Supporting Information for "Seasonal and hemispheric asymmetries in the cold ion outflow source region: Swarm and CHAMP observations of F-region polar cap plasma density"

S. M. Hatch[1] [*], S. Haaland[1,2], K. M. Laundal[1], T. Moretto[1], A. W. Yau[3], L.

Bjoland[4], J. P. Reistad[1], A. Ohma[1], K. Oksavik[1,4]

[1]Birkeland Centre for Space Science, University of Bergen, Norway

[2]Max-Planck Institute for Solar Systems Research, Göttingen, Germany

[3]Department of Physics and Astronomy, University of Calgary, Calgary, Alberta, Canada

[4]Department of Arctic Geophysics, University Centre in Svalbard, Longyearbyen, Norway

Contents of this file

————

Corresponding author: S. M. Hatch, Birkeland Centre for Space Science, University of Bergen,

Allégaten 55, Bergen, Norway. (Spencer.Hatch@uib.no)

[*]Department of Physics and Technology,

Allégaten 55, N-5007 Bergen, Norway

August 14, 2020, 7:00am

Introduction

In Text S1 and Figure S1 of this Supporting Information we provide a description and plot of the ratio of the Northern and Southern Hemisphere polar cap areas and circumferences.

In Text S2 and Figures S2–S3 we give a detailed description of the process for adjusting plasma densities measured by the Swarm and CHAMP satellites for variations in altitude and solar activity, which is summarized in the Appendix of the main article.

In Text S3 we describe the methodology we use for calculating the 95% confidence interval of the median.

Text S1.

Figure S1 displays the ratio of the Northern and Southern Hemisphere polar cap area (blue solid line) and circumference (orange dotted line) at 0-km altitude as a function of bounding Modified Apex magnetic latitude. Modified Apex coordinates are defined using a reference altitude of 110 km. Area and perimeter calculations are performed using the WGS84 ellipsoid and the `Geodesic` module of the `geographiclib` Python package (`https://geographiclib.sourceforge.io/html/python/`). Modified Apex coordinate conversions are performed using the `apexpy` Python package (see text).

August 14, 2020, 7:00am

Text S2.

As mentioned in section 2, we apply the Lomidze, Knudsen, Burchill, Kouznetsov, and Buchert (2018) calibrations

$$N_e = \nu_\alpha^2 \, N_{e,\alpha}. \tag{1}$$

to Swarm plasma density measurements. Here N_e is the flight-calibrated plasma density measurement, α indexes the three Swarm satellites, $N_{e,\alpha}$ is the plasma density measured by the LP instrument, and ν_α is one of the three satellite-dependent correction factors

$$\nu_A = 1.1067;$$

$$\nu_B = 1.0882;$$

$$\nu_C = 1.1157.$$

We additionally account for differences in altitude and solar activity via (i) application of an empirically derived scale height to N_e measurements that "maps" N_e to a common geodetic altitude of 500 km, and (ii) application of an empirically derived correction factor that accounts for the variation of N_e measurements with 10.7-cm solar radio flux (otherwise known as the F10.7 index).

S2.1 Adjustment of N_e for Altitude Variations

As pointed out in section 2 the CHAMP satellite and the three Swarm satellites cover different altitude ranges, and thus different portions of the F region. Figures S2a–b display the geodetic altitude coverage of each satellite in the NH and SH geomagnetic polar cap. The altitude range covered by the CHAMP satellite (315–455 km) is the lowest of all four satellites, while the altitude range covered by Swarm B (500–545 km) is mostly above those of the other satellites. The narrow range of geodetic altitudes 501–510 km in the

NH geomagnetic polar cap (518–527 km in the SH geomagnetic polar cap) where Swarm
B and Swarm C overlap represents the initial Swarm flight configuration when all three
satellites flew at approximately 500-km geodetic altitude in the NH (520-km geodetic
altitude in the SH). The orbit of Swarm B was later boosted to higher altitudes, while
the orbits of Swarm A and Swarm C were lowered. (The altitude distribution covered by
Swarm A is not shown since it and Swarm C have followed very similar orbital trajectories
since launch.)

