Geomagnetic activity and its sources during modern solar maximum

Reko Hynönen
2013

Supervisor: Eija Tanskanen
Examiners: Eija Tanskanen
Hannu Koskinen

UNIVERSITY OF HELSINKI
DEPARTMENT OF PHYSICS

P.O. Box 64 Gustaf Hällströmin katu 2
FI-00014 University of Helsinki
Geomagnetic activity and its sources during modern solar maximum

Geomagnetic storms and auroral substorms are manifestations of space weather. They are disturbances in the geomagnetic field caused by solar activity that consists of flare eruptions, coronal mass ejections, high-speed streams, corotating interaction regions and other disturbances within the solar wind. The occurrence rate and properties of these events vary greatly within the solar cycle some maximizing during the solar maximum and others during the declining phase of the solar cycle.

Any solar activity measure can be used to define the solar cycle though traditionally the sunspot number has been used. In addition to the sunspot number we have examined e.g. sunspot area, solar radio flux and solar X-ray flux. The solar cycle itself can be divided into four distinct phases: ascending, maximum, declining and minimum phases. Their properties depend on the solar activity measure they are based on.

Occurrence rates of geomagnetic storms, substorms and events of solar origin along with geomagnetic indices show that the most recent solar cycle, number 23, had its most disturbed time interval in 2003 in its declining phase. Though solar flares and CMEs were found to maximize in the solar maximum as expected, the slow CMEs and coronal hole originated structures like high-speed streams were found to maximize during the declining phase of the solar cycle. The same conclusion was confirmed studying the geomagnetic storm indices and the ultra-low frequency (ULF) fluctuations within solar wind and the magnetosphere, identified with the method of power spectra.

Ground-based Pc5 pulsations from three magnetic stations (KEV, OUJ and KIL) were identified and two maxima were found: the largest one in the declining phase of the solar cycle and the other one during the solar maximum. The ground Pc5 pulsations during the solar cycle 23 follow nicely the ULFs identified based on the ACE satellite measurements at the L1.

Records of storm indices show that the declining phase has been the most disturbed time interval in majority of the solar cycles during the modern solar maximum, not only during solar cycle 23.

Keywords: space weather, solar activity, geomagnetic pulsation
# Contents

Acknowledgements

1 Introduction 1
   1.1 Outline of the thesis  2

2 Background: Geomagnetic disturbances and their causes  4
   2.1 Solar activity  4
      2.1.1 Sunspots  5
      2.1.2 Solar flares  6
      2.1.3 Measures of solar activity  7
      2.1.4 Phases of the sunspot cycle  12
      2.1.5 Defining solar cycle and phases  13
   2.2 Solar wind and its disturbances  15
      2.2.1 Solar wind  15
      2.2.2 Corotating interaction regions  18
      2.2.3 Interplanetary coronal mass ejections  19
      2.2.4 High-speed solar wind streams  21
      2.2.5 Galactic cosmic rays  21
   2.3 Magnetic environment of Earth  23
      2.3.1 Magnetic field of Earth  24
      2.3.2 Magnetosphere  24
      2.3.3 Ring current  26
      2.3.4 Ionosphere  27
   2.4 Geomagnetic activity  27
      2.4.1 Geomagnetic storms  28
      2.4.2 Auroral substorms  29
      2.4.3 Geomagnetic pulsations  30
      2.4.4 Magnetic indices  31

3 Data analysis methods  37
   3.1 Statistical methods  37
      3.1.1 Mean and median  37
3.1.2 Variance and standard deviation .......................... 38
3.1.3 Skewness ................................................. 38
3.1.4 Kurtosis ................................................. 39
3.1.5 Correlation and regression analysis ......................... 40
3.1.6 MATLAB implementation of running average ............... 42
3.1.7 Interpolation of indefinite values .......................... 43
3.2 Fourier analysis ............................................. 43
  3.2.1 Discrete Fourier transform and its limitations ............ 44
  3.2.2 Power spectrum ......................................... 46
  3.2.3 Convolution theorem ..................................... 46
  3.2.4 Improving convergence of Fourier series .................. 46
  3.2.5 Filtering-squaring-averaging method of powers spectra .... 48
  3.2.6 FFT with MATLAB signal processing toolbox ............. 49
  3.2.7 Using Fourier analysis in wave detection ................. 49

4 Data analysis of the space weather activity measures 52
  4.1 Measures of solar activity ................................ 52
    4.1.1 International sunspot number .......................... 52
    4.1.2 Other sunspot numbers ................................ 53
    4.1.3 Sunspot area index .................................... 54
    4.1.4 Solar radio flux ....................................... 58
    4.1.5 Solar X-ray flux ....................................... 59
    4.1.6 Comparison of solar activity measures .................. 61
  4.2 Measures of geomagnetic activity .......................... 64
    4.2.1 Ap and aa indices ..................................... 64
    4.2.2 Dst, Dcx and Dxt indices .............................. 65
    4.2.3 Auroral electrojet indices ............................. 69
    4.2.4 Polar cap indices ..................................... 73

5 Analysis of the space weather events of solar cycle 23 75
  5.1 Data and its sources ...................................... 75
  5.2 Identifying fluctuations .................................... 76
  5.3 Solar disturbances ........................................ 77
    5.3.1 Solar flares .......................................... 77
    5.3.2 Coronal mass ejections ................................ 78
  5.4 Solar wind disturbances .................................... 79
    5.4.1 Interplanetary coronal mass ejections ................. 79
    5.4.2 High-speed streams ..................................... 81
    5.4.3 Corotating interaction regions ......................... 82
  5.5 Magnetospheric events ..................................... 82
    5.5.1 Geomagnetic storms .................................... 82
Acknowledgements

This thesis was a work of many months. First I would like to thank my supervisor, Dr. Eija Tanskanen, for all the help and tips I got while working on the thesis. I would also like to thank my other examiner, Prof. Hannu Koskinen for his advice.

Of course, many other people helped me along the way. For their advice and help I’d like to thank Kristian Snekvik on his advice on Fourier analysis and tapering functions. David Pérez-Suárez gave me valuable advice on formatting the figures. Also, I’d like to thank MATLAB for being such a great tool.

I have met some great people during my studies. Many of them I met through Resonanssi and Limes student organizations, and several others. Thank you all. However, with some of them I have shared my concerns on graduation and thesis writing, and for these talks and help I’d like to thank Olli Vartia, Jouni Kajasto, Erik Larismaa, Terhi Järvinen, Heidi Yli-Kankahila, Sanni Hoilijoki, Risto Karinkanta, Daniel Landau, Mikko Niemi, Tiina Lepistö, Pekka Rantala, Katri Leinonen, Sanna Särkikoski, Tommi Tenkanen, Jarkko Järvelä, Jari Rinta-aho, Olli Wilkman, Harri Waltari and Juho Saarinen.

Also, thanks to all who went to LiStraKe board game club when I ran it, and for Resonanssi LAN parties – those game nights were awesome. Hopefully, there will always be more.

Last, I’d like to thank my family and my wife Jenni Kaarela for always being there for me.
Chapter 1

Introduction

The northern lights have been a seed for stories and legends long before their physical origin was discovered (Holzworth, 1975, see e.g.). They are the visual manifestation of auroral substorms that are one of the two main appearances of solar storms. The other type is a geomagnetic storm, a global disturbance of the geomagnetic field.

It is still not known in detail how either of these special storms are generated in the solar convection zone. However, it became clear after the Carrington flare of summer 1859 (Carrington, 1860; Hodgson, 1860) and the subsequent geomagnetic superstorm that solar phenomena have a key role in the matter, even though Richard Carrington, one of the initial observers, himself noted first "One swallow does not make a summer" (Carrington, 1860).

The connection between solar phenomena and the space storms established, the next step was finding the mechanisms that drive the space storms. Even though the exact coupling is not fully understood, many individual parts of the geomagnetic environment and the activities within are nowadays known well as I will show in this thesis. In a world dependent on the constantly increasing number of electronics the hazard of space storms is also skyrocketing meaning that the ability to predict their occurrence has become a field of its own.

In this thesis my aim was to study the statistical properties of the phenomena occurring on the Sun, in the solar wind and within the magnetosphere. Focus was on the phenomena that are known to have influence on the space weather, manifesting itself through the occurrence of geomagnetic storms and auroral substorms. Special attention was given on the time period known as the modern solar maximum: solar cycles 15 to 23. Throughout this interval solar activity and, consequently, the space weather activity has been relatively high. Ground observations cover this period fully, except with the newer activity indices.
Since mere ground observations are spatially very limited measurements higher up in the atmosphere and especially in space are needed for better picture (Peitso, 2013). In the 1960s and 70s the basic structure of the Earth’s magnetosphere was quickly uncovered by several spacecraft of USSR and USA who were in the middle of the space race to the Moon. In particular in the 1990s several new missions introduced continuous measurements to the geospace environment. With the addition of ACE (Advanced Composition Explorer (Stone et al., 1998)) and Wind (Russell, 1995) missions there is enough data to commit detailed statistical analysis of different fluctuations in the geomagnetic environment. In this thesis the fluctuations were inspected during the solar cycle 23 along with the events of both solar origin and geomagnetic environment. For the earlier cycles a less thorough study was done lacking satellite data using a multitude of solar and geomagnetic indices and measures of activity.

1.1 Outline of the thesis

On these pages I attempt to find an answer to the following questions:

- When was the most active phase of the solar cycle 23 in terms of activity of the Sun?
- What about in terms of space weather?
- What can be said about the activity of Sun and space weather during any single cycle in general?

With the analysis in the Chapters 4 and 5 the answer to the first two is found. Because of the limited scope of this thesis but mostly because of the limited data due to the relative juvenility of space age only a rudimentary answer can be given to the third question.

In addition to this introductory chapter the thesis consists of five parts. First, in Chapter 2, the geomagnetic environment and its many different disturbance effects are reviewed starting from their origin, the Sun. The terms in the questions before are given precise definitions as for the measures of activity of the Sun and space weather. The basic properties of many solar-originated phenomena, such as solar wind, coronal mass ejections (CME) and high-speed streams (HSS) are examined and several wave phenomena (especially ULF fluctuations, ULF standing for ultra-low frequency) are considered for their significance for space weather. Solar cycle and its phases are given a precise, albeit statistical, definition. A brief review of the sunspot
cycle is given as well on both generally and specifically on the solar cycle number 23.

Tools to study these phenomena are provided in Chapter 3. The concepts and methods of data analysis, statistical analysis and Fourier analysis are introduced to use in the further chapters.

Statistical data analysis tools of previous chapter are put to use in Chapter 4, where the various measures of solar and geomagnetic activity are considered. The statistical properties of each index over solar cycles are derived and discussed to be later compared with the results on the event data in the next chapter.

The statistical analysis of the space weather disturbances and their frequency over the solar cycle number 23 is conducted in Chapter 5. This regards the data of the number and properties of interplanetary coronal mass ejections (ICME), high-speed streams and some other solar phenomena as well as galactic cosmic rays (GCR) during the cycle 23 and the associated geomagnetic storms and substorms. Also the analysis of the wave propagation in solar wind-magnetosphere-ionosphere system is regarded by looking at the Pc5 fluctuations, i.e. ULF waves. I use Fourier analysis to examine data of magnetic fields from ACE and Wind satellites and from the ground observations and find, which wave phenomena penetrate from the solar wind to the magnetosphere and from there to the ionosphere and when this happens.

Discussion and conclusions are in Chapter 6. Here the implications of the data analysis in Chapters 4 and 5 are discussed. I answer the questions raised in the beginning of this section. Finally, the appendices contain some useful mathematical tools, data and the MATLAB codes used in the thesis.
Chapter 2

Background: Geomagnetic disturbances and their causes

"I was suddenly surprised at the appearance of a very brilliant star of light, much brighter than the Sun’s surface, most dazzling to the protected eye, illuminating the upper edges of the adjacent spots and streaks — not unlike, in effect, the edging of the clouds at sunset."

– Excerpt from R. Hodgson’s publication on the first observation of a solar flare.(Hodgson, 1860)

We will first describe the Sun. It is the primary cause for almost everything concerning space weather on and around the Earth. Even the flux of galactic cosmic rays that originate outside the solar system is at least modulated by its activity (Ahluwalia, 2000)).

Therefore, it is natural to open this chapter with a section on the magnetic activity of the Sun. The journey will continue along solar wind through to the magnetic field of the Earth and respectively its magnetic activity.

2.1 Solar activity

Sun is a star, giant sphere of superheated highly ionized gas, i.e. plasma. Solar matter and most space weather related media consist of plasma that is often dubbed the fourth state of matter. While alien to the everyday experience, plasma is actually the most common state of matter in the solar system and the universe at large. In a plasma the matter is in a highly ionized state, enough to be dominated by any magnetic and electric fields it is exposed to. It consists of nearly equal amounts of negatively and positively charged particles, therefore dubbed quasi-neutral (Boyd and Sanderson, 2003).
Heat of the Sun is replenished by a fusion process deep within its core. Most importantly to our case here, Sun has a magnetic field that is constantly retained by a complicated dynamo process in the star’s convective layer. The dynamo has a cycle of 11 years, over which the magnetic activity peaks and the polarity of the field changes (Stix, 2004).

Magnetic field manifests itself most visually as sunspots. With the space instruments several other magnetically active phenomena, like solar flares and coronal mass ejections, have become common observations as well. In this Section, I give a brief introduction to these phenomena and the solar activity cycle itself and finally a definition to the solar cycle and its phases.

2.1.1 Sunspots

![Figure 2.1: Picture of a group of sunspots showing some key structures. This picture was taken by Hinode’s Solar Optical Telescope (SOT) in visible light on December 13th of 2006. Courtesy of NASA/JAXA. Image source: http://www.nasa.gov/mission_pages/solar-b/solar_022.html](image)

Figure 2.1 shows a picture of the solar surface showing several sunspots. The common structures are seen the clearest in the largest one. The dark central area is called the umbra, while the area looking like fuzzy radially inclined fibres is called the penumbra.
The Sun is bright because its surface is so hot. Against this background, cooler patches like sunspots show up as dark. The umbral area of the spot radiates only 20–30 % of the flux (integrated over wavelength) of the quiet Sun, while the penumbral area radiates 75–85 %, respectively, explaining the relative darkness of these areas (Thomas and Weiss, 2008).

The reason to their relatively low temperatures is their strong magnetic field. The sunspot is cooler than its surroundings because of the strong magnetic field that is suppressing the motion of convective eddies within (Biermann, 1941; Thomas and Weiss, 2008). The convective motions within the sunspot are principally channelled along the magnetic field lines, the transverse motion being very inefficient (Hoyle, 1949).

The average magnetic field within sunspots is 0.12–0.17 T and this does not vary a lot between different sunspots (Solanki, 2002). This observation justifies the use of sunspot number and similar quantities as indices of solar magnetic activity.

### 2.1.2 Solar flares

In 1859 a sudden brightening on the Sun was observed (Carrington, 1860; Hodgson, 1860). These eruptions are now known as solar flares. The Carrington flare was a white light flare, a relatively rare phenomenon of immense energy.

Since their initial discovery the flares have become a more frequent observation. For many decades the only way to detect flares was the direct visual observation particularly using emission in the Hα line that brightens up during solar flares. Atmosphere limits the observations at other wavelengths, so it was not until 1960s that regular measurements of solar X-ray flux began. X-ray flux is even more active than Hα emission during flares (Tandberg-Hanssen and Emslie, 1988).

Flares detected by X-ray spectroscopy can be classified by their intensity in the soft X-ray bands. This corresponds to the wavelength range 1–8 Å (0.1–0.8 nm) of the spectrum. According to the measured flux the flare thus

<table>
<thead>
<tr>
<th>Class</th>
<th>Intensity (erg cm⁻² s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>10⁻⁴</td>
</tr>
<tr>
<td>C</td>
<td>10⁻³</td>
</tr>
<tr>
<td>M</td>
<td>10⁻²</td>
</tr>
<tr>
<td>X</td>
<td>10⁻¹</td>
</tr>
</tbody>
</table>

Table 2.1: Soft X-ray flare classification
detected is assigned a classification of B, C, M or X indicating an order of magnitude and a number representing a multiplier of that magnitude (see Table 2.1, reconstructed from (Tandberg-Hanssen and Emslie, 1988, Section 1.2)).

2.1.3 Measures of solar activity

Sun’s activity can be monitored using several different measures. In this Section some of the most common ones are discussed. The data sources of the ones used in this study are also reported.

International sunspot number

The oldest and longest measure of the Sun’s activity is the number of sunspots. Every activity measure can and often will be compared to the sunspot number. There is a well-established rule that their number follows a cycle of approximately 11 years known as the Schwabe’s cycle, clearly seen in Figure 2.2. As already established the sunspots are magnetically active (Hale, 1908a). It turns out that the polarity of the magnetic field of the Sun and its sunspots switches during the solar cycle (Hale, 1908b), so that the true cycle is actually 22 years known as the Hale’s cycle (Hale and Nicholson, 1938; Babcock, 1961).

![Yearly average of the international sunspot number](image)

Figure 2.2: Sunspot cycle. Data source: (SIDC team, 1818–2013)

International sunspot number, or ISSN, is given by the international sunspot index $R$, formerly known as Wolf sunspot number for the observer
that collected sunspot data and constructed this time series. It is defined as:

\[ R = k(10g + f), \]

(2.1)

where \( f \) is the total number of sunspots, \( g \) is the number of sunspot groups and \( k \) is a normalization constant that depends on the observer (Brekke, 1997, p. 18).

Sunspot cycles have traditionally been numbered from the first full cycle accessible to James Wolf. The first recorded solar cycle, SC1, began approximately in the year 1755. Each 11-year cycle afterwards has its own number.

Sunspot number is used numerous times in this thesis as the measure of the magnetic activity of the Sun. For this purpose the data on the website of SIDC (Solar Influences Data Analysis Center) was accessed (SIDC team, 1818–2013, -> Sunspots -> Data). For years 1749–1817 only the monthly averaged data were available. Afterwards, also daily data were available. In this thesis the data of highest resolution available were used for any time period.

Being partly subjective in nature it can be argued that ISSN is not the best measure of the solar activity. For this purpose several other indices are considered next. The list is not exhaustive, as new activity indices are suggested every once in a while (for a recent example see for example (Lobzin et al., 2011)). However, ISSN is the longest time series of solar activity obtained by direct observations of the Sun. To compare, SAO/NASA Astrophysics Data System (ADS) at adsabs.harvard.edu found 3330 papers with a query on “sunspot number”, while any of the queries on “sunspot area”, “solar x-ray flux”, “solar radio flux” and “flare index” all had less than 1600 combined (as of 2013-04-18). Therefore it is not surprising that all other indicators are usually compared to the international sunspot number and it is considered the standard indicator of the magnetic activity of Sun.

Other sunspot numbers

Even though ISSN is the most common one, it is not the only time series of sunspot number recorded. For completeness several similar numbers are reviewed next.

*Boulder sunspot number* uses the same formula (2.1) as ISSN but incorporates data from different observatories. It is also a shorter time series starting in 1960s or 1970s (according to educational slides of (Biesecker, 2013)).

Another sunspot number is the NOAA *American relative sunspot number*, ARSN (Shapley, 1949). Originally ARSN was constructed because ISSN observations sent from Zürich, where the original ISSN observations were compiled, were constantly delayed to USA. Also, the scientists found that it
would be useful to have another series of the same thing instead of relying on a single time series that is constructed from only a few observations. The American sunspot number $R_A$ is defined

$$R_A = \frac{\sum_{i=1}^{n} w_i k_i R_i}{\sum_{i=1}^{n} w_i},$$

(2.2)

where $w_i$ is the statistical weight of the observer’s reports, $k_i$ is the observatory coefficient as in (2.1) and $R_i = 10g + f$ is the observed sunspot number as in (2.1). Due to statistical weights of different observatories, the individual observatories are more equal with $R_A$ than ISSN that is constructed using data from only a single primary observatory that differs from time to time (Hoyt and Schatten, 1998, Section 3). At present there are $n = 23$ different observatories included in calculating the index (NOAA Solar and Terrestrial Physics Division, 2013, Under point mark 3: American Relative Sunspot Numbers). Relatively short time series, $R_A$ has been constructed since year 1948.

