Optimization, evaluation and application of EEP measurements in the presence of contaminating protons

Astrid Haderlein

Master's Thesis in Physics

Supervisor: Dr. Hilde Nesse Tyssøy
Co-supervisor: Dr. Marit I. Sandanger

Department of Physics and Technology
University of Bergen

December, 2017

Abstract

It is well known that energetic particles can penetrate deep into the atmosphere and initiate a chain of chemical reactions, ultimately leading to enhanced production rates of HO_x and NO_x molecules and the depletion of local ozone abundances. This, in turn, alters Earth's radiation balance and might influence winter climate. In recent years, numerous studies have been conducted in order to examine the extent of ozone depletion due to precipitation of energetic protons during solar proton events (SPEs), which is mostly restricted to polar cap regions. The significance of a second source of energetic particles during these events, namely energetic electron precipitation (EEP) originating from the outer radiation belt, has however not been established yet. A key element regarding the necessary investigations is to obtain an estimate on the incoming electron flux from the MEPED detectors on board the NOAA POES satellites. Previous efforts have faced various obstacles including detector degradation and cross contamination. In order to correct the measured electron flux for falsely measured protons, the contaminating proton flux is determined and subtracted. During this process, statistical uncertainties arise making it necessary to test the remaining electron flux for significance.

The three objectives of this study are motivated by the community's lack of success in disentangling the EEP contribution, in particular during SPEs. Firstly, a novel criterion for the statistical significance of corrected electron fluxes is presented and implemented in order to improve the POES MEPED losscone dataset. Improvements resulting from the new criterion and general limits set by proton contamination and other technical issues are discussed. By applying a newly developed analysis technique, the losscone fluxes are estimated. Secondly, the improved losscone dataset is compared to the Ap-based EEP flux parametrization used in CMIP6. The modelled fluxes are tested with respect to their ability to reproduce general flux levels as well as temporal and spatial variability. The third aim of this work is to investigate the applicability of corrected electron fluxes measured by the MEPED instrument during periods of high contamination. Therefore, six SPEs are investigated with the aim to establish the degree of data rejection and data coverage throughout the events. Even after introducing a less strict contamination criterion, data rejection is high in regions of large proton fluxes. Low latitudes and MLTs corresponding to the post-midnight and morning sector, however, show promising data coverage and are used to determine whether SPE electron fluxes behave identically to regular storm fluxes defined by similar geomagnetic indices.

Acknowledgements

First and foremost, my gratitude goes to my supervisor and team leader of G3, Hilde Nesse Tyssøy, whose door stood always open for advice and guidance. I am extremely thankful for all the opportunity to experience the SOLARIS-HEPPA workshop in Paris and present parts of the results from my thesis project. She managed to find the perfect balance between giving me freedom during data analysis while nudging me into the right direction from time to time.

I want to thank my co-supervisor, Marit Sandanger, who provided great help with all Matlab-related issues and made sure I got all NOAA, POES and MEPED details right during the Methods write-up. Contributions by Professor Emeritus Johan Stadsnes regarding statistical derivations and the Results chapter as well as quite a number of discussions on flux corrections are deeply appreciated and were vital for this project's success.

Collective thanks are directed to all other G3 group members and Birkeland Centre members for providing a working atmosphere that lets everyone feel included and contributing. The biannual Birkeland meetings were always an event I was looking forward to, presenting my own results, receiving valuable feedback and listening to other members' achievements.

I thank the German Academic Exchange Service, DAAD, for providing financial support throughout my whole Master's degree and believing that my aspirations were worth the cost.

Last but not least, it is time to thank my family and friends, but especially my parents. From the very beginning, you got me interested in science and puzzling with explanations for unresolved problems. I will probably never forget what an Alto Cumulus Lenticularis is and continue becoming glad whenever I get to see one.
Since this is not your exact field of study (a bit too high up, I guess), I wrote an extra detailed chapter on the theoretical background. I hope that makes you feel sufficiently obligated to read the thesis.
I am very grateful for all the people I met during my studies and old friendships that survived the distance. Special credits go to Håkon for both his culinary and mental support and providing me with a home in this rainy and foreign country.

Contents

Chapter 1

Introduction

Whereas Northern lights presented a basis for countless spiritual theories and tales for many centuries, their scientific relevance was first noted in the end of the 18th century. More and more scientists committed to the new topic and research around the mysterious phenomenon blossomed. A large variety of theories regarding its origin and mechanisms were suggested, some of which were later confirmed by experiments while others were found out to be ill-founded in retrospect. The ample range of new theories and degree of disagreement between researchers at the time is nicely portrayed in a review booklet published in 1903 by German meteorologist Dr. A. Nippoldt. While the scientific community reached unity on Northern lights being glowing gases in the atmosphere that are excited by electric currents, the very origin of this electricity was still heavily debated. One explanation mentioned in the booklet was suggested by Norwegian physicist Kristian Birkeland who claimed electrically charged particles from the Sun caused the excitation of atmospheric gases leading to observable Northern lights.

FIGURE 1.1: Photograph of Northern lights. Courtesy: Jouni Jussila.

He later reinforced his statement by constructing his famous Terrella experiment, which consisted of a metallic sphere with a dipolar magnetic field, representing Earth, and an impinging beam of electrons which were redirected in the magnetic field forming glowing circles close to the surface around the magnetic poles. These discoveries, which were considered highly controversial in the beginning of the 19th century, are forming the very basis of today's knowledge in space science. Surprisingly enough, the little booklet from 1903 contained many theories and suggestions known to be justified by later research. The dependence of Northern light intensities on the 11-year solar cycle was resolved already in the beginning of Northern lights research and perturbations in the geomagnetic field were measured and associated with Northern lights occurrences. The book even reports on studies targeting the height of gas excitement in the atmosphere during a particularly strong auroral event on the 9th of September in 1898. Via parallax measurements, the heights of Northern lights observations throughout the event were found to range from 30 to 200 km, numbers which seem realistic based on modern measurement techniques (yielding > 90 km for visible aurora).

Years and years of research contributed to the vast resource of knowledge available today. In the year 1951, Biermann proposed the existence of a constant flow of plasma originating from the Sun, called the solar wind. The contained particles travel through space while the so-called Interplanetary Magnetic Field (IMF) is frozen into them. At a distance of one Astronomical Unit (AU), the solar wind and IMF hit Earth's magnetosphere. The mechanisms via which the IMF and geomagnetic field couple, allowing charged particles to fill Earth's magnetosphere, were introduced by Dungey in 1961 and details on the reconnection between the two magnetic systems are still subject to extensive research. The interaction between IMF and geomagnetic field leads to the transfer of energy from the ambient solar wind into the Earth magnetic field, resulting in the acceleration of particles. During periods of high geomagnetic activity, vast numbers of energetic particles penetrate into the atmosphere and deposit their energy by ionizing molecules and forming Northern lights. During very strong events, protons are known to penetrate as low as 30 km while electrons rarely reach altitudes below 50 km. As the booklet rightly claims, the found altitudes suggest meteorological impacts of the precipitating particles on the atmosphere. Due to limited knowledge and technological access, however, this idea was soon to be declared dead. Based on today's understanding of the atmospheric system, a measurable particle effect in the atmosphere does not seem that unlikely any more. Depending on the particles' energy deposition depth, different ionized molecules are produced. In an altitude corresponding to the middle atmosphere, locally increased $\mathrm{HO_x}$ and $\mathrm{NO_x}$ production rates can be the consequence of particle precipitation. The comparatively long $\mathrm{NO_x}$ life-span during high-latitude winter darkness and downward transport can lead to the depletion of stratospheric ozone. Ozone, however, is known to play a major role in stabilizing our planet's radiation balance and climate system by absorbing incoming solar short-wave and emitting long-wave infrared radiation. Changes in ozone concentration will eventually lead to local changes in temperature and stratosphere dynamics which may map onto surface climate.

The strength of this effect obviously depends on geomagnetic activity and the amount of particles reaching the atmosphere. Predominantly during solar maximum, spikes of high energetic solar protons are likely to reach stratospheric altitudes altering ozone concentrations both directly and indirectly, while the more continuous electron precipitation is restricted to higher altitudes and impacts ozone via secondary effects. Geomagnetic indices have long been used as proxies for geomagnetic activity but lack necessary spectral and spatial resolution. A needed remedy was provided by new technologies in the 1970s enabling in-situ measurements involving satellites. Detectors were mounted on spacecraft in order to measure electron and proton fluxes directly including their separation into different energy channels. While this presents a powerful tool when assessing particle precipitation, new obstacles continue to appear. In the end, electron fluxes during so-called solar proton events (SPE) are claimed to be highly unreliable due to possible proton contamination of the electron measurements. Whether nevertheless some information can be drawn from satellite data during these events and how truthful a predictive model solely based on the geomagnetic Ap index can be considered, are questions investigated in this work. Due to the particles' different origins, detector responses and atmospheric impacts, the background to this work is quite interdisciplinary, involving detector physics, space science and meteorology.

The objective of this work is to develop a statistically based criterion in order to judge the statistical significance of corrected electron fluxes. The introduction of a new proton contamination criterion reduces data rejection and thus improves the current electron losscone dataset. Subsequently, the improved dataset of the Northern Hemisphere (NH) will be used for inter-comparison with an electron flux parametrization which is solely based on one geomagnetic index (Ap). The predictive power of the Ap index with respect to electron fluxes will be judged by analysing general modelled flux levels as well as temporal and spatial flux variability. NH data rejection rates in different locations and time periods will give information on the applicability of the improved losscone dataset during SPEs and associated high contaminating proton fluxes. Data which is considered reliable can then be used in order to investigate whether the Ap index indeed evokes similar electron flux responses during SPE-related storms compared to regular geomagnetic disturbances as assumed by the Ap model.

Chapter 2

Theoretical Background

In order to understand how the Sun not only influences Earth's climate system via emitted radiation but also by constant ejection of plasma particles, it is important to look at the coupling between Sun and Earth magnetic fields. A short outline of the basic processes will illustrate the "bigger picture" and give a hint on how different plasma systems interact. In following sections, more details on the involved components, such as the Sun, solar wind and Earth, will be given.

Starting from information on the Sun and solar processes leading to plasma ejections, solar wind characteristics will be explained. A short description of near Earth space including the structure of the geomagnetic field will yield basic knowledge necessary for understanding how the solar wind couples to Earth's magnetic field. Terms like the Interplanetary Magnetic Field (IMF), the Dungey cycle and losscone fluxes will be defined. Subsequently, different regions of the Earth magnetosphere (e.g. radiation belts and plasmasphere) and the characteristics of particles they contain will be introduced leading to the topic of particle precipitation and energy deposition into the atmosphere. A description of different types of geomagnetic storms follows and basic geomagnetic indices are introduced as means of assigning quantitative values to geomagnetic activity. The final section covers the general atmospheric structure and describes the different mechanisms allowing solar variability to impact Earth climate.

2.1 The Sun

As the main source of energy, the Sun influences Earth's climate and nature, and thus the life and existence of every one of us. The largest part of the incoming energy is transported by radiation and originates from the approximated black-body spectrum of the Sun. A second part arrives in the form of a magnetized plasma, the solar wind. To which degree do the solar wind and incoming particles affect our climate and technological systems? In order to find answers to this question, the Sun presents a central object of interest. Understanding the mechanisms of ejection and acceleration turns out to be essential for any attempts of classifying and forecasting solar wind behaviour.

2.1.1 Structure and properties

Primarily, the Sun can be seen as a hot ball of plasma consisting mostly of hydrogen (~ 73 %) and helium (~ 25 %) as well as small abundances of heavier elements (Basu and Antia, 2008). Its radius, $R_{\odot}$, of approximately 700000 km exceeds the Earth radius by a factor of 100 and consequently its volume is roughly one million times the volume of the Earth. While the Sun is rotating, the plasma it consists of does not behave like a rigid body comparable to the Earth or its moon. This leads to a latitudinal dependency of the rotation velocity, so that plasma close to the solar equator rotates faster compared to polar regions (Newton and Nunn, 1951). This differential rotation gives rise to a twisting of magnetic field lines and results in spatial and temporal variations of magnetic flux and field topologies. Accordingly, the so-called solar activity oscillates between maxima and minima in an 11-year cycle. During the minimum of solar activity,

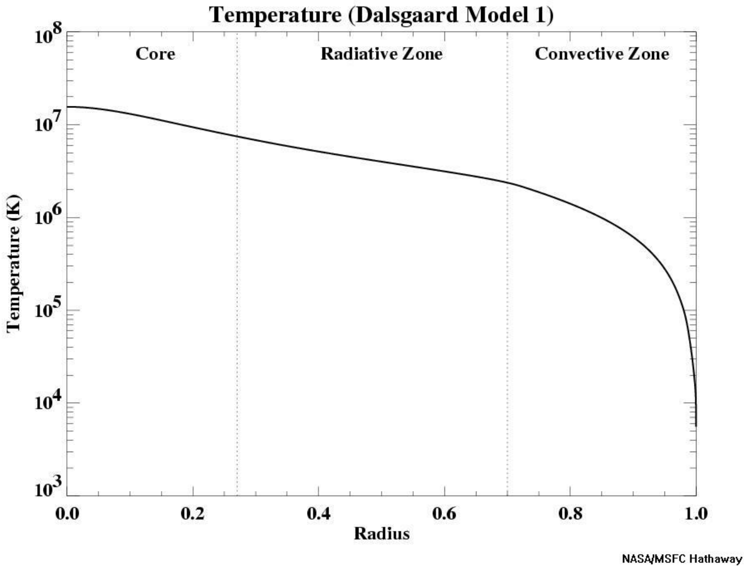


FIGURE 2.1: The temperature profile of the Sun within one solar radius. The different layers and their positions are indicated as well. From Christensen-Dalsgaard et al. (1996).

the Sun's magnetic field can be assumed to resemble a dipole configuration to a good degree. Increasing distortion of this initial field strengthens an overlaying quadrupole field and peaks at the maximum of solar activity (Banaszkiewicz et al., 1998). A subsequent period of relaxation back to a dipolar field and associated minimum in solar activity completes the cycle. It is worth noting, that the 11-year cycle describes the oscillation of the solar activity. When looking at the magnetic field configurations, a complete cycle comprises 22 years, as maxima occur with alternating polarities (positive/negative). In addition, the maximum solar activity of the 11-year solar cycle is subject to long-term variations with a periodicity of approximately 90 years, the so-called Gleissberg cycle (Gleissberg, 1939).

When it comes to the inner composition of the Sun and the structure of its atmosphere, researchers have faced many challenges. Without the possibility of in-situ measurements, models solving plasma equations of state could be combined with radiation measurements to determine three major layers within the Sun. Located at the very centre and extending to approximately 0.25 solar radii, $R_\odot$, is the core where the Sun's energy production takes place (Di Mauro et al., 2002). Nuclear fusion of highly abundant hydrogen to helium and further to carbon releases energy due to the equivalence of mass and energy (Russell, 2001). Since these processes are related to the heating of the core, heat has to be transported away towards the surface.

Between the core and surface, two further layers lie embedded like shells around the centre of the Sun. The innermost layer extends from 0.25 to about 0.75 $R_\odot$ (Basu and Antia, 2004) and consists of hot plasma which allows for the transport of heat via radiation. Accordingly, this region is called radiative zone. As the temperature decreases with distance from the core, the plasma becomes more opaque leading to decreasing efficiency of radiative transport. Below a temperature of roughly 3 million K (Figure 2.1), convective transport can be considered the dominating process, which marks the second layer called convection zone. This region is characterized by transport processes where hot plasma bubbles rise towards the cooler surface and colder plasma sinks down in return.

The visible surface of the Sun is called photosphere and its temperature of roughly 5700 K (Christensen-Dalsgaard et al., 1996) corresponds to the black-body radiation spectrum observed on Earth. With a thickness of only 100 km, the photosphere also represents the lowest layer of the solar atmosphere. The following atmospheric layer, the chromosphere, is characterized by a rise of temperature to approximately 20000 K leading to emission of red light by hot ions. The transition region acts as the junction between chromosphere and solar corona. With a temperature of well above 1 million K, the corona is extremely hot and the transition region is mainly characterized by a very steep temperature gradient (Hansteen, 1992).

2.1.2 Solar wind, coronal holes and sunspots

The existence of a continuous stream of charged particles from the Sun was first postulated by Biermann (1951) who linked the presence of a solar wind to the acceleration of molecules in comet tails. A later-on observed pattern consisting of slow and fast solar wind streams shows a 27-day periodicity indicating fixed and co-rotating source regions (Neugebauer and Snyder, 1966) and results from the Skylab mission in 1973 established so-called coronal holes (CH) as the sources of fast solar wind streams (Nolte et al., 1976). The light emitted by the corona usually falls within the EUV and X-Ray range. Coronal holes, however, appear dark in those frequencies as depicted in Figure 2.2, which shows a soft X-ray image of the Sun. Most coronal holes are located in high latitude regions where the coronal magnetic field lines are open and have only one footpoint in the photosphere (Figure 2.3a) but occasional low-latitudinal extensions occur predominantly in the declining phase of the solar cycle (Tsurutani et al., 2006). From coronal holes, high-speed solar wind can stream directly out into space carrying the magnetic field with it. Regarding slow solar wind, the low-latitudinal streamer belt could be established as source region (Gosling, 1997). Its position with respect to the ecliptic plane is shown in Figure 2.3b. Due to the Sun's rotation, the interplanetary magnetic field (IMF) embedded in the solar wind is formed like an Archimedian spiral, called Parker spiral (Parker, 1958; Thomas and Smith, 1980).

Also images of the photosphere can exhibit dark spots, so-called sunspots, where magnetic field lines bulge under increasing pressure during solar maxima. Due to the suppression of convection, sunspot plasma is colder in comparison to the surrounding photospheric plasma which makes sunspot regions look darker. Their polarity depends on which hemisphere they are located in and the assignment of polarities to hemispheres changes with each solar cycle (Hale, 1924). Sunspots do often occur in pairs in which case they show opposite magnetic polarity. The number of sunspots is indicative of the solar activity and was initially used to study the 11-year solar cycle and the occurrence of magnetospheric storms (Cortie, 1912). Groups of sunspots are called solar active

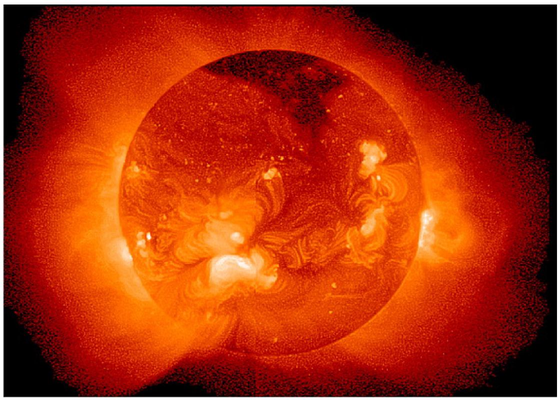
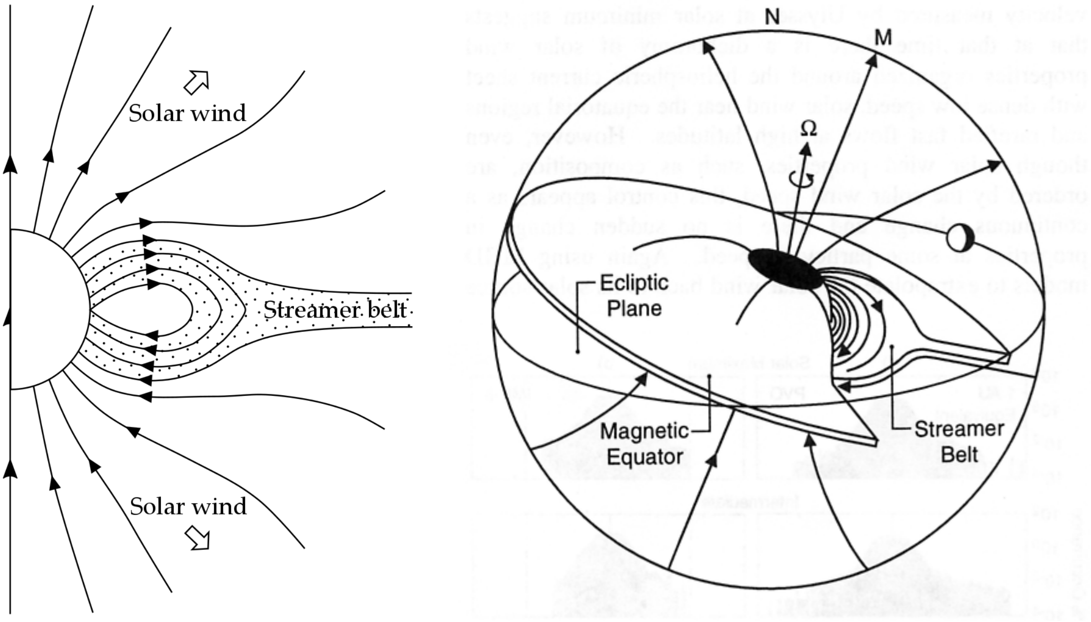


Figure 2.2: Soft X-Ray image of the Sun from Tsurutani et al. (2006). The dark regions at high latitudes are coronal holes.



(a) (b)

FIGURE 2.3: (a) Sketch of the solar coronal magnetic field with open field lines in high latitudes and closed field lines in low latitudes. Modified from Low and Hundhausen (1995). (b) Schematic showing the position of the IMF and streamer belt with regard to the ecliptic plane. From Russell (2001).

regions to which the occurrence of many secondary phenomena is linked.

In addition to the continuous solar wind originating from the corona, singular plasma ejections can occur when the magnetic pressure of sunspot field lines reaches a breaking point and magnetic field lines reconnect. The gain of electromagnetic energy is converted to kinetic energy by the radical acceleration of particles producing electromagnetic radiation as a side-effect (“Accelerated Particle Composition and Energetics and Ambient Abundances from Gamma-Ray Spectroscopy of the 1991 June 4 Solar Flare”). This burst of plasma through the chromosphere and corona is called solar flare. The accelerated particles continue propagating through space at high speeds and can reach Earth after approximately two days. Often but not always, solar flares are accompanied by Coronal Mass Ejections (CMEs) (Zhang et al., 2007), where plasma bursts out into space directly from the solar corona in a vast explosion. The occurrence rate of solar flares and CMEs peaks at solar maximum and follows in general the solar cycle, since their appearance is linked to sunspots and solar active regions.

2.2 Near Earth space

After having explained the Sun's structure and mechanisms responsible for the ejection of plasma in the previous section, this section's focus will lie on describing near Earth space, i.e. the geomagnetic field, magnetospheric currents and features as well as the mechanism leading to energy transfer from the solar wind to the magnetosphere. One also has to examine particle acceleration, trajectories and losscone distributions in order to account for the energy deposition in the upper atmosphere. Particle precipitation can originate from different sources and vary in latitude and energy spectra.

2.2.1 Bow shock, magnetosheath and magnetopause

The flow of solar wind is both supersonic and superalfvénic, which means that in case of a solid obstacle being located in its path neither acoustic density waves nor magnetic Alfvén waves could travel fast enough to deflect the solar wind flow around the obstacle (Baumjohann and Treumann, 1997, p. 164). As it is, Earth presents such an object and in order to redirect the solar wind flow around the planet, the bow shock forms in front of it (Baumjohann and Treumann, 1997, p. 181). Where in general, plasma exhibits a fluid-like behaviour which is described by magnetohydrodynamical (MHD) equations, its properties jump drastically in the vicinity of the shock. Thus, the solar wind becomes slower, denser and bent in its trajectory after passing the bow shock. The layer of compressed solar wind behind the bow shock is called magnetosheath and the border between the geomagnetic field and the IMF in the solar wind is the magnetopause. Figure 2.4 depicts a schematic drawing of the solar wind flow from the Sun as it hits the bow

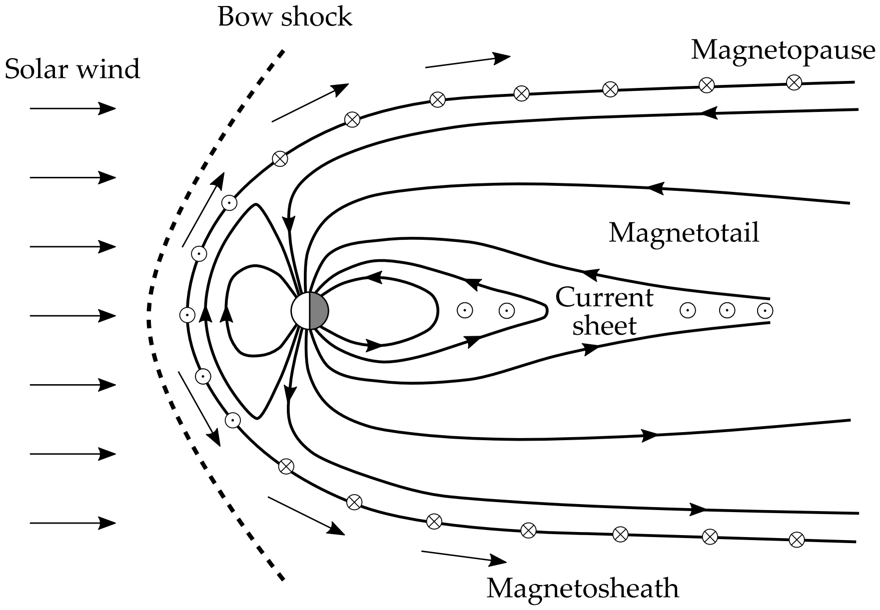


Figure 2.4: Schematic drawing of the bow shock, magnetopause and solar wind flow. Modified from Kivelson and Russell (1995, p. 229).

shock and the altered flow after passing it. The plasma flows around the magnetopause towards the night-side. The exact location and shape of bow shock and magnetopause are highly variable and depend first and foremost on properties of the impinging solar wind. The bow shock will adapt to the solar wind dynamic pressure, IMF and Mach numbers as necessary in order to adequately redirect the flow of plasma (Peredo et al., 1995). At the magnetopause the flow is tangential so that the dynamic pressure vanishes compared to the thermal pressure of the compressed plasma. The magnetopause location and shape follow therefore from the force balance between magnetic force and thermal pressure force. As a consequence, a current sheet on the magnetopause develops (Figure 2.4) resulting in an oppositely directed J × B force (Bothmer and Daglis, 2007, p. 113). A different and simplified way of explaining the magnetopause current is to look at what happens when electrons or protons penetrate the magnetopause due to their thermal movements. In such a case, they will suddenly be exposed to the Lorentz force of the geomagnetic field and reflected from the magnetopause. At the same time, electrons will be deflected towards dawn and protons towards dusk giving rise to a net current as shown in Figure 2.4. During periods with average solar wind conditions, the magnetopause and bow shock locations are given by approximately 10 and 13 Earth radii distance from the day-side surface, respectively (Gombosi, 1998, p. 283).

A sudden increase in large-scale solar wind dynamic pressure at the bow shock is often related to an arriving interplanetary shock front or high speed solar wind stream (HSSWS). The magnetopause will be pushed inwards resulting in a sudden increase of the magnetopause current (Gombosi, 1998, p. 308). In case of a following geomagnetic storm, this initial phase feature is called storm sudden commencement (SSC), otherwise one speaks of a sudden impulse (SI).

2.2.2 The geomagnetic field and Dungey cycle

It is basically the coupling of the IMF to the geomagnetic field that enables energy transfer from the ambient solar wind to magnetospheric particles. The geomagnetic field close to the surface can be approximated by a dipole field generated by a dynamo

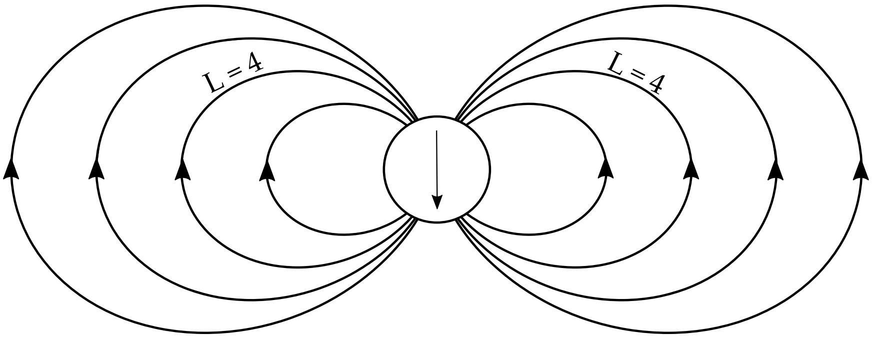


FIGURE 2.5: Plot of a magnetic dipole field. The magnetic field lines associated with the L = 4 shell are indicated as well.

inside the Earth core (Figure 2.5) (Merrill and McElhinny, 1983). It is worth noting that the corresponding dipole axis is shifted by 11° with regard to the Earth rotational axis. Thus, the geomagnetic North pole does not coincide with the geographic North pole and one differentiates between two different coordinate systems. One common method of visualizing the polar caps and auroral zones are azimuthal projection plots as shown in Figure 2.6. The plots are centred at the respective North poles and project the three-dimensional Earth surface on a flat circle. It can also be instructive to introduce local times (LTs) for geographic and magnetic local times (MLTs) for geomagnetic coordinates (Figure 2.6). The local time noon (12) always faces the Sun while local time midnight (00) lies on the night-side. The landmasses then rotate underneath the local time coordinate systems around the geographic axis.

A more correct model compared to the dipole approximation is given by a multipole expansion with a dominating dipole term and parameters fitted to the actual measured field. This approach reproduces the lower geomagnetic field with an accuracy of 0.05 % (Kallenrode, 2004). From the dipole approximation however, it is possible to define so-called L shells. The L value states the distance of a magnetic field line from the surface in the equatorial plane, $\mathrm{r_{eq}}$, with respect to the Earth radius, $\mathrm{R_E}$

$$L = \frac{r_{eq}}{R_E}\,. \tag{2.1}$$

One L shell therefore includes all magnetic field lines that have been assigned a certain L value and thus form a shell around Earth (McIlwain, 1961). It is also possible to calculate in which geomagnetic latitude, Φ_E, a given L shell intersects with the surface

$$\cos \Phi_E = \frac{1}{\sqrt{L}}\,. \tag{2.2}$$

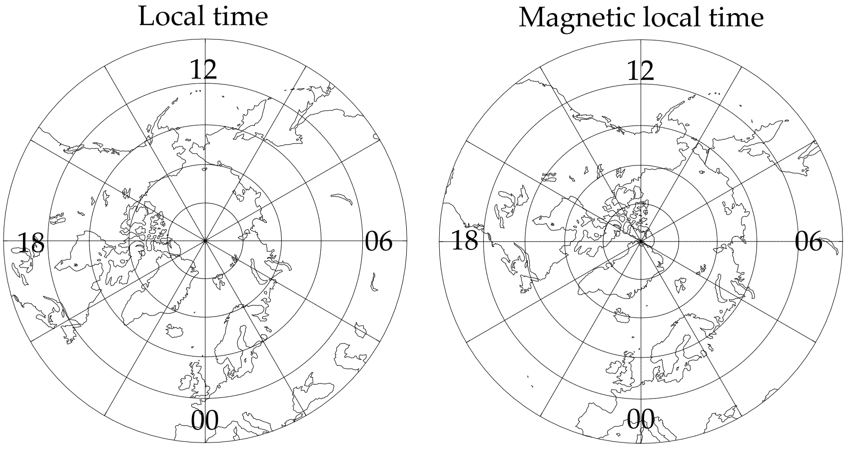


FIGURE 2.6: Azimuthal projection plots of the geographic (left) and geomagnetic (right) north pole including visualization of different local times.

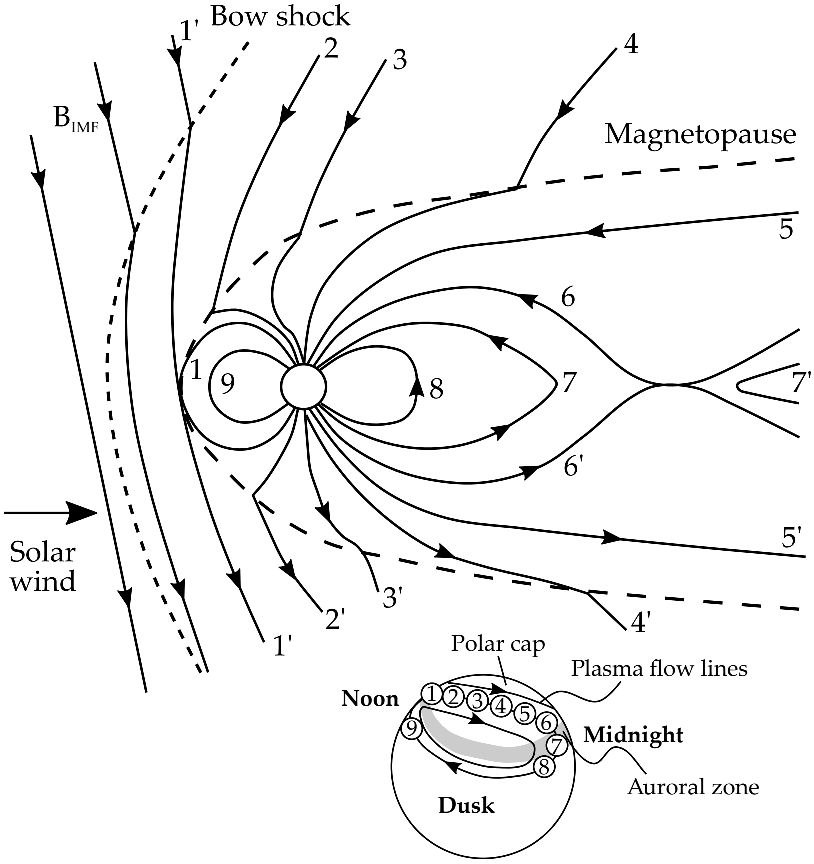


FIGURE 2.7: Not in scale drawing of the Dungey cycle and geomagnetic field. The lower drawing shows the movement of magnetic footlines throughout one pass of the Dungey cycle. Modified from Kivelson and Russell (1995, p.243).

Obviously, shells with a lower L value are located closer to Earth and intersect at lower latitudes, whereas shells with increasing L can be assigned to higher geomagnetic latitudes as shown in Figure 2.5. Deviations from the simple dipole-based models begin to arise at distances of few R_E from the Earth centre due to ionospheric and magnetospheric currents and the influence of the ambient solar wind (Kivelson and Russell, 1995, p. 289). The dynamic pressure of the solar wind compresses the geomagnetic field on the day-side and elongates it on the night-side, forming the magnetotail. A schematic drawing of the resulting field is shown in Figures 2.4 and 2.7. Please note, however, that neither depicts the geomagnetic field in scale for reasons of presentability. For example, the point in the magnetotail in Figure 2.7, where the field lines denoted by 6 and $6'$ intersect, may lie as far as 100 R_E away from Earth (Slavin et al., 1985). Asymmetries of the geomagnetic field may arise due to the tilt of the dipole axis with respect to the ecliptic plane and depend on the angle in which the solar wind hits the bow shock. In addition, the approximated dipole axis is not centred in Earth but slightly shifted, leading to an area of smaller magnetic flux density in the south Atlantic called South Atlantic Anomaly (SAA) (Heirtzler, 2002).

The question as to which mechanisms are responsible for driving geomagnetic dynamics has been subject to heavy debates for a long time. The commonly accepted picture nowadays is the one introduced by Dungey in 1961 (Dungey, 1961), which states that energy transfer from the solar wind is enabled by magnetic reconnection on the day-side magnetopause and in the magnetotail. When the magnetic pressure between two oppositely directed field lines overcomes the thermal pressure of the plasma, the field lines will reconnect. In the case of day-side reconnection at the magnetopause, this means that a previously closed geomagnetic field line will open up and reconnect to an interplanetary field line, forming a so-called open field line (points 1 and $1'$ in Figure 2.7). The IMF does not necessarily have to be southward directed in order to reconnect to the geomagnetic field, because reconnection for a northward IMF is still possible in higher latitudes. However, it is important to note, that the reconnection rate and efficiency of energy transfer peak for a strongly southward directed IMF. The new open field lines exhibit a strong curvature in the points where they reconnected which leads to the acceleration of local plasma and transport of the field line towards higher latitudes (points 2 and $2'$). In general, the magnetic energy of a field line can be assumed proportional to its length, a contraction thus referring to a transfer of magnetic energy to kinetic energy and associated acceleration of plasma. Outside of the magnetopause, the magnetic field lines are still frozen into the ambient solar wind and transported towards the night-side corresponding to points 3 and $3'$ as well as 4 and $4'$, respectively. The footpoints of open field lines lie in the ionosphere where conductivity is high and field line mobility rather low due to friction with the local ions. The open end of the field line, however, is dragged further towards the night-side by the movement of the ambient solar wind bending the field line (visible in points 3 and 4). The resulting kink is transported downwards by Alfvén waves until the magnetic stress of the curvature exceeds the frictional force in the ionosphere and the footpoint moves towards the night-side. After having passed the geomagnetic poles, open field lines are stretched extensively (points 5 and $5'$) with the associated necessary magnetic energy originating from the ambient solar wind kinetic energy. The more open field lines are stretched on the night-side, the closer oppositely directed lines are located in the magnetotail. On the one hand, this results in a magnetotail current sheet which is oriented from dusk to dawn as depicted in Figure 2.4. On the other hand, magnetic flux accumulates in the magnetotail until its density leads to magnetic reconnection and the conversion back to closed field lines (points 6 and $6'$). Again, the arising curvature at the reconnection point leads to particle acceleration and dipolarization of the magnetic field topology (points 7 and $7'$, 8 and $8'$). This corresponds to the respective closed field line moving closer to Earth associated with a footline movement towards lower latitudes. From there, closed field lines migrate back to the day-side where there is lower magnetic pressure due to prior reconnection (points 9 and $9'$). The footline movement associated with a magnetic field line passing once through this so-called Dungey cycle is depicted in the small drawing of Figure 2.7. Open field lines move across the polar cap from noon to midnight where they are reconnected. The formed closed field lines move back towards noon at lower latitudes below the auroral oval, which is indicated by the shaded area. As a net effect, kinetic energy is transferred from the ambient solar wind into the geomagnetic field and magnetosphere which is why the Dungey cycle acts as a magnetic dynamo.

2.2.3 The magnetosphere

The previous section described the mechanisms that couple the solar wind to the magnetosphere and allow for energy transfer and particle injections. In this section, we will take a closer look on the magnetosphere itself, i.e. how particle motion leads to the formation of current systems along with the internal magnetospheric structure and local particle populations.

Particle trajectories and magnetospheric currents

When charged particles move through electric and magnetic fields they can experience a variety of forces depending on the exact field configurations, e.g. the Lorentz force (Baumjohann and Treumann, 1997, Chapter 2). These forces can alter the direction of the particles' movements and in some cases their kinetic energy. Flows of charged particles can in turn lead to charge separation which creates electric fields, currents and magnetic fields. Therefore, particle motion and fields are strongly coupled, each influencing the other. A set of equations describing the behaviour of a plasma in the presence of electromagnetic fields is given by MHD. These equations are, apart from some exceptions, valid in the upper magnetosphere and can be used in order to investigate particle trajectories, currents and magnetic field perturbations. In the MHD

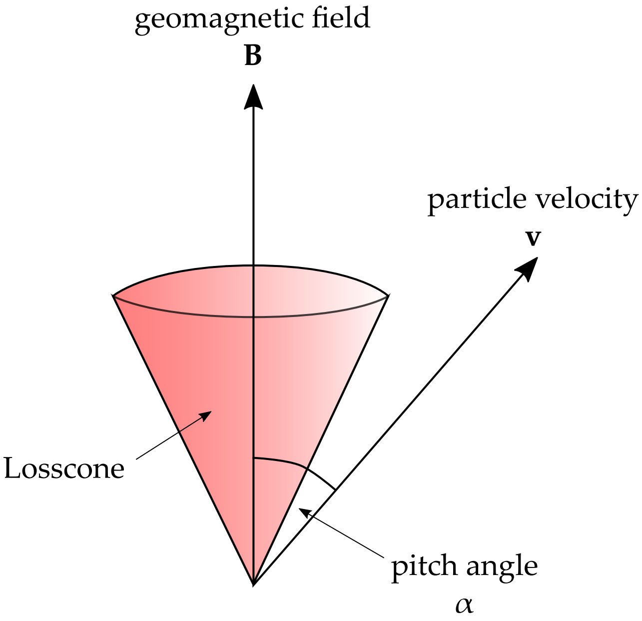


FIGURE 2.8: Schematic drawing indicating the losscone with respect to the background geomagnetic field and the pitch angle of a particle with velocity $\mathbf{v}$.

momentum equation, all electromagnetic forces are summarized in the J × B - force and the Maxwell stress tensor. It is, however, instructive to examine single particle motion in simplified settings to gain a good understanding of the basic processes happening when former solar wind particles move through the geomagnetic field.

A constant magnetic field, for example, will lead to a gyration motion of charged particles around the magnetic field lines, where positively and negatively charged particles gyrate in opposite directions. The resulting current resembles a magnetic dipole and an associated diamagnetic behaviour of this particle population is in fact measurable at the surface (see ring current). In addition, currents can be generated by other configurations, i.e. a gradient or curvature of the magnetic field. In terms of particle motion in a dipole field, one can differentiate between three categories of movements, each associated to different time scales and adiabatic invariants: the gyration of charged particles around magnetic field lines, a bounce motion between the two hemispheres due to the converging magnetic field and an azimuthal drift around the Earth (Baumjohann and Treumann, 1997, p. 22). Taking a closer look on the bounce motion, the particles move on a helix-like trajectory from one hemisphere to the other. The altitude at which particles are mirrored is differing depending on their energy and equatorial pitch angle. The equatorial pitch angle defines the angle between the particle velocity vector and the magnetic field in the equatorial plane as indicated in Figure 2.8. The smaller the equatorial pitch angle of a particle is, the deeper it can penetrate before it is reflected.

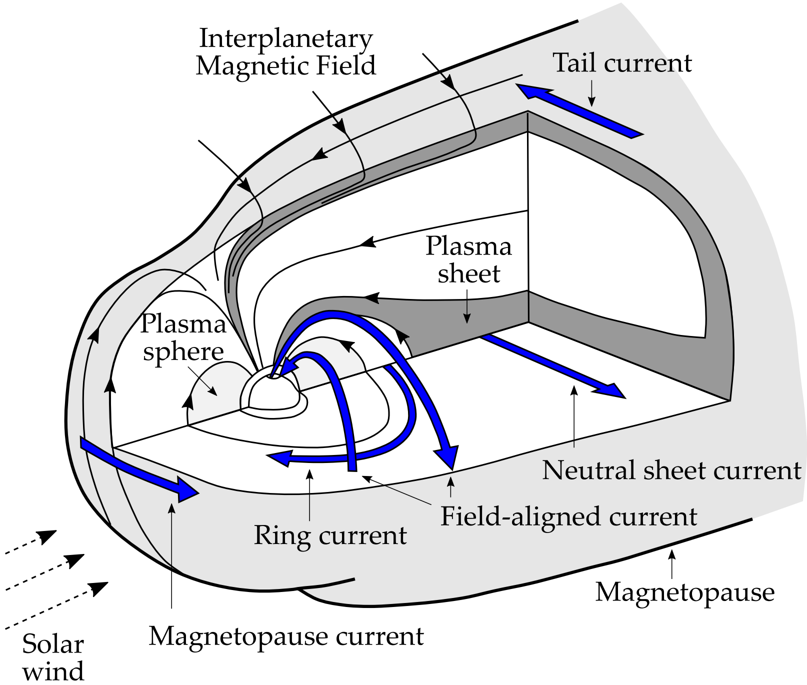


FIGURE 2.9: Schematic drawing of the magnetospheric current system showing the magnetopause current, ring current, tail current and neutral sheet current with coloured arrows. The auroral electrojet connects the field aligned currents in the ionosphere. Modified from Russell (1993).

Once a particle reaches the upper atmosphere, it can collide with neutrals, deposit its energy and be lost. The so-called losscone contains all pitch angles for which particles can be lost in the atmosphere. According to the conservation of magnetic momentum, the losscone is smaller in the equatorial plane and wider towards the poles. Following the logical conclusion, the losscone should be empty at all times because all contained particles will be lost in the atmosphere. On the contrary, the losscone is constantly refilled by pitch angle diffusion and interaction between particles and background plasma waves that can severely alter pitch angle distributions.

Figure 2.9 illustrates the major magnetospheric currents and their positions in near Earth space (Russell, 1993). The magnetopause current and neutral sheet current result from the impinging solar wind and long stretched magnetic field lines in the magnetotail, respectively, as described in the previous section. The tail current results from a process similarly to the magnetopause current as depicted in Figure 2.4 and closes the neutral sheet current. The ring current consists mainly of trapped electrons and ions that have an azimuthal drift while bouncing from one hemisphere to the other. The drift results in first order from the curvature and gradient in the magnetic field, which evokes opposite movements for electrons (eastward) and positive ions (westward) (Baumjohann and Treumann, 1997, p. 43). When calculating the magnetic field generated by the ring current, one also has to take into account the diamagnetic contribution from the particles' gyration. The existence of field-aligned currents was first postulated by Kristian Birkeland in 1908 (Birkeland, 1908). In general, they form a link between ionospheric and magnetospheric currents and arise partially from the junction between open and closed magnetic field lines. The strongest ionospheric currents are the eastward equatorial electrojet on the day-side and the auroral electrojets located further towards the night-side. The latter result from the convection of plasma as the footpoints of magnetic field lines move in the ionosphere. Electrons are generally more mobile and follow the moving field lines more easily, which leads to charge separation and an associated current opposite to field line movement (Hall current). At mid-latitudes, where closed field lines move towards the day-side, this leads to a westward current at dusk and eastward current around dawn. During strong magnetospheric storms and substorms, the field aligned currents can couple the ionosphere quite effectively to the neutral current sheet which strongly enhances the eastward electrojet. Magnetograms in the respective latitudes can be used in order to detect the transition from westward to eastward electrojet and its severity.

Magnetospheric structure

The magnetosphere consists of a number of different plasma regions as depicted in Figures 2.10 and 2.11. Located in the innermost part is the plasmasphere, a torus-shaped region around Earth which contains cold high-density plasma. The magnetic field in the plasmasphere consists of closed, dipole-like field lines. The existence of the plasmasphere was first discovered in 1963 by Don Carpenter who was studying Whistler waves and the associated electron density in the magnetosphere (Carpenter, 1963). A drop in the plasma density by roughly one order of magnitude marks the edge of the plasmasphere, the so-called plasmapause. Chappell et al. (1970) found that the plasmapause location is strongly dependent on geomagnetic activity, i.e. it moves closer to Earth with increasing activity. In general, its position varies between 4 - 7 Earth radii, depending on local time and activity.

The plasmasheet is located around field lines with footpoints in higher geomagnetic latitudes corresponding to the auroral zone (Figure 2.10). It is populated by low energy particles with electron energies lying in the keV range (Kivelson and Russell, 1995, p. 291). A continuous drizzle of particles emanating from the plasmasheet precipitates into the atmosphere and creates visible aurora (indicated by green arrows in Figure 2.10). The intensity of this so-called auroral precipitation has been found to vary in terms of local time and geomagnetic activity (Thorne, 1980) and short-term pulsations have been registered as a quite common phenomenon (Royrvik and Davis, 1977).

A very central magnetospheric structure are the Van Allen radiation belts which consist of trapped electrons and protons. The theoretical existence of magnetically trapped particles was suggested by a variety of space scientists and could indeed be confirmed in

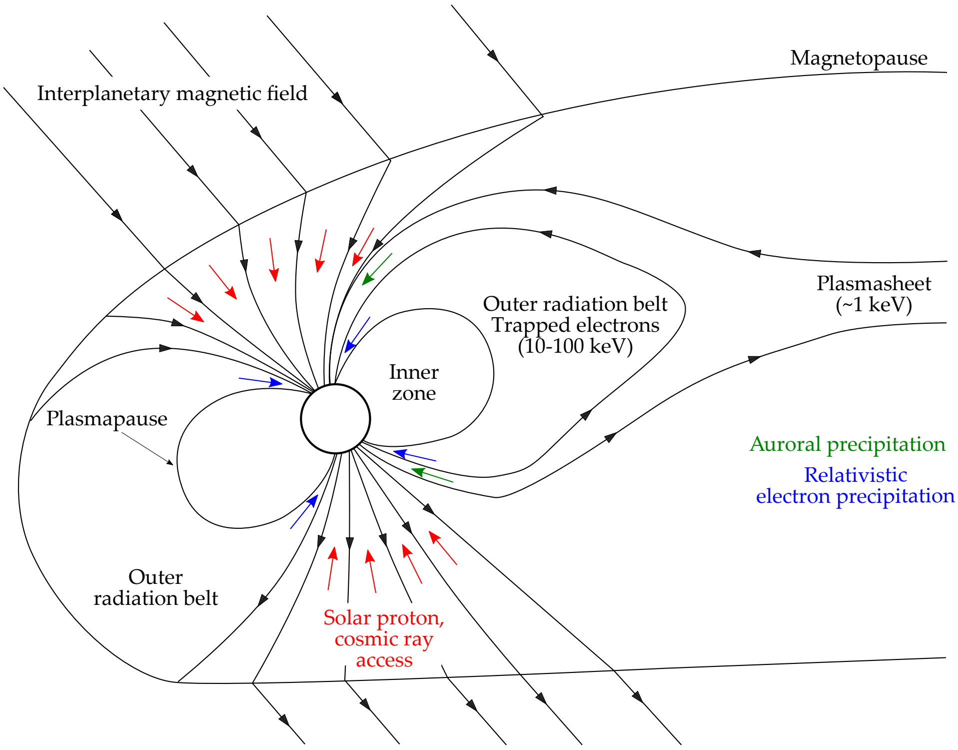


FIGURE 2.10: Schematic drawing depicting the different plasma regions within the magnetosphere. Green arrows indicate low energy auroral precipitation from the plasma sheet and blue arrows show relativistic electron precipitation originating from the outer radiation belt. General access regions for solar protons and cosmic rays are indicated by red arrows. Modified from Thorne (1980).

1958 by the Explorer 1 and Explorer 3 missions (Van Allen et al., 1958; Van Allen and Frank, 1959). The general radiation belt configuration as shown in Figure 2.11 consists of one inner belt, an empty slot region and one outer belt which supposedly splits up sporadically during strong solar events (Baker et al., 2013a). The high energy particle radiation contained in the belts is of great importance, not least since it presents a major source of potential radiation hazard for spacecraft and astronauts (e.g. Webb and Allen, 2004; Lopez et al., 2004). In addition, the outer radiation belt can be identified as one of the source regions for energetic particle precipitation (EPP) which influences atmospheric dynamics with impacts being transported all the way to the surface. In order to gain a better understanding of the inner magnetosphere, ongoing research deals with radiation belt dynamics, production and loss mechanisms and ambient particle populations.

The inner belt is located between 1.2 and 3 $\mathrm{R_E}$ in the inner zone of Figure 2.10 (Ganushkina et al., 2011) and reaches a minimum altitude of 250 km at the SAA. It contains mostly protons with an energy spectrum that has been determined to vary between 10 and several hundreds MeV (Freden and White, 1960). Albedo neutron decay is treated as the main source for energetic protons in the inner belt, including neutron production from both Galactic Cosmic Ray (GCR) and more transient solar protons in the upper atmosphere (Singer, 1958; Dragt et al., 1966). Incident GCRs interact with the Earth's atmosphere to create neutrons which are deflected towards higher altitudes, where they in turn decay into protons. Losses occur due to radial outward diffusion. The particle population can be considered rather stable and variations become visible out-of-phase with the 11-year solar cycle, i.e. decreasing population during solar maximum (Walt,

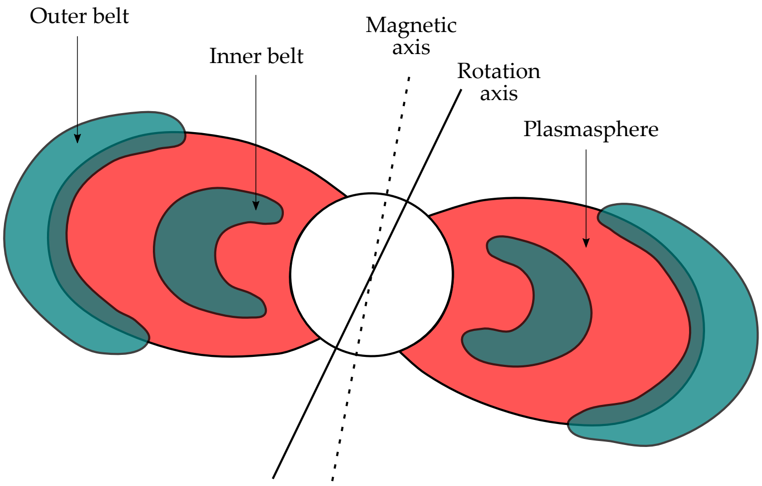


FIGURE 2.11: A cross section of the magnetosphere showing the Van Allen radiation belts in green and the plasmasphere in red. Based on an ESA illustration (`http://www.esa.int/spaceinimages/Images/2016/06/Illustration_showing_Van_Allen_radiation_belts`) and Rodger and Clilverd (2008).

