On the Considerations of Using Near Real Time Data for Space Weather Hazard Forecasting

A. W. Smith[1], C. Forsyth[1], I. J. Rae[2], T. M. Garton[3], C. M. Jackman[3], M. Bakrania[1], R. M. Shore[4], G. S. Richardson[5], C. D. Beggan[5], M. J. Heyns[6], J. P. Eastwood[6], A. W. P. Thomson[5], J. M. Johnson[7,8]

[1]Mullard Space Science Laboratory, UCL, Dorking, UK
[2]Department of Mathematics, Physics and Electrical Engineering, Northumbria University, Newcastle upon Tyne, UK
[3]School of Cosmic Physics, DIAS Dunsink Observatory, Dublin Institute for Advanced Studies, Dublin 15, Ireland
[4]British Antarctic Survey, Cambridge, UK
[5]British Geological Survey, Edinburgh, UK
[6]Imperial College London, London, UK
[7]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA
[8]Space Weather Prediction Center, National Oceanic and Atmospheric Administration, Boulder, CO, USA

Key Points:

- Near-Real-Time solar wind data show increased short-term variability and occasional anomalous values compared to post-processed data.
- Data gaps are most frequently short, meaning short interpolation schemes dramatically increase continuous data availability.
- Data gaps are more prevalent in the plasma moments than the magnetic field from ACE and DSCOVR, but DSCOVR is more continuous.

Corresponding author: A. W. Smith, andy.w.smith@ucl.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2022SW003098.

Abstract

Space weather represents a severe threat to ground-based infrastructure, satellites and communications. Accurately forecasting when such threats are likely (e.g. when we may see large induced currents) will help to mitigate the societal and financial costs. In recent years computational models have been created that can forecast hazardous intervals, however they generally use post-processed "science" solar wind data from upstream of the Earth. In this work we investigate the quality and continuity of the data that are available in Near-Real-Time (NRT) from the ACE and DSCOVR spacecraft. In general, the data available in NRT corresponds well with post-processed data, however there are three main areas of concern: greater short-term variability in the NRT data, occasional anomalous values and frequent data gaps. Some space weather models are able to compensate for these issues if they are also present in the data used to fit (or train) the model, while others will require extra checks to be implemented in order to produce high quality forecasts. We find that the DSCOVR NRT data are generally more continuous, though they have been available for small fraction of a solar cycle and therefore DSCOVR has experienced a limited range of solar wind conditions. We find that short gaps are the most common, and are most frequently found in the plasma data. To maximize forecast availability we suggest the implementation of limited interpolation if possible, e.g. for gaps of five minutes or less, which could increase the fraction of valid input data considerably.

Plain Language Summary

The variable plasma conditions in near-Earth space can create hazards for modern society. These include the generation of anomalous ground currents that pose a threat to the operation of infrastructure such as high voltage power grids. Forecasts of intervals when we are likely to be at risk generally use solar wind measurements gathered by satellites from upstream of the Earth. Various computational models have shown skill in predicting risk intervals; however, they are generally created using scientific quality data which are not available in near-real-time (NRT). To prepare for transitioning such models to operational use we assess the similarities and differences between the scientific quality and NRT data. We assess the properties and frequency of data gaps in the NRT data, to build an understanding of how to maximize the time for which forecasts can be successfully created.

1 Introduction

Space weather has the potential to pose a severe threat to modern society. The Earth's magnetosphere is constantly buffeted by the solar wind that emanates from the Sun. It is the variable nature of the solar wind that leads to a highly dynamic environment in near-Earth space. There are many different facets to space weather, including the changing radiation environment in near-Earth space, a hazard that faces Earth orbiting satellites (D. Baker et al., 1987; Iucci et al., 2005), and strong ground magnetic field variability that can induce damaging currents (Geomagnetically Induced Currents, GICs) in conductive infrastructure (Boteler et al., 1998; Boteler, 2021; Rajput et al., 2020). Forecasting intervals of risk in a timely manner is key; this allows necessary mitigating action to be taken.

Space weather forecasts are typically driven by data obtained upstream of the Earth at the L1 point (e.g. D. N. Baker et al., 1990; Wing et al., 2005; Forsyth et al., 2020; Keesee et al., 2020; A. W. Smith, Forsyth, Rae, Garton, et al., 2021; Chu et al., 2021), approximately 1.5 million km ahead of the Earth. Given the speed of the solar wind, this provides between 20 - 90 minutes of warning before solar wind plasma encounters the Earth (Baumann & McCloskey, 2021). To account for the variable time delay between the plasma measured at L1 and the arrival of that plasma at the Earth, many forecast models propagate the measurement to a fixed point relative to the Earth, the bow shock for example (e.g. Cash et al., 2016; Baumann & McCloskey, 2021). Such methods are present in commonly used scientific datasets, such as the OMNI database (https://omniweb.gsfc.nasa.gov/) (Weimer & King, 2008), and will have implications for data continuity.

Further, most current space weather forecasting models have been developed using post-processed, scientific quality data provided by the spacecraft at L1, or combined data products such as OMNI (e.g. Wintoft et al., 2017; Keesee et al., 2020; A. W. Smith, Forsyth, Rae, Garton, et al., 2021). Such data are generally only available after an extended period, beyond the interval in which a timely forecast must be made. Data are also available in near-real time (NRT), usually within five minutes of acquisition by the spacecraft (e.g. Zwickl et al., 1998). However, given limited telemetry and processing the NRT data are an approximation of the full scientific data that will later be available. This may lead to differences in the value that the data take, and also whether data are available for some intervals (e.g. Machol et al., 2013). Nonetheless, the NRT data must be used in order to provide actionable forecasts of upcoming hazardous space weather.

In this work we explore how we may best use the data available in NRT to produce space weather forecasts. We will use the example of the space weather threat to ground based infrastructure through GICs. GICs pose a severe threat to continuous and reliable power network operation in many countries around the world (e.g. Gaunt & Coetzee, 2007; Marshall et al., 2012), particularly at higher latitude locations (e.g. Bolduc, 2002). Even at mid-latitude locations, such as the UK, the estimated cost of a major geomagnetic storm - one that leads to a widespread and long-lasting interruption of electrical supply - has been estimated at billions of dollars a day (Oughton et al., 2017). However, with sufficient warning mitigating actions may be taken that would reduce this cost considerably (Eastwood et al., 2018; Oughton et al., 2019). Therefore, the forecasting and mitigation of large, damaging GICs is a critical endeavour.

GICs are driven by magnetic field variability at Earth's surface as a consequence of Faraday's law of induction. Though mainly discussed with respect to power networks (e.g. Pulkkinen et al., 2005; Mac Manus et al., 2017; Divett et al., 2018; Rajput et al., 2020), GICs may be induced in any large scale conducting infrastructure, such as pipelines (e.g. Campbell, 1980; Gummow & Eng, 2002; Viljanen et al., 2010; Dimmock et al., 2021) and railways (e.g. Wik et al., 2009; Liu et al., 2016; Love et al., 2019), with consequences ranging from increased corrosion to direct component failure. Ultimately strong magnetic field variability is driven by a myriad of magnetospheric processes in near-Earth space (e.g. Heyns et al., 2021; Rogers et al., 2020; Tsurutani & Hajra, 2021). Extreme field fluctuations are often linked to global scale transient phenomena such as geomagnetic storms (Kappenman & Albertson, 1990; Dimmock et al., 2019), substorms (Viljanen et al., 2006; Turnbull et al., 2009; Freeman et al., 2019) and sudden commencements (Rodger et al., 2017; A. W. Smith et al., 2019; A. W. Smith, Forsyth, Rae, Rodger, & Freeman, 2021).

Work in recent years has explored both empirical and first-principles methods to forecast the surface geomagnetic field or geomagnetic perturbations. These methods have included statistical techniques (e.g. Weigel et al., 2002; Weimer, 2013; Shore et al., 2017), global scale physics-based (Magneto-HydroDynamic, MHD) models (e.g. Pulkkinen et al., 2011, 2013; Tõth et al., 2014; Welling, 2019), machine learning-based techniques (e.g. Gleisner & Lundstedt, 2001; Wintoft et al., 2015, 2017; Keesee et al., 2020; A. W. Smith, Forsyth, Rae, Garton, et al., 2021; Blandin et al., 2022; Pinto et al., 2022; Upendran et al., 2022), or combinations of these methods (e.g. Camporeale et al., 2020). The forecast geomagnetic (or geoelectric) field predictions can then be used to drive models based on the local geology and properties of the power network to indirectly obtain GIC estimates (Beggan et al., 2013; Blake et al., 2016, 2018; Divett et al., 2018, 2020; Dimmock et al., 2021; Grawe & Makela, 2021; Mac Manus et al., 2022). Each model that is used to forecast the geomagnetic consequences of space weather will use the input solar wind data in a distinct fashion, and therefore may be impacted differently by the ways in which the NRT and scientific data differ.

In this study we explore the data that are available from the L1 point in NRT and investigate how they may be used by space weather forecasting models. In Section 2 we describe the data, while in Section 3 we evaluate the NRT data, including comparisons to the more commonly used scientific quality data. In Section 4 we show the results of training an example forecasting model on the NRT data, using a model based on A. W. Smith, Forsyth, Rae, Garton, et al. (2021). Section 5 then discusses the results in the context of producing space weather forecasts in near real time.

2 Data

In this study we use and explore the data made available by SWPC (the Space Weather Prediction Center) in near real time (NRT) from the ACE (Advanced Composition Explorer) and DSCOVR (Deep Space Climate Observatory) satellites located at L1. ACE launched in 1997 as a part of the NASA Explorer program and has provided solar wind observations since 1998 (Stone et al., 1998). Meanwhile, DSCOVR was a NASA/NOAA mission that launched more recently in 2015. The NRT data is provided at one minute resolution (Zwickl et al., 1998). For the magnetic field we consider the three GSM (Geocentric Solar Magnetospheric) components of the magnetic field (B_X^{GSM}, B_Y^{GSM}, B_Z^{GSM}) and the total field strength ($|B|$) (C. Smith et al., 1998). Meanwhile, for the plasma data we evaluate the derived proton density (n_p), solar wind bulk speed (V) and ion temperature (T_i) (McComas et al., 1998; Aellig et al., 2001).

The ACE spacecraft was the operational real time solar wind monitor at L1 until mid-2017, at which time it was replaced by DSCOVR. In this study we therefore use data from ACE between 1999 and 2015 (inclusive) and from DSCOVR between 2018 and 2020 (inclusive). This selection ensures we are using intervals in which each spacecraft was the primary real time operational solar wind monitor, and therefore prioritized for telemetry downlink. We note that even though DSCOVR is currently the operational real time solar wind monitor, data from ACE may occasionally be used to fill large missing intervals in the DSCOVR data. Due to issues with the spacecraft, DSCOVR data are largely absent between 06/26/19 and 02/26/20, however we do not perform any substitution with ACE during this period, and simply use the data that are available. Switching the data between spacecraft that are not co-located or cross-calibrated results in its own challenges that are beyond the scope of this work.

In Sections 3 and 4 we use scientific quality data for comparative purposes. For ACE we use H0 data (level 2), that have been re-sampled to the same one minute cadence as the NRT data. For the magnetic field data this means they are down sampled from a 16s cadence, while the plasma moments are up sampled from a 64s cadence. We note that if we instead down sampled the NRT measurements to 64s prior to running the analysis it would result in very small changes ($\leq 0.4\%$) to the quantitative comparisons performed. Meanwhile, for DSCOVR we use H0 magnetic field data (down-sampled from 1s resolution), and H1 plasma moments at a one minute resolution. The DSCOVR H1 plasma moments are currently only available until 06/27/2019, limiting the time interval for which the comparison with the NRT data can be evaluated.

3 Data Evaluation

The NRT data from the DSCOVR and ACE spacecraft at L1 are available within minutes of recording by the spacecraft and are therefore automatically processed. On the ground further processing steps, such as manual data review, data cleansing, more complex fitting methods and recalibration are used to produce science quality data. It is important to consider any differences between these two types of data when moving space weather models from the science quality data on which they were likely developed/evaluated to the operational NRT data. Though DSCOVR is the current operational solar wind monitor some models require large quantities of training data and thus the historical ACE NRT data are also of interest. Below, we evaluate the NRT data values for ACE and DSCOVR, as well as their continuity.

3.1 Data Validity

First, we compare the values returned by the NRT data and the equivalent science quality data that is released later after post-processing. Figure 1 shows a series of two-dimensional histograms of the occurrence of plasma moment values in these data sets for ACE (left) and DSCOVR (right). We note that if either data set (NRT or science) is missing then that interval is not represented in Figure 1. Additionally, in this analysis our base assumption is that the science data are the "correct" values, to which the NRT data represent an initial estimate. Therefore, in an ideal situation we would observe the case where both the science and NRT values are close, and lie along the red dashed lines (or the NRT data are within $\pm 10\%$ of the scientific values). However, we see deviations from this for both the ACE and DSCOVR data.

Inspecting the results for the ACE data (left of Figure 1) we see that the plasma velocity is best represented in the NRT data (Figure 1c), with the histogram most closely following the diagonal line of equality and the vast majority of the data (96%) being bounded by the orange $\pm 10\%$ lines. In contrast, the solar wind density (Figure 1a) and ion temperature (Figure 1e) are less well captured by the ACE NRT data, showing much larger spreads: only $\sim 20-30\%$ of both NRT data sets are within 10% of the science values. We also see evidence for a hard-coded lower limit to the NRT ion temperatures (Figure 1e).

