32nd URSI GASS, Montreal, 19-26 August 2017

Gaining the most utility from our geospace observational system: Network analysis of total electron content as a means to understand space weather to the point of prediction

Ryan M. McGranaghan[1,2], Anthony Mannucci[2], Olga Verkhoglyadova[2], and Nishant Malik[3]
(1) Cooperative Programs for the Advancement of Earth System Science (CPAESS), University Corporation for Atmospheric Research (UCAR), Boulder, CO, 80301, http://www.cpaess.ucar.edu
(2) NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109
(3) Department of Mathematics, Dartmouth College, Hanover, NH, 03755

Abstract

We present the first network analysis of interplanetary magnetic field (IMF) clock angle dependent, high-latitude, hemispheric-specific total electron content (TEC) data. We examine network parameters to describe spatio-temporal correlations in the TEC data for January 2016. We find that significant network structure exists distinguishing the dayside and nightside ionosphere, and specific features in the high-latitudes (cusp/ionospheric footpoints of magnetospheric boundary layers, polar cap, and auroral zone), and that these features vary with IMF clock angle. In this brief summary paper, we provide proof of concept results and identify important areas of future research, providing a basis for the discussion of network analysis and machine learning approaches for space weather applications.

1. Introduction

Global Positioning System (GPS) signals are one of the premier remote sensing tools to further understanding of space weather phenomena [1]. GPS signal delays due to propagation through the dispersive ionosphere can be used to yield Total Electron Content (TEC) data, which provide critical information about the Earth's ionosphere at higher cadence and over a larger portion of globe than any other single data source [2]. Large volumes of data augmented by the proliferation of Global Navigation Satellite Systems (GNSS) around the world and their relatively reliable future availability make these data critical to the future study of the Earth's space environment and its variability, or space weather.

Global, high-latitude response of TEC is the result of numerous complex geospatial processes, each with unique spatial and temporal scales [3]. Therefore, TEC data are rich with information about the Earth's space environment. However, the complex nature of these processes requires innovative and sophisticated approaches to: 1) understand the information content of these data; and 2) gain the most utility from them. In this paper, we attempt to understand the spatio-temporal characteristics of TEC in the high-latitude regime and assess the extent to which these data can be used to understand, and, ultimately, predict space weather phenomena.

We present a new, complex networks, approach to the analysis of high-latitude, hemispheric-specific, TEC data known as network analysis. The use of network analysis [4] originated in the social sciences, but has since expanded to biological, engineering, and, more recently, geophysical systems applications [5,6,7]. Networks are defined by nodes, or in geophysical applications, physical grid points, and connections, defined by spatio-temporal correlations that exceed a given threshold value. Given the nodes and connections, a network can be constructed and parameters calculated to describe the network topology [8]. In this paper we examine TEC network topology in the context of space weather phenomena and understanding. We perform separate network analyses on TEC data organized by interplanetary magnetic field (IMF) clock angle:

$$\theta = \tan^{-1}\left(\frac{B_{\mathrm{Y,GSM}}}{B_{\mathrm{Z,GSM}}}\right) \quad (1).$$

This work represents two significant firsts: 1) the first examination of IMF clock angle dependence of high-latitude, hemispheric-specific TEC in magnetic coordinates; and 2) the first time network methodology has been applied to TEC data. We conclude that our network analyses create a new means for discovery and understanding of the geospace system and space weather that complements existing approaches.

In Section 2 we describe the TEC data processing and network construction, Section 3 introduces our network analysis, Section 4 presents the results of our analyses and a discussion of future research, and Section 5 provides concluding remarks.

2. Data

We process TEC data from January 2016 obtained from the Madrigal upper atmospheric science database (http://madrigal3.haystack.mit.edu/), which compiles data from the worldwide system of ground-based GPS receivers into 1° latitude x 1° longitude geographic coordinate distributions at 5-minute cadence. We exclude all data for which the Madrigal TEC error estimate exceeds the 95th percentile. We use these data to form networks of high-latitude TEC data, using spatio-temporal correlations to determine connections between grid locations. Connected locations then exhibit a shared response to some ionospheric perturbing activity.