1996). This is due to the better shielding of GCRs by the solar wind and upper atmosphere during high solar activity, leading to smaller neutron production rates as fewer GCRs reach the necessary altitudes. The outer radiation belt consists mainly of energetic and relativistic electrons with energies ranging between 200 keV and up to 20 MeV (Baker et al., 2013b). It is subject to large spatial and temporal variations as it is strongly influenced by geomagnetic activity and the impinging solar wind. Its location varies between 3 - 7 R_E (Ganushkina et al., 2011) and a correlation between its innermost edge and the plasmapause position has been found, supporting the alleged role of geomagnetic activity (Li et al., 2006). When evaluating potential electron sources, one has to take into account radial movement from the ionosphere (with particle energies $< 1\,\mathrm{eV}$) and the plasmasheet (energies of roughly 1 keV). Comparison with typical radiation belt energies shows, however, that neither of the two source regions is able to sufficiently explain the high energetic electrons trapped in the radiation belt. Possible acceleration mechanisms have been heavily debated during previous decades. Acceleration by radial diffusion (e.g. Falthammar, 1965) seemed to be the theory which found most acceptance within the scientific community until Horne et al. (2005) showed that radial inward motion does not provide sufficient energies. Instead, interaction with electromagnetic waves (e.g. Chorus waves) propagating through plasma was suggested as the sought-for mechanism. In this case, particle injections during substorms form a seed particle population in the radiation belt, while the necessary wave activity is related to high-speed solar wind (Cully et al., 2008). The wave-particle interaction approach also accounts for observed depletion and fast reformation behaviour of the outer radiation belt (Horne et al., 2005; Chen et al., 2007). Due to the scattering process, electrons in the outer belt are accelerated and their pitch angle distribution is altered. Once they attain sufficient energy and the right pitch angle, they can penetrate into the atmosphere and deposit their energy at latitudes equatorward of regular auroral precipitation. This so-called Energetic Electron Precipitation (EEP) presents one of three different constituents of EPP with a distinct source region and energy spectrum. The presence of the empty slot region between the two radiation belt is still not completely understood, but wave-particle interaction is likely to be the key to the depletion of this region (Lyons et al., 1972).

The ionosphere, which has been mentioned in different contexts already, is strictly speaking not a magnetospheric region but the upper part of the atmosphere, located below the plasmasphere. In general, it presents the key link between space weather effects in the different regions, and its variability manifests itself in the form of changing conductivity and neutral winds. While only the general facts will be stated here, the interested reader is recommended a more comprehensive review of ionospheric physics provided by Hagfors and Schlegel (2001).

The otherwise neutral atmosphere is partially ionized above an altitude of roughly 60 km, mainly due to photo-ionization by solar radiation. A different source of ionization is given by particle precipitation and the associated bremsstrahlung, also called

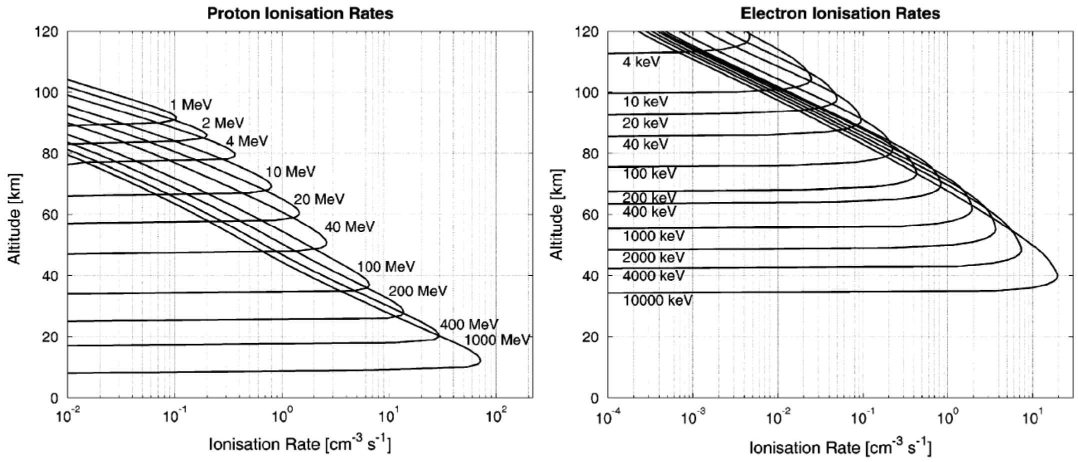


FIGURE 2.12: Ionization profiles of protons and electrons with kinetic energies ranging from 1 - 1000 MeV and 4 - 10000 keV, respectively. Adapted from Turunen et al. (2009).

impact ionization (Kivelson and Russell, 1995, p. 183). Whereas photo-ionization is largely restricted to the illuminated day-side, impact ionization affects also the night-side. Ion losses result from various recombination processes (radiative, dissociative) and may differ depending on altitude and local time. The different source and loss terms balance and yield equilibrium ion concentrations. The presence of neutrals and particle precipitation alters the local plasma conductivity and plays therefore an important role when investigating the motion of charged particles, for example due to field line movement in the Dungey cycle. The quiet-time ionosphere can be divided into different altitude regions according to its distinct electron density profile (Figure 2.17, right plot), starting with low density in the D-region and continuing upwards toward higher electron densities in the E-, F1- and F2-region (Gombosi, 1998, p. 176). During geomagnetic storms, precipitating particles change local ionization rates and smear out the otherwise regular electron density profile. Figure 2.12 shows the altitude dependent ionization rates of electrons and protons, separated in terms of their kinetic energy. As can be seen, ionization rates increase and peak ionization altitudes decrease with increasing particle energy. Comparing electrons and protons with the same energy, electron peak ionization is located lower. Typical upper proton energies, however, exceed electron energies significantly, rendering energetic protons as potentially dominant species at lower altitudes.

2.3 Geomagnetic storms

During the presence of a southward component in the IMF, magnetic reconnection with the Earth magnetic field is particularly strong and day-side reconnection rates are enhanced. The magnetic field is effectively opened up and open field lines are transported towards the night-side by the ambient solar wind. Once the magnetic pressure on the magnetotail becomes too large, open field lines reconnect and magnetic energy is converted to kinetic energy by the acceleration of local solar wind particles. The closed reconnected field lines are transported back towards the day-side, closing the Dungey cycle (Dungey, 1961). The trajectories of magnetospheric charged particles are determined by magnetic and electric fields. Differences in the motions of the ion and electron fluid result in the formation of a current system. According to Maxwell's equations, every current generates a magnetic field which is superimposed on the background Earth magnetic field and thus alters field strengths measured with ground-based magnetometers. For example, the magnetic field resulting from the ring current is opposite to the background geomagnetic field, leading to a negative deviation from the initial field. During geomagnetic storms, many particles are injected into the ring current which makes it stronger and increases its signature in magnetograms. Indices measuring the geomagnetic activity aim to reflect the extent of energy transfer into the geomagnetic field and thus categorize geomagnetic storm strengths. Although the general driver for the initial day-side reconnection relates to the velocity of solar wind particles and the direction of the $\mathrm{B_z}$ component, different features and phenomena can be responsible for the commencement of a geomagnetic storm. Essentially, one differentiates between storms driven by Co-rotating Interaction Regions (CIRs) and CME-driven storms (Zhang et al., 2007).

2.3.1 Definition of geomagnetic indices

Historically, scientists had to rely on ground-based measurement techniques to study geomagnetic activity. In-situ measurements were not possible for various reasons and many current systems were not entirely understood. In order to be able to interpret large datasets in a comparative and understandable fashion, a large number of geomagnetic indices was defined. This section will present the most common ones, which are still in use to classify geomagnetic storms, based on a review of Rostoker (1972).

The Kp index is a measure of the global geomagnetic activity which was first introduced in 1938 (Bartels et al., 1939). Thirteen different observatories located in sub-auroral latitudes measure the maximum variation of the horizontal magnetic field component throughout 3-hour intervals using magnetometer data. The obtained number is then corrected for natural latitudinal and diurnal variations and translated into an integer number between 0 (low) and 9 (high geomagnetic activity). In order to obtain a finer division, stages with "+", "-" and "0" are included for each number.

The daily Ap index is obtained by converting 3-hour Kp indices according to a set conversion scheme and taking daily averages.

Due to the locations of the observatories, the Kp and Ap index can not yield information on one particular source of geomagnetic deviation, because their measurements are influenced by at least two current systems, i.e. the ring current and the auroral electrojet (Figure 2.9). In order to get a better measure of the auroral-zone activity, i.e. the eastward and westward electrojets, the Auroral Electrojet (AE) index (Davis and Sugiura, 1966) is used. Observatories which are located in auroral latitudes record the perturbation of the horizontal magnetic field component in 1-minute intervals and normalize it to quiet-day values. From the results, the AU index can be defined as the maximum positive perturbation measured by any observatory for a given universal time and the AL index represents the lowest deflection, including mostly negative amplitudes. In a superimposed plot of all observatory measurements, the AU and AL act as upper and lower envelopes. The AU and AL can also be interpreted as estimates for the eastward and westward electrojet strengths, respectively. The total perturbation amplitude is then given by the AE index which is defined as $\mathrm{AE} = \mathrm{AU} - \mathrm{AL}$.

In order to get a grasp on the ring current strength, the Dst index is used (Sugiura, 1964). During a geomagnetic storm, injected particles enhance the ring current and the associated magnetic field which is directed oppositely to the undisturbed Earth magnetic field for ground-based measurements. Four low-latitude observatories measure the perturbation of the horizontal magnetic field component during 1-hour intervals. The Dst index is calculated as the average of the four measurements. In some cases it might be necessary to correct the Dst value for the disturbance originating from magnetopause currents, e.g. during SSCs. Most approaches for correction are based on a simple relationship provided by

$$Dst^{*} = Dst - bp^{a} + c \tag{2.3}$$

with the corrected Dst value, Dst*, and the solar wind pressure, p. The constants a, b and c are determined from data analysis and different studies yield different values, for example

$a = 0.5$, $b = 15.8\ \mathrm{nT\,nPa^{-0.5}}$ and $c = 20$ nT (Burton et al., 1975).

2.3.2 CIR-driven storms

Low-latitudinal CH extensions and a slight tilt of the streamer belt enable an interaction between streams of fast and slow solar wind. In this case, the fast solar wind stream ($\sim$ 750 - 800 $\mathrm{km\,s^{-1}}$) overtakes the preceding slow solar wind ($\sim$ 300 - 400 $\mathrm{km\,s^{-1}}$) as shown in Figure 2.13 leading to both plasma and magnetic field compressions (Tsurutani et al., 2006). Especially during the declining phase of the solar cycle, CH structures exhibit equatorward extensions which can be considered quite steady leading to a periodic occurrence of CIRs as CHs rotate with the Sun, where the rotational period close to the equator is given by approximately 27 days. Amongst others, Borovsky and Denton

(2006) managed to link periods of elevated geomagnetic activity measured in Kp during declining phases to a 27-day recurring pattern. For reasons discussed later, CIR-driven storms are rather weak and strengths of Dst $< -100\,\mathrm{nT}$ were considered impossible for a long time. Richardson et al. (2006), however, analysed in total 79 geomagnetic storms with a minimum Dst $< -100\,\mathrm{nT}$ from the period 1996 - 2004 and found nine cases, where the storm was exclusively generated by a CIR. In general, a CIR-driven storm can be divided into three phases, the initial, main and recovery phase, each having distinct features and differences from CME-driven storms. Upstream of the CIR, i.e. closer to Earth, propagates the Heliospheric Current Sheet Plasma Sheet (HCSPS), which is characterized by slow but high-density solar wind. Its impact on the magnetopause compresses the magnetosphere and moves magnetospheric currents closer to Earth resulting in a positive increase of the Dst during the initial phase (Tsurutani et al., 2006). In the main phase, the CIR itself impinges the Earth. Although the compression of the magnetic field generally leads to large field strengths, the $\mathrm{B_z}$ component remains highly variable. This means, that magnetic reconnection occurs rather sporadically associated with only short periods of energy transfer in the magnetotail. Therefore, CIR-driven

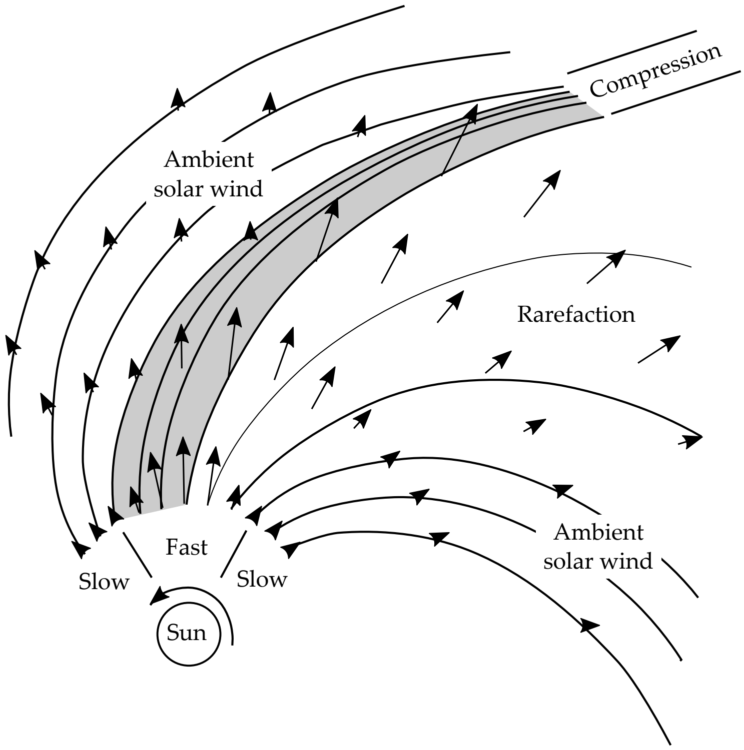


FIGURE 2.13: Schematic drawing showing the interaction between a fast solar wind stream (grey) and the ambient slow solar wind. Solar wind velocities are indicated by black arrows. When the fast solar wind catches up with the slow solar wind, a region of compression with high density and magnetic field strengths develops. Modified from Russell (2001).

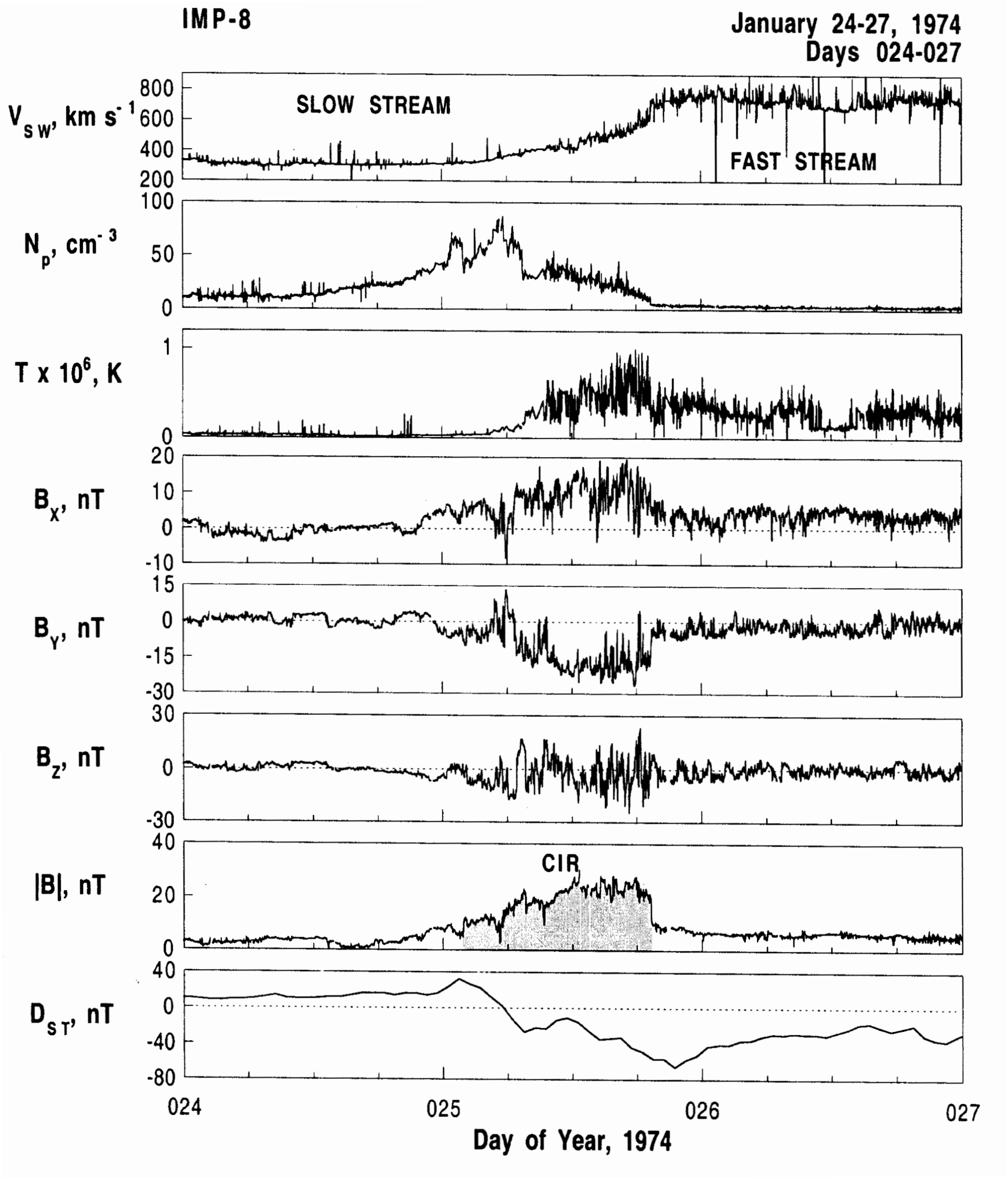


Figure 2.14: Example stack plot of a CIR-driven storm from January 1974. From Tsurutani (2000).

storms result in quite weak geomagnetic activity and irregular Dst profiles. The subsequent recovery phase is initiated by a sudden decrease in the magnetic field strength as the CIR passes Earth. Its duration, however, can easily exceed 24 hours. Occasionally, especially long recovery phases can lead to an elevated geomagnetic activity for several days, a so-called High-Intensity Long-Duration Continuous AE Activity (HILDCAA). In order to count as a HILDCAA, the AE index must exceed 1000 nT for more than two days and must not fall below 200 nT for more than two hours. During a HILDCAA, the magnetic B_z component is highly variable leading to small injections of particles and a coinciding increase in AE as well as decrease in Dst (Tsurutani et al., 2006). An example of a typical CIR-driven storm is given in Figure 2.14, which depicts solar wind properties measured by the ACE satellite and the Dst index. The panels show from top to bottom the solar wind velocity, proton density, temperature, magnetic x-, y- and z-component, absolute magnetic field strength and Dst. Looking at the upper panel, it can be nicely seen how the initial solar wind velocity ranges at roughly 350 $\mathrm{km\,h^{-1}}$, followed by an increase to 800 $\mathrm{km\,h^{-1}}$ at approximately 1800 UTC (coordinated universal time) on January 25th. This presents the junction between slow and fast solar wind stream. Prior to this, early on January 25th, the initial phase starts, marked by an increase in Dst (lower panel). As explained above, the high-density HCSPS drives the initial phase and an associated high proton density is evident from the second panel coinciding with the initial phase. During the main phase, where the CIR hits the magnetosphere, Figure 2.14 depicts a visible decrease in Dst and a strong absolute magnetic field strength. As previously claimed, the z-component is subject to strong variability. The beginning of the recovery phase is marked by the drastic drop in magnetic field strength at approximately 1800 UTC and the arrival of the fast solar wind stream. The Dst index remains negative for at least 1.5 days, even though the varying magnetic B_z component allows only short injections of plasma.

Due to their 27-day periodicity and long recovery phases, weak CIR-driven storms present a major source of energy input into the magnetosphere during declining phases. Also, they seem to play a significant role in the production of magnetospheric relativistic electrons during the recovery phases. Whereas the exact accelerating mechanism has not been accounted for yet, wave-particle interactions breaking either the first or third adiabatic invariant are considered (Tsurutani et al., 2006).

2.3.3 CME-driven storms

CME-driven storms form the second category of geomagnetic storms. When plasma is ejected from the Sun by a CME, it propagates through interplanetary space in different configurations, potentially crossing Earth's orbit and causing perturbations in the magnetosphere. In fact, roughly 87 % of catalogued geomagnetic storms during 1996 - 2005 with a minimum Dst $< -100\,\mathrm{nT}$ were associated with either one or multiple CMEs according to Zhang et al. (2007). Although the general development of a CME-driven storm resembles that of a CIR-driven storm, as depicted in Figure 2.15a, initial, main and recovery phases are caused by different processes and drivers.

CME ejecta in interplanetary space can appear in various configurations and are generally called interplanetary coronal mass ejections (ICMEs). The ICME speed at 1 AU ranges from 350 $\mathrm{km\,s^{-1}}$ up to 1350 $\mathrm{km\,s^{-1}}$, depending on their solar source region and ambient solar wind properties (Wang et al., 2004). Slow ICMEs are pushed by the faster solar wind downstream, whereas fast ICMEs are blocked upstream, leading to different compression regions and angles of deflection. Still, ejecta from front-side CMEs exhibit the highest probability of intersection with Earth including an east-west asymmetry from central meridian due to deflections mentioned above (Wang et al., 2004). With respect to source region longitudes, ICMEs usually originate from closed solar field lines, transporting a magnetic flux rope away from the Sun ("Coronal Mass Ejections and Magnetic Flux Ropes in Interplanetary Space"). A constant flux of electrons along these flux ropes suggests that they remain connected to the Sun at their footpoints ("Coronal Mass Ejections and Magnetic Flux Ropes in Interplanetary Space"). Assuming the ICME's velocity exceeds the ambient solar wind speed, magnetized plasma drapes around the ICME fluxtube forming a sheath region upstream of the ICME ejecta. If, in addition, the relative velocity between ICME and solar wind is larger than the present Alfvén velocity, a shock is produced in front of the sheath as shown in Figure 2.15b (Russell, 2001).

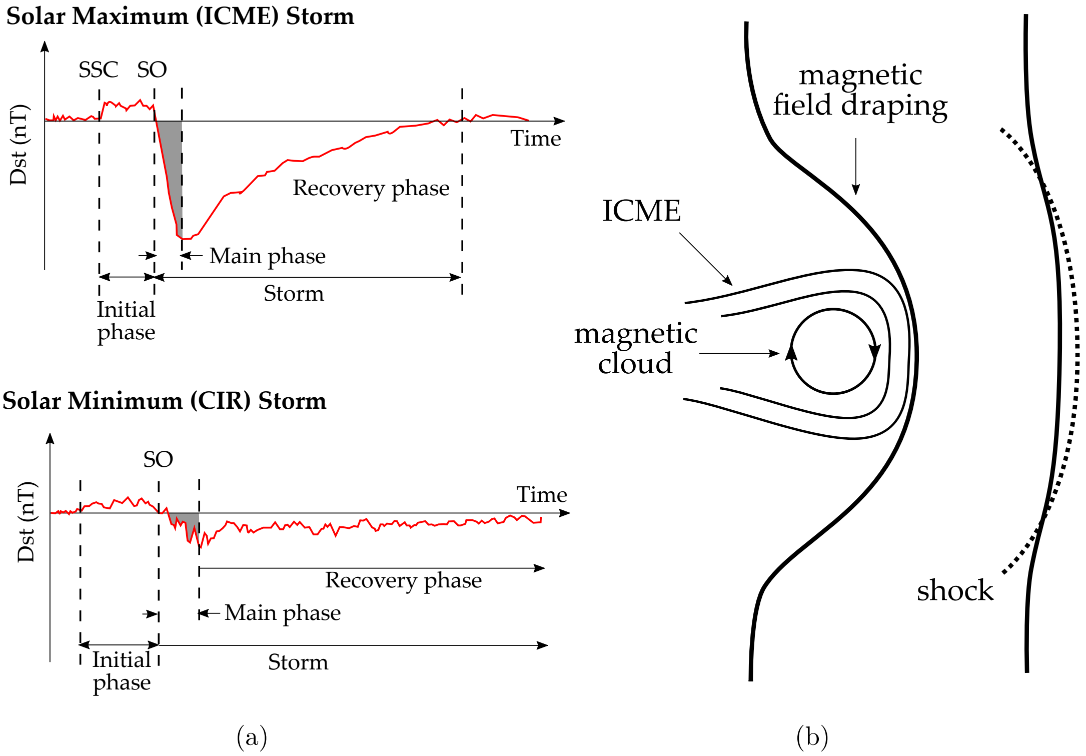


FIGURE 2.15: (a) Schematics of geomagnetic storms driven by an ICME (top) and a CIR (bottom). Lines mark SSC and storm onset (SO) times. (b) Schematic of an ICME in space with magnetic cloud, upstream sheath and shock. Based on Tsurutani (2000).

Approximately 30 % of ICMEs at 1 AU contain a so-called magnetic cloud (MC; “Coronal Mass Ejections and Magnetic Flux Ropes in Interplanetary Space”), a structure first proclaimed by Burlaga et al. (1981). It is characterized by a strong rotating magnetic field, which either forms closed loops or helix-like internal structures. The plasma density and temperature is very low inside the cloud, resulting in a dominating magnetic pressure. Strong magnetic fields both in the magnetic cloud and in the preceeding sheath present possible drivers of a geomagnetic storm. Although the sheath magnetic field is very turbulent, sheath-driven storms reach a maximum intensity similar to cloud-driven storms (Pulkkinen et al., 2007). As mentioned earlier, the number of CME-driven storms peaks during the maximum of the solar cycle (e.g. Webb, 1991), concurring with the relation between CMEs and solar active regions.

Figure 2.15a depicts schematic developments of a CIR-driven storm (bottom) and a CME-driven storm (top). Although both categories of magnetic storms can be divided into initial, main and recovery phase, their generators and features differ distinctively. In 70 % of Kp > 4+ CME-driven storms, the initial phase commences with a SSC (Borovsky and Denton, 2006), which describes a sudden increase in Dst as the CME shock arrives at the magnetosphere and compresses it. In general, CIR-driven storms lack this feature because they do not form an upstream shock and hence their initial phases depict a rather gradual increase in Dst. The CME-driven main phase is then characterized by long periods of a strong southward directed IMF (Tsurutani et al., 2006) resulting in a smooth and steep decrease of the Dst index as particles are injected into the ring current. During the recovery phase, particle losses in the ring current due to wave-particle interactions, Coulomb collisions and convection lead to an increase in Dst towards the zero line.

So far, it has been mentioned that CIR- and CME-driven storms differ in various aspects, i.e. their occurrence pattern, drivers, influence on the ring current and magnetosphere as well as general storm features. In addition, CME-driven storms are occasionally associated with strong solar proton events (Borovsky and Denton, 2006). SPE occurrence is usually linked to the presence of strong interplanetary shocks, so that CIR rarefaction shocks contribute only to very few and weak energetic protons (Mason and Sanderson, 1999).

2.3.4 Solar proton events

The term solar proton event describes a period of elevated energetic proton flux with at least 10 particles $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-1}$ of kinetic energies $\mathrm{E} \geq 10\,\mathrm{MeV}$ (Bothmer and Daglis, 2007, p. 143). SPEs are known by a variety of names, amongst others solar cosmic rays, polar cap absorption (PCA) and solar energetic particle (SEP) events, where historically each name refers to a specific observed phenomenon (Shea and Smart, 1990). The term "solar cosmic rays" originates from enhanced cosmic ray measurements in muon chambers due to radiation produced by precipitating energetic ions, whereas "PCA" stems from high latitude radio disturbances. SEP events describe the more general phenomenon of sub-relativistic and relativistic ions, including protons (Borovsky and Denton, 2006). Wild et al. (1963) suggested two categories of SEP events which are defined by different source regions and acceleration mechanisms. One therefore has to differentiate between impulsive and gradual SEP events. Impulsive events are related to solar flare eruptions where ions and electrons are strongly accelerated during reconnection and result mostly in the population of individual fluxtubes. Thus, their spatial extents and durations are rather limited. Gradual events on the contrary can fill up to half of the inner heliosphere with energetic particles and lead to long-lasting high intensities at 1 AU (Reames, 2013). The acceleration of particles takes place in the presence of strong interplanetary shocks, e.g. upstream of ICMEs, and ion abundances correspond in the widest sense to coronal compositions (Breneman and Stone, 1985) with protons being the dominating species. Since solar proton events describe a period of high energetic proton flux related to a gradual SEP, they form a subgroup of SEPs. Obviously, SPEs are connected to coronal and solar activity, so that the number of corresponding events mostly follows both the solar cycle (Shea and Smart, 1990) and the Gleissberg cycle (Reames, 2004). The occurrences and strengths of SPEs during the period 1975 - 2016 are shown in Figure 2.16. In addition, the diagram depicts the F10.7 radio flux, which is often used as a proxy for solar UV activity and the 11-year solar cycle. Although a

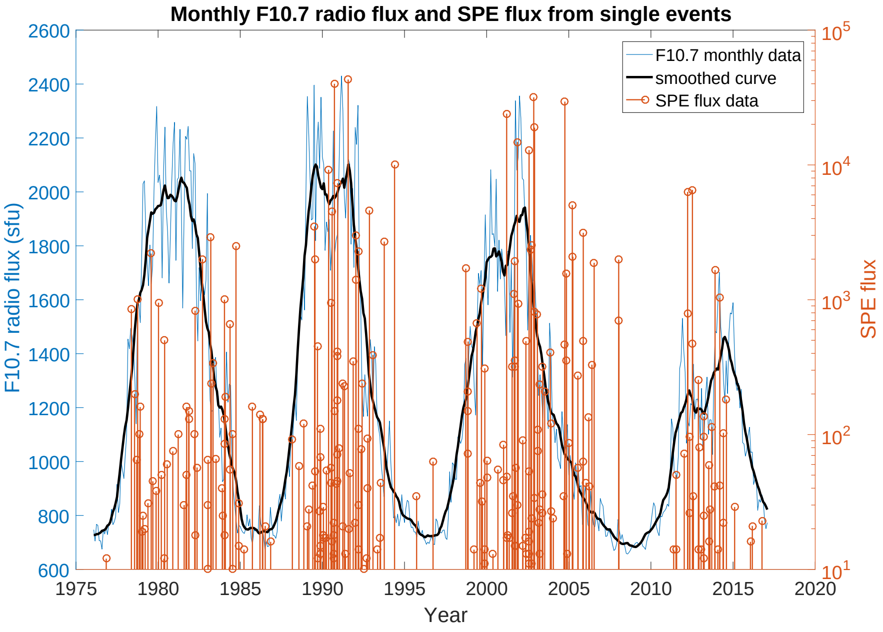


FIGURE 2.16: Diagram showing SPE occurrences and strengths (red stems) from 1975 to present. In addition, the F10.7 radio flux which is used as a proxy for solar activity is shown as monthly data (blue curve) and smoothed line (black curve). Data was obtained from `ftp://ftp.swpc.noaa.gov/pub/indices/SPE.txt` for SPE flux values and `ftp://ftp.ngdc.noaa.gov/STP/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_observed/listings/listing_drao_noontime-flux-observed_monthly.txt` for monthly F10.7 radio flux data.

large part of strong SPEs falls within times of solar maximum, their occurrence is clearly not restricted to them. Especially during the declining phase it seems as if a significant amount of SPEs was registered.

In general, the occurrence and strength of an SPE depends on a number of factors. Kahler and Reames (2003) found that the angular width of a CME seems to influence the formation of shocks, so that narrow CMEs do not produce energetic protons. In addition, the CME has to be fast enough to create a sufficiently strong shock, which is accomplished only by the fastest 1 - 2 % according to Reames (2013). The peak intensity and maximum energy depend on the speed of the shock and the size of the seed particle population (Kahler, 2001). Thus, strong SPEs can also be related to acceleration by several consecutive shocks.

The acceleration mechanism involving interplanetary shocks is based on equilibrium theories by Bell (1978) and Lee (1983) and is described in detail in Jones and Ellison (1991). Protons with sufficient energy can travel away from the shock and create Alfvén waves both upstream and downstream of the shock. Ions are then scattered back and forth between the resonant Alfvén waves, gaining energy each time they cross the shock. Eventually, their accumulated kinetic energy allows them to escape and travel away from the shock, which in turn amplifies the resonant Alfvén waves. Trapped ions are thereafter scattered even more effectively and reach higher speeds before escaping the shock. This circle of wave amplification and particle acceleration can lead to proton energies of up to 20 GeV (Kahler, 1994), which is enough to penetrate into spacecraft walls and deep into the polar atmosphere, where the magnetic field lines are open as shown in Figure 2.10 (Reames, 2004). Depending on the maximum energy, energetic proton precipitation can also affect latitudes with closed magnetic field lines down to approximately 55° magnetic latitude. Less energetic protons can act as seed particles for potential downstream shocks and intensify subsequent SPEs.

2.4 The atmosphere

After having discussed how the Sun couples to near Earth space via the solar wind and IMF, a new focus will lie on how solar variations present a driver for both global and local terrestrial climate variability. In order to account for atmospheric responses on solar variability, it is necessary to understand the underlying structures, dynamics and coupling mechanisms. Afterwards, different solar forcing and models for the feedback mechanisms they initiate will be presented. Although irradiance variations and particle forcing should ideally be treated separately, superposition and aliasing of effects present some of the major challenges when trying to assign observations to one specific forcing.

2.4.1 Atmospheric structure

Extensive descriptions of the vertical atmospheric structure and dynamics can be found in many basic climatology and meteorology publications (e.g. Marshall and Plumb, 2016). This summary of the most important features is based on a review provided by Smith (2012). In general, the vertical structure of the neutral atmosphere is defined by its temperature profile as depicted in Figure 2.17. In the left plot it is evident that the temperature decreases with altitude at first, but this behaviour reverses several times, at approximately 10, 50 and 100 km, respectively. The lowermost region, which extends from the surface to approximately 10 - 18 km, is called the troposphere followed by the stratosphere (up to ∼ 50 km), mesosphere (up to ∼ 90 - 100 km) and thermosphere (above ∼ 90 - 100 km). The respective transitions are called tropopause, stratopause and mesopause. Their exact positions exhibit a strong seasonal and latitudinal dependence. As additional terminology, the so-called middle atmosphere is here referred to as the region spanning from 50 to 120 km altitude which could be directly impacted

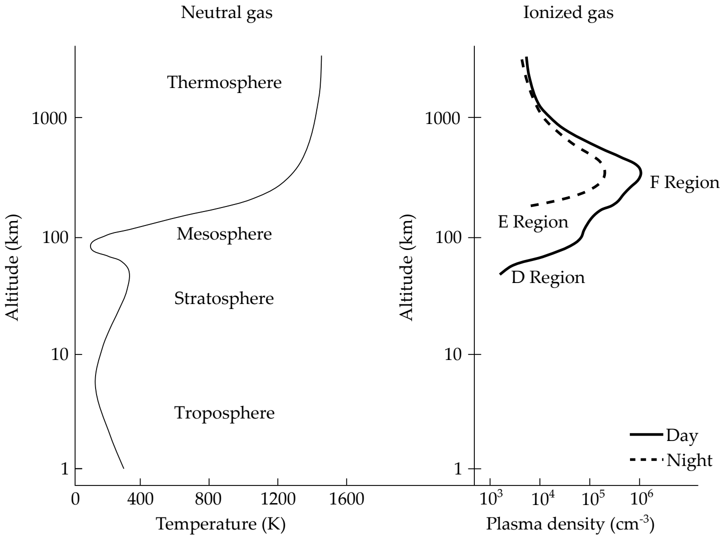


Figure 2.17: The temperature profile of the atmosphere (left) and electron density curve of the ionosphere for day and night. Adapted from Kelley (2009).

by EPP. It is also worth noting that the mesosphere and lower thermosphere region is often assigned the acronym MLT in professional literature. In this work, MLT will solely stand for magnetic local time and to prevent confusion, no abbreviation will be introduced for the mesosphere and lower thermosphere.

The troposphere is the atmospheric region where day-to-day weather takes place. Its high variability can in general be explained based on its characteristic temperature profile. The main heating mechanism in the troposphere is when a large part of the incoming solar radiation reaches the ground where it is either reflected or absorbed. This, in turn, leads to both evaporation and heating, impacting air temperatures close to the surface. Diabatic heating can be the cause of convective instability dependent on the ambient temperature gradient, $\frac{\partial \mathrm{T}}{\partial \mathrm{z}}$ (Smith, 2012, Equation 1). The general stability conditions are based on wet and dry adiabatic cooling (depending on the air's humidity) and the associated expansion of rising air parcels. The convective upward acceleration continues until the stability condition is fulfilled again by the ambient temperature profile. The presence of convective instabilities is closely linked to the negative temperature gradient in the troposphere. As this gradient reverses in the stratosphere, the tropopause forms a boundary for convective upward motion. Although the troposphere can be considered well mixed, especially with regard to vertical chemical distributions, local variability is present due to different heat capacities of land and ocean as well as varying ground albedo, i.e. reflectivity.

In contrast to the troposphere, where heating occurs first and foremost in connection with the surface, heating of the middle atmosphere happens internally. Different stratospheric layers are dominated by different atomic and molecular species which all own a unique response to incoming short-wave and outgoing long-wave radiation. The overall heating and cooling rate is therefore a superposition of the single species rates. As one of the dominating effects in the stratosphere, solar ultraviolet (UV) radiation is absorbed by ozone ($\mathrm{O_3}$) which leads to local heating (Matsumi and Kawasaki, 2003). Stratospheric heating rates are therefore closely linked to ozone concentrations in different altitudes. The more ozone which is present in a certain layer, the higher the probability of short-wave absorption. Thus, the observed temperature profile is strongly affected by the vertical ozone distribution. Molecular oxygen ($\mathrm{O_2}$) is a not quite so effective absorber but turns out to play an important role in the formation process of ozone, called Chapman cycle (F.R.S., 1930). Radiative cooling due to $\mathrm{CO_2}$, water vapour and ozone is common in the stratosphere as well (Oinas et al., 2001). Due to the positive lapse rate, the stratospheric layers can be considered very stable.

In the mesosphere, absorption of solar UV radiation by atomic oxygen (O) and water is the main source of heating. The photolysis of water does not contribute to the heating directly but via secondary effects caused by the associated products. Depending on their stability, potential transport has to be kept in mind so that heating does not necessarily occur at the site of absorption (Smith, 2012). Another example is the production of atomic oxygen in the thermosphere. Transport of O into the mesosphere can be considered quite common and the chemical reactions converting 2O into O_2 lead to local heating in the mesosphere. In contrast to the stratospheric composition, the mesosphere and thermosphere have to be considered part of the heterosphere where gases separate into their constituents according to their respective molecular masses. Atomic oxygen lifetimes increase with height giving it a central role in vertical energy redistribution. Heating in the mesosphere and lower thermosphere is dominated by species involving hydrogen and oxygen where many of the associated reactions cause heating of the ambient gas (Smith, 2012). As evident from Figure 2.17, concentrations of ions and charged particles begin to rise in the mesosphere and increase towards to thermosphere.

2.4.2 General circulation

To the very first degree, global temperature distributions may be ascribed to the radiative balance between absorption of incoming short-wave and emission of outgoing long-wave radiation. Observations, however, clearly contradict this simplified approach and suggest a more complex system of balance. Heating does occur largely in equatorial regions which are exposed to sunlight, while cooling and emission of long-wave radiation exhibits a more even latitudinal distribution (Figure 2.18). In order to achieve a meridionally balanced temperature distribution, heat and air masses must be transported from tropical regions towards higher latitudes. Otherwise low latitudes would continually heat up while higher latitudes would be characterized by constantly decreasing temperatures. When, however, considering the flow of air on a rotating planet, surface

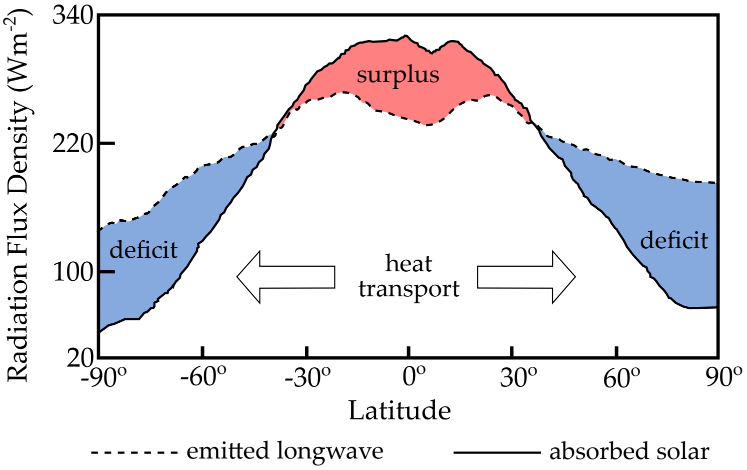


Figure 2.18: Annual mean absorbed short-wave and emitted long-wave radiation. In equatorial latitudes absorption exceeds emission leading to a surplus of energy (red). Closer to the poles more radiation is emitted than absorbed corresponding to an energy deficit (blue). Accordingly energy in the form of heat is transported from the equator towards the poles. Adapted from Marshall and Plumb (2016, Figure 5.5).

orography and pseudo forces, such as the Coriolis force, complicate the picture.
In order to describe the general circulation on Earth, reducing the variables to zonal means and a local perturbation term (Eulerian mean) turns out to be highly instructive. In case of the zonal velocity u, $\overline{u}$ corresponds to the background flow while u' presents the local deviation from the mean due to the presence of eddies and waves, which in turn implies the occurrence of different meridional circulations (Holton and Hakim, 2012). One of the most robust and basic equations for large-scale meteorology is the thermal wind balance

$$\frac{\partial u}{\partial z} = \frac{R}{fH}\frac{\partial T}{\partial y} \tag{2.4}$$

linking vertical wind shear to a meridional temperature gradient with the universal gas constant, R, atmospheric scale height, H, and Coriolis term, f, (Marshall and Plumb, 2016, p. 126, Equation 7-24). The latter accounts for the different strengths of the Coriolis force at different latitudes as it acts on the azimuthal projection. From the thermal wind balance it follows that high temperatures at the equator and low temperatures in higher latitudes will lead to westerly winds in both hemispheres with increasing speed at higher altitudes. Actual observed circulations in the different atmospheric regions can be resembled by this basic principle but deviate due to other influences.

Tropospheric circulations and the North Atlantic Oscillation

Figure 2.19 illustrates the circulations present in the troposphere. As can be seen, air rises in the tropics due to diabatic heating at the surface creating a high pressure region. In high altitudes, air will therefore flow towards regions of lower pressure at roughly 30° latitude creating a high pressure system at the surface and backflow of air towards the equator, closing the so-called Hadley cell (Marshall and Plumb, 2016, p. 73). As the air travels meridionally, it is deflected due to the Coriolis force creating

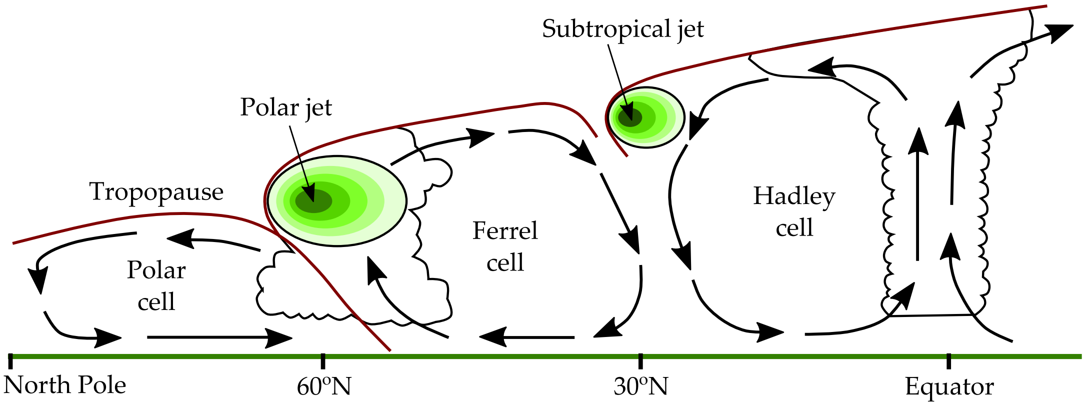


FIGURE 2.19: Illustration of the troposphere circulation. The Hadley cell is located in the tropics, followed by the Ferrel cell and the polar cell in higher latitudes. The tropopause altitude (red) decreases with latitude and the formation of jets (green) leads to tropopause foldings. Based on a figure from the U.S. National Weather Service (`http://www.srh.noaa.gov/jetstream/` `/global/images/jetstream3.jpg`).

the subtropical westerly jets, surface trade winds and Intertropical Convergence Zone (ITCZ). In addition, air at approximately 60° latitude rises, creating a high pressure region in tropopause altitudes and divergence both towards polar latitudes and the equator. Adiabatic warming by downward motion compensates the cooler temperatures in the polar regions and leads to a flow back towards the original convergence latitude. This polar cell, also called polar vortex, is mostly identified by a strong low pressure system at the poles, causing the presence of polar jet streams at mid-latitudes. These are particularly stable during winter and can be assigned different modes, depending on the strength of the westerly jet. One differentiates between the Arctic Oscillation (AO), which includes the entire polar jet and is measured with the AO index, and the North Atlantic Oscillation (NAO), which focuses on the westerly wind between the Iceland low and Azores high, reflected in the Northern Annular Mode (NAM) index. In addition, the 60° latitude up-welling is linked to the Hadley cell 30° latitude down-welling by the Ferrel cell (Holton and Hakim, 2012). As mentioned above, the troposphere can be considered well mixed whereas stratospheric layers are quite stable. In general, coupling between different atmospheric regions is rather weak. The presence of the three cells mentioned above, however, leads to tropopause foldings close to the jets which present efficient opportunities for mass exchange, i.e. chemical coupling, between the troposphere and stratosphere.

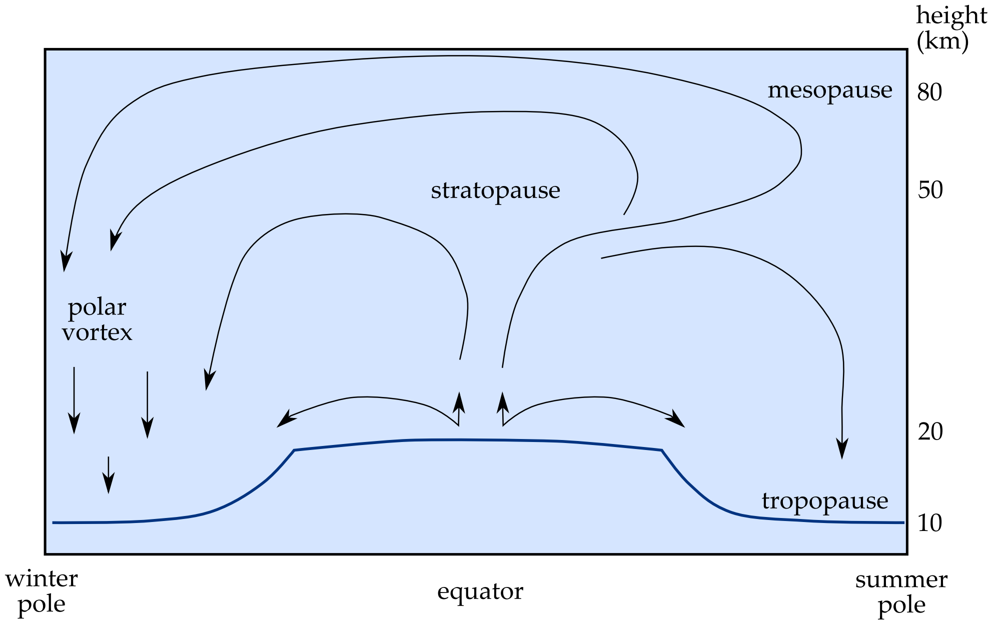


FIGURE 2.20: Illustration of circulations above the tropopause. Modified from Proedrou et al. (2016).

Circulations in the Stratosphere, Mesosphere and lower Thermosphere

When moving towards higher altitudes classifying as stratosphere and mesosphere, the major circulation drivers are gravity and planetary waves. Planetary waves, also known as Rossby waves, are large scale systems originating from the motion of air masses on the rotating planet, for example meanders of the jet streams. The phase velocity of planetary waves is always easterly but the observed group velocity can vary. Gravity waves emerge whenever gravitational or buoyant forces try to restore a displaced air parcel, while inertia and overshooting will lead to oscillations. Varying surface orography is a common generator for gravity wave activity but they may also result from air displacement in higher altitudes due to strong convection in the tropics. In order to quantify and track eddie heat and momentum flux, it is beneficial to shift to the Transformed Eulerian Mean formalism which separates advective and diffusive components in the eddy flux (Holton and Hakim, 2012). In general, wave amplitudes increase with altitude due to the decreasing air density. Limits for the propagation of waves are given by their amplitude and the background zonal wind. It can be assumed that background winds are not influenced by an ordinary linear wave, but that the interaction with the background flow requires the wave to break and deposit its momentum and energy. According to the Charney-Drazin criterion for atmospheric wave breaking (Charney and Drazin, 1961), a wave ceases to propagate when

i) it reaches a critical layer where the background flow equals the wave speed,
ii) the wave amplitude becomes too large leading to instabilities in the wave,
iii) its wave number is too large.

The propagation and subsequent breaking of a wave implies that the deposition of its energy and momentum occurs in a region different to its origin. Via this mechanism, convective instabilities in one atmospheric region can impact dynamics in a different region without chemically coupling to it. Thus, a planetary wave originating from the troposphere can propagate upwards until it breaks in the middle atmosphere altering the present circulation. Vice versa, zonal winds can act as wave filters. Depending on the background flow in a certain atmospheric layer, some waves are forced to break while others pass through it without effect. In the stratosphere, the equator-pole temperature gradient and thermal wind balance yield a certain initial background flow, which in turn acts as a wave filter for planetary waves. The associated planetary wave breaking, also called resolved wave forcing, drives and reinforces the Brewer-Dobson (BD) circulation with up-welling in the tropics, a flow towards the poles and down-welling there (Figure 2.20 below the stratopause). Breaking of gravity waves, i.e. unresolved wave forcing, turns out to be the main driver of the Murgatroyd-Singleton (MS) circulation in the mesosphere and lower thermosphere which transports air from the summer to the winter pole and presents a unique coupling between the two hemispheres (Figure 2.20 above the stratopause).

2.4.3 Solar forcing

As explained above, incoming solar radiation is one of the driving forces governing Earth climate. Solar activity, however, varies over a range of time scales, e.g. the 11-year solar cycle, and can exhibit single distinct perturbations like the Maunder minimum (Shindell et al., 2001). These variations can lead to three different categories of solar forcing (Gray et al., 2010, and references therein). Although sunspots appear dark, the total solar irradiance (TSI) peaks during solar maximum. An increase in TSI affects mainly surface heating rates. This in turn influences the formation of high convective clouds which are a possible generator for planetary waves. Thus, surface temperature anomalies can couple to higher atmospheric dynamics via wave forcing, resulting in a so-called bottom-up mechanism (Kodera and Kuroda, 2002). An increase in TSI will, for example, lead to higher temperatures and evaporation rates in the tropics. As more moisture is transported by the trade winds, the Hadley cell is significantly strengthened affecting wave filtering by the trade winds and wave breaking in higher altitudes. Increased upwelling in the ocean is also likely to decrease sea surface temperatures (SST).

The second form of solar forcing can be identified as variations in the solar spectral irradiance (SSI). In general, the frequency spectrum emitted by the Sun complies to a good degree with a black-body spectrum. During solar maximum, however, the percentage of UV radiation is significantly elevated. Such SSI variations affect primarily the stratosphere due to altered absorption by ozone and O_2 photolysis, leading to a warming

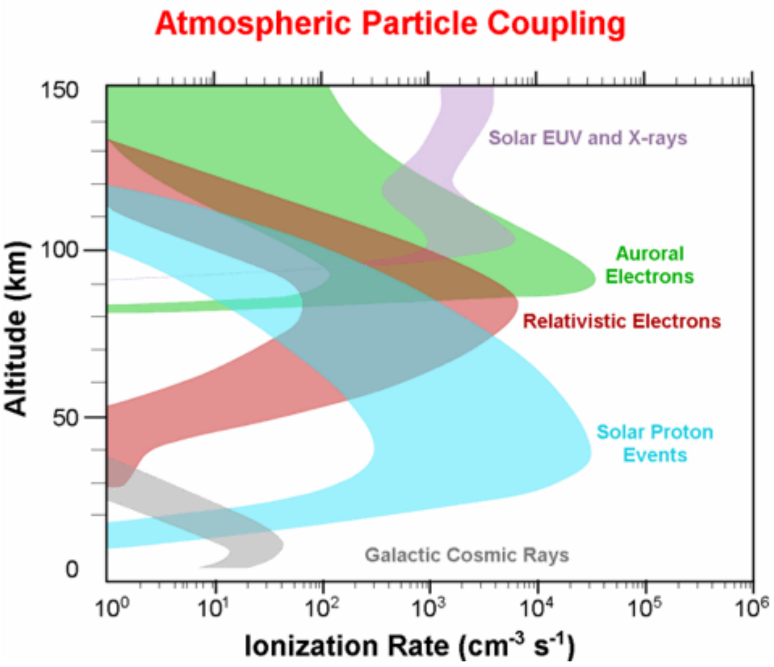


FIGURE 2.21: Altitude depenedent ionization rates of different forms of radiation, such as solar photons (violet), low energetic auroral electrons (green), relativistic electrons (red), solar protons (blue) and galactic cosmic rays (grey). Courtesy: LTPA Observer Project (`http://www.ltpaobserverproject.com/`).

effect at approximately 50 km altitude during solar maximum. Those elevated heating rates affect the meridional temperature gradient and the associated thermal wind balance (Holton and Hakim, 2012). The present circulation adapts to the new conditions by altering the meridional flow which is accompanied by zonal wind anomalies. Due to different interaction with planetary waves and changes in the BD circulation, subtropical anomalies propagate towards the winter pole where they alter local wave forcing as well as the polar vortex and AO strength. Thus, stratospheric dynamics might have an impact on surface winter climate via top-down mechanisms (Gray et al., 2010).