Considering the DSCOVR data (right of Figure 1), we once again find that the plasma velocity is best captured by the NRT data (Figure 1d). The vast majority of the DSCOVR data lie along the red dashed line, with 94% of the NRT data within 10% of the science values. We find similar spread in the DSCOVR density comparison (Figure 1b) as we did for the ACE data ($\sim 20\%$ within $\pm 10\%$). NRT data from both spacecraft appear to commonly underestimate the plasma density. Finally, considering the NRT temperatures reported by DSCOVR we find 20% of the data within the $\pm 10\%$ lines, slightly less than the 30% found for the ACE data. We note that the DSCOVR spacecraft has been operational for a shorter interval, and so has experienced fewer extreme events (e.g. at larger n_p or V).

Figure 2 shows a similar comparison of the magnetic field data, comparing the NRT and science data for ACE (left) and DSCOVR (right). Statistically, the ACE data are peaked around the diagonal line of gradient unity, though there is wide spread in values. This suggests that the magnetic field values are often corrected later on, sometimes by $10s$ of nT. There is noticeably less spread in the comparison of the B_X^{GSM} values (Figure 2a), though interestingly there are a series of values that seem to form a parallel distribution to the red line. This would correspond to a small offset of 5 nT or less, where the NRT data under-reported the value of B_X^{GSM}. There is also evidence of a distribution perpendicular to the line of equality in the results for B_X^{GSM} and B_Y^{GSM}, perhaps representing a rotation of the magnetic field in the X-Y plane. We note that despite the spread in the field components reported by ACE, 68% of the NRT measurements of the total field magnitude ($|B|$) are within $\pm 10\%$ of the science values.

When we compare the DSCOVR results (right of Figure 2), we again find that the NRT data are much more representative of the science data than was observed for ACE, with a clear diagonal distribution dominating the results. As with ACE, the NRT total field strength ($|B|$) is also very close to that later found within the science data, with 94% of the values being within 10%.

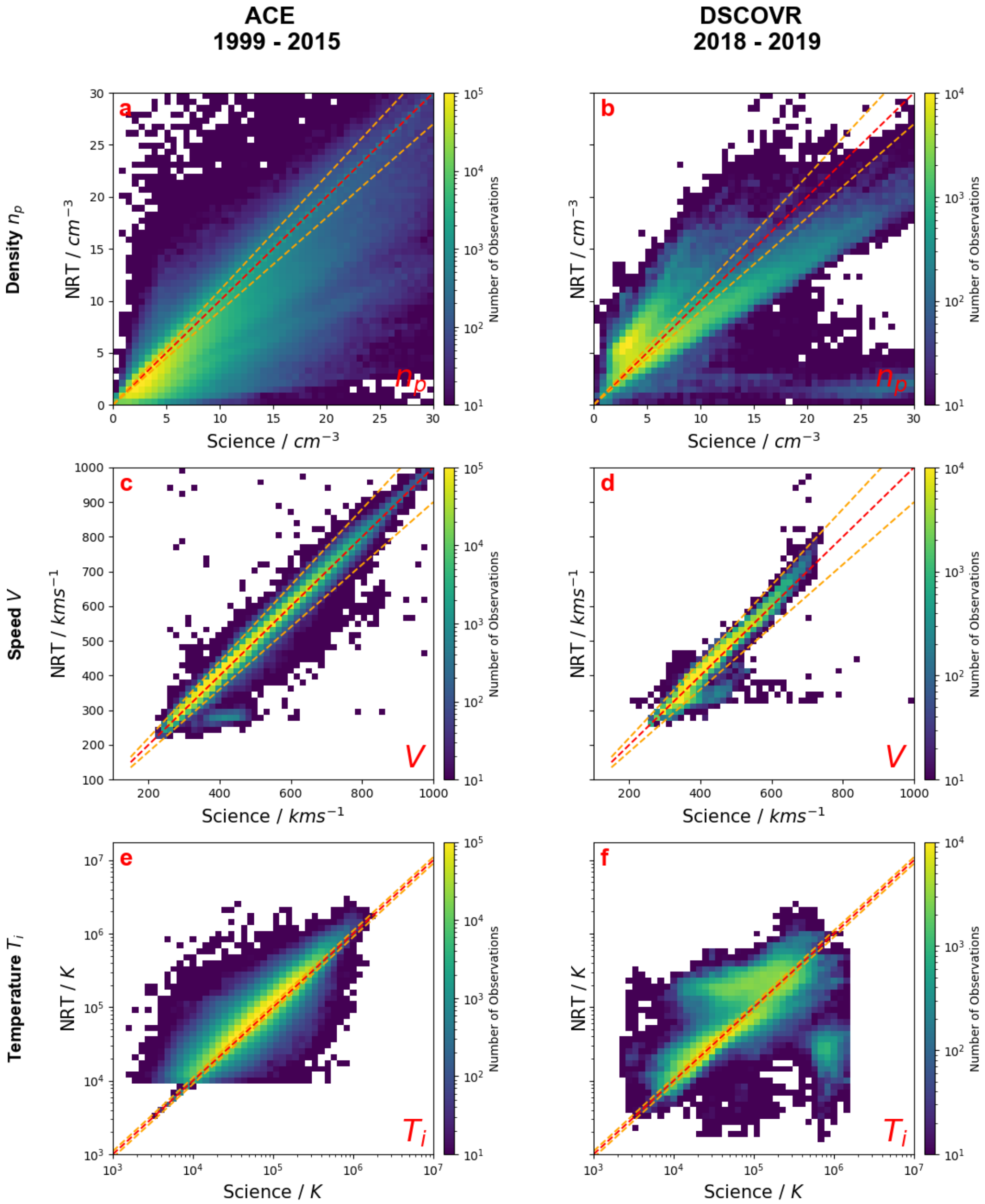

Figure 1. Comparison of the plasma moments in the NRT data with that recovered in the science quality data. Left column (a, c, e) shows ACE between 1999 and 2015 (inclusive), right column (b, d, f) shows DSCOVR between 2018 and mid-2019. The plasma moments shown are the plasma density (a, b), bulk velocity (c, d), and ion temperature (e, f). The diagonal red dashed line indicates where the data returned are equivalent, while the orange dashed lines indicate the region where the NRT values are within ±10% of the science data.

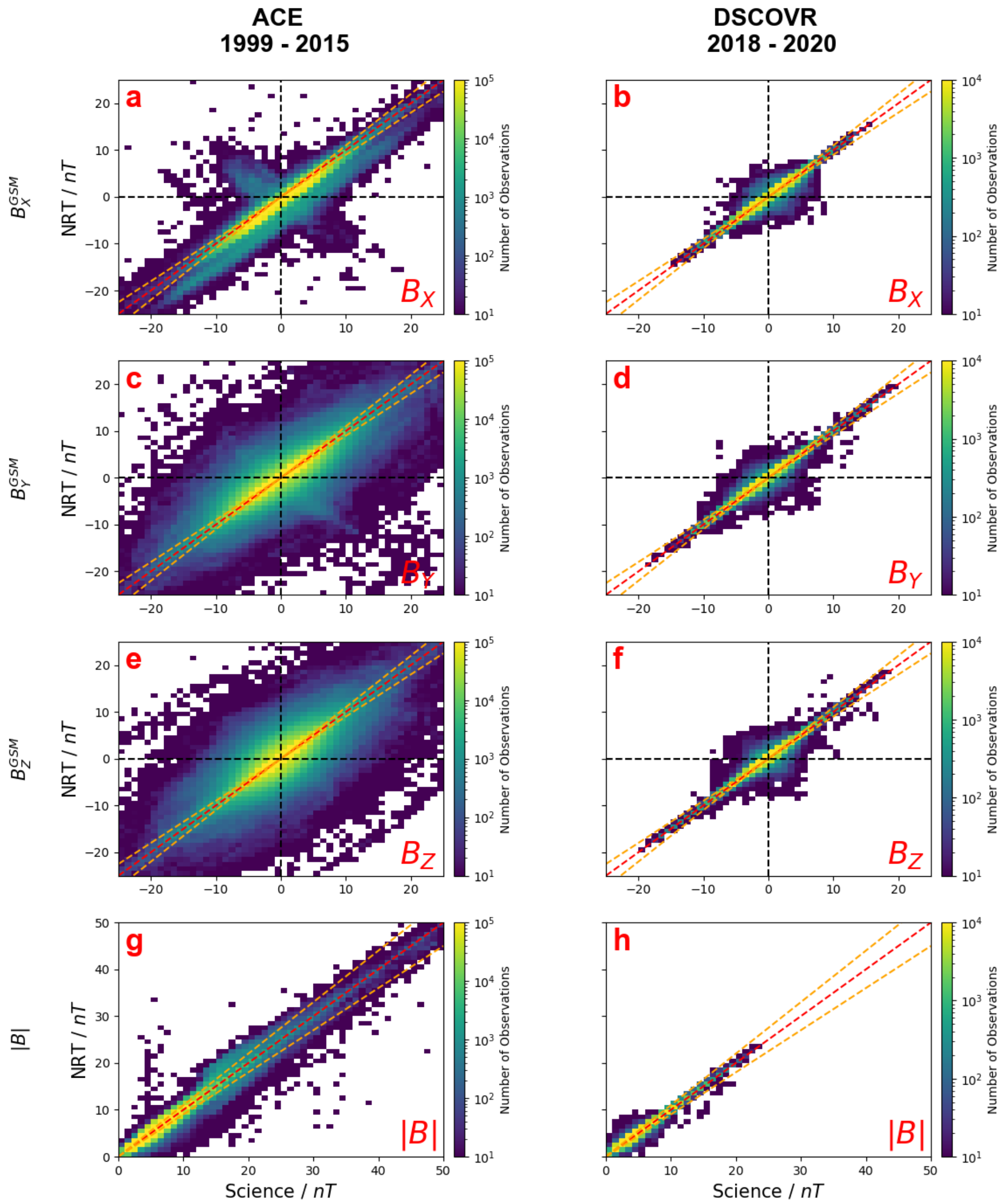

Figure 2. Comparison of the magnetic field values returned by the NRT data and science data. The format is similar to Figure 1, with the ACE (1999 to 2015) shown on the left (a, c, e, g) and DSCOVR (2018 to 2020) on the right (b, d, f, h). The top row shows the B_X^{GSM} component (a, b), the second row shows the B_Y^{GSM} (c, d), while the third and fourth rows show the B_Z^{GSM} (e, f) and total field $|B|$ (g, h).

3.2 Data Continuity

The magnetosphere is a highly dynamic system that couples to the solar wind on a range of time scales (e.g. Coxon et al., 2019; Shore et al., 2019; Borovsky, 2020). Space weather models therefore often require information regarding the preceding interval of solar wind, rather than a single measurement of the current conditions. Incomplete data - those with communication gaps or missing measurements - may pose a problem for this approach. Data gaps have also been inferred to cause errors in the derivation of coupling functions (Lockwood et al., 2019). There are many reasons why data may be missing from the real time data stream, or should be ignored. Missing data are generally flagged with a code according to the reason, be it an operational consideration (e.g. downlink issues) or problems/failure of the measurement or recording. It may be possible to develop bespoke solutions to account for different missing data flags, however in the following we treat all missing and/or flagged data identically.

3.2.1 Data Gap Occurrence

Figure 3 demonstrates the average yearly occurrence of data gaps with different lengths in the NRT data from ACE (top) and DSCOVR (bottom). The equivalent distributions for the science data sets are shown in black. Starting with ACE, we can see that the most common data gap length in both data sets is five minutes or less, i.e. five or fewer data points. We observe that there are nearly 200,000 gaps per year of less than five minutes in the plasma data (Figure 3b). The ACE NRT magnetic field data is more complete by comparison, and whilst the sub-five-minute data gaps are the most common type, there are only around 4000 of these per year. Comparing the scientific and NRT data, we see that the ACE scientific magnetic field data is much more continuous, with several orders of magnitude fewer gaps per year (of any duration). Meanwhile, for the ACE plasma data we see that there are generally fewer gaps shorter than ~ 60 minutes in the science data, but there are on average more gaps of greater duration, perhaps as longer windows of data are later flagged for removal.

DSCOVR NRT data has noticeably fewer data gaps per year than its ACE counterpart, particularly considering the plasma data. However, we do see the same patterns, with sub-five-minute data gaps being the most frequently observed, and the plasma data exhibiting more numerous gaps per year than the magnetic field. For instance, there were around 10,000 sub-five-minute data gaps per year in the plasma data, while there are approximately 4000 per year in the NRT magnetic field data. If we consider the equivalent distributions for both the DSCOVR scientific magnetic field and plasma data, we see fewer gaps of less than five minutes, but additional gaps of an hour or longer, suggesting that - as with ACE - longer periods of data are later discarded.

3.2.2 Windowed Data Validity

Space weather models often require continuous data as an input, and many numerical or computational techniques will not return a result should even a single entry be missing. For example, a model may require a window of 30 minutes of continuous solar wind data as an input (c.f. A. W. Smith, Forsyth, Rae, Garton, et al., 2021). It is therefore important to examine the fraction of input data windows that gaps would invalidate. This is examined in Figure 4, for the ACE (top) and DSCOVR (bottom) NRT data, plotted in green. In Figure 4 the windows are evaluated as they would present in an operational setting, such that the stride between adjacent windows of data is equal to the cadence of the data (one minute), and therefore the intervals of data tested overlap.

