Submitted to *Journal Geophysical Research*, Jan 6, 2004.

A Cumulant-Based Analysis of Nonlinear Magnetospheric Dynamics

Jay R. Johnson

Princeton University, Plasma Physics Laboratory, Princeton, NJ

Simon Wing

Johns Hopkins University, Applied Physics Laboratory, Laurel, MD

Abstract. Understanding magnetospheric dynamics and predicting future behavior of the magnetosphere is of great practical interest because it could potentially help to avert catastrophic loss of power and communications. In order to build good predictive models it is necessary to understand the most critical nonlinear dependencies among observed plasma and electromagnetic field variables in the coupled solar wind/magnetosphere system. In this work, we apply a cumulant-based information dynamical measure to characterize the nonlinear dynamics underlying the time evolution of the D_{st} and K_p geomagnetic indices, given solar wind magnetic field and plasma input. We examine the underlying dynamics of the system, the temporal statistical dependencies, the degree of nonlinearity, and the rate of information loss. We find a significant solar cycle dependence in the underlying dynamics of the system with greater nonlinearity for solar minimum. The cumulant-based approach also has the advantage that it is reliable even in the case of small data sets and therefore it is possible to avoid the assumption of stationarity, which allows for a measure of predictability even when the underlying system dynamics may change character. Evaluations of several leading K_p prediction models indicate that their performances are sub-optimal during active times. We discuss possible improvements of these models based on this nonparametric approach.

1. Introduction

The problem of greatest practical importance in the area of space physics is that of understanding magnetospheric response to solar wind input. This response is expected to be highly nonlinear because magnetic energy is stored in the magnetotail and then suddenly released during violent events termed substorms. During these violent releases of energy, energetic MeV electrons, which can damage satellite instrumentation, are injected into the ring current region. Power service and communications on the ground can also be interrupted due to induced currents generated during these massive events. It is therefore extremely important to be able to predict the magnetospheric response to solar wind input in order to be able to make provision for the protection of scientific, communication and defense satellite instrumentation as well as ground based power grids.

The most commonly used measure of magnetospheric activity are the geomagnetic indices obtained by statistically averaging magnetometer readings from ground-based stations located at various latitudes. The magnetic indices include the planetary index, K_p; the storm index, D_{st}; and the substorm indices AU, AL, AE, and AO. It is of great interest to understand and predict the behavior of these geomagnetic

indices. Because currents induced in power grids and killer electrons are commonly associated with intensification of the ring current associated with storms, they are also associated with the sharp dip in the D_{st} index which occurs at the onset of the storm. Accurate predictions of the K_p and D_{st} can be used as input for the Magnetospheric Specification and Forecast Models (MSFM) [*Freeman et al.*, 1995] to predict magnetospheric particle fluxes and electromagnetic fields in the ionosphere. Accurate knowledge of energetic particle fluxes and ionospheric fields could then be used to provide alerts so that precautionary measures could be taken to avoid catastrophic damage to power grids and satellites.

Satellites sitting between the sun and earth, e.g. Geotail, WIND, can be used to monitor the input solar wind parameters (density, velocity, magnetic field strength and orientation, etc.). The recently launched ACE satellite, which is located at the L1 Lagrangian point, has been dedicated to provide solar wind parameters approximately up to one hour before their arrival, allowing for short term forecasts of the magnetospheric activity based on these parameters. Therefore, there is a high demand for models that can predict geomagnetic activity accurately based on solar wind parameters as input, a demand that will likely to increase even more with the nation's increasing reliance on the space technologies. Predictability in this context means: given (a) a time series of solar wind parameters measured by a satellite sitting in the solar wind and (b) a time series of magnetic index measurements, can (A) the value of the geomagnetic index be predicted accurately at a future time and (B) if so, how far ahead can it be predicted?

Although progress has been made in recent years, comprehensive evaluations of the leading K_p and D_{st} models using 25 years of solar wind data still show that much improvement is still needed, particularly for K_p models, in order to obtain accurate forecasts that can be trusted by the users. In Figure 1 we show a comparison of predictions from the Costello Neural Network (*http://www.sec.noaa.gov/rpc/costello/*) and the actual value of K_p. Predictions during disturbed conditions, the periods of most interest to users, tend to be inaccurate in general.

Understanding the behavior of the geomagnetic indices also has important scientific merit because those indices in a convoluted way are embedded with rich information about the underlying dynamics of the coupled solar wind/magnetospheric system. For example, the D_{st} and AE indices are associated with the

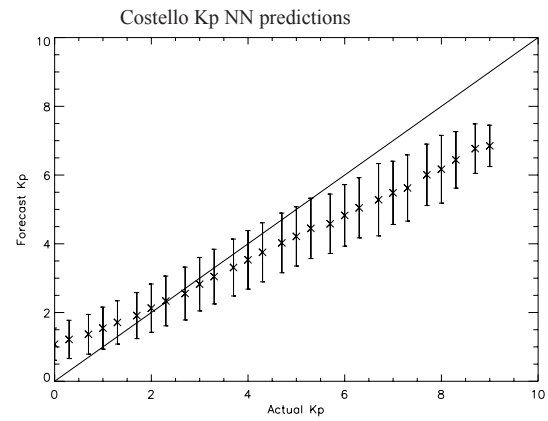

Figure 1. Comparison of actual K_p values and K_p forecasts using the Costello Neural Network over two solar cycles. Actual values of $K_p>5$ lie outside the standard deviation of predictions. This result illustrates that current K_p models can be significantly improved.

ring current and auroral electrojet current systems respectively which are intensified during times of storm or substorm activity, and therefore may provide useful information about the underlying physical processes involved in storms and substorms. Because the magnetosphere is driven by the solar wind, the magnetic indices have strong dependency on solar wind plasma and field variables. Identifying those dependencies is essential for understanding key physical processes that contribute to the nonlinear response of the magnetosphere to external and internal conditions. These dependencies lie at the heart of such questions as: what causes substorms and what physical processes relax the magnetosphere following storms?

Three basic approaches have commonly been applied to modeling the magnetospheric response to the solar wind: (a) physics based, (b) parametrically based, and (c) statistically based.

The *physics based* approach employs a set of physically derived equations describing the system. Examples of such models include global MHD simulations [*Raeder et al.*, 2001; *Berchem et al.*, 1998]) and simplified nonlinear circuit equations with inductors, capacitors, and voltage generators corresponding to physical quantities [*Horton and Doxas*, 1998]). Unfortunately, global MHD simulations fail to describe kinetic physics in boundary layers which determines to a great extent the transfer of mass, energy, and momentum to the magnetosphere while the nonlinear

circuit models oversimplify the physics.

Parametric models assume the response of the magnetosphere based on expected physical processes (such as particle injection due to solar wind/magnetosphere interaction and ring current decay due to charge exchange) which are modeled with parameters that are chosen to minimize the variance between the measurements and model predictions. Empirical forecasting models have been developed—some with a modest number of parameters [e.g. *Burton et al.*, 1975; *O'Brien and McPherron*, 2000] and others with a more extensive list of parameters [e.g. *Li et al.*, 2001; *Temerin and Li*, 2002]. This approach usually yields plausible results during times of low magnetospheric activity. Neural network models also provide predictability by assuming an n-step Markovian process which can be modeled based on past history. The last n steps are mapped with a parametric nonlinear function and then added with parametric weights. The parameters are trained on historical data [*Klimas et al.*, 1997, 1998; *Vassiliadis et al.*, 1995, 1999]. The models can then be used to predict system behavior.

Statistical modeling of time series is generally applied to nonlinear time series to understand the underlying statistical dependences of the data and to guide in developing an algorithm for accurate prediction. Progress has been made in magnetospheric systems by examining the dimensionality of the system and various correlation measures [*Vassiliadis et al.*, 1990, 1991; *Roberts et al.*, 1991; *Sharma*, 1995]. The information-theoretic work developed by *Prichard et al.* [1996] which demonstrated the predictability of the substorm dynamics is an innovative approach that has been widely recognized for its impact on information technology research and has been widely cited in recent work in the area of information-dynamics. More recent work has focused on the substorm as being a "phase transition" and its possible connection with the concept of "self-organized criticality" [*Sitnov et al.*, 2000; *Chang*, 1999; *Klimas et al.*, 2000; *Lui et al.*, 2000; *Chapman and Watkins*, 2001; *Chang et al.*, 2003]. However, to date the statistical approach has not been used as the basis of a predictive model.

While physics based models are well suited to describe and predict long term behavior of a system from physical principles, often the physics is too complex to be described accurately by such models. For example, some view that the ballooning instability is responsible for the onset of substorms which are a global phenomena [*Roux et al.*, 1991; *Cheng and Lui*, 1998]. Has this instability ever been identified in a global MHD simulation? If there were sufficient resolution, would it be possible to describe why measurements of the plasma β [*Lui et al.*, 1992] rise well above the threshold of the ballooning instability using the global MHD framework? Such issues could be important for threshold (onset/timing) of substorms and the dynamical evolution of AE.