Last but not least, EPP provides a source of energy input that has long been unaccounted for. Solar protons during SPEs can reach altitudes of down to 30 km while direct EEP impacts range between 50 and 120 km. Precipitating particles ionize molecules and alter the local chemistry. Due to differences in energy, species, temporal and spatial features, one commonly differentiates between SPEs, EEP and auroral precipitation (Matthes et al., 2017). Figure 2.21 shows ionization rates of different forms of radiation, including solar high energy radiation, different types of particle precipitation and galactic cosmic rays, which increase ionization rates in the troposphere. During SPEs, a vast number of energetic protons can enter the magnetosphere directly over the polar caps penetrating deep into the atmosphere (blue curve). Energetic (relativistic) electrons originating from the radiation belts precipitate also at lower latitudes during more frequent but shorter events. Low energetic auroral electrons are even more common, yet barely able to penetrate into the mesosphere and lower thermosphere. When reaching the middle atmosphere, EPP increases the local production rates of HO_x and NO_x molecules (H,

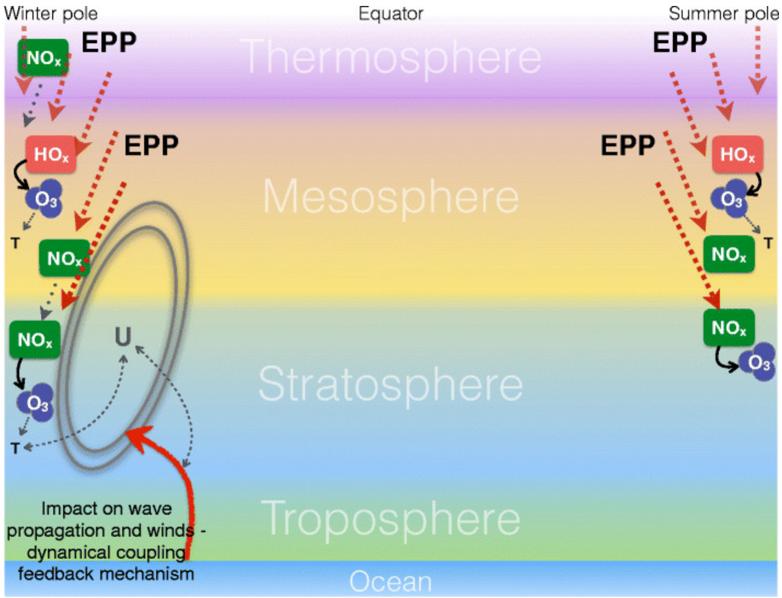


FIGURE 2.22: Pole to pole cross section showing direct and indirect impacts of energetic particle precipitation (EPP) in the atmosphere. Enhanced production and downward transport of HO_x and NO_x molecules in high latitudes lead to depletion of stratospheric ozone and zonal wind anomalies. Subsequent dynamical coupling enables alterations of surface temperatures. From Seppälä et al. (2014).

HO, HO_2 and NO, NO_2, NO_3) as well as members of the NO_y family, which act as a reservoir for NO_x (Thorne, 1980; Rozanov et al., 2012). While HO_x molecules are rather short-lived, NO_x and NO_y can survive long enough during winter to be transported towards lower altitudes where they can break down stratospheric ozone (Funke et al., 2014). The depletion of ozone leads to a top-down mechanism similar to the SSI mechanism described above where altered heating rates lead to zonal wind anomalies in the altitude of ozone depletion. As a result, vertical planetary wave propagation is altered affecting winds in the higher stratosphere and lower latitudes, effectively launching a stratospheric dynamical response between wave propagation and zonal wind anomalies that maps down to lower altitudes. Eventually, winter season oscillations such as the NAO and AO might be influenced leading to a strengthening of the NAO/AO and positive NAM (Seppälä et al., 2013). As winter continental climate and temperatures strongly depend on the air transported by the NAO, redistributions of local temperatures are the result. The impact of EPP on local surface temperatures is considered strongest during positive NAM and strong AO (Seppälä et al., 2013). These conditions correspond to a strong polar vortex which first enables the necessary stable downward transport of HO_x and NO_x molecules and reservoirs. Therefore, occurrences of other stratospheric forcing, which influence the formation and upholding of the winter polar vortex, can both affect the efficiency of EPP forcing and create an imprint on surface climate on their own. During the quasi-biennial oscillation (QBO), stratospheric zonal winds reverse, altering wave propagation and favouring either strong or weak AO (Yang et al., 2011). Although the QBO is confined to latitudes close to the equator it impact stratospheric wind stability in general. Sudden stratospheric warmings (SSWs) can lead

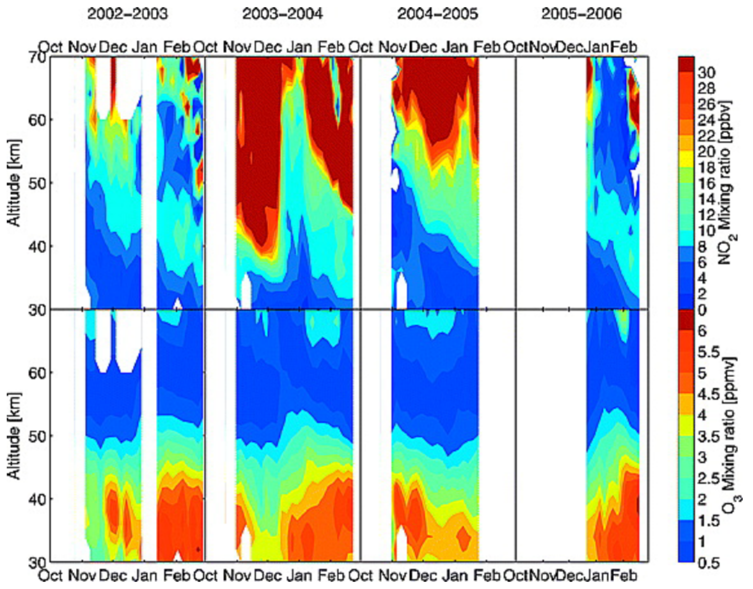


FIGURE 2.23: GOMOS Northern hemisphere winter NO_2 and O_3 mixing ratios from 2002 to 2006 during October - February. Data was averaged over 5 days and northward of 60° latitude. From Seppälä et al. (2007).

to an intermediate or even final break-down of the polar vortex due to a strong enhancement of the BD circulation (Limpasuvan et al., 2004). This can ultimately limit the downward transport of EPP generated HO_x and NO_x during the SSW itself and influence transport velocities after the polar vortex is restored (Turunen et al., 2009). Figure 2.23 shows spectrometer measurements of Northern hemisphere (NH) NO_2 and O_3 mixing ratios during winter months (October to February). NO_2 concentrations were clearly elevated after big SPEs in October/November 2003, called the Halloween event, February 2004 and December 2005. The subsequent downward propagation of the initial NO_2 anomaly initiates a visible depletion of stratospheric ozone at approximately 40 km.

2.5 Recent research

The impact of solar protons during SPEs on atmospheric ozone concentrations has been extensively studied with the conclusion that alterations in ozone can persist for several months after the SPE. According to Jackman et al. (2009), a long-term ozone effect is visible both in the mesosphere and the stratosphere, where stratospheric alterations are mainly caused by reservoir molecules from the NO_y family. Especially during the very active periods spanning years 2000 - 2004, the WACCM model suggested extensive ozone depletions in the polar stratosphere and mesosphere up to five months after SPE occurrence. Despite their short life-span, HO_x are named as dominating drivers for mesospheric effects. The point being made about SPEs, i.e. solar protons, causing major disruptions in the middle atmosphere, research on particle forcing extended by taking into account auroral electrons and EEP as well. In recent years, an increasing awareness of the importance of each particle group's specific effects has led to attempts of separating observed impacts. In Turunen et al. (2009), the SIC model was used to estimate NO_x production rates caused by solar protons, auroral electrons and relativistic electrons (EEP) in the NH winter in 2003/2004. According to the study, primary peak effects of solar protons (down to 50 km) are largest while EEP causes in turn higher initial NO_x production rates (roughly 70 km) than auroral precipitation (above 100 km). However, only the combination of all three effects reproduced observations by GOMOS to a satisfying degree.

While SPEs lead to a tremendous deposition of energy in quite low altitudes, their effect is limited to their sporadic occurrences. In comparison, auroral and energetic electrons might seem to yield only minor effects but several studies showed that the accumulated energy due to their higher repetition rate might well compete with the significance of SPEs (Turunen et al., 2009; Seppälä et al., 2013; Andersson et al., 2014). Andersson et al. (2014) showed, that the longitudinal distribution of mesospheric OH compares well to electron flux distributions measured by MEPED and falling into the range of EEP. In addition, decadal variability in the OH hotspot strengths follows geomagnetic activity variations. In Funke et al. (2014), EPP-produced NO_y was studies based on measurements by MIPAS. During winter, a peak in NO_y production lies in the mesosphere and analysing NO_y concentrations in vertical columns allows for the tracing of its downward transport. During spring-time, EPP-related NO_y peaks in the stratosphere, pointing out the significance of secondary effects. In Seppälä et al. (2013), re-analysis data was separated into years of high and low geomagnetic activity and differences in NH temperature and mean zonal wind during November - March months indicated that years of high geomagnetic activity exhibit a stronger polar vortex and positive NAM anomaly.

Although many studies which are based on the use of models only yield indications on the relevance of EEP in terms of surface air temperature patterns, its non-zero effect on atmospheric chemistry is well established. Efforts in a recent study (Verronen et al., 2015) were made to separate mesospheric electron and proton effects during two SPEs (Halloween event in 2003 and September 2005 event). The main focus in other SPE-related studies used to lie on impacts caused by solar protons, neglecting elevated electron flux levels completely. Compared to the extra-ordinarily large energy deposition by protons, electron effects might indeed seem minor in the stratosphere. In higher altitudes, however, electron fluxes may be able to compete with proton fluxes, at least during periods of moderate SPE strength. A combination of different models was used in order to reproduce fractions of electron and proton precipitation in the mesosphere as measured by EISCAT. The study's conclusion was, that electron precipitation during SPEs is the dominating driver for chemical alterations in altitudes higher than 90 km and that it contributes in the upper mesosphere in periods of moderate proton fluxes. As the used electron fluxes are strongly contaminated during peak SPE strength, EEP effects could not be studied during these periods.

Whereas SPE proton fluxes are included in most common chemistry climate models (CCMs) nowadays, technical issues like proton contamination complicate the use of satellite measured electron fluxes. Auroral precipitation is usually parametrized based on the Kp index (Matthes et al., 2017), but introducing geomagnetic indices as proxies for EEP is still somewhat debated. An attempt was made in Van de Kamp et al. (2016) which provides both a Dst and an Ap based parametrization for the EEP flux. In fact, the Ap based parametrization will be used as EEP input in the upcoming CMIP6 climate model in order to increase model goodness and concurrence with measurements.

Chapter 3

Methods

For many years, geomagnetic indices have been used as a measure for geomagnetic activity. Despite their many advantages, two of their major drawbacks are the spatial and temporal resolution. Although some local indices exist, the most common indices merely yield information valid on global scales, e.g. the global ring current strength. Details on for example precipitating particles, their energy and associated fluxes are, however, not reflected in the indices. Moreover, possible drivers of geomagnetic activity can vary with storm type and thus, signatures in the indices do not necessarily comply in a uniform manner with the particle precipitation input to the atmosphere. With the beginning of the space age, new possibilities arose as satellites and spacecraft could carry measurement devices into near Earth space in order to carry out in-situ measurements. Any measurements will, however, have their own limitations, e.g. noise thresholds and contamination, which present potential biases for the average data series.

3.1 The MEPED instrument on board the NOAA POES satellites

As part of the aim of this work is to assess the significance of EEP during SPEs, detailed information on the incoming particle fluxes are crucial. Ideally, one should be able to differentiate between different kinds of particles and energy ranges. A dataset which theoretically allows for exactly that is provided by the Medium-Energy Proton and Electron Detector (MEPED) which is mounted on board the National Oceanic and Atmospheric Administration Polar Orbiting Environmental Satellite (NOAA POES) and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) MetOp satellites. In the first part of this section, technical details on both satellite and detector telescopes will be given, followed by a concise description of the working principle of a solid state detector. As the usage of MEPED measurements has faced a number of challenges, methods of data correction and analysis will be presented towards the end of the current section.

3.1.1 Technical details

The MEPED instrument has so far been launched on 14 different satellites throughout the past four decades, whereof 12 were part of the NOAA POES programme and two belong to EUMETSAT MetOp (Ødegaard et al., 2016). Figure 3.1 shows the operational periods of all successfully launched MEPED instruments. The main focus of this work lies on years 2003 to 2008 where up to four NOAA POES spacecraft (15 - 18) and one EUMETSAT MetOp (2) were simultaneously active.

The NOAA POES and EUMETSAT MetOp satellites orbit Earth in an altitude of approximately 800 - 850 km on Sun-synchronous orbits, meaning that they stay in roughly constant LT zones. With a circulation time of approximately 100 minutes, the satellites circle around Earth 14 times per day (Evans and Greer, 2004). The MEPED detector, which is used to measure intensities of charged particle fluxes in the radiation belt regions, is one of the instruments of the so-called Space Environment Monitor (SEM). Other instruments included in the SEM are a Total Energy Detector (TED) and a device for measurements of alpha particles, which was only operating on board early POES satellites (Hill et al., 1985). The newer SEM-2 replaced its predecessor SEM-1 with the launch of NOAA 15 as shown in Figure 3.1. The two versions distinguish themselves with regard to the MEPED instrument mainly through the orientation of certain detectors and an additional proton channel.

The MEPED instrument in turn consists of three different detection systems. Two proton telescopes measure intensities of proton fluxes with energies ranging from 30 keV to

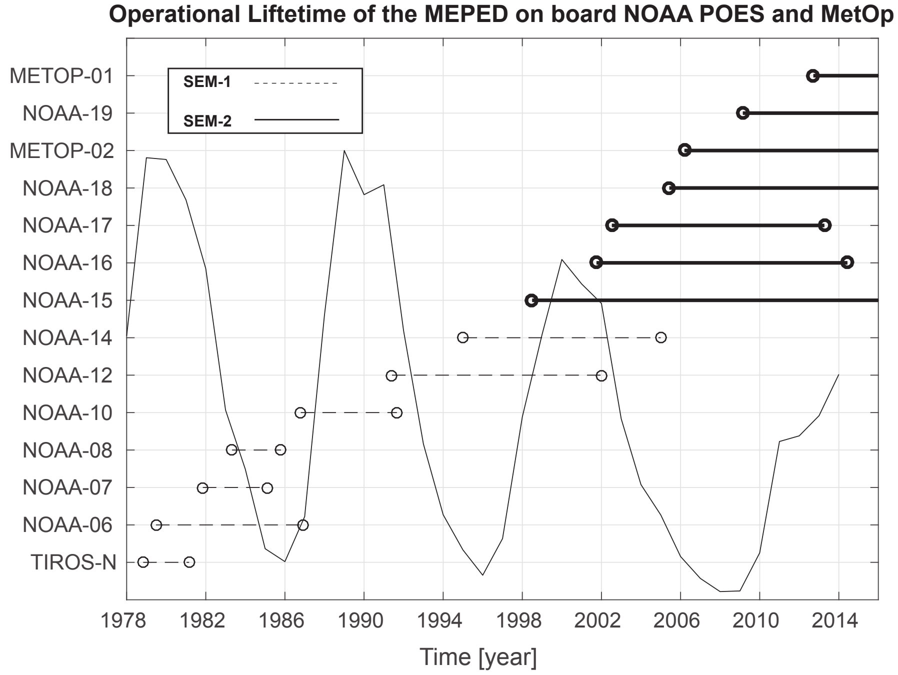


FIGURE 3.1: Figure depicting operational periods of successfully launched MEPED instruments. Dashed lines stand for satellites carrying SEM-1, solid lines indicate SEM-2 satellites. For comparison, the solar cycle is shown as well. From Ødegaard et al. (2016).

6.9 MeV in six differential energy bands. Two electron telescopes account for electron flux intensities covering a nominal energy range from 30 keV to 2.5 MeV in three integral energy channels (Evans and Greer, 2004). The nominal energy ranges of the different electron and proton channels are stated in Table 3.1. The optimized integral energy limits for channels E1, E2 and E3 are noted in brackets and the associated geometric factors found by are stated in the right column. The in total four detectors are arranged in a 0° and a 90° telescope, each consisting of one electron and one proton detector. Orientation-wise, the 0° telescope points radially outward along the Earth-satellite connecting axis. The 90° detectors are mounted perpendicular to the 0° telescopes. In case of SEM-1, the 90° detectors were orientated perpendicular to both the 0° telescope and the satellite's direction of travel (Galand and Evans, 2000). The SEM-2 90° detectors'

Channel	Nominal Energy Range (keV)	Contaminating energy range (keV)	Geometric factor (cm^2 sr)
P1	30 - 80	-	0.01 (0.0095)
P2	80 - 240	> 7000	0.01 (0.0095)
P3	240 - 800	> 7000	0.01 (0.0095)
P4	800 - 2500	-	0.01 (0.0095)
P5	2500 - 6900	-	0.01 (0.0095)
P6	> 6900	> 100	0.01 (0.0095)
E1	> 30 (> 43)	210 (223) - 2700	0.01 (0.0101)
E2	> 100 (> 114)	280 (294) - 2700	0.01 (0.0112)
E3	> 300 (> 292)	440 (432) - 2700	0.01 (0.0088)

TABLE 3.1: Nominal detector responses in the different channels of the SEM-2 MEPED electron and proton detectors (Evans and Greer, 2004). Optimized integral energy limits and geometric factors for channels E1 - E3 are noted in brackets (Ødegaard et al., 2017). Optimized geometric factors for channels P1 - P6 are provided by Yando et al. (2011) and stated in brackets as well. The third column states energy ranges of contaminating particle species, i.e. protons for the electron detectors and vice versa (Yando et al., 2011).

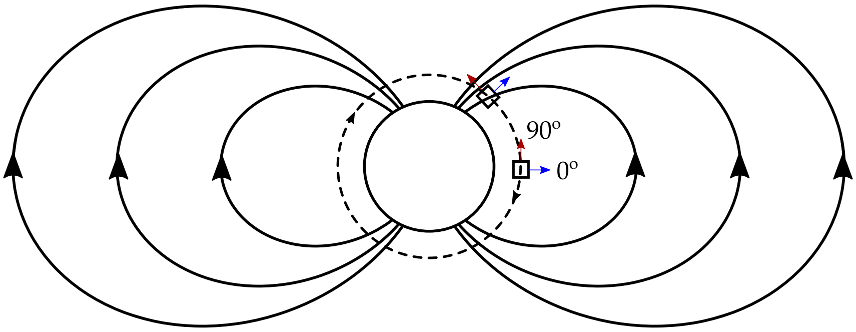


FIGURE 3.2: Sketch visualizing the satellite's position in the geomagnetic field. Field of view centres of the 0° and 90° MEPED telescopes are indicated by blue and red arrows, respectively.

field of view, however, lies anti-parallel to the direction of travel (Evans and Greer, 2004). A rotation of both telescopes by 9° away from the described axes is supposed to ensure a clear field of view (Evans and Greer, 2004). Figure 3.2 visualizes the SEM-2 set-up with respect to the background geomagnetic field. As explained earlier, charged particles gyrate around the magnetic field lines. Therefore, the 0° telescope will measure preferentially atmospheric losscone particles when the satellite crosses high geomagnetic latitudes (Rodger et al., 2010) while the 90° telescope measures high-latitude trapped radiation belt particles and equatorial losscone particles (Galand and Evans, 2000). A third "omni-directional" system measures energetic protons over a wide range of incident angles (Cayton, 2007).

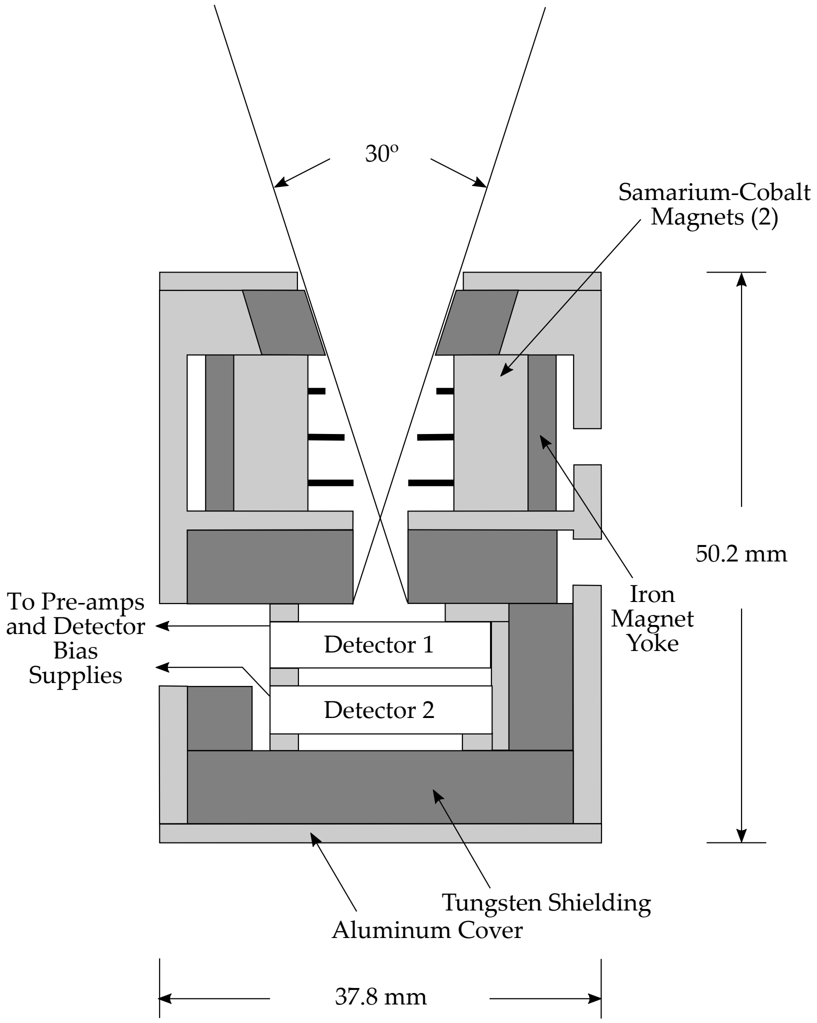


Figure 3.3: Schematic drawing of the SEM-2 MEPED proton detector. Modified from Evans and Greer (2004).

3.1.2 Solid state detectors

Both SEM-2 electron and proton detectors consist of an arrangement of solid-state detectors which is described in detail in Evans and Greer (2004). The basic working principle of a solid-state detector is that of a diode in reverse bias. Voltage is applied in a way that creates a depletion zone and prevents the flow of any current across. Incoming radiation can create electron-hole pairs in the depleted area which then travel to the respective electrodes and create a signal. The amount of created charge carriers is on the one hand dependent on the detector material. On the other hand, the energy of the incoming radiation determines the signal strength as well.
A schematic drawing of the proton telescope is shown in Figure 3.3. A 30° wide collimator restricts the field of view. Magnets in the aperture create a 0.2 T strong magnetic field in order to deflect electrons of energies less than 1 MeV. An aluminum and tungsten coating on the telescope walls prevents electrons with energies less than 6 MeV and

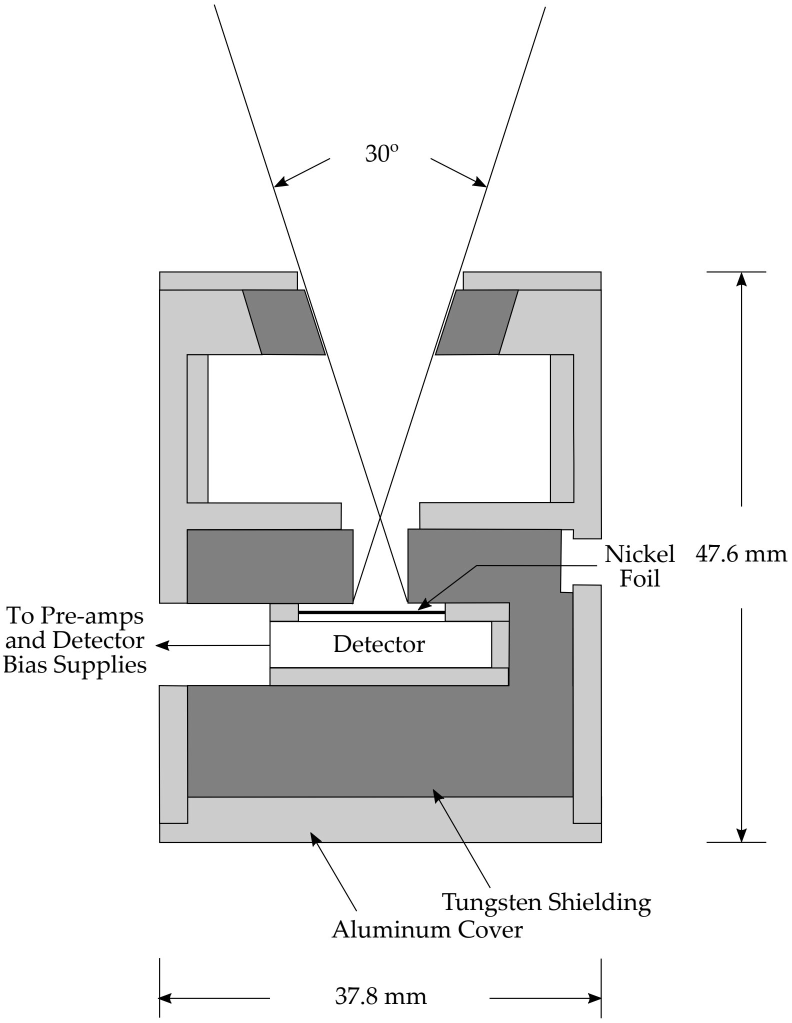


FIGURE 3.4: Schematic drawing of the SEM-2 MEPED electron detector. Modified from Evans and Greer (2004).

protons with energies less than 90 MeV from penetrating the detectors. The proton telescopes contain two 200 microns thick totally depleted silicon surface barrier detectors. Protons which are stopped in the front detector are assigned to one of five differential channels according to an electronic pulse height analysis. When a proton is stopped in the back detector after having passed the front detector it is sorted into a sixth (integral) energy channel. In each telescope, data is accumulated over a period of 1 second but as 0° and 90° telescopes share accumulation electronics, a total data set requires 2 seconds. Via the electronic pulse height, the telescopes determine an energy dependent count rate. In order to convert between count rate C (in particles s^{-1}) and directional particle flux J (in particles $cm^{-2}s^{-1}sr^{-1}$), a geometric factor G is introduced

$$J = \frac{C}{G} \tag{3.1}$$

(Yando et al., 2011). This factor takes into account the telescope's field of view and detector sensitivities. For the proton telescopes, a value of G = 0.0095 $cm^2\,sr$ is used (Yando et al., 2011).

Figure 3.4 depicts a schematic cross-section of the electron telescope. In analogy to the proton telescope, an aluminum and tungsten layer shields the detector against electrons and protons of energies less than 6 and 90 MeV, respectively. The collimator restricts the field of view to 30° and the aperture is covered with a thin nickel foil in order to prevent low energy protons from entering the detector. The telescope contains a 700 microns thick totally depleted silicon surface barrier detector. Pulse height analysis of the created electrical pulses allows for the classification of three integral energy channels. Anti-coincidence with one upper discriminator level limits the measurement population to electrons to below 2500 keV. A full data set containing data from both the 0° and 90° detector requires 2 seconds of accumulated measurement, 1 second per telescope. The nominal geometric factor for the electron telescope is 0.01 $cm^2\,sr$ (Evans and Greer, 2004). Ødegaard et al. (2017) found, however, that the geometric factor is dependent on the incoming energy spectra and optimized values and associated energy limits differ as shown in Table 3.1. Since the electron detectors measure count rates in integral energy channels, responses from E1 must be checked to be greater than responses in E2 and E3 (Evans and Greer, 2004).

3.1.3 Challenges and previous efforts

Research based on POES MEPED data has faced a number of challenges throughout the past decades. Discrepancies in a study from 1984 based on proton data from NOAA 6 hinted on a possible underestimation of fluxes in the 90° detector (Lyons and Evans, 1984). Following studies confirmed decreasing detector responses in both proton detectors due to radiation damage (Galand and Evans, 2000). The degradation proceeds significantly faster in the 90° detector compared to the 0° detector. This is probably due to the higher fluxes measured by the 90° detector, while measuring the trapped particle population. The rate of degradation depends also on the time of satellite launch with respect to the solar cycle (Ødegaard et al., 2016). The damage in the detector originates from the development of a dead layer where incoming particles are slowed down, and are not being measured to their full extent. In addition, created free electrons become less mobile in old detector material so that the likelihood for recombination on their way to the electrodes increases.

Although the nominal lifespan of a POES satellite is three years, most satellites remain operational much longer. NOAA 15, for example, has been operational for 19 years. The challenges with degradation become especially important when satellites remain operational after their nominal lifespan expired. For a long time, proton data from older NOAA satellites were considered untrustworthy and not fit for use. By now, different methods for correction have been developed (Asikainen and Mursula, 2011; Asikainen et al., 2012; Sandanger et al., 2015; Ødegaard et al., 2016). POES data used in this work are corrected for degradation using a method described in detail in Sandanger et al. (2015) and Ødegaard et al. (2016). The underlying principle is to compare measurements of a degraded satellite and a recently launched satellite while the temporal and spatial distances of the measurements are small. The factor by which the proton fluxes differ is fitted with regard to the accumulated flux of the degraded satellite and the Ap index. Knowing the two predictors, it is possible to determine the correction factor α for any POES satellite at any time.

As already mentioned, MEPED consists of a 0° and a 90° telescope, both measuring electron and proton fluxes. Depending on the satellite's location in terms of geomagnetic latitude, a given detector either measures losscone particles, trapped particles or a combination of both. Especially when analysing energy deposition into the atmosphere, it is of interest to know the exact electron losscone flux. A method to reconstruct the losscone flux using both the 0° and the 90° measurements is described in Nesse Tyssøy et al. (2016). The flux distribution inside the losscone is by no means constant but follows a curve which depends on pitch angle diffusion and scattering by wave-particle interaction. The ratio of the measured fluxes in the 0° and 90° detector can be used to determine the correct theoretical distribution function. The geomagnetic field strength of the respective latitude and longitude is provided by the International Geomagnetic Reference Field (IGRF) model. Knowing the resulting losscone width allows for the reconstruction of the losscone flux. Restrictions are given in situations where the two detectors measure the same part of the losscone and in magnetic latitudes greater 75° due to the used magnetic field model and exceeding computational efforts.

Another issue researchers had to face when working with MEPED data is potential cross-contamination in electron and proton detectors (Yando et al., 2011). In addition to detector shielding, the proton telescopes are largely insensitive to medium-energy electrons as the collimator magnets deflect electrons < 750 keV. The integral P6 channel, which normally measures > 6.9 MeV protons, however, shows a non-negligible response to > 750 MeV electrons. This relativistic electron flux can be assigned to a fourth integral electron channel E4 in times when no protons are measured in the P5 channel. These events are selected by an empirical condition presented in Nesse Tyssøy et al. (2016)

$$P6 > 10 \cdot P5 \cdot \Delta P5 \tag{3.2}$$

with the total flux measured in the P6 channels, the flux per keV in the P5 channels and the nominal bandwidth of the P5 channel, ΔP5, which is equal to 4400 keV (Nesse Tyssøy et al., 2016). Whenever the fluxes comply with the condition, the measured P6 flux is assumed to consist exclusively of relativistic electrons.

Also the electron detector is sensitive to contaminating protons in several energy channels. According to Rodger et al. (2010), proton contamination increases with increasing geomagnetic activity. The nominal contaminating energy ranges are stated in Table 3.1 and include 210 to 2700 keV protons. Since these energies are covered by the proton telescopes, it is possible to calculate and subtract contaminating proton fluxes and subtract them from the E1 - E3 fluxes so that only the portions originating from electron counts remains. The measured differential fluxes from the proton telescopes are converted to integral fluxes so that a piecewise cubic hermite interpolating polynomial (PCHIP) interpolation can be applied. PCHIP estimates a third degree polynomial function interpolating between each data point which provides a means to calculate fluxes in arbitrary energy ranges. Thus, the contaminating fluxes for each electron channel E1 - E3 can be determined and subtracted from the measured electron fluxes after accounting for different geometric factors. The measurements from the electron and proton telescopes must, however, be treated as two independent measurements of potentially the same quantity. Therefore, conditions as to when corrected electron fluxes can be considered reliable are vital. So far, a proton contamination of 50 % presented the limiting value and corrected fluxes must not have negative values (Nesse Tyssøy et al., 2016).

To sum up, the usual data evaluation routine consists of the conversion of measured count rates to directional fluxes using geometric factors stated in Table 3.1, followed by the application of α factors in order to correct proton measurements for detector degradation. Afterwards, 32-second averages and contaminating proton fluxes are calculated. Applying the criteria for maximum permitted proton contamination and checking the resulting electron energy spectrum, electron measurements are either corrected or discarded. This means, that after correcting for contaminating protons, the energy spectrum must still fulfil the requirement E1 > E2 > E3. The presence of a relativistic electron signal in P6 is checked and where appropriate registered as an E4 measurement. In cases where corrected electron measurements exist in both the 0° and the 90° detector, losscone fluxes are determined. In order to account for the detector noise level, LC fluxes are discarded whenever the associated corrected 0° electron flux drops below 250 $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$.

3.2 Statistical evaluations

In this work, the criterion for contaminating proton correction is investigated further. As explained above, the electron flux and the proton flux which is subtracted to correct the electron flux are two independent measurements of potentially the same quantity. In order to prevent false electron measurements, the current requirement is that the contaminating proton flux in the electron detector, $\mathrm{J_{pe}}$, does not account for more than 50 % of the total flux measured in the electron detector, $\mathrm{J_{te}}$. In the following, the first subscript letter of fluxes, count rates and geometric factors will refer to the measured species (p for proton, e for electron and t for total), whereas the second letter specifies the detector (p for proton detector and e for electron detector). Due to different geometric factors for measurements in the electron and proton detector, the contaminating proton flux measured in the proton detector, $\mathrm{J_{pp}}$, has to be corrected yielding

$$J_{pe} = \frac{G_{pe}}{G_{ee}} J_{pp}, \tag{3.3}$$

where the geometric factor for protons in the electron detector, $\mathrm{G_{pe}}$, can be assumed to be equal to the geometric factor for protons in the proton detector, $\mathrm{G_{pp}}$ (stated in Table 3.1) (Yando et al., 2011). With this, the current requirement can be written as

$$\frac{G_{pp}}{G_{ee}} J_{pp} < 0.5 \cdot J_{te}. \tag{3.4}$$

This requirement has to be fulfilled simultaneously in each electron channel E1 - E3 using the respective geometric factor, $\mathrm{G_{ee}}$. For obvious reasons, the inequality 3.4 will from now on be referred to as the "50%"-criterion. Drawbacks of this criterion are, that it has no physical basis and it is therefore difficult to interpret its statistical significance. It is also unclear which errors it permits and if its application results in an unequal treatment of different flux intensities. For example, in case of both high electron and proton fluxes during SPEs, the current criterion might reject statistically significant corrected electron fluxes and reduce the amount of data available to study these events.

3.2.1 Basic hypothesis testing

In order to obtain a statistically justified criterion, the statistical nature of the detection process is used in basic hypothesis testing. Given a Gaussian distributed dataset, hypothesis testing presents a means of determining whether the empirical mean or measurement outcome is likely to represent the true mean of the distribution. The comparison of two datasets or data points (defined by means μ_1 and μ_2 and standard deviations σ_1

and σ_2) is possible by examining their difference. Given normally distributed data

$$\begin{aligned} N_1 &\sim N(\mu_1, \sigma_1) \\ N_2 &\sim N(\mu_2, \sigma_2) \\ \implies N_d = N_1 - N_2 &\sim N\left(\mu_1 - \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right) \end{aligned} \tag{3.5}$$

their difference follows a Gaussian distribution as well. If, indeed, N_1 and N_2 are outcomes from identical distributions, the mean difference should be zero as illustrated in Figure 3.5a. Whether or not an obtained positive, non-zero difference, N_d, is significant or not can be determined based on an one-sided hypothesis test (Figure 3.5) with the hypotheses

$$\begin{aligned} \mathrm{H}_0: &\quad \mathrm{N}_\mathrm{d} = 0 \\ \mathrm{H}_a: &\quad \mathrm{N}_\mathrm{d} > 0. \end{aligned} \tag{3.6}$$

The null hypothesis H_0 assumes equality whereas the alternative hypothesis H_a presents the case where $\mu_1 > \mu_2$. Given a certain measurement outcome N_d, the null hypothesis will be rejected in favour of the alternative hypothesis if the measurement lies within the rejection zone (dark grey in Figure 3.5c). The basic principle is to assume that the null hypothesis is true and use the associated probability distribution to calculate a critical value with a set confidence level α

$$z_\mathrm{crit} = \mu_d + z_\alpha \cdot \sigma_d \tag{3.7}$$

where z_α can be obtained from a standard normal table. For a case as it is described by definitions 3.6, the null hypothesis is rejected if the obtained measurement outcome lies above the critical value, i.e.

$$N_d > \mu_d + z_\alpha \cdot \sigma_d. \tag{3.8}$$

A typical value for the confidence level is 5 %. This would correspond to falsely rejecting the null hypothesis in on average 5 % of all cases. This type of error is called type-I or α error in hypothesis testing. Its probability is given by α itself. The type-II or β error corresponds to retaining H_0 although H_a is true. In order to calculate the probability for a type-II error, the true mean value μ_d must be known, which one rarely does in praxis.

3.2.2 Hypothesis tests using count data

During data processing, 16 data points are added in order to calculate 32-second mean values (a measurement is taken every other second). The total number of counts used for averaging has thus been acquired over a measurement time span of 16 seconds. Assuming a true mean 16-second proton count, statistical temporal and spatial fluctuations throughout the measurement cycle are likely. The electron and proton detectors will always measure different selections of the overall proton population described by the same mean count rate (per 16 seconds). Therefore, it is quite unlikely that both detectors would measure the exact same amount of protons throughout one averaging cycle. The desired criterion must therefore meet several demands. If the number of electrons is close to zero, it must secure that falsely registered electron counts are improbable. This could happen if, by chance, the number of protons measured in the electron detector exceeds the number of counts in the proton detector. Subtraction would then result in a non-zero electron count rate although no electrons were present. On the other hand, if electrons are present, the criterion must minimize the chance of falsely rejecting the electron measurement. These demands are met by applying a hypothesis test defined by Equation 3.6.

In the following, the number of counts in the electron and proton detector, which has been accumulated throughout the averaging span of 16 seconds, will be compared. From Appendix A, it becomes clear that performing the statistical analysis based on the 16-second sum or the 16-second average yields identical results. The measured electron and proton detector counts will follow a Poisson distribution according to

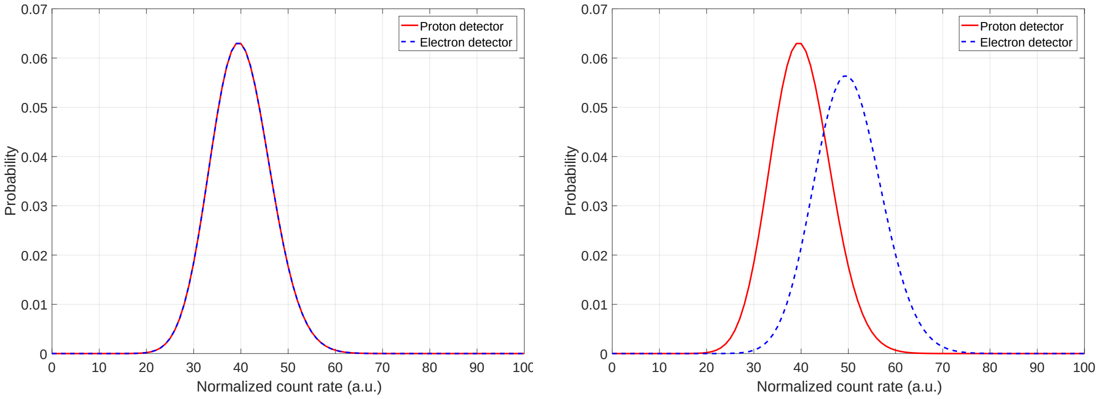
$$\begin{aligned} N_{te} &\sim P\left(\overline{N}_{te}\right) \\ N_{pp} &\sim P\left(\overline{N}_{pp}\right) \end{aligned} \tag{3.9}$$

so that measurements will be scattered around the mean values $\overline{\mathrm{N}}_{\mathrm{te}}$ and $\overline{\mathrm{N}}_{\mathrm{pp}}$ with a standard deviation of $\sqrt{\overline{\mathrm{N}}_{\mathrm{te}}}$ and $\sqrt{\overline{\mathrm{N}}_{\mathrm{pp}}}$, respectively. It should be noted for the record that if N is the sum of 16 independent Poisson distributed measurements (N_{i}, i = 1, ... , 16), N will still follow a Poisson distribution. The same logic applies for the total electron detector count rate which consist of electron and proton counts, each following a Poisson distribution.

As hypothesis testing is based on normal distributed data, a Gaussian approximation of the Poisson distributions is necessary which is legit for mean counts larger than 30. This condition is fulfilled for the fluxes considered in this study. A count rate of 30 counts/16 s would correspond to a flux of approximately 200 $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ which lies around the noise level of the detectors. The Gaussian approximation then yields

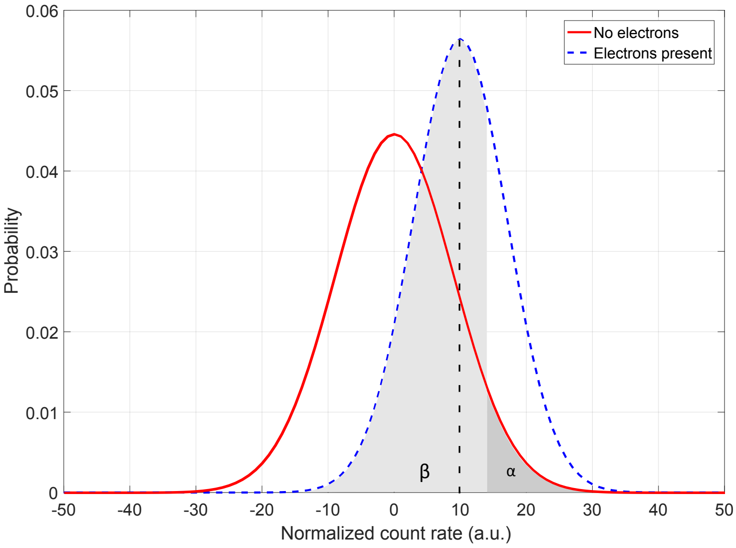
$$\begin{aligned} N_{te} &\sim N\left(\overline{N}_{te}, \sqrt{\overline{N}_{te}}\right) \\ N_{pp} &\sim N\left(\overline{N}_{pp}, \sqrt{\overline{N}_{pp}}\right). \end{aligned} \tag{3.10}$$

An example case is shown in Figure 3.5. Figure 3.5a depicts the case where no electrons are present. Consequently, both electron and proton detector distributions follow the same curve. Here, different measurement outcomes can occur: i) both detectors measure exactly the same number (not necessarily the true proton number) and the subtraction yields the true electron value (zero); ii) by chance, the proton detector measures a larger value than the electron detector, thus rendering a negative electron number (this value is rejected by default); iii) the proton detector measures a smaller value than the electron detector and subtraction yields a positive electron count rate although no electrons exist. A situation where the electron count rate differs from zero is shown in Figure 3.5b. Now,



(a) Electron and proton detector statistics if the true electron count rate is equal to zero.

(b) Electron and proton detector statistics if the true electron count rate is different from zero.



(c) Visualization of the associated hypothesis test.

FIGURE 3.5: Example for electron (blue, dashed) and proton (red, solid) detector statistics for a true proton count $N_p = 40$ and true electron count $N_e = 10$. (A) When the true electron count rate is equal to zero, probability distributions are equal for the electron and proton detectors. (B) With a non-zero electron count rate the electron and proton detector distributions differ. (C) Example for the associated hypothesis test. The case where no electrons are present (A) is depicted by the red line. The null hypothesis is rejected with a confidence level of $\alpha = 5\,\%$ if the actual measured difference between electron and proton detector exceeds a value of 14 as indicated by the rejection zone (α). The type-II (β) error zone is indicated using the case presented in (B) with a non-zero electron count of 10.

the electron detector distribution deviates from the proton detector distribution as it is likely to measure a larger number of counts. Again, negative values can occur after correction. In the remaining cases, subtraction will either yield the exact electron count rate, overestimations or underestimations. In order to judge the statistical significance of the corrected electron flux value via hypothesis testing, the probability distribution yielded by the subtraction must be investigated. As this evaluation is based on the amount of particle counts which is accumulated over 16 s, this averaging time has to be considered when converting between counts and fluxes. Equation 3.1 therefore has to be adapted by adding the time span $\Delta\mathrm{t} = 16\,\mathrm{s}$ yielding

$$J = \frac{N}{\Delta t \cdot G} \tag{3.11}$$

with the measured number of counts N.

Since the only requirement for applying hypothesis testing, i.e. normality of the underlying distributions, is fulfilled, a general condition for the electron detector measurement can be established based on Equation 3.8 which was deduced earlier. The difference of electron detector and proton detector measurements corresponds to the allegedly measured electrons in the electron detector and is therefore denoted by $\mathrm{N_{ee}}$. In order to statistically assess whether the obtained value for $\mathrm{N_{ee}}$ is trustworthy, the null hypothesis claims a zero electron count. Thus, $\mathrm{N_{ee}}$ is assumed to follow a Gaussian distribution

$$N_{ee} = N_{te} - N_{pp} \sim N\left(0, \sigma_{ee}\right). \tag{3.12}$$

In order to obtain an expression for the standard deviation σ_{ee}, the following assumptions have to be made:

$$\begin{align} \sigma_{ee} &\approx \sqrt{N_{te} + N_{pp}} \tag{3.13} \\ &\approx \sqrt{2N_{pp}} \tag{3.14} \end{align}$$

where the necessity of the second approximation will be elaborated later-on. Based on Equations 3.7 and 3.14, it is possible to formulate a statistical condition for the retention of the electron detector measurement $\mathrm{N_{ee}}$. The condition for rejecting the null hypothesis and considering the electron measurement trustworthy can be written as

$$\begin{aligned} N_{ee} > {} & 0 && + && z_\alpha \cdot \sqrt{2N_{pp}} \\ \Longleftrightarrow N_{te} - N_{pp} > {} & && && z_\alpha \cdot \sqrt{2N_{pp}} \\ \Longleftrightarrow \qquad N_{te} > {} & N_{pp} && + && \sqrt{2} z_\alpha \cdot \sqrt{N_{pp}}. \end{aligned} \tag{3.15}$$

Thus, the second approximation in Equation 3.14 yields a means of judging the electron detector measurement solely based on the simultaneous proton detector measurement and its associated statistical error. The confidence level will be determined by the chosen z_α. This would mean that in $\alpha\,\%$ of all measurements an electron measurement would be assumed even if there were no electrons present whatsoever. It is instructive to merge $\sqrt{2}$ and z_α in order to introduce the constant c_σ which contains information on the confidence level and states, how many σ the electron measurement must diverge from the proton measurement.

A numerical example is shown in Figure 3.5c, where the true electron count rate lies at 10 counts, corresponding to the situation depicted in Figure 3.5b. Given the normalized measurement outcome of the accumulated 16 s electron and proton count rate, N_{te} and N_{pp}, respectively, it is instructive to assume the measurements coincide with the true mean values of 50 and 40 counts, respectively. Hence, regarding the difference, the true electron count rate would be measured. The next step is to calculate the corresponding critical value for a given α of 5 %, which leads to a z_α of 1.64. The red line in Figure 3.5c indicates the probability distribution belonging to the null hypothesis. Its mean lies at zero ($\mu_{\mathrm{d}} = 0$) and Equation 3.14 is used in order to find its standard deviation. As N_{te} and N_{pp} are only measurement outcomes and not true means, the true standard deviations are unknown. Further calculations yield a critical value of 14.7 counts/16 s. Thus, although the true count rate is measured, the null hypothesis would not be rejected based on a 5 % confidence level, as 10 lies not within the rejection zone. The probability for this type-II (β) error can be determined by calculating the probability for obtaining an electron count N_{ee} which lies below the critical value based on the true electron probability distribution. This corresponds to the light grey area indicated in Figure 3.5c. The true mean value and associated probability distribution is in praxis, however, mostly unknown. The β error has therefore no role in the derivation of new criteria which is solely based on an appropriate α error. In order to determine how large the β error is for different flux levels, simulations where the true mean value is known are required.

3.2.3 Simulated datasets

Simulated datasets contain random data generated from Poisson distributions stated in Equation 3.9. If not declared otherwise, the datasets have a sample size of ten million data points. Generally, different true electron and proton count rates (per 16 seconds) can be chosen. In cases where the electron count rate is supposed to be equal to zero, two random datasets are generated from the same mean proton count rate, one resembling the proton detector measurement and the other relating to the measured counts in the electron detector. Whenever the true electron count rate differs from zero, the dataset resembling the electron detector measurements consists of the sum of random data from an electron and a proton Poisson distribution. As explained earlier, the sum of the two Poisson distributed datasets will form a quantity which is in turn following a Poisson distribution. Rejection rates and α errors can be obtained by comparing the electron and proton datasets while applying the respective criterion.

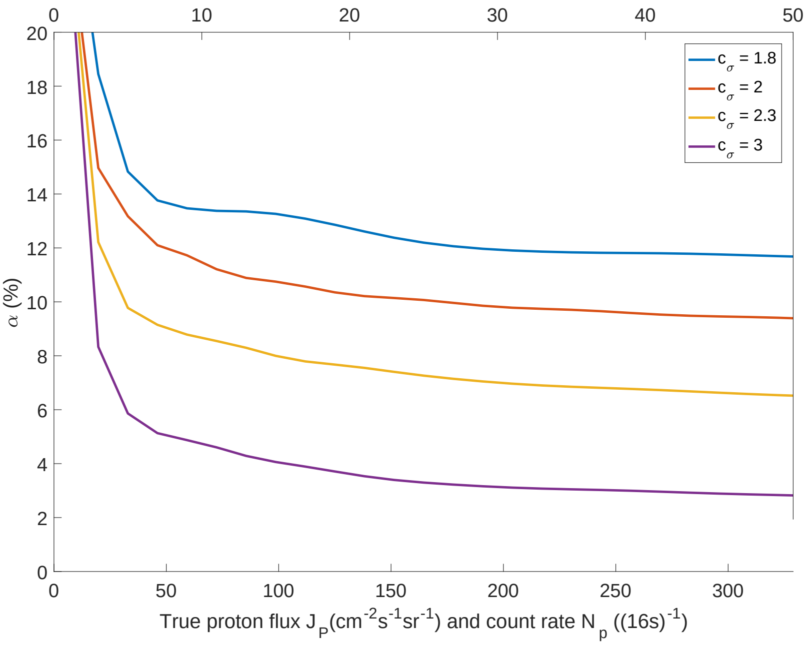


FIGURE 3.6: Calculated behaviour of the type I error probability (α) with respect to the simulated true proton number for different constants c_σ. The lower x-axis gives the proton flux while the upper x-axis states the corresponding count rate per 16 seconds. Conversion follows from Equation 3.11 and the geometric factor of the proton detector from Table 3.1.