Data treatment consists of the following steps and is illustrated in Figure 1:

1. 5-minute TEC data are accumulated over one hour and converted to altitude adjusted corrected geomagnetic

(AACGM) coordinates (Figure 1a shows data already converted to AACGM coordinates);
2. AACGM data are re-binned onto an equal area grid;
3. Relative perturbation data are computed (Equation 2);
4. Each 1-hour distribution of relative perturbations is binned by IMF clock angle;
5. A correlation matrix is computed for each clock angle bin;
6. An adjacency matrix is computed from each correlation matrix (Equation 3); and
7. The network is constructed, which can be visualized by the network nodes (AACGM grid points) and edges (connections).

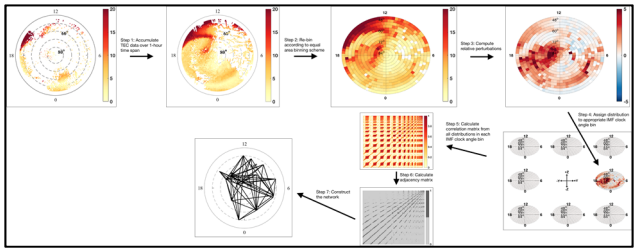

Figure 1. An illustration of the network construction process.

The equal area gridding scheme uses a constant 3° magnetic latitude (MLAT) resolution and variable magnetic local time (MLT) resolution (0.267 hours at 45° MLAT to 4 hours at the pole), yielding a total of 672 grid points. Relative perturbations (Step 4 in Figure 1) are calculated from:

$$TEC_{\text{rel. pert}} = \frac{TEC - TEC_b}{TEC_b} \qquad (2),$$

where TEC_b is a background TEC value, chosen to be the monthly median at each grid point for each IMF clock angle bin. A monthly background was chosen to remove long-term variations so that geomagnetic activity is the predominant driver of network connections in the perturbation data.

3. Complex networks analysis applied to high-latitude TEC data

The goal of network analysis in the geophysical context is to define physically meaningful connections between grid points from a spatio-temporal data set and to determine the physical implications of the resulting network topology. Here we provide further details about steps 5-7 from above (see Figure 1). To construct the network out of the correlation matrices we treat each grid point as a node of the network and correlations between nodes above a given threshold as a connection. Therefore, at its core, network analysis consists of quantifying when pairs of grid point locations are either connected or not connected. The threshold level used to define a connection is a central design consideration in network analysis. In this work we take the threshold to be those correlations in the top tenth percentile, which provides a measure of statistical significance, and identifies the most important connections. Each IMF clock angle bin is treated separately such that the threshold level varied across bins. Once this level was determined TEC networks could be formed and network parameters robustly identified to describe the network topology.

Given the correlation threshold level, we obtain the adjacency matrix, A:

$$A_{i,j} = \mathcal{H}\left[|c_{i,j}| - c_{\mathrm{T}}\right] \qquad (3),$$

where cij is the correlation coefficient between grid point i and grid point j, c_{T} is the threshold correlation level, and $\mathcal{H}$ is the Heaviside step function. From A we compute the degree centrality, which is the simplest measure of the spatio-temporal patterns in the network, to characterize the TEC networks:

$$C_{D_j} = \frac{\sum_{i=1, i\neq j}^{n} A_{i,j}}{(n-1)} \qquad (4),$$

where C_{D_j} is the degree centrality, or the normalized total number of connections, for grid point j, and n is the total number of grid points (672 for the equal area grid). Degree centrality indicates locations that have a higher influence on network function [6], providing a measure of grid points where critical processes responsible for TEC variation take place.

4. Results and discussion

4.1 High-latitude TEC dependence on IMF clock angle

We define eight IMF clock angle bins, centered at 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Figure 2 shows a histogram of the clock angles for January 2016.

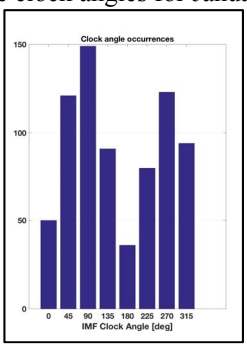

Figure 2. Histogram of IMF clock angle data for January 2016.

Despite fewer occurrences of purely northward and southward IMF (0° and 180°, respectively), which is typical of clock angle occurrence frequencies [9] data in the remaining directions are relatively evenly distributed.

