

Space Weather

New approaches towards quantifying and predicting extreme space weather events

J. Flynn, K. Foster, G. Watkins, and Y. Zhu

University of Warwick

Abstract. In this project we use data from a global network of ground-based magnetometers to investigate the influence of extreme space weather events on the dynamics of the Earth's surface-level magnetic field, and present new methods that contribute to the broader work being done to develop a predictive model for such events.

We build on the network-based approach described by Dods et al. (2015) to develop a methodology that incorporates temporal offsets when determining connections between pairs of magnetometer stations. Using the generated networks we investigate the network's dimensionless parameters and find that both average degree and average clustering coefficient are lower during a high-activity day compared to a quiet day. Generally the stations experience the effects of solar activity at around the same time, although when the levels of magnetic activity are greater there is a wider range of times over which these effects are felt.

We also apply methods adapted from the theory of collective motion. We show that the shape of the radial distribution of magnetometer readings roughly follows the amplitude of magnetometer readings and concentrates on higher values of r during a magnetic storm. Additionally, we show that the order parameter of magnetometer readings is smaller during a magnetic substorm than on a normal day.

This project outlines some completely new approaches and represents a starting point from which these methods could be applied to complement the current state-of-the-art techniques for the prediction and classification of extreme space weather events.

1 Introduction

Space weather can be defined as 'the conditions on the sun, in the solar wind, and within Earth's magnetosphere, ionosphere and thermosphere that can influence the performance and reliability of space-borne and ground-based tech-

nological systems and can endanger human life or health'. (Hanslmeier, 2010)

The importance of this field is demonstrated by the fact the U.S. has a National Space Weather Plan, which states that extreme space weather events can degrade or damage critical infrastructure, which may result in direct or cascading failures across key services such as electric power, communications, water supply, healthcare, and transportation. As such, an improved understanding of heliophysics - the field concerned with the Sun's effects on the Solar System - and the ability to anticipate and classify extreme space weather events would be enormously valuable, allowing appropriate mitigating strategies to be implemented in an effort to avoid or reduce the potentially disastrous impacts.

1.1 Physics Background

Without interference from the Sun, the Earth's magnetic field (EMF) would resemble that of a dipole magnet. However the Sun's influence causes disturbances in the EMF. The solar wind compresses the field on the Earth's day-side (i.e. the side facing the sun) and causes an elongated tail on the night-side (as shown in Figure 1) (Basavaiah, 2012).

The solar wind is composed of a stream of charged particles (plasma) in which the Sun's magnetic field is embedded; it is the medium through which disturbances from the Sun propagate (Schwenn, 2006). There are three main classes of solar event whose effects are carried by the solar wind and can be felt on or near Earth:

- **Solar flares**

Solar flares are brief, intense solar storms associated with sudden bursts of electromagnetic radiation. The radiation travels at the speed of light, meaning it takes just minutes Earth.

After 1859 (when Richard Carrington observed a solar flare that is considered to be among the most powerful ever seen, just 18 hours before the Earth experienced a huge geomagnetic storm), solar flares were believed

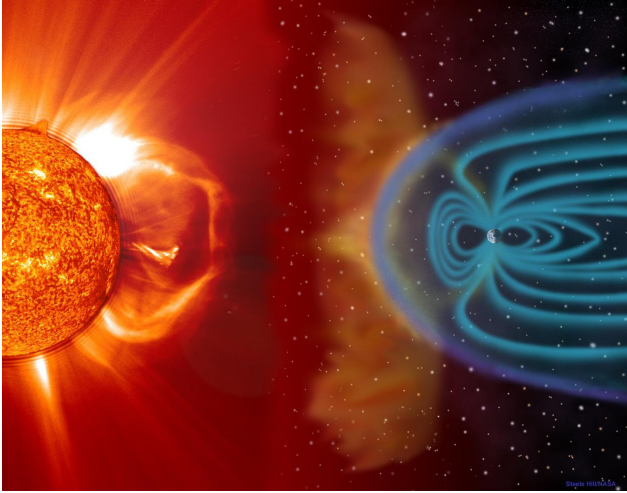


Fig. 1. Artist's impression of interaction between the solar wind and Earth's magnetic field, (NASA, 2009)

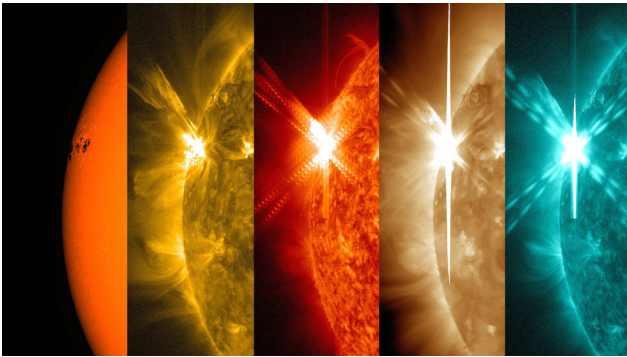


Fig. 2. Solar flare from 05/05/2015 photographed in different wavelengths (NASA, 2017a)

to play a significant role in the Sun's influence on the EMF. However, at the time Carrington warned against drawing such a conclusion (famously stating that "one swallow does not make a summer"), and in recent years the prevailing view that flares and geomagnetic activity have a cause-and-effect relationship has been largely discredited (Schwenn, 2006).

Instead, strong solar flares often occur around the same time as coronal mass ejections which are considered to be much more likely to cause (or at least be correlated with) geomagnetic activity.

– Coronal mass ejections

The Sun's corona is the outermost part of the Sun's atmosphere (which can be seen during a solar eclipse when the bright light of the Sun's surface is obscured, see Figure 3). A coronal mass ejection (CME) is defined as being an 'observable change in coronal struc-



Fig. 3. The Sun's corona visible during a solar eclipse (NASA, 2017b)

ture' (Hundhausen et al., 1984) and is accompanied by the ejection of gigantic clouds from the corona of ionized gas (Schwenn, 2006). The embedded magnetic field dwarfs the 'normal' solar wind magnetic field and thus the geoeffects are significantly more pronounced.

The plasma clouds emitted by CMEs can take as little as 15 hours to reach Earth, or as much as several days. Due to the (relatively) lower speeds of CMEs it is possible to register them using deep space satellites and relay advance warning of up to an hour to Earth. However this is still a very short time to take appropriate mitigating measures, and furthermore this does nothing to protect the satellite itself! For this reason, a data-driven predictive tool would still play an important role.

– Solar energetic particles

Solar energetic particle (SEP) events originate from solar flares and/or CMEs ploughing through the solar wind and producing high velocity charged particles (Papaioannou et al. (2016), Garner (2015)). The particles are accelerated to speeds so great they take just minutes to reach Earth (sometimes more than half the speed of light) with energies high enough to penetrate the outer shell of spacecrafts (and therefore pose a risk to the people and equipment on board).

Historically it was thought that solar flares cause *impulsive* SEP events while CMEs cause *gradual* SEP events (Reames, 2013) with the names referring to the timescales of the events (Schwenn, 2006). However recent research suggests that 'most SEP events do not conform to a simple two-class paradigm' (Papaioannou et al., 2016) which further shows how active this area of research is.

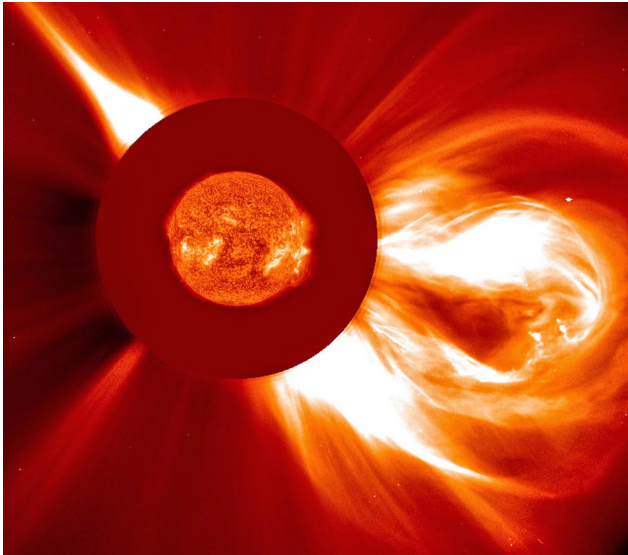


Fig. 4. A coronal mass ejection erupting from the Sun (NASA, 2015)

In general the EMF provides adequate protection from the solar wind by deflecting it around the planet as shown in Figure 1. However when there is an event large enough, the solar wind's interaction with the magnetosphere and ionosphere can cause geomagnetic storms which propagate the effects to the Earth's surface and may cause power outages for prolonged periods.

For example in 1989 a geomagnetic storm caused the entire Hydro-Quebec power grid to collapse in just 90 seconds, leaving the area without power for 9 hours. This event was caused by a CME, but it was almost superseded by an even more powerful event in 2012. This CME - considered the most powerful in more than 150 years (NASA, 2014) - marginally missed Earth but hit the STEREO-A satellite which was in orbit specifically to observe solar activity and was therefore designed to withstand such events. "If it had hit", said Daniel Baker (University of Colorado) in 2014, "we would still be picking up the pieces."

1.2 Potential Impact of Extreme Space Weather

Extreme space weather is of increasing concern to national governments due to the ever-growing societal reliance on technological systems that would be vulnerable to a major solar event. Specifically, power grids and satellites would be most affected, with impacts cascading through industries and systems falling like dominoes (Moldwin, 2008).

