

## Probability distribution invariance of 1-minute auroral-zone geomagnetic field fluctuations

R. S. Weigel and D. N. Baker

University of Colorado, Laboratory for Atmospheric and Space Physics, Boulder, Colorado, USA

Received 21 August 2003; revised 30 October 2003; accepted 6 November 2003; published 4 December 2003.

[1] A statistical model of short time-scale geomagnetic fluctuations is developed and used to evaluate how geomagnetic dynamics are influenced by different solar wind controlling parameters. The functional form of the probability distribution function (PDF) that describes extreme-value (greater than  $4\sigma$ ) minute-to-minute changes in the ground magnetic field ( $\Delta x$ ) at magnetometer station Sodankylä (geomagnetic latitude and longitude of [63.87,107.61]) is shown to be nearly independent of the variables solar wind (SW) forcing, local time (LT), and day of year (DOY). Instead of modifying the intrinsic dynamics, as characterized by the functional form of the PDF of  $\Delta x$ , these variables are shown either to amplify or reduce the absolute level of variability of the fluctuations: The primary difference in the PDF tail of  $\Delta x$  during weak and strong solar wind forcing is the standard deviation,  $\sigma$ ; the functional form of the PDF =  $f[\Delta x/\sigma(\text{DOY,LT,SW})]$  is nearly invariant. In a statistical interpretation, we conclude that differences in solar-generated conductivity, seasonal effects, strength of solar wind forcing and variability, and position of the magnetometer ground station in local time do not change the structure of the extreme-value dynamics, as characterized by the probability distribution of  $\Delta x$ , but they serve to amplify the intrinsic variability. **INDEX TERMS:** 2437 Ionosphere: Ionospheric dynamics; 2708 Magnetospheric Physics: Current systems (2409); 3210 Mathematical Geophysics: Modeling; 7839 Space Plasma Physics: Nonlinear phenomena; 7863 Space Plasma Physics: Turbulence. **Citation:** Weigel, R. S., and D. N. Baker, Probability distribution invariance of 1-minute auroral-zone geomagnetic field fluctuations, *Geophys. Res. Lett.*, 30(23), 2193, doi:10.1029/2003GL018470, 2003.

### 1. Introduction

[2] The earliest approaches in characterizing solar wind driving of geomagnetic dynamics were based on a linear circuit analogy for which the differential equation that describes changes in geomagnetic or magnetospheric variables  $x$  due to driving by the solar wind variable  $S$  has the general form [Clauer *et al.*, 1981]

$$\frac{dx}{dt} = A \cdot x + S(t). \quad (1)$$

The parameters in the matrix  $A$  are based on estimates of bulk electrical parameters of the magnetosphere. This model is able to capture the low-pass-filter (or “direct driving”) response of the magnetosphere to the solar wind [Vassiliadis

and Klimas, 1995]. The portion of the signal that this model does not capture is often substantial and highly non-Gaussian, however. The inadequacy of a linear model in reproducing the measured transfer function at high frequencies was shown clearly by Tsurutani *et al.* [1990], who found that the power spectral density ratio of  $x = AE$  to  $S = B_s$  fell off as  $1/f^{0.5}$  at high frequencies, as opposed to  $1/f^2$ , which is expected from Equation 1 with negative real eigenvalues.

[3] To better account for the measured nonlinear transfer function that connects auroral electrojet measurements and the solar wind driver, nonlinear models were introduced. These models accounted for the rapid changes in the field that are typically associated with a substorm. These models are generally of the form [Baker *et al.*, 1990; Klimas *et al.*, 1994; Horton and Doxas, 1998]

$$\frac{dx}{dt} = f(x) + \mathcal{S}(S(t), x, t) \cdot \Theta(x - x_o), \quad (2)$$

where  $x_o$  is a threshold value above which the magnetosphere suddenly unloads its stored energy onto the ionosphere. These equations are able to reproduce the feature of sudden changes in the ground magnetic field in the midnight sector that occur following an interval of extended southward IMF. However, for  $x$  that is not near  $x_o$ , the system has a transfer function similar to that of Equation 1, and thus will not capture large non-substorm geomagnetic fluctuations. The predictions from MHD models [Ridley *et al.*, 2001] and sum-of-sigmoid mapping functions [Gleisner and Lundstedt, 1997; Weigel *et al.*, 1999], which take solar wind variables as inputs, also have difficulty predicting short time-scale geomagnetic fluctuations.