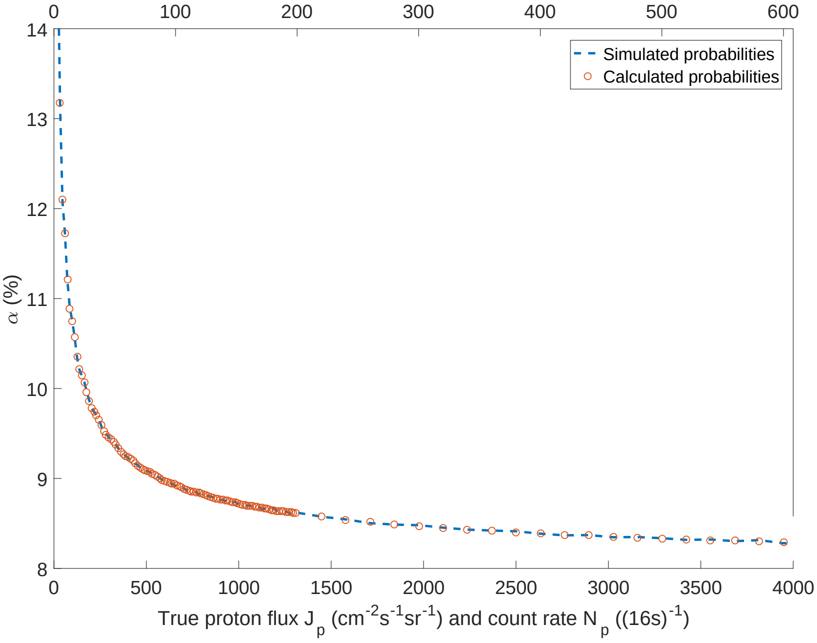


FIGURE 3.7: Evolution of the simulated (dashed line) and calculated (circles) α error for the 2σ-criterion. The upper x-axis states count rates (per 16 seconds) and the lower x-axis shows the corresponding proton fluxes.

3.2.4 Evaluating different σ-constants c_σ

The final step for obtaining new criteria is to determine which σ-constants, c_σ, should be considered. Figure 3.6 shows a case were the electron count rate was assumed to be zero under data simulation. Thus, the simulated electron and proton detector measurements stem from the same Poisson distribution. Comparing two data points at a time while applying different c_σ yields a specific rejection rate of the null hypothesis, i.e. the α error, for a given true proton count rate. Figure 3.6 depicts how the α error evolves with increasing simulated true proton flux for different values of c_σ. The lower x-axis states the true proton flux while the upper x-axis shows the associated simulated proton count rate per 16 seconds. Equation 3.11 and the geometric factor for the proton detector stated in Table 3.1 were used for conversion.

As can be seen in Figure 3.6, the error probability is quite high for low true count rates due to the inappropriateness of the Gaussian approximation and the standard deviation approximation in Equation 3.14. Both assumptions, however, become more accurate with increasing count rates. This results in a convergence of the error probability towards the theoretical value already at flux values corresponding to the detector noise level ($\sim 250\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$). The depicted values for c_σ (1.8, 2, 2.32 and 3) were chosen such that they correspond to confidence levels, i.e. minimal error probabilities, of 10, 8, 5 and 1.7 % for the proton counts, respectively. For a given true count rate N, the probability of an occurring α error was calculated with

$$\alpha = \sum_{k=0}^{\infty} P(k;N) \cdot (1 - F(k + c_\sigma \cdot k; N)) \tag{3.16}$$

with the Poisson probability function $\mathrm{P(k;N)}$ and the cumulative probability function $\mathrm{F(k;N)}$ for the outcome k.

Evaluating the curves of the different c_σ values, it is visible that the "1.8σ"-curve stays above the limiting value of 10 %. The corresponding criterion is therefore considered not strict enough. The "3σ"-curve obviously corresponds to a quite strict condition and the "2σ"-curve falls below 10 % already at count levels that classify as noise ($\sim 250\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$), more specifically at $\mathrm{N_{pp}} \geq 27$. Figure 3.7 shows how the 2σ error approaches levels as low as 8.5 % for only moderate fluxes which makes the "2σ"-criterion a promising candidate.

In the following Chapter 4, the 2σ- and 3σ-criterion are analysed and compared to their predecessor, the 50%-criterion introduced in Equation 3.4. A theoretical analysis based on the equations themselves and simulated data will show, which flux regions might be favoured by a certain condition and which might be treated more strictly. Later-on, the three criteria will be used during the complete POES data evaluation routine and it will be possible to examine how the different criteria influence the actual data rejection.

3.3 EEP parametrization used in CMIP6

POES MEPED data which is generated applying the new criteria will also be compared to datasets which result from an alternative approach. The newest chemistry-climate model CMIP6 will - besides radiative solar forcing - also include particle forcing to account for atmospheric ionization rates (Matthes et al., 2017). Instead of using POES data directly as input for EEP, electron flux data will be provided by a parametrization based on the Ap index. This model is described in detail in Van de Kamp et al. (2016). An alternative parametrization which is based on the Dst index is presented as well. The general principle of the model is to use modified POES electron data acquired in the period 2002 - 2012 to fit equations for the integral F_{30} (> 30 keV) electron flux and the spectral power law gradient k. The acquired 11-year dataset of experimental values is fitted using either Ap or Dst as predictor. Since a simple power law is assumed to yield the spectral density (Van de Kamp et al., 2016, Equation (1)), knowing both k and F_{30} allows for the determination of any desired integral or differential electron flux.

The modified POES dataset is based on electron data from the MEPED 0° detector which measures parts of the losscone while the satellite is located in the auroral zone. Under the assumption of a uniform losscone flux distribution, 0° data can be treated as losscone data. This assumption is known to work well during strong geomagnetic activity while underestimated fluxes are likely during moderate or low disturbances (Rodger et al., 2013). Counts are converted to fluxes based on geometric factors stated in Evans and Greer (2004). Data is corrected for proton contamination based on a method described in Lam et al. (2010) and entries which are likely to be contaminated by protons during SPEs are removed from the dataset. Note, however, that the proton measurements used for correction are not corrected for detector degradation.

The remaining flux data is binned with respect to their L value including L= 2 - 10 with a resolution of 0.5 in 3-hour UT intervals. There is no distinction between different MLT sectors. Median electron fluxes are determined for each bin and linearly averaged in order to calculate daily fluxes. This is done for each energy channel E1, E2 and E3. Proceeding from here, data points from the three channels are replaces by zero whenever the respective E1 electron flux lies below the noise level, which is identified to be 250 electrons $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$. The remaining data points can be used to fit a power law spectral function to the three integral electron fluxes for each day and L bin. Based on the obtained spectral gradient k, the F_{30} flux can be calculated. Data points with unphysical spectral gradients are removed in the process as well. Thus, the modified POES data consists of the "experimental" values for k and F_{30} for each day and L bin, covering a total of 11 years.

In the next step, two different models using either Ap or Dst as predictor are fitted to the obtained k and F_{30} data points by a simple least squares fitting routine. The fitting functions describe k and F_{30} entirely based on the respective daily geomagnetic index and the chosen L shell. The resulting equations and fit parameters which link

F_{30} and k to the Dst index are given in Equation (6) and (7), respectively in Van de Kamp et al. (2016). The model equations based on the Ap index are presented in Equations (8) and (9) in Van de Kamp et al. (2016) and given below:

$$F_{30} = \frac{\exp(A)}{\exp(-b(S_{pp} - s)) + \exp(c(S_{pp} - s)) + d}, \tag{3.17}$$

with

$$\begin{aligned}
A &= 8.2091\mathrm{Ap}^{0.16255} \\
b &= 1.3754\mathrm{Ap}^{0.33042} \\
c &= 0.13334\mathrm{Ap}^{0.42616} \\
s &= 2.2833\mathrm{Ap}^{-0.2299} \\
d &= 2.7563 \cdot 10^{-4}\mathrm{Ap}^{2.6116} \\
S_{pp} &= L - L_{pp} \\
L_{pp}(t) &= -0.743\ln(\max_{t-1,t}\mathrm{Ap}) + 6.5257
\end{aligned}$$

and

$$k = \frac{-1}{E\exp(-bS_{pp}) + 0.3045\cosh(0.20098(S_{pp} - s))} - 1, \tag{3.18}$$

with

$$\begin{aligned}
E &= 3.3777\mathrm{Ap}^{-1.7038} + 0.15 \\
b &= 3.7632\mathrm{Ap}^{-0.16034} \\
s &= 12.184\mathrm{Ap}^{-0.30111}
\end{aligned}$$

and S_{pp} as defined in Equation 3.17. It is important to mention that the Dst model is only defined for negative Dst values and yields zero electron flux otherwise. General advantages of the Ap- and Dst-based models are their long temporal coverage of up to one hundred years. Ambiguous situations, e.g. during SPEs when proton contamination of the original data is likely, are bypassed by relying on geomagnetic indices.

3.4 Datasources

Data of geomagnetic indices is general openly accessible and provided by different institutions. 3-hour Kp and Ap data is available on the NOAA web page (`ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/KP_AP/`) where it is also possible to obtain data on SPE occurrences and strengths (`ftp://ftp.swpc.noaa.gov/pub/indices/SPE.txt`) as well as the monthly F10.7 radio flux (`https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_observed/listings/listing_drao_noontime-flux-observed_monthly.txt`) which is used as a proxy for the solar cycle. Hourly Dst values are provided by the World Data System for Geomagnetism in Kyoto (`http://wdc.kugi.kyoto-u.ac.jp/dst_final/index.html`).

3.5 Reference model for auroral oval positions

After applying the various criteria throughout the data processing routine, the overall data rejection rate (RR) is likely to vary with different latitudes, MLTs and time periods. Especially during SPEs, temporarily and locally high rejection rates can be expected. Obviously, this influences the data coverage throughout these events and might result in gaps whenever the rejection rates grow too high. In order to obtain information on whether and to which extent POES data is available during SPEs in regions of elevated electron precipitation, it is important to estimate the auroral oval position and width. In periods of high rejections rates cases might occur where the dataset does not cover the entire oval and then it is beneficial to assess whether gaps are present close to the maximum or rather in regions of low flux intensity. The model used to obtain the electron flux intensities in the auroral oval is based on the Kp index and it provides global energy flux levels for any desired MLT or magnetic latitude. A detailed description can be found in Zhang and Paxton (2008). Four years (2002 - 2005) of auroral far ultraviolet (FUV) data measured by the Global Ultraviolet Imager (GUVI) on board the Thermosphere Ionosphere Mesosphere Energetics and Dynamics satellite (TIMED) were used to determine the mean energy and flux of precipitating electrons. A clear advantage with the chosen time period is that it contains a wide range of storm types, including strong SPEs. The TIMED satellite is located at an altitude of 630 km and moves on a circular orbit that covers all local time sectors with a period of 60 days. The available flux and energy data was then binned according to the concurrent Kp index into six bins (0 - 1.5, 1.5 - 3.0, 3.0 - 4.5, 4.5 - 6.0, 6.0 - 8.0, 8.0 - 10.0), where the authors

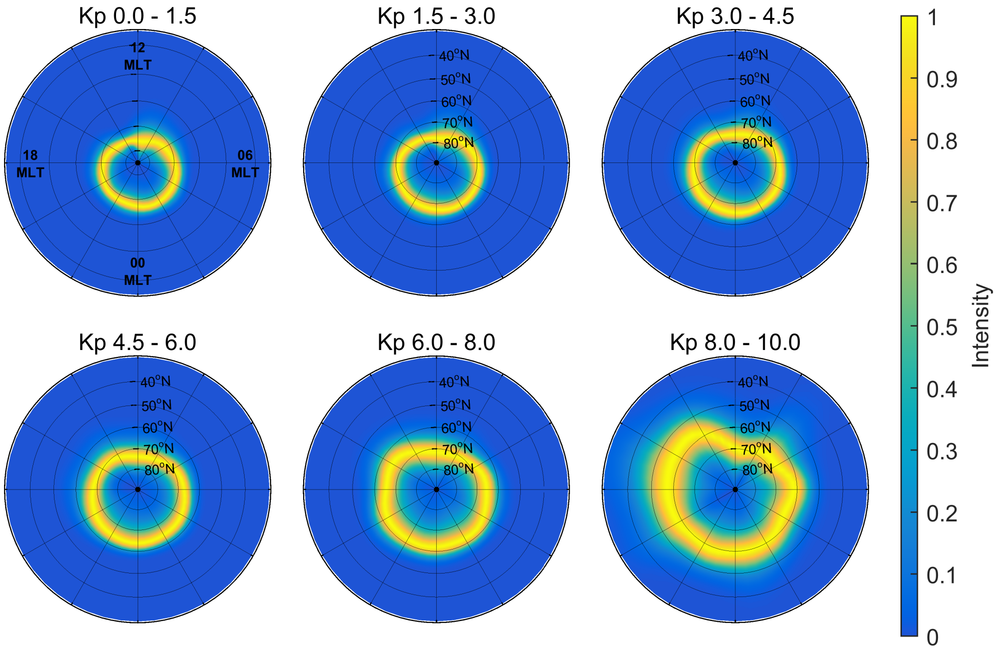


FIGURE 3.8: The auroral oval positions as predicted by the model in the six Kp categories.

assure a statistically significant amount of data in each MLT and magnetic latitude bin. In order to account for different equatorward and poleward slopes in the auroral oval, an Epstein function was chosen to represent the magnetic latitude dependence

$$F(x) = \frac{A \cdot \exp\left(\frac{x-B}{C}\right)}{\left(1 + \exp\left(\frac{x-B}{D}\right)\right)^2} \tag{3.19}$$

with the magnetic co-latitude x.

MLT dependent fitting parameters (A, B, C, D) where introduced as sixth order Fourier series containing sine and cosine functions with the MLT value (in hours) as argument

$$A(\mathrm{MLT}) = \sum_{n=0}^{6} \left(F_n \cdot \cos\left(n\pi \frac{\mathrm{MLT}}{12}\right) + G_n \cdot \sin\left(n\pi \frac{\mathrm{MLT}}{12}\right)\right). \tag{3.20}$$

The Fourier coefficients are Kp-dependent and provided by the publication (Zhang and Paxton, 2008).

Figure 3.8 depicts the different auroral oval flux intensities in the six Kp categories as an interpolated colour-coded map with a longitudinal and latitudinal resolution of 7.5 and 0.5°, respectively. The intensity is shown relative to the maximum flux in the respective meridional band, as day-night flux level asymmetries are of no concern in this application (night-side fluxes tend to exceed day-side fluxes by up to one order of magnitude during high Kp).

Chapter 4

Results and Discussion

So far, two new criteria for testing the statistical significance of corrected electron fluxes have been derived. In the following analysis, these new criteria are evaluated in terms of α errors (assuming false electron fluxes), β errors (rejecting a true electron measurement) and data rejection rates. Ideally, rejection rates should be minimized while sustaining a certain degree of data significance in order to prevent false data points. The criterion which turns out to deliver the best performance will be implemented in the data processing routine in order to obtain an improved losscone dataset.

In the next part, this LC dataset will be applied in a comparison to the proposed CMIP6 EEP parametrization which is based entirely on the Ap index as predictor for electron fluxes. The flux dataset yielded by the Ap model will be evaluated in terms of general flux strength, temporal and spatial variability and its aptitude for reproducing specific storm events will be tested.

The basic assumption upon which proxy-based models for electron fluxes rely is that the chosen proxy is equally indicative for fluxes during all types of geomagnetic events. In the case of the Ap model, it is assumed that a specific Ap value is connected to the same flux levels during an SPE and a regular, non-SPE related storm. In order to judge the adequacy of this assumption, the applicability of MEPED LC data during SPEs and associated high contaminating proton fluxes is tested. In regions which qualify as reliable, SPE fluxes are compared to equivalent non-SPE storms.

4.1 Evaluation of the new proton contamination criteria

In Chapter 3, statistically based criteria for corrected electron fluxes were derived. As the proton flux measured in one detector is subtracted from the contaminated electron flux of a different detector, statistical fluctuations throughout the measurement process may lead to varying statistical significance of the corrected fluxes. For example, in some cases higher proton fluxes might by chance be measured in the proton detector leading to an underestimation of the corrected electron flux in the electron detector after subtraction. As excessive data rejection leads to issues regarding data coverage and increasing biases, rejection should be minimized while sustaining a sufficient level of confidence for corrected data. In the following, different criteria will be compared theoretically by applying them to simulated datasets in order to assess their α and β errors. The α error corresponds to falsely assuming electrons where none are present while the β error corresponds to falsely rejecting real electron fluxes. Afterwards, the investigated criteria will be applied to real MEPED datasets from 2003 and their effects on data rejection throughout different time periods will be discussed.

4.1.1 Theoretical comparison of the proton contamination criteria

In the previous chapter, the 50%-criterion as well as the 2σ- and 3σ-criterion were introduced as means to compare contaminating proton count rates and the total count

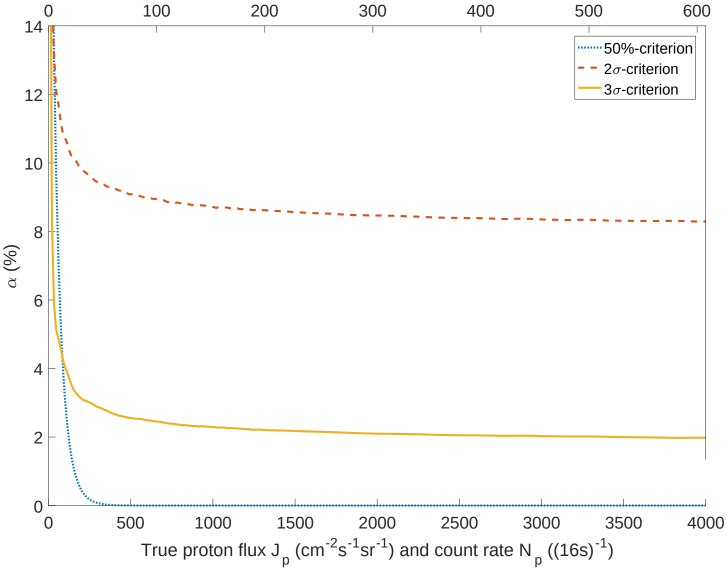


FIGURE 4.1: Simulated alpha errors of the three conditions (50%, 2σ and 3σ) for different true proton fluxes (lower x-axis) and count rates (upper x-axis).

rate measured in the electron detector. The respective criteria based on counts

$$N_{te} > 2 \cdot N_{pp} \tag{4.1}$$

$$N_{te} > N_{pp} + 2 \cdot \sqrt{N_{pp}} \tag{4.2}$$

$$N_{te} > N_{pp} + 3 \cdot \sqrt{N_{pp}} \tag{4.3}$$

can be converted into fluxes using Equation 3.11 which yields

$$J_{ee} > \frac{G_{pp}}{G_{ee}} J_{pp} \qquad \text{(50\%-criterion)} \tag{4.4}$$

$$J_{ee} > 2 \cdot \sqrt{\frac{G_{pp} J_{pp}}{16 G_{ee}^2}} \qquad (2\sigma\text{-criterion}) \tag{4.5}$$

$$J_{ee} > 3 \cdot \sqrt{\frac{G_{pp} J_{pp}}{16 G_{ee}^2}} \qquad (3\sigma\text{-criterion}) \tag{4.6}$$

α errors: keeping false electron measurements

Figure 4.1 depicts a simulation of how the α error (Equation 3.16) of these three criteria evolves with increasing true proton fluxes. Two Poisson distributed datasets are generated based on the true proton fluxes shown on the x-axis and the three conditions specified in Equations 4.4 - 4.6 are applied in order to find out how often a corrected electron measurement is discarded or kept. As both datasets are generated entirely from the proton fluxes with no electrons present, the number of kept data points corresponds to falsely registered electrons and thus to the α error of the respective condition.

For very low fluxes both the 2σ- and 3σ-criterion are stricter than the 50%-criterion. However, this changes for higher fluxes as the 50%-criterion approaches an α error of 0 % already for low to moderate fluxes, whereas the other criteria settle at their respective limits (approximately 8 and 2 %, respectively). Consequently, the two new criteria would be more likely to overestimate the electron flux by considering electron measurements reliable when no electrons could be measured.

In order to investigate this potential overestimation of electron fluxes and possibly set a limit to it, the characteristic normal distribution of the measured difference between electron and proton detector is exploited. The dashed lines in Figure 4.2 link the true proton flux on the x-axis to a limit for the falsely measured corrected electron flux (on the y-axis). The red dashed line depicts the border which only 5 % of all falsely registered electron fluxes surpasses. This means that with a certainty of 95 %, a wrongly registered electron measurement will lie below this line. Analogously, the blue dashed line depicts the 0.3 % border, meaning that 99.7 % of all false electron fluxes will lie below this limit. In addition, Figure 4.2 shows the limits set by the 2σ- and 3σ-condition in yellow and violet, respectively. By definition, all measurements located below the respective lines are discarded. The false fluxes which are not rejected by the conditions correspond to the α errors shown in Figure 4.1 and are located between the respective solid line and the blue dashed line. As an example, given a true proton flux of $2 \cdot 10^4$ cm^{-2}s^{-1}sr^{-1} there is an 8 % chance that electron fluxes between 600 and 1400 cm^{-2}s^{-1}sr^{-1} are falsely registered when using the 2σ-criterion. The probability of measuring an electron flux > 1400 cm^{-2}s^{-1}sr^{-1}, however, is approximately equal to zero.

β errors: rejecting real electron fluxes

Whereas relying on false electron fluxes leads to an overestimation of the true electron flux, rejection of real data results in an underestimation, which is related to the β error. For Figure 4.3, one dataset corresponding to the proton detector measurement was generated based on a true proton count number. A second dataset which is meant to resemble the electron detector measurements consists of two sub-datasets, one generated with the true proton count number and the other based on the true electron count number. The actual simulated measurement of the electron detector consists of the point-wise sum of the two sub-datasets. For each combination of true electron and proton count number shown in Figure 4.3 the rejection rate was determined by comparing the two datasets while applying the different rejection criteria. Note that in this case, the rejection rate corresponds to the β error. Subsequently, count rates were converted to particle fluxes using the geometric factors stated in Table 3.1 (channel E1 was used for the electron fluxes). The resulting β errors for each true electron and proton flux are shown in Figure 4.3 for every criterion in a separate panel. Note that the y-axes depict mere the true electron flux without the portion of measured protons (for measurements, this would correspond to the corrected electron fluxes). Data generated from the true

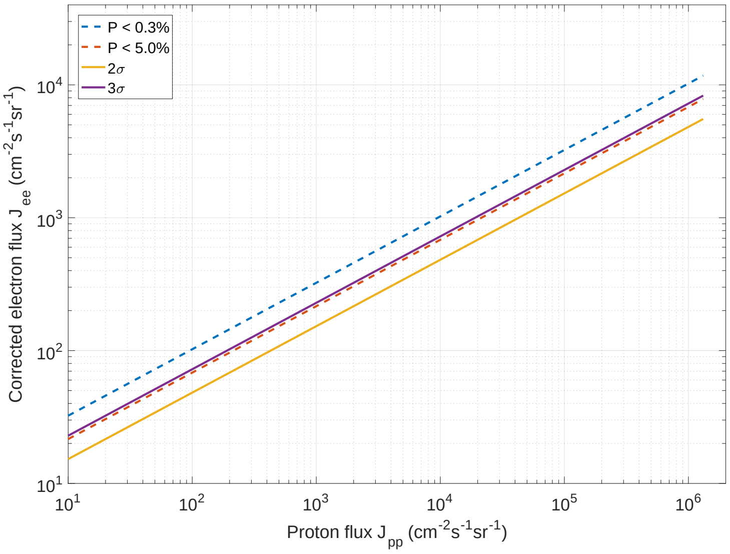


Figure 4.2: Approximated limits to overestimation based on a 5 and 0.3 % confidence. respectively (dashed lines). Limits as defined by the two new criteria (Equations 4.5 and 4.6) are indicated as well (solid lines).

values will have a certain spread, and thus it might happen that some data points are rejected even if the true values themselves would fulfil the condition and not be discarded. When comparing the three plots in Figure 4.3, the most prominent difference can be found for high proton fluxes. Here, the 50%-criterion is very strict and rejects electron fluxes of up to 10^6 cm^{-2}s^{-1}sr^{-1} quite distinctly. In reality, these high proton fluxes would be related to SPEs, meaning that in these cases simultaneous electron fluxes would have to be extremely high in order to be considered trustworthy. The less strict 2σ- and 3σ-criterion exhibit an ambivalent (cyan) area where measurements originating from these particular flux combinations are both rejected and kept, depending on the statistical outcome of the measurement. Even for very high proton fluxes of up to 10^6 cm^{-2}s^{-1}sr^{-1} most measurements of electron fluxes larger than 10^4 cm^{-2}s^{-1}sr^{-1} fulfil the conditions and are sustained. Therefore, the new criteria present a clear advantage with respect to keeping reliable electron measurements during e.g. SPEs. From mere visual inspection, the 2σ- and 3σ-criteria differ not considerably in the fluxes they reject but more in their severity. Whereas the 3σ-criterion rejects low electron fluxes quite strictly as indicated by the bright yellow area in the lower plot of Figure 4.3, the corresponding plot of the 2σ-criterion exhibits a darker yellow and lower rejection rates. In conclusion, the rejection of low electron fluxes is enhanced independently from the applied proton contamination criteron compared to high flux levels, resulting in a probable bias towards overestimated fluxes. In order to assess how strong this bias in praxis really is, the actual particle populations which are represented in the data have to be taken into consideration. In addition, it has to be considered whether ignoring rejected

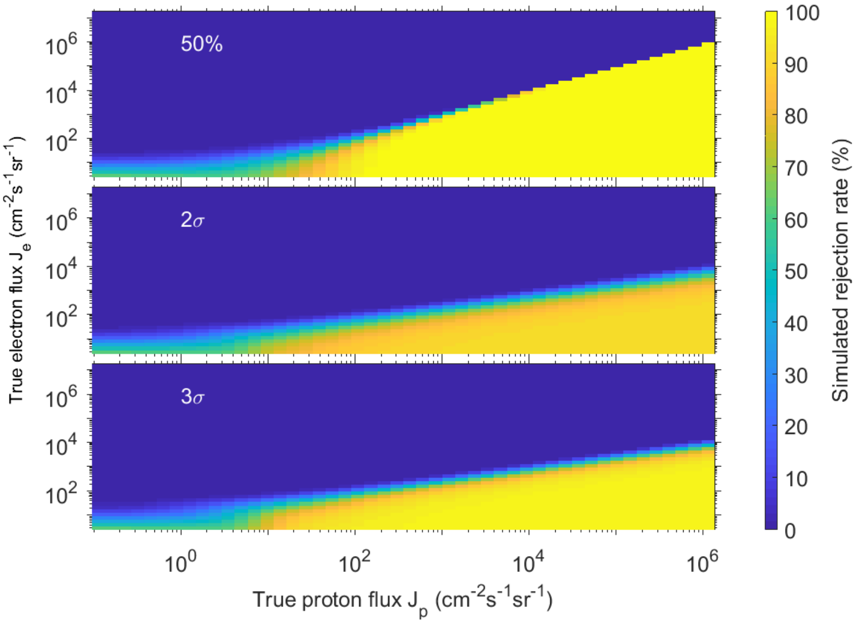


FIGURE 4.3: Simulated β error (colour-coded) of the three conditions (50%, 2σ and 3σ) for different true proton and electron fluxes (on x- and y-axis, respectively). For conversion between electron count rate and flux, the E1 geometric factor was assumed.

data or setting these values to zero reflects real fluxes levels best.

Applying the different criteria during the data processing routine will yield quantitative results on whether data is gained by introducing new criteria and if there are any temporal, latitudinal or longitudinal aspects which have to be considered.

4.1.2 Impact of the different criteria on POES MEPED data

The three different proton contamination criteria were tested on POES MEPED data from satellites NOAA 15 - 17 from the year 2003. A strong SPE in October/November of this particular year secures the presence of a variety of different proton flux strengths. The upper panel of Figure 4.4 shows the distribution of NOAA 15 0° detector E1 data during DOY (Day of Year) 302 - 304 in 2003 with respect to different contaminating proton and uncorrected electron fluxes. Clearly, two populations of protons become visible. The left population is characterized by unphysically low fluxes ($<$ 1 proton/16s). The origin of this proton population is explained in Appendix B. The right population is real and will be considered further here.

The two lower panels inspect how often data points are rejected in each bin using different criteria. Data in a dark blue bin is not rejected and the corresponding electron flux is considered trustworthy whereas a yellow bin indicates the rejection of all contained data points. In a transition area between the blue and yellow regions, rejection rates between 0 and 100 % occur. In all three panels, the black dotted line indicates equal uncorrected electron and contaminating proton fluxes. This means, all data points located below this line correspond to negative corrected electron fluxes and are rejected by default, independently from the used criterion in Equations 4.4 - 4.6. In accordance to the blue dashed line in Figure 4.2, the upper red dashed lines in Figure 4.4 show where the probability of an α error, i.e. falsely registering electron fluxes where no electrons are present, lies at approximately 0 %. The mid-panel depicts rejection rates based on the 50%-criterion and the lower panel presents rejection rates based on the 2σ-criterion. As can be seen, the former is less strict for lower fluxes where blue bins are located well below the upper red dashed line. Moving to higher fluxes it becomes stricter to the point where many yellow bins lie above the upper red line. Here, many data points are rejected although they should be considered trustworthy based on statistical estimations. The junction between rejected and sustained data points in the lower panel, however, orientates itself quite closely along the upper red line rejecting statistically doubtful electron measurements while ensuring that reliable measurements are kept. The 3σ-criterion is not included in this figure due to its similarity to the 2σ-criterion.

A critical remark must be made, however, with regard to the complete validity of the statistical considerations. Besides the made approximations during the deduction of the new criteria, the nature of the measurement error itself must be questioned. All calculations are based on the assumption that differences between measurements in the electron and proton detector arise exclusively due to statistical fluctuations. If this was indeed the case, the rejected data points located below the dotted black line in Figure 4.4 should not exhibit as large a spread as they do. The lower red dashed line illustrates a lower statistical limit for all obtainable electron fluxes. It must be emphasized that some data points lie below that limit indicating additional systematic aspects which are not covered by the deduced criteria. Possible origins of this potential bias are discussed in Appendix C.

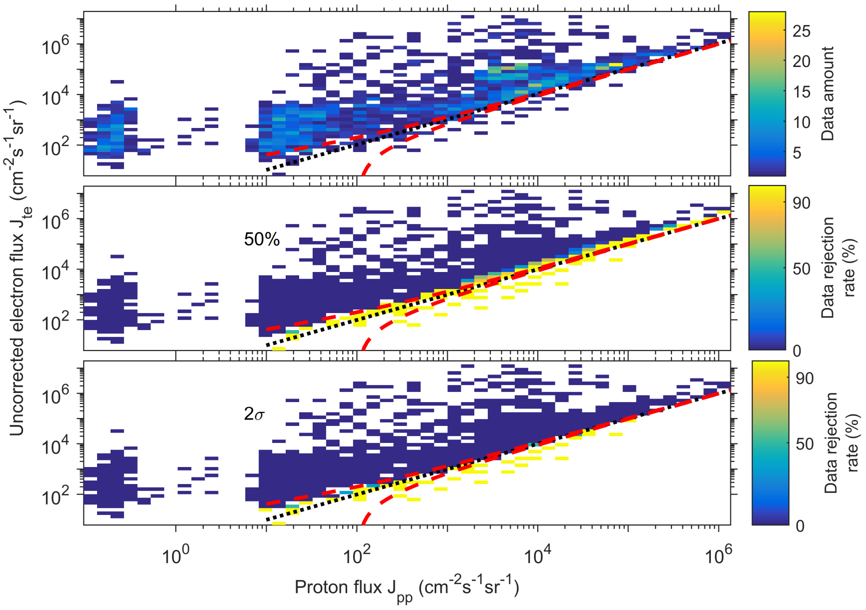


FIGURE 4.4: Data amount and rejection rates using the 50%- and 2σ-criterion (colour-coded) of different uncorrected electron and contaminating proton fluxes (on y- and x-axis, respectively). The plot is based on data from the E1 channel of the 0° detector on board NOAA 15 and covers DOY 302 - 304 in 2003. In addition, lines marking equal fluxes (black dashed) and the 3σ-confidence level (red dashed) are depicted.

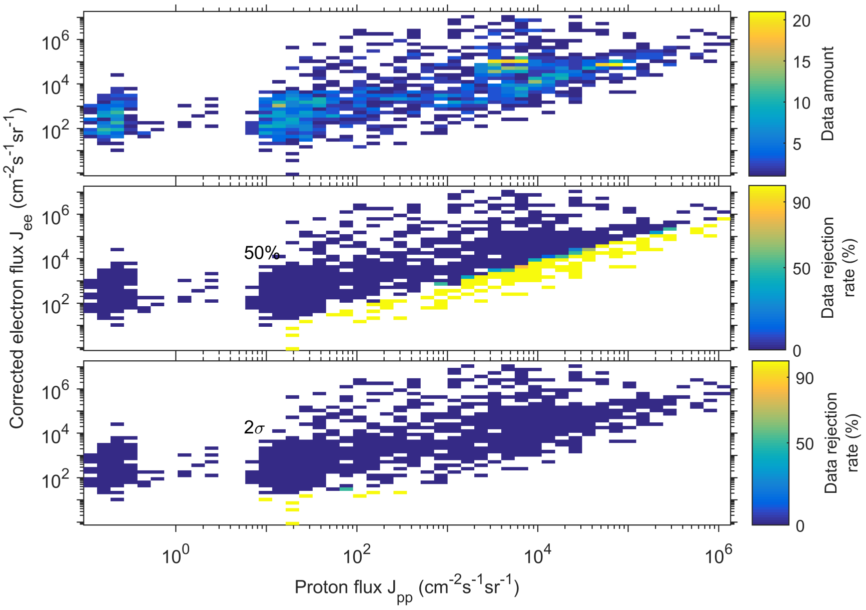


FIGURE 4.5: Data amount and rejection rates using the 50%- and 2σ-criterion (color-coded) of different corrected electron and contaminating proton fluxes (on y- and x-axis, respectively). The plot is based on data from the E1 channel of the 0° detector on board NOAA 15 and covers DOY 302 - 304 in 2003.

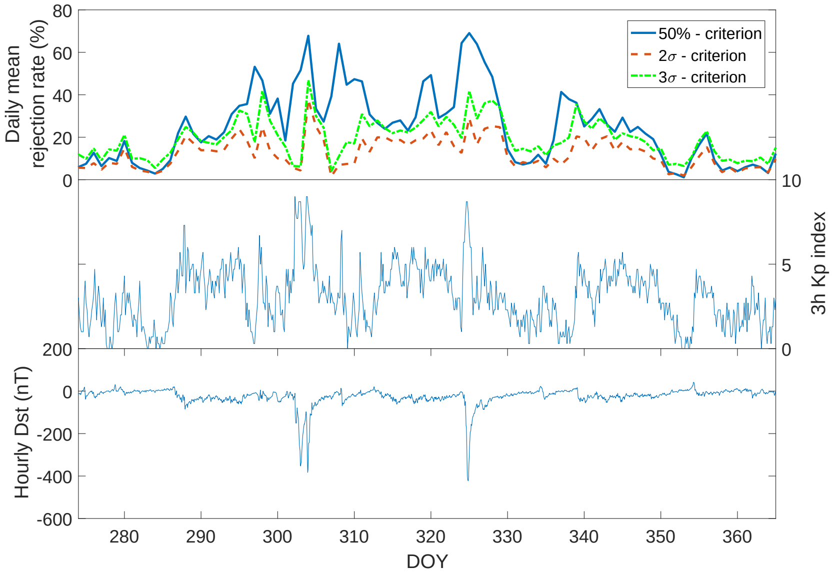


Figure 4.6: Daily mean rejection rates of the three investigated criteria (50%, 2σ and 3σ) of the NOAA 15 0° detector during DOY 274 - 365 in 2003. Rejection rates in each energy channel were superposed to form the daily mean rejection rate. The lower panels show the corresponding temporal evolution of the 3-hour Kp index and the hourly Dst.

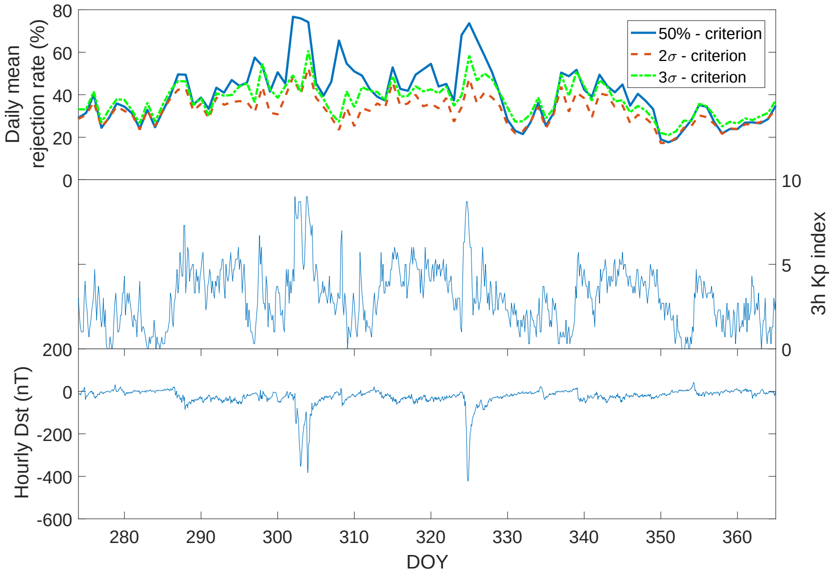


Figure 4.7: Daily mean rejection rates of the three investigated criteria (50%, 2σ and 3σ) of the NOAA 15 LC flux during DOY 274 - 365 in 2003. Rejection rates in each energy channel were superposed to form the daily mean rejection rate. The lower panels show the corresponding temporal evolution of the 3-hour Kp index and the hourly Dst.

Figure 4.5 shows corrected electron flux data from the NOAA 15 0° detector, again during DOY 302 - 304 in 2003 (analogous to Figure 4.4). A clear difference between the 50%-criterion and the 2σ-criterion is visible which is especially pronounced at higher flux levels. Thus, the previously used 50%-criterion potentially leads to unnecessary rejection of high electron fluxes leading to a bias and avoidable data gap during SPEs. Comparison with Figure 4.3 shows equal developments of the rejection rates for simulated and real data. Figure 4.3 was in the first place used to assess the β errors, i.e. the rejection of true data. Rejection of data points located in the rejection areas (yellow) might lead to an overestimation of the corrected flux given the rejected value was true. The upper panel in Figure 4.5, however, indicates that these rejection areas are barely covered by the satellite data. Especially for the 2σ-criterion, rejected data points are located at lower flux levels minimizing rejection of high flux data. Assuming that variations throughout the measurements are entirely based on statistical fluctuations, a possible bias would be reduced by applying the 2σ-criterion instead of the 50%-criterion. Rejection of data and the associated bias would therefore be a lesser issue with the 2σ-criterion.

Temporal behaviour of data rejection rates

An interesting aspect in the comparison of the different criteria is their temporal behaviour. Figure 4.6 depicts daily mean rejection rates of the NOAA 15 0° detector throughout DOY 274 - 365 in 2003 including all geomagnetic latitudes. Note that this rejection rate consists of the superposed rejection rates in all three energy channels. In order for a measurement to be kept, it must fulfil the chosen condition in each of the three energy channels as the final purpose is to create an energy spectrum. The three lines in the upper panel represent the different criteria while the lower panels show the 3-hour Kp and the hourly Dst index. During quiet periods, i.e. the beginning and end of the chosen time period, the 50%- and 2σ-criterion behave similarly while the 3σ-criterion yields higher rejection rates. This results from the fact that it is stricter at lower flux levels as shown in e.g. Figure 4.1. This behaviour changes completely during more active periods. Increased Kp values and a sharp drop in Dst indicate both the Halloween SPE and a second event at DOY 325. Both cases of elevated geomagnetic activity are accompanied by a sharp increase in the 50%-criterion rejection rates to values of almost 70 %. The 2σ- and 3σ-criteria, however, do not surpass 40 and 50 %, respectively. Thus, large differences between the new criteria are visible especially in the peak regions.

As described in Chapter 3, however, the proton contamination criteria are not the only conditions which have to be fulfilled in order to secure a measurement's validity. In addition, simultaneous measurements in the three electron energy channels must form a physically reasonable energy spectrum. So far, rejection rates in a single MEPED detector have been regarded. During the data processing routine, both the 0° and the associated 90° detector measurement must fulfil the criteria. The actual losscone flux rejection rate is composed of a superposition of all single requirement rejection rates and may therefore differ from the mere proton contamination condition rejection rate in the 0° detector as shown in Figure 4.6. In analogy to Figure 4.6, LC flux rejection rates are depicted in Figure 4.7. During quiet times the three criteria render almost equal rejection rates and the large differences between the criteria during elevated geomagnetic activity persists. The 2σ-criterion yields clearly the best performance in minimizing the data rejection rate during storm time. However, all rejection rates are significantly increased when compared to the 0° rejection rates. The rejection rate from the 2σ-criterion, for example, increases by up to 15 percentage points for the losscone flux highlighting the obviously important role of the other requirements.

Spatial behaviour of data rejection rates

The same conclusion can be drawn from Figures 4.8 and 4.9 which show Northern Hemisphere azimuthal projections of the 0° detector and losscone rejection rates, respectively, using Corrected Geomagnetic (CGM) coordinates. Again, data from the three NOAA satellites during DOY 302 - 304 in 2003 was used. As can be seen, the satellites cover quite different MLT ranges which is beneficial when investigating potential longitudinal effects. In general, rejection rates in the 0° detector are low both equatorward of 60° CGM latitude and poleward of 70° CGM latitude. An increase at approximately 60 to 70° CGM latitude results in the apparent formation of a ring where the auroral oval can be assumed to be located. Comparing how the three criteria influence the 0° rejection rates in Figure 4.8, differences are most visible in the mid-latitude high rejection rate areas. There, local rejection rates are decreased by up to 50 percentage points for the two new criteria. Differences between the 2σ- and 3σ-criterion are harder to spot and mainly restricted to latitudes lower than 60°.

Analysis of how the actual LC flux rejection rates are influenced by applying different proton contamination criteria yields smaller yet clearly visible differences than evaluations of the 0° detector data. Figure 4.9 does not show any data coverage over the polar cap because LC electron data is only available in latitudes $< 75°$ CGM latitudes. In latitudes lower than 75°, Figure 4.9 at first exhibits high rejection rates, followed by a decrease approximately southward of the auroral oval and a second increase at even lower latitudes. It is important to keep in mind that the time interval shown in Figures 4.8 and 4.9 are centred in the Halloween SPE in 2003, where the polar cap and high latitudes are filled with large amounts of protons. Naturally, this will influence the performance of both the proton contamination criteria and the requirements on the electron energy spectrum. High rejection rates in low latitudes are expected as electron fluxes are low in this region. Statistical fluctuations when the E2 and E3 channel reach the noise level can lead to the rejection of a considerable amount of data points as they yield unphysical energy spectra.

Where Figure 4.8 showed quite strong improvements for the two new criteria in auroral oval regions, improvements in the LC flux availability in Figure 4.9 are still visible, however less sharp. E.g. the bright yellow area in the 50% plot of NOAA 15 which is located at roughly MLT 18 in latitudes ranging from 55 to 75°, exhibits a significantly

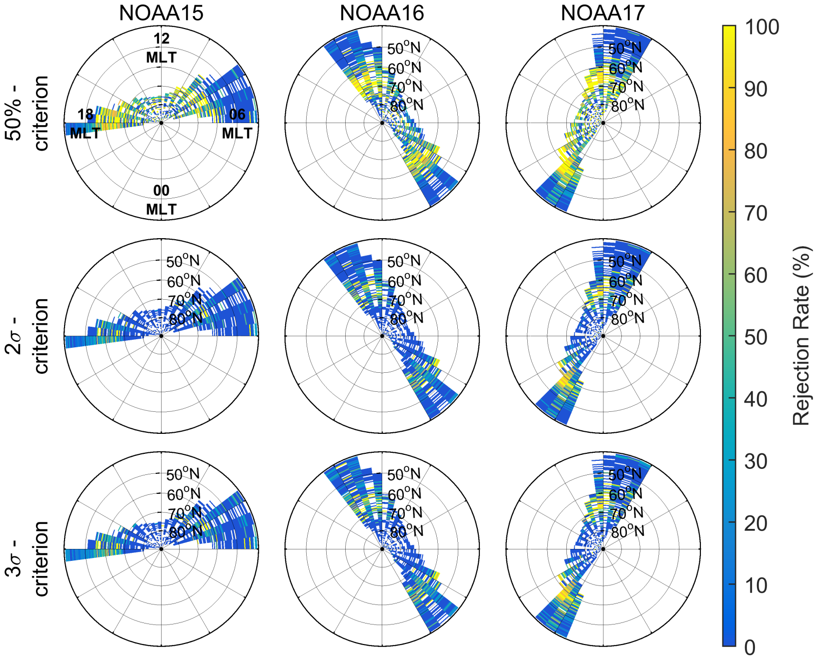


FIGURE 4.8: Azimuthal projections showing the geomagnetic latitude and longitude dependence of the E1 - E3 superposed rejection rate based on measurements by the 0° detector on board NOAA 15 - 17 during DOY 302 - 304 in 2003. The latitudinal and longitudinal resolutions are 0.5 and 7.5°, respectively.

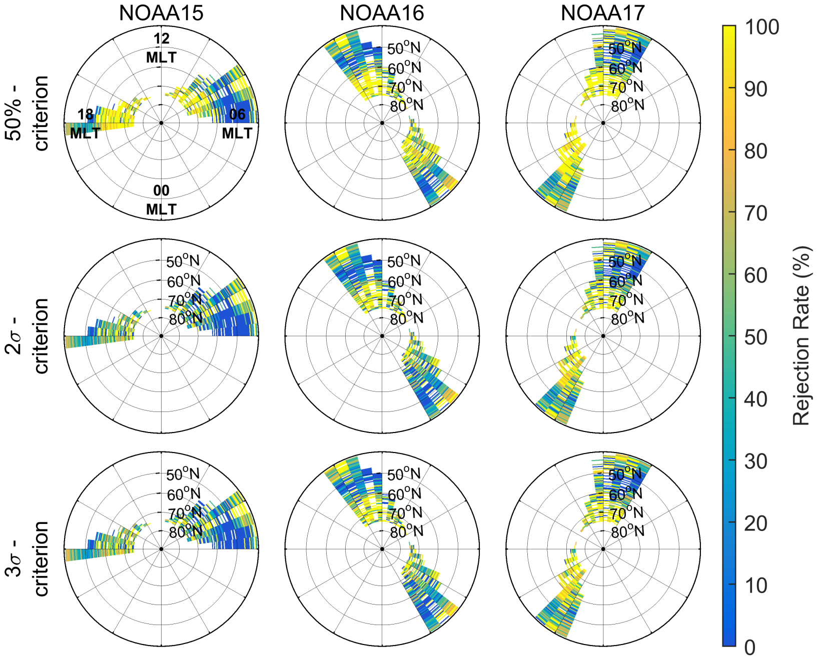


FIGURE 4.9: Azimuthal projections showing the geomagnetic latitude and longitude dependence of the E1 - E3 superposed LC flux rejection rate based on measurements by NOAA 15 - 17 during DOY 302 - 304 in 2003. The latitudinal and longitudinal resolutions are 0.5 and 7.5°, respectively.

darker colour on the corresponding 2σ and 3σ maps. In compliance with the conclusions drawn from Figures 4.6 and 4.7, the role of the other requirements seems to be central when judging improvements of the proton contamination requirement. If, for example, certain electron populations exhibit unphysical energy distributions, LC fluxes of all channels are discarded regardless of whether they fulfil the proton contamination criterion. The performance of the spectrum criteria therefore sets a limit to the possible improvement of data availability through adapting the proton contamination criterion. More azimuthal projections for rejection rates during other time periods are shown in Appendix D.

Quantitative numbers on the improvement of rejection rates due to the new criteria are given in Table 4.1, which states the mean rejection rates for the 0° and 90° detectors of the different satellites during DOY 302 - 304 in 2003 using the three criteria. This particular time period was chosen as it is located in the centre of the Halloween SPE which presents a challenging situation in terms of proton contamination and the associated correction of electron fluxes. In addition to 0° and 90° rejection rates, Table 4.1 states rejection rates derived from LC flux data as well. Again, substantial improvements are visible when comparing the 50%-criterion to the new criteria. In the 0° detector on board NOAA 15, for example, the rejection rate drops from 54.6 to 15.2 % when using the 2σ-criterion instead of the 50%-criterion. Even applying the stricter 3σ-criterion leads to a significant decrease down to 29.0 % which suggests improved applicability of the newer criteria with focus on SPEs. In addition, rejection rates in the 90° detector are probably smaller as it generally measures higher electron fluxes. Both Figure 4.8 and Table 4.1 suggest a worse performance of the criteria for NOAA 17 data. This might partly be due to its trajectory crossing the pre-midnight sector where electron fluxes are generally low leading to higher rejection rates in periods of elevated proton fluxes. When considering LC fluxes, rejection rates are significantly higher and the improvements made by the new criteria are, while still present, not as distinct as for the single detectors. This happens partly due to additional requirements on corrected electron data. Another cause are difficulties during the calculation of losscone fluxes, for instance when 0° and 90° measurements overlap.

In conclusion, a final limit to how far the rejection rates of the POES MEPED data can

Satellite	0° detector			90° detector			LC		
	50%	2σ	3σ	50%	2σ	3σ	50%	2σ	3σ
NOAA 15	54.55	15.18	28.98	36.15	2.09	2.94	75.64	46.82	50.05
NOAA 16	50.50	19.39	21.10	38.95	3.34	4.15	73.69	58.56	60.00
NOAA 17	62.80	35.16	37.57	49.54	15.58	17.81	83.69	71.55	73.31

TABLE 4.1: Different data rejection rates (in percent) during DOY 302 - 304 in 2003 separately for the 0° and 90° detector as well as the satellites NOAA 15 - 17. The shown rejection rates are composed of the single rejection rates in each energy channel. Rejection rates derived from losscone data are shown as well.

be improved, is not only given by the proton contamination requirement subject to this work but also by the various other requirements applied throughout the processing of flux data.

LC fluxes and other data quality requirements

During the processing routine, the condition for the correction of electron data with respect to contaminating protons is called Requirement 2. This implies, that simultaneous measurements in all three channels of one detector E1 - E3 fulfil either the 50%-, 2σ- or 3σ-criterion. Requirement 1 checks for the presence of relativistic electrons in the P6 channel which would be transferred to the fictional E4 channel. This requirement is not compulsory. Requirement 3 examines the energy spectrum of the corrected data in the integral channels E1 - E4 and is split into three sub-requirements 3.1 to 3.3. Requirements 3.1 to 3.3 check whether $E4 < E3$, $E3 < E2$ and $E2 < E1$, respectively. Requirement 3 unites the sub-requirements 3.1 to 3.3 in a way that makes Requirement 3.1 compulsory when relativistic electrons are present, i.e. Requirement 1 is fulfilled. Whenever this is the case, Requirements 3.1 to 3.3 have to be simultaneously fulfilled. Otherwise only Requirements 3.2 and 3.3. are necessary. The rejection rates associated with Requirements 1, 3.1, 3.2, 3.3 and 3 are shown in Table 4.2 for the 0° and 90° detector on NOAA 15 - 17. Again, the analysis is based on data from DOY 302 - 304 in 2003. Relativistic electrons are registered in only 13 - 16 % of all cases and Requirement 1 rejection rates do not differ largely for the two detectors. The requirements on the energy spectrum, however, are violated more often in the 0° detector when compared to the 90° detector. The determination of losscone fluxes requires the existence of simultaneous measurements in both detectors, such that cases where only one detector measurement fulfilled the previous requirements are rejected as well. In addition, losscone fluxes can only be calculated in CGM latitudes smaller than 75° and even then cases can occur where small flux measurements are impeding the reliability of the needed

Sat	Det	Req 1	Req 3.1	Req 3.2	Req 3.3	Req 3
NOAA 15	0°	87.48	49.91	12.33	4.74	18.22
	90°	84.06	20.02	0.09	0.28	0.57
NOAA 16	0°	87.38	41.84	10.00	4.78	16.77
	90°	85.39	17.67	0.27	0.54	0.81
NOAA 17	0°	88.50	57.24	21.24	14.29	28.01
	90°	84.42	35.16	0.56	0.83	2.04

TABLE 4.2: Rejection rates (in percent) during DOY 302 - 304 in 2003 of other criteria throughout data processing for different satellites (Sat) and detectors (Det). Requirement 1 (Req 1) checks for the presence of relativistic electrons and is not crucial for data retention. Requirement 2 checks for the trustworthiness of corrected electron data and is stated in Table 4.1. Requirement 3 examines the energy spectrum and is split up in three sub-requirements. Requirement 3.1 is only necessary during the presence of relativistic protons.

