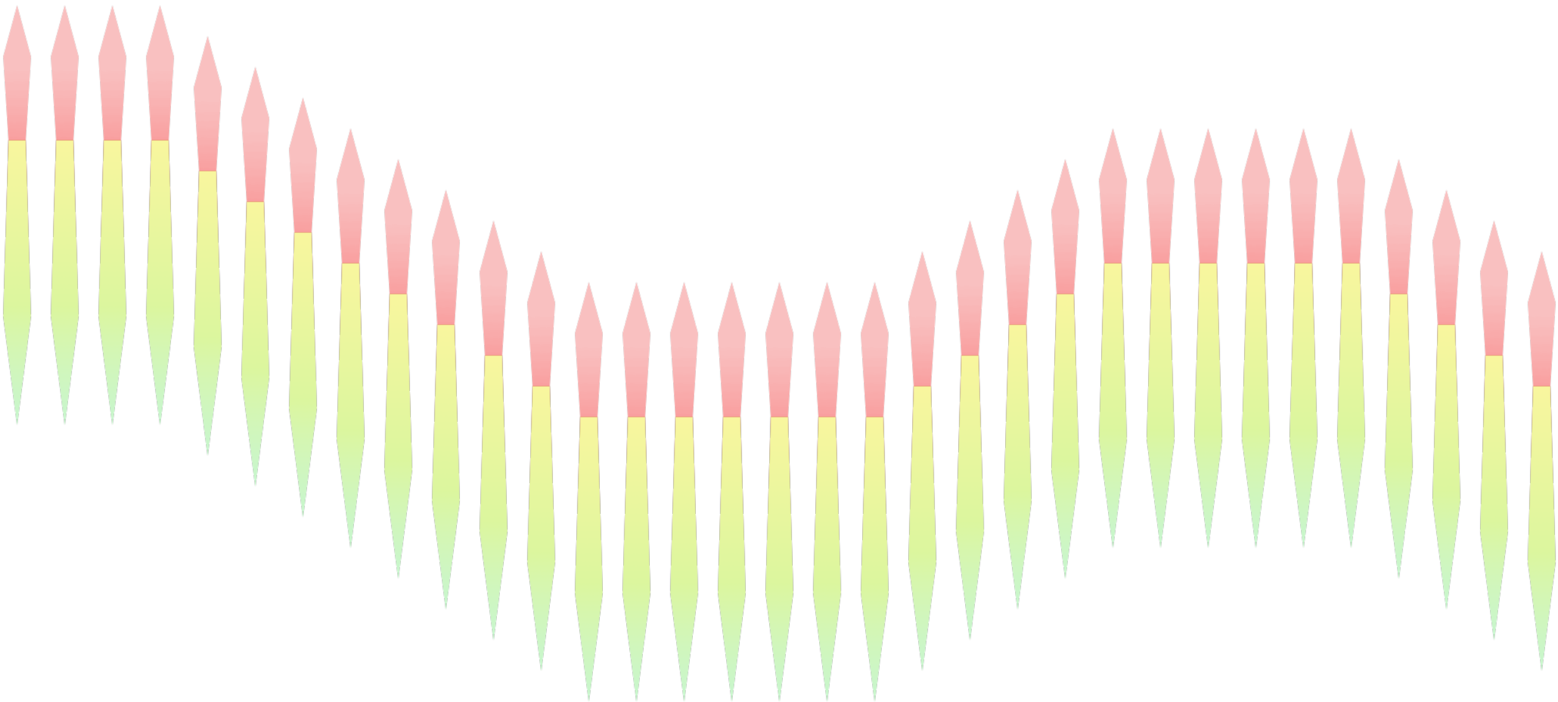
Long-term analysis of geomagnetic disturbances

June 5, 2015

Bachelor thesis
The Hague University of Applied Sciences
Swedish Institute of Space Physics



Author:
Bart Olsthoorn

Supervisor:
Dr. M. Yamauchi

Internship supervisors:
D.D. Land
R.A. Mantel

Summary

Earth's magnetic field, also known as the geomagnetic field, interacts with the solar wind, a stream of charged particles originating from the Sun. The interaction causes the geomagnetic field to undergo variations, called geomagnetic disturbances or geomagnetic activity. This bachelor thesis features a long-term trend in the geomagnetic activity, using data from individual geomagnetic stations since 1845.

The first objective is to reduce the raw recorded data of a selection of stations to the K_p-index, which provides information of geomagnetic disturbances on a planetary scale. The official K_p-index is only available from 1932, due to a lack of older data from the original K_p-index stations. The by now 83 years of K_p-index data is insufficient when examining centurial variation, which is why this thesis uses data from individual stations operating as far back as 1845 to extend the existing data.

The second objective is to use this new extended data set to look for long-term trends in geomagnetic disturbances. Geomagnetic disturbances are related to solar activity, which undergoes both periodic and irregular changes. The periodic changes are called solar cycles or sunspot cycles, with an average cycle duration of 11 years. The solar (sunspot) cycle is defined by the average number of sunspots on the solar surface, where sunspots have been observed for hundreds of years. The geomagnetic activity per solar cycle is compared to the number of sunspots to examine whether the geomagnetic activity is affected by the sunspot activities.

The results show that, thus far, the current solar cycle #24, being one of low solar activity, also features the lowest geomagnetic activity recorded since at least 1915. Furthermore, the declining phase of the examined solar cycles includes more geomagnetic activity for the same number of sunspots than the other phases of the cycle. When comparing individual solar cycles, some cycles also show different geomagnetic activity for the same number of sunspots, suggesting that some cycles are more geo-effective than others. The number of sunspots at the maximum of the solar cycle appears to be an indicator of this geo-effectiveness.

One of the consequences of geomagnetic activity is aurora, for which all-sky cameras are now commonly used all over the world. Since the relation between the aurora and geomagnetic activity is not clear, auroral activity must also be quantified. Such a quantification requires automatic detection of aurora from all-sky camera images. This thesis introduces an algorithm for detection of auroras using all-sky camera images.

Contents

1 Introduction

The Earth has a magnetic field, sometimes called the geomagnetic field, that undergoes both regular and irregular variations. Some irregular variations are accompanied by auroras (the northern and southern lights), and extremely large variations may cause strong induced current at the ground and long-distance power lines, causing electric power failures. At the Earth's surface, the geomagnetic variations have long been measured by magnetometers all over the world. These instruments measure the strength and direction of the magnetic field. Some magnetometers have been in operation since 1845 which provides a unique insight into the long-term trend of geomagnetic activity.

The disturbances in the geomagnetic field are a consequence of the Earth's interaction with the Sun. The Sun is continuously emitting charged particles, the solar wind, which travels toward the Earth and interact with the geomagnetic field. The solar wind is not constant, but follows a roughly 11-year cycle. The period ranges from 9 to 14 years, and not every cycle is the same [1]. In fact, the current solar cycle #24 features less activity than many previous cycles. How the Earth is influenced by the low solar activity is a major science issue because it might even affect the global climate. In this sense, it is ideal if one can examine the past similar cycles with low solar activity, that happened about 100 years ago. This bachelor thesis includes a long-term analysis of the geomagnetic disturbances during many solar cycles by using digitally available magnetometer data recorded around the Earth.

The first objective is to normalize the available data so that geomagnetic disturbances at different stations can be compared. After the data is normalized, a world-wide index is constructed, similar to the so-called K_p-index. The K_p-index provides information of geomagnetic disturbances on a planetary scale. The official K_p-index is only calculated from 1932, but the estimated K_p-index for the pre-1932 age would allow for analysis of continuous data for over 100 years. This thesis uses data from individual stations operating as far back as 1845 to extend the existing data.

The second objective is to use this new extended data set (first objective result) to look for long-term trends in geomagnetic disturbances, and comparing this for individual solar cycles. This gives an approximate relation between the solar activity and geomagnetic activity, i.e., how this coupling differs per cycle.

2 Space weather

The concept and terminology of *space weather* was introduced about 30 years ago, and became popular even to the general public during the 1990s [2]. Initially, the field was developed from auroral, ionospheric, geomagnetic, and magnetospheric researches because these activities are well known to be controlled by the solar wind. In the 1990s the knowledge and solar wind / solar surface monitors became mature enough to allow for predictions of large geomagnetic activity. Such prediction became more and more important in modern society because the largest geomagnetic activity (level of few times a decade) may cause satellite malfunction and electric power failure in big cities. Therefore, real-time predictions are required in addition to traditional studies of finding out the relation between the solar (wind) input and geomagnetic activities as output. This entire effect is now referred to as the space weather. The next sections will introduce several important elements in space weather.

2.1 Earth's magnetic field

The Earth has a magnetic field, or geomagnetic field, which can be approximated by a dipolar field of a bar magnet. In reality, the geomagnetic field is generated by the motion of iron alloys in the outer core of the Earth. The geomagnetic field reaches far beyond the ionosphere, forming a large region surrounding the Earth, called the magnetosphere. Figure 1 shows a schematic view of the structure of this system.

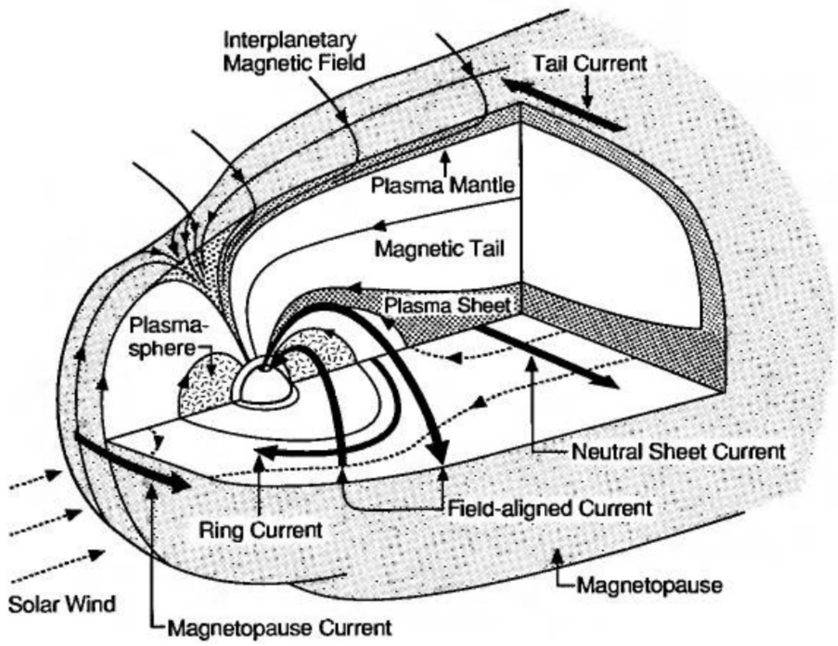


Figure 1: The structure of the magnetosphere, showing magnetic fields, plasma regions, and large-scale currents [3].

The solar wind, a stream of charged particles, is constantly pushing on the Earth's magnetic field. As the plasma tries to move across the geomagnetic field, it will experience the Lorentz force, causing the magnetosphere to distort and compress on the Sun-facing side. On the opposite side, it stretches the field out to a long tail, the magnetotail. The magnetosphere extends out to about 10 Earth radii (R_E) in the Sun-facing, and hundreds of Earth radii in the opposite direction [4]. Only the inner part of the magnetosphere (5 R_E) remains approximately dipolar. The following list addresses the different components displayed in Figure 1:

Solar wind
: A stream of charged particles emitted by the Sun. The solar wind is discussed in more detail in Section 2.2.2.

Magnetopause
: The magnetopause is the boundary of the magnetosphere to the solar wind, consisting of thin sheets (500 to 1000 km) of electrical currents.

Interplanetary magnetic field
: The interplanetary magnetic field, sometimes called the IMF, is the solar magnetic field carried by the solar wind (see Section 2.2.2).

Ring current
: Ring currents are formed by charged particles that are trapped in the magnetosphere. The Earth's ring current flows westward along the equator, at a distance of 3 to 5 R_E. These particles produce their own magnetic field, opposite of the Earth's field, and therefore slightly reduce the strength of the geomagnetic field. A strong ring current results in a weaker geomagnetic field around the equator.

Ionosphere
: The Sun's ultraviolet light ionizes particles in the Earth's upper atmosphere (mainly atomic O and molecular N_2) and frees electrons. The resulting plasma is called the ionosphere.

Plasmasphere
: The plasmasphere is located just outside the ionosphere. The electrons move outwards, along Earth's magnetic field lines, and the ionosphere becomes positively charged as more and more electrons leave. This leads to an electric field, causing the ions to move outward too, attracted by the negatively charged space around the ionosphere. On a large scale, this simply means that electrons and ions leave the ionosphere. Earth's magnetic field traps these charged particles, preventing them to reach interplanetary space [5].

Figure 1 shows the average state of the magnetosphere, but in reality, it is constantly changing due to the ever-changing solar wind and magnetospheric dynamics.

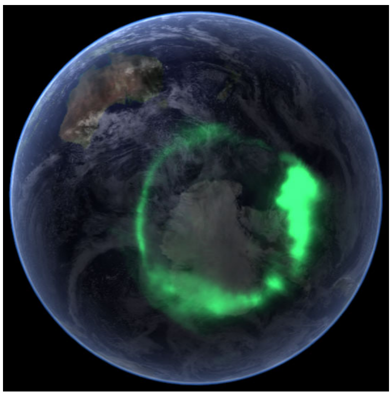


Figure 2: Aurora australis (southern lights) revealing the shape of the auroral oval. Captured by NASA's IMAGE satellite and overlaid onto Blue Marble image [6].

The auroral oval (Figure 2) marks the boundary between the polar cap and the low-latitude regions [4]. The polar cap is magnetically connected to the solar wind and magnetotail, whereas low-latitude regions are similar to dipole field lines and close on Earth. Around the polar caps, electrical currents (charged particles) flow into the ionosphere, at a height of 90 to 130 km, disturbing the geomagnetic field. The field-aligned flow of charged particles ionizes atmospheric particles leading to bright auroras. Concentrations of currents are also flowing through the ionosphere horizontally, commonly carrying more than 10^6 A. These currents are called auroral electrojets.

At lower latitudes, closer to the equator, it is the variation of the ring current that determines most of the geomagnetic activity. This is in contrast to the high-latitudes, where the solar wind and the magnetotail dictate the geomagnetic activity.

The two most characteristic features of geomagnetic activity are the geomagnetic storm and the geomagnetic substorm. The signature of a geomagnetic storm is an increased ring current, weakening the geomagnetic field at low- and mid-latitudes. A strong southward IMF and high solar-wind velocities lead to to a geomagnetic storm. The recovery takes several days, which is much longer than substorms, that only take two or three hours to recover. A geomagnetic substorm is caused by a different set of events, starting with a build-up of energy in the magnetotail (5-20 R_E). This is followed by a sudden release of energy toward the Earth. The sudden energy release causes large electric fields and aurora. What triggers such an energy release is on debate for more than 30 years.