Instead of calculating individual sunspot numbers it is possible to build a group sunspot number. It is designed to be less dependent upon seeing the tiniest spots and less noisy than ISSN, since it is constructed using only the number of sunspot groups instead of both groups and individual spots as its data set. The group sunspot number is defined

$$R_G = \frac{12.08}{N} \sum k_i' G_i,$$

(2.3)

where $G_i$ is the number of sunspot groups recorded by the observer $i$, $k_i'$ is the correction factor of the $i$th observer, $N$ is the number of observers used to get the daily value and 12.08 is a normalization number chosen so that the mean value of $R_G$ between 1874 and 1976 is the same as the mean value of ISSN as (2.1) measured by Royal Greenwich Observatory when they actively made observations (Hoyt and Schatten, 1998). While internally more consistent than ISSN, this time series unfortunately has only been constructed up until year 1995 and it is therefore inadequate for the purposes of studying the sunspot cycle 23 and the ones coming afterwards unless it is updated. Further analysis on group sunspot number would be possible for cycles before that but was omitted here.

ARSN data used here are the courtesy of American Association of Variable Star Observers and Boulder data of NOAA SWPC, both provided by Substorm Zoo http://www.substormzoo.org.
Sunspot area index

It can be argued that instead of counting individual sunspots, we should actually be determining the area of the spots covering solar surface. This is reasonable because solar magnetic flux is strongest at the sunspots. Index like this has been advocated by e.g. (Nagovitsyn, 2005).

The index based on the area is called sunspot area index. SSAI is constructed by summing the corrected areas of all observed sunspots and giving the answer in millionths of solar hemispherical surface area. The data used here was provided by the website of Marshall Space Flight Center, NASA (Hathaway and NASA, 2013).

Solar radio flux

When excluding sunspot observations, the longest direct observations of Sun are in the form of radio fluxes, i.e. measurements of solar radiation flux at scales of several centimetres or decimetres or, equivalently, several thousand megahertz. Line at 10.7 cm (2800 MHz) is the first of radio flux series beginning already in 1946 at Ottawa by the National Research Council of Canada. For this reason it is also known as the Ottawa number. In addition to the Ottawa number there are also 3.2, 8, 15 and 30 cm lines measured at Toyokawa since 1950s (Nicolet and Bossy, 1985). In this study only the band 10.7 cm was used. The data were acquired from National Geophysical Data Center (NOAA Solar–Terrestrial Physics Division, 2013).

The time series of solar radio flux have several advantages over sunspot number indices (Tapping, 1987). First, they are purely quantitative and nonsubjective, thus only prone to errors in calibration and various systematic errors. Second, being in the radio window of the Earth’s atmosphere, most radio wavelengths measured of the Sun have little interference from clouds or atmosphere in general. Therefore, they can be measured under almost any weather conditions, extreme local conditions excluded.

Solar radio fluxes are usually measured in the solar flux units: 1 sfu = $10^{-22}$ Wm$^{-2}$ Hz$^{-1}$. The flux is measured over the whole of solar disc using a radiotelescope. Quiet sun flux of the 10.7 cm line is approximately 50 sfu at average (Oster, 1983).

The sources of the radio fluxes vary. For the Ottawa number (10.7 cm line), the source flux is dominated by thermal free-free emission, i.e. bremsstrahlung, associated not only with sunspots but also hot complexes of activity on the Sun (Tapping and DeTracey, 1990). It mostly originates from the low corona and depends only on few quantities, the most important being the plasma density. This makes it a good independent index along with...
the other radio fluxes. It should be noted that while radio fluxes of different wavelengths correlate strongly with each other, there is still some distinction to each individual wavelength due to slightly different source mechanisms (Nicolet and Bossy, 1985).

Solar X-ray flux

Since the Earth’s atmosphere is mostly impenetrable to X-rays (Karttunen et al., 2010), the era of solar X-ray astronomy began slowly with balloons and rockets. These observations established the formation of the ionospheric layer by the ultraviolet and X-rays from the Sun (Friedman et al., 1951). Continuous observations had to wait until space-based instruments.

X-radiation contains all photons in the energy range 0.1 to 10 keV. Traditionally the X-rays are divided into two categories by energy. The soft X-rays have energies in the range 0.1–10 keV corresponding to wavelengths 10–0.1 nm. The hard X-rays have respectively higher energies 10–100 keV and lower wavelengths 0.1–0.01 nm (Karttunen et al., 2010). There are other conventions for the energy range though, the most notable for this study being the one used by GOES satellites.

The first space-based observations were made by satellites OSO III, OSO IV, OGO I and OGO III (Neupert, 1969) suggesting that solar X-rays originated from the corona. The longest continuous series of X-ray flux data has been collected by the series of the Geostationary Operational Environmental Satellites, GOES. For this study the data from GOES 5 to 12 were used, provided by NOAA GOES satellite data services at the website http://www.ngdc.noaa.gov/going/sem/getData/. This has been collected by the X-ray sensors, XRS, on-board the satellites that observe the disk-integrated X-ray flux of the Sun (Bornmann et al., 1996). The peak flux at 0.1–0.8 nm of this property is used to distinguish between different flares (see Table 2.1). Detecting flares through the X-ray flux works as the first warning of a possibility of CMEs and the associated geomagnetic storms.

Flare indices

Different flare indices can be created from the soft X-ray classification (Table 2.1) and the visual classification (Tandberg-Hanssen and Emslie, 1988, Section 1.2). Just calculating the number of flares inside a class in a given time period gives one index but a hybrid indices could be created by combining these numbers in different ways. In this study the indices constructed directly from the detected number of C, M and X type flares are used. The data were provided by Substorm Zoo http://www.substormzoo.org.
2.1.4 Phases of the sunspot cycle

In Figure 2.3 the sunspot cycle number 23 (SC23) composed of the international sunspot number is shown. In this figure the cycle is divided into distinct phases, which all have some typical phenomena associated with them. SC23 is shown here as an example, as it is also the most studied solar cycle as of present.

![Sunspot cycle graph](image)

Figure 2.3: Sunspot cycle 23 with its phases. Minimum phase is left partly out of the picture, spanning to the right after declining phase. Data source: (SIDC team, 1818–2013)

There are four distinct phases to each solar cycle, namely:

- Ascending phase
- Maximum phase
- Declining phase
- Minimum phase

The solar cycle begins with the ascending phase. This phase is characterized by increasing activity in the form of flares, coronal mass ejections (CMEs) and complexification of the magnetic field structure on the solar surface. SC23 was typical in many ways, e.g. the ascending phase was relatively short when compared to the declining phase, as shown in Chapter 4.

---

1Nomenclature here is varied. Both the terms "declining" and "descending" are used for this phase in the literature. For example Kamide et al. (1998) uses "declining phase" while Gonzalez et al. (1999) used "descending phase". Doing a query on the number of papers using these terms in scholar.google.com showed (as of August 2013) that "declining phase" is about twice as common as "descending phase". Because of this trend I have decided to use "declining phase" in this thesis.
The greatest number of CMEs was detected around the maximum of SC23. The CME occurrence rate correlates quite well with the sunspot number (see Section 5.3.2). Like a typical cycle, SC23 had a two-peaked maximum in the smoothed solar indices as shown in Section 4.1, a feature of many solar cycles (Gnevyshev, 1967).

In the declining phase the solar activity is decreasing until it reaches a solar minimum. Generally the declining phase is longer than the ascending phase. However, the SC23 was atypical with its abnormally long declining phase and minimum that defied any statistical or physical predictions. It has been suggested that Sun changed its behaviour during SC23 based on the change of correlation between sunspot number and 10.7 cm radio flux (Tapping and Valdés, 2011).

Minimum phase is associated with minimally suppressed bombardment of the Earth by galactic cosmic rays (GCRs) of high energy. During the more active phases of the Sun, the solar wind is modulating the GCR flux effectively suppressing a bulk of it while adding in its own cosmic rays (Ahluwalia and Dessler, 1962; Ahluwalia, 2003). As noted, the minimum phase of SC23 was abnormally long, SC24 beginning only in 2010.

The most intense geomagnetic storms of the cycle 23 happened in the October and November of 2003, dubbed Halloween storm 2003. This was placed in the declining phase of the cycle, contrary to the anticipation that the most intense geomagnetic activity would happen in the maximum phase of the solar cycle (Weaver et al., 2004).

2.1.5 Defining solar cycle and phases

To analyze the cycle and its phases further they need a definition.

Definition of a solar cycle: Calculate the running mean of the measure over a year using a resolution of 1 week, or 1/52 year. This is called the smoothed index. Then a cycle is a period between two minima of this new number. Local minima are excluded from this if the index is lower than the minimum somewhere in its vicinity, chosen to be within three years of the minimum.

This is a rather common way to define the minima between cycles. Almost the same method has been used in the historical records according to (Harvey and White, 1999). More elaborate measures have also been used defining the minimum using three additional parameters: the minimum of the monthly averaged sunspot number, the total number of active regions as well as the number of new and old active regions, and the number of spotless days (Harvey and White, 1999). Here, the proposed definition is used individually for each measure of solar activity and for simplicity any mutual
connections between them are disregarded.

Classifying the phases is a bit more tricky. I have not seen any rigorous attempt in the literature to do this before, though some papers (Bogdan et al., 1988, see e.g.) clearly use some hidden criteria. My definition here is rather arbitrary but lacking any standing definition it must suffice for this study. First the average and the standard deviation of the smoothed index are calculated.

- **Definition of the ascending phase:** This is the period situated between the points of time when the index is between one standard deviation of the average on the farther left side of the cycle.

- **Definition of the declining phase:** Like the ascending phase, except the farther right side of the cycle is picked.

- **Definition of the maximum phase:** Maximum is the period of time between ascending and declining phases.

- **Definition of the minimum phase:** Minimum is the period of time between the declining and ascending phases of two consecutive cycles.

This definition has some disadvantages. First of all, it cannot be used precisely for an ongoing cycle: the whole of cycle is needed for the full determination of phases unless extrapolation or predictive techniques are used. Especially the minimum period cannot be determined precisely before both the cycles before and after the cycle in question have been observed. The existence of minimum phase is also questionable for certain cycles or indices, as the new ascending phase could have begun so soon after the declining phase that there was no true minimum phase for a given cycle according to the definition. The usage of smoothed index is also arbitrary and for full data of a cycle an additional data set of at least half a year both before and after the cycle are required because of it. Using standard deviation is also rather arbitrary, because one could also use e.g. its multiples or the median in its place.

Advantages are nevertheless obvious. With a definition the phases can be compared within different indices. Most importantly the definition gives a precise period for each solar cycle and phase in regard to each index, and no observer is required to decide where to place the phases, so long as the measure is rigorous for the studying of the solar cycle.
2.2 Solar wind and its disturbances

Both the most powerful and numerous disturbances of the geomagnetic field are caused by the phenomena of the Sun. These include but are not strictly limited to interplanetary coronal mass ejections (ICME), shocks, high-speed solar wind streams (HSS, high-speed stream), corotating interaction regions (CIR), various magnetic field and plasma fluctuations and of course the solar wind itself. Only galactic cosmic rays (GCR) are associated with the outside of solar system and while their flux is modulated by the solar activity, Sun furthermore produces its own share of cosmic rays called solar energetic particles (SEP). The exact coupling between the solar wind phenomena and space weather remains unknown (Kamide et al., 1998, for review see).

2.2.1 Solar wind

Sun is constantly emitting a flow of matter in a highly ionized state, plasma. Embedded within is the solar magnetic field, which becomes interplanetary magnetic field (IMF) when observed in the interplanetary space by the magnetometric instruments of spacecraft. This particle flow is called the solar wind. It shoots out of the star because of the huge difference in the gas pressure between the solar corona and the interplanetary space (Kivelson and Russell, 1995, Chapter 4 by A. J. Hundhausen).

The solar wind consists mainly (∼95%) of protons and electrons and of a small number of helium and even smaller number of heavier elements (Kivelson and Russell, 1995, Chapter 4 by A. J. Hundhausen). The speed of solar wind ranges from anywhere as low as 200 km/s to higher than 1000 km/s.

Being an extension of the corona, the solar wind is formed in at least two different regions. The slow but dense solar wind is formed in the regions of closed magnetic field lines by presently unknown mechanism (for a review: Schwenn, 2007) while the fast but sparse solar wind, which evolves into high-speed streams, originates from the coronal holes that are regions of open field lines (e.g. Krieger et al., 1973). This is an important distinction between the two types of solar wind.

Alfvén theorem of frozen-in flux

To understand the dynamics of the solar wind and Sun’s ejecta in general it is important to know of the intrinsic connection of plasma to the solar magnetic field. Hannes Alfvén showed in (Alfvén, 1942) that given perfectly conducting plasma, the magnetic field lines are frozen-in the plasma or, equivalently,
the matter is forced to follow the lines of force. Originally Alfvén said the plasma was "fastened" to the magnetic field lines but since then the scientific community has found "frozen" to describe the phenomenon more clearly instead.

In the following I will prove the frozen-in flux theorem owing the presentation to (Boyd and Sanderson, 2003). The main equation of the theorem is the induction equation of magnetohydrodynamics, MHD, given

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\sigma \mu_0} \nabla^2 \mathbf{B} + \nabla \times (\mathbf{u} \times \mathbf{B}), \quad (2.4)$$

where \( \mathbf{B} \) is the magnetic field vector, \( \mathbf{u} \) is the velocity vector, \( t \) is time, \( \sigma \) is the conductivity of the plasma and \( \mu_0 \) is the permeability of vacuum. See the Appendix A for a brief introduction to the theory of magnetohydrodynamics.

In the case of ideal MHD the conductivity is assumed to be infinite, i.e. the first term in (2.4) is neglected and the resulting equation is

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}). \quad (2.5)$$

Alfvén’s theorem states that if the equation (2.5) is true, then the flux of magnetic field through any surface \( S \) bounded by a closed contour \( C \) moving with the fluid is constant. Proceeding to calculate the time derivative of the magnetic flux we get using the three-dimensional version of Leibniz’ integral rule (see Appendix B.2)

$$\frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{S} = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} + \oint_C \mathbf{B} \cdot \mathbf{v}_C \times d\mathbf{l}, \quad (2.6)$$

where \( d\mathbf{l} \) is a differential vector along the contour \( C \). The second term of RHS in the Leibniz’ rule disappears due to relation \( \nabla \cdot \mathbf{B} = 0 \) and the third term is found by using the rule of scalar triple product \( - (\mathbf{v} \times \mathbf{F}) \cdot d\mathbf{l} = \mathbf{F} \cdot (\mathbf{v} \times d\mathbf{l}) \).

Switching into the coordinates of the moving plasma we have \( \frac{d}{dt} \rightarrow \frac{D}{Dt} \) and \( \mathbf{v}_C \rightarrow \mathbf{u} \), resulting in

$$\frac{D}{Dt} \int_S \mathbf{B} \cdot d\mathbf{S} = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} + \oint_C \mathbf{B} \cdot \mathbf{u} \times d\mathbf{l}. \quad (2.7)$$

Here the first term of the right-hand side represents the change of flux due to the change of magnetic field in time and the second term represents the change of the surface area due to the movement of the bounding contour \( C \). Using the relation \( \mathbf{B} \times \mathbf{u} = -\mathbf{u} \times \mathbf{B} \) and the Stokes’ theorem (see Appendix B.1) Equation (2.7) becomes

$$\frac{D}{Dt} \int_S \mathbf{B} \cdot d\mathbf{S} = \int_S \left( \frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u} \times \mathbf{B}) \right) \cdot d\mathbf{S} = 0 \quad (2.8)$$
where the last equality stems from Equation (2.5). This completes the proof.

The frozen-in flux holds well in many dynamic phenomena of solar wind, failing only at the discontinuities like shocks and magnetospheric boundaries. It dictates the global shape of the solar wind system in the interplanetary space. Ultimately it dictates whether a given blob of plasma, say a coronal mass ejection, will have a trajectory towards the Earth.

**Global shape of the solar wind**

In 1958 in his famous paper Parker showed how solar wind would behave as an extension of the corona (Parker, 1958). He also showed that due to the magnetic field embedded within the plasma its global shape would be like an Archimedean spiral, shown in Figure 2.4, presently dubbed Parker spiral to honour the scientist who predicted it.

![Solar Wind](http://commons.wikimedia.org/wiki/File:Parker_spiral.png)

**Figure 2.4:** Global shape of the solar wind, i.e. the Parker spiral. Source: [http://commons.wikimedia.org/wiki/File:Parker_spiral.png](http://commons.wikimedia.org/wiki/File:Parker_spiral.png), originally from NASA, [http://sdo.gsfc.nasa.gov/mission/spaceweather.php](http://sdo.gsfc.nasa.gov/mission/spaceweather.php).
Solar wind plasma follows the magnetic field that is still tied to its solar base. Since the Sun is rotating, the plasma follows the rotation ultimately forming the spiral seen in Figure 2.4. Figure also shows how different initial velocities affect the formation of the spiral: the faster the plasma, the less it is affected by the rotation. Ultimately, the spiral depicts the path any plasma will take towards Earth helping in predicting any incoming magnetic disturbances.

### 2.2.2 Corotating interaction regions

Figure 2.5 shows an illustration of a corotating interaction region, CIR, forming. It is the result of the interaction of fast solar wind with slower solar wind ahead of it (Heber et al., 1999, see e.g.). Should the structure stay stable for several rotations it is said to be corotating, thus the name of the feature. CIRs normally occur in the low and middle latitudes in the heliosphere. Since fast solar wind originates from the coronal holes they often repeat about every 27 days, the rotation rate of the Sun, as the coronal hole faces the same direction once again. This behaviour is illustrated by the simulation in Figure 2.6 that shows how different CIRs develop farther downstream in solar wind (Akasofu and Hakamada, 1983).

![Figure 2.5: Schematic of the formation of corotating interaction region. Image source: (Pizzo, 1978)](image-url)
Shocks may form within the region depending on the differences in velocity and plasma pressure of the two streams. This normally occurs at a distance of $>\sim 1.5$ AU, also shown in Figure 2.6. These can function as particle acceleration regions that can have effects on the space weather at Earth. A decrease of cosmic ray intensity during several magnetic storms is observed (Forbush, 1938), which is not attributed to the atmosphere, a phenomenon now known as Forbush decrease. The Forbush decreases are due to interplanetary conditions like CMEs or CIRs refracting cosmic rays (Lockwood, 1971). CIRs may also induce geomagnetic storming and pose a threat to space-based instruments and people due to unsteady particle effects (Borovsky and Denton, 2006).

2.2.3 Interplanetary coronal mass ejections

Sun is constantly erupting material out into the interplanetary space but the most spectacular of these eruptions are the coronal mass ejections (CMEs). They were first detected in the coronagraph pictures of Sun like the one
pictured in Figure 2.7, an observation that also gave then their name. They consist of plasma embedded with solar magnetic field and generally have speeds varying from as few as 100 km/s to as rapid as 2000 km/s or more (Ivanov and Obridko, 2001, see e.g.). Generally the production rate of CMEs is the highest in the solar maximum with several eruptions per day with higher velocities and larger amounts of matter (Mullan, 2010). It has been thought that CMEs are main drivers of geomagnetic storms (Kamide et al., 1998).


Figure 2.7: LASCO C2 image showing a very large coronal mass ejection (CME) on 2.12.2002. Source: http://sohowww.nascom.nasa.gov/gallery/images/20021202c2cme.html, downloaded on 7.8.2013.