When using the ACE NRT data (Figure 4a, b), we can see that if five minutes or fewer of continuous data are required (i.e. the model requires an interval of five minutes

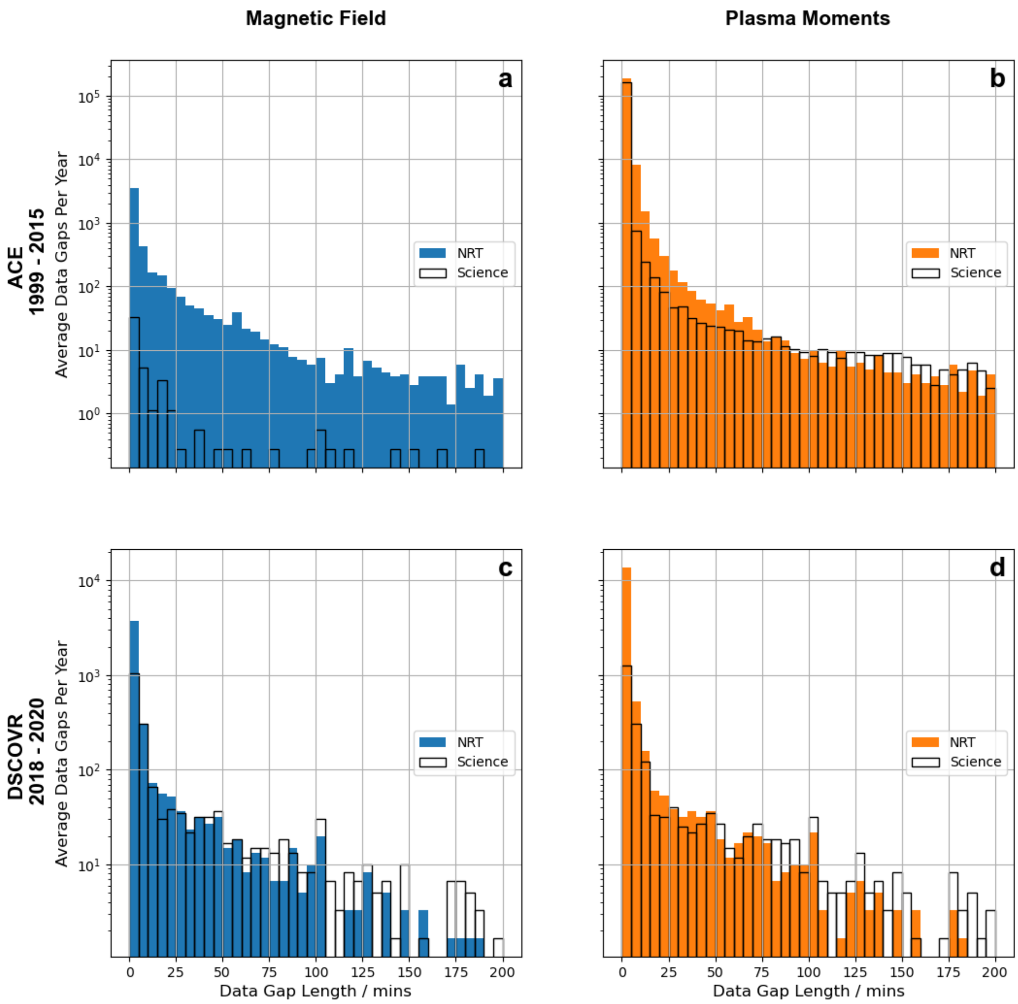

Figure 3. The average yearly occurrence of data gaps of different lengths for the ACE (top: a, b) and DSCOVR (bottom: c, d) NRT data, over 16 and 3 years respectively. This is shown for the magnetic field data (left: a, c) and plasma moments (right: b, d). The results for the science data are provided in black. Note the logarithmic y-axes.

of solar wind data as input) then around 90% of the magnetic field data will be available. However, should plasma data also be required then only 55% of input data are valid. As the required length of input window increases, the percentage of valid data intervals decreases. If two hours (120 minutes) of continuous input are required then only 75% of magnetic field data suffice, and approximately 1% of plasma data are available.

Comparatively, DSCOVR NRT data are more complete (Figure 4c, d), noting the different vertical axis scale. For DSCOVR, should a five minute window of data be required then 97% of magnetic field data and 96% of plasma data will be valid. Again, this decreases with longer input window lengths, reaching 86% of magnetic field and 72% of plasma data as the input window length reaches 120 minutes. While this represents a major improvement on ACE, it still has the potential to be a serious issue for forecasting models.

Given how common we have shown short data gaps to be in the NRT data (Figure 3), it is useful to consider how the use of interpolation schemes can increase the quantity of valid, continuous windows of data. Figure 4 shows the application of two maximum lengths of interpolation, in addition to the use of the "raw" data. We can see that if gaps of five minutes or fewer are interpolated (orange stars in Figure 4) then the fraction of valid data intervals increases significantly for both spacecraft and types of data. If the interpolation is permitted over larger gaps of 15 minutes or less (blue crosses in Figure 4) then the fractions increase again, though this is a smaller improvement than was found for the change from the raw data to interpolation over five minutes or less.

4 Example NRT Forecasting of Ground Magnetic Activity

We now qualitatively compare space weather forecast models using science quality data and NRT data. The example model we use is the CNN-based (Convolutional Neural Network) model developed by A. W. Smith, Forsyth, Rae, Garton, et al. (2021), designed to provide a probabilistic forecast as to whether the rate of change of the ground magnetic field (R) would exceed certain fixed thresholds at an observatory on the ground (i.e. may be linked to an enhanced GIC risk). In these examples we use the magnetic field at the LER (Lerwick) magnetometer station in Scotland as our target. For full model details the interested reader is directed to A. W. Smith, Forsyth, Rae, Garton, et al. (2021).

The models were trained on data between 2003 and 2014, with 1998 to 2002 being used as a validation set during training. This means that the data from 2015 and 2016 can be used as an unseen test of model performance, e.g. the selected storms discussed below. The models require copious amounts of data to train, covering as complete a record of possible solar wind conditions as can be sourced, and thus we are limited to the ACE datasets. Training a model on the combined ACE and DSCOVR datasets is beyond the scope of this work. Following on from Section 3.2, we interpolate linearly over gaps smaller than 15 minutes in order to maximize the data availability, and perform this consistently for both the ACE NRT and science data. We note that in an NRT data situation we will not necessarily know the next available value as the more recent time intervals may have no data, but we can repeat the last recorded value for a short interval of time. A forecast horizon of 180 minutes has been used for both models.

First, we show the 17th/18th March 2015 St Patrick's Day storm in Figure 5, which has been previously noted for its ionospheric (e.g. Astafyeva et al., 2015) and ground impacts (e.g. Carter et al., 2016; Kozyreva et al., 2018). We can compare the NRT and science data in the top three panels of Figure 5, from which we can observe some of the results previously discussed in Section 3. It is clear from the velocity and density that there is more rapid temporal variability in the NRT data than is seen in the post-processed ACE data (i.e. Figure 5a, c compared to Figure 5b, d). This short term-variability would explain some of the scatter in Figures 1. There are also significant data gaps in the sci-

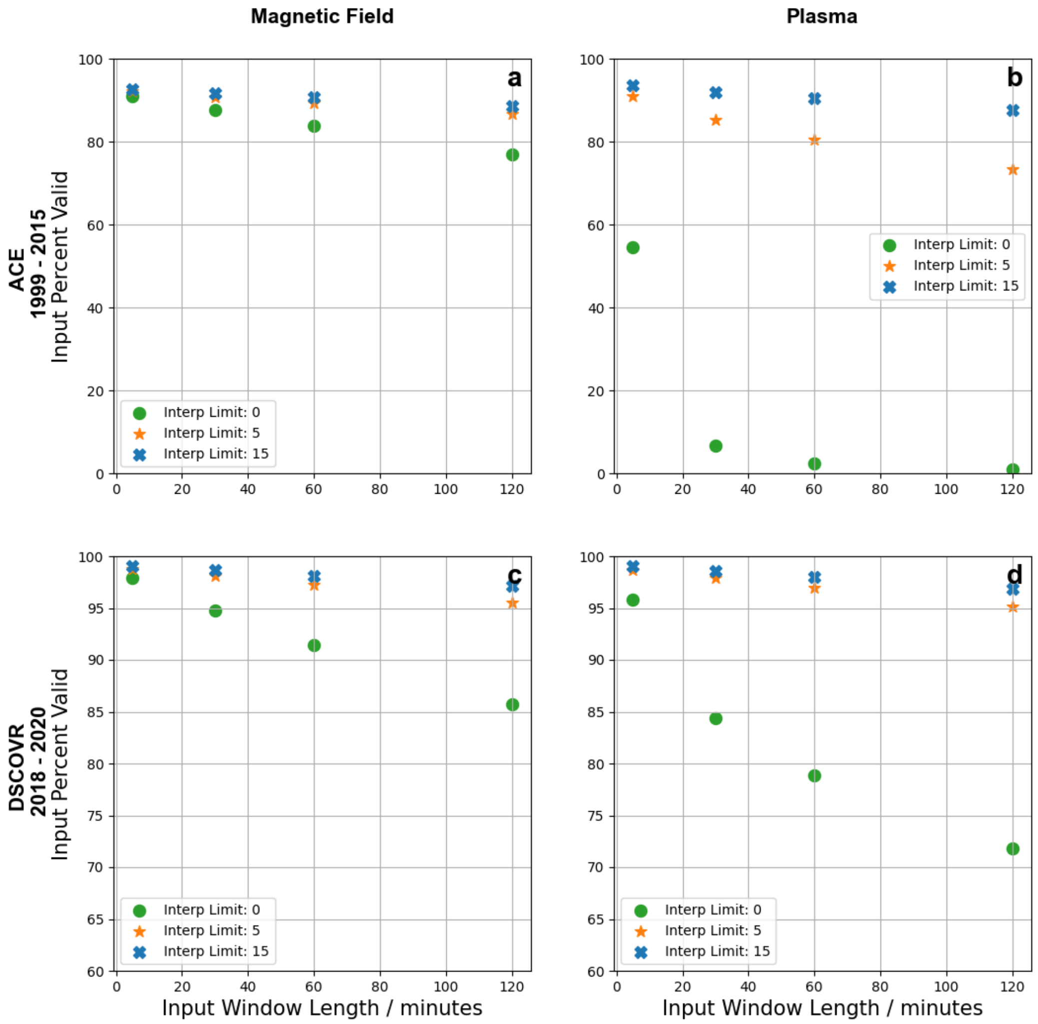

Figure 4. The fraction of data windows that are continuous, without data gaps for the ACE (top: a, b) and DSCOVR (bottom: c, d) NRT data. The results for the magnetic field (left: a, c) and plasma data (right: b, d) are shown. The fraction of complete data windows are provided as a function of input window length required to be continuous. Three different interpolation schemes are presented: no interpolation (green circles), interpolation of gaps 5 minutes or shorter (orange stars) and interpolation of gaps 15 minutes or shorter (blue crosses).

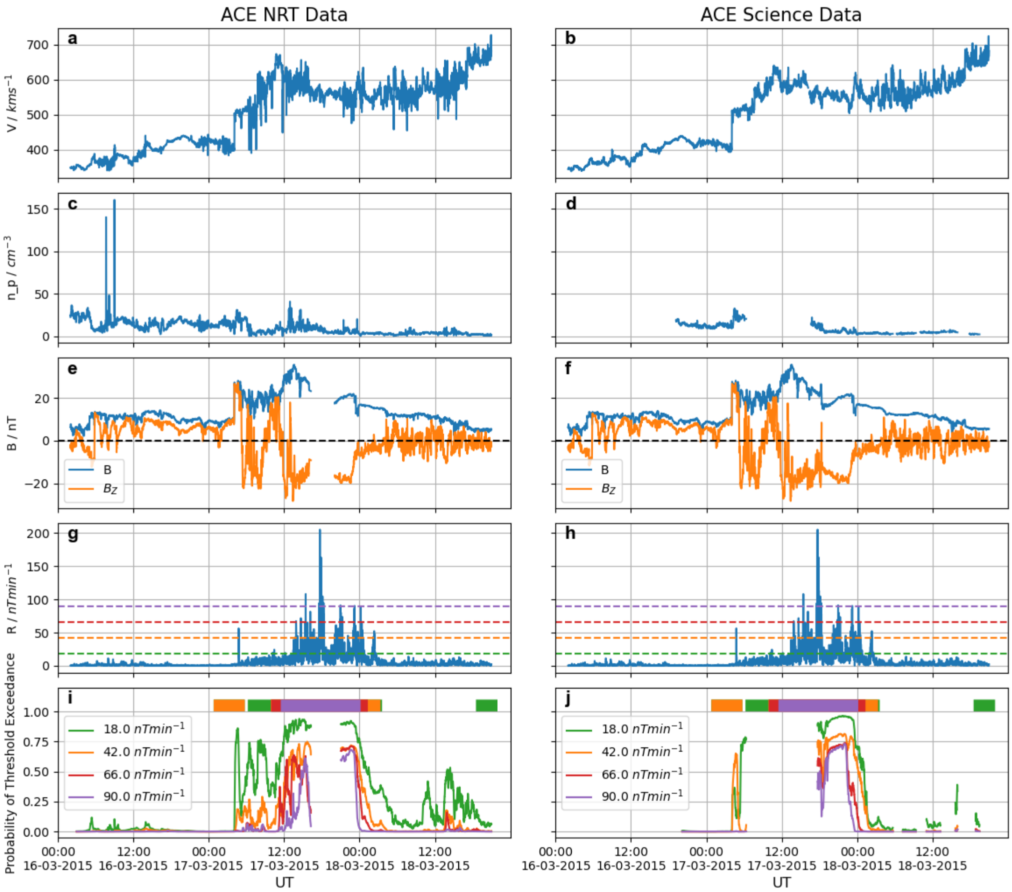

Figure 5. Space weather model forecasts produced during a geomagnetic storm in March 2015. ACE NRT data and the relevant model are shown on the left (a, c, e, g, i), while the scientific ACE data and relevant model are on the right (b, d, f, h, j). The solar wind velocity, density, magnetic field strength and B_Z^{GSM} are shown in the top three panels (a, b, c, d, e, f), while the rate of change of the horizontal ground magnetic field observed at LER (R) is shown in the fourth panel (g, h). Four model variants are shown, forecasting whether four thresholds of R will be exceeded (horizontal lines in panels g, h). The bottom panels (i, j) show the model predictions, while the horizontal bars indicate the perfect forecast and the 'maximum' threshold exceeded.

ence density data that are not present in the NRT data (e.g. a few hours around 1200 UT on the 17th March), suggesting that this data was later removed. On the other hand, there is a data gap in the NRT magnetic field data that appears to have been filled with further post-processing (e.g. around 1800 UT on the 17th March). Overall, there is an extended period during this storm when both the NRT and science-based models are unable to produce a forecast, corresponding to a gap in the NRT magnetic field data and the scientific density data, emphasizing the operational challenges in providing space weather forecasts. This appears to be precisely the kind of interval when predictions of large ground magnetic field variability would be desirable. Nonetheless, both models show increased probabilities during the storm, qualitatively reflecting the ground truth.