Empirical models tend to describe those dynamics prescribed to be important by the author of the model. Often they involve an extensive number of parameters—many of which may not really be that important. If the system does not have high dimensionality—and it is believed that the earth's magnetosphere exhibits low dimensionality [*Roberts et al.*, 1991; *Sharma*, 1995; *Vassiliadis et al.*, 1990]—then it should not require such an extensive number of parameters. Extensive lists of parameters also makes it difficult to extract the most important underlying physics.

Statistical models presume no a priori underlying dynamics and are therefore useful to flush out the critical nonlinear dependencies in the system. Moreover, the information gained from such an analysis can be invaluable for developing and constraining parametric models.

2. An Information Dynamical Approach

Most data gathered by satellites and ground based instruments are in the form of a time series. Analyzing these signals usually encompasses three fundamental tasks: characterization, forecasting, and modeling. **Characterization** involves determining what kind of system produced the signal. **Forecasting** involves predicting what the system will do next given its current state or past history. **Modeling** involves determining a set of governing equations which describe the evolution of the system [*Gershenfeld*, 1998].

For complex systems, modeling can be physically or computationally difficult. For some systems such as the brain, physical equations describing neural interactions are not well specified. For other systems, such as the magnetosphere, the underlying physical equations may be known at the most fundamental level (particle simulation), but global computations are beyond present and/or future computational capabilities without appropriate approximations. Empirical models that employ intuition assume a priori a dynamical framework that may or may not apply to the system. However, there is some danger in that it may be possible to fit the data by choosing enough

free parameters at the expense of loosing physical understanding. Because the magnetospheric system is complex and the nonlinear response of the magnetosphere during storms and substorms is not clearly understood, it seems appropriate to apply statistical techniques that are unbiased a priori.

The key to characterizing a system is to understand the dependencies in a system. For example, given data from the solar wind and a ground based magnetometer, is it possible to determine the degree to which the magnetometer data depends on the solar wind data? The typical method of choice for discovering such dependencies is the correlation function. However, in a highly nonlinear system, correlation functions are not very useful because nonlinear systems tend to have broadband power spectra and hence featureless correlation structure.

It is well known that the magnetosphere responds in a highly nonlinear way to solar wind input. Substorms involve loading and sudden release of energy which cannot be well described as a linear system. Therefore, it is necessary to go beyond typical correlation studies to understand the nonlinear dependencies between the solar wind driver and the magnetospheric response.

Information-theoretic quantities provide an elegant alternative that captures the essential features of the correlation function and more [*Gershenfeld*, 1998]. One commonly used information-theoretic quantity is mutual information which provides a statistical measure of dependency based on probability theory [*Prichard et al.*, 1996; *Gershenfeld*, 1998]. Although useful, it has basic limitations because of the need to compute a probability density which in many cases requires a large database to achieve good statistics. The cumulant-based significance and information flow are alternative information-theoretic quantities that can be used to detect dependencies in a system [*Deco and Schürmann*, 2000]. These quantities provide a measure of the cumulants that would normally vanish in the absence of dependencies. The significance and information flow are computed directly from the dataset in comparison with surrogate data sets. These measures have an advantage of providing good statistics for small data sets and reliable detection of dependencies when data is corrupted by noise. In §3 we define the cumulant-based information measures and provide examples of their utility. In §4 we apply these measures to geomagnetic indices and solar wind data to detect the presence of nonlinearities in the solar wind/magnetosphere system.

3. Cumulant-based measure of significance and information flow

3.1. System Dynamics

The Cumulant-based significance is a useful quantify for detecting nonlinearities in the underlying dynamics of a system. For purposes of explaining our approach, let us assume that the underlying dynamics of a system are described by the evolution of a state variable $\mathbf{x}$

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) + \nu \tag{1}$$

where ν is additive noise. The system dynamics are linear if

$$\mathbf{F}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \tag{2}$$

and nonlinear if F is not a linear function of $\mathbf{x}$.

In a real system, often only a subset of the state variables, $\mathbf{y}$, may be observed. Given a limited data set, it is often useful to consider an embedding vector of the system

$$\begin{aligned}\mathbf{c}(t) &= \{\mathbf{y}(t), \mathbf{y}(t-1), ..., \mathbf{y}(t-(m-1)\} \\ &= \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m\};\end{aligned} \tag{3}$$

The rationale for examining an embedding vector is that the dynamics of the original system can be captured in the dynamical evolution of the embedding vector [*Takens*, 1980; *Sauer et al.*, 1991]. The application of this concept to the magnetosphere should be obvious. State vectors are generally not known for the system as plasma and field measurements are not available for the entire system as a function of time (usually only a few single point measurements are available). These few variables are often combined into a single variable that is nonlinearly related to appropriate state variables. However, the time evolution of that single variable amazingly may contain much information about the dynamical evolution of the entire system.

3.2. Cost and Significance

To detect nonlinearities and understand the underlying dynamics of a system, we will examine realizations of an embedding vector extracted from the original data set (for example, the time history of D_{st}). We are interested in understanding the predictability of the system, so it is useful to understand the probability of finding system in a particular state given past history of the system. For purposes of illustration, we construct an embedding vec-

tor $(y_1, ..., y_m) \equiv (y_t, y_{t-\tau}, ..., y_{t-(m-1)\tau})$ from a single variable, y, in the system. A measure of the relevance to the past history on a current/future value of the system is captured addressed by the following equation

$$P(y_1, ..., y_m) \stackrel{?}{=} P(y_1)P(y_2, ..., y_m) \tag{4}$$

where P are probabilities. Equation 4 asks whether the probability of extracting the embedding vector $\mathbf{y} = (y_1, ..., y_m)$ depends on the past history $(y_2, ..., y_m)$. If Equation 4 were true, then there will be certain statistical relations between the higher-order correlation tensors

$$C_{i...j} = \int d\mathbf{y} P(\mathbf{y}) y_i ... y_j \equiv \langle y_i ... y_j \rangle \tag{5}$$

where $i, ..., j \in 1, ..., m$. In particular, the cumulants, $K_{1i_2...i_n}$, of the distribution defined by

$$\begin{aligned} K_i &= C_i = \langle y_i \rangle \\ K_{ij} &= C_{ij} - C_i C_j = \langle y_i y_j \rangle - \langle y_i \rangle \langle y_j \rangle \\ K_{ijk} &= C_{ijk} - C_{ij}C_k - C_{jk}C_i - C_{ik}C_j + 2C_iC_jC_k \\ K_{ijkl} &= C_{ijkl} - C_{ijk}C_l - C_{ijl}C_k - C_{ilk}C_j - C_{ljk}C_i \\ &\quad - C_{ij}C_{kl} - C_{il}C_{kj} - C_{ik}C_{jl} + 2(C_{ij}C_kC_l \\ &\quad + C_{ik}C_jC_l + C_{il}C_jC_k + C_{jk}C_iC_l + C_{jl}C_iC_k \\ &\quad + C_{kl}C_iC_j) - 6C_iC_jC_kC_l \end{aligned} \tag{6}$$

should vanish unless $i_2 = ... = i_n = 1$ where n is the order of the cumulant. Therefore a useful measure the statistical independence of the components of y_1 on $(y_2, ..., y_n)$, is the cost function defined as:

$$D = \sum_{n=1}^{\infty} \sum_{i_2,...,i_n=1}^{m} (1 - \delta_{1i_2...i_n}) \{K_{1i_2...i_n}\}^2 \tag{7}$$

where $\delta_{ij...n}$ is the Kronecker delta which eliminates the diagonal elements. In the absence of correlations, the cost function should vanish. To examine the significance of correlations, we employ the method of surrogates. The method consists of assuming a null hypothesis, constructing surrogate data consistent with that hypothesis, and then comparing the cost function for the original data set with the cost functions of the surrogate data sets. The significance is defined as:

$$S = \frac{|D_0 - \mu_S|}{\sigma_S} \tag{8}$$

where D_0 is the cost of the original data set μ_S and σ_S are the mean and variance of the costs computed with the surrogate data sets,

$$\mu_s = \frac{1}{N} \sum_{i=1}^{N} D_{S_i} \tag{9}$$

$$\sigma_S = \frac{1}{N-1} \sum_{i=1}^{N} (D_{S_i} - \mu_S)^2. \tag{10}$$

where N is the number of surrogate data sets.

