphabet \aleph_2 . The *ad hoc* time-shifted mutual entropy

$$\mathcal{M}_{ab}(\tau) \triangleq \sum_{\hat{a} \in \aleph_1, \hat{b} \in \aleph_2} p(\hat{a}(t + \tau), \hat{b}(t)) \log \left(\frac{p(\hat{a}(t + \tau), \hat{b}(t))}{p(\hat{a})p(\hat{b})} \right) \quad (3)$$

91 has been used as an indicator of causality, but suffers from the same problems as time-
92 shifted cross correlation when it has multiple peaks and long range correlations.



Similarly, examination of time-shifted cumulants could be used as an indicator of causality in a nonlinear system. In this case, we can define a discriminating statistic

$$D^C = \sum_{q=1}^{\infty} \sum_{i_1, \dots, i_q \in \Pi_q} K_{1i_2 \dots i_q}^2 \quad (4)$$

where

$$\begin{aligned} K_i &= C_i = \langle z_i \rangle \\ K_{ij} &= C_{ij} - C_i C_j = \langle z_i z_j \rangle - \langle z_i \rangle \langle z_j \rangle \\ K_{ijk} &= C_{ijk} - C_{ij} C_k - C_{jk} C_i - C_{ik} C_j + 2C_i C_j C_k \\ K_{ijkl} &= C_{ijkl} - C_{ijk} C_l - C_{ijl} C_k - C_{ilk} C_j - C_{ljk} C_i \\ &\quad - C_{ij} C_{kl} - C_{il} C_{kj} - C_{ik} C_{jl} + 2(C_{ij} C_k C_l \\ &\quad + C_{ik} C_j C_l + C_{il} C_j C_k + C_{jk} C_i C_l + C_{jl} C_i C_k \\ &\quad + C_{kl} C_i C_j) - 6C_i C_j C_k C_l \end{aligned} \quad (4)$$

are the cumulants

$$C_{i \dots j} = \int d\mathbf{z} P(\mathbf{z}) z_i \dots z_j \equiv \langle z_i \dots z_j \rangle \quad (5)$$

of the joint probability distribution for variables z_1, \dots, z_j .

With only two variables, a and b , defined above, we can consider the cost function

$$D_{a,b}^C(\tau) = D_C(a(t), b(t + \tau)) \quad (6)$$

The presence of nonlinear dependence has been identified by comparing the cumulant cost for a time series with the cumulant based cost of surrogate time series, which are constructed to have the same linear correlations as in [Johnson and Wing, 2005]). Significance measures the difference in the discriminating statistic from the mean of the discriminating statistic of the surrogates in terms of the spread of the surrogates, σ .

In Section 3, we will show an application of cumulant based analysis to the disturbance storm-time index (D_{st}). In principle, the cross-correlation, mutual information, and cumulant-based cost should be independent of the selection of measurement points if the system is stationary; therefore, time stationarity can be examined by comparing



103 these discriminating statistics for groups of measurements drawn from different windows
 104 of time as in [Johnson and Wing, 2005].

2.2. Transfer entropy

Another method for determining causality is the one-sided transfer entropy [Schreiber, 2000; ?; ?; Wing et al., 2016], which is based upon the conditional mutual information

$$\mathcal{M}_C(x, y|z) \triangleq \sum_{x \in \mathbb{N}_1} \sum_{y \in \mathbb{N}_2} \sum_{z \in \mathbb{N}_3} p(x, y, z) \log \left(\frac{p(x, y, z)p(z)}{p(x, z)p(y, z)} \right) \quad (7)$$

105 The conditional mutual information measures the dependence of two variables, x and y ,
 106 given a conditioner variable, z . If either x or y are dependent on z the mutual information
 107 between x and y is reduced, and this reduction of information provides a method to
 108 eliminate coincidental dependence, or conversely to identify causal dependence.