During a big geomagnetic storm or substorm, the magnetosphere experiences large variations and holds strong electric currents. This causes strong ionospheric currents, which then in turn can induce currents in conductors on Earth, leading to damage to electronic systems. Since the ionospheric currents can be monitored by geomagnetic disturbances, the geomagnetic field data is one of the most essential elements for space weather.

2.1.1 Geomagnetic field measurements

The Earth's magnetic field has been systematically measured at many points around the globe by ground-based observatories since 1844 [7]. This report uses data provided by the *World Data Center for Geomagnetism, Kyoto*, containing data of 295 observatories spread over the Earth (see Figure 13a).

The magnetic field at a certain point is described as a vector in one of two ways, the XYZ representation or the DHZ representation. These components are expressed in the units of nanotesla, with the exception of the D component, which is the angle between the geographic and geomagnetic north.

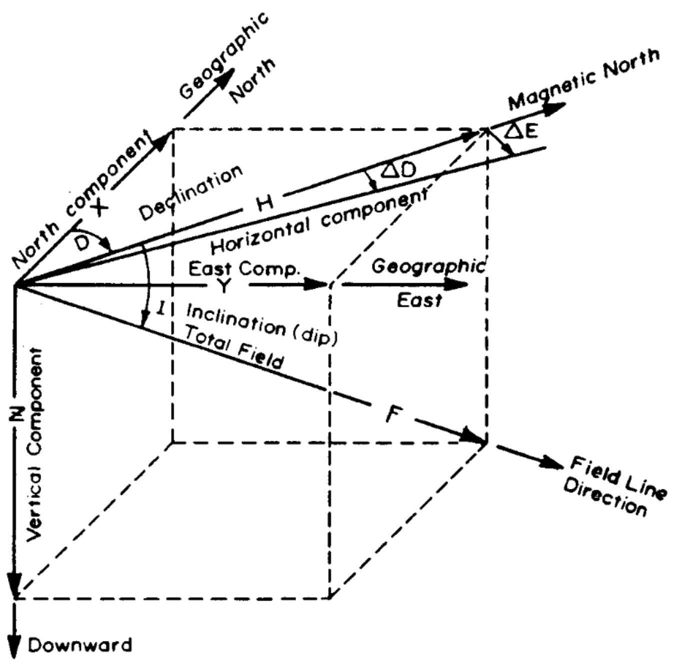


Figure 3: Components of the magnetic field measured at a certain point, showing the relationship between the XYZ and DHZ representation [8].

Old magnetogram data is often in the DHZ representation, but it is simple to convert it to XYZ using the following equations

$$X = H\cos(D) \quad Y = H\sin(D) \tag{1}$$

where D is the declination in degrees and H the horizontal component in nT.

The geomagnetic field has been expressed in different units over the years. Table 1 lists the equivalent values of the magnetic field strength (B). It is important to note that one gamma (γ) is equivalent to one nanotesla (nT).

Table 1: Equivalent magnetic field units [8]

$B = 10^4$ Gauss
$B = 1$ Weber/meter^2
$B = 10^9$ gamma
$B = 1$ Tesla

As Figure 3 already showed, the positions of Earth's geographic pole and geomagnetic pole do not match, resulting in two coordinate systems. This report uses data from ground-based stations that measure the geomagnetic field, which is why it is sometimes necessary to convert a geographic location (latitude and longitude) to a geomagnetic latitude/longitude.

2.1.2 Geomagnetic latitudes

The geomagnetic north pole is located approximately 80.4N 72.6W in 2015 [9]. The following rotation matrices can be used to move points in geomagnetic lat/long to geographic lat/long and vice versa.

$$
\begin{aligned}
R_x(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \\
R_z(\theta) &= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{2}
$$

These rotation matrices are used in combination with the geomagnetic pole location to calculate the geomagnetic latitudes displayed in Figure 4. In this case R_z(longitude) and $R_x(90° -$ latitude), with latitude $= 80.4°$ and longitude $= -72.6°$.

Converting geographic to geomagnetic locations also becomes increasingly important on long timescales, as the geomagnetic poles slowly wander away from their current position. As mentioned earlier, geomagnetic activity is largely dependent on geomagnetic latitude. It is therefore important to consider that the geographically fixed observatories are moving relative to the geomagnetic coordinate system.

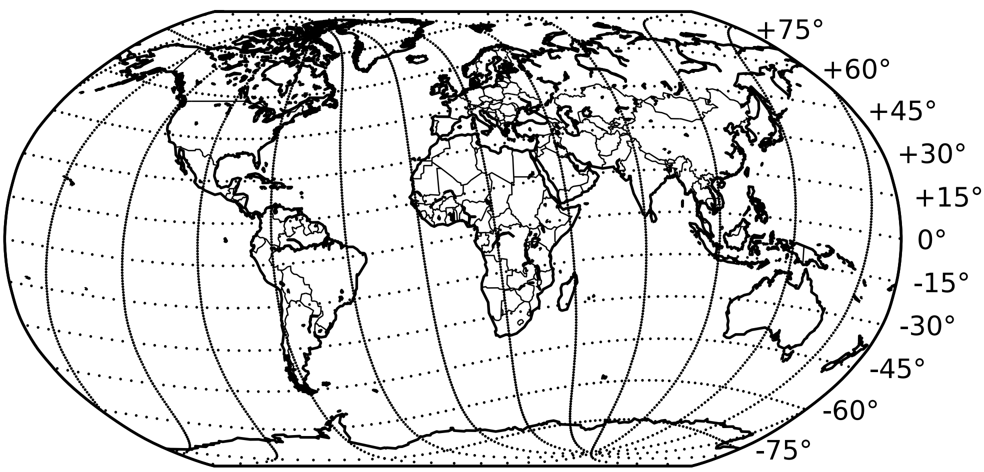


Figure 4: Geographic map with geomagnetic latitudes and longitudes (for 2015). The latitudes (in degrees) are indicated on the right. Geographic latitudes would appear as straight horizontal lines but are not displayed.

2.2 The Sun

The Sun is a main sequence star, fusing hydrogen into helium. The apparent surface of the Sun, called the photosphere, appears unchanging to the unaided eye, but is in fact undergoing constant changes. The Sun's surrounding atmosphere, the corona, travels outwards, creating the solar wind.

2.2.1 Solar activity

The solar cycle is the periodic change of the solar activity as defined by sunspot numbers (sunspot cycle). In modern ages when other parameter could be measured, it was found that other solar activities such as irradiance, solar wind, solar magnetic field, and radio activities also obey the same periodic cycle. The average duration of a single cycle is 11 years and the cycles have been observed for hundreds of years. Figure 5 shows the solar cycles indicated by the sunspot number.

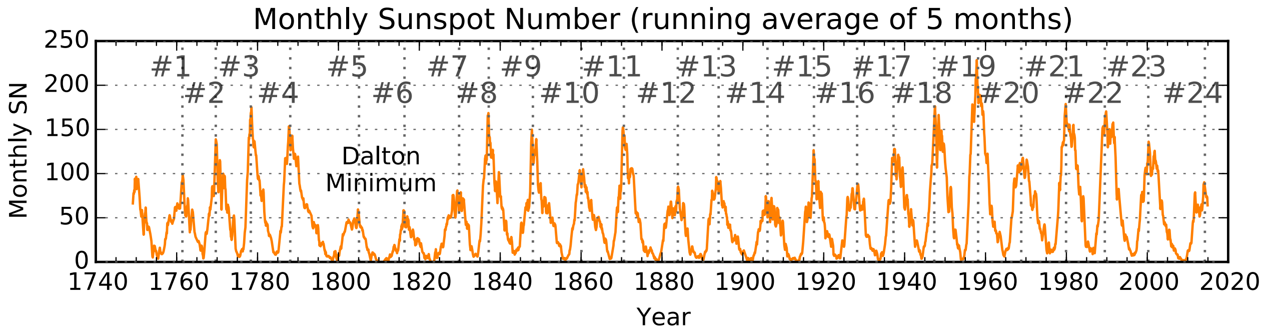
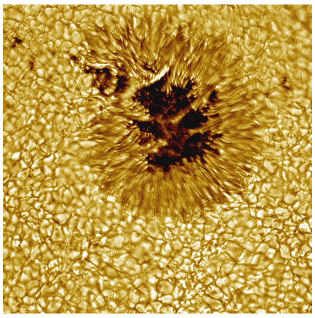


Figure 5: Monthly mean total sunspot number with a running average of 5 months. Data coverage starts at 1749. Sunspot data from the World Data Center SILSO, Royal Observatory of Belgium, Brussels [1].

From 1790 to 1830 there was a period of low solar activity, the Dalton Minimum, named after meteorologist John Dalton. Another period of inactivity occurred around the late 17th century, called the Maunder Minimum [10]. During this period there was no sunspot maximum for nearly 70 years, before the solar cycle restarted again. The Maunder Minimum is not shown in Figure 5 because it occurred before solar cycle #1, when extensive recording of solar sunspot activity started.

In the context of space weather, many aspects of solar activity are related to the magnetic field of the Sun. Sunspots are due to concentrations of the magnetic field, which causes a region to be cooler than the mean. Sometimes, magnetic energy is converted into kinetic energy, causing huge coronal mass ejections (CME) and flares (see Figure 6b). Also, the Sun's magnetic field polarity reverses completely on every solar maximum.



(a) Close-up of a sunspot [11].
Credit: Vacuum Tower Telescope, NSO, NOAO

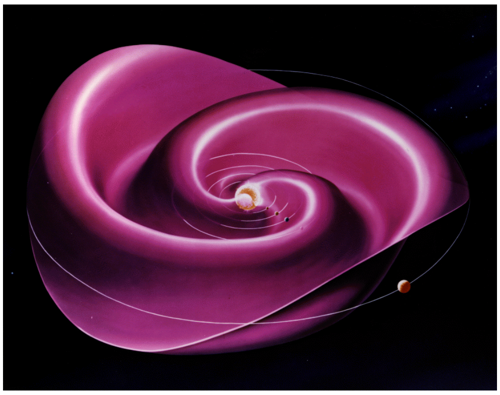
(b) Coronal mass ejection (CME) on August 31, 2012.
Credit: NASA's GSFC, SDO AIA Team

Figure 6: Solar activity related to the magnetic field of the Sun.

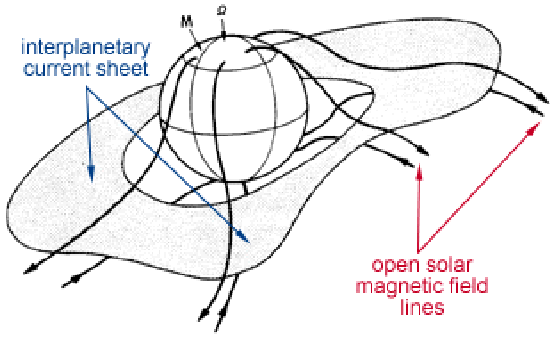
2.2.2 Solar wind

The Earth's magnetic field is disturbed by interactions with the solar wind, a stream of charged particles emitted by the Sun. These large-scale electrical currents affect the geomagnetic field, resulting in geomagnetic activity (disturbances).

The solar wind consists primarily of protons, electrons and alpha particles [3]. The observed densities for these particles (p, e and He^{2+}) are $6.6\,\mathrm{cm}^{-3}$, $7.1\,\mathrm{cm}^{-3}$ and $0.25\,\mathrm{cm}^{-3}$, respectively [3]. The Sun rotates at different speeds depending on latitude, with the fastest rotation being around the equator. The varying rate of rotation results in the solar wind travelling outward in a wavy spiral pattern (see Figure 7a). The solar wind carries the magnetic field toward the Earth, as the field lines are "frozen" in the solar wind plasma. This magnetic field is called the interplanetary magnetic field. The field direction for the northern hemisphere is opposite of the southern hemisphere (see Figure 7b).



(a) Heliospheric current sheet [12].



(b) The magnetic field travels outward along with the solar wind plasma [13].

Figure 7: The heliospheric current sheet.

The Earth is sometimes above and sometimes below the current sheet, causing periodal changes in polarity of the IMF, alternating between positive (toward the Sun) and negative (away from the Sun).

The IMF is described as a vector with three directional components: B_x, B_y and B_z. The B_z component is perpendicular to the Sun-Earth ecliptic plane, and is created by the disturbances and shocks in the solar wind. The B_x and B_y components are parallel to the ecliptic. The IMF is a weak field with strength ranging from 1 to 37 nT, and an average of around 6 nT (near the Earth) [14].

When the polarity of the IMF and magnetospheric field are opposite, they cancel each other, allowing solar wind plasma flowing through the magnetosphere due to the lack of protecting magnetic field (sometimes called reconnect). This process allows charged particles (solar wind) flow across the geomagnetic field, generating the electric field across the polar region and hence strong ionospheric electric current than can be detected as geomagnetic disturbance. Since a fast solar wind creates larger amount of electric field and ionospheric current, large geomagnetic activity is normally coupled with fast solar wind under specific direction of IMF. Such large activities are the main subject of this thesis.

2.3 Geomagnetic indices

To characterise geomagnetic activity, various geomagnetic indices have been introduced over the years. This report focuses primarily on the K-index, introduced in 1938 by J. Bartels [15], but several other indices will also be discussed in this section.

2.3.1 K-index

The K-index is an integer value assigned to each 3-hour interval for an individual station. For each 3-hour interval, the maximum recorded fluctuation R in any component of the magnetic field is expressed as an index, ranging from 0 (minimum) to 9 (maximum). The name, K-index, comes from the German word Kennziffer, meaning *characteristic digit*. Table 2 shows the conversion of the fluctuation R to K-index for the station in Niemegk, Germany. The right column indicates the lower bound of the ranges the K indices belong to.

Table 2: R to K-index conversion table for Niemegk, Germany [15].

K-index	Lower limit of ranges R (in nT)
0	0
1	5
2	10
3	20
4	40
5	70
6	120
7	200
8	330
9	500

The scale proceeds logarithmically up to $R = 40\,\mathrm{nT}$, but then continues to follow a less steep scale for the higher K indices. This choice was made because otherwise the high indices 8 and 9 would never have been reached.

It is not possible to use exactly the same conversion table (Table 2) for all observatories. The fluctuations at high-latitude observatories are always much larger than stations near the equator [15]. This is why the ranges are scaled for each observatory so that, ideally, the long-term K-index distribution becomes the same for all observatories. In other words, the number of 3-hour intervals for each K-index should be the same for all observatories, given a long period of time. In practice, this means that all the values in the right column of Table 2 are multiplied by a single factor, often a value between 0.5 and 3.0.