Next, we show a second example geomagnetic storm (Figure 6), this time on the 22nd/23rd June 2015, which was notably observed to have adverse impacts on a mid-

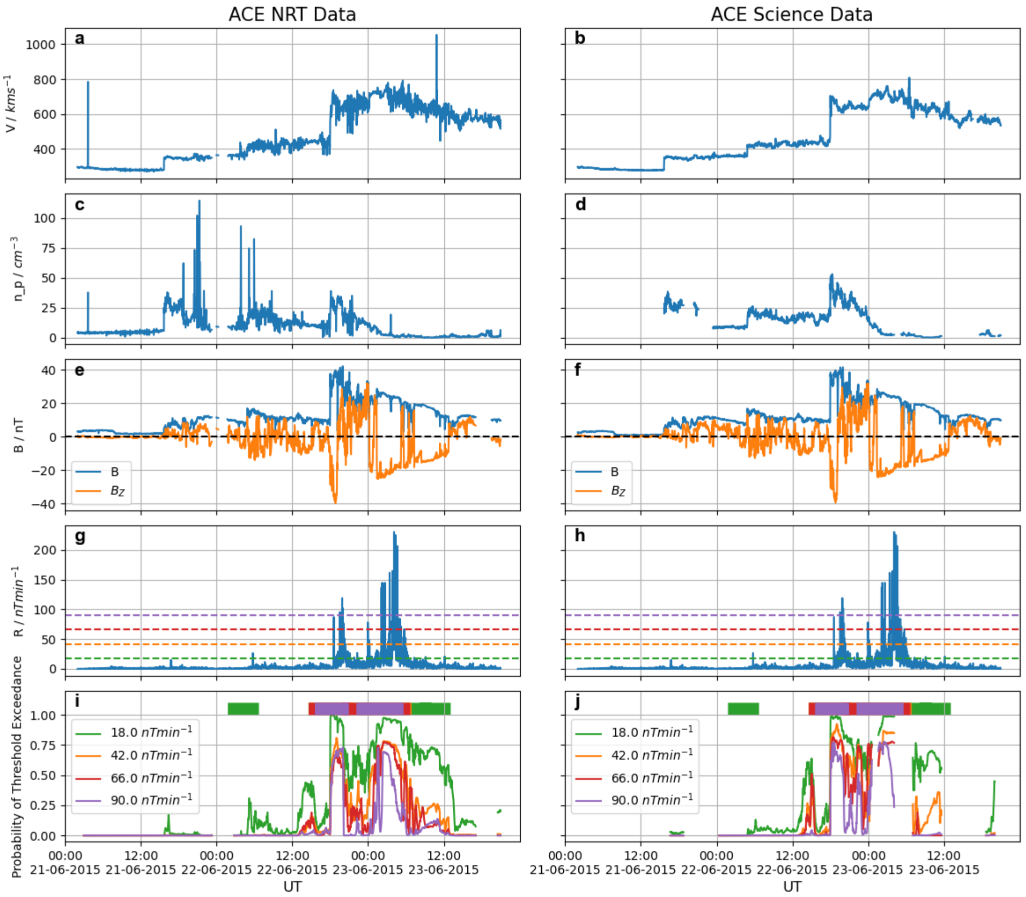

Figure 6. Space weather model forecasts produced during a geomagnetic storm in June 2015, with the same format as Figure 5.

latitude railway system (Liu et al., 2016). This storm offers opportunities to further explore several similarities and differences between the scientific and NRT driven models. The increased short-term variability present in the NRT data is once again apparent, and significant intervals of the density data on the 21st June are not present. Inspecting the NRT solar wind velocity data (Figure 6a) we see two anomalous spikes; at 0400 UT on the 21st June and at 1000 UT on the 23rd June. Naively in the NRT data these may initially resemble very large solar wind shocks, to which space weather models may respond. However, we see that the NRT model does not report increased probabilities around these intervals, presumably as the model has seen this anomalous behavior in the training data and only responds to more sustained changes in the solar wind that are corroborated by changes in other solar wind parameters.

Regarding the predicted model probabilities, from single case studies we cannot quantitatively assess the performance of the models, however we can see that during the period of disturbed ground magnetic field activity during the storm that the models are reporting elevated probabilities, as would be desirable, and the NRT and science-based model results qualitatively appear similar. We may compare the metrics returned by the models when applied to the "unseen" NRT and science data between 2015 and 2016. Figure 7 shows three metrics as a function of threshold of the ground magnetic field: the receiver-operator characteristic (ROC), precision-recall (PR) score and Brier skill score (BSS). In terms of ROC score, the scientific and NRT models perform very similarly, while

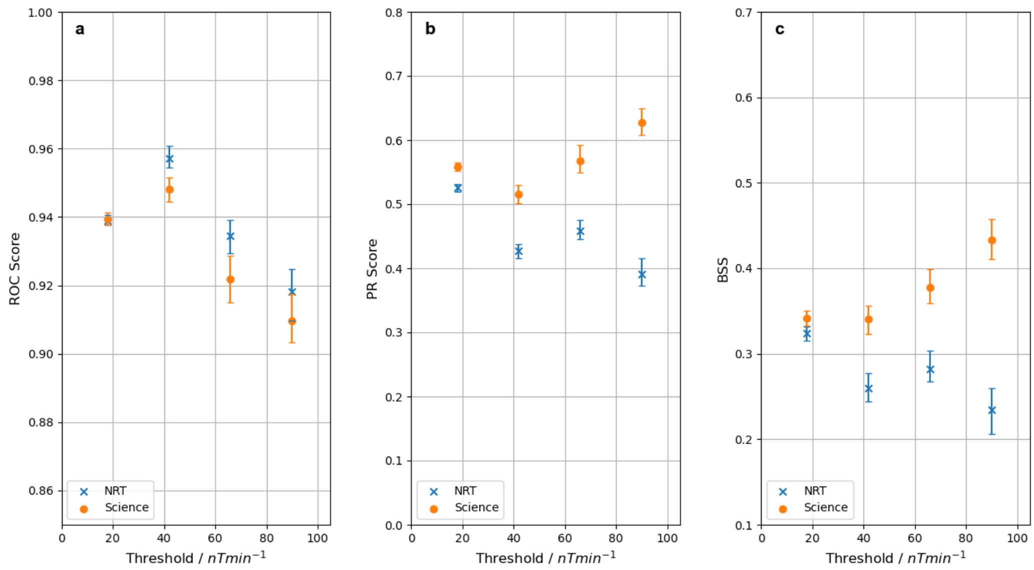

Figure 7. Metrics returned by the NRT and Science models when applied to the test data set (2015 - 2016) as a function of threshold of ground magnetic field variability. The receiver-operator characteristic (a), precision-recall (b), and Brier skill score (c) metrics are shown. An input of 60 minutes of solar wind data are used, along with a forecast horizon of 180 minutes (as in Figures 5 and 6). The error bars represent the 95% confidence intervals calculated from 100 iterations of a bootstrap method (with replacement).

in terms of PR and BSS the model based upon the scientific data performs slightly better. We note that due to the presence of different data gaps the metrics are not precisely comparable, and this will be most evident at the highest thresholds which are less frequently exceeded. Nonetheless, the NRT-based model is achieving at least comparable performance to the model employing the science quality data.

5 Discussion

The magnetosphere reacts to the impinging solar wind on a range of timescales, and the results are heavily dependent upon the history of the coupled system. For example, even the response to fast changes in solar wind dynamic pressure (e.g. Yue et al., 2010; Shinbori et al., 2012; Oliveira & Raeder, 2015; A. W. Smith, Forsyth, Rae, Rodger, & Freeman, 2021) may be modulated by the magnetospheric state that was induced in the preceding intervals (Zong et al., 2021). Further, characteristic timescales for the response of the currents linking the magnetosphere and ionosphere, and consequent ground magnetic field variability, range from 10s of minutes to several hours (Coxon et al., 2019; Shore et al., 2019). Magnetospheric physics therefore dictates that space weather models need to interpret - in some fashion - the history of the coupled solar wind magnetosphere system. There are two main methods in which models can infer the historical behavior: first, they can use statistical properties of a period of data, such as the variability or range of a solar wind property (e.g. Wintoft et al., 2017; Shprits et al., 2019; Zhelavskaya et al., 2019; Camporeale et al., 2020; A. W. Smith et al., 2020). Second, models can process entire time intervals of data and extract the necessary information themselves (e.g. Kunduri et al., 2020; Keesee et al., 2020; A. W. Smith, Forsyth, Rae, Garton, et al., 2021). Discontinuities and unreliable values in the NRT solar wind data thus pose potential risks and limitations to the validity and utility of these models.

Table 1. The percentage of magnetic field data with an inconsistent sign in the scientific and NRT data sets.

Magnetic Field Component	ACE	DSCOVR
B_X^{GSM}	6.9%	1.8%
B_Y^{GSM}	7.4%	1.9%
B_Z^{GSM}	12.7%	3.1%

5.1 Data Validity

Figures 1 and 2 clearly showed that there are often differences between the reported NRT data and the values that are later found in the post-processed science data. We also showed that the velocity is the most reliable plasma moment returned by the NRT data, compared to the density and temperature. Regarding the magnetic field, while there is considerable spread in the reported magnetic field components, the total magnetic field magnitude is mostly within 10% of the reported NRT value. Meanwhile, we note the limited interval for which DSCOVR data were available has constrained the solar wind conditions to which DSCOVR has been exposed. On inspection, we showed examples of two types of "error" that are present in the NRT data: first, the NRT data tend to exhibit short time scale variability that can be resolved with post-processing; second, there are occasionally very transient anomalous values present in the data (e.g. spikes, Figure 6).

These factors have strong implications for space weather model design and development. The differences between the science and NRT data suggest that - as is best practice - models should be trained upon the data type that will be used in the future. For example, the larger short term variability of the NRT data will need to be found in both the training and evaluation data. We showed that in Figures 5 and 6, the selected model was able to compensate for the additional short-timescale variability in the NRT data. Additionally, artifacts in the data (such as the anomalous spikes in Figure 6) may also be tolerated by some models. However, this is a particular problem if a statistical parameter such as the range of the solar wind velocity is evaluated (e.g. Zhelavskaya et al., 2019; A. W. Smith et al., 2020). This may require simple additional processing of the NRT data, e.g. through smoothing or averaging, prior to its use by a forecasting model.

Space weather models may employ solar wind-magnetosphere coupling functions (e.g. Newell et al., 2007; Milan et al., 2012) in order to assess the geoeffectiveness of the solar wind (e.g. Tan et al., 2018). We note that the selection of coupling function requires careful consideration of the magnetospheric prediction to be made (Lockwood & McWilliams, 2021; Lockwood, 2022). However, one commonality between the various coupling functions is their reliance on the orientation of the interplanetary magnetic field, and therefore the relative sign and value of B_Y^{GSM} and B_Z^{GSM}, in particular. As we have shown in Figure 2, particularly with the ACE NRT data, the reliability of these components must be considered with care. Of greatest concern are instances showing evidence of B_Y^{GSM} and B_Z^{GSM} changing sign upon further processing/calibration (e.g. data in the upper left or lower right quadrants of Figure 2). Table 1 shows the percentage of magnetic field data that has an inconsistent sign between the scientific and NRT data sets. We can see that the fractions are quite considerable, and that - as in the previous evaluations - the DSCOVR data are more consistent than that from ACE. The reliability (and short term variability) in other NRT solar wind parameters, such as the density, should also be considered. These would result in large fluctuations of the coupling functions that would not be present in the post-processed science data.