Transfer entropy considers the conditional mutual information between two variables using the past history of one of the variables as the conditioner.

$$\mathcal{T}_{a \rightarrow b}(\tau) = \sum_{\hat{a} \in \mathbb{N}_1} \sum_{\hat{a}^{(k)} \in \mathbb{N}_1^{(k)}} \sum_{\hat{b} \in \mathbb{N}_2} p(\hat{a}(t + \tau), \hat{a}^{(k)}(t), \hat{b}(t)) \log \left(\frac{p(\hat{a}(t + \tau) | \hat{a}^{(k)}(t), \hat{b}(t))}{p(\hat{a}(t + \tau) | \hat{a}^{(k)}(t))} \right) \quad (8)$$

109 where $\hat{a}^{(k)}(t) = [\hat{a}(t), \hat{a}(t - \Delta), \dots, \hat{a}(t - (k - 1)\Delta)]$. The standard definition of transfer
 110 entropy takes $k = 1$ (no lag), but keeping a higher embedding dimension could in prin-
 111 ciple provide a more precise measure (for example, if a has periodicity a dimension of 2
 112 may provide better prediction of future values of a from its past time series and therefore
 113 lower the transfer entropy. Transfer entropy as a discriminating statistic has the following
 114 advantages. First in the absence of information flow from a to b (i.e., $a(t + \tau)$ has no
 115 additional dependence from $b(t)$ beyond what is known from the past history of $a^{(k)}(t)$)
 116 $p(\hat{a}(t + \tau) | \hat{a}^{(k)}(t), \hat{b}(t)) = p(\hat{a}(t + \tau) | \hat{a}^{(k)}(t))$ and the transfer entropy vanishes. The transfer
 117 entropy is also highly directional so that $\mathcal{T}_{a \rightarrow b} \neq \mathcal{T}_{b \rightarrow a}$. The advantage can be clearly



118 seen for dynamical systems where variables are forward differenced and the transfer en-
119 tropy is clearly one-sided while mutual information and correlation functions can even be
120 symmetric [Schreiber, 2000]. This measure also accounts for static internal correlations,
121 which can be used to determine whether two variables are driven by a common driver or
122 whether the variable b is causally driving the variable a .

3. Application to space weather: D_{st} analysis

123 D_{st} (disturbance storm time index) is an hourly index that gives a measure of the
124 strength of the ring current that, in turn, provides a measure of the dynamics of geo-
125 magnetic storms [Dessler and Parker, 1959]. Because of its global nature, D_{st} is often
126 used as one of the several indices that represent the state of the magnetosphere. When
127 plasma sheet ions are injected into the Earth inner magnetosphere, they drift westward
128 around the Earth, forming the ring current. Studies have shown that the substorm occur-
129 rence rate increases with solar wind velocity (high speed streams) [e.g., Kissinger *et al.*,
130 2011; Newell *et al.*, 2016]. An increase in the solar wind electric field, VB_z , can increase
131 the dawn-dusk electric field in the magnetotail, which in turn determines the amount of
132 plasma sheet particles that move to the inner magnetosphere [e.g., Friedel *et al.*, 2001].

133 For the present study, we examine the relationships between solar wind velocity (V_{sw})
134 and VB_s ($V_{sw} \times$ southward IMF B_z) with D_{st} . We use D_{st} records in the period
135 1974 – 2001 obtained from Kyoto University World Data Center for Geomagnetism
136 (<http://swdcwww.kugi.kyoto-u.ac.jp/index.html>). The corresponding solar wind data
137 are obtained from IMP-8, ACE, WIND, ISEE1, and ISEE3 observations. The ACE
138 SWEPAM and MAG data; and the WIND MAG data are obtained from CDAWeb
139 (<http://cdaweb.gsfc.nasa.gov/>). The WIND 3DP data are obtained from the 3DP team



140 directly. The ISEE1 and ISEE3 data are obtained from UCLA (these datasets are also
141 available at NASA NSSDC [<http://nssdc.gsfc.nasa.gov/space/>]). The IMP8 data come
142 directly from the IMP teams. The solar wind is propagated with minimum variance tech-
143 nique [Weimer *et al.*, 2003] to GSM $(X, Y, Z) = (17, 0, 0) R_E$ to produce 1-min files,
144 from which hourly averaged solar wind parameters are constructed.

3.1. Cumulant based analysis

Section 2.1 presents the method of cumulant based cost. Here, we show an application of cumulant based cost to detect nonlinear dynamics in D_{st} . We consider the forward coupling between a solar wind variable such as VB_s and D_{st} , which characterizes the ring current response to the solar wind driver. We therefore consider the nonlinear cross-correlations of the vector

$$\mathbf{c}(t, \tau) = \{VB_s(t), D_{st}(t + \tau)\} = \{z_1, z_2\} \quad (9)$$

145 The generalization of cost is based on realizations of $\{z_1, z_2\}$. In this case, each variable
146 is Gaussianized with unit variance to eliminate static nonlinearities (i.e. higher order
147 self-correlations in VB_s and D_{st} are eliminated so that the cost measures only cross-
148 dependence between VB_s and D_{st}).