The surrogate data is chosen consistent with a null hypothesis which for our purposes will be that there is no causal relationship between the past and the present. The surrogate data may be obtained by taking random permutations of the original data set or by extracting data randomly from the same density distribution as the original dataset. If the significance is larger than $\hat{S} = 1.67$, there is a 95% chance that the null hypothesis is falsified and that there is a clear underlying dynamics governing the system (note that $\mathrm{erfc}(\hat{S}/2) = 0.05$) assuming Gaussianity in the surrogates distribution which is typically true if the number of surrogates is sufficiently large) [*Deco and Schürmann*, 2000]. If the significance remains roughly constant, the dynamics are stationary, but if the significance changes in magnitude during time, it suggests that the dynamics involved are non-stationary. The technique was tested on a number of datasets generated by mathematical equations for which the underlying dynamics is known prior to application to the magnetospheric datasets.

The significance can be useful for a number of tasks [*Deco and Schürmann*, 2000]:

- By comparing the significance when keeping only second-order cumulants with the significance including higher-order cumulants, we may determine whether the origin of the observed statistical correlations are linear or nonlinear. The second order cumulant is equivalent to the correlation function.

- As in a spectrogram, we can also consider windowed significance where data is sampled from a window of width N_w and a significance is computed for that data set. The window is then shifted and the significance recomputed. Time variations in the significance indicate changes in the underlying dynamics of the system.

- We can introduce a proxy for the information flow based on cumulant-based significance integrated over "look ahead." Information flow deals with changes in the information content of a system. The information flow can be used to detect changes in underlying dynamics and the loss of information in a system. The loss of

information in a system can be quantified and provides a measure of the predictability horizon of the system. Such calculations are practical because if there is a characteristic time for information loss in a system, it would be foolish to attempt to develop parametric models to predict system behavior on longer time scales.

3.3. Significance as a Measure of Nonlinearity

By comparing the significance when keeping only second-order cumulants with the significance including higher-order cumulants, we may also determine whether the origin of the observed statistical correlations are linear or nonlinear. We Gaussianize the data set in order to eliminate the effect of "static" nonlinearities in the original data set [*Kennel and Isabelle*, 1992] as described in § 3.5. Statistical studies are then performed on the Gaussianized data which exhibits no static nonlinearity (all cumulants vanish beyond second order).

We present two example which illustrate the utility of the method. The first example is the well-known Lorenz system [*Lorenz*, 1963] which satisfies the following equations:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x) \\ \frac{dy}{dt} &= -xz + rx - y \\ \frac{dz}{dt} &= xy - bz \end{aligned} \tag{11}$$

The Lorenz system exhibits a chaotic attractor for $r = 45.92$, $b = 4$, and $\sigma = 16$. We show a trajectory for the system in Figure 2. We consider an embedding dimension, $m = 2$, which appropriately captures the dynamics of the attractor. In Figure 3 we show the normalized significance as a function of time delay for the Lorenz system. The linear significance, S_L, is obtained by retaining only second order cumulants in Equation 7, while the nonlinear significance, S_{NL} includes cumulants to fourth order [*Deco and Schürmann*, 2000]. The linear significance is proportional to the correlation function. There is an obvious difference between the linear and nonlinear significance which indicates the importance of nonlinearities in the system dynamics. While the linear significance drops off rapidly, the nonlinear significance has a rich quasi-periodicity not seen in the correlation function. The nonlinear significance captures the nonlinearity of the system shown in Figure 2 where the system oscillates chaotically around two attractors. However, unlike a sinusoidal function, there is

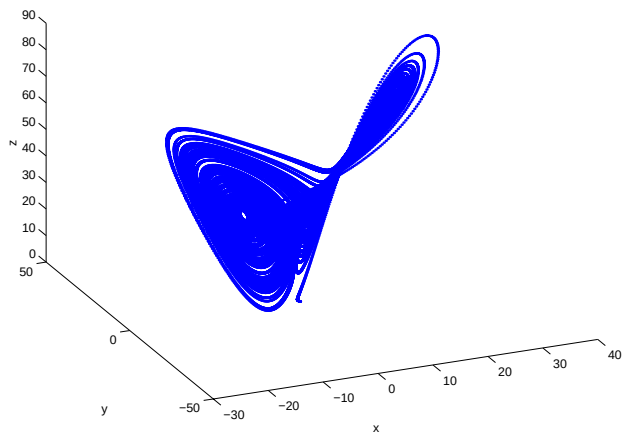

Figure 2. Phase space plot of the Lorenz system defined in Equation 11 for $r = 45.92$, $b = 4$, and $\sigma = 16$. This system is highly nonlinear and oscillates in a quasi-periodic manner about two fixed points.).

only quasi-periodicity. Although there is a long-term decreasing trend in the significance, a high level of nonlinearity remains even for long time separation as expected for a chaotic system for which information is not lost.

When performing nonlinear analysis of a system with an embedding vector, it is often useful to consider the most appropriate time delay. The most information about the system is obtained when the correlation is minimum. Usually, the first minimum of the correlation function is chosen. However, a more suitable choice would be the first minimum of the significance where an embedding of $m = 2$ is taken. In the case of the Lorenz attractor, the first minimum appears to be at approximately $\tau = 15$.

For comparison, in a linear system, the linear and nonlinear significance typically have the same general structure and may not differ due to the decreasing importance of the higher-order cumulants. As an example, consider a noisy oscillator defined by the following system of difference equations

$$\begin{aligned} x_{j+1} &= x_j + \alpha y_j + \nu_1 \\ y_{j+1} &= y_j - \alpha x_j + \nu_2 \end{aligned} \tag{12}$$

where ν_1 and ν_2 are a random, Gaussian distributed variable. In Figure 4 we take $\alpha = 0.01$ and the variance of ν is 0.01, which produces a noisy oscillator. The significance is shown in Figure 5 as a function of

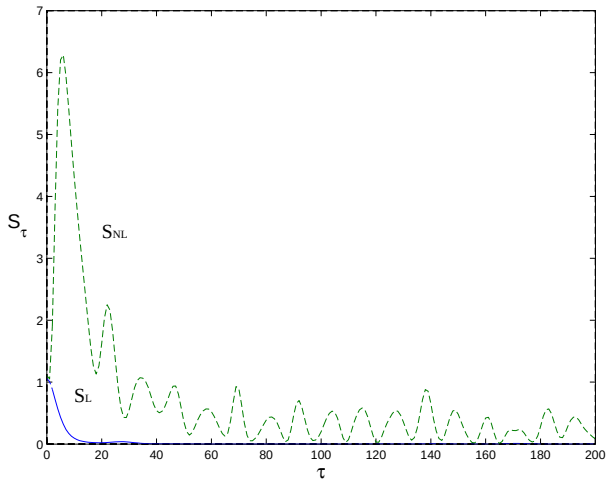

Figure 3. Normalized significance as a function of time delay, τ for the Lorenz system [Lorenz, 1963]. The time delay is measured in units of the sampling time (0.02). An embedding of m=2 is used in computing the significance. S_L: only linear correlations are considered; S_{NL}: Higher-order correlations are considered (up to fourth order). The existence of peaks with different amplitudes captures the quasiperiodic nature of the Lorenz attractor. The normalization factor is 10 for S_{NL}.

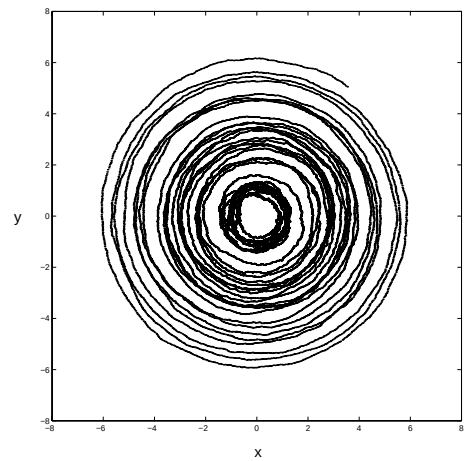

Figure 4. Noisy oscillator solution to Equation 4 for $\alpha = 0.01$.

time delay, τ, with an embedding vector of dimension $m = 2$. The linear and nonlinear significance do not show any significant differences.

3.4. Significance as an Indicator of Changes in Underlying Dynamics

As in a spectrogram, we can also consider windowed significance in which case data is sampled from a window of width N_w and a significance is computed for that data set. The window is then shifted and the significance recomputed. Time variations in the significance indicate changes in the underlying dynamics of the system. Because the significance is a function of time delay, it is useful to consider a time delay which is most suitable for the data. While it could be varied, a useful choice is the first minimum of the significance of the entire dataset which is considered to be the best time delay for analyzing the nonlinear dynamics of the system.