Figure 3 shows the northern (3a) and southern (3b) hemisphere median TEC distributions over January 2016. The influence of solar extra ultraviolet (EUV) irradiance on the median TEC distributions in both hemispheres is clear, and the seasonal difference between northern and southern hemispheres is evident from the MLT extent of dayside TEC levels. However, IMF clock angle dependencies are apparent as well. TEC enhancements during northward IMF conditions (top row in Figures 3a-b) are confined to higher MLATs (~70-80°) and tend toward dayside MLTs with respect to southward conditions (bottom row in

Figures 3a-b) where enhancements occur in the statistical auroral zone (~60-75º). Figure 4 shows the average TEC relative perturbation distributions to more clearly illustrate these trends.

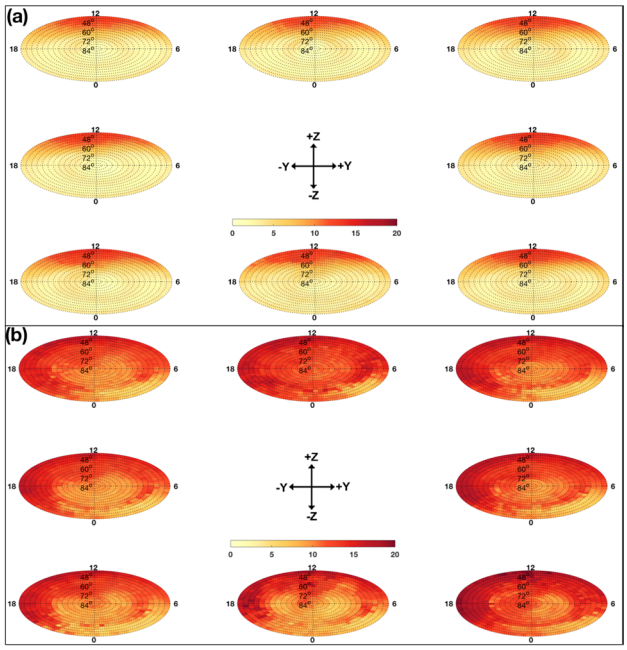

Figure 3. Median distributions of (a) northern and (b) southern hemisphere TEC data in January 2016, sorted by IMF clock angle. The data plotted are TEC units (TECU), 1 TECU = 10^{16} el/m^2. Data are shown in MLAT-MLT AACGM coordinates. The low-latitude limit on each polar plot is 45º. The magnetic latitude resolution is 3º, and the magnetic local time resolution is variable to yield equal area bins (0.267 hours at 45º MLAT to 4 hours at the pole). Noon local time is at the top of each polar plot with dawn to the right. Clock angles increase in 45º increments in the clockwise direction.

In Figure 4, the differences between northward and southward IMF are more distinct. Hemispheric differences are also apparent, notably in the MLAT-MLT locations of auroral region TEC enhancements. The northern hemisphere shows increased variability across clock angle bins, which might indicate different characteristics of accelerated particle precipitation between summer and winter hemispheres [10]. Further investigation of these patterns is the subject of ongoing research.

4.2 Spatio-temporal structuring of high-latitude TEC: Degree centrality

Figure 5 depicts the degree centrality distributions for the northern (4a) and southern (4b) hemispheres. The spatial regions of degree distribution are emphasized by dividing grid points into different bins given by $0 < C_{D_j} \leq 0.1$; $0.1 < C_{D_j} \leq 0.2$; $0.2 < C_{D_j} \leq 0.3$; $0.3 < C_{D_j}$. The spatial patterns of degree centrality are significantly different between the two hemispheres and clock angle bins. The distinction between dayside and nightside MLT sectors is primary among these differences. These two regimes represent the primary level of network connection, indicating that their responses are distinct and coordinated.