149 In Figure 1 we plot the significance obtained from the year 1999 as a function of time
150 delay, τ . Significance extracted from $\{VB_s(t), D_{st}(t + \tau)\}$ and $\{VB_s(t), VB_s(t + \tau)\}$
151 for 1999 are plotted in panels (a) and (b), respectively. It should be noted that there
152 is a strong linear response at around 3 hour time delay. As shown in Figure 1a, there
153 is a clear nonlinear response with peaking around 3–10, 25, 50 and 90 hours lasting for
154 approximately 1 week. In contrast, in Figure 1b, the nonlinearity only has one broad peak



155 around 3 – 12 hours in the self-significance for VBs , suggesting that the nonlinear and
156 linear peaks at $\tau = 3\text{--}12$ hours in in Figure 1a i may be associated with VBs . We will
157 revisit the solar wind causal relationship with D_{st} using transfer entropy in Section 3.2.

158 The absence of the nonlinear peaks at $\tau = 25, 50,$ and 90 hours in the self-significance
159 for VBs (Figure 1b) suggest that these nonlinearities in $\{VBs(t), D_{st}(t+\tau)\}$ are related to
160 internal magnetospheric dynamics. As the D_{st} index is thought to reflect storm activity,
161 it is reasonable that nonlinear significance would decay on the order of 1 week as storms
162 commonly last around that time. The strong nonlinear responses at $\tau = 25, 50,$ and 90
163 hours are likely related to multiple modes of relaxation of the ring current following the
164 commencement of storms. It should also be noted that other nonlinearities detected by
165 even higher order cumulants may also be present; however, the calculation demonstrates
166 the nonlinear nature of the underlying dynamics.

167 A common scenario for storm-ring current interaction is the following. A storm com-
168 presses the magnetosphere and intensifies the magnetic field in the magnetosphere and
169 energetic particles are injected into the ring current region. Conservation of magnetic mo-
170 ment implies that anisotropies develop in the ring current and plasma sheet. Anisotropy
171 drives the ring current plasma unstable to ion cyclotron waves. The ion cyclotron waves
172 scatter energetic ions into the loss cone so that they are lost from the ring current. Non-
173 linear interaction between waves and particles keeps the plasma near marginal instability
174 with a steady loss of energetic particles due to wave-particle scattering. The typical
175 time-scale for pitch-angle scattering in the ring current is the order of 24 hours. We can
176 speculate that the nonlinear response that is detected with the cumulant-based approach
177 is likely the relaxation of the ring current due to wave-particle interactions.



3.2. Transfer entropy

178 As mentioned in Section 2.2, transfer entropy gives a measure of how much information
179 is transferred from one variable to another. We have applied transfer entropy and mutual
180 information to the relationship between the V_{sw} and D_{st} for the period 1974 – 2001. The
181 result is shown in Figure 2. Note that the mutual information measure suggests strong
182 correlations between prior values of D_{st} and V_{sw} . This finding suggests that D_{st} could be
183 a driver of V_{sw} , which is counterintuitive. On the other hand, the transfer entropy clearly
184 shows that this information transfer in the backward direction ($D_{st} \rightarrow V_{sw}$) does not rise
185 above the noise level (the horizontal blue lines indicate mean and standard deviation of
186 100 surrogate data sets where the data was randomly reordered.) This result is expected
187 because it is the solar wind that drives the magnetosphere, not the other way around. The
188 transfer of information from V_{sw} to D_{st} peaks at $\tau = 8 - 11$ hours. The cumulant based
189 analysis in Section 3.1 shows that the response of D_{st} to VBs has similar time scale. The
190 analysis presented here illustrates the power of the transfer entropy for accessing causality.

4. Summary

191 We recently used mutual information, transfer entropy, and conditional mutual infor-
192 mation to discover the solar wind drivers of the outer radiation belt electrons [*Wing et al.*,
193 2016]. Because V_{sw} anticorrelates with solar wind density (n_{sw}), it is hard to isolate the
194 effects of V_{sw} on radiation belt electrons, given n_{sw} and vice versa. However, using condi-
195 tional mutual information, we were able to determine the information transfer from n_{sw}
196 or any other solar wind parameters to radiation belt electrons, given V_{sw} (or any other
197 solar wind parameters). We also showed that the triangle distribution in the radiation
198 belt electron vs. solar wind velocity plot [*Reeves et al.*, 2011] can be understood better



199 when we consider that V_{sw} and n_{sw} transfer information to radiation belt electrons with
200 2 days and 0 day (< 24 hr) lags, respectively.