When a CME is detected in the interplanetary space, it is called interplanetary CME (ICME). Given a suitable trajectory they sometimes hit the magnetosphere of the Earth. The hit does not need to be head on to be remarkable, for even a glancing blow can induce magnetic storming. Upon impact, a shock is formed in the front of the magnetosphere and the magnetopause is pushed closer to the Earth due to solar wind pressure. The hit may cause storms or substorms but the mechanism by which this happens is unclear. However it is known that ICMEs with southward component of the magnetic field seem to drive the storm-activity more efficiently (Kamide et al., 1998).

The strongest storms are often associated with several CMEs hitting magnetosphere in succession. This results in several storms of increasing intensity, since the first CME has already cleaned the solar wind of stray particles. The CME coming behind will then have a relatively free path to traverse result-
ing in faster CMEs. One such an event was the series of storms dubbed Halloween storm 2003 (Weaver et al., 2004). Two of the strongest storms (by Dst index) of solar cycle 23 occurred during this period. Similar case was the Carrington storm of 1859 when supposedly a flare-induced ICME hit the magnetosphere in short succession of another and produced the strongest geomagnetic storm since the invention of telegraph (Tsurutani et al., 2003; Green and Boardsen, 2006).

2.2.4 High-speed solar wind streams

High-speed streams, or HSS, are like bursts of wind but occur within the plasma of the solar wind. They can last from a few hours to several days or even weeks and as they carry the magnetic field of the Sun with them, they can drive the magnetospheric activity for relatively long periods. Like ICMEs, high-speed streams are more geoefficient when their magnetic fields have a southward component.

HSS event can be defined in several different ways, the most common ones being a minimum speed threshold or an minimum increase of speed over some period of time. The data I used in this thesis (Maris and Maris, 2013) has been acquired by looking for a speed increase of at least 100 km/s over one day and requiring this increase to last at least 2 days, a definition also used by (Lindblad and Lundstedt, 1981, 1983; Lindblad et al., 1989). Obviously a definition like this would miss any HSS that lasted shorter than 2 days.

Due to their long length of relatively steady conditions, it is possible that HSSs are more important to the space weather than other interplanetary phenomena at large. This is regarded in Chapter 5.

2.2.5 Galactic cosmic rays

Cosmic rays are high-energy particles that fill the space. They were found already in 1912 with balloon flights (Hess, 1912). Since then a large number of cosmic ray observatories have been established around the world. World Data Center for Cosmic Rays at http://center.stelab.nagoya-u.ac.jp/WDCCR/ was established in 1957 to collect and distribute this enormous set of data.

Cosmic rays can be divided into two classes by their origin. The Sun produces solar energetic particles, i.e. SEP that are most prominent in the maximum phase of the Sun’s magnetic cycle. The galactic cosmic rays (GCRs) originate from the outside of the solar system.

E. Cliwer has collected the the history of SEP research (Gopalswamy and Webb, 2009). The first SEP events were recorded by Lange and Forbush
(Lange and Forbush, 1942b,a) and finally interpreted as originating from solar flares (Forbush, 1946) with the help of additional observations in 1946, deemed the official beginning of SEP research (Gopalswamy and Webb, 2009).

![Energy Spectrum of Galactic Cosmic Rays](image)

**Figure 2.8:** Energy spectrum of galactic cosmic rays as a log Energy vs. log Flux. Adapted from (Nagano and Watson, 2000)

GCRs originate from the outside of the solar system. They are composed mostly (98%) of nuclei, including all from hydrogen to actinides, as well as electrons and positrons (2%) (Simpson, 1983). Of the nuclei part, 87% are hydrogen, 12% helium and the rest are heavier nuclei. Their energy spectrum (Figure 2.8) shows that their energy is an inverse power law and there are also ultra-high energy particles of $10^{20}$ eV. There is a compelling evidence that the bulk of GCRs with energies $\leq 10^{15}$ eV must be continually renewed in the galaxy lest their flux decrease in time and eventually cease to exist. Processes capable of this are believed to exist in the shock waves of expanding supernova remains though this has not been verified (Ginzburg
and Berezinskii, 1990; Hillas, 2005). Decrease has not been observed in the studies of the record of cosmic ray bombardment in meteorites (Forman and Schaeffer, 1979) nor in the Be-10 quantity in deep sea sediments (Inoue and Tanaka, 1979).

The existence of ultra-high energy GCRs are a mystery to the cosmic ray research, as due to the Greisen-Zatsepin-Kuzmin limit (Greisen, 1966; Zatsepin and Kuz'min, 1966) particles with energy over $5 \times 10^{19}$ eV would begin to interact with the cosmic microwave background and therefore they could not originate outside the galaxy unless the laws of physics are different for ultra-high energy particles. However, no one has been able to pin-point their origin.

Being charged particles, cosmic rays are modulated by the solar magnetic field through the solar wind resulting in a 11-year cycle (Forbush, 1954; Ahluwalia, 2000). As a consequence their route to the Earth is not a straight line but instead they are scattered by the magnetic field. Charged particles propagating through the outward-directed solar wind flow tend to be convected outward and also undergo adiabatic deceleration owing to the outward expansion of the magnetic fields carried by the solar wind (Parker, 1966). This way the IMF is protecting the Earth from the high-energy particles by removing part (about one third) of their initial energy.

### 2.3 Magnetic environment of Earth

Magnetic field of the Earth is the most important shield against the violent eruptions of the Sun. It also provides shielding against galactic cosmic rays. When solar wind interacts with the geomagnetic field, a cavity called the magnetosphere is formed. While this field is roughly dipolar, the large-scale structure of the field resembles more of a comet with a tail.

Several large current systems are formed within the magnetosphere. These constitute an essential part in the formation of geomagnetic storms. The auroral substorms on the other hand are associated with polar currents of ionospheric origin. Ionosphere forms when light from Sun ionizes the upper atmosphere of a planet. Ionization, along with the magnetic field, induces its own current systems in the ionosphere.

Here the basic properties of the magnetosphere and ionosphere systems are reviewed briefly along with the relevant physics.
2.3.1 Magnetic field of Earth

The Earth itself is an enormous magnet (Gilbert, 1958). At a large scale it is a dipolar magnet with a potential

$$V = \frac{\mu_0 m \cos \theta}{4\pi r^2},$$

(2.9)

where $\mu_0$ is the magnetic permeability of vacuum, $r$ is the distance from the center of the Earth and $m$ is the magnetic moment of the dipole (Lowrie, 1997, Section 5.2.3).

Seismic studies have revealed that the Earth has a solid inner core and outer fluid core. It is likely that the dipole field is produced by a dynamo process in the outer fluid core where the conducting fluid produces the magnetic field through convection. Due to magnetic minerals and inhomogeneities of the internal currents within the crust of the Earth the magnetic field is, nevertheless, not exactly dipolar but exhibits higher orders of structure expressed by the series expansion

$$V = R \sum_{n=1}^{\infty} \sum_{l=0}^{n} \left( \frac{R}{r} \right)^{n+1} \left( g_n^l \cos l\phi + h_n^l \sin l\phi \right) P_l^n(\cos \theta),$$

(2.10)

where $R$ is the radius of the Earth, $P_l^n(\cos \theta)$ are Schmidt polynomials and $g_n^l$ and $h_n^l$ are the associated Gauss coefficients (Lowrie, 1997, Section 5.4.4). Equation (2.10) is the multipole expression of the geomagnetic potential. It contains only the part of the potential intrinsic to the Earth. Space physics is mainly interested in the external part that originates from the Sun.

The dipole portion $n = 1$ of Equation (2.10) is the most important one. Equaling (2.10) with $n = 1$ and Equation (2.9) one also finds for the dipole magnetic potential

$$m = \frac{4\pi}{\mu_0} R^3 g_1^0,$$

(2.11)

where $g_1^0$ is the strongest component of the field, associated with the Earth’s rotation axis.

International Geomagnetic Reference Field, IGRF, is composed every year using the Equation (2.10) truncated to include coefficients up to $n = 10$. Removing the value given by IGRF the secular variation and other disturbances may be found, as well as the field of external origin (Lowrie, 1997, Section 5.4.4).

2.3.2 Magnetosphere

If the space was empty, Earth’s magnetic field would in theory reach an infinite distance only restricted by the speed of light and the age of the
planet. Obviously this is not the case since the Sun constantly emits charged particles in the form of the solar wind exhibiting its own magnetic field frozen within. The extent to which Earth’s magnetic field reaches is called the magnetosphere. Inside this region the magnetic field originating inside the Earth is the dominating magnetic field, although it is severely distorted by the solar wind.

In Figure 2.9 the structure of the Earth’s magnetosphere is presented. The speed of solar wind is supersonic and super-Alfvénic at 1 AU, meaning that the speed exceeds the speeds of both pressural and magnetic signals (Baumjohann and Treumann, 1997). Then a bow shock is formed in the front of the magnetosphere that is considered a hard sphere in regard to the solar wind. The bow shock slows down, heats and divides the solar wind into two components, one of which goes upstream around the magnetosphere and the one of which goes through the bowshock forming a region called magnetosheath.

Figure 2.9: Magnetosphere of the Earth shown from a direction perpendicular to the Sun–Earth line. Data source: Wikimedia Commons commons.wikimedia.org/wiki/File:Structure_of_the_magnetosphere.svg

The boundary between magnetosphere and solar wind is called magnetopause. The location of this region is determined primarily by pressure balance. Here the total pressure of the solar wind, i.e. the sum of magnetic and plasma pressures equals the magnetic and plasma pressures inside the magnetosphere. Therefore it represents the boundary to which the Earth’s
magnetic field reigns, since beyond this boundary the interplanetary magnetic field is the dominating field.

2.3.3 Ring current

Magnetic storms, as discussed in Section 2.4.1, are the result of the enhancement of the westward ring current in the equatorial region. The ring current itself results from the magnetic drift in the equator affecting charged particles there. A brief description of its physical origin is given below.

Magnetic drift

Ignoring the electric fields, a charged particle will experience a purely azimuthal magnetic drift, average velocity of which is given by

\[
\langle v_d \rangle \approx \frac{6L^2W}{qB_ER_E}(0.35 + 0.15\sin\alpha_{eq})
\] (2.12)

where \( L \) is a pure number \( L = r/R_E \), \( W \) the particle energy, \( R_E = 6370 \) km the Earth radius, \( B_E \) the dipolar magnetic field at the equator, \( q \) the particle charge and \( \alpha_{eq} \) the equatorial pitch angle (Baumjohann and Treumann, 1997). Given a suitable pitch angle, the particle is trapped in a trajectory between two latitudes, mirror points. The pitch angle is given by

\[
\sin^2\alpha_{eq} = \frac{\cos^6\lambda_m}{(1 + 3\sin^2\lambda_m)^{1/2}}.
\] (2.13)

where \( \lambda_m \) is the magnetic latitude of the mirror point. Despite being trapped, magnetic drift moves it azimuthally resulting in a westward net current.

Equatorial magnetic drift

Assuming \( \alpha_{eq} = 90^\circ \), i.e. \( \lambda_m = 0^\circ \) and using Equation (2.12) results in the equatorial magnetic drift velocity

\[
\langle v_d \rangle \approx \frac{3L^2W}{qB_ER_E},
\] (2.14)

or as a current of particles of charge \( q \) and density \( n \),

\[
\langle j_d \rangle \approx \frac{3nL^2W}{B_ER_E}.
\] (2.15)

This is the ring current: every charged particle contributes a tiny part of the ring current, so that when there are lots of charged particles input into the magnetosphere, the ring current is enhanced. When this happens, the time interval is called a geomagnetic storm.
2.3.4 Ionosphere

The ionized upper layer of the Earth’s atmosphere is called the ionosphere. Any planetary object with an atmosphere has one, ruling out barren planets like Mercury and most moons but not all cometary objects. Ultraviolet and higher energy rays from the Sun are absorbed by the planetary atmosphere exciting its molecules and creating conductive layers of fully or at least partly ionized gases.

Given time, the ionized molecules will recombine with electrons so the ionosphere has to be maintained constantly. However, ionosphere is maintained even in the high latitudes where the Sun’s rays have reduced effect on the formation of the ionosphere and in the night-time when the photoionization ceases altogether. The other component to the ionization is provided by the energetic particles of the magnetosphere, mainly electrons but also cosmic rays of both solar and interstellar origin. Being mostly charged particles they dominate the ionization of the high latitude ionosphere (Baumjohann and Treumann, 1997).

The auroral regions near the polar caps is the region where the substorms manifest themselves by disturbing the magnetic field through the auroral electrojets. The region is also the home to the northern lights, another manifestation of the substorms.

Auroral electrojets

The magnetic disturbances associated with the substorms are caused by strong ionospheric currents flowing within the auroral belt (Kamide, 1988). These currents are called auroral electrojets.

Generally the currents are directed towards night: they are eastward in the evening sector and westward in the morning sector. However, their configuration can be very dynamic despite these guidelines. Temporal changes in the auroral electrojets directed either eastward or westward can be monitored using the auroral electrojet (AE) indices (Davis and Sugiura, 1966). For further discussion see Section 2.4.4.

2.4 Geomagnetic activity

The geomagnetic environment is commonly affected by two kinds of major disturbances, namely the geomagnetic storms and auroral substorms. The naming convention here is very diverse for historical reasons, the same phenomena having multitudes of sometimes confusing names. In this thesis I
am using the term geomagnetic storm (or just storm) to call a major enhancement in westward ring current around the Earth. Respectively, auroral substorms (or just substorms) are the magnetic fluctuations happening within the auroral ovals. Both disturbances are commonly associated with the southward component of the interplanetary magnetic field within the solar wind or other interplanetary disturbance (e.g. coronal mass ejection) hitting the magnetosphere. A closer examination of the two kinds of storms follows.

Here, both kinds of major disturbances are reviewed, as are geomagnetic Pc5 pulsations. As a conclusion to the chapter several geomagnetic indices used to monitor the geomagnetic activity are introduced to be analyzed in the later chapters.

2.4.1 Geomagnetic storms

During geomagnetic storms the magnetic field in the magnetosphere and on the ground is strongly disturbed globally. The perturbation of the magnetic field during a storm is due to the enhancement of the equatorial ring current. This constant but time-dependent westward current consisting mainly of the westward drift of positively charged particles but also the eastward drift of negatively charged particles (see Section 2.3.3).

During stormtime the ring current is enhanced and moved spatially closer to the ground. This causes a disturbance in the $H$ component of the magnetic field. It can be detected on the ground, as is indeed done at the multiple magnetic observatories. The storms are then classified according to some criteria, most often a specific index calculated from the magnetic measurements of a subset of the observatories. Being global events, the disturbance of the magnetic field during magnetic storms has to be detected by several stations at once to qualify. The most common index to classify the magnitude of the storm is the Dst index (Gonzalez et al., 1994, see also Section 2.4.4).

Since no magnetic storm is quite the same due to their complicated origins and dynamics the statistical approach used here has to be used with care (Koskinen, 2011), any conclusions ultimately requiring statistical significance analysis. In this work storms were identified from the Dst index measured by four magnetometers close to the equator. As Dst is not the only index used to classify geomagnetic storms, a brief description of different magnetic indices is given in 2.4.4.
2.4.2 Auroral substorms

Owing their name to the early studies, auroral substorms situate themselves at the auroral ovals. At first the substorms were defined only by the occurrence of the aurora. Two distinct phases of the substorms were identified from the behaviour of the aurora, namely the expansion phase and the recovery phase, which both have their defining characteristics (Akasofu, 1964). Later, when the magnetic measurements became more important, the growth phase was identified to precede the expansion phase (McPherron et al., 1973). The reason for the addition was that the growth and expansion phases are clearly distinguishable in the satellite observations, while the growth phase is largely invisible in the ground observations, except in the case of isolated substorms, i.e. ones separated by at least 3 hours from other substorms. Both methods of observation are needed, as the recovery phase is not clearly visible in the satellite observations.

Substorms are local phenomena. Therefore, not all substorm phases may be detected in every part of the auroral zone. They are usually located in the nightside of the ionosphere and last only a few hours at maximum. However, no substorm looks exactly the same and the features of the phases explained next are not detected in every substorm.

The substorm growth phase begins with perturbations appearing in the electric and magnetic fields in the polar cap. Around the same time the field in the lobes starts increasing due to perturbations of cross-tail component of the lobe field. In the near tail the plasma sheet begins to thin with the magnitude of the field increasing (McPherron et al., 1973).

Expansion onset was originally considered the beginning of the substorm (Akasofu, 1964). During the rather short expansion phase lasting few tens of minutes the auroral arcs brighten up and also exhibit a poleward motion. At the same time the near-Earth plasma sheet thins up to almost nothing, only to start expanding at about ten times the speed of its initial thinning, a result of rapid changes in the cross-tail component of the magnetic field. These fluctuations accompany the sudden appearance of energetic particles. Electrojets in the auroral zone expand northward and westward and there’s an intense electron precipitation (McPherron et al., 1973).

Finally the recovery phase, lasting a few hours, starts with the decay of electrojet currents and the recovery of the magnetic field as it was during quiet time configuration, i.e. before the substorm. The aurora also start to faint out and they begin to return equatorward to the auroral zone of the quiet time.

Being very numerous, sometimes several substorms happening during a single night, the statistical approach here is more useful when considering
substorms than geomagnetic storms. In this work the substorm data collected provided by (Tanskanen, 2009).

2.4.3 Geomagnetic pulsations

Geomagnetic pulsations are short-term (0.2–600 seconds) fluctuations in the geomagnetic field (Saito, 1969). They are classified into two main types: continuous pulsations (Pc) and irregular pulsations (Pi) and further into seven subtypes, a classification scheme approved by the International Association of Geomagnetism and Aeronomy (IAGA). For completeness their properties are shown in Table 2.2 ((Jacobs et al., 1964)). However, only Pc5 were studied in this thesis.

Pc5 pulsations can readily be studied with data of 140-sec resolution or preferably higher using the methods of Fourier analysis. They can be grouped into two distinct categories by their nature: compressional Pc5 and toroidal Pc5 that is also known as the fundamental mode (Anderson, 1993).

Compressional Pc5

Compressional Pc5 pulsations are abundant in space and the dominant pulsation type occurring beyond $L = 8$. They are probably a manifestation of drift mirror waves associated with high-$\beta$ plasma in magnetosphere (Zhu and Kivelson, 1991). Occurring at afternoon and early evening in the geosynchronous orbit and correlating with the storm time development of the partial ring current has given them the name "storm-time Pc5". The more generic term of compressional Pc5 resides from the fact that they also occur in other regions, particularly in the morning where the geophysical conditions are very distinct from storm-time. Compressional Pc5 waves have a short azimuthal scale length and thus are not detectable by ground magnetometers.

Toroidal fundamental mode Pc5

Due to their long period and the abundance of ground magnetometer data the toroidal Pc5 pulsations are probably the most studied of the geomagnetic pulsations. Being toroidal the pulsation appears as nearly pure sinusoidal

<table>
<thead>
<tr>
<th>$T$ (s)</th>
<th>Pc 1</th>
<th>Pc 2</th>
<th>Pc 3</th>
<th>Pc 4</th>
<th>Pc 5</th>
<th>Pi 1</th>
<th>Pi 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$ (mHz)</td>
<td>0.2–5</td>
<td>5–10</td>
<td>10–45</td>
<td>45–150</td>
<td>150–600</td>
<td>1–40</td>
<td>40–150</td>
</tr>
</tbody>
</table>

Table 2.2: Types of geomagnetic pulsations
perturbation in the magnetic field but has a 90° rotation that makes the perturbation largest in the north-south component. The earlier ground-based studies on the toroidal Pc5 have shown that their local time distribution peaks at dawn and dusk. This has been taken as evidence that these pulsations are likely driven by energy sources at the flanks of the magnetopause by the Kelvin-Helmholtz instability (Anderson, 1993). Spacecraft studies support the energy source being located at high \( L \).

### 2.4.4 Magnetic indices

To measure the severity of storms and substorms several different indices have been derived. These are constructed from magnetic field measurements on the ground. The indices can be grouped roughly into categories of which kind of disturbances in the ionospheric currents they measure.

