Super Dual Auroral Radar Network Expansion and its Influence on the Derived Ionospheric Convection Pattern

Maria-Theresia Walach[1], Adrian Grocott[1], Frances A Staples[2], and Evan G. Thomas[3]

[1]Lancaster University
[2]Mullard Space Science Laboratory
[3]Dartmouth College

November 23, 2022

Abstract

The Super Dual Auroral Radar Network (SuperDARN) was built to study ionospheric convection and has in recent years been expanded geographically. Alongside software developments, this has resulted in many different versions of the convection maps dataset being available. Using data from 2012 to 2018, we produce five different versions of the widely used convection maps, using limited backscatter ranges, background models and the exclusion/inclusion of data from specific radar groups such as the mid-latitude radars. This enables us to simulate how much information was missing from previous decades of SuperDARN research. We study changes in the Heppner-Maynard boundary, the cross polar cap potential (CPCP), the number of backscatter echoes (n) and the χ-squared/n statistic which is a measure of the global agreement between the measured and fitted velocities. We find that the CPCP is reduced when the polar cap radars are introduced, but then increases again when the mid-latitude radars are added. When the background model is changed from the RG96 model, to the most recent TS18 model, the CPCP tends to decrease for lower values, but tends to increase for higher values. When comparing to geomagnetic indices, we find that there is on average a linear relationship between the Heppner-Maynard boundary and the geomagnetic indices, as well as n, which breaks at high values (e.g. HMB ˜50 degrees) due to the low observational density. We find that whilst n is important in constraining the maps (maps with n>400 are unlikely to change), is insufficient as the sole measure of quality.

Super Dual Auroral Radar Network Expansion and its Influence on the Derived Ionospheric Convection Pattern

M.-T. Walach[1], A. Grocott[1], F. Staples[2], E. G. Thomas[3]

[1]Lancaster University, Lancaster, LA1 4YW, UK
[2]Mullard Space Science Laboratory, University College London, Holmbury St. Mary, RH5 6NT, UK
[3]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA

Key Points:

- We identify changes in measurements when high- and mid-latitude radars are added to SuperDARN, and show the impact of different processing
- Measured convection parameters are highly susceptible to processing parameters and which radars are used
- We show how the number of backscatter echoes per map is critical to the convection maps, and discuss how this impacts map quality

Corresponding author: M.-T. Walach, `m.walach@lancaster.ac.uk`

Abstract

The Super Dual Auroral Radar Network (SuperDARN) was built to study ionospheric convection and has in recent years been expanded geographically. Alongside software developments, this has resulted in many different versions of the convection maps dataset being available. Using data from 2012 to 2018, we produce five different versions of the widely used convection maps, using limited backscatter ranges, background models and the exclusion/inclusion of data from specific radar groups such as the mid-latitude radars. This enables us to simulate how much information was missing from previous decades of SuperDARN research. We study changes in the Heppner-Maynard boundary, the cross polar cap potential (CPCP), the number of backscatter echoes (n) and the χ^2/n statistic which is a measure of the global agreement between the measured and fitted velocities. We find that the CPCP is reduced when the polar cap radars are introduced, but then increases again when the mid-latitude radars are added. When the background model is changed from the RG96 model, to the most recent TS18 model, the CPCP tends to decrease for lower values, but tends to increase for higher values. When comparing to geomagnetic indices, we find that there is on average a linear relationship between the Heppner-Maynard boundary and the geomagnetic indices, as well as n, which breaks at high values (e.g. HMB $\sim$50°) due to the low observational density. We find that whilst n is important in constraining the maps (maps with n >400 are unlikely to change), is insufficient as the sole measure of quality.

Plain Language Summary

The ionosphere, where space begins and the atmosphere ends, moves as a result of the Earth's magnetic field coupling with the Sun. The Super Dual Auroral Radar Network (SuperDARN) was built around the Earth's magnetic poles to study this phenomenon, known as ionospheric convection. Combining many line-of-sight convection measurements, we are able to build global maps of ionospheric convection using SuperDARN. This encapsulates dynamics which are central to space weather phenomena. SuperDARN, which has been gathering data for decades, has over time undergone numerous transformations, including the development of new processing software and more radars being added to the network. Using data from the years 2012 to 2018, we perform a statistical analysis on processed SuperDARN convection maps for the entire dataset and assess systematically how the dataset has changed over the years. We consider how the addition of more data and changes to the convection mapping procedures can affect scientific studies in the context of this large database.

1 Introduction

The Super Dual Auroral Radar Network (SuperDARN) consists of high-frequency coherent scatter radars built to study ionospheric convection by means of Doppler-shifted, pulse sequences and has been widely used in space physics and ionospheric research (e.g. Greenwald et al., 1995; Ruohoniemi & Greenwald, 1996; Chisham et al., 2007; Nishitani et al., 2019). SuperDARN data are continuously available since 1993, with the network having expanded over time from one radar (built in 1983) to 23 radars in the Northern hemisphere, 13 in the Southern hemisphere and more under construction (Nishitani et al., 2019). This expansion has allowed for a greater area to be covered by SuperDARN (i.e. down to magnetic latitudes of 40°) with at least 16 different azimuthal look directions (Nishitani et al., 2019) in the Northern hemisphere. Line-of-sight measurements by this large-scale network of radars can be combined and used to construct a picture of high-latitude ionospheric convection on time scales of 1-2 minutes (Ruohoniemi & Baker, 1998). The radars can be grouped into high-latitude radars, polar-latitude radars (or PolarDARN), and mid-latitude radars (or StormDARN). Nishitani et al. (2019) provides a summary from a historical northern hemisphere perspective: high-latitude radars, at magnetic latitudes of 50-70° were first built, starting in 1983 with the Goose Bay radar, followed by the polar radars (covering 70-90° magnetic latitude), and the expansion to mid-latitudes (∼40-50°), starting in 2005 with the Wallops Island radar. Over time new radars have improved global ionospheric convection mapping by increasing the number of measurements and look directions.

The most commonly used SuperDARN data product by the space science and ionospheric research community is the convection maps. Convection maps are large scale maps, showing ionospheric convection around the magnetic poles. In order to produce these maps, several data processing steps have to be undertaken. With the expansion of the dataset, as well as data processing software improvements, this data product has undergone several changes.

To make SuperDARN convection maps the raw data is processed using the Radar Software Toolkit (RST (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Bland, et al., 2018)):

1. An autocorrelation function is fitted to the raw radar data. This produces fitacf files, which store the line-of-sight velocity data.
2. The data is then gridded onto an equal area latitude-longitude grid (see equation 1 from Ruohoniemi & Baker, 1998) and split into typically one or two minute cadence records. Historically it has almost always been the case that all data from the radars were added to the grids. However, slow moving E-region scatter can and should be removed by setting the minimum range gate limit to 800 km (Forsythe & Makarevich, 2017; Thomas & Shepherd, 2018). It has recently become apparent that far range data beyond 2000 km can also be problematic owing to geolocation uncertainties in the range finding algorithm (Chisham et al., 2008).
3. Data from different radars are combined and the spherical harmonic fitting algorithm is applied which fits an electrostatic potential in terms of spherical harmonic functions to the data (Ruohoniemi & Greenwald, 1996; Ruohoniemi & Baker, 1998). To find the optimal solution for the spherical harmonic coefficients, a singular value decomposition (e.g. Press, W. H. and Teukolsky, S. A. and Vetterling W. T. and Flannery B. P., 2007) is minimised. When this fitting is performed, typically a background model, parameterised by solar wind conditions is used, to infill information in the case of data gaps. This method is also known as 'Map Potential' technique.

Several models are available for the fitting in step 3, most notably Ruohoniemi and Greenwald (1996) generated the most widely used statistical background model, which was subsequently implemented in the RST. This background model was thus used by most SuperDARN users to generate convection maps and used in many scientific studies. Ruohoniemi and Greenwald (1996) used the Goose Bay radar to create the background statistical model. Since then, however many more radars have been added to SuperDARN. This raises the question of how much of an effect changing the background model has on the convection map dataset, which was investigated by Shepherd and Ruohoniemi (2000). The main conclusion from Shepherd and Ruohoniemi (2000) was that the solution becomes insensitive to the choice of statistical model when the data coverage is high. Since then, Ruohoniemi and Greenwald (2005) produced an updated version of their statistical background model using data from 9 radars, but this was not implemented into RST, thus keeping the RG96-model the default which was used by the community. Since then, a number of updated background models, such as Pettigrew et al. (2010), Cousins and Shepherd (2010) and Thomas and Shepherd (2018) have been produced. The Pettigrew et al. (2010) and Cousins and Shepherd (2010) models were not implemented into RST until version 4.1 (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Bland, et al., 2018). Soon after, the statistical background model by Thomas and Shepherd (2018) was released, which is now standard in RST since version 4.2 (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Billett, et al., 2018). The RG96 and TS18 models are thus the most widely used and we will focus our analysis on these background models.

Alongside the use of a background model, a Heppner-Maynard boundary (HMB) (Heppner & Maynard, 1987), the low-latitude boundary of the convection pattern where the flows approach zero, can either be specified or be chosen using backscatter measurements. This is to constrain the convection pattern when the spherical harmonic fit is applied (Shepherd & Ruohoniemi, 2000). For typical two minute cadence convection maps, it is appropriate to find where three radar velocity measurements are greater than $100\mathrm{ms}^{-1}$ for the HMB (Imber et al., 2013). This boundary is circular around the nightside and cropped at the dayside to mimic the shape of the dayside magnetopause. Previous to Shepherd and Ruohoniemi (2000) however, a fully circular boundary was used, which was deemed to create unrealistic flows at lower latitudes when the radar network was expanded.

In this paper we conduct a large scale data analysis to assess systematically how the SuperDARN dataset has changed over the years and how this may have affected the dataset overall.

We specifically probe the effects of the following changes:

1. Inclusion of the backscatter range limits
2. Addition of the PolarDARN data
3. Addition of the StormDARN data
4. Updating of the background statistical model

2 Data and Method

To provide a meaningful large scale comparison of different versions of the SuperDARN dataset, we process Northern hemisphere data from the same time period (2012-2018) and create different versions of the SuperDARN convection maps. First, we create a baseline dataset (D0) with the high-latitude radars only, which is then modified by changing one aspect for each subsequent dataset. This allows us to contrast the changes in the dataset. Table 1 outlines the different datasets (D0 to D4) and how each one varies from the previous iteration. The basic data processing is the same for all the datasets, except with the changes outline in table 1. All raw SuperDARN data were obtained from the British Antarctic Survey's SuperDARN mirror and then processed using the Radar Software Toolkit version 4.3 (SuperDARN Data Analysis Working Group et al., 2019). The specific processing commands and options used for the data processing can be found in the appendix of this paper. The rawacf-files were converted into fitacf-files using the FITACF function (version 2.5). Two gridded map files were created to see how changing the backscatter range limit affects the dataset. One version of the gridded files was created with an added backscatter range limit. By only including data from a minimum range of 800 km and a maximum far range of 2000 km, we eliminate all possible E-Region scatter and all backscatter with higher uncertainties in their location (Chisham et al., 2008; Forsythe & Makarevich, 2017; Thomas & Shepherd, 2018). The version of gridded files with a backscatter range limit is used for D1-D4 and the one without a range limit is used for D0. The gridded map files were resolved into two minute records and used the Chisham virtual height model (Chisham et al., 2008).

Dataset versions D0 and D1 include the same radars, whereas for D2 and D3, more radars were included (see table 1). For this selection of PolarDARN and StormDARN groupings the list provided by table 1 in Thomas and Shepherd (2018) was used. As can be seen from the list provided in Thomas and Shepherd (2018), most of the StormDARN radars were built after the high-latitude and PolarDARN radars.

For D4, we keep the selection of radars the same as D3, but use the background model from Thomas and Shepherd (2018) instead of the one from Ruohoniemi and Greenwald (1996).

To make all the final convection maps (D0 to D4), using RST, the Heppner-Maynard boundary (Heppner & Maynard, 1987; Shepherd & Ruohoniemi, 2000) was chosen as the lowest possible latitude measured by a minimum of three LOS vectors with velocities greater than 100 m/s (Imber et al., 2013). To complete the map fitting algorithm, the model requires solar wind data to be selected. For this, we use solar wind data from the ACE spacecraft, which has been time-lagged to the magnetosphere using the algorithm from Khan and Cowley (1999) which takes magnetosheath transit time into account. Finally, we add the model, and use a fitting order of 6 with a 'light' doping level for the background solar wind model. This uses the technique from Ruohoniemi and Baker (1998) to fit electrostatic potentials to the measured velocity vectors as spherical harmonic functions.

Choosing these versions of the dataset allows for a large-scale analysis of systematic changes and in particular, how the introduction of new mid-latitude and polar data modifies the dataset on a large scale, which has implications for use of the maps in scientific studies. Having established this archive of 2-minute resolution convection map files, we then extract a set of measured parameters with which quantify ionospheric convection, such as the HMB latitude and cross polar cap potential (CPCP). These describe the spatial extent and strength of the convection and allow us to examine how changes in the processing might affect conclusions of scientific studies, whereas the number of backscatter echoes per map or the average number of backscatter points per radar allows us to study how changes affect coverage. In this study, we define the HMB latitude as the fitted latitudinal boundary on the nightside and we also investigate how this parameter changes alongside the minimum latitude where backscatter is obtained (Λ_{min}), which can be along any magnetic local time or longitude. We would thus expect the difference between the two parameters to be positive for well constrained maps (i.e. Λ_{min} is at a lower latitude than the HMB), but this can also be negative when either the minimum latitude of observations is on the dayside (where the HMB shifts to higher latitudes) or an indicator that the HMB is not constrained by data. We also show how the different processing affects the χ^2/n-statistic, which is a global measure of map quality. The χ^2 parameter is a result from the singular value decomposition, which is minimised when the spherical harmonic fitting is performed to find the optimal solution for the coefficients. χ^2/n was introduced by Ruohoniemi and Baker (1998) as an indicator how well the measured line-of-sight velocities match the fitted velocities, where a value of 1 would indicate a good match and higher values would indicate a worse match.

Additionally, we also discuss the relationship between the HMB latitude and measures of geomagnetic activity, such as the Auroral Lower index (AL), the Auroral Elec-

Table 1. Differences between the comparison datasets

Version	Introduced difference	Background model	high-latitude radars	range limit	PolarDARN radars	StormDARN radars
D0	High-latitude radars[a] only	RG96	yes	no	no	no
D1	added range limit: 800-2000 km	RG96	yes	yes	no	no
D2	added PolarDARN radars[b]	RG96	yes	yes	yes	no
D3	added all other (i.e. StormDARN radars)[c]	RG96	yes	yes	yes	yes
D4	changed the background model	TS18	yes	yes	yes	yes

[a] High-latitude radars (i.e. all other radars): King Salmon, Kodiak, Prince George, Saskatoon, Kapuskasing, Goose Bay, Stokkseyri, Pykkvibaer, Hankasalmi.

[b]PolarDARN radars include: Inuvik, Rankin Inlet, Clyde River, Longyearbyen.

[c]StormDARN radars include: Hokkaido West, Hokkaido East, Adak West, Adak East, Christmas Valley West, Christmas Valley East, Fort Hays West, Fort Hays East, Blackstone, Wallops Island.

trojet index (AE) and the Symmetric Horizontal index (Sym-H) (Davis & Sugiura, 1966; Iyemori, 1990). We also consider the relationship between the CPCP and Φ_D, the day-side reconnection rate, which is calculated from the IMF B_Z, solar wind speed and IMF clock angle (Milan et al., 2012; Walach et al., 2017).

3 Results

The timeseries data extracted from the SuperDARN convection maps is condensed into probability distribution functions. By showing the data as 3-dimensional data distributions, we are able to compare the effects of changing the dataset on various parameters, which is shown in this section alongside examples of convection maps illustrating the changes.

3.1 Restricting radar backscatter range

Figure 1 shows probability distribution functions for a number of parameters for the entire D0 and D1 datasets. With D1 we have introduced the use of a range limit, as described in section 2.

Fig. 1a shows the distribution of HMB latitudes in D0 against D1. As most datapoints lie above the line of unity, we see that the HMB generally retreats poleward when we introduce a backscatter range limit. By limiting the backscatter ranges the number of backscatter echoes is reduced and thus also always increasing the lowest latitude at which backscatter is observed. We also see a saturation of points at a HMB latitude of 60°, which is where the boundary is drawn if not enough data is available (due to low data coverage or no slow scatter being observed). Fig. 1b shows the difference between the HMB latitude and Λ_{min}. We see that this difference is mostly positive for both D0 and D1, which means that the HMB sits below Λ_{min} and is thus well constrained. This latitudinal difference tends to shrink as we change the dataset from D0 to D1, as would be expected with a limited backscatter range. For a number of observations (40%), this latitudinal difference changes from positive to negative. This occurs for maps where the HMB is either not well constrained or the minimum latitude of observations is obtained on the dayside. Fig. 1c shows the χ^2/n distribution. It shows that χ^2/n tends to increase when the range limit is introduced. The range limit is expected to remove slow-moving E-region scatter ($<$ 800 km ranges) or scatter that may be placed in the wrong location ($>$ 2000 km ranges), which is expected to eliminate noise and uncertainty. Sometimes, χ^2/n measured at higher values in D0 (15-30) decreases for D1 (0-10), indicating that the map fitting improves. Fig. 1d shows the distribution of the number of backscatter echoes per map, n. It is worth noting that for the majority of D0 and D1, n is below 200, which as we will see in sections 3.2 to 3.6, is fairly low. Fig. 1e shows the average number of backscatter echoes per radar. As expected, changing the dataset from D0 to D1 not only decreases n overall, but also decreases the average number of backscatter echoes per radar. Fig. 1f shows the distribution of the CPCP. We see that when a range limit is introduced, the CPCP can either increase or decrease and there is no preference either way.

Panels Fig. 1g and h show two example convection maps for the same date and time (21st December 2014 at 21:58 UT) from D0 and D1. In each case, the grid is geomagnetic latitude (which is in the AACGM-v2 coordinate system (Shepherd, 2014)) and magnetic local time, with noon towards the top, dusk towards the left, midnight towards the bottom and dawn towards the right. The coloured vectors show the gridded line-of-sight velocity vectors in locations where SuperDARN backscatter is available rather than the usual fitted vectors from Map Potential, which are usually shown in convection maps. The colours indicate the magnitudes of the vectors. The HMB is shown by the bright green line and the solid and dashed black lines show equipotentials in the electrostatic potential. To provide more context, this example map is indicated in the PDFs above by the light blue crosses. We see immediately that despite the high number of backscatter echoes and the low χ^2/n, there is a considerable difference in the potential patterns between D0 and D1. In D0 there are extra vectors in the dayside portion of the convection map, which provide fast flows, but also extra asymmetry that introduces an unphysical morphology. Adding in the range limit removes these and whilst it does not change the CPCP by much (4 kV), the convection maps themselves change considerably. Imposing the range limit removes fast vectors on the dayside and thus minimises the unphysical convection cells. This is an example where adding the range limit qualitatively improves the map and reduces the χ^2/n-statistic.

3.2 Adding PolarDARN

Figure 2 shows a comparison between D1 and D2 in the same format as in Fig. 1. In this comparison, we have introduced the Polar radars to the maps going from D1 to D2.

Fig. 2a shows the distribution of HMB latitudes. For 26.68%, the HMB moves to higher latitudes and for the majority of maps however (71.53%), the HMB does not change at all. The HMB moves to higher latitudes if it was not defined for D1 and adding more data can mean that the HMB is introduced at higher latitudes and the latitudinal difference between the HMB and the minimum latitude of observations thus becomes negative, shich is shown in Fig. 2b. This shows the difference between the HMB latitude and Λ_{min}. Fig. 2b shows that this distance tends to increase when we add the PolarDARN radars to the maps, which means the HMB is better constrained. The exception here are 1.79% of maps, where the minimum latitude HMB was already at high latitudes

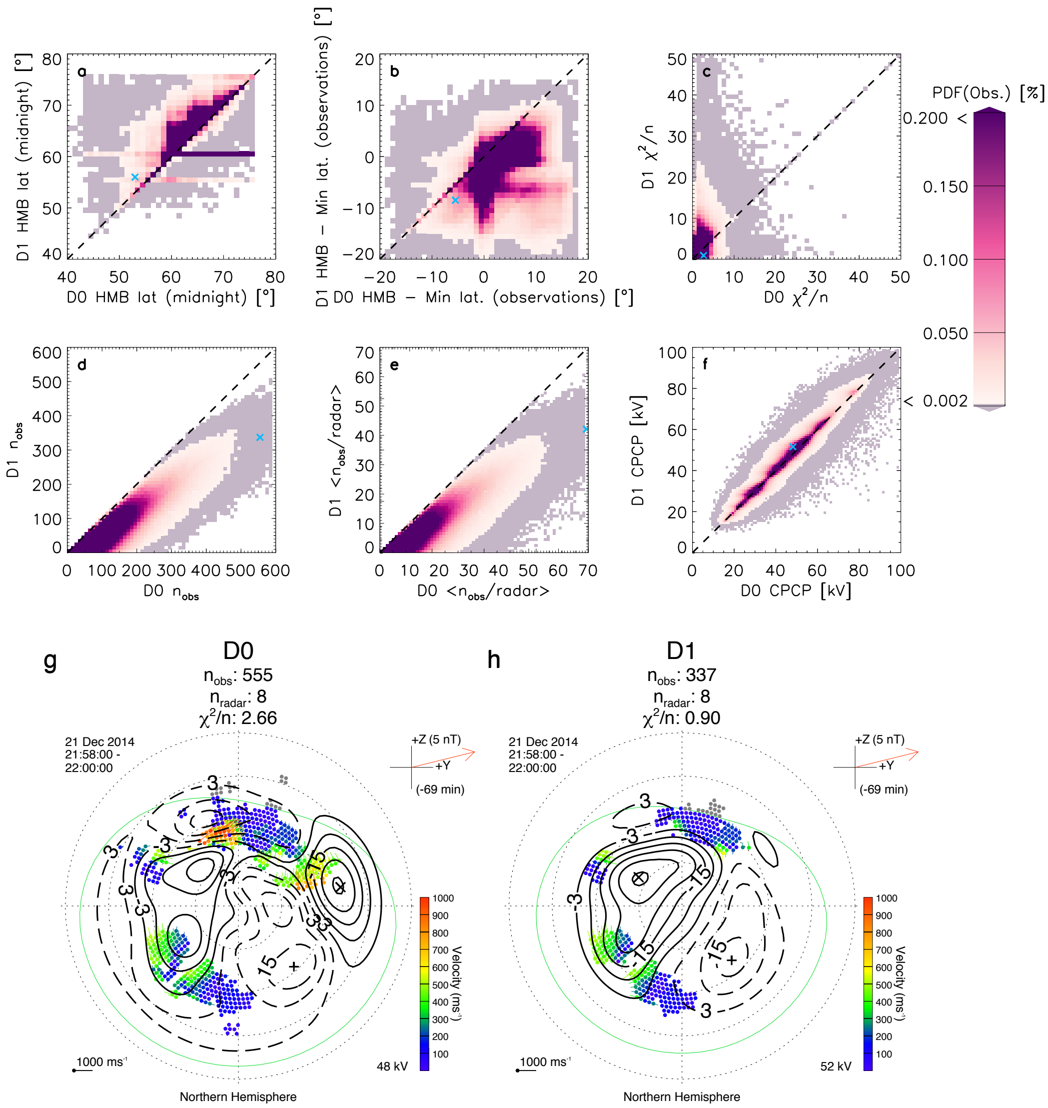

Figure 1. Probability distribution functions comparing the entire D0 and D1 datasets: (a) HMB latitude, (b) the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two panels show two example maps with the line-of-sight vectors from the same date and time (2014/12/21 21:58) for D0 (g) and D1, the convection map with the added range gate limit (h). These occurrences are indicated in the PDFs by blue crosses.

for D1 ($\geq$72°), suggesting low coverage in the first instance, and thus introducing new data at high latitudes moves the minimum latitude of observation to slightly lower latitudes. Fig. 2c shows the χ^2/n distribution. We see that χ^2/n sometimes increases and sometimes decreases: This split is approximately equal with 45.40% of χ^2/n increasing and 49.76% of χ^2/n decreasing.

Fig. 2d shows the distribution of n. As we are introducing new data, the number of backscatter observations always increases, independently of how much data were available in D1.

Fig. 2e shows the average number of backscatter observations per radar. We see that this is likely to increase when the PolarDARN data is added. This means that the polar radars observed on average more backscatter points than the older radars in the network.

