

# Multi-point data analysis using Swarm electron density data

Aksel Graneng



Thesis submitted for the degree of  
Master in Space physics and space technology  
60 credits

Department of Physics  
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

Spring 2021



# Multi-point data analysis using Swarm electron density data

Aksel Graneng

© 2021 Aksel Graneng

Multi-point data analysis using Swarm electron density data

<http://www.duo.uio.no/>

Printed: Reprosentralen, University of Oslo

## Abstract

In this paper, we attempt a multi-point analysis of electron density data gathered by the Swarm mission. The Swarm mission consists of three identical satellites launched by the European Space Agency in 2013. In December 2013, the Swarm satellites flew in a pearls-on-a-string configuration, giving us the opportunity to do a multi-point analysis. In our analysis, we introduce the comparison index, an index that tells us the difference between two electron density time series. We do a statistical analysis of computed comparison indices for the 9th through 31st of December, and see if we can differentiate between spatial and temporal effects. We do this by inspecting how the statistical moments of the comparison index change as the satellites drift apart. We find that the Swarm data is best suited for multi-point analysis in the high latitude polar cap regions. For the northern polar cap, we find that the comparison index tends to increase as the distance between satellites increases, but that we cannot differentiate between spatial and temporal effects.

# Acknowledgements

I would like to thank my supervisors, Lasse B.N Clausen and Andres Spicher, for their ceaseless support, advice and patience. In the middle of a pandemic, they always found the time to give me great advice and feedback.

I would also like to thank my fellow space physics students, Amund, Lise, Michael, Saida and Sean. Despite the reality of home office, they managed to keep me motivated and in good spirits while also providing solid discussions and valuable input on my thesis.

Swarm is a European Space Agency (ESA) project. Data used in this study are available from ESA [\[3\]](#).

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Theory</b>	<b>5</b>
2.1	The ionosphere . . . . .	5
2.1.1	Introducing the ionosphere . . . . .	5
2.1.2	Earth's magnetic field and the solar wind . . . . .	7
2.1.3	Scintillations and instabilities . . . . .	9
2.2	Spectral analysis . . . . .	11
<b>3</b>	<b>Instrumentation and Method</b>	<b>22</b>
3.1	The Swarm Mission . . . . .	22
3.2	Index-shifting . . . . .	25
3.3	Dealing with data gaps. . . . .	27
3.4	Histograms and linear regressions. . . . .	28
3.5	Assumptions and satellite drift . . . . .	29
<b>4</b>	<b>Results</b>	<b>31</b>
4.1	Comparison indices and statistical moments . . . . .	31
4.2	Assumptions and satellite drift . . . . .	37
<b>5</b>	<b>Discussion</b>	<b>43</b>
5.1	Discussing the method . . . . .	45
5.1.1	Index-shift . . . . .	46
5.1.2	Data gaps . . . . .	46
5.1.3	Dividing by geomagnetic latitude . . . . .	47
5.1.4	Parameters . . . . .	48
5.2	Interpretations . . . . .	49
5.2.1	Spatial and temporal changes . . . . .	49
5.2.2	Interpreting our results . . . . .	51
<b>6</b>	<b>Conclusion</b>	<b>54</b>
<b>7</b>	<b>Outlook</b>	<b>55</b>
<b>8</b>	<b>Appendix</b>	<b>56</b>
8.1	Codes . . . . .	56
8.2	Finding the along-track velocity of ionospheric structures. . . . .	56
8.3	Along-track velocity of large-scale ionospheric structures: an example	59
8.4	Other figures . . . . .	62

# 1 Introduction

As human activity becomes increasingly dependant on precise global navigation tools such as the Global Positioning System (GPS), understanding what factors may affect these becomes crucial. From a self-driving car crashing to a farmer plowing the wrong spot, relying on a un-reliable positioning system can have economic and potentially even fatal consequences.

One of the factors affecting a Global Navigational Satellite System (GNSS) such as GPS the spaceweather, and in particular the ionospheric conditions [17]. In order to reach a reciever on the ground, the signal from a GNSS satellite must pass through the ionosphere. The resulting position calculated by the GNSS is sensitive to changes in the electron density of the plasma which the signal passed through on its way through the ionosphere. As such, rapid changes in the electron density can effectively ruin the GNSS precision.

Understanding ionospheric structures and how these change is of great interest. If we could fully understand the ionospheric conditions, we could effectively make a space-weather forecast, which could be used to greatly enhance GNSS precision.

One way of gaining understanding of the ionosphere is through in-situ observations. By flying satellites and rockets mounted with instruments through the ionosphere, we can directly measure electron density [15] [12]. Flying a single satellite through an ionospheric structure gives us a picture of what the structure looked like at that time. In order to see how it changes over time, we need more than one satellite [2].

A problem arises when investigating how the ionospheric structures change over time with only two satellites. Say that we fly a satellite through a "bubble" of high electron density. We could quite easily find the width of the bubble by looking at the data. Now say that we fly a second satellite through the same path as the first satellite. We again find the width of the bubble, but now we find that the width is smaller than the first time. This could either be explained by the bubble shrinking over time; a temporal change, or by the bubble having moved so that the second satellite measured a different, thinner, part of the bubble; a spatial change.

In this paper, we will be investigating data analysis methods using data from ESA's Swarm mission; three satellites which in December of 2013 flew in a cross-polar pearls-on-a-string configuration [24]. The Swarm mission gives us an opportunity to see what, if anything, can be inferred about the differences in samplings of ionospheric structures, and whether we can say anything about their temporal and spatial nature. We will introduce the comparison index; an index that tells us the difference between the data gathered by the different satellites. We will compute comparison indices using electron density data from the Swarm mission. We will group these by latitudinal regions and do a statistical analysis in order to see if we can differentiate between spatial and temporal changes.

This paper is divided into 4 main sections. First we have the theory section. Here we will introduce the ionosphere and the comparison index. After this, we have instrumentation and method, where we introduce the Swarm satellites from which we got our data and the methods used to produce our results. In the results section, we present and describe our results. Following this, we discuss our results in the



discussion, focusing first on our method and then physical interpretations.

We round our paper off with concluding remarks and a few words on possible future work. We also include an appendix, where some extra methods and figures are shown.

## 2 Theory

In this section we introduce the ionosphere, as well as the spectral analysis methods we will use. We first lay down some background information on the ionosphere, giving a basic introduction to its structure and how it is relevant to us. We then introduce the Fourier transform and the comparison index, as well as building some intuition on what different comparison indices mean.

### 2.1 The ionosphere

A large part of space physics is the study of the ionosphere, a layer featuring electrically charged particles stretching from the upper atmosphere to the edge of space. In this paper, we investigate electron density data from the Swarm satellites who fly through the ionosphere as they orbit the Earth. In order to understand this data, we must first establish a basic understanding of the ionosphere and some of the phenomena therein.

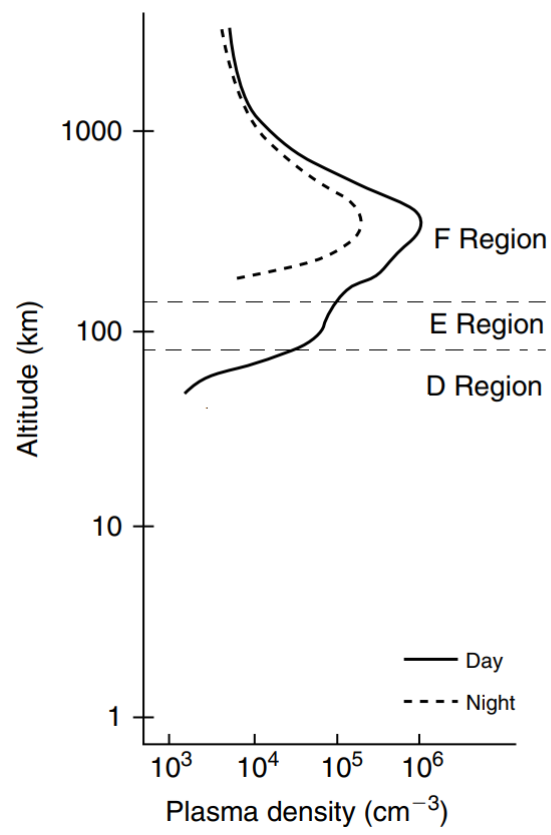
#### 2.1.1 Introducing the ionosphere

In 1925, physicists performed experiments where they sent short pulses of radiowaves into the sky and observed that they were reflected [1]. This could only be explained by the presence of a reflecting layer composed of electrons and positive ions, confirming the existence of the ionosphere [6]. The first ionospheric region discovered was named the E region for the electric vector of the reflected waves. Subsequently, later regions followed the same naming scheme with the lowest region being named the D region, and the upper region being named the F region.

Later, with the advent of space programs, in-situ measurements through the use of balloons, rockets and satellites became available. Combined with ground-based measurements, this has let us map out the general electron density profile of the ionosphere, as seen in Figure 1.

The main driving force behind the presence of plasma in the ionosphere is sunlight [6]. When gas particles in the atmosphere are hit by sufficiently energetic photons, they can be stripped of one or more electrons, becoming ionized. As the number of free electrons created is generally equal to the number of positive ions, we can use electron density as a parameter to see how ionized the gas is. As sunlight penetrates the atmosphere, some of it is absorbed, reducing the solar flux at lower altitudes. Combined with an increased atmospheric density which in turn increases the ion recombination rate, we get an explanation to why the electron density in Figure 1 decreases exponentially at lower altitudes.

With the main driving force behind the presence of plasma being sunlight, the ionosphere can naturally be divided into a dayside and a nightside. In Figure 1, we

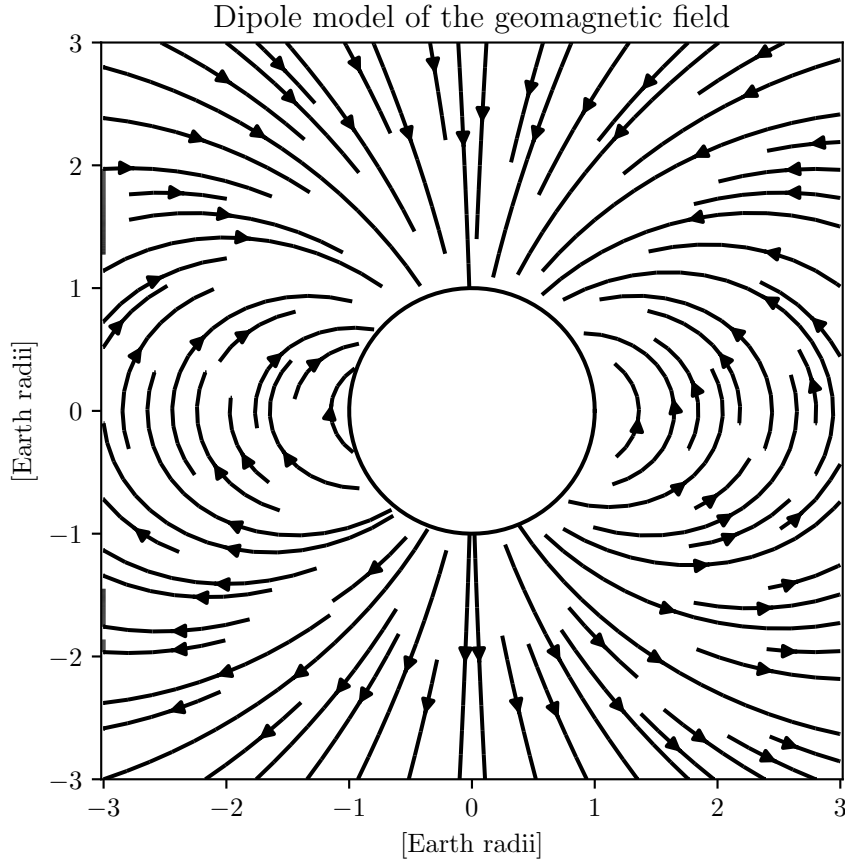


**Figure 1:** Ionospheric electron density profile. Adapted from Kelley (2009).

see the electron density profile for both the dayside and nightside. In the nightside ionosphere, the E and D regions practically disappear [6]. At lower altitudes, the plasma is mostly dominated by ionized molecules such as  $\text{O}_2^+$  and  $\text{NO}^+$ . At higher altitudes, from about where the F region starts, the dominating ions are ionized oxygen atoms  $\text{O}^+$ . Ionized atoms have a much slower recombination rate than ionized molecules. This leads to the plasma in the lower altitude regions quickly dissipating when the Sun sets as the molecular ions recombine, while the atomic ions in the F region slowly dissipate over the duration of the night, the plasma never fully neutralizing.

### 2.1.2 Earth's magnetic field and the solar wind

In order to fully understand the ionosphere and the phenomena we find there, we must understand Earth's magnetic field. The magnetic field is created by currents in the molten core of the Earth, and can be approximated as a dipole field above the surface [6]. In Figure 2 we see what such a model looks like, with magnetic field lines going from the south pole (down on the figure) to the north pole (up on the figure).

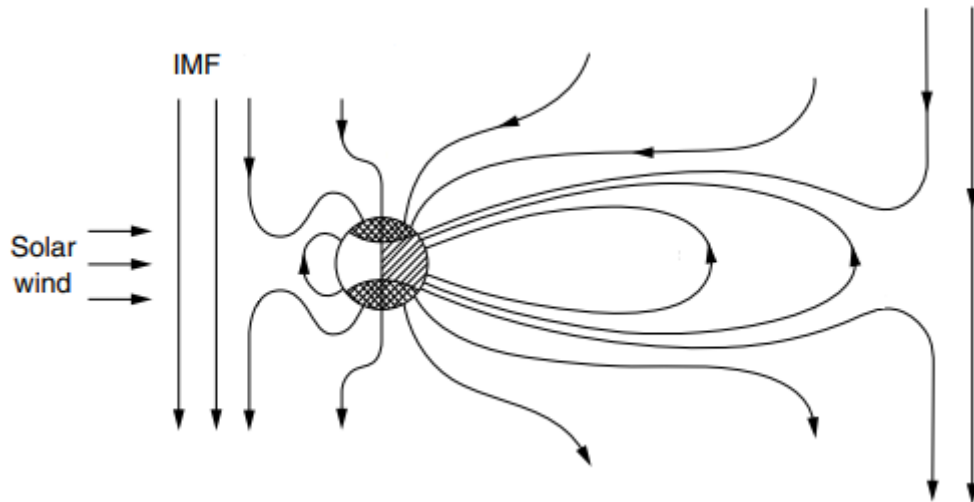


**Figure 2:** A dipole approximation to Earth's magnetic field.

As Earth orbits the Sun, it is constantly bombarded by charged particles in

what we call the solar wind. The solar wind originates in the Sun and carries the Sun's magnetic field, the interplanetary magnetic field (IMF). The solar wind travels at speeds on the order of hundreds of kilometers a second and has electron densities on the order of  $1 \text{ cm}^{-3}$  [1]. If the solar wind would be allowed to hit Earth's surface unopposed, it would be unlikely that life as we know it would have evolved. Thankfully, when the solar wind encounters Earth's magnetic field, it deflects away from the Earth's surface as the plasma follows the magnetic field lines.

If the Earth was isolated, Figure 2 would be a realistic representation of the geomagnetic field. This is not the case. Indeed, as Earth's magnetic field is hit by the solar wind and the IMF, it compresses and changes shapes depending on the direction and strength of the IMF. If the Z component of the IMF, the component parallel to the north-south axis of Earth's magnetic field, is strongly negative, Earth's magnetic field becomes strongly compressed on the dayside and Earth's magnetic field lines can connect with those of the IMF [6], as illustrated in Figure 3.



**Figure 3:** An illustration of Earth's magnetic field during a southward IMF. Adapted from Kelley (2009).

The interplay between the IMF and Earth's magnetic field is the driving force behind many of the ionospheric phenomena we observe [6]. When magnetic reconnection happens, the connection between field lines in the IMF and Earth's magnetic field happens, it can accelerate particles from the solar wind along the magnetic field lines and into the ionosphere. The particles hit the ionosphere in rings around the magnetic poles, in an area we call the auroral oval. This is what causes the northern lights. The position of the auroral oval is dependant on the properties of the IMF. If the IMF Z component is strongly negative, it can penetrate deeper the Earth's magnetic field. Then reconnection can happen with field lines closer to the equator, moving the auroral ovals accordingly. Historically, northern lights have been observed much further south than normal, to the astonishment of the observers. An example of this is the 1591 painting, seen in Figure 4, of northern lights over Nuremberg in modern south Germany, which lies at a latitude of  $49.5^\circ$ .



**Figure 4:** Nothern lights over Nuremberg, 1591. Source: The Norwegian Space Agency.

### 2.1.3 Scintillations and instabilities

When we gaze upon the night sky, we might notice the stars twinkling. This is a phenomena many people are familiar with, and even appears in the English lullaby "Twinkle, Twinkle, Little Star". The reason this flickering of the stars can be observed is that the starlight must pass through the atmosphere before hitting the Earth. On its way through a turbulent atmosphere, the amplitude and phase of the light changes [22]. This phenomenon was identified by Sir Isaac Newton, who in his book *Opticks* attributed the twinkling to "Tremors of the Atmosphere".

A similar effect to the twinkling of stars can be observed from satellite signals, except that they are more affected by the ionosphere than the atmosphere [22]. A GNSS finds your position by measuring your distance from a number of satellites orbiting the Earth. A reciever can calculate the distance between itself and GNSS satellites by timing how long an electromagnetic signal takes to reach it. Given that the signal travels at the speed of light  $c$ , this becomes a simple matter of multiplying the time  $t$  it took with the lightspeed. The distance  $x$  is then  $x = c \cdot t$ . However, we know that the speed of light is not constant when not in a vacuum. Indeed, it changes with the density of the medium it travels through. In particular, when traversing the ionosphere, the speed of light is affected by electron density [1]. If the electron density is constant, this is not a problem as we can compensate for it by adjusting the expected speed of light. If the electron density changes quickly however, it will look as if the distance between the reciever and the satellite also changes rapidly. Therefore, if the electron density the signal passes through experiences changes, we can see it from the distance measured by the satellite. This phenomena is known as ionospheric scintillation [22].

Investigations of ionospheric scintillations through analysis of radio signals from other galaxies and satellite data revealed the existance of ionospheric irregularities [22]. Over the years, much work has been done in understanding the various processes that can cause scintillation-inducing electron density irregularities in the ionosphere. The physical processes behind ionospheric irregularities are numerous and complicated [6]. They can depend on many factors, such as geographic location, the IMF, and whether it is day or night. Still, some irregularities have been

identified as driving factors behind ionospheric structures.

In the polar ionosphere, we believe that a dominant mechanism behind irregularities is the gradient drift instability (GDI) [11]. If an electron density gradient  $\nabla n_e$  is perpendicular to both a magnetic field and an electric field, who in turn are perpendicular to each other in such a way that we get a gradient drift  $\vec{V}_D = \vec{E} \times \vec{B}$  parallel to the electron density gradient, then a small perturbation in the electron density will be amplified [11]. These conditions can be found in the northern polar cap, where the magnetic field lines are pointing downwards, perpendicular to the ground and  $\vec{E}$  fields are parallel to the ground [6].

Instabilities can be characterized by a growth time, which gives us a time scale for how long it takes the instability to grow. For GDI, the growth time  $\tau_{GDI}$  is as in equation 1 [11].

$$\tau_{GDI} = \frac{L}{V} \quad (1)$$

Where  $V$  is the plasma drift velocity relative to the neutral gas, and  $L$  is the gradient scale length, given by:

$$L = N_0 \left( \frac{\Delta N}{\Delta x} \right)^{-1}$$

Where  $N_0$  is the background electron density and  $\Delta N/\Delta x$  is the density gradient parallel to the  $\vec{E} \times \vec{B}$  drift.

Knowing the growth time of a dominant instability can be useful in investigating whether a change in electron density is of a temporal nature. In this paper, we will be doing a multi-point data analysis of electron density data from the northern polar cap. With GDI being a dominant mechanism behind irregularities in this region, we will be able to infer whether temporal effects cause changes in our electron density measurements by comparing GDI growth times to the time between our measurements.

## 2.2 Spectral analysis

In 1807, the French mathematician and physicist Jean-Baptiste Joseph Fourier discovered that an arbitrary periodic function  $x(t)$  can be written as a linear combination of sines and cosines [4]:

$$x(t) = \sum_{k=-\infty}^{k=\infty} c_k e^{ik\Omega_0 t}$$

With  $c_k$  being the Fourier coefficients, and where each term in the summation is periodic with period  $T_0 = \frac{2\pi}{\Omega_0}$ .

In other words, we can rewrite any signal as a sum sines and cosines through a Fourier transform. This is also true for a discrete time series  $x[n]$ :

$$x[n] = \sum_{k=0}^{N-1} c_k e^{i\frac{2\pi}{N}kn}$$

The Fourier coefficients  $c_k$  form a Fourier series, defined by:

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\frac{2\pi}{N}kn}$$

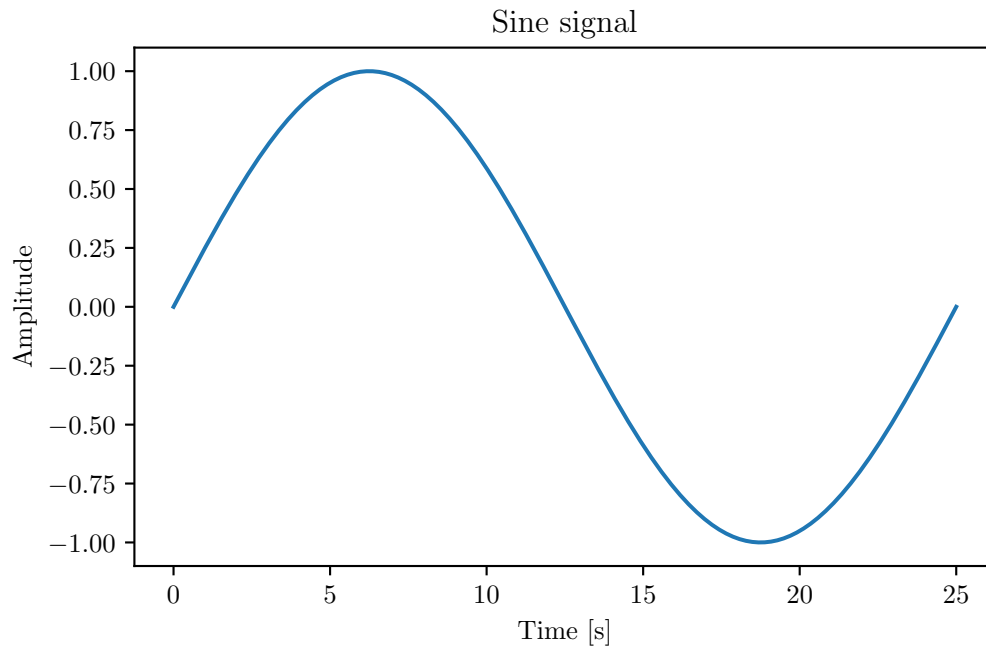
Every Fourier coefficient is a complex number  $a + ib$ . The real component relating to cosines, while the complex component relates to sines as in Euler's formula:

$$Re^{ix} = a \cos(x) + ib \sin(x), \quad R = |c_k|$$

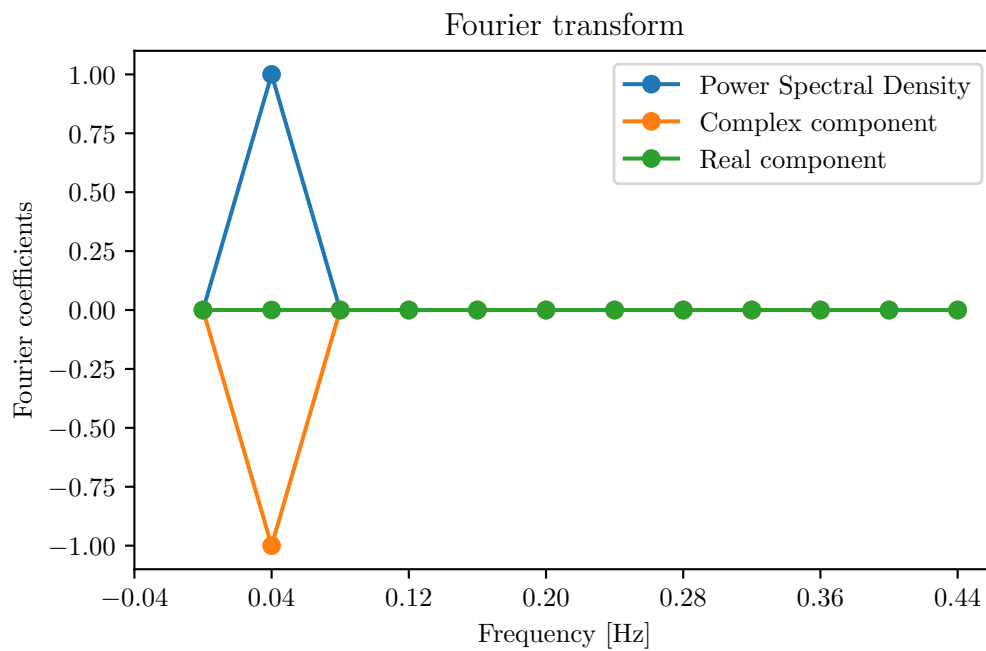
When interpreting Fourier transforms, it is common to look at the Power Spectral Density (PSD). This is done by looking at the absolute value of the Fourier coefficients  $|c_k|^2$ . Doing this lets us inspect what frequencies a signal is composed of, but truncates information on the phase.