5.2 Data Continuity

To provide a useful and practical NRT forecast of geomagnetic conditions it is important to maximize the length of time for which it is available. Many magnetospheric models require several hours of historical measurements order to provide a result (e.g. Wintoft et al., 2015; Bortnik et al., 2016; Keesee et al., 2020; A. W. Smith, Forsyth, Rae, Garton, et al., 2021; McGranaghan et al., 2020). The implementation of such models requires careful design, given the nature of the input data available. Additionally, many space weather models will also wish to maximize the data available to allow the largest possible training data sets.

In this work we have shown that data gaps are more numerous in the NRT data sourced from ACE than they are in the data from DSCOVR. Therefore, if only a limited quantity of training data is required then the DSCOVR data would appear to be a good choice. However, the ACE NRT data covers a much longer period of time. This is important as heliospheric conditions and space weather vary over the solar cycle (Luhmann et al., 2002; Chapman et al., 2018, 2020) and between cycles (Lockwood et al., 2014; Hajra et al., 2021; Reyes et al., 2021), while rare, extreme events are the most concerning (Thomson et al., 2011; Kilpua et al., 2015; Vennerstrom et al., 2016; Wintoft et al., 2016; Owens et al., 2021). For this reason, the longer time-span ACE NRT data is an appealing training data set despite its limitations.

We also showed that short data gaps are the most common, for both ACE and DSCOVR NRT data. These frequent but short data gaps significantly impair the continuity of the data. For example, if an hour of continuous data is required as a model input then only 3% of ACE NRT plasma data is valid (Figure 4b). A simple solution is to interpolate over short data gaps. Interpolating over gaps of five minutes or less significantly increases the portion of valid ACE NRT plasma data to 80% (if an hour of continuous data is required). Previous forecasting models, using the somewhat analogous one-minute resolution OMNI data, have applied simple linear interpolations over gaps of 10 (Keesee et al., 2020) or 15 minutes (Wintoft et al., 2015; A. W. Smith, Forsyth, Rae, Garton, et al., 2021) to good effect. Figure 4 also shows that the data gaps are far more prevalent in the plasma data, and so models have been developed that only use the magnetic field data (e.g. Wintoft et al., 2015). Such models may give slightly impaired performance, but they would allow forecasts to be produced during times when the more complete models would otherwise be unavailable. Additionally, techniques such as sub-sampling of the available data (e.g. McGranaghan et al., 2020) may be implemented in such a way as to reduce the effects of short data gaps.

While space weather forecasts have often employed simple (e.g. linear) interpolation methods, reconstructions of the solar wind data using statistical methods have also shown great promise (Kondrashov et al., 2014). A “pattern matching” approach could also be employed to identify historical analogous intervals to provide a surrogate input (c.f. Haines et al., 2021). Meanwhile, longer time-scale reconstruction of solar wind data has also been performed using ground based indices (Machol et al., 2013; Kataoka & Nakano, 2021) and magnetosheath data (Nabert et al., 2015). However, while ground magnetometer data may be available in near-real time, filling NRT data gaps with such reconstructions may be challenging in an operational forecast.

If interpolation is employed then it is useful to consider the solar wind autocorrelation timescales, in order to maximize the data available while minimizing the impact of the interpolation on the quality of the input data. The autocorrelation describes how similar a time series is with a lagged version of itself, and is evaluated for a series of different time offsets, or lags (Appendix A). Figures 8a and 8b show the autocorrelation and partial autocorrelation of four selected NRT solar wind parameters. We will assess these qualitatively. The precise autocorrelations depend on solar wind conditions, but such an analysis is beyond the scope of this work. The autocorrelations have been calculated for the longest continuous intervals of DSCOVR NRT data, though we note that the results do not change significantly based upon the time period selected (so long as it is of sufficient duration and includes both dynamic and quiescent solar wind).

The solar wind properties tested all show significant autocorrelations beyond the 60 minutes shown in Figure 8I, though we do see that the autocorrelation of B_Z^{GSM} has nearly reduced to the 95% confidence bound by 60 minutes. This suggests that - in a broad statistical sense - gaps shorter than 60 minutes could be reasonably interpolated. However, this completely neglects the importance of short-timescale variability in the coupling of the solar wind and magnetosphere, for example rapid changes in parameters found during phenomena such as interplanetary shocks. Therefore, while quiescent solar wind data may undoubtedly be interpolated over relatively large intervals, such large interpolation is unadvised during the active intervals of most importance to space weather models.

An additional complication is created if the data gap is present in the most recently acquired NRT data. Consequently there will be no opportunity to interpolate (e.g. linearly) as one side of the data gap is unknown. In this case techniques such as repeating the last recorded value can be performed, and we should consider the partial autocorrelation: it ignores the data within the lag period (i.e. it treats the lag period as a gap). For the solar wind velocity and density (Figure 8a II and b II) the partial autocorrelations diminish rapidly, and are within the 95% confidence intervals by around 15 and 5 minutes, respectively. Meanwhile, for the solar wind magnetic field magnitude and B_Z^{GSM} (Figure 8c II and d II) the significant partial autocorrelation timescales are short at around 8 and 5 minutes, respectively. From this we can conclude that - again a statistical sense - repeating the last recorded value for approximately five minutes will not hugely impact the data quality, while performing this for 15 minutes or more will begin to become problematic, particularly in active geomagnetic conditions.

It would be possible to use this kind of analysis to set variable interpolation limits separately for the magnetic field and plasma data, if models only required one type of data (c.f. Wintoft et al., 2015). We emphasize that these are statistical results on a long time interval of continuous data, and therefore likely not representative of more extreme or highly transient solar wind conditions under which data gaps may be more likely (for example due to instrument saturation effects). Additionally, these autocorrelations have been calculated from the NRT data which, given the short-time scale variability we have shown above, may cause this to underestimate the "true" solar wind autocorrelation timescales. Nonetheless, these provide a useful estimate of approximate interpolation timescales for the NRT data.

Generally, linear interpolations have been employed in the past (e.g. Wintoft et al., 2015; A. W. Smith, Forsyth, Rae, Garton, et al., 2021), however in the future more complex interpolation methods may give more confidence to interpolating over larger data gaps, for example using similar historical analogues (Haines et al., 2021), or the use of auto-regressive models, given the high level of autocorrelation observed. Nonetheless, future space weather missions should look to minimize NRT data gaps, those due to instrument saturation effects for example (e.g. Nicolaou et al., 2020).

6 Summary and Conclusions

Space weather hazards pose a severe threat to infrastructure such as electricity networks. Recognizing this threat, models have been developed to forecast magnetospheric parameters, such as indices (Tan et al., 2018; Tasistro-Hart et al., 2021; Chakraborty & Morley, 2020) or direct space weather consequences such as auroral-related particle precipitation (McGranaghan et al., 2020), radiation belt enhancements (Forsyth et al., 2020), or ground magnetic perturbations (Wintoft et al., 2015; Keesee et al., 2020; A. W. Smith,

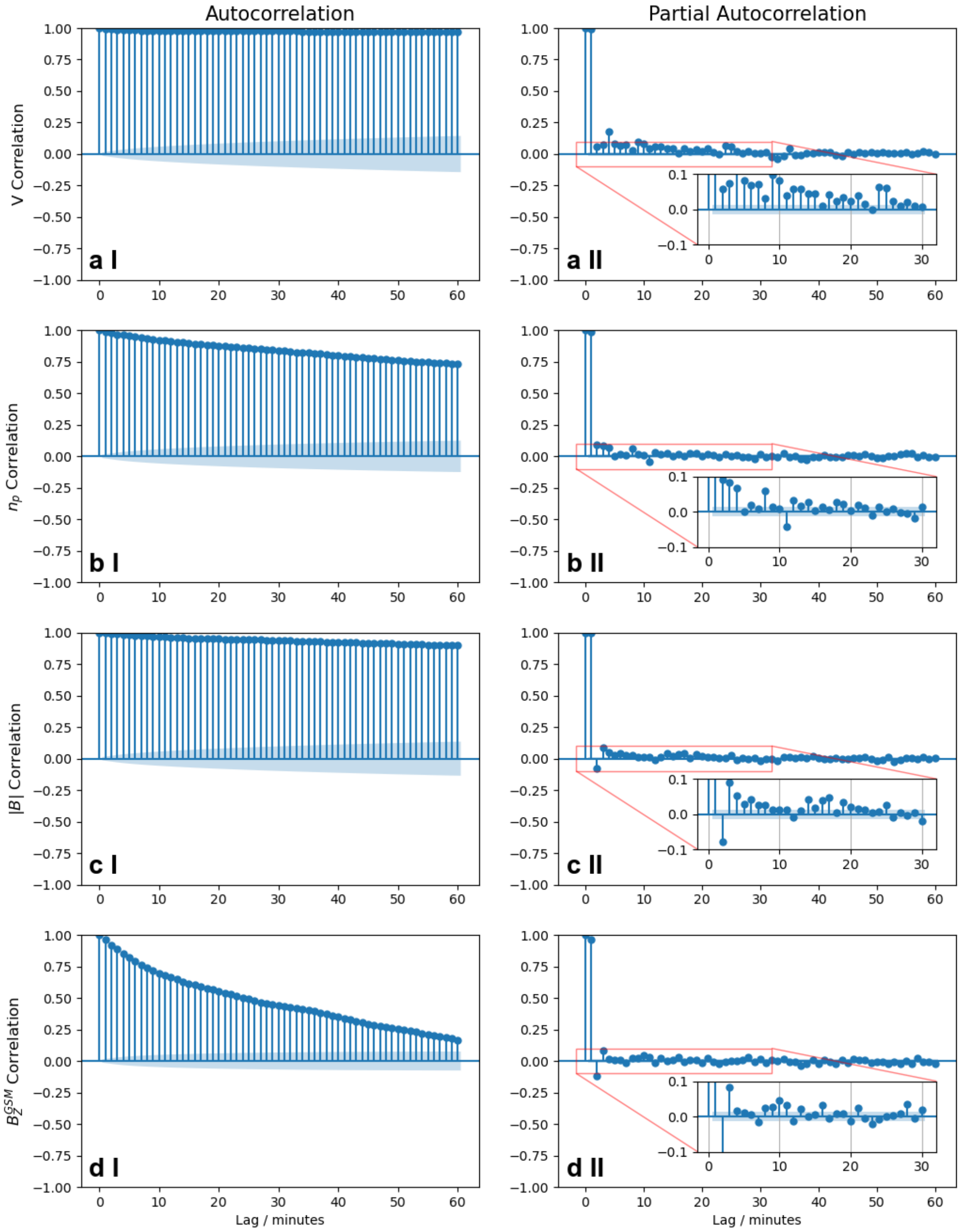

Figure 8. The autocorrelation (left, I) and partial autocorrelation (right, II) of four NRT solar wind measurements made by DSCOVR. The autocorrelations for the solar wind velocity (V, a), proton density (n_p, b), magnetic field magnitude ($|B|$, c) and B_Z^{GSM} component of the magnetic field (d) are presented. The shaded blue regions represent the 95% confidence bound within which points exhibit no significant autocorrelation. The autocorrelations were evaluated during the longest periods of continuous data present between 2018 and 2020 (inclusive). For the plasma properties (a, b) the autocorrelations were calculated on 350 hours of data between 01/05/18 and 16/05/18, while for the magnetic field properties (c, d) the autocorrelations were evaluated for 370 hours of data between 29/05/19 and 13/06/19.

Forsyth, Rae, Garton, et al., 2021). Such space weather models are most often trained and evaluated on post-processed scientific quality data from spacecraft upstream of the Earth at the L1 point. In this work we have assessed the validity and continuity of the Near-Real-Time (NRT) data that is available much more rapidly, and showed an example of a forecasting model using such data.

When the NRT data are compared to the post-processed, science quality data we found that the NRT data are subject to greater short-term variability and occasional anomalous values. Nonetheless, the solar wind velocity is mostly accurate to within ±10%. While the NRT solar wind density and temperature are generally close to the more processed values, they are also subject to greater uncertainty. Regarding the magnetic field, the total field magnitude is generally reported to within 10% of its processed value. Meanwhile, there is a greater spread in the individual components of the field, including some occasions when the sign of the B_Y^{GSM} and B_Z^{GSM} is later corrected. Some models may be able to compensate for the differences between the science quality and NRT data (if they are trained on the appropriate data source), while others may require additional data pre-processing.

When considering an operational space weather model it is important to maximize its availability. To this end we investigated the occurrence and lengths of data gaps that could impede an operational model. We found that short data gaps are frequently found in the NRT data, and are more numerous in the plasma data (compared to the magnetic field data). Additionally, DSCOVR NRT data shows gaps less frequently than the NRT data from ACE. Nonetheless, short interpolations (e.g. over gaps of five minutes or less) can dramatically increase data and hence model availability, even when long continuous windows are required.

Appendix A Autocorrelation Analysis

In Figure 8 we evaluate the autocorrelation functions of several solar wind parameters. Practically, we employ the python statsmodels package. Theoretically, if we let x_1, x_2, ... x_n be the observations of a sample of the time series, we take the mean of the time series as:

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \tag{A1}$$

The sample auto-covariance function for lag h ($h < n$) is then defined as:

$$\hat{\gamma}(h) = n^{-1}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x}) \tag{A2}$$

We then take the sample autocorrelation function for lag h as:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \tag{A3}$$

Meanwhile, the sample partial autocorrelation is the sample autocorrelation between x_t and x_{t+h} but with the linear dependence of x_t on x_{t+1} through x_{t+h-1} removed. For more details the interested reader is directed to Brockwell and Davis (2002).