Comparison of the geodetic altitude distributions in each hemisphere indicates a sys-
tematic difference in each satellite's range of altitudes: in each case the NH geomagnetic
polar cap distribution of altitudes is systematically offset by approximately 20 km relative
to the SH geomagnetic polar cap distribution. This is due to the combined effects of the
Earth's oblateness and the slightly elliptical orbit of each satellite.

From 380 conjunctions between Swarm B and either Swarm A or Swarm C for which
Swarm LP data is currently available (Dec 2013 through Feb 2020) we derive an em-
pirical scale height. These conjunctions were identified via the Query tool (`https://`
`sscweb.gsfc.nasa.gov/cgi-bin/Query.cgi`) at NASA's Satellite Situation Center Web
(`https://sscweb.gsfc.nasa.gov/`) by requiring a horizontal separation of less than 100
km after radially tracing the footpoint of each satellite to a common altitude.

For each conjunction identified by the Query tool, we use a 10-minute window to cal-
culate the time at which the angular separation

$$\Delta\zeta = \arccos\left[\sin\lambda_1^m \sin\lambda_2^m + \cos\lambda_1^m \cos\lambda_2^m \cos\left(\mid\theta_1 - \theta_2\mid\right)\right], \qquad (2)$$

between the two satellites is a minimum. Here λ^m and θ respectively denote MLat and
geomagnetic longitude in MA_{110} coordinates. We then use this more precise list of con-

August 14, 2020, 7:00am

junction times to calculate the Vertical Scale Height $\text{VSH} \equiv dh/(d \ln N_e)$ (e.g., Hu et al., 2019).

Each conjunction corresponds to a single point in Figure S2c, which shows the logarithm of the ratio of N_e measurements made by Swarm B and either Swarm A or Swarm C on the x axis, and the altitude separation Δh in kilometers between the satellite pair on the y axis. The circle and star markers respectively denote dayside (6–18 MLT) and nightside (18–6 MLT) conjunctions. The spread in the logarithm of density ratios on the x axis indicate overall significant variability in the plasma density at each altitude. This spread is not surprising given the various contributions to density made by, for example, plasma convection, polar cap patches, and auroral precipitation. On the other hand, individual points in Figure S2c also indicate the existence of a particular scale height, or an approximately linear relationship between altitude difference and the logarithm of the ratio of plasma density at each altitude.

We use so-called "robust regression" to estimate the VSH from measurements in Figure S2c. In specific we perform an iterative Huber-weighted least-squares linear regression (e.g., Huber (1973); Holland and Welsch (1977)) to data in Figure S2c with the function

$$\log \left(N_e^{\text{A,C}}/N_e^{\text{B}} \right) = \beta \Delta h, \tag{3}$$

where $\beta \equiv 1/\text{VSH}$. We use the `HuberRegressor` module of the `Scikit-learn` Python package (Pedregosa et al., 2011) with $\epsilon = 1.5$. The ϵ parameter in Huber-weighted iterative regression controls the degree to which the regression is sensitive to outlier points.

From this regression we obtain the purple and brown lines, which respectively correspond to VSH= 205 km on the "dayside" and VSH= 167 km on the "nightside." We use magnetic local time (MLT) in MA_{110} coordinates to define "dayside" as $6 \leq \text{MLT} < 18$,

and "nightside" as MLT < 6 and MLT ≥ 18. These VSH values are within the range

of typical estimates at geodetic altitudes of 350–500 km (see, e.g., Figure 2 in Hu et al.

(2019) and Figure 1B in Stankov and Jakowski (2006)). To each N_e measurement we then

apply a scaling factor

$$N_{e,h_0} = N_e \exp \left[(h - h_0) / \text{VSH} \right], \tag{4}$$

where h is the altitude at which the measurement is made and VSH is the empirical scale

height. We arbitrarily select a reference geodetic altitude $h_0 = 500$ km. This scaling

decreases the value of N_e for measurements made below h_0 and increases the value of N_e

for measurements made above h_0.

S2.2 Adjustment of N_e for $F10.7$ Variations

In addition to variation with altitude, polar cap N_e also varies with the intensity of

sunlight. We use an average of the $F10.7$ index during the preceding 27 days, denoted by

$\langle F10.7 \rangle_{27}$, as a proxy for solar EUV intensity. (The $F10.7$ index is publically available

via the NASA OMNI database at `https://omniweb.gsfc.nasa.gov/form/dx1.html`.)