The Fourier transform is a powerful tool, as it allows us to decompose a seemingly chaotic signal into its frequency components. In order to get a better understanding of the Fourier transform, we can look at some examples. In Figure 5, we see a sine wave with a frequency of 0.04 Hz sampled over one period, 25 seconds, at an arbitrarily high sampling frequency  $f_s$  so as to make the signal smooth.

In Figure 6, we see the lower frequency components of the Fourier transform of the signal in Figure 5. We see that the real component of the Fourier transform which corresponds to cosines, plotted in green, is zero across the board. This indicates that there are no cosine components in our signal. The complex component, corresponding to sines and plotted in orange, is zero everywhere except for at 0.04 Hz where it is -1. This indicates that our signal is indeed composed of a sine wave with a frequency of 0.04 Hz, with the minus sign indicating the signals phase. Plotted in blue, we see the PSD of the signal. It is also zero everywhere except for at 0.04 Hz where it is 1. Looking purely at the PSD, we can tell that the signal has a frequency component in 0.04 Hz, but information about the phase has been truncated.



**Figure 5:** A plain sine signal with a frequency of 0.04 Hz.

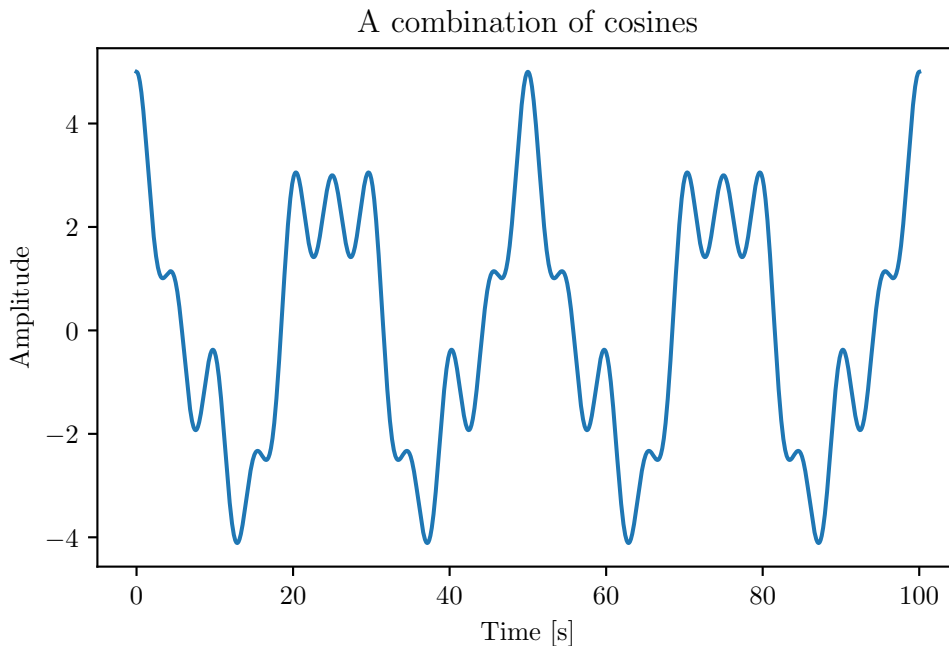


**Figure 6:** The Fourier transform of the signal in Figure 5.



While a good example, the information gained from the Fourier transform in Figure 6 could easily have been obtained by inspecting the signal itself. If a signal is composed of multiple frequencies however, the signal quickly becomes more complicated and the usefulness of the Fourier transform becomes apparent. As such, we can look at a more complex signal and its accompanying Fourier transform.

In Figure 7, we see a signal composed of cosines of multiple frequencies, sampled over 100 seconds. It is difficult to ascertain the frequency components of this signal by just looking at it. The Fourier transform now becomes a powerful tool in unraveling the composition of the signal.

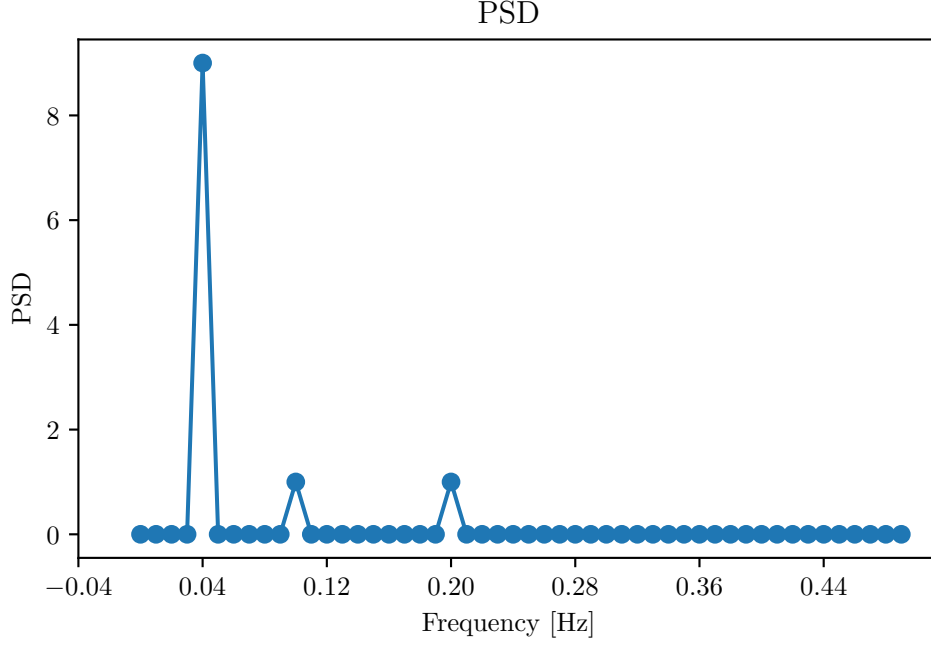


**Figure 7:** A more complicated signal, composed of multiple cosines.

In Figure 8, we see the PSD of the signal in Figure 7. We immediately see that there are three nonzero Fourier coefficients in the signal. They are at 0.04 Hz, 0.1 Hz and 0.2 Hz. The PSD value at 0.04 Hz is nine times larger than for 0.1 Hz and 0.2 Hz, meaning the sine or cosine with a frequency of 0.04 Hz has an amplitude three times larger than the other two. Since we are plotting the PSD of the signal, we cannot know the phase of the frequency components; they could be either sines, cosines or a combination.

It is worth noting that increasing the sampling frequency  $f_s$  when sampling a signal increases the sharpness of the signal in the time domain, but decreases the sharpness in the frequency domain. This is because the Fourier transform ranges from  $-f_s/2$  to  $f_s/2$ , stretching the x axis as the sampling frequency increases. In order to have both a smooth signal and an understandable Fourier transform, we have zoomed in on the lower frequency components in these examples.

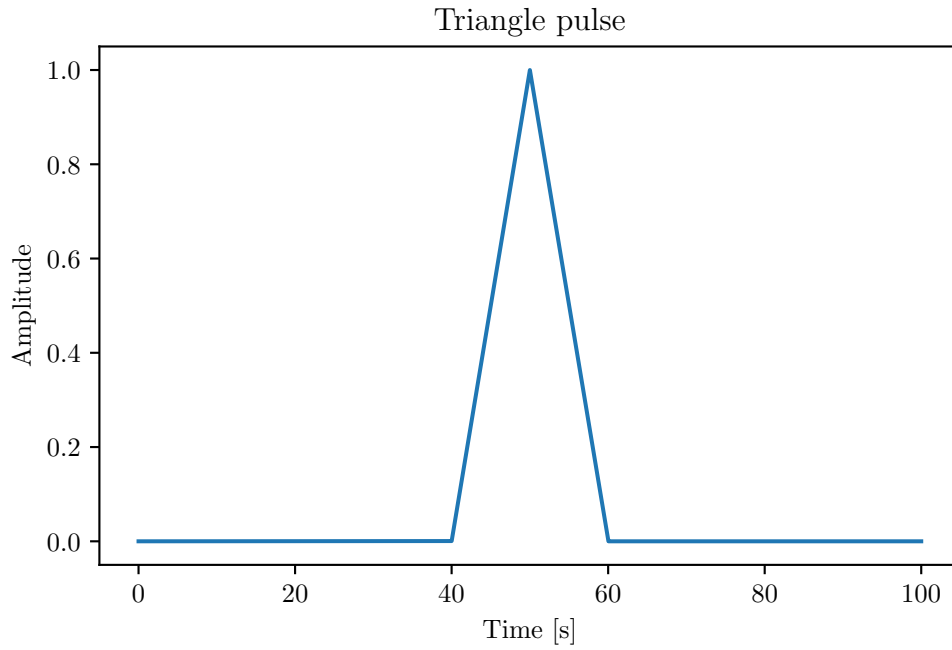
In this paper, we are interested in using the Fourier transform to inspect the



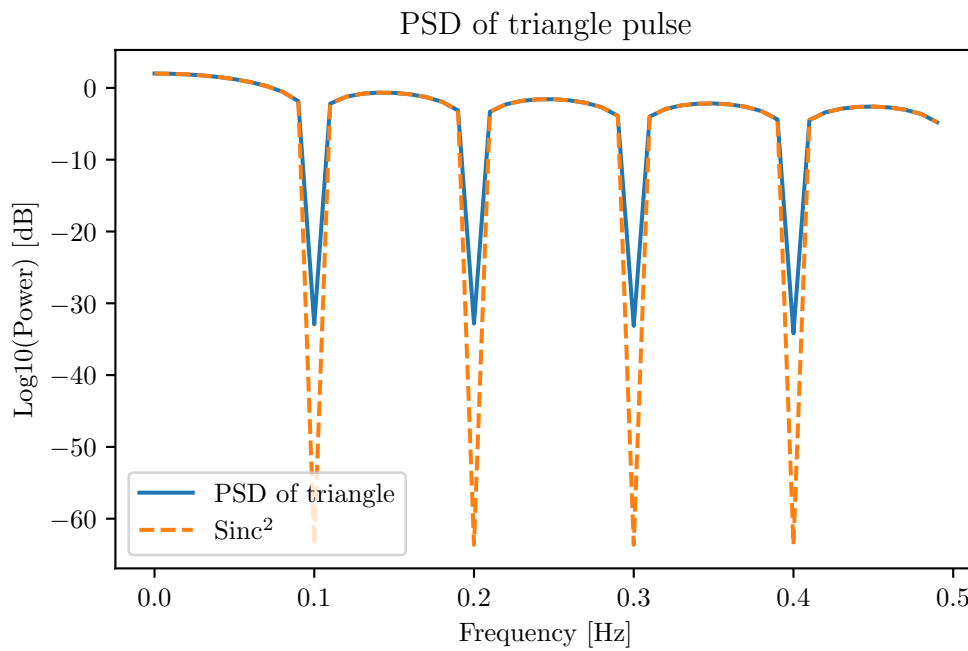
**Figure 8:** The Power Spectral Density of the signal in Figure 7. We see that the signal is composed of 3 frequencies; 0.04, 0.1 and 0.2 Hz. We also see that the amplitude of the 0.04 Hz component is 3 times larger than the other 2.

spatial and temporal evolution of large-scale ionospheric structures. These act more like triangles than sine waves, with increasing and decreasing density gradients. In order to get a better understanding of what the frequency domain of these large-scale ionospheric structures may look like, we inspect the spectral properties of triangle pulses.

In Figure 9, we see a triangle pulse with a width of 20 seconds and an amplitude of 1. Unlike sine waves, triangle pulses do not have a distinct peak in the frequency spectrum. Instead, the energy is divided among all frequencies, with the shape of a  $\text{sinc}^2$  function. We call the peaks of such a function "lobes", with the first peak being named the "mainlobe" and the other peaks being named "sidelobes". The width of the mainlobe is inverse proportional with the width of the triangle. In Figure 10 we see the PSD of the signal in Figure 9, along with a  $\text{sinc}^2$  function. The mainlobe carries the majority of the power of the signal, and ends at a frequency of  $f = 2/L = 0.1$  Hz, with  $L$  being the width of the triangle.

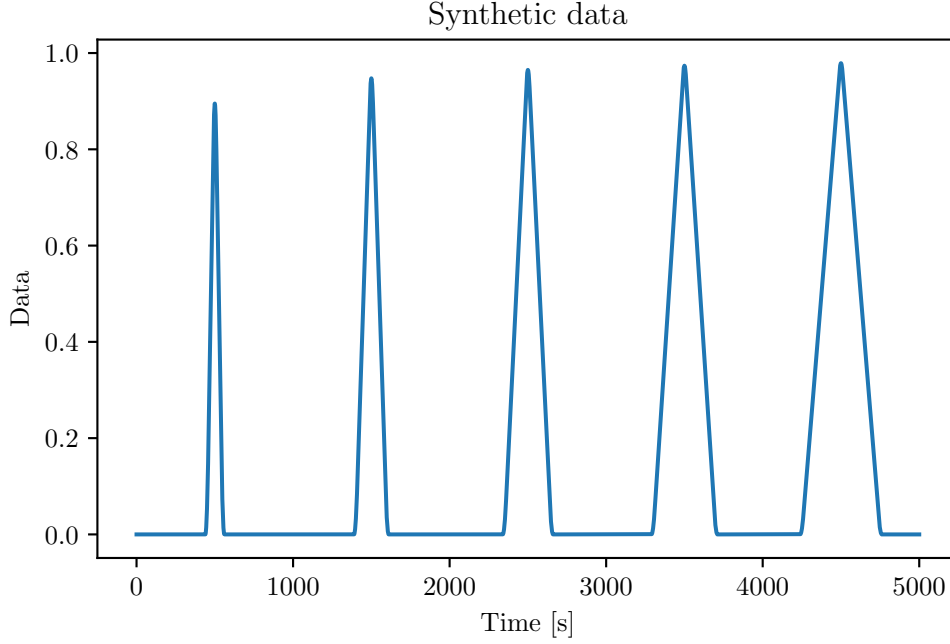


**Figure 9:** A triangle pulse signal.



**Figure 10:** The Fourier transform of a triangle pulse signal.

Unlike with sine waves, it becomes difficult to tell exactly what is going on by looking at the frequency spectrum of triangle pulses. Still, we can gain some insight by comparing multiple triangle pulses of different width. In Figure 11, we see a signal composed of 5 triangle pulses of different width, sampled over 5000 seconds. The first triangle pulse has a width of 100 seconds, with every following being 100 seconds wider than the last.

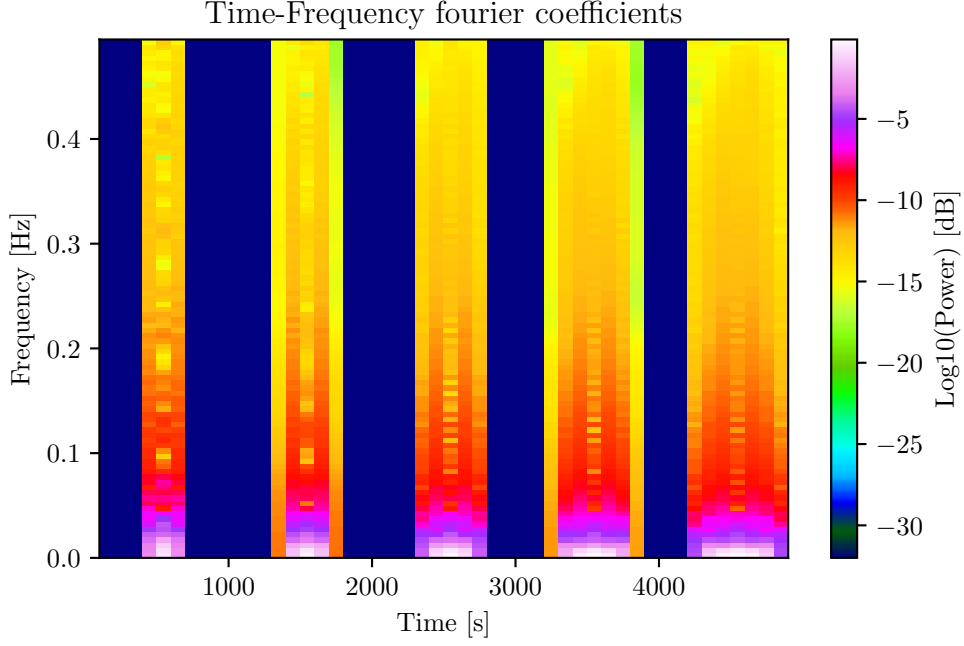


**Figure 11:** Multiple triangle pulse signals, each with a different width.

We would like to easily compare the triangles in Figure 11. One way of doing this is by dividing the signal into different time windows, calculating the PSD of each window, and plotting the PSD as functions of both frequency and time in a contour plot. In Figure 12, we see the time-frequency contour plot of the logarithmical PSD of the signal in Figure 11. A colormap with high contrast is used in order to more easily see the width of the mainlobe. We see that the thinner triangles have more power in the higher frequency bands.

From Figure 12, we see that the thinner triangles have more power in the higher frequency bands than the wider triangles. This is in accordance with the well known characteristic of the Fourier transform, that features that are sharp in the time domain are wide in the frequency domain. Consequently, this gives rise to the intuition that smaller structures will have higher frequency components than larger ones. This makes sense if one treats the time it takes to measure the structure as the structures period  $T$ . Then the frequency,  $f = \frac{1}{T}$  becomes larger as the structure becomes smaller. This is a useful characteristic if we are to for example inspect whether a structure has grown or shrunk, as we can look at whether the power of the PSD has moved into higher frequency bands.

Comparing structures such as the triangle pulses in Figure 11 would be easier to do if we could look at a single number, rather than having to visually compare their frequency spectra. As the power of the structures we are interested in are divided amongst the whole frequency spectrum, rather than being focused on a single frequency as is the case of sine waves, we will have to look at multiple frequency bands at the same time. This can be done by summing over all the Fourier coefficients between two frequencies. If we also normalize this sum by multiplying with the



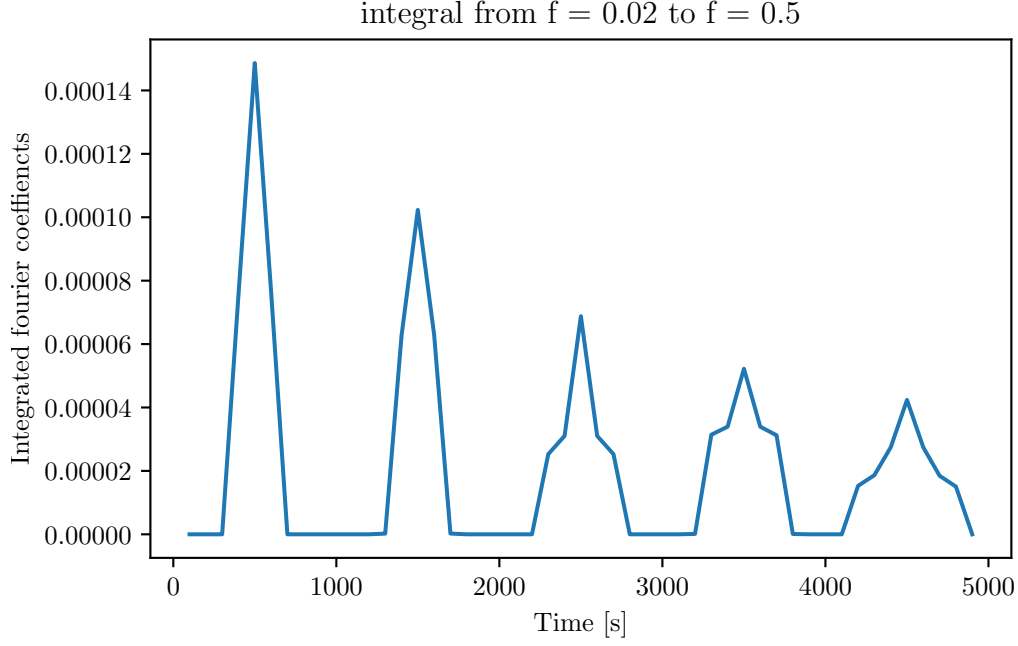
**Figure 12:** The frequency spectrum of the multiple triangle signal. We see here that the thinner the triangles are, the more power is in the higher frequency bands.

step length of the frequency  $\Delta f = f_s/N$  where  $f_s$  is the sampling frequency and  $N$  is the number of data points in the signal, we end up with a number equalling the numerical integral of the absolute value of Fourier coefficients between two frequencies:

$$C = \sum_{f=f_0}^{f_1} |c_f| \Delta f \quad (2)$$

Applying equation 2 to the Fourier transform shown by the PSD in Figure 12, we get Figure 13 where we see the integrated Fourier transform of each time window, integrated from the end of the largest mainlobe  $2/100 = 0.02$  Hz to the Nyquist frequency, half the sampling frequency,  $f_s/2 = 0.5$  Hz. We clearly see that the wider triangles have less total power in the higher frequency bands. We limit ourselves to frequencies past the widest mainlobe as this seems to give the clearest width-dependant difference in integrated PSD. If a mainlobe was included, it would dominate the integral and outshadow the power content in the higher frequency bands.