[4] Although these modeling approaches are able to predict or capture the gross, large-scale behavior (low-frequency and rapid unloading after extended southward IMF) of geomagnetic measurements, they tend not to reproduce or predict the high-frequency fluctuations whose amplitudes are often of the same magnitude as the background amplitude [Crowley and Hackert, 2001; Weigel *et al.*, 2003]. Without having an accurate or computationally efficient deterministic representation of the causes of these short time-scale fluctuations, we can take a stochastic modeling approach in which the unknown influences are regarded as a random forcing term in the dynamical equation [Sobczyk, 1990].

[5] The first step in building a stochastic representation of the dynamics is to determine the statistical properties of the measurements. The stochastic model we consider is for one-minute changes in the ground magnetic field  $dx/dt$  (denoted as  $\Delta x$ , with  $\Delta t = 1$  minute). The model specifies

the probability  $P$  of  $\Delta x$  as a function of spatial location and the average solar wind state,  $S$ ,

$$P(\Delta x) = F(\Delta x, DOY, LT, \langle S \rangle). \quad (3)$$

Because many of the drivers of geomagnetic activity are well-known, we must first attempt to evaluate and isolate their influence on the distribution function,  $F$ . We do this by taking an average of different known SW drivers of activity, represented by  $\langle S \rangle$  in Equation 3.

[6] Attempts have been made to interpret the dynamics of geomagnetic and magnetospheric fluctuations in terms of turbulence or self-organization by evaluating the PDF of long records of geomagnetic indices [Watkins *et al.*, 2001 and references therein]. Analysis has shown that many geomagnetic indices have a PDF of  $\Delta x$  or  $X$  that is heavy-tailed (PDF falls off slower than that of a Gaussian distribution). However, solar wind measurements also have heavy-tailed PDFs, and it has been noted that the geomagnetic statistics may only be a reflection of the solar wind statistics [Freeman *et al.*, 2000; Price and Newman, 2001; Hnat *et al.*, 2002]. Moreover, an additional difficulty exists in determining how index distributions relate to the geomagnetic dynamics that the indices are derived from. This difficulty is due to the fact that the indices are produced by taking the instantaneous maximum or minimum of a set of measurements. This indexing process itself will create a heavy-tailed distribution if the underlying measurements have any of the forms in a broad class of stable distribution functions [Bury, 1999]. Because of the difficulty in extracting the unique underlying distribution function of the underlying measurements from geomagnetic indices, individual geomagnetic and solar wind measurements are used in this work.

## 2. Analysis

[7] The analysis is performed on a 22-year data set that includes minute 1 of 1980 through the last minute of 2001 of the north-south component of the ground magnetic field  $X$  measured at the auroral-zone magnetometer station Sodankylä (SOD). This station has geomagnetic coordinates (in 1998) of [63.87, 107.61]. The hourly-averaged solar wind measurements of  $V_{sw}$  and  $B_z$  were obtained from OMNIWeb (Although preferred, shorter time averages of solar wind data are not consistently available in the years under consideration.).

[8] The fluctuations are characterized by using the minute-to-minute differences in the north-south component of the ground magnetic field  $\Delta x(t) \equiv X(t) - X(t-1)$  (analysis of the  $H$  and  $Y$  components yields the same results). This high-pass filter operation has low sensitivity to long-term trends in the data, and it is a method generally employed to study the scale of structures in turbulence [Frisch, 1995]. Moreover,  $\Delta x$  (as opposed to  $X$ ) is more relevant from a practical point of view: Although large  $X$  is associated (correlated) with large  $\Delta x$ ,  $\Delta x$  is a better proxy for geomagnetically induced currents (GICs) [Boteler *et al.*, 1998; Pirjola, 2000; Viljanen *et al.*, 2001]. (Faraday's law of induction states that a changing magnetic field implies the existence of a local electric field, which can drive currents in a conducting system.)

[9] The probability distribution function of  $\Delta x$  normalized to its standard deviation and partitioned by season, local time, strength of solar wind forcing, and solar wind variability ( $\sigma(B_z)$  and  $\sigma(V_{sw})$ ) is shown in Figure 1. The seasonal PDF was generated by selecting data in the range of  $\pm 40$  days around the spring equinox and the two solstices. The solar wind data are based on one-hour average values. The solar wind classification was created by selecting one-day intervals in which the average of the available (hourly-averaged) solar wind quantity satisfied the inequality in the legend of Figure 1. The partition limits were chosen so that each PDF has  $2.5 \pm 0.1 \cdot 10^6$  elements. If there were no valid solar wind measurements in an interval, the  $\Delta x$  data were omitted from the analysis.