There are several recent examples of even relatively tame events causing large-scale disruption. For example, in the autumn of 2003 solar storms caused power outages in Sweden, damaged Japanese satellites and forced aircrafts to change their flightpaths to avoid the elevated levels of radiation orig-

inating at the Sun (Parliamentary Office of Science and Technology, 2010).

The most violent storm 'on record' is considered to be that observed by Richard Carrington in 1859. Although it predated much of today's modern monitoring and recording equipment, energetic particles leave a record in nitrates in ice cores. So arctic ice holds the signature of historic space weather events; examining these historical records reveals the Carrington event to have been the largest in 500 years (NASA, 2008).

An investigation by Lloyd's in 2013 attempted to predict the impact of such an event, if it were to strike present-day North America. They concluded that there is 'potential for large-scale, long-term economic and societal chaos' (Lloyds, 2013); it was later revealed that this almost happened a year earlier as a huge CME only just missed the Earth (NASA, 2014). While the lack of actual evidence has led to some dispute as to the level of risk posed by extreme space weather events, figures as high as \$2 trillion (in the first year alone) have been proposed as estimates on the cost of a Carrington-style event.

Despite the infrequency of the most extreme events, and the lack of consensus on the exact extent of the impact, governments view space weather as a threat that cannot be ignored. The UK added space weather to the national risk register in 2012, indicating how damaging they believe an extreme event could be (Office, 2012), and the US published their National Space Weather Strategy in 2015. Furthermore as recently as June 2019 the UK, US and Europe announced they would be collaborating on a new space weather forecasting system (in the form of a plasma analyser which will be sent into space to monitor the solar wind), supported by £7m funding from the European Space Agency (Agency, 2019).

1.3 Current State of the Art

Nearly all modern day space weather forecasting systems make use of satellites such as DISCOVER and SOHO. These satellites record data such as solar wind speed and interplanetary magnetic field strength which are used to assess the solar wind and identify CMEs (which are slow-moving enough to allow the satellites to relay advance warning to Earth) (NOAA, 2019).

Many different methods are applied to use this data for the prediction of extreme space weather events, including the use of neural networks and statistical regression (Hapgood, 2018; Qin and Nishii, 2015). However the current suite of predictive tools still doesn't yield reliable results.

Despite the ongoing work, recent developments and government investment, space weather forecasting remains a major challenge (Schwenn, 2006); the predictions of current state-of-the-art methods are still too unreliable, especially given that the response to a suspected event may involve measures as costly as the temporary shut-down of an entire electricity grid. The challenge lies in the fact that solar activity

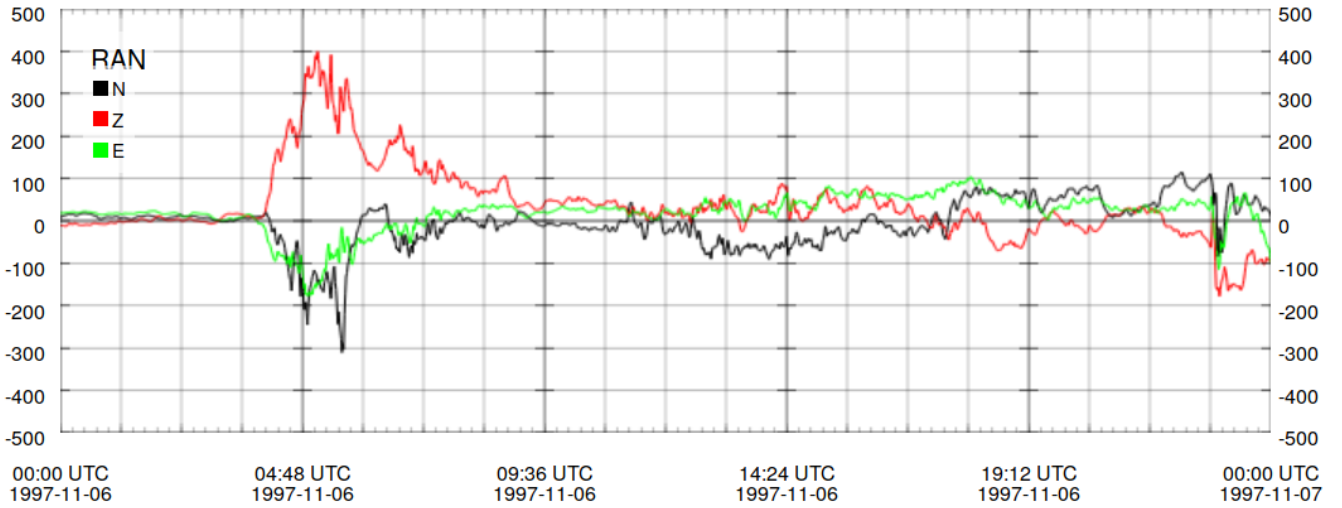


Fig. 5. Example of time series for a single magnetometer station as presented on the SuperMAG website.

appears to occur spontaneously, and the underlying physics is not sufficiently well understood to accurately model the system.

For this reason, many modern efforts to predict and characterise space weather events take a data-focused approach. Currently no unique signatures or precursors that might indicate an imminent event have been identified, but the hope is that with more data and more sophisticated methods, this might become possible.

The use of satellite data for this purpose is well documented in the literature. However the satellite orbits are often outside the Earth's magnetosphere, so while they get a 'clean' view of the solar wind, they do not experience the resulting geoactivity. As such it is important to supplement satellite data with ground-based measurements.

Readings of the Earth's magnetic field strength are taken around the world at magnetometer stations. Historically, conducting research using these data was painstaking and labour-intensive. Analysts who wanted to work with multiple magnetometers faced significant challenges: confusing or even unknown coordinate systems, a multitude of data artifacts and errors, unknown baselines and, given the global locations of the stations and sometimes even political barriers, difficulty even obtaining the data. These problems have resulted in a serious historical under-utilisation of magnetometer data. (Gjerloev, 2009).

2 The Data

The data used in this report are provided by SuperMAG, a worldwide collaboration of organizations and national agencies that together operate more than 300 ground-based magnetometers. SuperMAG aggregates the data, performs some pre-processing (including cleaning - both automatically and

manually - and removing baselines) and makes it available on their website so that 'scientists, teachers, students and the general public have easy access to measurements of the Earth's magnetic field' (Gjerloev, 2012).

When operational each station returns one data point every minute measuring the local magnetic field strength (in nanoteslas). A data point x_i is a vector with three components $x_i = (x_i[N], x_i[E], x_i[Z])$ where

$x_i[N]$ = component of magnetic field in **local magnetic north** direction

$x_i[E]$ = component of magnetic field in **local magnetic east** direction

$x_i[Z]$ = component of magnetic field in **vertically down** direction

Figure 5 gives an example of the data held in the SuperMAG database. The plot (which was generated by the SuperMAG website itself) shows data from a single station (RAN) for each of the three components over a 1 day time period. This particular day (6 November 1997) contains a number of interesting features.

The first phase (until around 04:00, although this is somewhat hard to discern precisely because of the unusual default scales used by SuperMAG plots) exhibits very little magnetic activity. During this period what we observe is a combination of the effects of 'normal' solar activity (i.e. in the absence of an extreme event) and background magnetic activity. Note that during pre-processing SuperMAG subtracts a baseline from the data (including the Earth's own magnetic field) but there is no consensus on how to remove background 'noise' entirely; indeed there are whole papers written on the subject (see for example Janzhura and Troshichev (2008), Takahashi and Anderson (1992) and Menvielle et al. (1995)). Gjerloev

writes that ‘in the SuperMAG data processing technique the baseline determination is unquestionably the most difficult and controversial problem’ (Gjerloev, 2012).

The second phase lasts from 04:00 until around 06:00 during which the Earth’s magnetic field is affected by extreme solar activity. It is unclear from the data alone what kind of event would have caused the large fluctuations observed; this is an area that would merit further work, and for which this project could serve as a good starting point.

From 06:00 the magnetic field strength reduces, but doesn’t completely revert to the same pre-event levels. This shows that there is residual activity following a solar event. This could be due to continued disturbances caused by lower (but still elevated relative to normal) levels of solar activity, or simply that it takes time for the Earth’s magnetic field to settle (analogous to a guitar string that continues to vibrate for some time after it has been plucked).

We can generate plots like this for every magnetometer on every day it was operation. As such the SuperMAG database contains a vast quantity of data. With more than 300 magnetometers, many of which have been operational for decades (and a few for over a century), there is a combined total of more than 15,000 years of data with 1 minute time resolution.

3 Network-Based Approach

One approach we have taken, building on the work of Dods et al. (2015), is to construct a network in which the vertices represent magnetometer stations, and two stations are connected if their respective time series are highly correlated. Rather than looking at long-term timescales (on the order of years) it is more interesting to compare shorter time windows (of a few hours) and thereby generate a dynamic network that evolves over time. This enables us to isolate solar events for comparison against days with normal magnetic activity.

3.1 Dynamic Network Construction and Analysis

Suppose we wish to analyse the dynamics of a set of stations $S = \{s_1, s_2, \dots, s_{|S|}\}$ over a certain n -minute time period $T = (t_1, t_2, \dots, t_n)$. Define also $\mathbb{T} = \{t_1, t_2, \dots, t_n\}$ to be the set of times in our time window.