Acknowledgments

We acknowledge and thank the ACE and DSCOVR teams for the solar wind data and NASA GSFC's Space Physics Data Facility's CDAWeb service for data availability.

Information on the NRT solar wind data may be found at https://www.swpc.noaa.gov/products/real-time-solar-wind, while we use data provided by SWPC but archived by the BGS (British Geological Survey) which may be found at the UK NGDC `https://webapps.bgs.ac.uk/services/ngdc/accessions/index.html#item172549`.

Scientific DSCOVR data may be accessed at https://cdaweb.gsfc.nasa.gov/pub/data/dscovr/, while the scientific ACE data may be accessed at https://cdaweb.gsfc.nasa.gov/pub/data/ace/.

The authors would like to thank Howard Singer for helpful discussions and suggestions.

AWS, CF and IJR were supported by NERC grants NE/P017150/1 and NE/V002724/1. CF was also supported by the NERC Independent Research Fellowship NE/N014480/1. JPE and MJH were supported by NERC grant NE/V003070/1. CMJ was supported by the Science Foundation Ireland (SFI) Grant 18/FRL/6199. AWPT, CDB and GSR are supported by NERC award NE/V002694/1 (SWIMMR Activities in Ground Effects, SAGE).

The analysis in this paper was performed using python, including the TensorFlow (Abadi et al., 2015), pandas (McKinney, 2010), numpy (Van Der Walt et al., 2011), scikit-learn (Pedregosa et al., 2011), scipy (Virtanen et al., 2020) and matplotlib (Hunter, 2007) libraries.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Research, G. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.. Retrieved from `www.tensorflow.org.`

Aellig, M. R., Lazarus, A. J., Kasper, J. C., & Ogilvie, K. W. (2001). Rapid Measurements of Solar Wind Ions with the Triana PlasMag Faraday Cup. *Astrophysics and Space Science*, *277*(1/2), 305–307. Retrieved from `http://link.springer.com/10.1023/A:1012229729242` doi: 10.1023/A:1012229729242

Astafyeva, E., Zakharenkova, I., & Förster, M. (2015, oct). Ionospheric response to the 2015 St. Patrick's Day storm: A global multi-instrumental overview. *Journal of Geophysical Research: Space Physics*, *120*(10), 9023–9037. Retrieved from `http://doi.wiley.com/10.1002/2015JA021629` doi: 10.1002/2015JA021629

Baker, D., Belian, R., Higbie, P., Klebesadel, R., & Blake, J. (1987, oct). Deep dielectric charging effects due to high-energy electrons in earth's outer magnetosphere. *Journal of Electrostatics*, *20*(1), 3–19. Retrieved from `https://www.sciencedirect.com/science/article/abs/pii/0304388687900829?via%3Dihub` doi: 10.1016/0304-3886(87)90082-9

Baker, D. N., McPherron, R. L., Cayton, T. E., & Klebesadel, R. W. (1990, sep). Linear prediction filter analysis of relativistic electron properties at 6.6 ¡i¿ R ¡sub¿E¡/sub¿ ¡/i¿. *Journal of Geophysical Research*, *95*(A9), 15133. Retrieved from `http://doi.wiley.com/10.1029/JA095iA09p15133` doi: 10.1029/JA095iA09p15133

Baumann, C., & McCloskey, A. E. (2021, aug). Timing of the solar wind propagation delay between L1 and Earth based on machine learning. *Journal of Space Weather and Space Climate*, *11*, 41. Retrieved from `https://www.swsc-journal.org/10.1051/swsc/2021026` doi: 10.1051/swsc/2021026

Beggan, C. D., Beamish, D., Richards, A., Kelly, G. S., & Alan, A. W. (2013, jul). Prediction of extreme geomagnetically induced currents in the UK high-voltage network. *Space Weather*, *11* (7), 407–419. Retrieved from https://onlinelibrary.wiley.com/doi/10.1002/swe.20065 doi: 10.1002/swe.20065

Blake, S. P., Gallagher, P. T., Campanyà, J., Hogg, C., Beggan, C. D., Thomson, A. W., ... Bell, D. (2018). A Detailed Model of the Irish High Voltage Power Network for Simulating GICs. *Space Weather*, *16* (11), 1770–1783. Retrieved from https://doi.org/10.1029/2018SW001926 doi: 10.1029/2018SW001926

Blake, S. P., Gallagher, P. T., McCauley, J., Jones, A. G., Hogg, C., Campanyà, J., ... Bell, D. (2016, dec). Geomagnetically induced currents in the Irish power network during geomagnetic storms. *Space Weather*, *14* (12), 1136–1154. Retrieved from http://doi.wiley.com/10.1002/2016SW001534 doi: 10.1002/2016SW001534

Blandin, M., Connor, H. K., Öztürk, D. S., Keesee, A. M., Pinto, V., Mahmud, M. S., ... Priyadarshi, S. (2022, may). Multi-Variate LSTM Prediction of Alaska Magnetometer Chain Utilizing a Coupled Model Approach. *Frontiers in Astronomy and Space Sciences*, *9*, 80. Retrieved from https://www.frontiersin.org/articles/10.3389/fspas.2022.846291/full doi: 10.3389/fspas.2022.846291

Bolduc, L. (2002, nov). GIC observations and studies in the Hydro-Québec power system. *Journal of Atmospheric and Solar-Terrestrial Physics*, *64* (16), 1793–1802. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1364682602001281 doi: 10.1016/S1364-6826(02)00128-1

Borovsky, J. E. (2020, aug). What magnetospheric and ionospheric researchers should know about the solar wind. *Journal of Atmospheric and Solar-Terrestrial Physics*, *204*, 105271. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S1364682620300882?via%3Dihub doi: 10.1016/J.JASTP.2020.105271

Bortnik, J., Li, W., Thorne, R. M., & Angelopoulos, V. (2016, mar). A unified approach to inner magnetospheric state prediction. *Journal of Geophysical Research A: Space Physics*, *121* (3), 2423–2430. Retrieved from http://doi.wiley.com/10.1002/2015JA021733 doi: 10.1002/2015JA021733

Boteler, D. H. (2021, nov). Modeling Geomagnetic Interference on Railway Signaling Track Circuits. *Space Weather*, *19* (1). Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002609 doi: 10.1029/2020SW002609

Boteler, D. H., Pirjola, R. J., & Nevanlinna, H. (1998, jan). The effects of geomagnetic disturbances on electrical systems at the Earth's surface. *Advances in Space Research*, *22* (1), 17–27. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S027311779701096X doi: 10.1016/S0273-1177(97)01096-X

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to Time Series and Forecasting - Second Edition.* Retrieved from http://home.iitj.ac.in/$\sim$parmod/document/introductiontimeseries.pdfhttp://books.google.com/books?id=9tv0taI8l6YC

Campbell, W. H. (1980, may). Observation of electric currents in the Alaska oil pipeline resulting from auroral electrojet current sources. *Geophysical Journal International*, *61* (2), 437–449. Retrieved from https://academic.oup.com/gji/article-lookup/doi/10.1111/j.1365-246X.1980.tb04325.x doi: 10.1111/j.1365-246X.1980.tb04325.x

Camporeale, E., Cash, M. D., Singer, H. J., Balch, C. C., Huang, Z., & Toth, G. (2020, oct). A gray-box model for a probabilistic estimate of regional ground magnetic perturbations: Enhancing the NOAA operational Geospace model with machine learning. *Journal of Geophysical Research: Space Physics*. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/

2019JA027684 doi: 10.1029/2019JA027684

Carter, B. A., Yizengaw, E., Pradipta, R., Weygand, J. M., Piersanti, M., Pulkkinen, A., . . . Zhang, K. (2016, oct). Geomagnetically induced currents around the world during the 17 March 2015 storm. *Journal of Geophysical Research: Space Physics*, *121*(10), 10,496–10,507. Retrieved from http://doi.wiley.com/10.1002/2016JA023344 doi: 10.1002/2016JA023344

Cash, M. D., Witters Hicks, S., Biesecker, D. A., Reinard, A. A., de Koning, C. A., & Weimer, D. R. (2016, feb). Validation of an operational product to determine L1 to Earth propagation time delays. *Space Weather*, *14*(2), 93–112. Retrieved from http://doi.wiley.com/10.1002/2015SW001321 doi: 10.1002/2015SW001321

Chakraborty, S., & Morley, S. K. (2020, jul). Probabilistic prediction of geomagnetic storms and the ¡i¿ K ¡sub¿p¡/sub¿ ¡/i¿ index. *Journal of Space Weather and Space Climate*, *10*, 36. Retrieved from https://www.swsc-journal.org/10.1051/swsc/2020037 doi: 10.1051/swsc/2020037

Chapman, S. C., McIntosh, S. W., Leamon, R. J., & Watkins, N. W. (2020, jun). Quantifying the Solar Cycle Modulation of Extreme Space Weather. *Geophysical Research Letters*, *47*(11), e2020GL087795. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020GL087795 doi: 10.1029/2020GL087795

Chapman, S. C., Watkins, N. W., & Tindale, E. (2018, aug). Reproducible Aspects of the Climate of Space Weather Over the Last Five Solar Cycles. *Space Weather*, *16*(8), 1128–1142. Retrieved from http://doi.wiley.com/10.1029/2018SW001884 doi: 10.1029/2018SW001884

Chu, X., Ma, D., Bortnik, J., Tobiska, W. K., Cruz, A., Bouwer, S. D., . . . Reeves, G. (2021, dec). Relativistic Electron Model in the Outer Radiation Belt Using a Neural Network Approach. *Space Weather*, *19*(12), e2021SW002808. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002808 doi: 10.1029/2021SW002808

Coxon, J. C., Shore, R. M., Freeman, M. P., Fear, R. C., Browett, S. D., Smith, A. W., . . . Anderson, B. J. (2019, jul). Timescales of Birkeland Currents Driven by the IMF. *Geophysical Research Letters*, *46*(14), 7893–7901. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2018GL081658 doi: 10.1029/2018GL081658

Dimmock, A. P., Rosenqvist, L., Hall, J. O., Viljanen, A., Yordanova, E., Honkonen, I., . . . Sjöberg, E. C. (2019, jun). The GIC and Geomagnetic Response Over Fennoscandia to the 7–8 September 2017 Geomagnetic Storm. *Space Weather*, *17*(7), 989–1010. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018SW002132 doi: 10.1029/2018SW002132

Dimmock, A. P., Welling, D. T., Rosenqvist, L., Forsyth, C., Freeman, M. P., Rae, I. J., . . . Yordanova, E. (2021, apr). Modeling the Geomagnetic Response to the September 2017 Space Weather Event Over Fennoscandia Using the Space Weather Modeling Framework: Studying the Impacts of Spatial Resolution. *Space Weather*, *19*(5), e2020SW002683. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002683 doi: 10.1029/2020SW002683

Divett, T., Mac Manus, D. H., Richardson, G. S., Beggan, C. D., Rodger, C. J., Ingham, M., . . . Obana, Y. (2020, jun). Geomagnetically Induced Current Model Validation From New Zealand's South Island. *Space Weather*, *18*(8). Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020SW002494 doi: 10.1029/2020SW002494

Divett, T., Richardson, G. S., Beggan, C. D., Rodger, C. J., Boteler, D. H., Ingham, M., . . . Dalzell, M. (2018, jun). Transformer-Level Modeling of Geomagnetically Induced Currents in New Zealand's South Island. *Space Weather*, *16*(6), 718–735. Retrieved from http://doi.wiley.com/10.1029/2018SW001814 doi: 10.1029/2018SW001814

Eastwood, J. P., Hapgood, M. A., Biffis, E., Benedetti, D., Bisi, M. M., Green, L., . . . Burnett, C. (2018, dec). Quantifying the Economic Value of Space Weather Forecasting for Power Grids: An Exploratory Study. *Space Weather*, *16*(12), 2052–2067. Retrieved from http://doi.wiley.com/10.1029/2018SW002003 doi: 10.1029/2018SW002003

Forsyth, C., Watt, C. E., Mooney, M. K., Rae, I. J., Walton, S. D., & Horne, R. B. (2020, jun). Forecasting GOES 15 ¿2 MeV Electron Fluxes From Solar Wind Data and Geomagnetic Indices. *Space Weather*, *18*(8). Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019SW002416 doi: 10.1029/2019SW002416

Freeman, M. P., Forsyth, C., & Rae, I. J. (2019, may). The influence of substorms on extreme rates of change of the surface horizontal magnetic field in the U.K. *Space Weather*, 2018SW002148. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018SW002148 doi: 10.1029/2018SW002148

Gaunt, C. T., & Coetzee, G. (2007). Transformer failures in regions incorrectly considered to have low GIC-risk. In *2007 ieee lausanne powertech, proceedings* (pp. 807–812). doi: 10.1109/PCT.2007.4538419

Gleisner, H., & Lundstedt, H. (2001, may). A neural network-based local model for prediction of geomagnetic disturbances. *Journal of Geophysical Research: Space Physics*, *106*(A5), 8425–8433. Retrieved from http://doi.wiley.com/10.1029/2000JA900142 doi: 10.1029/2000JA900142

Grawe, M. A., & Makela, J. J. (2021, jul). Predictability of Geomagnetically Induced Currents as a Function of Available Magnetic Field Information. *Space Weather*, *19*(8), e2021SW002747. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002747 doi: 10.1029/2021SW002747

