Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 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, Berrien Springs, MI, USA

Simon Wing

The Johns Hopkins University, Applied Physics Laboratory, Laurel, Maryland, USA

Enrico Camporeale

Center for Mathematics and Computer Science (CWI), Amsterdam, The 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@cwi.nl)

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

Abstract. It is well known that the magnetospheric response to the solar wind is nonlinear. Information theoretical tools such as mutual information, transfer entropy, and cumulant based analysis are able to characterize the nonlinearities in the system. Using cumulant based cost, we show that nonlinear significance of D_{st} peaks at $3 - 12$ hours lags that can be attributed to VBs, which also exhibit similar behavior. However, the nonlinear significance that peaks at lags 25, 50, and 90 hours can be attributed to internal dynamics, which may be related to the relaxation of the ring current. These peaks are absent in the linear and nonlinear self-significance of VBs. Our analysis with mutual information and transfer entropy show that both methods can establish that there are a strong correlation and transfer of information from V_{sw} to D_{st} at a time scale that is consistent with that obtained from the cumulant based analysis. However, mutual information also shows that there is a strong correlation in the backward direction, from D_{st} to V_{sw}, which is counterintuitive. In contrast, transfer entropy shows that there is no or little transfer of information from D_{st} to V_{sw}, as expected because it is the solar wind that drives the magnetosphere, not the other way around. Our case study demonstrates that these information theoretical tools are quite useful for space physics studies because these tools can uncover nonlinear dynamics that cannot be seen with the traditional analyses and models that assume linear relationships.

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

1. Introduction

One of the most practically important concepts in dynamical systems is the notion of causality. It is particularly useful to organize observational datasets according to causal relationships in order to identify variables that drive the dynamics. Understanding causal dependencies can also help to simplify descriptions of highly complex physical processes because it constrains the coupling functions between the dynamical variables. Analysis of those coupling functions can lead to simplification of the underlying physical processes that are most important for driving the system. It is particularly useful from a practical standpoint to understand causal dependencies in systems involving natural hazards 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 ouflows. Causal relationships are typically identified by considering a time-shifted correlation function

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

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

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

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

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

2. Linear vs Nonlinear Dependency

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

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

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

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

2.1. Mutual Information and Cumulant based cost

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

Correlation studies also only detect linear correlations, so if the feedback involves nonlinear processes (highly likely in this case) then their usefulness may be seriously limited. 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 probability 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 $\hat{a} \in \{a_1, a_2, ..., a_n\}$ of an alphabet $\aleph_1$ and $[b(t)]$ is assigned discrete values $\hat{b} \in \{b_1, b_2, ..., b_m\}$ of an alphabet $\aleph_2$. The *ad hoc* time-shifted mutual entropy

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

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

Annales Geophysicae Discussions

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

$$D^C = \sum_{q=1}^{\infty} \sum_{i_1,\ldots,i_q \in \Pi_q} K^2_{1 i_2 \ldots i_q} \tag{4}$$

where

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

are the cumulants

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

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

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

$$D^C_{a,b}(\tau) = D_C(a(t), b(t+\tau)) \tag{6}$$

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

these discriminating statistics for groups of measurements drawn from different windows 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

$$\mathcal{M}_C(x,y|z) \triangleq \sum_{x\in\aleph_1}\sum_{y\in\aleph_2}\sum_{z\in\aleph_3} p(x,y,z)\log\left(\frac{p(x,y,z)p(z)}{p(x,z)p(y,z)}\right) \tag{7}$$

The conditional mutual information measures the dependence of two variables, x and y, given a conditioner variable, z. If either x or y are dependent on z the mutual information between x and y is reduced, and this reduction of information provides a method to 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.

$$\mathcal{T}_{a\to b}(\tau) = \sum_{\hat{a}\in\aleph_1}\sum_{\hat{a}^{(k)}\in\aleph_1^{(k)}}\sum_{\hat{b}\in\aleph_2} p(\hat{a}(t+\tau),\hat{a}^{(k)}(t),\hat{b}(t))\log\left(\frac{p(\hat{a}(t+\tau)|\hat{a}^{(k)}(t),\hat{b}(t))}{p(\hat{a}(t+\tau)|\hat{a}^{(k)}(t))}\right) \tag{8}$$