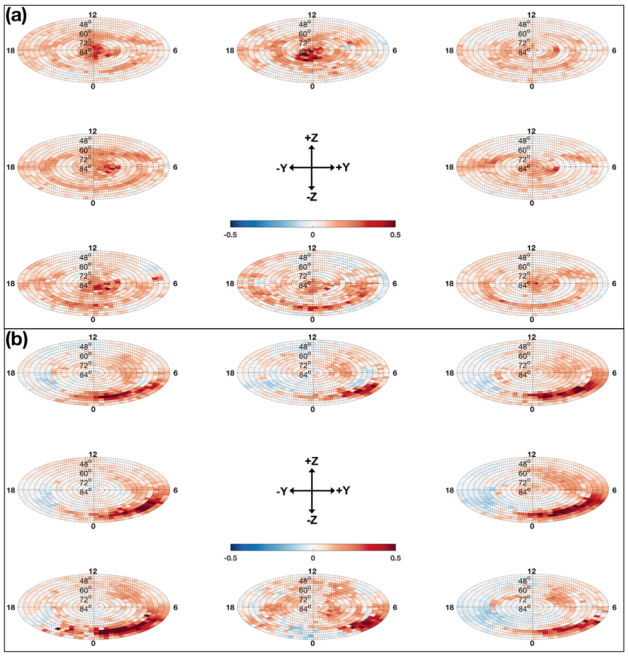

Figure 4. Distributions of average relative perturbations of (a) northern and (b) southern hemisphere TEC data in January 2016, following the format of Figure 2. A diverging color scheme is used to characterize these data.

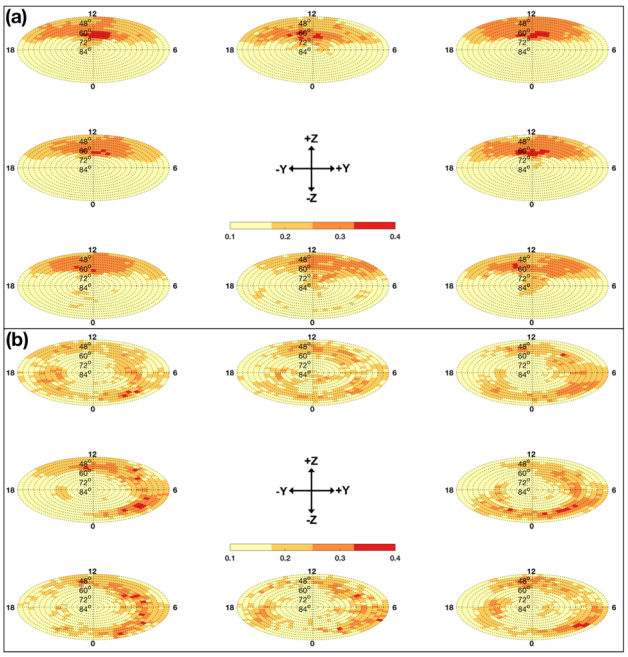

Figure 5. (a) Northern and (b) southern hemisphere degree centrality. Degree centrality data are organized into a discrete grid to emphasize spatial regions associated with degree distribution (see text).

In the northern hemisphere under northward IMF (top row, Figure 5a) the cusp in the noon MLT sector between 60-75º MLAT is a prominent feature. We note that these results are not able to discern that this is particularly the cusp and these features may also be representative of ionospheric footpoints of magnetospheric boundary layers (MBL) [11]. During northward IMF magnetic reconnection predominantly occurs at the high-latitude dayside

magnetopause and causes enhanced magnetosphere-ionosphere interactions at the ionospheric footpoints (e.g. the cusp and MBL). During southward IMF reconnection takes place in the magnetotail and maps to the nightside auroral region ionosphere. Figures 5a-b correspondingly show enhanced degree centrality at nighttime MLTs and in the statistical auroral oval for both hemispheres under southward IMF.

The southern hemisphere shows reduced variability across IMF clock angle bins, in general. This is perhaps due to the suppression of certain effects of geomagnetic activity in the summer hemisphere due to increased solar influence [10].

Greater degree centrality implies greater importance to the functioning of the network. Therefore, these results suggest that the winter, rather than summer, hemisphere cusp is more important in terms of dynamical processes affecting TEC. These relationships suggest the potential predictive insight revealed by network analyses.