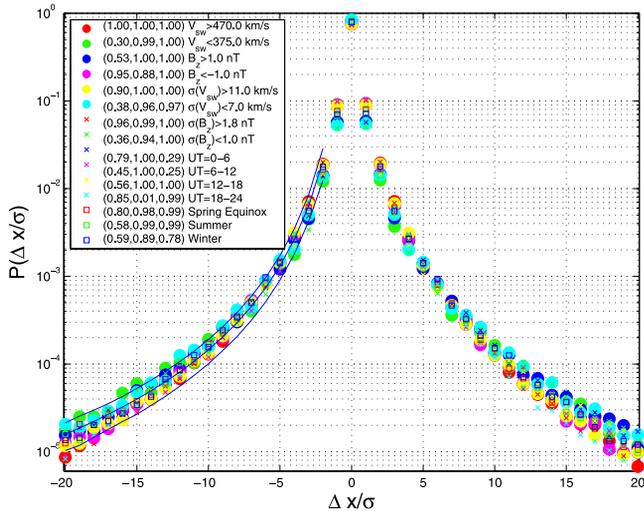
[10] Visually, all of the distributions in Figure 1 have nearly the same extreme-value behavior. To quantify this similarity, we use the Kolmogorov-Smirnov (K-S) test of significance. The significance level of the hypothesis that the distribution is the same as the reference distribution ( $V_{sw} > 433$  km/s) for  $\Delta x/\sigma < -B_i$  and  $\Delta x/\sigma > B_i$  and  $B_i = 4$  is listed as the second and third number in parenthesis, respectively. Note that in general the hypothesis can be accepted with high confidence, but exceptions exist. For example, for the UT = 18–24 distribution, the significance level is very low, indicating a low likelihood that its data were drawn from the same distribution as the reference distribution. However, if the range of evaluation is changed to  $\Delta x/\sigma < -8$ , we find that this significance increases to 0.99. There are 105 possible pair combinations of the 15 distributions, and in general the K-S statistic is high, but some exceptions exist. What is generally found is that the K-S statistic is high for all pairs when the inner boundary,  $B_i$ , is between 2 and 8 and for  $B_i < 2$  the differences become substantial. (The error lines in each bin of Figure 1 correspond to  $\pm 2$  standard deviations assuming the number of points in the bin is Gaussian distributed about its mean. Using this assumption can result in underestimates of the true error bars, especially for heavy-tailed distributions.) From Figure 1, we conclude that the distribution of  $\Delta x$  is nearly invariant with respect to the reference distribution when sorted by a single variable.

[11] Figure 2a shows the standard deviation of  $\Delta x$  as a function of LT and DOY. The standard deviation,  $\lambda$ , and  $E$  at each point were computed using data from the 22-year time series in the range of  $\pm 40$  days in DOY and  $\pm 2$  hours in MLT (216,480 total values). There is a well-defined seasonal and local time dependence, with equinox and the midnight sector having the highest variability.

[12] As an alternative to computing the K-S statistic between every pair of distributions on the grid, we use the single-parameter Frechet PDF as a reference, which has the form

$$P(z) = \lambda z^{-\lambda-1} e^{-z^{-\lambda}}. \quad (4)$$

The best-fit parameter  $\lambda$  is computed for each distribution on the grid along with the fitting error  $E$ . (Fitting Equation 4 to each distribution in Figure 1 gives values of  $\lambda$  ranging between 1.9 and 2.5 and an average of 2.2, all with approximately the same level of fitting error.) It is important to note that there are many processes which can generate a heavy-tailed distribution, and many PDFs with more parameters fit the data with approximately the same



**Figure 1.** PDF of one-minute fluctuations sorted by various parameters. The collection of  $\Delta x$  data for each class is scaled by its own standard deviation. The numbers in parenthesis represent, in order, the standard deviation with respect to  $\sigma_o = 12.4$  nT/min, the K-S statistic for negative  $\Delta x$ , and the K-S statistic for positive  $\Delta x$ . The center solid line is Equation 4 using the average value  $\lambda = 2.2$ . The upper and lower solid lines are the error limits described in the text.

accuracy as the above distribution. For example, the single-parameter power-law distribution ( $1/z^\lambda$ , with  $\lambda \simeq 3$ ) has a slightly higher average error than that the Fréchet distribution. Also, both a 2-parameter lognormal and a 3-parameter stretched exponential distribution fit with the same error in the considered range. Evaluating the K-S statistic leads to the conclusion that the data distributions are unlikely to have been drawn from any of these PDFs. Our motivation for using Equation 4 is as a method for showing that the PDFs have nearly the same extreme-value form. Without further analysis and hypothesis testing, it is difficult to infer with statistical significance which distribution is most appropriate, if any.