The publication introducing the K-index also describes a procedure for determining the ranges without the use of computers available today [15]. By counting the intervals with K $\geq$ 5 for each observatory, for a whole year, a scaling factor can be determined. Also, the lower range limit for K $=$ 1 should not be less than $3\,\mathrm{nT}$. At the time (1939), magnetograms were recorded on millimeter paper, sometimes at a scale of $8\,\mathrm{nT/mm}$. At this scale, it would be impractical to distinguish K $=$ 0 or 1.

More recently, computers have gradually taken over the recording task and many analog magnetograms have been digitised, which makes them readily available for computer processing. This allows for computations that were previously too impractical to perform, and it is possible to reshape the procedure previously used for calculating the conversion tables and K indices. The details of the procedure deriving the K indices using digital data will be described in Section 3.

2.3.2 K_p-index

The K_p index is a planetary index, describing the global state of the geomagnetic field. It is obtained as the mean value of the K indices collected by 13 selected subauroral stations. To allow for more precision, the K_p-index is sometimes given in a scale of thirds (adding + or − to the index, see Table 4).

To understand the significance of high K_p indices, the NOAA G-scale was introduced [16]. Table 3 shows the G-scale corresponding to the K_p. G5 events can cause widespread problems for electrical systems.

Table 3: NOAA G-scale [16].

G5	Extreme	K_p=9
G4	Severe	K_p=8
G3	Strong	K_p=7
G2	Moderate	K_p=6
G1	Minor	K_p=5

The NOAA G-scale also includes a list of possible effects caused by the events on power systems, spacecraft operations and other systems. For G5, this includes widespread voltage control problems and damage to transformers. Power grids may experience complete collapse or blackouts [16].

2.3.3 A-index

The daily average level of geomagnetic activity is expressed by the A-index. It is not always meaningful to take the average of a set of K indices, because the relationship between K and R is non-linear. To still be able to calculate a daily average, the K indices are converted to a linear scale, the a-index (Table 4). The A-index is the average of the set of a indices.

Table 4: Equivalent range a for a given K-index [17].

K	0	0+	1−	1	1+	2−	2	2+	3−	3	3+	4−	4	4+
a	0	2	3	4	5	6	7	9	12	15	18	22	27	32
K	5−	5	5+	6−	6	6+	7−	7	7+	8−	8	8+	9−	9
a	39	48	56	67	80	94	111	132	154	179	207	236	300	400

2.3.4 AE indices (high-latitude)

The term AE indices, where AE stands for Auroral Electrojet, is used to represent four related indices: AU, AL, AE, and AO [18]. These indices are a global measure of auroral zone magnetic activity and are given in the units of nT.

AU

The fluctuations in the horizontal component of the geomagnetic field are compared at each given time (UT) for a selection of observatories. The largest values are selected, resulting in a single value at each given time. This is the upper bound of the superposed plots of all the stations, explaining the U in AU.

AL

Similar to AU, but instead the lowest values at each given time are selected.

AE

The difference between AU and AL, mathematically: $AE = AU - AL$.

AO

The mean of AU and AL. $AO = (AU + AL)/2$.

2.3.5 Dst-index (low-latitude)

For the derivation of the Dst-index, the stations between the equator and auroral zone are used (15° to 35° latitude). The fluctuations in the horizontal component for selected observatories are averaged. Dst is an abbreviation of disturbance storm time and it gives information about the strength of Earth's ring current. This is the current around the equator caused by solar wind particles, weakening the Earth's magnetic field.

2.4 Sun–Earth coupling

The Earth is affected by the Sun through a set of interplanetary parameters, such as the solar wind, IMF, solar flux and solar radio bursts. Studies have found that long-term geomagnetic activities and climate are correlated to the long-term solar variation. For example, a study has shown that the centurial variation of global temperature shows a similar profile as that of the solar cycle length [19]. This was before the human impact on the environment became the major cause of rapid global warming [20].

It is clear that the solar wind and geomagnetic activity are linked, a fact that has inspired scientists to construct certain Sun-Earth coupling functions. These function use the solar wind parameters to predict geomagnetic activity, considering input parameters such as the sunspot number or F10.7 flux (the solar radio flux at 10.7 cm).

In the past, the geomagnetic activity during the solar minimum has been compared for various solar cycles. This showed that, for similar sunspot numbers, the geomagnetic activity is sometimes different. This suggests that there are more parameters involved than just the number of sunspots, such as the solar cycle strength. How effective solar wind is in causing geomagnetic activity is sometimes referred to as Sun–Earth coupling efficiency. A recent study has found that the last solar cycle has been one of decreased Sun-Earth coupling efficiency [21]. This means that for the same solar wind conditions, less geomagnetic activity has occurred.

To test the Sun-Earth coupling function, it must be compared to empirical results,i.e., for longer time scale that contains solar cycles with as low activity as the present cycle. Such a test can only done with data that has very long record, such as measurements by magnetometers. Magnetometers are operating all over the world, measuring the geomagnetic field at their location every time interval. These measurements are converted to K-indices, and subsequently K_p-indices. Considering that the major pathway of electromagnetic energy and plasma from the Sun is through high latitudes where the geomagnetic activity maximizes, it is primarily the data at high latitudes that needs to be analysed.

3 Long-term data sets of geomagnetic activity

Important global geomagnetic indices in the Sun–Earth coupling studies are Dst, K_p, and AE, which are official IAGA (International Association of Geomagnetism and Aeronomy, established 1873) endorsed indices. Dst and AE indices are provided by WDC (World Data Center) for Geomagnetism, Kyoto University, Japan (Dst and AL) [7]. K_p is provided by GFZ, Adolf-Schmidt-Observatory Niemegk, Germany (K_p). The international sunspot numbers (RI) is different number from Wolf number (normally about 40% less) and is provided by WDC-SILSO at the Royal Observatory of Belgium, Brussels (RI).

Among those, K_p index (see Section 2.3.2) has the longest history, and geomagnetic observatory that contribute to K_p has generally much longer operation history than those for Dst (equatorial region) and AE (auroral zone). The K_p indices are readily available since 1932. However, some magnetometers for the Kp stations have been operating before 1932, and the oldest one dates as far back as 1844. By using data of individual magnetic stations, it is possible to re-construct an extended K_p equivalent data set.

3.1 Generating K indices from digital magnetometer data

The systematic construction of an K_p equivalent dataset follows many steps. First, for an individual station, this means reducing raw magnetic field measurements to K indices, see Figure 8.

Individual station – From magnetogram to K-index

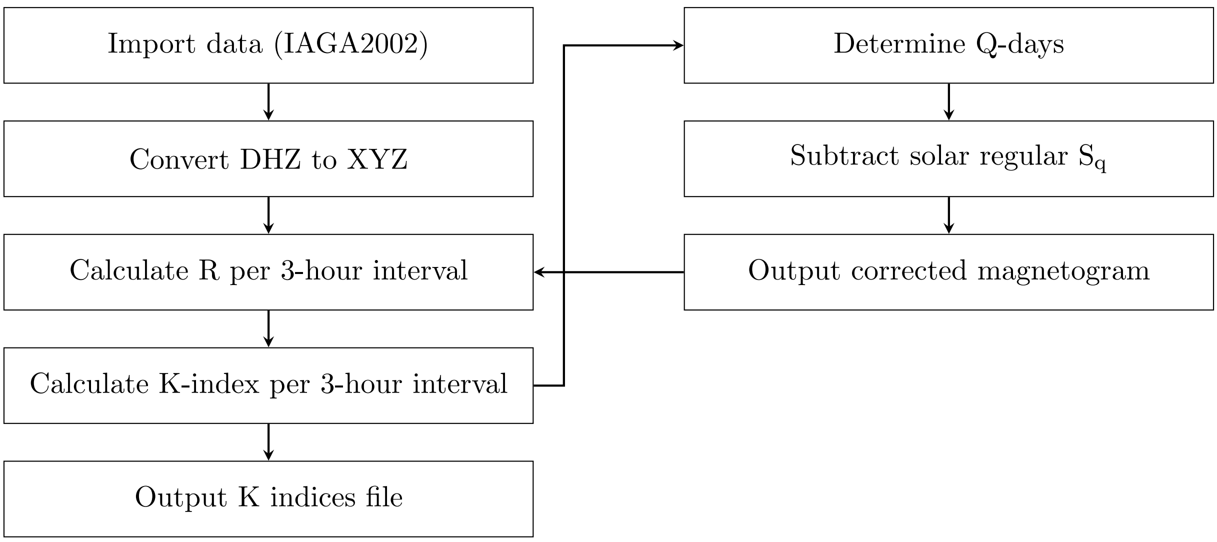


Figure 8: A step-by-step overview of the process from raw magnetometer data to K-indices data (for an individual station).

Raw magnetometer data is digitally available in the IAGA2002 format, which includes the magnetic field components at either hourly, minute or second resolution. Data before 1932 by the WDC Kyoto is only available in hourly resolution. Next, all magnetic field measurements are changed to XYZ representation (using Equation 1 in Section 2.1.1). For every 3-hour interval, the maximum recorded fluctuation (R) is detected. This fluctuation is then converted to a K-index using a conversion table for that specific station. The conversion tables (such as Table 2) are sometimes hard to find, which is why these are calculated as well. This is done by comparing the long-term K-index distribution of the station in Niemegk and the individual station.

At this point, all the K indices have been calculated (8 indices per day). The indices are now used to determine the geomagnetic quiet days (least activity), referred to as Q-days, which will be discussed in the next section. The magnetograms for the Q-days are averaged, resulting in a smooth daily curve, which must be subtracted from the original magnetogram. This results in a corrected magnetogram which is again used to calculate the fluctuations R and K indices, resulting in an accurate set of K indices. The next sections will discuss these steps in more detail.

3.1.1 Subtracting the solar regular curve

Regular changes in the geomagnetic field should not contribute to the K-index and need to be eliminated from the measurements. These regular variations, sometimes called non-K-variations, form the smooth solar quiet curve $\mathrm{S_q}$. This curve, alternatively called the solar regular $\mathrm{S_R}$, is often constructed from a selection of the quietest days (Q-days) of the current month.

The selection of the Q-days is deduced from the K-indices on the basis of three criteria for each day [22]:

1. The sum of the eight K-indices.
2. The sum of the squares of the eight K-indices.
3. The maximum of the eight K-indices.

These three criteria are used to rank each day of the month, resulting in three different order numbers per day. The days are then ordered by the average of these three order numbers, leading to the five quietest days (with the lowest mean order number), and the most disturbed days (with the highest mean order number). Appendix A shows the magnetograms of the five quietest days of January 1996 for the Niemegk observatory.

3.1.2 Effect of magnetometer sampling rate on K-index

Pre-1932 magnetometer data is often at hourly resolution, in contrast to post-1932 minute data. The K-index is determined from a 3-hour interval, which means that hourly data introduces a large drop in the number of data points per K-index compared to the minute data. More specifically, only 4 data points are used for every K-index, compared to the 181 data points for minute-data.

To determine the accuracy of the K-indices for hourly data, a statistical study is performed. Long-term data collected by the station at Niemegk (Germany) is used, as this was the first station for which the K-index scale was defined [15]. The results of the statistical study will be discussed in the next section, but Figure 9 shows the effect for a single day.

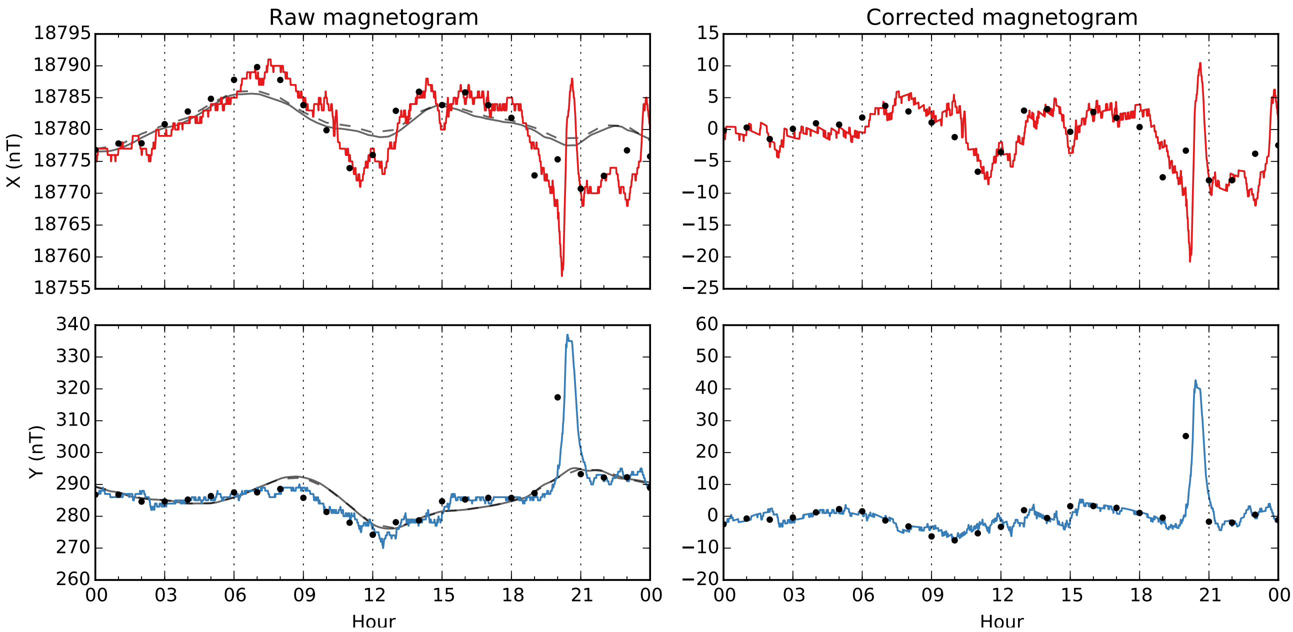


Figure 9: Niemegk, 01-01-1996 – Magnetogram of the X (top) and Y (bottom) component, before (left) and after (right) subtraction of the smooth solar regular. The solid and dashed grey line represent the S_q derived from hourly and minute data, respectively. The dots indicate hourly data points. The disturbance at 20-21 hour shows that hourly resolution is not always sufficient to determine the maximum 3-hour fluctuation.

3.1.3 Niemegk (NGK) – K-index distribution

In this section, long-term (2001-2010, 10 years) minute data and hourly data are compared. Figure 10 shows the distribution of both sets of K-indices. The same algorithms, for determining the K-indices and subtracting the smooth solar regular S_q, have been used for both the hourly and minute data.