For each station pair $(s_i, s_j) \in S^2$ ($i \neq j$) we need to compute the pairwise canonical correlations for each time window w_t of length ℓ , $w_t = (w_t[1], \dots, w_t[\ell]) \in \mathbb{T}^\ell$ completely contained within T (i.e. $w_t[1] \geq t_1$ and $w_t[\ell] \leq t_n$). This yields $n - \ell + 1$ time windows, $\{w_{t_1}, w_{t_2}, \dots, w_{t_{n-\ell+1}}\}$. Note that by convention we label a time window by its starting time in UTC.

We have chosen to focus on time windows of length 128 minutes (guided by the previous work of Dods et al. (2015) and discussions with Dr. Chapman), so for an N -minute time period T we will construct $N - 127$ time windows (and will

therefore compute $N - 127$ correlation coefficients for each station pair).

The existence of an edge for a station pair (s_i, s_j) in the network for a time window w_t depends on the magnitude of the canonical correlation between the stations’ time series for time window w_t (see Section 3.1.1 for an explanation of canonical correlation analysis). This is the technique used by (Dods et al., 2015), and we extend it by introducing time offsets (as outlined in Section 3.1.2). Once we have the canonical correlation between stations s_i and s_j for w_t , we then determine if this correlation is sufficiently high to connect the stations in the network; the process of calculating thresholds is discussed in Section 3.1.3. Finally we characterize the network in terms of certain parameters (as described in Section 3.1.4) with the goal of finding parameters that can be used to classify and predict extreme space weather events.

3.1.1 Canonical Correlation Analysis

Canonical correlation analysis provides a method for finding correlations between vectors of random variables, or time series in which the data points are themselves vectors. The coefficient that is ultimately returned is the correlation between two canonical variates, which are projections of the vectors into latent space in such a way that the resultant correlation is maximised.

The method, as applied to our magnetometer data, works as follows.

Start with two time series \mathbf{x} and \mathbf{y} for stations s_i and s_j in a time window w_t of length ℓ . Without loss of generality, set $t = 1$ for simplicity.

$$\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\ell)$$

$$\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_\ell)$$

Note that each of the \mathbf{x}_i and \mathbf{y}_j are 3-component vectors, i.e. $\mathbf{x}_i = (\mathbf{x}_i[N], \mathbf{x}_i[E], \mathbf{x}_i[Z])$

Define vectors

$$\mathbf{u} = (u_1, u_2, \dots, u_\ell)$$

$$\mathbf{v} = (v_1, v_2, \dots, v_\ell)$$

where

$$u_i = \mathbf{a} \cdot \mathbf{x}_i$$

$$v_j = \mathbf{b} \cdot \mathbf{y}_j$$

So \mathbf{u} and \mathbf{v} are projections of \mathbf{x} and \mathbf{y} into the \mathbf{a} and \mathbf{b} directions respectively.

The idea behind canonical correlation analysis is to find vectors \mathbf{a}_* and \mathbf{b}_* such that

$$\mathbf{a}_*, \mathbf{b}_* = \operatorname{argmax}_{\mathbf{a}, \mathbf{b}} \rho_{\mathbf{u}, \mathbf{v}}$$

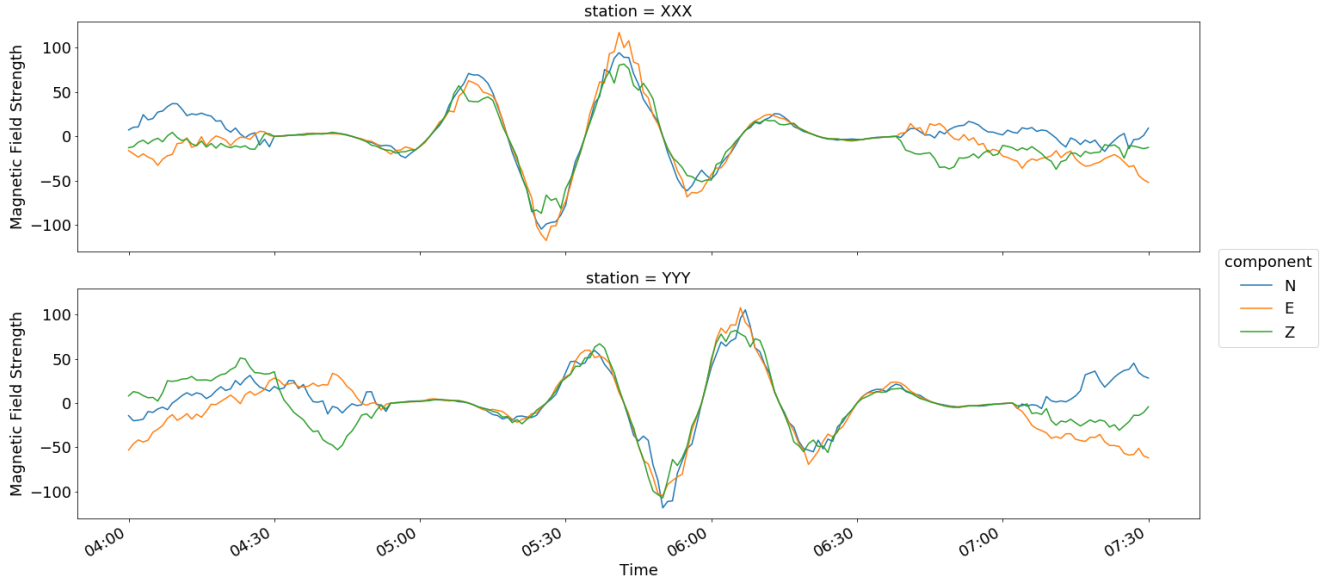


Fig. 6. Model-generated data for two magnetometer stations experiencing a solar event with a temporal offset.

where $\rho_{u,v}$ is the Pearson correlation coefficient between u and v . The maximum value of $\rho_{u,v}$ is called the first canonical correlation coefficient between x and y . In fact we could also find the second and third canonical coefficients (because each of the x_i and y_j has three components) but we will not require them for the purposes of constructing the network.

3.1.2 Time Offsets

When a solar event takes place we would expect to observe high correlations between stations' time series because they are experiencing the same event which dominates normal daily magnetic activity. This is particularly true for station pairs that are geographically proximate as it is even more likely that the stations are experiencing the solar event in the same manner. However even if the waves are coherent, the spatial differences in the magnetometer locations often cause a phase shift in the wave packets that may result in a lower canonical correlation. As such, for each station pair we need to investigate a range of time lags to find which lag value yields the maximum correlation. This process will inevitably give rise to some spurious maximal correlations, but these specious values will drop out when the network is evolved over time.

The following example demonstrates this point using artificially generated data designed to have the same wave packet properties as the real-world data. More information on how this data was generated is covered in Section 4.

Figure 6 shows two artificially generated time series with sine-like waves in the middle representing a solar event. Note

that without loss of generality the N, E and Z components are all based on the same underlying waveforms, with added noise. Since these waveforms are coherent we would expect to find a high correlation between stations XXX and YYY during this time period.

Station XXX starts experiencing magnetic activity caused by the solar event at 05:00, and station YYY starts to experience the same event 24 minutes later. This offset is greater than what we typically observe in real world data, but the lag has been exaggerated to demonstrate the principle. The solar event lasts for 128 minutes.

We focus on a time window between 05:24 and 07:08, during which both stations are experiencing the solar event. First we calculate the correlation in the case where the time series for both stations come from this time window. Then we introduce an offset of 24 minutes for station YYY, effectively removing the phase shift of the underlying sine waves, and calculate the new correlation.

Let

ρ_τ = the correlation between stations XXX and YYY using time window [05:24, 07:08] for XXX and an offset of τ minutes for YYY (i.e. [05:24+ τ , 07:08+ τ]).

Then we find:

$$\begin{aligned}\rho_0 &= 0.3833, \\ \rho_{24} &= 0.9922.\end{aligned}$$

This demonstrates the importance of including lag in our models. For each station pair $\{s_i, s_j\}$ and time window w_t ,

we fix the time series for station s_i and shift the time series for station s_j according to a given range of lags to find the maximum correlation between the two stations. The choice of which station to fix is relatively unimportant; by convention we choose the one with the higher latitude if that information is known. Based on historical data we have chosen the range of lags to be $\{-10, -9, \dots, 0, \dots, 9, 10\}$ minutes. From here on, when we discuss the correlation between two stations s_i and s_j for time window w_t we will be referring to the maximum canonical correlation coefficient obtained using this range of lags. We define $\tau_{i,j}$ as the lag at which maximum correlation is achieved for the station pair s_i, s_j , and we will refer to it as the optimal lag for $\{s_i, s_j\}$ for the time window w_t . Then the correlation coefficient will be denoted by $\rho(s_i, s_j, w_t; \tau_{i,j})$.

3.1.3 Thresholding and Adjacency

As discussed, in order to construct a network for a time window w_t we need to determine whether a station pair $\{s_i, s_j\}$ ought to be connected. Suppose that the maximum correlation for this station pair and time window is achieved using lag $\tau_{i,j}$, and is given by $\rho(s_i, s_j, w_t; \tau_{i,j})$. For this time window we define a threshold value for station pair $\{s_i, s_j\}$

$$C(s_i, s_j; \tau_{i,j}) = \rho(s_i, s_j, T; \tau_{i,j})$$

So having determined the optimal lag $\tau_{i,j}$ for time window w_t , we use this lag to compute the canonical correlation between s_i and s_j across the whole time period T , and use this as our threshold. Then if the correlation between stations s_i and s_j for time window w_t is greater than this threshold, we connect the vertices for s_i and s_j with an edge in the network for w_t .