Fig. 2f shows the CPCP distribution. When adding the PolarDARN data to the network, it is possible for the CPCP to increase or decrease. We see that the spread of points above the line of unity is larger than below it. This means that if the CPCP increases, it is possible to increase by more than 30 kV, though the majority of data lies below the unity line and is likely to decrease by less than ~30 kV.

As in Fig. 1, Fig. 2g and h show two example maps using D1(g) and D2 (h) for the same time (4$^{\text{th}}$ November 2014 at 20:08 UT), where the number of observations increases from 238 to 468. For this example the number of datapoints increases and this changes the pattern, despite the HMB still being constrained by the same datapoints. As high latitude datapoints are added however, the pattern is better constrained and a dawn cell appears due to fast flows being measured in the noon-morning region, leading to an increase in the CPCP from 27 kV to 54 kV.

3.3 Adding StormDARN

Figure 3 illustrates how the maps change when the mid-latitude (StormDARN) radars are added to the dataset. Fig. 3a shows the HMB distribution. This shows that the HMB is likely to stay at the same latitude or move closer to the equator. Fig. 3b shows the difference between the HMB latitude and Λ_{min}. As data from the mid-latitude radars are added, this latitudinal distance is likely to increase as would be expected. This dis-

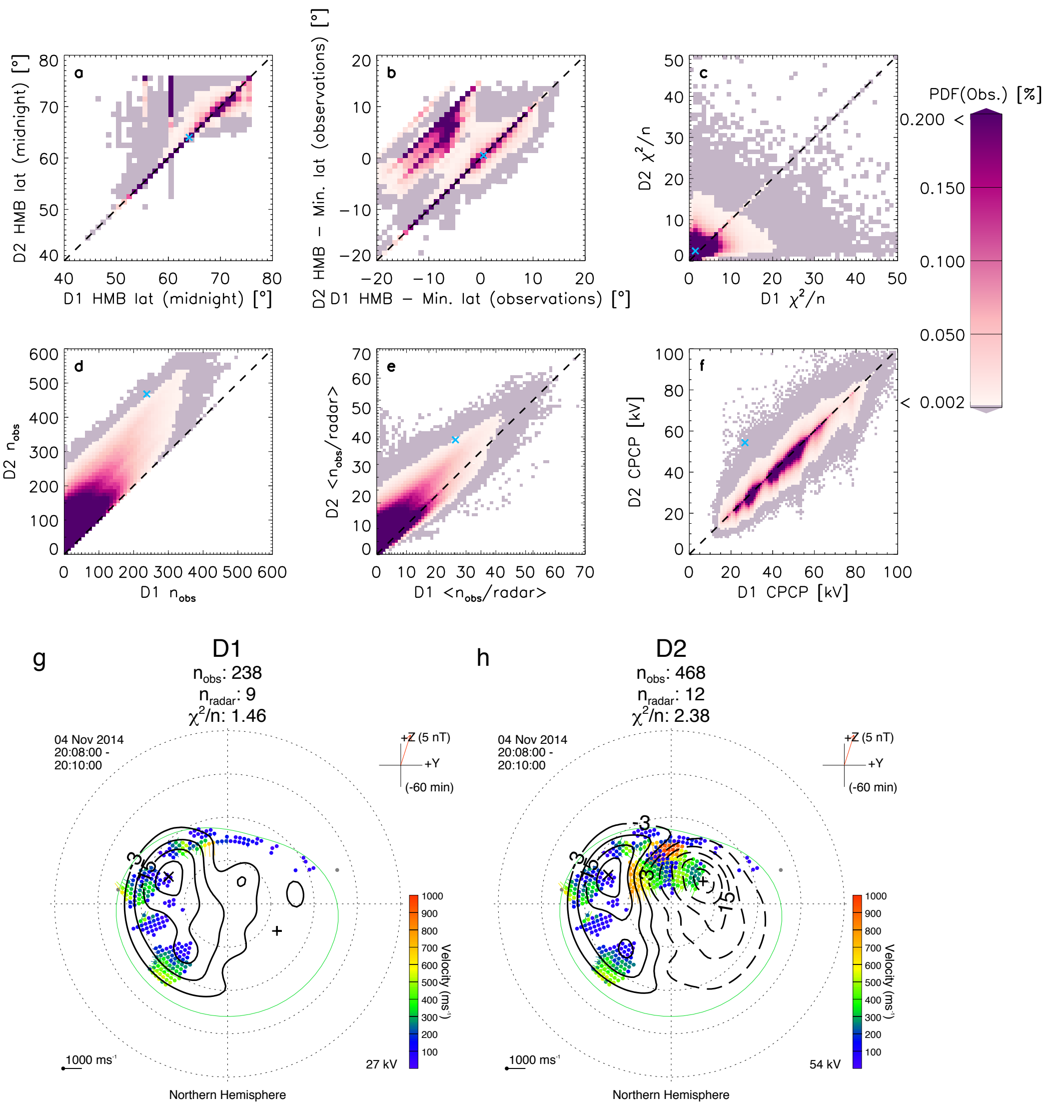

Figure 2. Probability distribution functions comparing the entire D1 and D2 datasets: (a) HMB latitude, (b) N the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two panels show two example maps with the line-of-sight vectors from the same date and time (2014/11/04 20:08) for D1 (g) and D2, the convection map with the added PolarDARN data (h). The example maps occurrences are indicated in the PDFs by blue crosses.

tance tends to be a positive one in the D3 dataset, meaning Λ_{min} tends to be closer to the equator than the HMB. This means the HMB is likely to be better constrained in D3 than D2. Fig. 3c shows the χ^2/n distribution. This value tends to decrease when we change the dataset from D2 to D3, which means that the background model fitting improves on average. Fig. 3d shows the number of backscatter observations, which increases as expected. Fig. 3e though shows that the number of gridded backscatter echoes per radar tends to decrease. This means that the average mid-latitude radar tends to observe fewer backscatter echoes than high-latitude radars. Fig. 3f shows the CPCP. We see from that the CPCP can increase or decrease, but the increases tend to be of a larger value than the decreases.

The bottom two rows in Fig. 3 show four example maps: The panels on the left (g and i) show example map of D2 from 9th November 2013 at 04:00 and the 8th February 2014 at 09:26, respectively. The two panels on the right (h and j) show the same date and time but using D3, where mid-latitude radars were included.

We see in panels g and h, that in this example adding these data increases the backscatter echoes by over 200 datapoints, even for this map, where the number of observations was already high previously. This moves the latitude of the HMB to lower latitudes from 62° to 52°. Furthermore, we see the convection cells change, in particular the dawn cell and the CPCP increases from 58 to 69 kV. All this will have a noticeable effect on any parameters extracted from the map. For example if we compute the convection velocity in D3 at the location where the HMB meets the midnight meridian for D2 (i.e. at 62° longitude and 00 MLT), the velocity would change from D2 to D3 from 0 m/s to 422 m/s.

Panels i and j of Fig. 3 however show an example of where adding mid-latitude data can make the convection maps look worse: Adding scatter at mid-latitudes almost doubles n, which increases from 326 to 613 here. Many of the measurements are however slow moving scatter, albeit not slow enough to fall below the HMB threshold, which results in the dawn convection cell almost disappearing. Initially this may seem like an extreme change in convection morphology, but the dawn cell only changes by ∼3 kV and the dusk cell is much better constrained by new mid-latitude vectors. The combination of these two changes causes an overall increase in the CPCP from 40 kV to 53 kV.

Computing the velocities for D3 at the HMB latitude location in D2 can be used as an indicator of how much the map has changed at specific locations and gives us an idea of how quantitatively different the convection maps might be without the mid-latitude radars. We explore this in more detail now.

Figure 4 shows the velocities, extracted from the D3 convection maps for the locations where the D2-HMB intersects with the noon, dusk, midnight and dawn meridians. We see that by adding the mid-latitude data, the maps change considerably at the locations where the HMB would have otherwise stipulated that there be zero flow. The curves show that at dawn, the effect is the least noticeable and that there is a 1 in 2 chance that the velocity measured in D3 has increased by 120 m/s or less, whereas this increases to 190 m/s for midnight and 220 m/s and 230 m/s for noon and dusk, respectively.

3.4 Changing the background model

In changing the dataset from D3 to D4, we are changing the background model from RG96 to TS18. This means that the observations which go into the convection maps stay constant, but the model fitting parameters (χ^2/n) change, as well as some of the resulting parameters, such as the CPCP.

Figure 5a shows the D3 versus D4 CPCP and we see that at the lower range (0-$\sim$50 kV), the CPCP is likely to decrease as we change the background model from RG96 to TS18 (this occurs 41.65% of the time as opposed to the increase which occurs 28.56% of the time). For the higher range (>50 kV) however, the CPCP is likely to increase when we change model from RG96 (D3) to TS18 (D4) (this occurs 16.46% of the time as opposed to the decrease which is 13.32%). Overall, TS18 thus provides a lower CPCP 54.97% of the time and a higher CPCP 45.02% of the time for the same data. Fig. 5b shows the CPCP difference against n. We see from this that the CPCP is in fact best constrained for maps with a high number of backscatter points, which means that there is a model dependency which decreases as n increases. For example, At n=200, the median and standard deviation are 0.87 kV and 8.88 kV, whereas at n=400, the median and standard deviation are 0.04 kV and 6.50 kV, respectively. Fig. 5c shows that χ^2/n, which can either decrease or increase when changing the background model. Although not immediately obvious, 63.81% of the data lie below the line of unity (in comparison to 36.15%

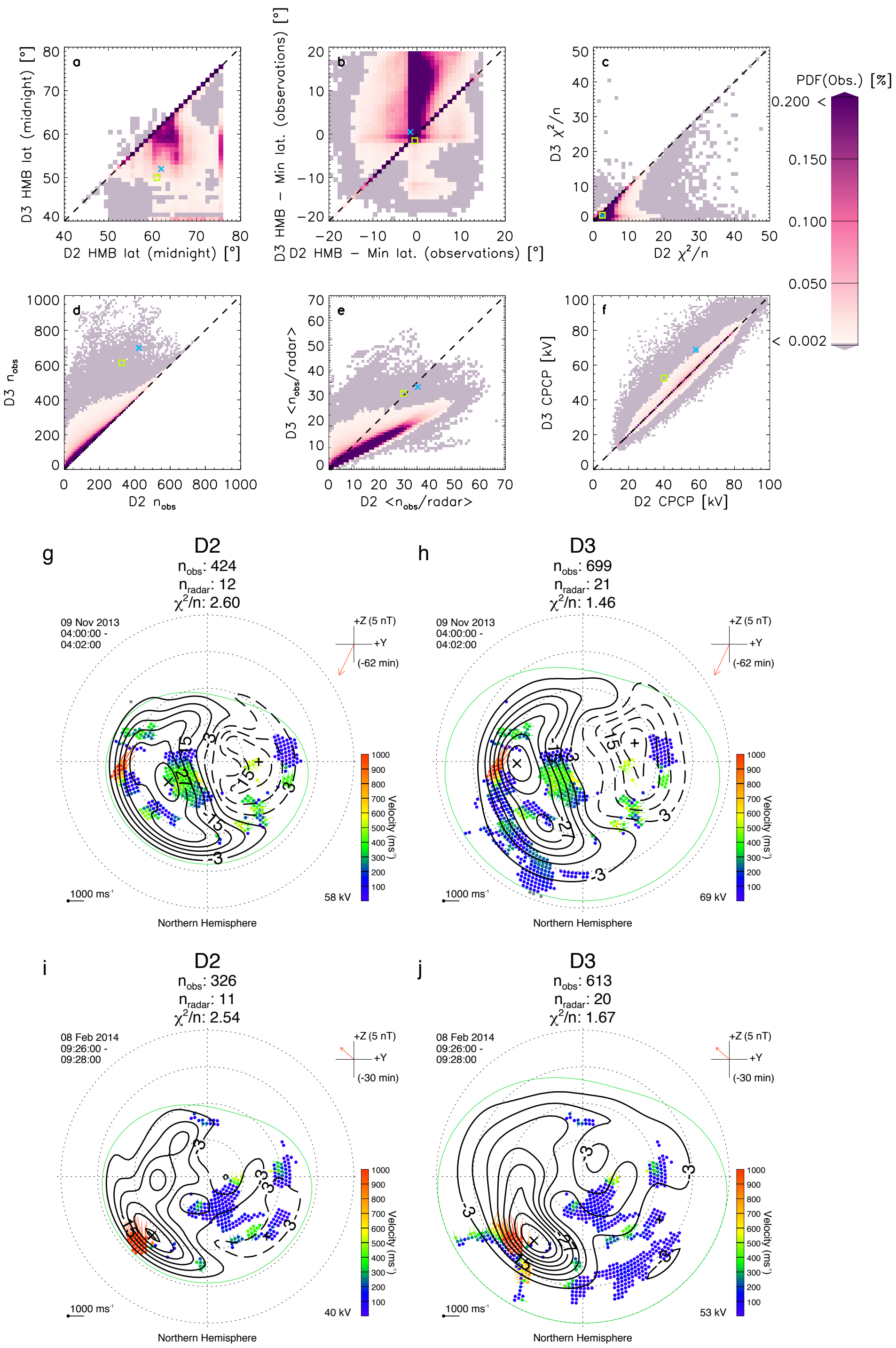

Figure 3. Probability distribution functions comparing the entire D2 and D3 datasets: (a) HMB latitude, (b) the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two rows show four example maps with the line-of-sight vectors from two different dates and times (2013/11/09 04:00 and 2014/02/08 09:26) for D2 (left, g and i) and D3, the convection map which includes the mid-latitude radar data (right, h and j). These occurrences are indicated in the PDFs by blue crosses and green squares.

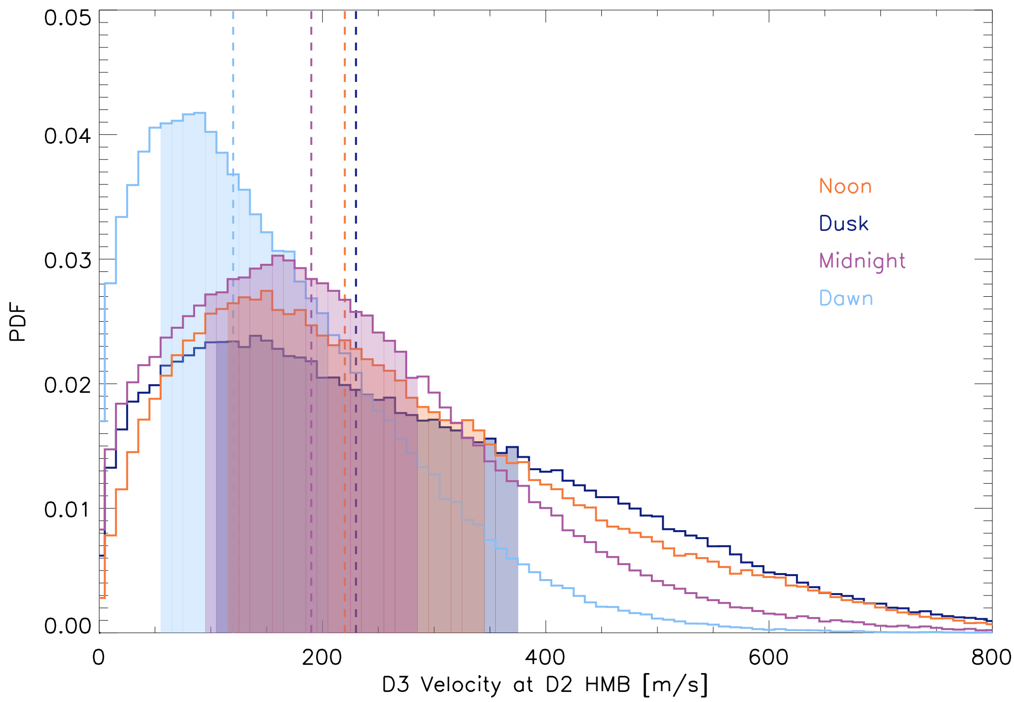

Figure 4. Probability distribution function of the velocity for D3, extracted at the noon, dusk, midnight and dawn locations where D2 would have had the HMB. Dashed lines show the medians for each distribution. Shaded regions indicate the boundaries of the lower and upper quartiles (25% and 75%).

of data above the line), meaning the fitting error is on average reduced when making the convection maps using TS18 in comparison to RG96.

As the input data does not change, the HMB values are largely the same for D3 and D4, except for times when the HMB cannot be defined. We have chosen not to show this plot, as these cases are extremely rare when we include the entire dataset (2.53% of cases). For D4, these cases will be defined by the background model and vary smoothly due to the interpolation in the background model between distinct bins, whereas for D3 (due to the parametrization in RG96), they will be defined as two distinct latitudes, as defined by the model: 60° (96.42% of instances) and 55° (3.57% of instances). Instead of showing the HMB latitude in D3 against D4, Fig. 5d thus shows the HMB latitude against n. It shows that the HMB is likely to move closer to the equator as the number of backscatter echoes increases.

Fig. 5e shows the HMB against AL. We see from this that the HMB is likely to move to lower latitudes as AL decreases, but this trend again breaks down at $\sim$50°. Similarly, in Fig. 5f we see a dependence in the HMB moving to lower latitudes as Sym-H becomes more negative, but this also breaks down at a HMB of 50 to 40°.

Panels d to f all show a seemingly linear trend with HMB, which seems to breaks down at low latitudes. As there are less occurrences for the extreme conditions, however this is difficult to establish.

Similar to previous figures, Fig. 5 shows two example maps in panels g and h, comparing D3 and D4. The map chosen as an example here is one of the best coverage maps, where n was the highest observed value with 1010. We see that having this much data coverage constrains the pattern very well and there are not many differences in the convection patterns: the CPCP only differs by 1 kV, the HMB is the same and the fitted convection potentials only differ very slighly in their morphology (e.g. noon-afernoon sector). This is to be expected, given the data distribution in Fig. 5b.

3.5 Changes to convection mapping since the first SuperDARN radar

Figure 6 provides a further comparison between the RG96 and TS18 datasets. Here we show comparisons between D0 and D4, providing a statistical viewpoint on how much

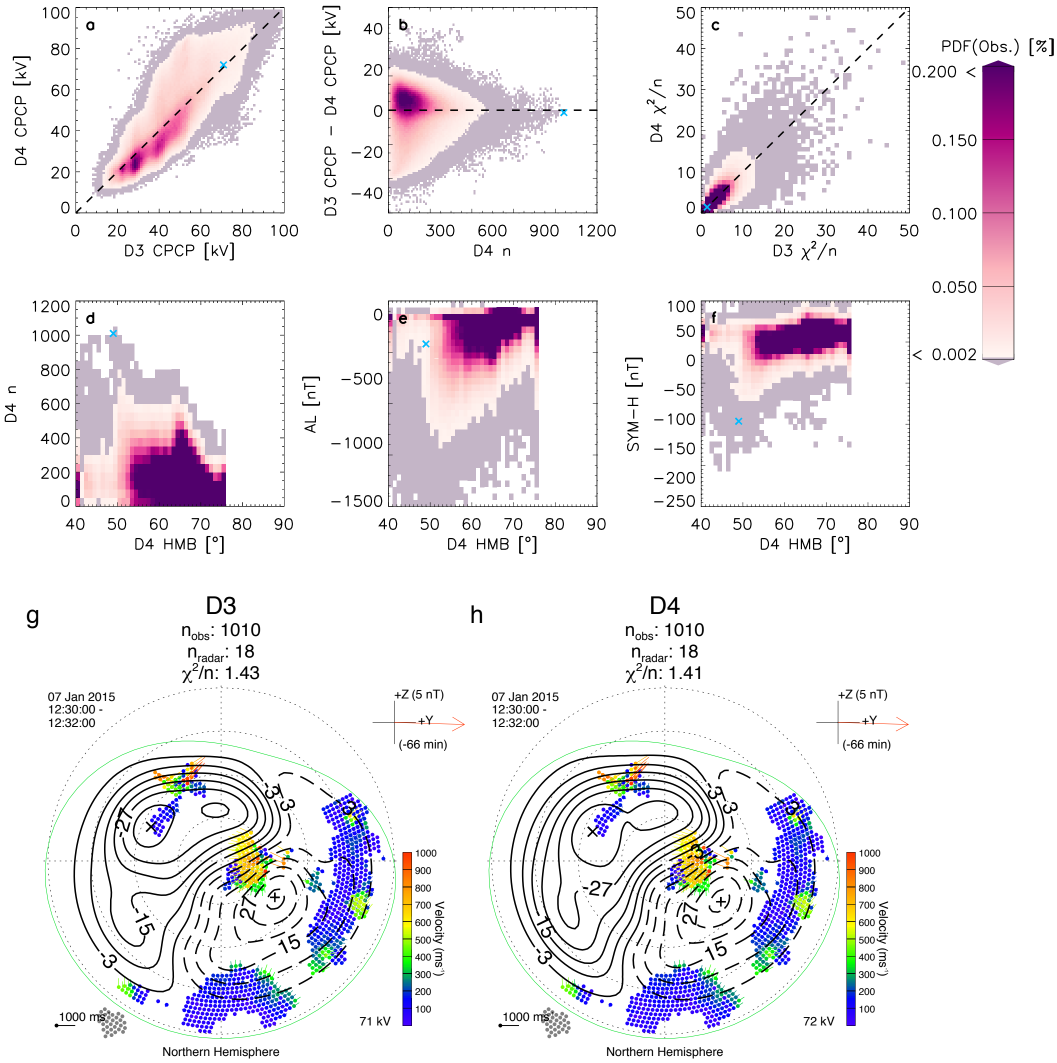

Figure 5. Probability distribution functions comparing the entire D3 and D4 datasets: (a) CPCP, (b) CPCP difference versus number of backscatter echoes, n, (c) χ^2/n distribution, (d) n versus HMB, (e) AL versus HMB, (f) Sym-H versus HMB. The bottom two panels show two example maps from the same date and time (2015/01/07 12:30) for D3 (g) and D4, the convection map which uses TS18 instead of RG96 (h). These occurrences are indicated in the PDFs by blue crosses.

has changed from the original SuperDARN convection map fitting to the most up-to-date version of datasets and fitting methods.

Fig. 6a shows the CPCP distribution. We see that the observed CPCP is on average smaller for D4 than D0 (54.28% of the time), but when the CPCP increases for D4, it increases by more on average (8.37 kV median; 10.45 mean; 92.08 kV maximum change) than it would otherwise decrease (6.87 kV median; 7.64 kV mean; 97.90 kV maximum change). Fig. 6b shows the χ^2/n, which can also increase (56.16%) or decrease (43.80%). By looking further at the statistical distribution, we find that for the times when χ^2/n is larger in D4 than D0, n for D4 tends to small (<200; 102 median; 123.13 mean). Fig. 6c shows the HMB distribution. Interestingly, this shows that the HMB is often higher (43.64% of the time) for D4 than for D0, despite the inclusion of mid-latitude data. This is mostly prominent when the HMB for D0 is above latitudes of 59° (39.76 % of the time), whereas the HMB is less likely to be at lower latitudes for D4 than D0 overall (18.00 % of the time). Fig. 6e shows the distribution of n, which carries a further surprise: n can increase, as well as decrease. We previously speculated that it would only increase, as the changes from D0 to D4 corresponds to the inclusion of polar and mid-latitude radars, but the distribution of n shows that it can also decrease due to the addition of the range limit, although this is less likely (31.63% of the time). The decrease in n scales consistently with n_{D4} and is on average a small change (-34.81 mean; -26.00 median and -349 maximum).

Fig. 6e shows the differences in the CPCP between D4 and D0 against the dayside reconnection rate, Φ_D. We see that the changes in the CPCP tend to be smaller for high solar wind driving (high Φ_D). Similarly, Fig. 6f shows the changes in the HMB against AE and Fig. 6g shows the changes in the HMB against AL. AE and AL, are the auroral electrojet indices, which are derived from ground-based magnetometer measurements and are a proxy for the magnetospheric activity in response to the dayside driving and internal dynamics (Davis & Sugiura, 1966; World Data Center for Geomagnetism in Kyoto et al., 2015). We see from panels f and g that changes in the HMB tend to be smaller when the auroral electrojet indices, AE and AL are enhanced.

Figs. 6h and i show the D4 and D0 HMB against AL. These include yellow and mint crosses that represent the median fits for each HMB bin, allowing us to compare D4 (yellow) with D0 (mint). This shows very clearly that when we use D0, we are less likely to observe a low HMB at enhanced (low) AL, which is not to mean that these occurrences do not exist, but simply that the SuperDARN fitting with the old dataset means we are less likely to observe them.