For example, in Figure 6 we examine a system governed by Lorenz dynamics for $t < 60$ with an abrupt change to sinusoidal dynamics for $t > 60$. Note that the significance remains roughly constant while

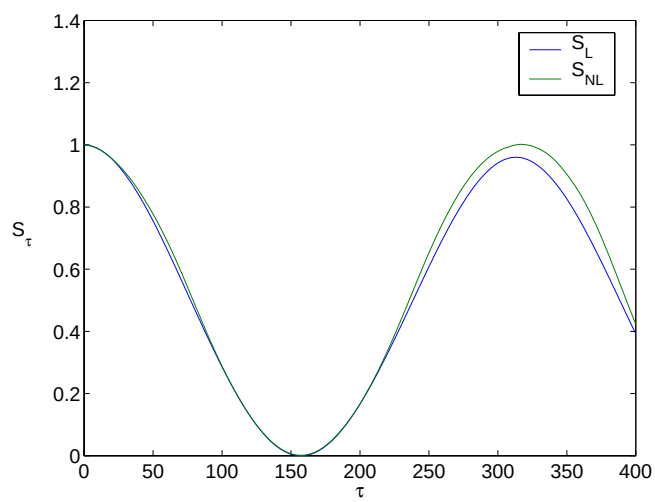

Figure 5. There is no significant difference in the linear and nonlinear significance for the linear system defined in Equation 12.

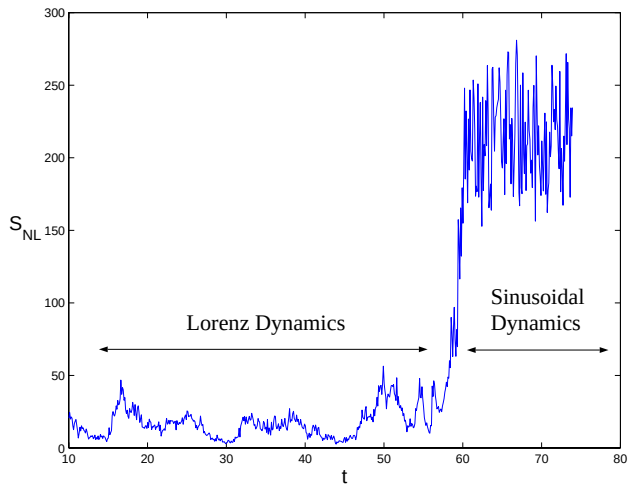

Figure 6. Windowed significance as a function of time. For $t < 60$ the system is governed by Lorenz dynamics [Lorenz, 1963]. Near $t = 60$ there is a transition to siniusoidal dyanmics. Note that an abrupt change in the significance clearly indicates the change in the underlying dynamics of the system.

the system dynamics are stationary, but the significance changes abruptly when the underlying dynamics change. Hence, the significance can be a good indicator of changes in the underlying dynamics of a nonlinear system. The technique of windowed significance will be applied to magnetospheric data to detect changes in the underlying dynamics of the magnetosphere. Any predictive model would need be able to account for such changes in the underlying model.

3.5. Gaussianization of Data

Measurement of a time series can involve static nonlinearities and dynamic nonlinearities. Static nonlinearities are those nonlinearities introduced simply by measuring a system variable and involve a nonlinear transformation of the observing process. Because we are only interested in dynamical nonlinearity inherent to the dynamics of the system not those static nonlinearities related to the measurement process, we have employed Gaussianization to the data to eliminate the contribution of the static nonlinearities to the calculation of cost and significance.

The Gaussianization procedure is described by [*Kennel and Isabelle*, 1992]. We generate a sequence of random Gaussian distributed numbers with unit variance the same length as the original data set. The data and random variable are then sorted numerically. This procedure defines a one-to-one mapping of the original data set to the random variable. The statistical procedures are then applied to the new time series defined by the reordered series of the random variable. The cumulants of the random variable beyond second order vanish because it is Gaussian distributed therefore eliminating any static nonlinearity in the original data. Any nonlinearities that are detected are the result of the dynamics (time-ordering). Surrogate data sets can also be drawn directly from the Gaussian distribution rather than by reshuffling the data.

3.6. Cumulants and Mutual Information

Mutual information is an elegant information-theoretic quantity that is commonly employed to detect nonlinear dependency in a complex system. Suppose measurements of two quantities are obtained (e.g. Solar Wind data and D_{st} measurements). The datasets will span a range of data which can be binned/quantized—the number of bins may be different if the variables require different resolution. After quantization, we have two variables, x and y, that will take on discrete values

$$x \in \{1, ..., N\} \equiv \aleph_1; y \in \{1, ..., M\} \equiv \aleph_2 \qquad (13)$$

The variables may be thought of as letters in alphabets $\aleph_1$ and $\aleph_2$ which have N and M letters respectively. If considering an embedding vector $(x_t, x_{t-\tau})$, we would use the same alphabet for both variables. The extracted data are then sequences of letters. The entropy associated with each of the variables is defined as

$$H(x) = -\sum_{\aleph_1} p(x) \log p(x); H(y) = -\sum_{\aleph_2} p(y) \log p(y); \qquad (14)$$

where $p(x)$ is the probability of finding letter x in the set of x-data and $p(y)$ is the probability of finding letter y in the set of y-data. To examine the relationship between the two variables, we extract a sequence of words (x, y) from the dataset. The joint entropy is defined by

$$H(x, y) = -\sum_{\aleph_1, \aleph_2} p(x, y) \log p(x, y); \qquad (15)$$

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

$$I(x, y) = H(x) + H(y) - H(x, y) \qquad (16)$$

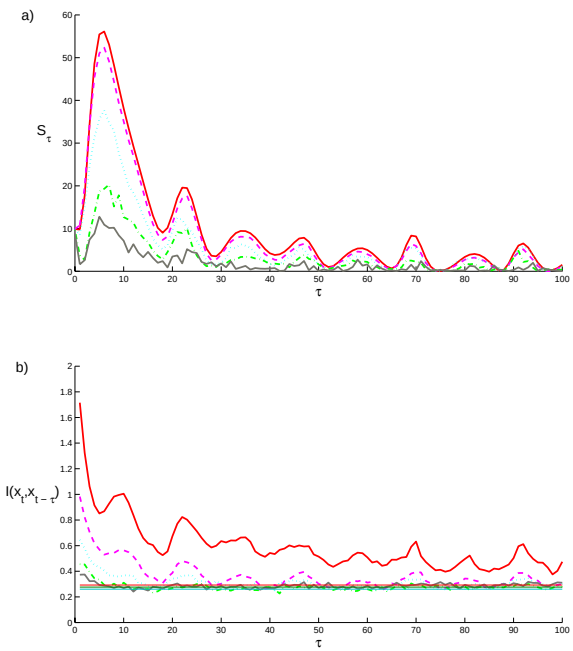

Figure 7. Comparison of (a) the cumulant-based significance and (b) the mutual information for the Lorenz system [Lorenz, 1963]. Shown are the cumulant based significance, S_τ, and mutual information I_τ verses time lag for one of the Lorenz variables. The solid curve is without additive noise. The dashed, dotted, dot-dashed, and grey curves show the results with additive noise with signal to noise ratio of 0.25, 0.50 , 0.75, and 1.

One prescription to compute the mutual information would be to first construct a database of $x \equiv (VBs(t-\tau))$ and $y \equiv Dst(t)$ and words $(x, y) \equiv$ (VBs(t - τ), Dst(t))). Secondly, discretize (bin) the data. Finally, obtain the probabilities by sorting the data pairs and counting. Mutual information is then obtained with simple summation over all possible bin combinations.

The mutual information may be used to assess appropriate sample resolution, system reversibility, and predictability horizon. Statistical limitations often hinder analysis based on the mutual information. For example, it is useful to examine the limit as the bin sizes become smaller, but the bins must be large enough to contain a statistically representative number of data points. If there is limited data, the mutual information cannot be accurately obtained. In contrast, the cumulant based significance may be applied to limited datasets.

It is useful to ask the question whether mutual information or cumulant based significance is more useful for detecting nonlinear dependencies. We applied the two measures to the well-known Lorenz system [*Lorenz*, 1963]. The solid curves in Figure 7a and Figure 7b show the cumulant based significance, S_τ, and mutual information I_τ verses time lag for one of the Lorenz variables. Note that the peaks and relative amplitude of the peaks are very similar for both measures indicating that they both detect the presence and nature of the nonlinearity of the Lorenz system.

3.7. Detecting Dynamics in a Noisy System

Measurements of physical quantities in the real world usually contain some noise or uncertainties. It is interesting to examine how the cumulant-based significance and mutual information perform in the presence of noise. Therefore, we also performed a similar analysis for the Lorenz system when the signal is artificially contaminated by random noise added to the signal. The dashed, dotted, dot-dashed, and grey curves show the results with additive noise with signal to noise ratio of 0.25, 0.50, 0.75, and 1. Surrogate datasets were used in evaluating the significance so that values of $S_\tau > 1.6$ are statistically significant. For the mutual information plot, we have also included three horizontal plots which show the central mean and standard deviation of the mutual information obtained from the surrogate datasets assuming no time ordering. The data contaminated by signal to noise ratio 0.75 and 1 cannot be distinguished from the surrogate data with any confidence with the mutual information measure. In contrast, even with signal to noise ratio of 1, the cumulant based method detects the nonlinearities. Hence, it would appear that the cumulant-based significance can be used to detect the presence of nonlinear dependencies in the underlying dynamics even when the signal is heavily contaminated by noise.