[13] Figure 2b shows that the best-fit  $\lambda$  has only two distinct regions; these regions have the same  $\lambda$ , while the regions between have a lower distribution index  $\lambda$ . Equation 4 has the property that the probability of large events increases with decreasing  $\lambda$ . Thus, this decrease in  $\lambda$  is an indication that relatively large-amplitude events are more likely in the transition region between the large-scale east-west and west-east electrojets. There are several possible causes for this difference. One cause may be geometrical: On average the equivalent large scale east-west current in these regions is zero. During strong magnetospheric driving, the eastward and westward current systems both grow rapidly; it is the intersection of these opposing currents that may cause a greater degree of relative variability.

[14] A key result of Figure 2 is that spatial locations which have drivers of long time-scale geomagnetic dynamics that are quite different in terms of solar wind coupling are similar with respect to their distribution of large-amplitude short time-scale fluctuations. For example, in the pre-noon sector Weigel *et al.* [2002] showed that 30-minute averages of one-minute values of  $|dB_x/dt|$  at Sodankylä (SOD) are mostly

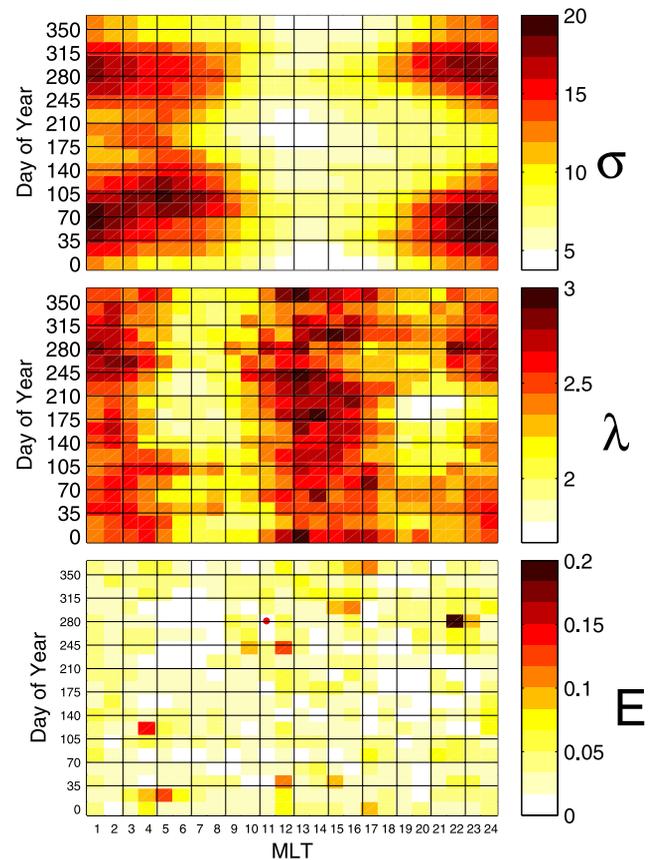
predicted by  $V_{sw}$ . This was interpreted as indicating that the model was capturing the driving of  $|dB_x/dt|$  by the Kelvin-Helmholtz instability process. In the midnight sector, both  $V_{sw}$  and  $B_z^{IMF}$  contributed to the data-model correlation, an indication that the model was capturing reconnection-based driving. Figures 2a and 2b clearly show that the PDF( $\Delta x$ ) in these spatial locations only differ by a scale factor,  $\sigma$ .

[15] Figure 2c shows that the error in the fit to Equation 4 is generally small and constant for all local times and DOYs, and there is no clear trend. This shows that the reference distribution of Equation 4 fits the data in each location with similar accuracy. As in the analysis of Figure 1, these errors are larger than what would be expected if they were drawn from a true Fréchet distribution.