In Figs. 6j and k, we provide a similar comparison for the D4 and D0 CPCP with respect to Φ_D. This comparison shows that for D4 we are more likely to observe a higher CPCP at higher values of Φ_D than for D0. In fact, at a Φ_D of 100 kV, the median CPCP for D4 is at ~75 kV and ~ 65 kV for D0. We also see that the median curve has a different shape for the two datasets: Both have a logarithmic shape to them and neither appear like a linear fit would suffice to describe the trend in the dataset. Finally in panel l, we show the ratio between the CPCP normalised by Φ_D for both datasets, for which we have also fitted the median per bin (shown by yellow crosses). This shows that the differences between the two versions of the CPCP are proportional to the dayside driving. It also shows that this is a linear trend and that the CPCP changes in D0 with respect to Φ_D are likely to be smaller than for D4.

3.6 Identification of minimum map reliability

When using SuperDARN maps in research, a frequent question is "How reliable is this map?" and often n is used to answer this question. If n is high, the maps are often deemed more reliable, but is there a universal limit for n, which can be used to select reliable convection maps?

To answer this question, we present in Figure 7a the PDF of the difference in χ^2/n between D4 and D0 against the difference in n. It shows that as the map becomes more constrained (i.e. the difference in χ^2/n is negative), the difference in n becomes very small. Similarly, as the difference in χ^2/n becomes larger, the difference in n is also very small. This means that a change in n does not necessarily translate to a better constrained map. In fact, changes in n are more likely to happen for maps that are already well constrained. We see from Fig. 7a and Fig. 6b and d that maps where χ^2/n does not change much tend to have a low χ^2/n to begin with. Figure 7b and c show the difference in χ^2/n versus n in D4 and n in D0. From this we see clearly that the changes in χ^2/n are most extreme when n is small (<100), but there is no clear uniform break-point in n, where χ^2/n is small and the maps are well constrained. We also find that as n increases, χ^2/n is less likely to change. We see that this trend is the same for D4 and D0, however, there is less

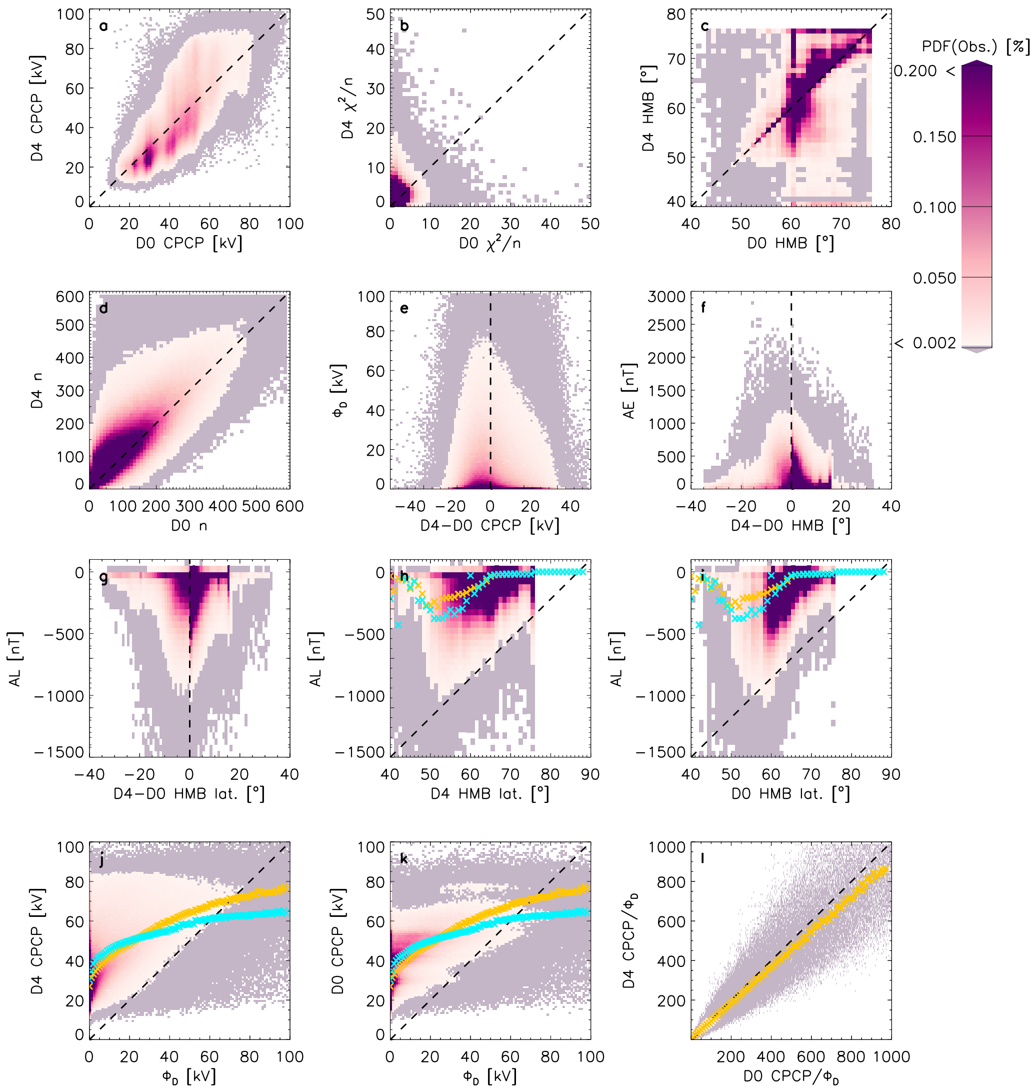

Figure 6. Probability distribution functions comparing the D0 and D4 datasets: (a) CPCP comparison, (b) χ^2/n comparison, (c) HMB comparison, (d) n comparison, (e) Φ_D versus the CPCP difference, (f) AE versus HMB difference, (g) AL versus HMB difference, (h)AL versus D4 HMB and (i) D0 HMB, (j) D4 CPCP versus Φ_D, (k) D0 CPCP versus Φ_D and (l) CPCP normalised by Φ_D. The crosses show the median in the y-direction for each x-bin (where applicable) with the yellow showing the fit for D4 and turquoise showing the fit for D0. Black dashed lines either show the lines of unity or the line at 0.

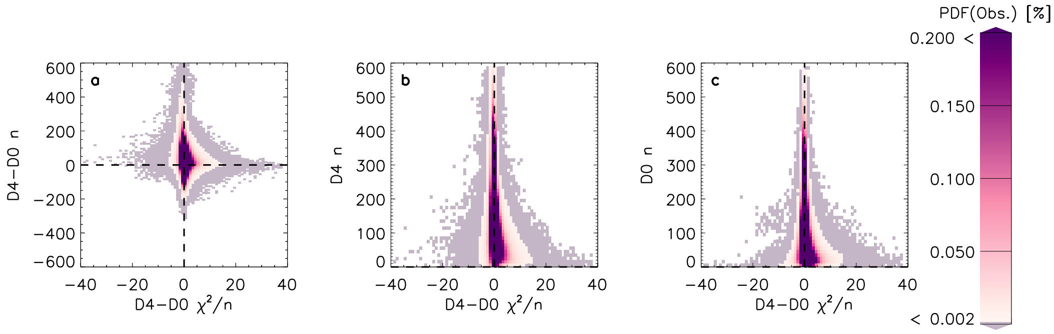

Figure 7. Probability distribution functions comparing the D0 and D4 datasets: The changes in χ^2/n versus (a) the changes in n, (b) D4 n and (c) D0 n. Black dashed lines show the line at 0.

spread and the peak is more pronounced for D0. We also note that the tail in the distributions of D4 n and D0 n versus the difference in χ^2/n are not symmetrical around 0. We will discuss these results further in the following section.

4 Discussion

4.1 How does changing the range limit affect the dataset?

Adding a range limit is intended to remove E-region scatter (i.e. slower moving scatter). This should increase convection in the maps and thus CPCP should increase. It also removes far-range scatter from slant range > 2000 km, which avoids potential errors in geolocation of LOS measurements at far range gates. Whilst this seems like should constrain the SHA solution, Thomas and Shepherd (2018) have shown that the opposite is true for a dataset that is limited in latitudinal coverage: Figure 11 in Thomas and Shepherd (2018) shows how the range limit impacts the data coverage afforded by the high-, polar-, and mid-latitude radars. For example, when data from beyond 2000 km slant range are removed from the high-latitude radar dataset, which is comparable to our D0 to D1 change, then the solution poleward of $\sim$76° magnetic latitude is purely constrained by the statistical model because no measurements are possible. This is to be expected and will be the same for our comparison. Reducing the range-limit will also reduce the number of backscatter echoes in the maps but we also see that the number of backscatter echoes are not solely responsible for map quality.

Chisham and Pinnock (2002) conclude that the contamination from non-F-region scatter does not usually have a large impact on the global characteristics of the SuperDARN convection maps. We find that for the analysed time period, the CPCP is $> 10\%$ different 4.86% of the time and the CPCP is $< 10\%$ different 95.13% of the time. Whilst less than 5% seems like a small set of observations, this does comprise more than 80000 maps, so it may be important on a case-study basis.

Chisham and Pinnock (2002) further showed that removing E-region scatter may not always result in more accurate convection maps. Whilst most E-region scatter is believed to move slower than F-region scatterm, this may not always be the case: Forsythe and Makarevich (2017) used SuperDARN data from the Southern hemisphere and showed that E-Region scatter can be of a similar order of magnitude as F-Region scatter ($\sim$200 m/s or larger). They also showed however that whilst F-Region scatter tends to have a Gaussian velocity profile, the E-Region velocity distribution is highly asymmetric, owing to the Farley-Buneman and gradient drift instabilities being the main drivers. This may be the reason why Chisham and Pinnock (2002) find that removing E-region scatter does not always improve convection maps, but the study by Forsythe and Makarevich (2017) provides clear evidence why removing this scatter makes scientific sense. Our method of adding the range limit follows the strategy of Thomas and Shepherd (2018), though they used this method for statistical convection maps and this may not always be practical for instantaneous convection maps. Whilst the method employed here to removing far range backscatter is a broad-brush approach, future alternatives could include the use of either calibrated elevation angles (which involves measuring the elevation angles using interferometry) or a more accurate virtual height model.

We thus remove both potential E-region scatter and scatter from far range gates. We find that by introducing this range limit, the normalised Chi-squared distribution of the map fitting procedure, χ^2/n is increased 73.61% of the time and decreased 25.54% of the time.

Sometimes, reducing the number of backscatter points by introducing a range limit will increase the HMB to higher latitudes due to removing lower-latitude scatter but more poignantly, this change will reduce E-region scatter at lower-latitudes and thus reduce the probability of choosing a HMB at too low a latitude, as is shown in the example maps in Fig. 1.

For the subset of observations where this is most likely the case (i.e. the difference between the HMB and Λ_{min} are greater in D0 than in D1 and the HMB is at a lower latitude in D0 than in D1), the median n is higher (D0: 128 and D1: 56) than the median for the entire dataset (D0: 93 and D1: 40). Other portions of the dataset which may indicate a worse map contain the population where χ^2/n increases: here, the median n is less (D0: 86 and D1: 38) than the medians for the entire dataset (D0: 93 and D1: 40). Both these statistics suggest, that n is not a good predictor for how good the fit is once the the range limit has been introduced if χ^2/n is used as a quality-of-fit indicator. Alternatively, we suggest that this illustrates a downfall of χ^2/n and that it may not be the perfect indicator for quality. We propose that in the future, a better indicator for map quality is sought.

4.2 How does the addition of the PolarDARN radars affect the dataset?

Adding the polar radars to the dataset increases the coverage, so we would expect the CPCP to be better constrained and n to increase.

We find that adding the PolarDARN radars reduces the CPCP on average, which could indicate that the CPCP is overestimated without good polar cap coverage or that adding PolarDARN causes an underestimation. This has also been shown by Mori et al. (2012), who compared the velocity measurements from PolarDARN radars to CADI ionosonde measurements, as well as comparing the CPCP. Adding the range limit to our processing will remove any slow-moving E-Region scatter, which may increase the CPCP. It is thus more likely that the CPCP is overestimated without good polar cap coverage, as we have added the range limit to our procedure prior to adding PolarDARN radars, which is also shown by the example maps in Fig. 2, as opposed to the latter.

We also find that the difference between the HMB and Λ_{min} either stays the same or tends to increase when the polar radars are added to the dataset. Whilst we would expect PolarDARN measurements mostly to be poleward of the observations from the original high-latitude radars (particularly after introducing the range limit), this does not seem to be the case, which is most likely due to the limited local time observations in these maps. We also see that the HMB tends to stay the same or increase to a higher latitude when adding the polar radars. This indicates that for a number of maps, the HMB was not well defined as we would not expect the introduction of PolarDARN data to move the HMB at all. Whilst this indicates that the HMB was not always necessarily well constrained prior to the introduction of the PolarDARN data, it also indicates that observations near the pole are important in constraining the maps.

Adding the PolarDARN radars to the dataset can increase or decrease χ^2/n. This parameter only tends to increase for D2 if it was low for D1 and tends to decrease for D2 if it was high for D1. This suggests that the maps where the fitting was not particularly good for D1, improve when adding PolarDARN data, but there are also a number of maps where the fit becomes less good. Overall however, we find that the difference between the HMB and Λ_{min} has a tendency to increase, which means the HMB is constrained by data at a lower latitude. The median n increases from 40 to 108 when adding the PolarDARN radars, which is a considerable increase in scatter.

4.3 How does the addition of the StormDARN radars affect the dataset?

Adding StormDARN radars improves the coverage of data at lower latitudes, so we expect HMB to move and CPCP to change.

We find that the mid-latitude radars add less data to the maps (on average), than the polar or high latitude radars, but nevertheless, adding their data to the maps generally improves the dataset. χ^2/n almost always decreases and the HMB tends to be better constrained.

Thomas and Shepherd (2018) made a new baseline model and showed that introducing the mid-latitude radars could increase the CPCP by as much as 40% (for the most strongly southward IMF conditions) due to the high-latitude radars only being able to image a proportion of the convection zone necessary to constrain the CPCP. It is worth noting that Thomas and Shepherd (2018) found very little change in the CPCP for weak to moderate solar wind driving because the low-latitude convection boundary remained within the FOV of the high-latitude radars. We find that, without using the TS18 model, but by simply including the mid-latitude radars, the CPCP does indeed increase more often (12.22% of times) than decrease (7.86% of times) but the maximum change seen is a 45% decrease when the CPCP changes from 34.70 kV in D2 to 19.19 kV in D3.

By investigating the D3 velocity measured at the HMB location of D2, we find that for 33.55% of cases the velocity change is less than 200 m/s, but for a considerable number of maps (7.90%, which equates to over 22000 maps), the velocity change is > 400 m/s at midnight, which indicates a considerable change to the convection pattern. This means that without the mid-latitude radars, the velocities at Λ_{HMB2} could be wrong by more than 190 m/s over half the time at midnight, which is considerable, assuming the HMB placing is constrained by data.

However, we have to consider the possibility that the HMB placing is not always correct: Fig. 3j shows large amounts of low velocity mid-latitude convection in the night-side ionosphere, which does not seem to improve the convection map. We postulate that these streams are associated with magnetic flux frozen into the plasmasphere (the inner part of the magnetosphere located just above the ionosphere) (Ribeiro et al., 2012). As the plasmasphere corotates with Earth, radars should not measure Doppler velocities associated with the rotation due to their fixed geographic location. However, if this co-rotation is not perfectly in sync with Earth's rotation then it may be possible to measure low Doppler velocities (tens-hundreds of ms^{-1}). While more transient in nature, over- or under-shielding scenarios may also lead to errors in the HMB latitude determination when including the mid-latitude radar data (e.g. Nishida, 1968; Nishitani et al., 2019): When this happens, the electric field formed at the inner edge of the plasma sheet and associated with the region 2 field-aligned currents counteracts the effects of the solar wind-driven magnetospheric convection at sub-auroral latitudes. Whilst these scenarios may lead to misidentification of the HMB, they are understood to be exceptional circumstances and not well enough understood to be explicitly taken into account when determining the HMB (Nishitani et al., 2019).

In either case, the HMB may need to be redefined. Currently, the HMB is calculated to be where velocity measurements suggest the electric field is zero, however low velocity measurements associated with imperfect co-rotation will also have an associated non-zero electric field. This suggests the HMB would not give the boundary of the convective regions associated with opening and closing of magnetic flux or that the boundary presents as a gradual change.

Walach and Grocott (2019) showed that during geomagnetic storms, which can also be described as extremely driven times, the HMB can move to latitudes as low as 40°, which SuperDARN radars prior to the mid-latitude expansion were not able to observe. Fogg et al. (2020) provide a fit for the HMB using AMPERE data, and show that the

HMB may be placed at too low latitudes when mid-latitude data are available. This might indicate that a changing HMB is not always an improvement when it moves equatorward in D3. It is however worth noting that the fitting by Fogg et al. (2020) does not include mid-latitude data and their fitting stops at 55°, so further analysis is necessary, which will be the subject of a future study.

Sub-auroral Polarization Streams (SAPS) are one of the main phenomenon studied with the mid-latitude radars (e.g. Kunduri et al., 2017, 2018). They consist of fast azimuthal streams, measured below auroral latitudes on the nightside (Kunduri et al., 2018). The possibility of the midlatitude radars observing either auroral flows in an expanded pattern, or sub-auroral flows in a smaller sized pattern, is an important distinction, which we have not studied in this paper but warrants further investigation. Kunduri et al. (2018) studied these flows in great detail and found that their occurrence and flow speed tends to increase with higher geomagnetic activity. To this date, SAPs have not been explicitly taken into account in the baseline SuperDARN models (e.g. RG96 and TS18) and it is thus likely that their effects are averaged over. We know that SAPs will occur at or near the lower latitudinal boundary of the convection patterns (e.g. Kunduri et al., 2018), but further investigation is necessary to understand how they fit in with the general convection pattern and in particular, how they affect HMB determination.

4.4 How does changing the background model affect the dataset?

When changing the background model from RG96 to TS18 we might expect a better fit due to a background model parametrization with more variables. Thomas and Shepherd (2018) not only use the IMF magnetic field strength and direction, their model parametrization also includes the solar wind's electric field and the Earth's dipole tilt, which results in 120 model bins that are trilinearly interpolated between to achieve smoother transitions, as opposed to the rigid 24 model bins chosen by Ruohoniemi and Greenwald (1996). The χ^2/n distribution indicates that sometimes this expected improvement is the case, however sometimes the fitting is worse, which is primarily the case for low n maps. Overall, we find (in Fig. 5) that the largest changes in the CPCP are produced when the CPCP was already high in D3 and these tend to occur when n is low. In fact, a higher n, means smaller likelihood of observing a change in CPCP. Thomas and Shepherd (2018) compared the changes in the baseline patterns and found that the CPCP can change by as much as 40%, when mid-latitude radars are included in the convection model, which is equivalent to a change of 32 kV for a CPCP of 80 kV without the mid-latitude radars. In comparison, we find that when using this model, the maximum observed percentage change in the CPCP is however a much larger change: a reduction of 63% for a CPCP of 48.84 kV in D3, which reduces to 17.91 kV in D4. The largest increases we see in CPCP when going from D3 to D4 is 59.38 kV, which happens for a CPCP of 59.38 kV in D3 and is a slightly larger change than the smallest decrease (57.11 kV), which happens for a CPCP of 33.41 kV in D3.

Fig. 5 shows that both AL and Sym-H show a linear trend in the likelihood of observations with HMB: As the HMB tends to lower latitudes, the values in AL and Sym-H tend to be enhanced until the HMB reaches a latitude of $\sim$50°, at which point the observational likelihood reduces drastically overall. We also see that at HMBs <50°, n is likely to be smaller in general also, which means the observations in this HMB range are less dense and less well constrained. This is not surprising, as not all radars are capable of measuring HMBs <50°. Furthermore, the coverage from radars at mid-latitudes is sparser as the radars tend to, on average, return less backscatter per radar than the higher latitude radars.

In Fig. 6 we further explore how changing the background model, as well as introducing the newest radars to the dataset, affects the dataset. This shows that the HMB is more likely to be found at lower latitudes (50-40°) for D4 due to the lower observational latitude limit of the data. This means that the HMB is more likely to be observed at lower latitudes when the auroral electrojet indices (AL and AE) are enhanced. It is possible that the observational peak in AL and HMB, which shifts from $\sim$-400nT in D0 to $\sim$-300nT in D4 and $\sim$66° in D0 to $\sim$50° in D4, respectively, is still limited by radar coverage and it is possible that the decreasing trend we see in the median should continue (see crosses in Fig. 6).

The RG96 model was built only using the data from the Goose Bay radar, which is located at a high-latitude and thus part of our D0 set. Whilst it is one of the oldest operating radars in the network (and thus a lot of data is available), the RG96 model was constrained in magnetic latitudes from 65-85° (Ruohoniemi & Greenwald, 1996). It is thus interesting to see χ^2/n reduced, when adding the mid-latitude radars. This shows that the data is important in generating the convection map files, but from comparing D3 and D4 we see that the model can also make a difference. It is however worth noting that due to its limited data ingestion, the RG96 model was not built to be used with a radar network that extends to mid-latitudes, whereas TS18 was. Regardless of the χ^2/n statistic not always decreasing for the change from D3 to D4, the RG96 model does not account for as wide a variety of solar wind driving, dipolar tilt and latitudinal changes of the pattern and it thus makes more sense to use the TS18 model for the extended dataset, especially when including data from the midlatitudes.

4.5 The importance of backscatter echoes

Historically, n has on average increased due to the expansion of SuperDARN. Nevertheless, when we compare our most historic version of the dataset (D0) with the version that includes all new radars, as well as updated processing techniques (TS18 and range limit), we see that sometimes n decreases (Fig. 6d). This is thus solely due to the range limit introduction. Whilst adding the newer radars to the dataset can in some cases increase n by 500 or more, adding a range limit can reduce n by 100. We have shown that n is an important parameter in constraining the convection pattern (e.g. HMB or CPCP): In particular, we find that if n is high, the CPCP is less likely to change (i.e. the maps are constrained well) and the HMB is more likely to be found at lower latitudes (see Fig. 5).

When using SuperDARN maps, the reliability of the map is important and often this has been tied to n. If n is high, the maps are often deemed more reliable (e.g. Imber et al. (2013) identified 200 to be a low threshold number for good convection maps but Fogg et al. (2020) chose 400 as threshold for an acceptable number of backscatter echoes). This raises the question of whether there is a universal threshold for n, which can be used to select reliable convection maps?

We show that when n changes by large amounts (>200), the maps tend to be already well constrained (χ^2/n changes by $\sim$10), but we also find that when n is large in D0 and D4, χ^2/n is unlikely to change by much, which means the map is well constrained (see Fig. 7). The in-between state, where n changes, but not by large amounts, contains the maps that are the least well constrained (χ^2/n changes by up to 40). As n approaches $\sim$200, χ^2/n is likely to vary by <20 and as n approaches $\sim$400, the changes in χ^2/n are approximately halved. For higher values of n (>400), the probability of observing a change in χ^2/n remain the same. We see that this trend is the same for D4 and D0, however, there is less spread and the peak is more pronounced for D0. This means that whether or not a threshold of 200 or 400 is chosen for D0 makes minimal difference to how well the map is constrained. There is no clear break, where n universally produces good convection maps, but we show that if we choose n >400, χ^2/n is unlikely to change by much and thus the map is as well constrained as it can be.

We also see from Fig. 7b-c that the spread of observations about 0 is not symmetrical. The left side of both distributions falls off much more abruptly than the right side, which implies that χ^2/n is larger in D4 than in D0 much more often and thus, for small n, the maps are less well-constrained for D4 than D0. This could be due to a number of reasons, but we suggest one main cause: D4 includes data over a larger spatial range but for a sixth order SHA, only 49 vectors are required to constrain it. As more vectors are added (e.g. from the midlatitude radars), more small-scale variability is added, which the 6th order SHA cannot resolve.

4.6 Geomagnetic conditions and SuperDARN observations

We have shown in Fig. 5d to f that when n is high, AL and Sym-H tend to enhance also and the HMB also tends to move to lower latitudes. It is worth considering the underlying physics and how these parameters are related as a result.

The expanding and contracting polar cap paradigm (e.g. Siscoe & Huang, 1985; Lockwood, 1991; Lockwood & Cowley, 1992; Milan, 2015; Walach et al., 2017, and references therein) requires the polar cap to increase in size when the dayside reconnection rate exceeds the nightside reconnection rate. This implies that the CPCP also increases when dayside driving is high. We have shown that this is mostly the case, although there are some deviations to this relationship, which we attribute to noise and errors in solar wind propagation. It has long been discussed whether or not the relationship between the dayside driving and the CPCP is linear and whether or not the CPCP saturates beyond a threshold (e.g. Hill et al., 1976; Reiff et al., 1981; Doyle & Burke, 1983; Wygant et al., 1983; Shepherd, 2007; Mori & Koustov, 2013, and references therein). Shepherd et al. (2002) and Shepherd (2007) discuss this in great detail and showed, using SuperDARN CPCP measurements, that during high solar wind driving (when the reconnection electric field is above 5.5 mV/m), the CPCP saturates.