Gummow, R., & Eng, P. (2002, nov). GIC effects on pipeline corrosion and corrosion control systems. *Journal of Atmospheric and Solar-Terrestrial Physics*, *64*(16), 1755–1764. Retrieved from https://www.sciencedirect.com/science/article/pii/S1364682602001256?via%3Dihub doi: 10.1016/S1364-6826(02)00125-6

Haines, C., Owens, M. J., Barnard, L., Lockwood, M., Ruffenach, A., Boykin, K., & McGranaghan, R. (2021, may). Forecasting Occurrence and Intensity of Geomagnetic Activity With Pattern-Matching Approaches. *Space Weather*, *19*(6), e2020SW002624. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002624 doi: 10.1029/2020SW002624

Hajra, R., Franco, A., Echer, E., & Bolzan, M. (2021, apr). Long-Term Variations of the Geomagnetic Activity: A Comparison Between the Strong and Weak Solar Activity Cycles and Implications for the Space Climate. *Journal of Geophysical Research: Space Physics*, *126*(4), e2020JA028695. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020JA028695 doi: 10.1029/2020JA028695

Heyns, M. J., Lotz, S. I., & Gaunt, C. T. (2021, nov). Geomagnetic Pulsations Driving Geomagnetically Induced Currents. *Space Weather*, *19*(2). Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002557 doi: 10.1029/2020SW002557

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95. Retrieved from http://ieeexplore.ieee.org/document/4160265/ doi: 10.1109/MCSE.2007.55

Iucci, N., Levitin, A. E., Belov, A. V., Eroshenko, E. A., Ptitsyna, N. G., Villoresi, G., . . . Yanke, V. G. (2005, jan). Space weather conditions and spacecraft anomalies in different orbits. *Space Weather*, *3*(1), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2003SW000056 doi:

10.1029/2003SW000056

Kappenman, J. G., & Albertson, D. (1990, mar). Bracing for the geomagnetic storms. *IEEE Spectrum*, *27*(3), 27–33. Retrieved from http://ieeexplore.ieee.org/document/48847/ doi: 10.1109/6.48847

Kataoka, R., & Nakano, S. (2021, nov). Reconstructing solar wind profiles associated with extreme magnetic storms: A machine learning approach. *Geophysical Research Letters*, e2021GL096275. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021GL096275 doi: 10.1029/2021GL096275

Keesee, A. M., Pinto, V., Coughlan, M., Lennox, C., Mahmud, M. S., & Connor, H. K. (2020, oct). Comparison of Deep Learning Techniques to Model Connections Between Solar Wind and Ground Magnetic Perturbations. *Frontiers in Astronomy and Space Sciences*, *7*, 72. Retrieved from https://www.frontiersin.org/article/10.3389/fspas.2020.550874/full doi: 10.3389/fspas.2020.550874

Kilpua, E. K., Olspert, N., Grigorievskiy, A., Käpylä, M. J., Tanskanen, E. I., Miyahara, H., . . . Liu, Y. D. (2015, jun). Statistical Study of Strong and Extreme Geomagnetic Disturbances and Solar Cycle Characteristics. *Astrophysical Journal*, *806*(2), 272. Retrieved from https://iopscience.iop.org/article/10.1088/0004-637X/806/2/272 doi: 10.1088/0004-637X/806/2/272

Kondrashov, D., Denton, R., Shprits, Y. Y., & Singer, H. J. (2014, apr). Reconstruction of gaps in the past history of solar wind parameters. *Geophysical Research Letters*, *41*(8), 2702–2707. Retrieved from http://doi.wiley.com/10.1002/2014GL059741 doi: 10.1002/2014GL059741

Kozyreva, O. V., Pilipenko, V. A., Belakhovsky, V. B., & Sakharov, Y. A. (2018, dec). Ground geomagnetic field and GIC response to March 17, 2015, storm. *Earth, Planets and Space*, *70*(1), 157. Retrieved from https://earth-planets-space.springeropen.com/articles/10.1186/s40623-018-0933-2 doi: 10.1186/s40623-018-0933-2

Kunduri, B. S., Maimaiti, M., Baker, J. B., Ruohoniemi, J. M., Anderson, B. J., & Vines, S. K. (2020, aug). A Deep Learning-Based Approach for Modeling the Dynamics of AMPERE Birkeland Currents. *Journal of Geophysical Research: Space Physics*, *125*(8), e2020JA027908. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020JA027908 doi: 10.1029/2020JA027908

Liu, L., Ge, X., Zong, W., Zhou, Y., & Liu, M. (2016, oct). Analysis of the monitoring data of geomagnetic storm interference in the electrification system of a high-speed railway. *Space Weather*, *14*(10), 754–763. Retrieved from http://doi.wiley.com/10.1002/2016SW001411 doi: 10.1002/2016SW001411

Lockwood, M. (2022, feb). Solar Wind—Magnetosphere Coupling Functions: Pitfalls, Limitations, and Applications. *Space Weather*, *20*(2), e2021SW002989. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002989 doi: 10.1029/2021SW002989

Lockwood, M., Bentley, S. N., Owens, M. J., Barnard, L. A., Scott, C. J., Watt, C. E., & Allanson, O. (2019, jan). The Development of a Space Climatology: 1. Solar Wind Magnetosphere Coupling as a Function of Timescale and the Effect of Data Gaps. *Space Weather*, *17*(1), 133–156. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018SW001856 doi: 10.1029/2018SW001856

Lockwood, M., & McWilliams, K. A. (2021, nov). On optimum solar wind – magnetosphere coupling functions for transpolar voltage and planetary geomagnetic activity. *Journal of Geophysical Research: Space Physics*, e2021JA029946. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021JA029946 doi: 10.1029/2021JA029946

Lockwood, M., Nevanlinna, H., Vokhmyanin, M., Ponyavin, D., Sokolov, S., Barnard, L., ... Scott, C. J. (2014). Reconstruction of geomagnetic activity and near-Earth interplanetary conditions over the past 167 yrs; Part 3: Improved representation of solar cycle 11. *Annales Geophysicae*, *32*(4), 367–381. Retrieved from www.ann-geophys.net/32/367/2014/ doi: 10.5194/angeo-32-367-2014

Love, J. J., Hayakawa, H., & Cliver, E. W. (2019, aug). Intensity and Impact of the New York Railroad Superstorm of May 1921. *Space Weather*, *17*(8), 1281–1292. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019SW002250 doi: 10.1029/2019SW002250

Luhmann, J. G., Li, Y., Arge, C. N., Gazis, P. R., & Ulrich, R. (2002, aug). Solar cycle changes in coronal holes and space weather cycles. *Journal of Geophysical Research: Space Physics*, *107*(A8), SMP 3–1–SMP 3–12. Retrieved from http://doi.wiley.com/10.1029/2001JA007550 doi: 10.1029/2001JA007550

Mac Manus, D. H., Rodger, C. J., Dalzell, M., Thomson, A. W. P., Clilverd, M. A., Petersen, T., ... Divett, T. (2017, aug). Long-term geomagnetically induced current observations in New Zealand: Earth return corrections and geomagnetic field driver. *Space Weather*, *15*(8), 1020–1038. Retrieved from http://doi.wiley.com/10.1002/2017SW001635 doi: 10.1002/2017SW001635

Mac Manus, D. H., Rodger, C. J., Ingham, M., Clilverd, M. A., Dalzell, M., Divett, T., ... Petersen, T. (2022). Geomagnetically Induced Current Model in New Zealand Across Multiple Disturbances: Validation and Extension to Non-Monitored Transformers. *Space Weather*, *20*(2). doi: 10.1029/2021SW002955

Machol, J. L., Reinard, A. A., Viereck, R. A., & Biesecker, D. A. (2013, jul). Identification and replacement of proton-contaminated real-time ACE solar wind measurements. *Space Weather*, *11*(7), 434–440. Retrieved from http://doi.wiley.com/10.1002/swe.20070 doi: 10.1002/swe.20070

Marshall, R. A., Dalzell, M., Waters, C. L., Goldthorpe, P., & Smith, E. A. (2012, aug). Geomagnetically induced currents in the New Zealand power network. *Space Weather*, *10*(8), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2012SW000806 doi: 10.1029/2012SW000806

McComas, D., Bame, S., Barker, P., Feldman, W., Phillips, J., Riley, P., & Griffee, J. (1998). Solar Wind Electron Proton Alpha Monitor (SWEPAM) for the Advanced Composition Explorer. *Space Science Reviews*, *86*(1/4), 563–612. Retrieved from http://link.springer.com/10.1023/A:1005040232597 doi: 10.1023/A:1005040232597

McGranaghan, R. M., Ziegler, J., Bloch, T., Hatch, S., Camporeale, E., Lynch, K., ... Skone, S. (2020, apr). Next generation particle precipitation: Mesoscale prediction through machine learning (a case study and framework for progress). *Space Weather*, e2020SW002684. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002684 doi: 10.1029/2020SW002684

McKinney, W. (2010). *Data Structures for Statistical Computing in Python.* Retrieved from http://conference.scipy.org/proceedings/scipy2010/mckinney.html

Milan, S. E., Gosling, J. S., & Hubert, B. (2012). Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap. *Journal of Geophysical Research: Space Physics*, *117*(3). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011JA017082 doi: 10.1029/2011JA017082

Nabert, C., Othmer, C., & Glassmeier, K.-H. (2015, dec). Solar wind reconstruction from magnetosheath data using an adjoint approach. *Annales Geophysicae*, *33*(12), 1513–1524. Retrieved from https://angeo.copernicus.org/articles/33/1513/2015/ doi: 10.5194/angeo-33-1513-2015

Newell, P. T., Sotirelis, T., Liou, K., Meng, C.-I., & Rich, F. J. (2007, jan). A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research: Space Physics*, *112*(A1). Retrieved from http://doi.wiley.com/10.1029/2006JA012015 doi: 10.1029/2006JA012015

Nicolaou, G., Wicks, R. T., Rae, I. J., & Kataria, D. O. (2020, dec). Evaluating the Performance of a Plasma Analyzer for a Space Weather Monitor Mission Concept. *Space Weather*, *18*(12), e2020SW002559. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002559 doi: 10.1029/2020SW002559

Oliveira, D. M., & Raeder, J. (2015, jun). Impact angle control of interplanetary shock geoeffectiveness: A statistical study. *Journal of Geophysical Research: Space Physics*, *120*(6), 4313–4323. Retrieved from http://doi.wiley.com/10.1002/2015JA021147 doi: 10.1002/2015JA021147

Oughton, E. J., Hapgood, M., Richardson, G. S., Beggan, C. D., Thomson, A. W., Gibbs, M., … Horne, R. B. (2019, may). A Risk Assessment Framework for the Socioeconomic Impacts of Electricity Transmission Infrastructure Failure Due to Space Weather: An Application to the United Kingdom. *Risk Analysis*, *39*(5), 1022–1043. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13229 doi: 10.1111/risa.13229

Oughton, E. J., Skelton, A., Horne, R. B., Thomson, A. W. P., & Gaunt, C. T. (2017, jan). Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure. *Space Weather*, *15*(1), 65–83. Retrieved from http://doi.wiley.com/10.1002/2016SW001491 doi: 10.1002/2016SW001491

Owens, M. J., Lockwood, M., Barnard, L. A., Scott, C. J., Haines, C., & Macneil, A. (2021, may). Extreme Space-Weather Events and the Solar Cycle. *Solar Physics*, *296*(5), 82. Retrieved from https://link.springer.com/10.1007/s11207-021-01831-3 doi: 10.1007/s11207-021-01831-3

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(Oct), 2825–2830. Retrieved from http://jmlr.org/papers/v12/pedregosa11a.html

Pinto, V. A., Keesee, A. M., Coughlan, M., Mukundan, R., Johnson, J. W., Ngwira, C. M., & Connor, H. K. (2022, may). Revisiting the Ground Magnetic Field Perturbations Challenge: A Machine Learning Perspective. *Frontiers in Astronomy and Space Sciences*, *9*, 123. Retrieved from https://www.frontiersin.org/articles/10.3389/fspas.2022.869740/full doi: 10.3389/fspas.2022.869740

Pulkkinen, A., Kuznetsova, M., Ridley, A., Raeder, J., Vapirev, A., Weimer, D., … Chulaki, A. (2011, feb). Geospace Environment Modeling 2008-2009 Challenge: Ground magnetic field perturbations. *Space Weather*, *9*(2), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2010SW000600 doi: 10.1029/2010SW000600

Pulkkinen, A., Lindahl, S., Viljanen, A., & Pirjola, R. (2005, aug). Geomagnetic storm of 29-31 October 2003: Geomagnetically induced currents and their relation to problems in the Swedish high-voltage power transmission system. *Space Weather*, *3*(8), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2004SW000123 doi: 10.1029/2004SW000123

Pulkkinen, A., Rastätter, L., Kuznetsova, M., Singer, H., Balch, C., Weimer, D., … Weigel, R. (2013, jun). Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather*, *11*(6), 369–385. Retrieved from http://doi.wiley.com/10.1002/swe.20056 doi: 10.1002/swe.20056

Rajput, V. N., Boteler, D. H., Rana, N., Saiyed, M., Anjana, S., & Shah, M.