Another common choice for averaging the $F10.7$ index is a centered 81-day window (e.g.,

Liu & Chen, 2009; Schunk & Nagy, 2009). We have elected to use $\langle F10.7 \rangle_{27}$ instead, since

we find that the RMS error between $\langle F10.7 \rangle_{27}$ and N_{e,h_0} is slightly ($\sim 5\%$) lower than the

RMS error between an 81-day centered average of the $F10.7$ index and N_{e,h_0}.

Figure S3a shows height-scaled polar cap N_{e,h_0} plotted versus $\langle F10.7 \rangle_{27}$. The orange

line indicates median N_{e,h_0} values within bins of 5 for $\langle F10.7 \rangle_{27}$ between 65 and 200. The

variation of median N_{e,h_0} with $\langle F10.7 \rangle_{27}$ is approximately linear.

To scale N_{e,h_0} measurements for variation with $F10.7$, we perform an iterative Huber-

weighted least-squares linear regression to all 3,456,077 N_{e,h_0} values with a model of the

form

$$N_{e,h_0} = A\langle F10.7\rangle_{27} - B = N\left(a\langle F10.7\rangle_{27} - 1\right). \tag{5}$$

In practice the actual best-fit parameters are not sensitive to our choice of initial parameter estimates, but for illustrative purposes we use the two points indicated in Figure S3a, $P_1 = (F_1, N_1) = (67.5, 37516)$ and $P_2 = (F_2, N_2) = (107.5, 77460)$ together with the model in Equation (5) to obtain the initial model estimates $\hat{A} = (N_2 - N_1)/(F_2 - F_1) = 999$ and $\hat{B} = \hat{N} = \hat{A}F_1 - N_1 = 29890$, such that $\hat{a} = \hat{A}/\hat{B} = 0.033$.

The resulting best-fit model parameters with weighting parameter $\epsilon = 1.5$ are $a = 0.02564$ and $N = 46,780$; the resulting model of the form given by Equation (5) is indicated by the black dashed line in Figure S3a. These fit parameters are obtained by requiring that the relative change in each model parameter be less than 10^{-8} after each iteration, which is generally acheived after 35–60 iterations. We apply the portion of the model in Equation (5) that is dependent on $\langle F10.7\rangle_{27}$, namely the parameter a, to each N_{e,h_0} value to finally obtain the height- and solar flux-adjusted density

$$N_e^* = N_{e,h_0}\frac{80a - 1}{a\langle F10.7\rangle_{27} - 1}. \tag{6}$$

The numerator in Equation (6) scales the final adjusted density N_e to a nominal solar flux level of $\langle F10.7\rangle_{27} = 80$. These final adjusted densities are shown in Figure S3b.

August 14, 2020, 7:00am

Text S3.

The 95% confidence interval of each median displayed in Figure 1a in the main article

is calculated in each DOY bin by first sorting the N_e^* values in that DOY bin, and then

identifying the value of N_e^* corresponding to the Lth and Uth sorted sample in that DOY

bin. These sorted sample numbers are given by (Conover, 1999)

$$L = \lfloor Q_b(0.025, N, 0.5) \rfloor;$$

$$U = \lceil Q_b(0.975, N, 0.5) \rceil;$$

(7)

where $\lfloor \rfloor$ and $\lceil \rceil$ are respectively floor and ceiling functions and $Q_b(p, N, q)$ is the quan-

tile function, otherwise known as the "inverse cumulative distribution" or "percent-point

function," of the binomial distribution. The parameters of this quantile function are the

probability p, the number of observations N, and the quantile of interest q. We use

$p = 0.025, 0.975$ corresponding to the 95% confidence interval, and $q = 0.5$ correspond-

ing to the 50% quantile or median. We calculate $Q_b(p, N, q)$ via the `stats.binom.ppf`

method of the `scipy` Python package (Virtanen et al., 2020).

August 14, 2020, 7:00am

References

Conover, W. J. (1999). *Practical Nonparametric Statistics* (3rd ed.). John Wiley & Sons.