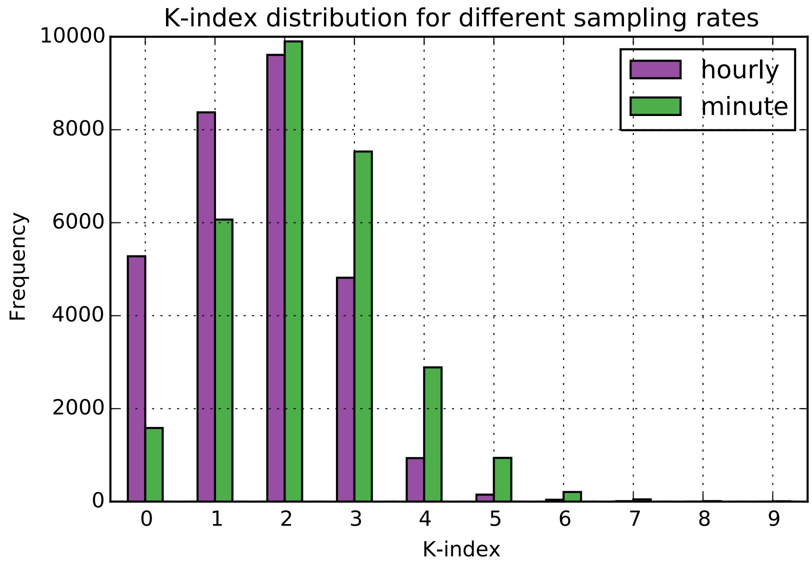


Figure 10: Niemegk, 2001-2010 – K-index distribution for both hourly and minute resolution.

Table 5: Niemegk, 2001-2010 – K-index distribution in numbers. Note that, in total, the minute data has 30 indices fewer than the hourly data. This is due to missing minute data, but this is negligible compared to the 29216 indices in total.

K-index	0	1	2	3	4	5	6	7	8	9
Minute data	1584	6067	9900	7531	2888	941	207	49	11	8
Hourly data	5278	8374	9610	4816	936	150	40	10	2	0

The distribution indicates that, for hourly data, the K-index is often underestimated. The agreement of K-indices for all 3 hour intervals is 35.4 %. This percentage expresses the agreement between the hourly and minute data for every 3-hour interval. Furthermore, the agreement allowing the K-index for hourly data to be one K-index off is 89.2 %. Allowing the K-index only to be lower by one, compared to minute data, results in an agreement of 85.2 %. This confirms that the K-index is often underestimated for hourly data.

However, Figure 10 indicates that it is possible to improve the accuracy of the K-index calculation of hourly data by introducing a separate R-to-K conversion table for hourly data from minute data. A different R-to-K conversion table allows to increase K-indices for the same R only for hourly value to make K-indices derived from hourly data closer to K-indices derived from minute data. This corrected conversion table will then be used to calculate K indices for hourly data.

Figure 11 shows the effect of multiplying the range limits by a single multiplication factor. The maximum achieved agreement is 54.8 %, at a multiplication factor of 0.611. This factor lowers the range limits for the hourly data, resulting in systematically higher K-indices.

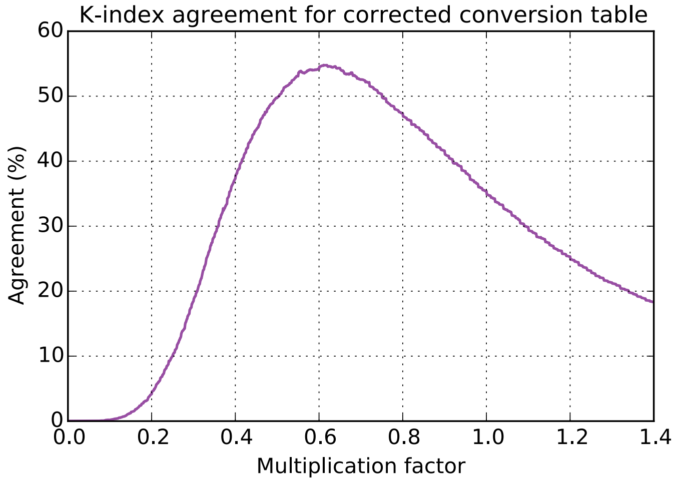
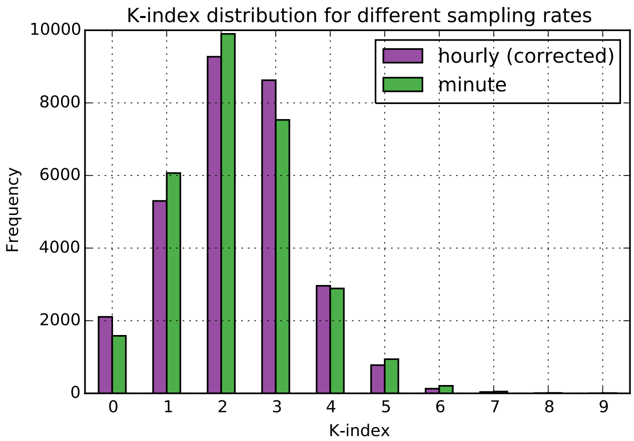


Figure 11: Niemegk, 2001-2010 – K-index agreement over changing conversion table ranges.

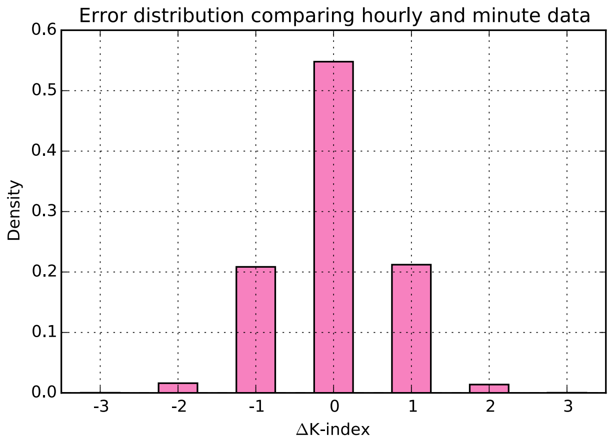
K-index	**R (nT)** before	after
0	0	0
1	5	3.1
2	10	6.1
3	20	12.2
4	40	24.4
5	70	42.8
6	120	73.3
7	200	122.2
8	330	201.6
9	500	305.5

Table 6: Corrected conversion table for Niemegk.

Table 6 shows the range limits for the corrected conversion table that have been multiplied by 0.611. The K indices are now recalculated using this new conversion table, leading to the K-index distribution as shown in Figure 12a. For every 3-hour interval, the K-index is compared to the correct K-index derived from minute data, resulting in the error distribution (Figure 12b). The error distribution confirms that for a multiplication factor of 0.611, the agreement (ΔK-index $= 0$) is 54.8 %. This error makes individual K indices unreliable, but this thesis will use the monthly averages for a statistical study.



(a) Comparing the K-index distribution reduced from hourly and minute data.



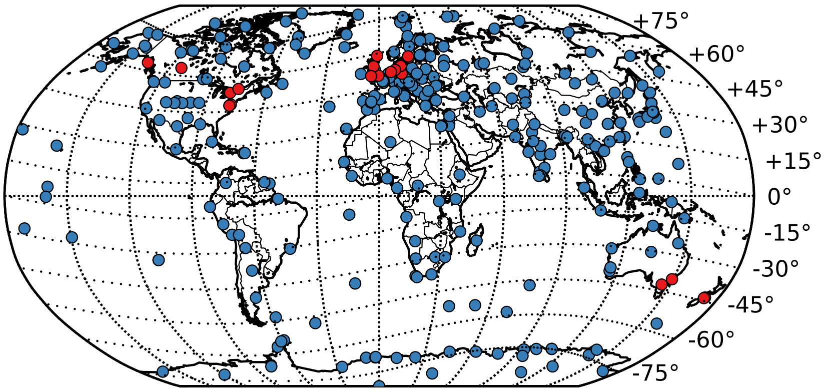
(b) The error distribution for minute - hourly data. The standard deviation of the errors $\sigma = 0.19$.

Figure 12: Niemegk, 2001-2010 – For these plots, the corrected conversion table (Table 6) is used.

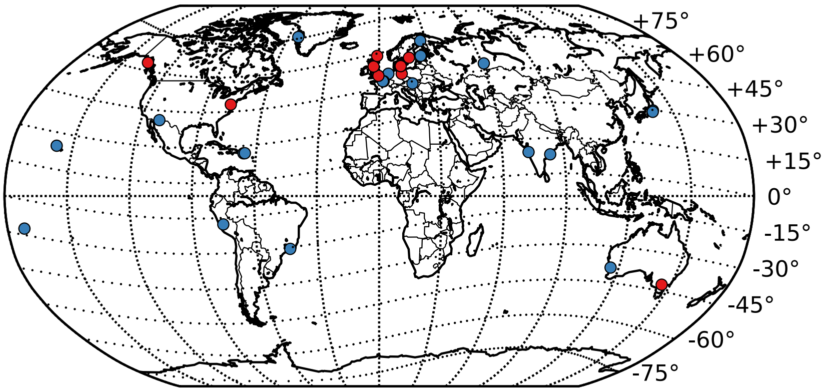
To summarize, by introducing a different R-to-K conversion table for hourly resolution data, it is possible to estimate K indices with greater accuracy, compared to the initial result. The Niemegk observatory has recorded hourly data as far back as 1890, thus featuring 42 years of additional data before 1932 (the start of the official K_p-index). There are several other observatories that were operating before 1932, which will be discussed in the next section.

3.2 Selection of geomagnetic field observatories

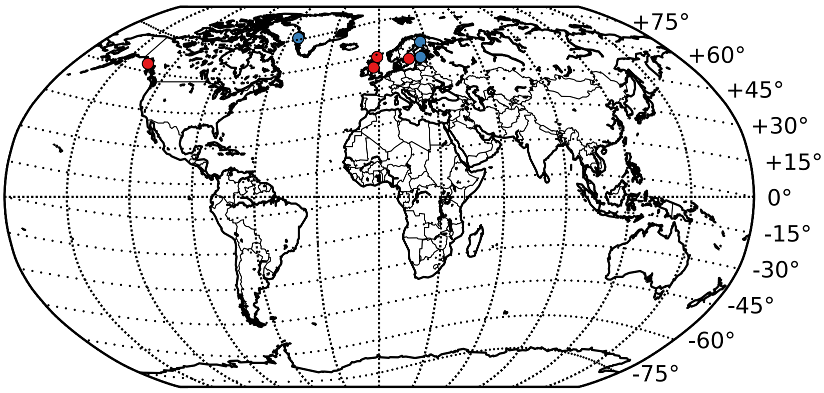
Figure 13 shows the locations of all the stations available by the World Data Center for Geomagnetism, Kyoto. The stations that are operating before 1932 at a high latitude are selected. The major pathway of electromagnetic energy and plasma from the Sun to the ionosphere is through high latitudes where the geomagnetic activity maximizes, which is why high latitude stations are chosen.



(a) All time (295 stations, 21 Kp)



(b) Stations operating before 1932 (29 stations, 9 Kp)



(c) Stations operating before 1932 and >55° (7 stations, 4 Kp)

Figure 13: Geomagnetic stations with data stored at WDC Kyoto. The curved horizontal dotted lines indicate the geomagnetic latitudes, with the latitude indicated on the right. Geographic latitudes would appear as straight horizontal lines, but geographic latitudes are not displayed, except for the equator (0° latitude). The vertical lines indicate geographic longitudes. The red dots indicate stations that are used for the official K_p-index.

Figure 13c shows the selection of stations that are used to construct a pre-1932 K_p-index equivalent. Table 7 shows the details of these individual stations.

Table 7: Details of the selected stations operating before 1932 at high latitudes. For reference, the Niemegk observatory and the K_p-index are also listed. GM stands for geomagnetic, and GG for geographic.

Station	ABB	Start	End	GMLat	GGLat	GGLong
Helsinki	HKI	1845	1909	57.41	60.170	24.980
Sitka	SIT	1902	2010	60.20	57.060	224.670
Eskdalemuir	ESK	1911	2010	57.42	55.317	356.800
Sodankyla	SOD	1914	2010	63.96	67.369	26.630
Godhavn	GDH	1926	2010	77.88	69.252	306.467
Lerwick	LER	1926	2010	61.67	60.133	358.817
Lovo	LOV	1929	2004	57.76	59.344	17.824
Niemegk	NGK	1890	2011	51.64	52.072	12.675
Official K_p		1932	2015			

The K indices for Sodankyla have already been acquired, which means that these will not be calculated again. For Helsinki, after 1897-05, only 5 hour resolution is available, which is not used.

Table 8: Properties of the data available for the selected stations. σ_R is the standard deviation of all the maximum fluctuations (R). For Sodankyla (SOD) the K indices are already available, which is why there is no correction factor presented.

Station	Missing data	σ_R (nT)	Correction	GMLat
Helsinki	16%	143	7.94	57.41
Sitka	5%	52	1.22	60.20
Eskdalemuir	20%	23	1.09	57.42
Sodankyla	1%	-	-	63.96
Godhavn	2%	141	3.86	77.88
Lerwick	10%	45	1.13	61.67
Lovo	20%	30	1.10	57.76
Niemegk	4%	17	1.00	51.64

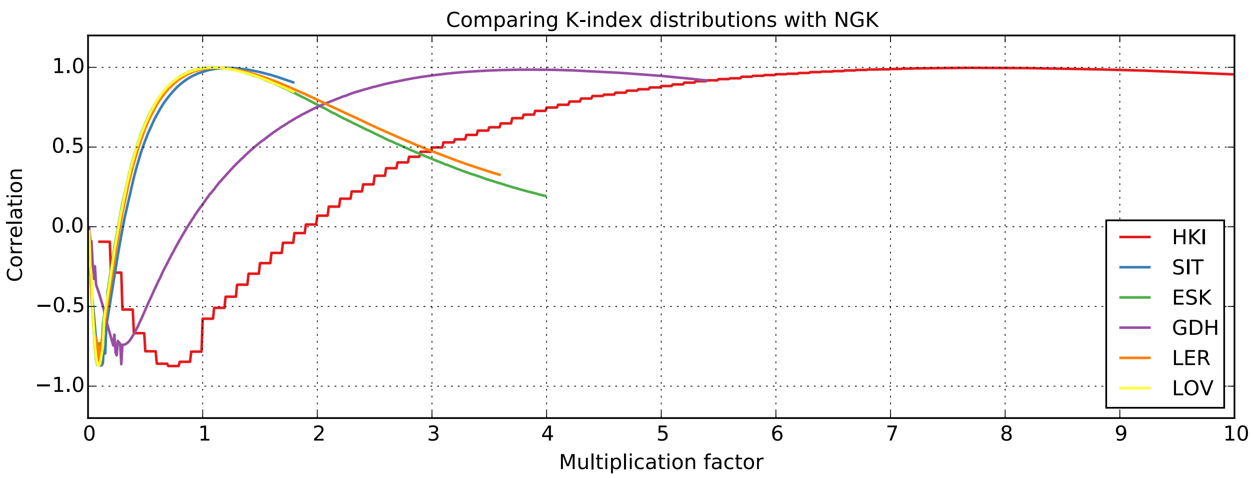


Figure 14: The conversion tables for the selected stations are calculated by multiplying the table for Niemegk a multiplication factor. The distributions are matched using a histogram correlation method.