3.8. Cumulants as a Measure of Information Flow

Entropy is commonly used to quantify the information content of a system. Information flow deals with changes in the information content of a system. Obviously, the information flow can be used to detect changes in underlying dynamics and the loss of information in a system. Understanding the loss of information in a system can yield important information about predictability of the system. For example, if there is a characteristic time for information loss

in a system, it would be foolish to attempt to develop parametric models to predict system behavior on longer time scales.

The flow of information in a system is best defined in terms of a conditional entropy, $\mathcal{H}$, which measures the uncertainty of a variable $x(t+p)$ given all possible preceding sequences of that variable.

$$\mathcal{I}_p = \lim_{n\to\infty}[\mathcal{H}(x(t+p)|x(t),...,x(t-n+1) - \mathcal{H}(x(t+p-1)|x(t),...,x(t-n+1)] \quad (17)$$

In the case of a chaotic system, there is no loss of memory and information flow is constant. On the other hand, for a noisy system information flow decreases with increasing look ahead. Development of predictive models is not practicable beyond the decay length of the information flow.

Unfortunately, it is usually difficult to compute a statistically meaningful information flow for a real system because of the limited size of the data set. However, cumulants which also carry information about the underlying system dynamics are readily obtained even from limited data sets. We can therefore introduce a proxy for for the information flow which we define as the cumulant-based information flow

$$I_C(p) = \sum_{n=1}^{\infty}\sum_{i_2,...,i_n=1}^{m}(1-\delta_{1i_2...i_n})\{K^{(p)}_{1i_2...i_n}\}^2 \quad (18)$$

where $K^(p)_{1i_2...i_n}$ are the cumulants associated with the vector $\{y1,...,y_m\} = \{y(t+p), y(t-\Delta),...,y(t-(m-1)\Delta)\}$. This quantity provides an estimate of how well a predictive model could estimate a future value of the time series p steps ahead given the past history of the time series. The minimal value of $I_c(p) = 0$ indicates statistical independence while increasing values of $I_C(p)$ point to increasing dependencies in the time series [*Deco and Schürmann*, 2000]. Ideally, m should be chosen large enough so that $I_C(p)$ becomes a measure of predictability given the entire past of the time series. For practical purposes, we choose the value of m such that the information flow does not change appreciably when m is increased. As for the case of significance, we limit the computation to the fourth-order cumulant.

While the information flow $I_C(p)$ provides an indication of how far into the future one should be able to predict the time series, when examining changes in the information flow, it is more practical to consider

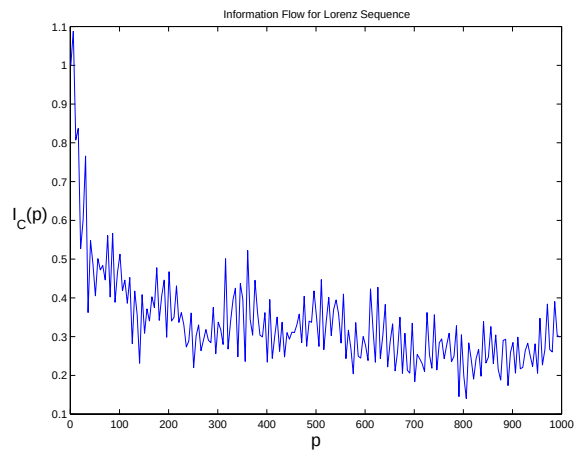

Figure 8. Information flow, $I_C(p)$ as a function of "look ahead", p. For this sequence, we have taken an embedding $m = 6$ and time spacing $\Delta = 15$ which corresponds to the first minima in Figure 3.

an integrated information flow,

$$\widehat{I_C} = \sum_{p=1}^{P}\sum_{n=1}^{\infty}\sum_{i_2,...,i_n=1}^{m}(1-\delta_{1i_2...i_n})\{K^{(p)}_{1i_2...i_n}\}^2 \quad (19)$$

The integrated information flow provides a measure of predictability up to a look ahead, P, which should be taken as large as the desired predictions keeping in mind that it is only of practical value if it is smaller than value of p which characterizes the falloff of the information flow $I_C(p)$

As an example, we consider the Lorenz attractor system defined in Equation 11. In Figure 8 we plot $I_C(p)$ for that system. We have taken an embedding vector, $m = 6$, with time separation $\Delta = 15$. Note that although the information flow does drop initially, it remains reasonably stationary over a long time. The constancy of the information flow is a property of a chaotic system and is more or less captured by the cumulant-based information flow.

4. Application to Magnetic Indices

4.1. Nonlinear response of K_p as a function of solar cycle

To illustrate how the cumulant-based method can be applied to the dynamics of geomagnetic indices, we consider the hourly averaged time series of K_p indices which are thought to gauge the level of magnetospheric activity. We have applied the statistical test

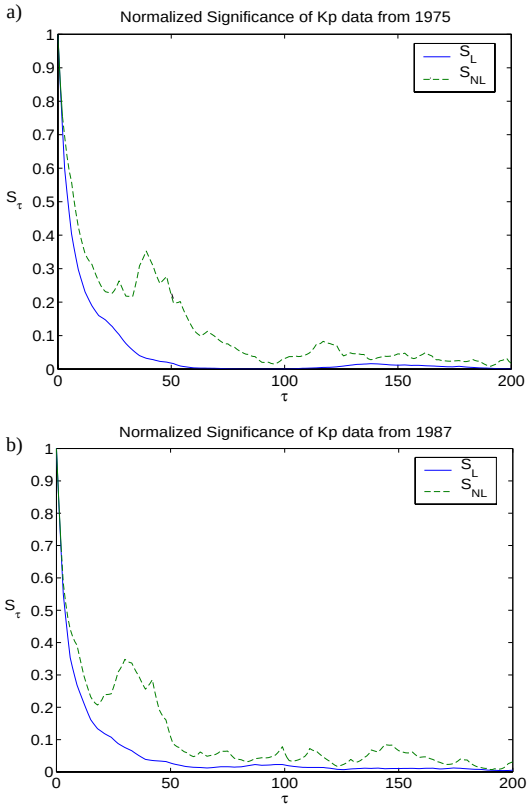

Figure 9. Linear (S_L) and nonlinear (S_{NL}) significance (up to fourth order cumulants) as a function of time-delay, τ, in hours for K_p data from two solar minima. We have used an embedding vector $m = 2$. Note that while the linear significance (linear correlation function) drops off with time, there is a strong nonlinear response with a time scale of around 40 hours.

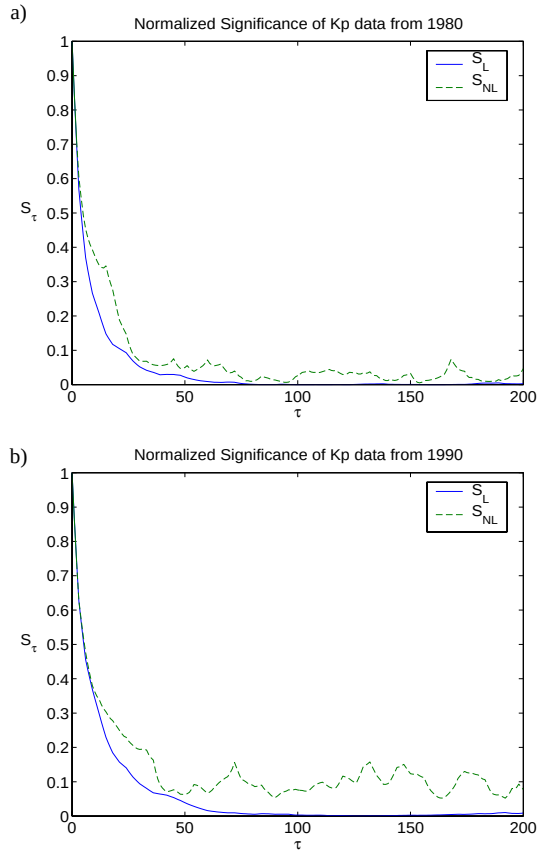

Figure 10. Linear (S_L) and nonlinear (S_{NL}) significance (up to fourth order cumulants) as a function of time-delay, τ, in hours for K_p data from two solar maxima. We have used an embedding vector $m = 2$. Note that both the linear and nonlinear significance drop off with time. For solar maxima, it there is not a strong nonlinear at 40 hours as in the case of solar minima.

to all years from 1974 to the 1999. In Figure 9 we show the normalized significance for two solar minima (1975 and 1987), and in Figure 10 we show the normalized significance for two solar maxima (1980 and 1990). Llinear significance is computed keeping cumulants to second order while nonlinear significance is computed using cumulants to fourth order. Interestingly, the nonlinear response is far more pronounced for solar minimum than solar maximum. This seems to apply also for years close to solar minimum (e.g. 1974-1976 and 1985-1987) exhibit strong nonlinear response. The nonlinear response is most likely related to internal magnetospheric dynamics. It would seem reasonable that nonlinearities would be more pronounced at solar minimum when the system is more dominated by internal dynamics than at solar maximum when the system may be driven by external dynamics (the response to the solar wind). A more recent solar minimum (1996-1999) does not seem to exhibit as strong a nonlinear response. However, this solar minimum exhibits significantly smaller solar wind velocity than usual which may be responsible for the reduced nonlinearity in the response.