Mori and Koustov (2013) talk about a SuperDARN "quantization" effect, whereby for high CPCP where the observational density is low and not all maps are well constrained, the CPCP oftentimes takes on the values of the underlying model (e.g. RG96). We see this quantization to some extent in Fig. 6 for RG96, but this problem is solved for TS18, which interpolates between solutions of the background model. Whilst this is not the focal point of our study, we find that as Φ_D increases, the CPCP also increases. Similar to Shepherd (2007), we note that observational density is an important factor when considering the behaviour of these parameters. We also find that depending on the dataset used (e.g. D0 or D4), the trend and steepness of the curve varies due to observational density of high CPCP for D0 being much lower than for D4. Furthermore, we find that the spread in values is much higher than observed by Shepherd (2007), which is due to a larger sample size (they only used equinox data for their study) and shorter sampling (they used 10 minute cadence for their map files whereas we use 2 minutes). We suggest that using the verb "saturate" to describe the behaviour of these parameters is misplaced, as even at high values of Φ_D the CPCP increases, whereas a saturation implies the gradient of the curve reaching 0.

Whilst n is high when AL, Sym-H and the HMB are enhanced, we are not suggesting that the correlation equates to a causal link. This was already discussed by Walach and Grocott (2019), who showed that the number of backscatter echoes tends to increase during geomagnetic storms (when Sym-H is enhanced), as dayside driving increases, the polar cap grows and the HMB moves to lower latitudes. Currie et al. (2016) showed however that during intense geomagnetic storms, a reduction of backscatter was observed in the Bruny Island radar in the middle- to far-ranges, and an increase in the amount of backscatter from close-ranges. Here we show statistically, that as Sym-H is enhanced, the HMB moves to lower latitudes and the number of backscatter echoes increases for mid-ranges (the far- and close- ranges were removed beyond D0 by the range limit). We thus find that the relationships found by Walach and Grocott (2019) hold statistically, though a large amount of variation is observed.

Wild and Grocott (2008) conducted a study (before the availability of mid-latitude radars) of regions where backscatter is lost during isolated substorms, and the progression through the phases of the substorm due to auroral absorption. They identify that backscatter reduction is greatest at ∼70-80° magnetic latitude region between ∼19 to 03 MLT. However, Wild and Grocott (2008) also observe that the main backscatter region shifts equatorward to lower latitudes (below ∼65°) across all local times. Our results support this statistically, as we find that the mid-latitude radars do on average observe more backscatter, and that the backscatter moves to lower latitudes when AL is enhanced (which is expected to be the case for substorms). We also find that this trend differs slightly for D0 and D4: due to better coverage with the mid-latitude radars, the HMB for D4 moves to lower latitudes than for D0. The trend of decreasing HMB with decreasing AL is a statistical one and thus breaks at a latitudes close to ∼40° due to low observational densities

5 Summary

We have investigated how the SuperDARN maps have changed historically by creating 5 different versions of the convection map files for a timespan of 6 years and comparing them statistically. By using different processing parameters and gradually introducing more data to the maps, we were able to investigate how the dataset changes with the inclusion of

- a backscatter range limit (as was used by Thomas and Shepherd (2018))
- the polar cap radars, PolarDARN
- the mid-latitude radars, StormDARN
- a different statistical background model (we compare Thomas and Shepherd (2018) and Ruohoniemi and Greenwald (1996))

We have shown that

- introducing a range limit does not always decrease χ^2/n,
- n is not a good predictor for how good the fit is once the range limit has been applied
- once the range limit has been applied the CPCP stays the same 29.71% of the time and the HMB stays constant most of the time (54.47%)
- the addition of PolarDARN data tends to reduce the CPCP,
- PolarDARN radars add the most data to the dataset (on average), but the mid-latitude radars are also important for constraining the maps,
- when introducing StormDARN radars to the maps, the χ^2/n values tend to decrease, the HMB becomes better constrained and the CPCP tends to increase

- when changing the background model to TS18, the CPCP tends to decrease for lower values of the CPCP in RG96, but is more likely to increase for larger values of the CPCP in RG96. If n is however high (> 400), the CPCP is less likely to change (changes $\sim$<20 kV).
- as n, AL and Sym-H all increase, the HMB tends to go to lower latitudes, which appears to be a linear trend, though a break is seen at HMB $\sim$50 degrees, where the observational density drops off sharply.
- if n is high, the CPCP is less likely to change and the HMB is more likely to be found at lower latitudes and χ^2/n tends to change by the least amount,
- there is no clear break, where n universally produces good convection maps, but we show that for $n > 400$, χ^2/n is unlikely to change by much and thus the map is as well constrained as it can be.

Naturally, assessing map quality has to include a qualitative discussion and there is currently no perfect quantitative method for this assessment. The current most simple way to assess map quality is to look at the χ^2/n statistic. If we sum χ^2 and divide by the sum of n for each dataset D0 to D4, we obtain the following average values: $<\chi^2/n>_{D0}$: 1.70; $<\chi^2/n>_{D1}$: 2.01; $<\chi^2/n>_{D2}$: 2.16; $<\chi^2/n>_{D3}$: 1.88; and $<\chi^2/n>_{D4}$: 1.81.

From this, we might conclude that D0 has overall the highest quality maps and is closest to the "good match" criterion (1) identified by Ruohoniemi and Baker (1998), but we have shown that whilst the map fitting may be better for D0, the missing data also equates to a qualitative penalty. We see from these values that most of the impact on χ^2/n are provided by the range limit and the addition of the mid-latitude radar data. This emphasizes the importance of good spatial coverage. We also see from these statistics, that overall, the TS18 model improves map fitting.

Overall, we have shown that the measured parameters (such as the CPCP and HMB) are highly susceptible to which processing parameters are used, as well as which radars are used when generating map files. This becomes particularly important when SuperDARN maps are used for studies of specific conditions or small case studies as a sampling bias can occur. A high number of SuperDARN backscatter echoes are particularly important when constraining maps, so it is important to include mid-latitude data in the generation of SuperDARN convection maps. We have also shown that the method of selecting the HMB is not always perfect and further work is necessary to generate a robust selection method, especially at lower latitudes.

Appendix A SuperDARN processing parameters

In the SuperDARN processing (see section 2), we use the following parameters and functions from RST:

- For fitting the autocorrelation function to the raw data: 'make_fit' with the option '-fitacf-version 2.5'.
- To make the gridded map files, the options '-i 120 -tl 120 -chisham -c' were added to 'make_grid'
- To add the range limit to the gridded files, the same options as above were used but in addition, the options '-minsrng 800 -maxsrng 2000' were added.
- The function 'map_grd' was used with 'map_addhmb -vel 100 -cnt 3'. Adding these options to 'map_addhmb' chooses the Heppner-Maynard boundary to the lowest possible latitude for which a minimum of three LOS vectors with velocities greater than 100 m/s lie along its boundary.
- To make the convection maps, we also use 'map_addimf -if' with the text file containing the IMF data and the option '-df' with the text file containing the IMF delay times.
- We then use 'map_addmodel -o 6 ' for a sixth order expansion and use '-d' to specify a light doping level.
- Finally, we add the model option '-rg96' to D0-D3 and '-ts18' to D4 and use the function 'map_fit' to make the convection map files.
- We also use the function 'cnvmaptomap' to convert the binary file to ASCII format and 'trim_map' with the options '-st', '-et', '-sd' and '-ed' to make two-hour long map files for our archive, but this is not necessary to obtain the results for this study.

Acknowledgments

All data used for this study are available opensource from nonprofit organizations. The authors acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Japan, South Africa, United Kingdom, and United States of America, and we thank the international PI team for providing the data. The authors acknowledge access to the SuperDARN database via the British Antarctic Survey (https://www.bas.ac.uk/project/superdarn/#data). Other data mirrors are hosted by the Virginia Tech SuperDARN group (http://vt.superdarn.org/) and the University of Saskatchewan (https://superdarn.ca/data-download). The Radar Software Toolkit (RST) to process the SuperDARN data can be downloaded from Zenodo (https://doi.org/10.5281/zenodo.1403226 and references). All solar wind data and geomagnetic indices were downloaded from NASA's SPDF Coordinated Data Analysis Web (https://cdaweb.gsfc.nasa.gov/index.html/). The AE data is also available from the WDC for Geomagnetism, Kyoto (http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html) who prepared this index. M.-T. W. and A. G. were supported by Natural Environments Research Council (NERC), UK, grant nos. NE/P001556/1 and NE/T000937/1. F. S. was supported by a Science and Technology Funding Council (STFC) studentship. E. G. T. thanks the National Science Foundation (NSF) for support under grants AGS-1934997 and OPP-1836426. We gratefully acknowledge the use of Lancaster University's High End Computing Cluster, which has facilitated the necessary dataprocessing for this study. M.-T. W. would like to thank LU's Women's Network Writing Group for providing a supportive virtual writing space and mentorship, which helped to forge this paper.

References

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., ... Walker, D. M. (2007). A decade of the Super Dual Auroral Radar Network (SuperDARN): Scientific achievements, new techniques and future directions. *Surveys in Geophysics*, *28*(1), 33–109. doi: 10.1007/s10712-007-9017-8

Chisham, G., & Pinnock, M. (2002). Assessing the contamination of SuperDARN global convection maps by non-F-region backscatter. *Annales Geophysicae*, *20*(1), 13-28. Retrieved from `https://hal.archives-ouvertes.fr/hal-00316917`

Chisham, G., Yeoman, T. K., & Sofko, G. J. (2008). Mapping ionospheric backscatter measured by the superdarn hf radars – part 1: A new empirical virtual height model. *Annales Geophysicae*, *26*(4), 823–841. doi: 10.5194/angeo-26-823-2008

Cousins, E. D. P., & Shepherd, S. G. (2010). A dynamical model of high-latitude convection derived from superdarn plasma drift measurements. *Journal of Geophysical Research: Space Physics*, *115*(A12). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JA016017 doi: 10.1029/2010JA016017

Currie, J. L., Waters, C. L., Menk, F. W., Sciffer, M. D., & Bristow, W. A. (2016). Superdarn backscatter during intense geomagnetic storms. *Radio Science*, *51*(6), 814-825. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016RS005960 doi: 10.1002/2016RS005960

Davis, T. N., & Sugiura, M. (1966). Auroral electrojet activity index ae and its universal time variations. *Journal of Geophysical Research (1896-1977)*, *71*(3), 785-801. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ071i003p00785 doi: https://doi.org/10.1029/JZ071i003p00785

Doyle, M. A., & Burke, W. J. (1983). S3-2 measurements of the polar cap potential. *Journal of Geophysical Research*, *88*(A11), 9125. doi: 10.1029/JA088iA11p09125

Fogg, A. R., Lester, M., Yeoman, T. K., Burrell, A. G., Imber, S. M., Milan, S. E., … Anderson, B. J. (2020). An improved estimation of superdarn heppner-maynard boundaries using ampere data. *Journal of Geophysical Research: Space Physics*, *125*(5), e2019JA027218. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JA027218 (e2019JA027218 10.1029/2019JA027218) doi: https://doi.org/10.1029/2019JA027218

Forsythe, V. V., & Makarevich, R. A. (2017). Global view of the e region irregularity and convection velocities in the high-latitude southern hemisphere. *Journal of Geophysical Research: Space Physics*, *122*(2), 2467-2483. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JA023711 doi: https://doi.org/10.1002/2016JA023711

Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., … Yamagishi, H. (1995). Darn/superdarn. *Space Science Reviews*, *71*(1), 761–796. doi: 10.1007/BF00751350

Heppner, J. P., & Maynard, N. C. (1987). Empirical high-latitude electric field models. *Journal of Geophysical Research*, *92*(A5), 4467–4489. doi: 10.1029/

JA092iA05p04467

Hill, T. W., Dessler, A. J., & Wolf, R. A. (1976). Mercury and mars: The role of ionospheric conductivity in the acceleration of magnetospheric particles. *Geophysical Research Letters*, *3*(8), 429-432. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL003i008p00429 doi: https://doi.org/10.1029/GL003i008p00429

Imber, S. M., Milan, S. E., & Lester, M. (2013). The heppner-maynard boundary measured by superdarn as a proxy for the latitude of the auroral oval. *Journal of Geophysical Research: Space Physics*, *118*(2), 685-697. doi: 10.1029/2012JA018222

Iyemori, T. (1990). Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations. *Journal of geomagnetism and geoelectricity*, *42*(11), 1249-1265. doi: 10.5636/jgg.42.1249

Khan, H., & Cowley, S. W. H. (1999, Sep 01). Observations of the response time of high-latitude ionospheric convection to variations in the interplanetary magnetic field using eiscat and imp-8 data. *Annales Geophysicae*, *17*(10), 1306–1335. doi: 10.1007/s00585-999-1306-8

Kunduri, B. S. R., Baker, J. B. H., Ruohoniemi, J. M., Nishitani, N., Oksavik, K., Erickson, P. J., … Miller, E. S. (2018). A new empirical model of the subauroral polarization stream. *Journal of Geophysical Research: Space Physics*, *123*(9), 7342-7357. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JA025690 doi: https://doi.org/10.1029/2018JA025690

Kunduri, B. S. R., Baker, J. B. H., Ruohoniemi, J. M., Thomas, E. G., Shepherd, S. G., & Sterne, K. T. (2017, June). Statistical characterization of the large-scale structure of the subauroral polarization stream. *J. Geophys. Res-Space Phys.*, *122*(6), 6035–6048. doi: 10.1002/2017JA024131

Lockwood, M. (1991). On flow reversal boundaries and transpolar voltage in average models of high-latitude convection. *Planetary and Space Science*, *39*(3), 397–409. doi: 10.1016/0032-0633(91)90002-r

Lockwood, M., & Cowley, S. W. H. (1992). Ionospheric convection and the substorm cycle. In *International conference on substorms.*

Milan, S. E. (2015). Magnetospheric Plasma Physics: The Impact of Jim Dungey's

Research. In D. Southwood, S. W. H. Cowley FRS, & S. Mitton (Eds.), *Magnetospheric plasma physics: The impact of jim dungey's research* (pp. 1–271). doi: 10.1007/978-3-319-18359-6_2

Milan, S. E., Gosling, J. S., & Hubert, B. (2012, mar). Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap. *Journal of Geophysical Research*, *117*(A3), A03226. doi: 10.1029/2011JA017082

Mori, D., & Koustov, A. (2013). Superdarn cross polar cap potential dependence on the solar wind conditions and comparisons with models. *Advances in Space Research*, *52*(6), 1155-1167. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0273117713003803` doi: https://doi.org/10.1016/j.asr.2013.06.019

Mori, D., Koustov, A. V., Jayachandran, P. T., & Nishitani, N. (2012). Resolute bay cadi ionosonde drifts, polardarn hf velocities, and cross polar cap potential. *Radio Science*, *47*(3). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011RS004947` doi: https://doi.org/10.1029/2011RS004947

Nishida, A. (1968). Coherence of geomagnetic dp 2 fluctuations with interplanetary magnetic variations. *Journal of Geophysical Research (1896-1977)*, *73*(17), 5549-5559. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA073i017p05549` doi: https://doi.org/10.1029/JA073i017p05549

Nishitani, N., Ruohoniemi, J. M., Lester, M., Baker, J. B. H., Koustov, A. V., Shepherd, S. G., ... Kikuchi, T. (2019). Review of the accomplishments of mid-latitude super dual auroral radar network (superdarn) hf radars. *Progress in Earth and Planetary Science*, *6*(1), 27. doi: 10.1186/s40645-019-0270-5

Pettigrew, E. D., Shepherd, S. G., & Ruohoniemi, J. M. (2010). Climatological patterns of high-latitude convection in the northern and southern hemispheres: Dipole tilt dependencies and interhemispheric comparisons. *Journal of Geophysical Research: Space Physics*, *115*(A7). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JA014956` doi: 10.1029/2009JA014956

Press, W. H. and Teukolsky, S. A. and Vetterling W. T. and Flannery B. P. (2007).

Numerical recipes: The art of scientific computing. Cambridge University Press.

Reiff, P. H., Spiro, R. W., & Hill, T. W. (1981). Dependence of polar cap potential drop on interplanetary parameters. *Journal of Geophysical Research*, *86*(A9), 7639–7648. doi: 10.1029/JA086iA09p07639

Ribeiro, A. J., Ruohoniemi, J. M., Baker, J. B. H., Clausen, L. B. N., Greenwald, R. A., & Lester, M. (2012). A survey of plasma irregularities as seen by the midlatitude blackstone superdarn radar. *Journal of Geophysical Research: Space Physics*, *117*(A2). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JA017207` doi: https://doi.org/10.1029/2011JA017207

Ruohoniemi, J. M., & Baker, K. B. (1998). Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations. *Journal of Geophysical Research*, *103*(A9), 20797. doi: 10.1029/98JA01288

Ruohoniemi, J. M., & Greenwald, R. A. (1996). Statistical patterns of high-latitude convection obtained from Goose Bay HF radar observations. *Journal of Geophysical Research*, *101*(A10), 21743. Retrieved from `http://doi.wiley.com/10.1029/96JA01584` doi: 10.1029/96JA01584

Ruohoniemi, J. M., & Greenwald, R. A. (2005). Dependencies of high-latitude plasma convection: Consideration of interplanetary magnetic field, seasonal, and universal time factors in statistical patterns. *Journal of Geophysical Research: Space Physics*, *110*(A09204). doi: 10.1029/2004JA010815

Shepherd, S. G. (2007). Polar cap potential saturation: Observations, theory, and modeling. *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*(3), 234-248. Retrieved from `https://www.sciencedirect.com/science/article/pii/S136468260600263X` (Global Aspects of Magnetosphere-Ionosphere Coupling) doi: https://doi.org/10.1016/j.jastp.2006.07.022

Shepherd, S. G. (2014). Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations. *Journal of Geophysical Research: Space Physics*, *119*(9), 7501-7521. doi: 10.1002/2014JA020264

Shepherd, S. G., Greenwald, R. A., & Ruohoniemi, J. M. (2002). Cross polar cap potentials measured with super dual auroral radar network during quasi-steady solar wind and interplanetary magnetic field conditions. *Journal of Geo-*

physical Research: Space Physics, *107*(A7), SMP 5-1-SMP 5-11. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JA000152 doi: https://doi.org/10.1029/2001JA000152

Shepherd, S. G., & Ruohoniemi, J. M. (2000). Electrostatic potential patterns in the high-latitude ionosphere constrained by superdarn measurements. *Journal of Geophysical Research: Space Physics*, *105*(A10), 23005-23014. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JA000171 doi: 10.1029/2000JA000171

Siscoe, G. L., & Huang, T. S. (1985). Polar cap inflation and deflation. *Journal of Geophysical Research*, *90*(A1), 543–547.

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Ponomarenko, P. V., Billett, D. D., Bland, E. C., Burrell, A. G., ... Walach, M.-T. (2018, August). *Superdarn radar software toolkit (rst) 4.2.* Retrieved from https://doi.org/10.5281/zenodo.1403226 doi: 10.5281/zenodo.1403226

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Ponomarenko, P. V., Bland, E. C., Burrell, A. G., Kotyk, K., ... Walach, M.-T. (2018, January). *Superdarn radar software toolkit (rst) 4.1.* Retrieved from https://doi.org/10.5281/zenodo.1143675 doi: 10.5281/zenodo.1143675

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Sterne, K. T., Shepherd, S. G., Kotyk, K., Schmidt, M. T., ... Billett, D. D. (2019, September). *Superdarn radar software toolkit (rst) 4.3.* Zenodo. Retrieved from https://doi.org/10.5281/zenodo.3401622 doi: 10.5281/zenodo.3401622

Thomas, E. G., & Shepherd, S. G. (2018, apr). Statistical Patterns of Ionospheric Convection Derived From Mid-latitude, High-Latitude, and Polar SuperDARN HF Radar Observations. *Journal of Geophysical Research: Space Physics*, *123*(4), 3196–3216. Retrieved from http://doi.wiley.com/10.1002/2018JA025280 doi: 10.1002/2018JA025280

Walach, M.-T., & Grocott, A. (2019). Superdarn observations during geomagnetic storms, geomagnetically active times, and enhanced solar wind driving. *Journal of Geophysical Research: Space Physics*, *124*(7), 5828-5847. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JA026816 doi: 10.1029/2019JA026816

Walach, M.-T., Milan, S. E., Yeoman, T. K., Hubert, B. A., & Hairston, M. R.

(2017). Testing nowcasts of the ionospheric convection from the expanding and contracting polar cap model. *Space Weather*, *15*(4), 623-636. doi: 10.1002/2017SW001615

Wild, J. A., & Grocott, A. (2008). The influence of magnetospheric substorms on superdarn radar backscatter. *Journal of Geophysical Research: Space Physics*, *113*(A4). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JA012910 doi: 10.1029/2007JA012910

World Data Center for Geomagnetism in Kyoto, Nose, M., Iyemori, T., Sugiura, M., & Kamei, T. (2015). *Geomagnetic ae index.* doi: 10.17593/15031-54800

Wygant, J. R., Torbert, R. B., & Mozer, F. S. (1983). Comparison of s3-3 polar cap potential drops with the interplanetary magnetic field and models of magnetopause reconnection. *Journal of Geophysical Research: Space Physics*, *88*(A7), 5727-5735. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA088iA07p05727 doi: https://doi.org/10.1029/JA088iA07p05727

Super Dual Auroral Radar Network Expansion and its Influence on the Derived Ionospheric Convection Pattern

By Maria-Theresia Walach

WORD COUNT	16300	TIME SUBMITTED	19-MAY-2021 02:47PM
		PAPER ID	72331383

Super Dual Auroral Radar Network Expansion and its Influence on the Derived Ionospheric Convection Pattern

M.-T. Walach[1], A. Grocott[1], F. Staples[2], E. G. Thomas[3]

[1]Lancaster University, Lancaster, LA1 4YW, UK
[2]Mullard Space Science Laboratory, University College London, Holmbury St. Mary, RH5 6NT, UK
[3]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA

Key Points:

- We identify changes in measurements when high- and mid-latitude radars are added to SuperDARN, and show the impact of different processing
- Measured convection parameters are highly susceptible to processing parameters and which radars are used
- We show how the number of backscatter echoes per map is critical to the convection maps, and discuss how this impacts map quality

Corresponding author: M.-T. Walach, `m.walach@lancaster.ac.uk`

Abstract

The Super Dual Auroral Radar Network (SuperDARN) was built to study ionospheric convection and has in recent years been expanded geographically. Alongside software developments, this has resulted in many different versions of the convection maps dataset being available. Using data from 2012 to 2018, we produce five different versions of the widely used convection maps, using limited backscatter ranges, background models and the exclusion/inclusion of data from specific radar groups such as the mid-latitude radars. This enables us to simulate how much information was missing from previous decades of SuperDARN research. We study changes in the Heppner-Maynard boundary, the cross polar cap potential (CPCP), the number of backscatter echoes (n) and the χ^2/n statistic which is a measure of the global agreement between the measured and fitted velocities. We find that the CPCP is reduced when the polar cap radars are introduced, but then increases again when the mid-latitude radars are added. When the background model is changed from the RG96 model, to the most recent TS18 model, the CPCP tends to decrease for lower values, but tends to increase for higher values. When comparing to geomagnetic indices, we find that there is on average a linear relationship between the Heppner-Maynard boundary and the geomagnetic indices, as well as n, which breaks at high values (e.g. HMB $\sim$50°) due to the low observational density. We find that whilst n is important in constraining the maps (maps with n >400 are unlikely to change), is insufficient as the sole measure of quality.

Plain Language Summary

The ionosphere, where space begins and the atmosphere ends, moves as a result of the Earth's magnetic field coupling with the Sun. The Super Dual Auroral Radar Network (SuperDARN) was built around the Earth's magnetic poles to study this phenomenon, known as ionospheric convection. Combining many line-of-sight convection measurements, we are able to build global maps of ionospheric convection using SuperDARN. This encapsulates dynamics which are central to space weather phenomena. SuperDARN, which has been gathering data for decades, has over time undergone numerous transformations, including the development of new processing software and more radars being added to the network. Using data from the years 2012 to 2018, we perform a statistical analysis on processed SuperDARN convection maps for the entire dataset and assess systematically how the dataset has changed over the years. We consider how the addition of more data and changes to the convection mapping procedures can affect scientific studies in the context of this large database.