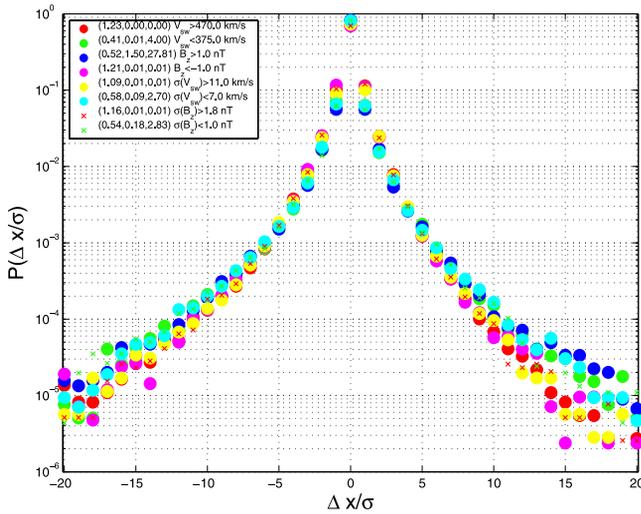
[16] Figure 3 shows that the form of the PDF is similar under very different solar wind forcing conditions in a MLT range that is highly responsive (in terms of amplitude) to solar wind driving. The error bars were computed in the same manner as in Figure 1, and the total number of points for each distribution is  $3 \pm 0.1 \cdot 10^5$ . Again, the K-S statistic between each distribution and the reference distribution is generally high.

### 3. Discussion and Conclusion

[17] Based on the observation that deterministic geomagnetic models have difficulty in predicting or capturing intermittent, high-frequency fluctuations, we have taken a



**Figure 2.** (a) Standard deviation, (b) best-fit PDF scaling exponent in the range  $2 \leq |\sigma| \leq 13$ , and (c) error in best-fit calculation as a function of day of year and magnetic local time (MLT).



**Figure 3.** PDF of one-minute fluctuations for magnetic local times (MLTs) in the range 21–24 sorted by average of solar wind measurements in the same interval. The numbers in parenthesis are in the same format as Figure 1.

statistical modeling approach in which the probability of  $\Delta x$  is computed as a function of the day of year, local time, and average solar wind state according to

$$P(\Delta x) = F(\Delta x, DOY, LT, \langle S \rangle). \quad (5)$$

We have shown that  $F$  can be approximately simplified as

$$F = f\left(\frac{\Delta x}{\sigma(DOY, LT, \langle S \rangle)}\right), \quad (6)$$

where  $f(z) = \lambda z^{-\lambda-1} e^{-z^{-\lambda}}$ , and  $\lambda$  has a slight dependence on local time and the day of year.

[18] As shown in this paper, the tails of the distributions of  $\Delta x/\sigma$  are nearly independent of the level of solar wind forcing and the level of solar wind variability. We use the term “nearly independent” in the sense that either (1) the significance level of the hypothesis that the data were drawn from the same distribution as a reference distribution is generally high or (2) that all of the sub-distributions fit a single-parameter reference distribution with similar accuracy but have parameter values that differ slightly. Because the average solar wind state primarily enters through the standard deviation, the solar wind input can be viewed as an amplifier of the signal  $\Delta x$ .

[19] The results indicate that the PDF tail of  $\Delta x$  is primarily a reflection of an intrinsic property of short time-scale auroral-zone geomagnetic fluctuations and is not a reflection of the solar wind driver state or the location of the magnetometer in local time. (Analysis of data from the Yellowknife magnetometer, which is also located in the auroral-zone, yields the same conclusion.) Perhaps most surprisingly, the form of the PDF tails for large  $|\Delta x/\sigma|$  in the midnight sector is similar to that in dayside local times. This indicates that there is no unique signature of magnetotail (local midnight) dynamics that is reflected in the shape of the PDF tails of one-minute magnetometer fluctuations.

[20] The heavy-tail distribution functions found in this work are consistent with *Viljanen et al.* [2001] who found that the distribution patterns of the averages of the horizontal field  $\mathbf{H}$  and  $d\mathbf{H}/dt$  implies the existence of small-scale

currents that are not explained by a sheet-type model of ionospheric currents. Although large-scale convection controls the background flow of the current systems, it is the two-dimensional nature of the ionospheric conductor that allows small-scale structures to exist.

[21] **Acknowledgments.** This material is based upon work supported by CISM which is funded by the STC Program of the National Science Foundation under Agreement Number ATM-0120950. The authors thank the maintainers of the Sodankylä magnetometer instrument and data and the World Data Center for providing access to the high-quality data set.