The existence of significant nonlinearity in K_p dynamics would complicate predictability of the system. In this section, we examine the predictions of the Costello neural network, which is commonly used to predict K_p based on solar wind V, IMF $|B|$, IMF B_z. We examine the performance of the neural network for two solar cycles. Performance is quantified using

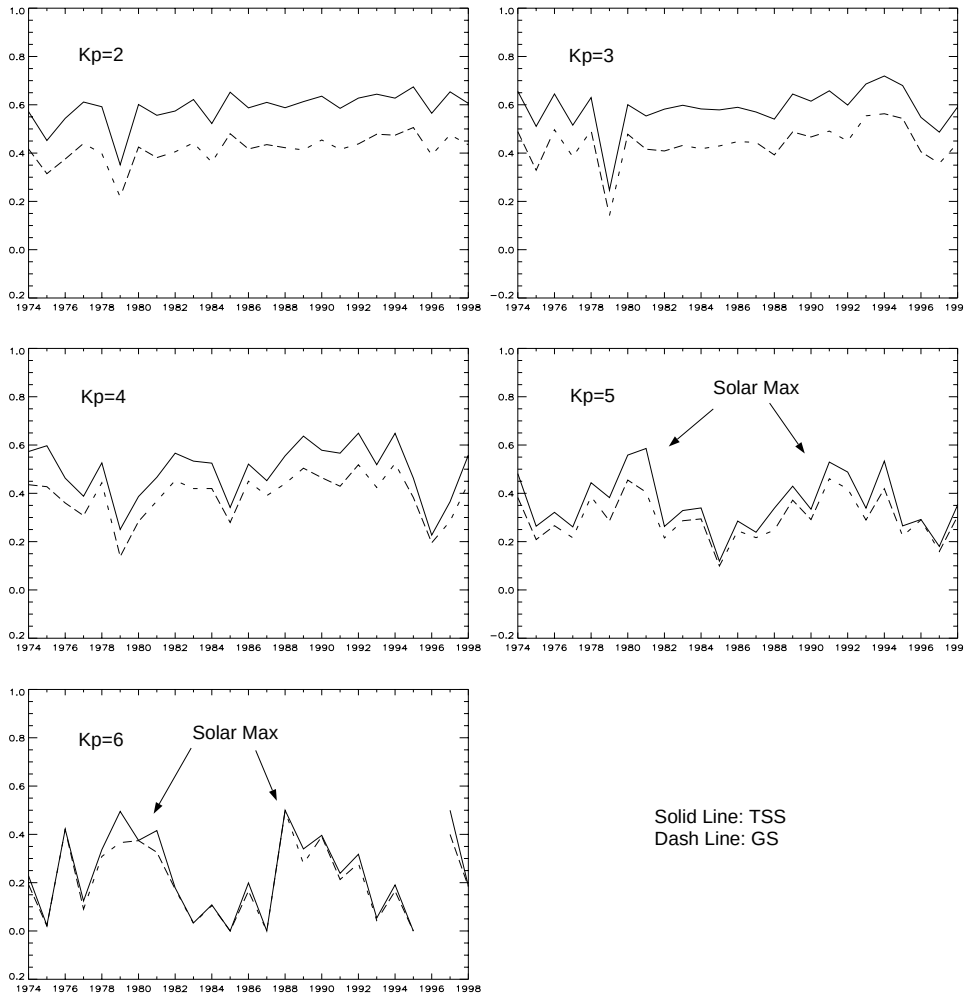

Figure 11. Performance of the Costello NN over two solar cycles from 1974 to 1998. The True Skill Statistic (TSS) and Gilbert Skill (GS) scores indicate that the Costello network could use improvement—especially for active conditions. Moreover, the variation of the scores over the solar cycle indicates that the neural network performs much better at solar maximum than at solar minimum. This result is not surprising given our finding that the dynamics involve more nonlinearity at solar minimum.

skill scores defined in [*Detman and Joselyn*, 1999].

In Figure 11 we show predictive model performance as a function of solar cycle for various values of K_p. We plot True Skill Statistics (TSS) and Gilbert Skill (GS) scores for the neural network model. While there is no significant difference in predictive ability for $K_p < 3$, there is a significant solar cycle dependence for $K_p > 3$. In either case, there is significant room for improvement. It is also clear that there is better predictability at solar maximum than at solar minimum. Although this result may be related to a training bias, it may also reflect the increased nonlinearity of the system dynamics for solar minimum.

4.2. Nonlinear response of D_{st} index

We also applied the same procedure to the D_{st} index which not only has better time resolution, but also can be interpreted physically in terms of the response of the ring current. In Figure 12 we plot the significance for solar minimum and solar maximum. It is also apparent that solar minimum exhibits a stronger nonlinear response, although in this case, solar maximum also exhibits some nonlinearity.

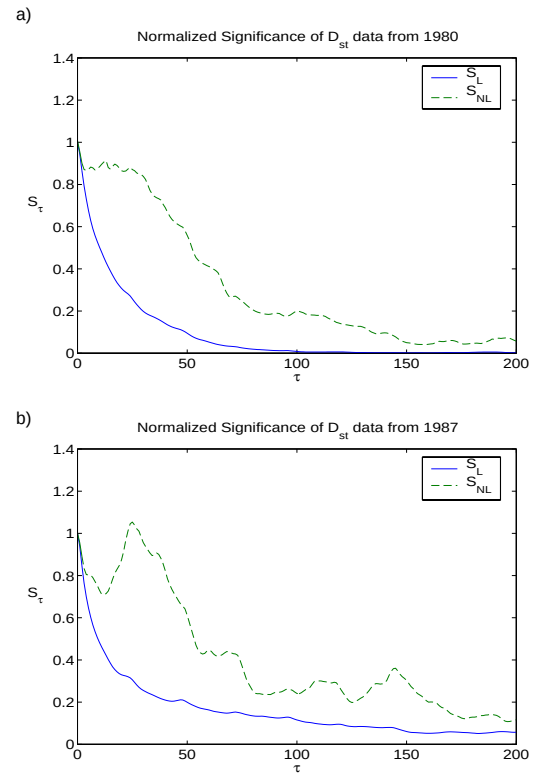

Figure 12. Linear and nonlinear significance as a function of time-delay, τ, in hours for D_{st} data from (a) solar maximum in 1980 and (b) solar minimum in 1987. Note that unlike K_p, D_{st} shows a significant nonlinear response at solar maximum; however, the nonlinear response is far more pronounced at solar minimum as for K_p. The characteristic time of the nonlinear response appears to be around 30 hours and is generally more rapid than K_p by around 10 hours. In this computation, we have used an embedding vector $m = 2$.

4.3. Windowed significance as an indicator of dynamical changes

In Figure 13, we plot the significance as a function of time for the nonlinear time series, K_p, over the course of 1974. The significance is computed in a window of width 300 hours every 5 hours over the course of the year. The significance is obtained in comparison with 100 surrogate datasets including terms up to fourth order cumulants.

The first item that can be noticed is that the significance is often quite large for extended periods of time (10-20 days). As such, the K_p time series should be reasonably predictable during those time periods. On the other hand, the evolution of the significance supports the existence of statistical correlations between past and future, which are probably of a nonstationary nature due to strong changes in the magnitude of the significance over time. As such, the underlying dynamics are probably changing during time, perhaps due to the sudden arrival of interplanetary shocks or CMEs; therefore, any predictive model would need to be able to accommodate such changes in the underlying dynamics.