Holland, P. W., & Welsch, R. E. (1977, jan). Robust regression using iteratively reweighted least-squares. *Communications in Statistics - Theory and Methods*, *6*(9), 813–827. doi: 10.1080/03610927708827533

Hu, A., Carter, B., Currie, J., Norman, R., Wu, S., Wang, X., & Zhang, K. (2019, jun). Modeling of Topside Ionospheric Vertical Scale Height Based on Ionospheric Radio Occultation Measurements. *Journal of Geophysical Research: Space Physics*, *124*(6), 4926–4942. doi: 10.1029/2018JA026280

Huber, P. J. (1973, sep). Robust Regression: Asymptotics, Conjectures and Monte Carlo. *The Annals of Statistics*, *1*(5), 799–821. doi: 10.1214/aos/1176342503

Liu, L., & Chen, Y. (2009, oct). Statistical analysis of solar activity variations of total electron content derived at Jet Propulsion Laboratory from GPS observations. *Journal of Geophysical Research: Space Physics*, *114*(A10). doi: 10.1029/2009JA014533

Lomidze, L., Knudsen, D. J., Burchill, J., Kouznetsov, A., & Buchert, S. C. (2018). Calibration and validation of Swarm plasma densities and electron temperatures using ground-based radars and satellite radio occultation measurements. *Radio Science*, *53*(1), 15–36. doi: 10.1002/2017RS006415

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Schunk, R., & Nagy, A. (2009). *Ionospheres* (Second ed.). Cambridge: Cambridge

University Press. doi: 10.1017/CBO9780511635342

Stankov, S., & Jakowski, N. (2006, jan). Topside ionospheric scale height analysis and modelling based on radio occultation measurements. *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(2), 134–162. doi: 10.1016/j.jastp.2005.10.003

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . Contributors, S. . . (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*. doi: https://doi.org/10.1038/s41592-019 -0686-2
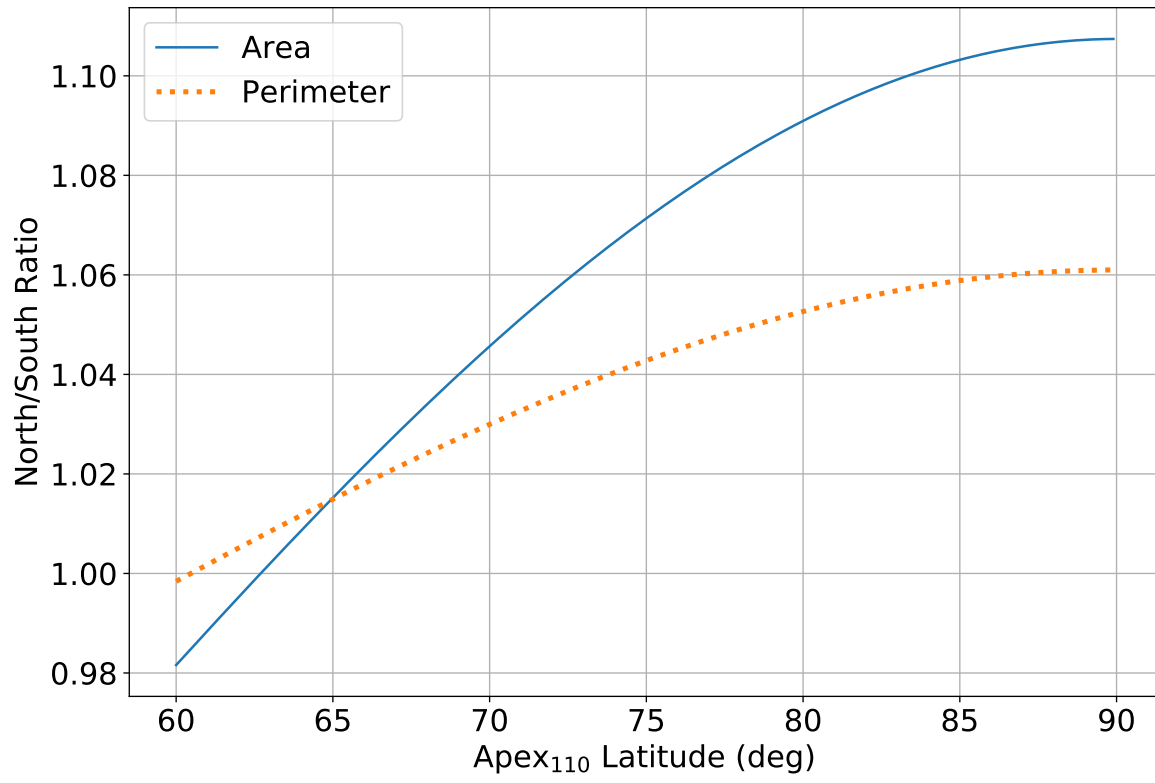
Figure S1. See Text S1.