From Figure 13 it now becomes easier to see that the smaller triangles are more pronounced in higher frequency bands. However, even after doing this we are still left with the problem of having to compare two numbers. When looking at ionospheric structures, it would be easier to compare large amounts of data if we had a single number that told us the difference in integrated power between two signals. As such, we introduce the comparison index; the relative difference



**Figure 13:** The integral of the PSD between frequencies 0.02 Hz and 0.5 Hz. These frequencies are just above the main lobe of the PSD for the wider triangles. We see that as the triangles get wider, the integral becomes smaller in this frequency band.

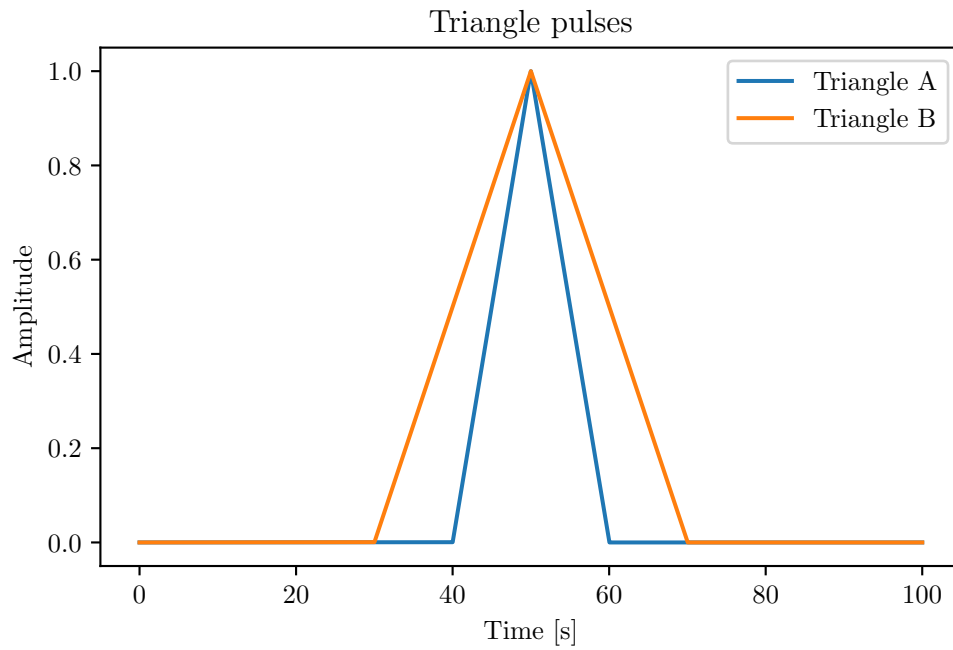
between two integrated Fourier transforms of signals A and B.

$$I_{A-B} = \frac{C_A - C_B}{\max\{C_A, C_B\}} \quad (3)$$

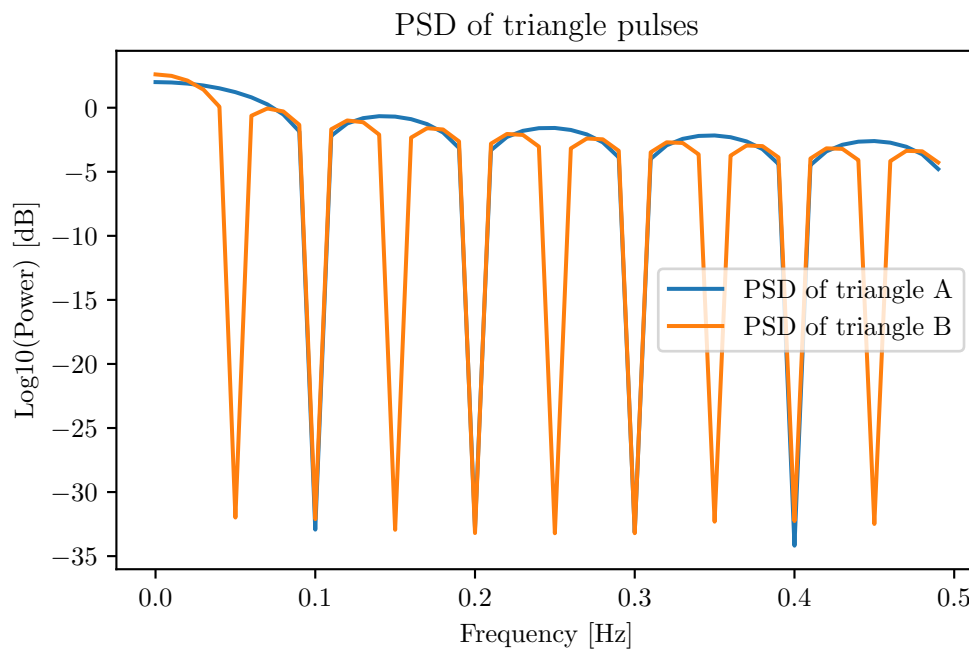
This index  $I_{A-B}$  lets us compare the activity of 2 signals within a frequency band.  $I_{A-B}$  is a number between -1 and 1. If one of the signals has a lot of activity in the frequency band we are looking at while the other one doesn't, the index  $I_{A-B}$  goes towards  $\pm 1$ ; negative if A is wider than B and positive if A is thinner than B. If the signals are quite similar,  $I_{A-B}$  will tend towards 0.

In order to further our understanding of the comparison index, we can look at an example of two triangle pulse signals with different width and their accompanying PSD. In Figure 14, we see two triangle pulse signals. Triangle A, plotted in blue, has a width of 20 seconds. Triangle B, plotted in orange, has a width of 40 seconds. In Figure 15, we see the PSD of the triangles in Figure 14. We see that the mainlobe of triangle B is shorter than the mainlobe of triangle A. Consequently, triangle A has more power in the higher frequency bands than triangle B. Calculating the comparison index  $I_{A-B}$  for these triangles, we find it to be approximately 0.5. This seems to be the case for all lower integral limits, as long as we do not include the mainlobes while holding the upper integral limit at the Nyquist frequency  $f_s/2 = 0.5$  Hz. If the lower integral limit is set to zero, the comparison index becomes approximately equal to the difference in mean value between the signals, in accordance with Parseval's theorem.

We can inspect the comparison index further by calculating comparison indices



**Figure 14:** Two triangle pulses, one is half the width of the other.



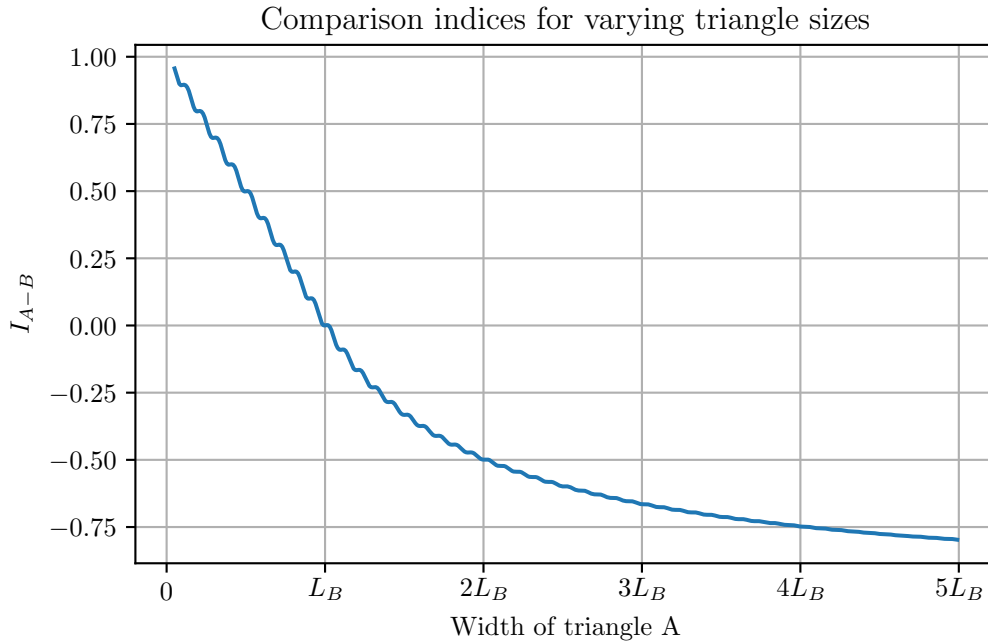
**Figure 15:** The PSD of the two triangle pulses from Figure 14.

for a multitude of cases by holding the width of one triangle stationary and changing the other.

In Figure 16, we see how the comparison index depends on the difference in 2 triangle pulse signals. The first thing we note is that the result is divided into a negative and a positive part, with zero where the signals are identical (width of A equals width of B). On the left side, the comparison index is positive. When triangle A is thinner we have a signal that has more power in the higher frequency bands. The integral of the Fourier transform then becomes larger, dominating equation 3. Consequently, we can say that a positive comparison index  $I_{A-B}$  means that the structure in signal A is thinner than the structure in signal B.

Around the width of 1, the comparison index is small, being zero at 1 exactly. Here, the structures in the signals are of about equal width, having about the same power in the higher frequency bands. The integrals then become about the same in size, giving us a small comparison index.

Beyond the width of 1, the comparison index becomes negative. Here, the structure in signal A is larger than the structure in signal B, having less power in the higher frequency bands. Between 0 and 1, there is a linear relationship between the comparison index and the difference in triangle widths on the form  $1 - \frac{\text{width}_A}{\text{width}_B}$ . After 1, the relationship goes as  $\frac{\text{width}_B}{\text{width}_A} - 1$ , approaching -1 as the width of triangle A becomes infinitely large.



**Figure 16:** Comparison indices for differing triangle widths. The width of triangle A is changed, while B is kept stationary. The comparison index of  $I_{A-B}$  is calculated. The width of triangle A is given as multiples of the width  $L_B$  of triangle B.

Looking at a given comparison index  $I_{A-B}$ , we can infer some of the properties



of  $A$  and  $B$  from it. If  $I_{A-B} < 0$ , the structure in signal  $B$  must be thinner than the structure in signal  $A$ . If  $I_{A-B} > 0$ , the opposite is true. Due to the asymptotic behavior of the comparison index when the structure in one signal is very wide, it becomes difficult to infer size comparisons between the structures based on the size of the comparison index. A large comparison index could mean that one structure is very wide, or it could mean that one is very thin. This is a bi-product of the normalization in equation 3. When used to look at changes in a structure, we can still infer that a large comparison index means that a signal has become a lot wider, while a large negative index means that a signal has become a lot thinner. This of course assumes that the structures we are looking at behave like triangle pulse signals. If this is not the case, then a large comparison index simply means that one signal has a lot of power in the frequency band we are looking at compared to the other.

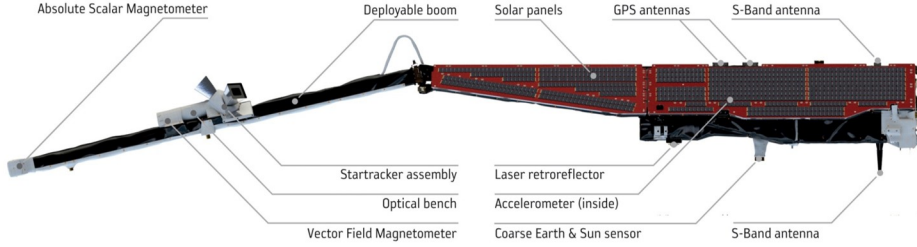
### 3 Instrumentation and Method

In this section we describe the instrumentation and methods used in this thesis. We begin by introducing the Swarm satellites from which we obtain electron density data. We then introduce the methods used when obtaining the results shown in the results section. We first introduce the general methods used to index-shift and deal with data gaps, then proceed to explain specifically what we did in order to obtain our different results.

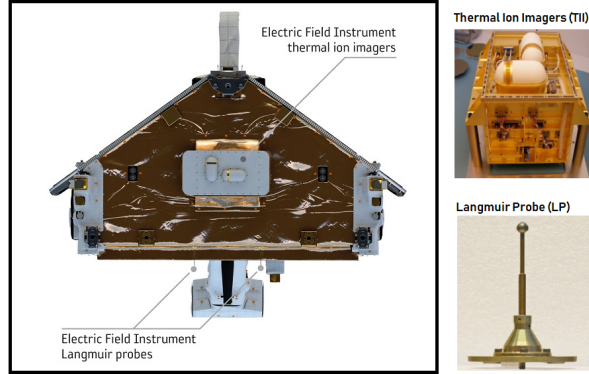
#### 3.1 The Swarm Mission

In this paper, we look at electron density and positional data from the European Space Agency (ESA)'s Swarm mission.

On 22 November 2013, ESA launched the Swarm mission. Swarm consists of 3 identical satellites (named A, B and C), carrying an array of instruments intended for creating a highly detailed survey of Earth's geomagnetic field [24] [3].



**Figure 17:** Side view of Swarm satellite. Source: ESA [3]



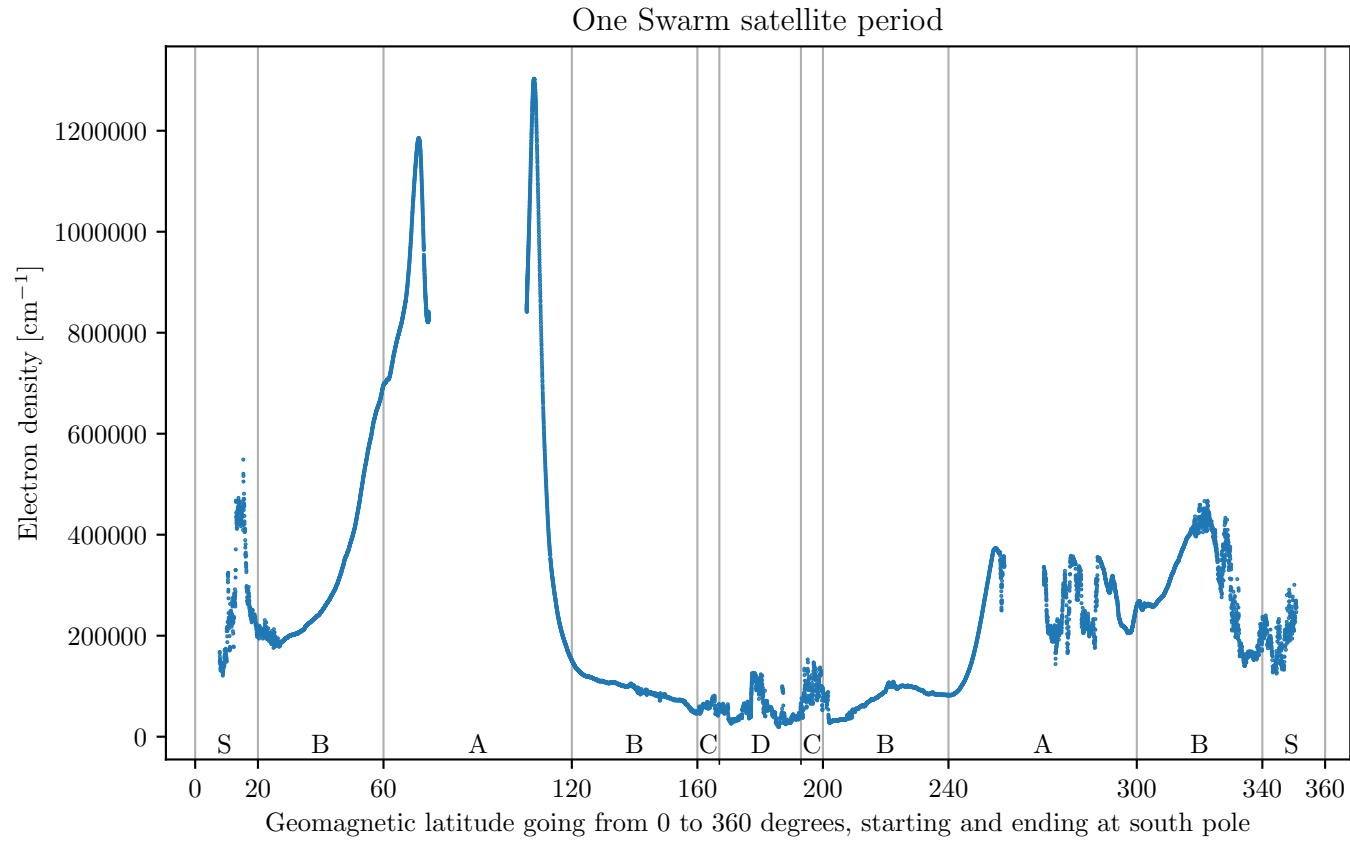
**Figure 18:** Rear view of Swarm satellite. Source: ESA [3]

Part of the instruments carried by the Swarm satellites is the Electric Field Instrument (EFI), as seen in Figure 18 [3]. The EFI includes Langmuir probes which measure electron density, electron temperature and the spacecraft potential. In our case, we are interested in the electron density measurements from the Langmuir probes in conjunction with positional measurements from the GPS. The electron density is measured in electrons per cubed centimeter [ $\text{cm}^{-3}$ ] while the positionals use spherical geocentric coordinates in the International Terrestrial Reference Frame (ITRF). Both are synchronized at a sampling frequency of 2 Hz.

During December 2013, the Swarm satellites orbited Earth in a cross polar pearls-on-a-string configuration, following approximately the same line of longitude. On the 11th of December, the satellites were flying at an altitude of 507 km, with an orbital velocity of 7600 m/s. On that day, the satellites flew in a configuration with satellite B in the lead, followed 28.5 seconds later by satellite A, followed in turn 51 seconds later by satellite C. This is equivalent to a distance of 215 km and 385 km respectively.

As time passed, the satellites drifted apart. On the 31st of December, the time difference between the satellites was 60.5 seconds and 110 seconds. This gives an average drift of about 12000 meters per day between the satellites B and A, and 22000 meters per day between satellites B and C.

In Figure 19, we can see the electron density measured by Swarm A during one orbital period, divided by geomagnetic latitudinal regions.



**Figure 19:** One period of a Swarm satellite. The x axis is geomagnetic latitude, going from 0 to 360, starting and ending at the south pole. The plot is divided into regions, with A noting low latitude, B noting mid latitude, C noting the approximate open-closed field-line boundary, D noting the north polar cap and S noting the south pole.

### 3.2 Index-shifting

The three Swarm satellites, Swarm A, B and C, measure electron density  $N_e$  and geographic positions  $\vec{x}$  twice every second. This results in 3 time series for electron density, one for each satellite A, B and C;  $N_{eA}[t]$ ,  $N_{eB}[t]$  and  $N_{eC}[t]$ , as well as 3 time series for position;  $\vec{x}_A[t]$ ,  $\vec{x}_B[t]$  and  $\vec{x}_C[t]$  where  $t$  is the time of samplings, and is on the form  $[t_0, t_0 + 0.5, t_0 + 1 \dots]$  with  $t_0$  being the time of the initial sampling. Our goal is to inspect the evolution of structures in the electron density time series, and investigate what we can tell about the temporal and spatial nature of changes in these.

In their initial form, the time series are synchronized in time, such that one data point  $N_{eA}[t_i]$  is sampled at the same time  $t_i$  as in the other time series  $N_{eB}[t_i]$  and  $N_{eC}[t_i]$ . Instead of looking at equal *times*, we want to look at equal *positions*, so that we can evaluate the evolution of electron density in one location at a time. Since the satellites are moving in a pearls-on-a-string configuration (with satellite B in front, followed by A then C), this can be done by **index-shifting** the time series for satellite A and C forward, so as to minimize the distance between data points  $|\vec{x}_B - \vec{x}_A|$  and  $|\vec{x}_B - \vec{x}_C|$ .

In Figure 20, we see an example of electron densities measured by the 3 Swarm satellites on the 21st of December 2013. Here, the electron density measured by satellite B is plotted in red, the electron density measured by satellite A is plotted in green and the electron density measured by satellite C is plotted in blue. We see that all 3 electron densities follow the same pattern, except for a time delay. This is most noticeable where the electron densities increase into the  $2 \cdot 10^5 \text{ cm}^{-3}$  range. This first happens to the electron densities measured by satellite B. The electron density measured by satellite A spikes in the same way around 45 seconds later, followed around 80 seconds later by satellite C. This delay is because of the distance between the satellites. They measure the same approximate density structure, but with some time difference  $\Delta t$ .

Index-shifting a time series  $\chi[t]$  can be done by discarding an appropriate amount of data points at the beginning of it. If we want to index-shift the time series  $\Delta t$  seconds forward, the number of data points to be discarded becomes  $\frac{\Delta t}{f_s}$ , where  $f_s$  is the sampling frequency of the time series. This is the same as adding a constant  $\Delta t$  to the time  $t$  of time series  $\chi[t]$ . This is shown in equation 4, where we make a new time series  $\chi'[t]$  which is the time series  $\chi[t]$  shifted forwards by a time difference  $\Delta t$ .