IAGA (International Association of Geomagnetism and Aeronomy) officially recognizes five magnetic indices, namely Dst, Kp, \( AE \), aa and am. In this work the first four were considered. In addition to the official indices several specialized or generalized indices have been derived. Both the utilized official magnetic indices and several unofficial ones are introduced and used in this thesis.

#### Storm indices

Geomagnetic storms are measured by several different indices. The storms are detected globally especially in the low latitudes due to their origin in the enhancement of the equatorial electrojet currents. Then a measure of the enhancement can be acquired by taking a group of geomagnetic observatories situated close to but not at the equator and derive an index out of their measurements. Observatories at the equator would measure only the equatorial electrojet, which is not as interesting in this context as the disturbance it causes. For measuring the disturbance the most common indices are Dst index, Kp index and the closely related Ap index (Rostoker, 1972). In addition the planetary aa index is used to monitor the global disturbance level.

#### Dst and related indices

The Dst index (Disturbance storm time) has been acquired since 1957 (Sugiura et al., 1991). For this particular index four geomagnetic observatories, Hermanus, Kakioka, Honolulu and San Juan are used, shown on the map in Figure C.1 in Appendix C. The four have been chosen so that they are
Table 2.3: Classification of storms by minimum Dst index (Loewe and Prölls, 1997)

<table>
<thead>
<tr>
<th>Storm type</th>
<th>Minimum Dst below</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak storm</td>
<td>$-30 \text{ nT}$</td>
</tr>
<tr>
<td>Moderate storm</td>
<td>$-50 \text{ nT}$</td>
</tr>
<tr>
<td>Strong storm</td>
<td>$-100 \text{ nT}$</td>
</tr>
<tr>
<td>Severe storm</td>
<td>$-200 \text{ nT}$</td>
</tr>
<tr>
<td>Great storm</td>
<td>$-350 \text{ nT}$</td>
</tr>
</tbody>
</table>

distributed as evenly as possible in longitude, and are at the same time sufficiently distant from both the auroral and equatorial electrojets so that they are not significantly influenced by either.

In determining the Dst index, the observatories determine the deviation of the horizontal components $H$ of the magnetic field from the baseline. The baseline is calculated separately for each observatory and year. The "five quietest days" for each month form the database for the baseline. Also the solar quiet daily variation, $S_q$, is removed from the $\delta H$ of each observatory to form the disturbance variation $D(t)$. Finally the Dst is acquired by averaging the results of all observatories and normalizing the result to the dipole equator by the division by the average of the cosines of the dipole latitudes of the observatories. The final result is commonly a negative number given in units of nT determined on hourly basis.

Dst index is used to determine the occurrence, power and phases of geomagnetic storms. One common classification of the storm severity is shown in Table 2.3 that categorizes storms by the minimum Dst during a storm (Loewe and Prölls, 1997).

There also exists a higher resolution version of Dst index, called SYM-H index. This index is constructed in a similar manner to Dst index, the main differences being the different determination of baseline and the higher time resolution of 1-minute. It has been shown (Wanliss and Showalter, 2006) that SYM-H can indeed be used as a higher-resolution analogue to Dst index. However, the resolution of Dst index is sufficiently accurate for this study.

An extended version of the Dst index has been constructed, called Dxt index, $x$ standing for extended (Karinen and Mursula, 2005). The extension of the 1941–1956 was constructed using the original four observatories and the nearby Cape Town station as a substitute for Hermanus station for the years 1932–1940. The whole range of 1932–2009 was acquired following the original procedure (Sugiura et al., 1991) as closely as possible.

There are two basic differences between the derivation of the two indices. First, in the Dst index the data gaps were filled with the data of various
other magnetic observatories, while in the Dxt the data in the original four observatories were used regardless. Secondly, the final part deriving the index was done differently: the disturbance variations were first normalized by dividing by the cosines of the latitudes and only then their average was taken to form the final Dxt index.

A further improvement on Dst index was made by removing the semi-annual seasonal variation from the Dxt index (Mursula and Karinen, 2005). The new index was called Dcx index. This new index was shown to have a better correlation with both the sunspot number and geomagnetic indices than the traditional Dst index had (Karinen and Mursula, 2006). Since the index is rather new, it has not been validated and therefore IAGA considers Dst index to be the best measure of the enhancement of ring current.

**Planetary Kp index and related indices**

Another common measure of geomagnetic activity is the planetary $K_p$ index, originally implemented to get a measure of the level of storminess within the geomagnetic environment. It is measured every 3 hours since 1932.

$K_p$ is computed in three stages (Rostoker, 1972). First, $K$ index is derived for each observatory by determining the maximum deviation $\delta_{\text{max}}$ for each component $H$, $D$ and $Z$ of the magnetic field. Each observatory has its own table for converting this value into the quasi-logarithmic $K$ index that ranges discretely from 0 to 9. The values of the table depend on the observatory’s latitude that is between 49° and 62° for northern or 46° and 48° for southern hemisphere observatories chosen for this particular index. The $K$ index is then normalized into $K_s$ index by taking into account diurnal and seasonal variation. $K_s$ index is more granular, reported with a resolution of thirds of an integer. Finally the planetary index $K_p$ is acquired by averaging the $K_s$ indices for all the observatories taking part in the measurement of the index.

Being quasi-logarithmic by nature $K_p$ is not very suitable for statistical analysis and is not used directly in this study. Instead a related linearized index called $A_p$ is used. First a 3-hour index $ap$ is obtained directly from the $K_p$ according to a conversion table (Rostoker, 1972). This is computed for only 8 observatories. Then $A_p$ index is the average of these 8 $ap$ indices. Like $K_p$, also $A_p$ index has been obtained since 1932.

Using higher-latitude observatories than for Dst, the $K_p$ and $A_p$ indices are highly affected by auroral currents as well as equatorial electrojet currents. Therefore they are suitable for monitoring the general state of space weather, obtaining an average of both geomagnetic storm and substorm activity. However, the resolution of 3 hours is a strong limitation. Luckily, for this study it is rather sufficient.
Planetary aa index

Also closely related to $Kp$ indices, the antipodal activity index $aa$ (Menvielle and Berthelier, 1991) is derived from two nearly antipodal observatories since 1959. Since the data itself has been available since 1868 its time series now spans almost 150 years.

Originally derived by (Mayaud, 1972), the index seeks to cancel the annual variation with a maximum in each hemisphere and the night maximum in local time. Using two antipodal observatories both effects can be accounted for and eliminated.

The observatories used in deriving the index are shown in Table C.1 in Appendix C. The $K$ index derived by the observatory is converted into nanoteslas and the average of the two stations from both hemispheres is used to derive the index after a couple of standardizing operations related to latitude of the station (Mayaud, 1980).

Indices of auroral substorm activity

The auroral substorms occur in the auroral zones of the Earth on both hemispheres and thus the measurements have to be made in the high latitudes. Along with indices to monitor geomagnetic storms, auroral electrojet index $AE$ is listed (Rostoker, 1972) as one created from such measurements and therefore suitable for monitoring auroral substorm activity in the northern hemisphere. Related indices $AU$ and $AL$ can also be used. With a very similar principle an index has been composed from the data of IMAGE network, which uses $I$ in place of $A$ in the index name.

$AE$ index was designed to act as a measure of global electrojet activity in the auroral zone (Davis and Sugiura, 1966). It is obtained through the cooperative effort of the observatories, the present ones listed in Table C.2 in Appendix. The index has been collected since 1932, though some of the observatories have changed since the initial launch. Figure C.2 in Appendix shows how they are situated on the map at the moment. The observatories measure the $H$ component of the perturbation field every 2.5 minutes using the average quiet-time baseline as a reference level (Rostoker, 1972).

From the measured perturbative field $H$ component the indices of $AU$ and $AL$ are determined simply by picking the maximum positive and negative values of all the stations, respectively. Then $AE$ index is composed using the formula $AE = AU - AL$.

Similarly the IMAGE stations take the minimum $X$ component of the magnetic field, where $X$ stands for the geographic north component. Again, $IU$ is the maximum value of this field and $IL$ the minimum and $IE$ is ob-
tained through the formula $IE = IU - IL$. However, IMAGE network has more stations than the one constructing $AE$ index, albeit longitudinally more restricted situating roughly in the longitudes 5°E to 35°E. Latitudinally, IMAGE covers a larger area that is favourable for electrojet studies. List of IMAGE stations is given in Table C.3 and their locations on the map are shown in Figure C.3 (International Monitor for Auroral Geomagnetic Effects, 2013) in Appendix.

**Polar cap index**

Polar cap index, or PC, has been designed to monitor the magnetic activity of the polar cap caused by the solar wind and especially its geoeffective properties (Troshichev et al., 1979, 1988). It is based on the data of a magnetic observatory near the geomagnetic pole, the chosen stations being Thule in Greenland for northern hemisphere and Vostok in Antarctica for southern hemisphere.

PC index has a good correlation with all solar wind parameters that include the southward component of the IMF and could in theory be used to monitor them on the ground (Troshichev and Andrezen, 1985). However, because the ionosphere’s conductivity varies greatly with seasons due to UV radiation of the Sun, it is necessary to have a station on both polar caps to minimize the variation.

For both stations the PC index is calculated by the formula group

$$\delta F = \Delta H \sin \beta \pm \Delta D \cos \beta$$

$$\beta = \lambda \pm D + U.T. + \phi$$ (2.17)

$$PC = \frac{\delta F}{\alpha}$$ (2.18)

where $\Delta H$ and $\Delta D$ are deviations of the $H$ and $D$ components of the magnetic field from the quiet level, $D$ in Equation (2.17) is the average value of the declination at the station ($D = -117^\circ$ for Vostok and $D = 285^\circ$ for Thule), $\lambda$ is the geographic longitude and $\phi$ is the optimal angle between the noon-midnight meridian and the equivalent current vector, i.e. the direction of the antisunward convection. Finally $\alpha$ is the best-fit regression coefficient derived from the expression

$$\delta F = \alpha v B_T \sin^2 \theta/2 + K,$$ (2.19)

where $K$ is another regression coefficient, $v$ is the speed of the solar wind and the toroidal magnetic field component $B_T = (B_y^2 + B_z^2)^{1/2}$.  

35
In the data used here the indices were given as 15-min average values of the magnetic perturbation.
Chapter 3
Data analysis methods

In this thesis, several statistical and mathematical methods are used in the data analysis. These methods are introduced in this chapter. In Section 3.1 the most important concepts of statistical analysis are reviewed. In Section 3.2 the methods of Fourier analysis are reviewed along with methods to attack the problems caused by sampling and other errors introduced by discrete Fourier transforms. In both sections the numerical challenges associated with these methods were also taken into consideration, especially using MATLAB environment and its signal analyzing toolbox. MATLAB code scripts used in the work are referred in the appendixes.

3.1 Statistical methods

In this Section a brief review of statistical analysis and methods is given. The various terms of statistics are defined to be used in Chapters 4 and 5. Only the discrete forms of the formulas are introduced, as the respectable integral forms cannot be used directly with experimental data.

3.1.1 Mean and median

The arithmetic mean (later just mean) of any time series of $N$ data points is defined as

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \quad (3.1)$$

where $x_i$ is the $i$th term of the time series. Mean describes the average value of the property in a given interval. Likewise a time mean can be defined as

$$\bar{x} = \frac{1}{N} \sum_{t_0 < t < t_f} x(t), \quad (3.2)$$
where the sum is over all values of \( x(t) \) for which \( t \in \left[ t_0, t_f \right] \) and \( N \) is the number of these values, assumed finite and discrete (Press et al., 2007, Chapter 14). For continuous functions \( f(x(t)) \) the sums would be replaced by integrals.

Similar to mean, median of a time series is the function value than which equal number of other values are higher and lower (Press et al., 2007, Chapter 14). It is the middle point of the data. It can be determined by sorting the data points into ascending or descending order. If there are odd number of points the median is

\[
x_{\text{med}} = x_{(N-1)/2}
\]

(3.3)

and for even number of points, respectively,

\[
x_{\text{med}} = \frac{1}{2} \left( x_{(N/2)-1} + x_{N/2} \right).
\]

(3.4)

### 3.1.2 Variance and standard deviation

Variance of the time series is related to the mean. It is defined by the equation

\[
\sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2,
\]

(3.5)

where the variables are the same as in 3.1.1. Square root of variance, \( \sigma \), is called the standard deviation. Both describe the variability of the data around the mean (Press et al., 2007, Chapter 14).

### 3.1.3 Skewness

The degree of asymmetry of data can be estimated by the property called skewness, \( \gamma \), defined as

\[
\gamma = \frac{1}{N} \sum_{i=1}^{N} \left[ \frac{x_i - \bar{x}}{\sigma} \right]^3,
\]

(3.6)

where the variables are the same as in 3.1.2. Unlike mean and variance this value is dimensionless. Skewness is positive, if the tail of the function’s distribution extends more toward positive than negative end of the series and the other way around for negative skewness. Naturally, being proportional to the third power of deviation of \( x \) from its mean value, this property should be used with care (Press et al., 2007, Chapter 14).

For illustration consider Figure 3.1 that shows three functions that resemble sunspot cycles. They differ only (beside offset in \( y \) axis for illustration) in
the length and slope of the tail. Figure shows the longer the declining phase and the minimum (or the ascending phase for that matter) in a solar cycle, the higher the value of skewness. This resides from the fact that then the mean is lower and more values are higher to it.

\[
\beta = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i - \bar{x}}{\sigma} \right)^4 \right\} - 3,
\]

(3.7)

where the variables are the same as in 3.1.2 and the constant \(-3\) makes the kurtosis zero for a normal distribution, to which the relative peakedness is compared with kurtosis. Like skewness, kurtosis should be used with care (Press et al., 2007, Chapter 14).

To illustrate the kurtosis we show it for several functions. Simple symmetric partially defined functions related to \(x, x^2, x^4, \sqrt{x}\) and \(\delta(0)\), all of which peak 1 at \(x = 0\), have been plotted in Figure 3.2. They consist of 201 points, 100 on each side of 0 for which kurtosis was calculated with MATLAB function 'kurtosis'. The stronger and higher the peak relative to the rest of the data, the higher the value of kurtosis.
However, it should be noted that ‘kurtosis’ function of MATLAB differs from the theoretical definition of Equation (3.7) in that it does not include the constant ‘−3’. MATLAB implementation is used throughout this thesis.

![Figure 3.2: A plot showing several symmetric functions in the range [-1 1]. The legend also shows their kurtosis values in this range. Each plot contains 201 data points, 100 data points on each side of 0.](image)

### 3.1.5 Correlation and regression analysis

To study the relationships between variables the correlation coefficients are a suitable tool. Due to issues of discrete Fourier analysis (Section 3.2.1) also regression analysis and the associated removal of linear trends from the data is very useful (Section 3.2.4). Theory of calculating them is largely reviewed in (Bendat and Piersol, 2010) and only briefly explained here.

**Correlation coefficients**

The expected value of a continuous function $g(x, y)$ of two random variables $x(k)$ and $y(k)$ is defined

$$E[g(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)p(x, y)dx\,dy,$$  \hspace{1cm} (3.8)
where \( p(x, y) \) is the probability distribution of the two variables often assumed normalized normal distribution. In turn covariance \( C_{xy} \) is defined

\[
C_{xy} = \begin{align*}
E[(x(k) - \mu_x)(y(k) - \mu_y)] \\
= E[x(k)y(k)] - E[x(k)]E[y(k)] \\
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y)p(x, y)\,dx\,dy,
\end{align*}
\]

(3.9)

where \( \mu_x \) and \( \mu_y \) are the means of \( x \) and \( y \) respectively.

Finally, correlation coefficient \( \rho_{xy} \) of two random variables \( x \) and \( y \) is defined through the covariance

\[
\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y},
\]

(3.10)

where \( \sigma_x \) and \( \sigma_y \) are the standard deviations associated with variables \( x \) and \( y \). This is a normalized quantity \(-1 \leq \rho_{xy} \leq +1\). If the variables are entirely uncorrelated, the coefficient is zero. Respectively they are said to be correlated if the coefficient is close to +1 and anticorrelated when close to \(-1\). It is also notable that uncorrelated variables are not necessarily independent though for physical variables being uncorrelated does imply independence.

**Method of least squares**

An effective way to fit a line into data is to use the method of least squares (Hefferon, 2012, e.g.). In this model the \( i \)th value of the curve \( y(x) \) is given by the equation

\[
y_i = a_0 + a_1 x_i,
\]

(3.11)

where \( a_0 \) and \( a_1 \) are constants. It can be expressed in a matrix form

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix} = \begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix} \begin{bmatrix}
a_0 \\
a_1
\end{bmatrix}
\]

(3.12)

or in short-hand notation

\[
\bar{y} = X\bar{a}.
\]

(3.13)

From this \( \bar{a} \) can be derived by

\[
\bar{a} = (X^T X)^{-1} X^T \bar{y}.
\]

(3.14)
The method can be extended to higher polynomials by rewriting Equation (3.11)

\[ y_i = a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m, \]  

so that Equation (3.12) becomes

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} =
\begin{bmatrix}
  1 & x_1 & x_1^2 & \cdots & x_1^m \\
  1 & x_2 & x_2^2 & \cdots & x_2^m \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & x_n & x_n^2 & \cdots & x_n^m
\end{bmatrix}
\begin{bmatrix}
  a_0 \\
  a_1 \\
  \vdots \\
  a_m
\end{bmatrix}
\]  

which reduces to Equation (3.13).

**MATLAB implementation**

In MATLAB environment the correlation coefficients of two column vectors \(x\) and \(y\) can be calculated with the command `corrcoef(x,y)`. The result is a two-by-two matrix, in which components \((1, 2)\) and \((2, 1)\) both contain the correlation coefficient between the two data sets. The probability distribution used to calculate the correlation coefficient is computed by transforming the correlation to create a t-statistic that has \(n - 2\) degrees of freedom, \(n\) being the number of rows in \(x\). The result thus-acquired is adequate for large samples (The Mathworks, 2013a).

Regression analysis through least-squares method is attained by the use of function `polyfit(x,y,n)`. Here \(y(x)\) is the original function expressed as column vectors \(x\) and \(y\), and \(n\) is the degree of the polynomial fitted to the data through the least squares method. The result is a vector of the coefficients of the polynomial (The Mathworks, 2013c). In this study the degree is usually \(n = 1\) (simple linear relationship) or \(n = 2\) (second-degree polynomial).

3.1.6 MATLAB implementation of running average

One means to smooth continuous experimental data is to calculate a running average. In this procedure, the data is divided into equal length samples. The samples may or may not overlap. Both the mean value of the time and function’s mean is calculated and the smoothed time series is constructed from these value pairs.

For this purpose I have constructed a MATLAB function to calculate the running mean, introduced in Appendix D.2. This code was used to produce Figure 2.2 of yearly average of the sunspot number and many of the figures in Chapters 4 and 5.
3.1.7 Interpolation of indefinite values

When doing data analysis on continuous data there are often intervals where the data is not defined, for example due to saturation, numerical corruption of the data or failure of the instrument. Downtime also results in undefined values. To issue discrete Fourier analysis these missing points have to be provided as well. Some form of interpolation has to be used to find out the data values that are left between proper data points.

In this thesis a linear interpolation between data points was used due to its simplicity when appropriate. Formally this is done through the following procedure, duplicated by my MATLAB code introduced in Appendix D.3:

- Identify intervals with indefinite values (in MATLAB these are NaN values, for 'not-a-number')
- Calculate the parameters $a$ and $b$ of a line $y = ax + b$ fitted between the proper function values on both sides of the interval.
- Replace the indefinite values in-between with the values given by the interpolation curve.

However, this procedure has a drawback, as polynomials in the series will introduce aliasing (see 3.2.1) that has to be taken into account when further analyzing the data with Fourier analysis. Sudden changes in the data can also be due to discontinuities that likewise will introduce error that has to be dealt with.