201 As a follow up to *Wing et al.* [2016], the present study demonstrates further how in-
202 formation theoretical tools can be useful for space physics and space weather studies.
203 Cumulant based analysis can be used to distinguish internal vs. external driving of the
204 system. Both mutual information and transfer entropy give a measure of shared infor-
205 mation between two variables (or vectors). However, unlike mutual information, transfer
206 entropy is highly directional. To illustrate, we apply mutual information, transfer entropy,
207 and cumulant based analysis to investigate the dynamics of D_{st} index.

208 Our analysis with mutual information and transfer entropy indicates that there are
209 strong linear and nonlinear correlations and transfer of information, respectively, in the
210 forward direction between V_{sw} and D_{st} ($V_{sw} \rightarrow D_{st}$). However, mutual information indi-
211 cates that there is also a strong correlation in the backward direction ($D_{st} \rightarrow V_{sw}$), which
212 is puzzling and counterintuitive. In contrast, the transfer entropy indicates that there is
213 no information transfer in the backward direction ($D_{st} \rightarrow V_{sw}$), as expected because it is
214 the solar wind that drives the magnetosphere, not the other way around. The transfer of
215 information from V_{sw} to D_{st} peaks at $\tau = 8 - 11$ hours.

216 Using the cumulant-based significance, we have established that the underlying dynam-
217 ics of D_{st} is in general nonlinear exhibiting a quasiperiodicity which is detectable only if
218 nonlinear correlations are taken into account. The strong nonlinear responses of D_{st} to
219 VBS at $\tau = 25, 50,$ and 90 hours are likely related to multiple modes of relaxation of the
220 ring current following the commencement of storms. The nonlinearities at $\tau = 3 - 12$
221 hours are not caused by internal dynamics but rather by the solar wind driver. This time



222 scale is consistent with the time scale for the information transfer from the solar wind to
223 D_{st} obtained from transfer entropy analysis.

224 Although linear models are useful, our results indicate that these models have to be
225 used with cautions because solar wind – magnetosphere system is inherently nonlinear.
226 Hence, nonlinearities generally need to be taken into account in order to describe the
227 system accurately. Local-linear models (which include slow evolution of parameters) may
228 be able to handle some nonlinearities, but it is expected that these local-linear models
229 would have difficulties if the dynamics suddenly and rapidly change.

230 **Acknowledgments.** All the derived data products in this paper are available upon
231 request by email (simon.wing@jhuapl.edu). Simon Wing acknowledges support from
232 JHU/APL Janney Fellowship. Jay R. Johnson acknowledges support from NASA
233 Grants (NNH11AR07I, NNX14AM27G, NNX14AY20I, NNX16AC39G), NSF Grants
234 (ATM0902730, AGS-1203299), and DOE contract DE-AC02-09CH11466. We thank
235 James M. Weygand for the solar wind data processing. The raw solar wind data from
236 ACE, Wind, ISEE1 and ISEE3 were obtained from NASA CDAW and NSSDC.

References

- 237 Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones, B. T. Tsurutani, E. J. Smith, and
238 S.-I. Akasofu (1983), An isee 3 high time resolution study of interplanetary parameter
239 correlations with magnetospheric activity, *Journal of Geophysical Research*, *88*(A8),
240 6230, doi:10.1029/ja088ia08p06230.
- 241 Bargatze, L. F., D. N. Baker, E. W. Hones, and R. L. McPherron (1985), Magnetospheric
242 impulse response for many levels of geomagnetic activity, *J. Geophys. Res.*, *90*, 6387–



243 6394.

244 Borovsky, J. E., M. F. Thomsen, and R. C. Elphic (1998), The driving of the plasma
245 sheet by the solar wind, *J. Geophys. Res.*, *103*, 17,617–17,640, doi:10.1029/97JA02986.

246 Clauer, C. R., R. L. McPherron, C. Searls, and M. G. Kivelson (1981), Solar wind control
247 of auroral zone geomagnetic activity, *Geophysical Research Letters*, *8*(8), 915–918, doi:
248 10.1029/g1008i008p00915.