August 14, 2020, 7:00am
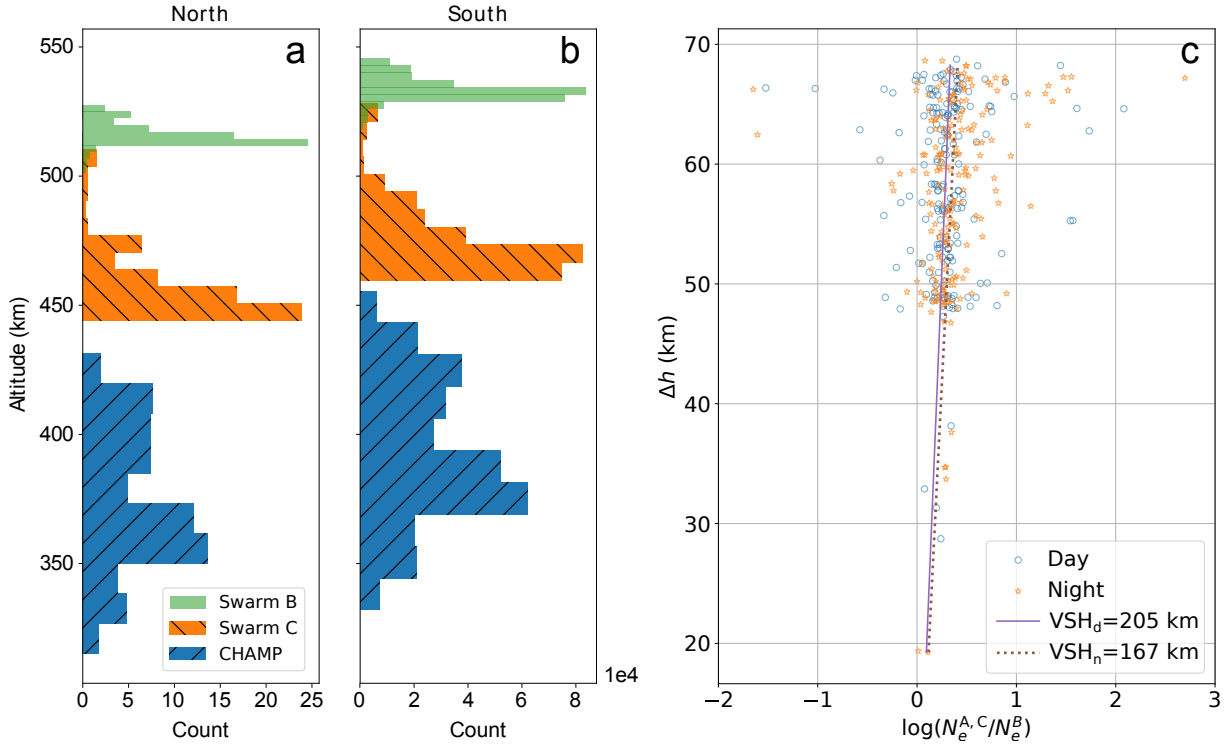
Figure S2. Panels a and b show geodetic altitude distributions of Swarm B (green), Swarm C (orange), and CHAMP (blue) satellites in the NH and SH geomagnetic polar caps above 80° MLat, respectively. (The altitude distribution of Swarm A is essentially identical to Swarm C and is not shown.) Panel c shows empirical Vertical Scale Height (VSH) estimates based on 380 conjunctions (see text for definition) between Swarm B and one of the two satellites at lower altitudes (either Swarm A or Swarm C) in the period from 10 Dec 2013 and 6 Feb 2020 for which Swarm LP measurements are available. The x axis indicates the logarithm of the ratio of Lomidze et al. (2018)-corrected plasma density measurements, and the y axis indicates the spacecraft altitude separation. The circle and star markers respectively denote dayside (6–18 MLT) and nightside (18–6 MLT) conjunctions.