3.2 Fourier analysis

Fourier analysis is a powerful method to find timewise repetitive phenomena. According to Fourier analysis, any function can be represented as a series sum of trigonometric functions now called Fourier series (Fourier, 1808). This method is used in analyzing wave patterns in Chapter 5.

A Fourier series of any function is defined as a sum of trigonometric sine and cosine functions in the following way: (Arfken et al., 2013, ch. 19)

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx.$$  \hspace{1cm} (3.17)

The coefficients $a_0$, $a_n$ and $b_n$ are functions of the definite integrals:

$$a_n = \frac{1}{\pi} \int_{0}^{2\pi} f(s) \cos nsds, n = 0, 1, 2, ...,$$  \hspace{1cm} (3.18)
\[ b_n = \frac{1}{\pi} \int_0^{2\pi} f(s) \sin n s \, ds, \quad n = 0, 1, 2, \ldots, \]  

(3.19)

so long as the integrals exist. The definitions are valid also for \( a_0 \), which is taken into account in the first term of 3.17. A sufficient condition for Equation (3.17) to be valid is that there are only a finite number of finite discontinuities in the considered interval \([0, 2\pi]\), a condition that is valid in the experimental data considered in this study.

If \( \cos nx \) and \( \sin nx \) are expressed in an exponential form, Equation (3.17) may be rewritten in a form

\[ f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \]  

(3.20)

where \( c_n \) are now the Fourier coefficients

\[ c_n = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-inx} \, dx. \]  

(3.21)

This form has the advantage of being simpler yet it might be more difficult to use in some cases.

A more useful representation of the series is the Fourier transformation \( \tilde{f}(k) \) of the function \( f(x) \), defined\(^1\) (Arfken et al., 2013, ch. 20)

\[ \tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{ikx} \, dx. \]  

(3.22)

Now \( f(x) \) is presented as a function of frequency. This way the periodic behaviour of the function is revealed. The original function \( f(x) \) can be obtained from the Fourier-transformed function by inverse Fourier transformation:

\[ f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{-ikx} \, dk. \]  

(3.23)

Experimental data is usually presented with discrete data points instead of continuous functions. For this purpose the method called discrete Fourier transformation was used, introduced below.

### 3.2.1 Discrete Fourier transform and its limitations

Similar to the Fourier transformation, discrete Fourier transform, or DFT, is defined by the following operation on the points \( x_k \) of the data:

\[ \hat{f}(x_p) = N^{-1/2} \sum_{k=0}^{N-1} e^{2\pi ikp/N} f(x_k), \]  

(3.24)

\(^1\)This is just one convention to normalize Fourier integral. See Section 3.2.6 for the convention used by MATLAB.
where \( N \) is the number of data points, \( f(x_k) \) are the values of the data and \( x_p \) are the frequency components.

Equation (3.24) is analogous to (3.22), and its inverse (analogous to Equation (3.23)) is

\[
f(x_k) = N^{-1/2} \sum_{k=0}^{N-1} e^{-2\pi i kp/N} \tilde{f}(x_p).
\] (3.25)

DFT equations are equivalent to the Fourier series when data sampling is imposed on the data. However, sampling also introduces an error not present in continuous Fourier transformation. This error, called aliasing, cannot be removed from the data once the sampling is used. Usually this has already happened when the data was acquired because writing down data values as function of time or similar is directly attributable to the act of sampling (Hamming, 1973, chapter 31). Because of this error it is important to know the sampling rate, the resolution, of the instrument in question.

In aliasing, some higher frequency components are added to the lower frequency components due to the data not having enough resolution. In fact, the highest observable frequency is given by the Nyquist sampling theorem, being one half of the frequency at which the sampling was done. However, the higher frequencies, not adequately represented by the sampling, are technically not lost but instead fold back onto the Fourier components at the Nyquist frequency adding error. This anomaly can be addressed by appropriate tapering methods that effectively filter out higher frequencies (see 3.2.4 below (Hamming, 1973, Section 31.6)).

Another error grows out of discontinuities in the data. Imposing Fourier transform on a data with a discontinuity introduces an overshooting error to the function (Gibbs, 1898, 1899). Known as Gibbs phenomenon, it can be reduced greatly by removing appropriate polynomials (see Section 3.2.4) and also using tapering methods at the same time (see Section 3.2.4) (Hamming, 1973, Sections 32.5. and 32.6).

Discrete Fourier transform, as also continuous Fourier transform, exhibits an issue with its strongest spectral components leaking power away into higher frequencies. This behaviour is the more pronounced the less there are data points taken into the transform. Long time series and high resolution are preferred to minimize this error. Also the methods introduced in Section 3.2.4 will help alleviate the problem though it cannot be completely removed with the methods herein.
3.2.2 Power spectrum

The coefficients $a_k$ and $b_k$ themselves are typically not physically relevant. Instead their square sum, $a_k^2 + b_k^2$, is invariant under translation (e.g. in time). This quantity represents the power at frequency $k$, and their plot as a function of $k$ is called the power spectrum (Hamming, 1973, Section 31.7).

It must be noted that the power spectrum does not conserve the phase angle of the periodic components. Fortunately the phase angle is usually not a physically interesting property and we can safely ignore it in this thesis as well.

3.2.3 Convolution theorem

For any arbitrary input function $x(t)$ and an impulse response function $h(t)$, also known as the weighting function, the system output $y(t)$ is given by the convolution integral

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau.$$  

(3.26)

For a system to be physically realizable, it may only respond to past inputs. Then the impulse response function has to have the property $h(\tau) = 0$ for $\tau < 0$. Hence, for physical systems, the effective lower limit of integration in the above Equation (3.26) must be zero instead of $-\infty$.

Using Fourier transformation on the both sides of Equation (3.26) results in the convolution theorem, only shown here without proof:

$$Y(k) = H(k)X(k),$$  

(3.27)

where $Y$, $H$ and $X$ are the Fourier-transformed functions $y(t)$, $h(t)$ and $x(t)$ (Bendat and Piersol, 2010).

This theorem will be of tremendous use to attack errors issued by sampling. The methods in the next Section 3.2.4 are based on the validity of the convolution theorem.

3.2.4 Improving convergence of Fourier series

...by removal of polynomials

To demonstrate the calamity that is the polynomial, we will calculate the Fourier transformations of polynomials. We will begin with the simplest one of a constant $C$

$$f(x) = C.$$  

(3.28)
Constant has the Fourier transformation

\[ \hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} C e^{ikx} \, dx \]  \\
= \frac{C}{ik\sqrt{2\pi}} e^{ikx} \]  \\
= \frac{C\sqrt{2}}{\sqrt{\pi}} \lim_{n \to \infty} \frac{\sin kn}{k} \]  \\
\tag{3.29}

which corresponds to a constant times a delta function, i.e. \( C\delta(0)/\sqrt{2\pi} \), when the definition \( \delta_n = \frac{1}{n} \int_{-n}^{n} e^{ix} \, dt \) is used with \( n \to \infty \) (Arfken et al., 2013). Therefore having a constant, i.e. a non-zero mean, in the time series corresponds to a spike in the zero-point of the Fourier series.

Given the Fourier transform of the constant it is simple to derive the Fourier transform of a polynomial of any order by taking \( n \) derivatives of Equation (3.31). Thus said,

\[ \frac{d^n \hat{f}(k)}{dk^n} = \frac{C}{\sqrt{2\pi}} \delta^{(n)}(0) \]  \\
= \frac{C}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{d}{dk} e^{ikx} \, dx \]  \\
= \frac{C}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (ix)^n e^{ikx} \, dx \]  \\
= \text{Fourier}(C(ix)^n), \]  \\
\tag{3.32}

or for polynomial \( f(x) = ax^n \) (\( a \) is constant)

\[ \hat{f}(k) = \frac{a}{i^n\sqrt{2\pi}} \cdot \delta^{(n)}(0), \]  \\
\tag{3.36}

in both of which \( \delta^{(n)} \) is the \( n \)th distribution derivative of Dirac delta function.

The convergence of the Fourier transform can therefore be improved by removing said polynomials (Bendat and Piersol, 2010). Often it is only practical to remove some of the lowest degree polynomials. For any data at least the mean (the constant function) should be removed but often also the linear or the second-degree polynomial is removed. However it is often impractical to remove higher degree polynomials due to difficulties of optimal polynomial fitting and numerics but also due to the fact that removing some of the higher polynomials may affect or eliminate some of the lower frequencies from the data.

In this study only low frequencies were studied and therefore it was practical to remove polynomials up to 2nd degree from any data, for which power spectrum analysis was used (see Section 3.1.5).
...through tapering functions

Problems of aliasing, overshooting and power leaking can be alleviated by using different tapering methods. They are based on the convolution theorem introduced above (Equation 3.27) and they will clear out some of the error introduced by sampling, though it cannot be removed entirely. Tapering window used here is the Blackman–Harris window that is a generalized tapering window based on the Hanning window, i.e. the cosine squared tapering window.

In the Hanning method the function to be Fourier transformed is multiplied by a tapering window, kernel, given by

\[ u_h(t) = \frac{1}{2} \left( 1 - \cos \left( \frac{2\pi t}{T} \right) \right), \quad 0 \leq t \leq T, \text{ otherwise } u_h(t) = 0 \] (3.37)

The Hanning window (3.37) will smooth out the plot of the Fourier transformation so that strong spikes are more elaborately preserved and the random noise is tuned down.

Blackman–Harris window is an optimized version of the Hanning window, defined

\[ w_n = a_0 - a_1 \cos \left( \frac{2\pi n}{N - 1} \right) + a_2 \cos \left( \frac{4\pi n}{N - 1} \right) - a_3 \cos \left( \frac{6\pi n}{N - 1} \right), \] (3.38)

where \( a_0 = 0.35875, \ a_1 = 0.48829, \ a_2 = 0.14128 \) and \( a_3 = 0.01168 \). It has been designed to minimize the error due to side-lobes (Harris, 1978).

3.2.5 Filtering-squaring-averaging method of powers spectra

A way to smooth the spectral data is to use filtering-squaring-averaging method. In this method, the following procedures are completed (Bendat and Piersol, 2010, Section 5.2.3):

- The signal of the power spectrum is filtered over certain bandwidth. In discretized data, certain number of data points.
- The values within filter are squared.
- The average of the squared signals is taken.

This procedure will result in smoother power spectra.
3.2.6 FFT with MATLAB signal processing toolbox

Heavy use of MATLAB is made in this thesis. Because of its special habits it is appropriate to look deeper into this toolbox, especially the signal processing pack that contains the functions to calculate fast Fourier transforms.

The fast Fourier transformation, or FFT, is dubbed fast because with this algorithm it takes time proportional to $N \log N$ to calculate the Fourier transformation of $N$ data points as compared to direct calculation that takes time proportional to $N^2$ (Cooley and Tukey, 1965). I also used this method in this study due to its speed and ease.

MATLAB uses different convention for the calculation of FFT than theory given in Section 3.2.1. Particularly the forms of Equations (3.24) and (3.25) used by MATLAB functions 'fft' and 'ifft' for DFT and inverse DFT are

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)} \quad (3.39)$$

for Fourier transform and

$$x(j) = (1/N) \sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)} \quad (3.40)$$

for its inverse, where $\omega_N = \exp(-2\pi i)/N$ is the $N$th root of unity (for full description, see (The Mathworks, 2013b)).

The Fourier transformation results in a function symmetric in $k$ by the middle of the frequency range, the first value starting at $k = 0$ and the last one being $k = N$. The function will be zero in the middle of the range. To derive a more elegant picture the two halves will be swapped and translated so that the two ranges meet at $k = 0$. Then the frequency distribution will be symmetric by $k = 0$.

3.2.7 Using Fourier analysis in wave detection

Fourier analysis extracts the periodic behaviour of the data. However, instead of the Fourier spectrum itself its power spectrum is used instead. The data is manipulated in several ways before and after acquiring power spectrum to get sharper results.

Methods were used in this order:

- Data is loaded into MATLAB environment. It contains date (year, month, day, hour, minute and second) of the event and the actual measured property, in here always the magnetic field.
• Date is turned into seconds.

• Undefined data points (NaN values) are identified from the data.

• Linear interpolation (see Appendix D.3) is used on the data.

• A certain hour is chosen for analysis. The following procedures are only done on the data of this hour. Generally the procedures were done for every hour of several years.

• If more than 1/4 of the data is undefined values, the data set is omitted, unsuitable for analysis.

• The mean, linear and second-order trends are removed from the hour of data.

• Fourier transformation of the data is calculated using Blackman–Harris window.

• Power spectrum is calculated of the Fourier transformed data.

• Squaring-filtering-averaging is used with a suitable number of overlapping data points (5 or 7).

• Maxima are identified in the power spectrum. The frequency, power and its standard deviation from the average of the data are noted as are the average and the standard deviation of the data.

• Finally, only potential events with certain threshold are chosen for suitable events. Only events within certain frequency range and having high enough power are chosen. In this study the frequency range was 2–7 mHz and the power threshold was arbitrarily chosen to be close to the average of all the maxima leaving out maxima of low amplitude.

Experimental data introduces additional layer of glitches. The motion and relative location of the instrument can change what is measured, especially in case of satellites that can move in and out of magnetosphere, but also in ground-based magnetometers that monitor the dynamic auroral oval. For full analysis it has to be taken into account.

Sometimes data are not entirely evenly based. This potential error-causing effect was omitted from the analysis. The effect is probably not high due to the rather short range (a single hour) of measurements used for the analysis.
Also the equipment itself can cause error. In ACE satellite the rotation rate of the satellite is readily seen in the power spectrum having a high maximum at 82.5 mHz corresponding to the satellite’s rotation rate of 5 rpm (Chiu et al., 1998).
Chapter 4

Data analysis of the space weather activity measures

In this Chapter I will show the results of the statistical analysis on the measures of space weather activity. These measures include the indices introduced in Sections 2.1 and 2.4.4.

4.1 Measures of solar activity

Sunspot number and other measures of the solar activity were analyzed over all solar cycles, for which data is easily available. Using the definition of a solar cycle and its phases given in Section 2.1.5 the solar cycle and the phase lengths and the time of the minima were determined for each solar activity index. For each solar cycle and activity measure also several statistical properties were calculated including the standard deviation (as defined in Section 3.1.2), skewness (3.1.3) and kurtosis (3.1.4). These are summarized in Section 4.1.6.

A special attention is given to the solar cycle 23 (SC23). This deems from this cycle bearing the largest amount of data on a sunspot cycle ever before. Its subtleties are therefore interesting to look into. Each index was separately compared to the international sunspot number, ISSN, which is used here as the reference index. Finally, properties of all the measures were compared to each other.

4.1.1 International sunspot number

Figure 4.1 shows the international sunspot number ISSN smoothed for a yearly average. Data used is of the courtesy of SIDC website (SIDC team,
The definition of the solar cycle in Section 2.1.5 identifies correctly the solar cycles as interpreted by Wolf. For the identified spot cycles 1 – 23 several key properties were calculated and are shown in the figures of Section 4.1.6. The average cycle length is 11.0 years confirming the 11-year cycle, with a standard deviation of 1.21 years.

SC23 is shown in Figure 4.2 with its phases marked. The solar cycle began in the summer 1996, shortly after which the ascending phase started. The first maximum occurred around summer 2000 with a local minimum in winter 2001 and another maximum in winter 2002. Only some half a year after this maximum the cycle started declining, a progress that lasted almost until 2007 before the minimum of the cycle. Consequently SC23 was the longest cycle known in the rather short history of recorded sunspot numbers, though SC5 and SC9 are quite close as well.

In the later subsections the other indices of solar activity are compared to this data. Therein sunspot number will always refer to the international sunspot number unless noted otherwise.

### 4.1.2 Other sunspot numbers

For completeness the American relative sunspot number ARSN and Boulder sunspot number are shown in Figure 4.3 alongside with ISSN. ARSN data is the courtesy of American Association of Variable Star Observers and
Figure 4.2: Smoothed international sunspot number during SC23. Changes of phase have been marked with vertical dashed lines, the minimum phase being left out of the picture to the right. Data source: (SIDC team, 1818–2013)

Boulder data of NOAA SWPC, provided by Substorm Zoo http://www.substormzoo.org.

Unfortunately, the Boulder sunspot number data available for this study comprised only years 1997 – 2011 and not many conclusions could be made of it. On the other hand ARSN was available since 1945 and therefore a full analysis was possible over solar cycles 19 – 23 and partly on SC18.

ISSN and ARSN correlate very well having a correlation coefficient 0.985 (the smoothed indices) for the whole series that they both are available. Figure 4.4 shows all three indices during SC23, when both ISSN and ARSN are nearly equal with just a small difference in 1998 – 1999 when ARSN peculiarly nudges upwards while ISSN raises more steadily. Otherwise they show very little difference. Boulder number is higher but otherwise follows the trend of the other two indices well. Curiously unlike ISSN both ARSN and Boulder number show that the second maximum of the cycle was higher than the first one, Boulder number being more pronounced in this conclusion.

### 4.1.3 Sunspot area index

Figure 4.5 shows the sunspot area index SSAI smoothed over a year. Data is of the courtesy of (Hathaway and NASA, 2013).

The data comprises solar cycles 12 – 23. The key properties of these
cycles are collected in the figures of Section 4.1.6. Average length of these cycles is 10.8 years, the same as for the sunspot number if calculated over the cycles 12 – 23.

SC23 for the sunspot area index is shown in Figure 4.6 with its phases marked. Comparing to the sunspot number the cycles are very similar but the method to locate the phase changes has resulted in every phase in SSAI beginning about 1/4 year later, and half a year for the switch to the minimum phase. Phase lengths are remarkably similar between the two indices. The worst coefficient is for the length of the maximum phase. This seems to be due to some cycles (especially SC14 and SC16) having exceptionally messy maximum phase in the area index resulting in longer maximum phase.
Figure 4.4: Smoothed international sunspot number alongside with American relative sunspot number and Boulder sunspot number during SC23. Data source of ARSN and Boulder data: http://www.substormzoo.org

Figure 4.5: Smoothed sunspot area index alongside with ISSN that has been multiplied by 10 to fit it in the picture with SSAI. Data source of SSAI: (Hathaway and NASA, 2013)
Figure 4.6: Smoothed sunspot area index during SC23 alongside with ISSN. Changes of phase have been marked with dashed lines for both indices with their own colours, the minimum phase being left out of the picture on the right. Data source of SSAI: (Hathaway and NASA, 2013)
4.1.4 Solar radio flux

Shown in Figure 4.7 is the solar radio flux at band 10.7 cm (2800 MHz) smoothed over a year. It is this band that is used throughout this analysis. Data is of the courtesy of (NOAA Solar–Terrestrial Physics Division, 2013).

![Smoothed solar radio flux in \(\lambda=10.7\) cm](image)

Figure 4.7: Smoothed solar radio flux at band 10.7 cm alongside with the sunspot number. Data source of radio flux data: (NOAA Solar–Terrestrial Physics Division, 2013)

Being relatively young index, data is only available for solar cycles 19 – 23. For the cycles shown the cycle length is 10.9 years at average, a result replicated by ISSN for the same cycles. Curiously, the properties of solar radio flux correlate very well with sunspot number with the exception of, perhaps, the length of minimum phase that (excluding the nonexistent value of SC23) has a correlation coefficient of only 0.65.

To make further comparisons the solar radio flux has been plotted during SC23 alongside ISSN in Figure 4.8. Figure shows that in radio flux the phases begin slightly later than in ISSN, especially the minimum phase that begins several months after indicated in ISSN. The radio flux follows ISSN very well, the most notable deviation in the shape of the two curves being that radio flux peaks higher during the second maximum instead of the first like ISSN does. This implies that because 10.7 cm radio flux is a product of the bremsstrahlung within the corona, the coronal processes are actually more active during the second maximum than the first one. The difference between them is approximately 20 sfu, which is to say the second maximum is 18% higher than the first one if both are compared to the level of the solar radio flux during solar minimum.
4.1.5 Solar X-ray flux

Solar-disc integrated X-ray flux is shown in Figure 4.9. The data has been retrieved from the NOAA Geostationary Operational Environmental Satellites (GOES) 6–10 and 12. Satellites 5 and 11 were omitted due to low amount of usable data. This data set spans solar cycles 22 and 23.