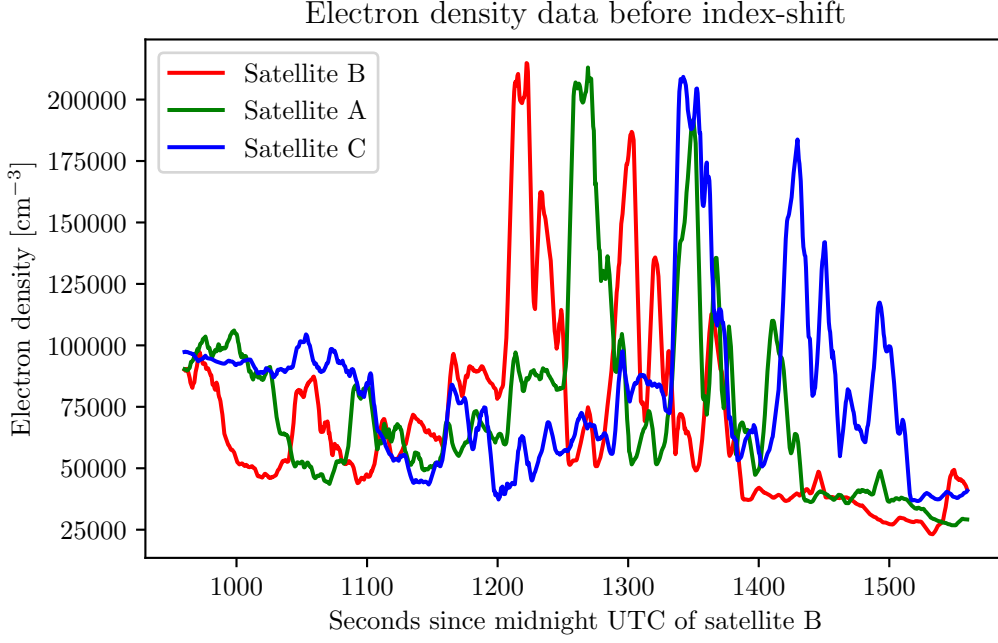
$$\chi'[t] = \chi[t + \Delta t] \quad (4)$$

We want to index-shift the time series for satellites A and C forward. Ideally, we could find the appropriate time differences  $\Delta t_{BA}$  and  $\Delta t_{BC}$  by solving equations 5 and 6.

$$\vec{x}_B[t] = \vec{x}_A[t] = \vec{x}_A[t + \Delta t_{BA}] \quad (5)$$

$$\vec{x}_B[t] = \vec{x}_C[t] = \vec{x}_C[t + \Delta t_{BC}] \quad (6)$$

It would be unreasonable to expect equations 5 and 6 to have solutions, as the satellites cannot follow the *exact* same orbit. There will never be two data points that have the exact same positions. Instead, for one day of data, we find the values

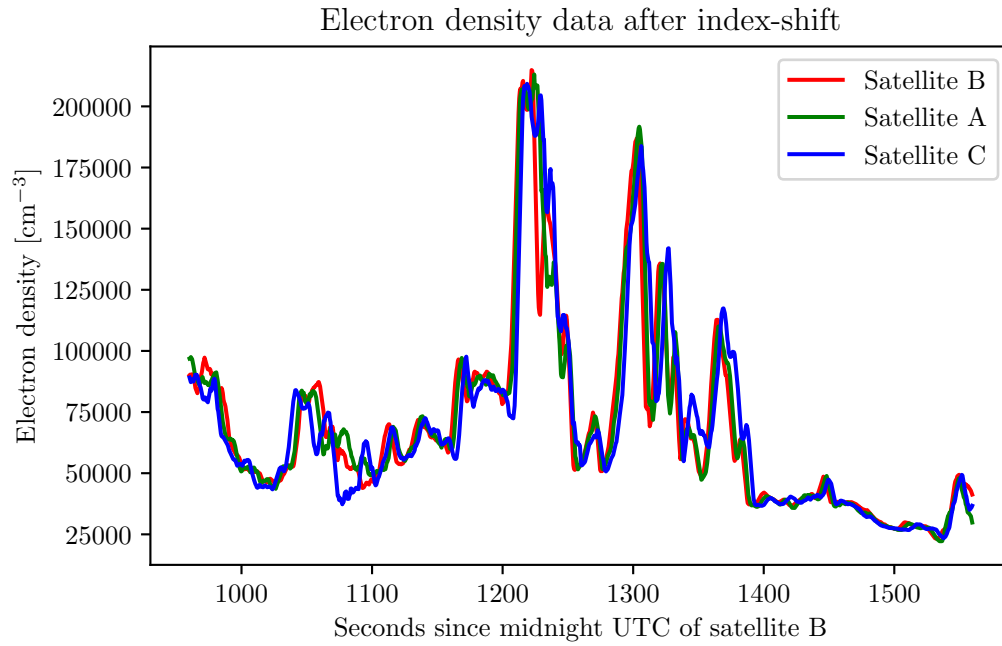


**Figure 20:** An example of electron densities measured by the 3 Swarm satellites for the 21st of December 2013. Red notes the electron density measured by satellite B, green notes the electron density measured by satellite A and blue notes the electron density measured by satellite C. We see a distinct pattern in the electron densities, which is first demonstrated by satellite B, before being repeated by first satellite A then satellite C.

for  $\Delta t_{BA}$  and  $\Delta t_{BC}$  that *minimizes* the distance between positions. Numerically, we do this by trying out many different values of  $\Delta t$ , calculating the average difference in latitude between data points for each value, then using the one that gives us the lowest distance. We can use latitude instead of actual distance since the satellites follow lines of longitude, making the difference in latitude the only meaningful index.

In Figure 21, we see the electron density time series of satellites A, B and C after index-shifting. The electron density measured by satellite B is plotted in red, the electron density measured by satellite A is plotted in green and the electron density measured by satellite C is plotted in blue. We clearly see that all 3 time series follow the same pattern with some variations, but without the time delay seen in Figure 20. In the x axis we have the time as measured by satellite B. We use the time of satellite B as we index-shift the time series of satellites A and C forward while holding B still. This time is akin to the time argument  $t$  as seen in equations 5 and 6. We could instead have used latitude, as the latitudes of the time series are now approximately equal.

Since the satellites slowly drift apart,  $\Delta t$  changes over time. Ideally, we would find  $\Delta t$  for every single data point. This is computationally exhaustive, so we settle with calculating  $\Delta t$  at the beginning of each day. This leads to an error of about 12000 m for satellites B-A and around 22000 m for satellites B-C, the average daily drift, at the end of each day.



**Figure 21:** An example of electron densities measured by the 3 Swarm satellites for the 21st of December 2013. Red notes the electron density measured by satellite B, green notes the electron density measured by satellite A and blue notes the electron density measured by satellite C.

### 3.3 Dealing with data gaps.

In order to compare the time series obtained by the Swarm satellites, we need them to be synchronized. Some data is missing, introducing gaps in the time series, ruining the synchronization. We can observe such a gap by spotting a leap in the UTC times of sampling. Usually, every sample is taken half a second after the previous, as the sampling frequency is 2 Hz. If the timegap between samples is larger than half a second, we have a hole in the data. If equivalent gaps are not present in all the time series we are comparing, a problem occurs.

If we are looking at the  $i$ 'th index of time series A and comparing it to the  $i$ 'th index of time series B, we must ensure that the times are synchronized so that  $t_{iA} = t_{iB}$ . If there are no gaps present in the data, this is always true. If a gap *is* present, this becomes untrue. If there is a gap at index  $j$  in the data set A, so that  $t_{jA} - t_{(j-1)A} > 0.5$ , we end up with the times of A being ahead of B after index  $j$ , so that  $t_{iA} - t_{iB} > 0$ ,  $i > j$ .

One simple solution is to make identical gaps in all the time series. Before index-shifting, we identify gaps in the time series by finding indices  $i$  where  $t_i - t_{i-1} > 0.5$ . We then fill in the gaps with zeros, making it so that the length of the time series becomes equal to what it would have been if no data was missing (172800; 2 times the number of seconds in a day). We can then find the indices that need to be deleted by searching for zeros. After doing this to all 3 time series A, B and C, we get a complete list of indices of all the times any data was missing. Then we go through all 3 time series and delete the elements with the index from our list.

The identification of gaps must be done by using the UTC time stamps recorded by the Swarm satellites. Once the list of indices to be deleted is retrieved, we go through the other data such as electron density and positionals and make equivalent deletions.

After synchronizing the data gaps, a problem arises when index-shifting. After index-shifting, the difference in timestamps should be constantly equal to the time-difference between the satellites. For satellites B and A,  $t_{iB} - t_{iA} = \Delta t_{BA}$ . This is not the case around the borders of data gaps. Here, we end up comparing timestamps before the gap in one time series with timestamps after the gap in the other. The difference in time becomes  $t_{iB} - t_{iA} = \Delta t_{BA} + L_t$  with  $L_t$  being the length of the gap, or the amount of time for which data was missing. To rectify this, we identify data points for which this is the case and delete them.

### 3.4 Histograms and linear regressions.

We could look at the Swarm data and calculate the comparison index (equation 3) between two samplings of an ionospheric structure. This would give us a number between -1 and 1, telling us how much the structure has changed between the two samplings. It would not, however, tell us *why* the structure changed. In particular, we would not be able to differentiate between a spatial or a temporal change. Is the comparison index a result of us measuring different places in the structure, or has the structure itself been altered somehow?

Instead of looking at individual comparison indices, we can take a statistical approach. We start out by producing histograms of comparison indices. We make one set of histograms per day, and divide each day of data into regions by geomagnetic latitude through altitude-adjusted corrected geomagnetic coordinates [10] by the use of the `aacgm2` library in Python. We start off with the equatorial region, going from  $-30^\circ$  magnetic latitude (MLAT) to  $30^\circ$  MLAT. We then have the mid-latitude regions ranging from  $\pm 30^\circ$  MLAT to  $\pm 70^\circ$  MLAT, choosing  $\pm 70^\circ$  as models have shown this to be a typical latitude for the dayside auroral oval boundary [8]. From  $70^\circ$  to  $77^\circ$  MLAT we have the region where we expect the auroral oval to be. We choose the limit of  $77^\circ$  in accordance with the Ionospheric Plasma IRregularities product [9], where the polar cap region is assumed to be poleward of  $77^\circ$  MLAT when the auroral oval is not well defined. The area  $> 77^\circ$  MLAT is assumed to be inside the northern polar cap.

After dividing the data by latitude, we split the data into windows of length 120, equaling one minute of data, with a 50% overlap. We then discard windows where the data is discontinuous (places where we made the split by latitude, or where there are data gaps). We then go through each window of data individually. First we calculate a rolling mean in order to minimize local variations, letting each data point be the average of itself and 5 neighbours in both directions. This effectively works as a lowpass filter, removing high frequency noise and letting us better inspect the evolution of the large-scale phenomena through their higher frequency components. After this, we calculate the comparison index for each window of data. This gives us three comparison indices for each window, one for each satellite pair  $I_{BA}$ ,  $I_{BC}$  and  $I_{AC}$ . We produce an example of one day of computed comparison indices, plotted by giving each data point a transparency based off the integrated fourier coefficients, shown in the results section. Finally, we combine the comparison indices in



histograms. This results in twelve histograms for each day of data, three histograms from the three combinations of satellites for each of the four latitudinal regions. We produce an example of one such histogram for the polar cap region along with a gaussian approximation for the 15th of December 2013, shown in the results section.

Each histogram produced has an accompanying time-difference  $\Delta t$  which we find while index-shifting. This time difference increases with every day. This gives us the opportunity to investigate how the histograms, and therefore the changes in structures, develop over time. We do this by plotting the four statistical moments mean, standard deviation, skewness and kurtosis of the comparison indices for each day, satellite pair and region as functions of the time difference  $\Delta t$ , and making linear regressions.

### 3.5 Assumptions and satellite drift

While doing our analysis, we assume that the data points have the same spatial coordinates after index-shifting. In order to check how good this assumption is, we inspect the distance between satellite pairs, and see what part of the distance is latitudinal or longitudinal.

As the satellites are moving in circles around the Earth, we use great circle distance to find a more realistic distance between the satellites, compared to drawing a straight line between them. The Earth is not perfectly spherical, but at short distances it is a good assumption. We calculate the great circle distance from latitudes  $\phi_1$  and  $\phi_2$ , and longitudes  $\theta_1$  and  $\theta_2$  at a distance from Earth's center  $R$  by:

$$\Delta x = R \arctan\left(\frac{\sqrt{(\cos \phi_2 \sin \Delta\theta)^2 + (\cos \phi_1 \sin \phi_2 - \sin \phi_1 \cos \phi_2 \cos \Delta\theta)^2}}{\sin \phi_1 \sin \phi_2 + \cos \phi_1 \cos \phi_2 \cos \Delta\theta}\right) \quad (7)$$

Where  $\Delta\theta = |\theta_1 - \theta_2|$ .

We first look at distances for a full day of data near the beginning of our dataset, then repeat the exercise for a day near the end of December 2013. For the first day, we choose the 10th of December as it is the first day in the data set without large data gaps. We index-shift the data and convert the UTC timestamps from date-times into seconds since midnight. We then calculate the great circle distance for all 3 satellite pairs,  $B - A$ ,  $B - C$  and  $A - C$ . This distance is then plotted first against time since midnight to give a view of how the distance develops over time, then against geographical latitude to give a view of how the distance is dependent on geographical position. Repeating this exercise, we choose the 31st of December. We again produce plots of great circle distance over time and geographical latitude, shown in the results section.

In addition to looking at how the distance between satellites evolve over the course of a day, we can look at how the mean distance per day changes as the time difference between satellites change. We expect the satellites to drift apart over time, and are interested in seeing how much.

We do this by calculating the mean great circle distance between each satellite pair for each day, and plot it against the time difference between the satellite pairs, as found when index-shifting. We do this once for all latitudes and once by limiting

ourselves to latitudes  $> 77^\circ$ . We then do a linear regression to see how correlated the great circle distances and the time differences are.

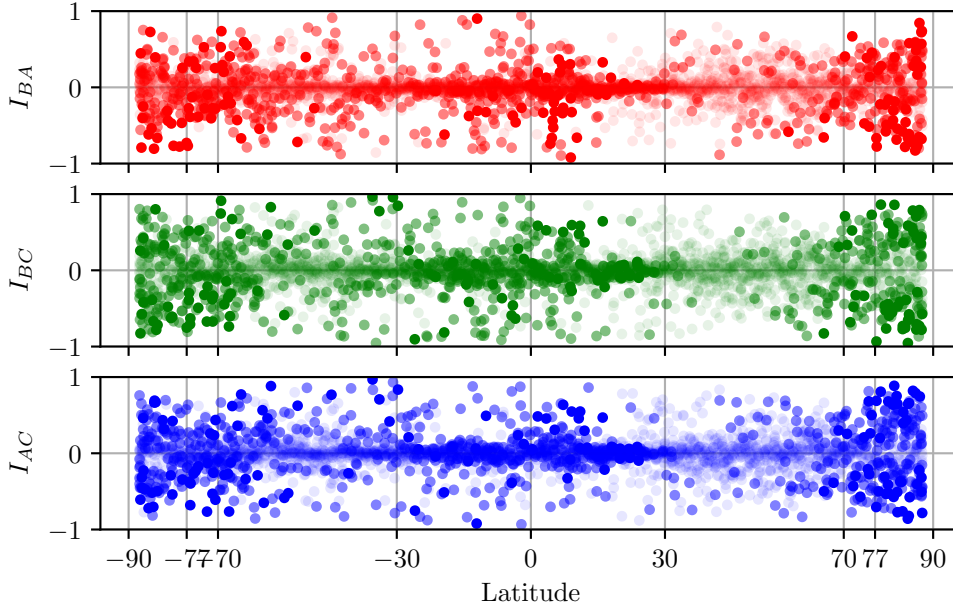
Assuming they are orbiting at the same altitude, the distance between satellites can be divided into a difference in latitude and a difference in longitude. The longitudinal distance is dependant on the geographical latitude, while the latitudinal distance is only dependant on the difference in latitude. We can easily find the latitudinal distance between satellites by looking at the difference in latitude between them. At the altitude of 507 km as the Swarm satellites are orbiting at, the orbital circumference is about 43200 km. Dividing this by 360, we see that one degree of latitude equals approximately 120 km. As such, we can find longitudinal distance by subtracting the difference in latitude times 120 km from the total satellite distance. In order to do this, we create plots of average absolute difference in latitude between each satellite pair per day against the time difference between each satellite pair per day. We do this first for all latitudes, then for latitudes  $> 77^\circ$ .

## 4 Results

### 4.1 Comparison indices and statistical moments

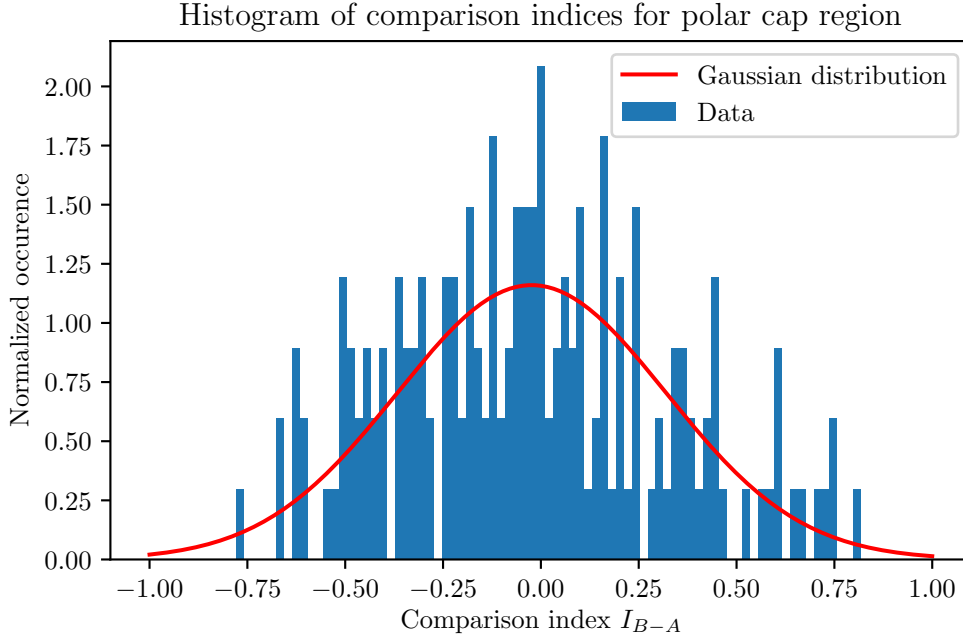
In this section we present computed comparison indices for all 3 satellite pairings B-A, B-C and A-C, for the 9th through 31st of December 2013. This was done using window sizes of one minute (or 120 indices), and integral limits in equation 2 from 0.1 to 1 Hz. We first include an example of computed comparison indices for one day of data, plotted by latitude. We then show a histogram of the comparison indices in the northern polar cap region. Following this, we show the 4 statistical moments of the distribution of comparison indices for each latitudinal region, plotted against the time difference  $\Delta t$  as found when index-shifting. We also include linear fits in these plots in order to see how well changes in the moments correlate with change in time, with the regression coefficients included in tables.

In Figure 22 we see comparison indices computed for the 15th of December 2013. Each data point plotted has been given a transparency dependant on the normalizing factor in equation 3. The 95th percentile and above have a transparency of 1, those between the 95th and 68th percentile have a transparency of 0.5 and the rest have a transparency of 0.1. The unit of the x axis is geographic latitude, and is not fully in line with the geomagnetic latitude shown in Figure 19. We see from the figure that the higher latitude regions have more rapid changes in structures with higher integrated power spectrum density when compared to lower latitudes.



**Figure 22:** Comparison indices for the 15th of December 2013. Each data point is given a transparency depending on the size of the normalizing factor in equation 3. The 95th percentile and above have a transparency of 1, those between the 95th and 68th percentile have a transparency of 0.5 and the rest have a transparency of 0.1.

In Figure 23, we see a histogram of comparison indices for the northern polar cap region on the 15th of December 2013, along with a gaussian bell curve based on the mean and standard deviation of the histogram. The data in the histogram is divided between 100 bins, with values ranging from -1 to 1. We see that the histogram follows the gaussian distribution decently well, with a mean of -0.006 and a standard deviation of 0.21. We see that in this instance, the lowest comparison index has a value of around -0.6, indicating one event where the structure measured thinned considerably, while the highest value was around 0.5, indicating a considerable increase in structure size.



**Figure 23:** Histogram of comparison indices in the northern polar cap region for the 15th of December 2013, along with a gaussian bell curve based on the mean and standard deviation of the histogram. The mean of the histogram is -0.006, and the standard deviation is 0.21.

In Figure 24, we see the mean of the comparison indices computed for all 4 latitudinal regions plotted against the time difference between the satellites, as found when index-shifting. Each data point is the mean of one day of data for one satellite pair. The black dots indicate data points, while the blue line is a linear regression of the data points as a function of the time difference.

In (A), we see the results for the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. Here, the means are between around -0.05 and 0.05.

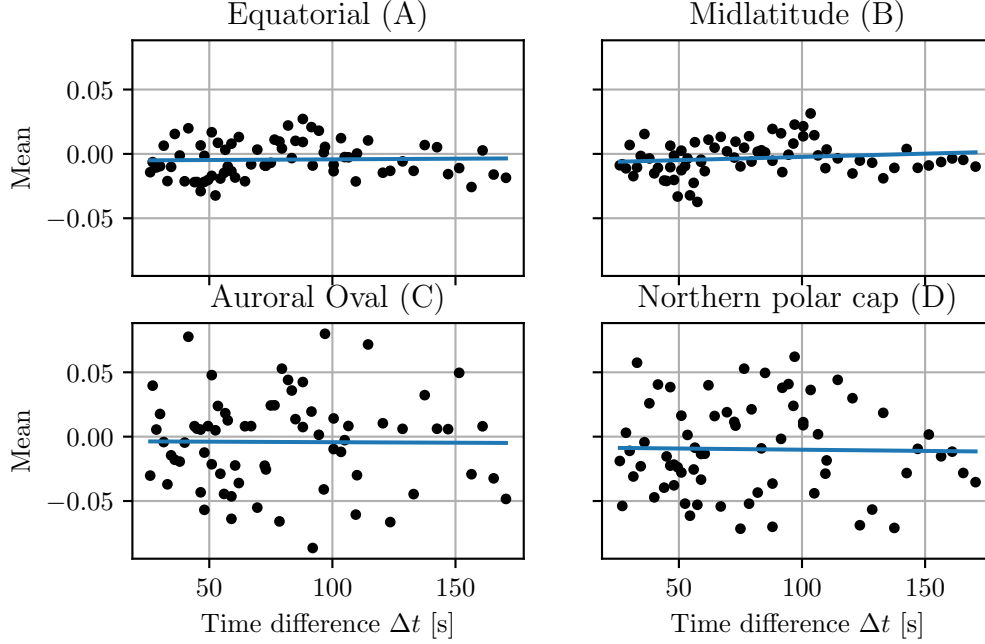
In (B) we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. Here, the means are between around -0.05 and 0.05. They start out negative, turn positive in the middle, then flatten out around zero towards the end.

In (C) we have the results for the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. Here, the means are spread out somewhat evenly between -0.1 and 0.1.

In (D), we have the results for the northern polar cap region, between  $77^\circ$  and  $90^\circ$  magnetic latitude. Here, the means are spread somewhat evenly between around

-0.09 and 0.08, with a slight negative tendency towards the end.

In Table 1, we see the regression coefficients for the linear regressions in Figure 24. In the first column we have the latitudinal regions, in the second column we have the regression slopes and in the third column we have the regression constants.



**Figure 24:** Mean of comparison indices plotted against time difference between satellite pairs. Black dots indicate data points, while the blue lines indicate linear regression. One black dot is the mean of one day of data for one satellite pair. In (A), we have the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. In (B), we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. In (C) we have the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. In (D), we have the northern polar cap region, ranging from  $77^\circ$  to  $90^\circ$  magnetic latitude.

**Table 1:** Table of regression coefficients for the linear regressions found in Figure 24. The first column has the latitudinal regions, the second column has the regression slopes and the third column has the regression constants.