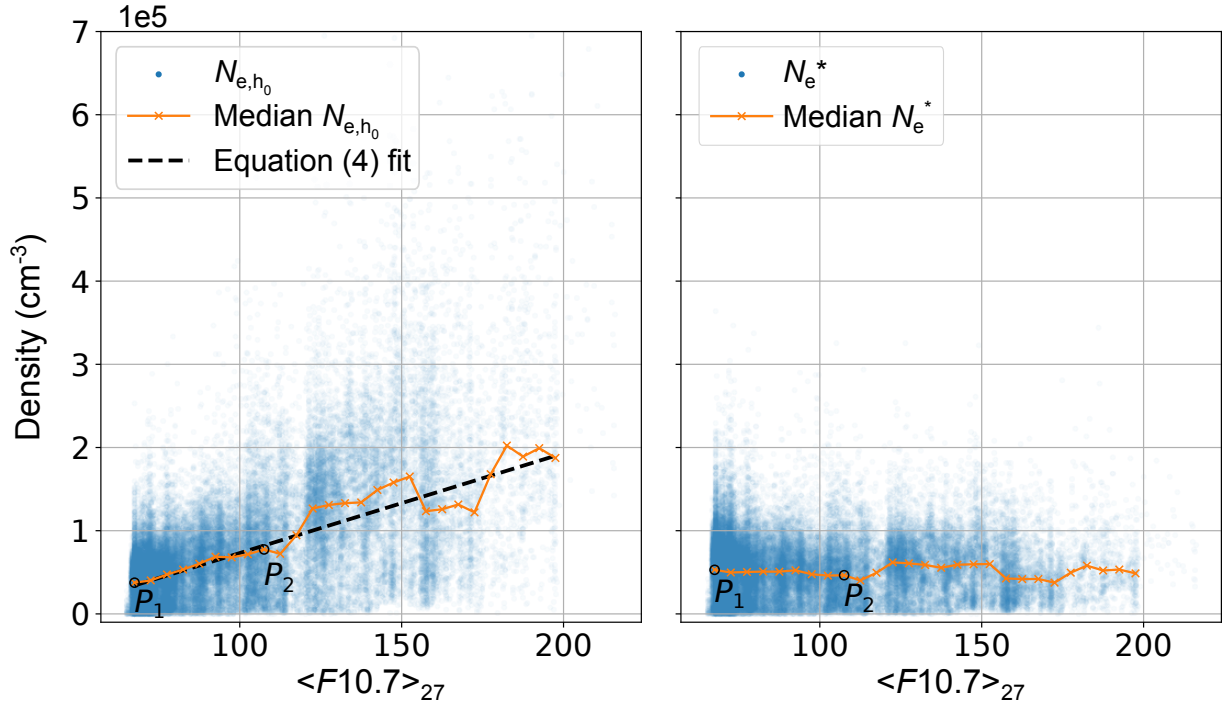
Figure S3. Height-scaled plasma densities before (left) and after (right) adjusting for variation in $\langle F10.7 \rangle_{27}$ using Equation (6). In the left panel the black dashed line indicates the model given by Equation (5) with best-fit parameters $a = 0.02564$ and $N = 46{,}780$ obtained by performing the iterative Huber-weighted least-squares nonlinear regression described in the text. The two points P_1 and P_2 indicate the two points used to obtain initial estimates for model parameters a and N, and are shown in the right panel to indicate the effect of applying Equation (6) to N_{e,h_0} values. In both panels the orange line indicates the median value, N_{e,h_0} in the left panel and N_e^{adj} in the right panel, within $\langle F10.7 \rangle_{27}$ bins of 5.

August 14, 2020, 7:00am