In SC22 X-ray flux follows the sunspot number well, having high peaks during both maximum. The flux as shown here is more quirky in SC23 though the double peaked structure can still be seen. Other studies have reached the same conclusion (Ramesh and Rohini, 2008). The bar-like appearances in hard X-ray flux in SC23 are the result of several very strong flares in a background of average activity. Particularly the ones of 2003 and 2005 are due to some of the most powerful flares recorded, the numbers 1 and 4 in the record listing (www.spaceweather.com/solarflares/topflares.html as of 21st Nov 2013).
Figure 4.8: Smoothed solar radio flux at band 10.7 cm during SC 23. Changes of phase have been marked with dashed lines, the minimum phase being left out of the picture on the right. Data source of radio flux data: (NOAA Solar–Terrestrial Physics Division, 2013)

Figure 4.9: Yearly smoothed solar X-ray flux showing two different channels in several GOES satellites. For this picture data from GOES satellites 6–10 and 12 was used. Short X-ray flux means channel 0.05 – 0.4 nm (red) for GOES 6–7 and 0.05 – 0.3 nm for GOES 8–12, long X-ray flux respectively 0.1 – 0.8 nm (blue) for all of them, single color presenting all the data sets of similar type. The sunspot number is plotted for comparison. Data source of X-ray data: NOAA http://www.ngdc.noaa.gov/goes/sem/getData
4.1.6 Comparison of solar activity measures

Figures 4.10, 4.11 and 4.12 show the solar cycle lengths, the lengths of activity phases and statistical variables of several activity indices during the recorded solar cycles. In each one the average of the associated property in ISSN is shown in dashed line.

Figure 4.10: Solar cycle length shown as a function of the cycle number for different activity indices. ISSN = International sunspot number, SSAI = Sunspot area index, ARSN = American relative sunspot number. The first ten cycles were omitted from the picture for sake of illustration because they were not represented by the other indices. The dashed black line shows the average of ISSN over all 23 cycles.

All designed to provide an estimate for the solar activity, it is no surprise that all of the indices (ISSN, SSAI, ARSN, solar radio flux) show close to the same length in the solar cycle. The first ARSN cycle seen is too short because the number was began to be recorded in the beginning of the ascending phase of the cycle. In all the figures here the SC18 of ARSN is thus biased. Other than that the highest difference between the lengths indicated by two indices is less than a year.

At average the ascending phase has been short for almost the whole modern maximum of cycles 15–23. On the other hand the maximum phase has been longer than average in the same time period. There is no such peculiarity in the declining and minimum phases. A behaviour akin to zigzag between cycles can be seen in all the solar phases and lengths with even and odd cycles taking turns being the longer ones.

Finally, skewness and kurtosis of the cycles were determined and shown in
Figure 4.11: Length of solar cycle phases as a function of the cycle number for different activity indices. ISSN = International sunspot number, SSAI = Sunspot area index, ARSN = American relative sunspot number. The black line on each figure shows the average of the associated phase length in ISSN.

Figure 4.12. The zigzag motion between the cycles is even more pronounced here. At average the absolute difference of the properties between two consecutive cycles is 0.24 for skewness and 0.26 for kurtosis. The behaviour can be interpreted as a consequence of the solar cycle actually being the one of Hall cycle, 22 years. However, the behaviour has exceptions as in SC3 and SC8 and maybe SC18 where values presented by ISSN and SSAI contradict each other.

However statistical quantity, kurtosis values of SC18–SC22 are remarkably close to each other having only 0.033 as the average of absolute difference in the kurtosis of ISSN. Between these cycles the ISSN skewness goes through two consecutive shifts in the parity that could indicate a change in the behaviour of the Sun.
Figure 4.12: Top figure: Skewness of a cycle for different activity indices. Bottom figure: The same for kurtosis. Colors are the same as in 4.10 and 4.11.
4.2 Measures of geomagnetic activity

Although developed for the analysis of solar activity, the definition of Section 2.1.5 may be used with other indices as well. The indices of geomagnetic activity follow the solar activity cycle indirectly. Using this observation as a guideline the minima defining the cycle can be picked manually and the corresponding properties of the cycle derived. However, geomagnetic indices look more irregular than solar activity measures, so the phases were omitted from the following analysis.

Definition of Section 2.1.5 is used on the geomagnetic indices introduced in Section 2.4.4 with the addition of the rule that only the closest minima to the minima of international sunspot number, ISSN, were chosen, the rest in the middle of the solar cycle omitted when defining the cycle. Otherwise there would have been several geomagnetic cycles inside a single solar cycle. It is however interesting to investigate its reasons and the topic is discussed within the sections below.

The same analysis as committed in Section 4.1 is done on each full cycle with the exceptions mentioned above. These indices are further compared to the space weather events in Chapter 5 to shed light on their causes. Also SC23 is inspected more closely.

4.2.1 Ap and aa indices

The related Ap and aa indices are plotted in Figure 4.13 smoothed over for a yearly average. Ap index data is from wdc.kugi.kyoto-u.ac.jp/kp/ and aa index data from ISGI website isgi_latmos_ipsl_fr/source/indices/aa/.

For Ap index the data comprises solar cycles 17 – 23, while aa index has been fully constructed since SC12.

To start, length of the cycles is 10.75 years at average for Ap index, which is remarkably close to the length of the sunspot number cycle over the same cycles (10.73 years). The respective number for aa index was 10.90 years and for ISSN for the same cycles 10.85 years. This was to be expected as the minima closest to the sunspot cycle minima were chosen. The cycle-defining minima of both indices were close to the ones of the sunspot number, the highest difference being for SC22, for which both indices were over 1 year late from the sunspot minimum.

The development of both indices in SC23 is shown in Figure 4.14. There was a single maximum year that did not coincide with the sunspot cycle. Comparing to the sunspot cycle the index has approximately twice the value than what it has during the minimum years. This is excluding the maximum period of 2003 when the index skyrocketted to the average value more than
The peak coincides with the period of Halloween storms of 2003 (Weaver et al., 2004), which is convenient because geomagnetic storms are one of the phenomena that both Ap and aa indices were created to monitor.

### 4.2.2 Dst, Dcx and Dxt indices

Shown in Figure 4.15 is the full Dst index 1957–2012 smoothed over a year. Similar indices Dcx and Dxt are also included for comparison. Dst index data used here is from Kyoto Dst index service, [http://wdc.kugi.kyoto-u.ac.jp/dstdir](http://wdc.kugi.kyoto-u.ac.jp/dstdir). Respectively the data of Dcx and Dxt indices are from Dcx index server, [http://dcx.oulu.fi/](http://dcx.oulu.fi/).

Dst index is highly peaked (the kurtosis ranges from 2.1 to 3.5 with deep local minima inside the sunspot cycles\(^1\)). Every solar cycle the index has a well-defined minimum around the sunspot minimum nevertheless. This is true of Dcx and Dxt indices as well but their kurtosises are even higher during some cycles.

On the other hand the skewness values seem to be quite random, as even between indices in the same cycle the sign of the property can vary. Skewness does not seem particularly useful for these indices.

---

\(^1\)Note that in Figure 4.15 the range has been reversed and the terminology here reflects that.
A closer look to SC23 is shown in Figure 4.16. The average Dst index stays between $-10$ and $-20$ nT for the whole cycle with the highest average disturbances reaching $-25$ nT. Dxt follows Dst quite well but the Dcx index, corrected for seasonal variation, goes its own way though still correlating slightly with the other two indices. Just like in Ap index, the Halloween storm period of 2003 is seen as a peak on the picture going to as low as $-25$ nT and it is indeed the lowest value of Dst and Dxt during this period. However, Dcx does not acknowledge this, the minimum staying at a measly $-15$ nT.
Figure 4.14: Smoothed Ap and aa indices during SC23. Sunspot cycle is plotted alongside them with its phases marked with vertical dashed lines. Data source of Ap index is Kyoto World Data Center [http://wdc.kugi.kyoto-u.ac.jp/kp/index.html](http://wdc.kugi.kyoto-u.ac.jp/kp/index.html) and for aa index ISGI website [isgi_latmos_ipsl_fr/source/indices/aa/](http://isgi_latmos_ipsl_fr/source/indices/aa/)

Figure 4.15: Smoothed Dst, Dcx and Dxt indices. ISSN is shown for comparison. Data sources are given in the text.
Figure 4.16: Smoothed Dst, Dxt and Dcx indices during SC23. ISSN is given for comparison. Data sources are given in the text.
4.2.3 Auroral electrojet indices

Figure 4.17 shows the three auroral electrojet indices AE, AU and AL alongside with the sunspot number. Unfortunately, these indices suffer from at least two major data gaps, the first one in 1977 situated in the minimum of SC20, and the second one in 1988 during maximum of SC22.

As with Ap and aa indices, all three of these show a correlation with the sunspot number having minima approximately at the same time as the solar index. The two-peaked structure of the sunspot cycle is at times dramatically powerful in the auroral electrojet indices that have quite a low local minima in the vicinity of the sunspot maximum.

![Auroral electrojet indices](image)

Figure 4.17: Smoothed AE, AU and AL indices alongside with ISSN. The sign of AL index was inverted to show better its correlation with the rest of the auroral electrojet indices. Two data gaps are shown with pointer arrows. Data source of AE indices: [wdc.kugi.kyoto-u.ac.jp/dstae/index.html](wdc.kugi.kyoto-u.ac.jp/dstae/index.html)

Also notable is that the indices are stronger in the declining phase of the sunspot cycle than in the ascending phase. In all of the solar cycles available to the data the indices stay strong far into the declining phase only to fall sharply near the sunspot minimum. Since the indices were designed to monitor the substorm activity it is not surprising that this trend also exhibits itself in the substorm occurrence rate in Figure 5.7 though admittedly it only shows SC23.

Figure 4.18 shows the auroral electrojet indices in more detail during SC23. A year after the ascending phase beginning in the sunspot number the auroral electrojet indices began to climb. They had a local minimum during the sunspot maximum but a very sharp maximum after that in the
declining phase telling of the high (sub)storm activity in 2003. There was another maximum in 2005 that was also known as a stormy year (Tanskanen, 2009). Finally the minimum was reached in the latter half of 2009. It is notable that this minimum was the lowest of all time in the series of auroral electrojet indices (omitting the data gaps in Figure 4.17), and indeed it was so low that only by 2012 did the average reach even the lowest values of the minimum of the previous cycle SC22.

Auroral electrojet indices composed of the data of the IMAGE network are shown in Figure 4.19. As the indices are quite new they are only computed for SC23. The smoothed index is quite similar to the smoothed AE index but there is a qualitative difference: inverted IL is lower than IU in contrast to the inverted AL being higher than AU. This is because the IMAGE indices measure the X (geographic north) component of the magnetic disturbance, while AE indices use the H (horizontal) component. In IMAGE index this relationship seems to flip shortly after the sunspot minimum but only future measurements can show if this is only a brief detour from normal. AE indices also begin to climb faster after the SC23 minimum but again, unfortunately, the IMAGE data is incomplete in this period to show if this is mutual to that index.
Figure 4.18: Smoothed AE, AU and AL indices over SC23 along with ISSN. The sign of AL index has been inverted to show better its correlation with the rest of the auroral electrojet indices. Data source: wdc.kugi.kyoto-u.ac.jp/dstae/index.html

Figure 4.19: Smoothed IE, IU and IL indices alongside with ISSN. The sign of IL index was inverted to show better its correlation with the rest of the auroral electrojet indices. Data source of IMAGE indices: http://substormzoo.org
Finally, it might be fruitful to study the differences of the two indices. Due to its longer latitudinal reach the IMAGE indices ought to see changes in the substorm activity earlier than AE indices during the solar cycle. IL index is the one used to monitor substorm activity in the IMAGE UT, so it is natural to study the difference between AL and IL indices. This is shown in Figure 4.20.

Figure 4.20: Smoothed property AL–IL, the sign reverted, over SC23. Sunspot number is shown in comparison.
4.2.4 Polar cap indices

Figure 4.21 shows the polar cap indices derived from the data of stations in both the northern (Thule) and the southern (Vostok) hemispheres. The data have been smoothed over for a yearly average, a result of which is two data sets both trapped between 0.5 and 1.75 nT. The 15-minute averages vary between $-12.8 - +33.8$ (Thule) and $-24.8 - +28.8$ (Vostok) implying that even though there are high negative disturbances, the positive ones outnumber them. However at average they do not fluctuate lot, having standard deviations 0.24 (Thule) and 0.15 (Vostok).

Vostok data has two large data gaps at 1980–1982 and 1992–1996. These have been removed from the plot. Thule data has one as well in 2010 seen as a discontinuity in Figure 4.21. Both data were provided by NOAA at ftp: //ftp.ngdc.noaa.gov/STP/SOLAR_DATA/RELATED_INDICES/PC_INDEX/.

![Polar cap indices](image)

Figure 4.21: Smoothed PC index of Thule (blue) and Vostok (red). Data source: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/RELATED_INDICES/PC_INDEX/

It is difficult to draw any conclusions of the PC Vostok index due to the inconvenient data gaps. Thule data shows two or more peaks every solar cycle (SC21–SC23) that the data comprises, as well as a minimum shortly after each sunspot minimum though the last one situates itself during a data gap.

A zoomed in Figure 4.22 of the PC indices during SC23 shows as its most prominent feature the maximum of PC Thule index in 2003. Maybe due to seasonal variance the Vostok index does not agree with the other index but instead seems to have a broken maximum during 2005. This is however not
conclusive as PC Vostok varies little during the cycle, the average staying between 0.7 and 1.3. It would still be interesting to compare these results with the solar wind velocity and IMF that the PC index has been designed to monitor.

Figure 4.22: Smoothed PC index of Thule (blue) and Vostok (red) over SC23 as defined by the sunspot number ISSN. Data source: ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/RELATED_INDICES/PC_INDEX/
Chapter 5

Analysis of the space weather events of solar cycle 23

In this Chapter I will present the results of the statistical analysis of space weather events during SC23. This corresponds roughly to years 1996–2010. First, I will introduce the data sources (Section 5.1) and the how the fluctuations were identified (Section 5.2). Second, I will show the occurrence rate of events identified on the Sun (Section 5.3), ones detected in the solar wind close to the Earth (Section 5.4), and finally in the magnetosphere and on the ground (Section 5.5). I will also show the results on identifying ULF fluctuations in the solar wind and Pc5 pulsations within the magnetosphere.

5.1 Data and its sources

Two kinds of data were analyzed in this Chapter: event data and in situ data. Event data consists of events of CMEs, ICMEs, HSSs, geomagnetic storms, auroral substorms and flares determined from observations, different magnetic indices and magnetic field data. It usually includes at least the time of event and several properties subject to the event, like e.g. magnetic field strength or the length of the event. In situ data consists of the actual magnetic field and cosmic ray data measured by satellites and ground stations. In Table 5.1 the data sources are listed for quick reference.

The CME catalog used is generated and maintained at the CDAW Data Center by NASA and The Catholic University of America in cooperation with the Naval Research LAboratory. SOHO is a spacecraft project of international cooperation between ESA and NASA. ICME data were acquired from two different sources. For the events of 1996–2007 the catalog on the website http://www.ssg.sr.unh.edu/mag/ace/ACElists/ICMEtable.
Data type | Data sources
---|---
ACE magnetic field | Goddard Space Flight Center (2013)
CIR events | Esequiel Esher (personal communication)
CME events | Gopalswamy et al. (2009)
Flare events | NOAA SWPC, through [http://www.substormzoo.org](http://www.substormzoo.org)
GCR events | Eskdalemuir observatory
HSS events | Maris and Maris (2013)
| and Emilia Kilpua 2007-2011
Ground magnetic field | FMI/IMAGE
Storms | FMI
Substorms | FMI
Wind magnetic field | Goddard Space Flight Center (2013)

Table 5.1: List of data sources

html was used (Richardson and Cane, 2007). Another catalog by Emilia Kilpua (personal communication) was used for the years 2008–2011 and the two were combined for the most encompassing results. Nevertheless the use of two different data sets has to be taken into account when analyzing the data.

High-speed streams were obtained from Maris and Maris (2013). Geomagnetic storms and auroral substorms were a courtesy of Substorm Zoo [http://www.substormzoo.org](http://www.substormzoo.org) (Tanskanen et al., 2011). For the GCR flux data from Eskdalemuir observatory was used. ACE (Stone et al., 1998) and Wind mission data (Russell, 1995) were acquired through NOAA CDAWeb (NASA Goddard Space Flight Center, 2013). The magnetic field data from Magnetic Fields Experiment (MAG) instrument on ACE (Smith et al., 1998) was used extensively, as was the magnetic field data from Magnetic Field Investigation (MFI) instrument on Wind (Lepping et al., 1995).

## 5.2 Identifying fluctuations

Spectral analysis was used on the magnetic field data of ACE and Wind satellites and magnetometer data of the magnetic observatories of Ouluunjärvi, Kevo and Kilpisjärvi (Section 5.5.3) to find periodic phenomena from the solar wind and magnetosphere. Task was to find out if and when the fluctuations would penetrate from solar wind through the magnetopause to the magnetosphere and onward through the ionosphere to the ground to be measured by magnetic observatories on the ground.
For this data the power spectra was determined using Blackman–Harris window with filtering-squaring-averaging over 7 data points for ACE data and 5 for the rest. An event was recorded if a specified fluctuation band maximum was found between 2 and 7 mHz during a single hour. Bandwidth 0.5 mHz was used when studying the years 1998–2008. To register as ULF the power maximum in the power spectrum has to be at least 1 that is near the average of the power maxima in ACE and Wind data. For Pc5 pulsations in the ground magnetometers the associated power maximum threshold was chosen to be 0.2, close to their power maxima. These thresholds were chosen because they were close to their averages and high enough to be considered real. For all the instruments the horizontal component of the magnetic field was used.

5.3 Solar disturbances

5.3.1 Solar flares

Figure 5.1: Monthly occurrence rate of different classes of flares during SC23. Data source: NOAA SWPC, downloaded from http://www.substormzoo.org.

Occurrence rate of different classes of flares is shown in Figure 5.1. The dominant feature of the figure is the number of C class flares that correlate well with the sunspot number over the solar cycle. Less obvious is the correlation of M and X class flares with the sunspot number as they are less numerous than the rather low energy flares of M class.
The ascending phase of SC23 began with an increasing rate of C and M class flares. The occurrence rate of both grew steadily with the solar cycle with sharp falls and rises during the maximum. Both local sunspot maxima are also maxima within the occurrence rates of C and M classes. However the X class flares, the ones with the highest energies, were equally prominent from the maximum until the end of the declining phase. Sun almost seizes all flare activity in its minimum phase when all the flares seem to come in bursts of several flares one month while none can be detected during other times.

The flare data shown here has been collected and interpreted from the large amount of X-ray sensor data of GOES satellites. Being geostationary the GOES satellites can only detect flares that burst from Sun on the Earthward face of the star. Using a window of 1/12 year as done here can be used as an approximation to the activity of the whole Sun as its rotation rate is approximately 1 month. Still it has to be considered that the activity on the unseen face of the Sun could be different from the seen face.

5.3.2 Coronal mass ejections

Occurrence rate of CMEs of different linear velocities during SC23 is shown in Figure 5.2. Because of earlier studies (Webb and Howard, 1994) it was to be expected that the rate was the highest during the sunspot maximum, as is verified by the figure.