Region	Mean regression slope $[\text{s}^{-1}]$	Mean regression constant $[\text{s}^{-1}]$
Equatorial (A)	$9.10 \cdot 10^{-6} \pm 4.4 \cdot 10^{-5}$	$-5.2 \cdot 10^{-3} \pm 3.9 \cdot 10^{-3}$
Midlatitude (B)	$5.0 \cdot 10^{-5} \pm 4.2 \cdot 10^{-5}$	$-7.4 \cdot 10^{-3} \pm 3.7 \cdot 10^{-3}$
Auroral Oval (C)	$-8.1 \cdot 10^{-6} \pm 1.1 \cdot 10^{-4}$	$-3.5 \cdot 10^{-3} \pm 1.0 \cdot 10^{-2}$
Northern polar cap (D)	$-1.9 \cdot 10^{-5} \pm 1.1 \cdot 10^{-4}$	$-8.3 \cdot 10^{-3} \pm 9.9 \cdot 10^{-3}$

In Figure 25, we see the standard deviation of the comparison indices computed for all 4 latitudinal regions plotted against the time difference between the satellites, as found when index-shifting. Each data point is the standard deviation of one day of data for one satellite pair. The black dots indicate data points, while the blue

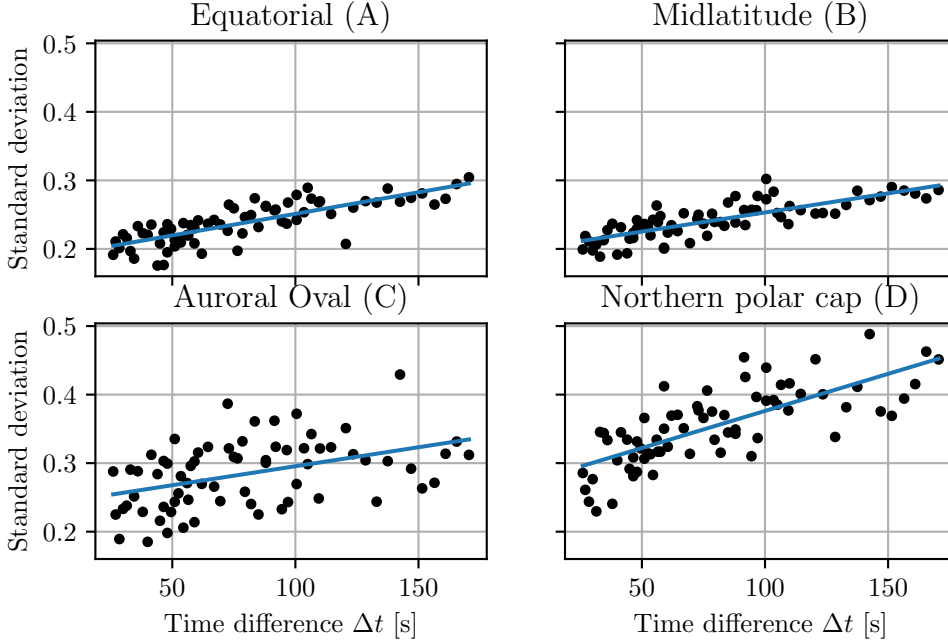
line is a linear regression of the data points as a function of the time difference. In (A), we see the results for the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. Here, the standard deviations start out between 0.15 and 0.25, and slope towards 0.3 over time.

In (B) we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. Here, the standard deviations start out between 0.15 and 0.25, and slope towards 0.3 over time.

In (C) we have the results for the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. Here, the standard deviations start out between 0.2 and 0.3 and increase slightly over time, flattening out at around 0.3 towards the end.

In (D), we have the results for the northern polar cap region, between  $77^\circ$  and  $90^\circ$  magnetic latitude. Here, standard deviations start out between around 0.2 and 0.35 and increase over time, ending up between 0.4 and 0.5 towards the end.

In Table 2, we see the regression coefficients for the linear regressions in Figure 25. In the first column we have the latitudinal regions, in the second column we have the regression slopes and in the third column we have the regression constants.



**Figure 25:** Standard deviation of comparison indices plotted against time difference between satellite pairs. Black dots indicate data points, while the blue lines indicate linear regression. One black dot is the standard deviation of one day of data for one satellite pair. In (A), we have the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. In (B), we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. In (C) we have the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. In (D), we have the northern polar cap region, ranging from  $77^\circ$  to  $90^\circ$  magnetic latitude.

In Figure 26, we see the skewness of the comparison indices computed for all 4 latitudinal regions plotted against the time difference between the satellite pairs, as found then index-shifting. Each data point is the skewness of one day of data for one satellite pair. The black dots indicate data points, while the blue lines are linear regressions of the data points as functions of the time difference.

**Table 2:** Table of regression coefficients for the linear regressions found in Figure 25. The first column has the latitudinal regions, the second column has the regression slopes and the third column has the regression constants.

Region	Std regression slope [s <sup>-1</sup> ]	Std regression constant [s <sup>-1</sup> ]
Equatorial (A)	$6.3 \cdot 10^{-4} \pm 5.8 \cdot 10^{-5}$	$1.9 \cdot 10^{-1} \pm 5.2 \cdot 10^{-3}$
Midlatitude (B)	$5.6 \cdot 10^{-4} \pm 5.2 \cdot 10^{-5}$	$1.10 \cdot 10^{-1} \pm 4.6 \cdot 10^{-3}$
Auroral Oval (C)	$5.5 \cdot 10^{-4} \pm 1.4 \cdot 10^{-4}$	$2.4 \cdot 10^{-1} \pm 1.2 \cdot 10^{-2}$
Northern polar cap (D)	$1.1 \cdot 10^{-3} \pm 1.2 \cdot 10^{-4}$	$2.7 \cdot 10^{-1} \pm 1.0 \cdot 10^{-2}$

In (A) we see the results for the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. Here, the skewness ranges from -0.5 to 0.5, spread evenly.

In (B), we see the results for the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. Here, the skewness ranges from around -0.5 to 0.5, sloping slightly towards the positive over time.

In (C) we have the results for the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. Here the skewness is somewhat evenly divided between -0.5 and 0.5, sloping slightly towards the positive.

In (D), we have the results for the northern polar cap region, between  $77^\circ$  and  $90^\circ$  magnetic latitude. Here, the skewness ranges from about -0.5 to 0.5, evenly divided around 0.

In Table 3, we see the regression coefficients for the linear regressions in Figure 26. In the first column we have the latitudinal regions, in the second column we have the regression slopes and in the third column we have the regression constants.

**Table 3:** Table of regression coefficients for the linear regressions found in Figure 26. The first column has the latitudinal regions, the second column has the regression slopes and the third column has the regression constants.

Region	Skewness regression slope [s <sup>-1</sup> ]	Skewness regression constant [s <sup>-1</sup> ]
Equatorial (A)	$9.8 \cdot 10^{-4} \pm 8.1 \cdot 10^{-4}$	$-8.9 \cdot 10^{-3} \pm 7.2 \cdot 10^{-2}$
Midlatitude (B)	$1.2 \cdot 10^{-3} \pm 4.9 \cdot 10^{-4}$	$-7.1 \cdot 10^{-2} \pm 4.3 \cdot 10^{-2}$
Auroral Oval (C)	$8.4 \cdot 10^{-4} \pm 7.2 \cdot 10^{-4}$	$-4.7 \cdot 10^{-2} \pm 6.4 \cdot 10^{-2}$
Northern polar cap (D)	$2.1 \cdot 10^{-4} \pm 4.6 \cdot 10^{-4}$	$-6.5 \cdot 10^{-3} \pm 4.1 \cdot 10^{-2}$

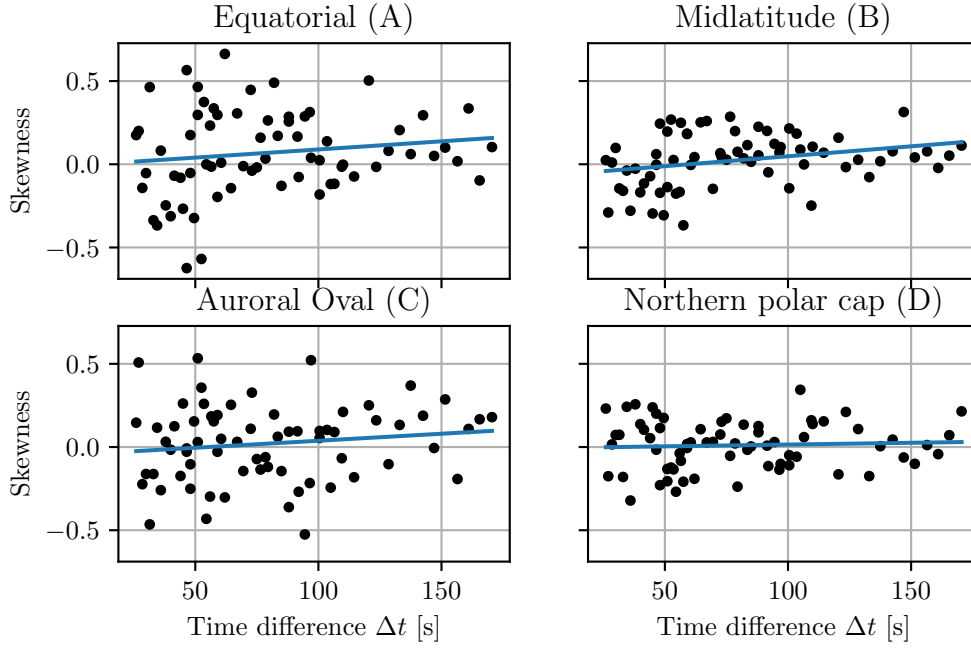
In Figure 27, we see the kurtosis of the comparison indices computed for all 4 latitudinal regions plotted against the time difference between the satellite pairs, as found then index-shifting. Each data point is the kurtosis of one day of data for one satellite pair. The black dots indicate data points, while the blue lines are linear regressions of the data points as functions of the time difference.

In (A) we see the results for the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. Here, the kurtosis ranges from around 2 to 5, decreasing over time.

In (B), we see the results for the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. Here, the kurtosis ranges from 0 to 4, decreasing over time.

In (C) we have the results for the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. Here, the kurtosis ranges from around 0 to 4 with a higher spread at the beginning, sloping slightly towards the negative.

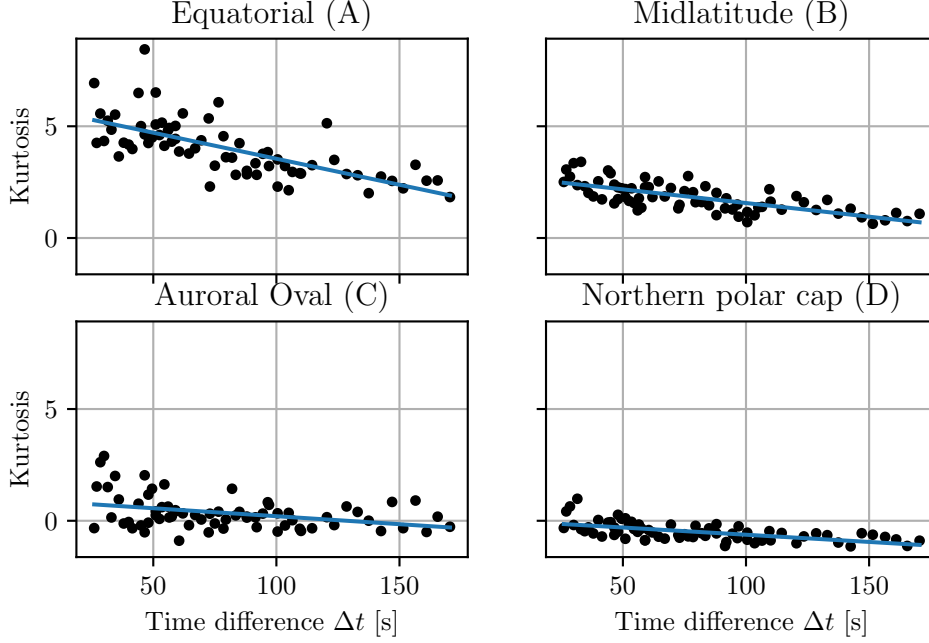
In (D), we have the results for the northern polar cap region, between  $77^\circ$  and  $90^\circ$  magnetic latitude. Here, the kurtosis starts out spread somewhat evenly around 0 and quickly falls below zero, ending at around -1.



**Figure 26:** Skewness of comparison indices plotted against time difference between satellite pairs. Black dots indicate data points, while the blue lines indicate linear regression. One black dot is the standard deviation of one day of data for one satellite pair. In (A), we have the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. In (B), we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. In (C) we have the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. In (D), we have the northern polar cap region, ranging from  $77^\circ$  to  $90^\circ$  magnetic latitude.



In Table 4, we see the regression coefficients for the linear regressions in Figure 27. In the first column we have the latitudinal regions, in the second column we have the regression slopes and in the third column we have the regression constants.



**Figure 27:** Kurtosis of comparison indices plotted against time difference between satellite pairs. Black dots indicate data points, while the blue lines indicate linear regression. One black dot is the standard deviation of one day of data for one satellite pair. In (A), we have the equatorial region, between  $-30^\circ$  and  $30^\circ$  magnetic latitude. In (B), we have the midlatitude region, between  $\pm 30^\circ$  and  $\pm 70^\circ$  magnetic latitude. In (C) we have the Auroral Oval region, between  $70^\circ$  and  $77^\circ$  magnetic latitude. In (D), we have the northern polar cap region, ranging from  $77^\circ$  to  $90^\circ$  magnetic latitude.

**Table 4:** Table of regression coefficients for the linear regressions found in Figure 27. The first column has the latitudinal regions, the second column has the regression slopes and the third column has the regression constants.

Region	Kurtosis regression slope [ $s^{-1}$ ]	Kurtosis regression constant [ $s^{-1}$ ]
Equatorial (A)	$-2.3 \cdot 10^{-2} \pm 2.9 \cdot 10^{-3}$	$5.9 \cdot 10^0 \pm 2.6 \cdot 10^{-1}$
Midlatitude (B)	$-1.2 \cdot 10^{-2} \pm 1.4 \cdot 10^{-3}$	$2.8 \cdot 10^0 \pm 1.3 \cdot 10^{-1}$
Auroral Oval (C)	$-7.1 \cdot 10^{-3} \pm 2.2 \cdot 10^{-3}$	$9.2 \cdot 10^{-1} \pm 1.10 \cdot 10^{-1}$
Northern polar cap (D)	$-6.5 \cdot 10^{-3} \pm 1.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-2} \pm 9.2 \cdot 10^{-2}$

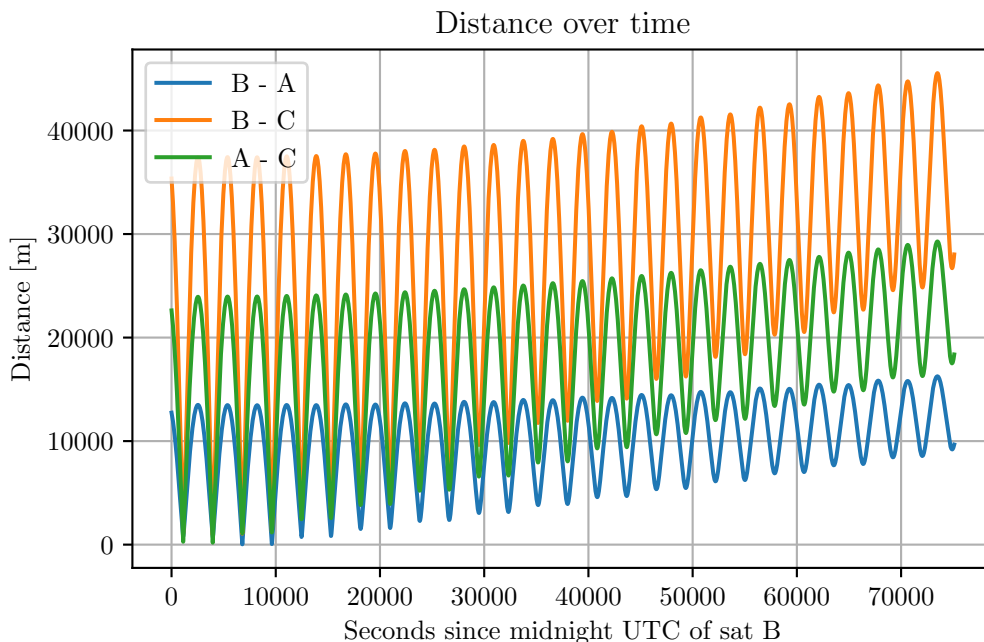
## 4.2 Assumptions and satellite drift

In this section, we present figures of the great circle distance between satellites after index-shifting. The distances shown are not the physical distances between the satellites, but rather the distance between the positions of the data points we

are comparing. First we present the great circle distance for the 10th of December and the 31st of December, plotted first against time then against latitude. We then present the average great circle distance for each satellite pair per day, first for all latitudes and then for the northern polar cap region, plotted against the time difference  $\Delta t$  as found while index-shifting. We also include linear regressions of this distance as a function of  $\Delta t$ . Lastly we present the mean absolute latitudinal difference per satellite pair per day, first for all latitudes then for the northern polar cap region, plotted against the time difference  $\Delta t$ .

In Figure 28, we see the great circle distance between where the satellite measurements were taken for the 10th of December 2013, after index-shifting, plotted against the time measured by satellite B. We see that the distance oscillates over time, with distances increasing over the duration of the day. Satellite pair B-A, shown with the blue line, has the lowest distances. Here, the values range from around 0 m to 12000 m in the beginning of the day, to between 10000 m and 15000 m at the end of the day. Satellite pair A-C, shown with green, has the second shortest distances. Here, the values range from around 0 m to 23000 m at the beginning of the day, increasing to 18000 m to 29000 m towards the end. Satellite pair B-C has the longest distances and is shown with the orange. Here, the values start out ranging from around 0 m to 37000 m in the beginning, ending up at around 27000 m to 45000 m.

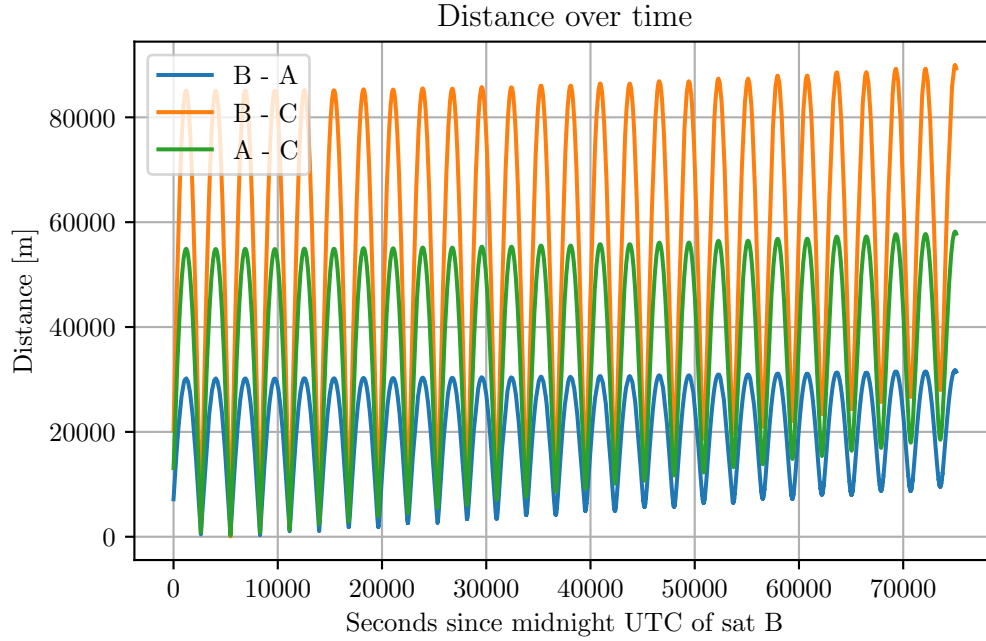
In Figure 29, we see the great circle distance between where the satellite mea-



**Figure 28:** Great circle distance between the data points on the 10th of December 2013, after index-shifting, plotted against the time measured by satellite B.

surements were taken for the 31st of December 2013, after index-shifting, plotted against the time measured by satellite B. The blue line notes satellite pair B-A, the orange line notes satellite pair B-C and the green line notes satellite pair A-C. The distance oscillates over time for all 3 satellite pairs, with values increasing over the duration of the day. For satellite pair B-A, the distances start out oscillating

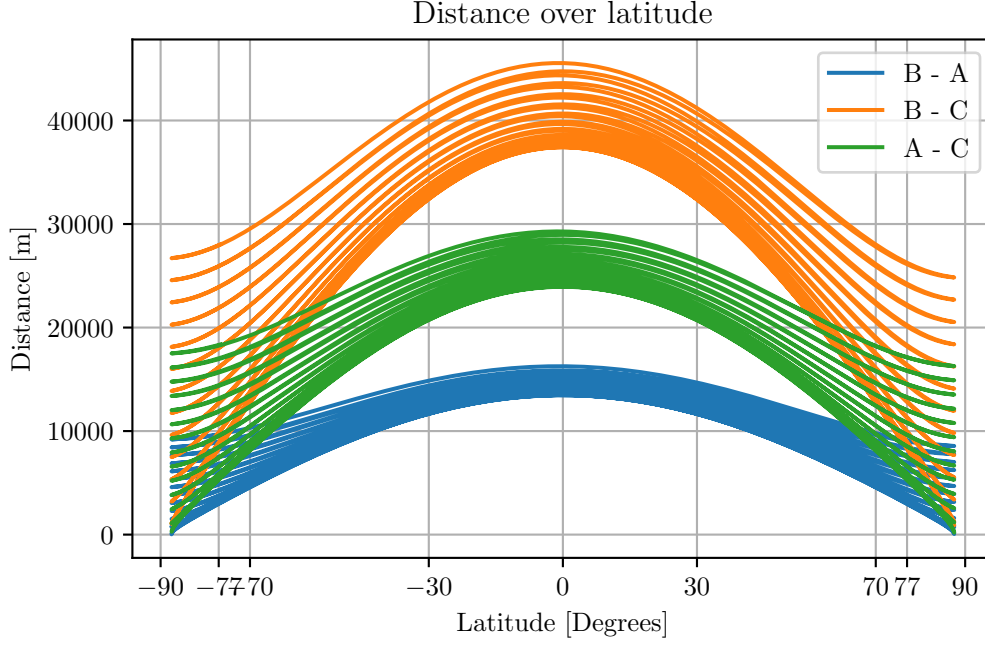
between around 0 m and 25000 m, increasing to oscillating between around 10000 m and 26000 m towards the end of the day. For satellite pair A-C, the distances start out oscillating between around 0 m and 55000 m, increasing to between around 20000 m and 59000 m towards the end of the day. For satellite pair B-C, the distances start out oscillating between around 0 m and 85000 m, increasing to between around 30000 m and 87000 m towards the end of the day.



**Figure 29:** Great circle distance between the data points on the 31st of December 2013, after index-shifting, plotted against the time measured by satellite B.

In Figure 30, we see the great circle distance between where the satellite measurements were taken for the 10th of December 2013, after index shifting, plotted against geographic latitude. The blue line notes satellite pair B-A, the orange line notes satellite pair B-C and the green line notes satellite pair A-C. The great circle distances show a clear dependence on latitude, being small around the poles and large around the equator. The distances increase with each orbital period. In the polar areas, the distances increase from almost zero to values almost equal to the initial distance at the equator. For satellite pair B-A the distance starts out at around 12000 m and ends at around 15000 m around the equator. At the pole, the distance starts out around 0 m and ends up around 10000 m. For satellite pair A-C the distance goes from around 25000 m to around 30000 m at the equator, and from around 0 to around 18000 m at the poles. For satellite pair B-C, the distance goes from around 38000 m to around 45000 m at the equator and goes from around 0 m to around 28000 m at the poles.