Hence the adjacency matrix A is formed such that

$$A_{i,j} = \begin{cases} 1 & \text{if } \rho(s_i, s_j, w_t; \tau_{i,j}) \geq C(s_i, s_j; \tau_{i,j}), \\ 0 & \text{otherwise.} \end{cases}$$

Note that this generates an undirected network. We can also generate directed networks in which the edge direction indicates positive lag (i.e. we have a directed edge from s_i to s_j if their correlation is above the threshold, and station s_i experiences the event before s_j). In this case, edge directions show how the event's effects are propagating.

$$\hat{A}_{i,j} = \begin{cases} 1 & \text{if } \rho(s_i, s_j, w_t; \tau_{i,j}) \geq C(s_i, s_j; \tau_{i,j}) \\ & \text{and } \tau_{i,j} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

3.1.4 Dimensionless Network Parameters

In order to quantify a space weather event, we calculate network parameters for the generated networks as they evolve over time. The goal is to find parameters that enable us

to characterise and perhaps predict different space weather events.

First we investigate the degrees of the stations. Degrees can be averaged over either time (for a fixed station) or stations (at a fixed time). When averaged over time (for fixed station s_i), we find the typical propensity for s_i to be connected in the network. On the other hand, averaging over stations (at a fixed time) shows how connected the network is (and therefore how correlated the time series for the stations are). Since we expect the connectivity of the network to increase during a solar event, we expect both the average degree (and consequently the edge count) to increase during periods of high activity (Dods et al., 2015). Naturally, we can also evolve this measure of connectivity over time to see how the connectedness of the network changes with respect to time.

The other type of network parameter that we calculate is the local clustering coefficient for one station over time. This parameter quantifies the connectivity of a station's neighbours, and as a result reveals highly connected sub-networks. We expect clustering among geographically proximate magnetometer stations because the magnetic field will be very similar for those stations, so the time series for those stations are likely to be correlated (and hence connected in our network). However, it is unclear a priori whether or not we should expect clustering across greater geospatial distances. Clustering in such a manner with zero lag would indicate either a very widespread and uniform magnetic disturbance or a potentially specious correlation. Including nonzero lags would allow the clustering coefficient to capture a general sweeping motion of the solar event's effects around the globe, an effect that we expect to observe in the real-world data.

3.2 Network Results

We constructed and analysed networks for a 'quiet day' with no space weather events and a day with a high levels of geomagnetic activity (Dods et al., 2015) in order to make comparisons. The days were chosen within three months of each other to provide temporal consistency. Moreover, these periods are in the winter months in the northern hemisphere so that sunlight is limited on the dayside (Dods et al., 2015). The quiet 'day' was chosen as 00:00 2 February 1998 to 13:59 3 February 1998 because no substorms occurred on that day, and it also shows relatively stable magnetometer readings. The high-activity day was chosen as 00:00 5 November 1997 to 23:59 5 November, during which we see the signature bulge of a substorm. Furthermore we only consider the magnetometer stations in a single chain (so they are positioned on very similar longitudes). This reduces the necessity of further complexities with the temporal lag since the magnetic disturbance travels in a single direction as it moves around the Earth.

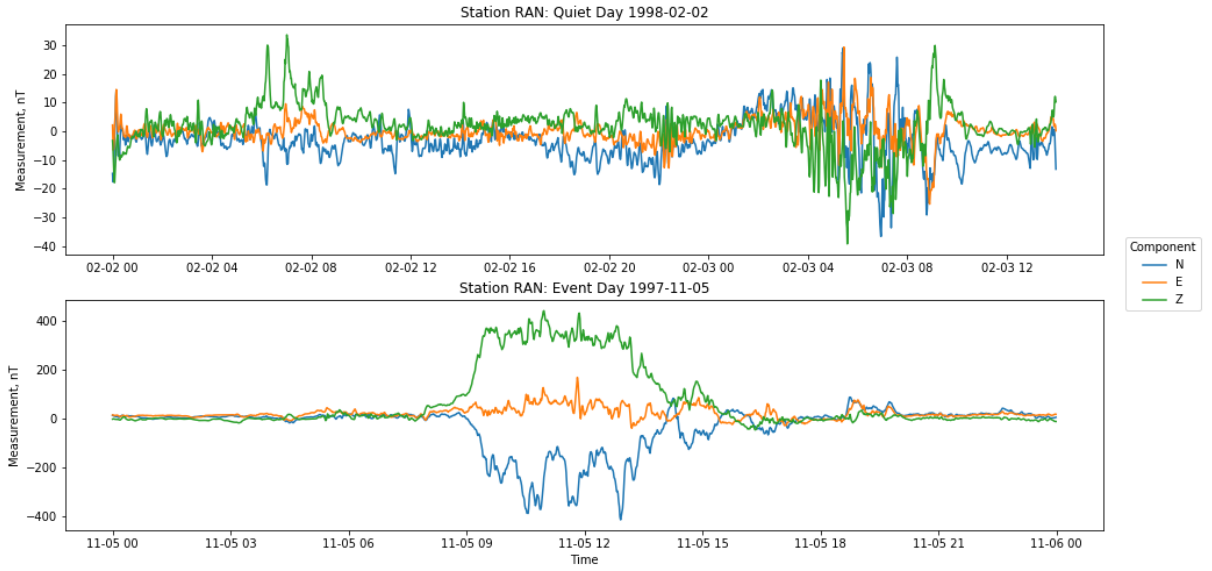


Fig. 7. Comparison of raw data using one magnetometer station, RAN. The top time series is from the quiet day, 2 February 1998, and the bottom time series is from the event day, 5 November 1997. Note the difference in scales on the y axis showing the different levels of activity.

The plots of the raw time series data for one magnetometer station on each day are shown in Figure 7. Note the difference in scale on the vertical axes: the quiet day time series has low-strength readings, which are effectively just noise; on the other hand the high-activity day has a similar pattern (with low magnitude noise) until the solar event hits around 09:00 and creates a significant spike in the magnetometer's readings.

3.2.1 Dimensionless Network Parameters

Our first step in differentiating between a quiet day and an event day is to investigate the differences in the network parameters discussed in Section 3.1.4. Table 1 shows the normalised time-averaged degree for each station and the normalised overall network average degree. The average degrees for the event day are much lower than that of the quiet day. This behaviour is expected because the magnetometers largely experience a similar magnetic field on the quiet day, leading to a high correlation. In comparison the stations record measurements that change rapidly during a substorm, and this magnetic activity leads to more sporadic connections.

In addition to the quiet day being more connected than the event day, we also observe that there is more clustering in the quiet day network than in that of the event day. This is due to the high connectivity in the quiet day network relative to the event day network. While we may expect small clusters of nearby stations to form during a magnetic substorm, spo-

radic connections between stations decrease the network average clustering coefficient because the strong magnetic activity that creates the clusters is so localised. These results are summarised in Table 2.

3.2.2 Correlation and Time Offsets

Now that we have the quantifiable network parameters, we can take a more qualitative approach to analysing the differences in network structure between a quiet day and an event day. When constructing each network, we consider time offsets- or time lag- as previously discussed in Section 3.1.2. We present two sample heat maps in Figure 8 that show which value of lag yields the maximum correlation between a pair of magnetometer stations at a given time. The correlation heat map for the quiet day is shown for the range 11:00-17:00, a six-hour time period with very minimal activity. The correlation heat map for the event day is shown for the range 09:00-15:00, a six-hour time period in which the solar event impacts the magnetometer readings.

The correlation heat map for the quiet day shows that the lag which produces the highest correlation is consistently around zero. This behaviour is what we expect from a geographically close station pair on a quiet day because there is no significant disturbance in the magnetic field that could cause the magnetometers' measurements to lag. Somewhat similarly to the quiet day, we see that the correlation heat map for the event day also suggests that the time lag should remain close to zero during the storm. However not only does

| | TAL | BLC | RAN | EKP | FCC | GIM | PIN | BSL | DLR | Mean Degree |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------------|
| Quiet Day | 0.881 | 0.942 | 0.971 | 0.683 | 0.777 | 0.814 | 0.819 | 0.935 | 0.832 | 0.851 |
| Event Day | 0.687 | 0.585 | 0.524 | 0.369 | 0.355 | 0.426 | 0.533 | 0.693 | 0.711 | 0.542 |

Table 1. The normalised time-averaged degree for each station and the normalised overall average degree for the network.

| | TAL | BLC | RAN | EKP | FCC | GIM | PIN | BSL | DLR | Mean Clustering Coefficient |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----------------------------|
| Quiet Day | 0.848 | 0.852 | 0.847 | 0.948 | 0.881 | 0.864 | 0.848 | 0.849 | 0.865 | 0.867 |
| Event Day | 0.492 | 0.509 | 0.623 | 0.741 | 0.789 | 0.603 | 0.575 | 0.496 | 0.501 | 0.592 |

Table 2. The time-averaged local clustering coefficient for each station and the overall time-averaged network clustering coefficient.

this band of high correlation have a larger variance than that of the quiet day, but the heat map also shows a massive increase in variance at the beginning and end of the solar event. These large increases in variance are indicative of a large shift in the magnetometer measurements. That is, we can observe a significant change in the correlation variance upon entering and exiting a solar event.