![CME count with different linear velocities](image)

Figure 5.2: Occurrence rate of CMEs of different velocities during SC23 averaged over a year. ISSN is shown for comparison. Data source of CME events: Gopalswamy et al. (2009)
The picture shown by the figure is very similar to earlier study of the same data set (Gopalswamy et al., 2003). An occurrence rate of 0.5 CME per day was reported during minimum of SC22 and 6 CME per day during the maximum years. The same information can be seen in Figure 5.2. From Figure 5.2 we see that CME rate never drops below 3 per day during the declining phase. Finally, in the minimum phase it falls to 2 CME per day but not below. This suggests that the CME activity of the minimum of SC23 was higher than the one of SC22, unless the method of CME detection has improved or changed. However, almost all the CMEs were of the slow type.

Figure 5.2 shows the CME rate of all velocities correlating reasonably well during the maximum years but several deviations occurred in the other phases. For example a peculiar time interval occurred in 2007 when the CME rate bulges to 4 CME per day. It is notable that this is almost entirely caused by CMEs of rather low velocity $v \leq 400$ km/s. In the last part of the data a sudden increase of CME rate was detected as a prelude to SC24. The occurrence rate has been particularly high, as it has already matched the rate during the maximum phase of SC23. However, most of these CMEs were slow. In the beginning of 2010 the rate of faster CMEs began climbing implying that the ascending phase was beginning during that year.

CMEs of the CME catalog used here have been identified manually from SOHO/LASCO data (Gopalswamy et al., 2009). This raises questions of how large a part of CMEs have been left unidentified and whether the identification practice has improved later on. If the latter has affected the building of the later parts of the catalog, it might explain at least part of the rather large CME rate difference between minima of SC22 and SC23. To check for this possibility a further study using other CME data sets would be useful.

5.4 Solar wind disturbances

Four kinds of events were examined in the solar wind: interplanetary coronal mass ejections (ICME), high-speed streams (HSS), corotating interaction regions (CIR) and magnetic field fluctuations of ultra-low frequency (ULF). Their occurrence rates were plotted and compared to the solar activity. ULFs are treated in Section 5.5.3 alongside with magnetospheric Pc5 pulsations.

5.4.1 Interplanetary coronal mass ejections

Occurrence rate of ICMEs is shown in Figure 5.3. In addition, the occurrence rate of ICMEs of different velocities, divided into portions of more or less 450 km/s are shown as well. Earlier study supports the results herein (Cane and
Richardson, 2003). For this study the event list (Cane and Richardson, 2003) was supplemented by a list of events in 2007–2010 provided by Emilia Kilpua (private communication).

Figure 5.3: Occurrence rate of ICMEs and their slow and fast portions at the Earth averaged over a year. Rate is given in ICMEs per year. Data source: Richardson and Cane (2007) and Emilia Kilpua’s catalog

Referring back to Section 5.3.2 and Figure 5.2 it is immediately noted that the ICME rate peaked at the sunspot maximum just like the CME rate. This is to be expected if ICMEs are indeed originally CMEs and their paths are not biased. Figure 5.3 shows a rather good correlation of ICME rate with both CME rate and sunspot number. In particular, the two peaks at both sunspot maxima show a higher number of ICMEs than at other times. There is also a local minimum in the rate between the two maxima not explained by the changes in the CME rate during this time. No ICMEs were observed for two months.

From the solar cycle definition’s point of view (see 2.1.5) it is especially interesting that the ICME rate goes up to 2/month basis just as the ascending phase starts and falls permanently under this right after the shift from the declining phase to the minimum of the sunspot number. During the minimum only 1 ICME per month at average was observed.

The slow ICMEs, i.e. the ones with velocity lower than 450 km/s, are the main portion of ICMEs in the ascending and minimum phases. However, both portions are almost equally important during the maximum phase. In the declining phase the faster portion seems to dominate only to dwindle to almost nonexistent close to the minimum phase. From mere statistical
analysis it is difficult to deduce what causes the higher portion of fast ICMEs in the declining phase. Interestingly, the ICME rate has not gone up at the beginning of the SC24 even though the change can be seen in the CME rate (Figure 5.2). The reason could have to do with the CMEs being too weak and slow to be detectable at 1 AU.

### 5.4.2 High-speed streams

Occurrence rate of HSS events is shown in Figure 5.4. Data begins in 1996 and lasts until 2008 (Maris and Maris, 2013). The occurrence rates of HSS events of different origin are shown as well.

![Number of HSS of different origins](image)

Figure 5.4: Occurrence rate of different HSS events at the Earth averaged over a year. The rate is given in events/year. An adjusted smoothed ISSN is shown for comparison. Data source: Maris and Maris (2013)

Just like ICMEs, HSS events were scarce before the ascending phase of the solar cycle, and during the next minimum phase they became rare again. Unlike ICMEs, there are a couple of events each month even during the time of lowest activity.

HSS activity remains strong throughout the solar cycle having a maximum during the declining phase in October–November 2003, the famous period of Halloween storm 2003 (Weaver et al., 2004). This maximum is due to both coronal hole originated HSS and less due to a minor peak in the HSS of flare origin. Overall during the solar cycle the rate of HSS of the coronal hole origin reached a maximum just before and after the two maxima in the sunspot number. The maximum in the flare originated HSS rate happened
expectedly during the sunspot maximum when the flares themselves are the most numerous.

Two smaller rate maxima happened during the summers of 2005 and 2007. In the beginning of 2005 an anomalous geomagnetic storm with storm main phase developing during northward interplanetary magnetic field (IMF) occurred (Du et al., 2008). According to data there was a flare-originated HSS event occurring at the exact same time with southward IMF (Maris and Maris, 2013).

The occurrence rate maximum of 2007 during the solar minimum correlates with the peak CME rate of 2007 (see Figure 5.2). As CMEs and HSSs have different points of origin, CMEs arising mainly from the active regions and HSSs from the coronal holes, this suggests a connection between the two types of events. Further study is needed to verify if this was not just a statistical fluctuation.

5.4.3 Corotating interaction regions

Figure 5.5 shows the occurrence rate of corotating interaction regions, CIRs, within solar wind from 1964 until the declining phase of SC23. The yearly number of CIRs has increased peculiarly after the minimum of SC22 (around 1996) to outnumber the associated number in any of the cycles before. The effect might be caused by the increased data of solar wind and the associated increase in the efficiency of CIR detection. Otherwise, the number of CIRs has been spectacular when compared to earlier solar cycles. It is seen that before SC23 the highest CIR counts were associated with the declining and minimum phases of the cycles.

5.5 Magnetospheric events

Magnetosphere is home to several fluctuation phenomena of magnetic field: geomagnetic storms, auroral substorms and geomagnetic pulsations (in here specifically Pc5 pulsations). In addition the flux of galactic cosmic rays is measured with ground-based cosmic ray observatories.

5.5.1 Geomagnetic storms

Mean number of geomagnetic storms during SC23 is shown in Figure 5.6. Figure shows the occurrence rate growing during the ascending phase, fluctuating over the cycle. In the declining phase in 2003 the number escalates to 15 per month making that the stormiest year of the cycle. This is in
accordance with all the geomagnetic activity indices showing high activity peaks in 2003.

Comparing to the occurrence rates of events of solar wind several observations can be made. Storms peak at both sunspot maxima and in 2005 when also CME (and ICME) numbers peak. However that does not explain the high maximum of 2003. That one corresponds to the peak in HSS rate which in part corresponds to the cycle maximum in the number of HSS originating from coronal holes. While not convincing, this behaviour implies a link between the two phenomena. However, there are peaks, albeit lower, in the number of HSS of coronal origin in 1999 and 2007 that actually do not appear to correlate with higher storm activity. No causality between the two can be claimed.

Peculiarly, the storm activity falls to almost none in the minimum of SC23, while never falling lower than 9/month (in the yearly average) in the minimum of SC22.

5.5.2 Auroral substorms

Mean number of geomagnetic storms during SC23 is shown in Figure 5.7. Because substorms often (but not always) occur at the same time as geomagnetic storms, it can be expected that substorm rate peaks at the same
Figure 5.6: Occurrence rate of geomagnetic storms during SC23. Sunspot number is shown as a reference. Data source: http://www.substormzoo.org

times as storm rate. Indeed, this seems to be the case: there are local peaks in the ascending phase, the both local sunspot maxima and the declining phase in 2005 and a grand maximum in 2003.

The same conclusions can be made of auroral substorms as geomagnetic storms if no further separation is made by separating substorms that occur within and without associated geomagnetic storms. However, substorm activity never ends even in the minimum: it is unclear what causes the substorms during this period of supposed inactivity of the Sun.

5.5.3 Pc5 pulsations and ULF fluctuations

ULF and Pc5 fluctuations (see Sections 2.4.3) were identified from the solar wind using the Fourier method described in Section 3.2.7. Data from ACE magnetometer MFE (Smith et al., 1998) and Wind MFI instrument (Lepping et al., 1995) was used for ULF fluctuations. For Pc5 pulsations data originated from magnetic observatories of Oulu (OUJ), Kilpisjärvi (KIL) and Kevo (KEV). Fluctuations were checked for every hour.

The highest resolution ACE/MFE data was used, i.e. a resolution of 1 second. For the Wind/MFI data only the 3-second data was used despite data of 0.1-second resolution been available. The higher resolution data was not used due to reasons of high memory usage and the longer computing times. Nyquist’s frequency for the ULFs is at most 14 mHz (twice the 7 mHz of Pc5 pulsations). This is a factor of 25 lower than the frequency used by
Figure 5.7: Occurrence rate of auroral substorms during SC23. Sunspot number is shown as a reference. Data source: http://www.substormzoo.org

Wind/MFI 3-second data and therefore it is very adequate for the task.

For the magnetic observatories the data resolution was 10 seconds. This is in theory adequate for detection of 100 mHz frequencies that is yet over 7 times higher than the Nyquist frequency of the highest ULF frequency. As at least a factor of 10 is preferred, it is less optimal than the spacecraft data described above but technically sufficient for the task. In the future, the analysis should be repeated for a data of higher resolution.

**Fluctuations**

Occurrence rate of potential ULF and magnetospheric Pc5 events is shown in Figure 5.8. The events were identified using the procedure described in Section 3.2.7 and the specifics in Section 5.2.

The ground-based magnetometers show internally consistent data with higher latitude stations ($\lambda$(KEV) = 69.76$^\circ$, $\lambda$(KIL) = 69.06$^\circ$, $\lambda$(OUJ) = 64.52$^\circ$) showing larger number of Pc5 events and thus higher power than their lower cousins due to their proximity with the auroral oval. Regrettably KIL data is missing the year 2002 leaving a gap in the data. Despite this shortcoming it still shows a good correlation with the data from KEV that is close by.

The most significant feature of Figure 5.8 is all of the five data sets peaking at approximately the same time in 2003. Consequently this was also the (sub)stormiest time period of SC23 indicating that ULF/Pc5 power plays a role in the geomagnetic storms and auroral substorms.
Figure 5.8: Occurrence rate of potential ULF events identified from ACE/MFE and Wind/MFI magnetometer data and Pc5 events identified from magnetometer data of Kilpisjärvi (KIL), Kevo (KEV) and Oulunjärvi (OUJ). Sunspot number is shown as a reference. Satellite data is from NASA CDAMWeb [http://cdaweb.gsfc.nasa.gov/istp_public/] and the ground-based magnetometer data from IMAGE website [http://www.ava.fmi.fi/image/].

This also implies that ULF waves in the solar wind and magnetosphere along with the magnetospheric Pc5 pulsations have an intrinsic connection other than mere similar frequencies. Either they are being created by different processes in all the regions at the same time or, which is an interesting possibility, the fluctuations traverse through from the solar wind into the magnetosphere and to the ground acting as a mechanism for the solar wind-magnetosphere-ionosphere coupling. Further study is required to confirm or compromise either of these possibilities. If they are indeed connected, the high latitude magnetic observatories of sufficient time resolution could be used to measure the approximate number of ULF fluctuations in space.

5.5.4 Galactic cosmic rays

Galactic cosmic ray flux from a single cosmic ray station (Eskdalemuir/ESK) is shown in Figure 5.9. Figure shows a definite anticorrelation between the sunspot number and cosmic ray flux confirming that the magnetic activity of the Sun effectively modulates the influx of galactic cosmic rays. The higher the sunspot number count during the cycle, the lower the GCR flux although
they do not have to reach their extrema at the same time (see especially SC21, i.e. the solar cycle around 1980).

Figure 5.9: Occurrence rate of galactic cosmic rays alongside with the smoothed sunspot number. Data source: courtesy of Eskdalemuir cosmic ray observatory.
Chapter 6

Conclusions

6.1 Results and their short-comings

In this study several solar and geomagnetic activity indices were analyzed along with the events of solar, solar wind and magnetospheric origin in the Chapters 4 and 5. The completeness of the analysis on each index and event type depended heavily on the completeness of the data set. This property was the single most important error source in the study. Data caps and bad data was usually easy to identify in the data.

Direct linear interpolation was used only with the analysis of ULF and Pc5 events in Section 5.5.3. The most severe cases were omitted resulting in periods of seemingly absent fluctuation activity when the missing data could have arisen from strong magnetic activity that rendered the instruments shut. In the future better and more stable instruments could work around this issue.

A small error is attributed to the MATLAB function \( y_{\text{frac}} \) (see Appendix D.1) that calculates the fraction of a year. When the period of time studied includes a leap year (i.e. always), that year’s fraction is a tiny bit different to other years due to slightly different number of days. The function practically makes every year of equal length resulting in an error of order 0.3 % in the calculation of averages of the event type data. It makes no practical difference in the figures but it could have some effect on the calculation of statistical properties. It was however disregarded because its significance was considered low in the scope of this study.

Some other events could have been studied as well but were omitted because of limited time and data. These include but are not strictly limited to group sunspot number, solar irradiance flux, sunspot numbers of different sunspot complexity types and the solar wind speed at 1 AU. The number of ground-level enhancements of galactic cosmic rays was also omitted because
of such a low amount of recorded events. Further studies on these and other subjects would improve the quality of the analysis.

6.2 Answers and further research

In Chapter 1 three questions were asked:

- When was the most active phase of the solar cycle 23 in terms of activity of the Sun?
- What about in terms of space weather?
- What can be said about the activity of Sun and space weather during any single cycle in general?

Regarding the data of solar activity indices and solar-originated events the first question was easy to answer: the most active phase was the maximum phase, specifically its two peaks and not the gap between.

The space weather side of the questions begged to differ: practically all the indices of geomagnetic activity screamed that space weather was the most stormy in the declining phase of the solar cycle. This was concluded also by the large number of storms and substorms during this time and the high amount of ULF and Pc5 events coinciding with the period.

Considering the other solar cycles it is not possible to say anything as conclusive mostly due to missing data. However, geomagnetic indices AE, PC, aa, Ap and Dst/Dxt/Dcx do suggest that the declining phase was the most active phase of the solar cycles SC20 and all but aa and Ap SC21 but not of SC20 and SC22 when the maximum phase was the most active instead.

Going further back gives us only aa, Ap and Dxt/Dcx until SC17 and even further back only aa and Ap on to rely. Using only the latter since SC15, the beginning of the modern solar maximum, will give five cycles with highest activity in the declining phase and three in the maximum phase respectively and a single one (SC21) with one in their transition. While not statistically conclusive, this implies that the declining phase is sometimes if not generally the stormiest period of the solar cycle. In the modern maximum this certainly was the case.

An explanation could be that in the declining phase and the late maximum phase there is still a large amount of active regions and sunspots associated with them. However they are closer to the solar equator due to Maunder’s butterfly effect of sunspots migrating towards the equator towards the end of the cycle. Active regions are often sites of flare eruptions
that can result in CME eruptions that could then more easily cause space weather effects.

6.3 Conclusions

In this thesis six different solar activity measures and 13 different related indices of geomagnetic activity were analyzed along with nine event types of either solar or space weather origin. Many relations between them were concluded and several known results were verified.

The definition of solar cycle phases here-in will make it possible to explicitly compare the phases of two different cycles. Our results have established that the solar phases are inherently different to each other and a classification scheme like this can be useful.

We also conclude that the declining phase might be the most active solar phase in a space weather point of view.

Last but not least was the identification of ULF fluctuations in space-based instruments and their tremendously good correlation with the Pc5 pulsations identified from Earth-based magnetometers. ULF waves could turn out to be one of the mechanisms coupling the solar wind-magnetosphere-ionosphere system together. Further study is required to establish this relation and find the possible relation as to how the pulsations are able to traverse in and through the system.
Appendix A

Principles of magnetohydrodynamics

Equations of magnetohydrodynamics (MHD) have been formulated in numerous books (Boyd and Sanderson, 2003; Schrijver and Siscoe, 2009; Craves, 1997, see e.g.).

In MHD several assumptions are made of the behaviour of the plasma (Boyd and Sanderson, 2003):

- Plasma is a fluid that can be treated as a continuous medium.
- Being a quasi-neutral charged fluid it is coupled by the Maxwell’s equations.
- Electromagnetic fields vary on the same time and length scales as the plasma variables.
- No relativistic motion is assumed.

First part allows the treatment of plasma through the hydrodynamic equations of continuity, conservation of momentum (Navier–Stokes equation) and conservation of energy. Second part provides constraints and introduces the force term in momentum equation but most importantly adds the induction equation to the mix. Third part lets to neglect the electrostatic force $qE$ and together with the fourth part also the displacement current $\epsilon_0\mu_0\partial E/\partial t$.

From the above principles the equations of MHD can be derived:

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{u}$$  \hspace{1cm} (A.1)

$$\rho \frac{D\mathbf{u}}{Dt} = (\nabla \times \mathbf{B}) \times \mathbf{B}/\mu_0 - \nabla P$$  \hspace{1cm} (A.2)
\[ \frac{DP}{Dt} = -\gamma P \nabla \cdot \mathbf{u} + (\gamma - 1) \nabla \cdot (\kappa \nabla T) + \frac{\gamma - 1}{\sigma} j^2 + (\gamma - 1) \mu \left( \frac{\partial u_i}{\partial r_j} + \frac{\partial u_j}{\partial r_i} - \frac{2}{3} \delta_{ij} \nabla \cdot \mathbf{u} \right) \frac{\partial u_i}{\partial r_j} \]  

(A.3)

\[ \frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\sigma \mu_0} \nabla^2 \mathbf{B} + \nabla \times (\mathbf{u} \times \mathbf{B}) \]  

(A.4)

where \( \rho \) and \( \mathbf{u} \) are the particle density and velocity, \( \mathbf{B} \) is the magnetic field, and \( P, \gamma, \sigma \) and \( \mu_0 \) are the pressure, ratio of specific heats, conductivity and the permeability of vacuum.

Equations (A.1) – (A.3) are better known as the equations of continuity and the conservation of momentum and energy, respectively. Equation (A.4) is the induction equation.

In practice more approximations, like the equation of state and the Ohm’s law, are required for meaningful calculations.
Appendix B

Some mathematical methods

B.1 Stokes theorem

Stokes’ theorem states (Arfken et al., 2013)

\[ \oint_C \mathbf{V} \cdot d\lambda = \int_S \nabla \times \mathbf{V} \cdot d\sigma, \]  
(B.1)

where \( \mathbf{V} \) is a vector field, \( d\sigma \) is a surface element on the surface \( S \) and \( \lambda \) a line segment around its boundary. Stokes’ theorem is useful in studying the behaviour of fluids like in magnetohydrodynamics.