In Figure 31, we see the great circle distance between where the satellite measurements were taken for the 31st of December 2013, after index shifting, plotted against geographic latitude. The blue line notes satellite pair B-A, the orange line notes satellite pair B-C and the green line notes satellite pair A-C. The great circle

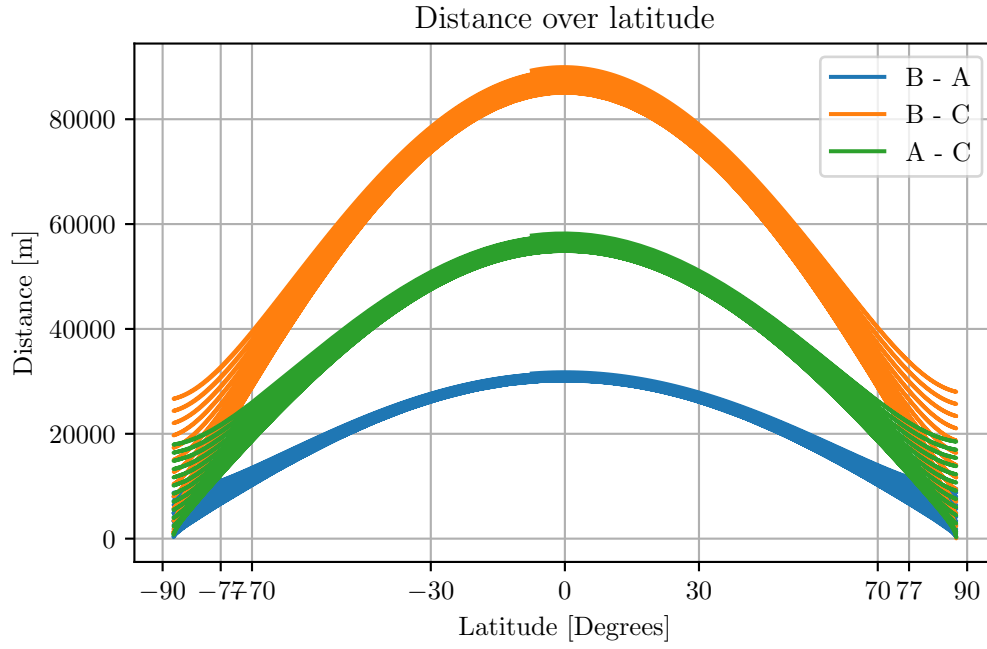


**Figure 30:** Great circle distance between the data points on the 10th of December 2013, after index-shifting, plotted against geographic latitude.

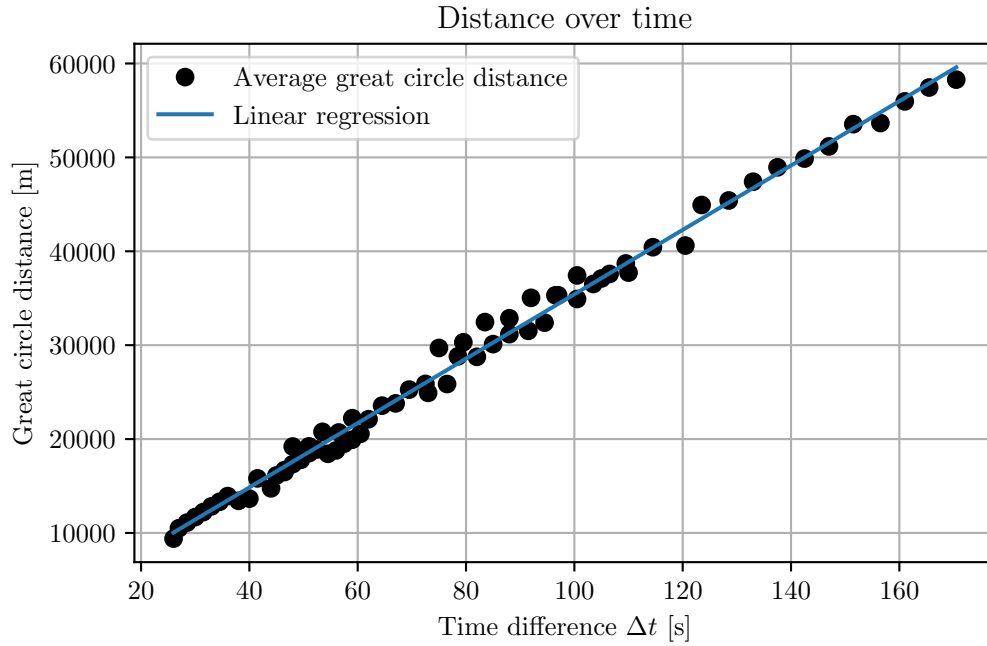
distances show a clear dependence on latitude, being small around the poles and large around the equator. The distances increase with each orbital period. The increase is much larger in the polar areas than at the equator. Satellite pair B-A has distances from around 25000 m to around 26000 m at the equator and distances from around 0 m to around 10000 m at the poles. Satellite pair A-C has distances ranging from around 55000 m to around 60000 m at the equator and distances going from around 0 m to around 20000 m at the poles. For satellite pair B-C, the distances range from around 85000 m to around 92000 m at the equator and from around 0 to around 30000 m at the poles.

In Figures 32 and 33 we see the average great circle distance between satellite pairs after index-shift plotted against the time difference  $\Delta t$  as found when index-shifting, along with a linear regression. The black dots indicate data points and the blue line indicates the linear regression. Figure 32 has the average distances for all latitudes, while Figure 33 is limited to geographic latitudes larger than  $77^\circ$ . We see a clear correlation between distance and time difference  $\Delta t$  in both figures. The linear regression in Figure 32 is found to have a slope of  $342.88 \pm 3.31 \frac{\text{m}}{\text{s}}$ . In Figure 33, the slope is  $126.15 \pm 5.34 \frac{\text{m}}{\text{s}}$ .

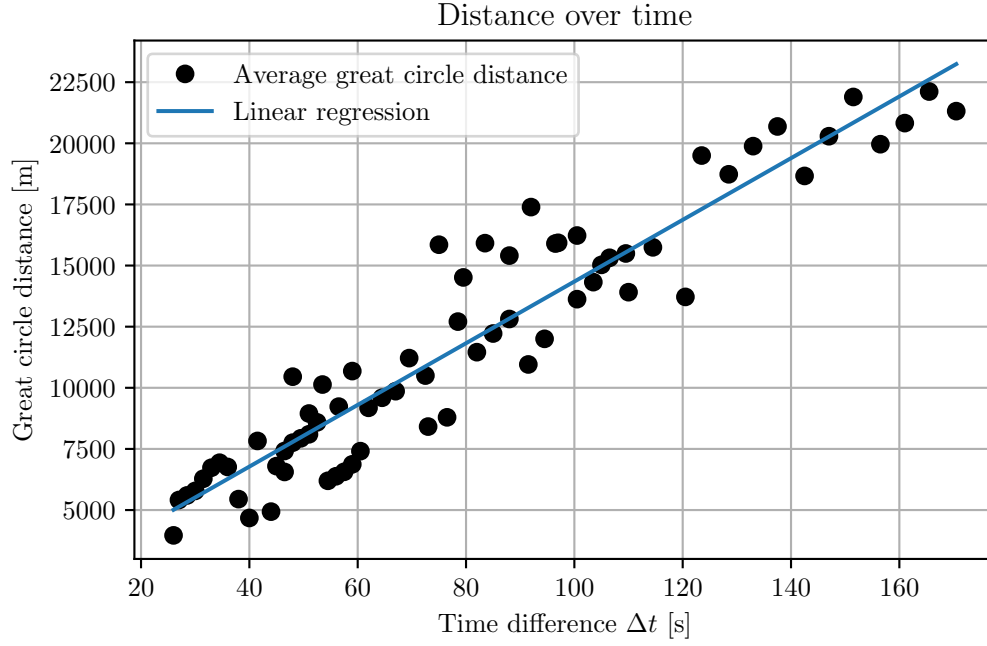
In Figures 34 and 35, we see the average absolute difference in latitude between satellite pairs after index-shift, plotted against the time difference  $\Delta t$  as found when index-shifting. Figure 34 has the average absolute difference in latitude for all latitudes, while in Figure 35 we are limited to latitudes larger than  $77^\circ$ .



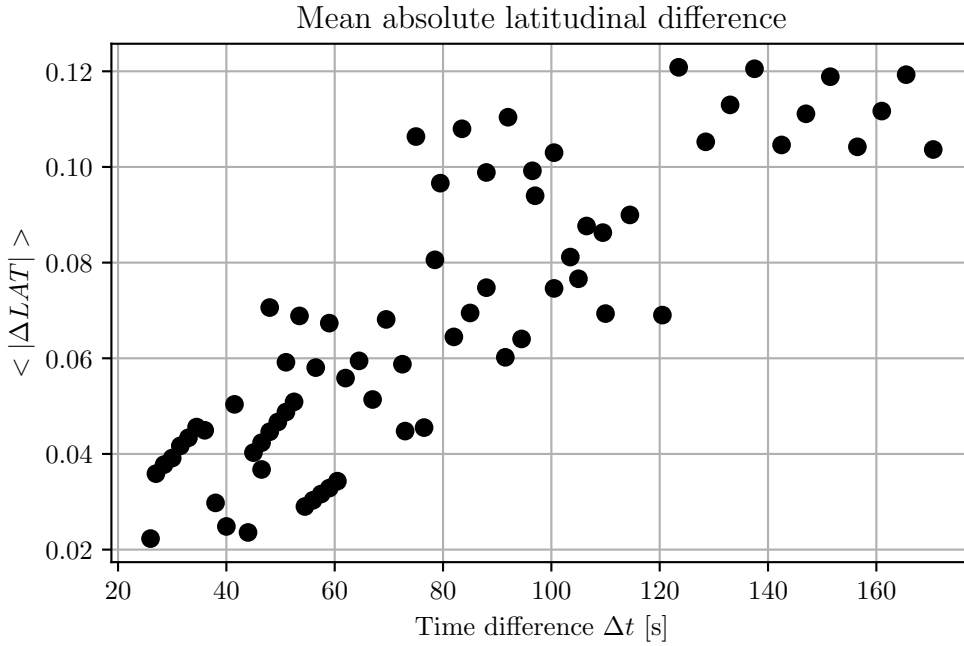
**Figure 31:** Great circle distance between the data points on the 31st of December 2013, after index-shifting, plotted against geographic latitude.



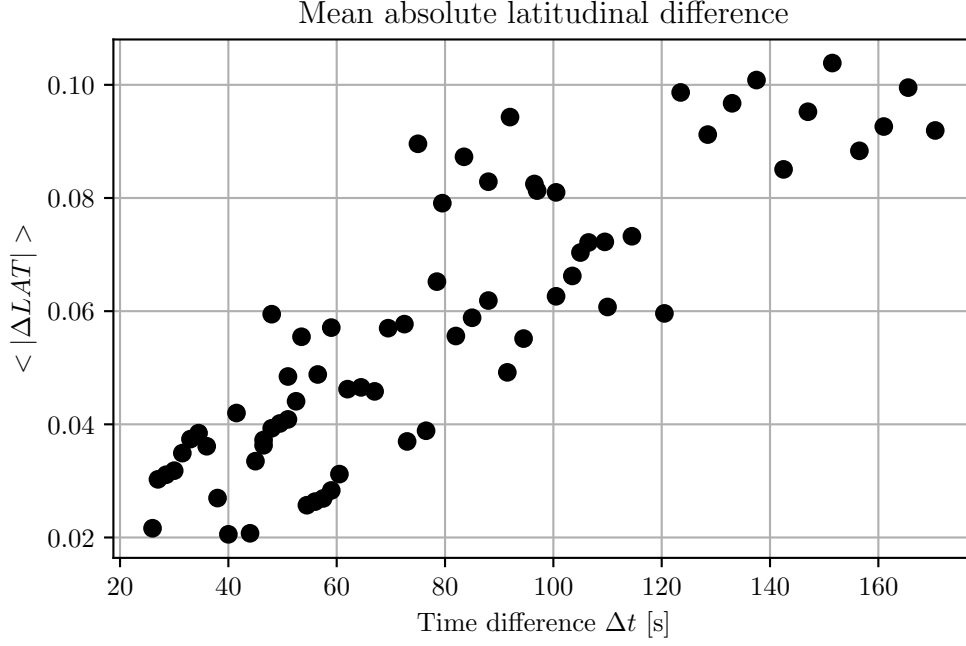
**Figure 32:** Daily average great circle distance between satellite pairs after index-shift for all latitudes, plotted against time difference between satellite pairs.



**Figure 33:** Daily average great circle distance between satellite pairs after index-shift for latitudes  $> 77^\circ$ , plotted against time difference between satellite pairs.



**Figure 34:** Daily average absolute latitudinal difference between satellite pairs for all latitudes, plotted against time difference between satellite pairs.



**Figure 35:** Daily average absolute latitudinal difference between satellite pairs for latitudes  $> 77^\circ$ , plotted against time difference between satellite pairs.

## 5 Discussion

It has long been known that electron density structures in the ionosphere are turbulent and can experience rapid changes [14] [11] [16]. Rapid changes in the electron density can affect trans-ionospheric radio waves, such as in the case of GNSS systems [17]. With society becoming increasingly dependant on such systems, understanding how electron density structures change is an important endeavor.

The ionosphere can be explored through ground-based measurements such as the measurement of scintillations [14] or in-situ through the use of sounding rockets [18] and orbiting satellites [15] such as the Swarm satellites. For in-situ measurements, one can directly measure electron density. Such a measurement yields a time series of electron density, and interpreting it is done by inspecting how the electron density changes.

When looking at two or more electron density time series collected at the same spatial location at different times, such as with the electron density data from the Swarm mission, we can observe that they are different from each other. The question is why. The obvious answer is that the electron density in the location we are looking at has changed. Where we might first have measured an electron density of say  $5 \cdot 10^4 \text{ cm}^{-3}$ , it might have increased to  $1 \cdot 10^5 \text{ cm}^{-3}$ . The question then becomes why the electron density changes.

We can divide the changes in electron density into two groups; spatial changes and temporal changes. If the density structure we are looking at changes shape somehow, such as an instability growing, it will naturally change over time. This is a temporal change, and will lead to a time dependant difference in the electron density measurements. If the density structure we are looking at is moving in rela-

tion to the satellites so that two satellites end up measuring different parts of the density structure, we get a spatial change. We could also have a combination of both, where an electron density structure is both moving and changing temporally at the same time.

Discovering whether changes are temporal or spatial in nature requires at least four spacecraft in a three-dimensional configuration [2]. In this paper, we try seeing what can be done with three spacecraft that follow the same orbit. In the following section, we will first discuss our methods and compare them to prior works. We will discuss the parameters of our analysis, such as data window sizes, the rolling mean and the size of the frequency band used in calculating the comparison index. Following the discussion on the method, we will attempt to interpret the physics behind the results shown in the results section. Here we will focus on the distance between data points, plasma drift velocity and known instability growth times.



## 5.1 Discussing the method

Fourier analysis has long been a staple of electron density structure analysis [14] [11] [19] [15]. In particular, it has been shown that in some cases one can characterize a PSD  $P(f)$  by a spectral index  $\beta$  through a power law  $P(f) \propto f^{-\beta}$  [14]. The spectral index  $\beta$  is the slope of the PSD on the log-log axes, and has been used to characterize irregularities [15].

In this paper, we introduced the comparison index. The comparison index  $I_{A-B}$  tells us the relative difference in power between two Fourier series, within a given frequency band. If we are looking at the Fourier transforms of signals  $A$  and  $B$  on the form of triangle pulses while excluding the lowest frequency bands, the comparison index tells us something about the difference in size between the triangles. Given that they have the same amplitude, a negative comparison index  $I_{A-B}$  would indicate that the triangle-like structure in signal  $A$  is larger/wider than the triangle-like structure in signal  $B$  as the thinner structure has more power in the higher frequency bands. Conversely,  $I_{A-B}$  being positive would indicate that  $A$  is thinner/smaller than  $B$ .

The spectral index and the comparison index serve the same purpose in presenting the large amount of information found in a PSD within a single number. A key difference is that the spectral index conveys information on one spectrum, while the comparison index was made to convey information on the difference between two spectra. One could perhaps achieve the same results as the comparison index by comparing two spectral indices. One drawback to this method would be that changes in the amplitude of large-scale structures could go unnoticed. When the amplitude of a large-scale structure increases, it is reflected equally in all frequency bands. The slope of the PSD would therefore remain unchanged, and a difference in spectral index would not reflect the difference in amplitudes.

This drawback is not present in the comparison index, since we compare the integral of the absolute value Fourier coefficients. This integral would naturally increase as the amplitude of the structure and therefore also magnitude of the Fourier coefficients increases. We are however dividing the comparison index by a normalizing factor, as seen in equation 3. While this limits the comparison index to the more intuitive values between -1 and 1, it truncates information on the size of the structures we are looking at, both in width and amplitude. In Figure 22, we compensate for this by giving each data point a transparency based on the size of the integrated Fourier coefficients. Weighting the data points in this way when doing our statistical analysis could perhaps have compensated for the normalization, and removed any outliers produced by sporadic changes in an otherwise calm electron density.

Another spectral analysis method that is utilized in multi-point analysis is the cross spectrum [23]. The cross spectrum between two Fourier series is found by multiplying one with the complex conjugate of the other, and dividing by a normalizing factor. This effectively gives us a spectrum that tells us the similarities between the two Fourier series. Frequency bands with activity on both Fourier series will be enhanced, while frequency bands with activity in only one or none of the Fourier series get diminished. In our case, we are interested in the differences between two signals which is why we introduced the comparison index. Since the cross spectrum gives us information on the similarities between two signals, we found it to be less useful in our analysis.

### 5.1.1 Index-shift

The first thing we did in order to compare the electron density time series gathered by the Swarm satellites was index-shift the time series of satellites A and C so that their coordinates were aligned with those of satellite B. Recognizing that the satellites flew in a cross-polar, pearls-on-a-string configuration making the difference in latitude the only meaningful index, we did this by minimizing the difference in latitude between the satellites. This was done at the beginning of each day of data. Since the satellites drift apart over time, this causes a difference in position to build up over the course of the day. As seen in Figures 28 and 29, the distance between data points is near to zero at the polar caps at the beginning of each day, but increases to 10000 - 30000 meters depending on the satellite pair over the course of a day.

In 2015, Park et al. calculated the along-track velocity of large-scale structures in the polar cap regions using data from the Swarm mission [7]. In doing so, they performed a similar index-shift to what we did. Instead of minimizing the difference in latitude as we did, Park et al. (2015) found the time of the satellites passing the poles. The difference in this time would then be  $\Delta t$  needed for the index-shift. This is potentially a better method than simply trying out many different values of  $\Delta t$ , as it would be less computationally exhaustive and be less dependant on input parameters.

The method used by Park et al. (2015) also brings to light a perhaps more natural way of dividing the data. Rather than dividing the data by days, we could divide the data by satellite periods, index-shifting and processing one satellite period worth of data at a time. This would remove the accumulative latitudinal error in distance between data points, as the satellites drift less per period than index-shifting by one index can account for. There are only room for four to five one minute windows with 50% overlap per pole-pass. If the data was divided by satellite periods, it would therefore still be advisable to for example combine all the satellite periods that have the same index-shift  $\Delta t$  when later doing statistical analyses. Said statistical analyses would benefit from an increased temporal resolution, but the data points might gain a more sporadic behavior as a consequence of a smaller sample size.

The benefits of dividing the data by satellite periods over dividing it by days is questionable. At the worst, a latitudinal difference of 30 kilometers is equivalent to less than eight indices. If a time window has a structure just on the edge, one time window could include eight more indices of the structure than the other. Since the Fourier transform assumes the given signal to be periodic [4], it will treat the cut structure as if it was thinner than the one that was not cut, changing the comparison index. This would be rectified by the next window, which overlaps the previous by 50%. While this effect is noticable when calculating individual comparison indices, it probably has no effect on the linear regression coefficients shown in the results section, as the latitudinal drift does not increase much over the course of the days included in our analysis.

### 5.1.2 Data gaps

When doing a multi-point analysis such as computing the comparison index, it is important that the time series are synchronized. This is mostly the case for the

data gathered by the Swarm satellites. Index-shifting effectively de-synchronizes the time series temporally in favor of attempting to synchronize them spatially. The validity of the spatial synchronization is investigated in section 4.2 of the results.

The presence of data gaps disturbs the synchronization of the time series gathered by the Swarm satellites. The way we dealt with these was to make equivalent data gaps in the other time series, and then ignoring any time windows where a data gap was present, as was done by Park et al. in 2015. This method ensures that synchronicity is kept, but discards a lot of data since we discard any time window where a gap is found. This could perhaps be avoided. There are multiple data gaps in the Swarm data that are on the order of a few seconds. An alternative way of dealing with this would be to replace the data by interpolation, as was done by Ritter et al. in 2013. Since a data gap in one time series would make us discard that time window for all time series, interpolating over the smallest data gaps would let us save much data. One can imagine that this would especially be the case when more than three measurement points are used, such as might be the case in future sounding rocket projects.

### 5.1.3 Dividing by geomagnetic latitude

Because the ionospheric conditions differ depending on latitude [6], it is natural to divide the ionosphere into regions by latitude. The different phenomena that govern the different latitudes can then be investigated separately. This is especially important when we calculate the four statistical moments of the distribution of comparison indices, as seen in the results section. We expect the different regions to yield different results. This is illustrated in Figure 22 where we can see the comparison indices behave differently depending on latitude.

Many ionospheric phenomena are better labeled by geomagnetic latitude rather than geographic latitude, as they are governed by the magnetic field [6]. It therefore makes sense to use geomagnetic coordinates over geographic ones. In our case, we computed the geomagnetic coordinates of our data points through the use of altitude-adjusted corrected geomagnetic coordinates (AACGMV) [10], implemented through the `aacgm2` package in python. A problem then appeared where some low latitude regions are undefined in AACGMV. We dealt with this problem by using geographic latitudes instead of geomagnetic latitudes where AACGMV was undefined. When plotting indices over geomagnetic latitude, this yields unrealistic results which is why we chose to use geographic latitudes in Figure 22. When grouping the data points by latitude, this is not a problem as the region of the geographic latitude coincide well with the undefined region of geomagnetic latitude. For Figure 22, geographic latitudes do not fully reflect the physics that govern the computed comparison indices. Especially in the southern polar region, geomagnetic and geographic latitudes do not coincide well. This could be why in Figure 22, we notice that the comparison indices in the northern polar region follow a latitudinal pattern while the comparison indices on the southern polar cap are more spread out latitudinally.

In 2018, Miloch et al developed the Ionospheric plasma irregularities (IPIR) product for characterization of ionospheric irregularities, for the Swarm mission [9]. Here they assigned the Swarm data to geomagnetic regions, with the equatorial region

corresponding to  $\pm 30^\circ$  MLAT and mid-latitudes between the equatorial region and the auroral oval boundary. This is what we based our geomagnetic regions on. In the IPIR product, the auroral oval region was calculated dynamically, with the polar cap region being poleward of the auroral oval. When the auroral oval region was not well defined, the polar cap region was set to  $> 77^\circ$  MLAT. Ideally, we would have used the IPIR latitude assignments in our work. However, we found that the IPIR assignments was not done for the dates we looked at. With calculations of the auroral oval boundary deemed outside the scope of our work, we decided to set hard limits to where we expected the auroral oval to be.