(2020, nov). Insight into impact of geomagnetically induced currents on power systems: Overview, challenges and mitigation. *Electric Power Systems Research*, 106927. Retrieved from https://www.sciencedirect.com/science/article/pii/S0378779620307252?casa_token=Qs4jRHozhEkAAAAA:4AV7PlSDLxVc7NkeVOhvhPk0qJFgVgo0NrejBL7VLzYer8fpctmlt2uGLtyVZsclpebG_U-imEi6 doi: 10.1016/J.EPSR.2020.106927

Reyes, P. I., Pinto, V. A., & Moya, P. S. (2021, sep). Geomagnetic Storm Occurrence and Their Relation With Solar Cycle Phases. *Space Weather*, *19*(9), e2021SW002766. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002766 doi: 10.1029/2021SW002766

Rodger, C. J., Mac Manus, D. H., Dalzell, M., Thomson, A. W. P., Clarke, E., Petersen, T., … Divett, T. (2017, nov). Long-Term Geomagnetically Induced Current Observations From New Zealand: Peak Current Estimates for Extreme Geomagnetic Storms. *Space Weather*, *15*(11), 1447–1460. Retrieved from http://doi.wiley.com/10.1002/2017SW001691 doi: 10.1002/2017SW001691

Rogers, N. C., Wild, J. A., Eastoe, E. F., Gjerloev, J. W., & Thomson, A. W. P. (2020, feb). A global climatological model of extreme geomagnetic field fluctuations. *Journal of Space Weather and Space Climate*, *10*, 5. Retrieved from https://www.swsc-journal.org/10.1051/swsc/2020008 doi: 10.1051/swsc/2020008

Shinbori, A., Tsuji, Y., Kikuchi, T., Araki, T., Ikeda, A., Uozumi, T., … Yumoto, K. (2012). Magnetic local time and latitude dependence of amplitude of the main impulse (MI) of geomagnetic sudden commencements and its seasonal variation. *Journal of Geophysical Research: Space Physics*, *117*(8), 8322. Retrieved from http://cdaweb.gsfc. doi: 10.1029/2012JA018006

Shore, R. M., Freeman, M. P., Coxon, J. C., Thomas, E. G., Gjerloev, J. W., & Olsen, N. (2019, jul). Spatial Variation in the Responses of the Surface External and Induced Magnetic Field to the Solar Wind. *Journal of Geophysical Research: Space Physics*, *124*(7), 6195–6211. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2019JA026543 doi: 10.1029/2019JA026543

Shore, R. M., Freeman, M. P., & Gjerloev, J. W. (2017, dec). An empirical orthogonal function reanalysis of the northern polar external and induced magnetic field during solar cycle 23. *Journal of Geophysical Research: Space Physics*. Retrieved from http://doi.wiley.com/10.1002/2017JA024420 doi: 10.1002/2017JA024420

Shprits, Y. Y., Vasile, R., & Zhelavskaya, I. S. (2019, aug). Nowcasting and Predicting the ¡i¿K¡/i¿ ¡i¿p¡/i¿ Index Using Historical Values and Real-Time Observations. *Space Weather*, *17*(8), 1219–1229. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2018SW002141 doi: 10.1029/2018SW002141

Smith, A. W., Forsyth, C., Rae, I. J., Garton, T. M., Bloch, T., Jackman, C. M., & Bakrania, M. (2021, aug). Forecasting the Probability of Large Rates of Change of the Geomagnetic Field in the UK: Timescales, Horizons and Thresholds. *Space Weather*, e2021SW002788. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002788 doi: 10.1029/2021SW002788

Smith, A. W., Forsyth, C., Rae, J., Rodger, C. J., & Freeman, M. P. (2021, jun). The Impact of Sudden Commencements on Ground Magnetic Field Variability: Immediate and Delayed Consequences. *Space Weather*, *19*(7), e2021SW002764. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2021SW002764 doi: 10.1029/2021SW002764

Smith, A. W., Freeman, M. P., Rae, I. J., & Forsyth, C. (2019, nov). The Influence of Sudden Commencements on the Rate of Change of the Surface Horizontal Magnetic Field in the United Kingdom. *Space Weather*, 2019SW002281. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019SW002281 doi: 10.1029/2019SW002281

Smith, A. W., Rae, I. J., Forsyth, C., Oliveira, D. M., Freeman, M. P., & Jackson, D. R. (2020, nov). Probabilistic Forecasts of Storm Sudden Commencements From Interplanetary Shocks Using Machine Learning. *Space Weather*, *18*(11). Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020SW002603 doi: 10.1029/2020SW002603

Smith, C., L'Heureux, J., Ness, N., Acuña, M., Burlaga, L., & Scheifele, J. (1998). The ACE Magnetic Fields Experiment. *Space Science Reviews*, *86*(1/4), 613–632. Retrieved from http://link.springer.com/10.1023/A:1005092216668 doi: 10.1023/A:1005092216668

Stone, E., Frandsen, A., Mewaldt, R., Christian, E., Margolies, D., Ormes, J., & Snow, F. (1998). The Advanced Composition Explorer. *Space Science Reviews*, *86*(1/4), 1–22. Retrieved from http://link.springer.com/10.1023/A:1005082526237 doi: 10.1023/A:1005082526237

Tan, Y., Hu, Q., Wang, Z., & Zhong, Q. (2018, apr). Geomagnetic Index ¡i¿Kp¡/i¿ Forecasting With LSTM. *Space Weather*, *16*(4), 406–416. Retrieved from http://doi.wiley.com/10.1002/2017SW001764 doi: 10.1002/2017SW001764

Tasistro-Hart, A., Grayver, A., & Kuvshinov, A. (2021, dec). Probabilistic Geomagnetic Storm Forecasting via Deep Learning. *Journal of Geophysical Research: Space Physics*, *126*(1). Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2020JA028228 doi: 10.1029/2020JA028228

Thomson, A. W., Dawson, E. B., & Reay, S. J. (2011, oct). Quantifying extreme behavior in geomagnetic activity. *Space Weather*, *9*(10). Retrieved from http://doi.wiley.com/10.1029/2011SW000696 doi: 10.1029/2011SW000696

Tõth, G., Meng, X., Gombosi, T. I., & Rastätter, L. (2014, jan). Predicting the time derivative of local magnetic perturbations. *Journal of Geophysical Research: Space Physics*, *119*(1), 310–321. Retrieved from http://doi.wiley.com/10.1002/2013JA019456 doi: 10.1002/2013JA019456

Tsurutani, B. T., & Hajra, R. (2021, mar). The Interplanetary and Magnetospheric causes of Geomagnetically Induced Currents (GICs) ¿ 10 A in the Mäntsälä Finland Pipeline: 1999 through 2019. *Journal of Space Weather and Space Climate*, *11*, 23. Retrieved from https://www.swsc-journal.org/10.1051/swsc/2021001 doi: 10.1051/swsc/2021001

Turnbull, K. L., Wild, J. A., Honary, F., Thomson, A. W. P., & McKay, A. J. (2009, sep). Characteristics of variations in the ground magnetic field during substorms at mid latitudes. *Annales Geophysicae*, *27*(9), 3421–3428. Retrieved from https://angeo.copernicus.org/articles/27/3421/2009/ doi: 10.5194/angeo-27-3421-2009

Upendran, V., Tigas, P., Ferdousi, B., Bloch, T., Cheung, M. C. M., Ganju, S., ... Gal, Y. (2022, may). Global geomagnetic perturbation forecasting using Deep Learning. *Space Weather*, e2022SW003045. Retrieved from https://onlinelibrary.wiley.com/doi/10.1029/2022SW003045 doi: 10.1029/2022SW003045

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011, mar). The NumPy array: A structure for efficient numerical computation. *Computing in Science and Engineering*, *13*(2), 22–30. Retrieved from http://ieeexplore.ieee.org/document/5725236/ doi: 10.1109/MCSE.2011.37

Vennerstrom, S., Lefevre, L., Dumbović, M., Crosby, N., Malandraki, O., Patsou, I., ... Moretto, T. (2016, may). Extreme Geomagnetic Storms – 1868–2010. *Solar Physics*, *291*(5), 1447–1481. Retrieved from http://link.springer.com/10.1007/s11207-016-0897-y doi: 10.1007/s11207-016-0897-y

Viljanen, A., Koistinen, A., Pajunpää, K., Pirjola, R., Posio, P., & Pulkkinen, A. (2010). Recordings of geomagnetically induced currents in the Finnish natural gas pipeline – summary of an 11-year period. *Geophysica*, *46*(1-2), 59–67. Retrieved from https://www.geophysica.fi/pdf/geophysica_2010_46_1-2_059_viljanen.pdf

Viljanen, A., Tanskanen, E. I., & Pulkkinen, A. (2006, mar). Relation between substorm characteristics and rapid temporal variations of the ground magnetic field. *Annales Geophysicae*, *24*(2), 725–733. Retrieved from https://angeo.copernicus.org/articles/24/725/2006/ doi: 10.5194/angeo-24-725-2006

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. doi: https://doi.org/10.1038/s41592-019-0686-2

Weigel, R. S., Vassiliadis, D., & Klimas, A. J. (2002, oct). Coupling of the solar wind to temporal fluctuations in ground magnetic fields. *Geophysical Research Letters*, *29*(19), 21–1–21–4. Retrieved from http://doi.wiley.com/10.1029/2002GL014740 doi: 10.1029/2002GL014740

Weimer, D. R. (2013, mar). An empirical model of ground-level geomagnetic perturbations. *Space Weather*, *11*(3), 107–120. Retrieved from http://doi.wiley.com/10.1002/swe.20030 doi: 10.1002/swe.20030

Weimer, D. R., & King, J. H. (2008, jan). Improved calculations of interplanetary magnetic field phase front angles and propagation time delays. *Journal of Geophysical Research: Space Physics*, *113*(A1), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2007JA012452 doi: 10.1029/2007JA012452

Welling, D. (2019, sep). Magnetohydrodynamic models of B and their use in GIC estimates. In *Geomagnetically induced currents from the sun to the power grid* (pp. 43–65). American Geophysical Union (AGU). Retrieved from https://onlinelibrary.wiley.com/doi/10.1002/9781119434412.ch3 doi: 10.1002/9781119434412.ch3

Wik, M., Pirjola, R., Lundstedt, H., Viljanen, A., Wintoft, P., & Pulkkinen, A. (2009, apr). Space weather events in July 1982 and October 2003 and the effects of geomagnetically induced currents on Swedish technical systems. *Annales Geophysicae*, *27*(4), 1775–1787. Retrieved from https://angeo.copernicus.org/articles/27/1775/2009/ doi: 10.5194/angeo-27-1775-2009

Wing, S., Johnson, J. R., Jen, J., Meng, C. I., Sibeck, D. G., Bechtold, K., … Takahashi, K. (2005, apr). Kp forecast models. *Journal of Geophysical Research: Space Physics*, *110*(A4). Retrieved from http://doi.wiley.com/10.1029/2004JA010500 doi: 10.1029/2004JA010500

Wintoft, P., Viljanen, A., & Wik, M. (2016, may). Extreme value analysis of the time derivative of the horizontal magnetic field and computed electric field. *Annales Geophysicae*, *34*(4), 485–491. Retrieved from https://angeo.copernicus.org/articles/34/485/2016/ doi: 10.5194/angeo-34-485-2016

Wintoft, P., Wik, M., Matzka, J., & Shprits, Y. (2017, nov). Forecasting ¡i¿Kp¡/i¿ from solar wind data: input parameter study using 3-hour averages and 3-hour range values. *Journal of Space Weather and Space Climate*, *7*, A29. Retrieved from http://www.swsc-journal.org/10.1051/swsc/2017027 doi: 10.1051/swsc/2017027

Wintoft, P., Wik, M., & Viljanen, A. (2015, mar). Solar wind driven empirical forecast models of the time derivative of the ground magnetic field. *Journal of Space Weather and Space Climate*, *5*, A7. Retrieved from http://www.swsc-journal.org/10.1051/swsc/2015008 doi: 10.1051/swsc/2015008

Yue, C., Zong, Q. G., Zhang, H., Wang, Y. F., Yuan, C. J., Pu, Z. Y., … Wang, C. R. (2010, may). Geomagnetic activity triggered by interplanetary shocks. *Journal of Geophysical Research: Space Physics*, *115*(A5), n/a–n/a. Retrieved from http://doi.wiley.com/10.1029/2010JA015356 doi: 10.1029/2010JA015356

Zhelavskaya, I. S., Vasile, R., Shprits, Y. Y., Stolle, C., & Matzka, J. (2019). Systematic Analysis of Machine Learning and Feature Selection Techniques for Prediction of the Kp Index. *Space Weather*, *17*(10), 1461–1486. Retrieved from https://doi.org/10.1029/2019SW002271 doi: 10.1029/2019SW002271

Zong, Q.-G., Yue, C., & Fu, S.-Y. (2021, mar). Shock Induced Strong Substorms and Super Substorms: Preconditions and Associated Oxygen Ion Dynamics. *Space Science Reviews*, *217*(2), 33. Retrieved from http://link.springer.com/10.1007/s11214-021-00806-x doi: 10.1007/s11214-021-00806-x

Zwickl, R., Doggett, K., Sahm, S., Barrett, W., Grubb, R., Detman, T., ... Maruyama, T. (1998). The NOAA Real-Time Solar-Wind (RTSW) System using ACE Data. *Space Science Reviews 1998 86:1*, *86*(1), 633–648. Retrieved from https://link.springer.com/article/10.1023/A:1005044300738 doi: 10.1023/A:1005044300738