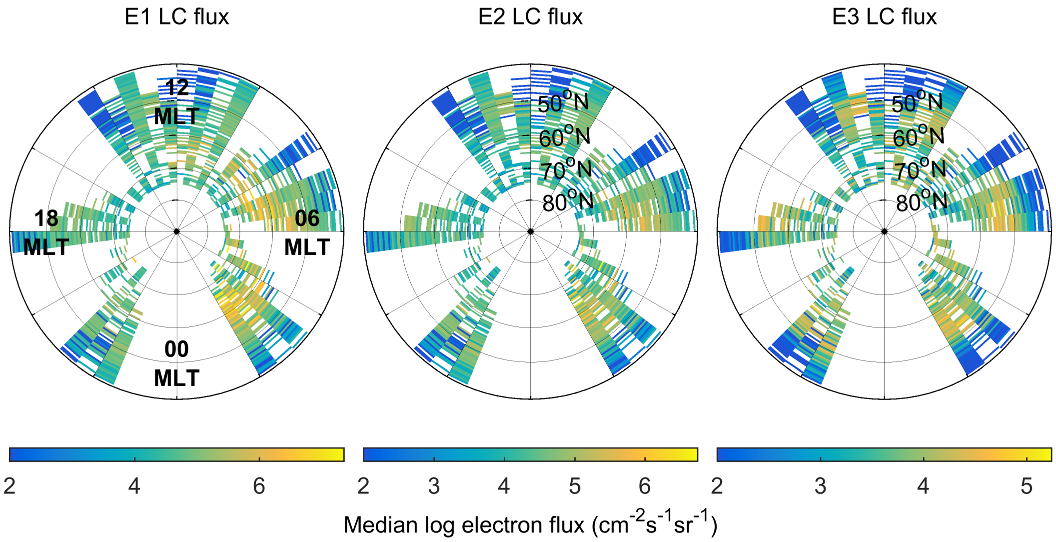


FIGURE 4.10: Maps showing the median corrected E1, E2 and E3 LC fluxes during DOY 302 - 304 in 2003 binned according to geomagnetic latitude and longitude (0.5 and 7.5° resolution, respectively). During proton correction, the 2σ-criterion was used.

ratio between 0° and 90° measurements. These additional requirements and challenges during LC flux calculations lead to the increased rejection rates of the losscone fluxes compared to the mere Requirement 2 rejection rates in the 0° and 90° detectors. Considering this, an improvement of the LC flux rejection rates of 12.1 - 28.8 percentage points during the example SPE in 2003 seems like a good achievement.

Concluding remarks on the considered proton contamination criteria

Concluding the evaluation of the different criteria for proton corrected electron data, it can be stated that both newly introduced criteria yield better results during SPEs than the previously applied 50%-criterion. Under quiet conditions, the 3σ-criterion seems overly strict suggesting the 2σ-criterion as the best candidate both during elevated and low geomagnetic activity. It appears to adequately reject false electron measurements while ensuring that reliable measurements are kept. Overestimation of electron fluxes is limited based on statistics as shown in Figure 4.2 and comparison with data distribution in Figure 4.5 shows that the maximum overestimation only accounts for a small fraction of the actually measured corrected electron flux. For a proton flux of $10^6\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$, overestimation of the corrected electron flux is limited to $10^4\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$. When checking which corrected electron fluxes occur in association with a $10^6\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$ proton flux, the minimal corrected electron flux turns out to be approximately $10^5\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$. This flux is one order of magnitude larger than the potentially over- or underestimated fraction. As the proton contamination criterion affects predominantly low electron fluxes, rejection of these fluxes might introduce a bias towards overestimated flux levels. In order to prevent this bias, rejection electron fluxes could be set to zero to force a bias towards underestimated electron fluxes.

Figure 4.10 shows electron LC fluxes in all three electron channels after applying the 2σ-criterion during DOY 302 - 304 in 2003. Data from NOAA 15 - 17 was binned according to CGM latitude and MLT value. Afterwards, the median flux in each bin was calculated. In general, a certain amount of gaps persists but the overall coverage seems balanced and satisfactory. In addition, typical structures like the stronger E1 fluxes in the midnight-dawn sector and its weakening in the other channels are visible. As can be seen, fluxes with higher energies exhibit a more symmetric longitudinal (MLT) distribution than lower energy fluxes. Analogous flux projections during other time intervals can be found in Appendix D.

4.2 Inter-comparison of the POES/MEPED LC dataset with the CMIP6 parametrization

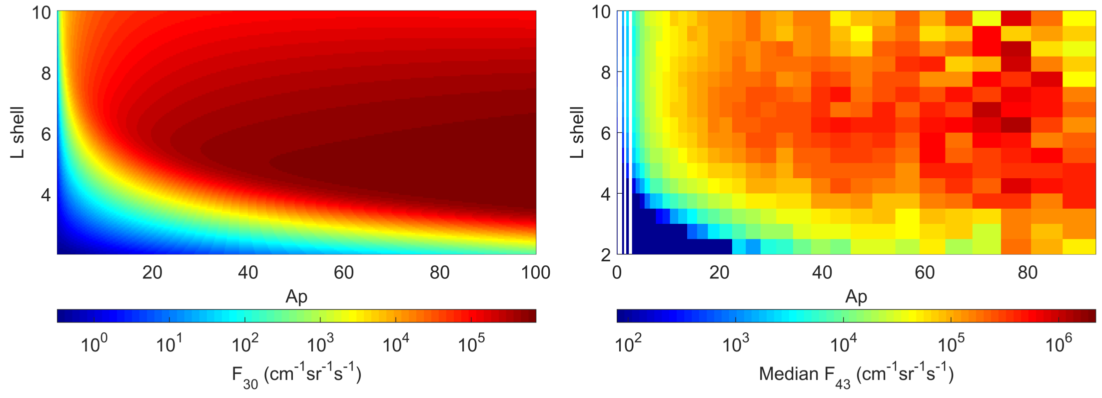
After concluding in the previous section that the 2σ-criterion seems to be the best approach for minimizing data rejection rates while maintaining a good degree of reliability, a 10-year POES/MEPED LC data set (2003 - 2012) is generated using the 2σ-criterion during correction for contaminating protons. So far, evaluation of data rejection rates has been focused on the Halloween SPE in 2003 which presents a worst case scenario with respect to proton contamination. In the general case, however, contaminating proton fluxes are lower than during sporadic SPEs, so that the associated rejection rates are more beneficial as well.

In this section, the 10-year LC dataset is used for comparison to the CMIP6 EEP parametrization described in Chapter 3.3. The applied model is based on modified 0° POES fluxes scaled by the Ap index as a proxy for geomagnetic activity. The aim of the following analysis is to examine how well the model reproduces the overall flux strength and flux variabilities at different latitudes, altitudes and MLT sectors. It will conclude with an investigation of model performances throughout different kinds of storms, i.e. CIR and CME storms, where the latter are often related to SPEs.

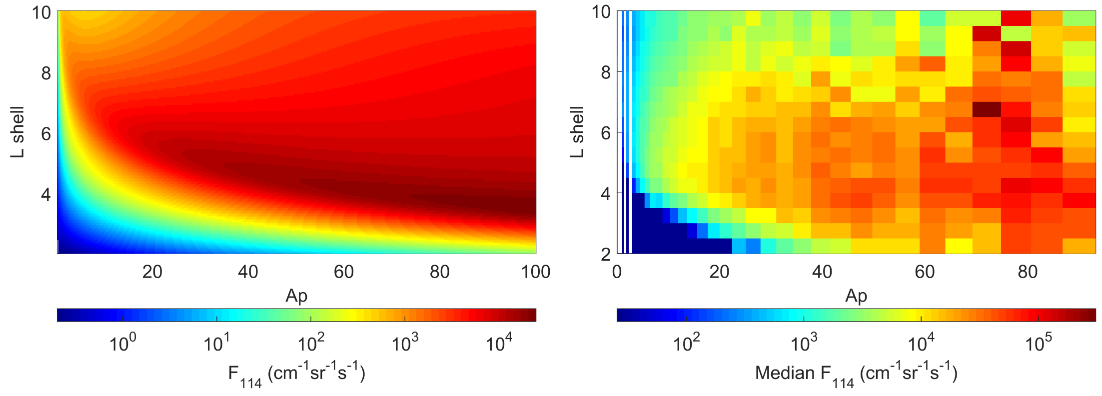
4.2.1 Overall flux strength

In the first step, the overall performance of the Ap model with regard to reproduing general flux strengths is tested. In this case it is important to keep in mind that the model derivation was based on MEPED 0° detector data while the dataset used for comparison contains losscone fluxes. The 0° detector measures only a portion of the losscone flux. Thus, during times of isotropic fluxes, 0° measurements resemble the actual losscone flux to a good degree. Issues, however, arise whenever the fluxes are distributed anisotropically leaving the losscone centre depleted. In these cases, the 0° measurements underestimate the true losscone flux while the 90° flux tends to overestimate. The degree of isotropy in the distribution is obviously affected by the strength of the refilling process. Electrons within the losscone are lost to the atmosphere, but wave-particle interactions can lead to pitch angle diffusion, effectively refilling the losscone with electrons. However, wave activity in the magnetosphere varies with different locations and times and the strength of the interaction between electrons and waves is very dependent on the electrons' energies. Thus, the basic assumption that the 0° flux can be used as an estimate for the losscone flux may not apply for all geomagnetic latitudes, longitudes and electron energies.

The left column of Figure 4.11 illustrates the electron flux distribution with respect to different L shells and Ap values as given by the CMIP6 EEP parametrization. The lower limit of the integral flux was increased to 43 and 114 keV in order to match the MEPED optimized energy thresholds using the energy spectral gradient provided by the model. As can be seen, modelled fluxes grow stronger with increasing Ap values reaching deep red colours. In order to compare the MEPED LC fluxes with the model, daily fluxes were determined in accordance with the approach described in Van de Kamp et al. (2016). Data was binned into 3-hour intervals and L shell bins (with a bin width of 0.5). For each bin, the median flux was calculated, corresponding to zonal averaging. Subsequently, eight 3-hour median fluxes were linearly averaged in order to obtain daily flux values. In the right column of Figure 4.11 the daily fluxes retrieved from the 10-year MEPED LC dataset are binned according to L shell values and Ap strengths. The median flux in each bin is determined and colour-coded in Figures 4.11b and 4.11d for the E1 and E2 channel, respectively. As can be seen, the overall colour distributions seem quite similar, with the model well reproducing the region of lower fluxes for low Ap and high L in the E2 channel. However, it is important to point out the different colour scales used for the model and the LC data. LC fluxes surpass modelled fluxes by one order of magnitude in high flux regions and the ratio between them is apparently not constant for high and low fluxes. In the low flux region, for instance, modelled fluxes are approximately two orders of magnitude smaller than low LC fluxes. Again, this relates to the different degrees of pitch angle diffusion depending on geomagnetic activity.



(a) Ap model. (b) MEPED E1 LC flux.



(c) Ap model. (d) MEPED E2 LC flux.

FIGURE 4.11: Comparison between the Ap model and MEPED LC fluxes in the E1 and E2 channel based on data from all MLT sectors of years 2003 - 2012. Note the different colour scaling of model and MEPED data plots.

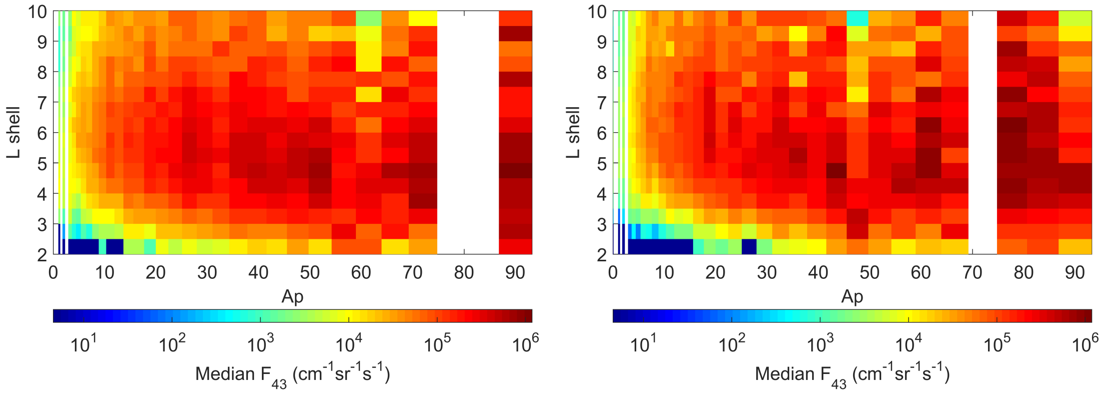
Thus, both in the E1 and E2 channel, LC fluxes exceed modelled fluxes significantly while the degree of underestimation varies with different flux levels.

4.2.2 Temporal variability

With the performance of the model being dependent on the general level of geomagnetic activity, i.e. the Ap index, there might exist tendencies in the model related to different periods within the solar cycle (decadal variability). In addition, it is important to establish how well the model includes short-term (day-to-day) variability and whether model and LC data evolve coherently. The model's ability to capture variations over multiple solar cycles will be analysed as well.

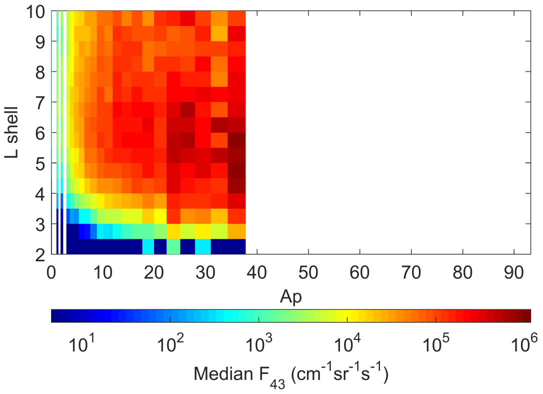
Decadal variability

In order to assess whether the modelled fluxes behave differently during periods of high and low geomagnetic activity, i.e. solar maximum and minimum, fluxes during 2003, 2005 and 2008 are examined separately and compared. The year 2003 was located in a solar



(a) 2003.

(b) 2005.



(c) 2008.

FIGURE 4.12: Comparison between MEPED E1 LC fluxes in different years (2003, 2005 and 2008) based on data from all MLT sectors of the respective year. Note the different colour-scaling compared to Figure 4.11.

cycle peak whereas 2005 and 2008 lay in the declining and minimum phase, respectively. Figure 4.12 shows POES/MEPED E1 LC data binned according to Ap index and L shell separately for the three years. One striking feature is the strong difference in available Ap values in 2008 compared to the other two years. Whereas the Ap index reaches values of up to 200 in 2003 and 100 in 2005 it does not surpass a value of 40 in 2008. The general colour distribution is almost undistinguishable in 2003 and 2005. In 2008, however, maximum flux values are already reached for rather low Ap values (< 37) whereas in 2003 and 2005 higher Ap values are as high as 50 and 60 before the same flux levels are obtained. In addition, low Ap fluxes in high L shells seem to be stronger in 2008 where they exhibit a strong red colour in Figure 4.12c. Furthermore, the dark blue area of low fluxes in low L shells stretches to significantly higher Ap values in 2008 compared to the figures showing 2003 and 2005.

As stated earlier, the general Ap level is significantly lower in 2008 than in 2003 and 2005. Since the modelled fluxes rely only on simultaneous daily Ap values, different flux levels can be expected in the three years. Table 4.3 contains values for median flux levels for the zero-mean E1 LC dataset and the Ap model fluxes in the different years. Indeed, the general flux level decreases from 2003 to 2008, but where the modelled F_{43} fluxes in 2008 are almost two orders of magnitude lower than in 2003, POES LC E1 fluxes decrease by only one order of magnitude. Table 4.3 also states the flux ranges of the two datasets as well as their median absolute difference (MEPED LC fluxes - Ap model fluxes) and ratio (Ap model fluxes/MEPED LC fluxes). As expected due to higher general flux values, the MEPED LC dataset contains fluxes with a significantly wider range than the Ap model dataset.

A noticeable feature is that the Ap model range for F_{114} electron fluxes is constant

Year	POES LC flux (fu)		Ap model flux (fu)		Difference (fu)	Ratio
	Median	Range	Median	Range	Median	Median
E1/F_{43}						
2003	135480	1650000	61750	290000	73300	0.32
2005	59066	1000000	5400	220000	35000	0.12
2008	21671	800000	1040	190000	20800	0.08
E2/F_{114}						
2003	11789	423000	6200	13000	6300	0.34
2005	4424	160000	1020	13000	2600	0.26
2008	1587	110000	338	13000	1100	0.26

TABLE 4.3: Median fluxes and flux ranges (max - min) of the POES and Ap model datasets in years 2003, 2005 and 2008 shown for channels E1 and E2. Median absolute differences (POES - model) and ratios (model/POES) are shown as well. The Ap model integral fluxes corresponding to E1 and E2 are F_{43} and F_{114}. Data from shells defined by an L value between 5 and 5.5 were considered. All MLT sectors were included. fu = flux units ($\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$).

throughout all three investigated years despite their large differences in geomagnetic activity and maximum Ap values. This observation will be referred to and explained later when the flux response with respect to Ap values is analysed. So far, it can be stated that the constant flux range in the E2/F_{114} channel is not confirmed by the LC dataset.

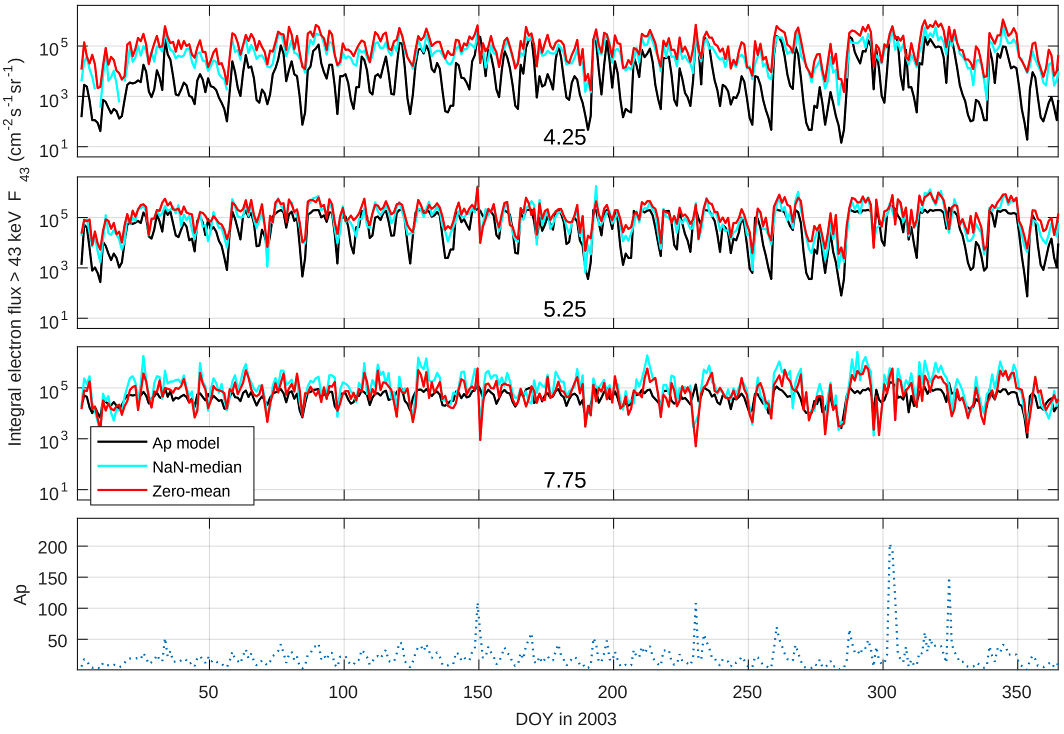
Analysing the median ratio between Ap model and LC fluxes in each year also confirms what has been hinted at earlier. Whereas both datasets contain a clear drop in flux levels from 2003 until 2008, LC fluxes do not decrease as strongly as suggested by the Ap model. Whereas the modelled fluxes were on average 30 % of the LC fluxes in 2003, they are approximately one order of magnitude lower in 2008. This might be explained by the differing colour features in Figure 4.12. If higher flux levels are obtained for lower Ap values in 2008 than during average years, the Ap model flux, which is based on a 11-year average in each bin, would give a larger underestimate compared to the 2008 LC flux. Hence it exaggerates the flux level difference between solar maximum and solar minimum.

Whereas until now only the general flux levels were discussed, it is likely that differences are present on other time scales as well. The strong and especially non-linear sensitivity of the model with respect to the Ap index will lead to different responses of modelled fluxes towards changes in Ap depending on the Ap level itself at which these changes are happening. Whether this short-term variability is affected by the position in the solar cycle will be covered in the next section.

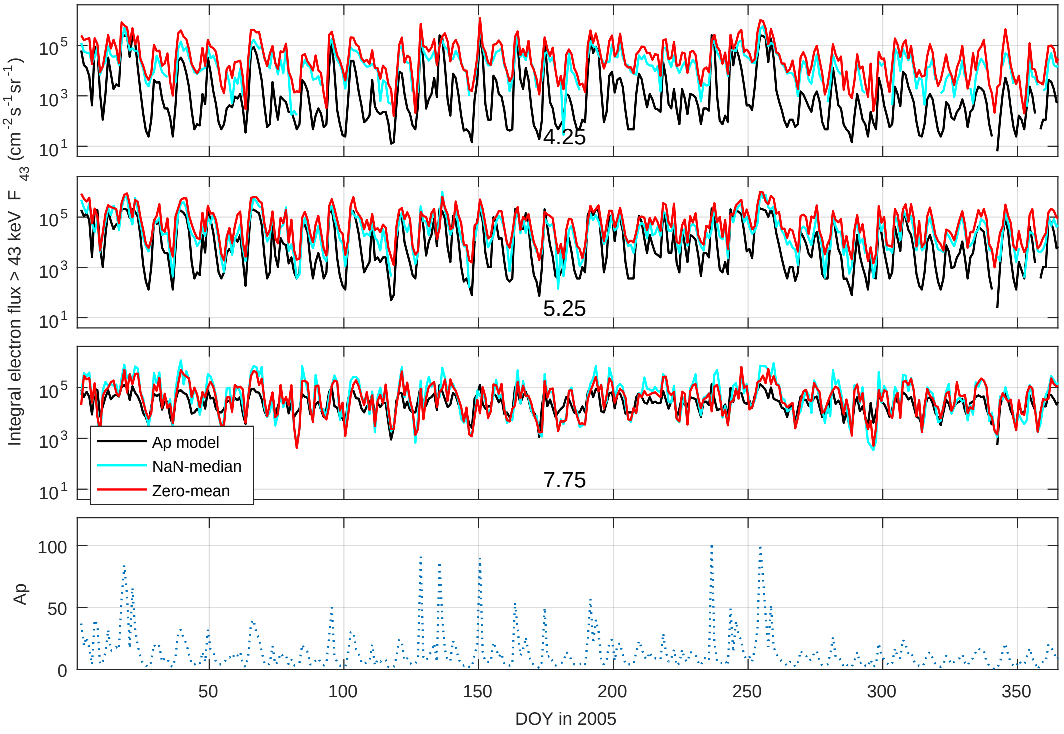
Short-term variability

In order to assess the Ap model's ability to reflect day-to-day flux variations, the temporal evolution of modelled fluxes are compared to MEPED LC fluxes on different L shells (4.25, 5.25 and 7.75 corresponding to 61, 64 and 69°, respectively) in the three years 2003, 2005 and 2008. Figure 4.13 contains separate plots showing fluxes in 2003, 2005 and 2008. The three upper panels in each sub-figure depict fluxes (on the y-axis) on the different L shells, with 4.24 in the upper panel and 5.25 and 7.75 in the middle and lower panel, respectively. The daily Ap index of the respective year is shown in the lowest plot of each figure. All panels share a common x-axis stating DOY in the three years.

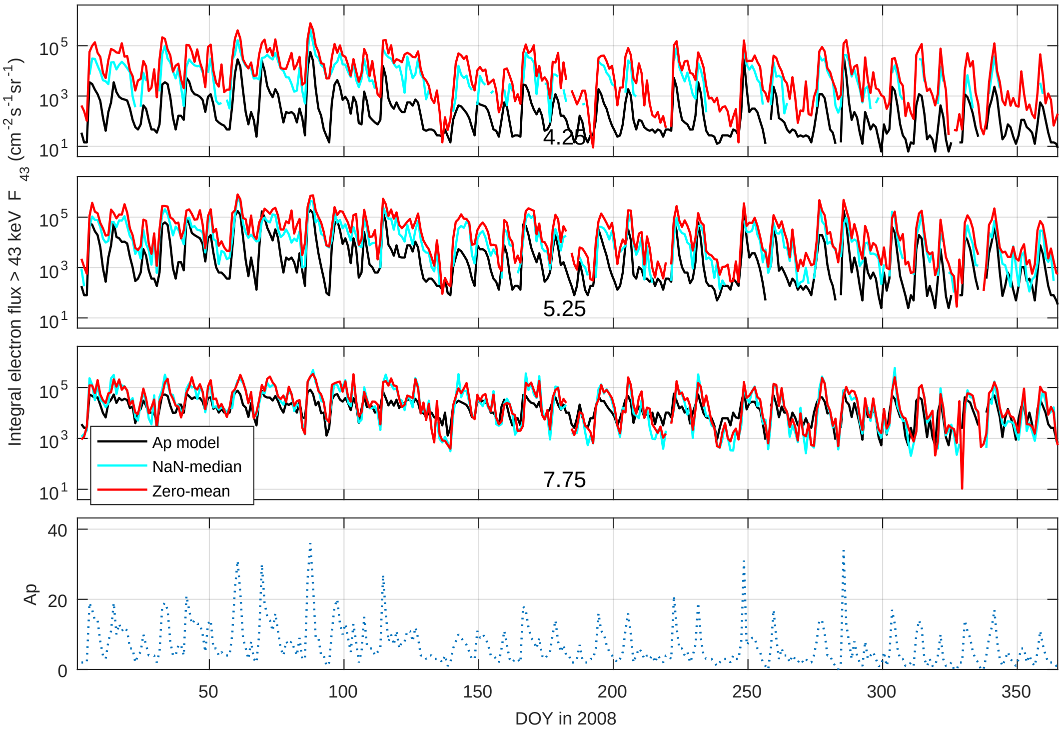
The black line depicts fluxes obtained from applying the Ap model to the respective daily Ap index. MEPED LC flux variations are depicted by two lines, one blue and the other red, called "nan-median" and "zero-mean", respectively. This differentiation is related to different approaches during data processing. At first, rejected data was ignored during further analyses by removing rejected values from the datasets. However, as the majority of rejected data corresponds to low electron fluxes, this might lead to an unfortunate bias towards overestimated electron fluxes. Pulling the bias into the safer direction yielding potentially underestimated electron fluxes can be obtained by replacing rejected data points by zero. In this case, calculating 3-hour median fluxes before converting them into daily fluxes proves troublesome as a zero median is returned


Integral electron flux > 43 keV F_{43} ($cm^{-2}s^{-1}sr^{-1}$)
10^5
10^3
10^1
4.25
5.25
7.75
Ap model
NaN-median
Zero-mean
Ap
200
150
100
50
50
100
150
200
250
300
350
DOY in 2003

(a)


Integral electron flux > 43 keV F_{43} ($cm^{-2}s^{-1}sr^{-1}$)
10^5
10^3
10^1
4.25
5.25
7.75
Ap model
NaN-median
Zero-mean
Ap
100
50
0
50
100
150
200
250
300
350
DOY in 2005

(b)



(c)

FIGURE 4.13: Comparison between POES/MEPED E1 LC electron fluxes and Ap model fluxes in years 2003, 2005 and 2008 for three different L shell values (4.25, 5.25 and 7.75). The daily Ap index is shown in the lower panel.

whenever the majority of data points in a bin is equal to zero. On a physical basis, however, continuing with a zero median neglects the non-zero effects of the minority high-flux data which are entirely ignored by applying the median. Therefore, determination of daily fluxes based on the dataset replacing rejected data by zero uses daily means instead of first calculating 3-hour medians which then are linearly averaged. The two approaches correspond to "nan-median" (removing data points and applying the median) and "zero-mean" (replacing data points by zero and continue with applying empirical means). A more detailed analysis of the potential flux strengths of data points rejected by the different criteria is given in Appendix E.

Apart from the general offset in flux levels, the fluxes shown in Figure 4.13 seem to evolve quite coherently. Short-term variability during high Ap which is not captured by the model can be seen in the two middle panels in Figure 4.13a which depicts fluxes in 2003. The entire year is characterized by high Ap values and strong geomagnetic activity. Especially during the Halloween SPE centred around DOY 300 long periods of strongly elevated Ap are present. Nonetheless, clear Ap peaks are visible. Examination of modelled fluxes in these periods shows plateau-like features with little variability. The corresponding MEPED LC fluxes, however, show variations both for the "nan-median" and the "zero-mean" dataset, which is a good indication of the reliability of the observed

Year	Correlation in E1/F_{43}			Correlation in E2/F_{114}			mean Ap
L shell	4.25 (61°)	5.25 (64°)	7.75 (69°)	4.25 (61°)	5.25 (64°)	7.75 (69°)	
2003	0.65	0.72	0.47	0.74	0.49	0.43	22
2005	0.77	0.75	0.49	0.77	0.65	0.41	13
2008	0.80	0.80	0.72	0.82	0.76	0.52	7

TABLE 4.4: Correlation coefficients between modelled and MEPED LC fluxes in different years, latitudes and energy channels. The mean Ap index of each year is stated as well.

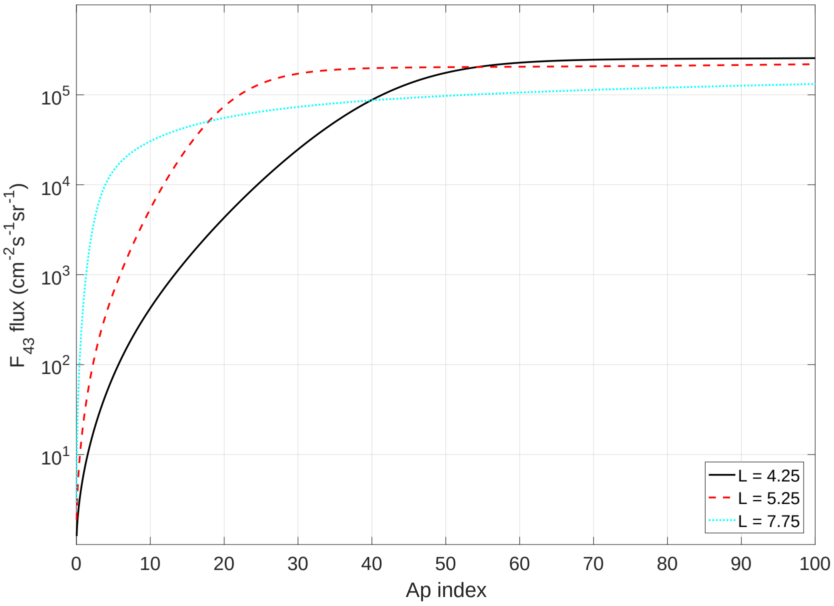
feature. The reason for this behaviour lies with the non-linearity of the Ap-dependence in the model. For elevated Ap values, modelled fluxes saturate which suppresses variations that are clearly observed in the LC data. Thus, the same flux levels in the 5.25 L shell are obtained for all Ap values above approximately 40. Looking specifically at the rise in Ap around DOY 285 in 2003, fluxes in this L shell reach maximum values and saturation immediately and stay there for many days despite a sudden drop of the Ap index by 50 % directly after the peak. This issue is less present in 2005 as the Ap index exhibits rather single peaks followed by drastic drops than extended periods of elevated activity. Due to the low Ap level in 2008, flux saturation does not occur at all.

In order to obtain numbers that state the degree of coherence between modelled and zero-mean LC flux data, the correlation coefficient for the two datasets is calculated for different years, latitudes and energies. The resulting values are stated in Table 4.4 together with the mean Ap index of the respective years. As the correlation coefficient is insensitive to constant offsets between two datasets and merely inspects how simultaneous they evolve in time, the general flux level offset between modelled and LC data is not contained in the correlation. Confirming a tendency visible in Figure 4.13, the correlation increases from 0.72 in 2003 to 0.80 in 2008 at the 5.25 L shell. Although the model generally captures short-term variability to a satisfactory degree, it yields better performances in doing so during solar minimum compared to solar maximum for all investigated L shells and energy channels.

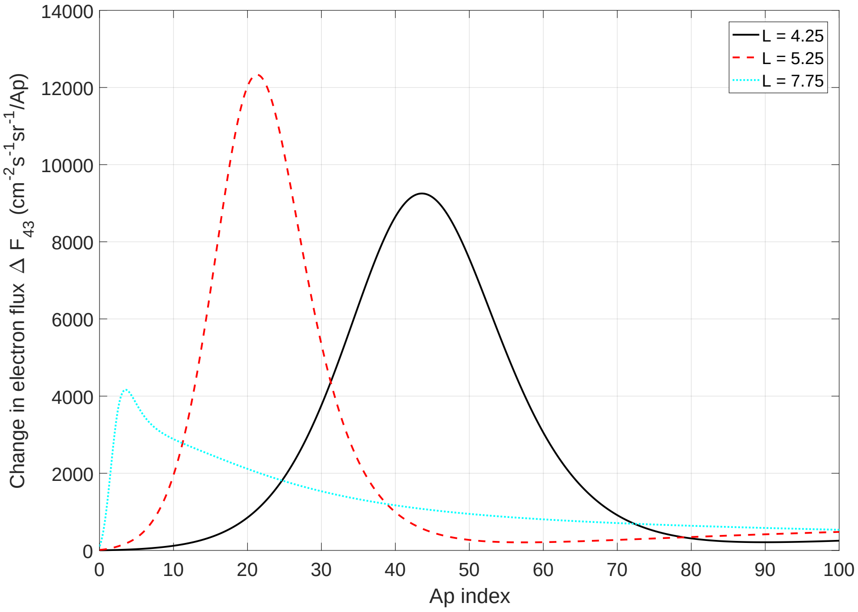
As shown so far, model performances with respect to general flux levels and short-term variations seem to depend heavily on the Ap level itself. Whereas periods of high Ap level, i.e. during solar maximum, are characterized by a better general flux level, they exhibit damped short-term variability. During solar minimum and related lower Ap levels, a higher correlation confirms better flux variability, but flux levels are strongly underestimated.

Ap sensitivity on different L shells

The non-linearity between modelled flux and Ap index has been mentioned before as a possible origin for discrepancies between modelled and LC fluxes. Figure 4.14a shows



(a) Logarithmic F_{43} flux dependent on Ap value.



(b) Absolute F_{43} flux response dependent on changing Ap value.

FIGURE 4.14: Illustrations of the Ap sensitivity of the flux levels themselves (a) and their response to changes in Ap (b).

how the modelled F_{43} flux depends on the Ap index on different L shells. It is evident that fluxes reach a limiting value for high Ap values, meaning that their sensitivity to variations in Ap is diminished. In order to investigate which Ap levels are concerned at different CGM latitudes, Figure 4.14b illustrates absolute flux responses to changes in the Ap index on a linear y-axis. As can be seen, fluxes on L shell 5.25 barely respond to changes in Ap which happen above a limiting Ap value of 40. Fluxes on lower L shells (4.25) respond on Ap levels higher than 40 but are ignorant of variations in high Ap levels (above 70). For high L shells (7.75), the amplitude of the peak response is strongly decreased and its location relates to very low Ap levels of approximately Ap 5. Its response falls off very flatly towards higher Ap values.

Analysing similar plots for the F_{114} electron flux yields an explanation for the constant range observed for the L shell 5.25 in Table 4.3. For all three years, 2003, 2005 and 2008, the modelled F_{114} fluxes exhibited equal ranges on this L shell. E2 electron fluxes from the LC dataset, however, had quite varying flux ranges with the 2003 range exceeding the 2008 range by one order of magnitude.

Figure 4.15 shows how modelled F_{114} depend on the Ap index for different L values analogous to Figure 4.14a depicting F_{43} fluxes. In this case, however, fluxes do not just saturate and reach a limiting flux level but show a decrease after passing a certain Ap value. The red dashed line, for example, depicts fluxes on the 5.25 L shell. At first, fluxes increase with increasing Ap until the index reaches a value of approximately 30. For higher Ap values than 30, modelled fluxes start to decrease again, meaning that a maximum flux level is reached for Ap 30. As this is a common value during both solar

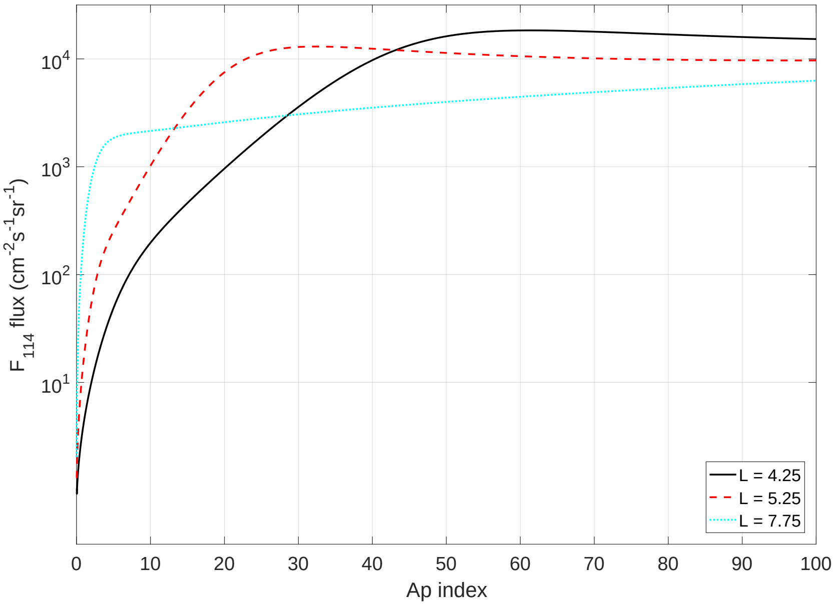


FIGURE 4.15: Logarithmic F_{114} flux dependent on Ap value.

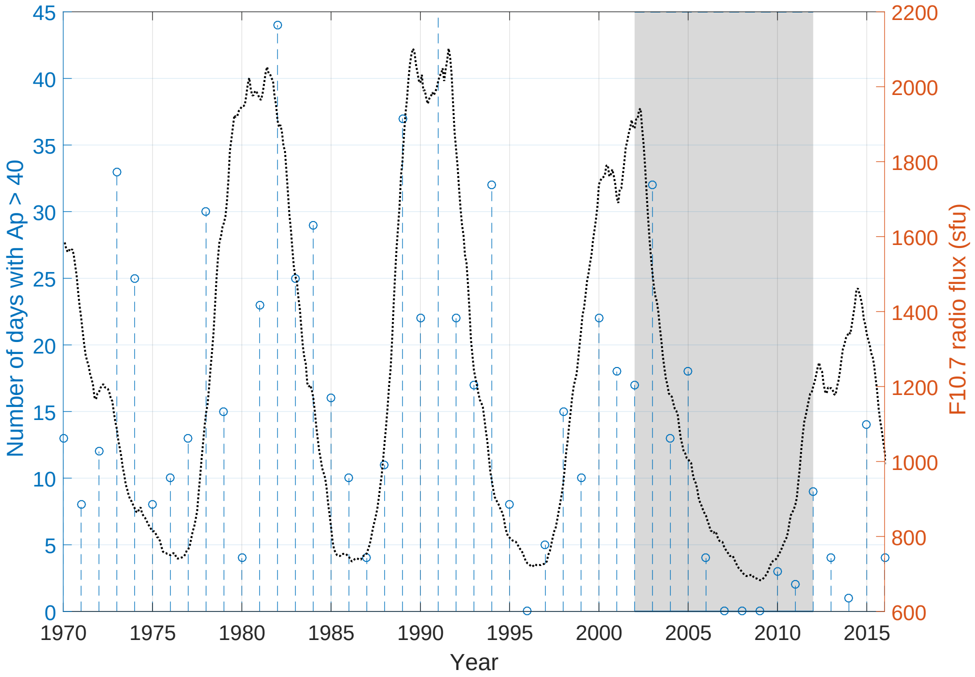


FIGURE 4.16: Number of days during each year with a daily Ap index > 40 from 1970 - 2016. The black dotted line indicates the solar cycle and the grey area marks years considered in the Ap model derivation.

maximum and solar minimum years, the maximum flux and flux range of the modelled F_{114} flux are equal in all three investigated years. Whether this behaviour, in addition to observed oscillations in the flux responses to changing Ap for higher energies, is adequate for describing real flux behaviour, remains quite doubtful.

Multiple solar cycles

One of the advantages of the Ap model is that it provides electron fluxes for as long into the past as the Ap index is available ($\sim$ 1850 according to Matthes et al. (2017)). The model itself, however, is only based on parts of solar cycle 23 and 24. As these are relatively weak cycles compared to previous solar cycles, the evident dampening of variability at times of high Ap index might impact the model's performance considering multiple solar cycles. The dampening effect appears for Ap values roughly larger than 40. This value is supported by the limit to flux response for the 5.25 L shell found from Figure 4.14b.

Figure 4.16 shows the number of days in each year from 1970 to 2016 with a daily Ap index > 40. These would be the days affected by dampened model sensitivity. The black dotted line indicates the solar cycle. As menationed earlier, the Ap model is based on MEPED electron data from 2002 to 2012. This time period is marked by the grey area. As can be seen, only few days throughout the whole period 2002 - 2012 are associated with Ap values > 40. Previous solar cycles, however, exhibit a larger amount of these days and are thus stronger affected by dampened short-term variability. Hence, it should be noted that parameters like flux strength and variability are not adequately represented by the Ap model in a long-term perspective.

4.2.3 Spatial variability

While establishing how well the model is able to capture temporal variability on different time scales, analyses was mainly focused on the E1 energy channel, a specific latitude and zonally averaged data. Whether or not there exist differences in model performance for different locations can be crucial with regard to judging the overall goodness of the model. It is also extremely important to be aware of local effects if using modelled fluxes at a specific height or location.

Energy impact in different altitudes

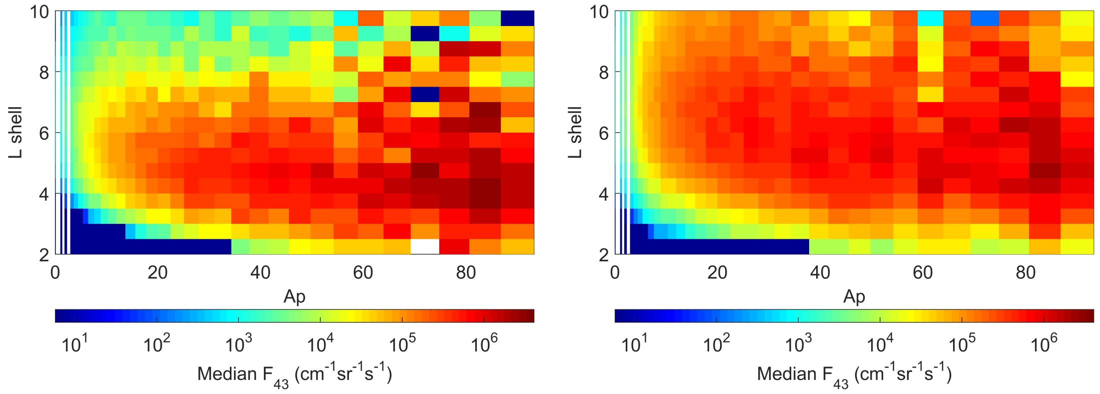
As explained earlier, electrons with higher energies will deposit their energies at lower altitudes than electrons with lower energies. Typical altitudes for peak ionization by electrons measured in the E1 channel (> 43 keV) range around 90 km as shown in Figure 2.12. The E2 channel, however, measures electrons with energies > 114 keV which lead to peak ionization at approximately 70 km. Energy deposition in these two regions will have different effects on atmospheric composition and dynamics. F_{43} fluxes (integral fluxes of electrons with energies > 43 keV) will impact the upper mesosphere, i.e. the D-region, altering D-region chemistry. When running a model including D-region parametrization, direct effects will be visible throughout the D-region and mesosphere. Due to downward transport, upper atmosphere observations will consist of a superposition of direct high energy electron impacts and secondary effects originating from higher altitudes. In order to separate the origins of superposed effects to understand dynamics and direct impacts on chemistry, it is important to have reliable information on which fluxes affect the different altitudes. In the Ap model, the F_{30} (> 30 keV) flux is fitted directly to the Ap index, all other integral fluxes, however, have to be calculated via the fitted spectral gradient. Assuming a given degree of uncertainty for the F_{30} flux, applying an estimated energy spectrum is likely to increase uncertainties for higher energy fluxes. According to Table 4.3, the median ratio between modelled data and zero-mean LC fluxes are higher in the E2 channel compared to the E1 channel. This means that E2 fluxes in general are captured better by the model, on the other hand, the previously named pitch angle diffusion might be stronger for electrons with higher energies. This would then lead to a distortion of the energy spectrum. Low energy fluxes would be underestimated by relying on 0° detector fluxes whereas higher energy fluxes appprarently comply better with losscone fluxes. Using an underestimated E1 flux for fitting a power law based on three data points (E1, E2 and E3) might result in a too flat energy spectrum which is used to predict the energy deposition throughout the atmosphere.
Correlation coefficients for the E2/F_{114} fluxes are stated in Table 4.4 and do not differ significantly from E1/F_{43} values. In some cases, however, there is a tendency towards lower correlation which is probably related to Ap sensitivities and the general model set-up, i.e. decreased sensitivity and even reduction of fluxes for high Ap.

Latitudinal variations

Already from mere visible examination of Figure 4.13, different behaviour of modelled fluxes in different latitude regions becomes obvious. The general flux level is captured best for high latitudes with large gaps being visible in low latitudes. The flux response to varying geomagnetic activity, however, tends to be overestimated in low latitudes while it seems to be underestimated in high latitudes. In the L = 7.75 panels of Figure 4.13, the modelled fluxes vary extremely little compared to variations visible in the LC flux data. This feature is also visible in numbers stated in Table 4.4, where the correlation decreases with increasing latitude in all three years and both energy channels. As both the nan-median and the zero-mean dataset exhibit these degrees of variability, it can be assumed to be an actual feature and not an artefact resulting from data processing.

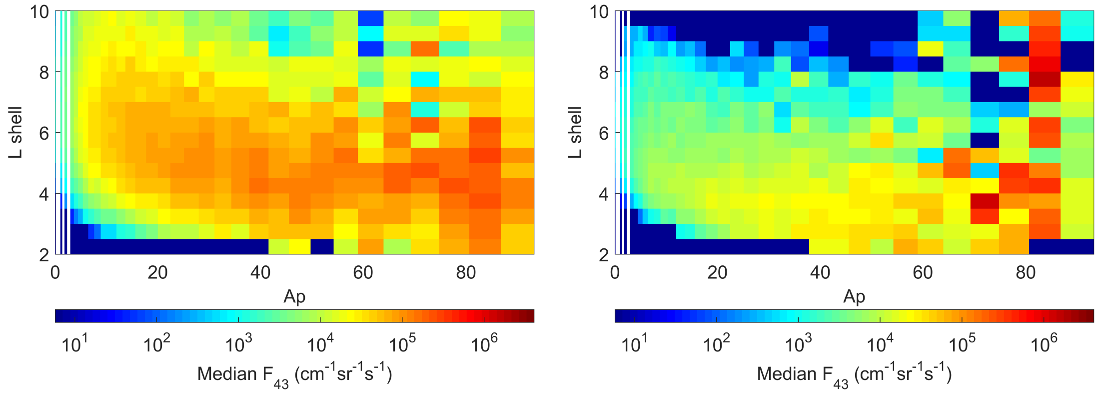
MLT variations

The main drawback of the Ap model is that it provides zonally averaged electron fluxes. When, however, taking into account electron drift directions and regions of increased wave-particle interaction, different levels of electron fluxes are to be expected in different MLT sectors. During geomagnetic storms, particle injections into lower L shells



(a) MLTs 00 - 06.

(b) MLTs 06 - 12.



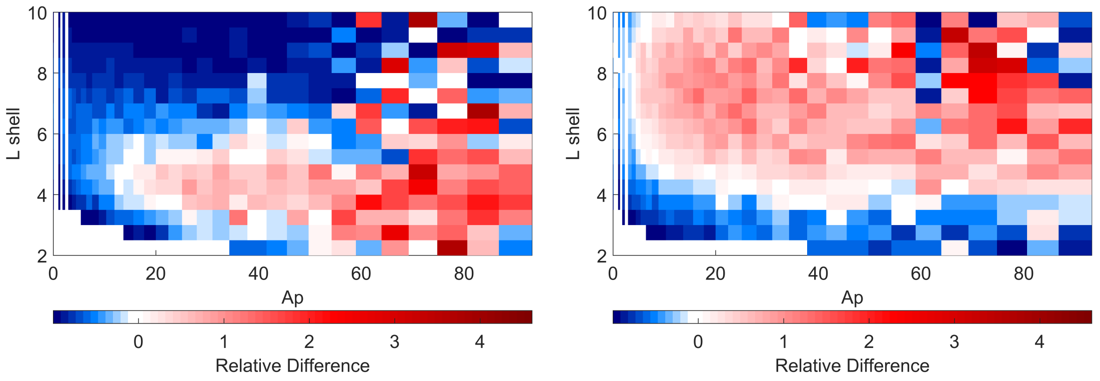
(c) MLTs 12 - 18.

(d) MLTs 18 - 24.

FIGURE 4.17: Comparison between the MEPED E1 LC fluxes in different MLT sectors based on data from years 2003 - 2012.

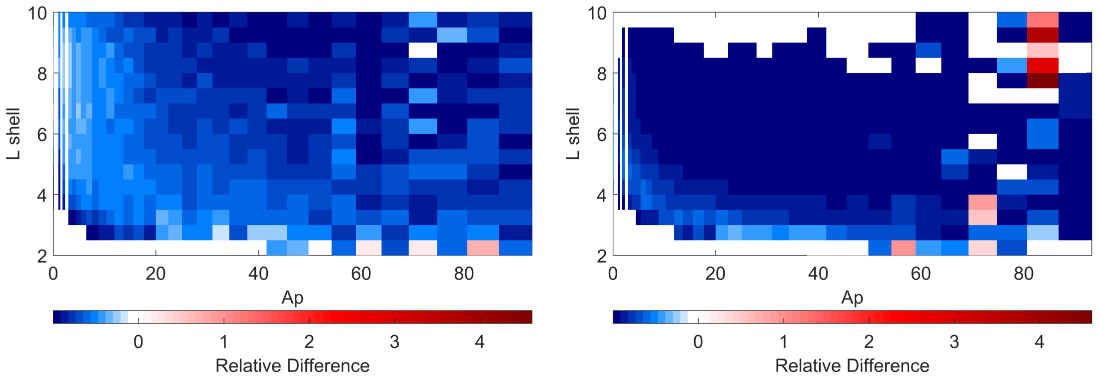
originate from the plasmasheet and are restricted to the night-side. While gyrating around field lines and bouncing between the two hemispheres, protons will drift westwards around Earth while electrons have an eastward drift, i.e. from midnight towards the morning sector. On their trajectory, wave activity increases, first in lower L shells and with increasing MLT also in higher L shells refilling the losscone and securing electron precipitation. Due to day-night asymmetries of the magnetic field (compression on the day-side, elongation on the night-side), particles on higher L shells are lost on the day-side (noon). A considerable amount of electrons is either lost in the atmosphere or drifted off on the day-side leading to lower ambient electron fluxes in the post-noon and evening sectors. Although wave activity is present, the electron population diminishes, leading to lower precipitating fluxes in the evening and night sector. Therefore, significantly lower fluxes are to be expected e.g. in the pre-midnight sector than in the pre-noon sector.

In order to establish how large these differences are, the 10-year MEPED E1 LC data, i.e. zero-mean daily fluxes, were binned according to the MLT value of their measurement into four different MLT categories: 00 - 06, 06 - 12, 12 - 18 and 18 - 24. The effective range of the fourth category end at 22 instead of 24 as no satellite covers MLTs 22 - 24.



(a) MLTs 00 - 06.

(b) MLTs 06 - 12.



(c) MLTs 12 - 18.

(d) MLTs 18 - 24.

Figure 4.18: Relative differences between MLT-separated fluxes and superposed-MLT fluxes associated with Figure 4.17.

Afterwards, data in each MLT sector was binned according to Ap and L shell and median fluxes in each bin were determined as done previously e.g. in Figure 4.11 or 4.12. The results as shown in Figure 4.17 confirm suspicions of different flux levels in different MLT sectors. The post-midnight/morning sector in Figure 4.17a shows high fluxes in lower L shells as wave activity there is strongest and starts earlier. Electrons in higher L shells have to drift further towards the day-side in order to experience wave-particle interaction resulting in lower fluxes in regions of high L value. In the second MLT sector, pitch angle diffusion is present in all investigated L shells. After crossing noon, particles in the upper L shells are lost leading to the observed decrease in fluxes in Figure 4.17c. Simultaneously, the general flux level decreases as less electrons are present. Thus, the electron flux is minimal in the evening sector as shown in Figure 4.17d, where barely any losscone electrons are present in high latitudes and also lower L shell fluxes are quite low. The visible strong fluxes between Ap values of 70 and 80 are likely to result from a distortion in the particle injection during high activity. Although injected particles originate from the night-side plasmasphere, this does not mean that injections are necessarily centred at midnight. Depending on field asymmetries in the y-direction, primary particle injections can be pushed towards the evening sector. A good confirmation that these high fluxes do not result from contaminating protons is that this feature persists even after removing SPE data.

The relative deviations of each MLT sector with respect to the analogous superposed MLT illustration in Figure 4.11b

$$\frac{\text{Flux}_{\text{specific MLT}} - \text{Flux}_{\text{all MLT}}}{\text{Flux}_{\text{specific MLT}}} \tag{4.7}$$

are shown in Figure 4.18. The relative difference is positive, when the flux in a specific MLT exceeds the zonally averaged flux.