249 Crooker, N. U., and K. I. Gringauz (1993), On the low correlation between long-term
250 averages of solar wind speed and geomagnetic activity after 1976, *Journal of Geophysical*
251 *Research*, *98*(A1), 59, doi:10.1029/92ja01978.

252 Deco, G., and B. Schürmann (2000), *Information Dynamics*, Springer.

253 Dessler, A. J., and E. N. Parker (1959), Hydromagnetic theory of magnetic storms, *J.*
254 *Geophys. Res.*, *64*(2239-2259).

255 Friedel, R. H. W., H. Korth, M. G. Henderson, M. F. Thomsen, and J. D. Scudder (2001),
256 Plasma sheet access to the inner magnetosphere, *Journal of Geophysical Research: Space*
257 *Physics*, *106*(A4), 5845–5858, doi:10.1029/2000ja003011.

258 Gershenfeld, N. (1998), *The Nature of Mathematical Modeling*, Cambridge University
259 Press, Cambridge, chapter.

260 Johnson, J. R., and S. Wing (2005), A solar cycle dependence of nonlinearity in magneto-
261 spheric activity, *Journal of Geophysical Research*, *110*(A4), doi:10.1029/2004ja010638.

262 Johnson, J. R., and S. Wing (2014), External versus internal triggering of substorms:
263 An information-theoretical approach, *Geophysical Research Letters*, *41*(16), 5748–5754,
264 doi:10.1002/2014gl060928.



- 265 Johnson, J. R., and S. Wing (2015), The dependence of the strength and thickness of
266 field-aligned currents on solar wind and ionospheric parameters, *Journal of Geophysical*
267 *Research: Space Physics*, 120(5), 3987?4008, doi:10.1002/2014ja020312.
- 268 Kissinger, J., R. L. McPherron, T.-S. Hsu, and V. Angelopoulos (2011), Steady magne-
269 topheric convection and stream interfaces: Relationship over a solar cycle, *Journal of*
270 *Geophysical Research: Space Physics*, 116(A5), n/a?n/a, doi:10.1029/2010ja015763.
- 271 Li, W. (1990), Mutual information functions versus correlation functions, *J. Stat. Phys.*,
272 60, 823.
- 273 Newell, P., K. Liou, J. Gjerloev, T. Sotirelis, S. Wing, and E. Mitchell (2016), Substorm
274 probabilities are best predicted from solar wind speed, *Journal of Atmospheric and*
275 *Solar-Terrestrial Physics*, 146, 28?37, doi:10.1016/j.jastp.2016.04.019.
- 276 Papitashvili, V. O., N. E. Papitashvili, and J. H. King (2000), Solar cycle effects in
277 planetary geomagnetic activity: Analysis of 36-year long omni dataset, *Geophysical*
278 *Research Letters*, 27(17), 2797?2800, doi:10.1029/2000gl000064.
- 279 Prichard, D., and J. Theiler (1995), Generalized redundancies for time series analysis,
280 *Physica D: Nonlinear Phenomena*, 84(3?4), 476 – 493, doi:10.1016/0167-2789(95)00041-
281 2.
- 282 Reeves, G. D., S. K. Morley, R. H. W. Friedel, M. G. Henderson, T. E. Cayton, G. Cun-
283 ningham, J. B. Blake, R. A. Christensen, and D. Thomsen (2011), On the relation-
284 ship between relativistic electron flux and solar wind velocity: Paulikas and blake
285 revisited, *Journal of Geophysical Research: Space Physics*, 116(A2), n/a?n/a, doi:
286 10.1029/2010ja015735.



287 Schreiber, T. (2000), Measuring information transfer, *Phys. Rev. Lett.*, *85*, 461–464, doi:

288 10.1103/PhysRevLett.85.461.

289 Strangeway, R., J. R. E. Ergun, Y.-J. Su, C. W. Carlson, and R. C. Elphic (2005),

290 Factors controlling ionospheric outflows as observed at intermediate altitudes, *Journal*

291 *of Geophysical Research*, *110*(A3), doi:10.1029/2004ja010829.

292 Weimer, D. R., D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness,

293 C. W. Smith, and J. Watermann (2003), Predicting interplanetary magnetic field (imf)

294 propagation delay times using the minimum variance technique, *Journal of Geophysical*

295 *Research*, *108*(A1), doi:10.1029/2002ja009405.