Once we have found the maximum correlation for each station pair, we can determine the adjacency for each station pair by comparing its maximum correlation to its threshold value. For each station pair, if its maximum correlation is greater than or equal to its threshold value, these two stations are connected and the adjacency matrix has a value of one for this pair; otherwise the pair is not connected and the adjacency matrix has a value of zero. Figure 9 illustrates this process and shows the difference in network structure between the quiet day and the event day. The sample times for each type were chosen to be within the six-hour time periods so that the network differences would be heightened.

We observe that the magnetometer station BSL appears to dominate the network for the quiet day, with most of its available connections safely above their respective pairwise thresholds. Additionally the station BLC has all of its available connections, yet those correlations are much closer to the threshold, indicating that the station BSL must be measuring a particularly strong magnetic field at the given time. The correlation-threshold difference matrix for the event day shows a much more even distribution of connections only moderately over their respective thresholds. We notice that there appears to be a small cluster of high correlations on the event day, which is unsurprising considering the amount of local magnetic activity occurring during the solar event. The resulting adjacency matrix for the event day reveals a large cluster of weaker connections and an overall more connected network compared to that of the quiet day.

Using the adjacency matrix, we can plot the network on a map. However, simply using the adjacency matrix loses information about how the network is actually constructed. Instead we will plot the network using the correlation-threshold

difference matrix because this method includes the information about the lag between stations. The correlation-threshold difference between two connected stations encodes information about the direction of the connection and the magnitude of the lag. These features can be seen by the direction and color of the arrows in Figure 10, where we show the directed network with lag for the quiet day and the event day. Note that these plots are for the same times used in the correlation-threshold difference matrices from Figure 9. By design, these stations lie along very similar longitudes because it makes their baseline measurements similar. However, we can see overall higher lag in the event day network than in the quiet day network: around ten minutes from the southern United States to northern Canada. This direction of lag indicates that the magnetic disturbance is traveling roughly from south to north, concurrent with the fact that in winter the southern hemisphere of the Earth is closer to the sun than the northern hemisphere. Furthermore the magnitude of the lag is about ten minutes, which is approximately the period of these magnetic disturbances. Given the number of connections with our maximal lag, we would include a wider range of lag values in future studies to better quantify such connections.

As these networks are dynamic and thus changing every minute, a static plot fails to illustrate how the differences in these networks are changing over time. Since we are unable to include animations in a PDF file, we have made the animations for each of these plots available in .gif format in a public Google Drive folder. In addition, we have provided lag network plots with a different set of stations to better demonstrate this effective visualisation. The link to this folder is located in Appendix A.

4 Generating artificial data

For the purposes of checking our functions and modelling the behaviour of the magnetosphere time series, it was useful to generate some artificial data of our own for which we knew the underlying structure.

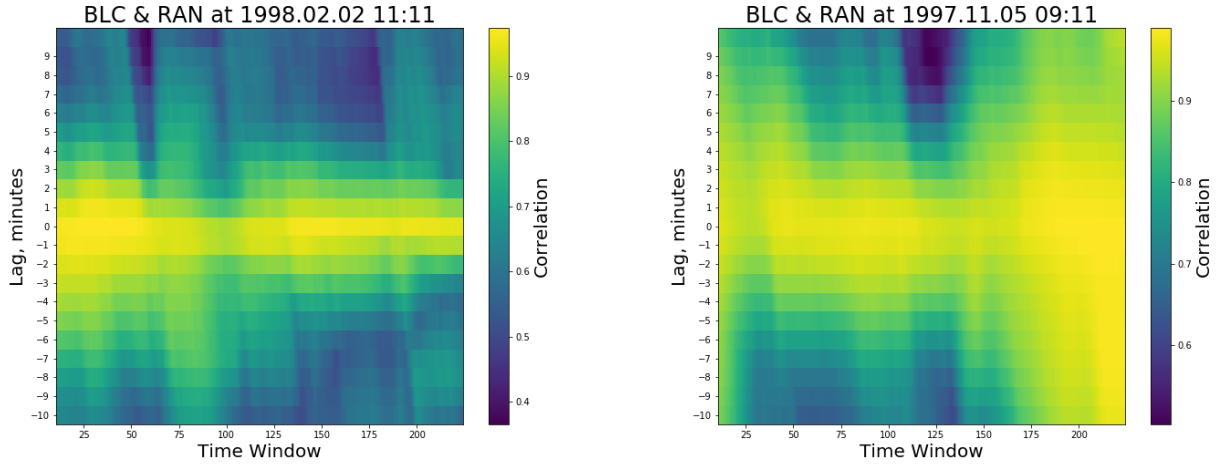


Fig. 8. Comparison of lag correlations over time using one station pair, BLC and RAN. The heat map on the left is for the quiet day, 2 February 1998, and heat map on the right is for the event day, 5 November 1997.

We created a model that can generate time series that resemble the real data but with known parameters. Figure 11 shows two plots for a day during which there was a large event, and relatively little magnetic activity otherwise. The top plot comes from real data while the bottom plot was generated using our model.

The two plots look qualitatively similar thanks to sensible choices of model parameters (notably the amplitude, duration and frequency of the waves). The model generates data in two 'phases': event and non-event.

To generate data resembling an event we need to specify the event start time and its duration. The magnetic field strength at the start and end of the event is set to 0, and in between the data is generated by the product of a Gaussian, a sine wave and random noise.

Outside of an event, the time series follows an Ornstein-Uhlenbeck process. This is a mean-reverting stochastic process, making it a suitable choice for the (seemingly) random fluctuations in the Earth's magnetic field. Note that there are in fact two OU processes in the generated data: one starting at the end of the event and another 'starting' at the beginning of the event, making use of the time-reversibility of the OU process.

5 Collective Motion

We also experimented briefly with the idea of examining the data in the framework of collective motion. Collective motion is perhaps most well known through flocking animals such birds or fish, but actually extends beyond living things into systems where units interact purely physically without communication (Vicsek and Zafeiris, 2012).

A system exhibiting collective motion is assumed to be made of units that are

- rather similar
- moving with a nearly constant absolute velocity and are capable of changing their direction
- interacting within a specific interaction range by changing their direction of motion, in a way involving an effective alignment
- subject to a noise of varying magnitude

5.1 Applying to magnetometer data

We can draw some parallels between such a system and our magnetic field data. Firstly, all of the magnetometers measure the flux of the Earth's magnetic field at their respective locations which we expect to show a rather high degree of similarity.

Secondly, the flux we measure certainly does change direction. The difficulty comes from the fact that the magnitudes of our data are not usually constant, especially when compared across different stations even if the stations are located within close proximity to each other. As alluded to in Section 2, some magnetometers are more sensitive than others and thus yield data with larger amplitude due to physical factors such as the conductivity of the soil on which the magnetometer stations are built. We can either circumvent this restriction by normalising the readings before analysing the data or by relaxing this assumption, and we will examine the effects of normalisation on the parameters.

Thirdly, since the flux is all driven by the behaviour of the Earth's magnetic field, we can reasonably expect some degree of alignment between the readings from different sta-