## References

- Baker, D. N., A. J. Klimas, R. L. McPherron, and J. Buchner, The evolution from weak to strong geomagnetic-activity - an interpretation in terms of deterministic chaos, *Geophys. Res. Lett.*, 17, 41–44, 1990.
- Boteler, D. H., R. J. Pirjola, and H. Nevanlinna, The effects of geomagnetic disturbances on electrical systems at the earth’s surface, *Advances in Space Research*, 22, 17–27, 1998.
- Bury, K., *Statistical Distributions in Engineering*, Cambridge University Press, 1999.
- Clauer, C. R., R. L. McPherron, C. Searls, and M. G. Kivelson, Solar-wind control of auroral-zone geomagnetic-activity, *Geophys. Res. Lett.*, 8, 915–918, 1981.
- Crowley, G., and C. Hackert, Quantification of high latitude electric field variability, *Geophys. Res. Lett.*, 28, 2783–2786, 2001.
- Freeman, M. P., N. W. Watkins, and D. J. Riley, Evidence for a solar wind origin of the power law burst lifetime distribution of the AE indices, *Geophys. Res. Lett.*, 27, 1087–1090, 2000.
- Frisch, U., *Turbulence. The legacy of A.N. Kolmogorov*, Cambridge University Press, 1995.
- Gleisner, H., and H. Lundstedt, Response of the auroral electrojets to the solar wind modeled with neural networks, *J. Geophys. Res.*, 102, 14,269–14,278, 1997.
- Hnat, B., S. C. Chapman, G. Rowlands, N. W. Watkins, and M. P. Freeman, Scaling of solar wind epsilon and the AU, AL and AE indices as seen by WIND, *Geophys. Res. Lett.*, 29, doi:10.1029/2001GL014587, 2002.
- 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.
- Klimas, A. J., D. N. Baker, D. Vassiliadis, and D. A. Roberts, Substorm recurrence during steady and variable solar-wind driving - evidence for a normal-mode in the unloading dynamics of the magnetosphere, *J. Geophys. Res.*, 99, 14,855–14,861, 1994.
- Pirjola, R., Geomagnetically induced currents during magnetic storms, *IEEE Trans. Plasma Sci.*, 28, 1867–1873, 2000.
- Price, C. P., and D. E. Newman, Using the R/S statistic to analyze AE data, *J. of Atmos. and Solar-Terrestrial Phys.*, 63, 1387–1397, 2001.
- Ridley, A. J., D. L. De Zeeuw, T. I. Gombosi, and K. G. Powell, Using steady state MHD results to predict the global state of the magnetosphere-ionosphere system, *J. Geophys. Res.*, 106, 30,067–30,076, 2001.
- Sobczyk, K., *Stochastic differential equations with applications to physics and engineering*, Kluwer Academic Publishers, 1990.
- Tsurutani, B. T., B. E. Goldstein, M. Sugiura, T. Iyemori, and W. D. Gonzalez, The nonlinear response of AE to the IMF Bs driver - a spectral break at 5 hours, *Geophys. Res. Lett.*, 17, 279–282, 1990.
- Vassiliadis, D., and A. Klimas, On the uniqueness of linear moving-average filters for the solar wind-auroral geomagnetic-activity coupling, *J. Geophys. Res.*, 100, 5637–5641, 1995.
- Viljanen, A., H. Nevanlinna, K. Pajunpää, and A. Pulkkinen, Time derivative of the horizontal geomagnetic field as an activity indicator, *Annales Geophysicae*, 19, 1107–1118, 2001.
- Watkins, N. W., M. P. Freeman, S. C. Chapman, and R. O. Dendy, Testing the SOC hypothesis for the magnetosphere, *J. of Atmos. and Solar-Terrestrial Phys.*, 63, 1435–1445, 2001.
- Weigel, R., W. Horton, T. Tajima, and T. Detman, Forecasting auroral electrojet activity from solar wind input with neural networks, *Geophys. Res. Lett.*, 26, 1353–1356, 1999.
- Weigel, R. S., D. Vassiliadis, and A. Klimas, Coupling of the solar wind to temporal fluctuations in ground magnetic fields, *Geophys. Res. Lett.*, 29, 2002.
- Weigel, R. S., A. Klimas, and D. Vassiliadis, Precursor analysis and prediction of large-amplitude relativistic electron fluxes, Accepted to *Space Weather*, 2003.