For each station, the K indices are calculated and the distribution is compared to the station in Niemegk. The definition of the K-index states that, over a long period, the distribution should be the same for all stations (see Section 2.3.1). Geomagnetic fluctuations are larger for higher latitudes, which is why a single R-to-K conversion table would not produce similar K-index distributions. To solve this issue, Niemegk's conversion table is again multiplied by a correction factor. This time, the K-index distributions of the individual station are compared to Niemegk's distribution. Ideally, these distribution must be exactly the same, but a similarity algorithm is used to optimize for the highest similarity (Appendix B). Figure 14 shows the result of this similarity test using all the data available. The figure below (Figure 15) shows the distribution for all the stations after the conversion tables for hourly data have been calculated. Table 9 shows all the hourly conversion table for all the stations.

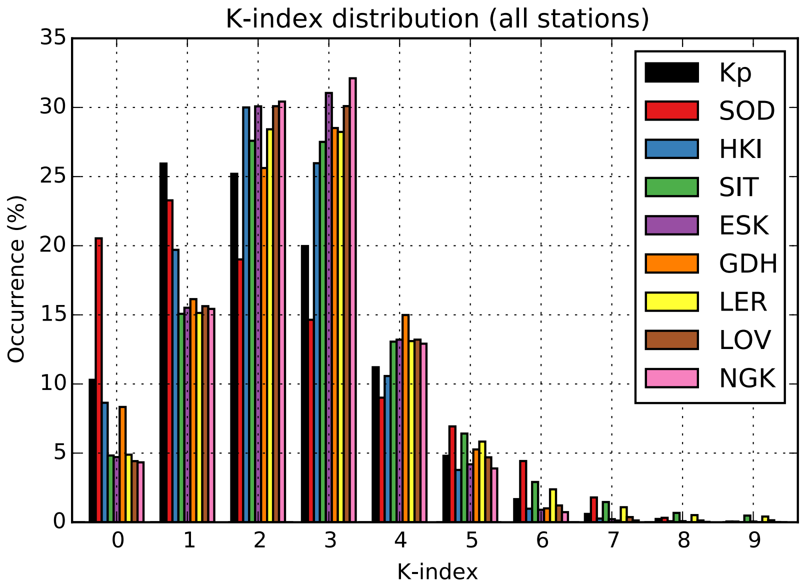


Figure 15: K-index distributions of all the stations.

By definition, the K-index distribution should be, *for the same time period*, similar for all stations on the long term. This is why, when comparing the K-index distributions, only the intersection of measurements are considered. Additionally, it is important to know how many stations are used along the long-term timescale since coverage varies, this is included in later analysis.

Table 9: Conversion tables for all stations.

	R (nT)							
	NGK		**HKI**	**SIT**	**ESK**	**GOD**	**LER**	**LOV**
K-index	minute	hourly	hourly	hourly	hourly	hourly	hourly	hourly
1	5	3.1	24.3	3.7	3.3	11.8	3.5	3.4
2	10	6.1	48.5	7.5	6.7	23.6	6.9	6.7
3	20	12.2	97.0	14.9	13.3	47.2	13.8	13.4
4	40	24.4	194.1	29.8	26.6	94.3	27.6	26.9
5	70	42.8	339.6	52.2	46.6	165.1	48.3	47.0
6	120	73.3	582.2	89.5	79.9	283.0	82.9	80.7
7	200	122.2	970.3	149.1	133.2	471.7	138.1	134.4
8	330	201.6	1600.9	246.0	219.8	778.3	227.8	221.8
9	500	305.5	2425.7	372.7	333.0	1179.2	345.2	336.1
Correction factor		0.611	7.94	1.22	1.09	3.86	1.13	1.10

3.3 Long-term geomagnetic activity trend

Figure 16 shows all the K indices converted to monthly A indices. The A indices for individual stations follow the same trend as the K_p-derived A-index. A indices are used because, as mentioned earlier, it is not meaningful to take the average of a set of K indices. The relationship between the K-index and the geomagnetic fluctuations is non-linear. For an alternative plot showing the monthly chance of $K \geq 4$ see Appendix C.

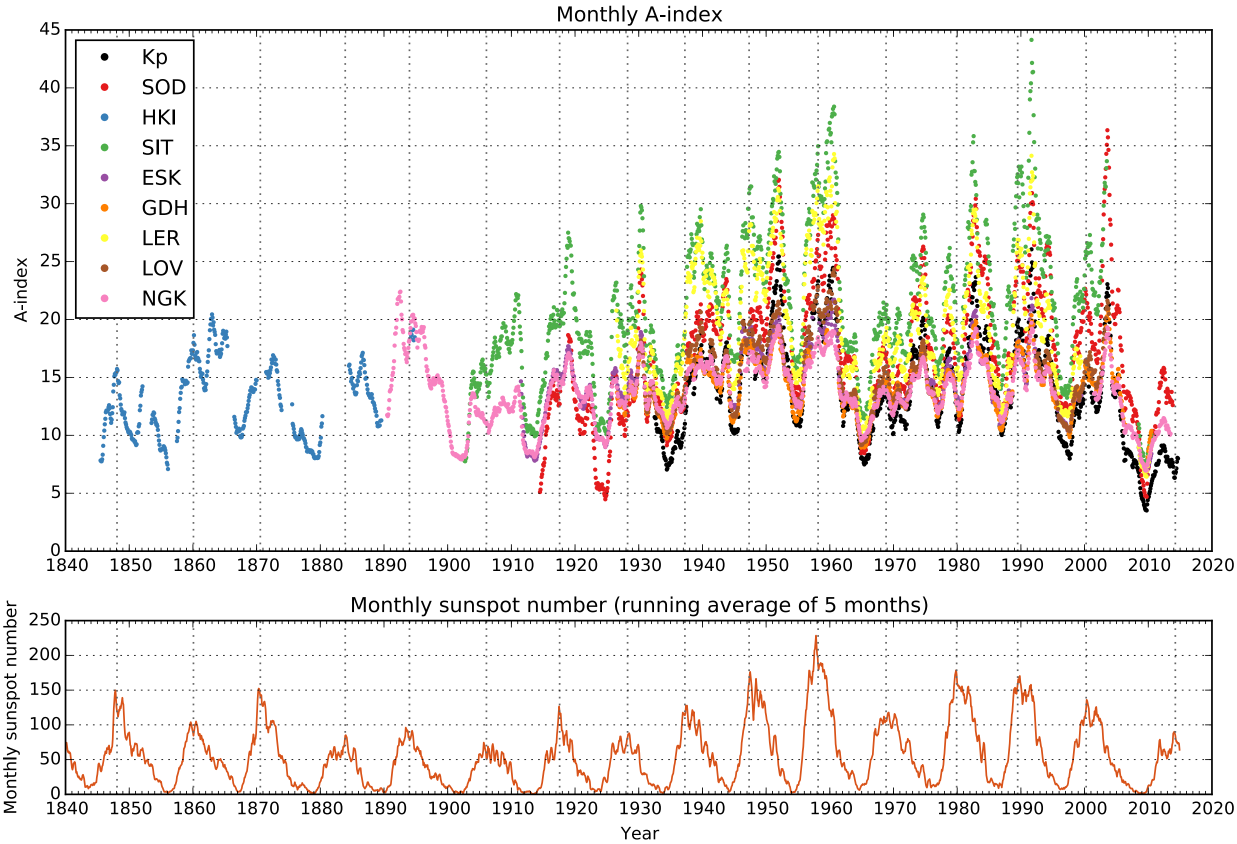


Figure 16: Monthly A indices with a rolling mean of 12 months. Every dot represents a single month. Note that the vertical grid lines indicate each solar cycle maximum instead of decades.

The average of all the stations produces a set of A indices that extends the K_p indices (converted to A indices), as shown by Figure 17. To maximize the fit, the entire equivalent series has been corrected with a multiplication factor of 0.856. The figures (Figure 16 and Figure 17) show that the geomagnetic activity roughly follows the 11-year solar cycles, and that the current solar cycle #24, thus far, features the lowest recorded geomagnetic activity since at least 1915.

This procedure does not take into account how many stations are used per A-index. Because the indices for all the stations (Figure 16) do not overlap perfectly, this might lead to deviations for the equivalent series, where there are very few stations, such as pre-1900. This is why the 19th century indices (Helsinki station) should be used with caution, it might be systematically too high or too low.

Figure 18 shows the error distribution of the equivalent A indices compared to the official A indices. The exact match is 22.7 % (ΔA-index = 0). The resulting extended a indices data can be converted back to K indices, producing a long-term K_p-index equivalent dataset. The next section will investigate the correlation of the long-term A indices with the sunspot number.

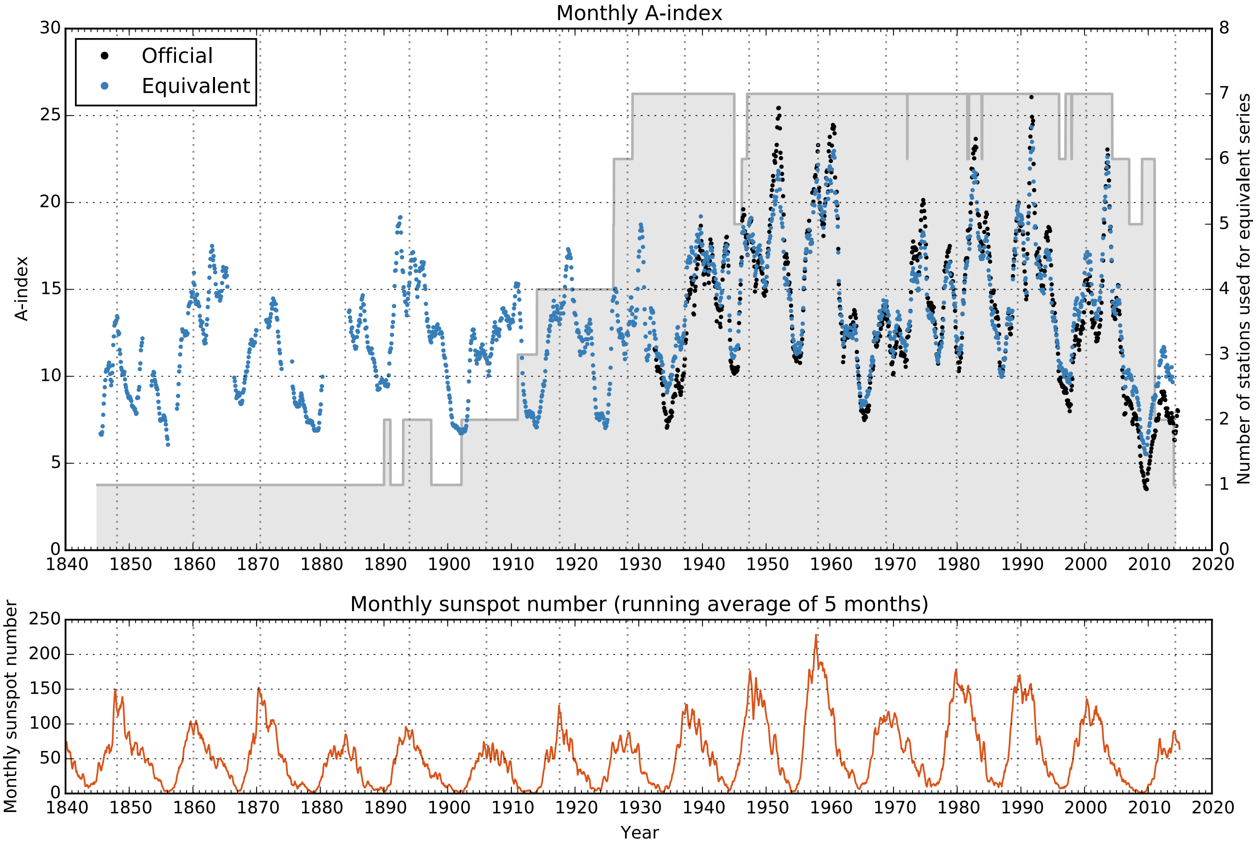


Figure 17: Comparing the monthly A indices to the K_p-derived A indices. (rolling mean of 12 months)

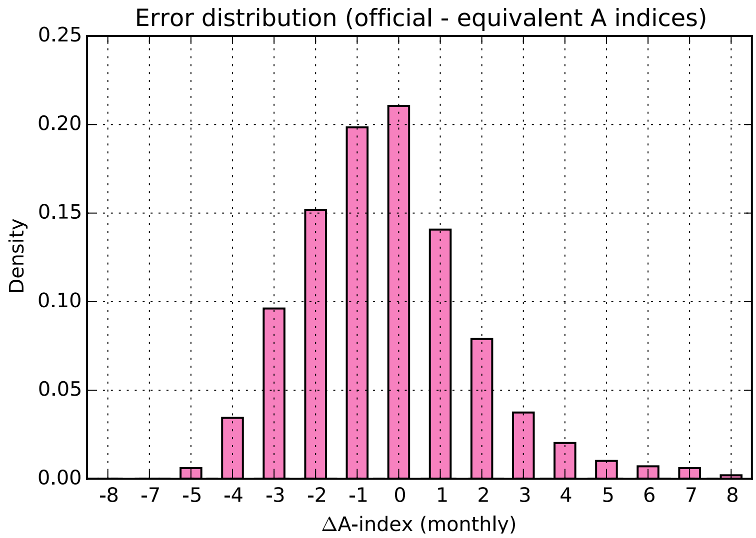


Figure 18: Error distribution of the equivalent A indices: official - equivalent. $\sigma = 2.05$

4 Correlating geomagnetic activity with the sunspot number

The indices can now be used to compare individual solar cycles. By comparing these two quantities, it is possible to determine the geo-effectiveness of the solar cycle. This indicates how effective the solar cycle was in causing geomagnetic activity. More detailed analyses will be performed in the future, below is just an example of the simplest analyses.

4.1 Comparing solar cycle phases

Solar cycles are never exactly the same between different cycles. Therefore, it is necessary to define some criteria to detect the different solar cycle phases. The bottom phase starts when the yearly average of the number of sunspots became equal to or less than 35. The maximum phase starts when the yearly average became equal to or more than 80 % of the number at maximum. The inclining phase and declining phases are defined as the period in between. Next, for each day, the daily A-index and sunspot number is plotted, leading to Figure 20.