1 Introduction

The Super Dual Auroral Radar Network (SuperDARN) consists of high-frequency coherent scatter radars built to study ionospheric convection by means of Doppler-shifted, pulse sequences and has been widely used in space physics and ionospheric research (e.g. Greenwald et al., 1995; Ruohoniemi & Greenwald, 1996; Chisham et al., 2007; Nishitani et al., 2019). SuperDARN data are continuously available since 1993, with the network having expanded over time from one radar (built in 1983) to 23 radars in the Northern hemisphere, 13 in the Southern hemisphere and more under construction (Nishitani et al., 2019). This expansion has allowed for a greater area to be covered by SuperDARN (i.e. down to magnetic latitudes of 40°) with at least 16 different azimuthal look directions (Nishitani et al., 2019) in the Northern hemisphere. Line-of-sight measurements by this large-scale network of radars can be combined and used to construct a picture of high-latitude ionospheric convection on time scales of 1-2 minutes (Ruohoniemi & Baker, 1998). The radars can be grouped into high-latitude radars, polar-latitude radars (or PolarDARN), and mid-latitude radars (or StormDARN). Nishitani et al. (2019) provides a summary from a historical northern hemisphere perspective: high-latitude radars, at magnetic latitudes of 50-70° were first built, starting in 1983 with the Goose Bay radar, followed by the polar radars (covering 70-90° magnetic latitude), and the expansion to mid-latitudes (∼40-50°), starting in 2005 with the Wallops Island radar. Over time new radars have improved global ionospheric convection mapping by increasing the number of measurements and look directions.

The most commonly used SuperDARN data product by the space science and ionospheric research community is the convection maps. Convection maps are large scale maps, showing ionospheric convection around the magnetic poles. In order to produce these maps, several data processing steps have to be undertaken. With the expansion of the dataset, as well as data processing software improvements, this data product has undergone several changes.

To make SuperDARN convection maps the raw data is processed using the Radar Software Toolkit (RST (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Bland, et al., 2018)):

1. An autocorrelation function is fitted to the raw radar data. This produces fitacf files, which store the line-of-sight velocity data.
2. The data is then gridded onto an equal area latitude-longitude grid (see equation 1 from Ruohoniemi & Baker, 1998) and split into typically one or two minute cadence records. Historically it has almost always been the case that all data from the radars were added to the grids. However, slow moving E-region scatter can and should be removed by setting the minimum range gate limit to 800 km (Forsythe & Makarevich, 2017; Thomas & Shepherd, 2018). It has recently become apparent that far range data beyond 2000 km can also be problematic owing to geolocation uncertainties in the range finding algorithm (Chisham et al., 2008).
3. Data from different radars are combined and the spherical harmonic fitting algorithm is applied which fits an electrostatic potential in terms of spherical harmonic functions to the data (Ruohoniemi & Greenwald, 1996; Ruohoniemi & Baker, 1998). To find the optimal solution for the spherical harmonic coefficients, a singular value decomposition (e.g. Press, W. H. and Teukolsky, S. A. and Vetterling W. T. and Flannery B. P., 2007) is minimised. When this fitting is performed, typically a background model, parameterised by solar wind conditions is used, to infill information in the case of data gaps. This method is also known as 'Map Potential' technique.

Several models are available for the fitting in step 3, most notably Ruohoniemi and Greenwald (1996) generated the most widely used statistical background model, which was subsequently implemented in the RST. This background model was thus used by most SuperDARN users to generate convection maps and used in many scientific studies. Ruohoniemi and Greenwald (1996) used the Goose Bay radar to create the background statistical model. Since then, however many more radars have been added to SuperDARN. This raises the question of how much of an effect changing the background model has on the convection map dataset, which was investigated by Shepherd and Ruohoniemi (2000). The main conclusion from Shepherd and Ruohoniemi (2000) was that the solution becomes insensitive to the choice of statistical model when the data coverage is high. Since then, Ruohoniemi and Greenwald (2005) produced an updated version of their statistical background model using data from 9 radars, but this was not implemented into RST, thus keeping the RG96-model the default which was used by the community. Since then, a number of updated background models, such as Pettigrew et al. (2010), Cousins and Shepherd (2010) and Thomas and Shepherd (2018) have been produced. The Pettigrew et al. (2010) and Cousins and Shepherd (2010) models were not implemented into RST until version 4.1 (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Bland, et al., 2018). Soon after, the statistical background model by Thomas and Shepherd (2018) was released, which is now standard in RST since version 4.2 (SuperDARN Data Analysis Working Group, Thomas, Ponomarenko, Billett, et al., 2018). The RG96 and TS18 models are thus the most widely used and we will focus our analysis on these background models.

Alongside the use of a background model, a Heppner-Maynard boundary (HMB) (Heppner & Maynard, 1987), the low-latitude boundary of the convection pattern where the flows approach zero, can either be specified or be chosen using backscatter measurements. This is to constrain the convection pattern when the spherical harmonic fit is applied (Shepherd & Ruohoniemi, 2000). For typical two minute cadence convection maps, it is appropriate to find where three radar velocity measurements are greater than 100ms^{-1} for the HMB (Imber et al., 2013). This boundary is circular around the nightside and cropped at the dayside to mimic the shape of the dayside magnetopause. Previous to Shepherd and Ruohoniemi (2000) however, a fully circular boundary was used, which was deemed to create unrealistic flows at lower latitudes when the radar network was expanded.

In this paper we conduct a large scale data analysis to assess systematically how the SuperDARN dataset has changed over the years and how this may have affected the dataset overall.

We specifically probe the effects of the following changes:

1. Inclusion of the backscatter range limits
2. Addition of the PolarDARN data
3. Addition of the StormDARN data
4. Updating of the background statistical model

2 Data and Method

To provide a meaningful large scale comparison of different versions of the SuperDARN dataset, we process Northern hemisphere data from the same time period (2012-2018) and create different versions of the SuperDARN convection maps. First, we create a baseline dataset (D0) with the high-latitude radars only, which is then modified by changing one aspect for each subsequent dataset. This allows us to contrast the changes in the dataset. Table 1 outlines the different datasets (D0 to D4) and how each one varies from the previous iteration. The basic data processing is the same for all the datasets, except with the changes outline in table 1. All raw SuperDARN data were obtained from the British Antarctic Survey's SuperDARN mirror and then processed using the Radar Software Toolkit version 4.3 (SuperDARN Data Analysis Working Group et al., 2019). The specific processing commands and options used for the data processing can be found in the appendix of this paper. The rawacf-files were converted into fitacf-files using the FITACF function (version 2.5). Two gridded map files were created to see how changing the backscatter range limit affects the dataset. One version of the gridded files was created with an added backscatter range limit. By only including data from a minimum range of 800 km and a maximum far range of 2000 km, we eliminate all possible E-Region scatter and all backscatter with higher uncertainties in their location (Chisham et al., 2008; Forsythe & Makarevich, 2017; Thomas & Shepherd, 2018). The version of gridded files with a backscatter range limit is used for D1-D4 and the one without a range limit is used for D0. The gridded map files were resolved into two minute records and used the Chisham virtual height model (Chisham et al., 2008).

Dataset versions D0 and D1 include the same radars, whereas for D2 and D3, more radars were included (see table 1). For this selection of PolarDARN and StormDARN groupings the list provided by table 1 in Thomas and Shepherd (2018) was used. As can be seen from the list provided in Thomas and Shepherd (2018), most of the StormDARN radars were built after the high-latitude and PolarDARN radars.

For D4, we keep the selection of radars the same as D3, but use the background model from Thomas and Shepherd (2018) instead of the one from Ruohoniemi and Greenwald (1996).

To make all the final convection maps (D0 to D4), using RST, the Heppner-Maynard boundary (Heppner & Maynard, 1987; Shepherd & Ruohoniemi, 2000) was chosen as the lowest possible latitude measured by a minimum of three LOS vectors with velocities greater than 100 m/s (Imber et al., 2013). To complete the map fitting algorithm, the model requires solar wind data to be selected. For this, we use solar wind data from the ACE spacecraft, which has been time-lagged to the magnetosphere using the algorithm from Khan and Cowley (1999) which takes magnetosheath transit time into account. Finally, we add the model, and use a fitting order of 6 with a 'light' doping level for the background solar wind model. This uses the technique from Ruohoniemi and Baker (1998) to fit electrostatic potentials to the measured velocity vectors as spherical harmonic functions.

Choosing these versions of the dataset allows for a large-scale analysis of systematic changes and in particular, how the introduction of new mid-latitude and polar data modifies the dataset on a large scale, which has implications for use of the maps in scientific studies. Having established this archive of 2-minute resolution convection map files, we then extract a set of measured parameters with which quantify ionospheric convection, such as the HMB latitude and cross polar cap potential (CPCP). These describe the spatial extent and strength of the convection and allow us to examine how changes in the processing might affect conclusions of scientific studies, whereas the number of backscatter echoes per map or the average number of backscatter points per radar allows us to study how changes affect coverage. In this study, we define the HMB latitude as the fitted latitudinal boundary on the nightside and we also investigate how this parameter changes alongside the minimum latitude where backscatter is obtained (Λ_{min}), which can be along any magnetic local time or longitude. We would thus expect the difference between the two parameters to be positive for well constrained maps (i.e. Λ_{min} is at a lower latitude than the HMB), but this can also be negative when either the minimum latitude of observations is on the dayside (where the HMB shifts to higher latitudes) or an indicator that the HMB is not constrained by data. We also show how the different processing affects the χ^2/n-statistic, which is a global measure of map quality. The χ^2 parameter is a result from the singular value decomposition, which is minimised when the spherical harmonic fitting is performed to find the optimal solution for the coefficients. χ^2/n was introduced by Ruohoniemi and Baker (1998) as an indicator how well the measured line-of-sight velocities match the fitted velocities, where a value of 1 would indicate a good match and higher values would indicate a worse match.

Additionally, we also discuss the relationship between the HMB latitude and measures of geomagnetic activity, such as the Auroral Lower index (AL), the Auroral Elec-

Table 1. Differences between the comparison datasets

Version	Introduced difference	Background model	high-latitude radars	range limit	PolarDARN radars	StormDARN radars
D0	High-latitude radars[a] only	RG96	yes	no	no	no
D1	added range limit: 800-2000 km	RG96	yes	yes	no	no
D2	added PolarDARN radars[b]	RG96	yes	yes	yes	no
D3	added all other (i.e. StormDARN radars)[c]	RG96	yes	yes	yes	yes
D4	changed the background model	TS18	yes	yes	yes	yes

[a] High-latitude radars (i.e. all other radars): King Salmon, Kodiak, Prince George, Saskatoon, Kapuskasing, Goose Bay, Stokkseyri, Pykkvibaer, Hankasalmi.

[b]PolarDARN radars include: Inuvik, Rankin Inlet, Clyde River, Longyearbyen.

[c]StormDARN radars include: Hokkaido West, Hokkaido East, Adak West, Adak East, Christmas Valley West, Christmas Valley East, Fort Hays West, Fort Hays East, Blackstone, Wallops Island.

trojet index (AE) and the Symmetric Horizontal index (Sym-H) (Davis & Sugiura, 1966; Iyemori, 1990). We also consider the relationship between the CPCP and Φ_D, the dayside reconnection rate, which is calculated from the IMF B_Z, solar wind speed and IMF clock angle (Milan et al., 2012; Walach et al., 2017).

3 Results

The timeseries data extracted from the SuperDARN convection maps is condensed into probability distribution functions. By showing the data as 3-dimensional data distributions, we are able to compare the effects of changing the dataset on various parameters, which is shown in this section alongside examples of convection maps illustrating the changes.

3.1 Restricting radar backscatter range

Figure 1 shows probability distribution functions for a number of parameters for the entire D0 and D1 datasets. With D1 we have introduced the use of a range limit, as described in section 2.

Fig. 1a shows the distribution of HMB latitudes in D0 against D1. As most datapoints lie above the line of unity, we see that the HMB generally retreats poleward when we introduce a backscatter range limit. By limiting the backscatter ranges the number of backscatter echoes is reduced and thus also always increasing the lowest latitude at which backscatter is observed. We also see a saturation of points at a HMB latitude of 60°, which is where the boundary is drawn if not enough data is available (due to low data coverage or no slow scatter being observed). Fig. 1b shows the difference between the HMB latitude and Λ_{min}. We see that this difference is mostly positive for both D0 and D1, which means that the HMB sits below Λ_{min} and is thus well constrained. This latitudinal difference tends to shrink as we change the dataset from D0 to D1, as would be expected with a limited backscatter range. For a number of observations (40%), this latitudinal difference changes from positive to negative. This occurs for maps where the HMB is either not well constrained or the minimum latitude of observations is obtained on the dayside. Fig. 1c shows the χ^2/n distribution. It shows that χ^2/n tends to increase when the range limit is introduced. The range limit is expected to remove slow-moving E-region scatter ($<$ 800 km ranges) or scatter that may be placed in the wrong location ($>$ 2000 km ranges), which is expected to eliminate noise and uncertainty. Sometimes, χ^2/n measured at higher values in D0 (15-30) decreases for D1 (0-10), indicating that the map fitting improves. Fig. 1d shows the distribution of the number of backscatter echoes per map, n. It is worth noting that for the majority of D0 and D1, n is below 200, which as we will see in sections 3.2 to 3.6, is fairly low. Fig. 1e shows the average number of backscatter echoes per radar. As expected, changing the dataset from D0 to D1 not only decreases n overall, but also decreases the average number of backscatter echoes per radar. Fig. 1f shows the distribution of the CPCP. We see that when a range limit is introduced, the CPCP can either increase or decrease and there is no preference either way.

Panels Fig. 1g and h show two example convection maps for the same date and time (21^{st} December 2014 at 21:58 UT) from D0 and D1. In each case, the grid is geomagnetic latitude (which is in the AACGM-v2 coordinate system (Shepherd, 2014)) and magnetic local time, with noon towards the top, dusk towards the left, midnight towards the bottom and dawn towards the right. The coloured vectors show the gridded line-of-sight velocity vectors in locations where SuperDARN backscatter is available rather than the usual fitted vectors from Map Potential, which are usually shown in convection maps. The colours indicate the magnitudes of the vectors. The HMB is shown by the bright green line and the solid and dashed black lines show equipotentials in the electrostatic potential. To provide more context, this example map is indicated in the PDFs above by the light blue crosses. We see immediately that despite the high number of backscatter echoes and the low χ^2/n, there is a considerable difference in the potential patterns between D0 and D1. In D0 there are extra vectors in the dayside portion of the convection map, which provide fast flows, but also extra asymmetry that introduces an unphysical morphology. Adding in the range limit removes these and whilst it does not change the CPCP by much (4 kV), the convection maps themselves change considerably. Imposing the range limit removes fast vectors on the dayside and thus minimises the unphysical convection cells. This is an example where adding the range limit qualitatively improves the map and reduces the χ^2/n-statistic.

3.2 Adding PolarDARN

Figure 2 shows a comparison between D1 and D2 in the same format as in Fig. 1. In this comparison, we have introduced the Polar radars to the maps going from D1 to D2.

Fig. 2a shows the distribution of HMB latitudes. For 26.68%, the HMB moves to higher latitudes and for the majority of maps however (71.53%), the HMB does not change at all. The HMB moves to higher latitudes if it was not defined for D1 and adding more data can mean that the HMB is introduced at higher latitudes and the latitudinal difference between the HMB and the minimum latitude of observations thus becomes negative, shich is shown in Fig. 2b. This shows the difference between the HMB latitude and Λ_{min}. Fig. 2b shows that this distance tends to increase when we add the PolarDARN radars to the maps, which means the HMB is better constrained. The exception here are 1.79% of maps, where the minimum latitude HMB was already at high latitudes

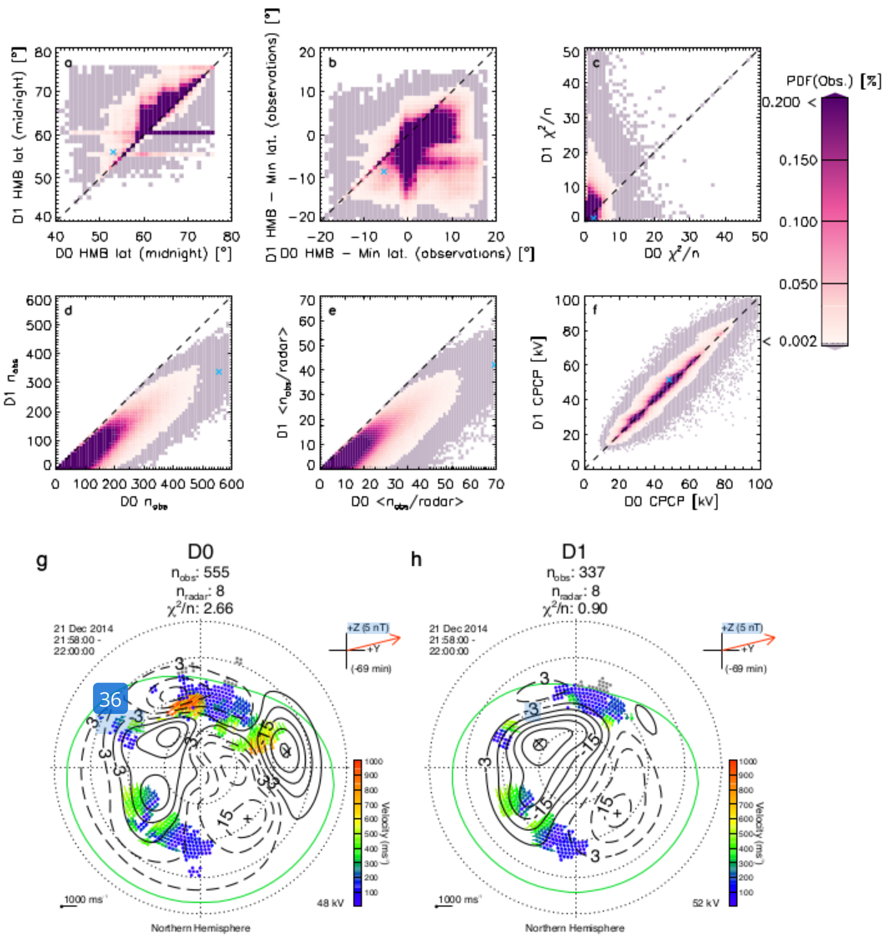

Figure 1. Probability distribution functions comparing the entire D0 and D1 datasets: (a) HMB latitude, (b) the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two panels show two example maps with the line-of-sight vectors from the same date and time (2014/12/21 21:58) for D0 (g) and D1, the convection map with the added range gate limit (h). These occurrences are indicated in the PDFs by blue crosses.

for D1 ($\geq 72°$), suggesting low coverage in the first instance, and thus introducing new data at high latitudes moves the minimum latitude of observation to slightly lower latitudes. Fig. 2c shows the χ^2/n distribution. We see that χ^2/n sometimes increases and sometimes decreases: This split is approximately equal with 45.40% of χ^2/n increasing and 49.76% of χ^2/n decreasing.

Fig. 2d shows the distribution of n. As we are introducing new data, the number of backscatter observations always increases, independently of how much data were available in D1.

Fig. 2e shows the average number of backscatter observations per radar. We see that this is likely to increase when the PolarDARN data is added. This means that the polar radars observed on average more backscatter points than the older radars in the network.

Fig. 2f shows the CPCP distribution. When adding the PolarDARN data to the network, it is possible for the CPCP to increase or decrease. We see that the spread of points above the line of unity is larger than below it. This means that if the CPCP increases, it is possible to increase by more than 30 kV, though the majority of data lies below the unity line and is likely to decrease by less than $\sim$30 kV.

As in Fig. 1, Fig. 2g and h show two example maps using D1(g) and D2 (h) for the same time (4$^{\text{th}}$ November 2014 at 20:08 UT), where the number of observations increases from 238 to 468. For this example the number of datapoints increases and this changes the pattern, despite the HMB still being constrained by the same datapoints. As high latitude datapoints are added however, the pattern is better constrained and a dawn cell appears due to fast flows being measured in the noon-morning region, leading to an increase in the CPCP from 27 kV to 54 kV.

3.3 Adding StormDARN

Figure 3 illustrates how the maps change when the mid-latitude (StormDARN) radars are added to the dataset. Fig. 3a shows the HMB distribution. This shows that the HMB is likely to stay at the same latitude or move closer to the equator. Fig. 3b shows the difference between the HMB latitude and Λ_{min}. As data from the mid-latitude radars are added, this latitudinal distance is likely to increase as would be expected. This dis-

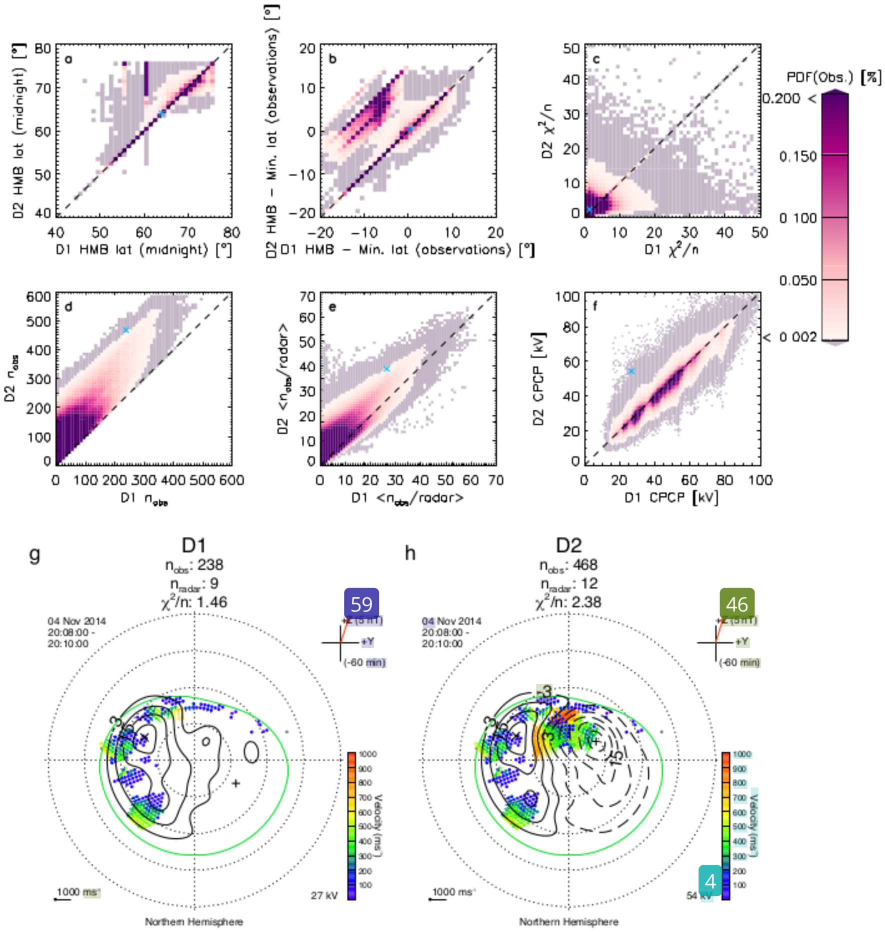

Figure 2. Probability distribution functions comparing the entire D1 and D2 datasets: (a) HMB latitude, (b) N the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two panels show two example maps with the line-of-sight vectors from the same date and time (2014/11/04 20:08) for D1 (g) and D2, the convection map with the added PolarDARN data (h). The example maps occurrences are indicated in the PDFs by blue crosses.

tance tends to be a positive one in the D3 dataset, meaning Λ_{min} tends to be closer to the equator than the HMB. This means the HMB is likely to be better constrained in D3 than D2. Fig. 3c shows the χ^2/n distribution. This value tends to decrease when we change the dataset from D2 to D3, which means that the background model fitting improves on average. Fig. 3d shows the number of backscatter observations, which increases as expected. Fig. 3e though shows that the number of gridded backscatter echoes per radar tends to decrease. This means that the average mid-latitude radar tends to observe fewer backscatter echoes than high-latitude radars. Fig. 3f shows the CPCP. We see from that the CPCP can increase or decrease, but the increases tend to be of a larger value than the decreases.

The bottom two rows in Fig. 3 show four example maps: The panels on the left (g and i) show example map of D2 from 9th November 2013 at 04:00 and the 8th February 2014 at 09:26, respectively. The two panels on the right (h and j) show the same date and time but using D3, where mid-latitude radars were included.

We see in panels g and h, that in this example adding these data increases the backscatter echoes by over 200 datapoints, even for this map, where the number of observations was already high previously. This moves the latitude of the HMB to lower latitudes from 62° to 52°. Furthermore, we see the convection cells change, in particular the dawn cell and the CPCP increases from 58 to 69 kV. All this will have a noticeable effect on any parameters extracted from the map. For example if we compute the convection velocity in D3 at the location where the HMB meets the midnight meridian for D2 (i.e. at 62° longitude and 00 MLT), the velocity would change from D2 to D3 from 0 m/s to 422 m/s.