4.4. Cross-significance of solar wind and magnetosphere data

So far, we have only considered the statistics of an embedding vector taken from a single variable. In this section, we generalize the approach to consider the nonlinear correlations between multiple variables. In this case, we are interested in understanding the nonlinear coupling between the solar wind and magnetospheric dynamical systems. We therefore consider coupling between a solar wind variable such as VBs or the dynamic pressure, P_{dyn} and D_{st} which characterizes the ring current response. We therefore consider

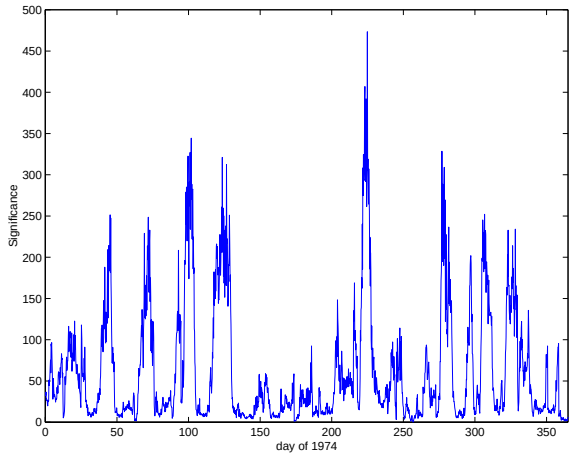

Figure 13. Significance as a function of time for K_p indices during 1974. An embedding dimension of m=3 is used with a time lag of 1 hour. All quantities were computed for a slicing window of 300 hours every 5 hours.

the nonlinear cross-correlations of the vector

$$\mathbf{c}(t) = \{VBs(t), D_{st}(t-\tau)\} = \{y_1, y_2\} \qquad (20)$$

The generalization of cost are based on realizations of $\{y_1, y_2\}$. In this case, each variable is Gaussianized with unit variance to eliminate static nonlinearities (higher order self-correlations are eliminated so that the cost measures only cross-correlations between VBs and D_{st}).

In Figure 14 we plot the significance obtained from the year 1999 as a function of time delay, τ. 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 peaks around 25, 50 and 90 hours lasting for approximately 1 week. The absence of these nonlinearities in the self-significance for VBs indicates that the nonlinearities are related to internal magnetospheric dynamics. As the D_{st} index is thought the 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 response at these peaks is 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.

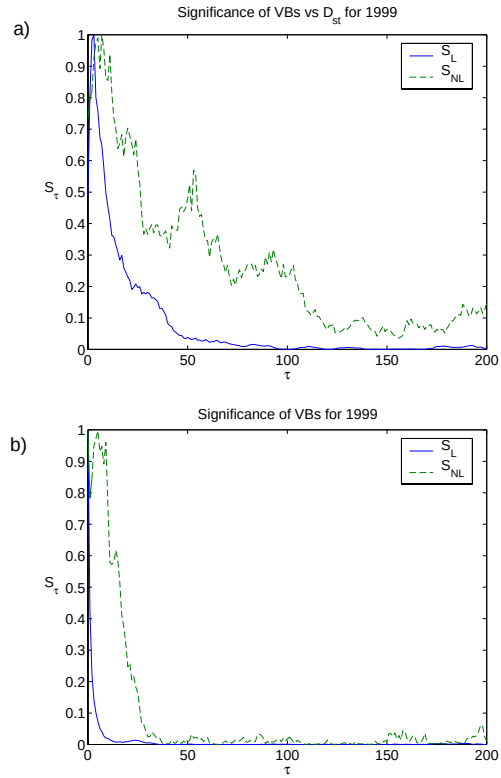

Figure 14. Significance extracted from (a) $\{VBs(t), D_{st}(t-\tau)\}$ and (b) $\{VBs(t), VBs(t-\tau)\}$ for 1999. It should be noted that there is a strong linear response at around 3 hour time delay. There is a clear nonlinear response with a strong peak around 50 hours lasting for approximately 1 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.

4.5. Cumulant-based information flow

The information flow is also a good measure of dynamical behavior of the system and especially predictability of the system. In Figure 15 we compute an information flow based on the multivariate vector $\{x_1, x_2, x_3, x_4\} = \{D_{st}(t+p), D_{st}(t), VBs(t), VBs(t-\Delta)\}$. The integrated information flow , $I_C(p, \Delta)$, is obtained as a function of look ahead and changes in the solar wind VBs data.

In Figure 16 we plot the integrated information computed in a sliding window of width 150 hours for the year 1999. Information is integrated up to P=20 hours and a time delay Δ=1 hour is considered in VBs. An important point is that there tend to be

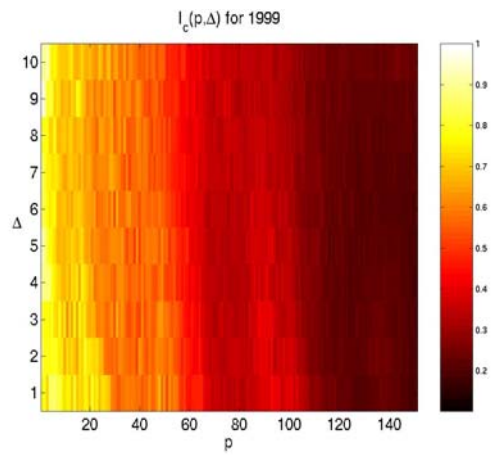

Figure 15. Information flow, $I_C(p, \Delta)$ as a function of look ahead, p, and solar wind delay time, Δ for 1999. Notice that the information decays on the order of 1 week which is a typical time scale for storm response. The information flow remains relatively high on the order of 20 hours and there is a strong nonlinear response at around 40-50 hours as detected based on the significance. In considering the integrated information, we will integrate up to 20 hours ahead.

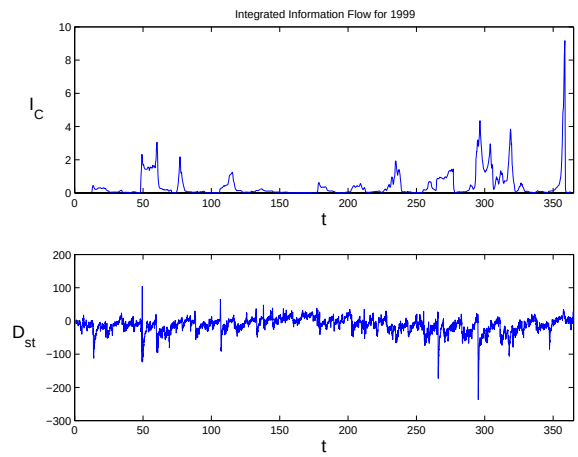

Figure 16. Integrated information flow verses day of 1999 based on a 20 hour look ahead. There are clearly times when there is high predictability. The integrated information flow appears to be largest during times of storm activity indicating increased information flow and better predictability. The most reproducible and predictable behavior of D_{st} appears to be the recovery of the index; however, the information flow does increase significantly prior to the onset of the storm indicating that the storm onset could also be predicted.

times of reasonable predictability.

5. Discussion

The non-parametric studies described above can be used to guide development of parametric models. The output from models depends heavily on the method by which the model is trained. If a training set is noisy, for example, the output is often spoiled or not well constrained. The windowed significance provides a measure of the significance of data within a dataset. Data with significance that does not exceed the surrogate data usually is corrupted by noise and is not useful for training purposes. Such data could result when the method to gather the data is corrupted by noise or by intrinsic noise in the dynamical system. On the other hand, data with high significance exhibits strong dependencies and is therefore more likely to be useful when trying to model those dependencies. Thus, the significance measure can be used to eliminate noisy data from a dataset. One consequence is that models trained with data that best captures the nonlinear dependencies should predict the dynamical evolution of the system better.

Another useful property of the significance is that it can be used to detect changes in the underlying dynamics of a system. When the system dynamics change, the significance detects this change. The magnetosphere is believed to be a multi-state system. Magnetospheric behavior is often classified as active or quiescent. The windowed cumulant-based significance is useful for detecting changes in the state of the system. Moreover, it is useful for classifying data according to the state of the system. The dynamics of those states are usually different, and therefore, it would be useful to apply different models to those two different states. By assessing the significance of data and separating it according to its state it is possible to fine tune models to describe the state of the system as illustrated in Figure 17. It is also possible to dynamically determine when to switch dynamical states using the significance measure for an incoming stream of data.

The cumulant-based information flow (computed by summing over the look-ahead significance) can be used as a proxy for information flow in the system. This information flow can be computed from the data as well as from the neural network. The network can

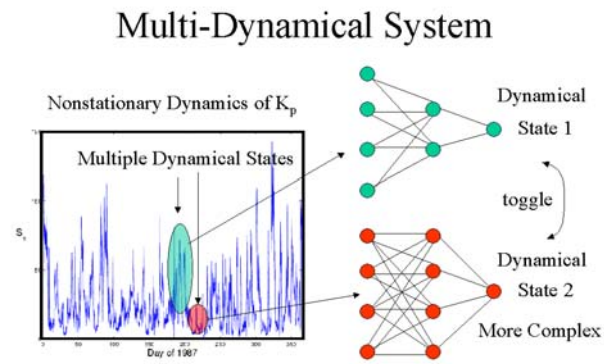

Figure 17. An illustration of how the windowed significance can be used to identify the state of a system. Larger significance usually indicates a system with simpler underlying dynamics while a state with lower significance is usually more complex. The data can be binned according its windowed significance so that a parametric neural network model can specialize in each state. The instantaneous significance measure can be used to identify whether to toggle between the two networks which predict the state of the system.

be trained so as to best approximate the information flow rather than to simply minimize the error in the training set (which does not account for the possibility of noise contamination). The network can be modified (i.e. new nodes added or removed) in a systematic way so as to best approximate the information flow with the minimum number of parameters. By requiring that information flow is well approximated by the parametric model, we are more assured that the underlying dynamics are faithfully modeled than with simple error minimization.