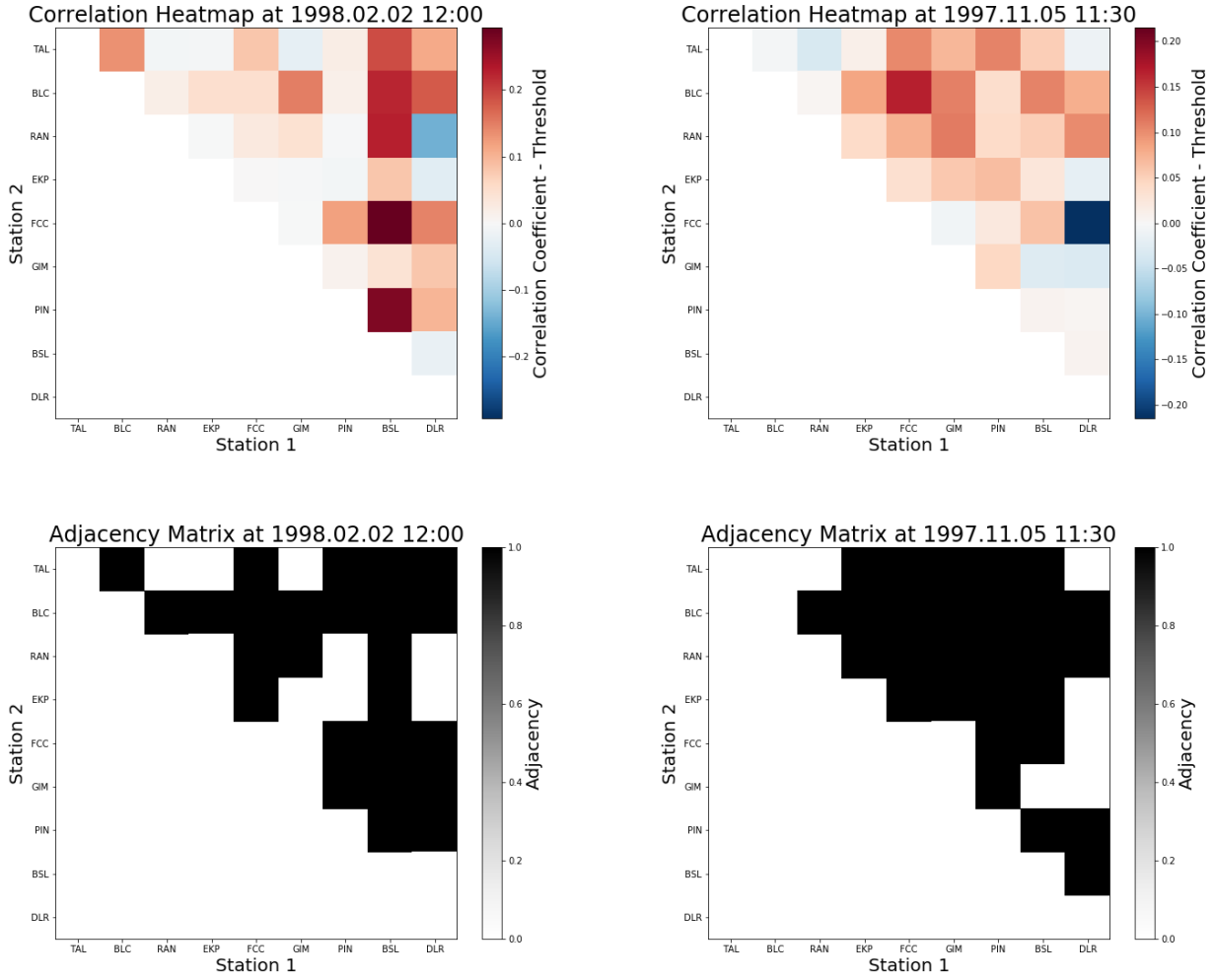


Fig. 9. The correlation-threshold difference matrices for the quiet day and the event day at the labeled time. The top row contains the matrices whose values are the thresholds subtracted from the correlations for each station pair. The bottom row contains the binary adjacency matrices; values greater than or equal to zero become one, and all other values become zero.

tions¹. Since the current understanding on the subject states that extreme solar events affect the Earth's entire magnetic field whose flux is measured by ground-based magnetometers, we will assume that the interaction range is infinite, i.e., we will disregard the physical locations of the stations during the analysis.

Finally, some noise does remain in the data despite efforts made by SuperMAG to subtract the baseline readings due to the fact that there is no consensus on what the correct baseline is, as previously explained.

Before applying the following tools, we detrended the time series by subtracting the mean of a running window on each component in the station readings. The window size is cho-

sen to be around 10 minutes, close to the period of oscillations in our data.

5.2 Order Parameter

The order parameter φ measures the average normalised velocity and is given by

$$\varphi = \frac{1}{Nv_0} \left| \sum_{i=1}^N \mathbf{v}_i \right|$$

where N is the number of units in the system, v_0 is the average absolute velocity, and \mathbf{v}_i is the vector velocity of unit i . As explained in (Vicsek and Zafeiris, 2012), if the motion is disordered, the velocities of the individual units point in random directions and average out to give a small magnitude vector, whereas for ordered motion the velocities all add up

¹Though this contrasts with flocks of birds, for example, where the driver of the behaviour comes from within the system

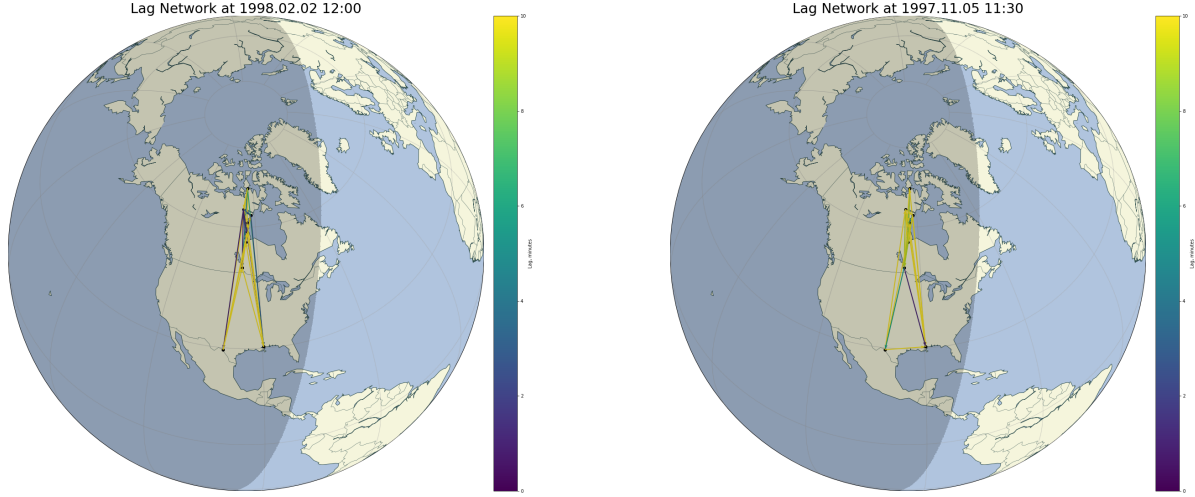


Fig. 10. The plot of the magnetometer stations on a map, including the lag information. The map on the left is for the quiet day, 2 February 1998, and the map on the right is for the event day, 5 November 1997. The direction of a connection is represented by the direction of the arrow, and the lag between two connected stations is represented by the color of the arrow. Note that there are no negative values of lag in this network as a negative lag simply reverses the direction of the arrow.

to a vector of absolute velocity close to Nv_0 . The parameter is normalised and thus $\varphi \in [0, 1]$, where $\varphi = 0$ indicates total disorder and $\varphi = 1$ indicates total order.

We were able to adapt this to the system that we are considering by substituting the required velocity vectors with magnetometer readings. Since the order parameter has in-built normalisation, we did not attempt to normalise the data beforehand.

5.3 Radial Distribution Function

The *radial distribution function*, $g(r)$, describes how the average unit density varies as a function of the distance from each unit. In other words, it is a measure of how clustered together the units are and can be applied to quantify coherence in the magnetometer reading vectors.

By considering the vector readings of magnetic flux from each station at each timestep as displacements from a common origin, we were able to simulate a system of units which are points dictated by the magnetometer readings.

We then used the following formula: for given r and dr ,

$$g(r) = \frac{\sum_i \sum_{j \neq i} (\mathbb{1}(\text{dist}(i, j) \in [r, r + dr]))}{N \cdot 4\pi r^2 dr \cdot (N/V)}$$

$$= \frac{V}{4\pi r^2 N^2 dr} \sum_i \sum_{j \neq i} (\mathbb{1}(\text{dist}(i, j) \in [r, r + dr]))$$

where V is the volume of the system, N is the number of units in the system, i and j are units in the system, and $\mathbb{1}$ is the indicator function.

This method works by iterating through all points in the system and summing the number of points that lie within a spherical shell of inner radius r and outer radius $r + dr$, then normalising over the number of units in the system, volume of the spherical shell, and the number density N/V .

There were several decisions to be made regarding the exact procedure.

Firstly, there was the issue of whether the magnetometer readings should be normalised prior to calculating the radial distribution. Normalisation essentially confines all points onto the unit sphere, thus eliminating their magnitudes from the calculation and shifts the focus to the directions of the readings. We show the effects of normalisation in Figure 12.

Secondly, we are able to choose the range of r that we iterate over, as well as the value of dr . After some brief preliminary experiments, we came to the conclusion that r only needs to be non-negative and as big as the distance between the two furthest points. On the other hand, dr has no obvious optimal value. Fundamentally, we are trying to cover the range of r , $[r_1, r_n]$, with smaller intervals of the form $[r_i, r_i + dr)$. The natural choice is to choose $dr = r_{i+1} - r_i$ in order to avoid overlaps. This isn't strictly the best choice, though, as we can use a wider interval to capture points that