Panels i and j of Fig. 3 however show an example of where adding mid-latitude data can make the convection maps look worse: Adding scatter at mid-latitudes almost doubles n, which increases from 326 to 613 here. Many of the measurements are however slow moving scatter, albeit not slow enough to fall below the HMB threshold, which results in the dawn convection cell almost disappearing. Initially this may seem like an extreme change in convection morphology, but the dawn cell only changes by ∼3 kV and the dusk cell is much better constrained by new mid-latitude vectors. The combination of these two changes causes an overall increase in the CPCP from 40 kV to 53 kV.

Computing the velocities for D3 at the HMB latitude location in D2 can be used as an indicator of how much the map has changed at specific locations and gives us an idea of how quantitatively different the convection maps might be without the mid-latitude radars. We explore this in more detail now.

Figure 4 shows the velocities, extracted from the D3 convection maps for the locations where the D2-HMB intersects with the noon, dusk, midnight and dawn meridians. We see that by adding the mid-latitude data, the maps change considerably at the locations where the HMB would have otherwise stipulated that there be zero flow. The curves show that at dawn, the effect is the least noticeable and that there is a 1 in 2 chance that the velocity measured in D3 has increased by 120 m/s or less, whereas this increases to 190 m/s for midnight and 220 m/s and 230 m/s for noon and dusk, respectively.

3.4 Changing the background model

In changing the dataset from D3 to D4, we are changing the background model from RG96 to TS18. This means that the observations which go into the convection maps stay constant, but the model fitting parameters (χ^2/n) change, as well as some of the resulting parameters, such as the CPCP.

Figure 5a shows the D3 versus D4 CPCP and we see that at the lower range (0-$\sim$50 kV), the CPCP is likely to decrease as we change the background model from RG96 to TS18 (this occurs 41.65% of the time as opposed to the increase which occurs 28.56% of the time). For the higher range (>50 kV) however, the CPCP is likely to increase when we change model from RG96 (D3) to TS18 (D4) (this occurs 16.46% of the time as opposed to the decrease which is 13.32%). Overall, TS18 thus provides a lower CPCP 54.97% of the time and a higher CPCP 45.02% of the time for the same data. Fig. 5b shows the CPCP difference against n. We see from this that the CPCP is in fact best constrained for maps with a high number of backscatter points, which means that there is a model dependency which decreases as n increases. For example, At n=200, the median and standard deviation are 0.87 kV and 8.88 kV, whereas at n=400, the median and standard deviation are 0.04 kV and 6.50 kV, respectively. Fig. 5c shows that χ^2/n, which can either decrease or increase when changing the background model. Although not immediately obvious, 63.81% of the data lie below the line of unity (in comparison to 36.15%

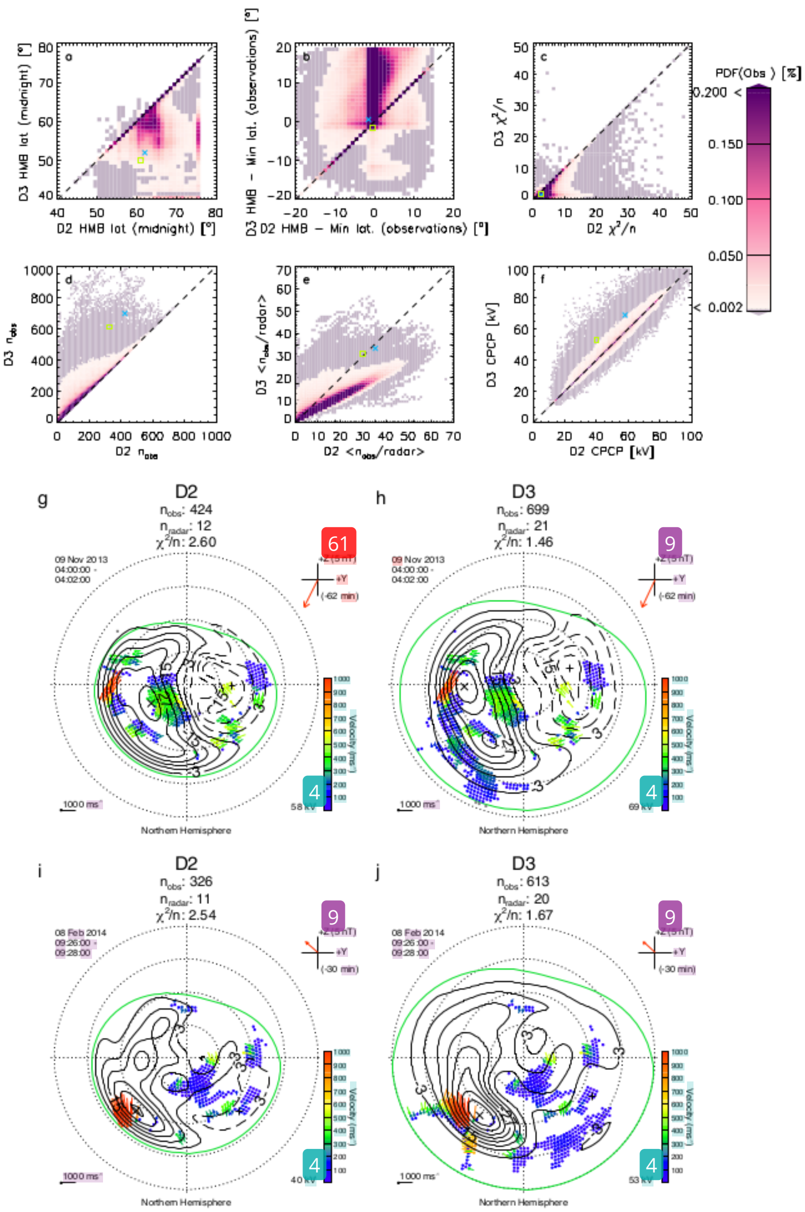

Figure 3. Probability distribution functions comparing the entire D2 and D3 datasets: (a) HMB latitude, (b) the difference between the HMB latitude and the minimum latitude where backscatter is observed, (c) χ^2/n distribution, (d) number of backscatter echoes, (e) average backscatter echoes per radar, (f) cross polar cap potential. The bottom two rows show four example maps with the line-of-sight vectors from two different dates and times (2013/11/09 04:00 and 2014/02/08 09:26) for D2 (left, g and i) and D3, the convection map which includes the mid-latitude radar data (right, h and j). These occurrences are indicated in the PDFs by blue crosses and green squares.

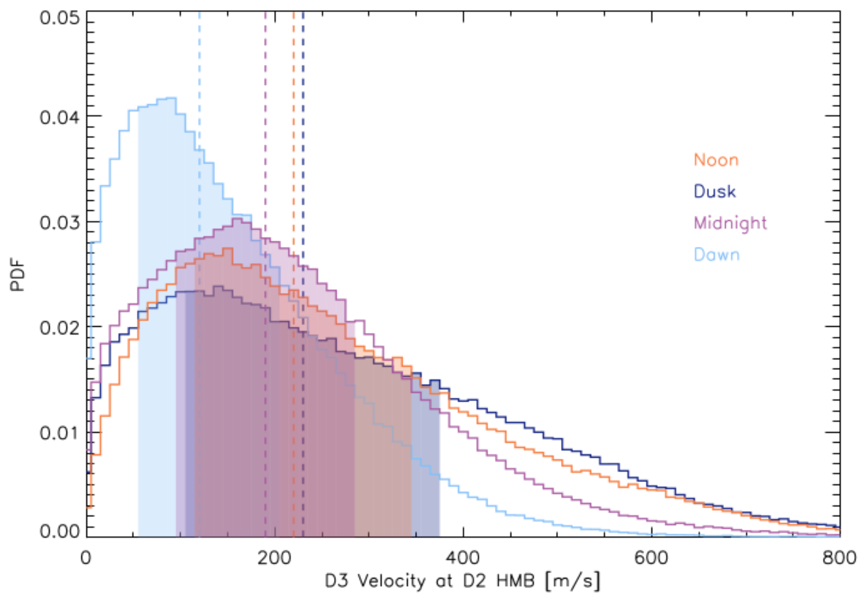

Figure 4. Probability distribution function of the velocity for D3, extracted at the noon, dusk, midnight and dawn locations where D2 would have had the HMB. Dashed lines show the medians for each distribution. Shaded regions indicate the boundaries of the lower and upper quartiles (25% and 75%).

of data above the line), meaning the fitting error is on average reduced when making the convection maps using TS18 in comparison to RG96.

As the input data does not change, the HMB values are largely the same for D3 and D4, except for times when the HMB cannot be defined. We have chosen not to show this plot, as these cases are extremely rare when we include the entire dataset (2.53% of cases). For D4, these cases will be defined by the background model and vary smoothly due to the interpolation in the background model between distinct bins, whereas for D3 (due to the parametrization in RG96), they will be defined as two distinct latitudes, as defined by the model: 60° (96.42% of instances) and 55° (3.57% of instances). Instead of showing the HMB latitude in D3 against D4, Fig. 5d thus shows the HMB latitude against n. It shows that the HMB is likely to move closer to the equator as the number of backscatter echoes increases.

Fig. 5e shows the HMB against AL. We see from this that the HMB is likely to move to lower latitudes as AL decreases, but this trend again breaks down at $\sim$50°. Similarly, in Fig. 5f we see a dependence in the HMB moving to lower latitudes as Sym-H becomes more negative, but this also breaks down at a HMB of 50 to 40°.

Panels d to f all show a seemingly linear trend with HMB, which seems to breaks down at low latitudes. As there are less occurrences for the extreme conditions, however this is difficult to establish.

Similar to previous figures, Fig. 5 shows two example maps in panels g and h, comparing D3 and D4. The map chosen as an example here is one of the best coverage maps, where n was the highest observed value with 1010. We see that having this much data coverage constrains the pattern very well and there are not many differences in the convection patterns: the CPCP only differs by 1 kV, the HMB is the same and the fitted convection potentials only differ very slighly in their morphology (e.g. noon-afernoon sector). This is to be expected, given the data distribution in Fig. 5b.

3.5 Changes to convection mapping since the first SuperDARN radar

Figure 6 provides a further comparison between the RG96 and TS18 datasets. Here we show comparisons between D0 and D4, providing a statistical viewpoint on how much

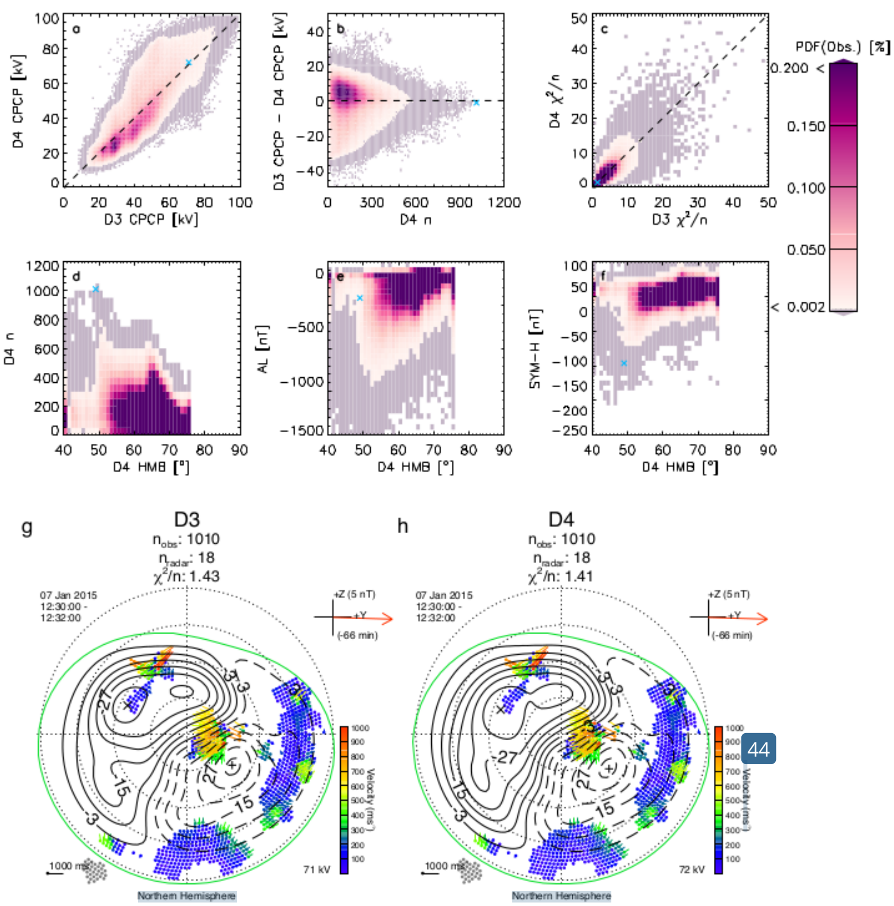

Figure 5. Probability distribution functions comparing the entire D3 and D4 datasets: (a) CPCP, (b) CPCP difference versus number of backscatter echoes, n, (c) χ^2/n distribution, (d) n versus HMB, (e) AL versus HMB, (f) Sym-H versus HMB. The bottom two panels show two example maps from the same date and time (2015/01/07 12:30) for D3 (g) and D4, the convection map which uses TS18 instead of RG96 (h). These occurrences are indicated in the PDFs by blue crosses.

has changed from the original SuperDARN convection map fitting to the most up-to-date version of datasets and fitting methods.

Fig. 6a shows the CPCP distribution. We see that the observed CPCP is on average smaller for D4 than D0 (54.28% of the time), but when the CPCP increases for D4, it increases by more on average (8.37 kV median; 10.45 mean; 92.08 kV maximum change) than it would otherwise decrease (6.87 kV median; 7.64 kV mean; 97.90 kV maximum change). Fig. 6b shows the χ^2/n, which can also increase (56.16%) or decrease (43.80%). By looking further at the statistical distribution, we find that for the times when χ^2/n is larger in D4 than D0, n for D4 tends to small (<200; 102 median; 123.13 mean). Fig. 6c shows the HMB distribution. Interestingly, this shows that the HMB is often higher (43.64% of the time) for D4 than for D0, despite the inclusion of mid-latitude data. This is mostly prominent when the HMB for D0 is above latitudes of 59° (39.76 % of the time), whereas the HMB is less likely to be at lower latitudes for D4 than D0 overall (18.00 % of the time). Fig. 6e shows the distribution of n, which carries a further surprise: n can increase, as well as decrease. We previously speculated that it would only increase, as the changes from D0 to D4 corresponds to the inclusion of polar and mid-latitude radars, but the distribution of n shows that it can also decrease due to the addition of the range limit, although this is less likely (31.63% of the time). The decrease in n scales consistently with n_{D4} and is on average a small change (-34.81 mean; -26.00 median and -349 maximum).

Fig. 6e shows the differences in the CPCP between D4 and D0 against the dayside reconnection rate, Φ_D. We see that the changes in the CPCP tend to be smaller for high solar wind driving (high Φ_D). Similarly, Fig. 6f shows the changes in the HMB against AE and Fig. 6g shows the changes in the HMB against AL. AE and AL, are the auroral electrojet indices, which are derived from ground-based magnetometer measurements and are a proxy for the magnetospheric activity in response to the dayside driving and internal dynamics (Davis & Sugiura, 1966; World Data Center for Geomagnetism in Kyoto et al., 2015). We see from panels f and g that changes in the HMB tend to be smaller when the auroral electrojet indices, AE and AL are enhanced.

Figs. 6h and i show the D4 and D0 HMB against AL. These include yellow and mint crosses that represent the median fits for each HMB bin, allowing us to compare D4 (yellow) with D0 (mint). This shows very clearly that when we use D0, we are less likely to observe a low HMB at enhanced (low) AL, which is not to mean that these occurrences do not exist, but simply that the SuperDARN fitting with the old dataset means we are less likely to observe them.

In Figs. 6j and k, we provide a similar comparison for the D4 and D0 CPCP with respect to Φ_D. This comparison shows that for D4 we are more likely to observe a higher CPCP at higher values of Φ_D than for D0. In fact, at a Φ_D of 100 kV, the median CPCP for D4 is at ~75 kV and ~ 65 kV for D0. We also see that the median curve has a different shape for the two datasets: Both have a logarithmic shape to them and neither appear like a linear fit would suffice to describe the trend in the dataset. Finally in panel l, we show the ratio between the CPCP normalised by Φ_D for both datasets, for which we have also fitted the median per bin (shown by yellow crosses). This shows that the differences between the two versions of the CPCP are proportional to the dayside driving. It also shows that this is a linear trend and that the CPCP changes in D0 with respect to Φ_D are likely to be smaller than for D4.

3.6 Identification of minimum map reliability

When using SuperDARN maps in research, a frequent question is "How reliable is this map?" and often n is used to answer this question. If n is high, the maps are often deemed more reliable, but is there a universal limit for n, which can be used to select reliable convection maps?

To answer this question, we present in Figure 7a the PDF of the difference in χ^2/n between D4 and D0 against the difference in n. It shows that as the map becomes more constrained (i.e. the difference in χ^2/n is negative), the difference in n becomes very small. Similarly, as the difference in χ^2/n becomes larger, the difference in n is also very small. This means that a change in n does not necessarily translate to a better constrained map. In fact, changes in n are more likely to happen for maps that are already well constrained. We see from Fig. 7a and Fig. 6b and d that maps where χ^2/n does not change much tend to have a low χ^2/n to begin with. Figure 7b and c show the difference in χ^2/n versus n in D4 and n in D0. From this we see clearly that the changes in χ^2/n are most extreme when n is small (<100), but there is no clear uniform break-point in n, where χ^2/n is small and the maps are well constrained. We also find that as n increases, χ^2/n is less likely to change. We see that this trend is the same for D4 and D0, however, there is less

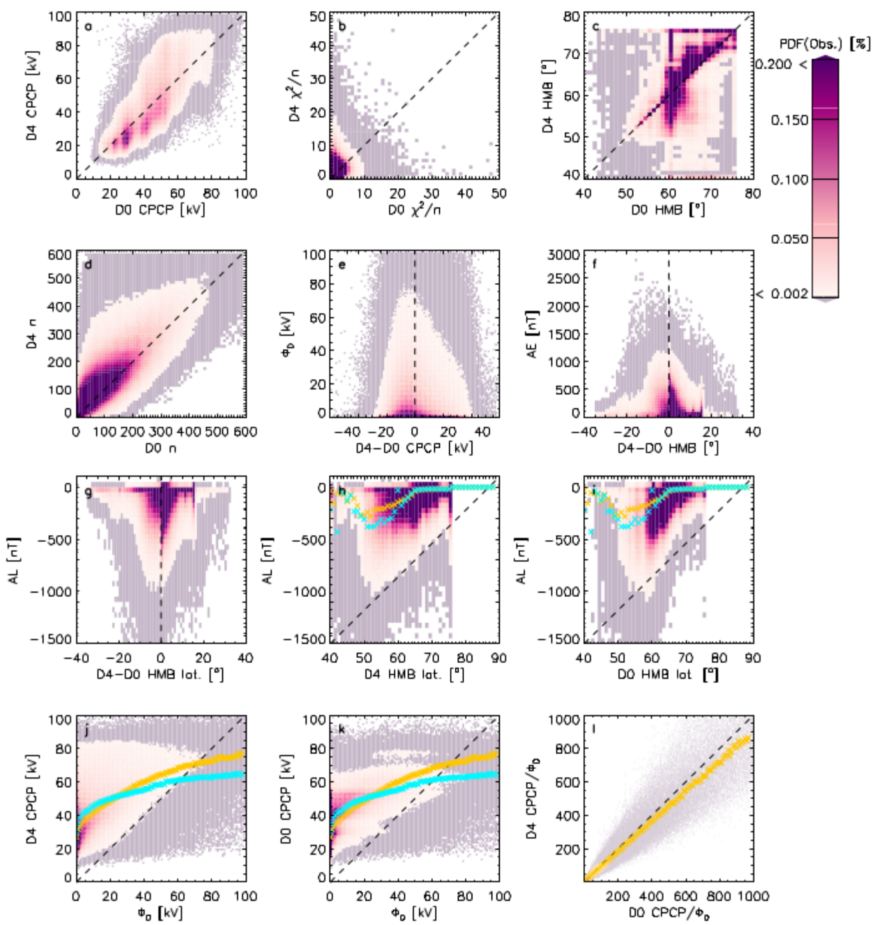

Figure 6. Probability distribution functions comparing the D0 and D4 datasets: (a) CPCP comparison, (b) χ^2/n comparison, (c) HMB comparison, (d) n comparison, (e) Φ_D versus the CPCP difference, (f) AE versus HMB difference, (g) AL versus HMB difference, (h)AL versus D4 HMB and (i) D0 HMB, (j) D4 CPCP versus Φ_D, (k) D0 CPCP versus Φ_D and (l) CPCP normalised by Φ_D. The crosses show the median in the y-direction for each x-bin (where applicable) with the yellow showing the fit for D4 and turquoise showing the fit for D0. Black dashed lines either show the lines of unity or the line at 0.

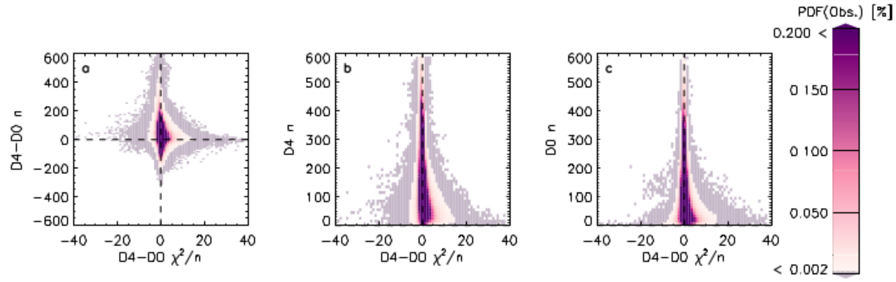

Figure 7. Probability distribution functions comparing the D0 and D4 datasets: The changes in χ^2/n versus (a) the changes in n, (b) D4 n and (c) D0 n. Black dashed lines show the line at 0.

spread and the peak is more pronounced for D0. We also note that the tail in the distributions of D4 n and D0 n versus the difference in χ^2/n are not symmetrical around 0. We will discuss these results further in the following section.

4 Discussion

4.1 How does changing the range limit affect the dataset?

Adding a range limit is intended to remove E-region scatter (i.e. slower moving scatter). This should increase convection in the maps and thus CPCP should increase. It also removes far-range scatter from slant range > 2000 km, which avoids potential errors in geolocation of LOS measurements at far range gates. Whilst this seems like should constrain the SHA solution, Thomas and Shepherd (2018) have shown that the opposite is true for a dataset that is limited in latitudinal coverage: Figure 11 in Thomas and Shepherd (2018) shows how the range limit impacts the data coverage afforded by the high-, polar-, and mid-latitude radars. For example, when data from beyond 2000 km slant range are removed from the high-latitude radar dataset, which is comparable to our D0 to D1 change, then the solution poleward of $\sim 76°$ magnetic latitude is purely constrained by the statistical model because no measurements are possible. This is to be expected and will be the same for our comparison. Reducing the range-limit will also reduce the number of backscatter echoes in the maps but we also see that the number of backscatter echoes are not solely responsible for map quality.

4 Chisham and Pinnock (2002) conclude that the contamination from non-F-region scatter does not usually have a large impact on the global characteristics of the SuperDARN convection maps. We find that for the analysed time period, the CPCP is > 10% different 4.86% of the time and the CPCP is < 10% different 95.13% of the time. Whilst less than 5% seems like a small set of observations, this does comprise more than 80000 maps, so it may be important on a case-study basis.

4 Chisham and Pinnock (2002) further showed that removing E-region scatter may not always result in more accurate convection maps. Whilst most E-region scatter is believed to move slower than F-region scatterm, this may not always be the case: Forsythe and Makarevich (2017) used SuperDARN data from the Southern hemisphere and showed that E-Region scatter can be of a similar order of magnitude as F-Region scatter (∼200 m/s or larger). They also showed however that whilst F-Region scatter tends to have a Gaussian velocity profile, the 21 E-Region velocity distribution is highly asymmetric, owing 21 to the Farley-Buneman and gradient drift instabilities being the main drivers. This may be the reason why Chisham and Pinnock (2002) find that removing E-region scatter does not always improve convection maps, but the study by Forsythe and Makarevich (2017) provides clear evidence why removing this scatter makes scientific sense. Our method of adding the range limit follows the strategy of Thomas and Shepherd (2018), though they used this method for statistical convection maps and this may not always be practical for instantaneous convection maps. Whilst the method employed here to removing far range backscatter is a broad-brush approach, future alternatives could include the use of either calibrated elevation angles (which involves measuring the elevation angles using interferometry) or a more accurate virtual height model.