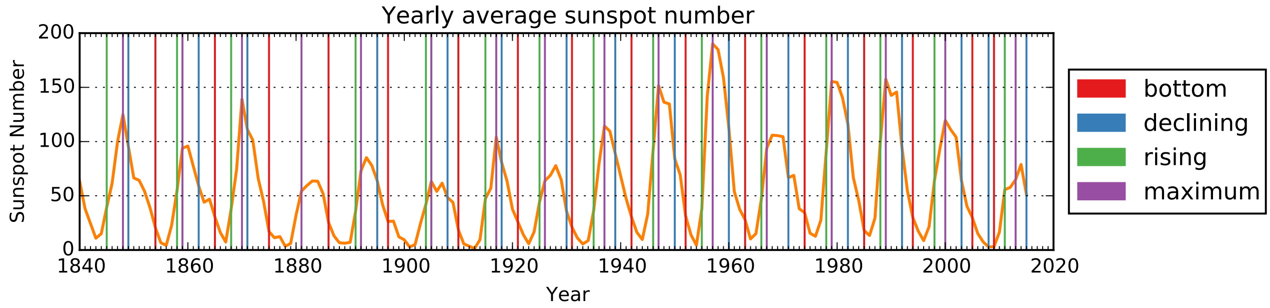


Figure 19: The different vertical lines announce the solar cycle phases.

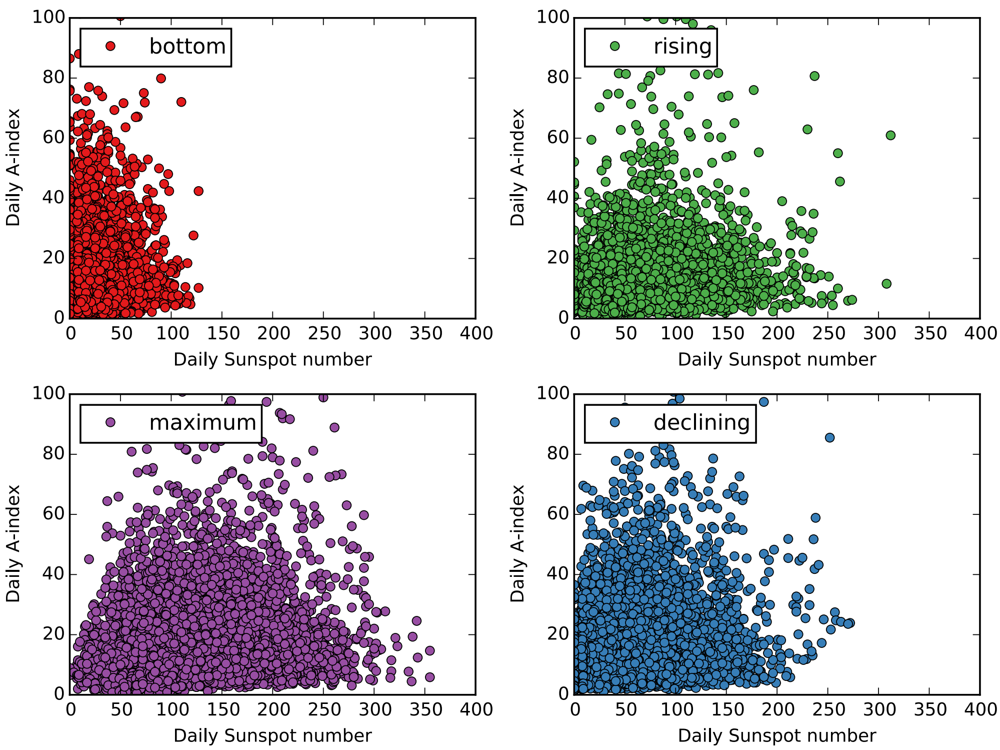


Figure 20: Scatter plot of each day since 1910 (solar cycle #15–#24), comparing the A-index and sunspot number. A two day delay between sunspot number and daily A-index is introduced, to account for the time it takes for the solar wind to reach the Earth.

The data points presented in the scatter plot are now binned per 5 daily sunspot number, as shown by Figure 21.

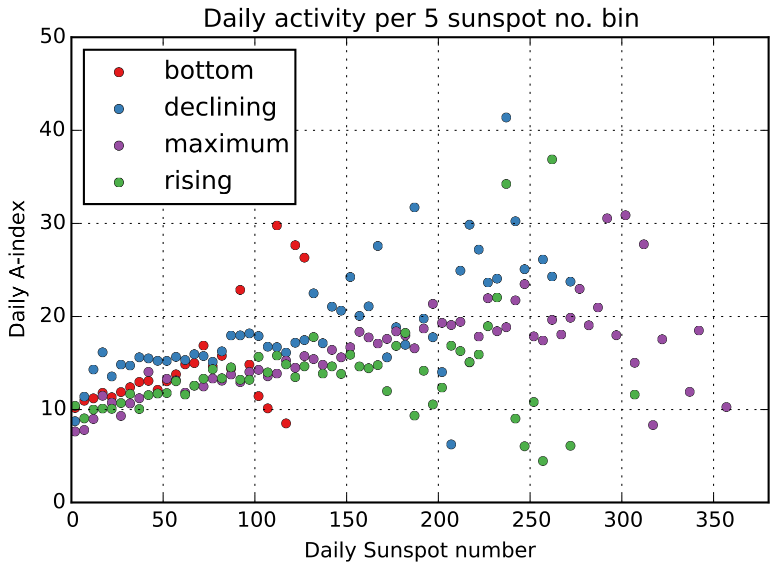


Figure 21: Binned plot of each day since 1910, comparing the A-index and sunspot number.

Since statistics for the sunspot number more than 150 is relatively poor (for the bottom phase, statistics more than 70 is relatively poor), the trend only below 150 of sunspot number is examined (days with many data points). Figure 21 reveals that the declining phase (blue) often shows higher A (and therefore Kp) indices than the other cycle phases. Both the inclining phase (green) and maximum phase (purple) show the low A indices for the same number of sunspots. This suggests that the declining phase is the most geo-effective phase, especially compared to the maximum phase.

4.2 Comparing individual solar cyles

The next page shows a similar analysis comparing daily A indices and sunspot numbers, but this time individual solar cycles are analysed. The solar cycles #13–#24 are examined (starting in 1890). Figure 22 and Figure 23 only differ in color coding, see the captions for details. Figure 22 shows that the cycle-to-cycle trend followed a brief upward trend (around cycle #19) before falling to the current low geomagnetic activity. Figure 23 shows that maximum number of sunspots for each cycle appears to be an indicator of geo-effectiveness. Solar cycles with a high maximum number of sunspots also cause higher A indices, for the same number of sunspots. This suggests a parameter such as the solar cycle strength for each solar cycle. More detailed analyses will be performed in the future (see Appendix F).

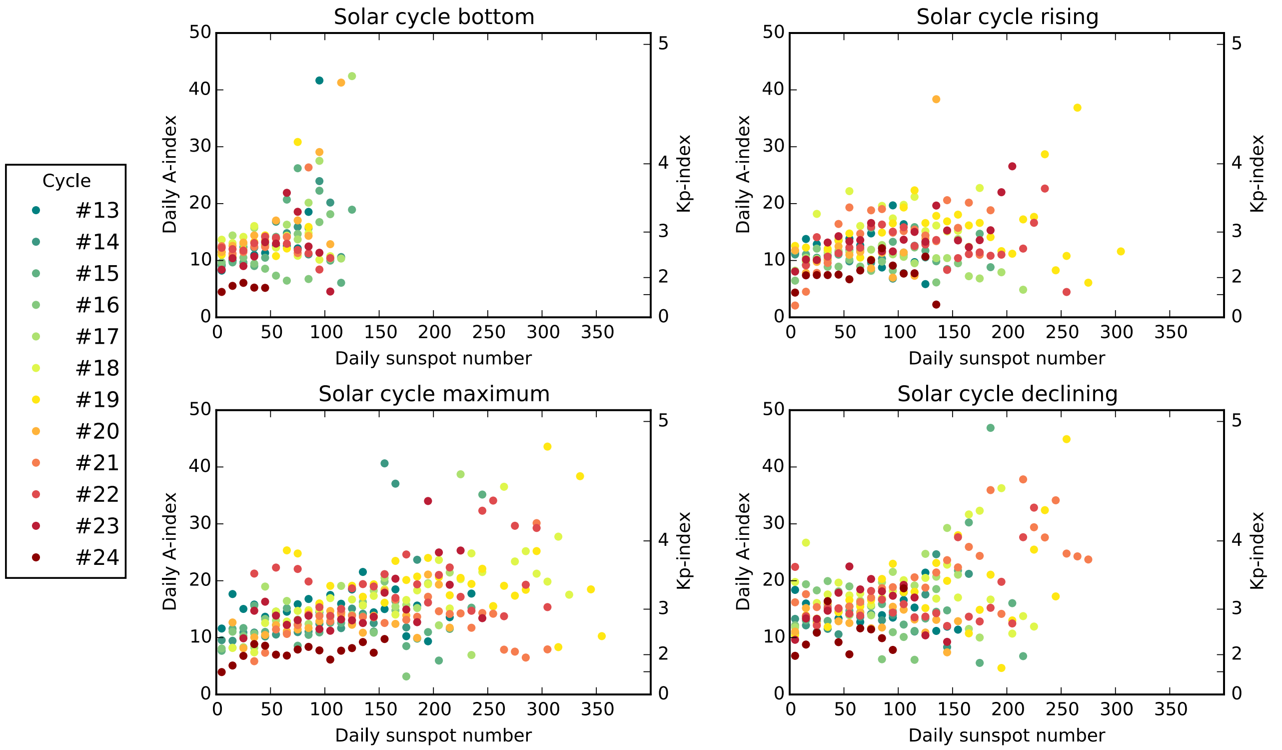


Figure 22: Color coding shows the long-term cycle-to-cycle trend. Binning by 10 sunspot no. bin. The right axis shows the equivalent K_p-index, following Table 4.

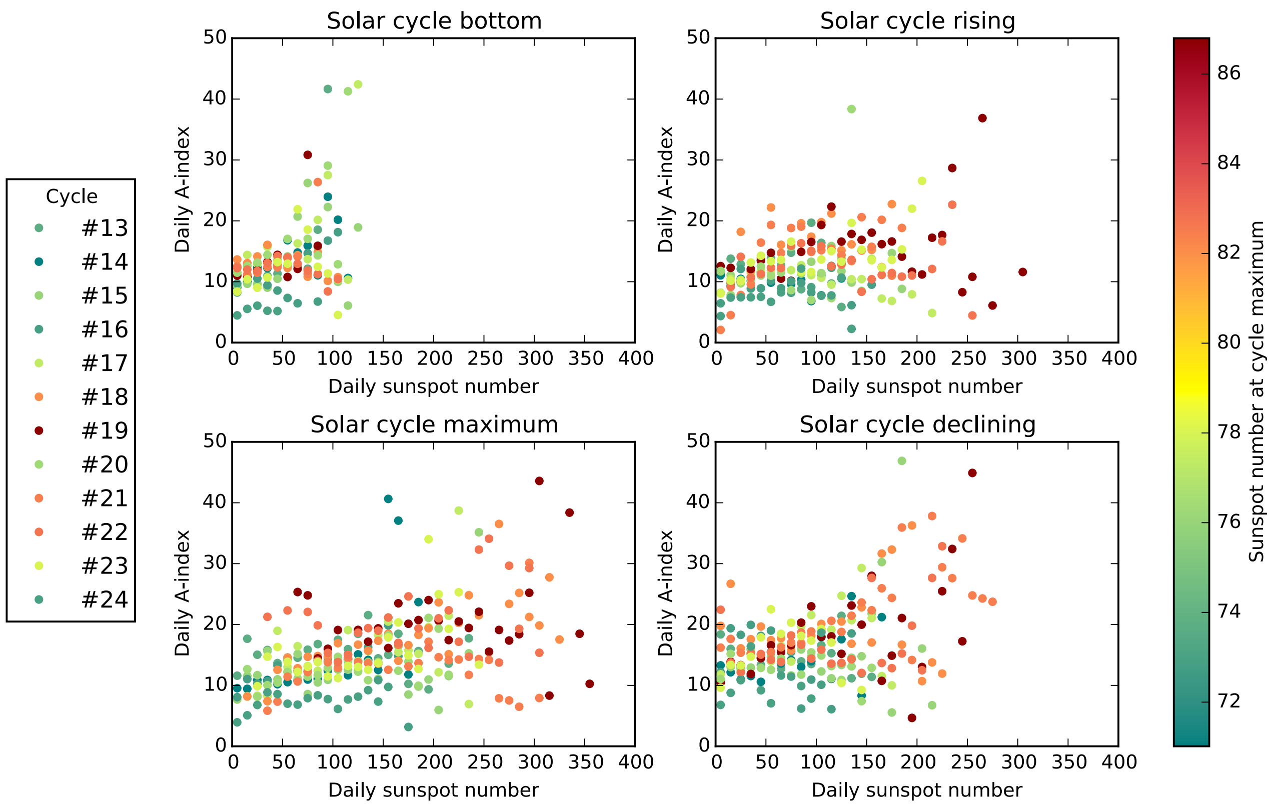


Figure 23: Color coding shows the height of the solar cycle, more specifically, the maximum yearly average of number of sunspots. Binning by 10 sunspot no. bin.

5 Conclusion

The K_p-index is a commonly used indicator of geomagnetic disturbances on a planetary scale. The official index is only available since 1932, but this report introduces a method to extend the dataset with 88 years, providing coverage since 1845. By using hourly data of individual geomagnetic observatories, a K_p-index equivalent data set is systematically calculated. The stations Helsinki, Sitka, Eskdalemuir, Sodankyla, Godhavn, Lerwick, Lovo and Niemegk are used.

This bachelor thesis also includes an analysis of the correlation of sunspot numbers and geomagnetic activity. The analysis shows that the current solar cycle #24 has the lowest geomagnetic activity recorded since at least 1915. Additionally, during declining phase of the solar cycles, more geomagnetic activity for the same number of sunspots is found compared to the other phases of the cycle. When comparing individual cycles, some cycles also show different geomagnetic activity for the same number of sunspots, suggesting that some cycles are more geo-effective than others. The number of sunspots at the maximum of the solar cycle appears to be an indicator of the strength of the solar cycle. More detailed analyses will be performed in the future.

The extended K_p-index provides insight into centurial trends and can be used to guide future research in Sun–Earth coupling functions. The software developed for this thesis can also be used in the future to calculate R-to-K conversion tables and K indices for other observatories.

Acknowledgement I'd like to thank Dr. Masatoshi Yamauchi for providing the opportunity to work for the Swedish Institute of Space Physics, and guiding me during this internship.

6 References

[1] SILSO World Data Center. The International Sunspot Number. *International Sunspot Number Monthly Bulletin and online catalogue*, 1749-2015.

[2] William B. Cade and Christina Chan-Park. The origin of "space weather". *Space Weather*, 13(2):99–103, 2015. 2014SW001141.

[3] M.G. Kivelson and C.T. Russell. *Introduction to Space Physics*. Cambridge atmospheric and space science series. Cambridge University Press, 1995.