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

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

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

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

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

directly. The ISEE1 and ISEE3 data are obtained from UCLA (these datasets are also available at NASA NSSDC [http://nssdc.gsfc.nasa.gov/space/]). The IMP8 data come directly from the IMP teams. The solar wind is propagated with minimum variance technique [*Weimer et al.*, 2003] to GSM (X, Y, Z) = (17, 0, 0) R_E to produce 1-min files, 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 VBs and D_{st}, which characterizes the ring current response to the solar wind driver. We therefore consider the nonlinear cross-correlations of the vector

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

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

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

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

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

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

3.2. Transfer entropy

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

4. Summary

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

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

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

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

Using the cumulant-based significance, we have established that the underlying dynamics of D_{st} is in general nonlinear exhibiting a quasiperiodicity which is detectable only if nonlinear correlations are taken into account. The strong nonlinear responses of D_{st} to VBs at $\tau = 25$, 50, and 90 hours are likely related to multiple modes of relaxation of the ring current following the commencement of storms. The nonlinearities at $\tau = 3 - 12$ hours are not caused by internal dynamics but rather by the solar wind driver. This time scale is consistent with the time scale for the information transfer from the solar wind to D_{st} obtained from transfer entropy analysis.

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

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. Jay R. Johnson acknowledges support from NASA Grants (NNH11AR07I, NNX14AM27G, NNH14AY20I, NNX16AC39G), NSF Grants (ATM0902730, AGS-1203299), and DOE contract DE-AC02-09CH11466. We thank James M. Weygand for the solar wind data processing. The raw solar wind data from ACE, Wind, ISEE1 and ISEE3 were obtained from NASA CDAW and NSSDC.

References

Baker, D. N., R. D. Zwickl, S. J. Bame, E. W. Hones, B. T. Tsurutani, E. J. Smith, and S.-I. Akasofu (1983), An isee 3 high time resolution study of interplanetary parameter correlations with magnetospheric activity, *Journal of Geophysical Research*, *88*(A8), 6230, doi:10.1029/ja088ia08p06230.

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

6394.

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.

Clauer, C. R., R. L. McPherron, C. Searls, and M. G. Kivelson (1981), Solar wind control of auroral zone geomagnetic activity, *Geophysical Research Letters*, *8*(8), 915?918, doi: 10.1029/gl008i008p00915.

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

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

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

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

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

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

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

Johnson, J. R., and S. Wing (2015), The dependence of the strength and thickness of field-aligned currents on solar wind and ionospheric parameters, *Journal of Geophysical Research: Space Physics*, *120*(5), 3987?4008, doi:10.1002/2014ja020312.

Kissinger, J., R. L. McPherron, T.-S. Hsu, and V. Angelopoulos (2011), Steady magnetospheric convection and stream interfaces: Relationship over a solar cycle, *Journal of Geophysical Research: Space Physics*, *116*(A5), n/a?n/a, doi:10.1029/2010ja015763.

Li, W. (1990), Mutual information functions versus correlation functions, *J. Stat. Phys.*, *60*, 823.

Newell, P., K. Liou, J. Gjerloev, T. Sotirelis, S. Wing, and E. Mitchell (2016), Substorm probabilities are best predicted from solar wind speed, *Journal of Atmospheric and Solar-Terrestrial Physics*, *146*, 28?37, doi:10.1016/j.jastp.2016.04.019.

Papitashvili, V. O., N. E. Papitashvili, and J. H. King (2000), Solar cycle effects in planetary geomagnetic activity: Analysis of 36-year long omni dataset, *Geophysical Research Letters*, *27*(17), 2797?2800, doi:10.1029/2000gl000064.

Prichard, D., and J. Theiler (1995), Generalized redundancies for time series analysis, *Physica D: Nonlinear Phenomena*, *84*(3?4), 476 – 493, doi:10.1016/0167-2789(95)00041-2.

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, *Journal of Geophysical Research: Space Physics*, *116*(A2), n/a?n/a, doi:10.1029/2010ja015735.

Schreiber, T. (2000), Measuring information transfer, *Phys. Rev. Lett.*, *85*, 461–464, doi: 10.1103/PhysRevLett.85.461.