We thus remove both potential E-region scatter and scatter from far range gates. We find that by introducing this range limit, the normalised Chi-squared distribution of the map fitting procedure, χ^2/n is increased 73.61% of the time and decreased 25.54% of the time.

Sometimes, reducing the number of backscatter points by introducing a range limit will increase the HMB to higher latitudes due to removing lower-latitude scatter but more poignantly, this change will reduce E-region scatter at lower-latitudes and thus reduce the probability of choosing a HMB at too low a latitude, as is shown in the example maps in Fig. 1.

For the subset of observations where this is most likely the case (i.e. the difference between the HMB and Λ_{min} are greater in D0 than in D1 and the HMB is at a lower latitude in D0 than in D1), the median n is higher (D0: 128 and D1: 56) than the median for the entire dataset (D0: 93 and D1: 40). Other portions of the dataset which may indicate a worse map contain the population where χ^2/n increases: here, the median n is less (D0: 86 and D1: 38) than the medians for the entire dataset (D0: 93 and D1: 40). Both these statistics suggest, that n is not a good predictor for how good the fit is once the the range limit has been introduced if χ^2/n is used as a quality-of-fit indicator. Alternatively, we suggest that this illustrates a downfall of χ^2/n and that it may not be the perfect indicator for quality. We propose that in the future, a better indicator for map quality is sought.

4.2 How does the addition of the PolarDARN radars affect the dataset?

Adding the polar radars to the dataset increases the coverage, so we would expect the CPCP to be better constrained and n to increase.

We find that adding the PolarDARN radars reduces the CPCP on average, which could indicate that the CPCP is overestimated without good polar cap coverage or that adding PolarDARN causes an underestimation. This has also been shown by Mori et al. (2012), who compared the velocity measurements from PolarDARN radars to CADI ionosonde measurements, as well as comparing the CPCP. Adding the range limit to our processing will remove any slow-moving E-Region scatter, which may increase the CPCP. It is thus more likely that the CPCP is overestimated without good polar cap coverage, as we have added the range limit to our procedure prior to adding PolarDARN radars, which is also shown by the example maps in Fig. 2, as opposed to the latter.

We also find that the difference between the HMB and Λ_{min} either stays the same or tends to increase when the polar radars are added to the dataset. Whilst we would expect PolarDARN measurements mostly to be poleward of the observations from the original high-latitude radars (particularly after introducing the range limit), this does not seem to be the case, which is most likely due to the limited local time observations in these maps. We also see that the HMB tends to stay the same or increase to a higher latitude when adding the polar radars. This indicates that for a number of maps, the HMB was not well defined as we would not expect the introduction of PolarDARN data to move the HMB at all. Whilst this indicates that the HMB was not always necessarily well constrained prior to the introduction of the PolarDARN data, it also indicates that observations near the pole are important in constraining the maps.

Adding the PolarDARN radars to the dataset can increase or decrease χ^2/n. This parameter only tends to increase for D2 if it was low for D1 and tends to decrease for D2 if it was high for D1. This suggests that the maps where the fitting was not particularly good for D1, improve when adding PolarDARN data, but there are also a number of maps where the fit becomes less good. Overall however, we find that the difference between the HMB and Λ_{min} has a tendency to increase, which means the HMB is constrained by data at a lower latitude. The median n increases from 40 to 108 when adding the PolarDARN radars, which is a considerable increase in scatter.

4.3 How does the addition of the StormDARN radars affect the dataset?

Adding StormDARN radars improves the coverage of data at lower latitudes, so we expect HMB to move and CPCP to change.

We find that the mid-latitude radars add less data to the maps (on average), than the polar or high latitude radars, but nevertheless, adding their data to the maps generally improves the dataset. χ^2/n almost always decreases and the HMB tends to be better constrained.

Thomas and Shepherd (2018) made a new baseline model and showed that introducing the mid-latitude radars could increase the CPCP by as much as 40% (for the most strongly southward IMF conditions) due to the high-latitude radars only being able to image a proportion of the convection zone necessary to constrain the CPCP. It is worth noting that Thomas and Shepherd (2018) found very little change in the CPCP for weak to moderate solar wind driving 5 because the low-latitude convection boundary remained within the FOV of the high-latitude radars. We find that, without using the TS18 model, but by simply including the mid-latitude radars, the CPCP does indeed increase more often (12.22% of times) than decrease (7.86% of times) but the maximum change seen is a 45% decrease when the CPCP changes from 34.70 kV in D2 to 19.19 kV in D3.

By investigating the D3 velocity measured at the HMB location of D2, we find that for 33.55% of cases the velocity change is less than 200 m/s, but for a considerable number of maps (7.90%, which equates to over 22000 maps), the velocity change is > 400 m/s at midnight, which indicates a considerable change to the convection pattern. This means that without the mid-latitude radars, the velocities at Λ_{HMB2} could be wrong by more than 190 m/s over half the time at midnight, which is considerable, assuming the HMB placing is constrained by data.

However, we have to consider the possibility that the HMB placing is not always correct: Fig. 3j shows large amounts of low velocity mid-latitude convection in the night-side ionosphere, which does not seem to improve the convection map. We postulate that these streams are associated with magnetic flux frozen into the plasmasphere (the inner part of the magnetosphere located just above the ionosphere) (Ribeiro et al., 2012). As the plasmasphere corotates with Earth, radars should not measure Doppler velocities associated with the rotation due to their fixed geographic location. However, if this co-rotation is not perfectly in sync with Earth's rotation then it may be possible to measure low Doppler velocities (tens-hundreds of ms^{-1}). While more transient in nature, over- or under-shielding scenarios may also lead to errors in the HMB latitude determination when including the mid-latitude radar data (e.g. Nishida, 1968; Nishitani et al., 2019): When this happens, the electric field formed at the inner edge of the plasma sheet and associated with the region 2 field-aligned currents counteracts the effects of the solar wind-driven magnetospheric convection at sub-auroral latitudes. Whilst these scenarios may lead to misidentification of the HMB, they are understood to be exceptional circumstances and not well enough understood to be explicitly taken into account when determining the HMB (Nishitani et al., 2019).

In either case, the HMB may need to be redefined. Currently, the HMB is calculated to be where velocity measurements suggest the electric field is zero, however low velocity measurements associated with imperfect co-rotation will also have an associated non-zero electric field. This suggests the HMB would not give the boundary of the convective regions associated with opening and closing of magnetic flux or that the boundary presents as a gradual change.

Walach and Grocott (2019) showed that during geomagnetic storms, which can also be described as extremely driven times, the HMB can move to latitudes as low as 40°, which SuperDARN radars prior to the mid-latitude expansion were not able to observe. Fogg et al. (2020) provide a fit for the HMB using AMPERE data, and show that the

HMB may be placed at too low latitudes when mid-latitude data are available. This might indicate that a changing HMB is not always an improvement when it moves equatorward in D3. It is however worth noting that the fitting by Fogg et al. (2020) does not include mid-latitude data and their fitting stops at 55°, so further analysis is necessary, which will be the subject of a future study.

Sub-auroral Polarization Streams (SAPS) are one of the main phenomenon studied with the mid-latitude radars (e.g. Kunduri et al., 2017, 2018). They consist of fast azimuthal streams, measured below auroral latitudes on the nightside (Kunduri et al., 2018). The possibility of the midlatitude radars observing either auroral flows in an expanded pattern, or sub-auroral flows in a smaller sized pattern, is an important distinction, which we have not studied in this paper but warrants further investigation. Kunduri et al. (2018) studied these flows in great detail and found that their occurrence and flow speed tends to increase with higher geomagnetic activity. To this date, SAPs have not been explicitly taken into account in the baseline SuperDARN models (e.g. RG96 and TS18) and it is thus likely that their effects are averaged over. We know that SAPs will occur at or near the lower latitudinal boundary of the convection patterns (e.g. Kunduri et al., 2018), but further investigation is necessary to understand how they fit in with the general convection pattern and in particular, how they affect HMB determination.

4.4 How does changing the background model affect the dataset?

When changing the background model from RG96 to TS18 we might expect a better fit due to a background model parametrization with more variables. Thomas and Shepherd (2018) not only use the IMF magnetic field strength and direction, their model parametrization also includes the solar wind's electric field and the Earth's dipole tilt, which results in 120 model bins that are trilinearly interpolated between to achieve smoother transitions, as opposed to the rigid 24 model bins chosen by Ruohoniemi and Greenwald (1996). The χ^2/n distribution indicates that sometimes this expected improvement is the case, however sometimes the fitting is worse, which is primarily the case for low n maps. Overall, we find (in Fig. 5) that the largest changes in the CPCP are produced when the CPCP was already high in D3 and these tend to occur when n is low. In fact, a higher n, means smaller likelihood of observing a change in CPCP. Thomas and Shepherd (2018) compared the changes in the baseline patterns and found that the CPCP can change by as much as 40%, when mid-latitude radars are included in the convection model, which is equivalent to a change of 32 kV for a CPCP of 80 kV without the mid-latitude radars. In comparison, we find that when using this model, the maximum observed percentage change in the CPCP is however a much larger change: a reduction of 63% for a CPCP of 48.84 kV in D3, which reduces to 17.91 kV in D4. The largest increases we see in CPCP when going from D3 to D4 is 59.38 kV, which happens for a CPCP of 59.38 kV in D3 and is a slightly larger change than the smallest decrease (57.11 kV), which happens for a CPCP of 33.41 kV in D3.

Fig. 5 shows that both AL and Sym-H show a linear trend in the likelihood of observations with HMB: As the HMB tends to lower latitudes, the values in AL and Sym-H tend to be enhanced until the HMB reaches a latitude of $\sim 50°$, at which point the observational likelihood reduces drastically overall. We also see that at HMBs $<50°$, n is likely to be smaller in general also, which means the observations in this HMB range are less dense and less well constrained. This is not surprising, as not all radars are capable of measuring HMBs $<50°$. Furthermore, the coverage from radars at mid-latitudes is sparser as the radars tend to, on average, return less backscatter per radar than the higher latitude radars.

In Fig. 6 we further explore how changing the background model, as well as introducing the newest radars to the dataset, affects the dataset. This shows that the HMB is more likely to be found at lower latitudes (50-40°) for D4 due to the lower observational latitude limit of the data. This means that the HMB is more likely to be observed at lower latitudes when the auroral electrojet indices (AL and AE) are enhanced. It is possible that the observational peak in AL and HMB, which shifts from ∼-400nT in D0 to ∼-300nT in D4 and ∼66° in D0 to ∼50° in D4, respectively, is still limited by radar coverage and it is possible that the decreasing trend we see in the median should continue (see crosses in Fig. 6).

The RG96 model was built only using the data from the Goose Bay radar, which is located at a high-latitude and thus part of our D0 set. Whilst it is one of the oldest operating radars in the network (and thus a lot of data is available), the RG96 model was constrained in magnetic latitudes from 65-85° (Ruohoniemi & Greenwald, 1996). It is thus interesting to see χ^2/n reduced, when adding the mid-latitude radars. This shows that the data is important in generating the convection map files, but from comparing D3 and D4 we see that the model can also make a difference. It is however worth noting that due to its limited data ingestion, the RG96 model was not built to be used with a radar network that extends to mid-latitudes, whereas TS18 was. Regardless of the χ^2/n statistic not always decreasing for the change from D3 to D4, the RG96 model does not account for as wide a variety of solar wind driving, dipolar tilt and latitudinal changes of the pattern and it thus makes more sense to use the TS18 model for the extended dataset, especially when including data from the midlatitudes.

4.5 The importance of backscatter echoes

Historically, n has on average increased due to the expansion of SuperDARN. Nevertheless, when we compare our most historic version of the dataset (D0) with the version that includes all new radars, as well as updated processing techniques (TS18 and range limit), we see that sometimes n decreases (Fig. 6d). This is thus solely due to the range limit introduction. Whilst adding the newer radars to the dataset can in some cases increase n by 500 or more, adding a range limit can reduce n by 100. We have shown that n is an important parameter in constraining the convection pattern (e.g. HMB or CPCP): In particular, we find that if n is high, the CPCP is less likely to change (i.e. the maps are constrained well) and the HMB is more likely to be found at lower latitudes (see Fig. 5).

When using SuperDARN maps, the reliability of the map is important and often this has been tied to n. If n is high, the maps are often deemed more reliable (e.g. Imber et al. (2013) identified 200 to be a low threshold number for good convection maps but Fogg et al. (2020) chose 400 as threshold for an acceptable number of backscatter echoes). This raises the question of whether there is a universal threshold for n, which can be used to select reliable convection maps?

We show that when n changes by large amounts (>200), the maps tend to be already well constrained (χ^2/n changes by $\sim$10), but we also find that when n is large in D0 and D4, χ^2/n is unlikely to change by much, which means the map is well constrained (see Fig. 7). The in-between state, where n changes, but not by large amounts, contains the maps that are the least well constrained (χ^2/n changes by up to 40). As n approaches $\sim$200, χ^2/n is likely to vary by <20 and as n approaches $\sim$400, the changes in χ^2/n are approximately halved. For higher values of n (>400), the probability of observing a change in χ^2/n remain the same. We see that this trend is the same for D4 and D0, however, there is less spread and the peak is more pronounced for D0. This means that whether or not a threshold of 200 or 400 is chosen for D0 makes minimal difference to how well the map is constrained. There is no clear break, where n universally produces good convection maps, but we show that if we choose $n > 400$, χ^2/n is unlikely to change by much and thus the map is as well constrained as it can be.

We also see from Fig. 7b-c that the spread of observations about 0 is not symmetrical. The left side of both distributions falls off much more abruptly than the right side, which implies that χ^2/n is larger in D4 than in D0 much more often and thus, for small n, the maps are less well-constrained for D4 than D0. This could be due to a number of reasons, but we suggest one main cause: D4 includes data over a larger spatial range but for a sixth order SHA, only 49 vectors are required to constrain it. As more vectors are added (e.g. from the midlatitude radars), more small-scale variability is added, which the 6th order SHA cannot resolve.

4.6 Geomagnetic conditions and SuperDARN observations

We have shown in Fig. 5d to f that when n is high, AL and Sym-H tend to enhance also and the HMB also tends to move to lower latitudes. It is worth considering the underlying physics and how these parameters are related as a result.

The expanding and contracting polar cap paradigm (e.g. Siscoe & Huang, 1985; Lockwood, 1991; Lockwood & Cowley, 1992; Milan, 2015; Walach et al., 2017, and references therein) requires the polar cap to increase in size when the dayside reconnection rate exceeds the nightside reconnection rate. This implies that the CPCP also increases when dayside driving is high. We have shown that this is mostly the case, although there are some deviations to this relationship, which we attribute to noise and errors in solar wind propagation. It has long been discussed whether or not the relationship between the dayside driving and the CPCP is linear and whether or not the CPCP saturates beyond a threshold (e.g. Hill et al., 1976; Reiff et al., 1981; Doyle & Burke, 1983; Wygant et al., 1983; Shepherd, 2007; Mori & Koustov, 2013, and references therein). Shepherd et al. (2002) and Shepherd (2007) discuss this in great detail and showed, using SuperDARN CPCP measurements, that during high solar wind driving (when the reconnection electric field is above 5.5 mV/m), the CPCP saturates.

Mori and Koustov (2013) talk about a SuperDARN "quantization" effect, whereby for high CPCP where the observational density is low and not all maps are well constrained, the CPCP oftentimes takes on the values of the underlying model (e.g. RG96). We see this quantization to some extent in Fig. 6 for RG96, but this problem is solved for TS18, which interpolates between solutions of the background model. Whilst this is not the focal point of our study, we find that as Φ_D increases, the CPCP also increases. Similar to Shepherd (2007), we note that observational density is an important factor when considering the behaviour of these parameters. We also find that depending on the dataset used (e.g. D0 or D4), the trend and steepness of the curve varies due to observational density of high CPCP for D0 being much lower than for D4. Furthermore, we find that the spread in values is much higher than observed by Shepherd (2007), which is due to a larger sample size (they only used equinox data for their study) and shorter sampling (they used 10 minute cadence for their map files whereas we use 2 minutes). We suggest that using the verb "saturate" to describe the behaviour of these parameters is misplaced, as even at high values of Φ_D the CPCP increases, whereas a saturation implies the gradient of the curve reaching 0.

Whilst n is high when AL, Sym-H and the HMB are enhanced, we are not suggesting that the correlation equates to a causal link. This was already discussed by Walach and Grocott (2019), who showed that the number of backscatter echoes tends to increase during geomagnetic storms (when Sym-H is enhanced), as dayside driving increases, the polar cap grows and the HMB moves to lower latitudes. Currie et al. (2016) showed however that during intense geomagnetic storms, a reduction of backscatter was observed in the Bruny Island radar in the middle- to far-ranges, and an increase in the amount of backscatter from close-ranges. Here we show statistically, that as Sym-H is enhanced, the HMB moves to lower latitudes and the number of backscatter echoes increases for mid-ranges (the far- and close- ranges were removed beyond D0 by the range limit). We thus find that the relationships found by Walach and Grocott (2019) hold statistically, though a large amount of variation is observed.

Wild and Grocott (2008) conducted a study (before the availability of mid-latitude radars) of regions where backscatter is lost during isolated substorms, and the progression through the phases of the substorm due to auroral absorption. They identify that backscatter reduction is greatest at ∼70-80° magnetic latitude region between ∼19 to 03 MLT. However, Wild and Grocott (2008) also observe that the main backscatter region shifts equatorward to lower latitudes (below ∼65°) across all local times. Our results support this statistically, as we find that the mid-latitude radars do on average observe more backscatter, and that the backscatter moves to lower latitudes when AL is enhanced (which is expected to be the case for substorms). We also find that this trend differs slightly for D0 and D4: due to better coverage with the mid-latitude radars, the HMB for D4 moves to lower latitudes than for D0. The trend of decreasing HMB with decreasing AL is a statistical one and thus breaks at a latitudes close to ∼40° due to low observational densities

5 Summary

We have investigated how the SuperDARN maps have changed historically by creating 5 different versions of the convection map files for a timespan of 6 years and comparing them statistically. By using different processing parameters and gradually introducing more data to the maps, we were able to investigate how the dataset changes with the inclusion of

- a backscatter range limit (as was used by Thomas and Shepherd (2018))
- the polar cap radars, PolarDARN
- the mid-latitude radars, StormDARN
- a different statistical background model (we compare Thomas and Shepherd (2018) and Ruohoniemi and Greenwald (1996))

We have shown that

- introducing a range limit does not always decrease χ^2/n,
- n is not a good predictor for how good the fit is once the range limit has been applied
- once the range limit has been applied the CPCP stays the same 29.71% of the time and the HMB stays constant most of the time (54.47%)
- the addition of PolarDARN data tends to reduce the CPCP,
- PolarDARN radars add the most data to the dataset (on average), but the mid-latitude radars are also important for constraining the maps,
- when introducing StormDARN radars to the maps, the χ^2/n values tend to decrease, the HMB becomes better constrained and the CPCP tends to increase

- when changing the background model to TS18, the CPCP tends to decrease for lower values of the CPCP in RG96, but is more likely to increase for larger values of the CPCP in RG96. If n is however high (> 400), the CPCP is less likely to change (changes $\sim<$20 kV).
- as n, AL and Sym-H all increase, the HMB tends to go to lower latitudes, which appears to be a linear trend, though a break is seen at HMB $\sim$50 degrees, where the observational density drops off sharply.
- if n is high, the CPCP is less likely to change and the HMB is more likely to be found at lower latitudes and χ^2/n tends to change by the least amount,
- there is no clear break, where n universally produces good convection maps, but we show that for $n > 400$, χ^2/n is unlikely to change by much and thus the map is as well constrained as it can be.

Naturally, assessing map quality has to include a qualitative discussion and there is currently no perfect quantitative method for this assessment. The current most simple way to assess map quality is to look at the χ^2/n statistic. If we sum χ^2 and divide by the sum of n for each dataset D0 to D4, we obtain the following average values: $<\chi^2/n>_{D0}$: 1.70; $<\chi^2/n>_{D1}$: 2.01; $<\chi^2/n>_{D2}$: 2.16; $<\chi^2/n>_{D3}$: 1.88; and $<\chi^2/n>_{D4}$: 1.81.

From this, we might conclude that D0 has overall the highest quality maps and is closest to the "good match" criterion (1) identified by Ruohoniemi and Baker (1998), but we have shown that whilst the map fitting may be better for D0, the missing data also equates to a qualitative penalty. We see from these values that most of the impact on χ^2/n are provided by the range limit and the addition of the mid-latitude radar data. This emphasizes the importance of good spatial coverage. We also see from these statistics, that overall, the TS18 model improves map fitting.

Overall, we have shown that the measured parameters (such as the CPCP and HMB) are highly susceptible to which processing parameters are used, as well as which radars are used when generating map files. This becomes particularly important when SuperDARN maps are used for studies of specific conditions or small case studies as a sampling bias can occur. A high number of SuperDARN backscatter echoes are particularly important when constraining maps, so it is important to include mid-latitude data in the generation of SuperDARN convection maps. We have also shown that the method of selecting the HMB is not always perfect and further work is necessary to generate a robust selection method, especially at lower latitudes.

Appendix A SuperDARN processing parameters

In the SuperDARN processing (see section 2), we use the following parameters and functions from RST:

- For fitting the autocorrelation function to the raw data: 'make_fit' with the option '-fitacf-version 2.5'.
- To make the gridded map files, the options '-i 120 -tl 120 -chisham -c' were added to 'make_grid'
- To add the range limit to the gridded files, the same options as above were used but in addition, the options '-minsrng 800 -maxsrng 2000' were added.
- The function 'map_grd' was used with 'map_addhmb -vel 100 -cnt 3'. Adding these options to 'map_addhmb' chooses the Heppner-Maynard boundary to the lowest possible latitude for which a minimum of three LOS vectors with velocities greater than 100 m/s lie along its boundary.
- To make the convection maps, we also use 'map_addimf -if' with the text file containing the IMF data and the option '-df' with the text file containing the IMF delay times.
- We then use 'map_addmodel -o 6 ' for a sixth order expansion and use '-d' to specify a light doping level.
- Finally, we add the model option '-rg96' to D0-D3 and '-ts18' to D4 and use the function 'map_fit' to make the convection map files.
- We also use the function 'cnvmaptomap' to convert the binary file to ASCII format and 'trim_map' with the options '-st', '-et', '-sd' and '-ed' to make two-hour long map files for our archive, but this is not necessary to obtain the results for this study.

Acknowledgments

All data used for this study are available opensource from nonprofit organizations. The authors acknowledge the use of SuperDARN data. SuperDARN is a collection of radars funded by national scientific funding agencies of Australia, Canada, China, France, Japan, South Africa, United Kingdom, and United States of America, and we thank the international PI team for providing the data. The authors acknowledge access to the SuperDARN database via the British Antarctic Survey (https://www.bas.ac.uk/project/superdarn/#data). Other data mirrors are hosted by the Virginia Tech SuperDARN group (http://vt.superdarn.org/) and the University of Saskatchewan (https://superdarn.ca/data-download). The Radar Software Toolkit (RST) to process the SuperDARN data can be downloaded from Zenodo (https://doi.org/10.5281/zenodo.1403226 and references). All solar wind data and geomagnetic indices were downloaded from NASA's SPDF Coordinated Data Analysis Web (https://cdaweb.gsfc.nasa.gov/index.html/). The AE data is also available from the WDC for Geomagnetism, Kyoto (http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html) who prepared this index. M.-T. W. and A. G. were supported by Natural Environments Research Council (NERC), UK, grant nos. NE/P001556/1 and NE/T000937/1. F. S. was supported by a Science and Technology Funding Council (STFC) studentship. E. G. T. thanks the National Science Foundation (NSF) for support under grants AGS-1934997 and OPP-1836426. We gratefully acknowledge the use of Lancaster University's High End Computing Cluster, which has facilitated the necessary dataprocessing for this study. M.-T. W. would like to thank LU's Women's Network Writing Group for providing a supportive virtual writing space and mentorship, which helped to forge this paper.