4.3 Discussion of future directions

We identify a several areas of important future research:

1. Large values of degree centrality emerge due to: 1) a higher volume of local connections or 2) longer spatial connections. Therefore, future work should analyze the local clustering coefficient and median length of connections [*Malik et al.*, 2012] to distinguish these two situations.
2. In order to understand the seasonal effects in these TEC networks additional months should be studied in the same manner.
3. Relationships between TEC variations and indicators of geomagnetic activity, such as field-aligned current and particle precipitation, can be studied with the network analysis outlined here, using multivariate correlations. Such network analyses will be important to the future use of TEC as a proxy for geomagnetic activity.

5. Conclusions

We conducted the first network analysis of IMF clock angle dependent, high-latitude, hemispheric-specific total electron content (TEC) data obtained from Global Positioning System (GPS) signals. Our primary conclusions are:

1. Dayside and nightside are the strongest network components, with the responses of these regimes distinct and coordinated.
2. Distinct areas emerge above the dayside and nightside systems, including the cusp/magnetospheric boundary layers, polar cap, and auroral zone, which react to geophysical processes together; One of the most strongly connected regions, when it exists distinctly, is the polar cusp/ionospheric footpoint of magnetospheric boundary layers.
3. Network activity indicated closer connections between MLATs poleward of ~70^{o} during northward IMF and within the statistical auroral zone region during southward IMF.

We have identified network methodology as a promising tool for the analysis of TEC data, and suggest that more systematic analyses should be conducted.

6. Acknowledgements

This research was supported by the NASA Living With a Star Jack Eddy Postdoctoral Fellowship Program, administered by the University Corporation for Atmospheric Research. Portions of this research were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.We gratefully acknowledge the Massachusetts Institute of Technology (MIT) Atmospheric Sciences Group for providing the TEC data (http://madrigal3.haystack.mit.edu/) and usage support (funded by NSF grants AGS-1242204 and AGS-1025467).

7. References

[1] Coster, A., and A. Komjathy (2008), Space weather and the global positioning system, *Space Weather*, *6*(6), S06D04, doi:10.1029/2008SW000400.

[2] Mannucci, A. J., B. D. Wilson, D. N. Yuan, C. H. Ho, U. J. Lindqwister, and T. F. Runge (1998), A global mapping technique for GPS-derived ionospheric total electron content measurements, *Radio Science*, *33*(3), 565–582, doi:10.1029/97RS02707.

[3] Shim, J. S. (2000), Analysis of total electron content (TEC) variations in the low- and middle-latitude ionosphere, dissertation, Utah State University.

[4] Boccaletti, S., V. Latora, Y. Moreno, M. Chavez, and D.U. Hwang (2006), Complex networks: Structure and dynamics, *Physics Reports*, *424*(45), 175–308, doi: http://dx.doi.org/10.1016/j.physrep.2005.10.009.

[5] Donges, J. F., Y. Zou, N. Marwan, and J. Kurths (2009), The backbone of the climate network, *EPL*, *87*(4), 48,007.

[6] Malik, N., B. Bookhagen, N. Marwan, and J. Kurths (2012), Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks", *Climate Dynamics*, *39*(3), 971–987, doi:10.1007/s00382-011-1156-4.

[7] Dods, J., S. C. Chapman, and J. W. Gjerloev (2015), Network analysis of geomagnetic substorms using the SuperMAG database of ground-based magnetometer stations, *JGR: Space Physics*, *120*(9), 2015JA021456, doi: 10.1002/2015JA021456.

[8] Newman, M. E. J. (2003), The structure and function of complex networks, *SIAM Review*, *45*(2), 167–256, doi:10.1137/S003614450342480.

[9] Fear, R. C., M. Palmroth, and S. E. Milan (2012), Seasonal and clock angle control of the location of flux transfer event signatures at the magnetopause, *Journal of Geophysical Research: Space Physics*, *117*(A4), A04202, doi:10.1029/2011JA017235.

[10] Newell, P. T., C.I. Meng, and K. M. Lyons (1996), Suppression of discrete aurorae by sunlight, *Nature*, *381*(6585), 766–767, doi:10.1038/381766a0.

[11] Vasyliunas, V. M. (1979), Interaction between the magnetospheric boundary layers and the ionosphere, in *Proceedings of the Magnetospheric Boundary Layers Conference*, ESA Special Publication, vol. 148, edited by B. Battrick, J. Mort, G. Haerendel, and J. Ortner.