In the night/morning sector, high latitude fluxes are lower than the zonal average. Areas corresponding to high Ap and low L shells exhibit higher fluxes. Electron flux levels are generally higher in the pre-noon sector while zonally averaged fluxes exceed post-noon and evening fluxes significantly in all L shells and Ap bins. As the relative differences are given in absolute numbers, large parts of the MLT-separated plots deviate from their zonally averaged counterpart by at least 100 %. Assuming averaged electron fluxes in the noon to midnight sector will for example drastically overestimate electron fluxes.

Another issue is that satellite trajectories are not distributed evenly over all MLTs but overlap in many cases. Fluxes of MLT sectors that contain more satellites than others will thus get a larger weight during averaging. In addition, the model derivation was based on MEPED measurements from 2002 until 2012. In the period 2002 - 2005, however, only three quite evenly distributed satellites were available (NOAA 15 - 17). With the launch of NOAA 18 in 2005 this number increased to four available satellites. The trajectory of NOAA 18, however, largely coincides with the NOAA 16 trajectory in the post-noon and post-midnight sector. This means that after the NOAA 19 launch, a heavier weight lay on small flux levels in high L shells and stronger flux levels during high geomagnetic activity as the positive relative deviation in Figure 4.18a exceeds the negative deviations in Figure 4.18c. With the launch of NOAA 19 in 2009 this tendency became even stronger as its trajectory coincides with the NOAA 18 trajectory as well. However, NOAA 16 had drifted significantly towards later MLTs somewhat damping but not annihilating this effect.

Thus, the goodness of the modelled fluxes can be expected to vary strongly with the investigated MLT sector and it is likely that they do not reflect proper zonal averages. In this respect the model can easily be improved by introducing a weighing function for data points originating from similar MLT sectors. In general, a model including an MLT resolution would present a very beneficial tool.

4.2.4 Model performance during different types of storms

As explained in Chapter 2, different types of geomagnetic storms exist and dominate during different periods in the solar cycle. While CME-driven occur predominantly during solar maximum, CIR-driven storms are more accentuated in the declining and minimum phase. Due to their differing origins and onset mechanisms, CIR and CME storms exhibit different characteristics in storm evolution, flux levels and magnetosphere states. Whether the Ap model, being based on solely one geomagnetic index, is able to capture the predominant storm behaviour for different types of storms will present a core issue when judging its applicability.

CME storms

An example for a typical CME-driven storm is shown in Figure 4.19. Strong geomagnetic perturbations lead to very high peak Ap values. The model (black line) which is based on the Ap index indicates therefore a comparatively high flux level during the CME which at peak storm times approaches nan-median and zero-mean daily LC fluxes (blue and red lines) consistent with more isotropic pitch angle distributions. During quiet conditions before and after the storm when the Ap is relatively low, the modelled fluxes are pulled towards lower flux levels due to their increased sensitivity at low Ap values. Although the correspondance between modelled and LC fluxes is improved during peak storm time, model fluxes are still underestimating losscone fluxes by a factor of approximately three. Even an Ap value surpassing 100 on DOY 254 does not

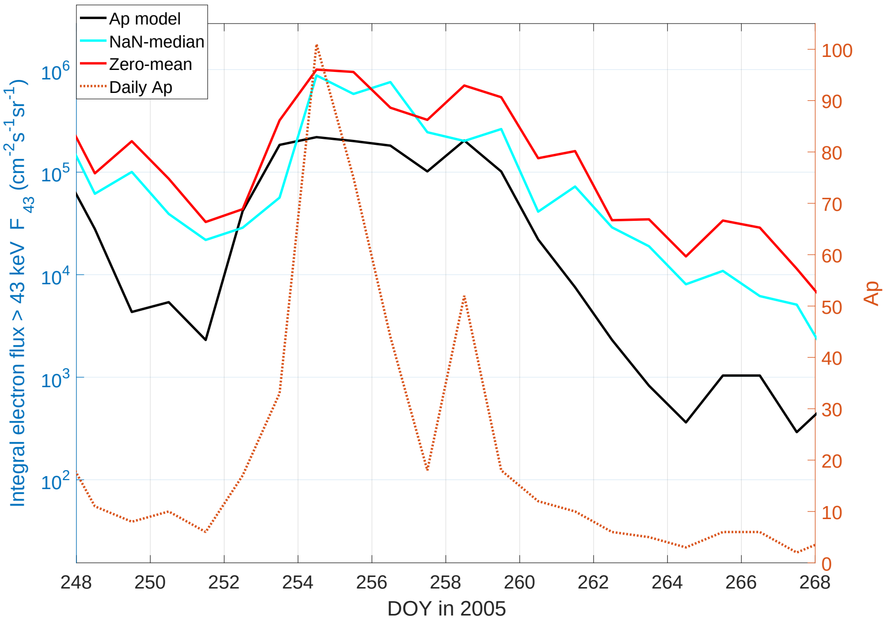


FIGURE 4.19: F_{43} Ap model, nan-median and zero-mean fluxes during a CME storm in 2005 (DOY 248 - 268) on the L shells 5 - 5.5. Daily Ap values are illustrated on the right-hand y-axis.

activate a significant increase in modelled fluxes, again underlining the central role of Ap sensitivity. In contrast, both nan-median and zero-mean fluxes clearly react to this large increase in Ap. Instead modelled fluxes saturate and the plateau-like flux level becomes visible during peak storm times.

As CME occurrence is often related to SPEs, the model's incapability of capturing elevated fluxes and their short-term variability during peak storm time will influence its overall performance during SPEs. In fact, the storm shown in Figure 4.19 is one of the stronger SPEs in the period 2003 - 2012 which was closely related to the arrival of an ICME. Although later discussion will show that LC fluxes exhibit larger uncertainties in connection with SPE proton contamination, both the zero-mean and the nan-median flux can be seen as lower estimates of the actual electron flux.

Statements valid with regard to both SPE-related CME storms and non-SPE related storms are that the same maximum flux level is reached for modelled fluxes independently from storm strength. During weaker storms (measured in Ap roughly $<$ 45 - 50) modelled and LC flux levels seem to comply but whereas LC fluxes increase with increasing Ap index, modelled fluxes are limited to their saturation level.

CIR storms

In contrast to CME-driven storms, CIR storms are characterized by significantly lower geomagnetic perturbations and lower Ap values. When taking into account the nonlinear sensitivity of the model with respect to the Ap predictive values, an increased

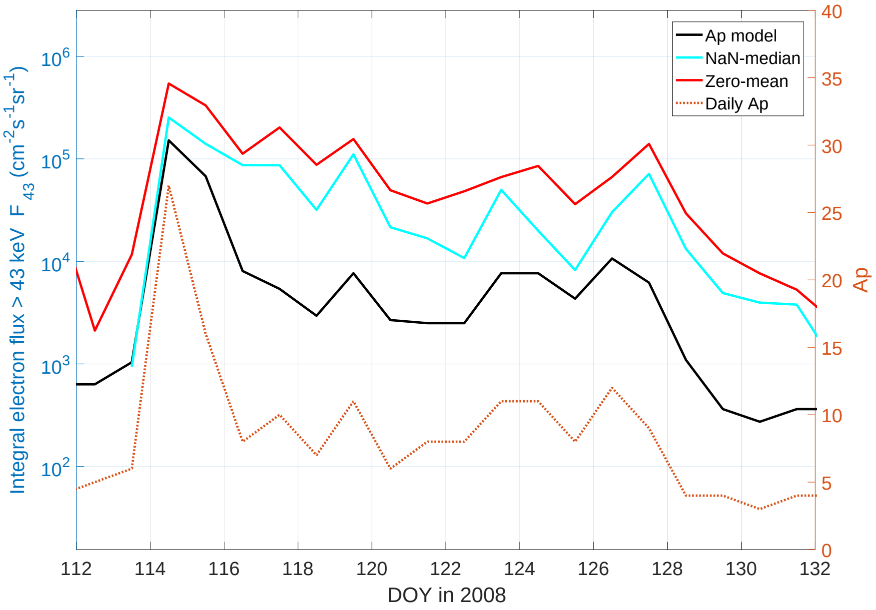


FIGURE 4.20: F_{43} Ap model, nan-median and zero-mean fluxes during a CIR storm in 2008 (DOY 112 - 132) on the L shells 5 - 5.5. Daily Ap values are illustrated on the right-hand y-axis.

sensitivity can be expected for CIR storms. Figure 4.20 illustrates F_{43} Ap model fluxes (black line) and E1 nan-median and zero-mean LC fluxes (blue and red lines) during a typical CIR-driven storm in 2008. As expected, the peak Ap value of approximately 27 is rather low, but as the Ap model is sensitive to Ap variations on this level, modelled fluxes follow the rise in Ap on DOY 114. Differences in peak flux levels between the Ap model and MEPED LC measurements persist but are expected due to the likely underestimation of losscone fluxes by the 0° detector. One issue occurring during CIR-driven storms which is visible in Figure 4.20, are quite elevated fluxes during the recovery phase although the Ap index has already restored itself to quiet time values. Bound to the Ap index, the modelled fluxes are drawn towards lower electron fluxes whereas both nan-median and zero-mean fluxes stay on higher flux levels and exhibit a slower decrease.

As this is difficult to observe in Figure 4.20 due to the logarithmic y-axis scale, a short evaluation routine was implemented which sorts out geomagnetic storms based on variations of the Ap index and examines the corresponding storm length of the elevated fluxes. Only storms which were preceded by a 2-day long quiet period were considered. The start and end of a storm were defined by the times fluxes dropped below values one order of magnitude lower than the storm peak fluxes (after removing flux base-lines). This routine worked well in terms of picking out and examining single storms in 2005 and 2008. The year 2003 was, however, marked by very elevated geomagnetic activity, eliminating most storms as they did not exhibit a quiet period in advance.

Applying the routine to Ap model and zero-mean LC fluxes in 2008 (nan-mean fluxes exhibit data gaps during quiet times and do not qualify) yields in total 18 identified storms one of which is the one depicted in Figure 4.20. According to the previously explained definition of storm length, LC fluxes imply a four day longer storm length than modelled fluxes in this case. As this corresponds to a 114 % longer period of elevated flux levels, significant differences in the overall energy input by EEP are to be expected. Indeed, LC fluxes suggest a 50 % longer storm time when averaging all 18 identified storms in 2008. For the L shell 5 - 5.5, eleven storms could be identified in 2005, yielding on average a 34 % longer storm time for LC fluxes than for modelled fluxes. It seems as if the increased Ap sensitivity forces modelled fluxes towards lower levels resulting in an underestimation of the integrated energy impact during the heavy tail CIR recovery phase. The effect is less present in the Dst-based model as the Dst index has longer recovery phases than the Ap index but the associated Dst model flux dataset contains many gaps due to positive Dst values in 2008. This behaviour is especially intriguing as this underestimation of EEP impact arises during the declining and minimum phase in the solar cycle, periods in which electron precipitation is of utter importance as it presents the dominant particle input into the atmosphere.

4.2.5 Discussion of model performances

As suggested by previous analyses, the presented Ap model which will be implemented as EEP parametrization in CMIP6 exhibits a number of issues which ought to be addressed.

- A difference in basic flux strength arises from utilizing 0° detector electron fluxes instead of actual losscone fluxes. This could be accounted for by modelling electron fluxes based on losscone fluxes instead of 0° detector measurements.
- The degree to which the model is able to reproduce general flux levels and short-term variabilities is very dependent on the chosen phase of the solar cycle.
- Flux levels are in all cases underestimated, but correlations between modelled datasets and LC flux datasets are generally high. The model succeeds in capturing short-term variations of E1/F_{43} fluxes, the range of the variability, however, are often underestimated. This might as well partly stem from the assumption of equivalent 0° detector and LC fluxes, as its precision varies with different L shells and energies.
- Answering how well the model works for higher energy fluxes is a difficult matter. Fluxes tend to be more evenly distributed in terms of longitude but the correlation between high energy fluxes and geomagnetic indices generally decreases. Trying to predict these fluxes using geomagnetic indices might therefore prove difficult. In addition, fluxes of energies higher than 30 keV have to be calculated using a modelled power law spectrum. Whether or not energy spectra always correspond to power laws, however, is still somewhat unclear and some studies also consider exponential spectra (Ødegaard et al., 2017), leaving the applicability of the power law gradient somewhat uncertain. The low range variability found in Table 4.3 and the saturation level found for low Ap values in Figure 4.15 suggest that there is room for improvements at higher energies.
- Taking zonally averaged fluxes as a basis might introduce a bias towards MLT regions with more traversing satellites and different (F_{43}) fluxes have to be expected in different MLT sectors. This issue can be eliminated by introducing a weighing function or modelling different MLT regions separately.
- A different issue are cases where the model struggles to reproduce storm features due to the incapability of the Ap index which it is based on and its non-linear Ap sensitivity. Clear examples are peak flux times during CME-driven storms where the Ap model fails to reproduce elevated flux levels and flux variability or typically heavy storm tails during the recovery phase of CIR-driven storms. Adapting the Ap sensitivity could improve model performances during CME storms by compensating for the lack of variability at high Ap levels as well as during CIR storms. It is also possible to add a longer Ap memory to the model in order to increase fluxes in CIR storm tails. Due to its longer recovery phase, the Dst index presents a better means of capturing CIR storm tails, and issues with data gaps during positive Dst index could be eliminated by using the pressure

corrected Dst* (Equation 2.3). This would, however, restrict the dataset to years after ∼ 1960.

- The decreased Ap sensitivity for Ap levels above approximately 40 might also question the model's applicability to previous solar cycles which were associated to generally higher Ap values throughout the entire cycle.
- In addition, contaminating proton fluxes were not corrected for detector degradation which would not be complicated to implement.

Given that the Ap-based parametrization is the first model to attempt ascribing electron fluxes to Ap (and Dst) values, its overall performance can be considered successful. It shows great potential with regard to improvements of temporal and longitudinal resolutions if some of its cavities are addressed. Its major advantage compared to satellite measurements is its very long time-span as the Ap index can be reconstructed until 1850 (Matthes et al., 2017), whereas satellite measurements cover merely the past four decades.

4.3 Applicability of POES/MEPED LC electron fluxes during SPEs

One of the main reasons why implementing a proxy-based parametrization as input for EEP in models might be preferable is the presumed unreliability of POES/MEPED fluxes during SPEs. Increased contamination of the electron detectors by medium energy protons presents a difficult issue that leaves corrected electron data unreliable and leads to data gaps and uncertainties. A basic assumption when using a parametrization like the Ap model instead of actual real-time flux data is that SPE electron fluxes behave identically to regular storm electrons. The Ap index, however, reflects the geomagnetic activity on a planetary scale. If increased SPE impact on the magnetosphere and the polar cap accounts for large parts of the Ap increase, electron fluxes might be effectively lower during SPEs when compared to a regular geomagnetic storm or sub-storm with identical Ap. In this section, the applicability of the POES/MEPED LC dataset during SPEs will be judged. The implementation of the new criterion for sorting out unreliable data points after correction for contaminating protons should improve the dataset's coverage and quality.

4.3.1 POES/MEPED LC data coverage during SPEs

During the investigated time period of 2003 - 2012, six strong SPEs occurred as stated in Table 4.5. The Halloween event, SPE1, has been mentioned before and SPE3 during September 2005 was shown in Figure 4.19. Two other SPEs took place in January 2005 (SPE2) and December 2006 (SPE4) and two strong SPEs occurred in 2012 (SPE5 and SPE6). The dates of the six SPEs' maximum strength as well as their overall strength are listed in Table 4.5 as well. As can be seen, the Halloween event is by far the strongest event in the examined time period followed by the two SPEs in 2012. During SPE5, NOAA 15 data was not available on DOY 24 and 25 and there is a data gap for all satellites prior to SPE6 onset covering DOY 60 - 67 (data is available starting DOY 68).

SPE	Maximum (DOY in year)	Strength (pfu)	mean Ap	mean Kp
SPE1	302 in 2003	29500	170	6.86
SPE2	17 in 2005	5040	67	5.39
SPE3	254 in 2005	1880	73	5.50
SPE4	341 in 2006	1980	19	3.07
SPE5	24 in 2012	6310	14	2.59
SPE6	68 in 2012	6539	44	4.00

Table 4.5: List of investigated SPEs and their characteristics (maximum, strength, mean Ap, mean Kp).

Four-day mean rejection rates during SPEs

In order to obtain first information on the degree of data rejection during the six SPEs, 4-day mean rejection rates were determined for different latitudes and longitudes. The four days cover one day prior to SPE maximum until two days after maximum strength. Data from these periods was binned according to CGM latitude (0.5° bin width) and longitude (7.5° width) for each of the six SPEs. In each bin, data rejection rates were calculated and plotted on azimuthal projections of the NH as shown in Figure 4.21. Yellow regions represent high degrees of data rejection whereas in blue areas most electron flux data points are considered reliable (analogous to e.g. Figure 4.9). The six SPEs exhibit in general similar patterns with higher rejection rates in low latitudes as electron fluxes are low in these regions. All six SPEs show a blue band where rejection rates are below 10 %. This band, however, is slightly weaker in the evening sector compared to the day-side. Also, the Halloween event seems to have generally higher rejection rates than the other, weaker SPEs. At high latitudes, yellow colours dominate as proton fluxes become increasingly strong towards the polar cap. In magnetic latitudes above 75°, the determination of LC fluxes is not possible, leaving the polar cap yellow. Differences in satellite coverage are visible when comparing the first four SPEs with the two 2012 events. Whereas NOAA 17 covers the pre-midnight sector in earlier years, a large gap spanning MLTs 19 - 01 is visible for 2012. Based on Figure 4.21, it seems as if rejection rates were generally low in the auroral zone, but analysing 4-day mean values might be a crude approximation. In order to get a better grasp on data coverage during the different SPEs, the varying auroral oval position should be taken into consideration and shorter time periods should be investigated.

Six-hour rejection rate timelines and associated auroral oval position

Geomagnetic activity is known to vary in shorter time scales than four days and satellite coverage might change throughout one SPE. In addition, it is important to gain intelligence on data coverage specifically in the auroral oval, where electron fluxes are high. For this purpose, the Kp parametrization for the auroral oval position (Section 3.5) yields an estimate for the location of high electron fluxes which are the actual subject of interest. If, for instance, rejection rates are high outside of the auroral oval, little harm is done.

For inter-comparison between rejection rates and the current auroral oval positions, 6-hour mean Kp values are used to assign each 6-hour interval to a Kp category. Based on the Kp parametrization, NH azimuthal plots of the auroral intensity relative to its meridional peak intensity, similar to Figure 3.8, are obtained. The relative auroral intensity is colour-coded with yellow standing for high electron fluxes and blue regions presenting low fluxes. In order to get a direct comparison, data rejection rates originating from NOAA 15, 16 and 17 are overlaid in the satellites' trajectories. Like before, blue stands for low rejection rates and yellow for high data rejection. This is done for

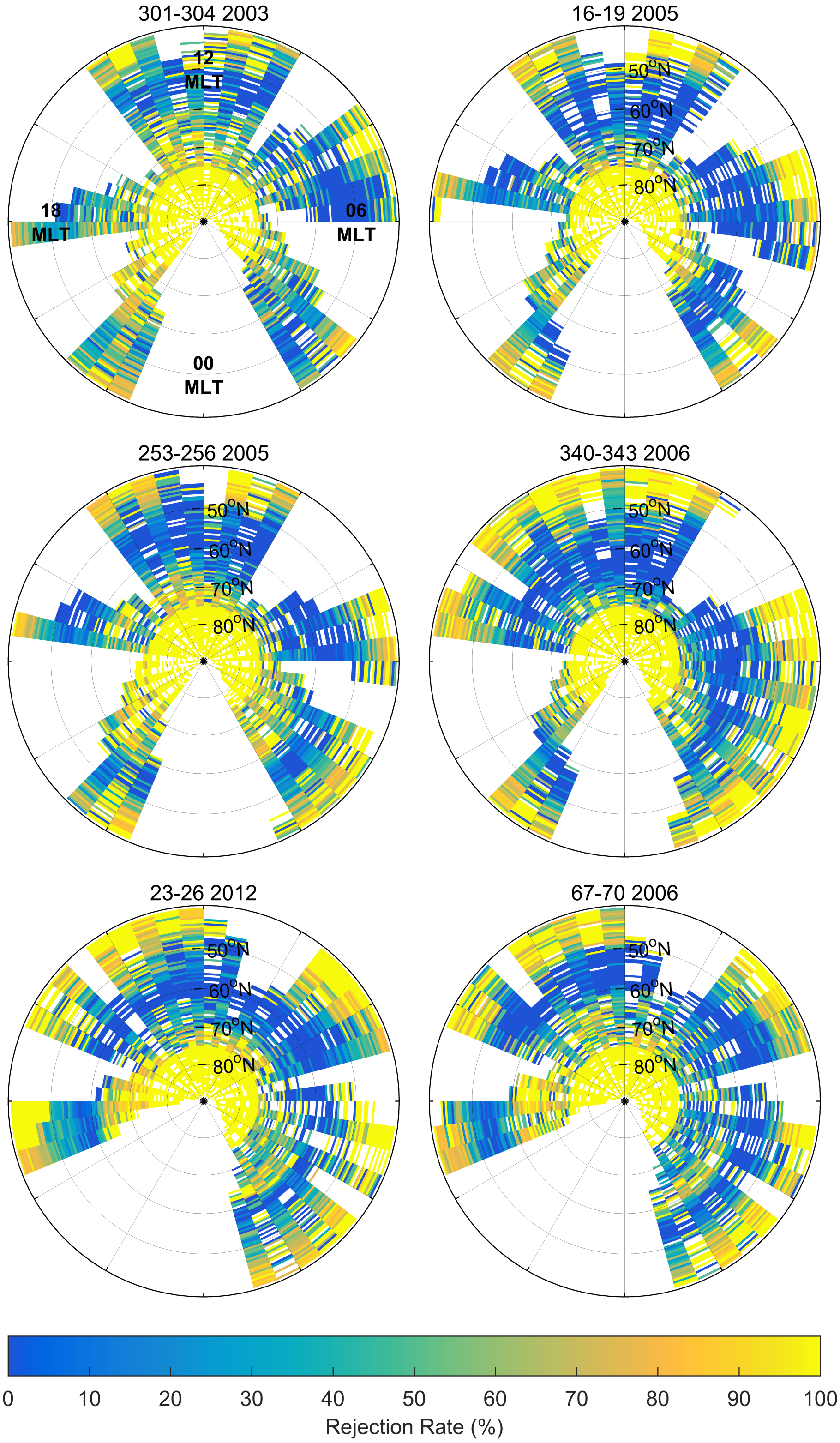
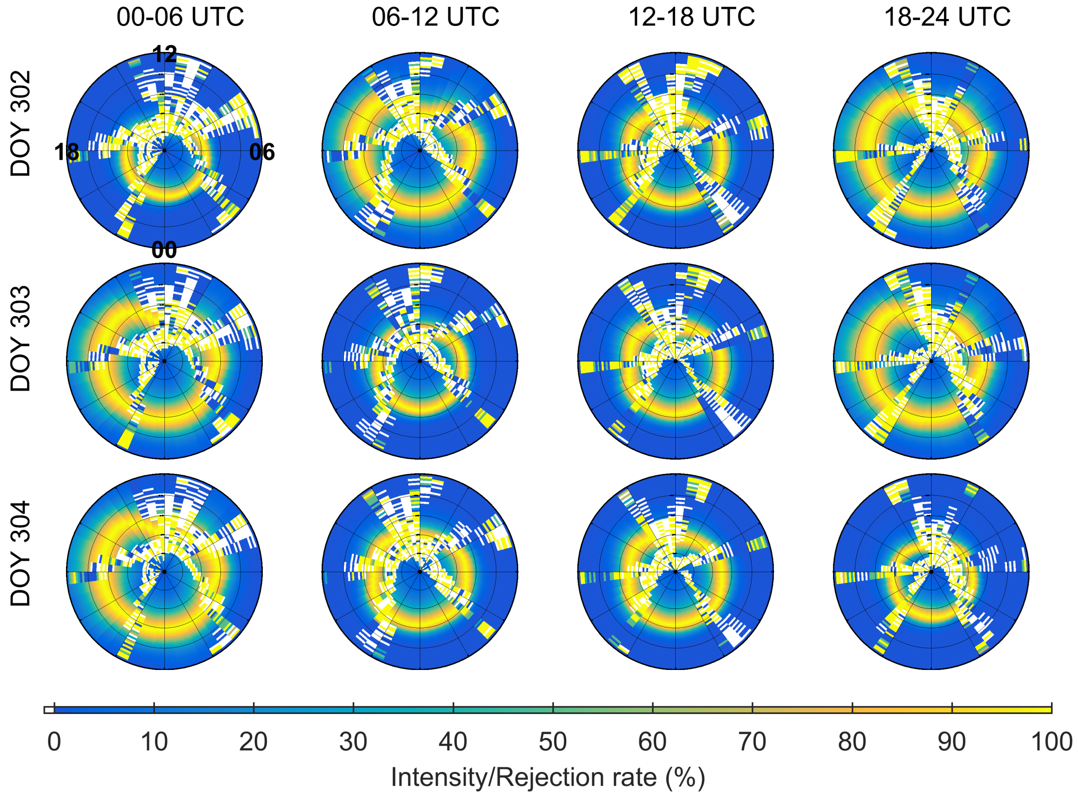
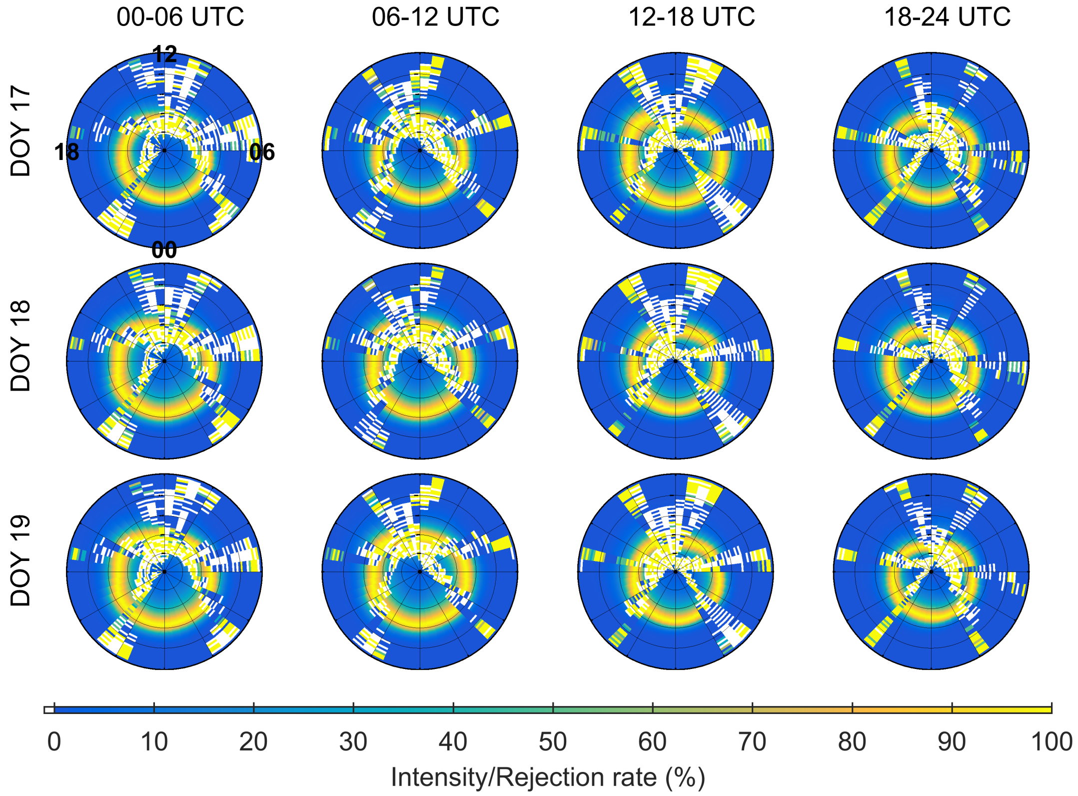


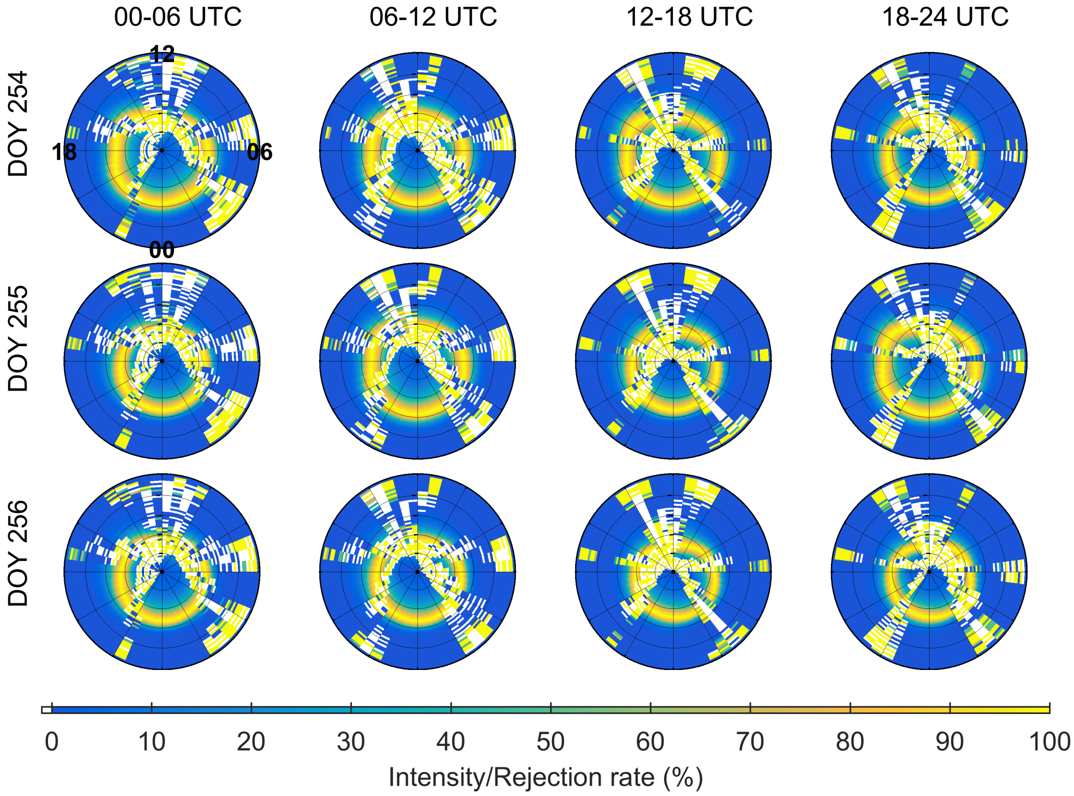
FIGURE 4.21: NH azimuthal projections of the 4-day mean LC flux rejection rate (in percent) during the six investigated SPEs. The CGM latitudinal and longitudinal resolutions are 0.5° and 7.5°, respectively.



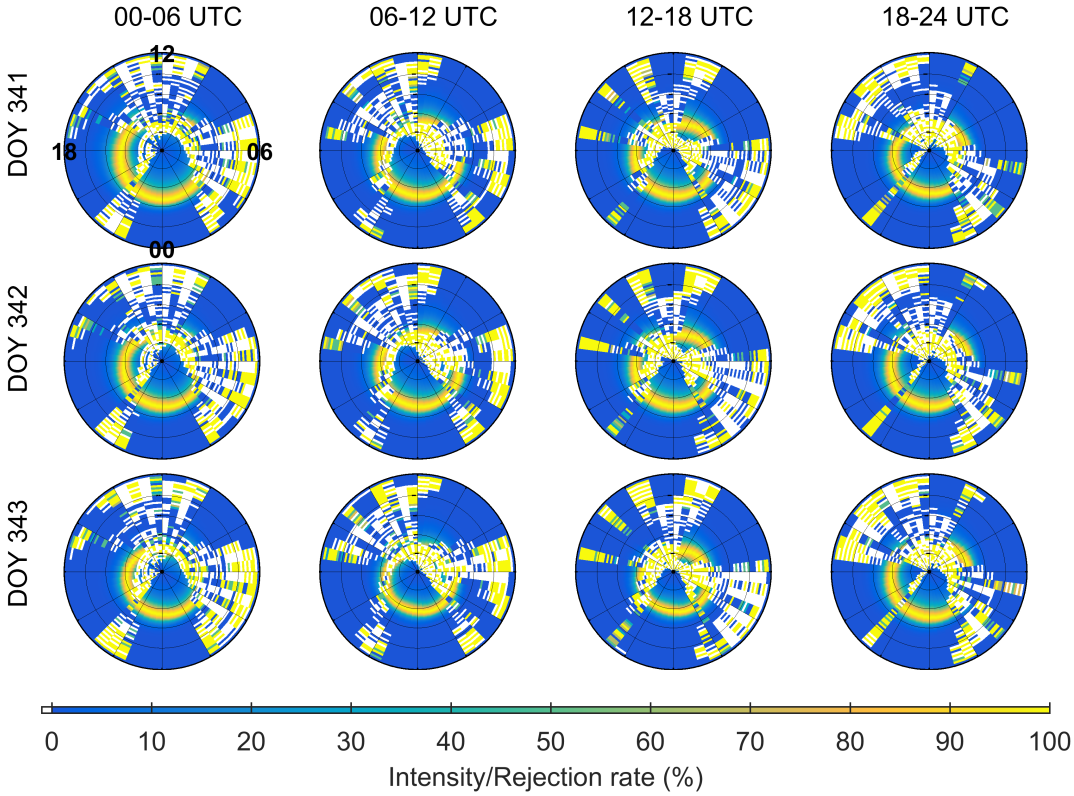
(a) DOY 302 - 304 in 2003.



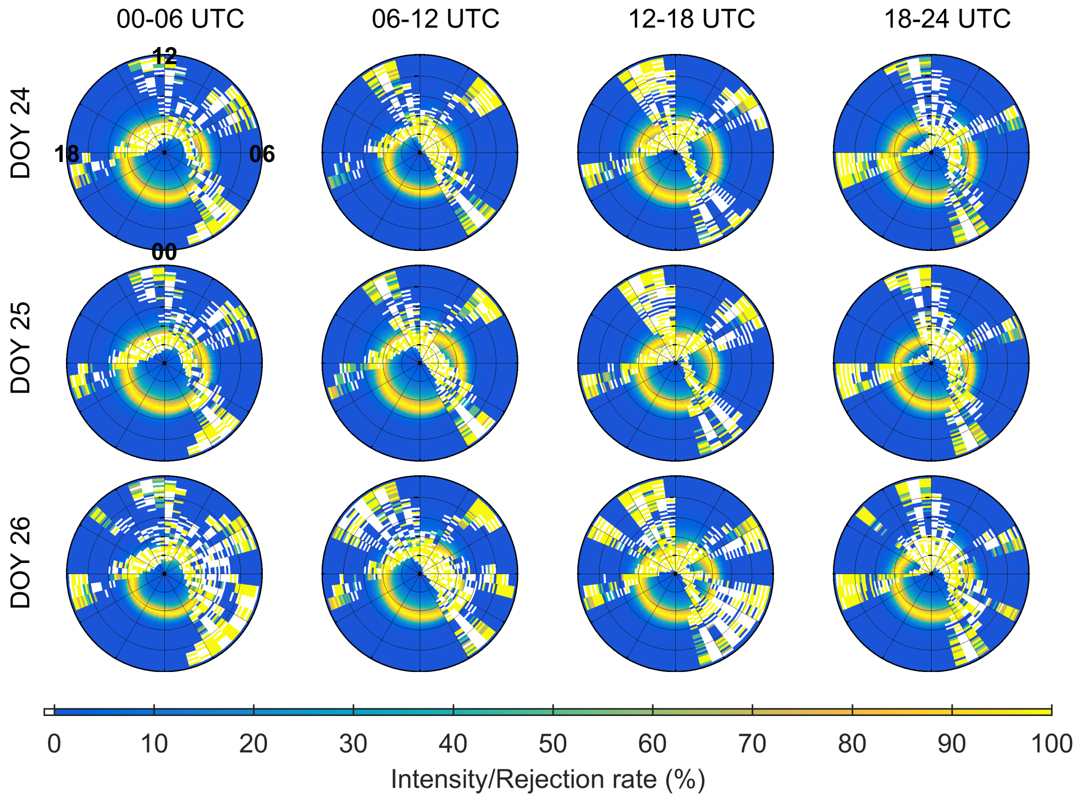
(b) DOY 17 - 19 in 2005.



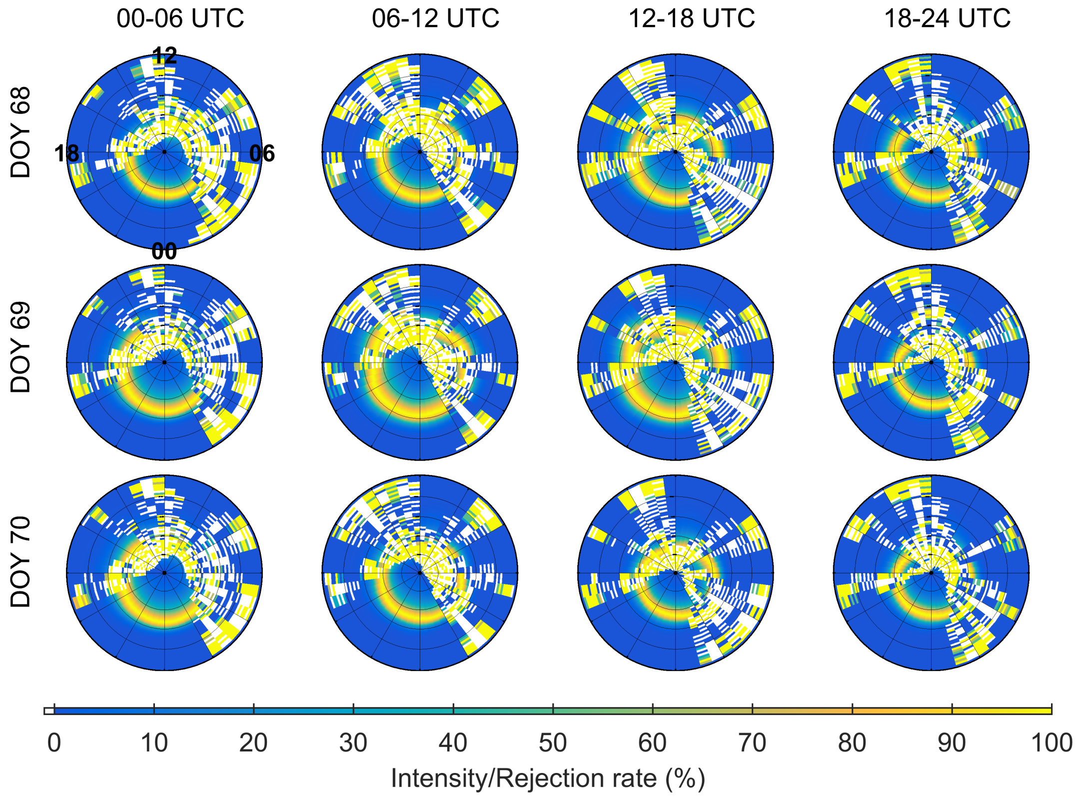
(c) DOY 254 - 256 in 2003.



(d) DOY 341 - 343 in 2005.



(e) DOY 24 - 26 in 2012.



(f) DOY 68 - 70 in 2012.

Figure 4.22: Time-line through the six SPEs showing the estimated auroral oval position overlaid by the LC flux rejection rate on the satellites' trajectories with a temporal resolution of 6 hours, a longitudinal bin width of 7.5° and a latitudinal resolution of 0.5 and 1° for the auroral oval position and rejection rate, respectively. A time period of three days after SPE onset is covered.

6-hour intervals starting on the day of maximum SPE strength covering in total three days. The resulting timelines for each of the six SPEs are shown in Figure 4.22. The Kp parametrization is shown with a CGM latitudinal and longitudinal resolution of 0.5 and 7.5°, respectively. For rejection rates, the latitudinal resolution was increased to 1°. White-coloured bins mark locations the satellites did not pass or at which no measurements were taken.

Ideally, data rejection should be low in areas of high auroral intensities. The wanted configuration in Figure 4.22 is therefore blue rejection rates on yellow auroral intensity. When examining the six timelines, large spread of data is visible where satellites cross different MLTs. Although 6-hour intervals and 1° latitudinal resolution were chosen, many white bins appear where no measurements are available. This results partly from taking 32-second averages during data processing to eliminate statistical fluctuations. As expected, rejection rates are quite high in lower latitudes but as the auroral oval is located north of these regions, no significant electron data is lost. Investigating rejection rates in the auroral oval, large variability between satellites and time intervals is visible. Satellites located in the evening sector seem to struggle more with large rejection rates compared to satellites in the morning sector due to high abundances of contaminating protons and lower electron fluxes. The morning sector, however, exhibits blue patches in the lower half of the auroral oval also during strong disturbances, whereas rejection rates rise with increasing latitudes.

Meridional rejection rate profiles for different satellites

The question how data rejection rates from different satellites, i.e. different locations, vary with latitude and Kp index is central to answering whether POES/MEPED electron fluxes can entirely or at least partly be considered usable during SPEs. Concerns are justified for evening or day-side measurements where contamination by protons is likely to be relatively high compared to the electron flux strength, but measurements from the morning sector might show potential as electron fluxes peak in these MLTs. Figure 4.23 shows meridional rejection rate profiles for pre-noon NOAA 15, post-midnight NOAA 16 and pre-midnight NOAA 17 data in the 3-hour Kp category 5. This corresponds to Kp values ranging from 6.0 to 8.0 which classifies as strong disturbances. Accordingly, the auroral intensity profile is located quite close to the equator. Note that due to the fact that NOAA 15, 16 and 17 pass different MLT sectors, the chosen auroral profile is estimated as the average of the single MLT profiles. Data from all six SPEs was superposed in order to increase data coverage in each latitude. This means that data originating from all six SPEs - considering a 3-day interval beginning with the SPE maximum day - was used. Whenever a 3-hour Kp value fell into one of the categories, the available satellite data was considered.

As can be seen, NOAA 17 rejection rates are equally high over the whole auroral oval

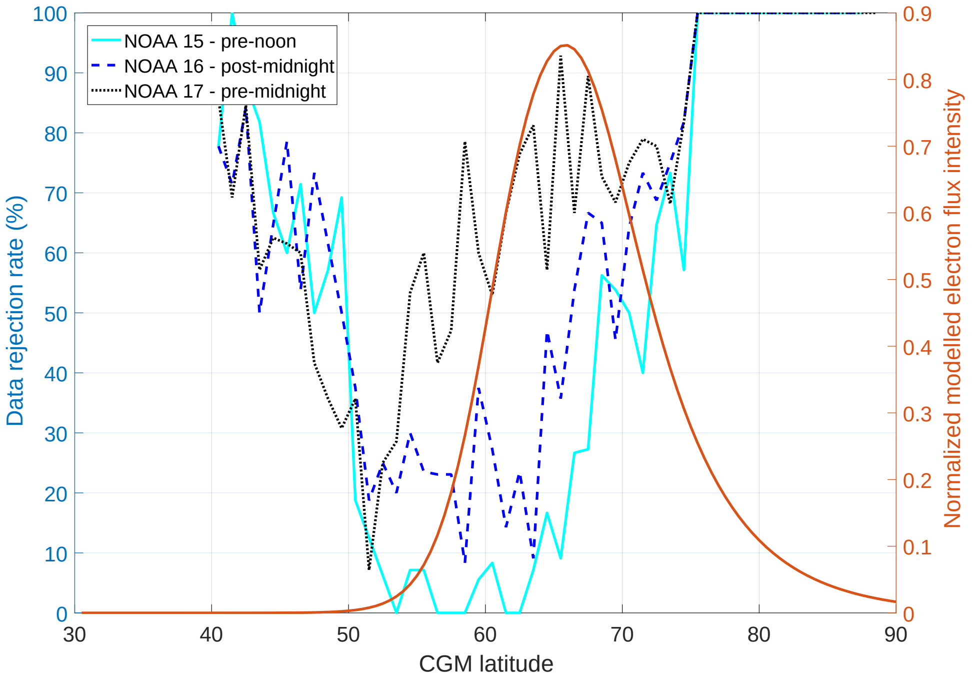


FIGURE 4.23: Kp category 5 rejection rates (3-hour mean) binned into 1° latitudinal bins and averaged. This is done for NOAA 15 pre-noon, NOAA 16 post-midnight and NOAA 17 pre-midnight data and shown together with the estimated auroral oval intensity from the Kp parametrization.

and beyond ranging at approximately 80 % in the centre of the profile. They drop at roughly 53° geomagnetic latitude just to rise again for lower latitudes and lower electron fluxes. NOAA 15 and 16 rejection rates, however, look acceptable being less than 10 and 30 %, respectively, in the equatorward half of the auroral oval despite the strong geomagnetic disturbance. As expected, rejection rates rise for higher latitudes where solar protons are able to penetrate open and closed field lines. Figures similar to the one shown in Figure 4.23 can be found in Appendix F for all six Kp categories. As can be seen there, NOAA 15 and 16 rejection rates in the auroral oval are relatively high during weak disturbances but their performance improves for higher Kp values and categories because the increase in Kp index is likely accompanied by rising electron fluxes. NOAA 17 datasets on the other hand seem to struggle with high rejection rates regardless of CGM latitude or Kp level. Whereas pre-noon NOAA 15 and post-midnight NOAA 16 rejection rates fall below 50 % for large portions of the auroral oval, pre-midnight NOAA 17 rejection rates barely reach values below 50 %.

Integrated rejection rates in the equatorward half of the auroral oval

One could argue that analysing the six SPEs simultaneously might distort the picture as all six events differ somewhat with regard to strength, year and Kp range. As a way of treating the investigated SPEs separately in a synoptic way, equatorward half auroral oval rejection rates and their evolution throughout the storms were examined. The equatorward half of the auroral oval was defined to range from the intensity peak to the equatorward position were the relative intensity dropped below 0.1. Afterwards, rejection rates in 3-hour intervals were integrated over the equatorward half of the current auroral oval position. This was done separately for pre-noon NOAA 15, post-midnight NOAA 16 and pre-midnight NOAA 17 data covering in total 24 three-hour time intervals per SPE, i.e. three days including the day of maximum strength. In addition, one treatment examined the equatorward half rejection rate of data from all available satellites. Then it was checked for how long the rejection rate stayed below

SPE	RR < 30 %				RR < 50 %				RR < 70 %			
	N15	N16	N17	all	N15	N16	N17	all	N15	N16	N17	all
SPE1	0.17	0.08	0.00	0.04	0.25	0.21	0.08	0.21	0.50	0.50	0.17	0.42
SPE2	0.46	0.33	0.13	0.25	0.71	0.54	0.33	0.58	0.96	0.79	0.67	0.83
SPE3	0.17	0.13	0.04	0.04	0.58	0.25	0.21	0.21	0.75	0.83	0.42	0.67
SPE4	0.63	0.08	0.17	0.21	0.96	0.54	0.29	0.75	1.00	0.83	0.88	1.00
SPE5	0.00	0.00	0.00	0.00	0.17	0.00	0.04	0.08	0.33	0.21	0.17	0.67
SPE6	0.25	0.04	0.04	0.04	0.50	0.08	0.08	0.33	0.75	0.42	0.21	0.71

Table 4.6: Absolute fraction of time during the central three days of SPEs 1 - 6 when integrated rejection rates in the equatorward half of the auroral oval were below 30, 50 and 70 %, respectively. Fractions were determined for post-midnight NOAA 15, pre-noon NOAA 16, pre-midnight NOAA 17 and a superposition of all available data.

30, 50 and 70 %, respectively. The resulting fractions with respect to total investigated time, i.e. three days, are stated in Table 4.6.

- During the Halloween event, for example, the equatorward half rejection rate of pre-noon NOAA 15 data stayed below 30 % for 17 % of the time. Considering data from all available satellites (NOAA 15 - 17) in all MLTs, this was the case only 4 % of the time. When asking for the rejection rate to stay below 50 %, pre-noon NOAA 15 data meets this demand 25 % of the time during **SPE1**. In general, post-midnight NOAA 16 rejection rates do not reach equally low levels as NOAA 15 rates. However, rejection rates from both satellites stay below 70 % throughout half the event.
- Numbers look entirely different for **SPE2** while the general trend - NOAA 15 being most promising while pre-midnight NOAA 17 being the satellite struggling most with high rejection rates - is preserved. Pre-noon NOAA 15 rejection rates stay below 30 % alomost during half of the investigated event.
- **SPE3** rejection rates look quite similar to numbers describing SPE1. Rejection rates are barely below 30 % for any of the three satellites but pre-noon NOAA 15 rejection rates are below 50 % throughout half of the event. In contrast to SPE2, where rejection rates based on all available data seemed promising, these numbers look worse for SPE3 (now NOAA 15 - 18). During 80 % of the investigated time, global rejetion rates were higher than 50 %.
- **SPE4** was one of the weaker SPEs regarding proton flux strength compared to the other five examined events while also being accompanied by quite low geomagnetic indices. As this would suggest low geomagnetic activity and electron fluxes, high rejection rates could be the consequence. Values shown in Table 4.6, however, show that pre-noon NOAA 15 rejection rates are below 30 % for 63 % of the time and below 50 % almost throughout the whole event. Whereas post-midnight NOAA 16 rejection rates do not fall below 30 % much, they lie below 50 % throughout the better part of the event. Pre-midnight NOAA 17 seems to struggle as seen in other events, reaching rejection rates below 50 % only one third of the time.
- When interpreting numbers referring to **SPE5**, it is important to keep in mind that NOAA 15 exhibits a data gap covering DOY 24 and 25. Thus, pre-noon NOAA 15 rejection rates stated in Table 4.6 are forced to higher values. None of the satellites measured a 3-hour interval in which the rejection rate fell below 30 % and rejection rates were barely below 50 % either. While rejection rates stayed below 70 % one third of the time for pre-noon NOAA 15 and roughly one fifth of the time for post-midnight NOAA 16 and pre-midnight NOAA 17, global rejection rates seem more promising. Although these rejection rates stay below 70 % for 67 % of the time, they barely fall below 50 % at all. Keep in mind that the NOAA 15 value for rejection rates below 70 % of 0.33 means that all

3-hour intervals for which measurements were available fall into this category (as data is not available 67 % of the time).

- Although comparably strong as SPE5, **SPE6** is related to better numbers for pre-noon NOAA 15 which measures with rejection rates below 50 % half of the event. Post-midnight NOAA 16 rejection rates, however, barely fall below 50 % at all.

As suggest earlier, Table 4.6 demonstrates the widely differing performance of different satellites in different SPEs. There seems to be a number of local effects which can lead to lower rejection rates in one satellite's path (confined to a certain MLT sector) while rejection intensifies in other regions.

Comparison between global LC and Ap model fluxes at 61° CGM latitude

The results shown in Table 4.6 indicate that the use of MEPED/POES electron fluxes during SPEs should be performed with caution. If data is predominantly rejected, investigated time intervals might contain only few actual measurements which might not be representative for the average flux level. One would therefore prefer to keep analyses to regions with increased abundance of reliable data such as middle latitudes, i.e. the equatorward part of the auroral oval, ideally avoiding the use of evening sector fluxes.

A different way of securing the reliability of measured fluxes is to evaluate the zero-mean and nan-median fluxes simultaneously. As rejected values are replaced by zero in one and ignored by the other, their impact can be considered small whenever the two lines comply. The inverse conclusion, that a difference between zero-mean and nan-median fluxes implies high rejection rates, however, is not true because differences can also result from low electron measurements in the 0° detector leading to a replacement of associated LC fluxes by zero.

An attempt was made to compare global LC electron fluxes to the Ap model fluxes. As the Ap model only yields zonal averaged fluxes, the LC dataset should not be restricted to fluxes from a certain MLT sector. Comparing modelled fluxes to only a selection of satellites and MLTs would imply a systematic bias towards e.g. higher LC fluxes and bigger differences between model and measurement, in the morning-noon sector (as discussed in Section 4.2.3).

Figure 4.24 shows the temporal evolution of modelled as well as zonally averaged nan-median and zero-mean LC fluxes in the 4 - 4.5 L shell throughout the six investigated SPEs. A 9-day period is covered in each plot, starting three days in advance of the SPE maximum and ending five days after maximum intensity. The LC dataset consist of all available data during the respective SPE not excluding any satellites or MLT sectors. The right y-axis and orange line indicate the rejection rate associated to the chosen L shells. The analysis was limited to low L shells (4 - 4.5), i.e. latitudes, as rejection rates and data coverage are more fortunate in these regions when compared to higher latitudes. Modelled fluxes are provided on a daily basis whereas LC fluxes were chosen

FIGURE 4.24: Timelines showing the Ap model and corrected zonal mean POES E1 LC fluxes during the four investigated SPEs in the L shell bin 4 - 4.5. The 3-hour averaged data rejection rate is depicted by the dashed line.