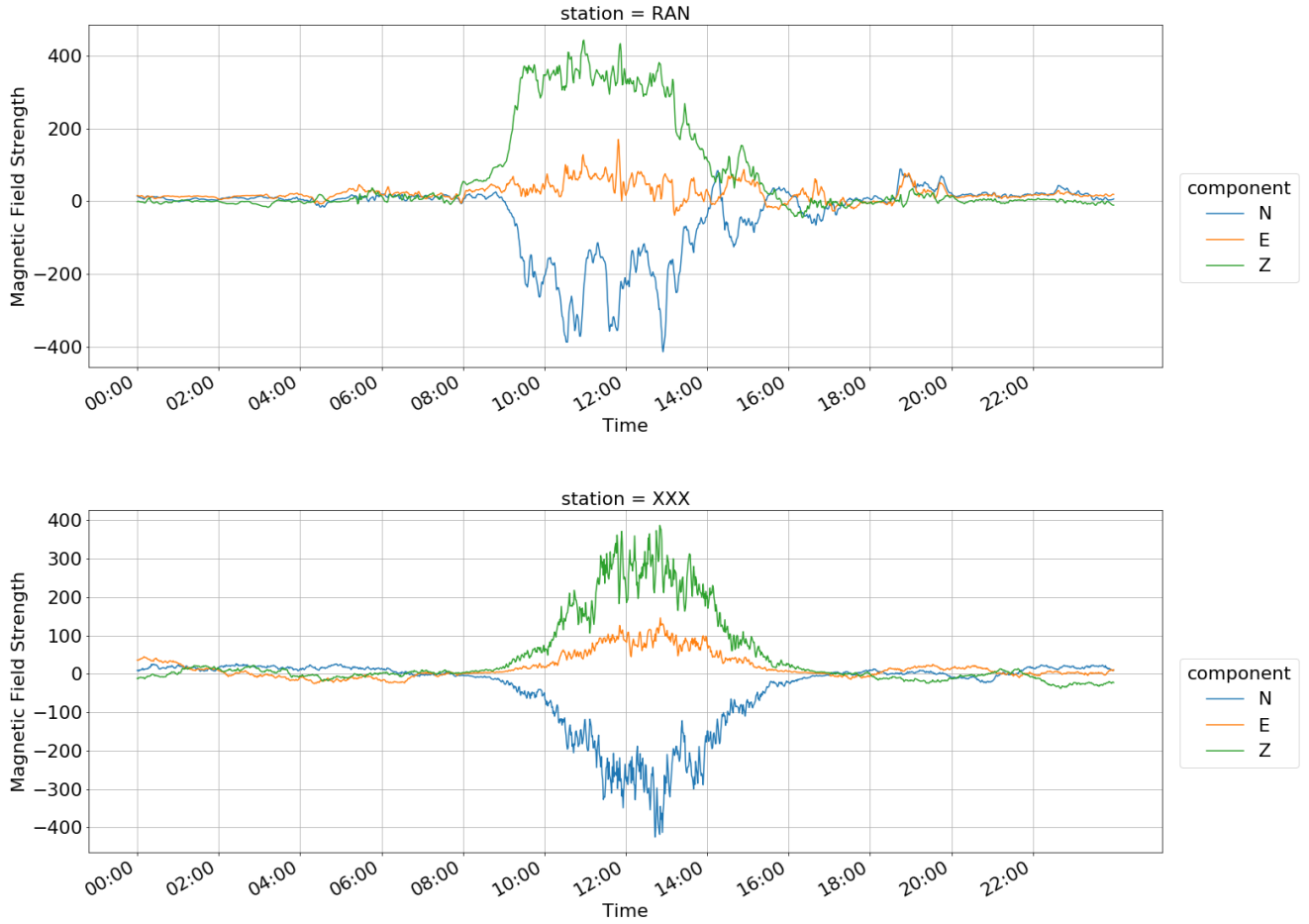


Fig. 11. Comparison of real data (top) and generated data (bottom).

are further away, which might actually prove beneficial given the variation in magnetometer sensitivity. This also has the ability to highlight sparser clusters of points. Ultimately, we decided to use $dr = r_{i+1} - r_i$ as we deemed the difference between lower dr values insufficient to compensate for the loss in information resultant from the overlap in intervals.

5.4 Results

5.4.1 Radial Distribution Function and the Effects of Vector Normalisation

The normalisation of vectors prior to the calculation of radial distribution actually turned out to affect the results significantly, as shown in Figure 12.

Without normalisation, the shape of the radial distribution corresponds roughly with that of the magnetometer readings. This is in line with our expectations as magnetometers have different sensitivities, so the amplitude of readings from different stations may vary significantly.

In contrast, when using normalised data, the density is almost non-existent in the $[0, 0.5]$ range and are concentrated towards higher values of r in both the stormy day and the quiet day. This points to a general lack of coherence in the direction of magnetometer readings which may be ameliorated through incorporating lag.

5.4.2 Mean of φ

The mean order parameter, $\bar{\varphi}$, across the 500 timestep windows is 0.52 for the day with high magnetic activity. This is significantly lower than that of the quiet day at 0.63 (both to 2 significant figures) and suggests that the flux of the magnetic field is less ordered during storms. This is once again within expectations and agrees with our result from network parameter analysis that the networks on event days exhibit lower connectivity, implying lower correlation.

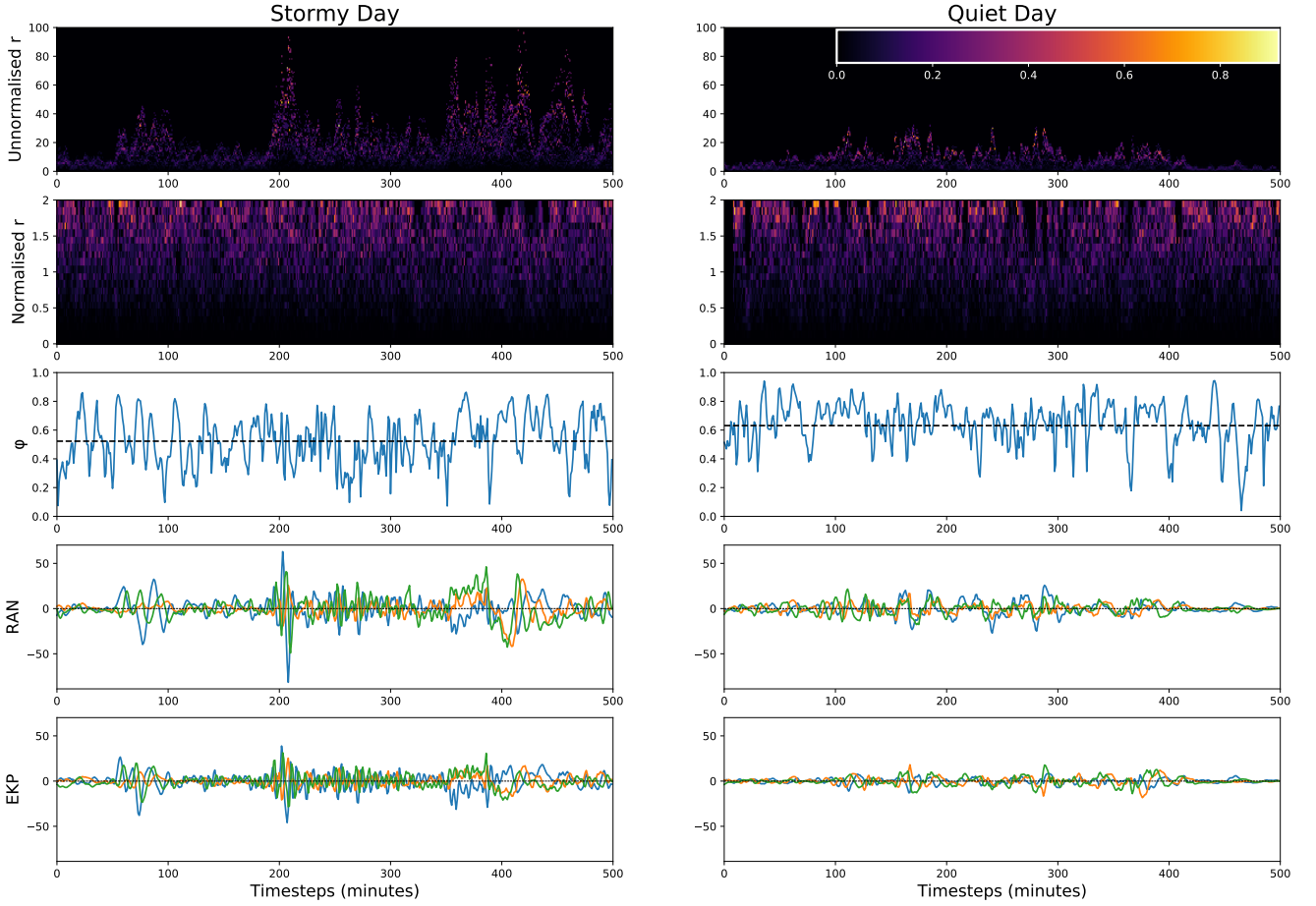


Fig. 12. Plots to showcase the results of the flocking parameters calculations done on stations from the North 330° chain. The left column contains 500 timesteps from a day with high magnetic activity (10:30, 3 April 2001), and the right column 500 timesteps from a quiet day (2:40, 3 February 1998). The rows (from top to bottom) are radial distribution $g(r)$ without and with normalised magnetometer readings, order parameter φ with a horizontal line indicating $\bar{\varphi}$, and detrended data from sample stations RAN and EKP.

5.5 Discussion

While we have demonstrated that collective motion techniques can be applied to systems of magnetometer readings and can potentially be useful for characterising solar events, there exist several difficulties in the context of utilising it to predict extreme space weather events.

For example, the calculation of the radial distribution function is highly computationally expensive, especially as we may want to scale this up to include more units (magnetometers) or run it in real time.

Moreover, $\bar{\varphi}$ is unlikely to be useful in prediction. Due to the rapidly fluctuating nature of φ , we need to calculate the mean over a long time window. Therefore any meaningful decrease in $\bar{\varphi}$ will not manifest until much later, making it unsuitable for prediction.

In addition, the simple mean subtraction technique we used for detrending the data is inadequate. The biggest prob-

lem comes from the fact that the Earth's magnetic field lines have different lengths, and thus natural frequencies, at different latitudes. When the Earth experiences an extreme solar event, the signals measured by magnetometers at different latitudes will resonate at different frequencies, meaning we need to detrend stations at different latitudes individually. Future endeavors may prove this analysis to be useful, but more work needs to be done to ensure the relevance of the results.

6 Conclusion

In this project we have investigated a number of approaches to predicting and quantifying extreme space weather events.

Our network-based methodology built upon the work of Dods et al. (2015) to construct a network of magnetometer stations that used the maximum correlation between stations, subject to a range of permissible time offsets, to determine

connections. This extension to the network-based approach had not been done previously, and we have been able to generate graphics that demonstrate how the connections between stations varies over time.