6. Summary

In this study, we have applied information dynamical techniques to the K_p and D_{st} magnetospheric indices. Using the cumulant-based significance, we have established that the underlying dynamics of K_p evolution is, in general, nonlinear exhibiting a quasi-periodicity which is detectable only if nonlinear correlations are taken into account. As such one expects that linear auto-regressive moving average (ARMA) models commonly used to predict magnetospheric response should be inaccurate. Local-linear models (which include slow evolution of parameters) are also likely to fail where the dynamics suddenly change which occurs regularly in Figures 13 and 16. A promising alternative would be to train a neural network with data identified to have large significance so as to avoid irrelevant noise which normally spoils the generalization characteristics of the neural network. Additional improvements could be realized by using the fact that magnetospheric dynamics are not stationary to separate the data into "states" and fine tune predictors for those states.

Our analysis of the information flow and integrated information flow provides additional insight. The system appears to have reasonable predictability properties on the time scale of 10 hours—a time scale that would be of practical value for protecting vulnerable satellite equipment. However, nonlinear effects already play a significant role even on that time scale. Longer term forecasting may also be possible as evidenced by peaks in the information flow. Improvements to parametric modeling could also be realized by training parametric models to approximate the information flow in the real system.

Acknowledgments The authors would like to thank J. Freeman for sharing the Costello NN code with us. This work was supported by NASA Grant NAG5-10971, NSF Grant ATM-9819705, NASA Grant W-19880, NSF Grant ATM-0218847, and DoE Contract No. DE-AC02-76-CHO-3073.

References

Berchem, J., J. Raeder, M. Ashour-Abdalla, L. A. Frank, W. R. Paterson, K. L. Ackerson, S. Kokubun, T. Yamamoto, and R. P. Lepping, The distant tail at 200 R_E: Comparison between Geotail observations and the results from a global magnetohydrodynamic simulation, *J. Geophys. Res.*, *103*, 9121–9142, 1998.

Burton, R. K., R. L. McPherron, and C. T. Russell, An emperical relationship between interplanetary conditions and Dst, *J. Geophys. Res.*, *80*, 4204–4214, 1975.

Chang, T., Self-organized criticality, multi-fractal spectra, sporadic localized reconnections and intermittent turbulence in the magnetotail, *Phys. Plasmas*, *6*, 4137–4145, 1999.

Chang, T., S. W. Y. Tam, C. Wu, and G. Consolini, Complexity, Forced and/or Self-Organized

Criticality, and Topological Phase Transitions in Space Plasmas, *Space Science Reviews*, *107*, 425–445, 2003.

Chapman, S., and N. Watkins, Avalanching and Self-Organised Criticality, a paradigm for geomagnetic activity?, *Space Science Reviews*, *95*, 293–307, 2001.

Cheng, C. Z., and A. T. Y. Lui, Kinetic ballooning instability for substorm onset and current disruption observed by AMPTE/CCE, *Geophys. Res. Lett.*, *25*, 4091–4094, 1998.

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

Detman, T., and J. A. Joselyn, Real-time Kp predictions from ACE real time solar wind, in *Solar Wind Nine*, edited by S. R. Habbal, R. Esser, J. V. Hollweg, and P. A. Isenberg, pp. 729–732. The American Institute of Physics, 1-56396-865-7, 1999.

Freeman, J. W., et al., The magnetospheric specification and forecast model, preprint available at http://hydra.rice.edu/ freeman/ding/www/msfm95/index.html, 1995.

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

Horton, W., and I. Doxas, A low-dimensional dynamical model for the solar wind driven geotail-ionosphere system, *J. Geophys. Res.*, *103*, 4561–4572, 1998.

Kennel, M. B., and S. Isabelle, Method to dinstinguish possible chaos from colored noise and to determine embedding parameters, *Phys. Rev. A*, *46*, 3111, 1992.

Klimas, A., J. A. Valdivia, D. Vassiliadis, D. N. Baker, M. Hesse, and J. Takalo, Self-organized criticality in the substorm phenomenon and its relation to localized reconnection in the magnetospheric plasma sheet, *J. Geophys. Res.*, *105*, 18,765–18,780, 2000.

Klimas, A. J., D. Vassiliadis, and D. N. Baker, Data-derived analogues of the magnetospheric dynamics, *J. Geophys. Res.*, *102*, 26993–27010, 1997.

Klimas, A. J., D. Vassiliadis, and D. N. Baker, Dst index prediction using data-derived analogues of the magnetospheric dynamics, *J. Geophys. Res.*, *103*, 20435–20448, 1998.

Li, X., M. Temerin, D. N. Baker, G. D. Reeves, and D. Larson, Quantitative Prediction of Radiation Belt Electrons at Geostationary Orbit Based on Solar Wind Measurements, *Geophys. Res. Lett.*, *28*, 1887, 2001.

Lorenz, E., Deterministic nonperiodic flows, *J. Atmosph. Sci.*, *96*, 5877–5882, 1963.

Lui, A. T. Y., et al., Current disruptions in the near-Earth neutral sheet region, *J. Geophys. Res.*, *97*, 1461, 1992.

Lui, A. T. Y., S. C. Chapman, K. Liou, P. T. Newell, C. I. Meng, M. Brittnacher, and G. K. Parks, Is the dynamic magnetosphere an avalanching system?, *Geophys. Res. Lett.*, *27*, 911, 2000.

O'Brien, T. P., and R. L. McPherron, An empirical phase space analysis of ring current dynamics: Solar wind control of injection and decay, *J. Geophys. Res.*, *105*, 7707–7720, 2000.

Prichard, D., J. E. Borovsky, P. M. Lemons, and C. P. Price, Time dependence of substorm recurrence: An information-theoretic analysis, *J. Geophys. Res.*, *101*, 15359–15370, 1996.

Raeder, J., et al., Global simulation of the Geospace Environment Modeling substorm challenge event, *J. Geophys. Res.*, *106*, 381–396, 2001.

Roberts, D. A., D. N. Baker, A. J. Klimas, and L. F. Bargatze, Indications of low dimensionality in magnetospheric dynamics, *Geophys. Res. Lett.*, *18*, 151–154, 1991.

Roux, A., S. Perraut, A. Morane, P. Robert, A. Korth, G. .Kremser, A. Pederson, R. Pellinen, and Z. Y. Pu, Plasma sheet instability related to the westward traveling surge, *J. Geophys. Res.*, *96*, 17697, 1991.

Sauer, T., J. Yorke, and M. Casdagli, Embedology, *Journal of Statistical Physics*, *65*, 579, 1991.

Sharma, A. S., Assessing the magnetosphere's nonlinear behavior: Its dimension is low, its predictibility, high, *Rev. Geophys.*, *33*, 645–650, 1995.

Sitnov, M. I., A. S. Sharma, K. Papadopoulos, D. Vassiliadis, J. A. Valdivia, A. J. Klimas, and D. N. Baker, Phase transition-like behavior of the magnetosphere during substorms, *J. Geophys. Res.*, *105*, 12955–12974, 2000.

Takens, F., Detecting strange attractors in turbulence., in *Dynamical Systems and Turbulence*, edited by D. A. Rand, and L. S. Young, vol. 898 of *Lecture Notes in Mathematics*, p. 366. Springer-Verlag, 1980.

Temerin, M., and X. Li, A new model for the prediction of Dst on the basis of the solar wind, *Journal of Geophysical Research (Space Physics)*, pp. 31–1, 2002.

Vassiliadis, D., A. S. Sharma, and K. Papadopoulos, Lyapunov exponent of magnetospheric activity from AL time series, *Geophys. Res. Lett.*, *18*, 1643–1646, 1991.

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

Vassiliadis, D., A. J. Klimas, J. A. Valdivia, and D. N. Baker, The D_{st} geomagnetic response as a function of storm phase and amplitude and the solar wind electric field, *J. Geophys. Res.*, *104*, 24957–24976, 1999.

Vassiliadis, D. V., A. S. Sharma, T. E. Eastman, and K. Papadopoulos, Low-dimensional chaos in magnetospheric activity from AE time series, *Geophys. Res. Lett.*, *17*, 1841–1844, 1990.

This preprint was prepared with the AGU LaTeX macros v3.0. File CUM formatted 2004 January 28.

With the extension package 'AGU++', version 1.2 from 1995/01/12