After looking at models of the open-closed field line boundary positions [8], we decided to put the lower auroral oval limit to the conservative  $70^\circ$  MLAT. A more fitting limit might have been  $60^\circ$  MLAT or  $65^\circ$  MLAT, as the northern twin cell convection pattern is found to stretch that far south quite regularly [13]. No matter what limit is set, the best analysis of the auroral oval region can only be made by calculating the auroral oval boundaries dynamically, as was done in the IPIR product [9].

In our case, our results from the auroral oval region should be taken with a grain of salt. The most important aspect of our latitudinal divisions is separating the northern polar cap region. As seen in the investigations of distance between data points in section 4.2 in the results, the polar cap regions are the regions where our assumption of inspecting the same spatial locations are most correct. This was also noted by Park et al [7], where they limited their investigations to geographic latitudes poleward of  $\pm 85^\circ$ .

#### 5.1.4 Parameters

When dividing the electron density time series into windows, we decided on a window length of one minute, or 120 indices, as was done by Park et al [7]. At an orbital velocity of about 7600 m/s, this is equivalent to 456 km which again corresponds to about 3.8 degrees of latitude. In the polar cap, we could potentially find structures with horizontal dimensions as large as 1000 km [11]. Our window size is therefore not large enough to analyze the largest structures. Regardless, some basic investigations pointed towards a window size of one minute giving us a good balance between having room for the larger structures while not including too many structures at the same time. A window size of one minute also gives us a decent sample size for the statistical analyses, as seen in the gaussian nature of the histogram in Figure 23. It is worth noting that the window size  $t$  decides the spectral resolution  $\Delta f$  through  $\Delta f = 1/t$  with  $t$  being measured in seconds. In our case, the spectral resolution does not matter much, except for deciding how small the bandwidth we integrate over can be.

When choosing the frequency band to integrate over when calculating comparison indices, we ended up integrating from 0.1 Hz to the Nyquist frequency of 1 Hz. When investigating what effect different lower integral limits had on the computed standard deviations, we found the standard deviation to increase until we hit a lower integral limit of about 0.1 Hz, as seen in Figure 42 in the appendix. It is worth noting that we found a discontinuity when going from the very lowest integral limit to a slightly higher one. This can be explained by us not detrending the electron density data before doing the Fourier transform. The lack of detrending

causes the zero'th Fourier coefficient to have a quite high value compared to the rest. This then completely dominates the computed comparison indices, drastically decreasing the standard deviation. It is also worth noting that the standard deviations follow the same trends, with satellite pair B-A consistently having the smallest standard deviation, followed by A-C and then B-C. This is the same order as the distance between satellite pairs, as seen in e.g Figure 30. We can therefore see a correlation between the distance between satellites and the standard deviations, something that is made quite clear in Figure 25.

It might be worthy of note that the discarding of the lowest frequency components is also done when calculating spectral indices [14]. In our case, we have calculated a rolling mean over the electron density time series. This has effectively removed the higher frequency components caused by smaller structures, leaving us with the higher frequency components of the large-scale structures themselves. If we were to calculate comparison indices for smaller scale instabilities, we could base the chosen frequency band off frequencies that can be approximated by a power law, while abstaining from taking a rolling mean.

In order to limit our investigation to large-scale structures, we calculated a rolling mean of the electron density in each window of data. By trial and error, we found that letting each data point be the average of itself and 5 neighbouring data points in both directions was a decent parameter. This process should ensure that the intuition we built in section 2.2 remains valid. With a rolling mean over 5.5 seconds, structures with scale sizes  $L$  smaller than around 40 km are effectively negated. This is equivalent to implementing a low pass filter that removes frequencies higher than 0.18 Hz. In order to fully understand the effect of a rolling mean on the comparison index, rigorous case studies of comparison indices of large-scale structures could be done. This was considered outside the scope of this paper. We can postulate that the minimum frequency set by the rolling mean should coincide with the lower integral limit set when calculating the comparison index. Further investigations would be required to test the effect this would have on computed comparison indices.

## 5.2 Interpretations

We have calculated comparison indices for electron density data gathered by the Swarm mission between the 9th and 31st of December 2013, as seen in Figure 22. After dividing the data by latitude and creating histograms as seen in Figure 23, we noticed a correlation between the time difference  $\Delta t$  between satellite pairs and the standard deviation of the histograms. As such, we got the idea to calculate the four statistical moments of the distribution of comparison indices for each day of data, and plot it against the time difference  $\Delta t$ . As  $\Delta t$  increases, the structures we are measuring have more time to change. Hence, the idea was that if we saw a trend between the statistical moments and  $\Delta t$ , it could indicate temporal changes in the structures.

### 5.2.1 Spatial and temporal changes

The main causes of spatial changes are differences in spatial location between data points and plasma drift. If the two satellites whose data we are comparing are following entirely different orbits around the Earth, the structures they are measuring will naturally not coincide. As for the Swarm mission, we are assuming

that the satellites are following a pearls-on-a-string configuration. This means that the longitudinal differences between the satellites must be limited. In order to test this assumptions, we plotted a number of figures, as seen in section 4.2 in the results.

In Figures 28 and 29 we see great circle distances between satellite pairs after index-shift for the 10th and the 31st of December 2013. We can see that the distances oscillate as the satellites orbit the Earth. This is as expected for satellites following slightly different lines of longitude, as the longitudinal difference goes as the cosine of the latitude. This causes the greatest distances to be at the equator and the shortest at the poles, as seen in Figures 30 and 31. Further, we see that the distances increase over the course of a day as discussed earlier. We can also note that the maximum great circle distances more than double over the course of December.

In Figures 30 and 31 we see great circle distance plotted against geographic latitude. We see that the distances decrease with latitude. In fact, while the distances at the polar areas are almost negligible, the distances at the equatorial regions are quite large. With distances as large as 80 kilometers, the impact of the distance between satellites becomes more noticable than the effect of plasma drift. Even with plasma drift velocities as high as 100 meters per second, the plasma drift can at most cause an error of around 15 kilometers.

Since we can quite easily control the difference in latitude, the main problem arises in longitudinal differences. If the longitudinal differences are large, we end up sampling slices of the density structures at different places. If the structures we are inspecting have spatial variations with spatial scales smaller than the longitudinal distance, the difference between the measured electron density structures and therefore also the computed comparison index could be quite large. In Figures 32 and 33, we see the average great circle distances between satellite pairs per day against the time difference between satellite pairs. In the first figure we see distances for all regions, while we have limited ourselves to the northern polar cap region in the second figure. We have also done a linear regression of distance as a function of time, and find a strong correlation in both figures, confirming that the distances indeed increase over the course of December. From the slopes of the linear regressions, we see that the average distance across all regions increase about three times faster than the average distance in the northern polar cap region.

In Figures 34 and 35 we see the average absolute difference in latitude between satellite pairs, plotted against the time difference between satellite pairs. In the first figure we have the average of all regions, while in the second figure we have the average for the northern polar cap region. We see that these figures are quite similar, meaning that the latitudinal differences do not depend on the latitude of the satellites. The difference in great circle distance between Figures 32 and 33 must therefore be longitudinal of nature. As such, the results for the northern polar cap must be more reliable than for the other regions. We can calculate the latitudinal distances by multiplying the difference in latitude by 120 km. For the northern polar cap, this gives us latitudinal differences of about half of what we see in Figure 33. As such, we can see that the average longitudinal differences in the northern polar cap region goes from around 2 km to around 10 km. With satellite velocities of about 7600 meters per second, this is not very much. The number would still continue to go down if we limited the northern polar cap to latitudes



even greater than  $77^\circ$ , such as was done by Park et al [7].

Other than differences in satellite orbit, the main cause of spatial changes are plasma drift. We choose to focus our efforts on the northern polar cap region, as it is the least affected by differences in satellite orbit while also being the most spatially active with large plasma drift velocities. It has long been known that the polar cap can experience plasma drift velocities on the order of several hundred to a thousand meters per second [13] [11] [5] [7]. The drift velocity of large-scale structures have been measured both from the ground through image tracking [5] and in-situ with Swarm satellites through use of a similar method as is explained in the appendix [7] [12]. The drift velocities vary, but values between 300 and 500 meters per second are not uncommon [7] [5].

The time differences between Swarm satellite pairs range from around 20 seconds to over 160 seconds. If the plasma is drifting across the satellite orbits, this can result in distances between measuring positions on the order of tens of kilometers. We can therefore say that the spatial effects of plasma drift is a lot more relevant than the spatial effects of differing satellite orbits in the northern polar cap region.

As for temporal changes, it is believed that a dominant mechanism for the creation of instabilities and therefore also changes in electron density in high latitude regions is the gradient drift instability [11]. We can perhaps say something about whether the gradient drift instability is affecting the comparison indices by looking at the time scales of gradient drift growth times. If the growth times are an order of magnitude larger than the tens of seconds of satellite time difference, we can be fairly confident that gradient drift instabilities are not contributing much to the comparison indices, and that any increase is of a spatial nature. In 2015, Spicher et al. did a case study of high latitude gradient drift instability growth times, and found growth times ranging from 23 seconds to 147 seconds, numbers that coincide surprisingly well with the time differences between satellites. In 2012, Oksavik et al. calculated gradient drift instability growth times using in-situ data from the ICI-2 sounding rocket and found values as low as 10 seconds. When we later look at our results, we can compare our time scales with these numbers.

### 5.2.2 Interpreting our results

In Figure 22, we see an example of comparison indices calculated for one day of data, plotted against geographic latitude. We see that the comparison indices are more prevalent in the equatorial and high latitude regions, and less prevalent in the mid-latitude regions. This is as expected, as the mid latitude region is typically more quiet than the high and low latitude regions [22]. In Figure 23, we see a histogram of comparison indices for the northern polar cap region for one day of data. As we can see, the histogram is fairly gaussian in shape, lending credence to the validity of a statistical analysis of the distribution of comparison indices. In order to interpret the statistical moments of our distributions of computed comparison indices, we must first understand what they indicate.

The four statistical moments of a probability distribution are the mean, standard deviation, skewness and kurtosis. The mean and skewness can tell us whether the structures tend to grow or shrink in size or amplitude. A negative mean would mean that the comparison indices tend to be negative more often than positive,

indicating that the structures are shrinking in size, that the gradients are getting steeper or that the amplitude is increasing. Conversely, a positive mean would mean that the structures tend to grow in size, that the gradients are becoming gentler or that the amplitude is decreasing. The skewness tells us the asymmetry of the distribution. If for example large structures tend to become a lot smaller, it would lead to a negative skewness and vice versa.

The standard deviation can tell us how large changes tend to be. A larger standard deviation means that the comparison indices tend more towards -1 and 1, telling us that the changes are more substantial. The standard deviation is perhaps the most important moment in our case. When inspecting the statistical moments as functions of the time difference between satellites, seeing the standard deviation grow would indicate that the structures change more as the time difference increases.

The kurtosis relates to the tails of the distribution, with more outliers in the data giving us a higher kurtosis. In our case, a high kurtosis could mean that while most of the comparison indices are centered around the mean, those that are not tend to be quite large. One would expect this to be the case if the structures we are looking at tend not to change much, but sometimes change a lot.

In Figure 24, we see the mean of the computed comparison indices plotted against the time difference between satellites, along with linear regressions. In table 1 we see the linear regression coefficients of the linear regressions. While all the regression constants are negative, indicating that the structures perhaps tend to shrink more than grow, the regression constants are not significantly different from zero. The means can therefore tell us nothing of significance. We also notice that the means are more spread out in the auroral oval and northern cap regions compared to the equatorial and midlatitude regions. This can simply be attributed to a higher sample size in the equatorial and midlatitude regions.

In Figure 25, we have the standard deviations of computed comparison indices plotted against the time difference between satellites, with accompanying linear regression coefficients found in table 2. These are our most interesting results. We can clearly see correlations between the time difference  $\Delta t$  and the standard deviation  $\sigma$ . This indicates that as more time passes, we are more likely to see larger differences in the electron density time series measured by the different satellites. The question then becomes why.

At first glance, the correlation is a clear indication of temporal changes, as it simply indicates that more time results in larger differences. However, as we have already discussed, the longitudinal difference between data points also increases over time, especially in the equatorial and midlatitude regions. Therefore, it becomes difficult to differentiate between changes caused by the change in distance and the change in time. In fact, if the correlation was entirely due to the increase in longitudinal distance it would explain why the equatorial and midlatitude region regressions have approximately the same slopes. As for the auroral oval, it is poorly defined in our case. It therefore becomes difficult to say anything about the physics governing the comparison indices in this region. As such, we will focus on the northern polar cap where the distance between satellites is low.

The northern polar cap features the steepest linear regression slope, indicating



that the structures tend to change more as more time passes. This can be explained either spatially through the structures moving, or temporally through an instability such as the GDI. As we have seen, the northern polar cap region has plasma drift velocities that can cause differences in measurement on the same order as the distance between satellites in the equatorial region. In addition, we have seen that gradient drift instabilities can have growth times on the same order of magnitude as the time differences  $\Delta t$ . We therefore can't dismiss either as the leading cause behind the increase in standard deviation, and we cannot easily differentiate between spatial and temporal changes.

A question comes to mind when looking at the effect of longitudinal plasma drift on the comparison index. Is the difference in structures measured dependant on how far the structure has drifted before being measured the second time? If the structure has a continous and smooth shape, like for example a spherical bubble, this would definitely be the case. A satellite measuring the width of the structure in the middle would find it to be wider than if it measured the width towards the edge. In this case, increasing the time between samplings and therefore also the distance the structure drifted across the satellite track would increase the computed comparison index. If the structure has a more sporadic shape where the width goes up and down along its length this will not be the case. If the structure has a more sporadic shape, the computed comparison index would be somewhat random, and no correlation between  $\Delta t$  and  $\sigma$  would be found. If this was the case for the density structures found in the polar cap region, the spatial changes would show up as a background in Figure 25, and would be characterized by the linear regression constant. Then the linear regression slope would fully be caused by temporal changes, and we could successfully differentiate between temporal and spatial changes. In such a case, we could for example have computed a characteristic time  $t_c = b/a$ , with  $b$  being the regression constant and  $a$  being the regression slope. This characteristic time would tell us at what time scales the temporal changes would be equally as impactful as the spatial ones. For the linear regression found in the northern polar cap region, the characteristic time would be 245.5 seconds.

In Figure 26 we find the skewness of the computed comparison indices plotted against the time difference between satellites, with the accompanying linear regression coefficients found in table 3. Much like with the mean, the regression constant is negative for all regions while not significantly different from zero. This indicates that the structures do not have a tendency to shrink more than they grow and vice versa.

In Figure 27 we find the kurtosis of the computed comparison indices plotted against the time difference between satellites, with the accompanying linear regression coefficients found in table 4. The first thing we notice is that all regions have significant negative linear regression slopes. This means that as time passes, we are more likely to see higher value comparison indices. This makes sense, as more time passing would give the structures more time to change. We also see from this figure that the equatorial region has quite high kurtosis compared to the other regions, in particular compared to the northern polar cap region. This indicates that changes do not appear quite as frequently in this region, but when they do they are quite large. This could be explained by one satellite measuring a structure that the other satellite misses due to a longitudinal distance. This is also the case for the midlatitude region. For the polar cap region, we see that the

kurtosis becomes negative over time. This indicates that there are few outliers in the computed comparison indices, a good indicator that we are not seeing too many sporadically high comparison indices.

Differentiating between temporal and spatial changes by doing a linear regression of the standard deviation of computed comparison indices over time differences turns out to not be straightforward. Perhaps the biggest issue is that while we can calculate the along-track velocity of any structures we find [7], we cannot find the cross-track velocity. It is the cross-track velocity that is important in this case, since latitudinal velocities will be reflected in the time series. When a structure moves along the satellite tracks we see it in the time series, but we cannot differentiate between a change caused by a cross-track movement and a temporal change. If we could somehow calculate the cross-track velocity of the plasma, we could do a multiple linear regression, separating out the increase in standard deviation caused by the plasma drift. An idea would be to combine the multi-point data with other sources. This is however outside the scope of this project.

When we combined one day worth of data into a statistical moment, we did it without regard for the ionospheric conditions at the time. Since the structures we encounter depend on ionospheric conditions, such as the strength of the Z component of the IMF [1], we could perhaps assign different indices to the data by what the ionospheric conditions were at the time. This in turn would let us inspect how the changes in electron density is affected by the different conditions. If for example we could expect that gradient drift growth times are larger when the solar wind is calm, we might be better able to identify their effect on the comparison index. One way of doing this would be to use data from the Advanced Composition Explorer [21], a satellite at the Sun-Earth libration point, built to among other things take measurements of the solar wind. This is however outside the scope of this project.

## 6 Conclusion

In this paper, we have undertaken a multi-point analysis of electron density data from the Swarm mission. In order to see if we could differentiate between spatial and temporal changes in ionospheric structures, we introduced the comparison index and performed a statistical analysis of comparison indices computed using data from the Swarm mission for the 9th through 31st of December 2013.

We can conclude that the Swarm data is well suited for multi-point data analysis in the higher latitudes, as the orbital paths are practically identical. For lower latitudes however, the difference in orbital paths quickly become non-negligible. Doing multi-point analyses using Swarm data are therefore less advisable in regions other than the polar regions.

We find that the comparison index could be well suited for studying changes in large-scale structures. However, we were unable to say much about the processes behind individual comparison indices. During our statistical analysis, we found that the comparison index tends to have higher values as the distance between satellites increases. For the regions other than the polar cap, this could be attributed to an increase in longitudinal distance between satellites. For the northern polar cap region, the longitudinal distance is negligible. Here, we inspected plasma drift and GDI as plausible causes for the increase in standard deviation of comparison

indices. We found that plasma drift can cause positional differences on the order of 80 kilometer, and we found GDI growth times on the same order as the time difference between satellites. We can therefore conclude that neither have a non-negligible impact on the increase in standard deviation. Thus, as expected, we cannot differentiate between a temporal or spatial cause in the increase of standard deviation purely from our multi-point analysis using 3 satellites.

## 7 Outlook

There are many parameters in our analysis, such as window size, frequency integral limits, how the data was grouped, the calculation of a rolling mean and the use of a window function. While we studied these in our analysis, we did not do so rigorously. As such, there is potential to further our understanding on how the comparison index behaves. There is also much potential in utilizing the comparison index. One idea would be to calculate spectral indices for two satellites, and then calculate the comparison index for the two where we set the integral limits in accordance to where the spectrum exhibited a power law. Another idea would be to group the data by ionospheric conditions. We could for example group the data by the strength of the Z component of the IMF by looking at data from the ACE satellite. We could then see how the statistical moments of the comparison indices in the northern polar cap change depending on different ionospheric conditions.

The main challenge we had when trying to differentiate between spatial and temporal changes in the northern polar cap is that we can't tell the cross-track plasma drift velocity from only looking at data from two or three satellites flying in a pearls-on-a-string configuration. If we could group the data by plasma drift velocity, we could do a multiple linear regression, and thus find how much of the increase in standard deviation is explained by spatial effects. One way to find the cross-track plasma drift velocity would be to change the orbital configuration of the Swarm satellites into a right-angled triangle, where two satellites have the same longitude and two satellites have the same latitude. The distance between the satellites would have to be smaller than the length scale of the structures we are interested in looking at. Some effects are spatio-temporal in nature. For example, GDI growth times are dependant on the velocity of the plasma. For these, this method would not work.

Work could be done on investigating the spatial effect on the standard deviation of comparison indices as the time difference between satellites increase. Does a greater distance between satellite necessarily increase the spatial effects? In order to answer this, we would have to investigate the shape of the structures in question. If the structures are irregular in nature, the distance between two slices of the structure would not necessarily impact the difference between the two slices. An increase in the time difference between satellites would therefore not cause an increase in spatial changes.

## 8 Appendix

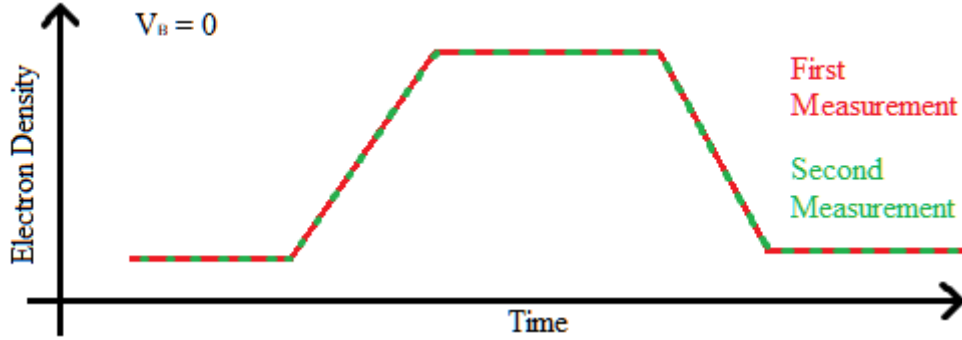
### 8.1 Codes

The analyses in this paper were performed in Python. The code used, along with a multitude of figures and a processed version of the Swarm data in .mat format, can be found on github through <https://github.com/Akselgra/Spacemaster>.

### 8.2 Finding the along-track velocity of ionospheric structures.

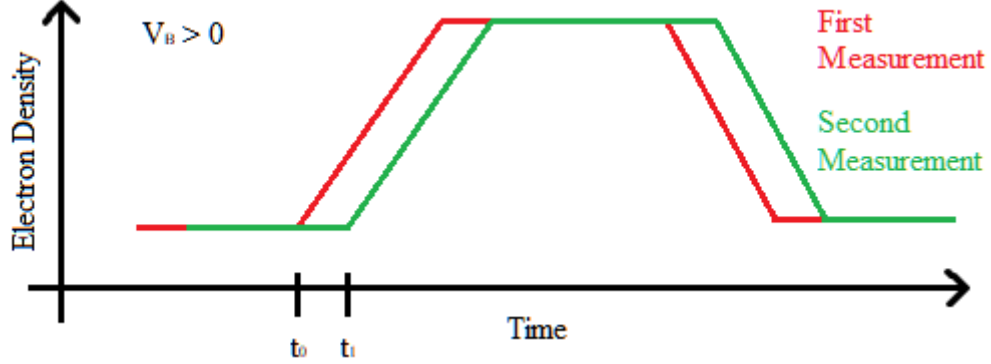
A well known exercise we can do with the data from the Swarm mission is finding the along-track velocity  $V_B$  of ionospheric structures. This is of interest because the movement of density structures is a major cause of spatial changes. After index-shifting, if the electron density structure is unchanged between the satellites measuring it, the time series will be identical. If a change in the time series is caused by the electron density structure moving relative to the orbit of the satellites, it will look akin to figures 36, 37 and 38 depending on whether the structure is moving with or against the satellites.

In Figure 36 we see what an electron density structure measured  $\Delta t$  seconds apart by two satellites may look like after index-shifting if the density structure has zero velocity along the track of the satellites, and if the density structure doesn't change over time. The time series would be identical.