[4] H. Gleisner. *Solar Wind and Geomagnetic Activity - Predictions Using Neural Networks*. Lund Observatory, 2000.

[5] NASA – The Earth's Plasmasphere. `http://plasmasphere.nasa.gov`. Accessed: 25-02-2015.

[6] NASA – IMAGE Spacecraft Pictures Aurora. `http://earthobservatory.nasa.gov/IOTD/view.php?id=6226`. Accessed: 27-04-2015.

[7] Kyoto World Data Center for Geomagnetism – Geomagnetic Data Service. `http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html`. Accessed: 16-02-2015.

[8] Wallace H. Campbell. *Introduction to geomagnetic fields*. Cambridge University Press, second edition, 2003.

[9] Kyoto World Data Center for Geomagnetism – Magnetic North, Geomagnetic and Magnetic Poles. `http://wdc.kugi.kyoto-u.ac.jp/poles/polesexp.html`. Accessed: 27-04-2015.

[10] John A. Eddy. The maunder minimum. *Science*, 192(4245):1189–1202, 1976.

[11] NASA – A Sunspot Up Close. `http://apod.nasa.gov/apod/ap051106.html`. Accessed: 27-04-2015.

[12] Werner Heil. NASA – The 3-D shape of the HCS. `http://lepmfi.gsfc.nasa.gov/mfi/hcs/hcs_shape.html`. Accessed: 21-04-2015.

[13] Bruce T. Tsurutani Edward J. Smith. Observations of the interplanetary sector structure up to heliographic latitudes of 16°: Pioneer 11. *Journal of Geophysical Research*, 83(A2), 1978.

[14] Southwest Research Institute – Interplanetary magnetic field (IMF). `http://pluto.space.swri.edu/image/glossary/IMF.html`.

[15] J. Bartels, N. H. Heck, and H. F. Johnston. The three-hour-range index measuring geomagnetic activity. *Journal of Geophysical Research*, 44(4):411–454, 1939.

[16] NOAA Space Weather Scales. `http://www.swpc.noaa.gov/noaa-scales-explanation`. Accessed: 27-03-2015.

[17] ISGI – Kp Indice. `http://isgi.latmos.ipsl.fr/des_kp_ind.html`.

[18] Kyoto World Data Center for Geomagnetism. *Auroral electrojet indices (Provisional AE) 1992*. Number 25. 1998.

[19] E. FRIIS-CHRISTENSEN and K. LASSEN. Length of the solar cycle: An indicator of solar activity closely associated with climate. *Science*, 254(5032):698–700, 1991.

[20] IPCC (2014) Working group I contribution to the IPCC fifth assessment report climate change 2013: the physical science basis. The Intergovernmental Panel on Climate Change. `http://www.ipcc.ch/report/ar5/wg1/`.

[21] M. Yamauchi. Decreased Sun-Earth energy-coupling efficiency starting from 2006. *Earth, Planets and Space*, 67(44), 2015.

[22] GFZ Potsdam Q/D-days. http://www.gfz-potsdam.de/en/research/organizational-units/departments/department-2/earths-magnetic-field/services/kp-index/explanation/qd-days. Accessed: 05-03-2015.

[23] OpenCV - Histogram Comparison. http://docs.opencv.org/doc/tutorials/imgproc/histograms/histogram_comparison/histogram_comparison.html. Accessed: 17-03-2015.

[24] Arnold Hanslmeier. *The Sun and Space Weather*. Springer, 2 edition.

A Quietest and most disturbed days of January 1996, Niemegk

Figure 24 shows the five quietest and most disturbed days of January 1996 at the Niemegk observatory. The mean of the quietest days forms the solar regular curve, which will be subtracted from the absolute magnetometer data.

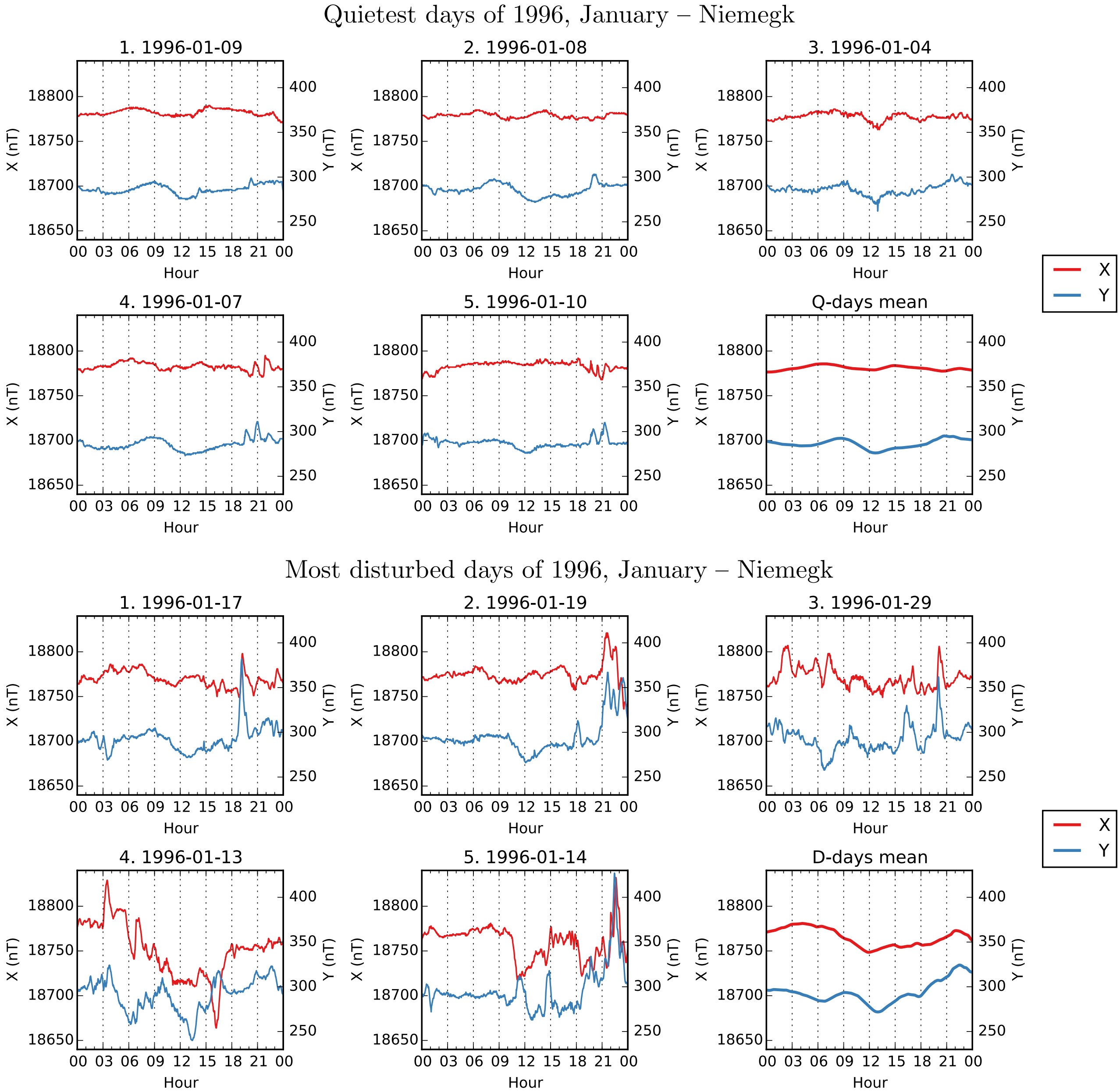


Figure 24: The five quietest and most disturbed days of 1996, January, in minute-data resolution. The dotted vertical lines show the 3-hour interval for the day. The mean magnetogram also includes a rolling mean of 120 minutes. Both Y axes display a range of 200 nT and show the X and Y component of the geomagnetic field.

B Histogram comparison

The equation below expresses how well two histograms (H_1 and H_2) correlate:

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H_1})(H_2(I) - \bar{H_2})}{\sqrt{\sum_I (H_1(I) - \bar{H_1})^2 \sum_I (H_2(I) - \bar{H_2})^2}} \tag{3}$$

where:

$$\bar{H_k} = \frac{1}{N} \sum_J H_k(J)$$

and N is the total number of histogram bins [23]. The output d ranges from -1 (negatively correlated) and 1 (identical).

Figure 25 (below) shows some example scenarios for a histogram with 4 bins.

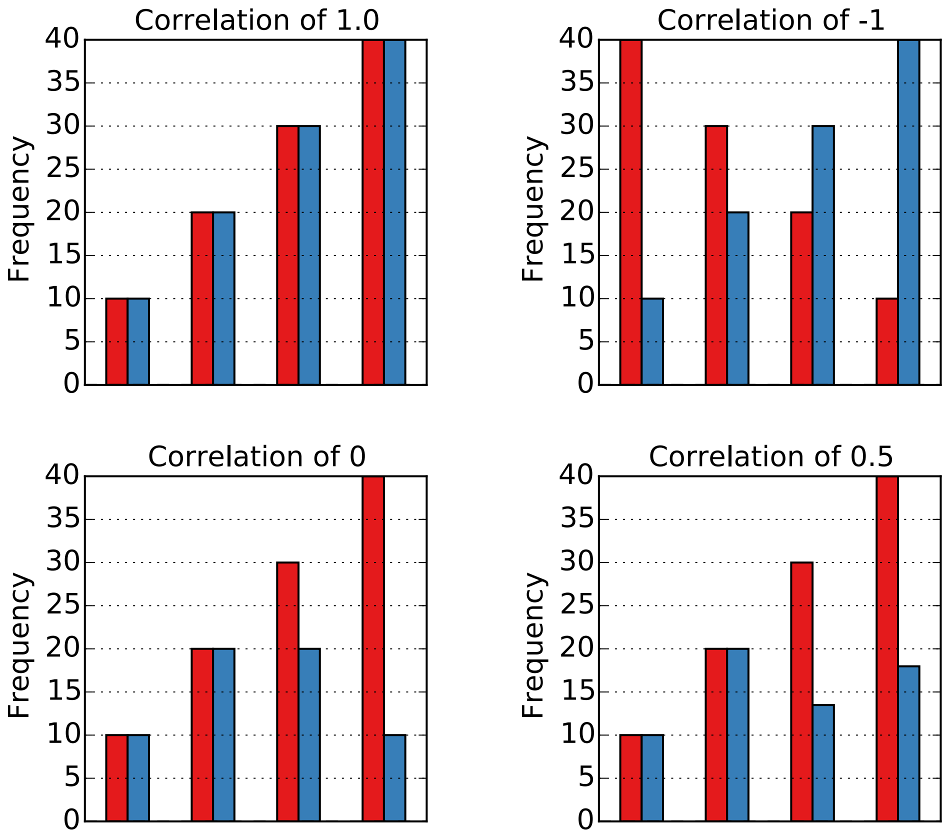


Figure 25: Examples of histogram comparison, with the calculated correlation.

C Monthly chance of K ≥ 4 for selected stations

Figure 26 shows the monthly chance of reaching a K-index of 4 or higher for the selected stations in this thesis. It is clear that the current solar cycle features very low chances of reaching high K indices.

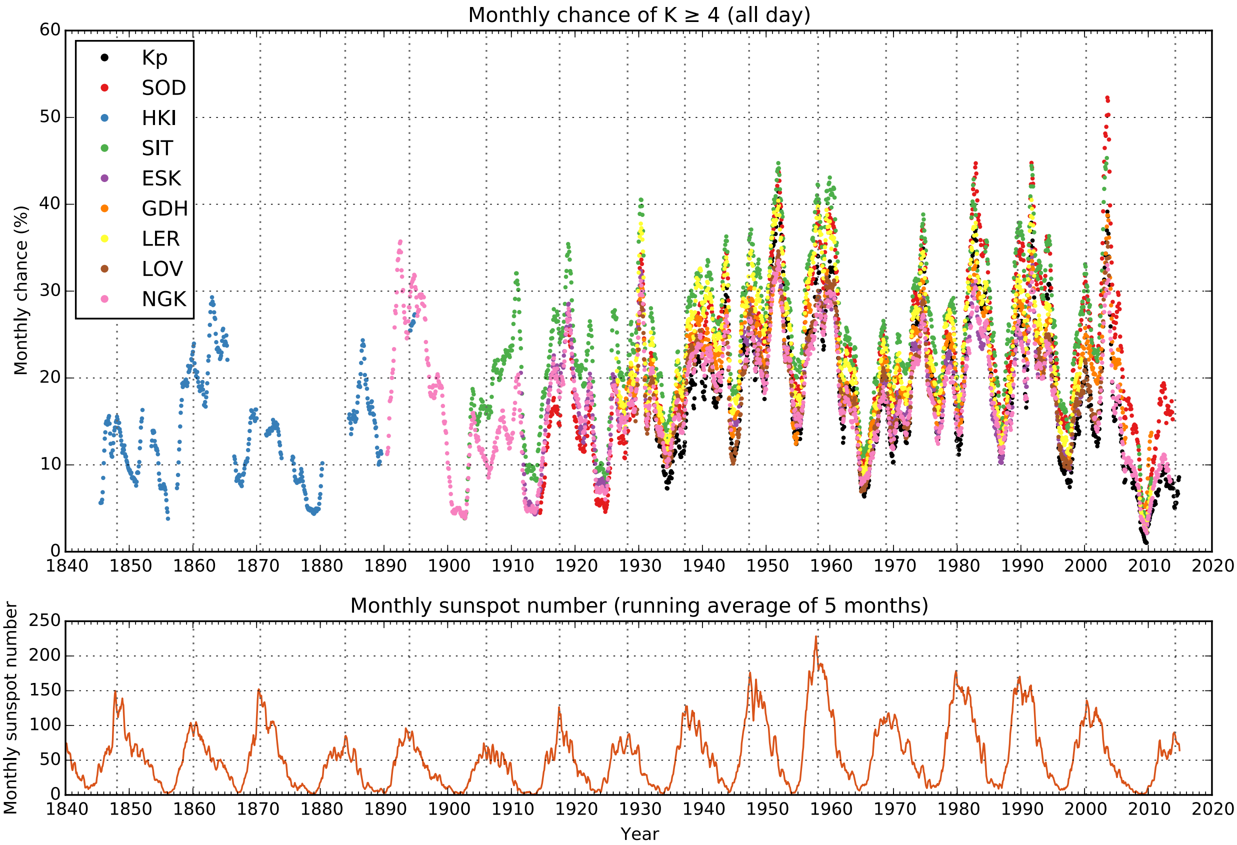


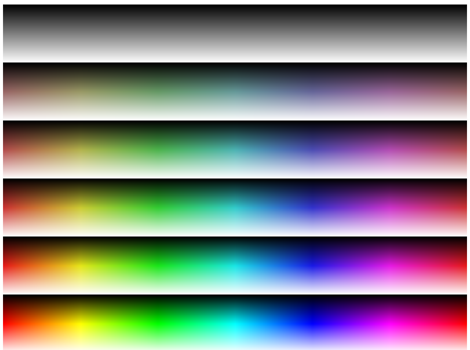
Figure 26: Monthly A indices with a rolling mean of 12 months. Every dot represents a single month. Note that the vertical grid lines indicate each solar cycle maximum instead of decades.