By constructing these networks, we have found two potential avenues for identifying a magnetic substorm: a change in network parameters such as average station degree and clustering coefficient, and a change in the variance of the temporal lag associated with the highest canonical correlation coefficients. The network parameters are easily calculated from the network structure that we have provided, and have been seen to change significantly from a quiet day to the midst of an event. The variance in temporal lag may prove to be more difficult to consistently quantify as a range of temporal lags must be fixed before performing the analysis. With more insight into the bounds of this range, this feature of our network-based approach could prove to be a reliable indicator of magnetic substorms.

Finally by applying methods inspired by flocking theory we were able to calculate the order parameter and radial density function for magnetometer readings. We found that these could potentially be used to characterise solar activities, but are unlikely to be useful in the prediction of space weather events.

In terms of further work, there is scope for the new approaches outlined in this project to complement the current state-of-the-art methods for prediction and quantification of extreme space weather events.

The work we have presented is primarily a proof-of-concept: we have outlined, developed and coded methodologies that could hold predictive power, but applying them extensively was not possible in the time available.

The next step would be to apply our methods across large subsets of the data. So far we have primarily been concerned with station chains (generally lying on the same line of longitude) but it would be interesting to expand this to observe what happens when our method are applied to stations that are more spread out, or to the entire global network.

Of course our results do not exist in isolation. There is satellite data available as well as imaging of the aurora (which are related to the ground-based magnetometer readings (Gjerloev, 2009)), and looking at our results in conjunction with the other data sources could reveal further connections that are not obvious when they are viewed in isolation.

6.1 Further work on the network-based approach

It would be interesting to explore some of the methodological choices we made in the application of the network-based approach. For example when calculating correlations we generally used a 128-minute time window. This was a sensible but somewhat arbitrary choice and it would be worth investigating the effects of varying this parameter.

Similarly it would be worth exploring other approaches to setting the threshold for determining whether stations ought

to be connected. Currently the threshold for a station pair depends only on the time series for those two stations. It would be worth investigating some relative thresholds (i.e. that depend on the data for all the stations in the network), and also factoring in the time of year (e.g. taking into account that March is typically the most active month) or the position of the 11-year solar cycle.

The network-based approach gives us the opportunity to investigate parameters that have never previously been analysed. Now that we have a method (and accompanying code) that takes the readings from a set of magnetometers as input, then builds a network and outputs the associated parameters, it would be interesting to see whether similarities between the parameters for different space weather events could be found using machine learning.

If a robust method is used to automatically classify events, supervised learning could be used identify whether there any network parameters that are predictive of future solar activity. Alternatively, unsupervised learning would allow us to group time series according to their network parameters, clustering those that exhibit similar features.

Given that there are so many open questions in this field, it is clear that there is significant scope for investigating the results of applying our methods.

Acknowledgements. Dr Sandra Chapman, SuperMAG, NASA

References

- J. Dods, S. Chapman, and J. Gjerloev. Network analysis of geomagnetic substorms using the supermag database of ground-based magnetometer stations. *Journal of Geophysical Research: Space Physics*, 120(9):7774–7784, 2015.
- A. Hanslmeier. The sun and space weather. In *Heliophysical Processes*, pages 233–249. Springer, 2010.
- N. Basavaiah. *Geomagnetism: solid earth and upper atmosphere perspectives*. Springer Science & Business Media, 2012.
- NASA. Solar storms and radiation exposure on commercial flights, 2009. URL <https://www.nasa.gov/topics/earth/features/airline-radiation.html>.
- R. Schwenn. Space weather: The solar perspective. *Living reviews in solar physics*, 3(1):2, 2006.
- NASA. ‘cinco de mayo’ solar flare, May 2017a. URL <https://www.nasa.gov/feature/goddard/nasas-sdo-observes-cinco-de-mayo-solar-flare>.
- NASA. 2017 total solar eclipse, Aug 2017b. URL <https://www.nasa.gov/image-feature/2017-total-solar-eclipse>.
- A. Hundhausen, C. Sawyer, L. House, R. Illing, and W. Wagner. Coronal mass ejections observed during the solar maximum mission: Latitude distribution and rate of occurrence. *Journal of Geophysical Research: Space Physics*, 89(A5):2639–2646, 1984.
- NASA. Nasa helps power grids weather geomagnetic storms, Nov 2015. URL <https://www.nasa.gov/feature/goddard/2016/nasa-helps-power-grids-weather-geomagnetic-storms>.

- Athanasios Papaioannou, Ingmar Sandberg, Anastasios Anastasiadis, Athanasios Kouloumvakos, Manolis K Georgoulis, Kostas Tziotziou, Georgia Tsiropoulou, Piers Jiggins, and Alain Hilgers. Solar flares, coronal mass ejections and solar energetic particle event characteristics. *Journal of Space Weather and Space Climate*, 6:A42, 2016.
- Rob Garner. Solar storm and space weather - frequently asked questions, Mar 2015. URL https://www.nasa.gov/mission_pages/sunearth/spaceweather/index.html#q5.
- Donald V Reames. The two sources of solar energetic particles. *Space Science Reviews*, 175(1-4):53–92, 2013.
- NASA, Jul 2014. URL <https://science.nasa.gov/science-news/science-at-nasa/2014/23jul.superstorm>.
- M. Moldwin. *An introduction to space weather*. Cambridge University Press, 2008.
- NASA. A super solar flare, May 2008. URL <https://science.nasa.gov/science-news/science-at-nasa/2008/06may.carringtonflare>.
- Lloyds. *Solar Storm Risk to the North American Electric Grid*. Lloyd's of London, 2013.
- Cabinet Office. *National Risk Register of Civil Emergencies*. UK Government, 2012.
- GOV.UK Space Agency. Uk and usa join forces to launch 'space weather' service, Jun 2019. URL <https://www.gov.uk/government/news/uk-and-usa-join-forces-to-launch-space-weather-service>.
- NOAA. Coronal mass ejections, 2019. URL <https://www.swpc.noaa.gov/phenomena/coronal-mass-ejections>.
- M. Hapgood. Societal and economic importance of space weather. In E. Camporeale, S. Wing, and J.R. Johnson, editors, *Machine Learning Techniques for Space Weather*, pages 3 – 26. Elsevier, 2018.
- Pan Qin and Ryuei Nishii. Statistical prediction of dst index by solar wind data and t -distributions. *IEEE Transactions on Plasma Science*, 43(11):3908–3915, 2015.
- JW Gjerloev. A global ground-based magnetometer initiative. *Eos, Transactions American Geophysical Union*, 90(27):230–231, 2009.
- J. Gjerloev. The supermag data processing technique. *Journal of Geophysical Research: Space Physics*, 117(A9), 2012.
- A. Janzhura and O. Troshichev. Determination of the running quiet daily geomagnetic variation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(7):962–972, 2008.
- K. Takahashi and B. Anderson. Distribution of ulf energy (f_i 80 mhz) in the inner magnetosphere: A statistical analysis of ampte cce magnetic field data. *Journal of Geophysical Research: Space Physics*, 97(A7):10751–10773, 1992.
- M. Menvielle, N. Papitashvili, L. Häkkinen, and C. Sucksdorff. Computer production of k indices: review and comparison of methods. *Geophysical Journal International*, 123(3):866–886, 1995.
- Tamás Vicsek and Anna Zafeiris. Collective motion. *Physics reports*, 517(3-4):71–140, 2012.
- Department for Business Innovation and Skills. *Space Weather Preparedness Strategy*. Cabinet Office, 2015.
- R. Van Der Linden. The solar weather/solar activity monitoring and forecast. In J. Liliensten, editor, *Space Weather*, chapter 10, pages 1–3. Springer, 2007.
- National Research Council. *Severe Space Weather Events: Understanding Societal and Economic Impacts*. The National Academic Press, 2008.
- V. Sergeev, V. Angelopoulos, and R. Nakamura. Recent advances in understanding substorm dynamics. *Geophysical Research Letters*, 39(5), 2012.
- J. Bortnik, X. Chu, Q. Ma, Q. Li, X. Zhang, R. Thorne, V. Angelopoulos, R. Denton, C. Kletzing, G. Hospodarsky, H. Spence, G. Reeves, and Baker D. Kanekal, S. Artificial neural networks for determining magnetospheric conditions. In E. Camporeale, S. Wing, and J.R. Johnson, editors, *Machine Learning Techniques for Space Weather*, pages 279–300. Elsevier, 2018.
- Parliamentary Office of Science and Technology. Space weather, Jul 2010. URL <https://www.parliament.uk/documents/post/postpn361-space-weather.pdf>.
- E. Underwood. A better way to predict space storms, Dec 2017. URL <https://eos.org/research-spotlights/a-better-way-to-predict-space-storms>.
- Office of Risk Management and Analysis. *OECD/IFP Futures Project on "Future Global Shocks" - "Space Weather"*. OECD, Jan 2011. URL <https://www.oecd.org/gov/risk/46891645.pdf>.

Appendix A**Google Drive Folder for gif Files**

The link to the Google Drive folder containing the gif files referenced at the end of Section 3.2.2 is <https://drive.google.com/drive/folders/15g8R4lXdoTKIr5j0xOWxskwDtybqPZfq?usp=sharing>.

Appendix B**GitHub Code**

Here is the link to the GitHub repository that hosts all of the code for this report: <https://github.com/kendalfoster/RSG-Space-Weather>.