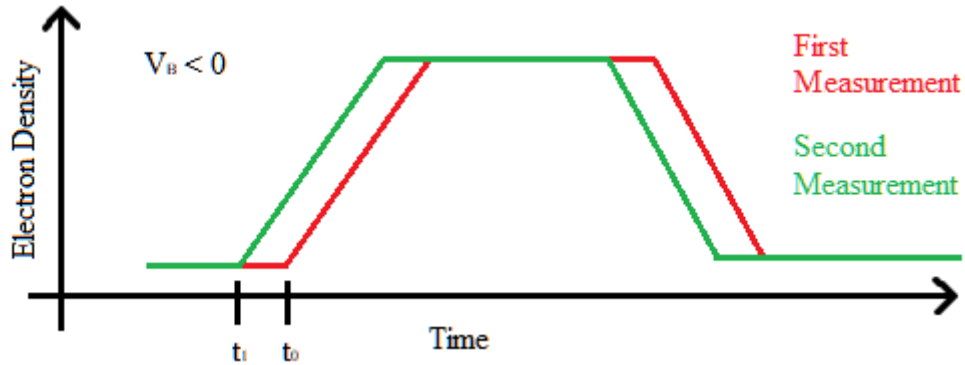
**Figure 36:** What an electron density structure measured  $\Delta t$  seconds apart by two satellites might look like if the density structure has zero velocity relative to the orbit of the satellites.

In Figure 37 we see what an electron density structure measured  $\Delta t$  seconds apart by two satellites may look like after index-shifting if the density structure has a positive velocity along the track of the satellites, and if the density structure doesn't change over time. Here, we see that the electron density in the second measurement increases later than for the first measurement, indicating a movement of the density structure. The satellites are moving at approximately the same velocity. As such, after index-shifting, a difference in time is equivalent to a difference in position. We name the time of the first satellite finding the density structure  $t_0$  and the time of the second satellite finding the density structure  $t_1$ .



**Figure 37:** What an electron density structure measured  $\Delta t$  seconds apart by two satellites might look like if the density structure has positive velocity relative to the orbit of the satellites.

In Figure 38 we see what an electron density structure measured  $\Delta t$  seconds apart by two satellites may look like after index-shifting if the density structure has a negative velocity along the track of the satellites, and if the density structure doesn't change over time. Here, we see that the electron density in the second measurement increases some time before that of the first measurement. This indicates that the density structure has moved in the opposite direction of the satellites, having a negative velocity. We again define  $t_0$  as the time where the first satellite discovers the density structure and  $t_1$  as the time where the second satellite discovers the density structure. Note that  $t_1 - t_0$  will have a different sign here than in the case of Figure 37.



**Figure 38:** What an electron density structure measured  $\Delta t$  seconds apart by two satellites might look like if the density structure has negative velocity relative to the orbit of the satellites.

We can name the time difference between the satellites discovering the density structure after index-shifting  $t_1 - t_0 = \Delta t_s$ . The along-track velocity of the density structure is equal to the distance  $X$  it has traveled, divided by the time  $t$  it took to travel that distance.

$$V_B = \frac{X}{t} \quad (8)$$

The distance  $X$  the structure has traveled by the time both satellites have found it is equal to the time difference between the two satellites finding it  $\Delta t_s$  times the satellite velocity  $V_s$ .

$$X = V_{sat} \Delta t_s$$

The total time the density structure spends traversing the distance  $X$  is equal to the time difference  $\Delta t$  found while index-shifting plus the extra time difference  $\Delta t_s$  between the satellites discovering the density structure.

$$t = \Delta t + \Delta t_s$$

Setting in for  $X$  and  $t$  in equation 8, we end up with equation 9 which tells us the along-track velocity  $V_B$  as a function of satellite velocity  $V_{sat}$ , the time difference  $\Delta t$  found while index-shifting and the extra time difference  $\Delta t_s$  between the satellites discovering the density structure.

$$V_B = \frac{V_{sat} \Delta t_s}{\Delta t + \Delta t_s} \quad (9)$$

We will be finding the velocity of one large-scale ionospheric structure measured by the Swarm satellites around 1150 seconds, or about 20 minutes, after midnight UTC on the 21st of December 2013. The exercise can be boiled down to finding the parameters  $\Delta t$  and  $\Delta t_s$  and putting them into equation 9 along with the satellite velocity  $V_{sat}$ .

The first thing we do is index-shift the time series. In doing so, we find the times  $\Delta t$ . Then, in order to lessen the impact of sporadic differences in electron density, we calculate the rolling mean by letting each data point be the mean of itself and 5 neighbouring data points both forwards and backwards.

In order to find  $\Delta t_s$ , we must find the position of the edge of the structure in each time series. The edge should be the place where the electron density suddenly increases. Consequently, the electron density gradient along the path of the satellite should have a maximum value right as the satellite enters the structure. Thus, in order to find the edge, and therefore also the time and position of the satellite entering the structure, we must first calculate the electron density gradient:

$$\nabla n_e = \frac{\Delta n_e}{\Delta x} \quad (10)$$

Where  $\Delta n_e$  is the difference in electron density between 2 points in the time series  $\Delta n_e = n_{ei} - n_{ei-1}$ , and  $\Delta x$  is the distance between measuring points in meters, and can be found by dividing the satellite velocity by the sampling frequency  $\Delta x = V_{sat}/f_s$ . In this case,  $\Delta x$  is about 3800 meters.

We can expect to discover velocities on the order of several hundred meters per second [5] [7]. With time differences around 50 to 150 seconds and satellite velocities around 7600 meters per second and a sampling frequency of 2 Hz, we can only expect a few indices to separate the edge of the structure. Missing the edge by one index can alter the result by as much as 75 meters per second. As such, we would like a better precision than simply assuming the maximum value of electron

density gradient to be the location of the edge. Instead, we approximate the density gradient with a gaussian fit, and assume the center of the gaussian to be the edge of the structure. We can then find  $\Delta t_s$  by taking the position of the mid value of a gaussian for a leading satellite, and subtracting the mid value of a gaussian for another. Assuming that all satellites are moving at the same speed, we can use time instead of position:  $n_{BA} = t_B - t_A$ .

Having obtained  $\Delta t_s$  and  $\Delta t$ , we calculate the velocity of the structure as seen in table 5. We also index-shift the time series obtained by satellites A and C backwards according to  $\Delta t$  in order to visually see how well our method fit.

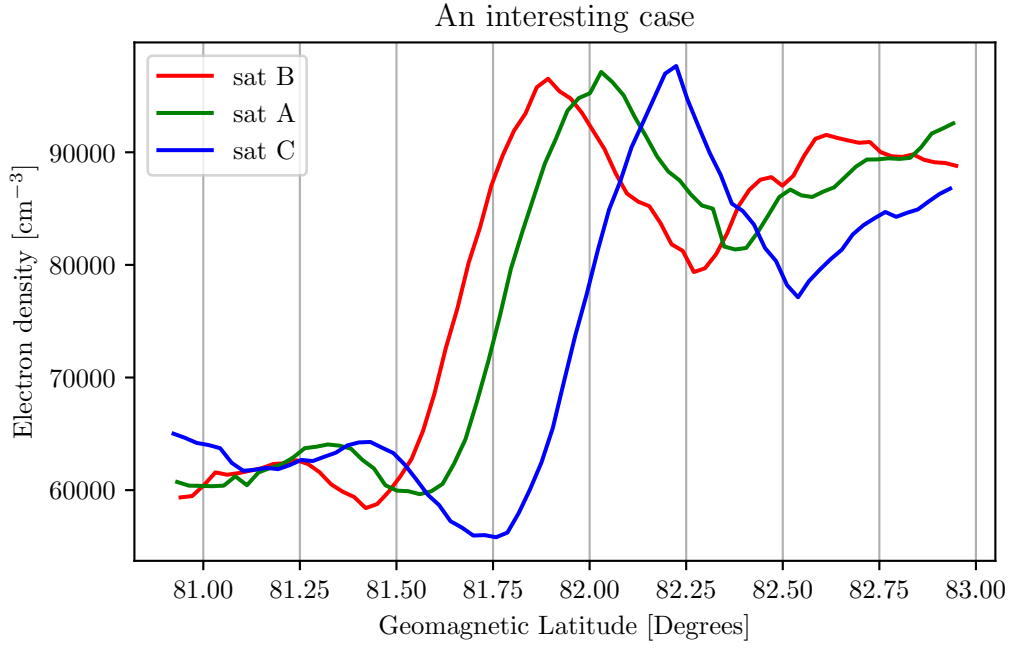
### 8.3 Along-track velocity of large-scale ionospheric structures: an example

In order to study the along-track velocity of a large-scale ionospheric structure, we look at an example from the northern polar cap region on the 21st of december. We use 35 seconds of electron density data, smoothened by calculating a rolling mean by letting each data point be the average of itself and 5 neighbours in both directions. We first present the electron density data, then electron density gradients with accompanying gaussian fits. Lastly we present the electron density data again, but index-shifted according to the results, along with the calculated along-track velocities.

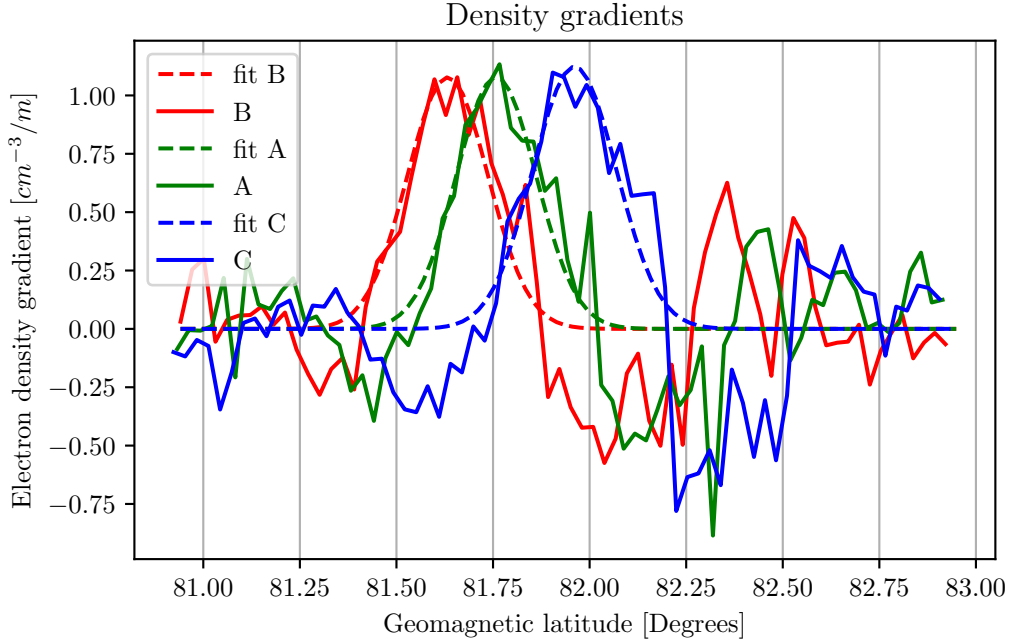
In Figure 39, we see the electron density data after smoothening. The electron densities start out at around  $60000 \text{ cm}^{-3}$ . The electron density measured by satellite B, shown as a red line, starts increasing at around  $81.5^\circ \text{ MLAT}$ , stopping at around  $90000 \text{ cm}^{-3}$  at around  $81.9^\circ \text{ MLAT}$ . The green line shows the electron density measured by satellite A. It follows the same pattern as the electron density measured by satellite B, but here the increase in density starts  $0.1^\circ \text{ MLAT}$  later, at around  $81.6^\circ \text{ MLAT}$  and ends at  $82.0^\circ \text{ MLAT}$ . The blue line shows the electron density measured by satellite C. It follows the same pattern as what satellites B and A measured, but here the increase starts at around  $81.75^\circ \text{ MLAT}$  and ends around  $82.25^\circ \text{ MLAT}$ .

In Figure 40 we see the electron density gradients and accompanying gaussian approximations. The color red is used for electron densities measured by satellite B, green is used for electron densities measured by satellite A and blue is used for electron densities measured by satellite C. The solid lines indicate the electron density gradients, while the dashed lines indicate the gaussian fits. The distance between the means of the gaussian fits indicates the along-track distance traveled by the ionospheric structure by the time the later satellites hit it. The mean of the fit for satellite B is at around  $81.625^\circ \text{ MLAT}$ , the mean of the fit for satellite A is at around  $81.75^\circ \text{ MLAT}$  and the mean of the fit for satellite C is at around  $81.9^\circ \text{ MLAT}$ . The calculated time difference in seconds is 2.16 seconds for satellite pair B-A, 5.64 for satellite pair B-C and 3.49 for satellite pair A-C.

In Figure 41, we have shifted the data from satellites A and C while holding B stationary, according to the results found in Figure 40, in order to get a visual representation of how well our method works. The red line represents satellite B,



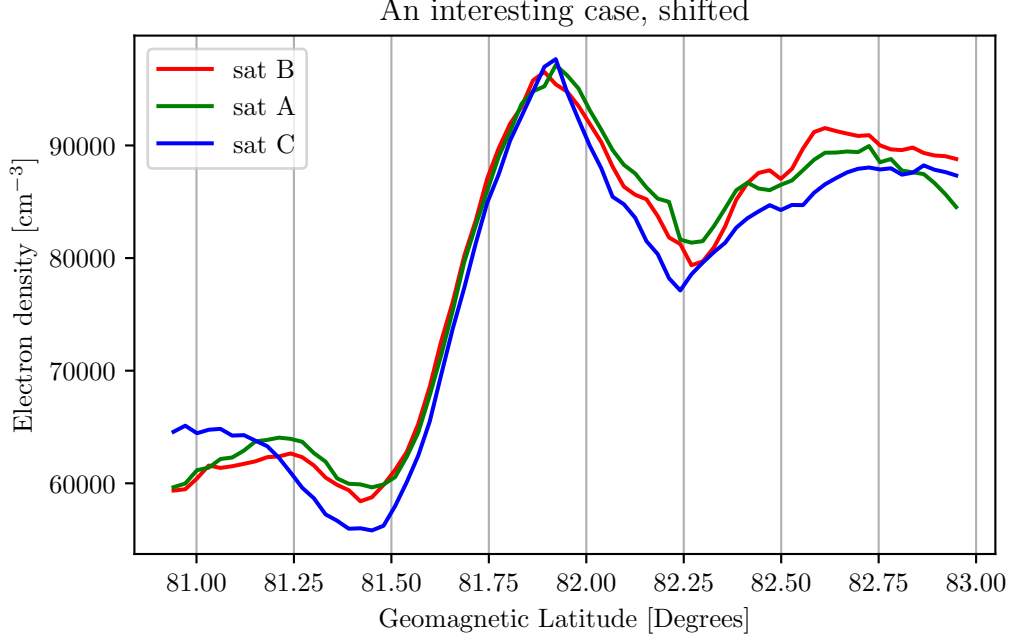
**Figure 39:** Electron density data after index-shifting from around 1150 seconds after midnight on the 21st of December, smoothened by taking a rolling mean. This is an interesting case, as we can clearly see the effects of the plasma structure moving in relation to the satellites orbit.



**Figure 40:** Electron density gradients, calculated from the electron densities in Figure 39 along the path of the satellites. The striped lines are gaussian fits of the gradients.



the green line represents satellite A and the blue line represents satellite C. We see that all 3 electron densities start out at around  $60000 \text{ cm}^{-3}$  and start increasing at around  $81.5^\circ$  MLAT. All 3 lines peak at around  $90000 \text{ cm}^{-3}$ , at around  $81.9^\circ$  MLAT. The 3 lines follow the same trend at the same magnetic latitudes, being barely distinguishable where they are increasing, between  $81.5^\circ$  MLAT and  $81.9^\circ$  MLAT.



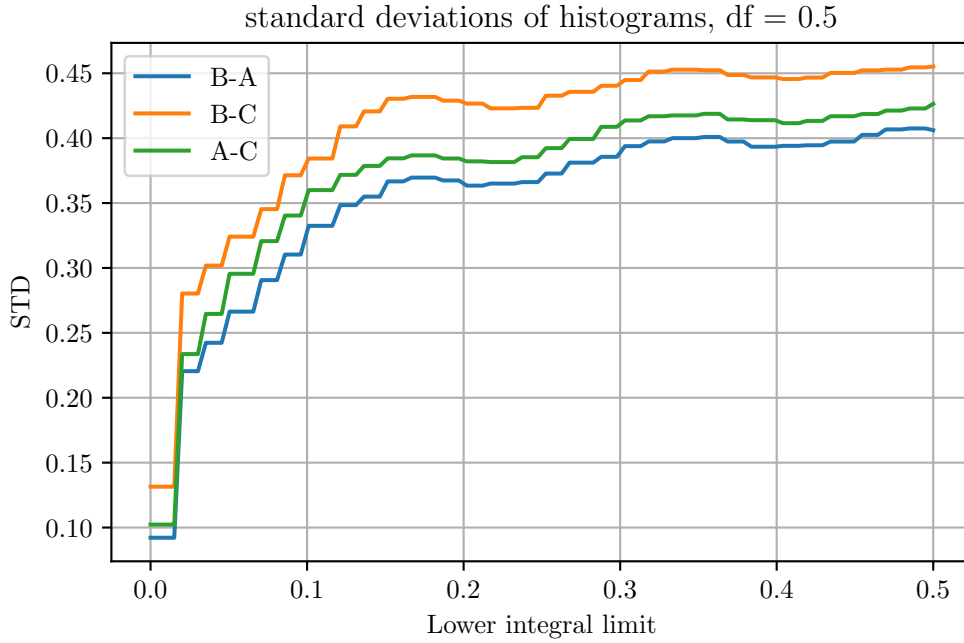
**Figure 41:** Electron density data from Figure 39 after shifting according to the distance  $n$  found in Figure 40. We now see that the data aligns almost perfectly.

In table 5, we see the numerical results for the distance the density structure has traveled by the time the different satellites measure it, given in units of seconds. Units of seconds are used as a consequence of the electron density data being time series. We can find the distance in meters by multiplying the distance in seconds with the common velocity of the satellites,  $7600 \frac{\text{m}}{\text{s}}$ . On the first column of the table we have the satellite pairs. In the second column, we see the distance traveled by the density structure between when a satellite and a following satellite measured it. In the third column, we see the time difference between the satellites, found while index-shifting. In the fourth column, we see the along-track velocity calculated through equation 9. We see that all 3 resulting along-track velocities are within 10% of each other, and on the order of  $300 \text{ ms}^{-1}$ .

**Table 5:** Numerical results for the along-track velocity of the ionospheric structure seen in Figure 39. In the first column we see satellite pairs. In the second column we see the distance traveled by the density structure between when a satellite and the following satellite measured it. In the third column, we see the time difference between satellites, as found while index-shifting. In the fourth column, we see the resulting along-track velocity as calculated using equation 9.

Satellite pair	$\Delta t_s$ [s]	$\Delta t$ [s]	$V$ [ $\frac{m}{s}$ ]
B-A	2.16	45.0	347.2
B-C	5.64	123.5	331.8
A-C	3.49	78.5	323.0

## 8.4 Other figures



**Figure 42:** Standard deviations of histograms as function of lower integral limit.

A multitude of figures are available on github through <https://github.com/Akselgra/Spacemaster/tree/master/SWARM/Figures>. In particular, one can find gifs that show the evolution of the cross spectral density (CSD) through <https://github.com/Akselgra/Spacemaster/tree/master/SWARM/Synthetic/Figures>.

## References

- [1] Russel, C.T et al (2016) *Space Physics, an introduction*. United Kingdom: Cambridge University Press
- [2] Paschmann, G & Daly, P.W (2000) *Analysis Methods for Multi-Spacecraft Data*. Noordwijk: ESA Publications Division
- [3] ESA (2021) *Swarm Mission* [internet]. Available from <https://earth.esa.int/eogateway/missions/swarm>
- [4] Manolakis, D.G & Ingle, V.K (2011) *Applied digital signal processing*. New York: Cambridge University Press
- [5] Hosokawa, K et al (2006) Estimating drift velocity of polar cap patches with all-sky airglow imager at Resolute Bay, Canada *Geophysical Research Letters*, vol 33 issue 15
- [6] Kelley, M.C (2009) *The Earth's Ionosphere, Plasma physics & Electrodynamics*. USA, Academic Press
- [7] Park, J et al (2015) Estimating along-track plasma drift speed from electron density measurements by the three Swarm satellites *Annales Geophysicae*, vol 33 issue 7
- [8] Kabin, K et al (2004) Open-closed field line boundary position: A parametric study using an MHD model *Journal of geophysical research*, vol 109 issue B5
- [9] Miloch, W.J et al (2018) Ionospheric plasma irregularities studied with Swarm satellites *E3S Web of conferences*, vol 62
- [10] Shepherd, S.G (2014) Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations *Journal of Geophysical Research*, vol 119 issue 9
- [11] Tsunoda, R. T (1988) *Reviews of Geophysics*, vol 26 issue 4
- [12] Spicher, A et al (2015) Observation of polar cap patches and calculation of gradient drift instability growth times: A Swarm case study *Geophysical Research Letters*, vol 42 issue 2
- [13] Rich, F.R & Hairston, M (1994) Large-scale convection patterns observed by DMSP *Journal of Geophysical Research*, vol 99 issue A3
- [14] Elkins, T.J & Papagiannis M.D (1969) Measurement and interpretation of Power Spectrums of Ionospheric Scintillation at a Sub-Auroral Location *Journal of Geophysical Research*, vol 74 issue 16
- [15] Dyson, P.L et al (1974) In Situ Measurements of the Spectral Characteristics of F Region Ionospheric Irregularities *Journal of Geophysical Research*, vol 79 issue 10
- [16] Wernik, A.W et al (2003) Ionospheric Irregularities and Scintillation *Advances in Space Research*, vol 31 issue 4
- [17] Jin, Y et al (2014) GPS scintillation effects associated with polar cap patches and substorm auroral activity: direct comparison *Journal of Space Weather and Space Climate*, vol 4 issue article A23
- [18] Oksavik, K et al (2012) In situ measurements of plasma irregularity growth in the cusp ionosphere *Journal of Geophysical Research*, vol 117 issue A11
- [19] De Michelis, P et al (2021) Looking for a proxy of the ionospheric turbulence with Swarm data *Scientific Reports*, vol 11 article 6183

- [20] Ritter, P et al (2013) Determining field-aligned currents with the Swarm constellation mission *Earth, Planets and Space*, vol 65 article 9
- [21] Stone, E et al (1998) The Advanced Composition Explorer *Space Science Reviews*, vol 86 page 1-22
- [22] Petrovski, I & Tuszii, T (2012) *Digital Satellite Navigation and Geophysics: A Practical Guide with GNSS Signal Simulator and Receiver Laboratory*. Cambridge: Cambridge University Press.
- [23] Labelle, J & Kintner, P.M (1989) The Measurement of Wavelength in Space Plasmas *Reviews of Geophysics*, vol 27 issue 4
- [24] Friis-Christensen, E (2006) Swarm: A constellation to study the Earth's magnetic field *Earth Planets Space*, vol 58 supplement 4