B.2 Leibniz integral rule

Leibniz’s rule for differentiating an integral is well known, given by

\[ \frac{d}{dt} \left( \int_{g(t)}^{h(t)} f(x,t)dx \right) = \left. \int_{g(t)}^{h(t)} \frac{\partial f(x,t)}{\partial t} \right|_{x=t} + \{ f[h(t),t] \cdot \dot{h}(t) - f[g(t),t] \cdot \dot{g}(t) \}. \]  
(B.2)

The formula can be generalized to three dimensions. Given vector quantity \( \mathbf{F}(\mathbf{r}, t) \), a direction vector \( d\mathbf{A} \) and a velocity \( \mathbf{v} \), Leibniz’ rule in three dimensions becomes (Flanders, 1973)

\[ \frac{d}{dt} \int_{S(t)} \mathbf{F}(\mathbf{r}, t) \cdot d\mathbf{A} = \int_{S(t)} \left( \frac{\partial}{\partial t} \mathbf{F}(\mathbf{r}, t) + (\nabla \cdot \mathbf{F}) \mathbf{v} \right) \cdot d\mathbf{A} - \oint_{\partial S} (\mathbf{v} \times \mathbf{F}) \cdot d\mathbf{l}, \]  
(B.3)

where \( \mathbf{S} \) is the surface area vector, \( \partial \mathbf{S} \) its edge and \( d\mathbf{l} \) the differential vector along the contour \( \partial \mathbf{S} \).
Appendix C

Information on data sources

Figure C.1: Map showing the locations of the observatories providing data for Dst index. Source: Sugiura et al. (1991)
Northern hemisphere

<table>
<thead>
<tr>
<th>Years</th>
<th>Observatory</th>
<th>Corrected geom. lat.($^\circ$ N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1868-1925</td>
<td>Greenwich</td>
<td></td>
</tr>
<tr>
<td>1926-1956</td>
<td>Abinger</td>
<td></td>
</tr>
<tr>
<td>1957-</td>
<td>Hartland</td>
<td>50.0$^\circ$</td>
</tr>
</tbody>
</table>

Southern hemisphere

<table>
<thead>
<tr>
<th>Years</th>
<th>Observatory</th>
<th>Corrected geom. lat.($^\circ$ S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1868-1919</td>
<td>Melbourne</td>
<td></td>
</tr>
<tr>
<td>1920-1979</td>
<td>Toolangui</td>
<td></td>
</tr>
<tr>
<td>1980-</td>
<td>Canberra</td>
<td>45.2$^\circ$</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Observatory</th>
<th>Code</th>
<th>Geogr. lat.($^\circ$ N)</th>
<th>Geogr. long.($^\circ$ E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abisko</td>
<td>ABK</td>
<td>68.36</td>
<td>18.82</td>
</tr>
<tr>
<td>Dixon Island</td>
<td>DIK</td>
<td>73.55</td>
<td>80.57</td>
</tr>
<tr>
<td>Cape Chelyuskin</td>
<td>CCS</td>
<td>77.72</td>
<td>104.28</td>
</tr>
<tr>
<td>Tixie Bay</td>
<td>TIK</td>
<td>71.58</td>
<td>129.00</td>
</tr>
<tr>
<td>Cape Wellen</td>
<td>CWE</td>
<td>66.17</td>
<td>190.17</td>
</tr>
<tr>
<td>Barrow</td>
<td>BRW</td>
<td>71.30</td>
<td>203.25</td>
</tr>
<tr>
<td>College</td>
<td>CMO</td>
<td>64.87</td>
<td>212.17</td>
</tr>
<tr>
<td>Yellowknife</td>
<td>YKC</td>
<td>62.40</td>
<td>245.60</td>
</tr>
<tr>
<td>Fort Churchkill</td>
<td>FCC</td>
<td>58.80</td>
<td>265.90</td>
</tr>
<tr>
<td>Poste-de-la-Baleine</td>
<td>PBQ</td>
<td>55.27</td>
<td>282.22</td>
</tr>
<tr>
<td>Narsasquaq (Narssarssuaq)</td>
<td>NAQ</td>
<td>61.20</td>
<td>314.16</td>
</tr>
<tr>
<td>Leirvogur</td>
<td>LRV</td>
<td>64.18</td>
<td>338.30</td>
</tr>
</tbody>
</table>

Table C.2: List of $AE$ stations. Source: Kyoto $AE$ index service (2013)
<table>
<thead>
<tr>
<th>Observatory</th>
<th>Code</th>
<th>Geogr. lat. (°N)</th>
<th>Geogr. long. (°E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ny Alesund</td>
<td>NAL</td>
<td>78.92</td>
<td>11.95</td>
</tr>
<tr>
<td>Longyearbyen</td>
<td>LYN</td>
<td>78.20</td>
<td>15.82</td>
</tr>
<tr>
<td>Hornsund</td>
<td>HOR</td>
<td>77.00</td>
<td>15.60</td>
</tr>
<tr>
<td>Hopen Island</td>
<td>HOP</td>
<td>76.51</td>
<td>25.01</td>
</tr>
<tr>
<td>Bear Island</td>
<td>BJN</td>
<td>74.50</td>
<td>19.20</td>
</tr>
<tr>
<td>Nordkapp</td>
<td>NOR</td>
<td>71.09</td>
<td>25.79</td>
</tr>
<tr>
<td>Sørøya</td>
<td>SOR</td>
<td>70.54</td>
<td>22.22</td>
</tr>
<tr>
<td>Alta</td>
<td>ALT</td>
<td>69.86</td>
<td>22.96</td>
</tr>
<tr>
<td>Kevo</td>
<td>KEV</td>
<td>69.76</td>
<td>27.01</td>
</tr>
<tr>
<td>Tromsø</td>
<td>TRO</td>
<td>69.66</td>
<td>18.94</td>
</tr>
<tr>
<td>Masi</td>
<td>MAS</td>
<td>69.45</td>
<td>23.70</td>
</tr>
<tr>
<td>Andenes</td>
<td>AND</td>
<td>69.30</td>
<td>16.03</td>
</tr>
<tr>
<td>Kilpisjärvi</td>
<td>KIL</td>
<td>69.06</td>
<td>20.77</td>
</tr>
<tr>
<td>Kautokeino</td>
<td>KAU</td>
<td>69.02</td>
<td>23.05</td>
</tr>
<tr>
<td>Ivalo</td>
<td>IVA</td>
<td>68.56</td>
<td>27.29</td>
</tr>
<tr>
<td>Abisko</td>
<td>ABK</td>
<td>68.35</td>
<td>18.82</td>
</tr>
<tr>
<td>Leknes</td>
<td>LEK</td>
<td>68.13</td>
<td>13.54</td>
</tr>
<tr>
<td>Muonio</td>
<td>MUO</td>
<td>68.02</td>
<td>23.53</td>
</tr>
<tr>
<td>Lovozero</td>
<td>LOZ</td>
<td>67.97</td>
<td>35.08</td>
</tr>
<tr>
<td>Kiruna</td>
<td>KIR</td>
<td>67.84</td>
<td>20.42</td>
</tr>
<tr>
<td>Sodankylä</td>
<td>SOD</td>
<td>67.37</td>
<td>26.63</td>
</tr>
<tr>
<td>Pello</td>
<td>PEL</td>
<td>66.90</td>
<td>24.08</td>
</tr>
<tr>
<td>Jäckvik</td>
<td>JCK</td>
<td>66.40</td>
<td>16.98</td>
</tr>
<tr>
<td>Dønna</td>
<td>DON</td>
<td>66.11</td>
<td>12.50</td>
</tr>
<tr>
<td>Rørvik</td>
<td>RVK</td>
<td>64.94</td>
<td>10.98</td>
</tr>
<tr>
<td>Lycksele</td>
<td>LYC</td>
<td>64.61</td>
<td>18.75</td>
</tr>
<tr>
<td>Oulujärvi</td>
<td>OUJ</td>
<td>64.52</td>
<td>27.23</td>
</tr>
<tr>
<td>Mekrijärvi</td>
<td>MEK</td>
<td>62.77</td>
<td>30.97</td>
</tr>
<tr>
<td>Hankasalmi</td>
<td>HAN</td>
<td>62.25</td>
<td>26.60</td>
</tr>
<tr>
<td>Dombås</td>
<td>DOB</td>
<td>62.07</td>
<td>9.11</td>
</tr>
<tr>
<td>Solund</td>
<td>SOL</td>
<td>61.08</td>
<td>4.84</td>
</tr>
<tr>
<td>Nurmijärvi</td>
<td>NUR</td>
<td>60.50</td>
<td>24.65</td>
</tr>
<tr>
<td>Uppsala</td>
<td>UPS</td>
<td>59.90</td>
<td>17.35</td>
</tr>
<tr>
<td>Karmøy</td>
<td>KAR</td>
<td>59.21</td>
<td>5.24</td>
</tr>
<tr>
<td>Tartu</td>
<td>TAR</td>
<td>58.26</td>
<td>26.46</td>
</tr>
</tbody>
</table>

Table C.3: List of IMAGE stations. Source: International Monitor for Auroral Geomagnetic Effects (2013, → Stations → Coordinates)
Figure C.2: Map showing the locations of the observatories providing data for AE index. Source: Kyoto AE index service (2013)

Figure C.3: Map showing the locations of the IMAGE observatories. Source: International Monitor for Auroral Geomagnetic Effects (2013, → Stations → Maps)
Appendix D

Matlab codes

In this Appendix I describe and introduce the MATLAB codes used in the data analysis in this work.

D.1 Fractionized year

Running mean function below requires all the data to be represented in the same time unit. This is achieved by converting the date into a fractionized year. This has been implemented in a MATLAB function y_frac.m. The full code is shown in Figures D.1 and D.2.
D.2 Running mean

To calculate the running mean I developed my own code. The MATLAB function, calc_time_mean.m, has been designed to be as versatile as possible. It can be used to calculate not only mean, but any other statistical property along the time series. If desired, it can be used to calculate means along other quantities than time as well though this property has not been used in this thesis. By using the last parameter it’s possible and indeed practical to calculate event density of a series of any kinds of events, a feature used extensively in the data analysis in Chapter 5.

The full code is shown in Figures D.3 and D.4.

D.3 Interpolation of indefinite values

Many MATLAB features do not work correctly with indefinite values, called NaN (not-a-number) in MATLAB terms. Yet it is not always practical to ignore data like this. To battle this problem a simple interpolation software was developed, usable with MATLAB and provided here as a reference.

The program interpolate_NaNs.m uses the procedure described in Section 3.1.7 for interpolation. The special case of NaNs in the boundary of the data, i.e. in the start or the end of a data vector, have to be treated separately. I have adopted the simplest possible solution and replaced these values with the value in the outermost non-NaN data point.

The full code is shown in Figures D.5 and D.6.
function fraction = y_frac(day, month, year, hour, minute, second)
% y_frac gives out the year and its fraction that has passed until the date.
% Give it the date in the format day, month, hour, minute, second, all
% of which are numbers. Day is the number of the day in the passing
% month, month is the number of the month, hour is the clock hour,
% minute is the clock minute and second is the clock second.

if (nargin==1)
% If user gives only the dates in a format dates(n rows, 6 columns),
% where 6 columns are day, month, year, hour, minute, second in that order
% then this loop is used. If you have seven columns, you can assign
% this result to the seventh by writing simply:
% day(:,7) = y_frac(day);
% where day contains the dates as explained above.
    number_of_dates = numel(day(:,1));
    fraction=zeros(number_of_dates,1);
    for k=1:number_of_dates
        if (rem(day(k,1), 400) == 0)
            skipyear = 1;
        elseif (rem(day(k,1), 4) == 0) && (rem(day(k,3),100) == 0))
            skipyear = 1;
        else
            skipyear = 0;
        end;

        if day(k,3) == 1 % January
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400;
        elseif day(k,3) == 2 % February
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 31;
        elseif day(k,3) == 3 % March
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 31 + skipyear;
        elseif day(k,3) == 4 % April
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 90 + skipyear;
        elseif day(k,3) == 5 % May
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 120 + skipyear;
        elseif day(k,3) == 6 % June
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 151 + skipyear;
        elseif day(k,3) == 7 % July
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 181 + skipyear;
        elseif day(k,3) == 8 % August
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 212 + skipyear;
        elseif day(k,3) == 9 % September
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 243 + skipyear;
        elseif day(k,3) == 10 % October
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 273 + skipyear;
        elseif day(k,3) == 11 % November
            sum = (day(k,1)-1) + (day(k,1)/24) + day(k,1)/1440 + day(k,1)/86400 + 304 + skipyear;
        else % December

    end
end

Figure D.1: MATLAB program y_frac.m, page 1
% End of the defining else.

end

else

% If user gives one date in the standard format of this function, % then this iteration is used.
if (rem(year, 400) == 0)
skipyear = 1;
elseif (rem(year, 4) == 0) && (rem(year, 100) == 0)
skipyear = 1;
else
skipyear = 0;
end;

if month == 0 % Not defined, only minutes given.
  if day==0
day=1;
end

sum = (day-1) + (hour/24) + minute/1440 + second/86400;
elseif month == 1 % January
sum = (day-1) + (hour/24) + minute/1440 + second/86400;
elseif month == 2 % February
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 31;
elseif month == 3 % March
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 59 + skipyear;
elseif month == 4 % April
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 90 + skipyear;
elseif month == 5 % May
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 120 + skipyear;
elseif month == 6 % June
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 151 + skipyear;
elseif month == 7 % July
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 181 + skipyear;
elseif month == 8 % August
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 212 + skipyear;
elseif month == 9 % September
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 243 + skipyear;
elseif month == 10 % October
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 273 + skipyear;
elseif month == 11 % November
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 304 + skipyear;
else % December
sum = (day-1) + (hour/24) + minute/1440 + second/86400 + 334 + skipyear;
end

fraction = sum/(365*skipyear)*year;
end; % End of the defining else.

end

Figure D.2: MATLAB program y_frac.m, page 2

101
function time_mean = calc_time_mean(timevalue, quantity, dT, timestep, function)

% Function calculates the time average of a certain quantity.
% Consequently the program has been constructed in such a way that one
% can also use it to calculate an average by some other values than
% time, like for example the magnetic field in certain intervals.

% Argument details:
% Quantity and timevalue are vectors. Quantity contains the value of
% certain physical quantity and timevalue is the moment of occurrence in
% chosen time units (anything as long as it's the same for all of them).
% Then dT is the width of the time window over which the average is taken.
% Finally timestep is the time between individual points of measurement.
% If you want to, you can use this function to give you something else than
% mean. This is done by providing some function as the fifth argument. If
% you do not provide the fifth one, the program will assume you mean mean.
% Do note that the function is passed by adding @ in front of it. This feature
% has to be used with care though as the function has not strictly been
% intended for such use.

% See the example at the bottom of this description.

% Output:
% Function saves the result as a 2-dimensional matrix. The first column is
% the value of time and the second column is the actual mean at this time.
% The first point is at firstpoint = min(timevalue)+dT/2.
% Then the next point is at nextpoint = firstpoint + timestep and so on,
% until nextpoint > max(timevalue). The last point is not taken into account.

% Example use:
% mean_of_data = calc_time_mean(array_of_time_values, array_of_measured_quantities, 10, 10);
% mode_of_data = calc_time_mean(array_of_time_values, array_of_measured_quantities, 10, 10, @mode);
% event_number_density = calc_time_mean(array_of_event_times, array_of_measured_quantities, 10, 1, @numel);
% Last one will calculate the number of events happening in a certain time period 10 and
% give the result.
% running every period of 1. Here only the array_of_event_times is important;
% array_of_measured quantities does not enter the equation, though it does have to be of equal length to
% array_of_event_times.

% Let's calculate the number of steps. This is got by taking the width of the data (max-min)
% and removing
% the values dT/2 from both ends (thus -dT)
number_of_steps = floor((max(timevalue)-min(timevalue)-dT)/timestep)+1;
time_mean = zeros(number_of_steps, 2);
firststep = min(timevalue)+dT/2;

% Take away NaNs from quantity data.
corrected_index_set = find(~isnan(quantity));
corrected_timevalue = timevalue(corrected_index_set);
corrected_quantity = quantity(corrected_index_set);

% Take away NaNs from time data. Mostly this is relevant only, if timevalue is something
% else than time.
corrected_index_set = find(~isnan(corrected_timevalue));
corrected_timevalue = corrected_timevalue(corrected_index_set);
corrected_quantity = corrected_quantity(corrected_index_set);

if nargin<5 % If the fifth argument is not passed, the program assumes it's a mean.
    function=@mean;
end;

for k=1:number_of_steps
    time_mean(k,1)=firststep+(k-1)*timestep;
    if time_mean(k,1)>max(timevalue)-dT/2
        break;
    end;
    % Calculate the time mean. Note that the lower limit is included into the calculation
    % and not the upper limit, thus every point is included only once.
    if isempty(corrected_quantity(corrected_timevalue>=firststep-dT/2*(k-1)*timestep &
                                   corrected_timevalue<firststep+dT/2*(k-1)*timestep))
        time_mean(k,2)=0;
    else
        time_mean(k,2)=function(corrected_quantity(corrected_timevalue>=firststep-dT/2*(k-1)*timestep &
                                                  corrected_timevalue<firststep+dT/2*(k-1)*timestep));
    end;
end;

% In case of @numel the average is calculated by dividing the result by dT,
% i.e. the time over which the average was calculated.
% The result's unit will be the # / time unit of the data. E.g. If the data is given in
% years, the unit will be years (e.g. CMEs per year).
if isequal(function,@numel)
    time_mean(:,2)=time_mean(:,2)/dT;
end;

% If there are points with no data (i.e. no data points inside a time step), they will be
% assigned NaN by default.
% This is not desired and thus these points will be assigned 0.
for k=1:number_of_steps
    if isnan(time_mean(k,2))
        time_mean(k,2)=0;
    end;
end;
end

Figure D.4: MATLAB program calc_time_mean.m, page 2
function output = interpolate_NaNs(data, to_concatenate_or_not);
% Data is supposedly 1-dimensional vector of scalar values.
% All NaN values in the data are replaced by linear interpolation.
% Argument details:
% Data is a column vector of evenly spaced terms of data. If the data
% were not evenly spaced, the function will give a false, though
% possibly an approximate interpolation.
% Parameter to_concatenate_or_not is either 'No' (default) or
% 'Yes'. If 'No', the function will "interpolate" also the start and end
% the end of the data if necessary, replacing the start and end
% NaN values with the closest non-NaN value. If 'Yes', it simply loses
% the missing data, concatenating the data if there's NaN values
% in the beginning or the end of the data vector.

% Check if the user gave an argument for to_concatenate_or_not.
% Defaults to 'No'.
if nargin < 2
    to_concatenate_or_not = 'No';
end;

% Prepare some variables.
how_many_NaNs_at_start = 0;
how_many_NaNs_at_end = 0;

% Check that the data vector is not empty.
if isempty(data)
    error('The data vector is empty.');
end;

% Check if the data vector contains actual data, not merely NaNs.
if numel(find(isnan(data(:,1)))) == numel(data(:,1))
    error('The data vector contains only NaN values: no interpolation can be made.');
end;

% Find the non-NaN beginning of the data.
while isnan(data(how_many_NaNs_at_start+1,:))
    how_many_NaNs_at_start = how_many_NaNs_at_start + 1;
end;

% Find the non-NaN end of the data.
while isnan(data(numel(data,:)-how_many_NaNs_at_end,:))
    how_many_NaNs_at_end = how_many_NaNs_at_end + 1;
end;

% Remove the start and the end NaNs if desired.
if isequal(to_concatenate_or_not, 'Yes')
    output = data(how_many_NaNs_at_start+end+how_many_NaNs_at_end,:);
else
    if how_many_NaNs_at_start > 0
        output(:,how_many_NaNs_at_start,:) = output(how_many_NaNs_at_start+1,:);
    end;
    if how_many_NaNs_at_end > 0
        output(end-how_many_NaNs_at_end+1:end,:) = output(end-how_many_NaNs_at_end,:);
    end;
end;

Figure D.5: MATLAB program interpolate_NaNs.m, page 1
% Interpolate the rest of the data.
for k=1:numel(output(:,1))
    % Check if the data point is NaN.
    if isnan(output(k,1))
        param=k;
        % Calculate how long is this NaN streak.
        while isnan(output(param,1))
            param=param+1;
        end
        % Calculate the slope coefficient.
        slopecoef=(output(param,1)-output(k-1))/(param-k+1);
        % The beginning point of the interpolation.
        origin=output(k-1,1);
        % Replace the NaNs by linear interpolation.
        for j=k:param-1
            output(j,1)=origin+slopecoef*(j-k+1);
        end
    end
end

Figure D.6: MATLAB program interpolate_NaNs.m, page 2
Bibliography