References

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., . . . Walker, D. M. (2007). A decade of the Super Dual Auroral Radar Network (SuperDARN): Scientific achievements, new techniques and future directions. *Surveys in Geophysics*, *28*(1), 33–109. doi: 10.1007/s10712-007-9017-8

Chisham, G., & Pinnock, M. (2002). Assessing the contamination of SuperDARN global convection maps by non-F-region backscatter. *Annales Geophysicae*, *20*(1), 13-28. Retrieved from `https://hal.archives-ouvertes.fr/hal-00316917`

Chisham, G., Yeoman, T. K., & Sofko, G. J. (2008). Mapping ionospheric backscatter measured by the superdarn hf radars – part 1: A new empirical virtual height model. *Annales Geophysicae*, *26*(4), 823–841. doi: 10.5194/angeo-26-823-2008

Cousins, E. D. P., & Shepherd, S. G. (2010). A dynamical model of high-latitude convection derived from superdarn plasma drift measurements. *Journal of Geophysical Research: Space Physics*, *115*(A12). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JA016017 doi: 10.1029/2010JA016017

Currie, J. L., Waters, C. L., Menk, F. W., Sciffer, M. D., & Bristow, W. A. (2016). Superdarn backscatter during intense geomagnetic storms. *Radio Science*, *51*(6), 814-825. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016RS005960 doi: 10.1002/2016RS005960

Davis, T. N., & Sugiura, M. (1966). Auroral electrojet activity index ae and its universal time variations. *Journal of Geophysical Research (1896-1977)*, *71*(3), 785-801. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ071i003p00785 doi: https://doi.org/10.1029/JZ071i003p00785

Doyle, M. A., & Burke, W. J. (1983). S3-2 measurements of the polar cap potential. *Journal of Geophysical Research*, *88*(A11), 9125. doi: 10.1029/JA088iA11p09125

Fogg, A. R., Lester, M., Yeoman, T. K., Burrell, A. G., Imber, S. M., Milan, S. E., ... Anderson, B. J. (2020). An improved estimation of superdarn heppner-maynard boundaries using ampere data. *Journal of Geophysical Research: Space Physics*, *125*(5), e2019JA027218. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JA027218 (e2019JA027218 10.1029/2019JA027218) doi: https://doi.org/10.1029/2019JA027218

Forsythe, V. V., & Makarevich, R. A. (2017). Global view of the e region irregularity and convection velocities in the high-latitude southern hemisphere. *Journal of Geophysical Research: Space Physics*, *122*(2), 2467-2483. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JA023711 doi: https://doi.org/10.1002/2016JA023711

Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., ... Yamagishi, H. (1995). Darn/superdarn. *Space Science Reviews*, *71*(1), 761–796. doi: 10.1007/BF00751350

Heppner, J. P., & Maynard, N. C. (1987). Empirical high-latitude electric field models. *Journal of Geophysical Research*, *92*(A5), 4467–4489. doi: 10.1029/

JA092iA05p04467

Hill, T. W., Dessler, A. J., & Wolf, R. A. (1976). Mercury and mars: The role of ionospheric conductivity in the acceleration of magnetospheric particles. *Geophysical Research Letters*, *3*(8), 429-432. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL003i008p00429` doi: https://doi.org/10.1029/GL003i008p00429

Imber, S. M., Milan, S. E., & Lester, M. (2013). The heppner-maynard boundary measured by superdarn as a proxy for the latitude of the auroral oval. *Journal of Geophysical Research: Space Physics*, *118*(2), 685-697. doi: 10.1029/2012JA018222

Iyemori, T. (1990). Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations. *Journal of geomagnetism and geoelectricity*, *42*(11), 1249-1265. doi: 10.5636/jgg.42.1249

Khan, H., & Cowley, S. W. H. (1999, Sep 01). Observations of the response time of high-latitude ionospheric convection to variations in the interplanetary magnetic field using eiscat and imp-8 data. *Annales Geophysicae*, *17*(10), 1306–1335. doi: 10.1007/s00585-999-1306-8

Kunduri, B. S. R., Baker, J. B. H., Ruohoniemi, J. M., Nishitani, N., Oksavik, K., Erickson, P. J., ... Miller, E. S. (2018). A new empirical model of the subauroral polarization stream. *Journal of Geophysical Research: Space Physics*, *123*(9), 7342-7357. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JA025690` doi: https://doi.org/10.1029/2018JA025690

Kunduri, B. S. R., Baker, J. B. H., Ruohoniemi, J. M., Thomas, E. G., Shepherd, S. G., & Sterne, K. T. (2017, June). Statistical characterization of the large-scale structure of the subauroral polarization stream. *J. Geophys. Res-Space Phys.*, *122*(6), 6035–6048. doi: 10.1002/2017JA024131

Lockwood, M. (1991). On flow reversal boundaries and transpolar voltage in average models of high-latitude convection. *Planetary and Space Science*, *39*(3), 397–409. doi: 10.1016/0032-0633(91)90002-r

Lockwood, M., & Cowley, S. W. H. (1992). Ionospheric convection and the substorm cycle. In *International conference on substorms.*

Milan, S. E. (2015). Magnetospheric Plasma Physics: The Impact of Jim Dungey's

Research. In D. Southwood, S. W. H. Cowley FRS, & S. Mitton (Eds.), *Magnetospheric plasma physics: The impact of jim dungey's research* (pp. 1–271). doi: 10.1007/978-3-319-18359-6_2

Milan, S. E., Gosling, J. S., & Hubert, B. (2012, mar). Relationship between interplanetary parameters and the magnetopause reconnection rate quantified from observations of the expanding polar cap. *Journal of Geophysical Research*, *117*(A3), A03226. doi: 10.1029/2011JA017082

Mori, D., & Koustov, A. (2013). Superdarn cross polar cap potential dependence on the solar wind conditions and comparisons with models. *Advances in Space Research*, *52*(6), 1155-1167. Retrieved from https://www.sciencedirect.com/science/article/pii/S0273117713003803 doi: https://doi.org/10.1016/j.asr.2013.06.019

Mori, D., Koustov, A. V., Jayachandran, P. T., & Nishitani, N. (2012). Resolute bay cadi ionosonde drifts, polardarn hf velocities, and cross polar cap potential. *Radio Science*, *47*(3). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011RS004947 doi: https://doi.org/10.1029/2011RS004947

Nishida, A. (1968). Coherence of geomagnetic dp 2 fluctuations with interplanetary magnetic variations. *Journal of Geophysical Research (1896-1977)*, *73*(17), 5549-5559. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA073i017p05549 doi: https://doi.org/10.1029/JA073i017p05549

Nishitani, N., Ruohoniemi, J. M., Lester, M., Baker, J. B. H., Koustov, A. V., Shepherd, S. G., ... Kikuchi, T. (2019). Review of the accomplishments of mid-latitude super dual auroral radar network (superdarn) hf radars. *Progress in Earth and Planetary Science*, *6*(1), 27. doi: 10.1186/s40645-019-0270-5

Pettigrew, E. D., Shepherd, S. G., & Ruohoniemi, J. M. (2010). Climatological patterns of high-latitude convection in the northern and southern hemispheres: Dipole tilt dependencies and interhemispheric comparisons. *Journal of Geophysical Research: Space Physics*, *115*(A7). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JA014956 doi: 10.1029/2009JA014956

Press, W. H. and Teukolsky, S. A. and Vetterling W. T. and Flannery B. P. (2007).

Numerical recipes: The art of scientific computing. Cambridge University Press.

Reiff, P. H., Spiro, R. W., & Hill, T. W. (1981). Dependence of polar cap potential drop on interplanetary parameters. *Journal of Geophysical Research*, *86*(A9), 7639–7648. doi: 10.1029/JA086iA09p07639

Ribeiro, A. J., Ruohoniemi, J. M., Baker, J. B. H., Clausen, L. B. N., Greenwald, R. A., & Lester, M. (2012). A survey of plasma irregularities as seen by the midlatitude blackstone superdarn radar. *Journal of Geophysical Research: Space Physics*, *117*(A2). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JA017207` doi: https://doi.org/10.1029/2011JA017207

Ruohoniemi, J. M., & Baker, K. B. (1998). Large-scale imaging of high-latitude convection with Super Dual Auroral Radar Network HF radar observations. *Journal of Geophysical Research*, *103*(A9), 20797. doi: 10.1029/98JA01288

Ruohoniemi, J. M., & Greenwald, R. A. (1996). Statistical patterns of high-latitude convection obtained from Goose Bay HF radar observations. *Journal of Geophysical Research*, *101*(A10), 21743. Retrieved from `http://doi.wiley.com/10.1029/96JA01584` doi: 10.1029/96JA01584

Ruohoniemi, J. M., & Greenwald, R. A. (2005). Dependencies of high-latitude plasma convection: Consideration of interplanetary magnetic field, seasonal, and universal time factors in statistical patterns. *Journal of Geophysical Research: Space Physics*, *110*(A09204). doi: 10.1029/2004JA010815

Shepherd, S. G. (2007). Polar cap potential saturation: Observations, theory, and modeling. *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*(3), 234-248. Retrieved from `https://www.sciencedirect.com/science/article/pii/S136468260600263X` (Global Aspects of Magnetosphere-Ionosphere Coupling) doi: https://doi.org/10.1016/j.jastp.2006.07.022

Shepherd, S. G. (2014). Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations. *Journal of Geophysical Research: Space Physics*, *119*(9), 7501-7521. doi: 10.1002/2014JA020264

Shepherd, S. G., Greenwald, R. A., & Ruohoniemi, J. M. (2002). Cross polar cap potentials measured with super dual auroral radar network during quasi-steady solar wind and interplanetary magnetic field conditions. *Journal of Geo-*

physical Research: Space Physics, *107*(A7), SMP 5-1-SMP 5-11. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JA000152 doi: https://doi.org/10.1029/2001JA000152

Shepherd, S. G., & Ruohoniemi, J. M. (2000). Electrostatic potential patterns in the high-latitude ionosphere constrained by superdarn measurements. *Journal of Geophysical Research: Space Physics*, *105*(A10), 23005-23014. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JA000171 doi: 10.1029/2000JA000171

Siscoe, G. L., & Huang, T. S. (1985). Polar cap inflation and deflation. *Journal of Geophysical Research*, *90*(A1), 543–547.

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Ponomarenko, P. V., Billett, D. D., Bland, E. C., Burrell, A. G., … Walach, M.-T. (2018, August). *Superdarn radar software toolkit (rst) 4.2.* Retrieved from https://doi.org/10.5281/zenodo.1403226 doi: 10.5281/zenodo.1403226

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Ponomarenko, P. V., Bland, E. C., Burrell, A. G., Kotyk, K., … Walach, M.-T. (2018, January). *Superdarn radar software toolkit (rst) 4.1.* Retrieved from https://doi.org/10.5281/zenodo.1143675 doi: 10.5281/zenodo.1143675

SuperDARN Data Analysis Working Group, P. m., Thomas, E. G., Sterne, K. T., Shepherd, S. G., Kotyk, K., Schmidt, M. T., … Billett, D. D. (2019, September). *Superdarn radar software toolkit (rst) 4.3.* Zenodo. Retrieved from https://doi.org/10.5281/zenodo.3401622 doi: 10.5281/zenodo.3401622

Thomas, E. G., & Shepherd, S. G. (2018, apr). Statistical Patterns of Ionospheric Convection Derived From Mid-latitude, High-Latitude, and Polar SuperDARN HF Radar Observations. *Journal of Geophysical Research: Space Physics*, *123*(4), 3196–3216. Retrieved from http://doi.wiley.com/10.1002/2018JA025280 doi: 10.1002/2018JA025280

Walach, M.-T., & Grocott, A. (2019). Superdarn observations during geomagnetic storms, geomagnetically active times, and enhanced solar wind driving. *Journal of Geophysical Research: Space Physics*, *124*(7), 5828-5847. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JA026816 doi: 10.1029/2019JA026816

Walach, M.-T., Milan, S. E., Yeoman, T. K., Hubert, B. A., & Hairston, M. R.

(2017). Testing nowcasts of the ionospheric convection from the expanding and contracting polar cap model. *Space Weather*, *15*(4), 623-636. doi: 10.1002/2017SW001615

Wild, J. A., & Grocott, A. (2008). The influence of magnetospheric substorms on superdarn radar backscatter. *Journal of Geophysical Research: Space Physics*, *113*(A4). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JA012910 doi: 10.1029/2007JA012910

World Data Center for Geomagnetism in Kyoto, Nose, M., Iyemori, T., Sugiura, M., & Kamei, T. (2015). *Geomagnetic ae index.* doi: 10.17593/15031-54800

Wygant, J. R., Torbert, R. B., & Mozer, F. S. (1983). Comparison of s3-3 polar cap potential drops with the interplanetary magnetic field and models of magnetopause reconnection. *Journal of Geophysical Research: Space Physics*, *88*(A7), 5727-5735. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA088iA07p05727 doi: https://doi.org/10.1029/JA088iA07p05727

Super Dual Auroral Radar Network Expansion and its Influence on the Derived Ionospheric Convection Pattern

ORIGINALITY REPORT

12%
SIMILARITY INDEX

PRIMARY SOURCES

1	M. - T. Walach, A. Grocott, S. E. Milan. "Average Ionospheric Electric Field Morphologies during Geomagnetic Storm Phases", Journal of Geophysical Research: Space Physics, 2021 Crossref	378 words — 3%
2	eprints.lancs.ac.uk Internet	204 words — 1%
3	agu.confex.com Internet	138 words — 1%
4	hal.archives-ouvertes.fr Internet	88 words — 1%
5	agupubs.onlinelibrary.wiley.com Internet	86 words — 1%
6	M. - T. Walach, A. Grocott. "SuperDARN Observations During Geomagnetic Storms, Geomagnetically Active Times, and Enhanced Solar Wind Driving", Journal of Geophysical Research: Space Physics, 2019 Crossref	63 words — < 1%
7	progearthplanetsci.springeropen.com Internet	58 words — < 1%

8	ecommons.usask.ca Internet	34 words — < 1%
9	www.ann-geophys.net Internet	32 words — < 1%
10	A. B. Workayehu, H. Vanhamäki, A. T. Aikio. "Seasonal effect on hemispheric asymmetry in ionospheric horizontal and field‐aligned currents", Journal of Geophysical Research: Space Physics, 2020 Crossref	31 words — < 1%
11	P.A. Inchin, J.B. Snively, M.D. Zettergren, A. Komjathy, O.P. Verkhoglyadova, S. Tulasi Ram. "Modeling of ionospheric responses to atmospheric acoustic and gravity waves driven by the 2015 Nepal Mw7.8 Gorkha earthquake", Journal of Geophysical Research: Space Physics, 2020 Crossref	25 words — < 1%
12	E. G. Thomas, S. G. Shepherd. "Statistical Patterns of Ionospheric Convection Derived From Mid-latitude, High-Latitude, and Polar SuperDARN HF Radar Observations", Journal of Geophysical Research: Space Physics, 2018 Crossref	24 words — < 1%
13	onlinelibrary.wiley.com Internet	24 words — < 1%
14	Erxiao Liu, Hongqiao Hu, Jianjun Liu, Lei Qiao. "Deep Learning Models for Estimation of the SuperDARN Cross Polar Cap Potential", Earth and Space Science, 2020 Crossref	22 words — < 1%

15	M. ‐ T. Walach, A. Grocott, S. E. Milan. "Average Ionospheric Electric Field Morphologies During Geomagnetic Storm Phases", Journal of Geophysical Research: Space Physics, 2021 Crossref	20 words — < 1%
16	eresources.uin-malang.ac.id Internet	20 words — < 1%
17	lra.le.ac.uk Internet	19 words — < 1%
18	munin.uit.no Internet	19 words — < 1%
19	A. R. Fogg, M. Lester, T. K. Yeoman, A. G. Burrell, S. M. Imber, S. E. Milan, E. G. Thomas, H. Sangha, B. J. Anderson. "An Improved Estimation of SuperDARN Heppner ‐ Maynard Boundaries Using AMPERE Data", Journal of Geophysical Research: Space Physics, 2020 Crossref	15 words — < 1%
20	Ningchao Wang, Jia Yue, Wenbin Wang, Liying Qian, Lan Jian, Jie Zhang. "A Comparison of the CIR ‐ and CME ‐ Induced Geomagnetic Activity Effects on Mesosphere and Lower Thermospheric Temperature", Journal of Geophysical Research: Space Physics, 2021 Crossref	15 words — < 1%
21	Victoriya V. Forsythe, Roman A. Makarevich. " Global view of the region irregularity and convection velocities in the high ‐ latitude Southern Hemisphere ", Journal of Geophysical Research: Space Physics, 2017 Crossref	15 words — < 1%

22	agupubs.pericles-prod.literatumonline.com Internet	15 words — < 1%
23	patents.google.com Internet	15 words — < 1%
24	Atsuki Shinbori, Yuichi Otsuka, Takuya Sori, Takuya Tsugawa, Michi Nishioka. "Temporal and Spatial Variations of Total Electron Content Enhancements During a Geomagnetic Storm on 27 and 28 September 2017", Journal of Geophysical Research: Space Physics, 2020 Crossref	14 words — < 1%
25	Currie, J. L., C. L. Waters, F. W. Menk, M. D. Sciffer, and W. A. Bristow. "SuperDARN Backscatter During Intense Geomagnetic Storms : STORMTIME SUPERDARN BACKSCATTER", Radio Science, 2016. Crossref	14 words — < 1%
26	export.arxiv.org Internet	12 words — < 1%
27	nipr.repo.nii.ac.jp Internet	12 words — < 1%
28	J. A. Wild. "The influence of magnetospheric substorms on SuperDARN radar backscatter", Journal of Geophysical Research, 04/25/2008 Crossref	11 words — < 1%
29	arxiv.org Internet	11 words — < 1%
30	earth-planets-space.springeropen.com Internet	11 words — < 1%
31	link.springer.com	

Internet

11 words — < 1%

32 A. Grocott. "Multi-instrument observations of the ionospheric counterpart of a bursty bulk flow in the near-Earth plasma sheet", Annales Geophysicae, 04/02/2004

Crossref

10 words — < 1%

33 Philip J. Erickson. "Mid-latitude ionospheric features: Natural complexity in action", Elsevier BV, 2020

Crossref

10 words — < 1%

34 Stephen E. Milan, Jennifer A. Carter, Harneet Sangha, Gemma E. Bower, Brian J. Anderson. "Magnetospheric Flux Throughput in the Dungey Cycle: Identification of Convection State During 2010", Journal of Geophysical Research: Space Physics, 2021

Crossref

10 words — < 1%

35 www.science.gov

Internet

10 words — < 1%

36 A. Grocott. "Ionospheric flow during extended intervals of northward but <i>B_y</i> - dominated IMF", Annales Geophysicae, 2003

Crossref

9 words — < 1%

37 Fiori, R.A.D., D.H. Boteler, D. Knudsen, J. Burchill, A.V. Koustov, E.D.P. Cousins, and C. Blais. "Potential impact of Swarm electric field data on global 2D convection mapping in combination with SuperDARN radar data", Journal of Atmospheric and Solar-Terrestrial Physics, 2013.

Crossref

9 words — < 1%

38 de Larquier, S., P. Ponomarenko, A. J. Ribeiro, J. M. Ruohoniemi, J. B. H. Baker, K. T. Sterne, and M. Lester. "On the spatial distribution of decameter-scale subauroral ionospheric irregularities observed by SuperDARN radars : SUBAURORAL IONOSPHERIC IRREGULARITIES", Journal of Geophysical Research Space Physics, 2013.
Crossref

9 words — < 1%

39 Aoi Nakamizo, Akimasa Yoshikawa. "Deformation of Ionospheric Potential Pattern by Ionospheric Hall Polarization", Journal of Geophysical Research: Space Physics, 2019
Crossref

8 words — < 1%

40 Chizurumoke M. Michael, Tim K. Yeoman, Darren M. Wright, Steve E. Milan, Matthew K. James. "A Ray Tracing Simulation of HF Ionospheric Radar Performance at African Equatorial Latitudes", Radio Science, 2020
Crossref

8 words — < 1%

41 Elvira Astafyeva, Mala S. Bagiya, Matthias Förster, Nozomu Nishitani. "Unprecedented Hemispheric Asymmetries During a Surprise Ionospheric Storm: A Game of Drivers", Journal of Geophysical Research: Space Physics, 2020
Crossref

8 words — < 1%

42 F. D. Wilder. "Reverse convection potential saturation during northward IMF", Geophysical Research Letters, 06/20/2008
Crossref

8 words — < 1%

43 Liu, E. X., H. Q. Hu, R. Y. Liu, Z. S. Wu, and M. Lester. "An adjusted location model for SuperDARN backscatter echoes", Annales Geophysicae, 2012.
Crossref

8 words — < 1%

44	M. Volwerk. "Interplanetary magnetic field rotations followed from L1 to the ground: the response of the Earth's magnetosphere as seen by multi-spacecraft and ground-based observations", Annales Geophysicae, 09/08/2011 Crossref	8 words — < 1%
45	Q.-H. Zhang, B.-C. Zhang, M. Lockwood, H.-Q. Hu et al. "Direct Observations of the Evolution of Polar Cap Ionization Patches", Science, 2013 Crossref	8 words — < 1%
46	Q.-H. Zhang, M. W. Dunlop, R.-Y. Liu, H.-G. Yang et al. "Coordinated Cluster/Double Star and ground-based observations of dayside reconnection signatures on 11 February 2004", Annales Geophysicae, 2011 Crossref	8 words — < 1%
47	eprints.soton.ac.uk Internet	8 words — < 1%
48	libtreasures.utdallas.edu Internet	8 words — < 1%
49	van der Meeren, Christer, Kjellmar Oksavik, Dag Lorentzen, Jøran Idar Moen, and Vincenzo Romano. "GPS scintillation and irregularities at the front of an ionization tongue in the nightside polar ionosphere : VAN DER MEEREN ET AL.", Journal of Geophysical Research Space Physics, 2014. Crossref	8 words — < 1%
50	A. Etemadi, S.W.H. Cowley, M. Lockwood, B.J.I. Bromage, D.M. Willis, H. Lühr. "The dependence of high-latitude dayside ionospheric flows on the North-South component of the IMF: A high time resolution correlation analysis using EISCAT “Polar” and AMPTE UKS and IRM data",	7 words — < 1%

Planetary and Space Science, 1988
Crossref

51 D.D. Billett, J.A. Wild, A. Grocott, A.L. Aruliah, A.M. Ronksley, M. ‐ T. Walach, M. Lester. "Spatially Resolved Neutral Wind Response Times During High Geomagnetic Activity Above Svalbard", Journal of Geophysical Research: Space Physics, 2019 — 7 words — < 1%
Crossref

52 J. A. Carter, S. E. Milan, A. R. Fogg, H. Sangha, M. Lester, L. J. Paxton, B. J. Anderson. "The evolution of long ‐ duration cusp spot emission during lobereconnection with respect to field ‐ aligned currents", Journal of Geophysical Research: Space Physics, 2020 — 7 words — < 1%
Crossref

53 J. E. Stawarz, L. Matteini, T. N. Parashar, L. Franci et al. "Comparative Analysis of the Various Generalized Ohm's Law Terms in Magnetosheath Turbulence as Observed by Magnetospheric Multiscale", Journal of Geophysical Research: Space Physics, 2020 — 7 words — < 1%
Crossref

54 Lopez, Ramon E.. "The integrated dayside merging rate is controlled primarily by the solar wind : The solar wind controls dayside merging", Journal of Geophysical Research Space Physics, 2016. — 7 words — < 1%
Crossref

55 Motoharu Nowada, Robert C. Fear, Adrian Grocott, Quan-Qi Shi et al. "Subsidence of Ionospheric Flows Triggered by Magnetotail Magnetic Reconnection During Transpolar Arc Brightening", Journal of Geophysical Research: Space Physics, 2018 — 7 words — < 1%
Crossref

56	S. E. Milan, J. A. Carter, G. E. Bower, S. M. Imber, L. J. Paxton, B. J. Anderson, M. R. Hairston, B. Hubert. "Dual‐lobe reconnection and horse‐collar auroras", Journal of Geophysical Research: Space Physics, 2020 Crossref	7 words — < 1%
57	S. Saito, S. Kurita, Y. Miyoshi, S. Kasahara et al. "Data‐driven simulation of rapid flux enhancement of energetic electrons with an upper‐band whistler burst", Journal of Geophysical Research: Space Physics, 2021 Crossref	7 words — < 1%
58	angeo.copernicus.org Internet	7 words — < 1%
59	A. Grocott. "The influence of IMF By on the nature of the nightside high-latitude ionospheric flow during intervals of positive IMF Bz", Annales Geophysicae, 04/08/2004 Crossref	6 words — < 1%
60	Astrophysics and Space Science Proceedings, 2015. Crossref	6 words — < 1%
61	Q.-H. Zhang. "Coordinated Cluster/Double Star and ground-based observations of dayside reconnection signatures on 11 February 2004", Annales Geophysicae, 10/24/2011 Crossref	6 words — < 1%
62	worldwidescience.org Internet	5 words — < 1%