Strangeway, R., J. R. E. Ergun, Y.-J. Su, C. W. Carlson, and R. C. Elphic (2005), Factors controlling ionospheric outflows as observed at intermediate altitudes, *Journal of Geophysical Research*, *110*(A3), doi:10.1029/2004ja010829.

Weimer, D. R., D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness, C. W. Smith, and J. Watermann (2003), Predicting interplanetary magnetic field (imf) propagation delay times using the minimum variance technique, *Journal of Geophysical Research*, *108*(A1), doi:10.1029/2002ja009405.

Wing, S., and J. R. Johnson (2015), Theory and observations of upward field-aligned currents at the magnetopause boundary layer, *Geophysical Research Letters*, *42*(21), 9149?9155, doi:10.1002/2015gl065464.

Wing, S., J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman, K. Costello, M. Balikhin, and K. Takahashi (2005), Kp forecast models, *Journal of Geophysical Research*, *110*(A4), doi:10.1029/2004ja010500.

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

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

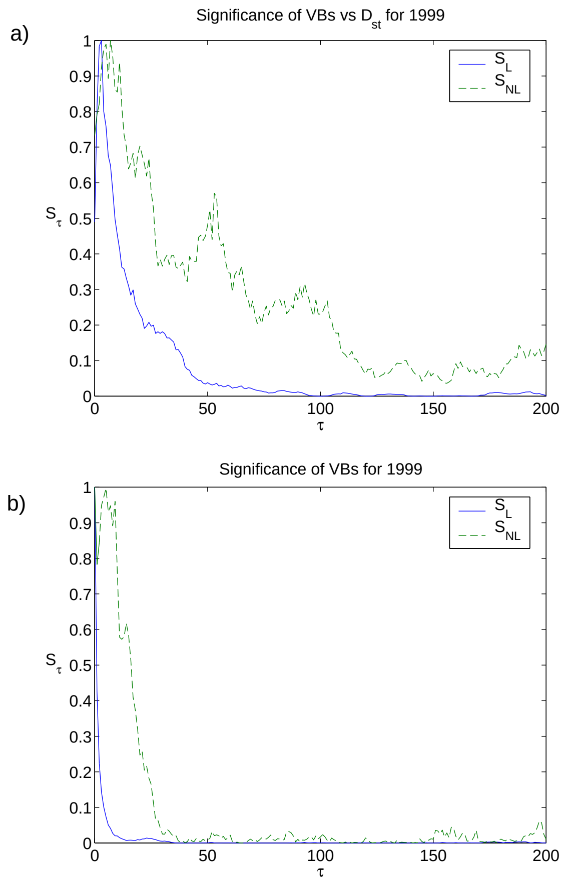

Figure 1. Significance extracted from (a) $\{VBs(t), D_{st}(t-\tau)\}$ and (b) $\{VBs(t), VBs(t-\tau)\}$ for 1999. It should be noted that there is a strong linear response at around 3 hour time delay. There is a clear nonlinear response with a strong peak around 50 hours lasting for approximately 1 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.

Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2018-7
Manuscript under review for journal Ann. Geophys.
Discussion started: 23 January 2018

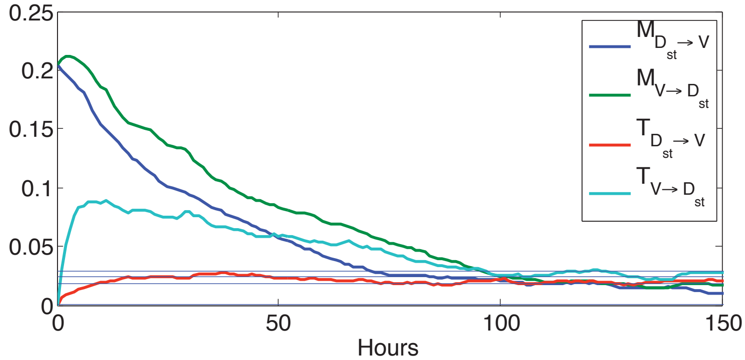

Figure 2. Comparison of mutual information and transfer entropy measures to determine causal driving of the magnetosphere as characterized by D_{st}. Note that causal driving appears to peak somewhat later (11 hours) than indicated by mutual information (2 hours) indicating that internal dynamics likely 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.