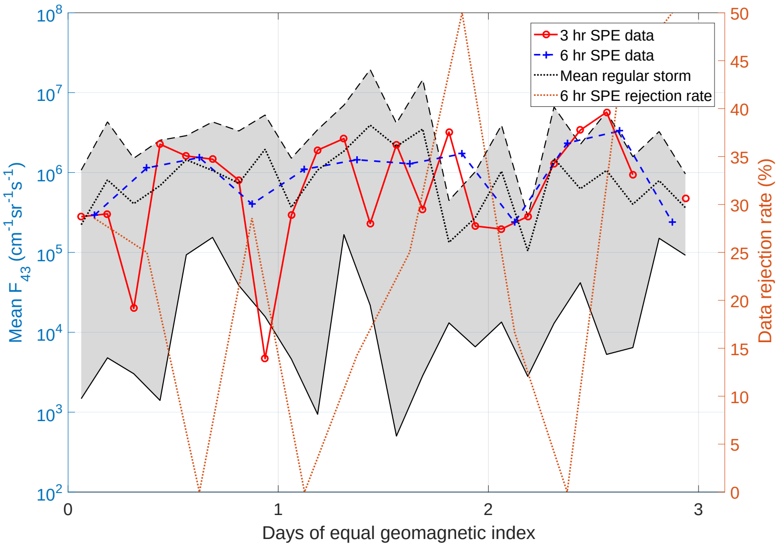
as 3-hour median and mean fluxes, respectively. Gaps in the nan-median LC flux occur in connection with incidents when the rejection rate approaches 100 % or whenever 0° fluxes qualify as noise. Comparing the six events, only SPE1 (Halloween event) exhibits extended periods of elevated rejection rates. For SPE2, SPE3, SPE5 and SPE6, rejection rates stay below 50 % for the better part of the main phase and rejection rates during SPE4 do not even reach 50 %. This indicates that the chosen latitudes are quite beneficial for examination as they are related to generally low rejection rates. As suggested by Table 4.6, rejection rates increase immediately when also taking into account higher latitudes up to the location of maximum auroral intensity.

The previously mentioned data gap in 2012 covering DOY 60 - 67 is visible in the lowest panel in Figure 4.24. SPE2, which took place during January 2005, was succeeded by a number of additional SPEs later that month, one of which is visible in increased rejection rates starting DOY 22.

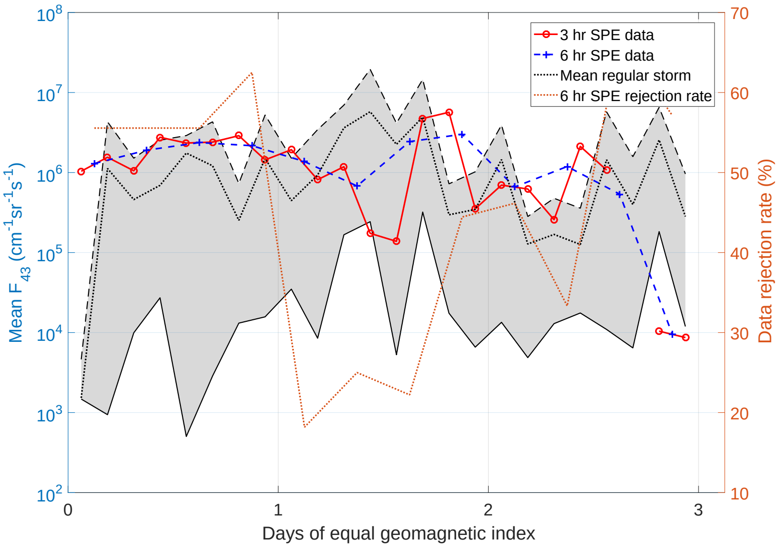
Apart from deviations associated with the general underestimation of quiet time fluxes previous to the initial phases of SPE1, SPE2 and SPE3, LC flux levels during the associated main phases are captured well by the modelled fluxes. In the SPE6 main phase, modelled and LC flux levels comply as well. Short-term variability which is visible in both nan-median and zero-mean LC data, is however not reproduced by modelled fluxes which seem to saturate. During SPE4 and SPE5, which were weak events based on geomagnetic activity, a systematic difference between modelled and LC fluxes persists. In these cases Ap model fluxes lie at least one order of magnitude below both nan-median and zero-mean LC fluxes which is related to the low Ap index associated with SPE4 and SPE5.

4.3.2 Comparison of electron fluxes during SPEs and regular storms

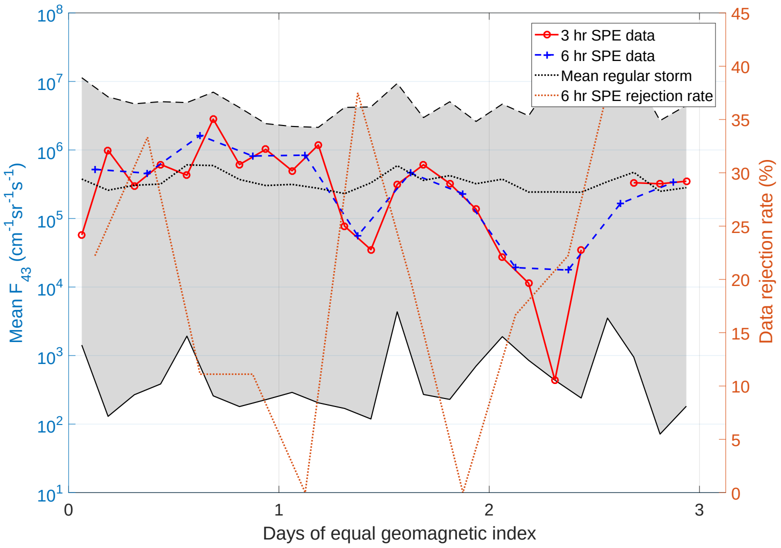
Having learnt about limitations to the application of LC fluxes during SPEs, the question remains whether electron fluxes during SPEs are of equal strength compared to electron fluxes during regular storms defined by the same level of geomagnetic activity. In order to minimize the effect of contaminating protons, only low L shell (5 - 5.5) and morning MLT (00 - 09) LC data was used during analysis. For each SPE, the 3-day mean Ap and Kp index were calculated (including the day of maximum strength until two days afterwards). The resulting values are listed in Table 4.5 along with other



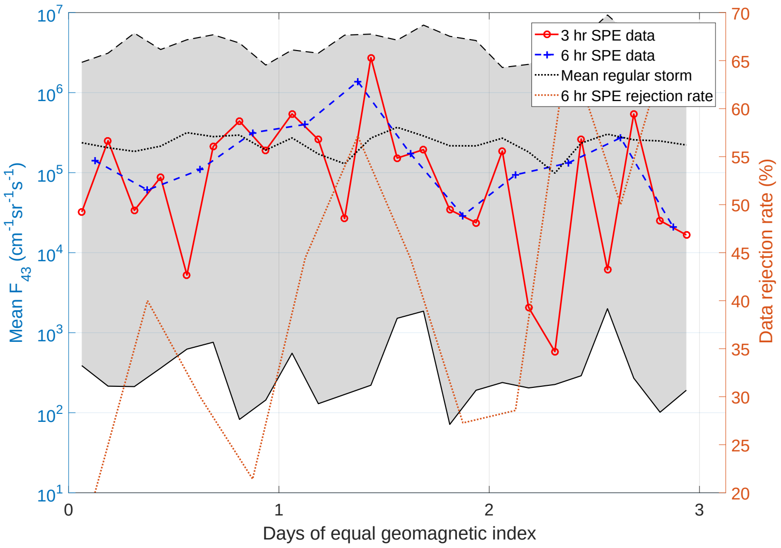
(a) SPE2.



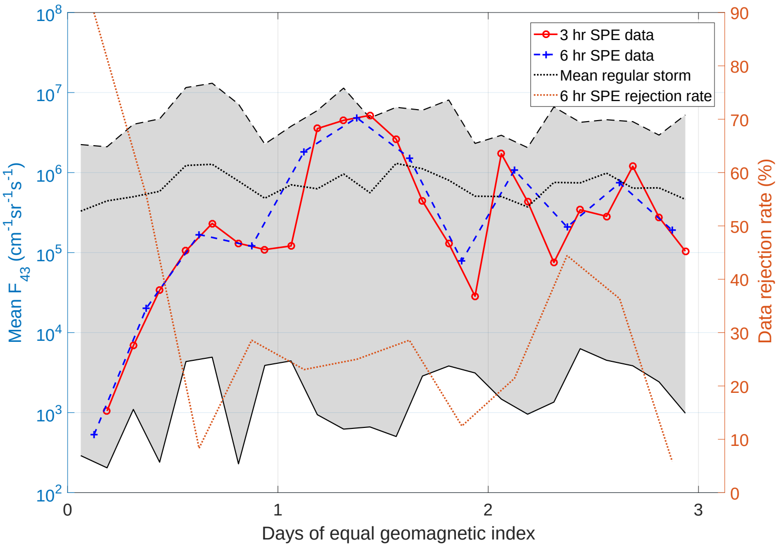
(b) SPE3.


Mean F$_{43}$ (cm^{-1}sr^{-1}s^{-1})
Data rejection rate (%)
Days of equal geomagnetic index
3 hr SPE data
6 hr SPE data
Mean regular storm
6 hr SPE rejection rate
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0
1
2
3
45
40
35
30
25
20
15
10
5
0

(c) SPE4.


Mean F$_{43}$ (cm^{-1}sr^{-1}s^{-1})
Data rejection rate (%)
Days of equal geomagnetic index
3 hr SPE data
6 hr SPE data
Mean regular storm
6 hr SPE rejection rate
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0
1
2
3
70
65
60
55
50
45
40
35
30
25
20

(d) SPE5.



(e) SPE6.

Figure 4.25: Comparison of E1 fluxes on L shells 5 - 5.25 during SPE2 - SPE6 with periods that were defined by equal 3-day running means. Six-hour and 3-hour SPE fluxes are shown (red and blue lines, respectively) on top of a grey background depicting regular storm variation. The dotted black line indicates the regular storm mean flux. Six-hour data rejection rate for the SPE data is shown as the orange dashed line.

characteristics. As mentioned earlier, SPE4 and SPE5 are by far the weakest events with regard to geomagnetic activity.

In order to assess if SPE electron flux levels are comparable to regular storm flux levels where the examined storms are defined by similar geomagnetic index values, 3-day running mean Ap and Kp values were determined for the period 2003 - 2012. A storm was considered regular as long as it was not listed as an SPE and 3-day periods were defined as similar to one of the six SPEs whenever their 3-day running means lay within a 10 % interval around the SPE's mean Ap or a 20 % interval around its mean Kp. Due to very high mean Ap and mean Kp values, no such events could be found for SPE1. The mean Ap value associated with SPE1 lies at 170 which is more than twice as high as the second largest mean Ap of 73 (SPE3). In fact, no non-SPE related 3-day period yielded an Ap value > 100 during the available 10-year time period.

However, the chosen regulations yielded six comparable 3-day intervals for SPE2, four for SPE3, 198 for SPE4, 309 for SPE5 and 55 for SPE6. As SPE4 and SPE5 were also defined by the lowest mean Ap and Kp, naturally more time periods qualified as comparable than for the other SPEs. Having found a list of comparable storms for SPE2 - SPE6, the 3-day intervals whose running mean index qualified them were separated into 24 three-hour intervals. For each interval, the mean flux was calculated. In order to obtain a boundary which reflects the spread of non-SPE related fluxes within each 3-hour interval, the maximum and non-zero minimum fluxes were chosen. Figure 4.25

contains similar timelines for each investigated SPE (2 - 6). The 3-hour zero-mean LC flux during the three central days of the SPE is indicated by the red line. Since it can contain gaps, i.e. zero fluxes, when rejection rates are high or electron fluxes are considered below the noise level, the 6-hour zero-mean LC flux is depicted as the blue dashed line as well. The underlying grey area marks the variability of the determined comparable regular storms. Their 3-hour mean flux is shown as the dotted black line and the grey area spans from the maximum flux to the non-zero minimum flux in each time interval. Of course, the considered regular storm fluxes were restricted to the same local region as the examined SPE fluxes. The data rejection rate associated with the chosen satellite data (L shells 5 - 5.5 and MLTs 00 - 09) is indicated by the orange line (right y-axis).

As can be seen, rejection rates can be considered rather manageable during all SPE central phases. 3-hour SPE fluxes are visibly smoothed by also taking into account 6-hour mean fluxes and data gaps are successfully filled. For all five SPEs the shown LC fluxes lie well within the upper and lower boundary of the grey area indicating regular storm behaviour. According to the LC data considered in this comparison, electron fluxes related to SPEs are not provably lower than flux levels during comparably strong storms. There are some cases where SPE flux peaks exceed the upper grey boundary which might be due to strong but distinct particle injections during the SPEs whose effect is averaged out by considering 3-day running mean activity.

Investigating fluxes at L shells 4 - 4.5 yields the same results, whereas considering L shells 7.5 - 8 proves difficult due to appearances of data gaps even for the 6-hour mean fluxes during SPE2. In the other cases, SPE LC electron fluxes lie well within the regular storm borders in these latitudes, as well. Data coverage is also very good when increasing the energy of the investigated electrons to E2 channel levels. Also for the E2 channel, SPE fluxes comply nicely to regular storm fluxes in low L shells (4 - 5.5) but exhibit too many data gaps in order to interpret flux levels in higher latitudes.

Thus, from the current analysis no evidence exists for SPE electron fluxes to be significantly lower than flux levels during regular storms with comparable geomagnetic indices. The very basic assumption upon which the Ap model and proxy-based electron flux models in general are based can therefore not be disproved. With regard to the usage of POES/MEPED fluxes during SPEs, contamination of electron fluxes by measured protons presents a serious restriction and should be handled with caution. If used at all, analysis should be restricted to regions of low contamination, i.e. low latitudes and morning sector MLTs.

Chapter 5

Conclusion and Outlook

This project's first objective was to find a way to improve the POES/MEPED dataset containing losscone electron fluxes by introducing a statistically based criterion for reliability of corrected electron data. The formerly used empirical criterion turned out to treat high flux data too strictly, rejecting statistically significant data. Data availability could indeed be improved by implementing a new criterion with an asymptotic confidence level of 8.5 %, but general limits to data coverage are set by other requirements and statistical uncertainties in periods of high proton fluxes. Nonetheless, LC flux availability during the strong Halloween SPE in 2003 (DOY 302 - 304) could be increased by up to 28 percentage points for NOAA 15.

In the study's second part, the obtained MEPED LC dataset was applied for inter-comparison with the EEP flux parametrization proposed for CMIP6. The modelled L shell dependent fluxes are solely predicted by the daily Ap index. Comparison with the LC dataset was focused on the general flux strength and the model's ability to reproduce both spatial and temporal variability on different scales. In conclusion, modelled fluxes exhibit a significantly lower flux level which is likely caused by the fact that the model is based on data from the MEPED 0° detector instead of losscone fluxes. As the 0° detector usually underestimates the actual losscone flux, this result is not entirely surprising.

Day-to-day variability in electron fluxes are well captured by the Ap model which manifests itself in good correlation coefficients. There seems, however, to be varying performance on a decadal basis with better flux levels but poorer variability during solar maximum and vice versa. Better resemblance of flux levels is owed to higher predicitive Ap values but as the model's sensitivity regarding changes in Ap decreases with increasing Ap and flux, it provides poorer fits concerning short-term variability. In addition, flux levels are captured better at higher latitudes, but short-term variability seems to be dampened.

One significant drawback of the model is that it does not contain any resolution in the longitudinal direction since it yields zonally averaged fluxes. Inter-comparison of LC data from different MLT sectors, however, clearly indicates the necessity of differentiating between fluxes in different MLTs. Evening fluxes are significantly lower on a general basis while local differences occur on the morning and day-side.

In addition, LC and modelled fluxes were compared regarding different types of storms.

During CME-driven storms, the Ap model struggles with capturing elevated flux levels in magnetic latitudes poleward of 64° where flux underestimations of half an order of magnitude are common. In addition, modelled fluxes saturate for the high Ap values associated with CME-driven storms forming plateau-like features during peak storm time and failing to reproduce short-term variability seen in losscone data. During CIR-driven storms, which are dominating in the declining and minimum phase of the solar cycle, the two dataset yield different storm lengths leading to an underestimation of model storm times by 50 % in 2008.

Eventually, the LC dataset's applicability during six SPEs was investigated. The risk of proton contamination is high during these events making the use of electron data a troublesome venture. Although the introduction of the new proton contamination criterion improved data coverage severely, large data gaps persist in certain MLTs and latitudes. The extent of data rejection in the vicinity of the auroral oval was examined by matching local data rejection rates to an estimated auroral oval position provided by a Kp-based model. As a result, evening and night-time electron rejection rates are high over the entire oval width, while post-midnight and pre-noon coverage is satisfying at least in the equator-ward half of the auroral oval. Data from corresponding MLTs and latitudes ranging between 61 and 64° was chosen in order to compare SPE electron flux behaviour to regular storm fluxes with similar accompanying geomagnetic indices. As the CMIP6 EEP parametrization relies on the assumption that the Ap index has the same predictive power for SPE and regular storm electron fluxes, studies supporting this assumption are vital. Indeed, based on the POES/MEPED LC fluxes investigated in this work, no significant differences could be established. Atmospheric impacts of SPE electrons should therefore be comparable to effects during regular storms. Although POES/MEPED electron data should be handled with care during SPEs and in regions of high proton fluxes, it is suitable for use in certain experiments. Due to high data rejection, in particular in high latitudes and the evening sector, it is adviced to restrict the application of this dataset to local regions, for example low latitudes or the morning sector. The evident issues with MEPED data on a global scale underline the necessity and importance of alternative approaches basing electron fluxes on proxies, such as e.g. an Ap model.

As mentioned in Chapter 4, the Ap model could easily be improved by using the LC dataset instead of data from the MEPED 0° detector. This would probably account for the evident differences in flux level and might counteract tendencies related to latitudinal, spectral and decadal variability. The evident issues with capturing CIR-storm lengths could be accounted for by introducing an Ap memory. It would also be important to test the differences in atmospheric ionization rates the two datasets evoke. Ionization does not necessarily exhibit a linear relation to flux strengths so that differences in flux levels might not be reflected by equally strong differences in ionization rates.

In addition, a model based on the pressure corrected Dst* (Equation 2.3) could account for issues which the current models exhibit during CIR-related storms.

The dataset upon which the Ap model is based consists of modified MEPED 0° fluxes from both hemispheres. The LC dataset used for comparison, however, containes only NH data. It would be instructive to check whether differences exist between the two hemispheres regarding latitudinal and spectral variability.

As the Ap model will be widely utilized via CMIP6, it is important to spread the results obtained from comparison with losscone fluxes. Researchers using EEP input in CMIP6 should be aware of potential biases due to local effects when studying the impact of EEP on the atmosphere and climate. Also results concerning the applicability of MEPED electron fluxes during SPEs are to be shared publicly as this is a topic which has been heavily debated throughout the past. The present study, however, is the first to put numbers on reliable data coverage in different polar regions. Data availability might still be improved in the lowest energy channel by allowing E1 data points to be kept although E2 and E3 fluxes identify as noise. Issues related to the calculation of losscone fluxes based on 0° and 90° measurements can lead to additional data rejection but can be accounted for by adapting the analysis routine, e.g. to differences in losscone width.

In conclusion, the MEPED LC dataset was successfully improved and used in order to identify drawbacks in the Ap model. The question as to whether MEPED data can be used during SPEs was answered and restrictions on the application of the dataset were proposed. Future work based on the obtained results will focus on the further improvement of data coverage, hopefully softening restrictions during SPEs, and spreading these results within the scientific community in order to suggest ways of improving the Ap model in the future and raise awareness during the current version's application.

List of Abbreviations

AO	**A**rctic **Oscillation**
AU	**A**stronomical **U**nit
BD	**B**rewer **D**obson
CCM	**C**hemistry **C**limate **M**odel
CGM	**C**orrected **G**eo**M**agnetic
CH	**C**oronal **H**ole
CIR	**C**orotating **I**nteraction **R**egion
CME	**C**oronal **M**ass **E**jection
DOY	**D**ay **O**f **Y**ear
EPP	**E**nergetic **P**article **P**recipitation
EUMETSAT	**Eu**ropean Organisation for the Exploitation of **Met**eorological **Sat**ellites
fu	**f**lux **u**nits
FUV	**F**ar **U**ltra**v**iolet
GCR	**G**alactic **C**osmic **R**ays
GUVI	**G**lobal **U**ltra**v**iolet **I**mager
HCSPS	**H**eliospheric **C**urrent **S**heat **P**lasma **S**heet
HILDCAA	**H**igh-**I**ntensity **L**ong-**D**uration **C**ontinuous **A**E **A**ctivity
HSSWS	**H**igh **S**peed **S**olar **W**ind **S**tream
ICME	**I**nterplanetary **C**oronal **M**ass **E**jection
IGRF	**I**nternational **G**eomagnetic **R**eference **F**ield
IMF	**I**nterplanetary **M**agentic **F**ield
ITCZ	**I**nter**t**ropical **C**onvergence **Z**one
LC	**L**oss**c**one
LT	**L**ocal **T**ime
MEPED	**M**edium **E**nergy **P**roton and **E**lectron **D**etector
MHD	**M**agneto**h**ydro**d**ynamics
MLT	**M**agnetic **L**ocal **T**ime
MS	**M**urgatroyd **S**ingleton
NAM	**N**orthern **A**nnular **M**ode
NAO	**N**orth **A**tlantic **O**scillation
NH	**N**orthern **H**emisphere
NOAA	**N**ational **O**ceanic and **A**tmospheric **A**dministration
pfu	**p**roton **f**lux **u**nits
POES	**P**olar **O**rbiting **E**nvironmental **S**atellite
PCA	**P**olar **C**ap **A**bsorption

PCHIP	**P**iecewise **C**ubic **H**ermite **I**nterpolating **P**olynomial
QBO	**Q**uasi **B**iennial **O**scillation
RR	**R**ejection **R**ate
SAA	**S**outh **A**tlantic **A**nomaly
SEM	**S**pace **E**nvironment **M**onitor
SEP	**S**olar **E**nergetic **P**articles
SI	**S**udden **I**mpuls
SO	**S**torm **O**nset
SPE	**S**olar **P**roton **E**vent
SSC	**S**torm **S**udden **C**ommencement
SSI	**S**olar **S**pectral **I**rradiance
SST	**S**ea **S**urface **T**emperature
SSW	**S**udden **S**tratospheric **W**arming
TED	**T**otal **E**nergy **D**etector
TIMED	**T**hermosphere **I**onosphere **M**esosphere **E**nergetics and **D**ynamics
TSI	**T**otal **S**olar **I**rradiance
UTC	**C**oordinated **U**niversal **T**ime
UV	**U**ltra**v**iolet

Appendix A

Explanatory note on 16-second accumulated counts and averages

Section 3.2.1 contains a statistical evaluation of the measurement process in the MEPED electron and proton detectors. During data processing, 16 measurements are added in order to calculate 32-second means. A full measurement requires 2 seconds because measurements are obtained every other second, which means that the resulting measurement period is 16 seconds. The temporal resolution of the calculated mean fluxes will however be 32 seconds. The 16 measurements used for averaging $\mathrm{N_i}$ each follow a Poisson distribution $\mathrm{P(N_{1s})}$ where $\mathrm{N_{1s}}$ is the true count rate per one second. The sum of the sixteen $\mathrm{N_i}$'s will yield the number of particles N measured during 16 seconds

$$N = \sum_{i=1}^{16} N_i \,. \tag{A.1}$$

As all 16 single measurements can be considered independent, it is valid that

$$N \sim P(16N_{1s} = N_{16s}) \tag{A.2}$$

with the true particle count number per 16 seconds, $\mathrm{N_{16s}}$. Applying the Gaussian approximation renders

$$N \sim N(\mu = N_{16s}, \sigma = \sqrt{N_{16s}}) \,. \tag{A.3}$$

The assumption that the electron detector count number, $\mathrm{N_{te}}$, and the proton detector count number, $\mathrm{N_{pp}}$, follow the above described normal distribution is valid for the investigated flux levels and equation 3.15 yields the condition

$$N_{te} > N_{pp} + c_\sigma \cdot \sqrt{N_{pp}} \,. \tag{A.4}$$

The corresponding criterion for particle fluxes instead of counts can be obtained by applying the conversion defined in Equation 3.11

$$J_{te} > \frac{G_{pp}}{G_{ee}} J_{pp} + c_\sigma \sqrt{\frac{G_{pp} J_{pp}}{16 G_{ee}^2}} \,. \tag{A.5}$$

Analogous calculations can be performed for the 16-second mean

$$\overline{N} = \frac{1}{16} \sum_{i=1}^{16} N_i \,. \tag{A.6}$$

As the statistical behaviour of a Poisson distributed quantity is not conserved under multiplication with a constant, a normal approximation is necessary before investigating the underlying distribution of the mean fluxes. For sufficiently high count rates, $\overline{\mathrm{N}}$ will follow a normal distribution with mean $\mathrm{N_{1s}}$ and standard deviation $\sqrt{\frac{\mathrm{N_{1s}}}{16}}$. It is important to note that the standard deviation has been improved compared to the original measurements $\mathrm{N_i}$. The resulting criterion based on count rates

$$\overline{N}_{te} > \overline{N}_{pp} + c_\sigma \cdot \sqrt{\frac{\overline{N}_{pp}}{16}} \tag{A.7}$$

can be converted to fluxes using Equations 3.10 and 3.1, yielding

$$J_{te} > \frac{G_{pp}}{G_{ee}} J_{pp} + c_\sigma \sqrt{\frac{G_{pp} J_{pp}}{16 G_{ee}^2}} \tag{A.8}$$

which is identical to the criterion deduced from the 16-second accumulated counts. As long as the correct conversion between count rate and flux is chosen, the two approaches yield identical results.

Appendix B

Origin of unphysically small contaminating proton fluxes

In Section 4.1.2, the impact of the different contamination criteria on POES/MEPED data was investigated. In this context, the distribution of flux data with respect to different electron and proton flux levels was illustrated in Figure 4.4 and Figure 4.4. Both clearly show a population of low proton fluxes which would correspond to less than one proton per 16-second measurement. As this is unphysical, the origin of these low proton fluxes has to be accounted for. During data processing, the contaminating proton flux is determined by applying a PCHIP interpolation to the integral proton fluxes in each of the five channels P1 - P5. PCHIP imposes a third degree polynomial between

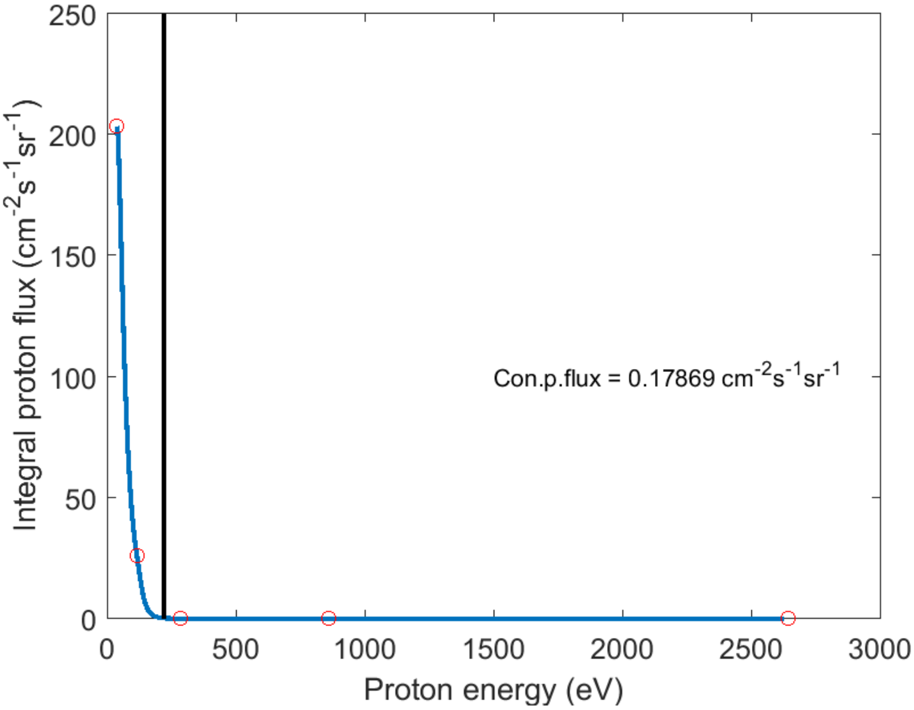


Figure B.1: Integral proton energy spectrum in a case that yields an unphysically low contaminating proton flux. Circles indicate the measurement positions, the line corresponds to the applied PCHIP interpolation and the vertical black line marks the proton contamination threshold in channel E1.

each measurement point. The obtained function corresponds to the presumed energy spectrum of the measured protons and might in some instances look like the case shown in Figure B.1. Since the higher energy channels do not measure any proton flux at the time, measurements in the two low-energy channels correspond to present low-energy protons. The lower energy threshold for contaminating protons in the E1 channel is marked in Figure B.1 by the black vertical line. Using the PCHIP value of this position and subtracting the PCHIP value of the upper limit will then render the contaminating proton flux. As the lower threshold is located between the P2 measurement of approximately four protons per 16 seconds and the P3 measurement which states zero protons per 16 seconds, the true proton number at the threshold is unknown but has to lie between zero and four. Since the slope to the right of the P2 measurement is quite steep and the PCHIP is not restricted to integer proton numbers, it claims a contaminating proton flux of 0.179 $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ which corresponds to 0.03 protons per 16 seconds. In cases like this, the proton thresholds lie within the non-zero tail of the integral spectrum. Whether or not the true integral flux at the threshold position is equal to zero or not is not possible to answer, as it only can be restricted to lying between the fluxes measured in the P2 and P3 channel. However, the occurrence of these low contaminating proton fluxes is mainly restricted to low P2 fluxes so that the true contaminating proton flux can also be considered low.

Appendix C

Possible origins of negative corrected flux values

Figure 4.4 exhibits the occurrence of measurements that will lead to negative corrected electron fluxes. To some extent, this has to be expected when electron fluxes are low and statistical variations can lead to higher proton fluxes in the electron detector compared to the proton detector. However, a certain number of measurements lies below the red dashed line which is supposed to set a lower limit on the statistically based negative values. As visible in Figure C.1, the occurrence of these measurements is not restricted to periods of enhanced fluxes, as unstatistical values are evident both in high activity years such as 2003 and years of lower geomagnetic activity, e.g. 2008. This is supported by values shown in Table C.1, which shows the fraction of unexplained data points with regard to the whole dataset in different time periods. The three upper rows correspond to SPEs and elevated flux levels of both protons and electrons. The lowest row shows the fraction for the whole year 2003. However, none of the values show a significant difference to each other. In addition, they occur both in the 0° and 90° detector, while to a lesser extent in the 90° detector. This suggests that the phenomenon does not depend entirely on differences in directionality. It is not an artefact from a single satellite as these negative values are observed for all satellites and their amount does not vary significantly between older and newer satellites. An example is shown in Figure C.2 depicting NOAA 19 measurements in 2010, one year after satellite launch. Thus, a defect in the correction for proton degradation can be ruled out as a potential cause as well. According to Figure C.3, influences by the energy level can be neglected

Time period	Fraction of unexplained negative data points
DOY 299 - 310 in 2003	2.3 %
DOY 15 - 21 in 2005	1.5 %
DOY 250 - 257 in 2005	1.4 %
DOY 1 - 365 in 2003	2.4 %

Table C.1: Fraction of unexplained negative data points of all available satellites in different time periods.

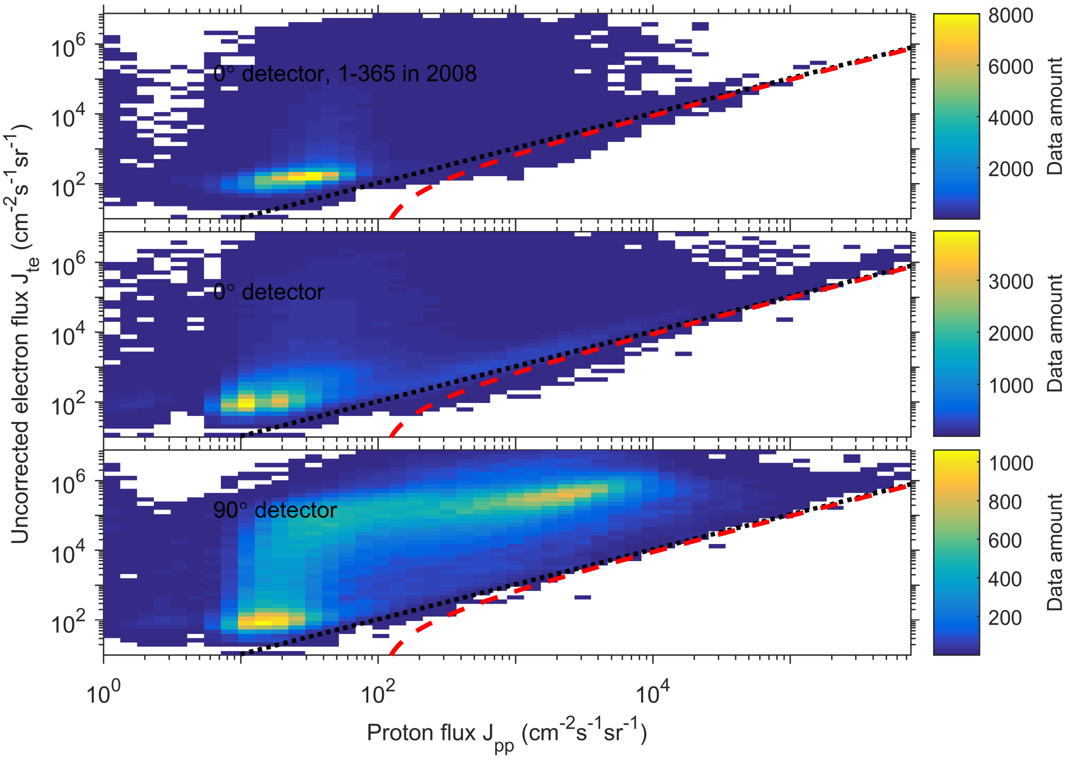


FIGURE C.1: The two upper panels correspond to data amount in the E1 channel of the 0° detector while the lower panel depicts the 90° detector of the MEPED instrument channel on board NOAA 15. The upper panel covers the time period of DOY 302 - 304 in 2003 and the two lower panels show DOY 1 - 365 in the same year.

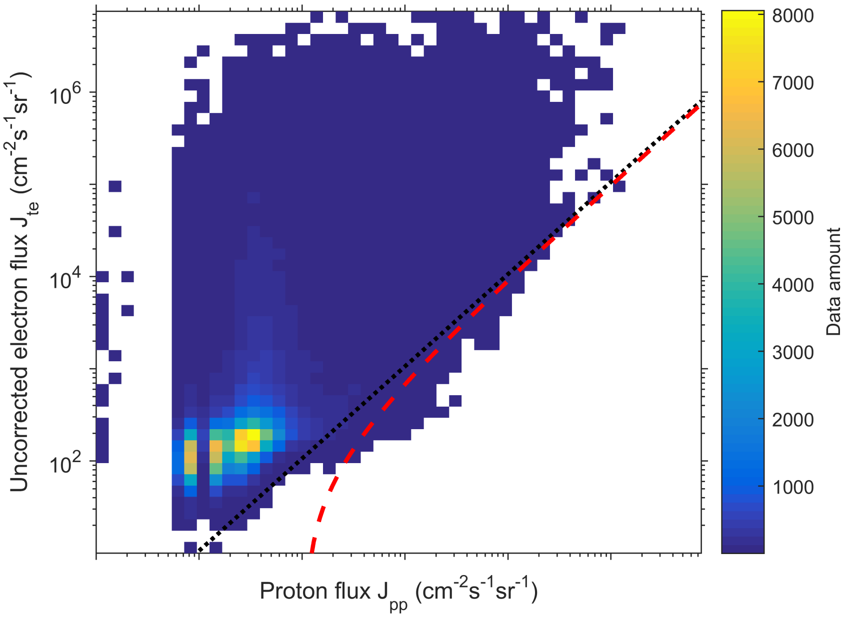


FIGURE C.2: Data amount and distribution in the E1 channel of the 0° detector on board NOAA 19 during DOY 1 - 365 in 2010.

as well, since similar amounts of unstatistically negative values are obtained in the E2 and E3 channel.

Another possibility concerns the geometric factors associated with the detectors' energy channels which are used during conversion from count rates to fluxes. Geometric factors are assumed to be constant on the energy interval where the detectors are sensitive to electrons and protons, respectively. If, indeed, the applied geometric factors were inaccurate, this could lead to an overestimation of the contaminating proton flux or an underestimation of the uncorrected electron flux resulting in a systematic bias towards smaller corrected electron fluxes. The geometric factors for electrons in the MEPED detector have been subject to extensive studies and are fairly well known (Yando et al., 2011). It is evident that the geometric factors will depend somewhat on the energy spectrum of the incident particles. Ødegaard et al. (2017) introduces optimized energy thresholds and associated geometric factors, shifting the threshold up in E1 and E2 (43 and 114 keV, respectively) and down in E3 (292 keV). The comparison of data distributions in E1, E2 and E3, however, does not suggest a underestimated electron flux in any particular channel. The geometric factor for protons in the electron detector investigated by Yando et al. (2011) coincides with its counterpart in the proton detector (0.0095 $\mathrm{cm^2 sr}$) for most proton energies. Differences persist in the lower end of the energy spectrum, where the geometric factor is lower than 0.0095 $\mathrm{cm^2 sr}$. Protons in these energy regions could then lead to an overestimation of the subtracted contaminating proton flux. This effect would be large for steep proton energy spectra and smaller for smaller slopes. The upper panel in Figure C.4 shows the data distribution in the E1 channel of the 0° detector on board NOAA 15 during 2003. The lower panel depicts the ratio between the proton fluxes measured in the P2 and P3 channel, respectively. A steep spectrum would be related to a large ratio. However, the ratios in the problematic bins is not significantly different from ratios of positive corrected electron fluxes.

When checking for characteristic behaviour with regard to flux levels, MLTs and latitudes, it turns out that high contaminating proton fluxes are more likely to be linked to these negative values. Also quite high uncorrected electron fluxes of up to $10^5\ \mathrm{cm^{-2}s^{-1}sr^{-1}}$ can be affected. The occurrence is mainly restricted to the pre-midnight sector and the day-side where proton fluxes are highest and electron fluxes decrease.

In conclusion, no clear origin of the unaccounted for negative data could be established. In case of a systematic error that leads to the overestimation of the contaminating proton flux or the underestimation of the uncorrected electron flux, the bias would be towards lower fluxes. This secures that the LC fluxes contained in the POES dataset are a lower estimate of the true fluxes, eliminating an overestimation of possible atmospheric effects.

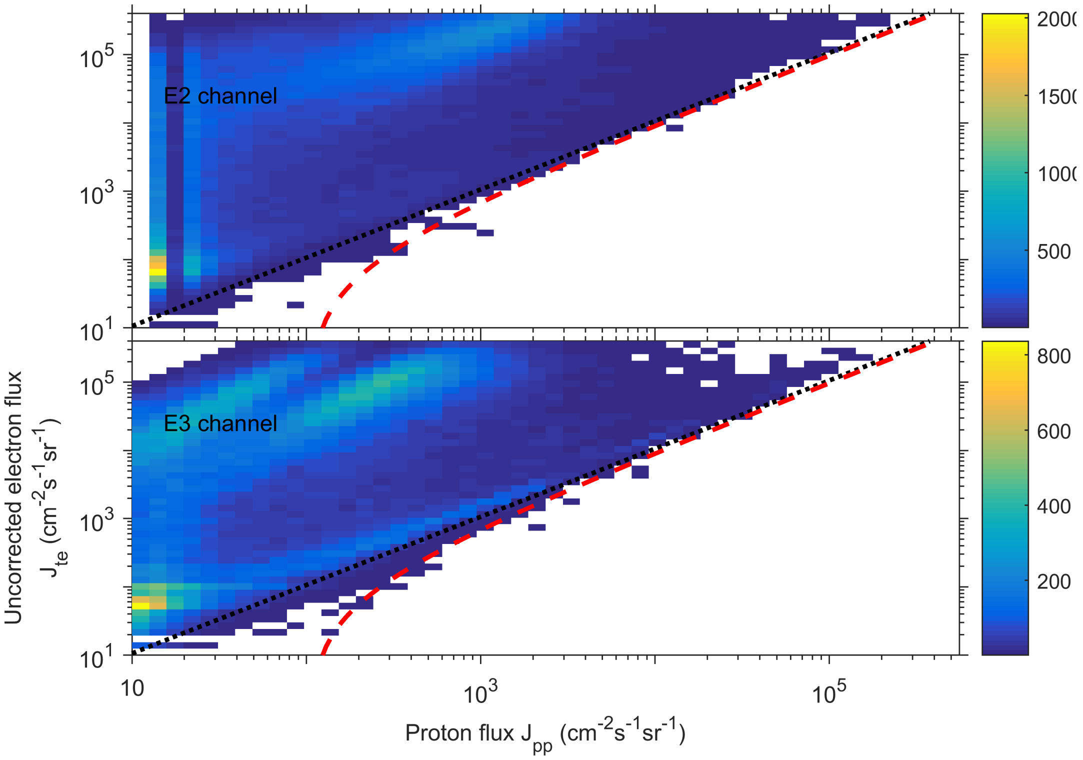


FIGURE C.3: Analoguous to Figure 4.4 for the E2 and E3 channel.

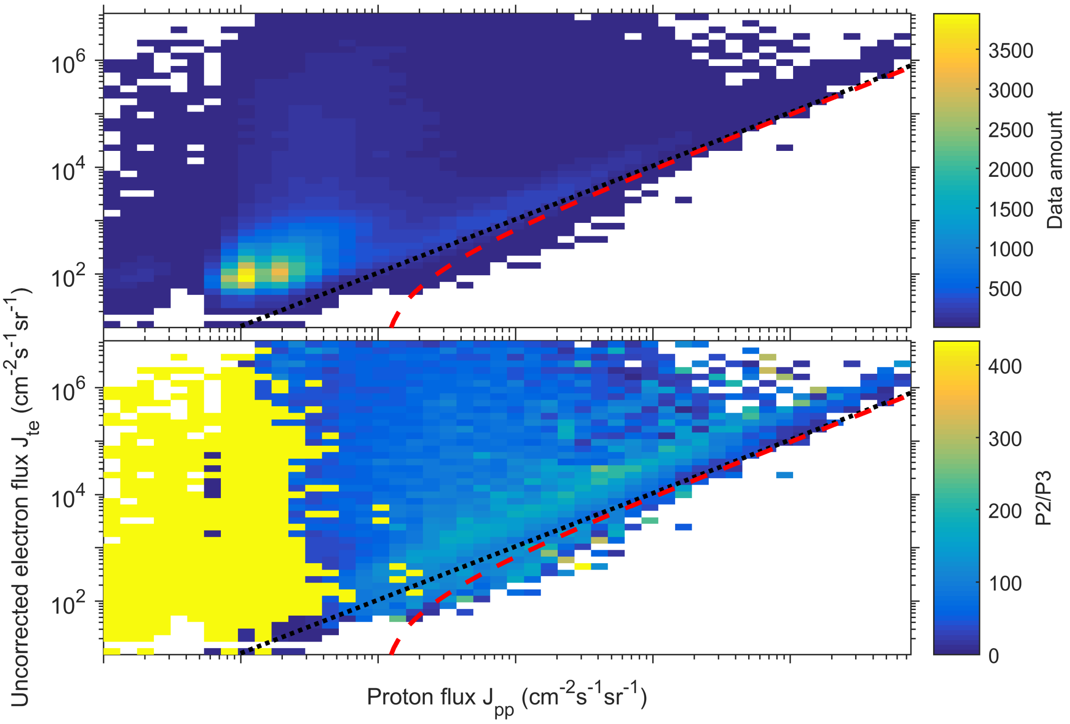


FIGURE C.4: The data amount in the E1 channel of the 0° detector on board NOAA 15 during DOY 1 - 365 in 2003 (upper panel) and associated ratio of the P2 and P3 proton fluxes in each bin (lower panel).

Appendix D

Maps showing local rejection rates and electron fluxes throughout different time periods in 2003

In Section 4.1 data rejection rates in the 0° detector and in the LC fluxes resulting from different proton contamination criteria (Requirement 2) were investigated during DOY 302 - 304 in 2003. Rejection rates are in general higher in the 0° detector when compared to the 90° detector due to different flux levels. This is evident from Table 4.1 and visible in Figure D.1 which shows 90° detector rejection rates during the same time

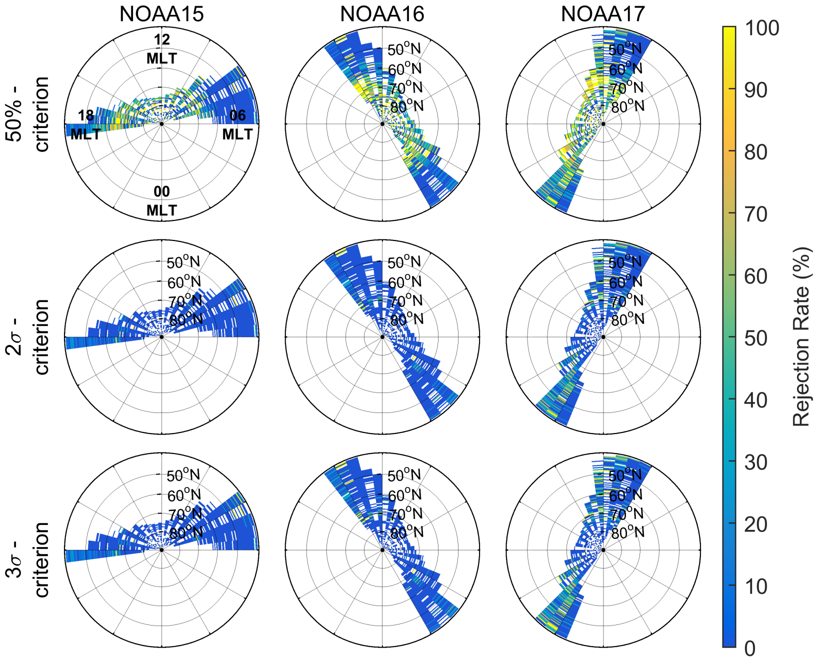


Figure D.1: Azimuthal projections showing the geomagnetic latitude and longitude dependence of the E1 - E3 superposed rejection rate based on measurements by the 90° detector on board NOAA 15 - 17 during DOY 302 - 304 in 2003. The latitudinal and longitudinal resolutions are 0.5 and 7.5°, respectively. (Analogous to Figure 4.8 and Figure 4.9.)

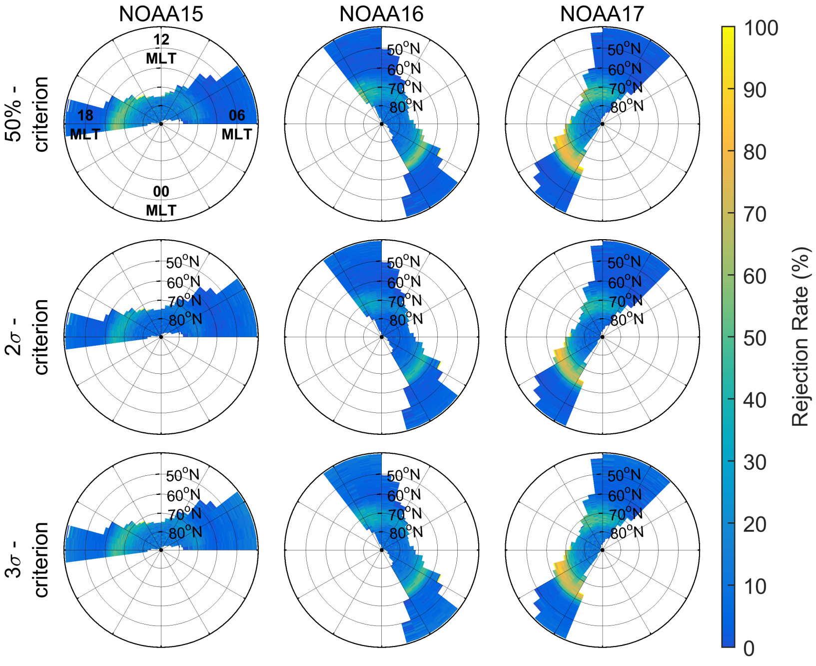


FIGURE D.2: Azimuthal projections showing the geomagnetic latitude and longitude dependence of the E1-E3 superposed rejection rate based on measurements by the 0° detector on board NOAA 15 - 17 during DOY 1 - 365 in 2003. The latitudinal and longitudinal resolutions are 0.5 and 7.5°, respectively.

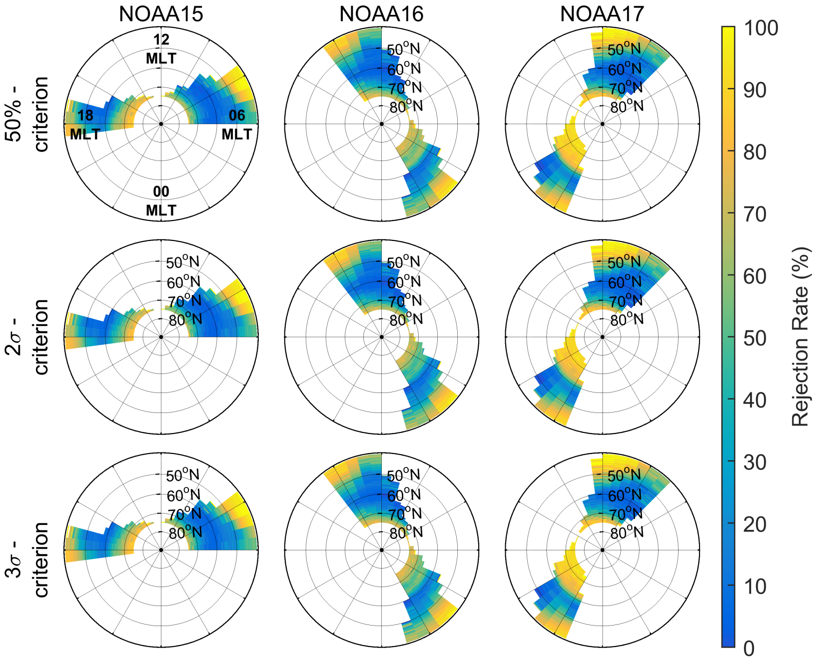


FIGURE D.3: Azimuthal projections showing the geomagnetic latitude and longitude dependence of the E1 - E3 superposed LC flux rejection rate based on measurements by NOAA 15 - 17 during DOY 1 - 365 in 2003. The latitudinal and longitudinal resolutions are 0.5 and 7.5°, respectively.

period as shown in Figure 4.8 for 0° fluxes and in Figure 4.9 for LC fluxes. Especially for NOAA 15 and 16 rejection rates are close to zero when applying the introduced criteria as marked by the large blue areas. The new criteria clearly profit from the higher flux levels measured by the 90° detector while the old criterion exhibits a weakness in treating high flux levels too strictly.

The rejection rate projections shown in Figures D.2 and D.3 are based on data from the whole year 2003 for the 0° and LC flux, respectively. As expected, rejection rates in the 0° detector are lower when averaged over a whole year compared to the Halloween SPE event. A region of elevated data rejection is still visible in the pre-midnight sector where proton populations are strong due to particle injections from the magnetotail and plasmasheet. As electrons drift eastwards and protons westwards, the percentage of present electrons compared to protons can be considered rather low. The comparison of LC rejection rates during the two chosen periods yields lower rejection rates in high latitudes when considering a whole year. During SPEs, solar protons precipitate over the polar cap and can also penetrate latitudes with closed field lines. The abundance of precipitating protons in high latitudes is therefore elevated during SPEs, presenting a challenge during electron flux correction and naturally leading to higher rejection rates in these regions.

In addition to the LC electron flux projections presented in Chapter 4, Figure D.4 shows electron fluxes in different locations averaged over the whole year 2003. While flux intensity distributions are similar, with a stronger asymmetry in lower energy channels,

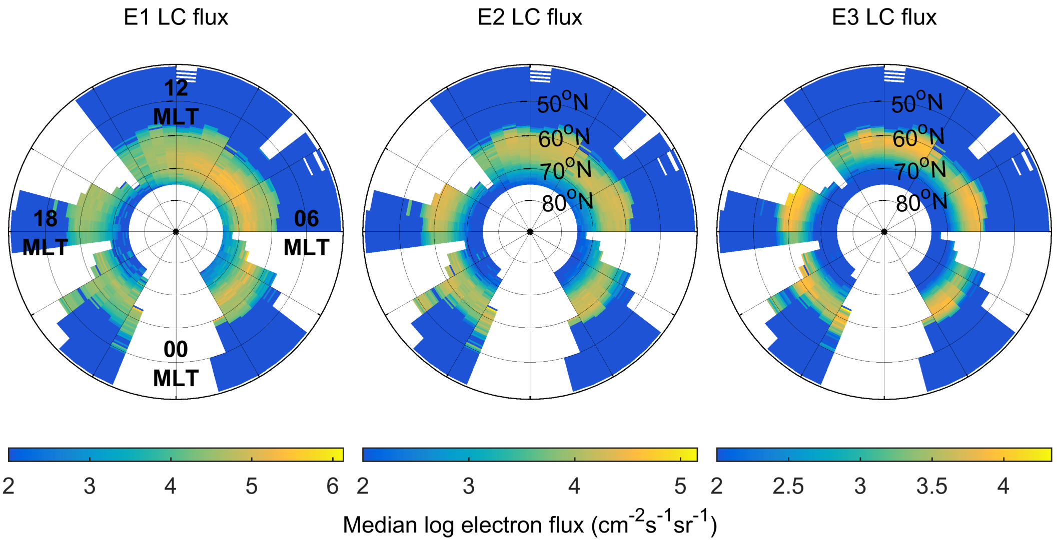


FIGURE D.4: Maps showing the median corrected E1, E2 and E3 LC fluxes during DOY 1 - 365 in 2003 binned according to geomagnetic latitude and longitude (0.5 and 7.5° resolution, respectively). During correction for contaminating protons, the 2σ-criterion was used.

the general flux strength is weaker compared to Figure 4.10 and the auroral oval is narrower. After all, Figure 4.10 concentrates on a SPE, which is a time period of generally elevated geomagnetic activity and higher fluxes.