296 Wing, S., and J. R. Johnson (2015), Theory and observations of upward field-aligned

297 currents at the magnetopause boundary layer, *Geophysical Research Letters*, *42*(21),

298 9149?9155, doi:10.1002/2015gl065464.

299 Wing, S., J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman,

300 K. Costello, M. Balikhin, and K. Takahashi (2005), Kp forecast models, *Journal of*

301 *Geophysical Research*, *110*(A4), doi:10.1029/2004ja010500.

302 Wing, S., J. R. Johnson, E. Camporeale, and G. D. Reeves (2016), Information theoret-

303 ical approach to discovering solar wind drivers of the outer radiation belt, *Journal of*

304 *Geophysical Research: Space Physics*, doi:10.1002/2016ja022711.

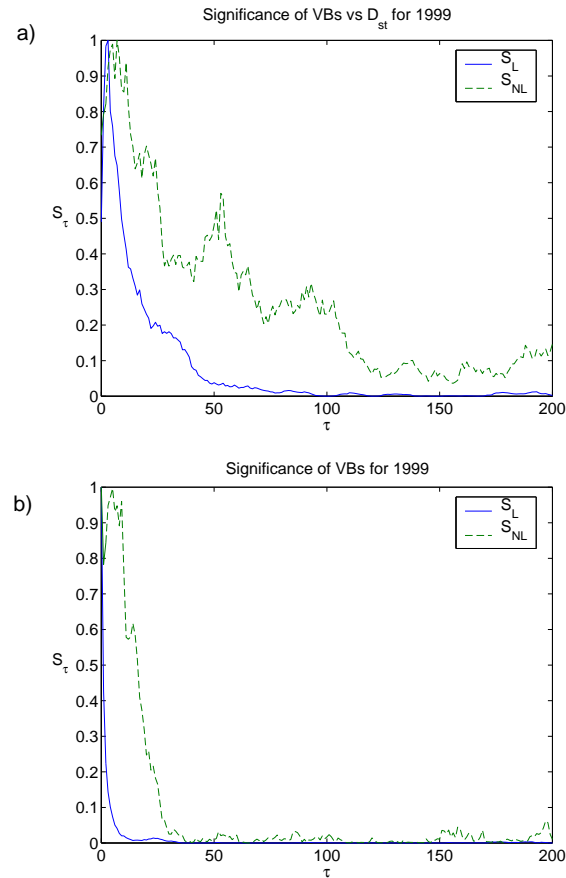


Figure 1. Significance extracted from (a) $\{VBs(t), D_{st}(t-\tau)\}$ and (b) $\{VBs(t), VBs(t-\tau)\}$ for 1999. It should be noted that there is a strong linear response at around 3 hour time delay. There is a clear nonlinear response with a strong peak around 50 hours lasting for approximately 1 week. The longterm nonlinear response is absent in the solar wind data indicating that the longterm nonlinear correlations between VBs and D_{st} are the result of internal magnetospheric dynamics.

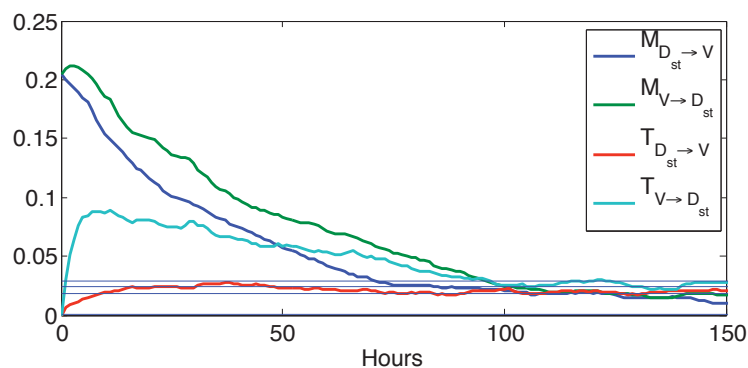


Figure 2. Comparison of mutual information and transfer entropy measures to determine causal driving of the magnetosphere as characterized by D_{st} . Note that causal driving appears to peak somewhat later (11 hours) than indicated by mutual information (2 hours) indicating that internal dynamics likely are very important initially. The backward transfer entropy is below the noise level for all values indicating that D_{st} in no way influences the upstream solar wind velocity. Such a conclusion could not be inferred from the mutual information measure.