D Visual detection of auroras using an all-sky camera

This appendix section introduces a heuristic algorithm for detection of auroras in all sky camera images.

The institute (IRF Kiruna) has an all-sky camera installed on the roof, which takes pictures every minute of the night sky. Auroras are systematically recorded, with colors varying from green, red and, on rare occasions, blue. These colors correspond with the constituents of the atmosphere, atomic oxygen (green/red) and nitrogen (red/blue) [24].

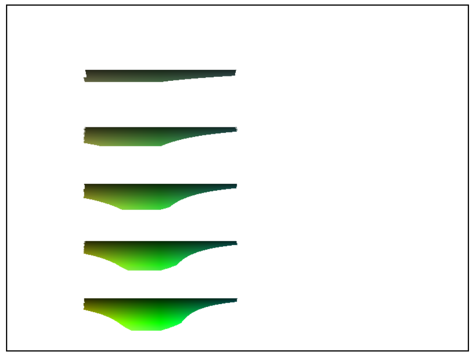
For a single pixel to be considered part of an aurora, it must follow all the rules (Figure 27).



(a) All colors

- $R \leq 138$
- $G \geq 40$
- $B \leq 68$
- $G > R$
- $G > B$

(b) Filter



(c) Matching colors

Figure 27: The selection of colors shown in the HSL (hue, saturation, lightness) color model. Hue and saturation are shown on the x and y axis, respectively. Vertically, the different bars present different values for the lightness parameter, from top to bottom: 0.0, 0.2, 0.4, 0.6, 0.8 and 1.0.

The values for the first three rules have been determined by picking various aurora colors. After implementing the first three rules, it was found that clouds also match these conditions, resulting in the two extra rules ($G > R$, $G > B$). Clouds rarely contain more green than red and more green than blue. The combination of these simple rules proves to work well and leads to detection of auroras ranging from very faint to strong.

So far the algorithm does not distinguish between strong and faint aurora. To add this functionality, every pixel (in RGB) is converted to the HSL color model, and the lightness component (ranging 0 to 0.1) is used as an indicator of aurora strength. The lightness component would be 1.0 for a white pixel, but white is never part of the matched colors (see Figure 27c). This is why the lightness component can be safely scaled so that the maximum achievable lightness is again 1.0.

The next page shows a demonstration of six different real all-sky camera images, and the result of the method presented here (Figure 28).

This algorithm provides an interface to statistical research of the auroras, and could be used to reduce images by multiple all-sky camera stations to a global map. This is beyond the scope of this thesis.

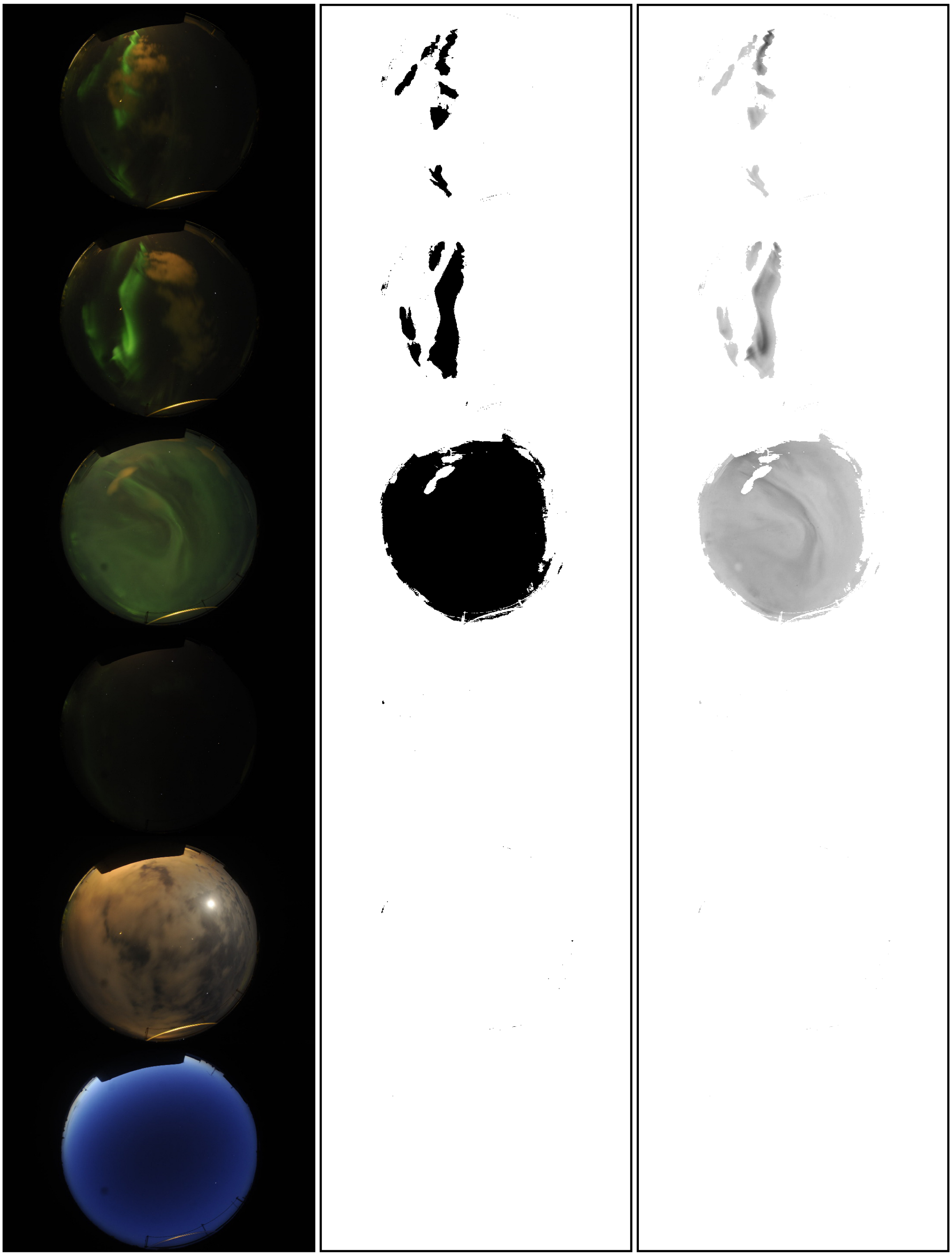


Figure 28: Demonstration of the capabilities of the presented method. The input (left), the captured pixels (middle) and the aurora intensity (right). The cases presented here are, from top to bottom, an aurora with clouds interfering, a strong aurora, faint aurora, black night sky, cloudy night, blue sky, respectively. The images are rotated 90° anticlockwise, resulting in the north pointing left.

E Developer's guide

The results and figures in this bachelor thesis are generated by scripts specifically developed for the task. This section explains how to get these programs running and reproduce the results.

The Python programming language is used because it is getting increasingly popular among scientists, and there are many libraries for scientific purposes available. Version 3.4.2 is used, but in practice, any 3.x version should suffice.

E.1 Installation

The official Python website (`python.org`) includes instructions on installing Python 3.x. After Python is installed, several packages need to be added. Python installations include *pip*, the default package manager. The following commands install the necessary packages using *pip*:

```
pip install matplotlib
pip install pandas
pip install "ipython[notebook]"
```

Matplotlib is responsible for drawing the plots using a syntax similar to Matlab. *Pandas* is a data analysis library, especially useful when working with data in table format. *Pandas* also adds features of the statistical programming language R to Python. The last package, *IPython Notebook*, adds an interactive mode to Python (hence the name), which will be discussed in the next section.

E.2 Notebooks

IPython Notebook provides a popular alternative for working with code, which offers more flexibility than the traditional approach. This is especially well suited to data analysis, . The traditional approach is to write code in a single file and rerun the entire code after every change. Notebooks allows for a more interactive approach, breaking up code in small pieces and providing immediate feedback. The following command starts IPython Notebook, which starts the interface in the user's browser. It is important to start this command from the same directory containing the notebook files (the *.ipynb* files).

```
ipython notebook
```

The interface should now be running locally on `http://localhost:8888`. A list of all the available notebooks will appear upon opening the web interface. Each notebook provides the step-by-step process of a task presented in this thesis.

E.3 World Data Center for geomagnetism, Kyoto

The code operates on data provided by the WDC for Geomagnetism, Kyoto. The data service web interface only provides access to limited amounts of data per request, which is why this download process is automated using a script. Execute the following commands to download all the hourly resolution raw magnetic data:

```
cd data
./wdc_hourly_longterm NGK 1800 2015
./wdc_hourly_longterm HKI 1800 2015
./wdc_hourly_longterm SIT 1800 2015
./wdc_hourly_longterm ESK 1800 2015
```

```
./wdc_hourly_longterm GDH 1800 2015
./wdc_hourly_longterm LER 1800 2015
./wdc_hourly_longterm LOV 1800 2015
```

The script will automatically skip years that are not provided by the WDC Kyoto, and fragment the request into small batches. The Sodankyla (SOD) station is not included because the K indices for this station were already available.

E.4 Reference list

This reference list shows the link between the figures in this thesis and the code/notebook that produced the result.

Table 10: Figures and codes.

Figure #	Notebook name	Description
Figure 4	Geomagnetic maps	Geomagnetic latitudes and longitudes
Figure 5	Solar cycles	Monthly sunspot number
Figure 9	Niemegk	Effect of magnetometer sampling rate
Figure 10	Niemegk	K-index distribution histogram
Table 5	Niemegk	K-index distribution numbers
Figure 11	Niemegk	Correction factor curve
Table 6	Niemegk	Corrected conversion table
Figure 12	Niemegk sampling	Corrected K-index distribution and error distribution
Figure 13	Geomagnetic maps	WDC Kyoto coverage
Table 7	Geomagnetic maps	Details of selected stations
Table 8	Geomagnetic maps	Properties of selected stations
Figure 14	Corrections	Determine correction factors
Figure 16	K indices	Monthly A-index per station
Figure 15	K indices	K-index distribution per station
Figure 17	K indices	Comparing stations with Kp
Figure 18	K indices	Comparing stations with Kp error distribution
Figure 19	Coupling	Cycle phase detection
Figure 20	Coupling	Daily A-index – Sunspots
Figure 21	Coupling	Daily A-index – Sunspots binned
Figure 22	Coupling	Coupling per cycle
Figure 23	Coupling	Coupling per cycle maximum
Figure 26	K indices	Monthly chance of K ≥ 4
Figure 27	Aurora	Aurora color filter
Figure 28	Aurora	Aurora detection

F Project description

This bachelor thesis is part of the following project:

Terrestrial consequences of the present extremely low solar maximum

Supervisor: Dr. M. Yamauchi
Postal address: Box 812, SE-981 28 Kiruna, Sweden

F.1 Purpose and aims

Question of how the Earth is influenced by the extremely low solar activity starting from the last solar minimum is a major science issue because it might impact space weather, magnetospheric/ionospheric physics, astrobiology, and even the climatology. Such a study is urged by the Scientific Committee on Solar Terrestrial Physics as "International Study of Earth-affecting Solar Transients/MinMax24" (SCOSTEP, 2013).

Along this purpose, we aim to quantify the difference in how the solar wind energy ultimately influences the Earth (the Sun-Earth coupling efficiency) between the current solar maximum and the previous several solar maximums (2000-2001, 1989-1990, 1979-1980, 1968-1969, 1957-1958 and so on), in the form of electromagnetic propagation and particle injection into the ionosphere and to the ground. The focus of this project is on the long-term variation of the Sun-Earth coupling efficiency over different solar cycles using the following data taken many decades and world-wide, including data by ourselves (IRF).

- Geomagnetic disturbance (magnetometer)
- Ionospheric conditions such as density, temperature, and disturbances (ionosondes and riometer)
- Auroral and substorm activities (optics)

F.2 Description

The task is to systematically correlate the solar input and terrestrial disturbances, mixing both short-term view and long-term view. Such a study will be different from the past ones if (1) both the long-term solar parameters (e.g., average over each solar cycle) and short-term solar parameters (instantaneous values) are considered in deriving the long-term variation of the Sun-Earth coupling efficiency, and if (2) different phases of a solar cycle are treated differently.

The project has two parts with different databases: years with in-situ solar wind data from spacecraft (past five solar cycles with 5 min resolution), and years without in-situ solar wind data. The former case is straightforward: for several different functions, we will first obtain the coefficient that appears in the proposed function separately at the different phases of the solar cycle. Then, we obtain the coefficient at a specific phase (e.g., inclining phase) of different solar cycles, e.g., after 2006 (when the extremely low solar activity started) and 1965-2005 ("normal solar cycles"). This naturally requires identifying of the solar condition that switches from one phase of a solar cycle to another with different Sun-Earth coupling efficiency.

We also plan to generalize this study to long-term data such as the ionospheric data (this requires making the database of the ionospheric parameters from the ionogram) and geomagnetic data high-latitude station with very long digital record (either hourly values or K index). Considering that the major pathway of electromagnetic energy and plasma from the Sun to the ionosphere/ground is through high latitudes where the geomagnetic activity maximizes, we should primarily analyze data at $> 60°$ geomagnetic latitude (GMlat).

At World Data Center in Kyoto, hourly values are available for (1) Sitka (SIT: 60.2° GMlat) since 1902, and (2) Sodankylä (SOD: 63.9° GMlat) since 1914. If we include stations at 55-60° GMlat, Helsinki (HKI: 57.4° GMLat) data exists for 1844-1909 and Eskdalemuir (ESK: 57.5° GMlat) data exists since 1911. Thus minimum dataset is available. By consulting individual observatories, it is possible to obtain even more data (we have actually obtained data from Sodankylä). As the solar parameters, we use the Smoothed Sunspot Number (SSN) and F10.7 flux in a similar way as we have done in the preliminary study, but with higher time resolution (using daily values instead for monthly values, such that averaging can be done after obtaining the correlation).

The correlation work should aim at reconstruction of the geomagnetic and ionospheric data from SSN and F10.7 flux. After such a general examination, we also need quantitative examinations, e.g., by obtaining values (or distribution) of geophysical parameters for given input SSN and F10.7 values.

Following the correlation study with geomagnetic activities, we examine the ionospheric and auroral conditions using ionosonde data and all-sky camera data. The ionosonde data that were taken in-house exists in digital image form since 1950's, while all-sky data that were taken in house exists in digital image form since 2001 and in 16 mm film format much longer (reading result on paper exists for these film data). Key issue for such a study is to parameterize these data that were basically taken as image data.

Figure 29: The location of Kiruna, Sweden.