Appendix E

Elaboration on "nan-median" and "zero-mean" fluxes

In order to establish that most rejected data points are related to low electron fluxes, the single requirements given in Table 4.1 and 4.2 are examined more detailed. The first requirement under investigation is Requirement 2, i.e. the 2σ-criterion, used after correction for contaminating protons. By definition of the criterion, lower electron fluxes are more likely to be rejected than higher fluxes. Although rejected fluxes might well exceed the noise level, the reason for rejection is that the probability of the measurement being related to a zero electron flux is considered too high. Referring to Figure 4.3, high electron fluxes are mainly rejected during periods of very elevated proton fluxes, for example during SPEs. During normal conditions, the probability of falsely rejecting a high or moderate electron flux are zero.

In the next step, fluxes rejected due to unphysical energy spectra are analysed (Requirement 3). Here, the integral energy spectrum is investigated and whenever a flux in a higher energy channel exceeds the flux measured in a lower energy channel, e.g. E2 > E1, all simultaneous measurements in E1, E2 and E3 are rejected. These cases occur when fluxes are as low as the noise level so that by chance the higher energy channel measures a larger flux than the lower energy channel. Thus, fluxes in the affected energy channels are very low and can be replaced by zero flux. However, with e.g. the E3 flux

Tested statement	DOY 1 - 365 in 2003: rate of fulfilment (%)	DOY 302 - 304 in 2003: rate of fulfilment (%)
I: E1 > E2	97.1	95.3
II: E2 < E3	13.5	11.3
III: I $\wedge$ II	11.5	9.7
III while E1 > 250 fu	4.1	9.0
III while E1 > 1000 fu	2.2	8.2

Table E.1: Examination of the energy spectrum criterion: How often are E1 data points from the 0° detector rejected due to noise level fluxes in E2 and E3 and which flux levels can the rejected data points be assigned to. Two periods with different average geomagnetic activity are tested. (fu = flux units).

exceeding the E2 flux, the electron flux measured in the E1 channel is not necessarily small. Table E.1 states how often certain statements are fulfilled (in percent) in two different time periods. Statement I checks whether the integral electron flux measured in the E1 channel is indeed larger than the flux measured by the E2 channel. More than 95 % of all considered data points fulfil this statement in both time periods. Statement II then checks, how often measured fluxes are bigger in the E3 channel than in the E2 channel. Data points fulfilling this statement are cause for the complete rejection of all three fluxes from E1, E2 and E3. Therefore, statement III tests, how often data is rejected due to low fluxes in E2 and E3 while the simultaneous flux measurement in E1 would fulfil the spectral requirement. Here it is important to note that the percentages for statements II and III do not differ significantly when comparing the two time periods. In addition, statement III accounts for almost all cases of statement II, meaning in 85 % of cases where measurements are rejected because E2 < E3, the E1 measurement would have fulfilled the requirement. The next two statements continue to investigate these "falsely" rejected E1 data points. During average conditions, 4.1 % of all E1 data points are "falsely" rejected and lie above the noise level. During the 2003 Halloween SPE, 9 % of all E1 data was affected in that manner. Checking for rejected E1 fluxes above 1000 $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ renders values of 2.2 and 8.2 % during average and elevated geomagnetic activity, respectively. Thus, while the overall fraction of rejected E1 data points remains fairly similar, rejected E1 fluxes are probably much higher during SPEs.

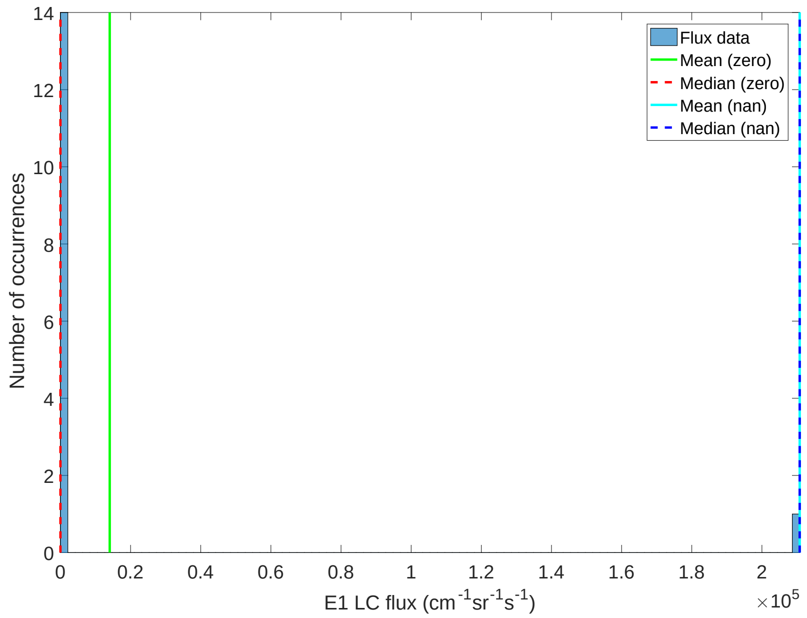


Figure E.1: Example for different averaging outcomes in a data bin where the "nan-median" flux exceeds the "zero-mean" flux by at least one order of magnitude.

The final cause for data rejection during processing occurs during losscone flux calculations. The best-fitting pitch angle distribution is chosen based on the ratio of the 90° and 0° detector fluxes. Whenever pitch angle diffusion is very low and the centre of the losscone is depleted, it can happen that the 0° detector measures zero flux. In these occasions, no losscone flux can be determined. A very depleted losscone may still contain some particles at its edges resulting in a non-zero losscone flux, but these flux levels can be assumed to be very low.

Having established that most rejected electron fluxes correspond to low flux values, the different interpretations of the mean and median are to be discussed. When for example analysing the most commonly appearing fluxes in a certain time bin or dependant on another variable, applying the median to the selection of data point will yield the answer. Thus, figures similar to Figure 4.11 where data is binned according to two variables (L shell and Ap index), calculating the median in each bin gives information on the flux most probable in the respective bin. When, however, analysing electron fluxes and their effect in the atmosphere over a certain time, it is misleading to apply the median. Very high fluxes will yield strong atmospheric effects, but are completely neglected by the median as long as they do not account for more than 50 % of the data. The same argumentation holds true for very low fluxes. With respect to their effect, all fluxes are of equal importance and one should therefore use a statistical method that reflects the whole flux spectrum. When calculating daily fluxes, linear averaging should be preferred. As the calculation of daily fluxes also implies zonal averaging, the

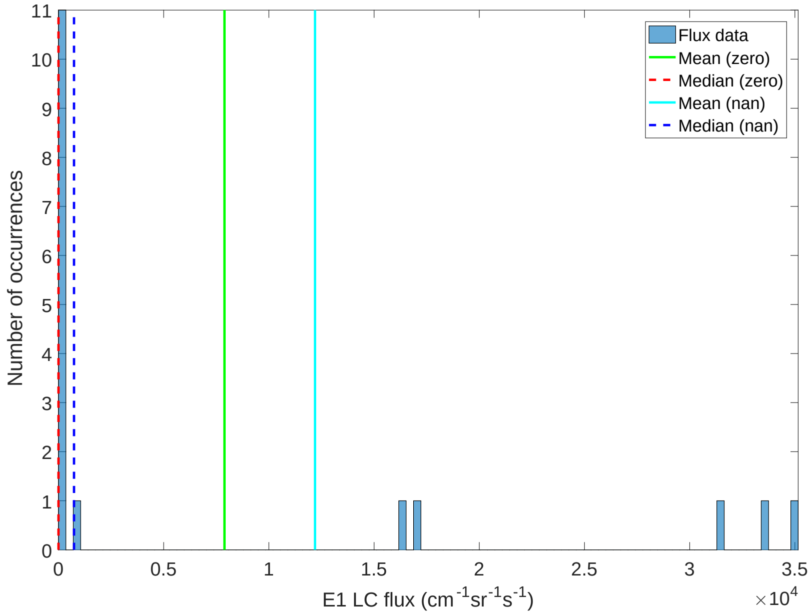


FIGURE E.2: Example for different averaging outcomes in a data bin where the "zero-mean" flux exceeds the "nan-median" flux by at least one order of magnitude.

determination of a zonal median in a 3-hour interval and a subsequent linear averaging to obtain daily fluxes is reasonable as well. The latter is, however, difficult to combine with the replacement of rejection fluxes by a zero flux. An example is given in Figure E.2. E1 LC data from NOAA 15 was binned according to time (DOY 299 - 310 in 2003 with a temporal resolution of six hours) and L shell (bin width of 0.5). After binning, four different data treatments were conducted. The median and mean were applied to both the dataset ignoring rejected data ("nan") and the dataset replacing these values by zero ("zero"). In 94 % of all investigated bins the "nan-median" and "zero-mean" values lie within one order of magnitude from each other. In 0.7 % of the bins the "nan-median" exceeds the "zero-mean" by at least one order of magnitude. Figure E.1 gives an example for the flux distribution of the "zero" dataset within such a bin. This dataset was chosen as it visualizes the number of rejected data as well. It is obvious that this particular been is characterized by a high rejection rate of 93 % which leads to a clear peak in data occurrences at zero flux. Zero flux data points can either result from the noise level correction which applies for both the "zero" and the "nan" dataset or from replacing rejected data points in the "zero" dataset. In addition, there exists only one flux measurement unlike zero. As the majority of data points is zero, applying the median to the "zero" dataset yields zero flux. Since apparently all zero flux data points originate from rejected data, they are ignored by applying the mean and median to the "nan" dataset. Thus, both the "nan-median" and "nan-mean" are centred in the one non-zero flux measurement. The mean value of the "zero" dataset however is located at approximately $1.5 \cdot 10^4\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$ reflecting the whole spectre of fluxes within the bin.

Whereas bins with very high rejection rates lead to higher "nan-medians", low and moderate rejection rates give rise to higher "zero-mean" values. In approximately 5 % of the examined bins, the "zero-mean" exceeds the "nan-median" by at least one order of magnitude. An example for flux data in such a bin in shown in Figure E.2. While there is a clear peak at zero flux, six flux measurements of higher fluxes are contained as well. Since the "nan-median" is not centred at zero, not all zero fluxes result from rejected data and the rejection rate in this bin lies at 35 %. Due to the large amount of small flux data however, both median values lie within the low flux region, completely neglecting the significantly higher fluxes at roughly $10^4\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}$. Both mean values, however, are pulled towards higher flux levels, reflecting the non-zero effect these fluxes would cause.

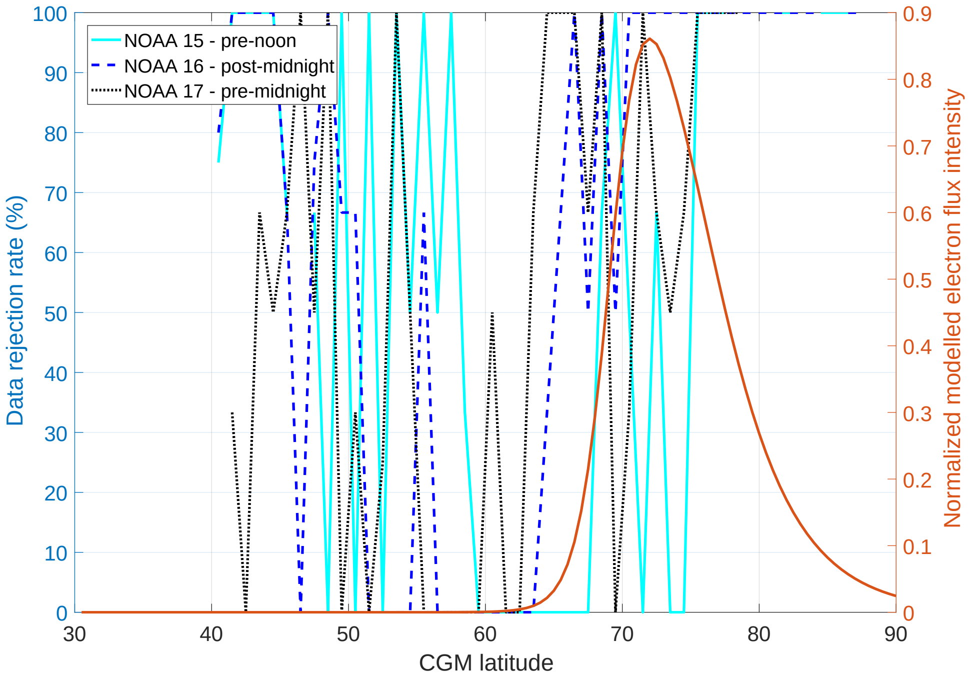
Appendix F

Meridional rejection rate profiles

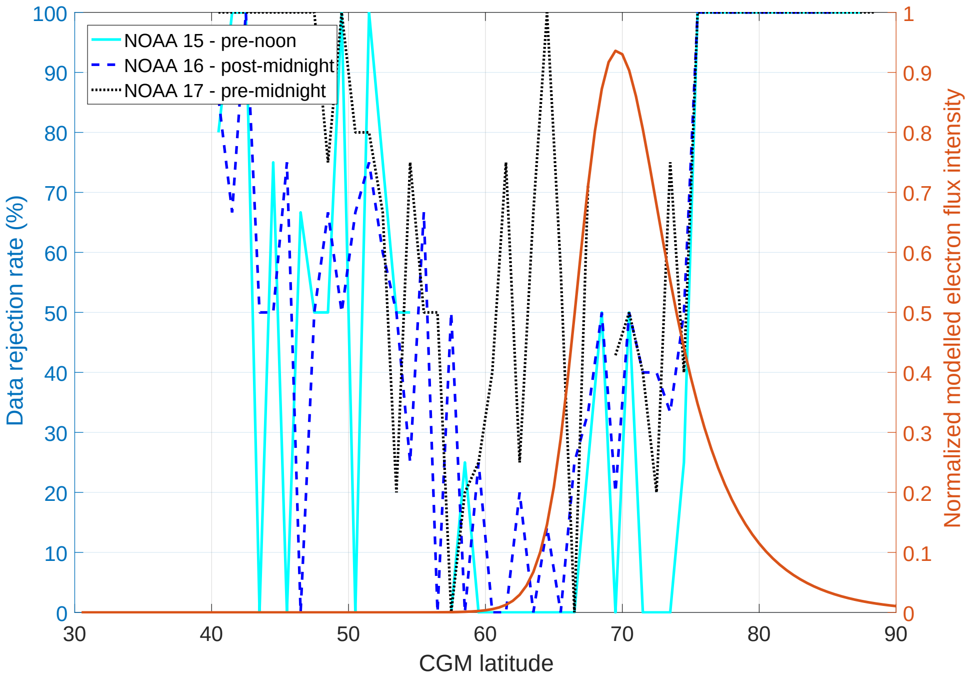
Three-hour intervals from the three peak days of the chosen SPEs are binned according to their 3-hour Kp value into categories specified in Table F.1. For each Kp bin, the latitudinal dependence of the data rejection rate is investigated. This is done for the pre-noon NOAA 15, post-midnight NOAA 16 and pre-midnight NOAA 17. An estimated auroral oval position is shown in Figure F.1 (orange line). Data in each Kp category and latitude bin is superposed in order to calculate rejection rates. The latitudinal resolution is 1°. The number of 3-hour intervals in each Kp bin is stated in the right column of Table F.1. As can be seen in Figure F.1, rejection rates are quite high in the auroral oval for all three investigated satellites during low Kp. In the Kp category 2, NOAA 15 and 16 rejection rates in the oval are decreasing to a value of approximately 50 %. A further increase in Kp is accompanied by lower NOAA 15 and 16 rejection rates, where especially NOAA 15 seems to yield good data coverage in the auroral oval. In the highest Kp category an increase in the rejection rate can be noticed. In general, rejection rates are lower in the equatorward end of the auroral oval profile and increase towards the polar end. NOAA 17 rejection rates are quite high in the whole oval.

Kp category	Included Kp values	Number of 3-hour intervals
1	$[0.0, 1.5[$	5
2	$[1.5, 3.0[$	27
3	$[3.0, 4.5[$	39
4	$[4.5, 6.0[$	33
5	$[6.0, 8.0[$	31
6	$[8.0, 10.0[$	9

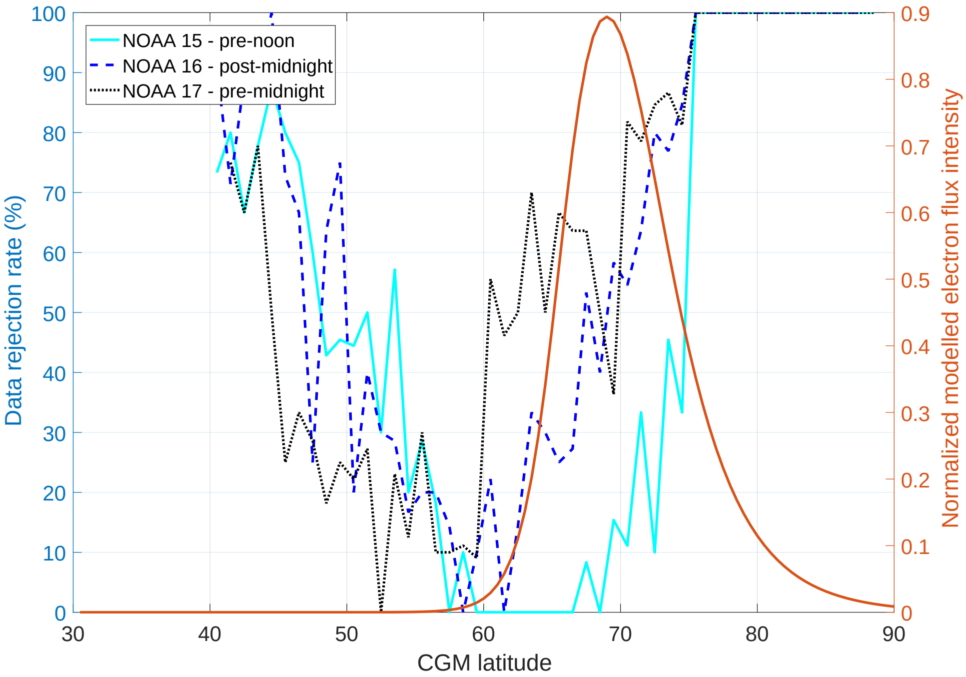
TABLE F.1: Amount of 3-hour intervals in each Kp category throughout the six chosen SPEs.



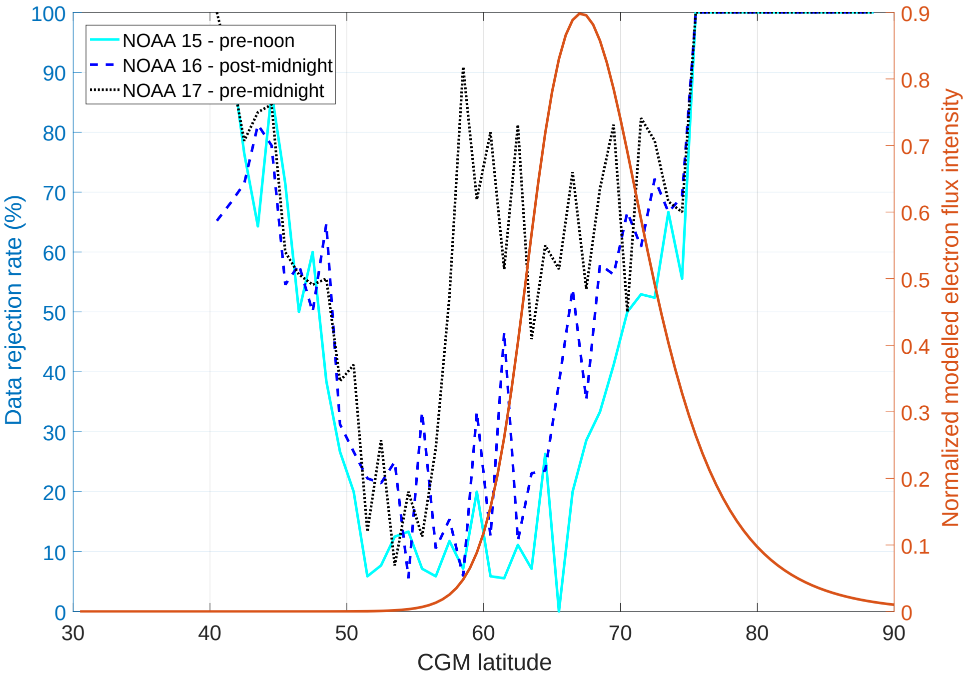
(a) Kp category 1.



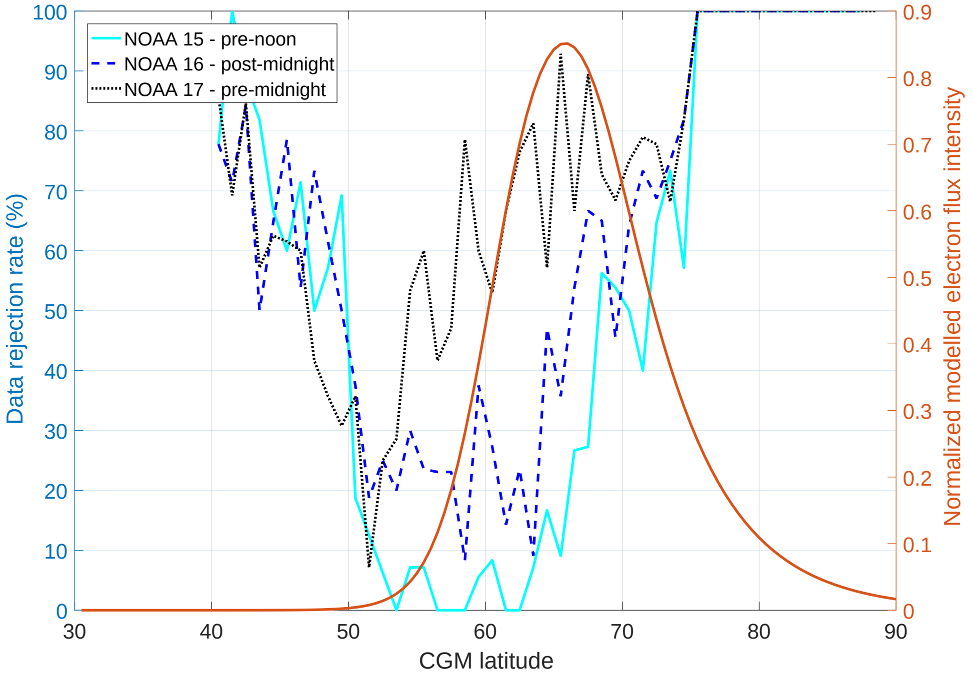
(b) Kp category 2.



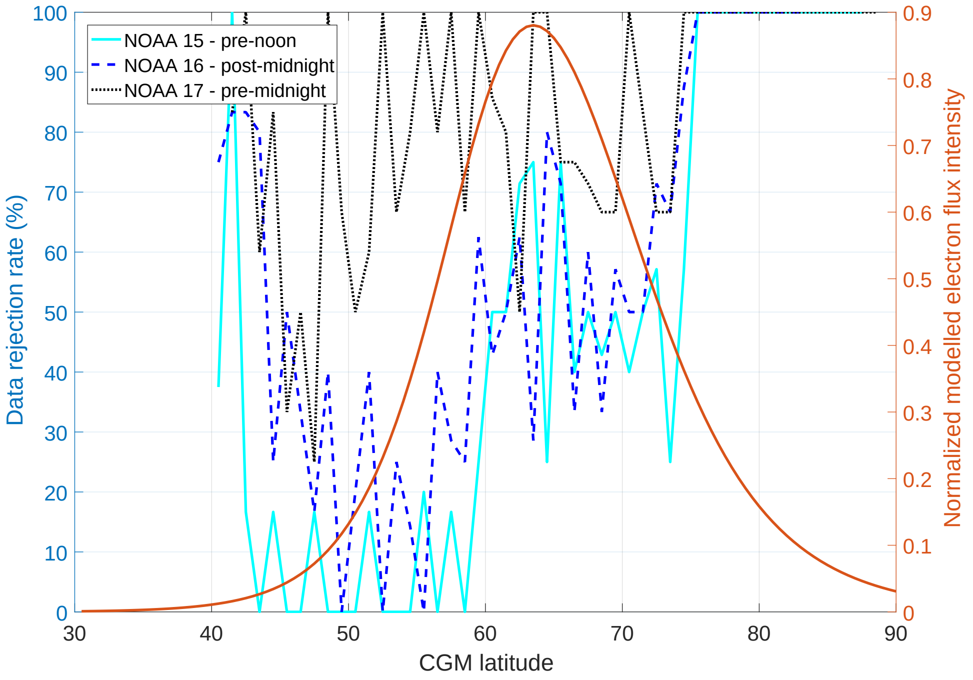
(c) Kp category 3.



(d) Kp category 4.



(e) Kp category 5.



(f) Kp category 6.

FIGURE F.1: Kp category 1 - 6 rejection rates (3-hour mean) binned into 1° latitudinal bins and averaged. This is done for NOAA 15 pre-noon, NOAA 16 post-midnight and NOAA 17 pre-midnight data and shown together with the estimated auroral oval intensity.

List of Figures

List of Tables

Bibliography

Andersson, M. E. et al. (2014). “Longitudinal hotspots in the mesospheric OH variations due to energetic electron precipitation”. In: *Atmospheric Chemistry and Physics* 14.2, pp. 1095–1105. ISSN: 16807316. DOI: `10.5194/acp-14-1095-2014`.

Asikainen, T. and K. Mursula (2011). “Recalibration of the long-term NOAA/MEPED energetic proton measurements”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 73.2-3, pp. 335–347. ISSN: 1364-6826. DOI: `10.1016/J.JASTP.2009.12.011`.

Asikainen, T., K. Mursula, and V. Maliniemi (2012). “Correction of detector noise and recalibration of NOAA/MEPED energetic proton fluxes”. In: *Journal of Geophysical Research: Space Physics* 117.9, pp. 1–16. ISSN: 21699402. DOI: `10.1029/2012JA017593`.

Baker, D. N. et al. (2013a). “A Long-Lived Relativistic Electron Storage Ring Embedded in Earth’s Outer Van Allen Belt”. In: *Science* 340, pp. 186 –190. DOI: `10.1126/science.1233518`.

Baker, D. N. et al. (2013b). “The Relativistic Electron-Proton Telescope (REPT) Instrument on Board the Radiation Belt Storm Probes (RBSP) Spacecraft : Characterization of Earth ’ s Radiation Belt High-Energy Particle Populations”. In: *Space Science Reviews* 179, pp. 337–381. DOI: `10.1007/s11214-012-9950-9`.

Banaszkiewicz, M, WI Axford, and JF McKenzie (1998). “An analytic solar magnetic field model”. In: *Astronomy and Astrophysics* 337, pp. 940–944.

Bartels, J., N. H. Heck, and H. F. Johnston (1939). “The three-hour-range index measuring geomagnetic activity”. In: *Journal of Geophysical Research* 44.4, p. 411. ISSN: 0148-0227. DOI: `10.1016/B978-1-4832-1304-0.50006-3`.

Basu, Sarbani and H M Antia (2004). “Constraining solar abundances using helioseismology”. In: *The Astrophysical Journal* 606.1, p. 12. ISSN: 0004-637X. DOI: `10.1086/421110`. arXiv: `0403485 [astro-ph]`.

Basu, Sarbani and H.M. Antia (2008). “Helioseismology and solar abundances”. In: *Physics Reports* 457.5, pp. 217–283. ISSN: 03701573. DOI: `10.1016/j.physrep.2007.12.002`.

Baumjohann, Wolfgang and Rudolf A Treumann (1997). *Basic space plasma physics*. World Scientific.

Bell, AR (1978). “The acceleration of cosmic rays in shock fronts–I”. In: *Monthly Notices of the Royal Astronomical Society* 182.2, pp. 147–156.

Biermann, Ludwig (1951). “Kometenschweife und solare Korpuskularstrahlung”. In: *Zeitschrift fur Astrophysik* 29, p. 274.

Birkeland, Kristian (1908). *The Norwegian aurora polaris expedition 1902-1903*. Vol. 1. H. Aschelhoug & Company.

Borovsky, Joseph E and Michael H Denton (2006). “Differences between CME-driven storms and CIR-driven storms”. In: *Journal of Geophysical Research* 111.March, pp. 1–17. DOI: `10.1029/2005JA011447`.

Bothmer, V. and I. A. Daglis (2007). *Space Weather: Physics and Effects*. Springer. ISBN: 978-3-540-34578-7.

Breneman, H Herbert and EC Stone (1985). “Solar coronal and photospheric abundances from solar energetic particle measurements”. In: *The Astrophysical Journal* 299, pp. L57–L61.

Burlaga, L et al. (1981). “Magnetic Loop Behind an Interplanetary Shock: Voyager, Helios, and IMP 8 Observations”. In: *Journal of Geophysical Research* 86.A8, pp. 6673–6684. ISSN: 0148-0227. DOI: `10.1029/JA086iA08p06673`.

Burton, R K, R L McPherron, and C T Russell (1975). “An empirical relationship between interplanetary conditions and Dst”. In: *Journal of Geophysical Research* 80.31, pp. 4204–4214. ISSN: 01480227. DOI: `10.1029/JA080i031p04204`.

Carpenter, D. L. (1963). “Whistler Measurements of Electron Density and Magnetic Field Strength in the Remote Magnetosphere”. In: *Journal of Geophysical Research* 68.13, pp. 3727–3730.

Cayton, Thomas E (2007). *Numerical Modeling of the Omnidirectional Spectrometer of the Medium Energy Proton and Electron Detector (MEPED) Subsystem of the Space Environment Monitor (SEM) that Flew Aboard Several Polar Orbiting Satellites*. Tech. rep. Lacanau, France.

Chappell, C R, K K Harris, and G W Sharp (1970). “A study of the influence of magnetic activity on the location of the plasmapause as measured by OGO 5”. In: *Journal of Geophysical Research* 7.1.

Charney, J G and P. G. Drazin (1961). “Propagation of Planetary-Scale Disturbances from the Lower into the Upper Atmosphere”. In: *Journal of Geophysical Research* 66.1, pp. 83–109. DOI: `10.1029/JZ066i001p00083`.

Chen, Yue, Geoffrey D Reeves, and Reiner HW Friedel (2007). “The energization of relativistic electrons in the outer Van Allen radiation belt”. In: *Nature Physics* 3.9, pp. 614–617.

Christensen-Dalsgaard, J. et al. (1996). “The current state of solar modeling”. English. In: *Science* 272.5266. Copyright - Copyright American Association for the Advancement of Science May 31, 1996; Last updated - 2010-06-08; CODEN - SCIEAS, p. 1286.

Cortie, AL (1912). “Sun-spots and terrestrial magnetic phenomena, 1898-1911”. In: *Monthly Notices of the Royal Astronomical Society* 73, p. 52.

Cully, C M, J W Bonnell, and R E Ergun (2008). “THEMIS observations of long-lived regions of large-amplitude whistler waves in the inner magnetosphere”. In: *Geophysical Research Letters* 35.April, pp. 1–5. DOI: `10.1029/2008GL033643`.

Davis, T. Neil and Masahisa Sugiura (1966). "Auroral electrojet activity index AE and its universal time variations". In: *Journal of Geophysical Research* 71.3, pp. 785–801. ISSN: 01480227. DOI: `10.1029/JZ071i003p00785`.

Di Mauro, M. P. et al. (2002). "Inferences on the solar envelope with high-degree modes". In: *A&A* 384, pp. 666–667.

Dragt, Alex J., Micheal M. Austin, and R Stephen White (1966). "Cosmic Ray and Solar Proton Albedo Neutron Decay Injection". In: *Journal of Geophysical Research* 71.5, pp. 1293–1304.

Dungey, J. W. (1961). "Interplanetary Magnetic Field and the Auroral Zones". In: *Phys. Rev. Lett.* 6 (2), pp. 47–48. DOI: `10.1103/PhysRevLett.6.47`.

Evans, D S and M S Greer (2004). "Polar Orbiting Environmental Satellite Space Environment Monitor - 2 : Instrument Descriptions and Archive Data Documentation Space Environment Center Boulder , Colorado December , 2000". In: *NOAA Technical Memorandum* version 1.4.

Falthammar, C.-G. (1965). "Effects of Time-Dependent Electric Fields on Geomagnetically Trapped Radiation". In: *Journal of Geophysical Research* 70.11, pp. 2503–2516.

Freden, Stanley C. and R Stephen White (1960). "Particle Fluxes in the Inner Radiation Belt". In: *Journal of Geophysical Research* 65.5, pp. 1377–1383.

F.R.S., S. Chapman (1930). "XXXV. On ozone and atomic oxygen in the upper atmosphere". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 10.64, pp. 369–383. DOI: `10.1080/14786443009461588`.

Funke, B, G P Stiller, and T Von Clarmann (2014). "Mesospheric and stratospheric NO y produced by energetic particle precipitation during 2002 – 2012". In: *Journal of Geophysical Research: Atmospheres* 119, pp. 4429–4446. DOI: `10.1002/2013JD021404`.

Galand, M and David Evans (2000). "Radiation Damage of the Proton MEPED Detector on POES (TIROS/NOAA) Satellites". In: *NOAA Technical Report OAR* 42.456.

Ganushkina, N Yu et al. (2011). "Locations of boundaries of outer and inner radiation belts as observed by Cluster and Double Star". In: *Journal of Geophysical Research* 116.June, pp. 1–18. DOI: `10.1029/2010JA016376`.

Gleissberg, W (1939). "A long-periodic fluctuation of the sun-spot numbers". In: *The Observatory* 62, pp. 158–159.

Gombosi, Tamas I (1998). *Physics of the space environment.* Cambridge University Press.

Gosling, J. T. "Coronal Mass Ejections and Magnetic Flux Ropes in Interplanetary Space". In: American Geophysical Union. ISBN: 9781118663868.

Gosling, J. T. (1997). "Physical nature of the low-speed solar wind". In: *AIP Conference Proceedings* 385.1, pp. 17–24. DOI: `10.1063/1.51743`.

Gray, L J et al. (2010). "Solar influences on climate". In: *Reviews of Geophysics* 48.RG4001. DOI: `10.1029/2009RG000282`.

Hagfors, Tor and Kristian Schlegel (2001). "Earth's ionosphere". In: *The Century of Space Science.* Ed. by Johan A. M. Bleeker, Johannes Geiss, and Martin C. E.

Huber. Dordrecht: Springer Netherlands, pp. 1559–1584. ISBN: 978-94-010-0320-9. DOI: `10.1007/978-94-010-0320-9_64`.

Hale, George E (1924). "Sun-spots as magnets and the periodic reversal of their polarity". In: *Nature* 113.2829, pp. 105–112.

Hansteen, V (1992). "A New Interpretation of the Red-Shift Observed in Optically Thin Transition Region Lines". In: *Solar wind seven: proceedings of the 3rd COSPAR Colloquium held in Goslar, Germany, 16-20 September 1991.* Vol. 3. Pergamon, p. 49.

Heirtzler, J.R (2002). "The future of the South Atlantic anomaly and implications for radiation damage in space". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 64.16. Space Weather Effects on Technological Systems, pp. 1701 –1708. ISSN: 1364-6826. DOI: `10.1016/S1364-6826(02)00120-7`.

Hill, V. J., D. S. Evans, and H. H. Sauer (1985). "TIROS/NOAA Satellites Space Environment Monitor Archive Tape Documentation". In: *NOAA Technical Memorandum ERL SEL-71.* Boulder, Colorado.

Holton, James R and Gregory J Hakim (2012). *An introduction to dynamic meteorology.* fifth edition. Vol. 88. Academic press.

Horne, Richard B et al. (2005). "Wave acceleration of electrons in the Van Allen radiation belts". In: *Nature* 437.September, pp. 227–230. DOI: `10.1038/nature03939`.

Jackman, Charles H et al. (2009). "Long-term middle atmospheric influence of very large solar proton events". In: *Journal of Geophysical Research* 114.D11304, pp. 1–14. DOI: `10.1029/2008JD011415`.

Jones, Frank C. and Donald C. Ellison (1991). "The plasma physics of shock acceleration". In: *Space Science Reviews* 58.1, pp. 259–346. ISSN: 1572-9672. DOI: `10.1007/BF01206003`.

Kahler, S (1994). "Injection profiles of solar energetic particles as functions of coronal mass ejection heights". In: *The Astrophysical Journal* 428, pp. 837–842.

Kahler, S. W. (2001). "The correlation between solar energetic particle peak intensities and speeds of coronal mass ejections: Effects of ambient particle intensities and energy spectra". In: *Journal of Geophysical Research* 106.A10, p. 20947. ISSN: 0148-0227. DOI: `10.1029/2000JA002231`.

Kahler, S. W. and D. V. Reames (2003). "Solar Energetic Particle Production by Coronal Mass Ejection–Driven Shocks in Solar Fast-Wind Regions". In: *The Astrophysical Journal*, pp. 1063–1070.

Kallenrode, May-Britt (2004). *Space physics: an introduction to plasmas and particles in the heliosphere and magnetospheres.* third edition. Springer Science & Business Media.

Kelley, Michael C. (2009). *The Earth's Ionosphere: Plasma Physics and Electrodynamics.* 2nd edition. San Diego: Academic Press. ISBN: 978-01-208-8425-4.

Kivelson, Margaret G and Christopher T Russell (1995). *Introduction to space physics.* Cambridge university press.

Kodera, Kunihiko and Yuhji Kuroda (2002). “Dynamical response to the solar cycle”. In: *Journal of Geophysical Research* 107.D24, pp. 4749–4761. DOI: `10.1029/2002JD002224`.

Lam, Mai Mai et al. (2010). “Origin of energetic electron precipitation >30 keV into the atmosphere”. In: *Journal of Geophysical Research A: Space Physics* 115.A4, pp. 1–15. ISSN: 21699402. DOI: `10.1029/2009JA014619`.

Lee, Martin A (1983). “Coupled hydromagnetic wave excitation and ion acceleration at interplanetary traveling shocks”. In: *Journal of Geophysical Research: Space Physics* 88.A8, pp. 6109–6119.

Li, Xinlin et al. (2006). “Correlation between the inner edge of outer radiation belt electrons and the innermost plasmapause location”. In: *Geophysical Research Letters* 33.May, pp. 2–5. DOI: `10.1029/2006GL026294`.

Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann (2004). “The Life Cycle of the Northern Hemisphere Sudden Stratospheric Warmings”. In: *Journal of Climate* 17, pp. 2584–2596.

Lopez, R. E., Daniel N Baker, and J. Allen (2004). “Sun Unleashes Halloween Storm”. In: *Eos* 85.11, pp. 105–108.

Low, BC and JR Hundhausen (1995). “Magnetostatic structures of the solar corona. 2: The magnetic topology of quiescent prominences”. In: *The Astrophysical Journal* 443, pp. 818–836.

Lyons, L. R. and D. S. Evans (1984). “An Association Between Discrete Aurora and Energetic Particle Boundaries”. In: *Journal of Geophysical Research: Space Physics (1978-2012)* 89.A4, pp. 2395–2400. ISSN: 0148-0227. DOI: `10.1029/JA089iA04p02395`.

Lyons, Lawrence R, Richard Mansergh Thorne, and Charles F Kennel (1972). “Pitch-Angle Diffusion of Radiation Belt Electrons within the Plasmasphere”. In: *Journal of Geophysical Research* 77.19, pp. 3455–3474.

Marshall, John and R Alan Plumb (2016). *Atmosphere, ocean and climate dynamics: an introductory text*. Vol. 21. Academic Press.

Mason, GM and TR Sanderson (1999). “CIR associated energetic particles in the inner and middle heliosphere”. In: *Corotating Interaction Regions*. Springer, pp. 77–90.

Matsumi, Yutaka and Masahiro Kawasaki (2003). “Photolysis of Atmospheric Ozone in the Ultraviolet Region”. In: *Chemical Reviews* 103, pp. 4767–4781. DOI: `10.1021/cr0205255`.

Matthes, Katja et al. (2017). “Solar forcing for CMIP6 (v3.2)”. In: *Geoscientific Model Development* 10.6, pp. 2247–2302. ISSN: 19919603. DOI: `10.5194/gmd-10-2247-2017`.

McIlwain, C E (1961). “Coordinates for mapping the distribution of magnetically trapped particles”. In: *Journal of Geophyisical Research* 66, pp. 3681–3691.

Merrill, Ronald T and Michael W McElhinny (1983). “The Earth’s magnetic field: Its history, origin and planetary perspective”. In: *International geophysics series* 32.

Murphy, R. J. et al. “Accelerated Particle Composition and Energetics and Ambient Abundances from Gamma-Ray Spectroscopy of the 1991 June 4 Solar Flare”. In: *The Astrophysical Journal* 490.2, p. 883.

Nesse Tyssøy, H et al. (2016). “Energetic electron precipitation into the middle atmosphere - Constructing the loss cone fluxes from MEPED POES”. In: *Journal of Geophysical Research: Space Physics* 121, pp. 5693–5707. DOI: `10.1002/2016JA022752`.

Neugebauer, M. and Conway W. Snyder (1966). “Mariner 2 observations of the solar wind: 1. Average properties”. In: *Journal of Geophysical Research* 71.19, pp. 4469–4484. ISSN: 01480227. DOI: `10.1029/JZ071i019p04469`.

Newton, H. W. and M. L. Nunn (1951). “The Sun’s Rotation Derived from Sunspots 1934–1944 and Additional Results”. In: *Monthly Notices of the Royal Astronomical Society* 111.4, p. 413. DOI: `10.1093/mnras/111.4.413`.

Nolte, J. T. et al. (1976). “Coronal holes as sources of solar wind”. In: *Solar Physics* 46.2, pp. 303–322. ISSN: 1573-093X. DOI: `10.1007/BF00149859`.

Ødegaard, L-K. G. et al. (2016). “Space Weather impact on the degradation of energetic proton detectors”. In: *Journal of Space Weather and Space Climate* 6, pp. 1–15. ISSN: 2115-7251. DOI: `10.1051/swsc/2016020`.

Ødegaard, Linn Kristine Glesnes et al. (2017). “Energetic electron precipitation in weak to moderate corotating interaction region-driven storms”. In: *Journal of Geophysical Research: Space Physics* 122.3, pp. 2900–2921. ISSN: 21699402. DOI: `10.1002/2016JA023096`.

Oinas, V. et al. (2001). “Radiative cooling by stratospheric water vapor: Big differences in GCM results”. In: *Geophysical Research Letters* 28.14, pp. 2791–2794. DOI: `10.1029/2001GL013137`.

Parker, Eugene N (1958). “Dynamics of the interplanetary gas and magnetic fields.” In: *The Astrophysical Journal* 128, p. 664.

Peredo, M. et al. (1995). “Three-dimensional position and shape of the bow shock and their variation with Alfvénic, sonic and magnetosonic Mach numbers and interplanetary magnetic field orientation”. In: *Journal of Geophysical Research* 100.A5, p. 7907. ISSN: 0148-0227. DOI: `10.1029/94JA02545`.

Proedrou, Elisavet, Klemens Hocke, and Peter Wurz (2016). “The middle atmospheric circulation of a tidally locked Earth-like planet and the role of the sea surface temperature”. In: *Progress in Earth and Planetary Science* 3.1, p. 22. ISSN: 2197-4284. DOI: `10.1186/s40645-016-0098-1`.

Pulkkinen, Tuija I. et al. (2007). “Differences in geomagnetic storms driven by magnetic clouds and ICME sheath regions”. In: *Geophysical Research Letters* 34.2, pp. 5–8. ISSN: 00948276. DOI: `10.1029/2006GL027775`.

Reames, D. V. (2004). “Solar energetic particle variations”. In: *Advances in Space Research* 34.2, pp. 381–390. ISSN: 02731177. DOI: `10.1016/j.asr.2003.02.046`.

Reames, Donald V. (2013). “The two sources of solar energetic particles”. In: *Space Science Reviews* 175.1-4, pp. 53–92. ISSN: 00386308. DOI: `10.1007/s11214-013-9958-9`. arXiv: `1306.3608`.

Richardson, I. G. et al. (2006). “Major geomagnetic storms (Dst <= -100 nT) generated by corotating interaction regions”. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–17. ISSN: 21699402. DOI: `10.1029/2005JA011476`.

Rodger, Craig J and Mark A Clilverd (2008). “Hiss from the chorus”. In: *Nature* 452.March, pp. 41–42. DOI: `10.1038/452041a`.

Rodger, Craig J et al. (2010). “Use of POES SEM-2 observations to examine radiation belt dynamics and energetic electron precipitation into the atmosphere”. In: *Journal of Geophysical Research* 115.A04202, pp. 1–13. DOI: `10.1029/2008JA014023`.

Rodger, Craig J. et al. (2013). “Comparison between POES energetic electron precipitation observations and riometer absorptions: Implications for determining true precipitation fluxes”. In: *Journal of Geophysical Research: Space Physics* 118.12, pp. 7810–7821. ISSN: 21699402. DOI: `10.1002/2013JA019439`.

Rostoker, Gordon (1972). “Geomagnetic indices”. In: *Reviews of Geophysics* 10.4, pp. 935–950. ISSN: 19449208. DOI: `10.1029/RG010i004p00935`.

Royrvik, O. and T N Davis (1977). “Pulsating Aurora: Local and Global Morphology”. In: *Journal of Geophysical Research* 82.29, pp. 4720–4740.

Rozanov, E. et al. (2012). “Influence of the Precipitating Energetic Particles on Atmospheric Chemistry and Climate”. In: *Surveys in Geophysics* 33.3, pp. 483–501. ISSN: 1573-0956. DOI: `10.1007/s10712-012-9192-0`.

Russell, C T (1993). “Magnetic Fields of the Terrestrial Planets”. In: *Journal of Geophysical Research* 98.E10, pp. 18,681–18,695. ISSN: 01480227. DOI: `10.1029/93JE00981`.

Russell, C. T. (2001). “Solar wind and interplanetary magnetic field: A tutorial”. In: *Space Weather*. Ed. by P. Song, H. J. Singer, and G. L. Siscoe. Americal Geophysical Union, pp. 73–88. DOI: `10.1029/GM125p0073`.

Sandanger, Marit Irene et al. (2015). “In-flight calibration of NOAA POES proton detectors - Derivation of the MEPED correction factors”. In: *Journal of Geophysical Research A: Space Physics* 120.11, pp. 9578–9593. ISSN: 21699402. DOI: `10.1002/2015JA021388`.

Seppälä, A et al. (2013). “Geomagnetic activity signatures in wintertime stratosphere wind , temperature , and wave response”. In: *Journal of Geophysical Research* 118, pp. 2169–2183. DOI: `10.1002/jgrd.50236`.

Seppälä, Annika et al. (2007). “Arctic and Antarctic polar winter NO x and energetic particle precipitation in 2002 – 2006”. In: *Geophysical Research Letters* 34.L12810. DOI: `10.1029/2007GL029733`.

Seppälä, Annika et al. (2014). “What is the solar influence on climate ? Overview of activities during CAWSES-II”. In: *Progress in Earth and Planetary Science* 1:24, pp. 1–12. DOI: `10.1186/s40645-014-0024-3`.

Shea, M. and D. F. Smart (1990). “A Summary of Major Solar Proton Events”. In: *Solar Physics* 127, pp. 297–320.

Shindell, Drew T. et al. (2001). “Solar Forcing of Regional Climate Change During the Maunder Minimum”. In: 294.5549, pp. 2149–2152. ISSN: 0036-8075. DOI: `10.1126/science.1064363`.

Singer, S. F. (1958). “Trapped Albedo Theory of the Radiation Belt”. In: *Physical Review Letters* 1.5, pp. 181–183.

Slavin, James A. et al. (1985). “An ISEE 3 Study of Average and Substorm Conditions in the Distant Magnetotail”. In: *Journal of Geophysical Research: Space Physics* 90.A11, pp. 10875–10895. ISSN: 0148-0227. DOI: `10.1029/JA090iA11p10875`.

Smith, Anne K (2012). “Interactions Between the Lower , Middle and Upper Atmosphere”. In: *Space Science Reviews* 168, pp. 1–21. DOI: `10.1007/s11214-011-9791-y`.

Sugiura, M (1964). *Hourly values of equatorial Dst for the IGY. Annual International Geophysical Year, vol. 35.*

Thomas, Barry T and Edward J Smith (1980). “The Parker spiral configuration of the interplanetary magnetic field between 1 and 8.5 AU”. In: *Journal of Geophysical Research: Space Physics* 85.A12, pp. 6861–6867. ISSN: 0148-0227. DOI: `10.1029/JA085iA12p06861`.

Thorne, Richard Mansergh (1980). “The importance of energetic particle precipitation on the chemical composition of the middle atmosphere”. In: *Pure and Applied Geophysics PAGEOPH* 118.1, pp. 128–151. ISSN: 00334553. DOI: `10.1007/BF01586448`.

Tsurutani, Bruce T. et al. (2006). “Corotating solar wind streams and recurrent geomagnetic activity: A review”. In: *Journal of Geophysical Research: Space Physics* 111.7, pp. 1–25. ISSN: 21699402. DOI: `10.1029/2005JA011273`.

Tsurutani, BT (2000). “Solar/interplanetary plasma phenomena causing geomagnetic activity at Earth”. In: *PROCEEDINGS-INTERNATIONAL SCHOOL OF PHYSICS ENRICO FERMI*. Vol. 142. IOS Press; Ohmsha; 1999, pp. 273–286.

Turunen, Esa et al. (2009). “Impact of different energies of precipitating particles on NOx generation in the middle and upper atmosphere during geomagneti storms”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 71.10-11, pp. 1176–1189. DOI: `10.1016/j.jastp.2008.07.005`.

Van Allen, J A and L A Frank (1959). “Radiation Around The Earth To A Radial Distance Of 107,400 Km”. In: *Nature* Vol: 183. DOI: `10.1038/183430a0`.

Van Allen, JA et al. (1958). “Observations of high intensity radiation by satellites 1958 alpha and gamma, lC Y. Satellite Report Series No. 3: Some preliminary reports of experiments in satellites 1958 alpha and gamma”. In: *NAat. Acad. Sci., Nat. Res. Council, Wash., DC*, pp. 767–769.

Van de Kamp, M. et al. (2016). “A model providing long-term data sets of energetic electron precipitation during geomagnetic storms”. In: *Journal of Geophysical Research: Atmospheres* 121, pp. 12.520–12.540. DOI: `10.1002/2015JD024212`.

Verronen, P T et al. (2015). "Contribution of proton and electron precipitation to the observed electron concentration in October – November 2003 and September 2005". In: *Annales Geophysicae* 33, pp. 381–394. DOI: `10.5194/angeo-33-381-2015`.

Walt, M (1996). "Source and loss processes for radiation belt particles". In: *Radiation belts: models and standards*, pp. 1–13.

Wang, Yuming et al. (2004). "Deflection of coronal mass ejection in the interplanetary medium". In: *Solar Physics* 222.2, pp. 329–343. ISSN: 00380938. DOI: `10.1023/B:SOLA.0000043576.21942.aa`.

Webb, David F. (1991). "The solar cycle variation of the rates of CMEs and related activity". In: *Advances in Space Research* 11.1, pp. 37–40. ISSN: 02731177. DOI: `10.1016/0273-1177(91)90086-Y`.

Webb, David F and Joe H Allen (2004). "Spacecraft and ground anomalies related to the October-November 2003 solar activity". In: *Space Weather* 2.3.

Wild, JP, SF Smerd, and AA Weiss (1963). "Solar bursts". In: *Annual Review of Astronomy and Astrophysics* 1.1, pp. 291–366.

Yando, Karl et al. (2011). "A Monte Carlo simulation of the NOAA POES Medium Energy Proton and Electron Detector instrument". In: *Journal of Geophysical Research* 116.A10231, pp. 1–13. DOI: `10.1029/2011JA016671`.

Yang, Gui-Ying, B. J. Hoskins, and J. M. Slingo (2011). "Equatorial Waves in Opposite QBO Phases". In: *Journal of Atmospheric Sciences* 68, pp. 839–862. DOI: `10.1175/2010JAS3514.1`.

Zhang, J. et al. (2007). "Solar and interplanetary sources of major geomagnetic storms (Dst <= -100 nT) during 1996-2005". In: *Journal of Geophysical Research: Space Physics* 112.10. ISSN: 21699402. DOI: `10.1029/2007JA012321`.

Zhang, Y. and L.J. Paxton (2008). "An empirical Kp-dependent global auroral model based on TIMED/GUVI FUV data". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 70.8-9, pp. 1231–1242. ISSN: 1364-6826. DOI: `10.1016/J.JASTP.2008.03.008`.